

Characterization, Long-Term Behavior, and Approximation of Koopman Operators

Dissertation

zur Erlangung des Doktorgrades der
Fakultät für Mathematik und Naturwissenschaften
der Bergischen Universität Wuppertal

vorgelegt von
Julian Hölz

Wuppertal, im April 2026

Erstgutachter: Jun.-Prof. Dr. Jochen Glück
Zweitgutachter: Prof. Balint Farkas
Drittgutachter: Prof. Dr. Tanja Eisner
Tag der mündlichen Prüfung: 13.04.2026

Acknowledgements

” *A mathematician is a blind man in a dark room
looking for a black cat which isn't even there.*

— **Charles Darwin**
(who probably never said that in the first place)

What this quote wonderfully illustrates is that doing research in mathematics or perhaps any other field is a path filled with a lot of uncertainties. And as a blind man needs guidance to find his way, I need support and advice from many people to pursue my research. As being industrious without any feedback can lead to going in circles or simply getting lost. Therefore, this section is here to express my gratitude to all the people who helped me along the way.

First and foremost, I thank my principal advisor Jochen Glück for his continuous support, patience, and guidance throughout my PhD journey. His expertise and insights have been invaluable in shaping my research and helping me navigate the complexities of my work. His unwavering willingness to answer my questions, no matter the topic, and his eagerness to discuss mathematical ideas at any time of the day have greatly contributed to my growth as a mathematician. In addition, I am grateful to the many opportunities he provide me to present my research at conferences and workshops, meet new and interesting people, and expand my academic network. It was the continuous exposure with other researchers and their ideas that led to some of the most interesting results.

During my PhD, I had the privilege to visit several researchers. I want to thank Markus Haase for inviting me to visit him at the university of Kiel to present some of the main results of this thesis. I also want to thank him for the interesting discussions that we had not just during my stay but also during the several times we met at conferences. Furthermore, I want to thank Anke Kalauch for inviting me to Dresden to discuss pre-Riesz spaces, and for Alexandre Mauroy and his working group for being fanatic hosts in Namur.

I also thank my co-author Alexander Dobrick. Spending time with him at conferences and in Kiel can only be described as a treat. I also want to thank him for providing me the this \LaTeX template and for suggesting to start the acknowledgements section with this wonderful and fitting quote. I also like to thank my co-authors Till Hauser and Markus Kunze for the fruitful collaboration on our joint projects.

I would like to extend my gratitude to Bálint Farkas and Henrik Kreidler, who served as my primary advisors when it comes to ergodic theory. Their expertise and guidance have been instrumental in shaping my understanding of the field and have significantly contributed to my research development. My sincere thanks especially goes to Henrik, as he had shown me an article that help shape my later research greatly.

I want to thank my colleagues who have created a fun and entertaining working environment within the research group. Mehmet Erbay, Merlin Schmitz, René Hosfeld, Lukas Vorberg, Jonathan Mui and Annika Meyer have contributed to some of the best memories during my PhD time, be it through mathematical discussions, extended coffee breaks, or social events. I am grateful for the camaraderie and support we shared, which made the challenging times more manageable and the good times even more enjoyable. I especially thank Mehmet and Merlin for reading parts of this thesis and providing valuable feedback.

I thank my long time partner Sophie for her endless support, patience, and understanding during the ups and downs of my PhD journey. Her encouragement and belief in me have been a necessary source of motivation. Her unwavering support has been instrumental in my success, and I am deeply appreciative of her presence in my life.

My flat mates, Max and Luca, also deserve a special thanks for creating a lively and enjoyable living environment. Their friendship and support have been incredibly valuable and made my stay in Wuppertal all the more memorable.

My family has been supportive throughout my life and I consider myself lucky to be a part of it. Their belief in me has been a constant source of motivation, and I am deeply appreciative of their presence in my life.

Contents

Introduction	1
Preliminaries and Notation	11
1 Topological Dynamical Systems	13
1.1 The Koopman and the Perron–Frobenius Operators	13
1.1.1 The Topology on the State Space	14
1.1.2 The Koopman Operator	16
1.1.3 The Perron–Frobenius Operator	20
1.2 Duality of the Perron–Frobenius and Koopman Operators	24
1.2.1 Norming Dual Pairs and Weak Topologies	24
1.2.2 The Koopman and Perron–Frobenius Operators as Adjoints .	33
1.2.3 Weakly Closed Ideals and Weakly Continuous Functionals . .	35
1.3 Characterizing Topological Dynamics with Operators	39
1.4 The Koopman Operator on the Bounded Measurable Functions . . .	46
1.5 Notes	49
2 Measure-Preserving Dynamical Systems	51
2.1 Measure-Preserving Systems and their Koopman Operators	51
2.2 Mean Ergodicity, Ergodicity and Mixing	58
2.3 From Topological to Measure-Preserving Dynamics	64
2.4 Notes	67
3 Entropies of Dynamical Systems	69
3.1 Measure-Theoretic Entropy	69
3.2 Approximation Entropy	73
3.2.1 Definition and Basic Properties	73
3.2.2 Lower Bounds for the Approximation Entropy	81
3.2.3 A Spectral-Theoretic Perspective on Approximation Entropy	93
3.3 Interplay between Measure-Theoretic and Approximation Entropy .	96
3.4 Topological Entropy	102
3.5 Notes	104
4 Data-Driven Approximation of Koopman Operators	105
4.1 The Extended Dynamic Mode Decomposition	107
4.1.1 How to Collect Data	108
4.1.2 A Description of the Algorithm	111
4.1.3 A Discussion on Ergodic Convergence Rates	120

4.1.4	The Structure of the Optimal Approximation	122
4.2	Lower Dimension Bound by Measure-Theoretic Entropy	126
4.3	Lower Dimension Bound by Approximation Entropy	131
4.4	Notes	134
5	Entropy via Banach Lattices	137
5.1	Abstract Entropy on Banach Lattices	137
5.2	Measure-Theoretic Entropy	141
5.3	Topological Entropy	148
5.4	Notes	153
6	Uniform Ergodicity for Koopman Operators	157
6.1	Uniform Ergodicity of Lattice Homomorphisms	158
6.2	Uniform Ergodicity for Topological Dynamics	163
6.3	Uniform Ergodicity of Measure-Preserving Dynamics	166
6.4	Quasi-Compactness and Compactness	169
6.5	Notes	175
	Bibliography	179
	Index	187

Introduction

” *Stückeschreiben ist wie Schach: Bei der Eröffnung ist man frei; dann bekommt die Partie ihre eigene Logik.*

— **Friedrich Dürrenmatt**

The theory of using linear operators to study dynamical systems has a long history dating back to the early 1930s when Koopman and von Neumann introduced what are now known as Koopman operators in [Koo31] and [KN32]. In their original papers, they considered time-continuous Hamiltonian dynamics $(\varphi_t)_{t \in \mathbb{R}}$ on a subset $\Omega \subseteq \mathbb{R}^d$, which is invertible and volume-preserving. They noticed that by composition with functions in $L^2(\Omega, \mathcal{B}(\Omega), \lambda_d)$, where λ_d denotes the d -dimensional Lebesgue measure, the resulting semigroup

$$\begin{aligned} T_{\varphi_t} : L^2(\Omega, \mathcal{B}(\Omega), \lambda_d) &\rightarrow L^2(\Omega, \mathcal{B}(\Omega), \lambda_d), \\ f &\mapsto f \circ \varphi_t \end{aligned}$$

becomes a strongly continuous group of unitary operators. This allowed them to apply spectral theory to study properties of the underlying dynamical system.

Since then, Koopman operators have been used extensively in ergodic theory [EFHN15], and have gained popularity as tools in the numerical analysis of dynamical systems [MMS20]. In this thesis, we will provide an overview of both a functional-analytic treatment of dynamical systems via Koopman operators and their applications in data-driven approximation methods. We will always assume that time is discrete, so we consider maps $\varphi: \Omega \rightarrow \Omega$ instead of flows $(\varphi_t)_{t \in \mathbb{R}}$. Let us discuss the types of dynamical systems that we will encounter throughout this thesis.

Continuous and Measure-Preserving Maps

It is a common problem in applied mathematics to encounter initial value problems of the form

$$\dot{x}(t) = F(x(t)), \quad x(0) = x_0 \in \Omega,$$

where Ω is some differential manifold, often just an open subset of \mathbb{R}^d , and $F: \Omega \rightarrow T\Omega$ is a vector field on Ω that maps into the tangent bundle $T\Omega$ of Ω . If well-posed, this initial value problem generates a flow $(\varphi_t)_{t \in \mathbb{R}}$ on Ω , namely a mapping $\varphi: \mathbb{R} \times \Omega \rightarrow \Omega$ such that $\varphi_0 = \text{id}_\Omega$ and $\varphi_{t+s} = \varphi_t \circ \varphi_s$ for all $t, s \in \mathbb{R}$. It is exactly these types of problems that we are interested in studying and approximating.

A first step of approximation is that of a discretization in time. Instead of considering the full flow $(\varphi_t)_{t \in \mathbb{R}}$, we will fix a sample time $\tau > 0$, which can be arbitrarily small, and only consider the time- τ map $\varphi_\tau: \Omega \rightarrow \Omega$. For notational convenience, we denote this map by φ . An approximation of the full flow $(\varphi_t)_{t \in \mathbb{R}}$ can then be obtained by the family $(\varphi^n)_{n \in \mathbb{N}_0}$ of iterates of φ , which coincide with the original flow for all times $t = n\tau$, as $\varphi^n = \varphi_{n\tau}$ holds for all $n \in \mathbb{N}_0$.

Although maps of the form $\varphi: \Omega \rightarrow \Omega$ arise naturally from initial value problems, we will not restrict ourselves to time-discretizations of flows. In fact, most examples that we consider are not time-discretizations. Instead, we will consider maps φ on a set Ω for their own sake. Two important classes of such maps are *topological* and *measure-preserving* maps.

In topological dynamics, we assume that Ω carries a topology and that φ is continuous with respect to this topology. In the time-discretized case, this is usually the case, as for many problems the flow $(\varphi_t)_{t \in \mathbb{R}}$ depends continuously on the initial value x_0 . The treatment of topological dynamical systems classically focuses on spaces Ω that are compact and metrizable or compact and Hausdorff. In the spirit of [FK20], we will consider a class of spaces that allows for a treatment of more general spaces. For our purposes, the requirement of complete regularity and the Hausdorff property will be sufficient.

In measure-preserving dynamics, we will use a probability space $X = (X, \Sigma, \mu)$ (which we denote by X instead of Ω) as the underlying state space and assume that $\varphi: X \rightarrow X$ is measurable with respect to Σ and preserves the measure of sets in the sense that $\mu(\varphi^{-1}(A)) = \mu(A)$ for all $A \in \Sigma$. In Koopman's original setting, the Hamiltonian flow preserves the volume measure on $\Omega \subseteq \mathbb{R}^d$, which is why the time- τ -discretization map φ is measure-preserving with respect to the Lebesgue measure. Measure-preserving maps, however, arise naturally in many other contexts as well. If $\varphi: \Omega \rightarrow \Omega$ is a continuous map on a compact topological space Ω , then the Krylov–Bogolyubov theorem (see Theorem 2.3.3) guarantees the existence of at least one Borel probability measure μ on Ω that is preserved by φ .

Observations and Observables

Only when we start taking measurements can we begin to learn properties of the dynamical systems. But how do we measure a dynamical system? We can imagine a measurement to be some kind of value, usually a real or complex number, that is returned when the system is in a certain state. In a system of climate dynamics, for example, this could be the temperature, wind speed, humidity, or air pressure at a certain location. In a mechanical system, this could be the position, velocity or acceleration of some component of the system. Mathematically, we model this using a function f that returns this measurement value $f(x)$ when the system is in state x .

In the topological setting, we will be using functions $f: \Omega \rightarrow \mathbb{K}$, where $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, and usually assume that f is bounded and continuous, so that small changes in the state x lead to small changes in the measurement value $f(x)$. When the situation requires it, we can occasionally use the less regular class of bounded measurable functions instead.

In the measure-preserving setting, we will be using functions $f: X \rightarrow \mathbb{K}$ that are measurable with respect to Σ and usually require some integrability assumption

with respect to the probability measure μ . We further identify functions that agree almost everywhere with respect to μ , and thus, work with equivalence classes of functions. The most common function spaces that we will be using are the spaces $L^p(X)$ for $p \in [1, \infty]$.

It is a common critique of choosing observables in $L^p(X)$ that point evaluations $f(x)$ of equivalence classes $f \in L^p(X)$ are not well-defined. However, in applications the situation is backwards. We usually begin with an everywhere defined function $f: X \rightarrow \mathbb{K}$ that we want to use as an observable and only later decide to work in the space $L^p(X)$ for some $p \in [1, \infty]$ and do a μ -almost everywhere analysis there.

The dynamical system φ allows us to measure not only current states but also future states. After one time step has passed, we observe the output $f(\varphi(x))$ when the state of the system is $\varphi(x)$; and after n time steps, we will measure the value $f(\varphi^n(x))$ at the new state $\varphi^n(x)$ of the system. The transition from f to $f \circ \varphi$ is modeled by the Koopman operator

$$T_\varphi f := f \circ \varphi.$$

In the topological setting, we define T_φ as an operator on $C_b(\Omega)$, the space of bounded continuous functions on Ω with values in \mathbb{K} , or on $B_b(\Omega)$, the space of bounded Borel measurable functions on Ω with values in \mathbb{K} . For measure-preserving systems, T_φ is an operator on $L^p(X)$ for some $p \in [1, \infty]$.

The Functional-Analytic Approach

The Koopman operators T_φ allow us to employ an entirely different set of mathematical tools to study dynamical systems. But before this, we will need to establish a way to translate from notions of dynamical systems into notions of operator theory. For topological dynamical systems, we will show correspondences of properties in Chapter 1. For the Koopman operators acting on $C_b(\Omega)$ and $B_b(\Omega)$, we obtain the following correspondences, which can be found as Theorem 1.3.3, Theorem 1.3.5, Theorem 1.4.4 and Theorem 1.4.5 as original results of this thesis. We provide an informal version of their statements here.

Theorem. *Let Ω be a completely regular Hausdorff space and $\varphi: \Omega \rightarrow \Omega$ a continuous map.*

- (i) *The map φ is injective if and only if the Koopman operator T_φ on $C_b(\Omega)$ (or $B_b(\Omega)$) has dense range with respect to a weak topology on $C_b(\Omega)$ (or $B_b(\Omega)$).*
- (ii) *The map φ has a dense image if and only if the Koopman operator T_φ on $C_b(\Omega)$ is injective.*
- (iii) *The map φ is surjective if and only if the Koopman operator T_φ on $B_b(\Omega)$ is injective.*

We will further discuss conditions that guarantee that φ is a homeomorphism in terms of properties of T_φ and extend these correspondences to the adjoint operator of T_φ , the so-called Perron–Frobenius operator P_φ , which acts on the space $M(\Omega)$ of complex-valued Radon measures on Ω , a space that is dual to $C_b(\Omega)$ in a weak sense.

In the measure-preserving setting, we will use an entirely different set of tools. From the classical ergodic theorem of Birkhoff (see Proposition 2.2.1), we know that the Koopman operator T_φ acting on $L^1(X)$ is pointwise mean ergodic, i.e., the time averages

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} (T_\varphi^k f)(x_0)$$

converge for μ -almost every $x_0 \in X$ and all $f \in L^1(X)$, which implies convergence in norm of the Cesàro averages $\frac{1}{n} \sum_{k=0}^{n-1} T_\varphi^k f$ in $L^p(X)$ for all $p \in (1, \infty)$ and $f \in L^p(X)$. A further powerful tool we will use frequently is the Jacobs–de Leeuw–Glicksberg decomposition (see Definition 2.2.6), which allows us to decompose the space $L^p(X)$ into

$$L^p(X) = L^p(X)_{\text{rev}} \oplus L^p(X)_{\text{aws}},$$

where $L^p(X)_{\text{rev}}$ and $L^p(X)_{\text{aws}}$ are closed T_φ -invariant subspaces, called the *reversible* and *almost weakly stable* part of $L^p(X)$, respectively.

On the reversible part, the operator T_φ behaves in a structured and predictable manner. The span of all eigenfunctions of T_φ corresponding to unimodular eigenvalues is dense in $L^p(X)_{\text{rev}}$ and the restriction of T_φ to $L^p(X)_{\text{rev}}$ is invertible. The almost weakly stable part $L^p(X)_{\text{aws}}$ of the operator T_φ captures the chaotic part of the system. If φ is invertible, it contains no non-trivial finite-dimensional invariant subspace. For each $f \in L^p(X)_{\text{aws}}$, the sequence of means $\left(\frac{1}{n} \sum_{k=0}^{n-1} T_\varphi^k f\right)_{n \in \mathbb{N}}$ converges weakly to zero. It is well-known that this decomposition allows us to translate the statistical properties of strong and weak mixing of a dynamical system into a decomposition of the Koopman operator. We will recall this in Proposition 2.2.10 and Proposition 2.2.13, since this will serve us as a key tool in characterizing the ergodic behavior of observables later on.

Approximation and Entropy

It is exactly the chaotic behavior of the almost weakly stable part of the Koopman operator that we aim to understand. One way to measure the chaotic behavior of a dynamical system is by determining its measure-theoretic entropy. The measure-theoretic entropy of a measure-preserving dynamical system (X, φ) is a non-negative number $h_\mu(\varphi)$ that quantifies the average information that is produced per time step when observing the system with a measurement that can differentiate finitely many states. A high entropy indicates that the system produces a lot of new information per time step and is thus more chaotic than a system with low entropy. We will provide a rather intuitive introduction to measure-theoretic entropy in Section 3.1.

We will also get to know another type of entropy, the so-called *approximation entropy* $h_{\text{apr}}(T_\varphi)$ of a Koopman operator T_φ , which was introduced in [Voi95]. This entropy quantifies how well the Koopman operator T_φ , in this case on the Hilbert space $L^2(X)$, can be approximated by finite-rank operators. A high approximation entropy indicates that the operator T_φ is hard to approximate by finite-rank operators, while a low approximation entropy indicates that good approximations by finite-rank operators exist. It turns out that both entropies are zero if $L^2(X)_{\text{aws}} = \{0\}$ (see Proposition 3.2.9 and Theorem 3.3.12). We also show that positive measure-theoretic

entropy always implies that the approximation entropy is unbounded (see Theorem 3.3.12), prove lower bounds on the approximation entropy (see Theorem 3.2.11), and introduce the concept of a fractional approximation entropy that will capture chaotic behavior that neither the measure-theoretic entropy nor the approximation entropy are able to detect (see Definition 3.2.21).

Approximation and Mean Ergodic Convergence Rates

As discussed, it is characteristic for observables $f \in L^p(X)_{\text{aws}}$ that the time averages $\frac{1}{n} \sum_{k=0}^{n-1} T_\varphi^k f$ converge weakly to zero. For $p = 2$, this gives us the convergence of the averaged correlation coefficients

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} |\langle f, T_\varphi^k f \rangle| = 0.$$

This convergence governs the rate at which Cesàro averages converge, as by Proposition 4.1.15 we have for each $f \in L^2(X)$ and $n \in \mathbb{N}$ that

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} T_\varphi^k f \right\|^2 \leq \frac{2}{n} \sum_{k=0}^{n-1} |\langle f, T_\varphi^k f \rangle|.$$

Thus, we may expect that fast convergence of the averaged correlation coefficients is good for statistical purposes, as it implies fast convergence of the Cesàro averages. We will find, however, that the fast convergence of Cesàro averages has some undesirable consequences. If this convergence is too fast, then we can show that the approximation entropy, which measures how well the Koopman operator can be approximated by finite-rank operators, becomes large.

For a finite set $\mathcal{O} \subseteq L^2(X)$, a finite-time horizon $n \in \mathbb{N}$ and an error bound $\delta > 0$, we let $H_{\text{apr}}(T_\varphi, \mathcal{O}, n, \delta)$ denote the minimal rank of a finite-rank orthogonal projection $P: L^2(X) \rightarrow L^2(X)$ that satisfies

$$\|PT_\varphi^k f - T_\varphi^k f\| \leq \delta$$

for all $k \in \{1, \dots, n-1\}$ and all $f \in \mathcal{O}$. Then we obtain the following growth rate of $H_{\text{apr}}(T_\varphi, \{f\}, n) := \sup_{\delta > 0} H_{\text{apr}}(T_\varphi, \{f\}, n, \delta)$ for observables $f \in L^2(X)_{\text{aws}}$ with sufficiently fast decay of correlation coefficients. This can be found as Theorem 3.2.11 and is original to this thesis.

Theorem. *Let $\varphi: X \rightarrow X$ be an invertible measure-preserving dynamical system on the probability space $X = (X, \Sigma, \mu)$, denote by T_φ its Koopman operator on $L^2(X)$, and let $f \in L^2(X)_{\text{aws}}$ be normalized. If*

$$C := \sum_{k=1}^{\infty} |\langle f, T_\varphi^k f \rangle| < \infty,$$

then $H_{\text{apr}}(T_\varphi, \{f\}, n) \geq n \frac{1}{1+2C}$ for all $n \in \mathbb{N}$ and all $\delta > 0$.

A finer analysis gives us lower bounds for fractional approximation entropies by the mean ergodic convergence rate of the observable f (see Theorem 3.2.18). This theorem has not been published or submitted for publication yet.

Theorem. Let $\varphi: X \rightarrow X$ be a measure-preserving dynamical system on the probability space $X = (X, \Sigma, \mu)$, denote by T_φ its Koopman operator on $L^2(X)$, and let $f \in L^2(X)_{\text{aws}}$ be normalized and $\alpha \in (0, 1)$. If

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} T_\varphi^k f \right\|^2 \leq C \frac{1}{n^\alpha}$$

holds for some $C > 0$ and all $n \in \mathbb{N}$, then $H_{\text{apr}}(T_\varphi, \{f\}, n) \geq n^{2\alpha-1} \frac{1}{C}$ for all $n \in \mathbb{N}$ and all $\delta > 0$.

Both results have direct implications for approximation methods for Koopman operators, or more generally, for any attempt to approximate the Koopman operator on a finite-dimensional subspace. If we let

$$h_{\text{apr}}(T_\varphi, F) := \sup_{\substack{\mathcal{O} \subseteq F \\ \mathcal{O} \text{ finite}}} \sup_{\delta > 0} \lim_{n \rightarrow \infty} \frac{1}{n} H_{\text{apr}}(T_\varphi, \mathcal{O}, n, \delta),$$

then we obtain in Theorem 4.3.1 the following result, which has been published in [HH25] by Hauser and the present author.

Theorem. Let $\varphi: X \rightarrow X$ be an invertible measure-preserving system on the probability space $X = (X, \Sigma, \mu)$, let $F \subseteq L^2(X)$ be a finite-dimensional linear subspace and $\varepsilon > 0$. Then there exists $K_0 \in \mathbb{N}$ and $\delta > 0$ that satisfy the following: Whenever for a linear operator $\hat{T}: F \rightarrow F$ and some $K \geq K_0$ the assertion

$$\|\hat{T}^k f - T_\varphi^k f\| \leq \delta \|f\|$$

holds for all $f \in F$ and all $k \in \{0, \dots, K-1\}$, we conclude that

$$\dim(F) \geq K(h_{\text{apr}}(T_\varphi, F) - \varepsilon).$$

Let us interpret the statement of the theorem. The finite-dimensional operator \hat{T} can be any approximation of the Koopman operator on the subspace F . This can be obtained by a data-driven method such as the extended Dynamic Mode Decomposition (eDMD) or by any other approximation technique. The only constraint is that the approximation that we obtain is linear and leaves a finite-dimensional space invariant (whereas the Koopman operator does not need to leave F invariant). Any matrix approximation of T_φ falls into this category. If we use this approximation to successfully predict the actual behavior of the Koopman operator on the subspace F for a finite time horizon $\{0, \dots, K-1\}$, then the theorem states that the dimension of the subspace F must have been large, when the approximation entropy of T_φ on F is large. This shows that chaotic behavior of the dynamical system, which is captured by a large approximation entropy, makes it hard to approximate the Koopman operator on finite-dimensional spaces.

Furthermore, we show that for certain types of subsets of observables the growth rate becomes exponential and is driven by measure-theoretic entropy $h_\mu(\varphi, \alpha)$ with respect to a finite Σ -measurable partition α of X . Note that $L^2(X | \alpha)$ denotes the space of all Σ -measurable functions that are constant on the elements of α . The following is Theorem 4.2.1 and has been previously published in [HH25].

Theorem. Let $\varphi: X \rightarrow X$ be a measure-preserving system on the probability space $X = (X, \Sigma, \mu)$, $\varepsilon > 0$, and let α be a finite Σ -measurable partition of X . Then there exist $K_0 \in \mathbb{N}$ and $\delta > 0$ that satisfy the following: Whenever a linear operator $\hat{T}: L^2(X | \alpha) \rightarrow L^2(X | \alpha)$ and some $K \geq K_0$ satisfy

$$\|\hat{T}^k f - T_\varphi^k f\| \leq \delta \|f\|$$

for all $f \in L^2(X | \alpha)$ and $k \in \{0, \dots, K - 1\}$, we conclude that

$$|\alpha| \geq \exp(K(h_\mu(\varphi, \alpha) - \varepsilon)).$$

Operator-Theoretic Entropies

The fact that measure-preserving entropy appears is due to the special structure of the subspace $F = L^2(X | \alpha)$. Orthogonal projections onto this subspace are so-called Markov projections. If we change our definition of the quantity $H_{\text{apr}}(T_\varphi, \mathcal{O}, n, \delta)$ to be the minimal rank of a finite-rank Markov projection $P: L^2(X) \rightarrow L^2(X)$ that satisfies

$$\|PT_\varphi^k f - T_\varphi^k f\| \leq \delta$$

for all $k \in \{1, \dots, n - 1\}$ and all $f \in \mathcal{O}$, then we obtain a different type of approximation entropy $h_{\text{bl}}(T_\varphi, F)$, which we call the *Banach lattice entropy* (see Definition 5.1.4). This entropy measures how well the Koopman operator can be approximated by finite-rank Markov projections on a subspace $F \subseteq L^2(X)$.

We will show in Theorem 5.2.1 that for ergodic measure-preserving dynamics $\varphi: X \rightarrow X$ the Banach lattice entropy of the Koopman operator T_φ on the entire space $L^2(X)$ coincides with the measure-theoretic entropy of φ . Since Banach lattice entropy admits generalizations to other Banach lattices such as $C(K)$ for some compact Hausdorff space K , we prove that for invertible topological dynamical systems $\varphi: K \rightarrow K$ the Banach lattice entropy coincides with the topological entropy $h_{\text{top}}(\varphi)$ of φ (see Theorem 5.3.1).

Uniform Ergodicity

We complete our discussion on mean ergodic rates by studying the measure-preserving and topological systems for which the Cesàro averages $\frac{1}{n} \sum_{k=0}^{n-1} T_\varphi^k f$ converge at the same rate for every observable f in the underlying function space. It then follows from the uniform boundedness principle that the Cesàro averages converge in operator norm; a property that is called *uniform ergodicity*. We will characterize uniform ergodicity for lattice homomorphisms (see Theorem 6.1.2) first and then apply our findings to the Koopman operators of topological and measure-preserving dynamical systems in Theorem 6.2.4 and Theorem 6.3.1, respectively. We find that uniform ergodicity of the Koopman operator is equivalent to eventual periodicity or periodicity of the underlying dynamical system. We will extend our investigation to the class of *quasi-compact* and *compact* Koopman operators in Section 6.4. This work is partly original to this thesis and partly published in [Höl25].

Previous and Other Work

While much of this work is new and has not been previously published, some of the results have appeared in articles that I co-authored during my time as a PhD student. The following articles contain results that are included in this dissertation.

- [HH25] Till Hauser and Julian Hölz. *Entropy based lower dimension bounds for finite-time prediction of Dynamic Mode Decomposition algorithms*. 2025. arXiv: 2504.20269.
- [Höl25] Julian Hölz. „A Note on the uniform ergodicity of dynamical systems“. *Israel J. Math.* (2025).

Additionally, the following articles, which I co-authored during my PhD phase, contain results that have not been included in this thesis.

- [DH24] Alexander Dobrick and Julian Hölz. „Uniform convergence of solutions to stochastic hybrid models of gene regulatory networks“. *J. Evol. Equ.* 24.3 (2024). Id/No 53, p. 17.
- [DHK24] Alexander Dobrick, Julian Hölz, and Markus Kunze. „Ultra Feller operators from a functional-analytic perspective“. *Stud. Math.* 279.3 (2024), pp. 243–271.
- [GH23] Jochen Glück and Julian Hölz. „Eventual cone invariance revisited“. *Linear Algebra Appl.* 675 (2023), pp. 274–293.

Organization of this Thesis

Let us conclude the introduction by providing an overview of the individual chapters of this thesis.

Chapter 1 In this chapter, we introduce the necessary background on topological dynamical systems on completely regular Hausdorff spaces. We discuss necessary conditions for the topology of the state space to guarantee the existence of sufficiently many continuous observables and introduce the Koopman and Perron–Frobenius operators as operators that are dual to each other. We further establish correspondences between properties of the dynamical system and properties of the associated operators.

Chapter 2 This chapter recalls notions for measure-preserving dynamical systems and their associated Koopman operators. We remind the reader of the Jacobs–de Leeuw–Glicksberg decomposition and show how it can characterize ergodic, weakly mixing and strongly mixing dynamical systems in terms of properties of the Koopman operator.

Chapter 3 Chaos in both measure-preserving and topological dynamics can be quantified by means of entropies. In this chapter we introduce the measure-theoretic and topological entropy and present the approximation entropy of Voiculescu. We relate the three notions of entropy to each other and study the lesser-known approximation entropy further. In particular, we provide lower bounds for it in terms of mean ergodic convergence rates.

Chapter 4 In this chapter, we describe the extended Dynamic Mode Decomposition (eDMD) algorithm for the data-driven approximation of Koopman operators in frame-theoretic language. We discuss convergence results for eDMD in both the

independent and identically distributed (i.i.d.) sampling setting and the ergodic sampling setting. We conclude by giving lower bounds on the dimension of the approximating finite-dimensional operator in terms of both the classical measure-theoretic entropy and the approximation entropy.

Chapter 5 We introduce the Banach lattice entropy of operators on Banach lattices as a twist on the approximation entropy that only allows for approximations by finite-rank Markov projections. We show that for Koopman operators on L^2 -spaces the Banach lattice entropy coincides with the measure-theoretic entropy of the underlying dynamical system. We further show that for Koopman operators on spaces of continuous functions over compact Hausdorff spaces the Banach lattice entropy coincides with the topological entropy of the underlying dynamical system.

Chapter 6 This chapter studies uniform ergodicity of Koopman operators in both the topological and measure-preserving setting. We show that uniform ergodicity is equivalent to eventual periodicity of the underlying topological dynamical system and to periodicity of the underlying measure-preserving dynamical system. We further study quasi-compact and compact Koopman operators and characterize the underlying dynamical systems in these cases.

Preliminaries and Notation

We denote by $\mathbb{N} := \{1, 2, 3, \dots\}$ the set of all natural numbers, and set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ as well as $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$. If X is some set, we denote by $\mathbb{1}_M : X \rightarrow \mathbb{R}$ the *indicator function* of some subset $M \subseteq X$. We use the abbreviation $\mathbb{1} := \mathbb{1}_X$, if the role of the set X is clear from the context.

Metric Spaces

If (M, d) is a metric space, then for each $\delta > 0$ and each subset $A \subseteq M$ we denote by $B_{\leq \delta}(A) := \{x \in M : \exists a \in A : d(x, a) \leq \delta\}$ and $B_{< \delta}(A) := \{x \in M : \exists a \in A : d(x, a) < \delta\}$ the *closed* and *open δ -neighborhood* of A , respectively. If $x \in M$ and $A = \{x\}$, we abbreviate with the notation $B_{\leq \delta}(x) := B_{\leq \delta}(\{x\})$ and $B_{< \delta}(x) := B_{< \delta}(\{x\})$.

Dualities

The *norm dual* of a Banach space E will be denoted by E^* . The abbreviation $X' = (X, \tau)'$ is used for the *topological dual* of a locally convex space (X, τ) for some locally convex topology τ on X and is only used when the topology τ is clear from the context. Depending on the context, we denote by $\langle \cdot, \cdot \rangle$ one of the following things:

- (i) The *duality pairing* of a norming dual pair in Chapter 1.
- (ii) In the context of Hilbert spaces, the *inner product*.
- (iii) In all other cases, the *duality pairing* of some Banach space E and its norm dual E^* .

Banach Spaces

Let E be a Banach space with either real or complex scalar field, which we typically denote by \mathbb{K} . For $A \subseteq E$ the *linear span* and the *closed linear span* of A in E is denoted by $\text{span}(A)$ and $\overline{\text{span}}(A)$, respectively.

Hilbert Spaces

If H is a Hilbert space, then we denote by $\langle \cdot, \cdot \rangle_H$ the inner product of H . We drop the subscript H if it is clear from the context. For complex Hilbert spaces, we use the convention that the inner product is linear in the first argument and antilinear in the second argument.

Banach Lattices

For a basic introduction to Banach lattice theory, we refer to [AB06, Chapter 9], [Mey91] or [Sch74].

Linear Operators

If E and F are Banach spaces, we denote the set of all *bounded, linear operators* from E to F by $\mathcal{L}(E; F)$ and endow it with the *operator norm* denoted by $\|\cdot\|$. Further, the abbreviation $\mathcal{L}(E) := \mathcal{L}(E; E)$ is used. The *identity operator* in $\mathcal{L}(E)$ will be denoted by I . We denote the subspace of $\mathcal{L}(E; F)$ of all *compact operators* by $\mathcal{K}(E; F)$ and again abbreviate $\mathcal{K}(E) := \mathcal{K}(E; E)$.

If $T \in \mathcal{L}(E)$, we denote its *norm adjoint operator* by $T^*: E^* \rightarrow E^*$ on its norm dual E^* and its adjoint with respect to a weaker locally convex topology by $T': E' \rightarrow E'$ on its continuous dual space E' . As usual, we denote by $\sigma(T)$ the *spectrum* of the bounded linear operator T . The *resolvent set* is denoted by $\rho(A)$. Moreover, for each $\lambda \in \rho(T)$ we denote the *resolvent* of T at λ by $R(\lambda, A) := (\lambda - T)^{-1}$.

List of Spaces

Let I be an arbitrary non-empty set, let $1 \leq p \leq \infty$ and $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Moreover, let Ω be a topological space, K be a compact Hausdorff space and $X = (X, \Sigma, \mu)$ be a measure space.

Spaces
$B_b(\Omega) := \{f: \Omega \rightarrow \mathbb{K} : f \text{ Borel measurable and bounded}\}$
$C_b(\Omega) := \{f: \Omega \rightarrow \mathbb{K} : f \text{ continuous and bounded}\}$
$M(\Omega) := \{\mu: \Sigma \rightarrow \mathbb{K} : \mu \text{ is a Radon measure}\}$
$\ell^p(I) := \{(x_n)_{n \in I} : \forall n \in \mathbb{N} : x_n \in \mathbb{K} \text{ and } \sum_{n \in I} x_n ^p < \infty\}$
$\ell^p := \ell^p(\mathbb{N})$
$\ell_n^p := \{(x_n)_{n \in \mathbb{N}} \in \ell^p : \forall k > n : x_k = 0\}$
$L^p(X) := \{f: X \rightarrow \mathbb{K} : f \text{ integrable and } \int_X f(x) ^p dx < \infty\}$

If we exclude the real or the complex field as scalar field \mathbb{K} , we mention this explicitly. In any other case, the spaces can be chosen either way.

We endow the spaces with their usual norms as follows: For $f \in C_b(\Omega)$ we set $\|f\|_\infty := \sup_{x \in \Omega} \|f(x)\|$ to be the supremum norm. If $p \in [1, \infty)$ for $(x_n)_{n \in \mathbb{N}} \in \ell^p(\mathbb{N})$ we set $\|(x_n)\|_p := (\sum_{n \in \mathbb{N}} \|x_n\|^p)^{1/p}$, for $f \in L^p(X)$ we set $\|f\|_{L^p} := (\int_\Omega \|f(x)\|^p dx)^{1/p}$ and if $p = \infty$ we set $\|(x_n)\|_\infty := \sup_{n \in \mathbb{N}} \|x_n\|$ for $(x_n)_{n \in \mathbb{N}} \in \ell^\infty(\mathbb{N})$ and $\|f\|_{L^\infty} := \text{esssup}_{x \in \Omega} \|f(x)\|$ for $f \in L^\infty(X)$. Finally, for $\mu \in M(\Omega)$ we set $\|\mu\|_1 := |\mu|(\Omega)$, where $|\mu|$ denotes the *total variation measure* of μ .

Let Ω be a completely regular Hausdorff space and $\varphi: \Omega \rightarrow \Omega$ be continuous. We denote the set of all *probability measures* in $M(\Omega)$ by $M^1(\Omega) := \{\mu \in M(\Omega) : \mu \geq 0, \|\mu\|_1 = 1\}$. We further denote by $M_\varphi^1(\Omega) := \{\mu \in M^1(\Omega) : \varphi_*\mu = \mu\}$ the set of all *φ -invariant probability measures*, where $\varphi_*\mu$ denotes the *pushforward measure* of μ with respect to φ .

This chapter links topological dynamical systems with linear operator theory via Koopman and Perron–Frobenius operators on spaces of continuous and measurable observables.

In the most abstract form, a topological dynamical system is given by a continuous map $\varphi: \Omega \rightarrow \Omega$ on a topological space Ω . In Section 1.1 we begin by discussing sensible assumptions on Ω in order for the space of bounded continuous functions $C_b(\Omega)$ to be rich enough to capture the dynamics of φ through composition. We then introduce the Koopman operator $T_\varphi: C_b(\Omega) \rightarrow C_b(\Omega)$, $f \mapsto f \circ \varphi$ and its dual, the Perron–Frobenius operator P_φ acting on Radon measures.

In order to establish a characterization of Koopman operators among all unital algebra and lattice homomorphisms, we establish a locally convex topology on $C_b(\Omega)$ that is weaker than the norm topology. In Section 1.2 we introduce norming dual pairs and use them to define this topology and to establish the duality between T_φ and P_φ .

In Section 1.3 we then provide operator-theoretic characterizations of topological properties of φ such as injectivity, density of the image, and homeomorphism in terms of properties of T_φ and P_φ such as injectivity, density of the range, isometry, and being a lattice homomorphism.

Finally, in Section 1.4 we extend the Koopman operator to the space $B_b(\Omega)$ of bounded measurable functions and show that this extension allows for a full characterization of surjectivity of φ in terms of properties of T_φ and P_φ .

The purpose of Sections 1.1 and 1.2 is to set up the necessary framework. The major new contributions of this chapter are Sections 1.3 and 1.4. The author has not published or submitted these results for publication yet.

1.1 The Koopman and the Perron–Frobenius Operators

Choosing the right topological requirements on the state space Ω of a topological dynamical system is crucial for studying its dynamics with Koopman operators. The purpose of Section 1.1.1 is to discuss reasonable assumptions.

In Section 1.1.2 we then introduce the Koopman operator T_φ associated with a continuous map $\varphi: \Omega \rightarrow \Omega$ on a topological space Ω as the composition operator acting on the space $C_b(\Omega)$ of bounded continuous functions on Ω . We introduce its dual operator P_φ , the Perron–Frobenius operator in Section 1.1.3, which acts on the space of Radon measures on Ω . We describe basic properties and frame Koopman and Perron–Frobenius operators as the image of functors from the category of topological dynamical systems to the category of Banach spaces with bounded linear operators.

1.1.1 The Topology on the State Space

In the next chapters, we will use continuous functions for the analysis of dynamical systems, or more generally, continuous maps $\varphi: \Omega_1 \rightarrow \Omega_2$ between two topological spaces. In this section we discuss assumptions that we need to pose on Ω_1 and Ω_2 in order for the space of continuous functions to be rich enough. That this is not always guaranteed is shown by the next simple example.

Example 1.1.1. Let $\Omega_1 = \Omega_2 = \mathbb{R}_{\text{ind}}$ be the real line endowed with the indiscrete topology, i.e., the only open sets are \emptyset and \mathbb{R} itself. Note that $\varphi: \Omega_1 \rightarrow \Omega_2$ can be any function, as it will always be continuous. Moreover, the space $C(\mathbb{R}_{\text{ind}})$ consists of the constant functions only and the Koopman operator T_φ is the rank-one identity operator on $C(\mathbb{R}_{\text{ind}})$. In this case, the Koopman operator does not carry any information about the dynamics of φ .

In the above simple example, the topology on \mathbb{R}_{ind} is clearly not rich enough to distinguish between different points in \mathbb{R}_{ind} . Somewhat surprisingly, the standard Hausdorff condition on the topology of Ω is not sufficient to guarantee that the space $C(\Omega)$ is rich enough to provide insight into the dynamics of φ . The relatively prime integer topology on $\Omega = \mathbb{Z}_+$ described in [SS95, Example 60] is an example of a topological space that is Hausdorff and *strongly connected*, i.e., the only continuous functions on X are the constant functions.

In order to effectively use T_φ to study the dynamics of φ , we need functional assumptions on the topological space Ω . There are two candidates that we will consider in this section:

Definition 1.1.2. A topological space Ω is called

- (i) *functionally Hausdorff* if for every $\omega_1, \omega_2 \in \Omega$ with $\omega_1 \neq \omega_2$ there exists a continuous function $f \in C(\Omega)$ such that $f(\omega_1) \neq f(\omega_2)$, and
- (ii) *completely regular* if for every closed subset $A \subseteq \Omega$ and every $\omega_1 \in \Omega \setminus A$ there exists a continuous function $f \in C(\Omega)$ such that $f(\omega_1) = 0$ and $f(\omega_2) = 1$ for all $\omega_2 \in A$.

Although both notions seem closely related, in general, they are independent. For a space that is functionally Hausdorff but not completely regular we refer to [SS95, Examples 68 or 69], and for a space that is completely regular but not functionally Hausdorff consider the indiscrete topology on a space with at least two elements. Notice that a completely regular Hausdorff space is always functionally Hausdorff, as in Hausdorff spaces for each $\omega \in \Omega$ the singleton $\{\omega\}$ is a closed subset of Ω .

The condition of being functionally Hausdorff seems artificial for the purpose of studying dynamical systems by composition with continuous functions. In fact, every Hausdorff topology generated as the initial topology of a family of continuous functions $\{f_i: \Omega \rightarrow \mathbb{R} : i \in I\}$ is completely regular [Eng89, Exercise 1.5.E (a)]. This makes completely regular spaces a more natural choice for our purposes. Thus, we will mostly work with completely regular spaces in the following. The necessity of complete regularity will further be noticeable in many subsequent proofs.

Completely regular spaces behave well under standard operations on topological spaces. Recall that a topological space Ω is called *normal* if every two disjoint closed subsets $A, B \subseteq \Omega$ have disjoint open neighborhoods.

- Remark 1.1.3.** (i) It follows from Urysohn’s lemma [Eng89, Theorem 1.5.11], that every normal space is completely regular. In particular, every metric space is completely regular and Hausdorff.
- (ii) A topological space Ω is completely regular if and only if there exists a compact space K and a mapping $\iota: \Omega \rightarrow K$ that is a homeomorphism onto its image $\iota(\Omega)$ [Eng89, Theorem 3.2.6].
- (iii) It is a simple consequence of (ii) that every subset of a completely regular (and Hausdorff) space is completely regular (and Hausdorff) with respect to the subspace topology.
- (iv) As every locally compact Hausdorff space embeds homeomorphically into its one-point compactification, it follows from (ii) that every locally compact Hausdorff space is completely regular.
- (v) It follows from (ii) and Tychonoff’s theorem [Eng89, Theorem 3.2.4] that the product of completely regular (and Hausdorff) spaces is completely regular (and Hausdorff) with respect to the product topology.

As in normal Hausdorff spaces, in completely regular spaces we can separate compact sets from disjoint closed sets.

Proposition 1.1.4. *Let Ω be a completely regular Hausdorff space and $K \subseteq \Omega$ be a compact subset. Let $A \subseteq \Omega$ be a closed subset with $K \cap A = \emptyset$. Then the following holds:*

- (i) *There exist disjoint open neighborhoods $U, V \subseteq \Omega$ with $A \subseteq U$ and $K \subseteq V$.*
- (ii) *There exists a continuous function $f \in C(\Omega)$ with $f(\Omega) \subseteq [0, 1]$ such that $f(x) = 0$ for all $x \in K$ and $f(x) = 1$ for all $x \in A$.*

Proof. “(i)”: This is [Eng89, Theorem 3.1.6].

“(ii)”: This is [Eng89, Theorem 3.1.7]. □

We finish this section by noting that on a completely regular Hausdorff space the continuity of a mapping can be checked using compositions with continuous functions.

Lemma 1.1.5. *Let Ω_1 be a topological space and Ω_2 be a completely regular Hausdorff space. Then a mapping $\varphi: \Omega_1 \rightarrow \Omega_2$ is continuous if and only if $f \circ \varphi$ is continuous for all $f \in C_b(\Omega_2)$.*

Proof. “ \Rightarrow ”: This is trivial.

“ \Leftarrow ”: Let $(\omega_\lambda)_\lambda$ be a net in Ω_1 that converges to a point $\omega \in \Omega_1$, and assume to show a contradiction that $(\varphi(\omega_\lambda))_\lambda$ does not converge to $\varphi(\omega)$. Then there exists an open neighbourhood U of $\varphi(\omega)$ and a subnet $(\omega_\theta)_\theta$ of $(\omega_\lambda)_\lambda$ such that $\varphi(\omega_\theta) \notin U$ for each index θ . As Ω_2 is completely regular, there exists a function $f \in C_b(\Omega_2)$ which is 1 at $\varphi(\omega)$ and vanishes outside of U . Hence,

$$0 = f(\varphi(\omega_\theta)) = (f \circ \varphi)(\omega_\theta) \rightarrow (f \circ \varphi)(\omega) = f(\varphi(\omega)) = 1,$$

which is a contradiction. □

1.1.2 The Koopman Operator

In the following, we let Ω be a completely regular Hausdorff space. We denote by $C_b(\Omega)$ the space of continuous functions $f: \Omega \rightarrow \mathbb{K}$ into a space of scalars $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, which we will specify only when necessary. If endowed with the supremum norm

$$\|\cdot\|_\infty: C_b(\Omega) \rightarrow \mathbb{R}_+, \quad \|f\|_\infty := \sup_{\omega \in \Omega} |f(\omega)|$$

it is well-known that $C_b(\Omega)$ becomes a Banach space.

Now let Ω_1 and Ω_2 both be completely regular Hausdorff spaces. We study continuous mappings $\varphi: \Omega_1 \rightarrow \Omega_2$. Note that we do not require $\Omega_1 \neq \Omega_2$ to obtain a mapping that we can iterate. As most results in this chapter will not deal with the iterates of dynamical systems, but rather with translation between categories, we will benefit from this more general setting.

Definition 1.1.6. Let $\varphi: \Omega_1 \rightarrow \Omega_2$ be a continuous mapping between topological spaces Ω_1 and Ω_2 . The *Koopman operator* or *composition operator* $T_\varphi: C_b(\Omega_2) \rightarrow C_b(\Omega_1)$ is defined by

$$(T_\varphi f)(\omega) = f(\varphi(\omega)),$$

for all $f \in C_b(\Omega_2)$ and all $\omega \in \Omega_1$.

Remark 1.1.7 (contravariance and backward action). (i) The transition from a continuous mapping to the Koopman operator defines a functor

$$F: \mathbf{regTopHaus} \rightarrow \mathbf{Ban}_1, \quad \varphi \mapsto T_\varphi$$

between the category **regTopHaus** of completely regular Hausdorff spaces as object with the continuous mappings as morphisms to the category **Ban₁** of Banach spaces as objects with the contractive linear operators as morphisms. We will see in Remark 1.1.10 that we can restrict the image category even further.

The functor F is *contravariant*, i.e., for completely regular Hausdorff spaces Ω_1 and Ω_2 and a continuous mapping $\varphi: \Omega_1 \rightarrow \Omega_2$ the diagram

$$\begin{array}{ccc} \Omega_1 & \xrightarrow{\varphi} & \Omega_2 \\ F \downarrow & & \downarrow F \\ C_b(\Omega_1) & \xleftarrow{T_\varphi} & C_b(\Omega_2) \end{array}$$

commutes. Notice that the lower arrow is backwards.

(ii) The contravariance of F has another consequence: The Koopman operator answers questions backwards in time. Let us illustrate this: let $U \subseteq \Omega_2$ be open and consider the vanishing ideal $I_{\Omega_2 \setminus U} := \{f \in C_b(\Omega_2) : \forall \omega \in \Omega_2 \setminus U : f(\omega) = 0\}$. Applying T_φ to some $f \in I_{\Omega_2 \setminus U}$ and evaluating at $y \in \Omega_1$ yields

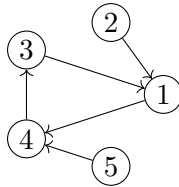
$$(T_\varphi f)(y) = f(\varphi(y)).$$

If $f(\varphi(y)) \neq 0$ this tells us that $\varphi(y) \in U$, or equivalently, that $y \in \varphi^{-1}(U)$. Conversely, if $f(\varphi(y)) = 0$ for all $f \in I_{\Omega_2 \setminus U}$, we may conclude that $\varphi(y) \notin U$, or equivalently, that $y \in \varphi^{-1}(\Omega_2 \setminus U)$. Hence, it tells us something about the action of φ backwards in “time”. (Of course the concept of time becomes more intuitive when $\Omega_1 = \Omega_2$.)

This becomes even clearer if we extend the domain of the Koopman operator to allow for indicator functions. If $M \subseteq \Omega_2$ is Borel measurable then $T_\varphi \mathbb{1}_M = \mathbb{1}_{\varphi^{-1}(M)}$. Therefore, the observable that we obtain describes the set M being moved backwards in time.

Let us consider an easy example of a Koopman operator of a finite graph, which we introduce for future use.

Example 1.1.8. Consider the following directed graph.



Note that each node has exactly one edge pointing outwards. Set $\Omega := \{1, \dots, 5\}$ and interpret the graph as the mapping

$$\varphi: \Omega \rightarrow \Omega, \quad \begin{cases} 1 \mapsto 4, & 2 \mapsto 1, \\ 3 \mapsto 1, & 4 \mapsto 3, \\ 5 \mapsto 4. \end{cases}$$

Then the space $C_b(\Omega)$ is isomorphic to \mathbb{K}^5 via the mapping $f \mapsto (f(\omega))_{\omega \in \Omega}$ and the Koopman operator T_φ can be written as a matrix \tilde{T}_φ given by

$$\tilde{T}_\varphi \begin{pmatrix} f(1) \\ \vdots \\ f(5) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} f(1) \\ f(2) \\ f(3) \\ f(4) \\ f(5) \end{pmatrix} = \begin{pmatrix} f(4) \\ f(1) \\ f(1) \\ f(3) \\ f(4) \end{pmatrix} = \begin{pmatrix} f(\varphi(1)) \\ f(\varphi(2)) \\ f(\varphi(3)) \\ f(\varphi(4)) \\ f(\varphi(5)) \end{pmatrix}$$

Remark 1.1.9. Let Ω be a completely regular Hausdorff space.

- (i) The space $C_b(\Omega)$ is a *Banach lattice* with respect to the *pointwise order*, where we say that two functions $f, g \in C_b(\Omega)$ satisfy $f \leq g$ if $f(\omega) \leq g(\omega)$ for all $\omega \in \Omega$. The modulus of a function $f \in C_b(\Omega)$ is given by composition with the complex absolute value function

$$|f|(\omega) = |f(\omega)|$$

for all $\omega \in \Omega$. As such, it is an *AM-space*, i.e., a space that satisfies

$$\|f \vee g\| = \|f\| \vee \|g\|$$

for all $0 \leq f, g \in C_b(\Omega)$.

- (ii) When $\mathbb{K} = \mathbb{C}$, the space $C_b(\Omega)$ is a commutative unital C^* -algebra with respect to the pointwise multiplication of functions defined by

$$(f \cdot g)(\omega) = f(\omega) \cdot g(\omega)$$

for all $f, g \in C_b(\Omega)$ and all $\omega \in \Omega$. The unit element of this algebra is the constant function $\mathbb{1} : \Omega \rightarrow \mathbb{C}$ defined by $\mathbb{1}(\omega) = 1$ for all $\omega \in \Omega$. The involution is given by the composition with the complex conjugation $\overline{f}(\omega) = \overline{f(\omega)}$ for all $\omega \in \Omega$.

Remark 1.1.10. Let Ω_1 and Ω_2 be completely regular Hausdorff spaces, $\varphi : \Omega_1 \rightarrow \Omega_2$ be continuous and $T_\varphi : C_b(\Omega_2) \rightarrow C_b(\Omega_1)$ the associated Koopman operator.

- (i) Recall from Remark 1.1.9 (i) that $C_b(\Omega_1)$ and $C_b(\Omega_2)$ are Banach lattices. The Koopman operator T_φ is a *lattice homomorphism*, i.e., it satisfies

$$|T_\varphi f| = T_\varphi |f|$$

for all $f \in C_b(\Omega_2)$.

- (ii) As mentioned in Remark 1.1.9 (ii), the spaces $C_b(\Omega_1)$ and $C_b(\Omega_2)$ are C^* -algebras. The Koopman operator is a **-homomorphism* on these algebras, i.e., it satisfies

$$T_\varphi(f \cdot g) = (T_\varphi f) \cdot (T_\varphi g) \quad \text{and} \quad T_\varphi \overline{f} = \overline{T_\varphi f}$$

for all $f, g \in C_b(\Omega_2)$.

As $T_\varphi \mathbb{1} = \mathbb{1}$, we say that T_φ is a *unital* lattice and **-homomorphism*.

This additional structure allows for the use of representation theorems to study the Koopman operator T_φ on $C_b(\Omega)$. In particular, the famous Gelfand–Naimark theorem [Ped89, Theorem 4.3.13] for commutative unital C^* -algebras and the Gelfand–Nadkarni [Mey91, Theorem 2.1.3] theorem for AM-spaces are applicable.

Proposition 1.1.11 (Stone–Čech compactification). *Let Ω be a completely regular Hausdorff space. Then*

- (i) *The space*

$$\beta\Omega := \{\alpha : C_b(\Omega) \rightarrow \mathbb{C} : \alpha \text{ is a unital algebra homomorphism}\}$$

is a compact Hausdorff space when endowed with the weak- topology.*

- (ii) *The mapping*

$$\Phi_\Omega : C_b(\Omega) \rightarrow C(\beta\Omega), \quad (\Phi_\Omega f)(\alpha) = \alpha(f)$$

is a unital algebra isomorphism.

- (iii) *The mapping*

$$\iota_\Omega : \Omega \rightarrow \beta\Omega, \quad \omega \mapsto \delta_\omega$$

is a homeomorphism onto its image $\iota_\Omega(\Omega)$, such that $\iota_\Omega(\Omega)$ is a dense subset of $\beta\Omega$ and the Koopman operator $T_{\iota_\Omega}: C(\beta\Omega) \rightarrow C_b(\Omega)$ is the inverse of the mapping Φ_Ω .

- (iv) If Ω_1 and Ω_2 are two completely regular Hausdorff spaces and $\varphi: \Omega_1 \rightarrow \Omega_2$ is a continuous mapping, then there exists a unique continuous mapping $\psi: \beta\Omega_1 \rightarrow \beta\Omega_2$ such that the Koopman operator T_φ is conjugate to the Koopman operator T_ψ , i.e., the diagram

$$\begin{array}{ccc} C_b(\Omega_2) & \xrightarrow{T_\varphi} & C_b(\Omega_1) \\ T_{\iota_{\Omega_2}} \uparrow & \downarrow \Phi_{\Omega_2} & T_{\iota_{\Omega_1}} \uparrow \downarrow \Phi_{\Omega_1} \\ C(\beta\Omega_2) & \xrightarrow{T_\psi} & C(\beta\Omega_1) \end{array}$$

is commutative.

Proof. “(i)”: This follows immediately from the fact that unital algebra homomorphisms are part of the unit ball of the dual of $C_b(\Omega)$ together with the theorem of Banach–Alaoglu.

“(ii)”: The fact that Ψ_Ω is a unital algebra isomorphism follows from the classical Gelfand–Naimark theorem.

“(iii)”: The continuity of ι_Ω follows readily from the continuity of the mappings

$$\Omega \ni \omega \mapsto \langle f, \iota_\Omega(\omega) \rangle = \langle f, \delta_\omega \rangle = f(\omega)$$

for each $f \in C_b(\Omega)$. As Ω is completely regular Hausdorff, the mapping ι_Ω is a homeomorphism onto its image $\iota_\Omega(\Omega)$. Indeed, the mapping ι_Ω is injective, as for $\omega_1, \omega_2 \in \Omega$ with $\omega_1 \neq \omega_2$ there exists a continuous function $f \in C_b(\Omega)$ such that $f(\omega_1) \neq f(\omega_2)$, and thus $\delta_{\omega_1} \neq \delta_{\omega_2}$. Moreover, the inverse mapping $\iota_\Omega^{-1}: \iota_\Omega(\Omega) \rightarrow \Omega$ is continuous. To see this, it suffices to show that for any net $(\omega_\lambda)_{\lambda \in \Lambda}$ in Ω the convergence $\delta_{\omega_\lambda} \rightarrow \delta_\omega$ for some $\omega \in \Omega$ implies that we have $\iota_\Omega^{-1}(\delta_{\omega_\lambda}) = \omega_\lambda \rightarrow \omega = \iota_\Omega^{-1}(\delta_\omega)$. Let $U \subseteq \Omega$ be an open neighborhood of ω . Then by complete regularity there exists a continuous function $f \in C_b(\Omega)$ such that $f(\omega) = 1$ and $f = 0$ on $\Omega \setminus U$. As $\delta_{\omega_\lambda} \rightarrow \delta_\omega$ in the weak-* topology, we have $\langle f, \delta_{\omega_\lambda} \rangle = f(\omega_\lambda) \rightarrow f(\omega) = 1$. Thus, for sufficiently large λ we have $f(\omega_\lambda) > 0$, which implies that $\omega_\lambda \in U$. Hence, $\omega_\lambda \rightarrow \omega$ and ι_Ω^{-1} is continuous. This shows that ι_X is a homeomorphism onto its image $\iota_\Omega(\Omega)$.

The Koopman operator $T_{\iota_\Omega}: C(\beta\Omega) \rightarrow C_b(\Omega)$ is the inverse of the algebra isomorphism Φ_Ω , as for all $f \in C_b(\Omega)$ and all $\omega \in \Omega$ we have

$$T_{\iota_\Omega}(\Phi_\Omega(f))(\omega) = \Phi_\Omega(f)(\delta_\omega) = \delta_\omega(f) = f(\omega).$$

Now, the fact that $\iota_\Omega(\Omega)$ is dense will follow from Theorem 1.3.5.

“(iv)”: This follows from the fact that the Koopman operator $T_\varphi: C_b(\Omega_2) \rightarrow C_b(\Omega_1)$ is a unital algebra homomorphism. Thus, the mapping

$$\Phi_{\Omega_1} \circ T_\varphi \circ \Phi_{\Omega_2}^{-1}: C(\beta\Omega_2) \rightarrow C(\beta\Omega_1)$$

is a unital algebra homomorphism. Now by [Sch74, Theorem III.9.1], there exists a continuous mapping $\psi: \beta\Omega_1 \rightarrow \beta\Omega_2$ such that

$$\Phi_{\Omega_1} \circ T_\varphi \circ \Phi_{\Omega_2}^{-1} = T_\psi. \quad \square$$

Remark 1.1.12 (lattice homomorphisms). Note that in the definition of $\beta\Omega$ the condition that α is a unital algebra homomorphism can be replaced with “ α is a unital lattice homomorphism”. Moreover, the mapping Ψ_Ω is even a lattice isomorphism.

Indeed, the Koopman operator T_{ι_Ω} clearly is a unital lattice homomorphism. Thus its inverse Φ_Ω is also a unital lattice isomorphism. In particular, this shows that each $\alpha \in \beta\Omega$ satisfies $\alpha(|f|) = (\Phi_\Omega|f|)(\alpha) = |\Phi_\Omega(f)|(\alpha) = |\Phi_\Omega(f)(\alpha)| = |\alpha(f)|$ for all $f \in C_b(\Omega)$, which implies that α is a unital lattice homomorphism. Thus, each element in $\beta\Omega$ is also a unital lattice homomorphism.

The proof of Proposition 1.1.11 (iv) heavily relies on the representation theorem [Sch74, Theorem III.9.1] for unital algebra or lattice homomorphisms on $C(K)$ for a compact Hausdorff space K . This theorem does not generalize to homomorphisms on $C_b(\Omega)$ for non-compact spaces Ω , as can be concluded from Example 1.3.13.

One property that is, in a sense made precise in Proposition 1.2.14, dual to being a lattice homomorphism is that of being interval preserving. The following example shows that a Koopman operator need not be interval preserving in general. We recall that the *order interval* between two functions $f_1, f_2 \in C_b(\Omega; \mathbb{R})$ is given by

$$[f_1, f_2] := \{f \in C_b(\Omega; \mathbb{R}) : f_1 \leq f \leq f_2\}.$$

Example 1.1.13. On the discrete space $\Omega = \{1, 2\}$, consider the mapping φ given by

$$\varphi(1) = \varphi(2) = 1.$$

Then the Koopman operator on $C_b(\Omega) \simeq \mathbb{C}^2$ is given by the matrix

$$T_\varphi = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

So we can see that T_φ is not *interval preserving*, i.e., it does not satisfy

$$T_\varphi[f_1, f_2] = [T_\varphi f_1, T_\varphi f_2]$$

for each $f_1, f_2 \in C_b(\Omega)$, since for $f_1 = (0, 0)^\top$ and $f_2 = (1, 0)^\top$ we have

$$T_\varphi[f_1, f_2] = T_\varphi \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] \neq \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right] = [T_\varphi f_1, T_\varphi f_2].$$

1.1.3 The Perron–Frobenius Operator

Another way to look at dynamical systems is to study how they transport ensembles or distributions of states from one time-step to the next. To model distributions we use the versatile concept of measures. More precisely, it is the Radon measures that are of interest to us.

Definition 1.1.14 (Radon measures). Let Ω be a completely regular Hausdorff space and denote by $\mathcal{B}(\Omega)$ its Borel- σ -algebra.

- (i) A mapping $\mu: \mathcal{B}(\Omega) \rightarrow \mathbb{K}$ is called a \mathbb{K} -valued *measure* on Ω if $\mu(\emptyset) = 0$ and for each sequence $(A_n)_{n \in \mathbb{N}}$ of pairwise disjoint sets in $\mathcal{B}(\Omega)$ we have

$$\mu \left(\bigcup_{n \in \mathbb{N}} A_n \right) = \sum_{n \in \mathbb{N}} \mu(A_n),$$

where the series converges unconditionally in \mathbb{K} .

- (ii) If μ is a measure, then the mapping $|\mu|: \mathcal{B}(\Omega) \rightarrow [0, \infty)$ defined by

$$|\mu|(A) := \sup \left\{ \sum_{k=1}^n |\mu(A_k)| : A_k \subseteq A \text{ measurable, pairwise disjoint, } n \in \mathbb{N} \right\}$$

is called the *total variation measure* of μ .

- (iii) A measure μ is called a *Radon measure* on Ω if it is *tight*, i.e., for each $\varepsilon > 0$ there exists a compact set $K \subseteq \Omega$ such that

$$|\mu|(\Omega \setminus K) < \varepsilon.$$

We denote the space of all complex-valued Radon measures by $M(\Omega)$ and write $M(\Omega; \mathbb{R})$ for the space of all real-valued Radon measures on Ω .

A measure μ is called *positive* if it maps into $[0, \infty)$. It follows from [Rud87, Theorem 6.2] that the total variation measure $|\mu|$ is always a positive measure. Moreover, it is easily checked that a \mathbb{K} -valued measure μ is positive if and only if $\mu = |\mu|$. We denote this by writing $\mu \geq 0$.

The literature also contains other definitions of Radon measures: Radon measures are often defined to map into the interval $[0, \infty]$ including the value ∞ . In this case, one usually requires the additional assumptions that the measure is *locally finite*, i.e., each point $\omega \in \Omega$ has an open neighborhood U with $\mu(U) < \infty$, and that μ is *outer regular*, i.e., for each measurable set $A \in \mathcal{B}(\Omega)$ the value $\mu(A)$ is the infimum of all $\mu(U)$ taken over all open neighborhoods of A . For measures mapping into $[0, \infty)$ both additional assumptions are automatic [AB06, Theorem 12.4].

We also note that Definition 1.1.14 extends the definition of a Radon measure to \mathbb{K} -valued measures by passing to the total variation measure. This is done in order to obtain a vector space. Observe that the space of all Radon measures $M(\Omega)$ forms a vector space when scalar multiplication and addition is defined pointwise. This space becomes a Banach space, when endowed with the right norm. The following is [AB06, Theorem 12.12].

Proposition 1.1.15 (total variation norm). *Let Ω be a completely regular Hausdorff space. Then the mapping*

$$\|\cdot\|_1: M(\Omega) \rightarrow [0, \infty), \quad \|\mu\|_1 := |\mu|(\Omega)$$

defines a norm under which $M(\Omega)$ is complete.

Moreover, the space of Radon measures $M(\Omega)$ forms a Banach lattice.

We use the symbol $\|\cdot\|_1$ here, as the properties of the total variation norm are quite reminiscent of the properties of the norm on an L^1 -space. One property that stands out is that of additivity on the cone of positive measures. More precisely, for any two measures $0 \leq \mu, \nu \in M(\Omega)$ we have

$$\|\mu + \nu\|_1 = \|\mu\|_1 + \|\nu\|_1.$$

Banach lattices with this property are usually referred to as *AL-spaces* [AB06, Corollary 12.13].

In addition to the Koopman operator, there is a second operator that captures the properties of a continuous mapping $\varphi: \Omega_1 \rightarrow \Omega_2$ between two completely regular Hausdorff spaces X and Y . It is called the *Perron–Frobenius operator*.

Definition 1.1.16 (Perron–Frobenius operator). Let Ω_1 and Ω_2 be completely regular Hausdorff spaces and $\varphi: \Omega_1 \rightarrow \Omega_2$ be a continuous mapping. The *Perron–Frobenius operator* $P_\varphi: M(\Omega_1) \rightarrow M(\Omega_2)$ is defined by

$$(P_\varphi\mu)(A) = \mu(\varphi^{-1}(A))$$

for all $\mu \in M(\Omega_1)$ and all Borel sets $A \subseteq \Omega_2$.

Note that the Perron–Frobenius operator is well-defined.

Remark 1.1.17 (well-definition). The complex-valued Radon measures are closed under taking pushforward measures with respect to continuous functions.

Indeed, if μ is a Radon measure on Ω_1 and $\varphi: \Omega_1 \rightarrow \Omega_2$ is continuous, then for each $\varepsilon > 0$ there exists a compact set $K \subseteq \Omega_1$ such that

$$|\mu|(\Omega_1 \setminus K) < \varepsilon.$$

As φ is continuous, the set $\varphi(K) \subseteq \Omega_2$ is compact as well. Moreover, we have

$$|P_\varphi\mu|(\Omega_2 \setminus \varphi(K)) \leq |\mu|(\varphi^{-1}(\Omega_2 \setminus \varphi(K))) \leq |\mu|(\Omega_1 \setminus K) < \varepsilon,$$

which shows that $P_\varphi\mu$ is also a Radon measure on Ω_2 .

Akin to the situation for the Koopman operator, the passing to the Perron–Frobenius operator has a categorical interpretation as a functor.

Remark 1.1.18 (covariance and forward action). (i) The transition from a continuous mapping to the Koopman operator defines a functor

$$G: \mathbf{regTopHaus} \rightarrow \mathbf{Ban}_1, \quad \varphi \mapsto P_\varphi$$

between the category **regTopHaus** of completely regular Hausdorff spaces as objects with the continuous mappings as morphisms to the category **Ban₁** of Banach spaces as objects with the contractive linear operators as morphisms.

The functor F is *covariant*, i.e., for completely regular Hausdorff spaces Ω_1 and Ω_2 and a continuous mapping $\varphi: \Omega_1 \rightarrow \Omega_2$ the diagram

$$\begin{array}{ccc} \Omega_1 & \xrightarrow{\varphi} & \Omega_2 \\ G \downarrow & & \downarrow G \\ C_b(\Omega_1) & \xrightarrow{P_\varphi} & C_b(\Omega_2) \end{array}$$

commutes.

- (ii) The covariance of G has the following consequence for the interpretation of the Perron–Frobenius operator: if we consider a measure $0 \leq \mu \in M(\Omega_1)$ as a mass distribution, then $P_\varphi \mu$ describes the result after this mass distribution has been moved forward by φ . This becomes particularly evident if we consider a point mass δ_ω located at a point $\omega \in \Omega_1$: for such a measure, we have

$$P_\varphi \delta_\omega = \delta_{\varphi(\omega)}.$$

We also point out that, as a consequence of this equality, the mapping φ is uniquely determined by the action of P_φ .

Example 1.1.19. Let us recall Example 1.1.8. We note that $M(\Omega)$ is isometrically isomorphic to \mathbb{K}^5 endowed with the 1-norm via the mapping $\mu \mapsto (\mu(\{\omega\}))_{\omega \in \Omega}$ and the Perron–Frobenius operator P_φ can be written as a matrix \tilde{P}_φ given by

$$\begin{aligned} \tilde{P}_\varphi \begin{pmatrix} \mu(\{1\}) \\ \vdots \\ \mu(\{5\}) \end{pmatrix} &= \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu(\{1\}) \\ \mu(\{2\}) \\ \mu(\{3\}) \\ \mu(\{4\}) \\ \mu(\{5\}) \end{pmatrix} = \begin{pmatrix} \mu(\{2\}) + \mu(\{3\}) \\ 0 \\ \mu(\{4\}) \\ \mu(\{1\}) + \mu(\{5\}) \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \mu(\{2, 3\}) \\ \mu(\emptyset) \\ \mu(\{4\}) \\ \mu(\{1, 5\}) \\ \mu(\emptyset) \end{pmatrix} = \begin{pmatrix} \mu(\varphi^{-1}(\{1\})) \\ \mu(\varphi^{-1}(\{2\})) \\ \mu(\varphi^{-1}(\{3\})) \\ \mu(\varphi^{-1}(\{4\})) \\ \mu(\varphi^{-1}(\{5\})) \end{pmatrix}. \end{aligned}$$

Note also the relationship $\tilde{P}_\varphi = \tilde{T}_\varphi^*$ between the Perron–Frobenius operator and the Koopman operator from Example 1.1.8. We will elaborate on this in Example 1.2.24 later.

The operator P_φ leaves the positive cone in $M(\Omega_1)$ invariant, i.e., P_φ is a *positive operator* between the Banach lattices $M(\Omega_1)$ and $M(\Omega_2)$.

Let us note next that the Perron–Frobenius operator is, in contrast to the Koopman operator, in general not a lattice homomorphism. We describe under which conditions P_φ is a lattice homomorphism in Theorem 1.3.3.

Example 1.1.20. Continuing Example 1.1.13, we see that the corresponding Perron–Frobenius operator on $M(\Omega) \simeq \mathbb{C}^2$ is given by the matrix

$$P_\varphi = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

So we can see that P_φ is not a lattice homomorphism since, for instance, the vector $\mu = (1, -1)^\top$ satisfies $|P_\varphi\mu| = 0 \neq P_\varphi|\mu|$.

We will show later in Theorem 1.3.3 that in certain situations the Perron–Frobenius operator happens to be a lattice homomorphism.

1.2 Duality of the Perron–Frobenius and Koopman Operators

In this section we describe the dual relationship between the Perron–Frobenius and the Koopman operator. To this end, we first need to introduce the concept of norming dual pairs and the associated weak topologies. This is due to the fact that on non-compact topological spaces Ω the space of Radon measures $M(\Omega)$ is in general not the norm dual of the space of continuous bounded functions $C_b(\Omega)$. Thus, we cannot expect a duality between the Perron–Frobenius on $M(\Omega)$ and the Koopman operator on $C_b(\Omega)$ in the classical sense of Banach space theory.

In Section 1.2.1 we first introduce the concept of norming dual pairs and the associated families of consistent weak topologies. We then describe in Section 1.2.2 how the Perron–Frobenius and the Koopman operator are dual to each other with respect to the norming dual pair $(C_b(\Omega), M(\Omega))$. Lastly in Section 1.2.3, we study the closed ideals of $C_b(\Omega)$ with respect to the weak topology induced by $M(\Omega)$ and their role in characterizing the Koopman operators as unital algebra and lattice homomorphisms.

1.2.1 Norming Dual Pairs and Weak Topologies

Let us first describe the concept of norming dual pairs as a special subclass of dual systems as presented in [SW99, Chapter IV] or [AB06, Section 5.15].

Definition 1.2.1 (norming dual pairs). Let X and Y be Banach spaces over the field \mathbb{K} and let

$$\langle \cdot, \cdot \rangle: X \times Y \rightarrow \mathbb{K}$$

be a \mathbb{K} -bilinear mapping. Then the triple $(X, Y, \langle \cdot, \cdot \rangle)$ is called *norming dual pair between X and Y* , if

$$\|x\| = \sup\{|\langle x, y \rangle| : \|y\| \leq 1\} \quad \text{and} \quad \|y\| = \sup\{|\langle x, y \rangle| : \|x\| \leq 1\}.$$

We often omit mentioning the bilinear mapping $\langle \cdot, \cdot \rangle$, if it is clear from the context, and merely say that (X, Y) is a norming dual pair. In case we want to eliminate ambiguity, we sometimes write $\langle \cdot, \cdot \rangle_{X, Y}$ instead of $\langle \cdot, \cdot \rangle$.

Due to the symmetry of the definition, (X, Y) is a norming dual pair if and only if (Y, X) is with the reverse duality pair defined by $\langle y, x \rangle_{Y, X} := \langle x, y \rangle_{X, Y}$ for each $x \in X$ and $y \in Y$. If not strictly necessary, we will not differentiate the duality pairs $\langle \cdot, \cdot \rangle_{X, Y}$ and $\langle \cdot, \cdot \rangle_{Y, X}$ explicitly.

Note that, if (X, Y) is a norming dual pair, then Y can be identified with a closed subspace of the norm dual X^* of X via the continuous embedding

$$Y \rightarrow X^*, \quad y \mapsto \langle \cdot, y \rangle.$$

Let us look at some norming dual pairs.

Example 1.2.2. (i) Let X be a Banach space and $Y = X^*$ denote its dual space. Then (X, X^*) together with its natural duality pairing

$$\langle x, x^* \rangle = x^*(x)$$

for all $x \in X$ and $x^* \in X^*$ forms a norming dual pair.

(ii) Let Ω be a completely regular Hausdorff space. Endowed with the bilinear mapping

$$\langle f, \mu \rangle := \int_{\Omega} f \, d\mu$$

for all $f \in C_b(\Omega)$ and $\mu \in M(\Omega)$ the pair $(C_b(\Omega), M(\Omega))$ is a norming dual pair. We refer to [Kun11, Example 2.4] for a proof.

We will see other norming dual pairs throughout this thesis, and introduce them as we need them.

If (X, Y) is a norming dual pair, then one can consider several natural locally convex topologies that can be defined in terms of the duality (see [SW99, Chapter IV] or [AB06, Section 5.15]).

Definition 1.2.3 (consistent topologies). Let (X, Y) be a norming dual pair. A locally convex topology τ on X is called *consistent* with the norming dual pair (X, Y) if the *continuous dual space* $(X, \tau)'$, i.e., the space of all τ -continuous linear functionals $\varphi: X \rightarrow \mathbb{K}$, is precisely the set of functionals of the form $\langle \cdot, y \rangle$ for some $y \in Y$.

The weak and the Mackey topology are standard examples of consistent topologies. Let us introduce them in the following. Recall that a subset $K \subseteq Y$ is called *circled*, if $\alpha \cdot K \subseteq K$ holds for all $\alpha \in \mathbb{K}$ with $|\alpha| \leq 1$.

Example 1.2.4. Let (X, Y) be a norming dual pair.

(i) The *weak topology* $\sigma(X, Y)$ is the coarsest locally convex topology on X that is consistent with (X, Y) . It is the initial topology of the family of seminorms given by $(p_y)_{y \in Y}$, where

$$p_y: X \rightarrow [0, \infty), \quad x \mapsto |\langle x, y \rangle|.$$

(ii) The *Mackey topology* $\mu(X, Y)$ is the initial topology on X of the family of seminorms $(p_K)_K$, where

$$p_K: X \rightarrow \mathbb{K}, \quad x \mapsto \sup_{y \in K} |\langle x, y \rangle|$$

and the family is indexed over all convex, circled, and $\sigma(Y, X)$ -compact subsets $K \subseteq Y$. It is a consequence of the Mackey–Arens theorem below that it is the finest locally convex topology on X that is consistent with (X, Y) .

The following theorem provides a characterization for all topologies consistent with a given norming dual pair [AB06, Theorem 5.112 and 5.113].

Theorem 1.2.5 (Mackey–Arens). *Let (X, Y) be a norming dual pair. A locally convex topology τ on X is consistent with (X, Y) if and only if there exists a family \mathfrak{G} of convex, circled and $\sigma(Y, X)$ -compact sets with $\bigcup_{K \in \mathfrak{G}} K = X$ such that τ is the initial topology of the family of seminorms $(p_K)_{K \in \mathfrak{G}}$ defined by*

$$p_K : X \rightarrow \mathbb{K}, \quad x \mapsto \sup_{y \in K} |\langle x, y \rangle|.$$

In particular, a locally convex topology τ on X is consistent with (X, Y) if and only if $\sigma(X, Y) \subseteq \tau \subseteq \mu(X, Y)$.

Notice that for the collection $\mathfrak{G} := \{\{\lambda \cdot y : |\lambda| \leq 1\} : y \in Y\}$, the family in the Mackey–Arens theorem generates the weak topology $\sigma(X, Y)$.

Let us illustrate the weak and Mackey topologies on an example.

Example 1.2.6. Let Ω be a compact Hausdorff space. Then the Mackey topology of the norming dual pair $(C(\Omega), M(\Omega))$ on $C(\Omega)$ coincides with the norm topology.

Indeed, for any $K \in \mathfrak{G}$ the seminorm p_K is trivially norm continuous, which shows that the Mackey topology $\mu(C(\Omega), M(\Omega))$ is coarser than the norm topology. Conversely, note that by Riesz’s representation theorem $M(\Omega)$ is isometrically isomorphic to the norm dual space of $C(\Omega)$. Thus, the closed unit ball $B_{\leq 1}(0) \subseteq M(\Omega)$ is an element in \mathfrak{G} , by the Banach–Alaoglu theorem. It follows that the seminorm $p_{B_{\leq 1}(0)}$ coincides with the norm $\|\cdot\|_\infty$, and thus, the Mackey topology is finer than the norm topology. Hence, both topologies coincide, as claimed.

Although different topologies that are consistent with (X, Y) obviously have different closed sets, it follows from an application of the Hahn–Banach separation theorem that this is not the case, if we additionally require convexity [AB06, Theorem 5.98].

Proposition 1.2.7 (closed convex sets). *Let (X, Y) be a norming dual pair. All topologies τ consistent with (X, Y) have the same collection of closed convex sets.*

Next let us take a look at operators and their adjoint from one norming dual pair into another. For this let (X_1, Y_1) and (X_2, Y_2) both be norming dual pairs and let $T : X_1 \rightarrow X_2$ be a linear mapping that is continuous from $\sigma(X_1, Y_1)$ to $\sigma(X_2, Y_2)$. Then for each $y \in Y_2$ the functional $X_1 \rightarrow \mathbb{K}$ defined by $x \mapsto \langle Tx, y \rangle$ is $\sigma(X_1, Y_1)$ -continuous. Hence, there exists a unique $y' \in Y_1$ such that $\langle Tx, y \rangle = \langle x, y' \rangle$. We call the mapping $T' : Y_2 \rightarrow Y_1$ that maps y to y' the *adjoint of T* . It is uniquely determined by the relation

$$\langle Tx, y \rangle = \langle x, T'y \rangle$$

for all $x \in X_1$ and $y \in Y_2$. From this duality, we can easily observe that T' is $\sigma(Y_2, X_2)$ - $\sigma(Y_1, X_1)$ -continuous.

Proposition 1.2.8. *Let (X_1, Y_1) and (X_2, Y_2) be norming dual pairs and $T: X_1 \rightarrow X_2$ a linear operator.*

- (i) *For $i \in \{1, 2\}$ let β_i be a locally convex topology consistent with (X_i, Y_i) . If T is β_1 - β_2 -continuous, then it is $\sigma(X_1, Y_1)$ - $\sigma(X_2, Y_2)$ -continuous.*
- (ii) *The operator T is $\sigma(X_1, Y_1)$ - $\sigma(X_2, Y_2)$ -continuous if and only if it is $\mu(X_1, Y_1)$ - $\mu(X_2, Y_2)$ -continuous.*

We could not find a reference for the elementary statement of Proposition 1.2.8 (i), so we provide a short proof. Proposition 1.2.8 (ii) can be found as a part of [SW99, Result IV.7.4].

Proof of Proposition 1.2.8. “(i)”: For each consistent topology β_i , denote by \mathfrak{G}_i the set of convex, circled and $\sigma(Y_i, X_i)$ -compact sets from the Mackey–Arens Theorem 1.2.5. Let $y_2 \in Y_2$. Then there exists $K_2 \in \mathfrak{G}_2$ such that $y_2 \in K_2$. Thus, the functional

$$\langle T(\cdot), y_2 \rangle: X_1 \rightarrow \mathbb{K}, \quad x_1 \mapsto \langle Tx_1, y_2 \rangle$$

is β_1 -continuous. Indeed, since $|\langle x_2, y_2 \rangle| \leq p_{K_2}(x_2)$ holds for all $x_2 \in X_2$, it follows that $\langle \cdot, y_2 \rangle$ is β_2 -continuous. Then, by composition, the β_1 - β_2 -continuity implies the β_1 -continuity of $\langle T(\cdot), y_2 \rangle$.

Since β_1 is consistent with (X_1, Y_1) , the functional $\langle T(\cdot), y_2 \rangle$ is $\sigma(X_1, Y_1)$ -continuous, which implies that T is $\sigma(X_1, Y_1)$ - $\sigma(X_2, Y_2)$ -continuous.

“(ii)”:

 This is the equivalence of (b₁) and (b₂) in [SW99, Result IV.7.4]. □

With more additional structure, the concept of norming dual pairs becomes more powerful. In the following, we assume the Banach spaces X and Y to be Banach lattices E and F . We additionally require that the duality pair respects the ordering on the Banach lattices.

Definition 1.2.9 (norming dual pairs of Banach lattices). Let E and F be Banach lattices over \mathbb{K} and let

$$\langle \cdot, \cdot \rangle: E \times F \rightarrow \mathbb{K}$$

be a \mathbb{K} -bilinear mapping. Then the triple $(E, F, \langle \cdot, \cdot \rangle)$ is called a *norming dual pair of Banach lattices* if it is a norming dual pair that satisfies

$$\begin{aligned} e \in E_+ &\Leftrightarrow \forall f \in F_+ : \langle e, f \rangle \geq 0 && \text{and} \\ f \in F_+ &\Leftrightarrow \forall e \in E_+ : \langle e, f \rangle \geq 0. \end{aligned}$$

Example 1.2.10. It is simple to verify that the norming dual pair $(C_b(\Omega), M(\Omega))$ from Example 1.2.2 (ii) is in fact a norming dual pair of Banach lattices.

A subset A of a Banach lattice is called *downward directed* if for each $e_1, e_2 \in A$ there exists $e \in A$ such that $e \leq e_1 \wedge e_2$. The norm on a Banach lattice E is said to be *order continuous* if every downward directed set A with $\inf(A) = 0$ satisfies $\inf\{\|e\| : e \in A\} = 0$.

There are many equivalent characterizations of Banach spaces with order continuous norm. Let us name a few that we will need later [Mey91, Theorem 2.4.2 and Corollary 2.4.4]. A *band projection* on a Banach lattice E is a positive projection $P: E \rightarrow E$ for which $I - P$ is also positive.

Proposition 1.2.11. *Let E be a Banach lattice. Then the following are equivalent.*

- (a) *The norm on E is order continuous.*
- (b) *Every monotone order bounded sequence of E is convergent.*
- (c) *Every norm closed ideal of E is the range of a positive projection.*
- (d) *Every norm closed ideal of E is the range of a band projection.*
- (e) *Every order interval*

$$[e_1, e_2] := \{e \in E : e_1 \leq e \leq e_2\}$$

is $\sigma(E, E^)$ -compact.*

Let us see some examples of Banach lattices that have order continuous norm.

Example 1.2.12. (i) If $p \in [1, \infty)$ and $X = (X, \Sigma, \mu)$ is a measure space, then the space $L^p(X)$ has order continuous norm.

(ii) The Banach lattice $M(\Omega)$ of Radon measures on a completely regular Hausdorff space Ω has order continuous norm.

(iii) Let Ω be a completely regular space with infinite cardinality. Then the Banach lattice $C_b(\Omega)$ does not have order continuous norm.

Proof. “(i)”: We prove a stronger statement and show that every norm bounded and monotone sequence converges. As every order bounded sequence is norm bounded, it follows from Proposition 1.2.11 (b) that the $\|\cdot\|_p$ -norm is order continuous.

In the case that $p = 1$, let $(g_n)_{n \in \mathbb{N}}$ be a norm bounded monotone decreasing sequence. By passing to $(g_n - g_1)_{n \in \mathbb{N}}$, we may assume that $g_1 \in L^1(X)_+$. Consider the sequence of differences $(f_n)_{n \in \mathbb{N}}$ with $f_n := g_n - g_{n+1}$, which lies in $L^1(X)_+$, by monotonicity. Hence, by induction and the additivity of the norm on the positive cone of $L^1(X)$ we obtain

$$\|g_1 - g_n\| = \left\| \sum_{k=1}^{n-1} f_k \right\| = \sum_{k=1}^{n-1} \|f_k\|$$

for all $n \in \mathbb{N}$. Since the sequence $(\|g_1 - g_n\|)_{n \in \mathbb{N}}$ is increasing and bounded, it follows that the series $\sum_{k \in \mathbb{N}} f_k$ converges absolutely in $L^1(X)$. It follows that $g_1 - \sum_{k \in \mathbb{N}} f_k$ is the limit of $(g_n)_{n \in \mathbb{N}}$.

In the case that $p \in (1, \infty)$, we note that $L^p(X)$ is reflexive. Thus, [Mey91, Theorem 2.4.15] shows that every monotone and norm bounded sequence in $L^p(X)$ converges.

“(ii)”: This is proved in analogy to the case of L^1 -spaces in (i), as the norm on $M(\Omega)$ is also additive on the positive cone.

“(iii)”: By Proposition 1.1.11 there exists a compact and non-finite Hausdorff space $K := \beta\Omega$ such that $C_b(\Omega)$ is isometrically lattice isomorphic to $C(K)$. Thus, we only need to show that $C(K)$ does not have order continuous norm. Take a non-isolated point $\omega \in K$, which always exists by the compactness and the non-finiteness of K . Let \mathcal{U} be a neighborhood filter of ω . For each $U \in \mathcal{U}$ let $f_U: K \rightarrow [0, 1]$ be continuous such that $f_U(\omega) = 1$ and f_U vanishes outside of U . Then consider the directed hull

$$A := \{f_{U_1} \wedge \cdots \wedge f_{U_n} : U_1, \dots, U_n \in \mathcal{U}, n \in \mathbb{N}\}$$

of these functions and notice that A is a directed set. As ω is not isolated, it follows that $\inf(A) = 0$ but $\inf\{\|e\| : e \in A\} = 1$. \square

Any complex Banach lattice E can be decomposed into its real and imaginary part $E_{\mathbb{R}}$ and $iE_{\mathbb{R}}$ by a direct sum of vector spaces

$$E = E_{\mathbb{R}} \oplus iE_{\mathbb{R}}.$$

We denote by $E_{\mathbb{R}}$ the *real part* of the Banach lattice E . Note that $E_{\mathbb{R}}$ is a real Banach lattice with respect to the inherited order and norm from E . By convention we set $E_{\mathbb{R}} := E$, if E is already a real Banach lattice. In a complex Banach lattice E we tacitly assume that $f_1, f_2 \in E_{\mathbb{R}}$, whenever we write $f \leq g$ for $f, g \in E$. Similarly, we assume $f_1, f_2 \in E_{\mathbb{R}}$ whenever we write

$$[f_1, f_2] := \{f \in F : f_1 \leq f \leq f_2\}$$

for the *order interval* bounded by $f_1, f_2 \in E$. We call a linear operator $T: F_1 \rightarrow F_2$ between two Banach lattices *interval preserving* if

$$T[f_1, f_2] = [Tf_1, Tf_2]$$

for all $f_1, f_2 \in F_1$.

Let E be a Banach lattice with bounded linear operator $T: E \rightarrow E$ and let $T^*: E^* \rightarrow E^*$ be the norm adjoint of T on the norm dual E^* of E . There is a classical duality between the properties of T being a lattice homomorphisms and T^* being interval preserving [Mey91, Theorem 1.4.19 i)]. This is proved by an application of the Riesz–Kantorovich formula for functionals on vector lattices, which follows from [Mey91, Theorem 1.3.2]. It states that the lattice operations on E can be expressed in terms of suprema and infima of linear functionals on E^* , i.e., for all $e \in E_{\mathbb{R}}$ and all $e^* \in E_{\uparrow}^*$ we have

$$\begin{aligned} \langle e^+, e^* \rangle &= \sup\{\langle e, \tilde{e}^* \rangle : \tilde{e}^* \in [0, e^*]\} \quad \text{and} \\ \langle e^-, e^* \rangle &= -\inf\{\langle e, \tilde{e}^* \rangle : \tilde{e}^* \in [0, e^*]\}. \end{aligned}$$

In general not every norming dual pairs of Banach lattices will satisfy the Riesz–Kantorovich formula. Therefore, we will aid their investigation by introducing the following notion.

Definition 1.2.13 (Riesz–Kantorovich property). Let (E, F) be a norming dual pair of Banach lattices. We say that E has the *Riesz–Kantorovich property* with respect to F if for all $e \in E_{\mathbb{R}}$ and $f \in F_+$ we have

$$\begin{aligned}\langle e^+, f \rangle &= \sup\{\langle e, \tilde{f} \rangle : \tilde{f} \in [0, f]\} \quad \text{and} \\ \langle e^-, f \rangle &= -\inf\{\langle e, \tilde{f} \rangle : \tilde{f} \in [0, f]\},\end{aligned}$$

where the intervals $[0, f]$ are taken in F .

Note that if one wants to show the Riesz–Kantorovich property for a norming dual pair (E, F) , it suffices to show one of the above equalities, as the other follows by application to $-e$.

Proposition 1.2.14. *Let (E_1, F_1) and (E_2, F_2) be norming dual pairs of Banach lattices. Suppose E_2 has the Riesz–Kantorovich property with respect to F_2 and that F_2 has order continuous norm. Let $T: E_1 \rightarrow E_2$ be a linear $\sigma(E_1, F_1)$ - $\sigma(E_2, F_2)$ -continuous operator. Then T is a lattice homomorphism if and only if T' is interval preserving.*

We follow the proof of [Mey91, Theorem 1.4.19 i)] and adapt it to the case of the norming dual pairs of Banach lattices.

Proof. “ \Rightarrow ”: Suppose that $T: E_1 \rightarrow E_2$ is a lattice homomorphism. We first remark that it suffices to show that $T'[0, f] = [0, T'f]$ for all $0 \leq f \in F_1$, since by linearity we have $T'[f_1, f_2] = T'[0, f_2 - f_1] + T'f_1$ and $[T'f_1, T'f_2] = [0, T'(f_2 - f_1)] + T'f_1$.

“ \subseteq ”: This inclusion is obvious and holds for all positive operators.

“ \supseteq ”: Since F_1 has order continuous norm, it follows from Proposition 1.2.11 (e) the order interval $[0, f]$ is $\sigma(F_1, (F_1)^*)$ -compact. As the $\sigma(F_1, E_1)$ -topology is coarser than $\sigma(F_1, (F_1)^*)$, it follows that $[0, f]$ is $\sigma(F_1, E_1)$ -compact. Hence, by continuity it follows that $T'[0, f]$ is $\sigma(F_2, E_2)$ -compact.

Now let $0 \leq f \in F_1$ and let $\hat{f} \notin T'[0, f]$. We show that $\hat{f} \notin [0, T'f]$, proving the inclusion. We may assume that $\hat{f} \geq 0$, otherwise the proof is trivial. Thus, suppose to show a contradiction that $\hat{f} \leq T'f$. By the Hahn-Banach separation theorem for locally convex spaces [AB06, Theorem 5.79], there exists a $\sigma(F_2, E_2)$ -continuous functional $\varphi: F_2 \rightarrow \mathbb{K}$ that separates \hat{f} from the convex $\sigma(F_2, E_2)$ -compact set $T'[0, f]$. Thus, by $\sigma(F_1, E_2)$ -continuity, there exists $e \in E_2$ such that $\varphi = \langle e, \cdot \rangle$ and

$$\langle e, \hat{f} \rangle < a := \inf\{\langle e, \tilde{f} \rangle : \tilde{f} \in T'[0, f]\}.$$

Hence, by the lattice homomorphism property of T and the Riesz–Kantorovich property of E_2 , it follows that

$$\begin{aligned}\langle e^-, T'f \rangle &= \langle Te^-, f \rangle = \langle (Te)^-, f \rangle \\ &= -\inf\{\langle Te, \tilde{f} \rangle : \tilde{f} \in [0, f]\} \\ &= -\inf\{\langle e, \tilde{f} \rangle : \tilde{f} \in T'[0, f]\} \\ &= -a < \langle -e, \hat{f} \rangle \leq \langle e^-, \hat{f} \rangle.\end{aligned}$$

Thus, $\langle e^-, T'f - \hat{f} \rangle < 0$. Since $\langle e^-, \cdot \rangle$ is a positive continuous functional, this is a contradiction to $\hat{f} \leq T'f$.

“ \Leftarrow ”: We assume that T' is interval preserving and let $e \in (E_1)_+$ and $f \in F_2$. Then the Riesz–Kantorovich property of E_2 implies that

$$\begin{aligned} \langle |Te|, f \rangle &= \langle (Te)^+, f \rangle + \langle (Te)^-, f \rangle \\ &= \sup\{\langle Te, \tilde{f} \rangle : \tilde{f} \in [0, f]\} - \inf\{\langle Te, \tilde{f} \rangle : \tilde{f} \in [0, f]\} \\ &= \sup\{\langle e, T'\tilde{f} \rangle : \tilde{f} \in [0, f]\} - \inf\{\langle e, T'\tilde{f} \rangle : \tilde{f} \in [0, f]\} \\ &= \sup\{\langle e, f \rangle : \tilde{f} \in [0, T'f]\} - \inf\{\langle e, f \rangle : \tilde{f} \in [0, T'f]\} \\ &= \langle e^+, T'f \rangle + \langle e^-, T'f \rangle = \langle |e|, T'f \rangle = \langle T|e|, f \rangle \end{aligned}$$

This shows that T is a lattice homomorphism. \square

Let us now discuss some sufficient conditions for the Riesz–Kantorovich property to hold. For this we need additional nomenclature. We call a Banach lattice E σ -order complete if every countable bounded subset of E has a supremum and an infimum in E . Two elements $e_1, e_2 \in E$ are called *disjoint* if $|e_1| \wedge |e_2| = 0$. The *disjoint complement* of a subset $A \subseteq E$ is defined as

$$A^\perp := \{e \in E : e \text{ is disjoint to every } a \in A\}.$$

Let us now introduce the notion of bands and projection bands. An ideal I of a Banach lattice E is called a *band* if every subset $A \subseteq I$ that has a supremum in E satisfies $\sup(A) \in I$. For a subset $A \subseteq E$ we denote by $B(A)$ the smallest band containing A . In case $A = \{e\}$ we simply write $B_e := B(\{e\})$ and call it the *principal band of e* . It follows from [Mey91, Proposition 1.2.7] that $B(A) = A^{\perp\perp}$ for every subset $A \subseteq E$. Note also that the range of a band projection is a band [Mey91, Lemma 1.2.8 ii)]. Bands that are the range of a band projection are called *projection bands*.

We will need the following fact about principal bands in σ -order complete Banach lattices from [Mey91, Proposition 1.2.11 iii)].

Lemma 1.2.15. *Let E be a Banach lattice that is σ -order complete. Then for every $e \in E_+$ the principal band B_e is a projection band and the band projection $P_e: E \rightarrow E$ onto B_e is given by*

$$P_e f = \sup_{n \in \mathbb{N}} (f \wedge ne)$$

for all $f \in E_+$.

With this knowledge we may conclude the following.

Proposition 1.2.16. *Let (E, F) be a norming dual pair of Banach lattices. Suppose that*

- (i) E is σ -order complete and
- (ii) for all $e \in E_+$ the band projection onto B_e is $\sigma(E, F)$ -continuous.

Then E has the Riesz–Kantorovich property with respect to F .

Proof. Let $e \in E_{\mathbb{R}}$ and $f \in F_+$. We only show the first equality of the Riesz–Kantorovich property, as the second follows analogously.

“ \geq ”: To prove the first inequality, let $\tilde{f} \in [0, f]$ and notice that $\langle e^+, \cdot \rangle$ defines a positive functional on F . Thus,

$$\langle e, \tilde{f} \rangle = \langle e^+, \tilde{f} \rangle - \langle e^-, \tilde{f} \rangle \leq \langle e^+, f \rangle.$$

“ \leq ”: For the converse inequality, let $\tilde{f} \in [0, f]$. It follows from Lemma 1.2.15 that the principal band B_{e^+} is a projection band with band projection. Then it follows from the $\sigma(E, F)$ -continuity of the band projection P_{e^+} that the adjoint $P'_{e^+}: F \rightarrow F$ exists and is $\sigma(F, E)$ -continuous. Moreover, since $0 \leq P'_{e^+}f \leq f$ for all $f \in F_+$, it is a band projection. Hence, $P'_{e^+}f \in [0, f]$ and

$$\begin{aligned} \langle e, P'_{e^+}f \rangle &= \langle e^+, P'_{e^+}f \rangle - \langle e^-, P'_{e^+}f \rangle \\ &= \langle e^+, P'_{e^+}f \rangle + \underbrace{\langle P_{e^+}e^-, f \rangle}_{=0} \leq \langle e^+, f \rangle. \end{aligned}$$

This shows the claim. □

We want to apply the above result to the pair $(C_b(\Omega), M(\Omega))$ for a completely regular Hausdorff space Ω . The space of bounded continuous functions fails to be σ -order complete in general but is embedded into the σ -order complete the space of bounded Borel measurable functions. This motivates the following remark that follows.

Remark 1.2.17. Let (E_1, F) be a norming dual pair of Banach lattices and E_2 be a Banach lattice that embeds into E_1 such that (E_2, F) is also a norming dual pair of Banach lattices. If the E_1 has the Riesz–Kantorovich property with respect to F , then E_2 has the Riesz–Kantorovich property with respect to F .

Let Ω be completely regular Hausdorff. We denote by $B_b(\Omega)$ the space of Borel measurable and bounded functions $f: \Omega \rightarrow \mathbb{K}$. If endowed with the supremum norm $\|f\|_\infty := \sup_{\omega \in \Omega} |f(\omega)|$, the space of bounded measurable functions becomes a Banach space; and with pointwise defined lattice operations it becomes a Banach lattice.

Proposition 1.2.18. *Let Ω be a completely regular Hausdorff space. The tuple $(B_b(\Omega), M(\Omega))$ together with the bilinear mapping*

$$\langle f, \mu \rangle := \int_{\Omega} f \, d\mu$$

is a norming dual pair of Banach lattices.

We omit the measure theoretic details of the proof and focus on examples of norming dual pairs of Banach lattices and the Riesz–Kantorovich property.

Example 1.2.19. (i) Let E be a Banach lattice. Consider the norming dual pair of Banach lattices (E, E^*) , where E^* is the norm dual of E . Then E has the Riesz–Kantorovich property with respect to E^* .

(ii) Let Ω be a completely regular Hausdorff space. Then the Banach lattice $B_b(\Omega)$ has the Riesz–Kantorovich property with respect to $M(\Omega)$.

- (iii) Let Ω be a completely regular Hausdorff space. Then the Banach lattice $C_b(\Omega)$ has the Riesz–Kantorovich property with respect to $M(\Omega)$.

Proof. “(i)”: This follows directly from the classical Riesz–Kantorovich formula for Banach lattices, see [Mey91, Theorem 1.3.2].

“(ii)”: Clearly, the Banach lattice $B_b(\Omega)$ is σ -order complete. Moreover, let $f \in B_b(\Omega)_+$. Then by Lemma 1.2.15 the band projection $P_f: B_b(\Omega) \rightarrow B_b(\Omega)$ onto the principal band B_f is given by

$$P_f g = \sup_{n \in \mathbb{N}} (g \wedge n f)$$

for all $g \in B_b(\Omega)$. Notice that the lattice supremum here coincides with the pointwise supremum. Set $S := \{\omega \in \Omega : f(\omega) = 0\}$. We claim that $P_f g = g \cdot \mathbb{1}_{\Omega \setminus S}$ for all $g \in B_b(\Omega)$. Indeed, for every $\omega \in \Omega \setminus S$ there exists $n_0 \in \mathbb{N}$ such that $n_0 f(\omega) > g(\omega)$, and thus, we have $P_f g(\omega) \geq (g \wedge n_0 f)(\omega) = g(\omega)$. Conversely, for $n \geq n_0$ we have $(g \wedge n f)(\omega) = g(\omega)$, and thus, $P_f g(\omega) = g(\omega)$. For $\omega \in S$ we have $(g \wedge n f)(\omega) = 0$ for all $n \in \mathbb{N}$ and thus, $P_f g(\omega) = 0$. This shows the claim.

Consider the operator $Q_f: M(\Omega) \rightarrow M(\Omega)$ defined by

$$(Q_f \mu)(A) := \mu(A \setminus S)$$

for every measurable set $A \subseteq \Omega$. It is easily seen that Q_f is the adjoint operator of P_f , since

$$\langle P_f g, \mu \rangle = \int_{\Omega \setminus S} g \, d\mu = \int_{\Omega} g \, d(Q_f \mu) = \langle g, Q_f \mu \rangle$$

holds for every $g \in B_b(\Omega)$ and every $\mu \in M(\Omega)$. Thus, P_f is $\sigma(B_b(\Omega), M(\Omega))$ -continuous and the assertion follows from Proposition 1.2.16.

“(iii)”: This follows from (ii) and Remark 1.2.17, since $C_b(\Omega)$ embeds into $B_b(\Omega)$. \square

1.2.2 The Koopman and Perron–Frobenius Operators as Adjoints

Let us now apply the abstract results from the previous subsection to the setting of the Koopman and Perron–Frobenius operators. Recall that we have seen in Example 1.2.2 (ii) and Example 1.2.10 that for a completely regular Hausdorff space Ω the tuple $(C_b(\Omega), M(\Omega))$ is a norming dual pair of Banach lattices. It is now easily seen that the Koopman and Perron–Frobenius operators are mutually adjoint operators.

Proposition 1.2.20. *Let Ω_1 and Ω_2 be completely regular Hausdorff spaces and $\varphi: \Omega_1 \rightarrow \Omega_2$ be continuous. Denote by $T_\varphi: C_b(\Omega_2) \rightarrow C_b(\Omega_1)$ the Koopman operator and by $P_\varphi: M(\Omega_1) \rightarrow M(\Omega_2)$ the Perron–Frobenius operator of φ . Then it holds that*

$$\langle T_\varphi f, \mu \rangle_{C_b(\Omega_1), M(\Omega_1)} = \langle f, P_\varphi \mu \rangle_{C_b(\Omega_2), M(\Omega_2)}$$

for all $f \in C_b(\Omega_2)$ and all $\mu \in M(\Omega_1)$. In particular, T_φ and P_φ are their mutual adjoint operators.

Proof. This is the transformation theorem for pushforward measures. \square

Proposition 1.2.21. *Let Ω_1 and Ω_2 be completely regular Hausdorff spaces and $\varphi: \Omega_1 \rightarrow \Omega_2$ be continuous. Then*

- (i) *the Koopman operator $T_\varphi: C_b(\Omega_2) \rightarrow C_b(\Omega_1)$ is continuous with respect to the weak topologies $\sigma(C_b(\Omega_2), M(\Omega_2))$ and $\sigma(C_b(\Omega_1), M(\Omega_1))$, and*
- (ii) *the Perron–Frobenius operator $P_\varphi: M(\Omega_1) \rightarrow M(\Omega_2)$ is continuous with respect to the weak topologies $\sigma(M(\Omega_1), C_b(\Omega_1))$ and $\sigma(M(\Omega_2), C_b(\Omega_2))$.*

Proof. We only show the $\sigma(C_b(\Omega_2), M(\Omega_2))$ - $\sigma(C_b(\Omega_1), M(\Omega_1))$ -continuity of T_φ . The continuity of P_φ follows analogously. Let $\mu \in M(\Omega_2)$. Note that the functional

$$\langle T_\varphi(\cdot), \mu \rangle: C_b(\Omega_2) \rightarrow \mathbb{K}, \quad f \mapsto \langle T_\varphi f, \mu \rangle$$

is $\sigma(C_b(\Omega_2), M(\Omega_2))$ -continuous. Indeed, this follows from Proposition 1.2.20, since

$$\langle T_\varphi(\cdot), \mu \rangle = \langle \cdot, P_\varphi \mu \rangle$$

and the right-hand side is clearly $\sigma(C_b(\Omega_2), M(\Omega_2))$ -continuous. \square

It is now an easy corollary of Example 1.2.12 (ii), Proposition 1.2.14, Example 1.2.19 (iii), Proposition 1.2.20 and Proposition 1.2.21 that Perron–Frobenius operators are always interval preserving.

Corollary 1.2.22. *Let Ω_1 and Ω_2 be completely regular Hausdorff spaces and $\varphi: \Omega_1 \rightarrow \Omega_2$ continuous. The Perron–Frobenius operator P_φ is interval preserving, i.e., if $\mu, \nu \in M(\Omega_1)$ are real-valued measures, then*

$$P_\varphi[\mu, \nu] = [P_\varphi \mu, P_\varphi \nu].$$

We can use the duality for a further property of the Perron–Frobenius operator. It is, by duality, equivalent to the property that the Koopman operator maps the constant function $\mathbb{1}_{\Omega_2}$ to $\mathbb{1}_{\Omega_1}$.

Proposition 1.2.23. *Let Ω be a completely regular Hausdorff space. Then the Perron–Frobenius operator is norm-preserving on the positive cone of $M(\Omega_1)$, i.e., we have*

$$\|P_\varphi \mu\|_1 = \|\mu\|_1$$

for all $0 \leq \mu \in M(\Omega_1)$.

Proof. Let $0 \leq \mu \in M(\Omega_1)$, then $P_\varphi \mu \geq 0$, and thus, by Proposition 1.2.20

$$\|P_\varphi \mu\|_1 = (P_\varphi \mu)(\Omega_2) = \langle \mathbb{1}_{\Omega_2}, P_\varphi \mu \rangle = \langle T_\varphi \mathbb{1}_{\Omega_2}, \mu \rangle = \langle \mathbb{1}_{\Omega_1}, \mu \rangle = \mu(\Omega_1) = \|\mu\|_1.$$

This concludes the proof. \square

Example 1.2.24. We recall Examples 1.1.8 and 1.1.19. Then the duality of the Koopman operator T_φ and the Perron–Frobenius operator can be seen in their

matrix representation. For $f \in C_b(\Omega)$ and $\mu \in M(\Omega)$ write \tilde{f} for the column vector $(f(1), \dots, f(5))^\top$ and $\tilde{\mu}$ for the column vector $(\mu(\{1\}), \dots, \mu(\{5\}))^\top$. Note that we have

$$\langle f, \mu \rangle = \sum_{\omega \in \Omega} f(\omega) \mu(\{\omega\}) = \tilde{\mu}^* \tilde{f}$$

for each $f \in C_b(\Omega)$ and each $\mu \in M(\Omega)$. Then

$$\langle T_\varphi f, \mu \rangle = \tilde{\mu}^* \tilde{T}_\varphi \tilde{f} = (\tilde{T}_\varphi^* \tilde{\mu})^* \tilde{f} = (\tilde{P}_\varphi \tilde{\mu})^* \tilde{f} = \langle f, P_\varphi \mu \rangle.$$

1.2.3 Weakly Closed Ideals and Weakly Continuous Functionals

To obtain a characterization of the injectivity and surjectivity of the Koopman operator T , we require other topologies on the spaces $C_b(X)$ and $C_b(Y)$ than the norm topologies. In Section 1.2.1 we have already gotten to know the weak topology $\sigma(C_b(\Omega), M(\Omega))$. In order to prepare further results, we will take a closer look at notions of closed ideals in $C_b(X)$ endowed with topologies consistent with $(C_b(\Omega), M(\Omega))$.

If Ω is completely regular Hausdorff, then by Remark 1.1.3 (iii) every subset $M \subseteq \Omega$, if endowed with the subspace topology, is completely regular Hausdorff as well. We are now going to describe subsets M of Ω in terms of subspaces of $C_b(\Omega)$. To this end, we need some terminology from algebra and lattice theory first:

- Definition 1.2.25** (Ideals). (i) Let A be an algebra. A linear subspace I of A is called an (*algebra*) *ideal* if for all $f \in I$ and all $g \in A$ the product $f \cdot g$ and $g \cdot f$ are contained in I .
- (ii) Let E be a vector lattice. A linear subspace I of E is called a (*lattice*) *ideal* if $f \in I$, $g \in E$ and $|g| \leq f$ implies $g \in I$.

For $M \subseteq \Omega$ we define

$$I_M := \{f \in C_b(\Omega) : f(\omega) = 0 \text{ for all } \omega \in M\}.$$

The ideal I_M is called the *vanishing ideal* of M in $C_b(X)$. It is easily verified that I_M has indeed an algebra and a lattice ideal structure in $C_b(X)$.

Remark 1.2.26. Let τ be a locally convex topology on $C_b(\Omega)$ that is consistent with $(C_b(\Omega), M(\Omega))$. For each subset $M \subseteq \Omega$ the vanishing ideal I_M is τ -closed. Indeed, for $M \subseteq \Omega$ it holds that

$$I_M = \bigcap_{\omega \in M} \{f \in C_b(\Omega) : f(\omega) = 0\} = \bigcap_{\omega \in M} I_{\{\omega\}}.$$

So it suffices to show that $I_{\{\omega\}}$ is a τ -closed ideal in $C_b(\Omega)$ for all $\omega \in \Omega$. As τ is a finer topology than $\sigma(C_b(\Omega), M(\Omega))$ it suffices to show closedness in the latter topology.

Fix $\omega \in \Omega$ and denote by $\delta_\omega \in M(\Omega)$ the Dirac measure of ω . Then for every $\sigma(C_b(\Omega), M(\Omega))$ -convergent net $(f_\lambda)_{\lambda \in \Lambda}$ in $I_{\{\omega\}}$ with limit $f \in C_b(\Omega)$ we have

$0 = \langle f_\lambda, \delta_\omega \rangle \rightarrow \langle f, \delta_\omega \rangle = f(\omega)$. Hence, $f \in I_M$ and the ideal is $\sigma(C_b(\Omega), M(\Omega))$ -closed.

In fact, we can describe all τ -closed ideals in $C_b(\Omega)$ as vanishing ideals.

Proposition 1.2.27 (weakly closed ideals). *Let Ω be a completely regular Hausdorff space and let τ be a locally convex topology on $C_b(\Omega)$ that is consistent with $(C_b(\Omega), M(\Omega))$. Then the τ -closed algebra and τ -closed lattice ideals coincide and the mappings*

$$\begin{aligned} M &\mapsto I_M := \{f \in C_b(\Omega) : f(\omega) = 0 \text{ for all } \omega \in M\}, \\ I &\mapsto M_I := \{\omega \in \Omega : f(\omega) = 0 \text{ for all } f \in I\} \end{aligned}$$

are mutually inverse bijections between the set of τ -closed ideals of $C_b(\Omega)$ and the set of closed subsets of Ω .

This follows readily from [Coo71, Proposition I.2.7], as the mixed topology there is consistent with $(C_b(\Omega), M(\Omega))$. For convenience we provide an elementary proof.

Proof. Note that by Proposition 1.2.7 we may assume that $\tau = \sigma(C_b(\Omega), M(\Omega))$.

“*Well-definedness of $M \mapsto I_M$* ”: It is Remark 1.2.26 that states that the mapping $M \mapsto I_M$ maps subsets $M \subseteq X$ into τ -closed ideals in $C_b(X)$. So the mapping $M \mapsto I_M$ is well-defined.

“*Surjectivity of $M \mapsto I_M$* ”: Let I be a τ -closed algebra or lattice ideal in $C_b(X)$ and let $M := M_I$. Then clearly, $I \subseteq I_M$. To show the converse, fix $f \in I_M$ and assume without restriction that $\|f\|_\infty = 1$. We construct a net $(f_\lambda)_{\lambda \in \Lambda}$ in I that converges to f with respect to τ , as this shows that $f \in I$ by the τ -closedness of I . Let $\Lambda := M(\Omega)_+ \times (0, 1)$ ordered such that $(\mu_1, \varepsilon_1) \leq (\mu_2, \varepsilon_2)$ if and only if $\mu_1 \leq \mu_2$ and $\varepsilon_1 \geq \varepsilon_2$. Then Λ is a directed set. Fix $(\mu, \varepsilon) \in \Lambda$ and assume that $\mu \neq 0$. As μ is Radon, there exists a compact set $K \subseteq \Omega$ such that $\mu(\Omega \setminus K) < \varepsilon$. Define further

$$N_\varepsilon := \{\omega \in K : |f(\omega)| \geq \|\mu\|_1^{-1} \cdot \varepsilon\}$$

and notice that N_ε is a closed subset of Ω that lies in $K \setminus M$, and thus, it is compact. Moreover, for each point $\omega \in N_\varepsilon$ there exists a function $f_\omega \in I$ such that $f_\omega(\omega) > 0$. By passing to $|f_\omega| \in I$ and scaling, we may assume that $f_\omega \geq 0$ and $f_\omega(\omega) > 1$. Then the collection of open sets $\{\hat{\omega} \in \Omega : f_\omega(\hat{\omega}) > 1\}$ indexed over all $\omega \in N_\varepsilon$ covers N_ε . Thus, by compactness of N_ε there exist $0 \leq f_1, \dots, f_k \in I$ such that

$$N_\varepsilon \subseteq \bigcup_{i=1}^k f_i^{-1}((1, \infty)).$$

Set $\hat{h} := f_1 + \dots + f_k \in I$. For some to be specified $n \in \mathbb{N}$ define the function

$$h := \frac{n\hat{h}}{1+n\hat{h}}f.$$

If I is an algebra ideal, then clearly $h \in I$ as $\hat{h} \in I$ and $\frac{n}{1+n\hat{h}}f \in C_b(X)$. If I is a lattice ideal, then it follows from $\left| \frac{n}{1+n\hat{h}}f \right| \leq n$ that $|h| \leq n\hat{h}$, and thus, $h \in I$.

Now notice that for $\omega \in N_\varepsilon$ we have $\hat{h}(\omega) \geq 1$ and $\frac{1}{1+n} \leq \|\mu\|_1^{-1}\varepsilon$ if we chose n large enough. For $\omega \in K \setminus N_\varepsilon$ we have $|f(\omega)| \leq \|\mu\|_1^{-1}\varepsilon$. Thus,

$$\begin{aligned} \langle |h - f|, \mu \rangle &= \int_{\Omega} \left| \frac{n\hat{h}}{1+n\hat{h}}f - f \right| d\mu = \int_{\Omega} \left| \frac{f}{1+n\hat{h}} \right| d\mu \\ &\leq \int_{N_\varepsilon} \frac{|f|}{1+n} d\mu + \int_{K \setminus N_\varepsilon} \frac{\|\mu\|_1^{-1}\varepsilon}{1+n\hat{h}} d\mu + \int_{\Omega \setminus K} \|f\|_\infty d\mu \\ &\leq \|\mu\|_1^{-1}\varepsilon \int_{N_\varepsilon} |f| d\mu + \|\mu\|_1^{-1}\varepsilon \int_{K \setminus N_\varepsilon} \mathbb{1} d\mu + \mu(\Omega \setminus K) \\ &\leq \|\mu\|_1^{-1}\|f\|_\infty \|\mu\|_1 \varepsilon + \|\mu\|_1^{-1}\mu(K \setminus N_\varepsilon)\varepsilon + \varepsilon \\ &\leq 3\varepsilon. \end{aligned}$$

Setting $f_{(\mu, \varepsilon)} := h$ we obtain a net $(f_\lambda)_{\lambda \in \Lambda}$ in I that converges to f with respect to τ . This shows that $f \in I$, which proves the surjectivity of the mapping $M \mapsto I_M$.

“*Injectivity of $M \mapsto I_M$* ”: We show that $I \mapsto M_I$ is its left-inverse, i.e., that $M = M_{I_M}$, as then injectivity follows. Let $M \subseteq \Omega$ be a closed subset of Ω and $\omega \in M$. Then for all $f \in I_M$ we have $f(\omega) = 0$, which shows that $\omega \in M_{I_M}$. Conversely, let that $\omega \notin M$. Then by complete regularity there exists a function $f \in I_M$ with $f(\omega) = 1$. Thus, $\omega \notin M_{I_M}$. This shows that $M = M_{I_M}$. \square

In general the claim of Proposition 1.2.27 is false for ideals closed in the topology induced by the supremum norm on $C_b(\Omega)$.

Example 1.2.28. There exist a completely regular Hausdorff space Ω and a norm closed ideal in $C_b(\Omega)$ that is neither the entire space nor a vanishing ideal.

Let $\Omega := \mathbb{N}$. We provide a concrete and a more abstract example.

- (i) Consider the subset $c_0 \subseteq C_b(\Omega)$ of all sequences that converge to zero. The space c_0 is a closed subspace of $C_b(\mathbb{N})$ and thus a Banach lattice. Moreover, it is easily seen to be a lattice ideal of $C_b(\Omega)$. However, it is not a vanishing ideal, as there is no point $\omega \in \Omega$ such that all sequences in c_0 vanish at ω .
- (ii) Recall the notation from Proposition 1.1.11. Then the embedding $\iota_\Omega: \Omega \rightarrow \beta\Omega$ into its Stone–Čech compactification is not surjective. If $\hat{\omega} \in \beta\Omega \setminus \iota_\Omega(\Omega)$, then the vanishing ideal $I_{\{\hat{\omega}\}} \subseteq C(\beta\Omega)$ is norm closed. Note that the image $J := T_{\iota_\Omega}(I_{\{\hat{\omega}\}})$ is also norm closed in $C_b(\Omega)$. However, the ideal J is neither the entire space nor a vanishing ideal.

Indeed, for $\omega \in X$ by Urysohn’s lemma there exists a function $f \in C(\beta\Omega)$ such that $f(\iota_\Omega(\omega)) = 1$ and $f(\hat{\omega}) = 0$. So $T_{\iota_\Omega}(f)(\omega) = 1$ and $T_{\iota_\Omega}(f) \in J$. It follows that for each $\omega \in \Omega$ there exists a function $f \in J$ such that $f(\omega) \neq 0$. Thus, J is not a vanishing ideal.

Lemma 1.2.29. *Let Ω be a completely regular Hausdorff space. For a linear functional $\psi: C_b(\Omega) \rightarrow \mathbb{K}$ the following are equivalent.*

- (a) ψ is a Dirac functional

$$\delta_\omega: C_b(\Omega) \rightarrow \mathbb{K}, \quad f \mapsto f(\omega)$$

for some $\omega \in \Omega$.

(b) ψ is a $\sigma(C_b(\Omega), M(\Omega))$ -continuous unital $*$ -homomorphism.

(c) ψ is a $\sigma(C_b(\Omega), M(\Omega))$ -continuous unital lattice homomorphism.

Proof. “(a) \Rightarrow (b) and (c)”: Let $\omega \in \Omega$. Then clearly $\delta_\omega(\mathbb{1}) = 1$, is multiplicative, commutes with the involution on $C_b(\Omega)$ and with the modulus function. Hence, δ_ω is a unital $*$ -homomorphism and a unital lattice homomorphism.

The $\sigma(C_b(\Omega), M(\Omega))$ -continuity of δ_ω follow directly from the definition of the topology.

“(b) or (c) \Rightarrow (a)”: By the continuity of ψ it follows that $\ker(\psi)$ is a $\sigma(C_b(\Omega), M(\Omega))$ -closed algebra or lattice ideal, respectively. It follows from Proposition 1.2.27 that there is a unique closed set $M \subseteq \Omega$ such that $\ker(\psi) = I_M$. Note that $\psi \neq 0$ as it is unital. By the isomorphism theorem for vector spaces, the quotient space $C_b(\Omega)/\ker(\psi)$ is isomorphic as a vector space to $\text{Im}(\psi) = \mathbb{C}$. So $\ker(\psi)$ has codimension one.

We next claim that M is merely a singleton $\{\omega\}$ for some $\omega \in \Omega$. Suppose to show a contradiction that M contains at least two distinct points $\omega_1, \omega_2 \in M$. Then by complete regularity and the Hausdorff property, there exist two functions $f_1, f_2 \in C_b(\Omega)$ such that

$$\begin{aligned} f_1(\omega_1) = f_2(\omega_2) = 1 & \quad \text{and} \\ f_1(\omega_2) = f_2(\omega_1) = 0. \end{aligned}$$

Then $f_1 + I_M$ and $f_2 + I_M$ are linearly independent in the quotient space $C_b(X)/I_M$. Since the quotient space is one-dimensional, this yields a contradiction and it follows that $M = \{\omega\}$ for some $\omega \in \Omega$.

Let $\omega \in \Omega$ be such that $\ker(\psi) = I_{\{\omega\}}$. For $f \in C_b(\Omega)$ we always have that $f(\omega) \cdot \mathbb{1} - f \in I_{\{\omega\}}$, and thus, it follows from unitality of ψ that

$$f(\omega) - \psi(f) = \psi(f(\omega) \cdot \mathbb{1}) - \psi(f) = \psi(f(\omega) \cdot \mathbb{1} - f) = 0.$$

Hence, $\psi = \delta_\omega$ is a Dirac functional. □

Following the idea of Example 1.2.28 we can construct a norm continuous unital $*$ -homomorphism, or equivalently a norm continuous unital lattice homomorphism (see Remark 1.1.12), $\psi: C_b(\Omega) \rightarrow \mathbb{K}$ that are not of the form δ_ω for some $\omega \in \Omega$. Consequently, if Ω is not compact, there will always be norm continuous functionals that are not Dirac.

Example 1.2.30. There exist a completely regular Hausdorff space Ω and a norm continuous unital $*$ -homomorphism, or equivalently, a norm continuous unital lattice homomorphism, $\psi: C_b(\Omega) \rightarrow \mathbb{K}$ that is not of the form δ_ω for some $\omega \in \Omega$.

Take $\Omega := \mathbb{N}$ and recall the notation from Proposition 1.1.11. Then there exists $\hat{\omega} \in \beta\Omega \setminus \iota_\Omega(\Omega)$. Consider the Dirac functional $\delta_{\hat{\omega}}$ on $C(\beta\Omega)$ and note that $\psi := \delta_{\hat{\omega}} \circ T_{\iota_\Omega}$ defines a norm continuous functional on $C_b(\Omega)$. However, ψ is not of the form δ_ω for some $\omega \in \Omega$.

Indeed, the kernel of ψ is of the form $\ker(\psi) = T_{\iota_\Omega}(I_{\{\hat{\omega}\}})$, which is no vanishing ideal by Example 1.2.28, whereas $\ker(\delta_\omega) = I_{\{\omega\}}$.

We can now apply Lemma 1.2.29 to characterize the Koopman operators. Note that Lemma 1.2.29 is a special case of the following proposition by considering the spaces Ω_1 with a single element.

Proposition 1.2.31. *Let Ω_1 and Ω_2 be completely regular Hausdorff spaces and denote by $\tau_i := \sigma(C_b(\Omega_i), M(\Omega_i))$ the weak topologies for $i \in \{1, 2\}$. For a linear operator $T: C_b(\Omega_2) \rightarrow C_b(\Omega_1)$ the following are equivalent.*

- (a) *There exists a continuous map $\varphi: \Omega_1 \rightarrow \Omega_2$ such that $T = T_\varphi$.*
- (b) *The operator T is a unital algebra homomorphism and τ_2 - τ_1 -continuous.*
- (c) *The operator T is a unital lattice homomorphism and τ_2 - τ_1 -continuous.*

Proof. “(a) \Rightarrow (b) and (c)”: By Remark 1.1.10 it follows that the Koopman operator is a unital algebra and lattice homomorphism. The τ_2 - τ_1 -continuity follows readily from Proposition 1.2.21.

“(b) or (c) \Rightarrow (a)”: Let $\omega_1 \in \Omega$. Then the mapping

$$\delta_{\omega_1} \circ T: C_b(\Omega) \rightarrow \mathbb{K}, \quad f \mapsto (Tf)(\omega_1)$$

is an algebra or lattice homomorphism, respectively. Moreover, it is unital and τ_2 -continuous. Thus, by Lemma 1.2.29 there exists an element $\omega_2 \in \Omega_2$ such that

$$\delta_{\omega_1} \circ T = \delta_{\omega_2}. \tag{1.1}$$

Let $\varphi: \Omega_1 \rightarrow \Omega_2$ be the mapping that maps ω_1 to ω_2 satisfying (1.1). The continuity of φ follows from Lemma 1.1.5, since it follows from $f \circ \varphi = Tf$ that $f \circ \varphi$ is continuous for all $f \in C_b(\Omega_2)$. \square

1.3 Characterizing Topological Dynamics with Operators

The injectivity and surjectivity of a mapping $\varphi: K_1 \rightarrow K_2$ between two compact Hausdorff spaces K_1 and K_2 can be characterized by the properties of the induced Koopman operator $T_\varphi: C(K_2) \rightarrow C(K_1)$.

Proposition 1.3.1. *Let K_1 and K_2 be compact Hausdorff spaces and $\varphi: K_1 \rightarrow K_2$ be a continuous mapping. Then the following holds:*

- (i) *The Koopman operator $T_\varphi: C(K_2) \rightarrow C(K_1)$ is surjective if and only if φ is injective.*
- (ii) *The Koopman operator $T_\varphi: C(K_2) \rightarrow C(K_1)$ is injective if and only if φ is surjective.*

We will skip this proof as this result will follow from the more general results in this section. It can be found as an exercise in [EFHN15, Lemma 4.14].

For mappings $\varphi: \Omega_1 \rightarrow \Omega_2$ on completely regular Hausdorff spaces Ω_1 and Ω_2 the Proposition 1.3.1 (i) fails to hold.

Example 1.3.2. There exist completely regular Hausdorff spaces Ω_1 and Ω_2 and a continuous mapping $\varphi: \Omega_1 \rightarrow \Omega_2$ that is injective such that T_φ is not surjective.

Let $\Omega_1 = [0, 1]_d$ be the unit interval endowed with the discrete topology and $\Omega_2 = [0, 1]$ the usual unit interval endowed with the Euclidean topology. Consider the identity mapping $\varphi = \text{id}_{[0,1]}$ and note that φ is continuous and injective. The space $C_b([0, 1]_d)$ is the set of all functions $f: [0, 1] \rightarrow \mathbb{K}$ and the range of the Koopman operator T_φ is the space of all continuous functions $C_b([0, 1])$ on the unit interval $[0, 1]$. Hence, T_φ is not surjective.

This example motivates to look for condition under which it can be guaranteed that φ is injective.

Theorem 1.3.3. *Let Ω_1 and Ω_2 be completely regular Hausdorff spaces and $\varphi: \Omega_1 \rightarrow \Omega_2$ be continuous. Denote by τ a locally convex topology on $C_b(\Omega_1)$ that is consistent with $(C_b(\Omega_1), M(\Omega_1))$. The following statements are equivalent.*

- (a) φ is injective.
- (b) T_φ has τ -dense range in $C_b(\Omega_1)$.
- (c) P_φ is a lattice homomorphism.
- (d) P_φ is injective.

Proof. “(a) \Rightarrow (c)”: We show that for each $\mu \in M(\Omega_1)$ we have $|P_\varphi\mu| = P_\varphi|\mu|$. Since P_φ is a positive operator, the inequality \leq always holds. To see the converse inequality, let $A \subseteq \Omega_2$ be measurable and note that

$$P_\varphi|\mu|(A) = \sup \left\{ \sum_{k=1}^n |\mu(B_k)| : B_k \subseteq \varphi^{-1}(A) \text{ meas.}, \text{ pairwise disjoint, } n \in \mathbb{N} \right\},$$

and analogously,

$$|P_\varphi\mu|(A) = \sup \left\{ \sum_{k=1}^n |\mu(\varphi^{-1}(A_k))| : A_k \subseteq A \text{ meas.}, \text{ pairwise disjoint, } n \in \mathbb{N} \right\}.$$

Now fix $n \in \mathbb{N}$ and let B_1, \dots, B_n be a family of measurable and pairwise disjoint subsets of $\varphi^{-1}(A)$. Then the sets $A_k := \varphi(B_k)$ for $k \in \{1, \dots, n\}$ are measurable subsets of $\varphi(\varphi^{-1}(A)) \subseteq A$. It follows from the injectivity of φ that the B_k 's are pairwise disjoint and that $\varphi^{-1}(A_k) = B_k$. Hence,

$$\sum_{k=1}^n |\mu(B_k)| = \sum_{k=1}^n |\mu(\varphi^{-1}(A_k))|$$

and taking the supremum over all possible families yields that $|P_\varphi\mu|(A) \geq P_\varphi|\mu|(A)$. This shows that the inequality \geq holds.

“(c) \Rightarrow (d)”: Suppose P_φ is not injective. Then there exists a non-zero Radon measure $\mu \in M(\Omega_1)$ in the kernel of P_φ . Since P_φ is assumed to be a lattice homomorphism it follows that $P_\varphi|\mu| = |P_\varphi\mu| = 0$, and thus, $|\mu|$ is a positive non-zero measure in the kernel of P_φ . Thus,

$$0 < \langle \mathbb{1}, |\mu| \rangle = \langle T_\varphi \mathbb{1}, |\mu| \rangle = \langle \mathbb{1}, P_\varphi|\mu| \rangle = 0.$$

This is a contradiction.

“(d) \Rightarrow (b)”: Denote by $F \subseteq C_b(\Omega_2)$ the τ -closure of $\text{ran}(T_\varphi)$ and suppose that F is a proper subspace of $C_b(\Omega_2)$. Then, by the Hahn–Banach theorem for locally convex spaces [AB06, Theorem 5.87], there exists a non-zero τ -continuous functional $\mu: C_b(\Omega_1) \rightarrow \mathbb{K}$ that vanishes on F . As τ is consistent with $(C_b(\Omega_1), M(\Omega_1))$ it follows that $\mu \in M(\Omega_1)$. Thus,

$$0 = \langle T_\varphi f, \mu \rangle = \langle f, P_\varphi \mu \rangle$$

for every $f \in C_b(\Omega_2)$. This implies that $P_\varphi \mu = 0$, and so, the Perron–Frobenius operator is not injective.

“(b) \Rightarrow (a)”: Let $\omega_1, \omega_2 \in \Omega_1$ be distinct points that satisfy $\varphi(\omega_1) = \varphi(\omega_2)$. As Ω_1 is completely regular Hausdorff, there exists a continuous function $f \in C_b(\Omega_1)$ with $f(\omega_1) \neq f(\omega_2)$. Since $\text{ran}(T_\varphi)$ is τ -dense, there exists a net $(f_\lambda)_\lambda$ in $C_b(\Omega_2)$ such that $(T_\varphi f_\lambda)_\lambda$ converges to f with respect to τ . As $\sigma(C_b(\Omega_1), M(\Omega_1))$ is coarser than τ the convergence also holds weakly. Hence, setting $\mu := \delta_{\omega_1} - \delta_{\omega_2} \in M(\Omega_1)$ we obtain

$$0 = f_\lambda(\varphi(\omega_1)) - f_\lambda(\varphi(\omega_2)) = \langle T_\varphi f_\lambda, \mu \rangle \rightarrow \langle f, \mu \rangle = f(\omega_1) - f(\omega_2) \neq 0.$$

This is a contradiction. □

For mappings $\varphi: \Omega_1 \rightarrow \Omega_2$ on completely regular Hausdorff spaces Ω_1 and Ω_2 the Proposition 1.3.1 (ii) fails to hold as well.

Example 1.3.4. There exist completely regular Hausdorff spaces Ω_1 and Ω_2 and a continuous mapping $\varphi: \Omega_1 \rightarrow \Omega_2$ that is not a surjective such that T_φ is injective.

Let $\Omega := \mathbb{N}$ and $\alpha\Omega := \mathbb{N} \cup \{\infty\}$ be its one-point compactification. Consider the inclusion mapping $\varphi: \Omega \rightarrow \alpha\Omega$. The Koopman operator $T_\varphi: C_b(\alpha\Omega) \rightarrow C_b(\Omega)$ is given by the restriction of functions from $\alpha\Omega$ to Ω and its image is the space c of all functions $f \in C_b(\Omega)$ for which $\lim_{n \rightarrow \infty} f(n)$ exists. In particular, T_φ is injective, as any function $f \in c$ has a unique extension $\hat{f} \in C_b(\alpha\Omega)$ by setting $\hat{f}(\infty) := \lim_{n \rightarrow \infty} f(n)$. However, φ is not surjective as $\infty \in \alpha\Omega$ is not in the image of φ .

We explore conditions under which T_φ is injective.

Theorem 1.3.5. *Let Ω_1 and Ω_2 be completely regular Hausdorff spaces and $\varphi: \Omega_1 \rightarrow \Omega_2$ be continuous. Denote by τ a locally convex topology on $M(\Omega_2)$ that is consistent with $(M(\Omega_2), C_b(\Omega_2))$. The following statements are equivalent.*

- (a) φ has dense image in Ω_2 .
- (b) T_φ is injective.
- (c) T_φ is an isometry with respect to the supremum norms on $C_b(\Omega_2)$ and $C_b(\Omega_1)$.
- (d) P_φ has τ -dense range in $M(\Omega_2)$.

Proof. “(a) \Rightarrow (c)”: Clearly, $\|f\|_\infty \geq \|f \circ \varphi\|_\infty = \|T_\varphi f\|_\infty$. So we need to show the converse inequality. Let $f \in C_b(\Omega_2)$ with $M := \sup\{|f(\omega_2)| : \omega_2 \in \Omega_2\}$ and

let $\varepsilon > 0$. By continuity the set $U_\varepsilon := \{\omega_2 \in \Omega_2 : |f(\omega_2)| > M - \varepsilon\}$ is open and non-empty. Thus, by density of φ in Ω_2 there exists $\omega_1 \in \Omega_1$ such that $\varphi(\omega_1) \in U_\varepsilon$. Hence, $\|f\|_\infty < |f(\varphi(\omega_1))| + \varepsilon \leq \|T_\varphi f\|_\infty + \varepsilon$. Since ε was chosen arbitrarily it follows that $\|f\|_\infty \leq \|T_\varphi f\|_\infty$.

“(c) \Rightarrow (b)”: Every isometry between two metric spaces is injective.

“(b) \Rightarrow (d)”: Denote by $F \subseteq M(\Omega_2)$ the τ -closure of $\text{ran}(P_\varphi)$ and suppose that F is a proper subset of $M(\Omega_2)$. By the Hahn-Banach theorem [AB06, Theorem 5.87] there exists a non-zero τ -continuous functional $f := M(\Omega_2) \rightarrow \mathbb{K}$ that vanishes on F . Since τ is consistent with $(M(\Omega_2), C_b(\Omega_2))$ it follows that $f \in C_b(\Omega_2)$ and

$$0 = \langle f, P_\varphi \mu \rangle = \langle T_\varphi f, \mu \rangle$$

for all $\mu \in M(\Omega_1)$. This implies that $T_\varphi f = 0$ and shows that T_φ is not injective.

“(d) \Rightarrow (a)”: Suppose that $\varphi(\Omega_1)$ is not dense in Ω_2 . Then there exists a non-empty open set U in Ω_2 that is disjoint of the image $\varphi(\Omega_1)$. Then $(P_\varphi \mu)(U) = \mu(\varphi^{-1}(U)) = \mu(\emptyset) = 0$ for all $\mu \in M(\Omega_1)$. By the complete regularity of Ω_2 there exists a continuous functions $0 \leq f \in C_b(\Omega_2)$ that is 0 on $\Omega_2 \setminus U$ and satisfies $f(\omega_2) = 1$ for some $\omega_2 \in \Omega_2$. Thus,

$$\langle f, P_\varphi \mu \rangle = 0 \quad \text{and} \quad \langle f, \delta_{\omega_2} \rangle = 1$$

hold for all $\mu \in M(\Omega_2)$ simultaneously. This shows that $\text{ran}(P_\varphi)$ is not dense in $M(\Omega_2)$ with respect to $\sigma(M(\Omega_2), C_b(\Omega_2))$. Thus, by Proposition 1.2.7 $\text{ran}(P_\varphi)$ can also not be τ -dense. \square

We note that a trivial corollary of Proposition 1.3.1 is the following.

Corollary 1.3.6. *Let K_1 and K_2 be compact Hausdorff spaces and $\varphi: K_1 \rightarrow K_2$ be a continuous mapping. Then the Koopman operator T_φ is invertible if and only if φ is a homeomorphism.*

That this is not true for continuous maps on completely regular Hausdorff spaces is demonstrated in the following example.

Example 1.3.7. There exist completely regular Hausdorff spaces Ω_1 and Ω_2 and a continuous mapping $\varphi: \Omega_1 \rightarrow \Omega_2$ that is not a homeomorphism such that T_φ is bijective.

Let $\Omega := \mathbb{N}$ and recall the notation from Proposition 1.1.11. Consider the embedding $\iota_\Omega: \Omega \rightarrow \beta\Omega$ into its Stone–Čech compactification. The Koopman operator $T_{\iota_\Omega}: C(\beta\Omega) \rightarrow C_b(\Omega)$ is an algebra and lattice isomorphism. Recall that ι is a homeomorphism onto its image. Hence, ι is not surjective, as Ω is not compact whereas $\beta\Omega$ is.

We will now determine conditions under which the bijectivity of the Koopman operator T_φ guarantee that φ is a homeomorphism.

Theorem 1.3.8. *Let Ω_1 and Ω_2 be completely regular Hausdorff spaces and $\varphi: \Omega_1 \rightarrow \Omega_2$ be continuous. If φ is a homeomorphism, then T_φ is invertible.*

If Ω_1 is normal and Ω_2 is metrizable, then T_φ is invertible if and only if φ is a homeomorphism.

If φ is a homeomorphism, then the inverse of the Koopman operator T_φ is given by the Koopman operator $T_{\varphi^{-1}}: C_b(\Omega_1) \rightarrow C_b(\Omega_2)$.

To prove this, we need the following two lemmas.

Lemma 1.3.9. *Let Ω be a metrizable topological space and $(\omega_n)_{n \in \mathbb{N}}$ be a sequence in Ω that has no convergent subsequence. Then the set of all elements*

$$A := \{\omega_n : n \in \mathbb{N}\}$$

is a closed subset of Ω .

Proof. Without loss of generality, we may assume that the elements of the sequence $(\omega_n)_{n \in \mathbb{N}}$ are pairwise distinct. Let $(\hat{\omega}_n)_{n \in \mathbb{N}}$ be a sequence in A that converges to some $\hat{\omega} \in \Omega$. We show that $(\hat{\omega}_n)_{n \in \mathbb{N}}$ only takes finitely many values, and hence, is eventually constant. This implies that $\hat{\omega} \in A$ and in turn that A is closed.

If we suppose that $(\hat{\omega}_n)_{n \in \mathbb{N}}$ takes infinitely many values, then we may pick a subsequence $(\tilde{\omega}_n)_{n \in \mathbb{N}}$ of $(\hat{\omega}_n)_{n \in \mathbb{N}}$ that has pairwise distinct elements. Let $h: \mathbb{N} \rightarrow \mathbb{N}$ be the mapping defined by

$$h(n) := m, \quad \text{such that } \tilde{\omega}_n = \omega_m$$

for all $n \in \mathbb{N}$. This is well-defined, since $(\omega_n)_{n \in \mathbb{N}}$ was assumed to have pairwise distinct elements. Moreover, h is injective, since $(\tilde{\omega}_n)_{n \in \mathbb{N}}$ has pairwise distinct elements. It follows that $h(n) \rightarrow \infty$ as $n \rightarrow \infty$. Therefore, we may pick a strictly monotone subsequence $(m_n)_{n \in \mathbb{N}}$ of $(h(n))_{n \in \mathbb{N}}$ and obtain that $(\omega_{m_n})_{n \in \mathbb{N}}$ is a subsequence of $(\tilde{\omega}_n)_{n \in \mathbb{N}}$. As a subsequence of $(\omega_n)_{n \in \mathbb{N}}$ the sequence $(\omega_{m_n})_{n \in \mathbb{N}}$ can not converge, but as a subsequence of $(\tilde{\omega}_n)_{n \in \mathbb{N}}$ the sequence $(\omega_{m_n})_{n \in \mathbb{N}}$ must converge. This is a contradiction.

It follows that $(\hat{\omega}_n)_{n \in \mathbb{N}}$ may only take finitely many values and that A is closed. \square

Lemma 1.3.10. *Let Ω_1 be a normal space, Ω_2 be metrizable and $\varphi: \Omega_1 \rightarrow \Omega_2$ be continuous. If the Koopman operator T_φ is surjective, then $\varphi(\Omega_1)$ is closed.*

Proof. Suppose that there exists an element $\omega \in \overline{\varphi(\Omega_1)} \setminus \varphi(\Omega_1)$. Since Ω_2 is metrizable also $\overline{\varphi(\Omega_1)}$ is. Hence, there exists a sequence $(\omega_n)_{n \in \mathbb{N}}$ in Ω_1 such that $(\varphi(\omega_n))_{n \in \mathbb{N}}$ converges to ω . Without loss of generality, we may assume that $(\varphi(\omega_n))_{n \in \mathbb{N}}$ has pairwise distinct elements. Since $\omega \notin \varphi(\Omega_1)$ it follows that the sequence $(\varphi(\omega_n))_{n \in \mathbb{N}}$ does not have any convergent subsequence in the metrizable space $\varphi(\Omega_1)$. Thus, it follows from Lemma 1.3.9 that the sets

$$A_1 := \{\varphi(\omega_{2n}) : n \in \mathbb{N}\} \quad \text{and} \quad A_2 := \{\varphi(\omega_{2n+1}) : n \in \mathbb{N}\}$$

are closed in $\varphi(\Omega_1)$. Notice also that A_1 and A_2 are disjoint. By the continuity of $\varphi: \Omega_1 \rightarrow \varphi(\Omega_1)$ it follows that the sets B_1 and B_2 , defined by $B_i := \varphi^{-1}(A_i)$ for $i \in \{1, 2\}$, are disjoint and closed in Ω_1 . Since Ω_1 is normal, it follows from Urysohn's lemma that there exists a continuous function $f \in C_b(\Omega_1)$ such that

$f|_{B_1} \equiv 0$ and $f|_{B_2} = 1$. Since $\text{ran}(T_\varphi) = C_b(\Omega_1)$ there exists $g \in C_b(\Omega_2)$ such that $g \circ \varphi = T_\varphi g = f$. Then the sequence $((g \circ \varphi)(\omega_n))_{n \in \mathbb{N}}$ takes the values 0 and 1 infinitely many times and does not converge. However, by the continuity of g the limit

$$\lim_{n \rightarrow \infty} (g \circ \varphi)(\omega_n) = \lim_{n \rightarrow \infty} g(\varphi(\omega_n)) = g(\omega)$$

exists. This is a contradiction, which shows that $\varphi(\Omega_1)$ must have been closed in the first place. \square

In the setting of Lemma 1.3.10 it does not suffice to assume that T_φ has closed range to conclude that $\varphi(\Omega_1)$ is closed.

Example 1.3.11. There exist a normal space Ω_1 , a metrizable space Ω_2 and a continuous mapping $\varphi: \Omega_1 \rightarrow \Omega_2$ such that the Koopman operator T_φ has closed range but $\varphi(\Omega_1)$ is not closed.

Let $\Omega = \mathbb{N}$. Recall from Example 1.3.4 the inclusion mapping $\varphi: \Omega \rightarrow \alpha\Omega$ into the one-point compactification $\alpha\Omega$. As shown in that example, the range of the Koopman operator T_φ is given by the closed subspace c . However, the image $\varphi(\Omega)$ is not closed in $\alpha\Omega$.

We are now able to prove Theorem 1.3.8.

Proof of Theorem 1.3.8. “ \Rightarrow ”: If φ is a homeomorphism, then the operator $T_{\varphi^{-1}}$ is easily checked to be the inverse of T_φ . Hence, T_φ is invertible.

“ \Leftarrow ”: We first show that φ^{-1} exists. From Theorem 1.3.3 and Theorem 1.3.5 it follows that φ is injective and has dense image.

If Ω_1 is normal and Ω_2 is metrizable, then Lemma 1.3.10 implies that $\varphi(\Omega_1)$ is closed, and thus, φ is surjective. It also follows that φ is surjective, and thus, we obtain that φ is bijective and φ^{-1} exists.

Now we show that φ^{-1} is continuous. Let $f \in C_b(\Omega_1)$ and set $g := T_\varphi^{-1}f$. Then we obtain

$$f \circ \varphi^{-1} = (T_\varphi g) \circ \varphi^{-1} = g \circ \varphi \circ \varphi^{-1} = g \in C_b(\Omega_2).$$

It follows that $f \circ \varphi^{-1}$ is continuous for all $f \in C_b(\Omega_1)$. Now Lemma 1.1.5 implies that φ^{-1} is continuous. \square

Note that the inverse of an invertible unital algebra or lattice homomorphism is again an unital algebra or lattice homomorphism, respectively. So in light of Proposition 1.2.31, if we want to conclude from the bijectivity of T_φ that φ is a homeomorphism, it suffices to show that the inverse is also weakly continuous with respect to the Radon measures. Let us make this more precise.

Proposition 1.3.12. *Let Ω_1 and Ω_2 be completely regular Hausdorff spaces and $\varphi: \Omega_1 \rightarrow \Omega_2$ be continuous. Suppose that the Koopman operator $T_\varphi: C_b(\Omega_2) \rightarrow C_b(\Omega_1)$ is invertible. Then T_φ^{-1} is continuous with respect to the weak topologies $\sigma(C_b(\Omega_1), M(\Omega_1))$ and $\sigma(C_b(\Omega_2), M(\Omega_2))$ if and only if φ is a homeomorphism.*

Proof. “ \Rightarrow ”: If φ is a homeomorphism, then by Theorem 1.3.8 the inverse of the Koopman operator is given by $T_{\varphi^{-1}}$. By Proposition 1.2.21 the operator $T_{\varphi^{-1}}$ is weakly continuous.

“ \Leftarrow ”: Suppose that T_{φ}^{-1} is weakly continuous. Then, by Proposition 1.2.31 the mapping T_{φ}^{-1} is a Koopman operator. Thus, there exists a continuous mapping $\psi: \Omega_2 \rightarrow \Omega_1$ such that $T_{\varphi}^{-1} = T_{\psi}$. It follows that $\psi = \varphi^{-1}$ and that φ is a homeomorphism. \square

For the norm continuity the continuity of the inverse is automatic and follows from the open mapping theorem.

Let us revisit Example 1.3.7 and verify that the weak continuity of the inverse actually fails to hold (without using the insight of Proposition 1.3.12).

Example 1.3.13. There exist completely regular Hausdorff spaces Ω_1 and Ω_2 and a continuous map $\varphi: \Omega_1 \rightarrow \Omega_2$ such that the Koopman operator T_{φ} is invertible but T_{φ}^{-1} is not continuous with respect to the weak topologies $\sigma(C_b(\Omega_i), M(\Omega_i))$ for $i \in \{1, 2\}$.

Recall the notation from Example 1.3.4, let $\Omega = \mathbb{N}$ and recall from Proposition 1.1.11 that the inverse of the Koopman operator $T_{\iota_{\Omega}}$ is the mapping $\Phi_{\Omega}: C_b(\Omega) \rightarrow C(\beta\Omega)$, where $\beta\Omega$ is the space of all unital algebra homomorphisms endowed with the weak*-topology. Let $\alpha \in \beta\Omega \setminus \iota_{\Omega}(\Omega)$ and consider the Dirac measure $\delta_{\alpha} \in M(\beta\Omega)$. To show that Φ_{Ω} is not weakly continuous, we need to show that for every $\mu \in M(\Omega)$ there exists a function $f \in C_b(\Omega)$ such that

$$|\langle \Phi_{\Omega} f, \delta_{\alpha} \rangle| > |\langle f, \mu \rangle|.$$

By tightness there exists a compact set $K \subseteq \Omega$ such that $|\mu|(\Omega \setminus K) < 1/2$. Note that for every open neighborhood $U \subseteq \beta\Omega$ of α the preimage $\iota_{\Omega}^{-1}(U)$ is also open. Now there exists an open neighborhood U of α such that $\iota_{\Omega}^{-1}(U) \cap K = \emptyset$. Otherwise we can construct a net $(\omega_U)_U$ in Ω indexed by the neighborhood basis of α such that $\omega_U \in \iota_{\Omega}^{-1}(U) \cap K$ for every neighborhood U of α . By compactness of K and by passing to a subnet, we may assume that the net converges to some $\omega \in K$. Thus, we have by continuity $\iota_{\Omega}(\omega_U) \rightarrow \iota_{\Omega}(\omega)$ and $\iota_{\Omega}(\omega_U) \in U$ for every neighborhood U of α . It follows that $\iota_{\Omega}(\omega) = \alpha$ contradicting the assumption that $\alpha \notin \iota_{\Omega}(\Omega)$.

By complete regularity of Ω and Proposition 1.1.4 (ii) there exists a continuous function $f: \Omega \rightarrow [0, 1]$ such that $f(\omega) = 0$ for all $\omega \in K$ and $f(\omega) = 1$ for all $\omega \in \iota_{\Omega}^{-1}(U)$. Let $(\omega_{\lambda})_{\lambda}$ be a net in Ω such that $(\iota_{\Omega}(\omega_{\lambda}))_{\lambda}$ converges to α . Then eventually for large enough λ we have $\iota_{\Omega}(\omega_{\lambda}) \in U$ and

$$1 = |\langle \Phi_{\Omega} f, \delta_{\omega_{\lambda}} \rangle| \rightarrow |\langle \Phi_{\Omega} f, \delta_{\alpha} \rangle|,$$

whereas

$$|\langle f, \mu \rangle| < \int_K f \, d\mu + \mu(\Omega \setminus K) < 1/2.$$

This concludes the proof.

1.4 The Koopman Operator on the Bounded Measurable Functions

For some investigations the continuous functions do not provide sufficient details to fully study the underlying dynamical system. The attentive readers might have noticed that Section 1.3 does not provide an operator theoretic description of the surjectivity of φ . To do this we need the Koopman operator to be defined on a richer space of functions. One natural choice is to allow for functions that are Borel measurable.

We let Ω be completely regular Hausdorff and denote by $B_b(\Omega)$ the space of Borel measurable and bounded functions $f: \Omega \rightarrow \mathbb{K}$. If endowed with the supremum norm $\|f\|_\infty := \sup_{\omega \in \Omega} |f(\omega)|$, the space of bounded measurable functions becomes a Banach space. In fact, it has similar algebraic properties to those described in Remark 1.1.9 for the space of bounded continuous functions. If the order, the modulus function, the multiplication and the involution are defined analogously, then $B_b(\Omega)$ is

- (i) a Banach lattice that is also an AM-space
- (ii) a Banach algebra.

We define the Koopman operator on $B_b(\Omega)$ in complete analogy to the Koopman operator on $C_b(\Omega)$.

Definition 1.4.1 (Koopman operator on bounded measurable functions). Let Ω_1 and Ω_2 be completely regular Hausdorff spaces and $\varphi: \Omega_1 \rightarrow \Omega_2$ be continuous. The *Koopman operator* $T_\varphi: B_b(\Omega_2) \rightarrow B_b(\Omega_1)$ on the bounded measurable functions is defined by

$$(T_\varphi f)(\omega) = f(\varphi(\omega))$$

for all $f \in B_b(\Omega_2)$ and all $\omega \in \Omega_1$.

Since $f \circ \varphi$ is measurable for all $f \in B_b(\Omega_2)$, the Koopman operator is well-defined. Notice that we use the same notation for the Koopman operator on the continuous functions as for the Koopman operator on the measurable functions. In the following this shall not lead to ambiguities.

Let us now introduce the weak topologies on $B_b(\Omega)$. For this recall from Proposition 1.2.18 that $(B_b(\Omega), M(\Omega))$ is a norming dual pair of Banach lattices. Akin to Proposition 1.2.20, let us remark that the Koopman operator T_φ on $B_b(\Omega)$ is also dual to the Perron–Frobenius operator P_φ on the Radon measures.

Proposition 1.4.2. *Let Ω_1 and Ω_2 be completely regular Hausdorff spaces and $\varphi: \Omega_1 \rightarrow \Omega_2$ be continuous. Denote by $T_\varphi: B_b(\Omega_2) \rightarrow B_b(\Omega_1)$ the Koopman operator and by $P_\varphi: M(\Omega_1) \rightarrow M(\Omega_2)$ the Perron–Frobenius operator of φ . Then it holds that*

$$\langle T_\varphi f, \mu \rangle_{B_b(\Omega_1), M(\Omega_1)} = \langle f, P_\varphi \mu \rangle_{B_b(\Omega_2), M(\Omega_2)}$$

for all $f \in B_b(\Omega_2)$ and all $\mu \in M(\Omega_1)$. In particular, T_φ and P_φ are their mutual adjoint operators.

Proof. This is the transformation theorem for pushforwards of measures applied to the measurable functions. \square

As we have seen in Section 1.2.1 the setting of norming dual pairs allows us to consider a variety of weak topologies $\sigma(\mathbb{B}_b(\Omega), \mathbb{M}(\Omega)) \subseteq \tau \subseteq \mu(\mathbb{B}_b(\Omega), \mathbb{M}(\Omega))$ sandwiched between the weak and the Mackey topology that respect the dual space $\mathbb{M}(\Omega)$. As in Theorem 1.3.3, we will use this topology to characterize the injectivity of φ .

Furthermore, we need the weaker *topology of pointwise convergence* τ_{ptw} on $\mathbb{B}_b(\Omega)$, which is the coarsest topology on $\mathbb{B}_b(\Omega)$ for which the Dirac functionals

$$\delta_\omega: \mathbb{B}_b(\Omega) \rightarrow \mathbb{K}, \quad f \mapsto f(\omega)$$

are continuous. A net $(f_\lambda)_\lambda$ converges in this topology if and only if it converges pointwise, i.e., there exists $f \in \mathbb{B}_b(\Omega)$ such that $f_\lambda(\omega) \rightarrow f(\omega)$ for all $\omega \in \Omega$. Clearly, the topology $\sigma(\mathbb{B}_b(\Omega), \mathbb{M}(\Omega))$ is stronger than τ_{ptw} , i.e., convergence with respect to $\sigma(\mathbb{B}_b(\Omega), \mathbb{M}(\Omega))$ implies pointwise convergence. The converse does not hold in general. A common counterexample is given by the net $(\mathbb{1}_F)_F$ of indicator functions $\mathbb{1}_F: [0, 1] \rightarrow \mathbb{R}$ indexed by the finite subsets F of $[0, 1]$ ordered by reverse inclusion, which converges to the constant one function pointwisely, but not in the weak topology $\sigma(\mathbb{B}_b([0, 1]), \mathbb{M}([0, 1]))$, as can be seen by testing against the Lebesgue measure.

Note that the Koopman operator T_φ is continuous with respect to the topology of pointwise convergence.

Proposition 1.4.3. *Let Ω_1 and Ω_2 be completely regular Hausdorff spaces and $\varphi: \Omega_1 \rightarrow \Omega_2$ be continuous. Then the Koopman operator $T_\varphi: \mathbb{B}_b(\Omega_2) \rightarrow \mathbb{B}_b(\Omega_1)$ is continuous with respect to the weak topologies $\sigma(\mathbb{B}_b(\Omega_2), \mathbb{M}(\Omega_2))$ and $\sigma(\mathbb{B}_b(\Omega_1), \mathbb{M}(\Omega_1))$, and continuous with respect to the topology of pointwise convergence on both spaces.*

Proof. The $\sigma(\mathbb{B}_b(\Omega_2), \mathbb{M}(\Omega_2))$ - $\sigma(\mathbb{B}_b(\Omega_1), \mathbb{M}(\Omega_1))$ -continuity follows exactly as in Proposition 1.2.21. To show that T_φ is continuous with respect to the topology of pointwise convergence, let $(f_\lambda)_\lambda$ be a net in $\mathbb{B}_b(\Omega_2)$ that converges pointwise to some $f \in \mathbb{B}_b(\Omega_2)$. Then, for every $\omega \in \Omega_1$ we have

$$(T_\varphi f_\lambda)(\omega) = f_\lambda(\varphi(\omega)) \rightarrow f(\varphi(\omega)) = (T_\varphi f)(\omega)$$

as λ tends to its limit. This shows the continuity of T_φ with respect to the topology of pointwise convergence. \square

Let us now characterize the injectivity of φ in terms of the Koopman operator on the bounded measurable functions.

Theorem 1.4.4. *Let Ω_1 and Ω_2 be completely regular Hausdorff spaces and let $\varphi: \Omega_1 \rightarrow \Omega_2$ be continuous. Denote by T_φ the Koopman operator on the bounded measurable functions and by τ a locally convex topology on $\mathbb{B}_b(\Omega_1)$ that is consistent with $(\mathbb{B}_b(\Omega_1), \mathbb{M}(\Omega_1))$. The following statements are equivalent.*

- (a) φ is injective.
- (b) T_φ has τ -dense range in $\mathbb{B}_b(\Omega_1)$.

(c) T_φ has dense range with respect to the topology of pointwise convergence.

Proof. “(a) \Rightarrow (b)”: Let F denote the τ closure of $\text{ran}(T_\varphi)$ in $B_b(\Omega_1)$. Suppose that F is a proper subset of $B_b(\Omega_1)$. By the Hahn–Banach theorem [AB06, Theorem 5.87] there exists a non-zero τ -continuous linear functional $\mu: M(\Omega_1) \rightarrow \mathbb{K}$ that vanishes on F . By consistency of τ with $(B_b(\Omega_1), M(\Omega_1))$ it follows that $\mu \in M(\Omega_1)$ and

$$0 = \langle T_\varphi f, \mu \rangle = \langle P_\varphi \mu, f \rangle$$

for all $f \in B_b(\Omega_2)$. This implies $P_\varphi \mu = 0$, and thus, P_φ is not injective. Now it follows from Theorem 1.3.3 that φ is not injective.

“(b) \Rightarrow (c)”: This is clear, as the convergence with respect to τ necessarily implies pointwise convergence.

“(c) \Rightarrow (a)”: If φ is not injective, there exist two distinct points $\omega_1, \omega_2 \in \Omega_1$ such that $T_\varphi f(\omega_1) = T_\varphi f(\omega_2)$ for all $f \in B_b(\Omega_2)$. Thus, the indicator function $f = \mathbb{1}_{\{\omega_1\}}$ does not lie in the closure of the range of T_φ with respect to the topology of pointwise convergence. \square

Let us now turn to a characterization of the surjectivity of φ . For this we need the consistent topologies of the reversed pair $(M(\Omega), B_b(\Omega))$.

Theorem 1.4.5. *Let Ω_1 and Ω_2 be completely regular Hausdorff spaces and $\varphi: \Omega_1 \rightarrow \Omega_2$ be continuous. Denote T_φ the Koopman operator on the bounded measurable functions and by τ a locally convex topology on $M(\Omega_2)$ that is consistent with $(M(\Omega_2), B_b(\Omega_2))$. The following statements are equivalent.*

- (a) φ is surjective
- (b) T_φ is injective.
- (c) T_φ is an isometry with respect to the supremum norms on $B_b(\Omega_2)$ and $B_b(\Omega_1)$.
- (d) P_φ has τ -dense range in $M(\Omega_2)$.

Proof. The implications “(a) \Rightarrow (c) \Rightarrow (b) \Rightarrow (d)” are proved in complete analogy to the same implications in the proof of Theorem 1.3.5. We show the remaining implication.

“(d) \Rightarrow (a)”: Suppose to show a contradiction that φ is not surjective. Then there exists $\omega \in \Omega_2 \setminus \varphi(\Omega_1)$. Setting $f := \mathbb{1}_{\{\omega\}}$ we find that for every $\mu \in M(\Omega_2)$ we have

$$\langle f, P_\varphi \mu \rangle = \mu(\varphi^{-1}(\{\omega\})) = \mu(\emptyset) = 0$$

and

$$\langle f, \delta_\omega \rangle = \delta_\omega(\{\omega\}) = 1.$$

We conclude that δ_ω is not in the $\sigma(M(\Omega_2), B_b(\Omega_2))$ -closure of $\text{ran}(P_\varphi)$. Since τ is finer than the weak topology, it follows that δ_ω is also not in the τ -closure of $\text{ran}(P_\varphi)$. \square

Let us return to Example 1.3.4.

Example 1.4.6. There exist completely regular Hausdorff spaces Ω_1 and Ω_2 and a continuous map $\varphi: \Omega_1 \rightarrow \Omega_2$ such that the Koopman operator $T_\varphi: C_b(\Omega_2) \rightarrow C_b(\Omega_1)$ is injective but the Koopman operator $T_\varphi: B_b(\Omega_2) \rightarrow B_b(\Omega_1)$ is not injective.

Let $\Omega = \mathbb{N}$ and $\alpha\Omega := \mathbb{N} \cup \{\infty\}$ and consider the embedding $\varphi: \Omega \rightarrow \alpha\Omega$. We have seen in Example 1.3.4 that the Koopman operator $T_\varphi: C_b(\alpha\Omega) \rightarrow C_b(\Omega)$ is injective. However, the Koopman operator $T_\varphi: B_b(\alpha\Omega) \rightarrow B_b(\Omega)$ is not injective, as the subspace spanned by the indicator function $\mathbb{1}_{\{\infty\}}$ lies in its kernel.

1.5 Notes

By passing from the possibly nonlinear dynamics φ to the linear Koopman and Perron–Frobenius operators T_φ and P_φ , it is a priori unclear which properties of φ translate to properties of the linear operators and vice versa. Therefore, results on the relation between properties have been studied extensively in the literature. Our contribution in this chapter has been to generalize known results from the setting of compact Hausdorff spaces to completely regular Hausdorff spaces and do so using the tools provided by the theory of norming dual pairs. We managed to describe the injectivity and surjectivity of φ in terms of density and injectivity properties of T_φ and P_φ with respect to weak topologies. Moreover, we characterized when φ is a homeomorphism in terms of the bijectivity of T_φ and the weak continuity of its inverse.

An extensive survey on how ergodic properties of dynamical systems on compact Hausdorff spaces translate to properties of T_φ and P_φ can be found in [EFHN15] and more recently in [EF25]. In [Küh21] it is shown how almost weak, weak and strong attractors translate to convergence properties of the corresponding Koopman operators on a closed ideal. A similar investigation for Perron–Frobenius operators has been carried out in [Ger18].

In [Küs21a] the level sets of the fixed space of T_φ have been described in terms of superorbits of φ . This investigation has been extended to include suffix functions, which form the so-called Lyapunov algebra, in [Küs21b, Section II.3].

The spectrum of Koopman operators on $C(K)$ with K compact Hausdorff has been fully described in [Sch71a, Theorem 2.7] and can take only a few forms: either it is the entire unit ball or a subgroup of the unit circle potentially with zero. Sufficient criteria for either case are given. In the second case 1 might be isolated in the spectrum. We show in Chapter 6 that this is equivalent to the ergodicity and the eventual periodicity of Koopman operators.

An introduction to Koopman semigroups of dynamical systems $\varphi = (\varphi_t)_{t \geq 0}$ on completely regular Hausdorff spaces that are time-continuous can be found in [FK20]. The long-term behavior, ergodicity and stability properties of such semigroups have been studied in [Eis10].

In Section 1.2, we have developed the duality between the Koopman and Perron–Frobenius operators and used it to characterize the class of Koopman operators. Dualities of this type have been studied first by Jean Dieudonné in 1940 and George W. Mackey in 1945 (see [Pie07, Section 3.3.2.4 on p. 70]). The exposition given in Section 1.2 closely follows the presentation in [SW99, Section IV.5], [AB06, Section 5.15] and more recently [Kun11], who also requires that the dual pairs are norming. Kunze also studied the norming dual pair $(C_b(\Omega), M(\Omega))$ in Example 2.4.

For studying Koopman semigroups on $C_b(\Omega)$ that are strongly continuous with respect to the weak topology $\sigma(C_b(\Omega), M(\Omega))$ it has been noted in [FK20] that one benefits from a topology that is quasi-complete. A locally convex topological vector space is called *quasi-complete* if every bounded closed set in it is complete. It has been noted in [Kun11, Remark 4.5] and follows from [Köt69, Section 18.4 (4)] that a quasi-complete consistent topology on $C_b(\Omega)$ exists if and only if the Mackey topology is quasi-complete. This is the case if and only if Ω is compactly generated, i.e., a function $f: \Omega \rightarrow \mathbb{K}$ is continuous if and only if its restriction to every compact subset of Ω is continuous, see [Coo87, Corollary II.1.9. 1)]. The topology that is shown to be quasi-complete in this case is the so-called *strict topology*, which can be equivalently described as the *mixed topology*, namely the finest locally convex topology that coincides with the compact-open topology on norm-bounded sets, see [Coo87, Section II.1], where it is called β_K , and [Sen72], where it is referred to as the substrict topology β_0 . A generating family of seminorms of the strict topology is presented in [Coo87, Proposition II.1.11] and it has been shown that the continuous dual of the strict topology are the Radon measures $M(\Omega)$ [Coo87, Proposition II.3.3]. Hence, the strict topology is consistent with the norming dual pair $(C_b(\Omega), M(\Omega))$.

Measure-Preserving Dynamical Systems

This chapter introduces measure-preserving dynamical systems as the central object of study in ergodic theory. We begin in Section 2.1 with the definition of measure-preserving maps and Koopman operators. We then describe alternative ways to view measure-preserving systems, such as measure algebra homomorphisms and unital lattice homomorphisms on L^p -spaces.

In Section 2.2 we introduce standard notions of ergodic theory. We recall Birkhoff's pointwise ergodic theorem, which we will need in the proof of convergence of the extended Dynamical Mode decomposition further later in this thesis. When describing mixing properties of dynamical systems, we place particular emphasis on the characterizations in terms of the Jacob–de Leeuw–Glicksberg decomposition of its associated Koopman operator. This decomposition will be a recurrent tool throughout many parts of this thesis.

Finally, in Section 2.3 we show how measure-preserving dynamics arise naturally as subsystems of topological dynamics, when restricted to the supports of invariant measures. Fixed points of the Perron–Frobenius operator yield topological measure-preserving systems and the Krylov–Bogoliubov theorem shows the existence of such.

The purpose of this chapter is to recall notations from the literature, to frame ergodic notions in terms of the Jacobs–de Leeuw–Glicksberg decomposition, and to prepare the ground for later chapters. It does not contain new and original results.

2.1 Measure-Preserving Systems and their Koopman Operators

Let us begin by defining measure-preserving maps and measure-preserving dynamical systems.

Definition 2.1.1. Let $X = (X, \Sigma_X, \mu)$ and $Y = (Y, \Sigma_Y, \nu)$ be probability spaces.

- (i) A Σ_X - Σ_Y -measurable mapping $\varphi: X \rightarrow Y$ is called *measure-preserving* if

$$\mu(\varphi^{-1}(A)) = \nu(A)$$

for all $A \in \Sigma_Y$. The mapping φ is called *invertible* if there exists a measure-preserving mapping $\varphi^{-1}: Y \rightarrow X$ such that $\varphi \circ \varphi^{-1} = \text{id}_Y$ holds ν -almost surely and $\varphi^{-1} \circ \varphi = \text{id}_X$ holds μ -almost surely.

- (ii) If $X = Y$, a measure-preserving mapping $\varphi: X \rightarrow X$ is referred to as a *measure-preserving dynamical system* or just *measure-preserving system*.

Example 2.1.2. (i) Fix a probability space $Y = (Y, \Sigma_Y, \nu)$ and consider the infinite product space $X = (X, \Sigma_X, \mu)$, where

$$X := \prod_{k \in \mathbb{N}} Y, \quad \Sigma_X := \bigotimes_{k \in \mathbb{N}} \Sigma_Y, \quad \text{and} \quad \mu := \bigotimes_{k \in \mathbb{N}} \nu.$$

The shift mapping

$$\varphi: X \rightarrow X, \quad (x_k)_{k \in \mathbb{N}} \mapsto (x_{k+1})_{k \in \mathbb{N}}$$

is a measure-preserving system on X and called a *(one-sided) Bernoulli shift*. If we replace \mathbb{N} by \mathbb{Z} we call the resulting measure-preserving system a *two-sided Bernoulli shift*.

Often the probability space Y is assumed to be finite, i.e., of the form $Y = \{0, \dots, N-1\}$ with $\Sigma_Y = 2^Y$ and $\nu(A) = \sum_{a \in A} p_a$ for some probability vector $p = (p_0, \dots, p_{N-1})$.

(ii) Denote by $\mathbb{T} = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$ the complex unit circle, let \mathbf{m} be the normalized Lebesgue measure on the Borel- σ -algebra $\mathcal{B}(\mathbb{T})$ of \mathbb{T} , and let $\alpha \in (0, 1)$ be irrational. Then the rotation

$$\varphi: \mathbb{T} \rightarrow \mathbb{T}, \quad \lambda \mapsto e^{2\pi i \alpha} \lambda$$

defines a measure-preserving system on $(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathbf{m})$, which we refer to as an *irrational rotation*.

(iii) On the probability space X from (ii) the square mapping

$$\varphi: \mathbb{T} \rightarrow \mathbb{T}, \quad \lambda \mapsto \lambda^2$$

is a measure-preserving system on X . Indeed, denoting by $\varphi_* \mathbf{m}$ the *pushforward measure* defined by $\varphi_* \mathbf{m}(A) = \mathbf{m}(\varphi^{-1}(A))$ for all $A \in \mathcal{B}(\mathbb{T})$ we have

$$\begin{aligned} \int_{\mathbb{T}} f \, d\varphi_* \mathbf{m} &= \int_{\mathbb{T}} f \circ \varphi \, d\mathbf{m} \\ &= \int_0^\pi f(\varphi(e^{\lambda i})) \, \mathbf{m}(d\lambda) + \int_\pi^{2\pi} f(\varphi(e^{\lambda i})) \, \mathbf{m}(d\lambda) \\ &= \int_0^{2\pi} \frac{1}{2} f(e^{\lambda i}) \, \mathbf{m}(d\lambda) + \int_{2\pi}^{4\pi} \frac{1}{2} f(e^{\lambda i}) \, \mathbf{m}(d\lambda) \\ &= \int_0^{2\pi} f(e^{\lambda i}) \, \mathbf{m}(d\lambda) = \int_{\mathbb{T}} f \, d\mathbf{m} \end{aligned}$$

for all bounded measurable functions $f: \mathbb{T} \rightarrow \mathbb{R}$. Note that this system is isomorphic as a measure-preserving system to the doubling map $\psi: [0, 1] \rightarrow [0, 1]$ preserving the Lebesgue measure in [EFHN15, Section 5.1], i.e., there exists an invertible measure-preserving mapping $h: \mathbb{T} \rightarrow [0, 1]$ intertwining both systems as $h \circ \varphi = \psi \circ h$.

For more examples, we refer the interested reader to [EFHN15, Section 5.1].

Definition 2.1.3 (Koopman operator). Let $\varphi: X \rightarrow Y$ be a measure-preserving mapping between the probability spaces $X = (X, \Sigma_X, \mu)$ and $Y = (Y, \Sigma_Y, \nu)$ and $p \in [1, \infty]$. We call the operator

$$T_\varphi: L^p(Y) \rightarrow L^p(X), \quad f \mapsto f \circ \varphi$$

the *Koopman operator* associated with φ .

Note the similarity of the Koopman operator to the Koopman operator for topological dynamical system defined in Definition 1.1.6. Both operators are defined via composition with the underlying dynamical system.

The Koopman operator is well-defined in that it maps p -integrable functions to p -integrable functions. This is a consequence of the measure-preserving property of φ . It has some further properties.

Remark 2.1.4. Let $\varphi: X \rightarrow Y$ be a measure-preserving mapping between the probability spaces $X = (X, \Sigma_X, \mu)$ and $Y = (Y, \Sigma_Y, \nu)$ and $p \in [1, \infty]$. Denote by T_φ the Koopman operator of φ .

- (i) The Koopman operator T_φ is an isometry, i.e., it satisfies $\|T_\varphi f\|_{L^p(X)} = \|f\|_{L^p(Y)}$ for all $f \in L^p(Y)$.
- (ii) It is a *lattice homomorphism*, i.e., it satisfies $|T_\varphi f| = T_\varphi |f|$ for all $f \in L^p(Y)$.
- (iii) The space $L^\infty(X)$ is a Banach algebra with respect to pointwise multiplication. Moreover, the Koopman operator T_φ is an *algebra homomorphism*, i.e., it satisfies $T_\varphi(f \cdot g) = (T_\varphi f) \cdot (T_\varphi g)$ for all $f, g \in L^\infty(Y)$.

Let us remark the following about the adjoint of the Koopman operator.

Remark 2.1.5 (Perron–Frobenius operator). Let $\varphi: X \rightarrow Y$ be a measure-preserving mapping between the probability spaces $X = (X, \Sigma_X, \mu)$ and $Y = (Y, \Sigma_Y, \nu)$ and $p \in [1, \infty]$. We denote by $q \in [1, \infty]$ the *Hölder adjoint* of p that satisfies $\frac{1}{p} + \frac{1}{q} = 1$. Consider the norming dual pair $(L^p(X), L^q(X))$ with the duality pairing

$$\langle f, g \rangle_{L^p(X), L^q(X)} = \int_X fg \, d\mu$$

for each $f \in L^p(X)$ and each $g \in L^q(X)$ (cf. Example 1.2.2 (i)). It is easily verified that T_φ is $\sigma(L^p(X), L^q(X))$ -continuous. Hence, by the discussion preceding Proposition 1.2.8, it follows that it has an adjoint $P_\varphi: L^q(X) \rightarrow L^q(Y)$ satisfying

$$\langle T_\varphi f, g \rangle_{L^p(X), L^q(X)} = \langle f, P_\varphi g \rangle_{L^p(Y), L^q(Y)}$$

for each $f \in L^p(Y)$ and each $g \in L^q(X)$ is called the *Perron–Frobenius operator*. If φ is invertible, then it follows from an application of the transformation theorem for measures that

$$P_\varphi: L^q(X) \rightarrow L^q(Y), \quad g \mapsto g \circ \varphi^{-1}.$$

This shows that T_φ is invertible with inverse $T_\varphi^* = P_\varphi$.

In particular, if $X = Y$, $p = q = 2$ and φ is invertible, then T_φ is unitary.

Recall from Definition 1.1.16 that the Perron–Frobenius operator for topological dynamical systems is defined on a space of measures. Here the interpretation is similar: the space $L^q(X)$ can be densely embedded into a space of measures on X that are absolutely continuous with respect to μ via the Radon–Nikodym theorem.

The Koopman operators are elements of a larger class of operators on Banach lattices called *Markov operators*. We will give a brief introduction to these operators on L^p -spaces and show how these operators can be utilized to characterize factors of measure-preserving dynamical systems. We refer to [EFHN15, Section 13] for a more thorough introduction into Markov operators on L^p -spaces.

Definition 2.1.6 (Markov operator).

- (i) Let $X = (X, \Sigma_X, \mu)$ and $Y = (Y, \Sigma_Y, \nu)$ be probability spaces and $p \in [1, \infty]$. A linear operator $T: L^p(X) \rightarrow L^p(Y)$ is called a *Markov operator* if
 - (a) $\int_Y Tf \, d\nu = \int_X f \, d\mu$ for all $f \in L^p(X)$,
 - (b) $Tf \geq 0$ for all $0 \leq f \in L^p(X)$ and
 - (c) $T\mathbb{1}_X = \mathbb{1}_Y$.
- (ii) We call a Markov operator T a *Markov embedding* if additionally

$$T|f| = |Tf|$$

holds for all $f \in L^p(X)$, i.e., T is a lattice homomorphism.

- (iii) A Markov embedding is called a *Markov isomorphism* if it is surjective.

Note that every Markov isomorphism T is invertible and its inverse T^{-1} is easily checked to be a Markov embedding as well. It is easily seen that every Koopman operator is a Markov embedding.

We will now introduce measure-algebra homomorphisms as another way of viewing measure-preserving dynamical systems. We have already used terminology from Banach lattice theory above. In what follows we need more general notions from lattice theory which we briefly repeat here. We follow [EFHN15, Section 12.2] in our presentation.

A lattice is a partially ordered set (V, \leq) such that

$$x \wedge y := \inf\{x, y\} \quad \text{and} \quad x \vee y := \sup\{x, y\}$$

exist in V for all $x, y \in V$. It is called *distributive* if

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z) \quad \text{and} \quad x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$$

for all $x, y, z \in V$. It is further called *complete* if every set $A \subseteq V$ has a supremum and an infimum.

A *Boolean algebra* is a distributive lattice (V, \leq) with a greatest element 1 called the *unit element* and a least element 0 called the *zero element* such that for every element $x \in V$ there exists an element $y \in V$ with

$$x \vee y = 1 \quad \text{and} \quad x \wedge y = 0,$$

called the *complement* of x . The complement in a Boolean algebra is always uniquely determined [Bir67, Theorem I.16]. We denote it by x^c . We call two elements x, y in a Boolean algebra V *disjoint* if $x \wedge y = 0$.

Definition 2.1.7 (measure-algebras and homomorphisms).

- (i) A (*probability*) *measure-algebra* is a pair (V, μ) consisting of a complete Boolean algebra (V, \leq) together with a mapping $\mu: V \rightarrow [0, \infty)$ such that
- (a) $\mu(1) = 1$,
 - (b) $\mu(x) = 0$ if and only if $x = 0$, and
 - (c) for each sequence $(x_n)_{n \in \mathbb{N}}$ of pairwise disjoint elements in V , we have

$$\mu \left(\bigvee_{n \in \mathbb{N}} x_n \right) = \sum_{n \in \mathbb{N}} \mu(x_n).$$

- (ii) Let (V_1, μ_1) and (V_2, μ_2) be measure-algebras. A mapping $\Phi: V_1 \rightarrow V_2$ is called a *measure-algebra homomorphism* if
- (a) $\Phi(x \vee y) = \Phi(x) \vee \Phi(y)$,
 - (b) $\Phi(x^c) = \Phi(x)^c$,
 - (c) $\mu_1(x) = \mu_2(\Phi(x))$
- holds for all $x, y \in V_1$.

Our main example of a measure-algebra and a measure-algebra homomorphism is the following.

Example 2.1.8. (i) If $X = (X, \Sigma, \mu)$ is a probability space, we define an equivalence relation on Σ by setting

$$A \sim B \quad :\Leftrightarrow \quad \mu(A \Delta B) = 0.$$

The resulting space

$$\tilde{\Sigma} := \{[A]_{\sim} : A \in \Sigma\}$$

of equivalence classes together with the shorthand $0 := [\emptyset]_{\sim}$ and $1 := [X]_{\sim}$ for the zero and unit element and the mapping

$$\tilde{\mu}: \tilde{\Sigma} \rightarrow [0, \infty), \quad [A]_{\sim} \mapsto \mu(A)$$

is a measure-algebra.

- (ii) If $\varphi: X \rightarrow Y$ is a measure-preserving mapping between two probability spaces $X = (X, \Sigma_X, \mu)$ and $Y = (Y, \Sigma_Y, \nu)$, then the mapping

$$\Phi: \tilde{\Sigma}_Y \rightarrow \tilde{\Sigma}_X, \quad [A]_{\sim} \mapsto [\varphi^{-1}(A)]_{\sim}$$

is a measure-algebra homomorphism between the measure-algebra $(\tilde{\Sigma}_Y, \tilde{\nu})$ and $(\tilde{\Sigma}_X, \tilde{\mu})$.

For convenience, we will not differentiate between the equivalence class $[A]_{\sim}$ and A in $\tilde{\Sigma}$.

There exists a one-to-one relation between unital Banach lattice homomorphisms on L^p and measure-algebra homomorphisms. For a proof we refer to [EFHN15, Theorem 12.10].

Proposition 2.1.9. *Let $p \in [1, \infty)$ and let $X = (X, \Sigma_X, \mu)$ and $Y = (Y, \Sigma_Y, \nu)$ be probability spaces and denote by $(\tilde{\Sigma}_X, \tilde{\mu})$ and $(\tilde{\Sigma}_Y, \tilde{\nu})$ their associated measure-algebras as defined in Example 2.1.8 (i).*

- (i) *Let $\Phi: \tilde{\Sigma}_X \rightarrow \tilde{\Sigma}_Y$ be a measure-algebra homomorphism. Then there is a unique Markov embedding $T_\Phi: L^p(X) \rightarrow L^p(Y)$ that maps indicator functions as $\mathbb{1}_A \mapsto \mathbb{1}_{\Phi(A)}$ for all $A \in \tilde{\Sigma}_X$.*
- (ii) *Conversely, let $T: L^p(X) \rightarrow L^p(Y)$ be a Markov embedding. Then there exists a measure-algebra homomorphism*

$$\Phi_T: \tilde{\Sigma}_X \rightarrow \tilde{\Sigma}_Y, \quad A \mapsto (T \mathbb{1}_A)^{-1}(\{1\}).$$

Moreover, the identities $\Phi = \Phi_{T_\Phi}$ and $T = T_{\Phi_T}$ hold, and Φ is surjective if and only if T_Φ is a Markov isomorphism.

Example 2.1.10 (Continuation of Example 2.1.8). Recall the measure-algebra homomorphism $\Phi: \tilde{\Sigma}_Y \rightarrow \tilde{\Sigma}_X$ from Example 2.1.8 (ii) and denote by $T_\varphi: L^p(Y) \rightarrow L^p(X)$ the Koopman operator of the underlying measure-preserving mapping for some $p \in [1, \infty)$. Then the associated unital algebra homomorphism $T_\Phi: L^p(Y) \rightarrow L^p(X)$ satisfies

$$T_\Phi \mathbb{1}_A = \mathbb{1}_{\Phi(A)} = \mathbb{1}_{\varphi^{-1}(A)} = T_\varphi \mathbb{1}_A$$

for all $A \in \tilde{\Sigma}_Y$, where $\mathbb{1}_A$ denotes the indicator function of any representative of A . By density of the indicator functions in $L^p(Y)$, we obtain that $T_\Phi = T_\varphi$.

To obtain a similar one-to-one characterization between Koopman operators and measure-preserving dynamics, we need to restrict ourselves to a special class of probability spaces. Following [Gla03, Section 2.1] we introduce standard Lebesgue spaces, as they form a central assumption in most of the literature on ergodic theory.

- Definition 2.1.11** (standard Lebesgue system). (i) A measurable space (X, Σ) , where X is a completely metrizable and second countable topological space and $\Sigma = \mathcal{B}(X)$ is the Borel- σ -algebra of X , is called a *standard Borel space*.
- (ii) A measure space (X, Σ, μ) is called a *standard Lebesgue space* if (X, Σ) is a standard Borel and additionally μ is a Borel probability measure on X , i.e., a probability measure defined on the Borel- σ -algebra $\Sigma = \mathcal{B}(X)$.
 - (iii) A measure-preserving system $\varphi: X \rightarrow X$ on a standard Lebesgue space (X, Σ, μ) is called *standard Lebesgue*.

A completely metrizable and second countable topological space is sometimes referred to as a *Polish space*. A Borel probability measure on a Polish space is called

(inner) regular [AB06, Theorem 12.7] or tight, if $\mu(A) = \sup\{\mu(K) : K \in \Sigma, K \subseteq A \text{ is compact}\}$. As it is a Borel probability measure, the tightness automatically implies the regularity [AB06, Theorem 12.4].

If K is compact and metrizable, then it is second countable and complete, so $(K, \mathcal{B}(K), \mu)$ is standard Lebesgue for each probability measure $\mu: \mathcal{B}(K) \rightarrow [0, 1]$.

Proposition 2.1.12 (standard Lebesgue systems). *Let $X = (X, \Sigma, \mu)$ be a probability space and $p \in [1, \infty)$. If X is standard Lebesgue, then $L^p(X)$ is separable.*

Conversely, let $\varphi: X \rightarrow X$ be a measure-preserving dynamical system on the probability space $X = (X, \Sigma, \mu)$. Suppose that $L^p(X)$ is separable, then there exists

- (i) a compact metric space K with Borel- σ -algebra $\mathcal{B}(K)$,
- (ii) a Borel probability measure $\nu: \mathcal{B}(K) \rightarrow [0, 1]$ with $\text{supp}(\nu) = K$,
- (iii) a continuous and measure-preserving mapping $\psi: K \rightarrow K$ and
- (iv) a Markov isomorphism $T: L^p(K, \mathcal{B}(K), \nu) \rightarrow L^p(X)$

such that T intertwines the Koopman operators T_φ and T_ψ on $L^p(X)$ and $L^p(K, \mathcal{B}(K), \nu)$, respectively, as

$$T_\varphi T = T T_\psi.$$

In particular, the probability space $(K, \mathcal{B}(K), \nu)$ is standard Lebesgue.

Proof. “ \Rightarrow ”: Let \mathcal{U} be a countable basis for the topology of X . Then \mathcal{U} is a countable generator of the σ -algebra Σ . Define the countable set

$$\tilde{\mathcal{U}} := \{U_1 \cap U_n : U_1, \dots, U_n \in \mathcal{U}, n \in \mathbb{N}\}.$$

Now it follows readily from a monotone class argument that

$$\text{span}\{1_U : U \in \tilde{\mathcal{U}}\} \subseteq L^p(X)$$

is dense in $L^p(X)$.

“ \Leftarrow ”: As $L^p(X)$ is separable and dense in $L^1(X)$, it follows readily that $L^1(X)$ is separable. Thus the statement follows from [EFHN15, Theorem 12.22]. \square

Proposition 2.1.13 (Von Neumann). *Let $X = (X, \Sigma_X, \mu)$ and $Y = (Y, \Sigma_Y, \nu)$ be standard Lebesgue probability spaces, $p \in [1, \infty)$ and let $T: L^p(Y) \rightarrow L^p(Y)$ be a Markov embedding. Then there exists a μ -almost everywhere unique measure-preserving mapping $\varphi: X \rightarrow Y$ such that $T = T_\varphi$, where T_φ denotes the Koopman operator of φ on L^p .*

Proof. By [EFHN15, Proposition 13.6] the Markov operator T on L^p can be uniquely extended to a Markov embedding on L^1 . Now the statement follows from [EFHN15, Theorem F.9] and the discussion in [EFHN15, Remark F.3] on the definition of a standard Lebesgue space. \square

For later reference, we will need the notion of a factor. So let us finish this section with a discussion on factor mappings of measure-preserving dynamical systems.

Definition 2.1.14 (factor). Let $p \in [1, \infty)$ and $\varphi: X \rightarrow X$ be a measure-preserving map on the probability space $X = (X, \Sigma_X, \mu)$. We call a measure-preserving map $\psi: Y \rightarrow Y$ on the probability space $Y = (Y, \Sigma_Y, \nu)$ a *factor of φ* , if there exists a Markov embedding $T: L^p(Y) \rightarrow L^p(X)$ such that $T_\varphi T = TT_\psi$.

Note that by [EFHN15, Proposition 13.6] this definition is independent of the p chosen. A factor in the sense of the above definition is sometimes also referred to as a *semi-conjugate image* of φ [Wal82, Definition 2.7].

If X and Y are both standard Lebesgue spaces, then Proposition 2.1.13 guarantees the existence of a measure-preserving mapping $h: X \rightarrow Y$ such that the Markov embedding T in Definition 2.1.14 is of the form $T = T_h$. In this case, h intertwines as $h \circ \psi = \varphi \circ h$. Most literature on ergodic theory, including [Wal82, Definition 2.6] and [Gla03, Section 2.2 on p. 55], define ψ to be a factor of φ if such a measure-preserving mapping h exists. As we often work with spaces that are not assumed to be standard Lebesgue we find it more helpful to define a factor as we have done in Definition 2.1.14.

2.2 Mean Ergodicity, Ergodicity and Mixing

In this section we provide a quick overview of some ergodic and mixing notions that will appear throughout this thesis.

Let $T: E \rightarrow E$ be a bounded linear operator on a Banach space E . We denote for each $n \in \mathbb{N}$ by

$$A_n[T]f := \frac{1}{n} \sum_{k=0}^{n-1} T^k f$$

the *n th-Cesàro average* of T .

We call an operator $T: E \rightarrow E$ on a Banach space E *mean ergodic* if the sequence $(A_n[T]f)_{n \in \mathbb{N}}$ of Cesàro averages converges in E for each $f \in E$. In this case, there exists a bounded projection $P: E \rightarrow E$ onto $\text{fix}(T)$ such that $A_n[T]f \rightarrow Pf$ for all $f \in E$ as $n \rightarrow \infty$ [EFHN15, Theorem 8.5]. We call P the *mean ergodic projection*. The next result is classical and can be found in [EFHN15, Theorem 11.1].

Proposition 2.2.1 (Birkhoff's pointwise ergodic theorem). *Let $\varphi: X \rightarrow X$ be a measure-preserving system on a probability space $X = (X, \Sigma, \mu)$ and let T_φ be the Koopman operator on $L^1(X)$. Then for each $f \in L^1(X)$ the limit*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} (T_\varphi^k f)(x)$$

exists for μ -almost every $x \in X$.

By employing Lebesgue's dominated convergence theorem we obtain the following.

Proposition 2.2.2. *Let $\varphi: X \rightarrow X$ be a measure-preserving system on a probability space $X = (X, \Sigma, \mu)$. Then the Koopman operator T_φ on $L^p(X)$ is mean ergodic for all $p \in [1, \infty)$.*

Note that usually Birkhoff's pointwise ergodic theorem is a consequence of the mean ergodicity of the Koopman operator on L^p -spaces and a maximal inequality. Here we present how the mean ergodicity follows from Birkhoff's pointwise ergodic theorem.

Proof. Let $f \in L^\infty(X)$. Let f^* denote the μ -almost everywhere pointwise-defined limit of the Cesàro sums $(A_n[T_\varphi]f)_{n \in \mathbb{N}}$, which exists by Birkhoff's pointwise ergodic theorem. As the sequence of Cesàro sums are order bounded by $\|f\|_\infty \cdot \mathbb{1}$, it follows that $f^* \in L^\infty(X)$. In particular, the mapping $f \mapsto f^*$ defines a bounded linear operator on $L^\infty(X)$. From Lebesgue's theorem on dominated convergence we obtain that

$$\lim_{n \rightarrow \infty} \int_X \left(\frac{1}{n} \sum_{k=0}^{n-1} T_\varphi^k f - f^* \right)^p d\mu = 0.$$

As $L^\infty(X)$ is dense in $L^p(X)$ and the Cesàro averages $(A_n[T_\varphi])_{n \in \mathbb{N}}$ are bounded in operator norm, the statement follows by approximation. \square

In case $p = \infty$ the above result fails to hold. In fact a Koopman operator T_φ is mean ergodic in $L^\infty(X)$ if and only if the $L^\infty(X)$ is finite dimensional [EFHN15, Proposition 12.28].

For Koopman operators $T = T_\varphi$ an additional assumption on φ yields more structure of the mean ergodic projection.

Definition 2.2.3. Let $\varphi: X \rightarrow X$ be a measure-preserving system on the probability space $X = (X, \Sigma, \mu)$. We call φ *ergodic* if every set $A \in \Sigma$ with $A \subseteq \varphi^{-1}(A)$ satisfies $\mu(A) \in \{0, 1\}$.

In terms of the Koopman operator, the property of ergodicity can be described as follows [EFHN15, Proposition 7.15].

Proposition 2.2.4. Let $\varphi: X \rightarrow X$ be a measure-preserving system on the probability space $X = (X, \Sigma, \mu)$, let $p \in [1, \infty]$ and denote by T_φ the Koopman operator on $L^p(X)$. Then φ is ergodic if and only if $\text{fix}(T_\varphi) = \text{span}\{\mathbb{1}\}$.

In this case, the mean ergodic projection P_φ of T_φ takes the form

$$P_\varphi f = \int_X f d\mu$$

for all $f \in L^p(X)$.

Proof. The first assertion is [EFHN15, Proposition 7.15].

It remains to show the second assertion. Let $f \in L^p(X)$. Since $\text{fix}(T_\varphi) = \{\mathbb{1}\}$ there exists $c_f \in \mathbb{K}$ such that $Pf = c_f \cdot \mathbb{1}$. To determine c_f notice that for essentially bounded f , using Lebesgue's dominated convergence theorem and the fact that φ is measure-preserving, we have

$$c_f = \int_X P_\varphi f d\mu = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \int_X T_\varphi^k f d\mu = \int_X f d\mu.$$

Now the statement follows from the fact that $L^\infty(X)$ is dense in $L^p(X)$. \square

Using the form of the mean ergodic projection for ergodic systems and the fact that L^p -limits and pointwise almost everywhere limits coincide almost everywhere, we obtain Birkhoff's pointwise ergodic theorem for ergodic systems [EFHN15, Corollary 11.2].

Corollary 2.2.5. *Let $\varphi: X \rightarrow X$ be an ergodic measure-preserving system on a probability space $X = (X, \Sigma, \mu)$ and let T_φ be the Koopman operator on $L^1(X)$. Then for each $f \in L^1(X)$ we have*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} (T_\varphi^k f)(x) = \int_X f \, d\mu$$

for μ -almost every $x \in X$.

Let again $T: E \rightarrow E$ be a bounded linear operator on a general Banach space E . The property of mean ergodicity can also be described by means of a decomposition of the space E into the T -invariant and closed subspaces

$$E = \text{fix}(T) \oplus \overline{\text{ran}}(I - T),$$

which we call the *mean ergodic decomposition*. We refer to [EFHN15, Theorem 8.20] for a proof and further equivalent properties. We note that $f \in \text{fix}(T)$ if and only if $Tf = f$; and $f \in \overline{\text{ran}}(I - T)$ if and only if $Pf = 0$, or equivalently, $\lim_{n \rightarrow \infty} A_n[T]f = 0$.

A similar decomposition of operators is given by the Jacob–de Leeuw–Glicksberg decomposition.

Definition 2.2.6 (Jacob–de Leeuw–Glicksberg decomposition). *Let $T: E \rightarrow E$ be a bounded linear operator on a Banach space E . We call T *Jacob–de Leeuw–Glicksberg decomposable* or *JdLG-decomposable* if E decomposes into two T -invariant and closed subspaces*

$$E = E_{\text{rev}} \oplus E_{\text{aws}}$$

with

- (i) $E_{\text{rev}} := \overline{\text{span}}\{f \in E : \exists \lambda \in \mathbb{T} : Tf = \lambda f\}$ and
- (ii) $E_{\text{aws}} := \left\{ f \in E : \forall g \in E' : \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} |\langle T^k f, g \rangle| = 0 \right\}$.

We call E_{rev} the *reversible part* and E_{aws} the *almost weakly stable part*.

Instead of using the fixed space as a factor, the JdLG-decomposition has the closed linear span of the unimodular eigenvalues as one factor of the decomposition. Let us prove that every JdLG-decomposable operator is also mean ergodic.

Proposition 2.2.7. *Let $T: E \rightarrow E$ be a bounded linear operator on a Banach space E . If T is JdLG-decomposable, then T is mean ergodic.*

Proof. To see this show the convergence of the Cesàro averages $(A_n[T])_{n \in \mathbb{N}}$ in the weak operator topology. Then the mean ergodicity follows using [EFHN15, Theorem 8.20].

For $f \in E_{\text{aws}}$ and all $g \in E'$ it follows readily that

$$\left| \left\langle \frac{1}{n} \sum_{k=0}^{n-1} T^k f, g \right\rangle \right| \leq \frac{1}{n} \sum_{k=0}^{n-1} |\langle T^k f, g \rangle| \rightarrow 0.$$

If $f \in E_{\text{rev}}$ with $Tf = \lambda f$ for some $\lambda \in \mathbb{T}$ we obtain from the geometric sum identity that

$$\frac{1}{n} \sum_{k=0}^{n-1} T^k f = \frac{1}{n} \sum_{k=0}^{n-1} \lambda^k f = \begin{cases} \frac{1}{n} \frac{1-\lambda^n}{1-\lambda} f, & \lambda \neq 1, \\ f, & \lambda = 1 \end{cases} \rightarrow \begin{cases} 0, & \lambda \neq 1, \\ f, & \lambda = 1 \end{cases},$$

since $\left| \frac{1-\lambda^n}{1-\lambda} \right|$ is bounded by $\frac{2}{1-\lambda}$ for all $n \in \mathbb{N}$. A simple approximation shows that the Cesàro means $(A_n[T])_{n \in \mathbb{N}}$ converge strongly to a projection P onto $\text{fix}(T)$. This shows that T is mean ergodic. \square

For a more abstract proof, we refer to [EFHN15, Remark 16.26].

Let us return to dynamical systems and their Koopman operators and investigate their JdLG-decomposability.

Proposition 2.2.8. *Let $\varphi: X \rightarrow X$ be a measure-preserving dynamical system on the probability space $X = (X, \Sigma, \mu)$, let $p \in [1, \infty)$. Then its Koopman operator T_φ on $L^p(X)$ is Jacob–de Leeuw–Glicksberg decomposable.*

Proof. By Theorem 16.33 and Theorem 16.34 in [EFHN15] it suffices to argue that $\{T_\varphi^k : k \in \mathbb{N}_0\}$ is relatively compact in the weak operator topology on $\mathcal{L}(L^p(X))$. For $p = 1$ this follows from [EFHN15, Theorem 16.19] and the fact that T_φ is a Markov operator. For $p \in (1, \infty)$ the space $L^p(X)$ is reflexive. So the relative weak compactness follows from [EFHN15, Corollary 16.21]. \square

With certain extra assumptions, we are able to directly compare the mean ergodic decomposition to the Jacob–de Leeuw–Glicksberg decomposition of a Koopman operator. For this we need the following notions. For a subset $J \subseteq \mathbb{N}_0$ we define its *asymptotic density* to be

$$d(J) := \lim_{n \rightarrow \infty} \frac{|J \cap \{1, \dots, n\}|}{n}$$

if the limit exists. For a sequence $(x_n)_{n \in \mathbb{N}}$ in a metric space M we say that $x_n \rightarrow x \in M$ *in density* as $n \rightarrow \infty$, if there exists a subsequence $(n_k)_{k \in \mathbb{N}}$ of $(n)_{n \in \mathbb{N}}$ such that $d(\{n_k : k \in \mathbb{N}\}) = 1$ and $x_{n_k} \rightarrow x$ in M .

Definition 2.2.9 (weakly mixing system). Let $\varphi: X \rightarrow X$ be a measure-preserving dynamical system on the probability space $X = (X, \Sigma, \mu)$. We say that φ is *weakly mixing* if

$$\mu(\varphi^{-n}(A) \cap B) \rightarrow \mu(A) \cdot \mu(B) \quad \text{in density as } n \rightarrow \infty$$

for all $A, B \in \Sigma$.

In terms of the Jacob–de Leeuw–Glicksberg decomposition of the Koopman operator T_φ the weakly mixing property can be characterized as follows.

Proposition 2.2.10. *Let $\varphi: X \rightarrow X$ be a measure-preserving dynamical system on the probability space $X = (X, \Sigma, \mu)$, let $p \in [1, \infty)$ and $q \in (1, \infty]$ be its Hölder conjugate, and denote by T_φ the Koopman operator of φ on $L^p(X)$. Then the following are equivalent.*

- (a) *The system φ is weakly mixing.*
- (b) *We have*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \left| \langle T_\varphi^k f, g \rangle - \langle f, \mathbb{1} \rangle \cdot \langle \mathbb{1}, g \rangle \right| = 0$$

for all $f \in L^p(X)$ and all $g \in L^q(X)$.

- (c) *We have $\text{fix}(T_\varphi) = \text{span}\{\mathbb{1}\}$ and $\overline{\text{ran}}(I - T_\varphi) = E_{\text{aws}}$.*
- (d) *We have $E_{\text{rev}} = \text{span}\{\mathbb{1}\}$.*
- (e) *The product system $\varphi^2: X^2 \rightarrow X^2$, $(x_1, x_2) \mapsto (\varphi(x_1), \varphi(x_2))$ is ergodic.*

In this case, the JdLG-decomposition and the mean ergodic decomposition coincide and take the form

$$L^p(X) = \text{span}\{\mathbb{1}\} \oplus E_{\text{aws}}.$$

In particular, it follows that every weakly mixing measure-preserving system φ is ergodic.

Proof. “(a) \Leftrightarrow (b)”: This follows from [EFHN15, Theorem 9.19 (i) \Leftrightarrow (ii)] together with [EFHN15, Theorem 9.15].

“(a) \Leftrightarrow (e)”: This follows from [EFHN15, Theorem 9.23].

“(c) \Rightarrow (a)”: This is [EFHN15, Theorem 9.19 (v) \Rightarrow (i)].

“(a) \Rightarrow (c)”: This is [EFHN15, Theorem 9.19 (i) \Rightarrow (v)] we obtain that $\text{fix}(T_\varphi) = \text{span}\{\mathbb{1}\}$ and $\text{ran}(I - T) \subseteq E_{\text{aws}}$. Since E_{aws} is closed it follows that $\overline{\text{ran}}(I - T_\varphi) \subseteq E_{\text{aws}}$. To see the converse inclusion let P_φ be the mean ergodic projection of T_φ . Then $I - P_\varphi$ is the projection onto $\overline{\text{ran}}(I - T_\varphi)$. Take $f \in E_{\text{aws}}$ and note that $P_\varphi f = f$. Thus, $f = (I - P_\varphi)f \in \overline{\text{ran}}(I - T_\varphi)$.

“(a) \Leftrightarrow (d)”: This follows from [EFHN15, Theorem 9.19 (i) \Leftrightarrow (vi)]. □

Example 2.2.11. There exist ergodic measure-preserving systems that are not weakly mixing.

Indeed, recall the irrational rotation $\varphi: \mathbb{T} \rightarrow \mathbb{T}$, $\lambda \mapsto e^{2\pi\alpha}\lambda$ from Example 2.1.2 (ii). This is a classical example of an ergodic system. We include the proof of the ergodicity. Let $A \in \mathcal{B}(\mathbb{T})$ with $A \subseteq \varphi^{-1}(A)$ and suppose that $\mathbf{m}(A) > 0$. Then there exists $\lambda \in \mathbb{T}$ and an open neighborhood U of λ with $U \subseteq A$. Then, by induction $U \subseteq \varphi^{-k}(A)$, and thus, $\varphi^k(U) \subset A$ for all $k \in \mathbb{N}$. Moreover, by Kronecker’s theorem (see [EFHN15, Theorem 2.26 and Example 2.37]) the set $\{e^{-2\pi k\alpha}\lambda : k \in \mathbb{N}_0\}$ is dense in \mathbb{T} . So it follows that $\bigcup_{k \in \mathbb{N}} \varphi^{-k}(U) = \mathbb{T}$. Hence, $A = \mathbb{T}$ and $\mathbf{m}(A) = 1$. This shows that φ is ergodic.

To see that φ is not weakly mixing, let T_φ be the associated Koopman operator on $L^p(X)$ for some $p \in [1, \infty]$ and consider the family of functions

$$f_k: \mathbb{T} \rightarrow \mathbb{C}, \quad \lambda \mapsto \lambda^k$$

for $k \in \mathbb{Z}$. Note that $\text{span}\{f_k : k \in \mathbb{Z}\}$ is dense in $L^p(X)$ and that $T_\varphi f_k = e^{2\pi k \alpha} f_k$ for every $k \in \mathbb{Z}$. Hence, $L^p(X) = L^p(X)_{\text{rev}}$. Now it follows from Proposition 2.2.10 that φ is not mixing, as $L^p(X)_{\text{aws}} = \{0\}$ and $L^p(X)_{\text{rev}} \neq \text{fix}(T_\varphi)$.

Definition 2.2.12 (strongly mixing system). Let $\varphi: X \rightarrow X$ be a measure-preserving dynamical system on the probability space $X = (X, \Sigma, \mu)$. We say that φ is *strongly mixing* if

$$\mu(\varphi^{-n}(A) \cap B) \rightarrow \mu(A) \cdot \mu(B) \quad \text{as } n \rightarrow \infty$$

for all $A, B \in \Sigma$.

In strongly mixing systems the vectors in the almost weakly stable E_{aws} part of the JdLG-decomposition are weakly stable, that is, they are contained in the set

$$E_{\text{ws}} := \{f \in E : \forall g \in E' : \lim_{n \rightarrow \infty} \langle T^n f, g \rangle = 0\}$$

called the *weakly stable part* of $T: E \rightarrow E$.

Proposition 2.2.13. *Let $\varphi: X \rightarrow X$ be a measure-preserving dynamical system on the probability space $X = (X, \Sigma, \mu)$, let $p \in [1, \infty)$ and $q \in (1, \infty]$ be its Hölder conjugate, and denote by T_φ the Koopman operator of φ on $L^p(X)$. Then the following are equivalent.*

- (a) *The system φ is strongly mixing.*
- (b) *We have*

$$\lim_{n \rightarrow \infty} \langle T_\varphi^n f, g \rangle = \langle f, \mathbb{1} \rangle \cdot \langle \mathbb{1}, g \rangle$$

for all $f \in L^p(X)$ and all $g \in L^q(X)$.

- (c) *We have $\text{fix}(T_\varphi) = \text{span}\{\mathbb{1}\}$ and $\overline{\text{ran}}(I - T_\varphi) = E_{\text{ws}}$.*
- (d) *We have $E_{\text{rev}} = \text{span}\{\mathbb{1}\}$ and $E_{\text{aws}} = E_{\text{ws}}$.*

In this case, the JdLG-decomposition and the mean ergodic decomposition coincide and take the form

$$L^p(X) = \text{span}\{\mathbb{1}\} \oplus E_{\text{ws}}.$$

In particular, it follows that every strongly mixing measure-preserving system φ is weakly mixing.

Proof. “(a) \Leftrightarrow (b)”: This is [EFHN15, Theorem 9.4].

“(a) \Leftrightarrow (c)”: This is easily deduced from [EFHN15, Theorem 9.6 (i) \Leftrightarrow (v)].

“(a) \Rightarrow (d)”: That $E_{\text{rev}} = \text{span}\{\mathbb{1}\}$ follows from the fact that every strongly mixing system is weakly mixing and Proposition 2.2.10 (a) \Rightarrow (d). That $E_{\text{aws}} = E_{\text{ws}}$ follows

from [EFHN15, Theorem 9.6 (i) \Rightarrow (iv)], since any function $f \in E_{\text{aws}}$ satisfies $\int_X f \, d\mu = 0$.

“(d) \Rightarrow (a)”: This follows from [EFHN15, Theorem 9.6 (vi) \Rightarrow (i)]. \square

Remark 2.2.14. Historically, the question whether there exist weakly mixing systems that are not strongly mixing was not answered by providing a concrete example. Instead, the following category argument was used to show the existence of such systems.

If we consider the set G to be the set of all invertible measure-preserving mappings $\varphi: [0, 1] \rightarrow [0, 1]$, where we endow $[0, 1]$ with its standard Borel- σ -algebra $\mathcal{B} := \mathcal{B}([0, 1])$ and the standard Lebesgue measure λ on $[0, 1]$. We endow G with the topology induced by the family of pseudometrics $\{d_A : A \in \mathcal{B}\}$, where $d_A(\varphi_1, \varphi_2) := \mu(\varphi_1^{-1}(A) \Delta \varphi_2^{-1}(A))$. This topology is called the *weak topology*. A moment’s thought reveals that in the weak topology a net of mappings $(\varphi_\lambda)_{\lambda \in \Lambda}$ in G converges to $\varphi \in G$ if and only if the net of Koopman operators $(T_{\varphi_\lambda})_{\lambda \in \Lambda}$ on $L^2(X)$ converges to T_φ in the strong operator topology. It follows that, since $([0, 1], \mathcal{B}, \lambda)$ is a standard Lebesgue space, the space G is metrizable. It can be shown [DNP87, Discussion X.D.3] that G is complete.

The following has been shown:

- (i) In [Hal44] it has been shown that the set of all weakly mixing maps φ is of second Baire category in G .
- (ii) In [Rok48] it was proved that the set of all strongly mixing maps φ is of first Baire category in G .

Therefore, there is a system that is weakly mixing but not strongly mixing.

An explicit example of a weakly mixing system is given in the main result of [Cha67]. Note that Chacon refers to the weakly mixing property as “only having continuous spectrum and no roots of all orders ≥ 2 ”. Having no roots of order ≥ 2 refers to the fact that a weakly mixing Koopman operator has no eigenvalues that are roots of unity of order greater than 2; in fact, it has no eigenvalues other than 1. The construction uses a stacking method and is rather involved. For a more efficient display of a counterexample we refer the interested reader to [DNP87, Theorem X.9].

2.3 From Topological to Measure-Preserving Dynamics

In order to find examples of measure-preserving systems, let us return to topological dynamics once again. Suppose that Ω is a completely regular Hausdorff space and $\varphi: \Omega \rightarrow \Omega$ is continuous. Consider its associated Perron–Frobenius operator

$$P_\varphi: M(\Omega) \rightarrow M(\Omega), \quad \mu \mapsto \varphi_*\mu,$$

where $\varphi_*\mu$ is the *pushforward measure* of μ with respect to φ defined by $(\varphi_*\mu)(A) = \mu(\varphi^{-1}(A))$ for all Borel sets $A \subseteq \Omega$.

If $\mu \in M(\Omega)$ is a non-zero fixed point of P_φ , then

$$\mu(\varphi^{-1}(A)) = \mu(A)$$

holds for all Borel sets $A \subseteq \Omega$. From the definition of the modulus in Definition 1.1.14 and the transformation theorem for measures it can be shown that

$$|\mu|(\varphi^{-1}(A)) = |\mu|(A)$$

for all Borel sets $A \subseteq \Omega$. So after rescaling we can assume that μ is a probability measure. It follows that $\varphi: \Omega \rightarrow \Omega$ is a measure-preserving system on the probability space $X = (\Omega, \mathcal{B}(\Omega), \mu)$.

Let

$$I_\mu := \{f \in C_b(\Omega) : \langle |f|, |\mu| \rangle_{C_b(\Omega), M(\Omega)} = 0\}$$

and notice that I_μ is a non-trivial $\sigma(C_b(\Omega), M(\Omega))$ -closed lattice ideal, so by Proposition 1.2.27 there exists a non-empty closed subset $M \subseteq \Omega$ such that I_μ coincides with the vanishing ideal $I_M = \{f \in C_b(\Omega) : \forall x \in M : f(x) = 0\}$. We call M the *support of μ* and denote it by $\text{supp}(\mu)$.

Definition 2.3.1. Let Ω be a completely regular Hausdorff space, $\varphi: \Omega \rightarrow \Omega$ be continuous and $\mu \in \text{fix}(P_\varphi)$ be a probability measure. Then (φ, μ) is called a *topological measure-preserving system* on the probability space $(\Omega, \mathcal{B}(\Omega), \mu)$.

If $\text{supp}(\mu) = \Omega$, we call the topological measure-preserving system *faithful*.

A moment's thought reveals that $\varphi(\text{supp}(\mu)) = \text{supp}(\mu)$. Hence, when we restrict φ to $\text{supp}(\mu)$, we always obtain a faithful topological measure-preserving system.

Unfortunately, the fixed space $\text{fix}(P_\varphi)$ of the Perron–Frobenius operator can be trivial as the following example shows.

Example 2.3.2. Let $\Omega = \mathbb{Z}$ and $\varphi: \Omega \rightarrow \Omega$ be defined as the shift $\varphi(k) = k + 1$. Suppose that $\mu \in \text{fix}(P_\varphi)$ and assume without loss of generality that μ is a probability measure. By countability of Ω there exists $k \in \mathbb{Z}$ with $\mu(\{k\}) > 0$. Then $\mu(\varphi^{-n}(\{k\})) = \mu(\{k\}) > 0$, and therefore,

$$\mu(\{k, k-1, k-2, \dots\}) = \sum_{n \in \mathbb{N}_0} \mu(\varphi^{-n}(\{k\})) = \infty.$$

This is a contradiction.

The lack of an invariant probability measure in $\text{fix}(P_\varphi)$ is an artefact of the non-compactness of Ω . In fact, if Ω is compact, then there always exists an invariant probability measure in $\text{fix}(P_\varphi)$ as the following theorem shows. We provide a direct functional analytic proof, although it can be derived from Markov–Kakutani's fixed point theorem [EFHN15, Theorems 10.1 and 10.2].

Theorem 2.3.3 (Krylov–Bogoliubov). *Let K be a compact Hausdorff space and $\varphi: K \rightarrow K$ be continuous. For every $0 \neq f \in \text{fix}(T_\varphi)$ there exists a probability measure $\mu \in \text{fix}(P_\varphi)$ such that $\langle f, \mu \rangle_{C(K), M(K)} \neq 0$.*

Proof. Let $f \in \text{fix}(T_\varphi)$ and $\omega \in K$ with $f(\omega) \neq 0$. Consider for $n \in \mathbb{N}$ the sequence $(\mu_n)_{n \in \mathbb{N}}$ defined by the Dirac combs

$$\mu_n := \frac{1}{n} \sum_{k=0}^{n-1} (\varphi^k)_* \delta_\omega = \frac{1}{n} \sum_{k=0}^{n-1} \delta_{\varphi^k(\omega)}$$

for all $n \in \mathbb{N}$. Note that by the Riesz's representation theorem the space of Radon measures $M(K)$ can be identified with the continuous Banach space dual of $C(K)$. So by the Banach–Alaoglu theorem, there exists a subnet $(n_\lambda)_\lambda$ of $(n)_{n \in \mathbb{N}}$ such that $(\mu_{n_\lambda})_\lambda$ converges with respect to the $\sigma(M(K), C(K))$ -topology, which is the weak*-topology on $M(K)$. Denote by $\mu \in M(K)$ its limit.

As $\mu_n \geq 0$ and $\langle \mathbb{1}, \mu_n \rangle = 1$, it follows that $\mu \geq 0$ and $\langle \mathbb{1}, \mu \rangle = 1$. Thus, μ is a probability measure. Moreover,

$$\begin{aligned} P_\varphi \mu_n &= \frac{1}{n} \sum_{k=1}^n \delta_{\varphi^k(\omega)} = -\frac{1}{n} \delta_\omega + \frac{1}{n} \sum_{k=0}^{n-1} \delta_{\varphi^k(\omega)} + \frac{1}{n} \delta_{\varphi^n(\omega)} \\ &= \mu_n + \frac{1}{n} (\delta_{\varphi^n(\omega)} - \delta_\omega). \end{aligned}$$

So by $(M(K), C(K))$ -continuity of P_φ we obtain $P_\varphi \mu = \lim_\lambda P_\varphi \mu_{n_\lambda} = \lim_\lambda \mu_{n_\lambda} + \lim_\lambda \frac{1}{n_\lambda} (\delta_{\varphi^{n_\lambda}(\omega)} - \delta_\omega) = \mu$. This shows that $\mu \in \text{fix}(P_\varphi)$. Finally we note that

$$\begin{aligned} \langle f, \mu \rangle &= \lim_\lambda \langle f, \mu_{n_\lambda} \rangle = \lim_\lambda \frac{1}{n_\lambda} \sum_{k=0}^{n_\lambda-1} \langle f, P_\varphi^k \delta_\omega \rangle \\ &= \lim_\lambda \frac{1}{n_\lambda} \sum_{k=0}^{n_\lambda-1} \langle T_\varphi^k f, \delta_\omega \rangle = \langle f, \delta_\omega \rangle \neq 0. \quad \square \end{aligned}$$

Example 2.3.4. If we consider the rotation mapping from Example 2.1.2 (ii) as a topological dynamical system on the compact Hausdorff space \mathbb{T} , we obtain that the normalized Lebesgue measure \mathbf{m} is the only measure probability measure in $\text{fix}(P_\varphi)$.

Indeed, let $f \in C(\mathbb{T})$ and $\mu \in M_\varphi^1(\mathbb{T})$. Then from Weyl's equidistribution theorem [EFHN15, Theorem 10.20 and Proposition 10.22] and Lebesgue's dominated convergence theorem we obtain

$$\langle f, \mu \rangle = \left\langle f, \frac{1}{n} \sum_{k=0}^{n-1} P_\varphi^k \mu \right\rangle = \int_{\mathbb{T}} \frac{1}{n} \sum_{k=0}^{n-1} f \circ \varphi^k \, d\mu \rightarrow \int_{\mathbb{T}} \int_{\mathbb{T}} f \, d\mathbf{m} \, d\mu = \langle f, \mathbf{m} \rangle$$

as $n \rightarrow \infty$. Hence, it follows that $\mu = \mathbf{m}$.

If $\varphi: K \rightarrow K$ is a continuous mapping on a compact Hausdorff space, then we say that the measures in $\text{fix}(P_\varphi)$ are φ -invariant. Let us denote the non-empty set of invariant probability measures that we endow with the topology inherited from the weak*-star topology on $M(K)$ by $M_\varphi^1(K)$. A simple computation shows that $M_\varphi^1(K)$ is closed, and therefore, by the Banach–Alaoglu theorem, compact. Moreover, it is a convex subset of $M(K)$. Let us further study the structure of $M_\varphi^1(K)$.

Recall that we call a point p in a convex subset A of a vector space an *extreme point* if whenever some points $a, b \in A$ and $t \in (0, 1)$ satisfy $ta + (1-t)b = p$ it

follows that $a = b = p$. It follows from the theorem of Krein–Milman that the set of extreme point $\text{ext}(M_\varphi^1(K))$ is non-empty. It is easy to see that $\text{ext}(M_\varphi^1(K))$ is closed, and thus, also compact Hausdorff. The extremal points of $M_\varphi^1(K)$ have an ergodic-theoretic interpretation [EFHN15, Proposition 10.4].

Proposition 2.3.5. *Let $\varphi: K \rightarrow K$ be continuous on a compact Hausdorff space K . Then the extreme points of $M_\varphi^1(K)$ are exactly measures in $M_\varphi^1(K)$ for which the topological measure-preserving system $\varphi: K \rightarrow K$ on $(K, \mathcal{B}(K), \mu)$ is ergodic.*

This observation allows us to employ Choquet–Bishop–de Leeuw’s existence theorem [Phe01, Sections 4 and 12] in order to deduce the existence of a mapping

$$M_\varphi^1(K) \rightarrow M(\text{ext}(M_\varphi^1(K))), \quad \mu \mapsto \lambda_\mu$$

such that λ_μ is a probability measure and

$$\int_K f \, d\mu = \int_{\text{ext}(M_\varphi^1(K))} \int_{\text{supp}(\nu)} f \, d\nu \, \lambda_\mu(d\nu).$$

holds for all functions $f \in C(K)$. Intuitively, this states that we can write μ as an “infinite” convex combination of ergodic measures of φ .

Let us close with the following simple remark.

Remark 2.3.6 (standard Lebesgue property). If the compact Hausdorff space K is metrizable (or equivalently second-countable) and $\varphi: K \rightarrow K$ continuous, then it follows directly from the definition of a standard probability space that for every $\mu \in M_\varphi^1(K)$ the associated topological measure-preserving system $\varphi: K \rightarrow K$ on $(K, \mathcal{B}(K), \mu)$ is standard.

2.4 Notes

The Koopman operator on $L^2(X)$ has first been described by Bernhard Koopman in [Koo31] for Hamiltonian systems and further studied in [KN32] jointly with John von Neumann.

The treatment of measure-preserving dynamical systems via the measure-algebra homomorphisms and the associated Koopman operators is loosely inspired by [EFHN15] and the ISem28 lecture notes by Henrik Kreidler and Asgar Jamneshan [JK25].

The order of the exposition in Section 2.2 does not coincide with the historical development of ergodic theory. The mean ergodic theorem has first been proved by von Neumann in [Neu32] for unitary operators on Hilbert spaces and later extended to power-bounded operators on reflexive Banach spaces by Kosaku Yoshida and Shizuo Kakutani in [Yos38] and [YK38]. Another exposition where Birkhoff’s pointwise ergodic theorem is treated as a corollary of the ergodic theorem of von Neumann can be found in [Wal82, Corollary 1.14.1]. From the perspective of statistical mechanics, the pointwise ergodic theorem has more importance. It has been proved only shortly after the von Neumann theorem by George Birkhoff in [Bir31].

The treatment of the strongly and weakly mixing properties follows the exposition in [Hal60] and [EFHN15, Chapter 9]. Although the study of mixing properties goes back much further. In the original work of Bernhard Koopman and John von

Neumann [KN32] the property of having continuous spectrum has already been studied. For unitary operators on Hilbert spaces it can be shown that this property is equivalent to weak mixing [EFHN15, Corollary 18.25].

We have put a particular focus on the functional analytic approach to the mixing properties via the JdLG-decomposition. This approach is helpful for our purposes, as it divides the observables into subspaces with distinct asymptotic behaviour, which can be studied separately. We will use the JdLG-decomposition again throughout the rest of the thesis. So understanding mixing properties from this perspective is rather helpful. Originally the JdLG-decomposition has been developed by Jacobs in [Jac56], de Leeuw and Glicksberg in [LG59] and in [LG61]. Generalizations to semigroups of operators can be found in [EFHN15, Chapter 16].

For unitary operators on Hilbert spaces the JdLG-decomposition coincides with the spectral decomposition into the point spectrum and the continuous spectrum [EFHN15, Corollary 18.18]. We will use this fact in Section 3.2.3.

This chapter presents three notions of entropy for dynamical systems and explains how they relate to each other.

We begin in Section 3.1 with measure-theoretic entropy, defined via finite measurable partitions and their dynamic refinements along an orbit, and we provide a detailed description of why it measures information learned while looking at trajectories through the resolution of a given partition.

We then introduce Voiculescu’s approximation entropy in Section 3.2 for unitary operators, and in particular for Koopman operators. This unitary invariant measures the asymptotic dimension required to approximate finite sets of orbit snapshots to a prescribed tolerance. In Section 3.2.1 we record some of its core properties.

In Section 3.2.2 a newly developed quantitative connection between decay of correlation and approximability is obtained through Toeplitz matrix techniques and Fejér-type approximations. More precisely, we show that rates of decay of correlations and mean-ergodic rates yield lower bounds on (fractional) approximation entropy.

We then establish a connection between the spectral theory of unitary operators and the approximation entropy in Section 3.2.3. We recall from [Voi95] that approximation entropy is induced by the absolutely continuous part of the spectrum of such operators.

The two entropy notions are then linked in Section 3.3: We discover that positive measure-theoretic entropy forces a countable Lebesgue component in the spectrum of the Koopman operator, which in turn implies infinite approximation entropy. Conversely, purely reversible dynamics have zero approximation entropy.

Finally, in Section 3.4 we briefly discuss topological entropy and its relation to measure-theoretic entropy via the variational principle, as we will later use both notions in Chapter 5. There we describe both notions as special cases of a Banach lattice entropy.

In Sections 3.1 and 3.4 we recall known notions from ergodic theory. In Section 3.2, the definition of approximation entropy is taken from [Voi95, Section 7] and has been further studied in this thesis. Proposition 3.2.1 (except (viii)), Proposition 3.2.9 and Corollary 3.2.8 are original to this thesis. Section 3.2.2 is also unpublished work nor has it been submitted for publication. The results in Section 3.2.3 as well as Section 3.3 have been published in [HH25].

3.1 Measure-Theoretic Entropy

We begin by describing the classical notion of measure-theoretic entropy for a measure-preserving dynamical system $\varphi: X \rightarrow X$ on a probability space $X = (X, \Sigma, \mu)$.

Static Entropy of Measurable Partitions

Definition 3.1.1. Let $X = (X, \Sigma, \mu)$ be a finite measure space. We call a finite subset $\beta = \{B_1, \dots, B_d\}$ of Σ a Σ -measurable partition of X if

$$\mu\left(X \setminus \bigcup \beta\right) = 0 \quad \text{and} \quad \mu(B_k \cap B_n) = 0$$

for all $k, n \in \{1, \dots, d\}$ with $k \neq n$.

The latter condition states that the elements of the partition must be essentially pairwise disjoint. Note also that we do not bother to cover the entire space X but only everything up to a null set.

A Σ -measurable partition α of X provides information about the probability space $X = (X, \Sigma, \mu)$: we view two points $x, y \in X$ as *distinguishable* if and only if they lie in distinct sets of α . Therefore, given only the information provided by a partition α and an element $x \in X$, the only thing we can know about x is in which set $A \in \alpha$ it is located. In more mathematical terms, we only ever view X through the observable

$$\bar{\alpha}: X \rightarrow \alpha, \quad x \mapsto A, \quad \text{such that } x \in A.$$

Next, we want to quantify the information that we obtain from learning that a random point $x \in X$, which is chosen with respect to the distribution of μ , lies in a measurable set $A \subseteq X$. We aim to find a function $I_\mu: \Sigma \rightarrow [0, \infty)$ that returns a quantification of the amount of information we obtain from learning that $x \in A$ for each set $A \in \Sigma$. Such a function should satisfy the following properties:

- (i) The function I_μ depends continuously on the probability $\mu(A)$ of an event $A \in \Sigma$ and only on this probability. In other words, the map I_μ factors through the map $A \mapsto \mu(A)$, i.e., there exists a continuous function $i: [0, 1] \rightarrow [0, \infty)$ such that $i(\mu(A)) = I_\mu(A)$ for all $A \in \Sigma$.
- (ii) The surprisal to learn that x is in A should be higher if $\mu(A)$ is small, as this happens with a lower likelihood. Hence, the quantity of information learned decreases as $\mu(A)$ increases. This means that i is monotonically decreasing.
- (iii) If $\mu(A) = 1$, then learning that $x \in A$ is not surprising, and therefore, we gain no information. In this case we want that $I_\mu(A) = i(\mu(A)) = 0$.
- (iv) If events $A, B \in \Sigma$ are stochastically independent, i.e., $\mu(A \cap B) = \mu(A) \cdot \mu(B)$, we want the quantity of information that we gain from learning that $x \in A \cap B$ and $x \in B$ to be the sum of information that we gain from learning that $x \in A$ and $x \in B$. In mathematical terms, we require that for independent sets $A, B \in \Sigma$ we have $i(\mu(A) \cdot \mu(B)) = i(\mu(A \cap B)) = i(\mu(A)) + i(\mu(B))$.

To summarize, we seek a continuous monotonically decreasing function $i: [0, 1] \rightarrow [0, \infty)$ that satisfies $i(1) = 0$ and $i(a \cdot b) = i(a) + i(b)$ for all $a, b \in [0, 1]$. The latter is the fundamental functional equation for the logarithm. Thus, up to a choice in basis for the logarithm, the only such function to satisfy the above criteria is $i(a) = -\log(a)$, and therefore,

$$I_\mu: \Sigma \rightarrow [0, \infty), \quad A \mapsto -\log(\mu(A)),$$

which we call the *information content* or *surprisal function* of X .

For a finite Σ -measurable partition α the composition $I_\mu \circ \bar{\alpha}: X \rightarrow [0, \infty)$ assigns to each $x \in X$ the information content that we gain from learning which set $A \in \alpha$ contains x . We call the expectation of the information content the μ -entropy of α , denoted by

$$H_\mu(\alpha) = \int_X I_\mu \circ \bar{\alpha} d\mu = \sum_{A \in \alpha} -\mu(A) \log(\mu(A)).$$

Dynamic Entropy of Measure-Preserving Systems For two finite Σ -measurable partitions α and β of X , we set

$$\alpha \vee \beta := \{A \cap B : A \in \alpha, B \in \beta\}$$

to be the *common refinement* of α and β . Notice that \vee is associative and commutative, so we may use the shorthand

$$\bigvee_{k=1}^n \alpha_k := \alpha_1 \vee \cdots \vee \alpha_n$$

for a finite sequence of finite Σ -measurable partitions $\alpha_1, \dots, \alpha_n$.

For the rest of this section let us suppose that φ is a measure-preserving mapping on the probability space $X = (X, \Sigma, \mu)$ such that the map $\varphi: X \rightarrow X$ is invertible μ -almost everywhere (or equivalently, that there exists a Σ -measurable mapping $\varphi^{-1}: X \rightarrow X$ such that the equalities $\varphi^{-1} \circ \varphi = \varphi \circ \varphi^{-1} = \text{id}_X$ hold μ -almost surely).

Let α be a finite Σ -measurable partition of X and $x_0 \in X$ and $n, m \in \mathbb{Z}$ with $n \leq m$. If we look at the trajectory $\varphi^n(x_0), \varphi^{n+1}(x_0), \dots, \varphi^m(x_0)$ through the lens of the observable $\bar{\alpha}$, we notice that for x_0 we learn that

$$\varphi^n(x_0) \in A_n, \dots, \varphi^m(x_0) \in A_m \tag{3.1}$$

for some sequence of sets $A_n, \dots, A_m \in \alpha$. In other words, we learn that

$$x_0 \in \varphi^{-n}(A_n) \cap \cdots \cap \varphi^{-m}(A_m).$$

The collection of all such sets can be described in the language of common refinements of partitions: the *dynamic refinement of α and φ along $[n, m] \subseteq \mathbb{Z}$* is given by

$$\varphi_n^m \alpha := \bigvee_{k=n}^m \varphi^{-k}(\alpha),$$

where $\varphi^k(\alpha) := \{\varphi^k(A) : A \in \alpha\}$ is again a finite Σ -measurable partition of X . It follows that evaluating the observable $\bar{\alpha}$ for all points in the trajectory (3.1) is equivalent to evaluating the observable $\overline{\varphi_n^m \alpha}$ at x_0 .

The entropy of a measure-preserving system $\varphi: X \rightarrow X$ with respect to α is defined to be the exponential rate of growth of expected surprisal that we experience when

we observe a trajectory of increasing lengths. In more mathematical terms, we define the *measure-theoretic entropy* of φ on a finite Σ -measurable partition α to be

$$h_\mu(\varphi, \alpha) := \lim_{m \rightarrow \infty} \frac{1}{m} H_\mu(\varphi_0^{m-1} \alpha). \quad (3.2)$$

It becomes the *measure-theoretic entropy* of φ by taking the supremum over all finite and Σ -measurable partitions α , i.e.,

$$h_\mu(\varphi) := \sup_\alpha h_\mu(\varphi, \alpha).$$

The existence of the limit in (3.2) follows from Fekete's lemma [Wal82, Corollary 4.9.1] or more generally from the Ornstein–Weiss lemma [LW00, Theorem 6.1]. Moreover, if φ is invertible, then it does not matter whether we take refinements forward or backward in time; or in other words, the measure-theoretic entropy can also be defined by

$$h_\mu(\varphi) = \sup_\alpha \lim_{n \rightarrow \infty} \frac{1}{n} h_\mu(\varphi_{n-1}^0 \alpha).$$

This is proved by noticing that $h_\mu(\varphi) = h_\mu(\varphi^{-1})$ (see, e.g., [DNP87, Proposition XII.4 (v)] or [Wal82, Theorem 4.12 (vii)]).

Conditional Entropy and Measurable Subspaces Assume we are given two finite Σ -measurable partitions α and β of X and suppose we can already distinguish points according to the partition β . If we additionally learn to distinguish points according to the partition α , we will now be able to distinguish the points according to the common refinement $\alpha \vee \beta$, or in other words, the quantity of information that we have now is $H_\mu(\alpha \vee \beta)$. To quantify what we have learned from the additional information provided by α , we will use the quantity

$$H_\mu(\alpha \mid \beta) := H_\mu(\alpha \vee \beta) - H_\mu(\beta)$$

called the *conditional entropy* of α given β . A simple calculation shows that

$$\begin{aligned} H_\mu(\alpha \mid \beta) &= - \sum_{\substack{B \in \beta \\ \mu(B) \neq 0}} \mu(B) \sum_{A \in \alpha} \frac{\mu(A \cap B)}{\mu(B)} \log \frac{\mu(A \cap B)}{\mu(B)} \\ &= - \sum_{\substack{A, B \\ \mu(B) \neq 0}} \mu(A \cap B) \log \frac{\mu(A \cap B)}{\mu(B)}. \end{aligned}$$

To each Σ -measurable partition α of X , we may associate a vector space of functions that are “ α -measurable”, by setting

$$L^2(X \mid \alpha) := \overline{\text{span}}\{\mathbb{1}_A : A \in \alpha\},$$

where the closure is taken in $L^2(X)$. This space contains all the functions $f \in L^2(X)$ that are measurable with respect to the sub- σ -algebra $\Sigma(\alpha) \subseteq \Sigma$ generated by α . Furthermore, $L^2(X \mid \alpha)$ forms a Banach sublattice of $L^2(X)$ (cf. [Sch74, Proposition III.11.2] or [EFHN15, Proposition 13.19 b]).

3.2 Approximation Entropy

We introduce another entropy concept for an invertible measure-preserving system $\varphi: X \rightarrow X$ on a probability space $X = (X, \Sigma, \mu)$. This entropy concept utilizes the representation of φ as a unitary operator on a Hilbert space via the Koopman representation T_φ on $L^2(X)$. Therefore we will introduce the approximation entropy for the more general class of unitary operators and follow [Voi95, Section 7]. Much of the results presented in Sections 3.2.1 and 3.2.3 has previously been published in the article [HH25] by Hauser and the present author. Section 3.2.2 contains new results.

3.2.1 Definition and Basic Properties

Let H be a Hilbert space, denote by $\mathcal{F}_{\text{fin}}(H)$ the space of all finite-dimensional subspaces $F \subseteq H$ and by $\mathcal{P}_{\text{orth}}(H)$ the space of all orthogonal projections $P: H \rightarrow H$. Consider for $\delta > 0$ the quantity measuring the complexity of approximating a finite point set $\mathcal{O} \subseteq H$ by a finite-dimensional subspace $F \subseteq H$, which is given by

$$\begin{aligned} H_{\text{apr}}(\mathcal{O}, \delta) &:= \min\{\dim F : F \in \mathcal{F}_{\text{fin}}(H), \forall x \in \mathcal{O} : \exists y \in F : \|x - y\| < \delta\} \\ &= \min\{\dim(PH) : P \in \mathcal{P}_{\text{orth}}(H), \forall x \in \mathcal{O} : \|Px - x\| < \delta\}. \end{aligned}$$

Considering a unitary operator $T: H \rightarrow H$, a standard subadditivity argument and Fekete's lemma yield the existence of the following limit. For details see [Voi95, Proposition 7.13] or Proposition 3.2.1 (viii) below. We denote

$$h_{\text{apr}}(T, \mathcal{O}, \delta) := \lim_{n \rightarrow \infty} \frac{1}{n} H_{\text{apr}}\left(\bigcup_{k=0}^{n-1} T^k \mathcal{O}, \delta\right) \quad (3.3)$$

and define the *approximation entropy of \mathcal{O} with respect to T* by

$$h_{\text{apr}}(T, \mathcal{O}) := \sup_{\delta > 0} h_{\text{apr}}(T, \mathcal{O}, \delta).$$

For a (not necessarily T -invariant) closed linear subspace $F \subseteq H$ we define the *approximation entropy of F with respect to T* by

$$h_{\text{apr}}(T, F) := \sup_{\substack{\mathcal{O} \subseteq F \\ \text{finite}}} h_{\text{apr}}(T, \mathcal{O}). \quad (3.4)$$

The *approximation entropy of T* is given by $h_{\text{apr}}(T) := h_{\text{apr}}(T, H)$.

We begin by proving some properties of the approximation entropy that will aid in proving our main results in this chapter later. For further properties we refer to [Voi95, Section 7]. Note that (viii) can be found as [Voi95, Proposition 7.13].

Proposition 3.2.1. *Let $T: H \rightarrow H$ be a unitary operator on a Hilbert space H . The following assertions are true.*

- (i) *For each finite $\mathcal{O}_1, \mathcal{O}_2 \subseteq H$, $\delta > 0$ we have $H_{\text{apr}}(\mathcal{O}_1 \cup \mathcal{O}_2, \delta) \leq H_{\text{apr}}(\mathcal{O}_1, \delta) + H_{\text{apr}}(\mathcal{O}_2, \delta)$.*
- (ii) *Let $\mathcal{O}_1, \mathcal{O}_2 \subseteq H$ be finite with $\mathcal{O}_1 \subseteq \mathcal{O}_2$ and $\delta > 0$. Then $H_{\text{apr}}(\mathcal{O}_1, \delta) \leq H_{\text{apr}}(\mathcal{O}_2, \delta)$, and consequently, $h_{\text{apr}}(T, \mathcal{O}_1, \delta) \leq h_{\text{apr}}(T, \mathcal{O}_2, \delta)$.*

- (iii) Let $\mathcal{O} \subseteq H$ be finite and $\delta_1, \delta_2 > 0$ with $\delta_1 \leq \delta_2$. Then $H_{\text{apr}}(\mathcal{O}, \delta_1) \geq H_{\text{apr}}(\mathcal{O}, \delta_2)$, and consequently, $h_{\text{apr}}(T, \mathcal{O}, \delta_1) \geq h_{\text{apr}}(T, \mathcal{O}, \delta_2)$.
- (iv) For every finite subset $\mathcal{O} \subseteq H$ and $\delta > 0$ we have $H_{\text{apr}}(\mathcal{O}, \delta) \leq |\mathcal{O}|$ and $h_{\text{apr}}(T, \mathcal{O}) \leq |\mathcal{O}|$.
- (v) For $M, N \in \mathbb{N}$ with $M \leq N$, finite $\mathcal{O} \subseteq H$ and $\delta > 0$ we have $H_{\text{apr}}(\mathcal{O}, \delta) = H_{\text{apr}}(T^k \mathcal{O}, \delta)$ for all $k \in \mathbb{Z}$ and

$$h_{\text{apr}}(T, \mathcal{O}, \delta) = h_{\text{apr}}\left(\bigcup_{k=M}^N T^k \mathcal{O}, \delta\right).$$

In particular, it follows that $h_{\text{apr}}(T, F) = h_{\text{apr}}\left(T, \bigcup_{k=M}^N T^k F\right)$ for every closed linear subspace $F \subseteq H$.

- (vi) Let $\mathcal{O}_1, \mathcal{O}_2 \subseteq H$ be finite and $\delta, \varepsilon > 0$ such that $\mathcal{O}_1 \subseteq B_{<\varepsilon}(\mathcal{O}_2)$. Then $h_{\text{apr}}(T, \mathcal{O}_1, \delta + \varepsilon) \leq h_{\text{apr}}(T, \mathcal{O}_2, \delta)$.
- (vii) If \mathcal{O}_1 and \mathcal{O}_2 are finite subsets of H with $\mathcal{O}_1 \subseteq \text{span } \mathcal{O}_2$, then

$$h_{\text{apr}}(T, \mathcal{O}_1) \leq h_{\text{apr}}(T, \mathcal{O}_2).$$

- (viii) The limit in (3.3) exists.

Proof. “(i)”: Let $\delta > 0$. For $i \in \{1, 2\}$, let $F_i \in \mathcal{F}_{\text{fin}}(H)$ be such that for each $x \in \mathcal{O}_i$ there exists $y \in F_i$ such that $\|x - y\| < \delta$. Set $F := \text{span}(F_1 \cup F_2)$. Then $\dim(F) \leq \dim(F_1) + \dim(F_2)$ and for all $x \in \mathcal{O}_1 \cup \mathcal{O}_2$ there exists $y \in F$ such that $\|x - y\| < \delta$. This shows the subadditivity claim. The existence of the limit now follows from Fekete’s lemma [Wal82, Theorem 4.9].

“(ii)”: This follows immediately from the definitions.

“(iii)”: This follows immediately from the definitions.

“(iv)”: The first inequality follows from the fact that $F = \text{span}(\mathcal{O})$ satisfies the condition of the set over which we minimize in the definition of $H_{\text{apr}}(\mathcal{O}, \delta)$. The second inequality follows by noting that $h_{\text{apr}}(T, \mathcal{O}, \delta) = \lim_{n \rightarrow \infty} \frac{1}{n} H_{\text{apr}}(T, \mathcal{O}, \delta) \leq |\mathcal{O}|$ and by taking the supremum over δ .

“(v)”: It follows readily from the fact that T is unitary that $H_{\text{apr}}(\mathcal{O}, \delta) = H_{\text{apr}}(T^k \mathcal{O}, \delta)$ for all $k \in \mathbb{Z}$. Thus, we have

$$\begin{aligned} h_{\text{apr}}\left(T, \bigcup_{k=M}^N T^k \mathcal{O}, \delta\right) &= \lim_{n \rightarrow \infty} \frac{1}{n} H_{\text{apr}}\left(\bigcup_{\bar{k}=0}^{n-1} \bigcup_{k=M}^N T^{k+\bar{k}} \mathcal{O}, \delta\right) \\ &= \lim_{n \rightarrow \infty} \frac{n + (N - M) + 1}{n} \cdot \frac{1}{n + (N - M) + 1} H_{\text{apr}}\left(\bigcup_{k=0}^{n+(N-M)} T^k \mathcal{O}, \delta\right) \\ &= h_{\text{apr}}(T, \mathcal{O}, \delta). \end{aligned}$$

Notice that the inequality “ \leq ” of the claimed consequence is trivial. To see the converse inequality let $\mathcal{O} \subseteq \bigcup_{k=M}^N T^k F$ be finite and note that there exists a finite $\mathcal{O}_0 \subseteq F$ such that $\mathcal{O} \subseteq \bigcup_{k=M}^N T^k \mathcal{O}_0$. Thus,

$$h_{\text{apr}}(T, F) \geq h_{\text{apr}}(T, \mathcal{O}_0) = h_{\text{apr}}\left(T, \bigcup_{k=M}^N T^k \mathcal{O}_0\right) \geq h_{\text{apr}}(T, \mathcal{O}).$$

Taking the supremum over all $\mathcal{O} \subseteq \bigcup_{k=M}^N T^k F$ yields the claim.

“(vi)”: If $F \in \mathcal{F}_{\text{fin}}(H)$ is such that for all $x_2 \in \mathcal{O}_2$ and all $k \in \{0, \dots, n-1\}$ there exists $y \in F$ such that $\|T^k x_2 - y\| < \delta$, then we have for all $x_1 \in \mathcal{O}_1$ and all $k \in \{0, \dots, n-1\}$ that there exists $x_2 \in \mathcal{O}_2$ and $y \in F$ such that

$$\begin{aligned} \|T^k x_1 - y\| &\leq \|T^k x_1 - T^k x_2\| + \|T^k x_2 - y\| \\ &= \|x_1 - x_2\| + \|T^k x_2 - y\| < \varepsilon + \delta. \end{aligned}$$

Thus, we obtain $h_{\text{apr}}(T, \mathcal{O}_1, \delta + \varepsilon) \leq h_{\text{apr}}(T, \mathcal{O}_2, \delta)$.

“(vii)”: For $x_1 \in \mathcal{O}_1$ we write $x_1 = \sum_{x_2 \in \mathcal{O}_2} \lambda_{x_1, x_2} x_2$ and set

$$\lambda := \max_{x_1 \in \mathcal{O}_1} \sum_{x_2 \in \mathcal{O}_2} |\lambda_{x_1, x_2}|.$$

Let $n \in \mathbb{N}$ and let $F \in \mathcal{F}_{\text{fin}}(H)$ be such that for all $x_2 \in \mathcal{O}_2$ and all $k \in \{0, \dots, n-1\}$ there exists $y_{x_2} \in F$ such that $\|T^k x_2 - y_{x_2}\| < \delta/\lambda$. Let $\tilde{x}_1 \in \bigcup_{k=0}^{n-1} T^k \mathcal{O}_1$, then there exists $x_1 \in \mathcal{O}_1$ and $k \in \{0, \dots, n-1\}$ such that $\tilde{x}_1 = T^k x_1 = \sum_{x_2 \in \mathcal{O}_2} \lambda_{x_1, x_2} T^k x_2$. Let $\tilde{y} := \sum_{x_2 \in \mathcal{O}_2} \lambda_{x_1, x_2} y_{x_2} \in F$. Then

$$\|\tilde{x}_1 - \tilde{y}\| \leq \sum_{x_2 \in \mathcal{O}_2} |\lambda_{x_1, x_2}| \cdot \|T^k x_2 - y_{x_2}\| < \lambda \cdot \frac{\delta}{\lambda} = \delta.$$

Hence, it follows that for all $n \in \mathbb{N}$ we have

$$H_{\text{apr}}\left(\bigcup_{k=0}^{n-1} T^k \mathcal{O}_1, \delta\right) \leq H_{\text{apr}}\left(\bigcup_{k=0}^{n-1} T^k \mathcal{O}_2, \frac{\delta}{\lambda}\right)$$

and the claim follows.

“(i)”: It follows from (i) and (v) that the sequence

$$a_n := H_{\text{apr}}\left(\bigcup_{k=0}^{n-1} T^k \mathcal{O}, \delta\right)$$

is subadditive. Hence, the limit $\lim_{n \rightarrow \infty} \frac{a_n}{n}$ exists by Fekete’s lemma [Wal82, Theorem 4.9]. \square

A subset $\mathcal{O} \subseteq H$ is called a *T-generator*, whenever

$$H = \overline{\text{span}} \bigcup_{k \in \mathbb{Z}} T^k \mathcal{O}.$$

For more information on the approximation entropy we refer to [Voi95, Section 7] and the references therein. The following is a slight generalization of [Voi95, Proposition 7.4]. As no proof was provided there, we prove the result.

Proposition 3.2.2. *Let $T: H \rightarrow H$ be a unitary operator on a Hilbert space H . If $(\mathcal{O}_\lambda)_{\lambda \in \Lambda}$ is an increasing net of finite subsets of H such that $\bigcup_{\lambda \in \Lambda} \mathcal{O}_\lambda$ is a T -generator, then*

$$h_{\text{apr}}(T) = \sup_{\lambda \in \Lambda} h_{\text{apr}}(T, \mathcal{O}_\lambda).$$

Proof. Denote $\mathcal{O} := \bigcup_{\lambda \in \Lambda} \mathcal{O}_\lambda$ and note that \mathcal{O} is a T -generator. Let $\varepsilon > 0$ and choose $\delta > 0$, and let $\tilde{\mathcal{O}}_1 \subseteq H$ be a finite subset such that

$$h_{\text{apr}}(T) - \varepsilon/2 \leq h_{\text{apr}}(T, \tilde{\mathcal{O}}_1, \delta).$$

Since \mathcal{O} is a T -generator, there exists a finite $\tilde{\mathcal{O}}_2 \subseteq \text{span} \bigcup_{k \in \mathbb{Z}} T^k \mathcal{O}$ such that $\tilde{\mathcal{O}}_1 \subseteq B_{\leq \delta/2}(\tilde{\mathcal{O}}_2)$. Hence, by Proposition 3.2.1 (vi) we obtain $h_{\text{apr}}(T, \tilde{\mathcal{O}}_1, \delta) \leq h_{\text{apr}}(T, \tilde{\mathcal{O}}_2, \delta/2)$.

Since $\tilde{\mathcal{O}}_2$ is finite, there exists $N \in \mathbb{N}$ and a finite subset $\Lambda_0 \subseteq \Lambda$ such that $\tilde{\mathcal{O}}_2 \subseteq \text{span} \bigcup_{k=-N}^N \bigcup_{\lambda \in \Lambda_0} T^k \mathcal{O}_\lambda$. As Λ is directed there exists $\lambda_0 \in \Lambda$ such that $\lambda \leq \lambda_0$ for all $\lambda \in \Lambda_0$. Then

$$\tilde{\mathcal{O}}_2 \subseteq \text{span} \bigcup_{k=-N}^N \bigcup_{\lambda \in \Lambda_0} T^k \mathcal{O}_\lambda \subseteq \mathcal{O}_2 \bigcup_{k=-N}^N T^k \mathcal{O}_{\lambda_0}$$

and it follows from Proposition 3.2.1 (ii) and (v) that

$$h_{\text{apr}}(T, \tilde{\mathcal{O}}_2) \leq h_{\text{apr}} \left(T, \bigcup_{k=-N}^N T^k \mathcal{O}_{\lambda_0} \right) = h_{\text{apr}}(T, \mathcal{O}_{\lambda_0}).$$

In summary, we obtain

$$\begin{aligned} h_{\text{apr}}(T) - \varepsilon &\leq h_{\text{apr}}(T, \tilde{\mathcal{O}}_1, \delta) \leq h_{\text{apr}}(T, \tilde{\mathcal{O}}_2, \delta/2) \\ &\leq h_{\text{apr}}(T, \tilde{\mathcal{O}}_2) \leq h_{\text{apr}}(T, \mathcal{O}_{\lambda_0}) \leq \sup_{\lambda \in \Lambda} h_{\text{apr}}(T, \mathcal{O}_\lambda). \end{aligned}$$

As ε was arbitrarily chosen, we obtain $h_{\text{apr}}(T) \leq \sup_{\lambda \in \Lambda} h_{\text{apr}}(T, \mathcal{O}_\lambda)$. The reverse inequality follows directly from the definition, and thus, the equality is proved. \square

As an immediate consequence of Proposition 3.2.2 we observe the following, which can be found as [HH25, Proposition 3.2].

Proposition 3.2.3. *Let $F \subseteq H$ be a finite-dimensional subspace, $\mathcal{O} \subseteq F$ be a finite subset with $F = \text{span}(\mathcal{O})$, and denote $G := \overline{\text{span}}(\bigcup_{k \in \mathbb{Z}} T^k F)$. Then we have*

$$h_{\text{apr}}(T, F) = h_{\text{apr}}(T, \mathcal{O}) = h_{\text{apr}}(T|_G) = h_{\text{apr}}(T, G).$$

Note that in this context \mathcal{O} is a $T|_G$ -generator. Moreover,

$$h_{\text{apr}}(T) = \sup_F h_{\text{apr}}(T, F),$$

where the supremum is taken over all finite-dimensional subspaces F of H .

The following upper bound for the approximation entropy on finite-dimensional subspaces follows easily from Proposition 3.2.1 (iv).

Corollary 3.2.4. *For finite-dimensional subspaces $F \subseteq H$ we have*

$$h_{\text{apr}}(T, F) \leq \dim(F).$$

We will now illustrate how we can use Gram matrices to estimate the approximation entropy in certain cases.

Let $\mathcal{O} = \{x_1, \dots, x_{|\mathcal{O}|}\} \subseteq L^2(X)$ be finite and such that the tuple $(x_1, \dots, x_{|\mathcal{O}|})$ is linearly independent, let $F := \text{span}(\mathcal{O})$ and let $C: F \rightarrow \mathbb{K}^{|\mathcal{O}|}$ be the coefficient mapping $Cx = (\langle x, x_j \rangle)_{j=1}^{|\mathcal{O}|}$, which is also called the *analysis operator* of \mathcal{O} (cf. Section 4.1.2). Then the *Gram matrix* of \mathcal{O} can be expressed as

$$G_{\mathcal{O}} = CC^*: \mathbb{K}^{|\mathcal{O}|} \rightarrow \mathbb{K}^{|\mathcal{O}|}, \quad (a_j)_{j=1}^{|\mathcal{O}|} \mapsto \left(\sum_{j=1}^{|\mathcal{O}|} a_j \langle x_j, x_i \rangle \right)_{i=1}^{|\mathcal{O}|}.$$

We refer to Section 4.1.2 for more information on Gram matrices and their role in frame theory. A simple computation shows that the Gram matrix is Hermitian and positive semidefinite. Moreover, it is invertible, and thus, positive definite, if and only if the vectors $x_1, \dots, x_{|\mathcal{O}|}$ are linearly independent. In matrix representation $G_{\mathcal{O}}$ takes the form

$$G_{\mathcal{O}} = \begin{pmatrix} \|x_1\|^2 & \langle x_2, x_1 \rangle & \cdots & \langle x_{|\mathcal{O}|}, x_1 \rangle \\ \langle x_1, x_2 \rangle & \|x_2\|^2 & \cdots & \langle x_{|\mathcal{O}|}, x_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_1, x_{|\mathcal{O}|} \rangle & \langle x_2, x_{|\mathcal{O}|} \rangle & \cdots & \|x_{|\mathcal{O}|}\|^2 \end{pmatrix} \in \mathbb{K}^{|\mathcal{O}| \times |\mathcal{O}|}.$$

Remark 3.2.5. By the assumed linear independence it follows that C is an isomorphism and the eigenvalues of $G_{\mathcal{O}}$ coincide with the eigenvalues of

$$S := C^*C: F \rightarrow F, \quad x \mapsto \sum_{j=1}^{|\mathcal{O}|} \langle x, x_j \rangle x_j,$$

which is called the *frame operator* of \mathcal{O} (cf. Section 4.1.2).

Indeed, if λ is an eigenvalue of C^*C with eigenvector $x \in H$, then Cx is an eigenvector of $G_{\mathcal{O}} = CC^*$ to the eigenvalue λ . Conversely, let λ be an eigenvalue of $G_{\mathcal{O}}$ to the eigenvector $x \in \mathbb{K}^{|\mathcal{O}|}$. Since the vectors in \mathcal{O} are linearly independent, C is invertible, and thus, $C^{-1}x$ is an eigenvector of C^*C for the eigenvalue λ . Hence, the eigenvalue structure of C^*C and $G_{\mathcal{O}}$ coincide.

Lemma 3.2.6. *Let $\mathcal{O} \subseteq H$ be a finite set of normalized linearly independent vectors in a Hilbert space H . Let $\delta > 0$ and suppose that $P: H \rightarrow H$ is a positive definite and self-adjoint operator such that*

$$\|Px - x\| \leq \delta$$

for all $x \in \mathcal{O}$. Let further $\lambda_1 \geq \dots \geq \lambda_{|\mathcal{O}|}$ be the eigenvalues of $G_{\mathcal{O}}$. Then

$$|\mathcal{O}| \cdot \delta^2 \geq \sum_{k=\dim(PH)+1}^{|\mathcal{O}|} \lambda_k = |\mathcal{O}| - \sum_{k=1}^{\dim(PH)} \lambda_k.$$

Proof. Let Q denote the orthogonal projection onto the subspace $F \subseteq H$ spanned by \mathcal{O} , let $V := QPH \subseteq H$, and let \tilde{P} be the orthogonal projection onto V . Then $\dim(\tilde{P}H) = \dim(V) \leq \dim(PH)$ and, since for $x \in F$ we have $\|x - \tilde{P}x\| \leq \|x - Px\|$ for all $y \in V$, we obtain

$$\|x - \tilde{P}x\| \leq \|x - QPx\| = \|Qx - QPx\| \leq \|Q\| \|x - Px\| < \delta$$

for all $x \in \mathcal{O}$. Hence, by replacing P with \tilde{P} , we may assume that $PH \subseteq F$.

Since $I - P$ is a projection, the eigenvalues $\kappa_1 \geq \dots \geq \kappa_{|\mathcal{O}|}$ of $I - P$ satisfy $\kappa_1 = \dots = \kappa_{|\mathcal{O}| - \dim(PH)} = 1$ and $\kappa_{|\mathcal{O}| - \dim(PH) + 1} = \dots = \kappa_{|\mathcal{O}|} = 0$.

Moreover, from Remark 3.2.5 we learn that $\lambda_1, \dots, \lambda_{|\mathcal{O}|}$ are exactly the eigenvalues of C^*C .

Denote by e_k the k -th unit vector in $\mathbb{K}^{|\mathcal{O}|}$ and note that

$$\begin{aligned} |\mathcal{O}| \delta^2 &\geq \sum_{x \in \mathcal{O}} \|(I - P)x\|^2 = \sum_{k=1}^{|\mathcal{O}|} \|(I - P)C^*e_k\|^2 \\ &= \sum_{k=1}^{|\mathcal{O}|} \langle (I - P)C^*e_k, (I - P)C^*e_k \rangle = \sum_{k=1}^{|\mathcal{O}|} \langle C(I - P)C^*e_k, e_k \rangle \\ &= \text{trace}_{\mathbb{K}^{|\mathcal{O}|}}(C(I - P)C^*) = \text{trace}_F(C^*C(I - P)) \end{aligned}$$

Now a trace inequality [MOA11, Theorem 9.H.1.h on p. 341] for positive semidefinite Hermitian operators, which states that

$$\text{trace}(AB) \geq \sum_{k=1}^n \alpha_k \beta_{n-k+1},$$

holds for every pair of Hermitian positive semidefinite operators $A, B: \mathbb{K}^n \rightarrow \mathbb{K}^n$ with eigenvalues $\alpha_1 \geq \dots \geq \alpha_n$ and $\beta_1 \geq \dots \geq \beta_n$, respectively, yields that

$$\text{trace}_F(C^*C(I - P)) \geq \sum_{k=1}^{|\mathcal{O}|} \lambda_k \cdot \kappa_{|\mathcal{O}| - k + 1} = \sum_{k=\dim(PH)+1}^{|\mathcal{O}|} \lambda_k,$$

which shows the inequality. To see the equality note that

$$\begin{aligned} \sum_{k=\dim(PH)+1}^{|\mathcal{O}|} \lambda_k &= \sum_{k=1}^{|\mathcal{O}|} \lambda_k - \sum_{k=1}^{\dim(PH)} \lambda_k \\ &= \text{trace}(G_{\mathcal{O}}) - \sum_{k=1}^{\dim(PH)} \lambda_k = |\mathcal{O}| - \sum_{k=1}^{\dim(PH)} \lambda_k, \end{aligned}$$

where $\text{trace}(G_{\mathcal{O}}) = |\mathcal{O}|$ holds true, since the vectors in \mathcal{O} are normalized. \square

We use the above lemma to deduce the next proposition that relates the smallest eigenvalue of the Gram matrix $G_{\mathcal{O}}$ to the dimension of a subspace that approximately contains \mathcal{O} .

Proposition 3.2.7. *Let $\mathcal{O} \subseteq H$ be a finite set of normalized linearly independent vectors in a Hilbert space H . Let $\delta > 0$ and suppose that $P: H \rightarrow H$ be a positive definite and self-adjoint operator such that*

$$\|Px - x\| \leq \delta$$

for all $x \in \mathcal{O}$. Denote by $\lambda_{\min}(G_{\mathcal{O}})$ the smallest and by $\lambda_{\max}(G_{\mathcal{O}})$ the largest eigenvalue of $G_{\mathcal{O}}$. Then

$$\dim(PH) \geq \left(1 - \frac{\delta^2}{\lambda_{\min}(G_{\mathcal{O}})}\right) |\mathcal{O}|$$

and

$$\dim(PH) \geq \left(\frac{1 - \delta^2}{\lambda_{\max}(G_{\mathcal{O}})}\right) |\mathcal{O}|.$$

Proof. From Lemma 3.2.6 we obtain that

$$|\mathcal{O}| \cdot \delta^2 \geq \sum_{k=\dim(PH)+1}^{|\mathcal{O}|} \lambda_k \geq (|\mathcal{O}| - \dim(PH)) \cdot \lambda_{\min}(G_{\mathcal{O}}).$$

Rearranging yields the first claim. Also notice that we obtain from Lemma 3.2.6 that

$$|\mathcal{O}| \cdot \delta^2 \geq |\mathcal{O}| - \sum_{k=1}^{\dim(PH)} \lambda_k \geq |\mathcal{O}| - \dim(PH) \cdot \lambda_{\max}(G_{\mathcal{O}}).$$

Rearranging yields the second claim. \square

By noticing that for an orthonormal system $\mathcal{O} \subseteq H$ the Gram matrix $G_{\mathcal{O}}$ is the identity matrix, we obtain the following consequence of Proposition 3.2.7, which can also be found as [Voi95, Lemma 7.8].

Corollary 3.2.8. *Let $\mathcal{O} \subseteq H$ be a finite set of normalized linearly independent vectors in a Hilbert space H and let $\delta > 0$. Then it follows that*

$$H_{\text{apr}}(\mathcal{O}, \delta) \geq \left(1 - \frac{\delta^2}{\lambda_{\min}(G_{\mathcal{O}})}\right) |\mathcal{O}|$$

and

$$H_{\text{apr}}(\mathcal{O}, \delta) \geq \frac{1 - \delta^2}{\lambda_{\max}(G_{\mathcal{O}})} |\mathcal{O}|,$$

where $\lambda_{\min}(G_{\mathcal{O}})$ and $\lambda_{\max}(G_{\mathcal{O}})$ denote the minimal and maximal eigenvalue of $G_{\mathcal{O}}$.

In particular, if \mathcal{O} is a orthonormal system, then

$$H_{\text{apr}}(\mathcal{O}, \delta) \geq |\mathcal{O}|(1 - \delta^2).$$

Let $T: H \rightarrow H$ be a unitary operator on the Hilbert space H . By [EFHN15, Example 16.25] T is Jacob–de Leeuw–Glicksberg decomposable. Recall from Definition 2.2.6 that this means the existence of T -invariant subspaces

$$H = H_{\text{rev}} \oplus H_{\text{aws}}.$$

We recall that $H_{\text{rev}} = \overline{\text{span}}\{x \in H : \exists \lambda \in \mathbb{T} : Tx = \lambda x\}$. It can be easily shown that $T|_{H_{\text{rev}}}$ has no approximation entropy. Note that this is a special case of [Voi95, Lemma 7.7].

Proposition 3.2.9. *Let $T: H \rightarrow H$ be a unitary operator on the Hilbert space H and suppose that $H_{\text{aws}} = \{0\}$. Then $h_{\text{apr}}(T) = 0$.*

Proof. Note that $H = H_{\text{rev}}$ and let

$$\Omega := \{x \in H : \exists \lambda \in \mathbb{T} : Tx = \lambda x, \|x\| = 1\}.$$

Let $(\mathcal{O}_\lambda)_\lambda$ be a monotone net of finite subsets of Ω with $\bigcup_\lambda \mathcal{O}_\lambda = \Omega$. Then clearly, $\bigcup_\lambda \mathcal{O}_\lambda$ is a T -generator. Hence, by Proposition 3.2.2 we obtain $h_{\text{apr}}(T) = \sup_\lambda h_{\text{apr}}(T, \mathcal{O}_\lambda)$. So it suffices to show that $h_{\text{apr}}(T, \mathcal{O}_\lambda) = 0$ for all λ . To this end, fix λ and notice that $F_\lambda := \text{span } \mathcal{O}_\lambda$ is T -invariant. Hence,

$$H_{\text{apr}}\left(\bigcup_{k=0}^{n-1} T^k \mathcal{O}_\lambda, \delta\right) \leq \dim(F_\lambda)$$

for each $\delta > 0$. So $h_{\text{apr}}(T, \mathcal{O}_\lambda, \delta) = 0$ for each $\delta > 0$, which shows the claim. \square

Let us illustrate how one can calculate the approximation entropy of an operator.

Example 3.2.10. (i) Let $X = (\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathfrak{m})$ be the torus \mathbb{T} endowed with the normalized Lebesgue measure. Consider the unitary operator

$$T: L^2(X) \rightarrow L^2(X), \quad (Tf)(z) \mapsto zf(z).$$

Then starting with the constant function $\mathbb{1}$, the trajectory

$$\mathcal{O}_N := \{\mathbb{1}, T\mathbb{1}, \dots, T^{N-1}\mathbb{1}\}$$

consists of the functions $(T^k \mathbb{1})(z) = z^k$, and is therefore an orthonormal system. We obtain from Corollary 3.2.8 that

$$h_{\text{apr}}(T, \{\mathbb{1}\}, \delta) = \lim_{N \rightarrow \infty} \frac{1}{N} H_{\text{apr}}(\mathcal{O}_N, \delta) \geq (1 - \delta^2).$$

So $h_{\text{apr}}(T, \{\mathbb{1}\}) \geq 1$. Using Proposition 3.2.3 and Corollary 3.2.4 we obtain $h_{\text{apr}}(T, \{\mathbb{1}\}) = h_{\text{apr}}(T, \text{span}\{\mathbb{1}\}) \leq 1$.

Now note that $\text{span} \bigcup_{k \in \mathbb{Z}} T^k \{\mathbb{1}\}$ is a dense subspace of $L^2(X)$. Hence, $\{\mathbb{1}\}$ is a T -generator. It follows from Proposition 3.2.2 that $h_{\text{apr}}(T) = h_{\text{apr}}(T, \{\mathbb{1}\}) = 1$.

- (ii) Recall the irrational rotation system from Example 2.1.2 (ii). Let T_φ be the Koopman operator on $H = L^2(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathfrak{m})$. Then it follows from arguments in Example 2.2.11 that

$$H_{\text{rev}} = \overline{\text{span}}\{f_k : k \in \mathbb{Z}\} = H,$$

where $f_k: \mathbb{T} \rightarrow \mathbb{K}$, $\lambda \rightarrow \lambda^k$ are eigenfunctions to the eigenvalue $e^{2\pi k\alpha}$. Hence, $H_{\text{aws}} = \{0\}$, and thus, it follows from Proposition 3.2.9 that $h_{\text{apr}}(T_\varphi) = 0$.

3.2.2 Lower Bounds for the Approximation Entropy

We will now provide assumptions under which we can deduce the presence of approximation entropy $h_{\text{apr}}(T)$. One assumption that suffices is that for some vector $x \in H$ we have fast enough *decay of correlation*, i.e., the sequence

$$\langle x, Tx \rangle, \langle x, T^2x \rangle, \langle x, T^3x \rangle, \dots$$

converges to 0 sufficiently fast.

We recall from Definition 2.2.6 that the *almost weakly stable part* of T is given by

$$H_{\text{aws}} := \left\{ x_1 \in H : \forall x_2 \in H : \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} |\langle x_1, T^k x_2 \rangle| = 0 \right\}.$$

For vectors in H_{aws} we can prove the following.

Theorem 3.2.11. *Let $T: H \rightarrow H$ be a unitary operator on a Hilbert space H . Let $x \in H_{\text{aws}}$ be normalized and such that*

$$C := \sum_{k=1}^{\infty} |\langle x, T^k x \rangle| < \infty. \tag{3.5}$$

Then $h_{\text{apr}}(T, \{x\}) \geq \frac{1}{1+2C}$.

Remark 3.2.12. (i) Note that the assumption (3.5) implies by Proposition 4.1.15 that

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} T^k x \right\|^2 \leq \frac{2C}{n}.$$

(ii) If we have assumption (3.5) for all $x \in H_{\text{aws}}$, then it follows from the polarization identity applied to the sesquilinear form $a(x_1, x_2) := \langle x_1, T x_2 \rangle$ defined for $x_1, x_2 \in H_{\text{aws}}$ that

$$\sum_{k=1}^{\infty} |\langle x_1, T^k x_2 \rangle| \leq \frac{1}{4} \sum_{k=1}^{\infty} \sum_{l=0}^3 |\langle x_1 + i^l x_2, T^k (x_1 + i^l x_2) \rangle| < \infty$$

in the complex case. Hence, it follows from [Wei89, Theorem 1] that $r(T|_{H_{\text{aws}}}) < 1$. As $T|_{H_{\text{aws}}}$ is unitary, this means that $H_{\text{aws}} = \{0\}$. So whenever $H_{\text{aws}} \neq \{0\}$ one can not expect the assumption (3.5) to be satisfied for all $x \in H_{\text{aws}}$.

In the proof of Theorem 3.2.11 we will use some theory about Toeplitz matrices. Let $n \in \mathbb{N}$. A matrix $A \in \mathbb{K}^{n \times n}$ is called *Toeplitz* if there exists a sequence $(a_k)_{k=-(n-1)}^{n-1} \in \mathbb{K}^{2n-1}$ such that

$$A = (a_{i-j})_{i,j=1}^n = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{-1} & a_0 & \cdots & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{-(n-1)} & a_{-(n-2)} & \cdots & a_0 \end{pmatrix}$$

Conversely, given a sequence $(a_k)_{k \in \mathbb{Z}}$ in \mathbb{K} we can associate a sequence of Toeplitz matrices $(A_N)_{N \in \mathbb{N}}$ with $A_N = (a_{i-j})_{i,j=1}^N \in \mathbb{K}^{N \times N}$. The matrices are Hermitian if and only if $a_k = \overline{a_{-k}}$ for all $k \in \mathbb{N}_0$ (which implies $a_0 \in \mathbb{R}$).

If $(a_k)_{k \in \mathbb{Z}}$ is absolutely summable, i.e., $\sum_{k \in \mathbb{Z}} |a_k| < \infty$, then the sequence of Toeplitz matrices is said to be of *Wiener class*. In this case, we can associate to the sequences a function

$$f: \mathbb{T} \rightarrow \mathbb{C}, \quad z \mapsto \sum_{k=-\infty}^{\infty} a_k z^k$$

called the *symbol* of the Toeplitz sequence. It follows readily from the summability assumption on $(a_k)_{k \in \mathbb{Z}}$ that the series converges uniformly and that f is bounded and continuous. A moment's thought reveals that the Toeplitz matrices A_N are Hermitian if and only if f is real-valued. We can recover the Toeplitz matrix from a symbol f by a simple Fourier transform as

$$a_k = \int_{\mathbb{T}} z^{-k} f(z) \mathbf{m}(dz)$$

for all $k \in \mathbb{Z}$, where \mathbf{m} denotes the normalized Lebesgue measure on \mathbb{T} .

For a real-valued symbol, we denote by

$$M_f := \operatorname{ess\,sup}_{z \in \mathbb{T}} f(z), \quad m_f := \operatorname{ess\,inf}_{z \in \mathbb{T}} f(z) \quad (3.6)$$

the *essential supremum* and *essential infimum* of f .

Suppose from now on that the sequence $(A_N)_{N \in \mathbb{N}}$ is Hermitian and denote by $\lambda_{1,N} \geq \dots \geq \lambda_{N,N}$ the eigenvalues of A_N for each $N \in \mathbb{N}$. We can use the symbol of $(a_k)_{k \in \mathbb{Z}}$ to bound the eigenvalues of the corresponding Toeplitz matrices [Gra05, Lemma 4.1].

Lemma 3.2.13. *Let $(a_k)_{k \in \mathbb{Z}}$ be an absolutely summable sequence in \mathbb{K} with $a_k = \overline{a_{-k}}$ for all $k \in \mathbb{N}_0$ and let $(A_N)_{N \in \mathbb{N}}$ be the associated sequence of Toeplitz matrices. Denote by $\lambda_{1,N} \geq \dots \geq \lambda_{N,N}$ the eigenvalues of A_N for each $N \in \mathbb{N}$. Then*

$$m_f \leq \lambda_{k,N} \leq M_f$$

for all $k \in \{1, \dots, N\}$ and all $N \in \mathbb{N}$, where m_f and M_f are given by (3.6).

Proof of Theorem 3.2.11. Note that T is JdLG-decomposable [EFHN15, Theorem 16.19 2)], and therefore H decomposes into two invariant subspaces $H = H_{\text{rev}} \oplus H_{\text{aws}}$ as in Definition 2.2.6.

By assumption we have $x \in H_{\text{aws}}$. Now assume for a contradiction that for some $N \in \mathbb{N}$ the vectors in

$$\mathcal{O}_N := \{x, Tx, \dots, T^{N-1}x\}$$

are not linearly independent and let N be the smallest such number. Then there exist non-zero coefficients $a_0, \dots, a_{N-1} \in \mathbb{K}$ such that $a_0x + a_1Tx + \dots + a_{N-1}T^{N-1}x = 0$. By the choice of N we obtain that $a_{N-1} \neq 0$. By rearranging it follows that $T^{N-1}x \in \operatorname{span}(\{x, Tx, \dots, T^{N-2}x\})$, and thus,

$$T^N x \in \operatorname{span}(\{Tx, \dots, T^{N-1}x\}) \subseteq \operatorname{span}(\{x, \dots, T^{N-2}x\}).$$

By induction we obtain that $T^k x \in \operatorname{span}(\{x, \dots, T^{N-2}x\})$ for all $k \geq 0$. This shows that $\{T^k x : k \geq 0\}$ is a bounded subset of a finite dimensional space, and hence, it is relatively compact. Now [EFHN15, Theorem 16.33 a)] implies that $x \in H_{\text{rev}}$. This is a contradiction to $x \in H_{\text{aws}}$, since $x \neq 0$.

For each $N \in \mathbb{N}$ and $\delta > 0$ we can now apply Corollary 3.2.8 and obtain

$$H_{\text{apr}}(\mathcal{O}_N, \delta) \geq \frac{1 - \delta^2}{\lambda_{\max}(G_{\mathcal{O}_N})} N. \quad (3.7)$$

Denote $a_k := \langle x, T^k x \rangle$ for all $k \in \mathbb{Z}$ and let $(A_N)_{N \in \mathbb{N}}$ be the sequence of Toeplitz matrices corresponding to $(a_k)_{k \in \mathbb{Z}}$. Then we notice that $A_N = G_{\mathcal{O}_N}$ is Hermitian and obtain from Lemma 3.2.13 that $\lambda_{\max}(G_{\mathcal{O}_N}) \leq M_f$.

Note that $M_f \leq \sum_{k=-\infty}^{\infty} |a_k| = 1 + 2C$, since $a_k = \overline{a_{-k}}$ for all $k \in \mathbb{N}_0$, and thus, $\lambda_{\max}(G_{\mathcal{O}_N}) \leq 1 + 2C$. Then (3.7) yields

$$h_{\text{apr}}(T, \{x\}, \delta) = \lim_{N \rightarrow \infty} \frac{1}{N} H_{\text{apr}}(\mathcal{O}_N, \delta) \geq \frac{1 - \delta^2}{1 + 2C}.$$

Finally we obtain $h_{\text{apr}}(T, \{x\}) = \sup_{\delta > 0} h_{\text{apr}}(T, \{x\}, \delta) \geq \frac{1}{1 + 2C}$ as claimed. \square

Remark 3.2.14. We can more easily obtain the inequality $\lambda_{\max}(G_{\mathcal{O}_N}) \leq 1 + 2C$ by noting that for the spectral radius of G we have

$$r(G_{\mathcal{O}_N}) \leq \|G_{\mathcal{O}_N}\|_{\infty} \leq \sum_{k \in \mathbb{Z}} |\langle x, T^k x \rangle|,$$

as $\|G_{\mathcal{O}_N}\|_{\infty}$ is the maximum of the absolute sums of each line of the Gram matrix $G_{\mathcal{O}_N}$. As we will use Toeplitz theory in the proofs of Proposition 3.2.17 and Theorem 3.2.18, we preferred the above arguments.

The author currently does not know whether assumption (3.5) is “optimal” for similar results to hold. Note, however, that it was essential for the proof of Theorem 3.2.11 that (3.5), as otherwise we could not have defined the symbol f . For sequences that do not satisfy the absolute summability condition, however, there is something akin to a symbol. We refer to [EFHN15, Theorem 18.6] for a proof of the following.

Proposition 3.2.15 (Bochner–Herglotz). *For a sequence $(a_k)_{k \in \mathbb{Z}}$ in \mathbb{K} the following are equivalent.*

- (a) *There is a Hilbert space H , a unitary operator $T: H \rightarrow H$ and $x \in H$ such that $a_k = \langle x, T^k x \rangle$ for all $k \in \mathbb{Z}$.*
- (b) *There is a positive measure $\mu \in \mathcal{M}(\mathbb{T})$ such that*

$$a_k = \int_{\mathbb{T}} z^{-k} \mu(dz).$$

- (c) *Each Toeplitz matrix in the associated sequence $(A_N)_{N \in \mathbb{N}}$ is Hermitian and positive semidefinite.*

In this case, the measure μ is uniquely determined.

The uniqueness of the measure follows readily from the fact that by Weierstraß’s theorem the subspace $\text{span}\{z \mapsto z^k : k \in \mathbb{Z}\}$ forms a dense subset of $C_b(\mathbb{T})$.

So the “symbol” in this case is a positive measure μ that is not necessarily absolutely continuous with respect to the Lebesgue measure \mathbf{m} on \mathbb{T} . It is easily seen that, in the weak- $*$ -topology, μ can be approximated by absolutely continuous measures with bounded Radon–Nikodým derivatives with respect to the Lebesgue measure \mathbf{m} . The following is a slight generalization of [Gla03, Proposition 5.9 on p. 118].

Proposition 3.2.16. *Let $T: H \rightarrow H$ be a unitary operator on a Hilbert space H and let $x \in H$. Let $\mu_x \in \mathcal{M}(\mathbb{T})$ be the measure given by*

$$\langle x, T^k x \rangle = \int_{\mathbb{T}} z^{-k} \mu_x(dz).$$

Set

$$f_{n,x}: \mathbb{T} \rightarrow [0, \infty), \quad z \mapsto \frac{1}{n} \left\| \sum_{k=0}^{n-1} z^{-k} T^k x \right\|^2$$

and let $\nu_{n,x}$ be a measure on \mathbb{T} defined by $\nu_{n,x}(A) := \int_A f_{n,x} \, d\mathbf{m}$ for all measurable $A \subseteq \mathbb{T}$, where \mathbf{m} is the normalized Lebesgue measure on \mathbb{T} . Then $\nu_{n,x} \rightarrow \mu_x$ as $n \rightarrow \infty$ in the weak- $*$ -topology.

Moreover,

$$\int_{\mathbb{T}} z^{-k} f_{n,x}(z) \, \mathbf{m}(dz) = \left(1 - \frac{|k|}{n}\right)_+ \langle x, T^k x \rangle$$

for all $k \in \mathbb{Z}$ and $n \in \mathbb{N}$, where $a_+ := \max\{a, 0\}$ for $a \in \mathbb{R}$.

The measure μ_x is usually referred to as *spectral measure* of T at x . The approximation $(\nu_{n,x})_{n \in \mathbb{N}}$ of μ_x is called a *Fejér approximation*.

Proof. By the density of $\text{span}\{z \mapsto z^k : k \in \mathbb{Z}\}$ in $C_b(\mathbb{T})$ it suffices to show that

$$\int_{\mathbb{T}} z^{-k} f_{n,x} \, \mathbf{m}(dz) \rightarrow \int_{\mathbb{T}} z^{-k} \mu_x(dz)$$

for all $k \in \mathbb{Z}$. To this end, note that by Proposition 4.1.15 we have

$$\begin{aligned} \frac{1}{n} \left\| \sum_{k=0}^{n-1} z^{-k} T^k x \right\|^2 &= n \left\| \frac{1}{n} \sum_{k=0}^{n-1} z^{-k} T^k x \right\|^2 \\ &= 2 \sum_{k=0}^{n-1} \left(1 - \frac{k}{n}\right) \text{Re} \langle x, z^{-k} T^k x \rangle - \|x\|^2 \\ &= \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) (\langle x, z^{-k} T^k x \rangle + \langle z^{-k} T^k x, x \rangle) + \langle x, x \rangle \\ &= \sum_{k \in \mathbb{Z}} \left(1 - \frac{|k|}{n}\right)_+ \langle x, z^{-k} T^k x \rangle. \end{aligned} \tag{3.8}$$

Thus, multiplying with $z^{-\bar{k}}$ and integrating with respect to \mathbf{m} yields

$$\begin{aligned} \int_{\mathbb{T}} z^{-\bar{k}} f_{n,x}(z) \, \mathbf{m}(dz) &= \sum_{k \in \mathbb{Z}} \left(1 - \frac{|k|}{n}\right)_+ \langle x, T^k x \rangle \int_{\mathbb{T}} z^{k-\bar{k}} \, \mathbf{m}(dz) \\ &= \left(1 - \frac{|\bar{k}|}{n}\right)_+ \langle x, T^{\bar{k}} x \rangle \rightarrow \langle x, T^{\bar{k}} x \rangle = \int_{\mathbb{T}} z^{-\bar{k}} \mu_x(dz) \end{aligned}$$

as $n \rightarrow \infty$. □

Let us further break down the proof of Theorem 3.2.11. If we replace the assumption of the theorem and assume instead that for $x \in H$ there exists $C > 0$ and a positive null sequence $(b_n)_{n \in \mathbb{N}}$ such that $(\frac{1}{nb_n})_{n \in \mathbb{N}}$ is also an null sequence and we have

$$b_n^{-1} \left\| \frac{1}{n} \sum_{k=0}^{n-1} T^k x \right\|^2 < C$$

for all $n \in \mathbb{N}$. In other words, the rate of convergence of the Cesàro means in L^2 -norm is in $O(\sqrt{b_n})$ as $n \rightarrow \infty$. Then we obtain from a calculation akin to (3.8) that

$$\frac{b_n^{-1}}{n} \sum_{k \in \mathbb{Z}} (1 - \frac{|k|}{n})_+ \langle x, T^k x \rangle < C$$

for all $n \in \mathbb{N}$.

We can now argue as in the proof of Theorem 3.2.11 that for each $N \in \mathbb{N}$ and $\delta > 0$ we have (3.7). Now denote $a_{k,n} := (1 - \frac{|k|}{n})_+ \langle x, T^k x \rangle$ for all $k \in \mathbb{Z}$ and all $n \in \mathbb{N}$ and let $(A_{N,n})_{N,n \in \mathbb{N}}$ be the associated Toeplitz matrices. We notice that $A_{N,n}$ is not the Gram matrix $G_{\mathcal{O}_N}$. Instead, the sequence converges to $G_{\mathcal{O}_N}$. Indeed, fix $N \in \mathbb{N}$ and notice that $A_{N,n} \rightarrow G_{\mathcal{O}_N}$ as $n \rightarrow \infty$. By the continuity of eigenvalues of matrices we obtain $\lambda_{\max}(A_{N,n}) \rightarrow \lambda_{\max}(G_{\mathcal{O}_N})$ as $n \rightarrow \infty$.

Now note that

$$M_{f_{n,x}} \leq \sum_{k \in \mathbb{Z}} (1 - \frac{|k|}{n})_+ \langle x, T^k x \rangle$$

holds for all $n \in \mathbb{N}$. Thus, by Lemma 3.2.13 we have $\lambda_{\max}(A_{N,n}) \leq M_{f_{n,x}} \leq Cnb_n$ for all $n \in \mathbb{N}$.

Let $\varepsilon > 0$ and note that for each $N \in \mathbb{N}$ there exists $n_N \in \mathbb{N}$ with $n_N \rightarrow \infty$ such that $|\lambda_{\max}(A_{N,n_N}) - \lambda_{\max}(G_{\mathcal{O}_N})| < \varepsilon$. In particular, this shows that

$$\lambda_{\max}(G_{\mathcal{O}_N}) \leq Cn_N b_{n_N} + \varepsilon$$

for all $n \in \mathbb{N}$. Now (3.7) yields

$$\liminf_{N \rightarrow \infty} \frac{n_N b_{n_N}}{N} H_{\text{apr}}(\mathcal{O}_N, \delta) \geq \frac{1 - \delta^2}{C}.$$

Or in other terms, the approximation entropy $H_{\text{apr}}(\mathcal{O}_N, \delta)$ grows faster when the convergence of the Cesàro means

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} T^k x \right\| \rightarrow 0.$$

is faster.

Let us summarize.

Proposition 3.2.17. *Let $T: H \rightarrow H$ be a unitary operator on a Hilbert space H . Let $x \in H_{\text{aws}}$ be normalized and set*

$$\mathcal{O}_N := \{x, Tx, \dots, T^{N-1}x\}$$

for all $N \in \mathbb{N}$. Suppose that there exists a nonnegative null sequence $(b_n)_{n \in \mathbb{N}}$ such that $(\frac{1}{nb_n})_{n \in \mathbb{N}}$ is also a null sequence and there exists $C > 0$ such that

$$b_n^{-1} \left\| \frac{1}{n} \sum_{k=0}^{n-1} T^k x \right\|^2 < C$$

for all $n \in \mathbb{N}$. Then there is a sequence $(n_N)_{N \in \mathbb{N}}$ in \mathbb{N} with $n_N \rightarrow \infty$ such that for all $\delta > 0$ we have

$$\liminf_{N \rightarrow \infty} \frac{n_N b_{n_N}}{N} H_{\text{apr}}(\mathcal{O}_N, \delta) \geq \frac{1 - \delta^2}{C}.$$

If we could choose $n_N = N$ for all $N \in \mathbb{N}$, we would be able to deduce that

$$\liminf_{N \rightarrow \infty} b_N H_{\text{apr}}(\mathcal{O}_N, \delta) \geq \frac{1 - \delta^2}{C}.$$

In this case, we would have a precise asymptotic rate of growth for $H_{\text{apr}}(\mathcal{O}_N, \delta)$ in terms of the order of convergence of the Cesàro means. If we could show that $n_N \leq \tilde{C}N$ for some $\tilde{C} > 0$ for all $N \in \mathbb{N}$, and $b_n = n^{-\alpha}$ for some $\alpha \in (0, 1)$, then $n_N b_{n_N} = (n_N)^{1-\alpha} \leq \tilde{C}^{1-\alpha} N^{1-\alpha}$ and we would be able to deduce that

$$\liminf_{N \rightarrow \infty} \frac{1}{N^\alpha} H_{\text{apr}}(\mathcal{O}_N, \delta) \geq \frac{1 - \delta^2}{C \tilde{C}^{1-\alpha}}.$$

Unfortunately the author was not able show these bounds on the growth rate of $(n_N)_{N \in \mathbb{N}}$. Instead we prove the following result that further analyzes the convergence of $\lim_{n \rightarrow \infty} \lambda_{\max}(A_{N,n}) = \lambda_{\max}(G_{\mathcal{O}_N})$ for each $N \in \mathbb{N}$.

Theorem 3.2.18. *Let $T: H \rightarrow H$ be a unitary operator on a Hilbert space H . Let $x \in H_{\text{aws}}$ be normalized and set*

$$\mathcal{O}_N := \{x, Tx, \dots, T^{N-1}x\}$$

for all $N \in \mathbb{N}$.

(i) *Suppose that there is $\alpha \in (\frac{1}{2}, 1)$ and $C > 0$ such that*

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} T^k x \right\|^2 < C \frac{1}{n^\alpha} \tag{3.9}$$

for all $n \in \mathbb{N}$. Then we have for all $\delta > 0$ that

$$\liminf_{N \rightarrow \infty} \frac{1}{N^{2\alpha-1}} H_{\text{apr}}(\mathcal{O}_N, \delta) \geq \frac{1 - \delta^2}{C}.$$

(ii) Suppose that there is $\alpha \in (0, 1)$ such that

$$\lim_{n \rightarrow \infty} \frac{1}{n^{1-\alpha}} \sum_{k=1}^n |\langle x, T^k x \rangle| = 0. \quad (3.10)$$

Then we have for all $\delta \in (0, 1)$ that

$$\liminf_{N \rightarrow \infty} \frac{1}{N^{\alpha^2}} H_{\text{apr}}(\mathcal{O}_N, \delta) = \infty.$$

Remark 3.2.19. (i) The assumption (3.9) of Theorem 3.2.18 (i) implies the assumption (3.10) of (ii) by Proposition 4.1.15.

(ii) The assumption (3.10) of Theorem 3.2.18 (ii) is in turn implied by asking for the sequence of correlation coefficients $(\langle x, T^k x \rangle)_{k \in \mathbb{Z}}$ to be in $\ell^p(\mathbb{Z})$ for some $p \in (1/\alpha, \infty)$, as the Hölder conjugate q of p satisfies $1/p < 1 - \alpha$ and by Hölder's inequality we then have

$$\frac{1}{n^{1-\alpha}} \sum_{l=1}^n |\langle x, T^k x \rangle| \leq \frac{1}{n^{1-\alpha}} \cdot n^{1/p} \cdot \left(\sum_{l=1}^n |\langle x, T^k x \rangle|^p \right)^{1/p} \rightarrow 0$$

as $n \rightarrow \infty$.

We need the following simple lemma for the proof of Theorem 3.2.18.

Lemma 3.2.20. Let $A, B \in \mathbb{K}^{n \times n}$ be Hermitian matrices. Denote by $\lambda_{\max}(A)$ and $\lambda_{\max}(B)$ the maximal eigenvalue of A and B . Then we have for the maximal eigenvalues that

$$|\lambda_{\max}(A) - \lambda_{\max}(B)| \leq \|A - B\|_2,$$

where $\|\cdot\|_2$ denotes the Euclidean operator norm.

Proof. From [MOA11, Theorem G.1 on p. 329] one obtains that for Hermitian matrices G and H we have

$$\lambda_{\max}(G + H) \leq \lambda_{\max}(G) + \lambda_{\max}(H).$$

Since $\|G\|_2$ coincides with the spectral bound for Hermitian matrices, we obtain for $G := A - B$ and $H := B$ that

$$\begin{aligned} \lambda_{\max}(A) - \lambda_{\max}(B) &= \lambda_{\max}(G + H) - \lambda_{\max}(H) \\ &\leq \lambda_{\max}(G) = \lambda_{\max}(A - B) \leq \|A - B\|_2 \end{aligned}$$

Exchanging the roles of A and B yields the claim. \square

Proof of Theorem 3.2.18. We proceed by presenting arguments that are needed in the proof of both claims.

We can argue as in the proof of Theorem 3.2.11 that \mathcal{O}_N is a linearly independent set for each $N \in \mathbb{N}$. Recall the notation $(a_{k,n})_{k \in \mathbb{Z}}$, $(A_{N,n})_{N \in \mathbb{N}}$ and $f_{n,x}$ from the

discussion before Proposition 3.2.17. Consider for each $n \in \mathbb{N}$ the sequence $(b_{k,n})_{n \in \mathbb{Z}}$ defined by

$$a_{n,k} := \begin{cases} \frac{|k|}{n} \langle x, T^k x \rangle, & |k| \leq n, \\ 0, & |k| > n \end{cases}$$

and let $(B_{N,n})$ be the corresponding sequence of Toeplitz matrices with $B_{N,n} \in \mathbb{K}^N$. For $n \in \mathbb{N}$ we denote by f_n the symbol of $(b_{k,n})_{n \in \mathbb{Z}}$ and let $\lambda_{1,N,n} \geq \dots \geq \lambda_{N,N,n}$ be the eigenvalues of $B_{N,n}$. From Lemma 3.2.13 we obtain that

$$m_{f_n} \leq \lambda_{k,N,n} \leq M_{f_n}$$

for all $k \in \{1, \dots, N\}$ and all $n \in \mathbb{N}$. Now notice that

$$\begin{aligned} \|B_{N,n}\|_2 &= \max_{1 \leq k \leq N} |\lambda_{k,N,n}| \leq \max\{m_{f_n}, M_{f_n}\} \\ &\leq \sum_{k \in \mathbb{Z}} |a_{k,n}| = \sum_{k=-N}^N \frac{|k|}{n} \langle x, T^k x \rangle. \end{aligned}$$

“(i)”: We show that the sequence $(n_N)_{N \in \mathbb{N}}$ in Proposition 3.2.17 can be chosen to be $n_N = N^2$. Then, since $b_n = n^{-\alpha}$ we obtain $n_N b_{n_N} = N^2 N^{-2\alpha} = N^{2-2\alpha}$, and thus,

$$\liminf_{N \rightarrow \infty} \frac{1}{N^{2\alpha-1}} H_{\text{apr}}(\mathcal{O}_N, \delta) > \frac{1 - \delta^2}{C},$$

as claimed.

Recall from the discussion before Proposition 3.2.17 that, under the given assumptions, we have proved the bound $\lambda_{\max}(A_{N,n}) \leq M_{f_{n,x}} \leq Cn^{1-\alpha}$ for all $n, N \in \mathbb{N}$. By choosing $n = N^2$ we obtain

$$\|B_{N,N^2}\|_2 \leq \frac{1}{N} \sum_{k=-N}^N \frac{|k|}{N} \langle x, T^k x \rangle = \frac{1}{N} \sum_{k=-N}^N |\langle x, T^k x \rangle| \rightarrow 0$$

as $N \rightarrow \infty$. Let $\varepsilon > 0$, and pick $N_\varepsilon \in \mathbb{N}$ such that for all $N \geq N_\varepsilon$ we have $\|B_{N,N^2}\|_2 \leq \varepsilon$

Notice that for $n \geq N$ we have $G_{\mathcal{O}} - A_{N,n} = B_{N,n}$. So Lemma 3.2.20 yields

$$|\lambda_{\max}(A_{N,N^2}) - \lambda_{\max}(G_{\mathcal{O}_N})| \leq \|B_{N,N^2}\|_2 \leq \varepsilon$$

for all $N \geq N_\varepsilon$.

Applying Corollary 3.2.8 yields

$$H_{\text{apr}}(\mathcal{O}_N, \delta) \geq \frac{1 - \delta^2}{\lambda_{\max}(G_{\mathcal{O}_N})} N \geq \frac{1 - \delta^2}{\lambda_{\max}(A_{N,N^2}) + \varepsilon} N \geq \frac{1 - \delta^2}{CN^{2-2\alpha} + \varepsilon} N$$

for each $\delta > 0$. By rearranging and taking the lim inf we obtain

$$\liminf_{N \rightarrow \infty} \frac{1}{N^{2\alpha-1}} H_{\text{apr}}(\mathcal{O}_N, \delta) \geq \frac{1 - \delta^2}{C}$$

as claimed.

“(ii)”: Let $C > 0$. Notice that by Proposition 4.1.15 the assumption implies that there exists $n_C \in \mathbb{N}$

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} T^k x \right\| \leq \frac{2}{n} \sum_{k=1}^n |\langle x, T^k x \rangle| \leq C \frac{1}{n^\alpha}$$

for all $n \geq n_C$. Hence, we can deduce the bound $\lambda_{\max}(A_{N,n}) \leq M_{f_{n,x}} \leq Cn^{1-\alpha}$ for all $N \in \mathbb{N}$ and $n \geq n_C$, as before.

By choosing $n = \lceil N^{1+\alpha} \rceil$ or $n = \lceil N^{2-\alpha} \rceil$, we obtain

$$\|B_{N, \lceil N^{1+\alpha} \rceil}\|_2 \leq \frac{1}{N^\alpha} \sum_{k=-N}^N \frac{|k|}{N} \langle x, T^k x \rangle = \frac{2}{N^\alpha} \sum_{k=1}^N |\langle x, T^k x \rangle| + \frac{1}{N^\alpha} \rightarrow 0.$$

Let $\varepsilon > 0$ and choose $N_\varepsilon \in \mathbb{N}$ such that $\lceil N^{1+\alpha} \rceil \geq n_C$ and that for each $N \geq N_\varepsilon$ we have $\|B_{N, \lceil N^{1+\alpha} \rceil}\|_2 \leq \varepsilon$.

Notice that, as before, for $n \geq N$ we have $G_{\mathcal{O}} - A_{N,n} = B_{N,n}$. So Lemma 3.2.20 yields

$$|\lambda_{\max}(A_{N, \lceil N^{1+\alpha} \rceil}) - \lambda_{\max}(G_{\mathcal{O}_N})| \leq \|B_{N, \lceil N^{1+\alpha} \rceil}\|_2 \leq \varepsilon$$

for all $N \geq N_\varepsilon$.

Applying Corollary 3.2.8 yields

$$\begin{aligned} H_{\text{apr}}(\mathcal{O}_N, \delta) &\geq \frac{1 - \delta^2}{\lambda_{\max}(G_{\mathcal{O}_N})} N \geq \frac{1 - \delta^2}{\lambda_{\max}(A_{N, \lceil N^{1+\alpha} \rceil}) + \varepsilon} N \\ &\geq \frac{1 - \delta^2}{C \lceil N^{1+\alpha} \rceil^{1-\alpha} + \varepsilon} N \geq \frac{1 - \delta^2}{CN^{(1+\alpha)(1-\alpha)} + \varepsilon} N \\ &= \frac{1 - \delta^2}{CN^{1-\alpha^2} + \varepsilon} N = \frac{1 - \delta^2}{C + N^{\alpha^2-1}\varepsilon} N^{\alpha^2} \end{aligned}$$

for each $\delta > 0$. By rearranging and taking the lim inf we obtain

$$\liminf_{N \rightarrow \infty} \frac{1}{N^{\alpha^2}} H_{\text{apr}}(\mathcal{O}_N, \delta) \geq \frac{1 - \delta^2}{C}.$$

As $C > 0$ was arbitrarily chosen, the claim follows. \square

Theorem 3.2.18 justifies the definition of a *fractional approximation entropy*.

Definition 3.2.21. Let $T: H \rightarrow H$ be a unitary operator on a Hilbert space H and let $\mathcal{O} \subset H$ be a finite set of observables. For $\alpha \in (0, 1)$ we define

$$h_{\text{apr}}^\alpha(T, \mathcal{O}) := \sup_{\delta > 0} \liminf_{n \rightarrow \infty} \frac{1}{n^\alpha} H_{\text{apr}} \left(\bigcup_{k=0}^{n-1} T^k \mathcal{O}, \delta \right)$$

to be the *lower α -approximation entropy* and

$$\bar{h}_{\text{apr}}^\alpha(T, \mathcal{O}) := \sup_{\delta > 0} \limsup_{n \rightarrow \infty} \frac{1}{n^\alpha} H_{\text{apr}} \left(\bigcup_{k=0}^{n-1} T^k \mathcal{O}, \delta \right)$$

to be the *upper α -approximation entropy* for $\alpha \in (0, 1)$.

The author does currently not know whether both concepts coincide, as it is the case for the standard approximation entropy. Note that for $\alpha = 1$ one recovers the standard approximation entropy and there Proposition 3.2.1 (viii) shows that the limit superior and the limit inferior coincide.

Let us illustrate the results of Theorem 3.2.11 and Theorem 3.2.18 by considering as an example the two-sided Bernoulli shift that we got to know in Example 2.1.2 (i).

Example 3.2.22. Let $X = (\{-1, 1\}^{\mathbb{N}}, \mathcal{A}, \mu)$ be the infinite sequence space with the product σ -algebra \mathcal{A} and μ be the product measure of the Rademacher measures on $\{-1, 1\}$. Consider the left shift mapping $\varphi: X \rightarrow X$ defined by $(x_m)_{m \in \mathbb{Z}} \mapsto (x_{m+1})_{m \in \mathbb{Z}}$. As φ is an invertible measure-preserving mapping, the Koopman operator T_φ on $L^2(X)$ defines a unitary operator.

Note that φ is strongly mixing, and thus, the reversible part $L^2(X)_{\text{rev}}$ of the JdLG-decomposition is precisely the subspace of the constant functions.

Let us first consider the observables $f \in L^2(X)$ defined by $f((x_m)_{m \in \mathbb{Z}}) = x_0$. A straightforward calculation shows $\langle f, \mathbb{1} \rangle = 0$, and thus, $f \in L^2(X)_{\text{aws}}$. Moreover, the correlation coefficients satisfy

$$\langle f, T_\varphi^k f \rangle = \delta_{0,k}$$

for all $k \in \mathbb{Z}$. As a result, the assumption of Theorem 3.2.11 is satisfied with $C = 0$ and we conclude that the approximation entropy $h_{\text{apr}}(T_\varphi, \{f\})$ is lower bounded by 1. While at the same time the mean ergodic convergence rate satisfies

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} T_\varphi^k f \right\|^2 = \frac{2}{n}$$

for all $n \in \mathbb{N}$, by Proposition 4.1.15. Note that the mean ergodic convergence in this case is also guaranteed by the strong law of large numbers, as the sequence $(T_\varphi^k f)_{k \in \mathbb{N}}$ is a sequence of independent and identically distributed random variables with mean zero.

Next, consider the observable $g \in L^2(X)$ defined by $g((x_m)_{m \in \mathbb{Z}}) = \sum_{m=1}^{\infty} \frac{x_m}{m}$. Then $g \in L^2(X)$ with $\|g\|_{L^2} = (\sum_{m=1}^{\infty} \frac{1}{m^2})^{1/2}$. A straightforward calculation shows that $\langle g, \mathbb{1} \rangle = 0$, and thus, $g \in L^2(X)_{\text{aws}}$. Moreover, the correlation coefficients satisfy

$$\langle g, T_\varphi^k g \rangle = \sum_{m=1}^{\infty} \frac{1}{(m+k)m} = \frac{1}{k} \sum_{m=1}^{\infty} \left(\frac{1}{m} - \frac{1}{m+k} \right) = \frac{1}{k} \sum_{m=1}^k \frac{1}{m}$$

for all $k \in \mathbb{Z}$. As a result, we obtain for any $\alpha > 0$ that there is some constant $C > 0$ such that

$$\begin{aligned} \frac{1}{n^\alpha} \sum_{k=1}^n |\langle g, T_\varphi^k g \rangle| &= \frac{1}{n^\alpha} \sum_{k=1}^n \frac{1}{k} \sum_{m=1}^k \frac{1}{m} \leq \frac{1}{n^\alpha} \sum_{k=1}^n \frac{1}{k} (1 + \log(k)) \\ &\leq \frac{1}{n^\alpha} \int_1^n \frac{1 + \log(x)}{x} dx + \frac{C}{n^\alpha} \\ &= \frac{1}{n^\alpha} \left(\frac{(\log(n))^2}{2} + \log(n) \right) + \frac{C}{n^\alpha} \rightarrow 0. \end{aligned}$$

So the assumption of Theorem 3.2.18 (ii) is satisfied for all $\alpha \in (0, 1)$ and we conclude that the lower α^2 -approximation entropy $\underline{h}_{\text{apr}}^{\alpha^2}(T_\varphi, \{g\}) = \infty$ for all $\alpha \in (0, 1)$.

Note that we also obtain from Proposition 4.1.15 that the mean ergodic convergence rate satisfies

$$\begin{aligned} \left\| \frac{1}{n} \sum_{k=0}^{n-1} T_\varphi^k g \right\|^2 &= \frac{2}{n} \sum_{k=1}^n \left(1 - \frac{k}{n}\right) \langle g, T_\varphi^k g \rangle + \frac{1}{n} \|g\|^2 \\ &\leq \frac{2}{n} \left(\frac{(\log(n))^2}{2} + \log(n) + \|g\|^2 + C \right) \rightarrow 0. \end{aligned}$$

for all $n \in \mathbb{N}$. So we might not expect the mean ergodic convergence to be as quick as for the observable f .

Intuitively, we can say that g contains much more information on the future evolution of the system than f does. So g is “much more invariant” under the action of T_φ than f is, which is reflected in the slower mean ergodic convergence rate and the higher growth rate of the approximation entropy. This hints towards an interesting trade-off between mean ergodic convergence rates and approximation entropy growth rates, which is worth further investigation. Here, we have only explored one direction of this trade-off, so it would be interesting to study whether faster growth rates of the approximation entropy really imply faster mean ergodic convergence rates.

By fine tuning and putting more weight on the observations far in the future, one could try to construct observables with slower mean ergodic convergence rates but better approximation entropy growth rates. Possible candidates to study are the members of the family of functions

$$g_\beta: \{0, 1\}^{\mathbb{Z}} \rightarrow \mathbb{R}, \quad g_\beta((x_m)_{m \in \mathbb{Z}}) = \sum_{m=1}^{\infty} \frac{x_m}{m^\beta}$$

for $\beta > \frac{1}{2}$. As proofs of the above type become more complicated for functions of this family, we declare this to be beyond the scope of this thesis. The above example suffices to illustrate our point.

3.2.3 A Spectral-Theoretic Perspective on Approximation Entropy

In this section, we present details and a slight generalization of [Voi95, Proposition 7.11] on the relationship between the approximation entropy and the spectral theory of unitary operators [Con85; EFHN15; Kri86; Voi95].

Let $T: H \rightarrow H$ be a unitary operator on a complex and separable Hilbert space and assume that $\mathbb{K} = \mathbb{C}$. For $x \in H$ we call the unique measure $\mu_x: \mathcal{B}(\mathbb{T}) \rightarrow \mathbb{R}$ satisfying the property

$$\langle x, T^k x \rangle = \int_{\mathbb{T}} z^{-k} \mu_x(dz)$$

for all $k \in \mathbb{Z}$ a *spectral measure* of T . That such a measure exists follows from the Bochner–Herglotz theorem (see Proposition 3.2.15). A spectral measure μ_x can be decomposed by means of a Lebesgue decomposition (see [RS80, p. 22–25] or [EFHN15, Theorem B.25]) into mutually singular measures

$$\mu_x = \mu_{x,d} + \mu_{x,ac} + \mu_{x,sc},$$

where $\mu_{x,d}$ is a discrete measure, $\mu_{x,ac}$ is absolutely continuous with respect to the Lebesgue measure \mathfrak{m} on \mathbb{T} (in short $\mu_{x,ac} \ll \mathfrak{m}$), and $\mu_{x,sc}$ is non-atomic and singular to the Lebesgue measure (in short $\mu_{x,sc} \perp \mathfrak{m}$). Note that all three parts are absolutely continuous with respect to μ_x , and therefore, there exists elements $x_d, x_{ac}, x_{sc} \in H$ such that $\mu_{x,*} = \mu_{x_*}$ for $* \in \{d, ac, sc\}$ [EFHN15, Lemma 18.9]. This yields a decomposition of the Hilbert space $H = H_d \oplus H_{ac} \oplus H_{sc}$ into T -invariant subspaces

$$\begin{aligned} H_d &:= \{x \in H : \mu_x \text{ is discrete}\}, & H_{ac} &:= \{x \in H : \mu_x \ll \mathfrak{m}\}, \\ H_{sc} &:= \{x \in H : \mu_x \text{ is non-atomic and } \mu_x \perp \mathfrak{m}\}. \end{aligned}$$

The T -invariance follows readily from the fact that $\mu_{Tx} = \mu_x$ for all $x \in H$, as T is unitary. Furthermore, the decomposition into T -invariant subspaces yields a decomposition of T into

$$T = T_d \oplus T_{ac} \oplus T_{sc}$$

with $T_*: H_* \rightarrow H_*$ for all $* \in \{d, ac, sc\}$.

Remark 3.2.23 (Jacob–de Leeuw–Glicksberg decomposition). (i) The subspace H_d is given by the closure of the linear span of all eigenvectors of T [EFHN15, Proposition 18.17].

(ii) If $H = H_{\text{rev}} \oplus H_{\text{aws}}$ is a Jacob–de Leeuw–Glicksberg decomposition of T (see Definition 2.2.6) into a *reversible part* H_{rev} and an *almost weakly stable part* H_{aws} , then H_d and H_{rev} coincide [EFHN15, Corollary 18.18].

- (iii) The space $H_c := H_{ac} \oplus H_{sc}$ coincides with the almost weakly stable part H_{aws} of the Jacob–de Leeuw–Glicksberg decomposition [EFHN15, Corollary 18.18]. The elements $x \in H_c$ have the property that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} |\langle T^k x, y \rangle|^p = 0$$

holds for all $y \in H$ and all $p \in [1, \infty)$ [EFHN15, Theorem 16.34]. The subspace H_c is also often referred to as the *continuous part* of T [EFHN15, Chapter 18.2].

Next we define the concept of the multiplicity function by following [Kri86, Section 2]. Let $\ell^2 := \ell^2(\mathbb{N}; \mathbb{C})$ denote the Hilbert space of complex square-summable sequence indexed by the natural numbers. Further let $\nu: \mathcal{B}(\mathbb{T}) \rightarrow [0, \infty)$ be a Borel measure on \mathbb{T} , write $\mathbb{N} := (\mathbb{T}, \mathcal{B}(\mathbb{T}), \nu)$ for the corresponding measure space and denote by $L^2(\mathbb{N}; \ell^2)$ the space of all functions $f: \mathbb{T} \rightarrow \ell^2$ such that $\int_{\mathbb{T}} \|f\|_{\ell^2}^2 d\nu < \infty$. The scalar product $\langle f, g \rangle = \int_{\mathbb{T}} \langle f(\lambda), g(\lambda) \rangle_{\ell^2} \nu(d\lambda)$ on $L^2(\mathbb{N}; \ell^2)$ makes it into a Hilbert space.

Now let $m: \mathbb{T} \rightarrow \mathbb{N}_0 \cup \{\infty\}$ be a Borel-measurable function with $\nu(\{m = 0\}) = 0$ and set $\ell_n^2 := \{x = (x_n)_{n \in \mathbb{N}} \in \ell^2 : \forall k > n : x_k = 0\}$ for all $n \in \mathbb{N}_0 \cup \{\infty\}$. We denote

$$H_{\nu, m} := \{f \in L^2(\mathbb{N}; \ell^2) : f(\lambda) \in \ell_{m(\lambda)}^2 \text{ for } \nu\text{-almost all } \lambda \in \mathbb{T}\}.$$

Then $H_{\nu, m}$ is a closed subspace of $L^2(\mathbb{N}; \ell^2)$. Consider the multiplication operator

$$T_{\nu, m}: H_{\nu, m} \rightarrow H_{\nu, m}, \quad (T_{\nu, m} f)(\lambda) \mapsto \lambda \cdot f(\lambda)$$

by the identity function.

The representation theorem for unitary operators states that every unitary operator $T: H \rightarrow H$ on a separable Hilbert space is unitarily equivalent to $T_{\nu, m}$ for some Borel measure ν and some Borel function m as above (see [Con81, Theorem II.9.18] or [Con85, Theorem IX.10.20]). Moreover, if $\tilde{\nu}$ is another Borel measure and \tilde{m} another Borel function with this property, then ν and $\tilde{\nu}$ have the same null sets (or equivalently, are mutually absolutely continuous) and ν -almost everywhere we have $m = \tilde{m}$ [Con81, Corollary II.9.12 b)]. In this case, the measure ν is called a *maximal spectral measure* and the function m is called a *multiplicity function* of T .

The maximal spectral measure ν is indeed a spectral measure of T . To see this denote by x the constant function in $L^2(\mathbb{N}; \ell^2)$ that maps to the first unit vector $e_1 \in \ell^2$. Since $\nu(\{m = 0\}) = 0$, it follows that $x \in H_{\nu, m}$. Then

$$\langle x, T_{\nu, m}^k x \rangle_{H_{\nu, m}} = \int_{\mathbb{T}} \langle x, T_{\nu, m} x \rangle_{\ell^2} d\nu = \int_{\mathbb{T}} z^{-k} \nu(dz)$$

holds for all $k \in \mathbb{Z}$, from which follows that $\nu = \mu_x$. Thus, by unitary equivalence of T and $T_{\nu, m}$ the statement follows.

Suppose now that $T = T_{ac}$ and let ν be a maximal spectral measure and m be a multiplicity function of T . Then there exists $x \in H = H_{ac}$ such that $\nu = \mu_x$. In particular, ν is absolutely continuous with respect to the Lebesgue measure \mathfrak{m} on \mathbb{T} . Let $\frac{d\nu}{d\mathfrak{m}} \in L^1(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathfrak{m})$ be its Radon–Nikodým derivative. Denote by $\Delta_T := \{\lambda \in \mathbb{T} : \frac{d\nu}{d\mathfrak{m}}(\lambda) \neq 0\}$ its set of non-vanishing points, which is uniquely

defined up to a \mathbf{m} -null set and does not depend on the maximal spectral measure ν chosen. Then ν and $\mathbf{m}|_{\Delta_T}$ are mutually absolutely continuous and can therefore be exchanged in the representation of T .

The following was proved in [Voi95, Theorem 7.11] for the case in which the multiplicity function for the absolutely continuous part T_{ac} is finite. We present the short argument showing how the case of a general multiplicity function follows from this, which can also be found as [HH25, Proposition 5.2].

Proposition 3.2.24. *Let $T: H \rightarrow H$ be a unitary operator on a complex and separable Hilbert space and denote by m the multiplicity function of its absolutely continuous part T_{ac} . Then*

$$h_{\text{apr}}(T) = \int_{\Delta_{T_{\text{ac}}}} m \, d\mathbf{m}.$$

In particular, if $H_{\text{ac}} = \{0\}$, then $h_{\text{apr}}(T) = 0$.

Before providing a proof of Proposition 3.2.24, we illustrate why the restriction $\Delta_{T_{\text{ac}}}$ is necessary for the integral on the right-hand side in Proposition 3.2.24 to be well-defined.

Example 3.2.25. Let $\tilde{\mathbf{m}}$ be the Lebesgue measure supported on $\mathbb{T}_+ := \{\lambda \in \mathbb{T} : \text{Re}(\lambda) \geq 0\}$. The multiplication operator $T: L^2(\mathbb{T}, \mathcal{B}(\mathbb{T}), \tilde{\mathbf{m}}) \rightarrow L^2(\mathbb{T}, \mathcal{B}(\mathbb{T}), \tilde{\mathbf{m}})$ defined by $f \mapsto \text{id} \cdot f$ has maximal spectral measure $\tilde{\mathbf{m}}$, and thus, both functions $m_1 := \mathbb{1}_{\mathbb{T}}$ and $m_2 := \mathbb{1}_{\mathbb{T}_+}$ are multiplicity functions for T .

Proof of Proposition 3.2.24. As shown in the proof of [Voi95, Theorem 7.11] we may assume that $T = T_{\text{ac}}$. Let m be a multiplicity function of T .

If $m < \infty$, then the statement of Proposition 3.2.24 is proved in [Voi95, Theorem 7.11] by showing that on the additive semigroup \mathcal{C} of functions $m: \mathbb{T} \rightarrow \mathbb{N}_0$ the functional

$$\mu: \mathcal{C} \rightarrow [0, \infty), \quad m \mapsto h_{\text{apr}}(T),$$

where $T = T_{\nu, m}$ and $\nu = \mathbf{m}|_{\{m \neq 0\}}$, coincides with the integral $\int_{\mathbb{T}} m \, d\mathbf{m} = \int_{\Delta_T} m \, d\mathbf{m}$.

Now consider the general case and let $T_{\nu, m}$ be a representation of T with $\nu = \mathbf{m}|_{\{m \neq 0\}}$. Consider the closed subspaces

$$H_n := \{f \in H_{\nu, m} : f \in \ell_n^2\}$$

for each $n \in \mathbb{N}$. Then $H_n \subseteq H_{n+1}$ and $\bigcup_{n \in \mathbb{N}} H_n = H_{\nu, m}$. Moreover, $T_n := T|_{H_n}$ has the multiplicity function $m_n := \min(m, n)$ with ν being a maximal spectral measure of T_n . Thus, by the first part of the proof

$$h_{\text{apr}}(T, H_n) = h_{\text{apr}}(T_n) = \int_{\Delta_T} m_n \, d\mathbf{m}.$$

Now increasing $n \rightarrow \infty$, we obtain on the right-hand side of the equation from the monotone convergence theorem that $\sup_{n \in \mathbb{N}} \int_{\Delta_T} m_n \, d\mathbf{m} = \int_{\Delta_T} m \, d\mathbf{m}$. On the left-hand side we have $h_{\text{apr}}(T, H_n) = h_{\text{apr}}(T_n)$ and obtain from Proposition 3.2.2 together with Proposition 3.2.3 that $\sup_{n \in \mathbb{N}} h_{\text{apr}}(T_n) = h_{\text{apr}}(T_{\nu, m}) = h_{\text{apr}}(T)$. \square

For different constructions of a representation of normal operators using Hahn–Hellinger theory we refer to [Con85, Theorems IX.10.20 and IX.10.21], [Nad20, Theorems 1.34 and 1.36] or [Que10, Theorems II.4 and II.11]. Spectral theory of dynamical systems is surveyed in [KT06]; more specifically to Section 2.1 therein for more information on dynamical systems with countable Lebesgue part. The latter is still an active field of research [Lem24] and has been for some time [CFS82, Part III].

3.3 Interplay between Measure-Theoretic and Approximation Entropy

In this section we will show that any measure-preserving dynamical system with positive measure-theoretic entropy has infinite approximation entropy. For this we will need to introduce the following property from [EFHN15, Section 18.3].

Definition 3.3.1. Let $T: H \rightarrow H$ be a unitary operator on a separable Hilbert space H . Then T is said to have *countable Lebesgue spectrum* if and only if there exists an orthonormal sequence $(x_n)_{n \in \mathbb{N}}$ of H such that for each $n \in \mathbb{N}$ the sequence $(T^k x_n)_{k \in \mathbb{Z}}$ is an orthonormal basis of the *cyclic subspace* $Z(x_n; T) := \overline{\text{span}}\{T^k x_n : k \in \mathbb{Z}\}$ and

$$H = \bigoplus_{n=1}^{\infty} Z(x_n; T).$$

Remark 3.3.2. The reader familiar with spectral theory might be used to define countable Lebesgue spectrum of a unitary operator $T: H \rightarrow H$ in the context of separable Hilbert spaces by the equivalence of the Lebesgue measure on \mathbb{T} to the spectral measures appearing in the Hahn–Hellinger theorem. This definition is equivalent to our definition as shown in [EFHN15, Proposition 18.32].

Whenever (X, φ) is a measure-preserving dynamical system, then the Koopman operator T_φ fixes the constant functions. Thus, no Koopman operator can have countable Lebesgue spectrum without restricting it to certain non-trivial subspaces. It is thus common to define that (X, φ) has *countable Lebesgue spectrum* if the Koopman operator restricted to $\text{span}\{1\}^\perp$ has countable Lebesgue spectrum [EFHN15, Section 18.4.2]. More generally, we can study the following notion.

Definition 3.3.3. A unitary operator $T: H \rightarrow H$ on a Hilbert space H is said to have a *countable Lebesgue component* if there exists a closed separable subspace $H_0 \subseteq H$ with $TH_0 = H_0$ such that T restricted to H_0 has countable Lebesgue spectrum.

We purposefully omit the separability assumption on H , as it is not necessary for our results. Note that for operators on separable Hilbert spaces, this definition is equivalent to the notion of countable Lebesgue components given in [HK02, Section 3.7.j].

We next show that unitary operators with countable Lebesgue components have infinite approximation entropy. This is [HH25, Proposition 4.4].

Proposition 3.3.4. *Let $T: H \rightarrow H$ be a unitary operator on a Hilbert space H that has a countable Lebesgue component. Then $h_{\text{apr}}(T) = \infty$.*

Proof. For closed separable subspaces $H_0 \subseteq H$ with $TH_0 = H_0$ it is straightforward to observe that $h_{\text{apr}}(T|_{H_0}) \leq h_{\text{apr}}(T)$. We can thus assume without loss of generality that H is separable and that T has countable Lebesgue spectrum.

Let $(x_n)_{n \in \mathbb{N}}$ be the orthonormal sequence from Definition 3.3.1 and set $\mathcal{O}_j := \{x_1, \dots, x_j\}$ for each $j \in \mathbb{N}$. Then for each $n \in \mathbb{N}$ the set $\bigcup_{k=0}^{n-1} T^k \mathcal{O}$ is an orthonormal system with cardinality $j \cdot n$ and from Corollary 3.2.8 we obtain

$$H_{\text{apr}} \left(\bigcup_{k=0}^{n-1} T^k \mathcal{O}, \delta \right) \geq (j \cdot n)(1 - \delta^2)$$

for every $\delta > 0$. Then $h_{\text{apr}}(T, \mathcal{O}_j) \geq j$ holds for all $j \in \mathbb{N}$. Taking the supremum over all \mathcal{O}_j , we obtain $h_{\text{apr}}(T) = \infty$. \square

From [Rok67, Result 14.3] and also [HK02, Theorem 3.7.13] we know the following proposition in the case of X being a standard Lebesgue space (see Definition 2.1.11). We will present below how this follows for general invertible measure-preserving systems.

Proposition 3.3.5. *Let φ be an invertible measure-preserving dynamical system on the probability space $X = (X, \Sigma, \mu)$ with $h_\mu(\varphi) > 0$. Then T_φ has a countable Lebesgue component.*

As the literature on ergodic theory usually assumes φ to be standard Lebesgue, we need the next two lemmas so that we can reduce the general case to the standard Lebesgue case. It allows us to restrict the proof of Proposition 3.3.5 to standard Lebesgue systems only.

Lemma 3.3.6. *Let $\varphi: X \rightarrow X$ be an invertible measure-preserving dynamical system on the probability space $X = (X, \Sigma_X, \mu)$. Then $h_\mu(\varphi) = \sup_\psi h_\nu(\psi)$, where the supremum is taken over all standard Lebesgue factors $\psi: Y \rightarrow Y$ on $Y = (Y, \Sigma_Y, \nu)$ of φ .*

Proof. “ \leq ”: This is a straightforward consequence of the monotonicity of measure-theoretic entropy under factors. For details see [Wal82, Section 4.4 on p. 89].

“ \geq ”: Let α be a finite Σ_X -measurable partition of X . We proceed by choosing a factor ψ of φ . Let Σ_Y be the σ -algebra generated by the countable set $\bigcup_{n \in \mathbb{N}} \bigvee_{k=-n}^n \varphi^k(\alpha)$. Then it follows from a monotone class argument that

$$L^2(X, \Sigma_Y, \nu) = \overline{\text{span}} \bigcup_{n \in \mathbb{N}} \left\{ \mathbb{1}_A : A \in \bigvee_{k=-n}^n \varphi^k(\alpha) \right\}, \quad (3.11)$$

where ν denotes the restriction of μ to Σ_Y , let $\psi := \varphi$ and denote $Y = (X, \Sigma_Y, \nu)$. It follows directly from (3.11) that $L^2(Y)$ is separable.

From Proposition 2.1.12 it follows that there exists a measure space $Z = (K, \mathcal{B}(K), \xi)$, where K is compact metric and ξ is a probability measure on $\mathcal{B}(K)$, a measure-preserving dynamical system κ on Z , and a Markov isomorphism $T_1: L^2(Z) \rightarrow L^2(Y)$

that intertwines as $T_\psi T_1 = T_1 T_\kappa$. Moreover, the identity operator $T_2: L^2(Y) \rightarrow L^2(X)$ is a Markov embedding that trivially intertwines as $T_\varphi T_2 = T_2 T_\psi$. In summary, the operator $T := T_2 T_1: L^2(Z) \rightarrow L^2(X)$ is a Markov embedding that intertwines as $T_\varphi T = T T_\kappa$. This shows that κ is a standard Lebesgue factor of φ .

Furthermore, since (Z, κ) and (Y, ψ) are factors of each other via T_1 and T_1^{-1} it follows from monotonicity that $h_\xi(\kappa) = h_\nu(\psi)$. Moreover, the measure-theoretic entropy satisfies $h_\mu(\varphi, \alpha) = h_\nu(\psi, \alpha) = h_\nu(\psi)$ by the Kolmogorov–Sinai generator theorem [Wal82, Theorem 4.17]. This proves the claim. \square

Lemma 3.3.7. *Let $\varphi: X \rightarrow X$ be a measure-preserving dynamical system on the probability space $X = (X, \Sigma_X, \mu)$ and let $\psi: Y \rightarrow Y$ be a factor of φ on the probability space $Y = (Y, \Sigma_Y, \nu)$. Denote by T_φ and T_ψ the Koopman operators on $L^2(X)$ and $L^2(Y)$, respectively. When T_ψ has a countable Lebesgue component, then so does T_φ .*

Proof. Denote by $T: L^2(Y) \rightarrow L^2(X)$ the associated Markov embedding that intertwines T_φ and T_ψ as $T_\varphi T = T T_\psi$ and $H_0 \subseteq L^2(Y)$ for a closed separable subspace with $T_\psi H_0 = H_0$, such that T_ψ restricted to H_0 has countable Lebesgue spectrum. Clearly $H_1 := T H_0$ is a closed separable subspace of $L^2(X)$ with $T_\varphi H_1 = H_1$. Since any Markov embedding is isometric by [EFHN15, Theorem 13.9] it preserves orthonormality. Hence, T_φ restricted to H_1 has a countable Lebesgue spectrum. \square

Let $\varphi: X \rightarrow X$ be a measure-preserving system on the probability space $X = (X, \Sigma, \mu)$. The main tool in our proof will be the following result by Rohlin and Sinai that provides us with a σ -algebra \mathcal{F} that will serve us as a “generator of randomness”. For this to make sense, we need a notion of something that we do not consider random. One criterion of randomness can be the presence of measure-theoretic entropy. Hence, we define the “non-random” part to be

$$\Pi := \{A \in \Sigma : h_\mu(\varphi, \{A, X \setminus A\}) = 0\},$$

which we call the *Pinsker algebra* of φ and is in fact a σ -algebra. We refer to [Gla03, Theorem 18.8] for some properties of the Pinsker algebra. Most notably is the fact that if α is a finite Π -measurable partition, then $h_\mu(\varphi, \alpha) = 0$. The Pinsker algebra captures the “structured part” of a dynamical system. We say that φ has *completely positive entropy* if $\Pi = \{\emptyset, X\}$.

The Rohlin–Sinai theorem provides a way to decompose the information in a measure-preserving dynamical system into “predictable” and “random” parts. The sub- σ -algebra \mathcal{F} allows the observer to capture more and more information going forward in time, while going backward in time we do not gain more information but only arrive at the structured part.

Lemma 3.3.8 (Rohlin–Sinai theorem). *Let $\varphi: X \rightarrow X$ be a measure-preserving dynamical system on the probability space $X = (X, \Sigma, \mu)$ that is standard Lebesgue. Then there exists a sub- σ -algebra $\mathcal{F} \subseteq \Sigma$ such that*

- (i) $\varphi^{-1}(\mathcal{F}) \subseteq \mathcal{F}$,
- (ii) $\bigcup_{n=0}^{\infty} \varphi^n(\mathcal{F})$ generates Σ as a σ -algebra, and
- (iii) $\bigcap_{n=1}^{\infty} \varphi^{-n}(\mathcal{F}) = \Pi$.

Proof. Combine [Gla03, Theorem 18.7.2] and [Gla03, Theorem 18.9]. \square

Remark 3.3.9. In the situation of the above lemma and if additionally φ is invertible and has positive measure-theoretic entropy $h_\mu(\varphi) > 0$, we observe that $\Pi \subsetneq \varphi^n(\mathcal{F}) \subsetneq \varphi^{n+1}(\mathcal{F}) \subsetneq \Sigma$ for all $n \in \mathbb{Z}$. Indeed, if we assume that $\varphi^n(\mathcal{F}) = \varphi^{n+1}(\mathcal{F})$ for some $n \in \mathbb{Z}$, then by induction $\varphi^n(\mathcal{F}) = \varphi^{n+1}(\mathcal{F})$ for all $n \in \mathbb{Z}$. It follows that $\Pi = \Sigma$, which contradicts the assumption that φ has positive entropy.

If V is a closed subset of the Hilbert space H then we write $H \ominus V$ for the subspace $W := V^\perp \subseteq H$. Note that then $H = W \oplus V$ and $W \perp V$. This notation is useful, as it leaves no ambiguity in which Hilbert space H we take the orthogonally complemented subspace V^\perp .

Let (X, Σ, μ) be a measure space and $A \subseteq X$ is Σ -measurable. Then we denote by $(A, \Sigma_A, \mu|_A)$ the restricted measure space. Note that $L^2(A, \Sigma_A, \mu|_A)$ embeds into $L^2(X, \Sigma, \mu)$ by extension with the value 0 outside of A , and similarly, if Σ' is a sub- Σ -algebra of Σ the difference $L^2(A, \Sigma_A, \mu|_A) \ominus L^2(A, \Sigma'_A, \mu|_A)$ embeds into $L^2(X, \Sigma, \mu) \ominus L^2(X, \Sigma', \mu)$. We say that $A \in \Sigma$ is an *atom* if for every $B \in \Sigma$ with $B \subseteq A$ we have $\mu(B) \in \{0, \mu(A)\}$. A measure space (X, Σ, μ) is called *atomless* if Σ contains no atoms.

The following is [Rok67, Result 14.1].

Lemma 3.3.10. *Let (X, Σ, μ) be an atomless measure space with measure $\mu: \Sigma \rightarrow [0, \infty)$. Then for any distinct sub- σ -algebra $\mathcal{F} \subsetneq \Sigma$ we have that $L^2(X, \Sigma, \mu) \ominus L^2(X, \mathcal{F}, \mu)$ is infinite-dimensional.*

Proof. Since $\mathcal{F} \neq \Sigma$ we observe that $L^2(X, \Sigma, \mu) \ominus L^2(X, \mathcal{F}, \mu)$ is not trivial, and hence, there exists $0 \neq f \in L^2(X, \Sigma, \mu)$ that is orthogonal to $L^2(X, \mathcal{F}, \mu)$. For a Σ -measurable version of f consider the set $A := \{x \in X : f(x) \neq 0\}$ and notice that $\mu(A) > 0$ and notice that the positive measure space $(A, \Sigma_A, \mu|_A)$ is also atomless.

By the above discussion it suffices to show that $L^2(A, \Sigma_A, \mu|_A) \ominus L^2(A, \mathcal{F}_A, \mu|_A)$ is not finite-dimensional. This is clear if $L^2(A, \mathcal{F}_A, \mu|_A)$ is finite-dimensional, as by the lack of atoms in $(A, \Sigma_A, \mu|_A)$ the space $L^2(A, \Sigma_A, \mu|_A)$ is infinite-dimensional, and thus, the statement is trivially true. So we may suppose that $L^2(A, \mathcal{F}_A, \mu|_A)$ is infinite-dimensional. Then there exists a \mathcal{F}_A -measurable partition β of A into at least countable infinite sets, and thus, the set of indicator functions $\{\mathbb{1}_B : B \in \beta\}$ forms an infinite system of pairwise orthogonal elements in $L^2(A, \mathcal{F}_A, \mu|_A)$. Notice that $\{f \cdot \mathbb{1}_B : B \in \beta\}$ is also an infinite system of pairwise orthogonal functions, as none of the functions $f \cdot \mathbb{1}_B$ vanishes. Moreover, for each $B \in \beta$ we have $f \cdot \mathbb{1}_B \in L^2(A, \Sigma_A, \mu|_A) \ominus L^2(A, \mathcal{F}_A, \mu|_A)$, since for each $h \in L^2(A, \mathcal{F}_A, \mu|_A)$ we have $h \cdot \mathbb{1}_B \in L^2(A, \mathcal{F}_A, \mu|_A)$, and thus,

$$\langle f \cdot \mathbb{1}_B, h \rangle = \langle f, h \cdot \mathbb{1}_B \rangle = 0.$$

by the choice of f . This shows that $L^2(A, \Sigma_A, \mu|_A) \ominus L^2(A, \mathcal{F}_A, \mu|_A)$ is infinite-dimensional and concludes the proof. \square

Theorem 3.3.11. *Let $\varphi: X \rightarrow X$ be an invertible measure-preserving system on the probability space $\mathbf{X} = (X, \Sigma, \mu)$ that is standard Lebesgue. If φ has positive measure-theoretic entropy $h_\mu(\varphi) > 0$, then it has a countable Lebesgue component.*

For reference see [Rok67, Result 14.4] and also [HK02, Theorem 3.7.13].

Proof. Denote by $\mathcal{A}(X) \subseteq \Sigma$ the set of atoms in X with positive measure μ . Then by Poincaré's recurrence theorem, for each atom $A \in \mathcal{A}(X)$ there exists $k \in \mathbb{N}$ such that $\mu(\varphi^{-k}(A) \cap A) = \mu(A)$, and thus, $\mu(\varphi^{-k}(A) \Delta A) = 0$. Hence, the elements in the sequence of common refinements $\left(\bigvee_{j=0}^{n \cdot k} \varphi^{-j}(\{A, X \setminus A\}) \right)_{n \in \mathbb{N}}$ differ only by μ -null sets from each other, which implies that

$$h_\mu(\varphi, \{A, X \setminus A\}) = k \cdot \lim_{n \rightarrow \infty} \frac{1}{n \cdot k} H_\mu \left(\bigvee_{j=0}^{n \cdot k} \varphi^{-j}(\{A, X \setminus A\}) \right) = k \cdot 0 = 0$$

It follows that $\mathcal{A}(X)$ is contained in the Pinsker algebra Π of φ .

Since φ maps atoms to atoms, the set $A := \bigcup \mathcal{A}(X)$ is φ -invariant. Consequently, the complement $X \setminus A$ is also φ -invariant. By construction of A there exists an at most countable Σ -measurable partition α into atoms of positive measure.

Denote by ν_A the normalization of the non-zero measure $\mu|_A$. We notice that the restriction $\varphi|_A$ has zero h_{ν_A} -entropy. Indeed, for any finite Σ_A -measurable partition β of A we have that $\alpha \vee \beta$ consists of atoms of ν_A only, and thus, $h_{\nu_A}(\varphi|_A, \beta) \leq h_{\nu_A}(\varphi|_A, \alpha \vee \beta) = 0$ by [Gla03, Theorem 18.1.1]. As we assume that $h_\mu(\varphi) > 0$, it follows that $\mu(X \setminus A) > 0$.

Thus, let $\nu_{X \setminus A}$ denote the normalization of the measure $\mu|_{X \setminus A}$ and note that $(X \setminus A, \Sigma_{X \setminus A}, \nu_{X \setminus A})$ is atomless. We show that $\varphi|_{X \setminus A}$ has positive measure-theoretic entropy $h_{\nu_{X \setminus A}}(\varphi|_{X \setminus A})$. For any Σ -measurable partition β we have

$$\beta \vee \{A, X \setminus A\} = \{B \cap A : B \in \beta\} \cup \{B \setminus A : B \in \beta\} =: \tilde{\beta}_A \cup \tilde{\beta}_{X \setminus A}.$$

Then $\tilde{\beta}_A$ and $\tilde{\beta}_{X \setminus A}$ are partitions of A and $X \setminus A$, respectively. Moreover, $\beta_A := \tilde{\beta}_A \cup \{X \setminus A\}$ and $\beta_{X \setminus A} := \tilde{\beta}_{X \setminus A} \cup \{A\}$ are partition of X with $\beta_A \vee \beta_{X \setminus A} = \beta \vee \{A, X \setminus A\}$, and thus,

$$h_\mu(\varphi, \beta) \leq h_\mu(\varphi, \beta \vee \{A, X \setminus A\}) = h_\mu(\varphi, \beta_A \vee \beta_{X \setminus A}) \leq h_\mu(\varphi, \beta_A) + h_\mu(\varphi, \beta_{X \setminus A}).$$

Set

$$\beta_A^n := \bigvee_{k=0}^{n-1} \varphi^{-k}(\beta_A) \quad \text{and} \quad \tilde{\beta}_A^n := \bigvee_{k=0}^{n-1} \varphi^{-k}(\tilde{\beta}_A)$$

and note that by φ -invariance of A and $X \setminus A$ we have $\beta_A^n = \tilde{\beta}_A^n \cup \{X \setminus A\}$. Thus, we have

$$\begin{aligned} h_\mu(\varphi, \beta_A) &= - \lim_{n \rightarrow \infty} \frac{1}{n} H_\mu \left(\bigvee_{k=0}^{n-1} \varphi^{-k}(\beta_A) \right) = - \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{B \in \beta_A^n} \mu(B) \log(\mu(B)) \\ &= -c_A \cdot \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{B \in \tilde{\beta}_A^n} \nu_A(B) (\log(\nu_A(B)) - \log(c_A)) \\ &\quad - \lim_{n \rightarrow \infty} \frac{1}{n} \mu(X \setminus A) \log(\mu(X \setminus A)) \\ &= c_A \cdot h_{\nu_A}(\varphi|_A, \tilde{\beta}_A), \end{aligned}$$

and similarly,

$$h_\mu(\varphi, \beta_{X \setminus A}) = c_{X \setminus A} \cdot h_{\nu_{X \setminus A}}(\varphi|_{X \setminus A}, \tilde{\beta}_{X \setminus A}),$$

where $c_A := \mu(A)^{-1}$ and $c_{X \setminus A} := \mu(X \setminus A)^{-1}$ are the normalization constants of the measures $\mu|_A$ and $\mu|_{X \setminus A}$. Thus,

$$\begin{aligned} h_\mu(\varphi, \beta) &\leq c_A \cdot h_{\nu_A}(\varphi|_A, \tilde{\beta}_A) + c_{X \setminus A} \cdot h_{\nu_{X \setminus A}}(\varphi|_{X \setminus A}, \tilde{\beta}_{X \setminus A}) \\ &= c_{X \setminus A} \cdot h_{\nu_{X \setminus A}}(\varphi|_{X \setminus A}, \tilde{\beta}_{X \setminus A}). \end{aligned}$$

It follows that $h_{\nu_{X \setminus A}}(\varphi|_{X \setminus A}) > 0$. So by passing to $X \setminus A$, we may thus assume without loss of generality that $A = \emptyset$, and thus, (X, Σ, μ) is atomless. In particular, this means that (X, Σ', μ) is atomless for any sub- σ -algebra $\Sigma' \subseteq \Sigma$.

Now consider a sub- σ -algebra $\mathcal{F} \subseteq \Sigma$ as in Lemma 3.3.8 and note that Remark 3.3.9 implies

$$T_\varphi \mathbb{L}^2(X, \mathcal{F}, \mu) = \mathbb{L}^2(X, \varphi^{-1}(\mathcal{F}), \mu) \subsetneq \mathbb{L}^2(X, \mathcal{F}, \mu).$$

Thus $H_0 := \mathbb{L}^2(X, \mathcal{F}, \mu) \ominus T_\varphi \mathbb{L}^2(X, \mathcal{F}, \mu)$ is infinite-dimensional by Lemma 3.3.10.

Moreover for $k \in \mathbb{Z} \setminus \{0\}$ we have $H_0 \perp T_\varphi^k H_0$. Indeed, let $x, y \in H_0$ and suppose first that $k < 0$. Then we have $T_\varphi^k x \in T_\varphi^k H_0 = T_\varphi^k \mathbb{L}^2(X, \mathcal{F}, \mu) \ominus T_\varphi^{k+1} \mathbb{L}^2(X, \mathcal{F}, \mu)$, and since $\mathbb{L}^2(X, \mathcal{F}, \mu) \subseteq T_\varphi^{k+1} \mathbb{L}^2(X, \mathcal{F}, \mu)$ it follows that $\langle T_\varphi^k x, y \rangle = 0$. For $k > 0$ this follows by noticing that $\langle T_\varphi^k x, y \rangle = \langle x, T_\varphi^{-k} y \rangle = 0$, since T_φ is unitary.

Since it is easily verified that

$$H \ominus \mathbb{L}^2(X, \Pi, \mu) = \bigoplus_{k \in \mathbb{Z}} T_\varphi^k H_0,$$

it follows from [EFHN15, Proposition 18.32] that the operator T_φ has countable Lebesgue spectrum on $H \ominus \mathbb{L}^2(X, \Pi, \mu)$. Hence, T_φ has a countable Lebesgue component. \square

Proof of Proposition 3.3.5: By Lemma 3.3.6 there exists a factor $\psi: Y \rightarrow Y$ on a standard Lebesgue probability space $Y = (Y, \Sigma_Y, \nu)$ with $h_\nu(\psi) > 0$. Now by Theorem 3.3.11, we know that the Koopman operator T_ψ has a countable Lebesgue component. As ψ is a factor of φ we observe that also T_φ has a countable Lebesgue component from Lemma 3.3.7. \square

The following theorem now follows from Propositions 3.3.4 and 3.3.5 and can be found as [HH25, Theorem 4.9] by Hauser and the present author.

Theorem 3.3.12. *Let (X, φ) be an invertible measure-preserving dynamical system with $h_\mu(\varphi) > 0$. Then $h_{\text{apr}}(T_\varphi) = \infty$.*

There exist invertible measure-preserving dynamical systems φ with $h_\mu(\varphi) = 0$ such that T_φ has a countable Lebesgue component (and hence $h_{\text{apr}}(T_\varphi) = \infty$), as the following example demonstrates.

Example 3.3.13. Let $\alpha \in \mathbb{R}$ be irrational. Consider the skew-rotation on the two-dimensional torus

$$\varphi: \mathbb{T}^2 \rightarrow \mathbb{T}^2, \quad (x, y) \mapsto (e^{\pi i \alpha} x, xy),$$

and equip \mathbb{T} with the Lebesgue measure μ . The mapping φ becomes an invertible measure-preserving system. Moreover, it has a countable Lebesgue component [EFHN15, Proposition 18.36]. Hence, $h_{\text{apr}}(T_\varphi) = \infty$.

However, the measure-theoretic entropy $h_\mu(\varphi)$ is zero. Indeed, as presented in [Aus88, Chapter 7] the respective topological dynamical system is distal and minimal and we observe from [Yan15, Corollary 6.7] or [Gla03, Corollary 18.21] that it has zero topological entropy. From the variational principle [Wal82, Theorem 8.6] it follows that $h_\mu(\varphi) = 0$.

3.4 Topological Entropy

We can also define the classical notion of topological entropy for continuous mappings $\varphi: K \rightarrow K$ on compact Hausdorff spaces and, similar to the measure-theoretic case, it has an interpretation as a measure of information that is produced along the paths of dynamical systems.

Let us recall some basic notions. We call a collection α of open sets in K an *open cover* of K if $\bigcup_{U \in \alpha} U = K$. We can compare two open covers α and β of K by saying that α is a *subcover* of β if $\beta \subseteq \alpha$ and α is a *refinement* of β if for each $U \in \alpha$ there exists $V \in \beta$ such that $U \subseteq V$. We write $\alpha \preceq \beta$ in this case. The *join* or *common refinement* of two open covers $\alpha \vee \beta$ is defined to be

$$\alpha \vee \beta := \{U \cap V : U \in \alpha, V \in \beta\}.$$

Note that the constructions are similar to the ones for measurable partitions in Section 3.1.

For an open cover α of K we define its *covering number* $N(\alpha)$ to be the minimal cardinality of a subcover of α , namely

$$N(\alpha) := \min\{|\beta| : \beta \text{ is subcover of } \alpha\}.$$

Note that by compactness of K the covering number is always finite and bounded by $|\alpha|$. We then define the topological entropy of a continuous mapping as follows.

Definition 3.4.1 (topological entropy). Let $\varphi: K \rightarrow K$ be a continuous mapping on the compact Hausdorff space K . The *topological entropy* of φ is defined to be

$$h_{\text{top}}(\varphi) := \sup_{\alpha} \lim_{n \rightarrow \infty} \frac{1}{n} \log N \left(\bigvee_{k=0}^{n-1} \varphi^{-k}(\alpha) \right),$$

where the supremum is taken over all open covers α of K .

Note that it follows using a similar argument as in the measure-theoretic case involving Fekete's lemma that the limit in the above definition exists [Wal82, Theorem 7.1]. We refer to [Wal82, Section 7.1] for more details on topological entropy.

Similarly to the measure-theoretic entropy, the topological entropy has an interpretation as a measure of information produced along the paths of the dynamical system. Indeed, when we consider an open cover α of K as states that we can distinguish in that we can tell for each $x \in K$ and $U \in \alpha$ whether $x \in U$ or not, then the quality of $N(\alpha)$ tells us how many different states we can distinguish. The cardinality of $|\alpha|$ is not as meaningful in this regard, as for example learning that for two sets $U_1 \subseteq U_2$ with $U_1, U_2 \in \alpha$ that $x \notin U_2$ also tells us that $x \notin U_1$. So we don't actually learn much by having U_1 in the covering and can discard it, and thus, obtain a refinement of the original α . Hence, the covering number $N(\alpha)$ captures the number of distinguishable states more accurately.

If we now observe trajectories of φ starting from some $x \in K$, then after n steps we can tell for each $k \in \{0, \dots, n-1\}$ in which element of the cover $\varphi^k(x)$ lies. This is equivalent to being able to decide whether x lies in the elements of the common refinement $\bigvee_{k=0}^{n-1} \varphi^{-k}(\alpha)$. Thus, the covering number $N \left(\bigvee_{k=0}^{n-1} \varphi^{-k}(\alpha) \right)$ tells us how many different paths of length n we can distinguish when starting from different points in K . Hence, the quantity $\frac{1}{n} \log N \left(\bigvee_{k=0}^{n-1} \varphi^{-k}(\alpha) \right)$ captures the average amount of information that is produced per step along the trajectories of φ when we observe the system through the open cover α . Taking the supremum over all open covers then gives us the maximal amount of information that is produced per step along the trajectories of φ .

Recall from Section 2.3 that each topological dynamical system $\varphi: K \rightarrow K$ on a compact Hausdorff space K has invariant probability measures. We can therefore compare the topological entropy of φ with the measure-theoretic entropy of φ with respect to each invariant probability measure. The following result is classical and can be found as [Wal82, Theorem 8.6].

Theorem 3.4.2 (variational principle). *Let $\varphi: K \rightarrow K$ be a continuous mapping on the compact metric space K . Then*

$$h_{\text{top}}(\varphi) = \sup_{\mu} h_{\mu}(\varphi),$$

where the supremum is taken over all φ -invariant probability measures μ on K .

3.5 Notes

Parts of this chapter are based on the article [HH25]. The presentation of measure-theoretic entropy follows [Wal82] and [DNP87]. We have adapted their presentation to deduce the form of measure-theoretic entropy from first principles.

The concept of approximation entropy for unitary operators has been introduced in [Voi95] and does not appear to have been studied further. Our treatment of approximation entropy extends the original article by Voiculescu, in that it provides additional elementary properties and provides further tools for its computation. We especially extend his treatment by connecting fast mean ergodic convergence rates to growth of (fractional) approximation entropies in Theorem 3.2.18 (i). This theorem uses fast convergence rates as a sufficient criterion for positive fractional approximation entropy. It is open whether there exists a connection in the opposite direction.

It would be interesting to extend the concept of approximation entropy to more general operators on Banach spaces. The definition generalizes straightforwardly by replacing the Hilbert space norm by the Banach space norm. However finding lower bounds of the approximation entropy seems more challenging in this setting. One guess would be to replace the eigenvalues of the Gram matrix by s -numbers or other approximation numbers of the frame operator S for operators on Banach spaces as introduced in [Pie74]. An extension of approximation entropy to operators on Banach spaces could then be used to study the dynamics of Koopman operators on $C_b(\Omega)$ for completely regular Hausdorff spaces Ω .

The connection of positive measure-theoretic entropy to the existence of a countable Lebesgue component in the Koopman operator has been first established by Rokhlin in [Rok67]. Our presentation follows Rokhlin's presentation closely, but we have added additional details to the proof of Theorem 3.3.11 for clarity. The role of Lemma 3.3.6 in the proof of Proposition 3.3.5 is solely to extend the result from standard Lebesgue spaces to general measure-preserving dynamical systems. We did not find a comparable result in the literature.

Data-Driven Approximation of Koopman Operators

The extended Dynamic Mode Decomposition algorithm (see Section 4.4 for an explanation of the terminology) aims to approximate T_φ on a closed linear subspace $F \subseteq L^2(X)$, which is usually finite-dimensional. More precisely, it seeks to construct an operator $\hat{T}_\varphi: F \rightarrow F$ that minimizes the L^2 -error

$$\sup_{f \in F, \|f\|_2 \leq 1} \|\hat{T}_\varphi f - T_\varphi f\|_2. \quad (4.1)$$

Let $P_F: L^2(X) \rightarrow L^2(X)$ be the orthogonal projection onto F . Since for all $f, g \in F$ the inequality

$$\|P_F T_\varphi f - T_\varphi f\|_2 \leq \|g - T_\varphi f\|_2$$

holds, an exact solution \hat{T}_φ to this optimization problem exists and is given by the operator $\hat{T}_\varphi := P_F T_\varphi|_F$. It has been shown in [KM18, Theorem 2] that, under some assumptions on a spanning set \mathcal{O} of F (called the *dictionary*), the extended Dynamical Mode Decomposition algorithm converges to the linear operator $P_F T_\varphi|_F$ in the time-series limit.

Note that with limited data the eDMD algorithm will not return \hat{T}_φ exactly but only an approximation of \hat{T}_φ . We denote this approximation by \hat{T} and use it as an approximation of the original Koopman operator T_φ on the subspace F . An algorithm that computes the best approximation \hat{T}_φ exactly is known as the “analytic Dynamic Mode Decomposition” [KM18, Section 7].

It is a fundamental epistemological question in approximation theory, whether an algorithm that tries to approximate an operator on a finite-dimensional subspace can do so. And if an operator can be approximated, we can ask how well this can be done. To be more specific, we want to understand how well an approximation $\hat{T}: F \rightarrow F$ of T_φ on a subspace $F \subseteq L^2(X)$ (not necessarily an approximation from a DMD algorithm) is capable of predicting trajectories of the Koopman operator T_φ . In order to study this problem we use \hat{T} to approximate trajectories $f, T_\varphi f, \dots, T_\varphi^{K-1} f$ of the Koopman operator with trajectories of the approximation $f, \hat{T}f, \dots, \hat{T}^{K-1} f$. Ideally, we would like the error

$$\|\hat{T}^k f - T_\varphi^k f\|_2$$

to be small for all times k in a given finite time horizon $\{0, \dots, K-1\}$. One purpose of this chapter is to show the limits of such approximations and to discuss conditions under which such approximations are hard to obtain.

Before presenting negative results, let us reflect on some positive error bounds. It is fairly simple to provide a rough error bound for trajectory errors. If the approximation \hat{T} is assumed to be “good enough”, meaning that there is an $\varepsilon > 0$

such that $\|\hat{T}f - T_\varphi f\| \leq \varepsilon\|f\|$ for each $f \in F$ and if \hat{T} is a contraction, we may conclude from an induction argument that

$$\|\hat{T}^k f - T_\varphi^k f\| = \left\| \sum_{j=0}^{k-1} T_\varphi^j (\hat{T} - T_\varphi) \underbrace{\hat{T}^{(k-1)-j} f}_{\in F} \right\| \leq k\varepsilon\|f\|$$

for all $f \in F$ and all $k \in \mathbb{N}$. For small times this estimate provides bounds on the prediction error; however, for large k this estimate does not provide any value as the bound $\|\hat{T}^k f - T_\varphi^k f\| \leq 2\|f\|$ is trivial for every $f \in F$.

Instead of finding a bound that grows linearly in k , we may change perspectives and ask for conditions on a subspace $F \subseteq L^2(X)$ such that for some $\varepsilon > 0$ and some time horizon $K \in \mathbb{N}$ we have that

$$\|\hat{T}^k f - T_\varphi^k f\| \leq \varepsilon\|f\| \quad (4.2)$$

holds for all $f \in F$ and all $k \in \{0, \dots, K-1\}$. Note that in this framework the dynamical system φ , the desired error bound ε and the time horizon K are fixed, and the only freedom we have is to choose the subspace F and the approximation $\hat{T}: F \rightarrow F$. As we can always restrict attention to the case that \hat{T} is the optimal approximation $\hat{T}_\varphi = P_F T_\varphi|_F$, the only freedom we have is to choose the subspace F .

Note that for the optimal solution $\hat{T} = \hat{T}_\varphi$ it has been shown in [KM18, Theorem 5] that for any error bound $\varepsilon > 0$ and any finite time horizon $K \in \mathbb{N}$ such a subspace exists. If given an orthonormal basis $(f_n)_{n \in \mathbb{N}}$ of $L^2(X)$, then this subspace F can be chosen to be the linear span of f_1, \dots, f_N for some large enough $N \in \mathbb{N}$.

One way to describe a limitation on the choice of the subspace F is to derive a lower bound on the minimal dimension of F such that the error bound in (4.2) can be satisfied. Sections 4.2 and 4.3 will be the culmination of this chapter. There we will show in Theorem 3.2.11 that for any approximation $\hat{T}: F \rightarrow F$ that satisfies the error (4.2), the dimension of F needs to grow at least linearly with a rate governed by an approximation entropy; and in Theorem 4.2.1 we show that the dimension grows exponentially with a rate governed by the measure-theoretic entropy of the underlying dynamical system φ , in case F is a sublattice of $L^2(X)$.

Before this, we introduce the extended Dynamical Mode Decomposition algorithm and describe it in the language of functional analysis in Section 4.1. We first show that the dictionary does not need to be limited to the span of some set of given observables but that it can be chosen in a variety of other ways. Then we use synthesis and reconstruction operators to obtain a description of the best approximation $\hat{T}_\varphi = P_F T_\varphi|_F$ using a spanning dictionary. We then show how data enters the picture in approximating inner products of the form $\langle T_\varphi f_i, f_j \rangle$ and $\langle f_i, f_j \rangle$ of f_i 's in a spanning set of F . Finally, we discuss convergence rates for ergodic sampling of data and describe the structure of the optimal approximations in some special cases.

Section 4.1.1 where we describe how data can be obtained from a measure-preserving dynamical system using observables, although very basis, is original to this thesis and has not been presented in the literature before. The description of the Algorithm in Section 4.1.2 is a reformulation of the extended Dynamic Mode Decomposition algorithm as presented in [KM18]. Theorem 4.1.12 on the convergence of the eDMD

algorithm has been proved in [KM18, Theorem 2]. There some assumptions had to be used, which we managed to drop. The discussion of ergodic convergence rates in Section 4.1.3 is a very brief survey of the literature. The main results of this chapter are Theorems 3.2.11 and 4.2.1, which provide lower bounds on the dimension of subspaces F that allow for good approximations of the Koopman operator. Both results have been published by Hauser and the present author in [HH25].

4.1 The Extended Dynamic Mode Decomposition

In the following, we are going to describe the extended Dynamical Mode Decomposition in the language of functional analysis. Given a probability space $X = (X, \Sigma_X, \mu)$ and a measure-preserving system $\varphi: X \rightarrow X$ and a finite-dimensional subspace $F \subseteq L^2(X)$, we recall that the objective of the algorithm is to find a linear operator $\hat{T}_\varphi: F \rightarrow F$ that minimizes the error

$$\|\hat{T}_\varphi f - T_\varphi f\|_2$$

for all $f \in F$ simultaneously. Let us keep that in mind while we discuss the ingredients that go into the computation of the approximation. Since the DMD algorithms fall into the category of data-driven algorithms, we will need some data that we gather from the underlying dynamical system φ . The structure of the data and how the data is gathered and processed will later help us to study the quality of the approximation.

In Section 4.1.1 we formalize how we gain data from observing a measure-preserving dynamical system. We discuss two types of sampling: iid and ergodic sampling, and we introduce the concept of a dictionary \mathcal{O} , which will serve as a theoretical analog of a set of sensors. We continue to describe the subspaces $L^2(X, O^{-1}(\Sigma_Y), \mu)$ that will serve as the natural information hull, which is characterized by the property that we can approximate the Koopman operator on an finite-dimensional subspace $F \subseteq L^2(X, O^{-1}(\Sigma_Y), \mu)$ by the data provided by the dictionary \mathcal{O} (see Proposition 4.1.2).

Next we develop the extended Dynamic Mode Decomposition algorithm in Section 4.1.2 using a functional-analytic, frame-theoretic language. The optimal approximation $\hat{T}_\varphi = P_F T_\varphi|_F$ is represented by a computable surrogate $O_\varphi = M_\varphi G^\dagger$, where M_φ encodes the action of φ on some spanning set of F and G is the Gram matrix of the spanning set, which describes the geometric relation between the spanning vectors. We derive data-driven estimators for M_φ and G under iid and ergodic sampling and prove almost-sure convergence (see Theorem 4.1.12).

As the convergence for ergodic sampling in Theorem 4.1.12 is dependent on mean ergodic convergence rates, we discuss these in Section 4.1.3. While no universal rate exists for the mean ergodic theorem, we obtain quantitative bounds in special cases: $O(1/n)$ for coboundaries and an identity that controls Cesàro average errors by the decay of correlation of a trajectory. We provide references for further reading.

In Section 4.1.4 we identify the explicit form of the optimal approximation $\hat{T}_\varphi = P_F T_\varphi$ in two basic settings: (i) for $F = \text{span}\{\mathbb{1}, f\}$ the eDMD estimator reduces to an ordinary least squares regression $f \circ \varphi \approx a_1 + a_2 f$, with coefficients given in Proposition 4.1.16; (ii) if F is a finite-dimensional sublattice of $L^2(X)$, the optimal approximation \hat{T}_φ gives rise to a Markov process $(X_n)_{n \in \mathbb{N}}$, which describes transition

probabilities from $A \in \alpha$ to the preimage $\varphi^{-1}(B)$ of some other set B in a Σ -measurable partition α of X . Reversibility of the Markov chain occurs precisely when F is T_φ -invariant.

4.1.1 How to Collect Data

We begin by assuming that we have a measure-preserving dynamical system $\varphi: X \rightarrow X$ on the probability space $X = (X, \Sigma, \mu)$ and treat this as a black box. This means that we do not assume any structure on X apart from it being a probability space and on φ apart from it being measure-preserving. We need the latter assumption to obtain a bounded Koopman operator T_φ on $L^2(X)$.

The Available Data We are able to observe a dynamical system using an *observable*, which in the broadest sense is a measurable function $O: (X, \Sigma_X) \rightarrow (Y, \Sigma_Y)$, where (Y, Σ_Y) is a measurable space. We often assume that an observable is of the form

$$O: X \rightarrow \mathbb{K}^d, \quad x \mapsto (f_i(x))_{i=1}^d \quad (4.3)$$

for a set of measurable “sensors” $\mathcal{O} := \{f_i: X \rightarrow \mathbb{K} : i \in \{1, \dots, d\}\}$ that monitor the system. We also refer to these sensors as (*scalarly-valued*) *observables*. The collection of all (scalarly-valued) observables available to us is called a *dictionary*, denoted by \mathcal{O} , and is assumed to be finite.

The data that we have available is of one of two forms:

- (i) In *independent identically distributed sampling*, we assume that we have access to an independent and identically distributed (iid) sequence of measurable random variables $x_0, x_1, \dots : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (X, \Sigma)$ such that $x_0 \sim \mu$. For a given realization $\omega \in \Omega$, we treat the sequence $x_0(\omega), x_1(\omega), \dots \in X$ as initial states of the dynamical system. Up to some n we then record the dictionary’s output at these initial states

$$O(x_0(\omega)), \dots, O(x_{n-1}(\omega))$$

Note that this information does not yet contain any information on the dynamics of φ . To obtain some information about φ we record the output of the observable once again but after the initial states have been iterated a single time step

$$O(\varphi(x_0(\omega))), \dots, O(\varphi(x_{n-1}(\omega))).$$

These recordings put together serve as a dataset that feeds the algorithm.

One obvious problem with this sampling type is that in applications it is nearly impossible to reinitialize the system with initial values distributed according to μ . Note that knowledge about μ violates the above assumption that X and φ are treated as a black box in this data-driven algorithm, and thus, μ is also unknown.

- (ii) In *ergodic sampling*, we sample a time-series. We are given an initial point $x_0 \in X$, which we treat as part of the black box, that is, we don’t assume any knowledge of x_0 except for the values $O(x_0)$. Then we record the values of

the observable after each time step and continue until we reach a predefined number of time steps. The recorded data is then of the form

$$O(x_0), O(\varphi(x_0)), \dots, O(\varphi^{n-1}(x_0)).$$

As opposed to the iid sampling methods, the ergodic sampling does not need to reinitialize the system after each recording. In fact, we do not need to actively initialize the system at all. We simply start recording from the state in which the system is at the beginning of our recording. By the ergodic assumption, it becomes irrelevant from which initial point we started recording. We will show later that this almost surely leads to the desired result.

Predictable Subspaces To complete the description of the problem, it remains to specify the subspace $F \subseteq L^2(X)$ on which the approximation will be defined. Let us quickly go over sensible restrictions that we impose on F . Loosely speaking, F should be related to the observable O in that it does not carry more “information”. Otherwise the output of the algorithm is not based on the observed data.

What is it that we view as information? Informally, we say that F does not provide more information if we can predict the output of square-integrable functions in F from the values of the observable O . More formally, we could say that there is a function $p: Y \rightarrow \mathbb{K}^F$ that maps $O(x)$ to $(f(x))_{f \in F}$ for every $x \in X$. Unfortunately, there is an immediate problem with this definition, as the elements in F are equivalence classes and cannot be properly evaluated at a point. We can circumvent this with the following definition.

Definition 4.1.1. Let $X = (X, \Sigma_X, \mu)$ be a probability space, (Y, Σ_Y) be a measurable space, $O: X \rightarrow Y$ be measurable and $F \subseteq L^2(X)$ be a closed subspace. We say that O *predicts* F if there exists a family of measurable and square-integrable functions $\{f_i: X \rightarrow \mathbb{K} : i \in I\}$ and a family of measurable mappings $\{p_i: Y \rightarrow \mathbb{K} : i \in I\}$ such that

- (i) the span of the equivalence classes of the f_i 's in $L^2(X)$ is a dense subset of F and
- (ii) for every $i \in I$ the diagram

$$\begin{array}{ccc} X & \xrightarrow{O} & Y \\ & \searrow f_i & \downarrow p_i \\ & & \mathbb{K} \end{array}$$

is commutative.

Although from this definition it becomes clear how the space F does not provide more information about the system than the observable O , the definition is rather cumbersome to work with. There is another approach to predictability. In probability theory this is typically done via sub- σ -algebras. Consider the initial σ -algebra

$$O^{-1}(\Sigma_Y) := \{O^{-1}(A) : A \in \Sigma_Y\}$$

of the observable O , which by measurability is a sub- σ -algebra of Σ_X . This σ -algebra provides a “resolution” with which we can observe the state space X . The finer such a resolution is, the more information is provided; the coarser, the less.

By $L^2(X, O^{-1}(\Sigma_Y), \mu)$ we denote the space of all $O^{-1}(\Sigma_Y)$ -measurable functions in $L^2(X)$. It is easily checked that the space $L^2(X, O^{-1}(\Sigma_Y), \mu)$ is a closed sublattice of $L^2(X)$ and coincides with the range of the composition operator

$$T_O: L^2(Y, \Sigma_Y, O_*\mu) \rightarrow L^2(X, \Sigma_X, \mu), \quad f \mapsto f \circ O,$$

where $O_*\mu$ denotes the *pushforward measure* of μ under the mapping O , i.e., the measure defined by $(O_*\mu)(A) := \mu(O^{-1}(A))$ for all $A \in \Sigma_Y$.

Proposition 4.1.2. *Let $X = (X, \Sigma_X, \mu)$ be a probability space, (Y, Σ_Y) be a measurable space, $O: X \rightarrow Y$ be measurable and $F \subseteq L^2(X)$ be a closed subspace. Then O predicts F if and only if $F \subseteq L^2(X, O^{-1}(\Sigma_Y), \mu)$.*

The intuition here is the following: Since for each $A \in \Sigma_Y$, we can decide whether $O(x) \in A$, we can “predict” the observable $\mathbb{1}_{O^{-1}(A)}$. All the observables $\mathbb{1}_{O^{-1}(A)}$ for $A \in \Sigma_Y$ put together span a dense subset of $L^2(X, O^{-1}(\Sigma_Y), \mu)$. We will therefore call the space $L^2(X, O^{-1}(\Sigma_Y), \mu)$ the *information hull* of O .

Proof of Proposition 4.1.2. “ \Rightarrow ”: Suppose that O predicts F and let $\{f_i : i \in I\}$ and $\{p_i : i \in I\}$ be as in Definition 4.1.1. Since $f_i = p_i \circ O$ and f_i is square-integrable for each $i \in I$, it follows that the equivalence class of f_i in $L^2(X)$ lies in $\text{ran}(T_O) = L^2(X, O^{-1}(\Sigma_Y), \mu)$. Since these equivalence classes span a dense subspace of F it follows that $F \subseteq L^2(X, O^{-1}(\Sigma_Y), \mu)$.

“ \Leftarrow ”: Let $\{\tilde{f}_i : X \rightarrow \mathbb{K} : i \in I\}$ be a family in $L^2(X)$ that spans a dense subspace of F . Since $F \subseteq L^2(X, O^{-1}(\Sigma_Y), \mu) = \text{ran}(T_O)$ there exist $\tilde{p}_i \in L^2(Y, \Sigma_Y, O_*\mu)$ such that $\tilde{f}_i = T_O(\tilde{p}_i) = \tilde{p}_i \circ O$. For each $i \in I$, we may pick a Σ_X -measurable representative \hat{f}_i of \tilde{f}_i and a Σ_Y -measurable representative \hat{p}_i of \tilde{p}_i . Then we have μ -almost surely that $\hat{f}_i = \hat{p}_i \circ O$. Let $M_i \in \Sigma_X$ be a set with $\mu(M_i) = 1$ on which the equality holds. Set $f_i := \hat{f}_i \cdot \mathbb{1}_{M_i}$ and $p_i := \hat{p}_i \cdot \mathbb{1}_{O^{-1}(M_i)}$. Then the families $\{f_i : i \in I\}$ and $\{p_i : i \in I\}$ satisfy the conditions of Definition 4.1.1. \square

Remark 4.1.3. Note that the “ \Leftarrow ” direction of the proof shows more: If $\{\tilde{f}_i : i \in I\}$ is any family in $L^2(X)$ that spans a dense subset of F , then there exist everywhere defined Σ_X -measurable representatives $f_i: X \rightarrow \mathbb{K}$ and Σ_Y -measurable maps $p_i: Y \rightarrow \mathbb{K}$ such that $f_i = p_i \circ O$.

Proposition 4.1.2 allows us to find general conditions on which subspaces F of $L^2(X)$ the extended Dynamic Mode Decomposition can find approximations. Let us go over some examples of common subspaces.

Example 4.1.4. (i) If O is of the form (4.3) for some finite dictionary \mathcal{O} of square-integrable functions, then F can be chosen to be the span of the equivalence classes in $L^2(X)$ of the observables in \mathcal{O} .

(ii) Consider the probability space $Y = (Y, \Sigma_Y, O_*\mu)$. Suppose we have a Σ_Y -measurable partition \mathcal{P} of Y (see Definition 3.1.1) and consider the Σ_X -measurable partition $O^{-1}(\mathcal{P}) := \{O^{-1}(A) : A \in \mathcal{P}\}$. Since we can decide for

each set $A \in \mathcal{P}$, whether $O(x) \in A$ holds, we are able to decide whether $x \in B$ for each $B \in O^{-1}(\mathcal{P})$. This allows us to predict the subspace

$$F = \text{span}\{\mathbb{1}_B : B \in O^{-1}(\mathcal{P})\} \subseteq L^2(X).$$

This subspace forms a finite-dimensional sublattice of $L^2(X, O^{-1}(\Sigma_Y), \mu)$ that contains all functions that are measurable with respect to the σ -algebra generated by $O^{-1}(\mathcal{P})$. We will consider approximations of subspaces of this type in Section 4.2.

- (iii) If we can predict the subspace $F \subseteq L^2(X)$, we can always extend F by applying any lattice operation to functions in F . We may take the maximum or the minimum of multiple functions, or the modulus, positive or negative part of a single function. If the product or powers of some functions in F are square-integrable, then we may extend F with these, since both operations conserve the $O^{-1}(\Sigma_Y)$ -measurability.

Delay Observability Often the subspace F or even the sublattice $L^2(X, O^{-1}(\Sigma_Y), \mu)$ are not invariant under the Koopman operator T_φ . This allows us to extend the space of observables by delay observables.

Remark 4.1.5 (Delay observability). Let $X = (X, \Sigma_X, \mu)$ be a probability space, (Y, Σ_Y) be a measurable space, $O: X \rightarrow Y$ be a measurable observable and let $n \in \mathbb{N}_0$.

- (i) If we store the past n observations of O , then we can construct a new observable

$$O_n: X \rightarrow Y^{n+1}, \quad x \mapsto (O(x), O(\varphi(x)), \dots, O(\varphi^n(x))),$$

which we call the n -th delay observable of O . We have that for every $n \in \mathbb{N}$ that

$$L^2(X, O^{-1}(\Sigma_Y), \mu) \subseteq L^2(X, O_n^{-1}(\Sigma_Y), \mu) \subseteq L^2(X, O_{n+1}^{-1}(\Sigma_Y), \mu),$$

so the space of predictable subspaces F increases as the delay n increases.

- (ii) If O predicts the closed subspace $F \subseteq L^2(X)$ or equivalently if $F \subseteq L^2(X, O^{-1}(\Sigma_Y), \mu)$, then the subspace

$$F_n := \text{span} \bigcup_{k=0}^n T_\varphi^k F$$

is called the n -th delay subspace of F . It is easily checked that $F_n \subseteq L^2(X, O_n^{-1}(\Sigma_Y), \mu)$ for every $n \in \mathbb{N}_0$.

4.1.2 A Description of the Algorithm

We begin by describing the algorithm called “extended Dynamic Mode Decomposition” that yields a candidate for the approximation \hat{T}_φ of T_φ .

Let φ be a measure-preserving dynamical system on the probability space $X = (X, \Sigma_X, \mu)$ and let $O: X \rightarrow Y$ be a measurable observable that predicts the finite-dimensional subspace $F \subseteq L^2(X)$. Let $\mathcal{O} := \{f_1, \dots, f_d\}$ be a dictionary of everywhere defined Σ_X -measurable and square-integrable functions $f_i: X \rightarrow \mathbb{K}$ whose

equivalence classes in $L^2(X)$ span F . Then by Definition 4.1.1 there exist Σ_Y -measurable functions $\{p_1, \dots, p_d\}$ such that $f_i = p_i \circ O$ for all $i \in \{1, \dots, d\}$.

Definition 4.1.1 only yields the existence of such p_i 's and does not give any instructions on how to choose them. However, in many situations it is rather obvious how these functions need to be chosen, as the following continuation of Example 4.1.4 shows.

Example 4.1.6. (i) Continuing Example 4.1.4 (i), consider the observable

$$O: X \rightarrow \mathbb{K}^d, \quad x \mapsto (f_i(x))_{i=1}^d$$

as defined in (4.3) and let F be spanned by the equivalence classes in $L^2(X)$ of the functions in \mathcal{O} . Then

$$p_i: \mathbb{K}^d \rightarrow \mathbb{K}, \quad (z_j)_{j=1}^d \mapsto z_i$$

clearly satisfies $f_i = p_i \circ O$.

(ii) Continuing Example 4.1.4 (ii), if \mathcal{P} is a finite Σ_Y -measurable partition of Y and if $F = \text{span}\{\mathbb{1}_B : B \in \mathcal{O}^{-1}(\mathcal{P})\}$, then we may choose $\mathcal{O} = \{\mathbb{1}_B : B \in \mathcal{O}^{-1}(\mathcal{P})\} = \{\mathbb{1}_{O^{-1}(A)} : A \in \mathcal{P}\}$ and for each $A \in \mathcal{P}$ choosing $p_A := \mathbb{1}_A$ yields that $\mathbb{1}_{O^{-1}(A)} = p_A \circ O$.

By abuse of notation we will not distinguish between f_i and its equivalence class in $L^2(X)$. Whenever we evaluate f_i at a point, the reader will know that we refer to the everywhere defined function.

Since we have recorded data about the output of O , by composition with the p_i 's, we obtain the following data: In case we sample iid, we have

$$f_i(x_0(\omega)), \dots, f_i(x_{n-1}(\omega)) \quad \text{and} \quad f_i(\varphi(x_0(\omega))), \dots, f_i(\varphi(x_{n-1}(\omega))) \quad (4.4)$$

for every $i \in \{1, \dots, d\}$ and some realization $\omega \in \Omega$ of the iid random variables $x_0, x_1, \dots : \Omega \rightarrow X$ with $x_0 \sim \mu$. In case we sample ergodically, we obtain

$$f_i(x_0), f_i(\varphi(x_0)) \dots, f_i(\varphi^{n-1}(x_0)) \quad (4.5)$$

for all $i \in \{1, \dots, d\}$ and some initial state $x_0 \in X$.

Let us now explore how we can use the data to approximate the least squares solution $\hat{T}_\varphi = P_F T_\varphi$ of problem (4.1). To this end, we will use some theory from finite-dimensional frame theory. We note that the spanning set \mathcal{O} of the subspace F is a frame in the sense of [Chr16, Definition 1.1.1 and Corollary 1.1.3]. We call the operator

$$C: F \rightarrow \mathbb{K}^d, \quad f \mapsto (\langle f, f_i \rangle)_{i=1}^d = \left(\int_X f \overline{f_i} \, d\mu \right)_{i=1}^d \quad (4.6)$$

the *analysis operator* of the dictionary \mathcal{O} and its adjoint

$$C^*: \mathbb{K}^d \rightarrow F, \quad (a_i)_{i=1}^d \mapsto \sum_{i=1}^d a_i f_i$$

the *synthesis operator* of \mathcal{O} . The composition

$$S := C^*C: F \rightarrow F, \quad f \mapsto \sum_{i=1}^d \langle f, f_i \rangle f_i$$

is called the *frame operator* of \mathcal{O} and the composition

$$G := CC^*: \mathbb{K}^d \rightarrow \mathbb{K}^d, \quad (a_j)_{j=1}^d \mapsto \left(\sum_{i=1}^d a_i \langle f_i, f_j \rangle \right)_{j=1}^d \quad (4.7)$$

is called the *Gram matrix* of \mathcal{O} . In case \mathcal{O} is an orthonormal basis of F it is simple to see that S and G become identity operators.

It follows from [Chr16, Theorem 1.1.5] that the frame operator S is a self-adjoint and invertible operator on F . By using the inverse of S we can construct a left-inverse of C as

$$R := S^{-1}C^*: \mathbb{K}^d \rightarrow F, \quad (a_i)_{i=1}^d \mapsto \sum_{i=1}^d a_i S^{-1} f_i. \quad (4.8)$$

A more general statement is true: The orthogonal projection P_F onto F takes the form

$$P_F f = \sum_{i=1}^d \langle f, f_i \rangle S^{-1} f_i = \sum_{i=1}^d \langle f, S^{-1} f_i \rangle f_i$$

for all $f \in L^2(X)$. The first equality follows from noticing that the sum coincides with $S^{-1}C^*Cf = f$ for all $f \in F$ and vanishes on the annihilator F^\perp of F . The equality between the first and last term is [Chr16, Lemma 5.2.3]. Let us state some consequences of this observation.

Remark 4.1.7. (i) The optimal solution $\hat{T}_\varphi: F \rightarrow F$ of (4.1) is given by

$$\hat{T}_\varphi f = P_F T_\varphi|_F f = \sum_{i=1}^d \langle T_\varphi f, S^{-1} f_i \rangle f_i = \sum_{i=1}^d \langle T_\varphi f, f_i \rangle S^{-1} f_i$$

for every $f \in F$.

(ii) The operator R is also a right-inverse on the image of C . Indeed, let $(a_i)_{i=1}^d \in \text{ran}(C)$. Then there exists $f \in F$ such that $\langle f, f_i \rangle = a_i$ for all $i \in \{1, \dots, d\}$. Thus,

$$\begin{aligned} (C \circ R)(a_i)_{i=1}^d &= (C \circ S^{-1} \circ C^*)(\langle f, f_i \rangle)_{i=1}^d = C \sum_{i=1}^d \langle f, f_i \rangle S^{-1} f_i \\ &= \left(\left\langle \sum_{i=1}^d \langle f, f_i \rangle S^{-1} f_i, f_j \right\rangle \right)_{j=1}^d = (\langle f, f_j \rangle)_{j=1}^d = (a_j)_{j=1}^d. \end{aligned}$$

Having found a way to decompose vectors in F into coefficients in \mathbb{K}^d via the mapping C and to reconstruct the vectors using their coefficients via the mapping R , we now aim at finding a representation $O: \mathbb{K}^d \rightarrow \mathbb{K}^d$ of \hat{T}_φ such that

$$\hat{T}_\varphi = R \circ O_\varphi \circ C, \quad (4.9)$$

or equivalently, that the diagram

$$\begin{array}{ccc} F & \xrightarrow{\hat{T}_\varphi} & F \\ \downarrow C & & \uparrow R \\ \mathbb{K}^d & \xrightarrow{O_\varphi} & \mathbb{K}^d \end{array}$$

is commutative. We note that the operator O_φ is not uniquely determined.

Remark 4.1.8. Let $\hat{T}: F \rightarrow F$ be a linear operator and consider the set

$$\mathfrak{D}(\hat{T}) := \{O: \mathbb{K}^d \rightarrow \mathbb{K}^d : \hat{T} = R \circ O \circ C\}.$$

The operator

$$[O]: \text{ran}(C) \rightarrow \mathbb{K}^d / \ker(R), \quad x \mapsto Ox + \ker(R)$$

does not depend on the exact $O \in \mathfrak{D}(\hat{T})$ chosen and uniquely determines the set $\mathfrak{D}(\hat{T})$. In particular, if \mathcal{O} is a basis of F it follows that $\mathfrak{D}(\hat{T})$ only contains a single element. Notice that since R is the left-inverse of C , it follows that $\text{ran}(C)$ is isomorphic to $\mathbb{K}^d / \ker(R)$ as vector spaces.

It is easy to verify that the operator

$$O_\varphi: \mathbb{K}^d \rightarrow \mathbb{K}^d, \quad (a_i)_{i=1}^d \mapsto \left(\sum_{i=1}^d a_i \langle T_\varphi S^{-1} f_i, f_j \rangle \right)_{j=1}^d$$

satisfies the requirement of being in $\mathfrak{D}(\hat{T}_\varphi)$. Indeed, we have for $f \in F$ that

$$\begin{aligned} (R \circ O_\varphi \circ C)f &= R \left(\left\langle T_\varphi \sum_{i=1}^d \langle f, f_i \rangle S^{-1} f_i, f_j \right\rangle \right)_{j=1}^d \\ &= R(\langle T_\varphi f, f_j \rangle)_{j=1}^d \\ &= \sum_{j=1}^d \langle T_\varphi f, f_j \rangle S^{-1} f_j \\ &= P_F T_\varphi f = \hat{T}_\varphi f. \end{aligned} \quad (4.10)$$

Algorithmically this is a problem, as we do not yet have a means to estimate the values of $\langle T_\varphi S^{-1} f_i, f_j \rangle$ for all $i, j \in \{1, \dots, d\}$. This only works well if \mathcal{O} is a tight frame.

Remark 4.1.9. We say that \mathcal{O} is a *frame* for F if there exist $A, B > 0$ such that

$$A\|f\|^2 \leq \sum_{i=1}^d |\langle f, f_i \rangle|^2 \leq B\|f\|^2$$

for all $f \in F$. We call this frame *tight* if we can choose $A = B$. In this case it follows from [Chr16, Corollary 5.1.7] that $S = A \cdot \text{id}$; and therefore, we have for all $i, j \in \{1, \dots, d\}$ that

$$\sum_{i=1}^d a_i \langle T_\varphi S^{-1} f_i, f_j \rangle = \frac{1}{A} \sum_{i=1}^d a_i \langle T_\varphi f_i, f_j \rangle.$$

So it suffices to estimate the coefficients $\langle T_\varphi f_i, f_j \rangle$ for all $i, j \in \{1, \dots, d\}$ from data. We will see in Proposition 4.1.11 how this is done.

If \mathcal{O} is not a tight frame, we will need to estimate the coefficients $\langle T_\varphi S^{-1} f_i, f_j \rangle$. We recover them by splitting the operator O into two operators. For this let us calculate the pseudoinverse of the Gram matrix G . Recall that a *pseudoinverse* of an operator $G: \mathbb{K}^{d_1} \rightarrow \mathbb{K}^{d_2}$ is an operator $G^\dagger: \mathbb{K}^{d_2} \rightarrow \mathbb{K}^{d_1}$ such that $G \circ G^\dagger$ is the orthogonal projection onto $\text{ran}(G)$ and $G^\dagger \circ G$ is the orthogonal projection onto $\text{ran}(G^\dagger)$. We refer to [Chr16, Proposition 1.6.1] to see that this definition of a pseudoinverse is equivalent to the classical definition.

Lemma 4.1.10. *Let \mathcal{O} be a spanning set of F . Then the pseudoinverse G^\dagger of the Gram matrix G as defined in (4.7) is given by*

$$G^\dagger: \mathbb{K}^d \rightarrow \mathbb{K}^d, \quad (b_j)_{j=1}^d \mapsto \left(\sum_{j=1}^d b_j \langle S^{-1} f_j, S^{-1} f_i \rangle \right)_{i=1}^d.$$

Proof. Indeed, since \mathcal{O} spans F , we have $\text{ran}(G) = \text{ran}(C)$; and thus, it follows from Remark 4.1.7 (ii) that for $(b_j)_{j=1}^d \in \text{ran}(G)$ we have the identity

$$\begin{aligned} (G \circ G^\dagger)(b_j)_{j=1}^d &= G \left(\left\langle \sum_{k=1}^d b_k S^{-1} f_k, S^{-1} f_i \right\rangle \right)_{i=1}^d \\ &= \left(\sum_{i=1}^d \left\langle \sum_{k=1}^d b_k S^{-1} f_k, S^{-1} f_i \right\rangle \langle f_i, f_j \rangle \right)_{j=1}^d \\ &= \left(\left\langle \sum_{k=1}^d b_k S^{-1} f_k, \sum_{i=1}^d \langle f_j, f_i \rangle S^{-1} f_i \right\rangle \right)_{j=1}^d \\ &= \left(\left\langle \sum_{k=1}^d b_k S^{-1} f_k, P_F f_j \right\rangle \right)_{j=1}^d = \left(\left\langle \sum_{k=1}^d b_k S^{-1} f_k, f_j \right\rangle \right)_{j=1}^d \\ &= \left(\langle R(b_k)_{k=1}^d, f_j \rangle \right)_{j=1}^d = (C \circ R)(b_k)_{k=1}^d = (b_j)_{j=1}^d, \end{aligned}$$

whereas we obtain from the same calculations and the self-adjointness of S that for $(b_i)_{i=1}^d \in \text{ran}(G)^\perp = \text{ran}(C)^\perp$ we have

$$\begin{aligned} (G \circ G^\dagger)(b_j)_{j=1}^d &= \left(\left\langle \sum_{k=1}^d b_k S^{-1} f_k, f_j \right\rangle \right)_{j=1}^d \\ &= \left(\sum_{k=1}^d b_k \langle f_k, S^{-1} f_j \rangle \right)_{j=1}^d = 0. \end{aligned}$$

Thus, $G \circ G^\dagger$ is the orthogonal projection onto $\text{ran}(G)$.

It follows that $G^\dagger \circ G$ is an orthogonal projection onto $\text{ran}(G^\dagger)$ by replacing \mathcal{O} with the canonical dual frame $\tilde{\mathcal{O}} := \{S^{-1}f_i : i \in \{1, \dots, n\}\}$ (see [Chr16, page 124]). Indeed, denote by \tilde{S} , \tilde{G} and \tilde{G}^\dagger the respective operators for the dual frame $\tilde{\mathcal{O}}$. Then, clearly, $\tilde{G} = G^\dagger$ and since

$$(S \circ \tilde{S})f = S \sum_{i=1}^d \langle f, S^{-1}f_i \rangle S^{-1}f_i = \sum_{i=1}^d \langle f, S^{-1}f_i \rangle f_i = P_F f = f$$

for all $f \in F$, we have that

$$\tilde{G}^\dagger(b_j)_{j=1}^d = \left(\sum_{j=1}^d b_j \langle \tilde{S}^{-1}S^{-1}f_j, \tilde{S}^{-1}S^{-1}f_i \rangle \right)_{i=1}^d = \left(\sum_{j=1}^d b_j \langle f_j, f_i \rangle \right)_{i=1}^d = G(b_j)_{j=1}^d$$

for all $(b_j)_{j=1}^d \in \mathbb{K}^d$. Thus, it follows from the first part of the proof that $G^\dagger \circ G = \tilde{G} \circ \tilde{G}^\dagger$ is the orthogonal projection onto $\text{ran}(G^\dagger) = \text{ran}(\tilde{G})$. \square

In order to include the action of the dynamical system φ we further consider the operator

$$M_\varphi: \mathbb{K}^d \rightarrow \mathbb{K}^d, \quad (a_i)_{i=1}^d \mapsto \left(\sum_{i=1}^d a_i \langle T_\varphi f_i, f_j \rangle \right)_{i=1}^d.$$

Then we split the operator O_φ into $M_\varphi \circ G^\dagger$ and notice that for each $(b_j)_{j=1}^d \in \mathbb{K}^d$ we have

$$\begin{aligned}
(M_\varphi \circ G^\dagger)(b_j)_{j=1}^d &= M_\varphi \left(\sum_{k=1}^d b_k \langle S^{-1} f_k, S^{-1} f_i \rangle \right)_{i=1}^d \\
&= \left(\sum_{i=1}^d \left\langle \sum_{k=1}^d b_k S^{-1} f_k, S^{-1} f_i \right\rangle \langle T_\varphi f_i, f_j \rangle \right)_{j=1}^d \\
&= \left(\left\langle \sum_{k=1}^d b_k S^{-1} f_k, \sum_{i=1}^d \langle f_j, T_\varphi f_i \rangle S^{-1} f_i \right\rangle \right)_{j=1}^d \quad (4.11) \\
&= \left(\left\langle \sum_{k=1}^d b_k S^{-1} f_k, P_F T_\varphi^* f_j \right\rangle \right)_{j=1}^d \\
&= \left(\sum_{k=1}^d b_k \langle T_\varphi S^{-1} f_k, f_j \rangle \right)_{j=1}^d = O_\varphi (b_j)_{j=1}^d.
\end{aligned}$$

So obtaining the operator O_φ from data can be reduced to estimating the coefficients

$$\langle T_\varphi f_i, f_j \rangle \quad \text{and} \quad \langle f_i, f_j \rangle$$

for each $i, j \in \{1, \dots, d\}$ from data in order to construct approximations of M_φ and G .

Using the data from (4.4) observed through the observables while sampling independently identically distributed, the strong law of large numbers justifies the use of the approximations

$$\begin{aligned}
M_{\varphi, n, \omega} : \mathbb{K}^d &\rightarrow \mathbb{K}^d, & (a_i)_{i=1}^d &\mapsto \left(\sum_{i=1}^d \frac{a_i}{n} \sum_{k=0}^{n-1} f_i(\varphi(x_k(\omega))) \overline{f_j(x_k(\omega))} \right)_{j=1}^d \\
G_{n, \omega} : \mathbb{K}^d &\rightarrow \mathbb{K}^d, & (a_i)_{i=1}^d &\mapsto \left(\sum_{i=1}^d \frac{a_i}{n} \sum_{k=0}^{n-1} f_i(x_k(\omega)) \overline{f_j(x_k(\omega))} \right)_{j=1}^d
\end{aligned}$$

for some realization $\omega \in \Omega$ of the operators M_φ and G .

Similarly using the data from (4.5) observed through ergodic sampling, Birkhoff's pointwise ergodic theorem (see Proposition 2.2.1) gives justification to use the approximations

$$\begin{aligned}
M_{\varphi, n, x_0} : \mathbb{K}^d &\rightarrow \mathbb{K}^d, & (a_i)_{i=1}^d &\mapsto \left(\sum_{i=1}^d \frac{a_i}{n} \sum_{k=0}^{n-1} f_i(\varphi^{k+1}(x_0)) \overline{f_j(\varphi^k(x_0))} \right)_{j=1}^d \\
G_{n, x_0} : \mathbb{K}^d &\rightarrow \mathbb{K}^d, & (a_i)_{i=1}^d &\mapsto \left(\sum_{i=1}^d \frac{a_i}{n} \sum_{k=0}^{n-1} f_i(\varphi^k(x_0)) \overline{f_j(\varphi^k(x_0))} \right)_{j=1}^d
\end{aligned}$$

for some initial state $x_0 \in X$.

Let us note their convergence in the next lemma.

Proposition 4.1.11. *Let φ be a measure-preserving dynamical system on the probability space $X = (X, \Sigma_X, \mu)$.*

- (i) *Let $x_0, x_1, \dots : \Omega \rightarrow X$ be an independent identically distributed sequence of random variables with $x_0 \sim \mu$. Then*

$$\lim_{n \rightarrow \infty} M_{\varphi, n, \omega} = M_{\varphi} \quad \text{and} \quad \lim_{n \rightarrow \infty} G_{n, \omega} = G$$

for \mathbb{P} -almost every $\omega \in \Omega$. If additionally every observable $f \in \mathcal{O}$ is bounded, then there is $\Omega_0 \in \mathcal{A}$ with $\mathbb{P}(\Omega_0) = 1$ such that for all $\delta > 0$ we have

$$\lim_{n \rightarrow \infty} n^{1/2-\delta} \|M_{\varphi, n, \omega} - M_{\varphi}\| = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} n^{1/2-\delta} \|G_{n, \omega} - G\| = 0.$$

- (ii) *If φ is ergodic, then*

$$\lim_{n \rightarrow \infty} M_{\varphi, n, x_0} = M_{\varphi} \quad \text{and} \quad \lim_{n \rightarrow \infty} G_{n, x_0} = G$$

for μ -almost every $x_0 \in X$.

Proof. “(i)”: By the strong law of large numbers we have for each $i, j \in \{1, \dots, d\}$ that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f_i(\varphi(x_k(\omega))) \overline{f_j(x_k(\omega))} = \int_X f_i(\varphi(x)) \overline{f_j(x)} \mu(dx) = \langle T_{\varphi} f_i, f_j \rangle$$

So for \mathbb{P} -almost every $\omega \in \Omega$ each entry of $M_{\varphi, n, \omega}$ converges to the respective entry of M_{φ} , which by finite-dimensionality implies the claimed convergence.

To obtain the convergence rate, suppose that the observables in \mathcal{O} are bounded. Then, since the f_i 's are assumed bounded integrable, we have that $(f_i \circ \varphi) \cdot f_j$ is square-integrable, and thus, by [MNR12, Satz 3.5] for each $i, j \in \{1, \dots, d\}$ there exists $\Omega_{i,j} \in \mathcal{A}$ with $\mathbb{P}(\Omega_{i,j}) = 1$ such that we have

$$\lim_{n \rightarrow \infty} n^{-1/2+\delta} \left| \sum_{k=0}^{n-1} f_i(\varphi(x_k(\omega))) \overline{f_j(x_k(\omega))} - \langle T_{\varphi} f_i, f_j \rangle \right| = 0.$$

for all $\delta > 0$ and all $\omega \in \Omega_{i,j}$. Setting $\Omega_0 := \bigcup_{i,j=1}^d \Omega_{i,j}$ yields $\mathbb{P}(\Omega_0) = 1$ and for all $\omega \in \Omega_0$ we have

$$\lim_{n \rightarrow \infty} n^{-1/2+\delta} \|M_{\varphi, n, \omega} - M_{\varphi}\|_{\mathbb{F}} = 0,$$

where $\|\cdot\|_{\mathbb{F}}$ denotes the Frobenius norm. By the equivalence of norms in finite-dimensions, the statement follows.

The proof for $G_{n, \omega} \rightarrow G$ is analogous.

“(ii)”: Since the f_i 's are assumed square-integrable, we have that $(f_i \circ \varphi) \cdot \overline{f_j}$ is integrable for all $i, j \in \{1, \dots, d\}$. Thus, we obtain from Birkhoff's pointwise ergodic theorem (see Theorem 2.2.1) that for each $i, j \in \{1, \dots, d\}$ we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f_i(\varphi^{k+1}(x_0)) \overline{f_j(\varphi^k(x_0))} = \int_X f_i(\varphi(x)) \overline{f_j(x)} \mu(dx) = \langle T_\varphi f_i, f_j \rangle$$

for μ -almost every $x_0 \in X_0$. So it follows that M_{φ, n, x_0} converges to M_φ for μ -almost every $x_0 \in X$. \square

We will now see that the compositions

$$M_{\varphi, n, \omega} \circ G_{n, \omega}^\dagger \quad \text{and} \quad M_{\varphi, n, x_0} \circ G_{n, x_0}^\dagger$$

both approximate $O_\varphi = M_\varphi \circ G^\dagger$ as $n \rightarrow \infty$ in some sense that is to be specified. One of the difficulties we face in the proof is that the operation of taking the pseudoinverse is not continuous in general. The following theorem generalizes part of the assertion of [KM18, Theorem 2] as we may drop the assumption of μ -independence of the observables $\{f_1, \dots, f_d\}$.

Theorem 4.1.12. *Let φ be a measure-preserving dynamical system on the probability space $X = (X, \Sigma_X, \mu)$.*

- (i) *Let $x_0, x_1, \dots : \Omega \rightarrow X$ be an independent identically distributed sequence of random variables with $x_0 \sim \mu$. Then*

$$\lim_{n \rightarrow \infty} (R \circ M_{\varphi, n, \omega} \circ G_{n, \omega}^\dagger \circ C) = \hat{T}_\varphi$$

for \mathbb{P} -almost every $\omega \in \Omega$.

- (ii) *If φ is ergodic, then*

$$\lim_{n \rightarrow \infty} (R \circ M_{\varphi, n, x_0} \circ G_{n, x_0}^\dagger \circ C) = \hat{T}_\varphi$$

for μ -almost every $x_0 \in X$.

Proof. By means of Proposition 4.1.11 we may treat both cases of sampling simultaneously and write G_n instead of $G_{n, \omega}$ or G_{n, x_0} as well as $M_{\varphi, n}$ instead of $M_{\varphi, n, \omega}$ or M_{φ, n, x_0} . By choosing ω or x_0 outside of the respective null set, we may also assume that $G_n \rightarrow G$ and $M_{\varphi, n} \rightarrow M_\varphi$.

We begin by showing that $G_n^\dagger \circ G \circ G^\dagger$ converges to $G^\dagger \circ G \circ G^\dagger = G^\dagger$ as $n \rightarrow \infty$. Since $\text{ran}(G \circ G^\dagger \circ G_n) \subseteq \text{ran}(G \circ G^\dagger)$ and $G \circ G^\dagger$ is the orthogonal projection onto $\text{ran}(G)$ it follows that $\text{rank}(G \circ G^\dagger \circ G_n) \leq \text{rank}(G)$ for all $n \in \mathbb{N}$. Conversely, since $G \circ G^\dagger \circ G_n \rightarrow G \circ G^\dagger \circ G = G$ and the set $\{K : \mathbb{K}^d \rightarrow \mathbb{K}^d : \text{rank}(K) \leq k\}$ is closed for all $k \in \mathbb{N}$ it follows that $\text{rank}(G \circ G^\dagger \circ G_n) \geq \text{rank}(G)$ for large enough $n \in \mathbb{N}$. It follows that there exists $n_0 \in \mathbb{N}$ such that

$$\text{rank}(G \circ G^\dagger \circ G_n) = \text{rank}(G)$$

for all $n \geq n_0$. Since $G \circ G^\dagger$ is an orthogonal projection, we obtain $(G \circ G^\dagger \circ G_n)^\dagger = G_n^\dagger \circ (G \circ G^\dagger)^\dagger = G_n^\dagger \circ (G^\dagger)^\dagger \circ G^\dagger = G_n^\dagger \circ G \circ G^\dagger$ for each $n \in \mathbb{N}$. Now applying [Ste69, Theorem 5.2] yields $G_n^\dagger \circ G \circ G^\dagger = (G \circ G^\dagger \circ G_n)^\dagger \rightarrow G^\dagger$.

Since $\text{ran}(C) = \text{ran}(G)$ we further obtain the convergence $G_n^\dagger \circ C = G_n^\dagger \circ G \circ G^\dagger \circ C \rightarrow G^\dagger \circ C$. Since also $M_{\varphi,n} \rightarrow M_\varphi$ it follows that

$$R \circ M_{\varphi,n} \circ G_n^\dagger \circ C \rightarrow R \circ M_\varphi \circ G^\dagger \circ C.$$

The claim now follows from the identities (4.11) and (4.10). \square

Remark 4.1.13. (i) Note that in general the mapping that maps a matrix to its pseudoinverse is not continuous [Rak97, Example 4.1]. The fact that we obtain convergence of G_n^\dagger to G^\dagger relies heavily on the fact that we restrict the operator to $\text{ran}(C) = \text{ran}(G)$. The proof makes no statement on the behaviour of G_n^\dagger on $\text{ran}(C)^\perp$. Therefore, we cannot guarantee that the sequence $(G_n^\dagger)_{n \in \mathbb{N}}$ is bounded, as this already implies the convergence of G_n^\dagger to G^\dagger on the entire space \mathbb{K}^d [Rak97, Theorem 4.3].

When \mathcal{O} is a basis of F this problem vanishes, as then $\text{ran}(C) = \mathbb{K}^d$.

(ii) The algorithm heavily simplifies when we know that the \mathcal{O} is an orthonormal basis of F . In this case we know that $G = I$ and therefore the estimation of G^\dagger via G_n^\dagger need not take place. We obtain readily from Proposition 4.1.11 that

$$\lim_{n \rightarrow \infty} (R \circ M_{\varphi,n} \circ C) = \hat{T}_\varphi$$

where $M_{\varphi,n}$ stands for $M_{\varphi,n,\omega}$ or M_{φ,n,x_0} with $\omega \in \Omega$ and $x_0 \in X$ outside of the corresponding exceptional null set.

This frame perspective of the extended Dynamic Mode Decomposition algorithm may offer new possibilities for approximating G^\dagger , and therefore, methods of developing new DMD algorithms, for example by approximating S^{-1} using established methods for frames (see [Chr16, Section 23]).

4.1.3 A Discussion on Ergodic Convergence Rates

For ergodic sampling, the convergence rates are more difficult to obtain. Part of the problem here is that there is no general convergence rate for the mean ergodic theorem. Nonetheless we can provide convergence rates in special cases.

Proposition 4.1.14. *Let φ be a measure-preserving dynamical system on the probability space $X = (X, \Sigma_X, \mu)$. Fix $f \in L^2(X)$ and consider the Cesàro averages*

$$A_n[T](f) := \frac{1}{n} \sum_{k=0}^{n-1} T_\varphi^k f$$

for each $n \in \mathbb{N}$.

(i) *Suppose that φ is strongly mixing. Then there exists $g \in L^2(X)$ such that $f = g - T_\varphi g$ if and only if there exists $C > 0$ such that $\|A_n[T](f)\|_{L^2} \leq C/n$ holds for all $n \in \mathbb{N}$.*

- (ii) Suppose that φ is ergodic. Then there exists $g \in L^\infty(X)$ such that $f = g - T_\varphi g$ if and only if there exists a constant $C > 0$ such that $|A_n[T](f)| \leq C/n$ holds μ -almost everywhere for all $n \in \mathbb{N}$.

The first statement of Proposition 4.1.14 is [Kac96, Remark 1 in Section 1.1], and the second is [Kac96, Theorem 19]. The author is not aware whether (i) in the above proposition holds without assuming that φ is strongly mixing or (ii) holds without assuming that φ is ergodic.

The following proposition links the convergence in $L^2(X)$ -norm of the Cesàro means to the decay of correlation coefficients $(\langle x, T^k x \rangle)_{k \in \mathbb{N}}$. It is in particular applicable to Koopman operators T_φ of a measure-preserving system φ on the probability space X . Note that this result is related to Theorem 2 in [Kac96].

Proposition 4.1.15. *Let $T: H \rightarrow H$ be an isometry on a Hilbert space H and let $x \in H$. Then*

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} T^k x \right\|^2 = \frac{2}{n} \sum_{k=0}^{n-1} \left(1 - \frac{k}{n}\right) \operatorname{Re} \langle x, T^k x \rangle - \frac{1}{n} \|x\|^2 \leq \frac{2}{n} \sum_{k=0}^{n-1} |\langle x, T^k x \rangle|$$

Proof. Since T is an isometry, we obtain for all $x \in H$ and all $n \in \mathbb{N}$ that

$$\begin{aligned} \left\| \frac{1}{n} \sum_{k=0}^{n-1} T^k x \right\|_{L^2}^2 &= \left\langle \frac{1}{n} \sum_{k=0}^{n-1} T^k x, \frac{1}{n} \sum_{k=0}^{n-1} T^k x \right\rangle \\ &= \frac{1}{n^2} \sum_{k=0}^{n-1} \langle T^k x, T^k x \rangle + \frac{1}{n^2} \sum_{k=0}^{n-1} \sum_{l=k+1}^{n-1} (\langle T^k x, T^l x \rangle + \langle T^l x, T^k x \rangle) \\ &= \frac{1}{n^2} \sum_{k=0}^{n-1} \|x\|^2 + \frac{1}{n^2} \sum_{k=0}^{n-1} \sum_{l=1}^{n-k-1} (\langle T^k x, T^{l+k} x \rangle + \langle T^{k+l} x, T^k x \rangle) \\ &= \frac{1}{n} \|x\|^2 + \frac{1}{n^2} \sum_{k=0}^{n-1} \sum_{l=1}^{n-k-1} (\langle x, T^l x \rangle + \langle T^l x, x \rangle) \\ &= \frac{2}{n} \|x\|^2 + \frac{2}{n^2} \sum_{l=1}^{n-1} (n-l) \operatorname{Re} \langle x, T^l x \rangle - \frac{1}{n} \|x\|^2 \\ &= \frac{2}{n} \operatorname{Re} \langle x, x \rangle + \frac{2}{n} \sum_{l=1}^{n-1} \left(1 - \frac{l}{n}\right) \operatorname{Re} \langle x, T^l x \rangle - \frac{1}{n} \|x\|^2 \\ &= \frac{2}{n} \sum_{l=0}^{n-1} \left(1 - \frac{l}{n}\right) \operatorname{Re} \langle x, T^l x \rangle - \frac{1}{n} \|x\|^2. \end{aligned}$$

The remaining inequality is trivial. □

The power of this result lies in the statements that we may conclude from it: If given a convergence rate of the correlation coefficients $(\langle x, T^k x \rangle)_{k \in \mathbb{N}}$ we may deduce from it a rate for the convergence of the Cesàro means $\left\| \frac{1}{n} \sum_{k=0}^{n-1} T^k x \right\|$. We refer to [Kac96, Theorem 6] for an example of such a result.

The convergence rate in the von Neumann's ergodic theorem can further be described by the behaviour of the spectral measure around the point 1 of the spectrum of an invertible operator as is described in [Kac96, Theorem 3] and [ACOS24, Lemma 1]. This thesis will introduce spectral measures in Section 3.2.3.

The convergence rate of the Cesàro means of observables can be characterized by the containment of observables in certain fractional coboundaries. This has been shown in [DL01] and extended for more general coboundaries in [GHT11].

Lastly, suppose there exists a rate $(a_n)_{n \in \mathbb{N}}$ with $a_n \geq 0$ and $a_n \rightarrow 0$ such that for each $f \in L^2(X)$ there exists $C_f > 0$ with

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} T_\varphi^k f - Pf \right\| \leq C_f a_n$$

for all $n \in \mathbb{N}$, where P denotes the mean ergodic projection of T_φ onto $\text{fix}(T_\varphi)$. Then the uniform boundedness principle implies that there exists $C > 0$ such that

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} T_\varphi^k - P \right\| \leq C a_n.$$

Therefore, the convergence of the Cesàro means occurs in the uniform operator topology. This leads to uniform ergodicity of T_φ , which will be characterized for lattice homomorphisms and specifically for Koopman operators in Chapter 6.

4.1.4 The Structure of the Optimal Approximation

Let $\varphi: X \rightarrow X$ be a measure-preserving dynamical system on the probability space $X = (X, \Sigma, \mu)$. In the previous sections we have seen how to approximate the Koopman operator $T_\varphi: L^2(X) \rightarrow L^2(X)$ on a finite-dimensional subspace $F \subseteq L^2(X)$ using data sampled from φ and have proved in Theorem 4.1.12 that the approximation converges to the best approximation of T_φ given by

$$\hat{T}_\varphi = P_F T_\varphi$$

on the subspace F , where $P_F: L^2(X) \rightarrow L^2(X)$ denotes the orthogonal projection onto F .

In this section we will discuss the structure of this best approximation in the case where F is

- (i) given by $F := \text{span}\{\mathbb{1}, f\}$ for a single non-constant and real-valued observable $f \in L^2(X)$, and
- (ii) given by $F := L^2(X | \alpha) := \text{span}\{\mathbb{1}_A : A \in \alpha\}$ for some finite Σ -measurable partition α of X .

Let us begin with the first case.

Proposition 4.1.16. *Let $X = (X, \Sigma_X, \mu)$ be a probability space and let $\varphi: X \rightarrow X$ be a measure-preserving dynamical system on X . Suppose that $f \in L^2(X)$ is a*

non-constant and real-valued observable and set $F := \text{span}\{\mathbb{1}, f\}$. Then the optimal approximation $\hat{T}_\varphi = P_F T_\varphi$ of the Koopman operator T_φ on F is given by

$$\hat{T}_\varphi g = \frac{\langle g, \mathbb{1} \rangle \langle f, f \rangle - \langle T_\varphi g, f \rangle \langle f, \mathbb{1} \rangle}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \cdot \mathbb{1} + \frac{\langle T_\varphi g, f \rangle - \langle g, \mathbb{1} \rangle \langle f, \mathbb{1} \rangle}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \cdot f$$

for all $g \in F$. In particular, the matrix representation O_φ of $\hat{T}_\varphi = R \circ O_\varphi \circ C$ with respect to the basis $(\mathbb{1}, f)$ is given

$$O_\varphi = \begin{pmatrix} 1 & a_1 \\ 0 & a_2 \end{pmatrix},$$

where

$$a_1 := \frac{\langle f, \mathbb{1} \rangle (\langle f, f \rangle - \langle T_\varphi f, f \rangle)}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \quad \text{and} \quad a_2 := \frac{\langle T_\varphi f, f \rangle - \langle f, \mathbb{1} \rangle^2}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2}.$$

and R and C are the coordinate and reconstruction maps with respect to the basis $(\mathbb{1}, f)$ as defined in (4.6) and (4.8).

The extended Dynamic Mode Decomposition algorithms will therefore compute approximations of the coefficients a_1 and a_2 from data sampled from φ . In particular, we obtain a formula for the predicted value of the observable f along trajectories of φ given its current value by

$$f(\varphi(x)) \approx \hat{T}_\varphi f(x) = a_1 + a_2 \cdot f(x).$$

Hence, the extended Dynamic Mode Decomposition algorithm calculates an ordinary least squares linear regression of the observable f against its time-shifted version $f \circ \varphi$. Its error in L^2 -norm is given by the limit of the R^2 -number of the regression

$$\|T_\varphi f - \hat{T}_\varphi f\|_2^2 = \int_X (f \circ \varphi - a_1 - a_2 \cdot f)^2 \, d\mu$$

and is in fact minimized by $\hat{T}_\varphi f$ among all affine linear functions of f .

Proof of the Proposition 4.1.16. In order to compute the orthogonal projection onto F we need to orthogonalize the set $\{\mathbb{1}, f\}$. Noting that $\langle f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}, \mathbb{1} \rangle = 0$ we obtain that the set $F = \text{span}\{\mathbb{1}, f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\}$ is an orthogonal basis of F . Thus, the orthogonal projection $P_F: L^2(X) \rightarrow L^2(X)$ onto F is given by

$$\begin{aligned} P_F g &= \langle g, \mathbb{1} \rangle \cdot \mathbb{1} + \frac{\langle g, f \rangle - \langle g, \mathbb{1} \rangle \langle f, \mathbb{1} \rangle}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \cdot (f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}) \\ &= \frac{\langle g, \mathbb{1} \rangle \langle f, f \rangle - \langle g, f \rangle \langle f, \mathbb{1} \rangle}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \cdot \mathbb{1} + \frac{\langle g, f \rangle - \langle g, \mathbb{1} \rangle \langle f, \mathbb{1} \rangle}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \cdot f \end{aligned}$$

for all $g \in L^2(X)$. Therefore, we obtain for the optimal approximation $\hat{T}_\varphi = P_F T_\varphi$ of the Koopman operator T_φ on F that

$$\begin{aligned}\hat{T}_\varphi g &= \frac{\langle T_\varphi g, \mathbb{1} \rangle \langle f, f \rangle - \langle T_\varphi g, f \rangle \langle f, \mathbb{1} \rangle}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \cdot \mathbb{1} + \frac{\langle T_\varphi g, f \rangle - \langle T_\varphi g, \mathbb{1} \rangle \langle f, \mathbb{1} \rangle}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \cdot f \\ &= \frac{\langle g, \mathbb{1} \rangle \langle f, f \rangle - \langle T_\varphi g, f \rangle \langle f, \mathbb{1} \rangle}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \cdot \mathbb{1} + \frac{\langle T_\varphi g, f \rangle - \langle g, \mathbb{1} \rangle \langle f, \mathbb{1} \rangle}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \cdot f\end{aligned}$$

for all $g \in L^2(X)$. Let us evaluate the action of \hat{T}_φ on the basis elements of F to obtain the matrix representation O_φ of \hat{T}_φ with respect to the basis $(\mathbb{1}, f)$ to obtain

$$\begin{aligned}\hat{T}_\varphi \mathbb{1} &= \frac{\langle f, f \rangle - \langle \mathbb{1}, f \rangle \langle f, \mathbb{1} \rangle}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \cdot \mathbb{1} + \frac{\langle \mathbb{1}, f \rangle - \langle \mathbb{1}, \mathbb{1} \rangle \langle f, \mathbb{1} \rangle}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \cdot f \\ &= \mathbb{1} + \frac{\langle \mathbb{1}, f \rangle - \langle \mathbb{1}, f \rangle}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \cdot f = \mathbb{1}\end{aligned}$$

as expected and

$$\begin{aligned}\hat{T}_\varphi f &= \frac{\langle f, \mathbb{1} \rangle \langle f, f \rangle - \langle T_\varphi f, f \rangle \langle f, \mathbb{1} \rangle}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \cdot \mathbb{1} + \frac{\langle T_\varphi f, f \rangle - \langle f, \mathbb{1} \rangle^2}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2} \cdot f \\ &= \underbrace{\frac{\langle f, \mathbb{1} \rangle (\langle f, f \rangle - \langle T_\varphi f, f \rangle)}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2}}_{a_1 :=} \cdot \mathbb{1} + \underbrace{\frac{\langle T_\varphi f, f \rangle - \langle f, \mathbb{1} \rangle^2}{\|f - \langle f, \mathbb{1} \rangle \cdot \mathbb{1}\|_2^2}}_{a_2 :=} \cdot f.\end{aligned}$$

Thus, the assertion follows. \square

Let us now turn to the second case, where we approximate the Koopman operator T_φ on the subspace $F = L^2(X | \alpha)$ for a finite Σ -measurable partition α of X .

Proposition 4.1.17. *Let $X = (X, \Sigma_X, \mu)$ be a probability space and let $\varphi: X \rightarrow X$ be a measure-preserving dynamical system on X . Suppose that α is a finite Σ -measurable partition of X with $\mu(A) > 0$ for all $A \in \alpha$ and set $F := L^2(X | \alpha) = \text{span}\{\mathbb{1}_A : A \in \alpha\}$. Then the optimal approximation $\hat{T}_\varphi = P_F T_\varphi$ of the Koopman operator T_φ on F is given by*

$$\hat{T}_\varphi \mathbb{1}_B = \sum_{A \in \alpha} \frac{\mu(A \cap \varphi^{-1}(B))}{\mu(A)} \cdot \mathbb{1}_A = \sum_{A \in \alpha} \mu(\varphi^{-1}(B) | A) \cdot \mathbb{1}_A \quad (4.12)$$

for all $B \in \alpha$. In particular, the matrix representation O_φ of $\hat{T}_\varphi = R \circ O_\varphi \circ C$ with respect to the basis $(\mathbb{1}_A)_{A \in \alpha}$ is of a column-stochastic matrix $O_\varphi = \{(p_{A,B})\}_{B,A \in \alpha}$, where B denotes the line and A the column index, its entries are the transition probabilities

$$p_{A,B} := \frac{\mu(A \cap \varphi^{-1}(B))}{\mu(A)} = \mu(\varphi^{-1}(B) | A),$$

and R and C are the coordinate and reconstruction maps with respect to the basis $(\mathbb{1}_A)_{A \in \alpha}$ as defined in (4.6) and (4.8). Together with the steady-state $(\pi_A)_{A \in \alpha}$ given by $\pi_A := \mu(A)$ for all $A \in \alpha$, this defines a time-homogeneous Markov chain $(X_n)_{n \in \mathbb{N}}$ with values in the finite state space α .

Note that O_φ is the transpose of the row-stochastic matrix that is usually used to denote the transition matrix of a Markov chain. Here, we use the convention that the columns sum to one in order to be consistent with our previous use of this notation. Therefore, we denote the fact that π is a steady-state of the Markov chain by $O_\varphi \pi = \left(\sum_{A \in \alpha} \pi_A p_{A,B} \right)_{B \in \alpha} = \pi$ and treat π as a column vector.

Since we approximate the time-reversing Koopman operator, the Markov chain answers questions going backwards in time along trajectories of φ . This allows us to answer questions like “What is the probability of having started a trajectory of φ in B , if we know that after n time steps the trajectory lies in A ?” or formally

$$\mathbb{P}(X_n = B \mid X_0 = A) = \sum_{A_1, \dots, A_{n-1} \in \alpha} \prod_{k=0}^{n-1} \mu(\varphi^{-1}(A_{k+1}) \mid A_k),$$

where $A_0 := A$ and $A_{-n} := B$. As we can only use information about containment of trajectories of φ in sets of α in order to construct the Markov chain, this value need not coincide with the coefficients $\mu(\varphi^{-n}(B) \mid A)$ of the approximation of the iterated Koopman operator

$$P_F T_\varphi^n \mathbb{1}_B = \sum_{A \in \alpha} \frac{\mu(A \cap \varphi^{-n}(B))}{\mu(A)} \cdot \mathbb{1}_A = \sum_{A \in \alpha} \mu(\varphi^{-n}(B) \mid A) \cdot \mathbb{1}_A,$$

which give a more accurate answer to the question. We may however view the probability $\mathbb{P}(X_n = B \mid X_0 = A)$ as an approximation of $\mu(\varphi^{-n}(B) \mid A)$. More precisely, since we have

$$(P_F T_\varphi)^n \mathbb{1}_B = \sum_{A_0, \dots, A_{n-1} \in \alpha} \prod_{k=0}^{n-1} \mu(\varphi^{-1}(A_{k+1}) \mid A_k) \cdot \mathbb{1}_{A_0},$$

for all $B \in \alpha$, we may deduce the following error estimate

$$\begin{aligned} |\mathbb{P}(X_n = B \mid X_0 = A) - \mu(\varphi^{-n}(B) \mid A)| &\leq | \langle (P_F T_\varphi)^n \mathbb{1}_B - P F T_\varphi^n \mathbb{1}_B, \mu(A)^{-2} \cdot \mathbb{1}_A \rangle | \\ &\leq \| (P_F T_\varphi)^n \mathbb{1}_B - P F T_\varphi^n \mathbb{1}_B \| \leq \| \hat{T}_\varphi^{n-1} \mathbb{1}_B - T_\varphi^{n-1} \mathbb{1}_B \| \end{aligned}$$

for all $n \in \mathbb{N}$.

Proof of the Proposition 4.1.17. Notice that equation (4.12) follows from the fact that the orthogonal projection $P_F: L^2(X) \rightarrow L^2(X)$ onto $F = L^2(X \mid \alpha)$ is given by the conditional expectation operator (see Lemma 4.2.6). The fact that O_φ defines a time-homogeneous Markov chain with steady-state $(\pi_A)_{A \in \alpha}$ follows from the fact that O_φ is a column-stochastic matrix, since

$$\sum_{B \in \alpha} p_{A,B} = \sum_{B \in \alpha} \frac{\mu(A \cap \varphi^{-1}(B))}{\mu(A)} = \sum_{B \in \alpha} \mu(\varphi^{-1}(B) \mid A) = 1.$$

That $(\pi_A)_{A \in \alpha}$ is a steady-state follows from the measure-preserving property of φ , as

$$\begin{aligned} \sum_{A \in \alpha} \pi_A p_{A,B} &= \sum_{A \in \alpha} \mu(A) \frac{\mu(A \cap \varphi^{-1}(B))}{\mu(A)} = \sum_{A \in \alpha} \mu(A \cap \varphi^{-1}(B)) \\ &= \mu(\varphi^{-1}(B)) = \mu(B) = \pi_B \end{aligned}$$

for all $B \in \alpha$. □

Remark 4.1.18 (time-reversal). One might wonder why the Markov chain is one that answers questions going backwards in time along trajectories of φ . Suppose that φ is invertible. If we were to approximate the forward Perron–Frobenius operator

$$P_\varphi: L^2(X) \rightarrow L^2(X), \quad P_\varphi f = f \circ \varphi^{-1}$$

on the subspace $F = L^2(X | \alpha)$ instead, the optimal approximation $\hat{P}_\varphi = P_F P_\varphi$ would be given by

$$\hat{P}_\varphi \mathbb{1}_A = \sum_{B \in \alpha} \frac{\mu(\varphi(A) \cap B)}{\mu(B)} \cdot \mathbb{1}_B = \sum_{B \in \alpha} \mu(\varphi(A) | B) \cdot \mathbb{1}_B$$

for all $A \in \alpha$. This would define a time-homogeneous Markov chain $(Y_n)_{n \in \mathbb{N}}$ with values in the finites state space α and transition probabilities

$$q_{B,A} := \frac{\mu(\varphi(A) \cap B)}{\mu(B)} = \mu(\varphi(A) | B)$$

for all $A, B \in \alpha$. Note that π is also a stationary distribution for $(Y_n)_{n \in \mathbb{N}}$. The duality $\langle T_\varphi \mathbb{1}_B, \mathbb{1}_A \rangle = \langle \mathbb{1}_B, P_\varphi \mathbb{1}_A \rangle$ between the operators leads to the relation

$$\mu(A) \cdot p_{A,B} = \mu(B) \cdot q_{B,A}$$

for all $A, B \in \alpha$. This duality shows that the Markov chains $(X_n)_{n \in \mathbb{N}}$ and $(Y_n)_{n \in \mathbb{N}}$ are *time-reversals* of each other.

In case T_φ (or equivalently P_φ) leaves the subspace $F = L^2(X | \alpha)$ invariant, the two Markov chains $(X_n)_{n \in \mathbb{N}}$ and $(Y_n)_{n \in \mathbb{N}}$ are *time reversals* of each other in the sense that they satisfy $p_{A,B} = q_{B,A}$ for all $A, B \in \alpha$. Indeed, this follows since $\varphi(A), \varphi^{-1}(A) \in \alpha$ for all $A \in \alpha$, and thus, $p_{A,B} = 1$ if and only if $A = \varphi^{-1}(B)$ and $q_{B,A} = 1$ if and only if $B = \varphi(A)$ for all $A, B \in \alpha$, and from the measure-preserving property of φ . In particular, in this case the transition probabilities are either 0 or 1 and both Markov chains are deterministic.

4.2 Lower Dimension Bound by Measure-Theoretic Entropy

Consider a measure-preserving system $\varphi: X \rightarrow X$ on the probability space $X = (X, \Sigma, \mu)$. Measure-theoretic entropy $h_\mu(\varphi, \alpha)$ of a finite Σ -measurable partition α with respect to φ was introduced by Kolmogorov and Sinai and has since then become a cornerstone of ergodic theory. We will give a precise definition of this concept in Section 3.1 and additionally recommend [Wal82] for an excellent exposition. Roughly

speaking, the measure-theoretical entropy $h_\mu(\varphi, \alpha)$ of α is a quantification of the unpredictability present in a measure-preserving system or equivalently a measure of the information provided by trajectories, when observing at the resolution α . It is thus natural to ask, whether the presence of measure-theoretical entropy for a finite Σ -measurable partition α limits the possibilities of long-term predictions via DMD algorithms. For this we assume that the subspace F is of the form

$$F = L^2(X | \alpha) := \text{span}\{\mathbb{1}_A : A \in \alpha\}$$

for some finite Σ -measurable partition α of X , i.e., we choose the dictionary to consist only of functions that tell us whether the state of the system is currently in some set $A \in \alpha$. In this case, we can derive a lower bound for the dimension of F , which depends on the measure-theoretic entropy $h_\mu(\varphi, \alpha)$ of the partition α with respect to φ . Note that this choice of a subspace F is used in the multiplicative Dynamic Mode Decomposition algorithm (see [BC24, Section 3]) and has been considered in [SKK16, Sections 3.2 and 4.1]. For the original publication of the following result we refer to [HH25, Theorem 1.1].

Theorem 4.2.1. *Let $\varphi: X \rightarrow X$ be a measure-preserving system on the probability space $X = (X, \Sigma, \mu)$, $\varepsilon > 0$ and α be a finite Σ -measurable partition of X . Then there exist $K_0 \in \mathbb{N}$ and $\delta > 0$ that satisfy the following: Whenever for a linear operator $\hat{T}: L^2(X | \alpha) \rightarrow L^2(X | \alpha)$ and some $K \geq K_0$ the assertion*

$$\|\hat{T}^k f - T_\varphi^k f\| \leq \delta \|f\|$$

holds for all $f \in L^2(X | \alpha)$ and $k \in \{0, \dots, K - 1\}$, we conclude that

$$|\alpha| \geq \exp(K(h_\mu(\varphi, \alpha) - \varepsilon)). \quad (4.13)$$

Here δ only depends on the system φ , the choice of ε and the cardinality of α .

Note that the measure-theoretic entropy $h_\mu(\varphi)$ of a measure-preserving system (X, φ) is defined as the supremum over all $h_\mu(\varphi, \alpha)$ ranging over all finite Σ -measurable partitions α . Thus, from the theorem we observe that whenever the measure-preserving system (X, φ) has non-zero measure-theoretic entropy, then care has to be taken while choosing α . Note that $h_\mu(\varphi, \alpha) = 0$ if and only if α is measurable with respect to the Pinsker- σ -algebra Π introduced in Section 3.3 of φ (see [Gla03, Section 18.1]).

We will now prepare the proof of Theorem 4.2.1. We begin with the incredibly simple observation that no approximation of an operator on a finite-dimensional subspace can be better than the best approximation, which is given by the operator composed with an orthogonal projection.

Lemma 4.2.2. *Let $T: L^2(X) \rightarrow L^2(X)$ be bounded linear operator and denote by $F \subseteq L^2(X)$ a finite-dimensional linear subspace. Let further $\hat{T}: F \rightarrow F$ be linear, P_F be the orthogonal projection onto F and $f \in F$ as well as $k \in \mathbb{N}$. If there exists $\delta > 0$ such that $\|\hat{T}^k f - T^k f\|_2 \leq \delta \|f\|_2$, then also $\|P_F T^k f - T^k f\|_2 \leq \delta \|f\|_2$.*

Proof. Since P_F is an orthogonal projection onto F , the inequality

$$\|P_F T^k f - T^k f\| \leq \|g - T^k f\|$$

holds for all $f, g \in F$. Hence, setting $g := \hat{T}^k f$ we obtain

$$\|P_F T^k f - T^k f\| \leq \|\hat{T}^k f - T^k f\| \leq \delta \|f\|.$$

Thus, the statement follows. \square

In $L^2(X)$ the orthogonal projections onto the subspace $L^2(X | \alpha)$ have a simple characterization.

For a finite Σ -measurable partition the subspace

$$L^2(X) = \{\mathbb{1}_A : A \in \alpha\}$$

consists precisely of the functions that are measurable with respect to the finite σ -algebra generated by α . Furthermore, it forms a finite-dimensional unital sublattice of $L^2(X)$. We refer to [Sch74, Proposition III.11.2] or [EFHN15, Proposition 13.19 b)].

Definition 4.2.3 (conditional expectation). Let $X = (X, \Sigma, \mu)$ be a probability space and α a finite Σ -measurable partition of X . We call an operator $P: L^2(X) \rightarrow L^2(X | \alpha)$ the *conditional expectation* of α , if

$$\int_A \mathbb{E}[f | \alpha] f \, d\mu = \int_A f \, d\mu$$

for all $f \in L^2(X)$ and all $A \in \alpha$ (or equivalently all A in the σ -algebra generated by α). We denote the conditional expectation of α by $\mathbb{E}[\cdot | \alpha] := P$.

We refer to [Bil95, Section 34] for a thorough introduction to conditional expectations. A moment's thought reveals that a conditional expectation is uniquely defined by the partition α . Consider the operator

$$\mathbb{E}[\cdot | \alpha]: L^2(X) \rightarrow L^2(X), \quad f \mapsto \sum_{A \in \alpha} \frac{1}{\mu(A)} \int_A f \, d\mu \cdot \mathbb{1}_A$$

and notice that it maps to $L^2(X)$. It is easy to check that it is in fact the operator satisfying Definition 4.2.3.

Recall further the notion of a Markov operator from Definition 2.1.6.

Definition 4.2.4 (Markov projection). Let $X = (X, \Sigma, \mu)$ be a probability space and $p \in [1, \infty)$. A Markov operator $P: L^p(X) \rightarrow L^p(X)$ is called a *Markov projection* if P is a projection, i.e., $P^2 = P$.

For the proof of Theorem 4.2.1 we require the following lemma, which is a consequence of [EFHN15, Lemma 13.16 and Remark 13.21]

Lemma 4.2.5. *Let α be a Σ -measurable partition of X and let $P: L^2(X) \rightarrow L^2(X)$ be a bounded linear operator. Then the following statements are equivalent.*

- (a) *The operator P is the orthogonal projection onto $L^2(X | \alpha)$.*
- (b) *The operator P is a Markov projection onto $L^2(X | \alpha)$.*
- (c) *The operator P is the conditional expectation $\mathbb{E}[\cdot | \alpha]$.*

In the next two lemmas, we establish a connection between the error of the conditional expectation and the conditional entropy of two partitions.

Lemma 4.2.6. *For any finite Σ -measurable partition β of X and any measurable $A \subseteq X$ we have*

$$\|\mathbb{E}[\mathbb{1}_A | \beta] - \mathbb{1}_A\|_1 = 2 \sum_{B \in \beta} \frac{\mu(A \cap B)\mu(B \setminus A)}{\mu(B)}.$$

Proof. We compute

$$\begin{aligned} \mathbb{E}[\mathbb{1}_A | \beta] - \mathbb{1}_A &= \sum_{B \in \beta} \frac{\mu(A \cap B)}{\mu(B)} \mathbb{1}_B - \sum_{B \in \beta} \mathbb{1}_A \cdot \mathbb{1}_B \\ &= \sum_{B \in \beta} \left(\left(\frac{\mu(A \cap B)}{\mu(B)} - 1 \right) \mathbb{1}_{A \cap B} + \frac{\mu(A \cap B)}{\mu(B)} \mathbb{1}_{B \setminus A} \right). \end{aligned}$$

Thus,

$$\begin{aligned} |\mathbb{E}[\mathbb{1}_A | \beta] - \mathbb{1}_A| &= \sum_{B \in \beta} \left(\left(1 - \frac{\mu(A \cap B)}{\mu(B)} \right) \mathbb{1}_{A \cap B} + \frac{\mu(A \cap B)}{\mu(B)} \mathbb{1}_{B \setminus A} \right) \\ &= \sum_{B \in \beta} \left(\frac{\mu(B \setminus A)}{\mu(B)} \mathbb{1}_{A \cap B} + \frac{\mu(A \cap B)}{\mu(B)} \mathbb{1}_{B \setminus A} \right) \end{aligned}$$

and we observe that

$$\|\mathbb{E}[\mathbb{1}_A | \beta] - \mathbb{1}_A\|_1 = 2 \sum_{B \in \beta} \frac{\mu(A \cap B)\mu(B \setminus A)}{\mu(B)}. \quad \square$$

For the following lemma we adapt and extend an argument given in [Wal82, Lemma 4.15] to our context.

Lemma 4.2.7. *Let $\kappa \in \mathbb{N}$, $p \in [1, \infty)$ and $\varepsilon > 0$. Then there exists $\delta > 0$, such that for any finite Σ -measurable partitions α and β of X with $|\alpha| \leq \kappa$ and $\sup_{A \in \alpha} \|\mathbb{E}[\mathbb{1}_A | \beta] - \mathbb{1}_A\|_p < \delta$, we obtain that $H_\mu(\alpha | \beta) < \varepsilon$.*

Proof. Let $c := \|\mathbb{1}\|_q$, where $q \in (1, \infty]$ is the Hölder conjugate of p . Choose $\delta \in (0, \frac{1}{4c\kappa})$ such that $-\kappa c \delta \log(c\delta) - (1 - \kappa c \delta) \log(1 - \kappa c \delta) < \varepsilon$.

Now consider finite Σ -measurable partitions α and β of X with $|\alpha| \leq \kappa$ and $\sup_{A \in \alpha} \|\mathbb{E}[\mathbb{1}_A | \beta] - \mathbb{1}_A\|_p < \delta$. For $A \in \alpha$ we denote by β_A the collection of sets $B \in \beta$ for which we have

$$\frac{\mu(B \cap A)}{\mu(B)} > \frac{1}{2}.$$

Then $\beta \setminus \beta_A$ contains all the sets B such that

$$\frac{\mu(B \setminus A)}{\mu(B)} = 1 - \frac{\mu(B \cap A)}{\mu(B)} \geq \frac{1}{2}.$$

We thus observe for $C_A := \bigcup_{B \in \beta \setminus \beta_A} (A \cap B)$ that

$$\begin{aligned} \mu(C_A) &\leq \sum_{B \in \beta_A} \mu(B \setminus A) + \sum_{B \in \beta \setminus \beta_A} \mu(A \cap B) \\ &\leq 2 \sum_{B \in \beta} \frac{\mu(B \setminus A) \mu(A \cap B)}{\mu(B)}. \end{aligned}$$

Thus Lemma 4.2.6 and the Hölder inequality imply

$$\mu(C_A) \leq \|\mathbb{E}[\mathbb{1}_A \mid \beta] - \mathbb{1}_A\|_1 \leq c \|\mathbb{E}[\mathbb{1}_A \mid \beta] - \mathbb{1}_A\|_p < c\delta \quad (4.14)$$

for all $A \in \alpha$. Define the index set $\Lambda := \{(A, B) : B \in \beta_A, A \in \alpha\}$ and set $C := \bigcup_{(A, B) \in \Lambda} A \cap B$. Then it is straightforward to observe that $\gamma := \{C_A : A \in \alpha\} \cup \{C\}$ is a finite Σ -measurable partition of X . Notice that $\alpha \vee \beta \subseteq \gamma \vee \beta$. Indeed, if $A \in \alpha$ and $B \in \beta \setminus \beta_A$, then

$$A \cap B = \left(\bigcup_{\tilde{B} \in \beta \setminus \beta_A} (A \cap \tilde{B}) \right) \cap B \in \gamma \vee \beta.$$

And if $A \in \alpha$ and $B \in \beta_A$, then

$$\gamma \vee \beta \ni \left(\bigcup_{(\tilde{A}, \tilde{B}) \in \Lambda} (\tilde{A} \cap \tilde{B}) \right) \cap B = \bigcup \{ \tilde{A} \cap B : \tilde{A} \in \alpha \text{ such that } B \in \beta_{\tilde{A}} \}. \quad (4.15)$$

Notice that B is only an element of $\beta_{\tilde{A}}$ for at most one $\tilde{A} \in \alpha$. Otherwise, if there are two $A_1, A_2 \in \alpha$ such that $B \in \beta_{A_1} \cap \beta_{A_2}$, then by disjointness of \tilde{A}_1 and \tilde{A}_2 we obtain

$$\frac{\mu(B \cap (\tilde{A}_1 \cup \tilde{A}_2))}{\mu(B)} = \frac{\mu(B \cap \tilde{A}_1)}{\mu(B)} + \frac{\mu(B \cap \tilde{A}_2)}{\mu(B)} > \frac{1}{2} + \frac{1}{2} > 1.$$

This is a contradiction. Continuing (4.15), it then follows that

$$\gamma \vee \beta \ni \bigcup \{ \tilde{A} \cap B : \tilde{A} \in \alpha \text{ such that } B \in \beta_{\tilde{A}} \} = A \cap B.$$

Hence, the inclusion $\alpha \vee \beta \subseteq \gamma \vee \beta$ holds.

Now recall from (4.14) that $\mu(C_A) < c\delta$ for all $A \in \alpha$. This implies that $\mu(C) \geq 1 - |\alpha|c\delta \geq 1 - \kappa c\delta$ and we observe

$$\begin{aligned} H_\mu(\gamma) &\leq -|\alpha|c\delta \log(c\delta) - (1 - \kappa c\delta) \log(1 - \kappa c\delta) \\ &\leq -\kappa c\delta \log(c\delta) - (1 - \kappa c\delta) \log(1 - \kappa c\delta) \\ &< \varepsilon. \end{aligned}$$

Thus, since $\alpha \vee \beta \subseteq \gamma \vee \beta$, we obtain from [Wal82, Theorem 4.3] that

$$H_\mu(\beta) + H_\mu(\alpha \mid \beta) = H_\mu(\alpha \vee \beta) = H_\mu(\gamma \vee \beta) \leq H_\mu(\gamma) + H_\mu(\beta) < H_\mu(\beta) + \varepsilon,$$

which implies $H_\mu(\alpha \mid \beta) < \varepsilon$. \square

We are now ready to prove the main result of this section.

Proof of Theorem 4.2.1. Let $\varepsilon > 0$ and α be the finite Σ -measurable partition of X . Let $F := L^2(X | \alpha)$ and denote by P_F the orthogonal projection onto F . It follows from Lemma 4.2.5 that P_F coincides with the conditional expectation $\mathbb{E}[\cdot | \alpha]$. So by Lemma 4.2.7 there exists $\delta > 0$ such that whenever $\|P_F T_\varphi^k \mathbb{1}_A - T_\varphi^k \mathbb{1}_A\| < \delta$ for some $k \in \mathbb{N}$, we obtain that $H_\mu(\varphi^{-k}(\alpha) | \alpha) < \varepsilon/2$. Also notice that there exists $K_0 \in \mathbb{N}$ such that

$$\left| h_\mu(\varphi, \alpha) - \frac{1}{K} H_\mu \left(\bigvee_{k=0}^{K-1} \varphi^{-k}(\alpha) \right) \right| < \varepsilon/2$$

for all $K \geq K_0$.

Let $f \in F$ and fix $K \geq K_0$ and assume that $\|\hat{T}^k f - T_\varphi^k f\|_2 \leq \delta \|f\|_2$ holds for all $f \in F$ and all $k \in \{0, \dots, K-1\}$. From Lemma 4.2.2 we observe that $\|P_F T_\varphi^k f - T_\varphi^k f\| \leq \delta \|f\|$ for all $f \in F$ and all $k \in \{0, \dots, K-1\}$. In particular, this implies that $\|P_F T_\varphi^k \mathbb{1}_A - T_\varphi^k \mathbb{1}_A\| \leq \delta$ holds for all $k \in \{0, \dots, K-1\}$ and all $A \in \alpha$. Hence, Lemma 4.2.7 yields that

$$H_\mu(\varphi^{-k}(\alpha) | \alpha) \leq \varepsilon/2$$

for all $k \in \{0, \dots, K-1\}$. Thus,

$$\begin{aligned} H_\mu \left(\bigvee_{k=0}^{K-1} \varphi^{-k}(\alpha) \right) &\leq H_\mu \left(\alpha \vee \left(\bigvee_{k=1}^{K-1} \varphi^{-k}(\alpha) \right) \right) \\ &\leq H_\mu(\alpha) + H_\mu \left(\bigvee_{k=1}^{K-1} \varphi^{-k}(\alpha) \middle| \alpha \right) \\ &\leq H_\mu(\alpha) + \sum_{k=1}^{K-1} H_\mu(\varphi^{-k}(\alpha) | \alpha) \\ &\leq \log(|\alpha|) + \frac{(K-1)\varepsilon}{2}. \end{aligned}$$

Now dividing by K yields

$$h_\mu(\varphi, \alpha) - \varepsilon \leq \frac{1}{K} H_\mu \left(\bigvee_{k=0}^{K-1} \varphi^{-k}(\alpha) \right) - \varepsilon/2 \leq \frac{1}{K} \log(|\alpha|),$$

and the statement follows. \square

4.3 Lower Dimension Bound by Approximation Entropy

Recall from Section 4.2 that Theorem 4.2.1 applies to finite-dimensional subspaces of the form $L^2(X | \alpha)$ for a finite Σ -measurable partition α of X . It is natural to ask, whether it can be generalized to arbitrary finite-dimensional subspaces F of $L^2(X)$. For this we consider the *approximation entropy* $h_{\text{apr}}(T_\varphi, F)$ of T_φ with respect to a

subspace F as introduced in Section 3.2, which gives another quantification of the unpredictability of a measure-preserving system.

The main result that complements Theorem 4.2.1 is the following. Note that this is [HH25, Theorem 1.2].

Theorem 4.3.1. *Let $\varphi: X \rightarrow X$ be an invertible measure-preserving system on the probability space $X = (X, \Sigma, \mu)$, let $F \subseteq L^2(X)$ be a finite-dimensional linear subspace and $\varepsilon > 0$. Then there exist $K_0 \in \mathbb{N}$ and $\delta > 0$ that satisfy the following: Whenever for a linear operator $\hat{T}: F \rightarrow F$ and some $K \geq K_0$ the assertion*

$$\|\hat{T}^k f - T_\varphi^k f\| \leq \delta \|f\|$$

holds for all $f \in F$ and all $k \in \{0, \dots, K-1\}$, we conclude that

$$\dim(F) \geq K(h_{\text{apr}}(T_\varphi, F) - \varepsilon).$$

We have shown in Theorem 3.3.12 that whenever φ has non-zero measure-theoretic entropy (and in Proposition 3.3.4 that if T_φ has a countable Lebesgue component), then there exist finite-dimensional subspaces F of $L^2(X)$, which allow for arbitrarily large values of $h_{\text{apr}}(T_\varphi, F)$. Thus, in Theorem 4.2.1, care has to be taken by choosing F in this context.

Let $F \subseteq L^2(X)$ be a finite-dimensional subspace and $n \in \mathbb{N}$. The n -th delay subspace for (X, φ) is given by

$$F_n := \text{span} \bigcup_{k=0}^n T_\varphi^k F.$$

Note that its dimension is upper bounded by $n \cdot \dim(F)$. It follows from Proposition 3.2.1 (v) that we have $h_{\text{apr}}(T_\varphi, F_n) = h_{\text{apr}}(T_\varphi, F)$. It is thus natural to ask for a stronger statement of Theorem 4.3.1 by replacing an approximation on F with one on F_n . The author did not manage to prove this for Theorem 4.2.1 as there the δ depended on the cardinality of α , and thus, on the dimension of F . It is possible for lower bounds by the approximation entropy as the following theorem shows. This is [HH25, Theorem 3.4].

Theorem 4.3.2. *Let $\varphi: X \rightarrow X$ be an invertible measure-preserving system on the probability space $X = (X, \Sigma, \mu)$, let $F \subseteq L^2(X)$ be a finite-dimensional linear subspace and $\varepsilon > 0$. Then there exist $K_0 \in \mathbb{N}$ and $\delta > 0$ that satisfy the following: Whenever for some $n \in \mathbb{N}$, a linear operator $\hat{T}_n: F_n \rightarrow F_n$ and $K \geq K_0$ the assertion*

$$\|\hat{T}_n^k f - T_\varphi^k f\| \leq \delta \|f\|$$

holds for all $f \in F_n$ and all $k \in \{0, \dots, K-1\}$, we conclude that

$$\dim(F_n) \geq K(h_{\text{apr}}(T_\varphi, F) - \varepsilon).$$

Note that Theorem 4.3.2 is a special case of Theorem 4.3.1 if we set $n = 0$. It therefore suffices to prove Theorem 4.3.2.

As the approximation entropy has been defined for unitary operators, it feels natural to prove Theorem 4.3.2 in a form that holds for all unitary operators. So let

$T: H \rightarrow H$ be a unitary operator on a Hilbert space H . For a finite-dimensional subspace $F \subseteq H$ we analogously say that

$$F_n := \text{span} \bigcup_{k=0}^n T^k F$$

is the n -th delay subspace of F . Note that the dimension of F_n is upper bounded by $n \cdot \dim(F)$ and that it follows from Proposition 3.2.1 (v) that $h_{\text{apr}}(T, F) = h_{\text{apr}}(T, F_n)$.

Theorem 4.3.3. *Let H be a Hilbert space and $T: H \rightarrow H$ be unitary, let $F \subseteq H$ be a finite-dimensional linear subspace and $\varepsilon > 0$. Then there exists $K_0 \in \mathbb{N}$ and $\delta > 0$ such that whenever for some $n \in \mathbb{N}$, for a linear operator $\hat{T}_n: F_n \rightarrow F_n$ and $K \geq K_0$ the assertion*

$$\|\hat{T}_n^k x - T^k x\| \leq \delta \|x\|$$

holds for all $x \in F_n$ and all $k \in \{0, \dots, K-1\}$, we conclude that

$$\dim(F_n) \geq K(h_{\text{apr}}(T, F) - \varepsilon).$$

Proof. Let \mathcal{O} be an orthonormal basis of F . From Proposition 3.2.3 we observe that $h_{\text{apr}}(T, F) = h_{\text{apr}}(T, \mathcal{O})$. In particular, there exists $\delta > 0$, such that $h_{\text{apr}}(T, F) - \frac{\varepsilon}{2} \leq h(T, \mathcal{O}, \delta)$. Furthermore, there exists $K_0 \in \mathbb{N}$, such that for all $K \geq K_0$ we have

$$h_{\text{apr}}(T, F) - \varepsilon \leq h(T, \mathcal{O}, \delta) - \frac{\varepsilon}{2} \leq \frac{1}{K} H_{\text{apr}} \left(\bigcup_{k=0}^{K-1} T^k \mathcal{O}, \delta \right).$$

Now consider $K \geq K_0$ and $n \in \mathbb{N}$, such that the assertion

$$\|\hat{T}_n^k x - T^k x\| \leq \delta \|x\|$$

holds for all $x \in F_n$ and all $k \in \{0, \dots, K-1\}$. Denote by P_{F_n} the orthogonal projection onto the n -th delay subspace F_n . From Lemma 4.2.2 we observe that

$$\|P_{F_n} T^k x - T^k x\| \leq \delta \|x\| = \delta$$

for all $x \in \mathcal{O}$ and all $k \in \{0, \dots, K-1\}$. In particular, it follows that

$$H_{\text{apr}} \left(\bigcup_{k=0}^{K-1} T^k \mathcal{O}, \delta \right) \leq \dim(F_n).$$

We observe

$$h_{\text{apr}}(T, F) - \varepsilon \leq \frac{1}{K} H_{\text{apr}} \left(\bigcup_{k=0}^{K-1} T^k \mathcal{O}, \delta \right) \leq \frac{1}{K} \dim(F_n),$$

and the statement follows. \square

4.4 Notes

The study of statistical analysis via Dynamic Mode Decomposition algorithms has gained significant attention in recent years. The foundational work by Schmid [Sch10] introduced the basis variant of DMD as a powerful tool for analyzing complex dynamical systems, particularly in fluid dynamics. The fundamental idea of this algorithm is to assume that we have data

$$x_0, x_1, \dots, x_K \in \mathbb{R}^d \quad \text{and} \quad y_0, y_1, \dots, y_K \in \mathbb{R}^d$$

such that there exists a linear operator $A: \mathbb{R}^d \rightarrow \mathbb{R}^d$ satisfying $y_k = Ax_k$ for all $k \in \{0, \dots, K\}$. We can then write $Y \approx AX$ with

$$X := \begin{bmatrix} | & & | \\ x_0 & \cdots & x_K \\ | & & | \end{bmatrix} \in \mathbb{R}^{d \times (K+1)} \quad \text{and}$$

$$Y := \begin{bmatrix} | & & | \\ y_0 & \cdots & y_K \\ | & & | \end{bmatrix} \in \mathbb{R}^{d \times (K+1)}$$

and find an approximation $\hat{A} \in \mathbb{R}^d$ of A by solving the least-squares problem $\min_{\hat{A} \in \mathbb{R}^{d \times d}} \|Y - \hat{A}X\|_F$ with respect to the Frobenius norm $\|\cdot\|_F$. The solution is given by $\hat{A} = YX^\dagger$, where X^\dagger denotes the Moore-Penrose pseudoinverse of X . We note the similarity of this approach to the description of the algorithm in Section 4.1.2. Later, the assumed correspondence of data has been replaced by a correspondence via a possibly nonlinear dynamical system $\varphi: \Omega \rightarrow \Omega$ on a subset $\Omega \subseteq \mathbb{R}^d$. The assumption of linear correspondence has then been interpreted as $y_k = \varphi(x_k)$ for all $k \in \{0, \dots, K\}$. In order to interpret this in the context of Koopman operator theory, the data points x_k and y_k have been considered as evaluations of observables at different time instances. More precisely, one obtains the collection of *full state observables* given by $f_i: \Omega \rightarrow \mathbb{R}$ given by $f_i(x) = x^{(i)}$ for $i \in \{1, \dots, d\}$, where $x^{(i)}$ denotes the i -th component of x , that

$$y_k = \begin{bmatrix} f_1(\varphi(x_k)) \\ \vdots \\ f_d(\varphi(x_k)) \end{bmatrix} = \begin{bmatrix} (T_\varphi f_1)(x_k) \\ \vdots \\ (T_\varphi f_d)(x_k) \end{bmatrix} \approx \hat{A} \begin{bmatrix} f_1(x_k) \\ \vdots \\ f_d(x_k) \end{bmatrix} = \hat{A}x_k$$

for all $k \in \{0, \dots, K\}$. Here \hat{A} can be found by solving the least-squares problem as before. This works well if $\text{span}\{f_1, \dots, f_d\} \subseteq L^2(\Omega, \mu)$ happens to be an invariant subspace of the Koopman operator $T_\varphi: L^2(\Omega, \mu) \rightarrow L^2(\Omega, \mu)$ associated with φ . In this case, one can interpret \hat{A} as the matrix representation of the restriction $T_\varphi|_{\text{span}\{f_1, \dots, f_d\}}$ with respect to the basis $\{f_1, \dots, f_d\}$. In general, however, $\text{span}\{f_1, \dots, f_d\}$ is not invariant under T_φ , and thus, \hat{A} only approximates the action of T_φ on this subspace.

This observation has led to the development of the extended DMD in [WKR15], where the observables are chosen from a larger dictionary in order to better approximate the action of the Koopman operator. The choice of observables is crucial for the performance of DMD algorithms, and various strategies have been proposed to select appropriate observables based on the underlying dynamics. The choice of observables

can be influenced in various ways, one may design the experiment to measure specific observables, or one may use data-driven techniques to identify relevant observables from the data itself. Simpler methods include the use of time-delay observables of the form $f \circ \varphi^k$ for some observable f and $k \in \mathbb{N}$ or by taking products/maxima/minima of observables to generate a larger dictionary.

For us, the non-invariance of the span of the dictionary was the fundamental observation that led to the development of lower dimension bounds for DMD approximations of Koopman operators in this chapter.

The assumption that the data is sampled ergodically or iid has been added in [KM18] when they first used the strong law of large numbers and Birkhoff's ergodic theorem to show convergence of the DMD algorithm to the Koopman operator in the large data limit.

The utility of DMD algorithms for predicting time-series has led to a variety of variants of DMD including but not limited to the extended DMD [WKR15], Kernel DMD [KRW15], Measure-preserving DMD [Col23a], Multiplicative DMD [BC24], Residual DMD [CAS23] and many more. We refer to [Col23b] for a recent and extensive overview of the different DMD algorithms and their properties. A similar frame-theoretic description and asymptotic analysis as given in this chapter of DMD variants would be interesting.

We further note that the form of the best approximation of the Koopman operator that has been described in Proposition 4.1.17 has been noted before in [SKK16] and is used to construct finite-rank approximations of transfer operators by Ulam's method.

Typically, the error bounds of DMD algorithms are given as bounds of the one step error $\|\hat{T}_\varphi - T_\varphi\|$. The propagated error $\|\hat{T}_\varphi^k - T_\varphi^k\|$ for $k > 1$ has been considered in [KM18, Theorem 5] although only asymptotically where the dimension of F is increased. In the main results of Section 4.2 and Section 4.3, we take a different approach to analyze the error by proving necessary conditions in terms of a lower bound of the dimension of F that needs to be satisfied in order to achieve a certain bound for the propagated error. This perspective seems to be new.

The fractional approximation entropies that were introduced at the end of Section 3.2.2 could be used in order to obtain fractional (instead of linear) growth rates of lower bounds for the dimension of DMD approximations of Koopman operators. We might be able to extend the result of Theorem 4.3.1 or Theorem 4.3.3 in this way. This has not been explored in this work, but could be an interesting direction for future research.

Entropy via Banach Lattices

Classical frameworks of dynamical entropy divide into measure-theoretic entropy and topological entropy for measure-preserving and topological dynamical systems, respectively. The former measures the expected surprisal of learning that trajectories evolve through measurable partitions, and the latter counts orbits that are distinguishable under open covers. Despite their roles as quantitative measures of the complexity of dynamical evolution, these notions arise from different structural ingredients: measure theory in one case and topological arguments in the other. This chapter develops a unified entropy formalism based on Banach lattices, showing that both classical invariants are recoverable as specializations of a single operator-theoretic construction.

We encode the growth of finite data necessary to understand trajectories not by partitions or covers, but by finite subsets \mathcal{O} of a Banach lattice. We look at trajectory data $\bigcup_{k=0}^{n-1} T^k \mathcal{O}$, where T is the evolution operator, and then quantify their compressibility or approximability through positive contractive projections fixing a quasi-interior point. The rank of such a projection then becomes the measure of complexity. This leads to the notion of Banach lattice entropy, which we introduce in Definition 5.1.4.

In Section 5.1 we introduce the Banach lattice entropy for Markov operators on Banach lattices with a quasi-interior point and derive fundamental properties.

Section 5.2 relates the Banach lattice entropy of Koopman operators on L^p to measure-theoretic entropy. Using conditional expectations and the Shannon–McMillan–Breiman theorem, it is shown that for ergodic measure-preserving transformations the Banach lattice entropy coincides with Kolmogorov–Sinai entropy (Theorem 5.2.1, proved via Propositions 5.2.7 and 5.2.10). Subadditivity in the case $p = 2$ ensures the existence of the defining limit (Corollary 5.2.6).

In Section 5.3 we treat continuous dynamics on compact Hausdorff spaces. Here, partitions of unity replace measurable partitions. The Banach lattice entropy of an invertible Koopman operator agrees with topological entropy (Theorem 5.3.1), with the two inequalities established in Propositions 5.3.7 and 5.3.8.

The main contributions of this chapter are Definition 5.1.4, Theorem 5.2.1 and Theorem 5.3.1. They have not been published or submitted for publication yet.

5.1 Abstract Entropy on Banach Lattices

To obtain a common generalization of measure-theoretic and topological entropy, we introduce an abstract entropy notion on Banach lattices. Positive and contractive projections on Banach lattices that leave a quasi-interior point invariant play a central role in the definition of Banach lattice entropy.

Let E be a Banach lattice and let $e \in E_+$ be a *quasi-interior point* of E , i.e., the principal ideal

$$I_e := \{x \in E : \exists \lambda > 0 : |x| \leq \lambda e\}$$

is dense in E . We assume by normalization that $\|e\| = 1$.

Definition 5.1.1 (Markov operators and projections). Let E be a Banach lattice and let $e \in E_+$ be a quasi-interior point of E . We call an operator $T: E \rightarrow E$ a *Markov operator* on (E, e) if T is positive and contractive and satisfies $Te = e$. A Markov operator on (E, e) is called a *Markov projection* on (E, e) if it is a projection.

It is instructive to compare the notion of Markov operators and projections on L^p -spaces introduced in Definition 2.1.6 and 4.2.4 with the abstract notions on $(L^p(X), \mathbb{1})$. Note that a Markov operator on $L^p(X)$ is a positive linear operator T that satisfies $T\mathbb{1} = \mathbb{1}$ and $\int_X Tf \, d\mu = \int_X f \, d\mu$ for all $f \in L^p(X)$. The latter condition is easily seen to be equivalent to $T^*\mathbb{1} = \mathbb{1}$.

Remark 5.1.2 (Disambiguation). Let $X = (X, \Sigma, \mu)$ be a probability space and let $p \in [1, \infty)$. Then the constant function $\mathbb{1} \in L^p(X)$ is a quasi-interior point of the Banach lattice $L^p(X)$. A linear operator $T: L^p(X) \rightarrow L^p(X)$ is a Markov operator on $(L^p(X), \mathbb{1})$ if and only if T is a Markov operator according to Definition 2.1.6.

In particular, the Markov projections on $(L^p(X), \mathbb{1})$ are exactly the Markov projections according to Definition 4.2.4.

Proof. “ \Rightarrow ”: It suffices to show that $\int_X Tf \, d\mu = \int_X f \, d\mu$ for all $f \in L^p(X)$. Note that this property is equivalent to the property that $T^*\mathbb{1} = \mathbb{1}$ for the adjoint operator $T^*: L^q(X) \rightarrow L^q(X)$, where $q \in (1, \infty]$ is the Hölder conjugate of p . Since T is positive and contractive, so is T^* .

Let us first prove the statement for $p \in (1, \infty)$. Following an idea in the proof of [And66, Lemma 1], we obtain from Hölder’s inequality and the fact that T^* is also contractive that

$$\begin{aligned} 1 &= \int_X \mathbb{1} \cdot \mathbb{1} \, d\mu = \int_X (T\mathbb{1}) \cdot \mathbb{1} \, d\mu \\ &= \int_X \mathbb{1} \cdot (T^*\mathbb{1}) \, d\mu \leq \|\mathbb{1}\|_p \|T^*\mathbb{1}\|_q \leq 1. \end{aligned}$$

In particular, Hölder’s inequality above is an equality. It follows that $\mathbb{1}$ and $T^*\mathbb{1}$ are linearly dependent. Hence, $T^*\mathbb{1} = c\mathbb{1}$ for some constant $c \geq 0$. Applying the norm on both sides yields $c = \|T^*\mathbb{1}\|_q = 1$ by the above inequality.

Let us now prove the statement for $p = 1$. Note that $T^*: L^\infty(X) \rightarrow L^\infty(X)$ is also contractive and it follows that $\|T^*\mathbb{1}\|_\infty \leq 1$. This implies that $T^*\mathbb{1} \leq \mathbb{1}$. Thus,

$$0 \geq \langle \mathbb{1}, T^*\mathbb{1} - \mathbb{1} \rangle_{L^1(X), L^\infty(X)} = \langle T\mathbb{1} - \mathbb{1}, \mathbb{1} \rangle_{L^1(X), L^\infty(X)} = 0,$$

implies that $T^*\mathbb{1} = \mathbb{1}$.

“ \Leftarrow ”: It follows from [EFHN15, Proposition 13.6] that every Markov operator on L^p can be extended to one on L^1 , and from [EFHN15, Theorem 13.2] that its restriction to L^p is a contraction. \square

The definition of a Markov operator on an arbitrary Banach lattice (E, e) allows us to introduce Markov operators on $(L^\infty(X), \mathbb{1})$ and $(C_b(\Omega), \mathbb{1})$ as well. On both spaces, the Markov operators are simply positive contractions leaving the constant function $\mathbb{1}$ invariant.

Let us now introduce a special class of Markov operators that will play an important role in the following.

Definition 5.1.3 (Markov automorphism). Let E be a Banach lattice and let $e \in E_+$ be a quasi-interior point of E . We call a Markov operator T on (E, e) a *Markov automorphism* on (E, e) if T is bijective and its inverse $T^{-1}: E \rightarrow E$ is also a Markov operator on (E, e) .

Not every invertible Markov operator is a Markov automorphism, since the inverse does not need to be positive (see [EFHN15, Section 13.2]). However, one can show that a Markov operator T on (E, e) is Markov automorphism on (E, e) if and only if T is a lattice isomorphism [EFHN15, Theorem 13.9, Theorem 13.12 and Corollary 13.14]. The Koopman operator on $L^p(X)$ for $p \in [1, \infty)$ induced by an invertible measure-preserving system $\varphi: X \rightarrow X$ and every Koopman operator on $C_b(\Omega)$ induced by a homeomorphism $\varphi: \Omega \rightarrow \Omega$ is a Markov automorphism on $(L^p(X), \mathbb{1})$ and $(C_b(\Omega), \mathbb{1})$, respectively.

Having established a notion of Markov projections on Banach lattices, we can now introduce the Banach lattice entropy.

Definition 5.1.4 (Banach lattice entropy). Let E be a Banach lattice, let $e \in E_+$ be a quasi-interior point of E and let $T: E \rightarrow E$ be a Markov operator on (E, e) . For a finite subset $\mathcal{O} \subseteq E$ and $\delta > 0$ we define the δ -*information* $H_{\text{bl}}(\mathcal{O}, \delta)$ of \mathcal{O} as the minimum of $\dim(PE)$, where P runs through all Markov projections on (E, e) satisfying

$$\|Px - x\| \leq \delta \quad \text{for all } x \in \mathcal{O}.$$

The *Banach lattice entropy* of T with respect to \mathcal{O} and δ is defined as

$$h_{\text{bl}}(T, \mathcal{O}, \delta) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log \left(H_{\text{bl}} \left(\bigcup_{k=0}^{n-1} T^k \mathcal{O}, \delta \right) \right). \quad (5.1)$$

Finally, the *Banach lattice entropy* of T (with respect to \mathcal{O}) is defined as

$$h_{\text{bl}}(T, \mathcal{O}) := \sup_{\delta > 0} h_{\text{bl}}(T, \mathcal{O}, \delta) \quad \text{and} \quad h_{\text{bl}}(T) := \sup_{\substack{\mathcal{O} \subseteq E \\ |\mathcal{O}| < \infty}} h_{\text{bl}}(T, \mathcal{O}).$$

The author does not know whether this limit superior is in fact a limit. We will discuss this in Corollary 5.2.6 for Markov operators on L^p -spaces with $p \in [1, \infty)$.

Let us record some basic properties of the Banach lattice entropy. Note the analogy to Proposition 3.2.1.

Proposition 5.1.5. *Let E be a Banach lattice and let $e \in E_+$ be a quasi-interior point of E and $T: E \rightarrow E$ be a Markov operator on (E, e) . Then the following assertions hold:*

- (i) Let $\mathcal{O}_1, \mathcal{O}_2 \subseteq E$ be finite with $\mathcal{O}_1 \subseteq \mathcal{O}_2$ and $\delta > 0$. Then $H_{\text{bl}}(\mathcal{O}_1, \delta) \leq H_{\text{bl}}(\mathcal{O}_2, \delta)$, and consequently, $h_{\text{bl}}(T, \mathcal{O}_1, \delta) \leq h_{\text{bl}}(T, \mathcal{O}_2, \delta)$.
- (ii) Let $\mathcal{O} \subseteq E$ be finite and $\delta_1, \delta_2 > 0$ with $\delta_1 \leq \delta_2$. Then $H_{\text{bl}}(\mathcal{O}, \delta_1) \geq H_{\text{bl}}(\mathcal{O}, \delta_2)$, and consequently, $h_{\text{bl}}(T, \mathcal{O}, \delta_1) \geq h_{\text{bl}}(T, \mathcal{O}, \delta_2)$.
- (iii) Suppose that T is a Markov automorphism on (E, e) . For $M, N \in \mathbb{N}$ with $M \leq N$, finite $\mathcal{O} \subseteq E$ and $\delta > 0$ we have $H_{\text{bl}}(\mathcal{O}, \delta) = H_{\text{bl}}(T^k \mathcal{O}, \delta)$ for all $k \in \mathbb{Z}$ and

$$h_{\text{bl}}(T, \mathcal{O}, \delta) = h_{\text{bl}}\left(\bigcup_{k=M}^N T^k \mathcal{O}, \delta\right).$$

- (iv) Let $\mathcal{O}_1, \mathcal{O}_2 \subseteq E$ be finite and $\delta, \varepsilon > 0$ such that $\mathcal{O}_1 \subseteq B_{<\varepsilon}(\mathcal{O}_2)$. Then $h_{\text{bl}}(T, \mathcal{O}_1, \delta + 2\varepsilon) \leq h_{\text{bl}}(T, \mathcal{O}_2, \delta)$.
- (v) If \mathcal{O}_1 and \mathcal{O}_2 are finite subsets of E with $\mathcal{O}_1 \subseteq \text{span } \mathcal{O}_2$, then

$$h_{\text{bl}}(T, \mathcal{O}_1) \leq h_{\text{bl}}(T, \mathcal{O}_2).$$

Proof. “(i)”: This follows immediately from the definition.

“(ii)”: This follows immediately from the definition.

“(iii)”: Fix $k \in \mathbb{Z}$. Let P be a Markov projection on (E, e) with $\|PT^k x - T^k x\| \leq \delta$ for all $x \in \mathcal{O}$ with minimal rank. Set $Q := T^{-k}PT^k$. Then, for each $x \in \mathcal{O}$ we have

$$\|Qx - x\| = \|T^{-k}PT^k x - T^{-k}T^k x\| = \|PT^k x - T^k x\| \leq \delta.$$

Since Q is again a Markov projection on (E, e) with the same rank as P , we obtain that

$$H_{\text{bl}}(T^k \mathcal{O}, \delta) \geq H_{\text{bl}}(\mathcal{O}, \delta).$$

The reverse inequality follows analogously by interchanging the roles of \mathcal{O} and $T^k \mathcal{O}$.

Thus,

$$\begin{aligned} h_{\text{bl}}(T, \bigcup_{k=M}^N T^k \mathcal{O}, \delta) &= \limsup_{n \rightarrow \infty} \frac{1}{n} \log \left(H_{\text{bl}} \left(\bigcup_{j=M}^{N+n-1} T^j \mathcal{O}, \delta \right) \right) \\ &= \limsup_{n \rightarrow \infty} \frac{n + (N - M) + 1}{n} \cdot \frac{1}{n + (N - M) + 1} \log \left(H_{\text{bl}} \left(\bigcup_{j=0}^{n+(N-M)} T^j \mathcal{O}, \delta \right) \right) \\ &= h_{\text{bl}}(T, \mathcal{O}, \delta). \end{aligned}$$

“(iv)”: Let P be a Markov projection on (E, e) with $\|PT^k x_2 - T^k x_2\| \leq \delta \|x_2\|$ for all $x_2 \in \mathcal{O}_2$ and all $k \in \{0, \dots, n-1\}$. Then, for each $x_1 \in \mathcal{O}_1$ there exists $x_2 \in \mathcal{O}_2$ with $\|x_1 - x_2\| < \varepsilon$, and thus,

$$\begin{aligned} \|PT^k x_1 - T^k x_1\| &\leq \|PT^k x_1 - PT^k x_2\| + \|PT^k x_2 - T^k x_2\| + \|T^k x_2 - T^k x_1\| \\ &\leq \|x_1 - x_2\| + \delta + \|x_2 - x_1\| < \delta + 2\varepsilon. \end{aligned}$$

Hence, we obtain that $h_{\text{bl}}(T, \mathcal{O}_1, \delta + 2\varepsilon) \leq h_{\text{bl}}(T, \mathcal{O}_2, \delta)$.

“(v)”: Let $x_1 \in \mathcal{O}_1$. We write $x_1 = \sum_{x_2 \in \mathcal{O}_2} \lambda_{x_1, x_2} x_2$ with $\lambda_{x_1, x_2} \in \mathbb{K}$ and set

$$\lambda := \max_{x_1 \in \mathcal{O}_1} \sum_{x_2 \in \mathcal{O}_2} |\lambda_{x_1, x_2}|.$$

Let P be a Markov projection on (E, e) with $\|PT^k x_2 - T^k x_2\| \leq \delta/\lambda$ for all $x_2 \in \mathcal{O}_2$ and all $k \in \{0, \dots, n-1\}$. Then

$$\|PT^k x_1 - T^k x_1\| \leq \sum_{x_2 \in \mathcal{O}_2} |\lambda_{x_1, x_2}| \|PT^k x_2 - T^k x_2\| \leq \sum_{x_2 \in \mathcal{O}_2} |\lambda_{x_1, x_2}| \frac{\delta}{\lambda} \leq \delta.$$

Hence, it follows that for all $n \in \mathbb{N}$ we have

$$H_{\text{bl}} \left(\bigcup_{k=0}^{n-1} T^k \mathcal{O}_1, \delta \right) \leq H_{\text{bl}} \left(\bigcup_{k=0}^{n-1} T^k \mathcal{O}_2, \frac{\delta}{\lambda} \right),$$

and the claim follows. \square

We remark that not every assertion of Proposition 3.2.1 finds its counterpart in the above Proposition 5.1.5. In particular, we do not have an analogue of Proposition 3.2.1 (i), (iv) and (viii). This is due to the fact that the proofs of these items do not easily transfer to the general setting of Markov projections on Banach lattices.

However, for Markov operators of special pairs (E, e) we can establish further properties of the Banach lattice entropy, which we will do in the following sections before we prove the main results.

5.2 Measure-Theoretic Entropy

Consider $E = L^p(X)$ for a probability space $X = (X, \Sigma, \mu)$, $p \in [1, \infty)$ and consider the quasi-interior point $e = \mathbb{1}$. We recall the definition of measure-theoretic entropy from Section 3.1 to be

$$h_\mu(\varphi) := \sup_\alpha \lim_{n \rightarrow \infty} \frac{1}{n} H_\mu(\varphi_0^{n-1} \alpha),$$

where $H_\mu(\alpha) := -\sum_{A \in \alpha} \mu(A) \log \mu(A)$ for a finite Σ -measurable partition α of X and

$$\varphi_0^{n-1} \alpha := \bigvee_{k=0}^{n-1} \varphi^{-k} \alpha.$$

For measure-preserving systems $\varphi: X \rightarrow X$ we can relate the Banach lattice entropy of the associated Koopman operator $T_\varphi: L^p(X) \rightarrow L^p(X)$, $f \mapsto f \circ \varphi$ to the measure-theoretic entropy of φ in the following way.

Theorem 5.2.1. *Let $\varphi: X \rightarrow X$ be an ergodic measure-preserving system on the probability space $X = (X, \Sigma, \mu)$, let $p \in [1, \infty)$ and let $T_\varphi: L^p(X) \rightarrow L^p(X)$, $f \mapsto f \circ \varphi$ be the associated Koopman operator. Then*

$$h_{\text{bl}}(T_\varphi) = h_\mu(\varphi).$$

We will divide the proof into two parts, showing the two inequalities separately in Proposition 5.2.7 and Proposition 5.2.10.

We will first see how the Banach lattice entropy compares with an entropy notion for automorphisms of von Neumann algebras introduced by Voiculescu in [Voi95].

Remark 5.2.2. In [Voi95, Section 1], Voiculescu introduces an entropy notion for automorphisms of von Neumann algebras that generalizes the classical measure-theoretic entropy. We adapt his definition to our situation.

For a probability space $X = (X, \Sigma, \mu)$, the von Neumann algebra $\mathcal{M} := L^\infty(X)$ is equipped with the normal faithful trace $\tau: L^\infty(X) \rightarrow \mathbb{C}$, $f \mapsto \int_X f \, d\mu$. An automorphism α of $L^\infty(X)$ is called τ -preserving if $\tau \circ \alpha = \tau$. Every invertible measure-preserving system $\varphi: X \rightarrow X$ induces a τ -preserving automorphism $\alpha_\varphi: L^\infty(X) \rightarrow L^\infty(X)$, $f \mapsto f \circ \varphi$.

Voiculescu's entropy $ha_\tau(\alpha)$ of a τ -preserving automorphism α of a von Neumann algebra (\mathcal{M}, τ) is defined via approximations of finite subsets of \mathcal{M} by finite-dimensional subalgebras of \mathcal{M} . The error is measured in the L^2 -norm. As the finite-dimensional subalgebras of \mathcal{M} correspond to ranges of finite-rank Markov projections on $(L^1(X), \mathbb{1})$, it can be seen that for an invertible measure-preserving system $\varphi: X \rightarrow X$ we have $h_{\text{bl}}(T_\varphi) = ha_\tau(\alpha_\varphi)$.

The validity of Theorem 5.2.1 for $p = 2$ can therefore be seen as a special case of [Voi95, Proposition 1.7]. However, we provide a proof that is independent of von Neumann algebra theory and that is valid for Koopman operators on L^p for all $p \in [1, \infty)$.

Let us begin preparing the proof of the main result by discussing the structure of Markov projections on L^p -spaces, which we briefly mentioned in Lemma 4.2.5.

Lemma 5.2.3. *Let $X = (X, \Sigma, \mu)$ be a probability space and $p \in [1, \infty)$ and let $P: L^p(X) \rightarrow L^p(X)$ be a linear projection. Then the following are equivalent.*

- (a) *The operator P is a Markov projection on $(L^p(X), \mathbb{1})$.*
- (b) *The operator P extends uniquely to a Markov projection \hat{P} on $(L^1(X), \mathbb{1})$.*
- (c) *The operator $\tilde{P} := \hat{P}|_{L^2(X)}$ is an orthogonal projection satisfying $\tilde{P} \geq 0$ and $\tilde{P} \mathbb{1} = \mathbb{1}$.*
- (d) *There exists a sub- σ -algebra $\Sigma_0 \subseteq \Sigma$ such that P is the conditional expectation operator $\mathbb{E}[\cdot \mid \Sigma_0]$, i.e., P is a projection onto $L^p(X_0)$, where $X_0 := (X, \Sigma_0, \mu|_{\Sigma_0})$, and*

$$\int_A P f \, d\mu = \int_A f \, d\mu \tag{5.2}$$

for all $f \in L^p(X)$ and all $A \in \Sigma_0$.

Proof. “(a) \Leftrightarrow (b)”: It follows from [EFHN15, Proposition 13.6] that P extends uniquely to a Markov operator \hat{P} on $L^1(X)$. Since P is a projection, so is \hat{P} . The restriction of a Markov projection \hat{P} to $L^p(X)$ is then clearly again a Markov projection.

“(b) \Leftrightarrow (c)”: This is [EFHN15, Theorem 13.20].

“(b) \Leftrightarrow (d)”: It follows from [And66, Lemma 4] that every Markov projection on $L^1(X)$ is a conditional expectation operator onto $L^1(X_0)$ for a suitable sub- σ -algebra $\Sigma_0 \subseteq \Sigma$. The restriction of such a conditional expectation operator to $L^p(X)$ is then a conditional expectation operator onto $L^p(X_0)$. Conversely, the conditional expectation operator onto $L^p(X_0)$ extends to a conditional expectation operator onto $L^1(X_0)$, as (5.2) implies that P is contractive in the $L^1(X)$ -norm. Clearly, this extension is a conditional expectation operator mapping to $L^1(X_0)$, and thus, a Markov projection on $L^1(X)$ by the discussion before [And66, Lemma 4]. \square

If a Markov projection $P: L^p(X) \rightarrow L^p(X)$ has finite-dimensional range, then the corresponding sub- σ -algebra $\Sigma_0 \subseteq \Sigma$ from Lemma 5.2.3 (d) is one that is generated by a finite Σ_0 -measurable partition α of X . In this case, the conditional expectation operator $\mathbb{E}[\cdot | \alpha] := P = \mathbb{E}[\cdot | \Sigma_0]$ is given by

$$\mathbb{E}[f | \alpha] = \sum_{A \in \alpha} \frac{1}{\mu(A)} \left(\int_A f \, d\mu \right) \cdot \mathbb{1}_A \quad (5.3)$$

for all $f \in L^p(X)$. In particular, the range of P is of the form

$$\text{ran}(P) = L^p(X_0) = \text{span}\{\mathbb{1}_A : A \in \alpha\},$$

and thus, $\dim(PL^p(X)) = |\alpha|$.

Using this observation, we can prove the following subadditivity property of the δ -information on L^2 -spaces.

Proposition 5.2.4. *Let $X = (X, \Sigma, \mu)$ be a probability space and consider the Banach lattice with quasi-interior point $(L^2(X), \mathbb{1})$. For each finite $\mathcal{O}_1, \mathcal{O}_2 \subseteq L^2(X)$, $\delta > 0$ we have $\log(H_{\text{bl}}(\mathcal{O}_1 \cup \mathcal{O}_2, \delta)) \leq \log(H_{\text{bl}}(\mathcal{O}_1, \delta)) + \log(H_{\text{bl}}(\mathcal{O}_2, \delta))$.*

Proof. Let P_1 and P_2 be Markov projections on $(L^2(X), \mathbb{1})$ with

$$\|P_i x - x\| \leq \delta$$

for all $x \in \mathcal{O}_i$ and all $i \in \{1, 2\}$ and minimal rank. Let $\Sigma_i \subseteq \Sigma$ be the sub- σ -algebra corresponding to P_i according to Lemma 5.2.3 (d) and let α_i be a finite Σ_i -measurable partition of X generating Σ_i . Consider the partition $\alpha := \alpha_1 \vee \alpha_2 := \{A_1 \cap A_2 : A_1 \in \alpha_1, A_2 \in \alpha_2\}$ and let $\Sigma_0 \subseteq \Sigma$ be the sub- σ -algebra generated by α . Further, let $P: L^2(X) \rightarrow L^2(X)$ be the conditional expectation operator onto $L^2(X_0)$ that is defined via (5.3), where $X_0 := (X, \Sigma_0, \mu|_{\Sigma_0})$. Then, by Lemma 5.2.3, P is a Markov projection on $(L^2(X), \mathbb{1})$ satisfying

$$\dim(PL^2(X)) = |\alpha_1 \vee \alpha_2| \leq |\alpha_1| \cdot |\alpha_2| = \dim(P_1 L^2(X)) \cdot \dim(P_2 L^2(X)).$$

Moreover, we have that $PP_i = P_iP = P$ for all $i \in \{1, 2\}$, which follows from the fact that Σ_0 is finer than both Σ_1 and Σ_2 . Thus, for each $x \in \mathcal{O}_i$ and $i \in \{1, 2\}$ we have

$$Px - x = Px - P_i x + P_i x - x = P(x - P_i x) + (P_i x - x) = (I - P)(P_i x - x).$$

Since Lemma 5.2.3 implies that P is an orthogonal projection, the operator $I - P$ is also an orthogonal projection, and thus, $\|I - P\| \leq 1$ follows. We then obtain that $\|Px - x\| \leq \|P_i x - x\| \leq \delta$ for all $x \in \mathcal{O}_i$ and all $i \in \{1, 2\}$. This yields that

$$\begin{aligned} H_{\text{bl}}(\mathcal{O}_1 \cup \mathcal{O}_2, \delta) &\leq \dim(PL^2(X)) \leq \dim(P_1L^2(X)) \cdot \dim(P_2L^2(X)) \\ &= H_{\text{bl}}(\mathcal{O}_1, \delta) \cdot H_{\text{bl}}(\mathcal{O}_2, \delta), \end{aligned}$$

and the claim follows by taking logarithms. \square

Remark 5.2.5. For $p \in [1, \infty) \setminus \{2\}$ the above proof can be modified by replacing the bound $\|I - P\|_2 \leq 1$ by the trivial bound $\|I - P\|_p \leq 2$. We then obtain the slightly weaker inequality

$$H_{\text{bl}}(\mathcal{O}_1 \cup \mathcal{O}_2, \delta) \leq H_{\text{bl}}(\mathcal{O}_1, \delta/2) \cdot H_{\text{bl}}(\mathcal{O}_2, \delta/2)$$

for all finite sets $\mathcal{O}_1, \mathcal{O}_2 \subseteq L^p(X)$ and all $\delta > 0$.

The last proposition allows us to conclude the following. Recall that every Koopman operator T_φ on $L^p(X)$ associated with an invertible measure-preserving system $\varphi: X \rightarrow X$ is a Markov automorphism on $(L^p(X), \mathbb{1})$.

Corollary 5.2.6. *Let $X = (X, \Sigma, \mu)$ be a probability space and consider the Banach lattice with quasi-interior point $(L^2(X), \mathbb{1})$. Let T be a Markov automorphism on $(L^2(X), \mathbb{1})$. Then the limit in (5.1) exists for each finite $\mathcal{O} \subseteq L^2(X)$ and $\delta > 0$.*

Note that this answers a question that was left open in [Voi95, Remark 1.6].

Proof. Define the sequence $(a_n)_{n \in \mathbb{N}}$ via

$$a_n := H_{\text{bl}}\left(\bigcup_{k=0}^{n-1} T^k \mathcal{O}, \delta\right)$$

for each $n \in \mathbb{N}$. Then for $m, n \in \mathbb{N}$ we have by Proposition 5.2.4 and Proposition 5.1.5 (iii) that

$$\begin{aligned} a_{m+n} &= H_{\text{bl}}\left(\bigcup_{k=0}^{m+n-1} T^k \mathcal{O}, \delta\right) = H_{\text{bl}}\left(\left(\bigcup_{k=0}^{m-1} T^k \mathcal{O}\right) \cup \left(\bigcup_{k=m}^{m+n-1} T^k \mathcal{O}\right), \delta\right) \\ &\leq H_{\text{bl}}\left(\bigcup_{k=0}^{m-1} T^k \mathcal{O}, \delta\right) \cdot H_{\text{bl}}\left(\bigcup_{k=0}^{n-1} T^k \mathcal{O}, \delta\right) = a_m \cdot a_n. \end{aligned}$$

Thus, the sequence $(\log(a_n))_{n \in \mathbb{N}}$ is subadditive and the claim follows from Fekete's lemma [Wal82, Corollary 4.9.1]. \square

Let us now prove the first inequality of Theorem 5.2.1.

Proposition 5.2.7. *Let $\varphi: X \rightarrow X$ be a measure-preserving system on the probability space $X = (X, \Sigma, \mu)$, let $p \in [1, \infty)$ and T_φ denote the associated Koopman operator on $L^p(X)$. Then*

$$h_{\text{bl}}(T_\varphi) \geq h_\mu(\varphi).$$

The proof is, in essence, similar to that of Theorem 4.2.1. We will repeat it here for the sake of completeness and refer to the necessary lemmas that we have already established in Section 4.2.

Proof. Let $\varepsilon > 0$ and α be a finite Σ -measurable partition of X and fix $K \in \mathbb{N}$. For $\kappa := |\alpha|$ let $\delta > 0$ be such that the conclusion of Lemma 4.2.7 is satisfied. Let P be a Markov projection on $(L^p(X), \mathbb{1})$ such that

$$\|P \mathbb{1}_{\varphi^{-k}(A)} - \mathbb{1}_{\varphi^{-k}(A)}\|_p = \|PT_\varphi^k \mathbb{1}_A - T_\varphi^k \mathbb{1}_A\|_p \leq \delta$$

for all $A \in \alpha$ and all $k \in \{0, \dots, K-1\}$. We may assume that P has minimal rank among all such projections, and hence, by a simple approximation argument, we see that it has finite-rank. So by Lemma 5.2.3 and (5.3) there exists a finite Σ -measurable partition β of X such that $P = \mathbb{E}[\cdot \mid \beta]$. Note that $\dim(PL^p(X)) = |\beta|$. By our choice of δ we then have $H_\mu(\varphi^{-k}(\alpha) \mid \beta) < \varepsilon/2$ for all $k \in \{0, \dots, K-1\}$, where $\varphi^{-k}(\alpha) := \{\varphi^{-k}(A) : A \in \alpha\}$. Thus,

$$\begin{aligned} H_\mu \left(\bigvee_{k=0}^{K-1} \varphi^{-k}(\alpha) \right) &\leq H_\mu \left(\beta \vee \left(\bigvee_{k=0}^{K-1} \varphi^{-k}(\alpha) \right) \right) \\ &= H_\mu(\beta) + h_\mu \left(\bigvee_{k=0}^{K-1} \varphi^{-k}(\alpha) \mid \beta \right) \\ &= H_\mu(\beta) + \sum_{k=0}^{K-1} H_\mu \left(\varphi^{-k}(\alpha) \mid \beta \right) \\ &< \log(|\beta|) + K \cdot \varepsilon \\ &= \log \left(H_{\text{bl}} \left(\bigcup_{k=0}^{K-1} T^k \mathcal{O}_\alpha, \delta \right) \right) + K \cdot \varepsilon, \end{aligned}$$

where $\mathcal{O}_\alpha := \{\mathbb{1}_A : A \in \alpha\}$. Dividing by K and taking the limit superior as $K \rightarrow \infty$ yields

$$\begin{aligned} h_\mu(\varphi, \alpha) &= \lim_{K \rightarrow \infty} \frac{1}{K} H_\mu \left(\bigvee_{k=0}^{K-1} \varphi^{-k}(\alpha) \right) \\ &\leq \limsup_{K \rightarrow \infty} \frac{1}{K} \log \left(H_{\text{bl}} \left(\bigcup_{k=0}^{K-1} T^k \mathcal{O}_\alpha, \delta \right) \right) + \varepsilon \\ &= h_{\text{bl}}(T, \mathcal{O}_\alpha, \delta) + \varepsilon \leq h_{\text{bl}}(T) + \varepsilon. \end{aligned}$$

Now the statement follows by first noting that this inequality holds for all $\varepsilon > 0$ and then taking the supremum over all finite Σ -measurable partitions α of X . \square

By approximating arbitrary finite sets in L^p -spaces by simple functions, we can further reduce the computation of the Banach lattice entropy on L^p -spaces to sets of indicator functions. This adapts [Voi95, Proposition 1.4] to our setting.

Lemma 5.2.8. *Let $X = (X, \Sigma, \mu)$ be a probability space, $p \in [1, \infty)$ and consider the Banach lattice with quasi-interior point $(L^p(X), \mathbb{1})$. Let $T: L^p(X) \rightarrow L^p(X)$ be a Markov operator on $(L^p(X), \mathbb{1})$. Then*

$$h_{\text{bl}}(T) = \sup_{\alpha} h_{\text{bl}}(T, \{\mathbb{1}_A : A \in \alpha\})$$

where the supremum is taken over all finite Σ -measurable partitions α of X .

Proof. Note that it suffices to show the inequality “ \leq ” as the other direction follows directly from the definition of $h_{\text{bl}}(T)$.

Let $\mathcal{O} = \{f_1, \dots, f_m\} \subseteq L^p(X)$ be finite and $\delta_1, \delta_2 > 0$. For each $i \in \{1, \dots, m\}$ there exist finite Σ -measurable partitions α_i such that the simple function $s_i = \sum_{A \in \alpha_i} \lambda_{i,A} \mathbb{1}_A$ satisfies $\|f_i - s_i\| < \delta_2/2$. Consider the common refinement $\alpha := \alpha_1 \vee \dots \vee \alpha_m$. Then each s_i is contained in $\text{span}\{\mathbb{1}_A : A \in \alpha\}$ and thus, by Proposition 5.1.5 (iv) and (v), we obtain that

$$h_{\text{bl}}(T, \mathcal{O}, \delta_1 + \delta_2) \leq h_{\text{bl}}(T, \hat{\mathcal{O}}, \delta_1) \leq h_{\text{bl}}(T, \{\mathbb{1}_A : A \in \alpha\}),$$

where $\hat{\mathcal{O}} := \{s_1, \dots, s_m\}$. Taking the supremum over all $\delta_2 > 0$, we obtain from Proposition 5.1.5 (ii) that

$$h_{\text{bl}}(T, \mathcal{O}, \delta_1) \leq \sup_{\alpha} h_{\text{bl}}(T, \{\mathbb{1}_A : A \in \alpha\}).$$

Now taking the supremum over all $\delta_1 > 0$ and all finite $\mathcal{O} \subseteq L^p(X)$ yields the other inequality. \square

Let $X = (X, \Sigma, \mu)$ be a probability space and let α be a finite Σ -measurable partition of X . Recall from Section 3.1 the definition of the *information function of α with respect to μ* as

$$I_{\mu}(\alpha) := I_{\mu} \circ \alpha = - \sum_{A \in \alpha} \mathbb{1}_A \log(\mu(A)).$$

The Shannon–McMillan–Breiman theorem now describes the asymptotic behavior of the information functions of the partitions $\varphi_0^{n-1}\alpha$ as n tends to infinity.

Lemma 5.2.9 (Shannon–McMillan–Breiman). *Let $\varphi: X \rightarrow X$ be an ergodic measure-preserving system on the probability space $X = (X, \Sigma, \mu)$ and α be a finite Σ -measurable partition of X . Then for μ -almost every $x \in X$ we have*

$$\lim_{n \rightarrow \infty} \frac{1}{n} I_{\mu}(\varphi_0^{n-1}\alpha)(x) = h_{\mu}(\varphi, \alpha).$$

The convergence also holds in $L^1(X)$.

This version of the theorem follows from [Par69, Theorem 2.5 and the discussion in Section 3.1].

The following proposition establishes the second inequality between the Banach lattice entropy and the measure-theoretic entropy from Theorem 5.2.1 and is proved similarly to [Voi95, Proposition 1.7].

Proposition 5.2.10. *Let $\varphi: X \rightarrow X$ be an ergodic measure-preserving system on the probability space $X = (X, \Sigma, \mu)$ and let T_φ denote the associated Koopman operator on $L^p(X)$. Then*

$$h_{\text{bl}}(T_\varphi) \leq h_\mu(\varphi).$$

Proof. By Lemma 5.2.8 it suffices to show that for each finite Σ -measurable partition α we have

$$h_{\text{bl}}(T_\varphi, \mathcal{O}_\alpha) \leq h_\mu(\varphi, \alpha),$$

where $\mathcal{O}_\alpha := \{\mathbb{1}_A : A \in \alpha\}$. So let α be such a partition and let $\varepsilon > 0$. Then by the Shannon–McMillan–Breiman Theorem 5.2.9 and the Markov inequality there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ we have

$$\begin{aligned} & \mu \left(\underbrace{\left\{ x \in X : \left| \frac{1}{n} I_\mu(\varphi_0^{n-1} \alpha)(x) - h_\mu(\varphi, \alpha) \right| > \varepsilon \right\}}_{B:=} \right) \\ & \leq \frac{1}{\varepsilon} \int_X \left| \frac{1}{n} I_\mu(\varphi_0^{n-1} \alpha) - h_\mu(\varphi, \alpha) \right| d\mu < \varepsilon. \end{aligned}$$

Fix such an $n \geq n_0$. Note that $I_\mu(\varphi_0^{n-1} \alpha)$ is constant on each set $A \in \varphi_0^{n-1} \alpha$, and thus, B is a union of such sets. So let $\alpha_1 \subseteq \varphi_0^{n-1} \alpha$ be such that $B = \bigcup_{A \in \alpha_1} A$ and set $\alpha_2 := \varphi_0^{n-1} \alpha \setminus \alpha_1$. Then for every $A \in \alpha_2$ we have

$$\left| \frac{-\log(\mu(A))}{n} - h_\mu(\varphi, \alpha) \right| < \varepsilon.$$

Rearranging this yields

$$e^{-n(h_\mu(\varphi, \alpha) + \varepsilon)} < \mu(A) < e^{-n(h_\mu(\varphi, \alpha) - \varepsilon)}$$

for all $A \in \alpha_2$. From the lower bound we conclude that the number of sets in α_2 is at most $e^{n(h_\mu(\varphi, \alpha) + \varepsilon)}$. We define a new partition β of X by setting

$$\beta := \alpha_2 \cup \{B\},$$

which satisfies $|\beta| \leq e^{n(h_\mu(\varphi, \alpha) + \varepsilon)} + 1$. Let P be the Markov projection on $(L^p(X), \mathbb{1})$ corresponding to the sub- σ -algebra generated by β given by (5.3). Then $\dim(PL^p(X)) = |\beta|$ and for each $A \in \alpha$ and each $k \in \{0, \dots, n-1\}$ we have

$$\|PT_\varphi^k \mathbb{1}_A - T_\varphi^k \mathbb{1}_A\|_p^p = \int_X \left| \mathbb{E}[\mathbb{1}_{\varphi^{-k}(A)} \mid \beta] - \mathbb{1}_{\varphi^{-k}(A)} \right|^p d\mu = 0,$$

if $\varphi^{-k}(A) \in \alpha_2$ and

$$\|PT_\varphi^k \mathbb{1}_A - T_\varphi^k \mathbb{1}_A\|_p^p = \int_X \left| \mathbb{E}[\mathbb{1}_{\varphi^{-k}(A)} \mid \beta] - \mathbb{1}_{\varphi^{-k}(A)} \right|^p d\mu \leq \mu(B) < \varepsilon,$$

if $\varphi^{-k}(A) \in \alpha_1$. Thus, we obtain

$$\begin{aligned} H_{\text{bl}}\left(\bigcup_{k=0}^{n-1} T_\varphi^k \mathcal{O}_\alpha, \varepsilon^{1/p}\right) &\leq \dim(PL^p(X)) \leq |\beta| \\ &\leq e^{n(h_\mu(\varphi, \alpha) + \varepsilon)} + 1. \end{aligned}$$

Taking logarithms, dividing by n and taking the limit superior for $n \rightarrow \infty$ yields

$$h_{\text{bl}}(T_\varphi, \mathcal{O}_\alpha, \varepsilon^{1/p}) \leq h_\mu(\varphi, \alpha) + \varepsilon.$$

Now the claim follows by letting ε tend to zero. \square

The author currently does not know whether the ergodicity assumption in Proposition 5.2.10 can be removed, but conjectures that this is the case. Formulations of the Shannon–McMillan–Breiman theorem for non-ergodic systems exist (see, e.g. [Par69, Theorem 2.5] or [Kre85, Theorem 9.2.5]), but their limit function is not necessarily constant; we only know that it is measurable with respect to the sub- σ -algebra of φ -invariant sets, or in other words, the limit function is in $\text{fix}(T_\varphi)$.

Proving Proposition 5.2.10 in the non-ergodic case would thus require constructing suitable finite partitions that approximate the information function well enough on each ergodic component. Figuring out the details of this construction might be part of future research.

5.3 Topological Entropy

Having explored the equivalence of Banach lattice entropy to the measure-theoretic entropy for Koopman operators on $L^p(X)$, our next goal is to establish a similar relationship between Banach lattice entropy and the topological entropy of topological dynamical systems on compact spaces.

Throughout this section, we let $\varphi: K \rightarrow K$ be a continuous map on a compact Hausdorff space K and denote by T_φ the associated Koopman operator on $C(K)$. The main result of this section is the following theorem.

Theorem 5.3.1. *Let $\varphi: K \rightarrow K$ be an invertible and continuous map on the compact Hausdorff space K and let T_φ denote the associated Koopman operator on $C(K)$. Then*

$$h_{\text{bl}}(T_\varphi) = h_{\text{top}}(\varphi).$$

We will divide the proof of this theorem into Proposition 5.3.8 and Proposition 5.3.7, which show the two inequalities separately. Note that the invertibility assumption is only necessary for the inequality “ \geq ”.

Recall from Section 3.4 that the topological entropy $h_{\text{top}}(\varphi)$ of φ is defined via open covers of K . In order to relate this notion to our operator-theoretic approximation entropies, we need to approximate open covers by partitions of unity.

Definition 5.3.2. Let K be a compact Hausdorff space.

- (i) A (finite) partition of unity on K is a finite set $\mathcal{F} \subseteq C(K)_+$ such that $\sum_{f \in \mathcal{F}} f = \mathbb{1}$. We denote by

$$\alpha_{\mathcal{F}} := \{\{x \in K : f(x) > 0\} : f \in \mathcal{F}\}$$

its associated open cover of K .

- (ii) Let α be an open cover of K . A partition of unity \mathcal{F} is said to be *subordinate to the open cover α* of K if $\alpha_{\mathcal{F}} \preceq \alpha$.
- (iii) A partition of unity \mathcal{F} is called *minimal* if for every $f \in \mathcal{F}$ there exists $x \in K$ such that $f(x) = 1$.

Note that, in compact Hausdorff spaces, partitions of unity always exist. Moreover, for each finite open cover α there exists a partition of unity \mathcal{F} that is subordinate to α . Indeed, this follows from Urysohn's lemma.

Lemma 5.3.3. *Let K be a compact Hausdorff space and α be a finite open cover of K . Then there exists a minimal partition of unity \mathcal{F} on K that is subordinate to α .*

Proof. By passing to a minimal subcover we may assume that α is minimal, i.e., that no proper subset of α is a cover of K . For each $U \in \alpha$ denote by \tilde{U} the union of all sets in $\alpha \setminus \{U\}$. By Urysohn's lemma there exists a continuous function $g_U : K \rightarrow [0, 1]$ such that $g_U = 0$ on $K \setminus U$ and $g_U = 1$ on $K \setminus \tilde{U}$. Then the function $h := \sum_{U \in \alpha} g_U$ is strictly positive on K as α is a finite cover of K . Thus, the set

$$\mathcal{F} := \left\{ \frac{g_U}{h} : U \in \alpha \right\}$$

is a partition of unity that is subordinate to α . It follows from the minimality of α that \mathcal{F} is minimal. \square

Partitions of unity exist for more general spaces. We refer the interested reader to [Eng89, Theorem 5.1.9] for more information on this topic.

Note that a partition of unity \mathcal{F} is minimal if and only if the covering $\alpha_{\mathcal{F}}$ fails to be a cover when we take away a single element. An even stronger characterization is given in the following lemma.

Lemma 5.3.4. *Let K be a compact Hausdorff space and \mathcal{F} be a partition of unity on K . Then \mathcal{F} is minimal if and only if for each open covering α with $\alpha \preceq \alpha_{\mathcal{F}}$ we have $|\alpha| \geq |\mathcal{F}|$.*

Proof. “ \Rightarrow ”: Let \mathcal{F} be minimal and let α be an open covering with $\alpha \preceq \alpha_{\mathcal{F}}$. Fix $f \in \mathcal{F}$ and choose $x_f \in K$ such that $f(x_f) = 1$. Then for $A \in \alpha$ with $x_f \in A$ it follows that $A \supseteq U_f := \{x \in K : f(x) > 0\}$, as $x_f \notin U_g$ for $f \neq g$ with $g \in \mathcal{F}$. Thus, different x_f belong to different elements of α and hence, $|\alpha| \geq |\mathcal{F}|$.

“ \Leftarrow ”: Assume that \mathcal{F} is not minimal. Then there exists $g \in \mathcal{F}$ such that $\sum_{f \neq g} f$ is positive everywhere. Hence, the collection

$$\{\{x \in K : f(x) > 0\} : f \in \mathcal{F}, f \neq g\}$$

is an open covering that is a refinement of $\alpha_{\mathcal{F}}$ but has only $|\mathcal{F}| - 1$ elements. \square

The following lemma shows that each finite-dimensional subspace of $C(K)$ that contains the constant functions has a basis that is a partition of unity.

Lemma 5.3.5. *Let K be a compact Hausdorff space and let $F \subseteq C(K)$ be a finite-dimensional subspace that contains the constant functions. Then F has a basis \mathcal{F} that is a partition of unity on K .*

Proof. Note that the positive cone $F_+ := F \cap C(K)_+$ of F satisfies $F = F_+ - F_+$. Indeed, the inclusion “ \supseteq ” is clear. For the reverse inclusion take $f \in F$. Then $f + \|f\| \cdot \mathbb{1} \in F_+$, and thus, $f = (f + \|f\| \cdot \mathbb{1}) - \|f\| \cdot \mathbb{1} \in F_+ - F_+$. Hence, the cone F_+ linearly spans F . Thus, we can choose a basis B of F that is contained in F_+ with $\mathbb{1} \in B$. Consider the function

$$h := \sum_{\substack{f \in B \\ f \neq \mathbb{1}}} f.$$

Set $\mathcal{F} := \left\{ \frac{f}{\|h\|_\infty} : f \in B, f \neq \mathbb{1} \right\} \cup \left\{ \mathbb{1} - \frac{h}{\|h\|_\infty} \right\}$. Then \mathcal{F} is a partition of unity on K and by construction $\text{span } \mathcal{F} = F$. \square

Unfortunately, the proof does not yield minimality of the partition of unity. For ranges of Markov projections this additional property follows similarly to [JP83, Theorem 2]. We use this to prove the following proposition on the structure of Markov projections on $(C(K), \mathbb{1})$.

Proposition 5.3.6. *Let K be a compact Hausdorff space and let $(C(K), \mathbb{1})$ be the associated Banach lattice with quasi-interior point. Then an operator $P: C(K) \rightarrow C(K)$ is a Markov projection on $(C(K), \mathbb{1})$ if and only if there exists a minimal partition of unity \mathcal{F} on K and a family of probability measures $\{\mu_f : f \in \mathcal{F}\}$ such that $\langle g, \mu_f \rangle = \delta_{f,g}$ for all $f, g \in \mathcal{F}$ and*

$$Ph = \sum_{f \in \mathcal{F}} \langle h, \mu_f \rangle f \tag{5.4}$$

for all $h \in C(K)$. In particular, the probability measures have pairwise disjoint supports with $\text{supp}(\mu_f) \subseteq \{x \in K : f(x) = 1\}$.

Proof. “ \Rightarrow ”: By Lemma 5.3.5 there exists a partition of unity \mathcal{F} that is a basis of $PC(K)$. Moreover, there are functionals $\{\mu_f : f \in \mathcal{F}\}$ on $C(K)$ such that

$$Ph = \sum_{f \in \mathcal{F}} \mu_f(h) \cdot f$$

for all $h \in C(K)$. Since P is positive, the functionals are also positive. Moreover, $Pf = f$ and the fact that \mathcal{F} is a basis show that $\langle g, \mu_f \rangle = \delta_{f,g}$ for all $f, g \in \mathcal{F}$. As the μ_f 's are probability measures and $f \leq \mathbb{1}$, the last assertion follows readily.

“ \Leftarrow ”: If P is of the form (5.4), then it is straightforward to verify that P is a Markov projection on $(C(K), \mathbb{1})$. \square

The following proposition proves the first direction that shows how the Banach lattice entropy of the Koopman operator T_φ is bounded by the topological entropy of

φ . This is proved analogously to [Voi95, Proposition 4.8] and is the easier inequality to show.

Proposition 5.3.7. *Let $\varphi: K \rightarrow K$ be a continuous map on a compact Hausdorff space K and let T_φ denote the associated Koopman operator on $C(K)$. Then*

$$h_{\text{bl}}(T_\varphi) \leq h_{\text{top}}(\varphi).$$

Proof. Let $\mathcal{O} \subseteq C(K)$ be finite and let $\delta > 0$. Fix $n \in \mathbb{N}$. For $f \in \mathcal{O}$ and $x \in K$ we set

$$U_x := \{y \in K : \forall f \in \mathcal{O} : |f(x) - f(y)| < \delta\}.$$

Let α be a finite subcover of $\{U_x : x \in K\}$ (which is an open cover of K) and let β be a subcover of minimal cardinality of

$$\varphi_0^{n-1}\alpha = \bigvee_{k=0}^{n-1} \varphi^{-k}(\alpha),$$

where $\varphi^{-k}(\alpha) := \{\varphi^{-k}(U) : U \in \alpha\}$. Note that β is finite. Let \mathcal{G} be a partition of unity subordinate to β . Then by Lemma 5.3.4 there exists a point $x_g \in K$ for each $g \in \mathcal{G}$ such that $g(x_h) = \delta_{g,h}$ for all $g, h \in \mathcal{G}$. Consider the linear operator

$$P: C(K) \rightarrow C(K), \quad h \mapsto \sum_{g \in \mathcal{G}} h(x_g) \cdot g.$$

Then P is a Markov projection by Proposition 5.3.6. Note that $g(x) > 0$ for some $x \in K$ and $g \in \mathcal{G}$ implies that $x, x_g \in U$ for some common open set $U \in \beta$. As this means that there are sets $V_0, \dots, V_{n-1} \in \alpha$ with $x, x_g \in \varphi^{-k}(V_k)$ for all $k \in \{0, \dots, n-1\}$, it follows that we have for said k and all $f \in \mathcal{O}$ that

$$|f(\varphi^k(x)) - f(\varphi^k(x_g))| \leq \delta.$$

Hence, we have for $k \in \{0, \dots, n-1\}$, all $f \in \mathcal{O}$ and $x \in K$ that

$$\begin{aligned} |(PT_\varphi^k f)(x) - (T_\varphi^k f)(x)| &\leq \sum_{g \in \mathcal{G}} g(x) |(T_\varphi^k f)(x_g) - (T_\varphi^k f)(x)| \\ &= \sum_{\substack{g \in \mathcal{G} \\ g(x) > 0}} g(x) |(T_\varphi^k f)(x_g) - (T_\varphi^k f)(x)| \leq \delta. \end{aligned}$$

It follows that

$$H_{\text{bl}} \left(\bigcup_{k=0}^{n-1} T_\varphi^k \mathcal{O}, \delta \right) \leq \dim(PC(K)) = |\beta| = N \left(\bigvee_{k=0}^{n-1} \varphi^{-k}(\alpha) \right),$$

where N denotes the covering number from Section 3.4. Applying the logarithm, dividing by n , passing to the limit (superior) as $n \rightarrow \infty$ and then taking the supremum over all finite open covers α of K yields $h_{\text{bl}}(T_\varphi, \mathcal{O}, \delta) \leq h_{\text{top}}(\varphi)$. Now taking the supremum over all δ and all finite subsets $\mathcal{O} \subseteq C(K)$ yields the claim. \square

Proposition 5.3.8. *Let $\varphi: K \rightarrow K$ be an invertible and continuous map on a compact Hausdorff space K and let T_φ denote the associated Koopman operator on $C(K)$. Then*

$$h_{\text{bl}}(T_\varphi) \geq h_{\text{top}}(\varphi).$$

In [KL05] a similar entropy notion called *contractive approximation entropy* for continuous maps on compact metric spaces is introduced and shown to coincide with the topological entropy. We introduce this approximation entropy (for the special case in which $X = Y$ and $\gamma = \text{id}_X$).

Let E be a Banach space and $T: E \rightarrow E$ be a bounded linear operator. For a finite set $\mathcal{O} \subseteq E$ and $\delta > 0$ we say that (Ψ_1, Ψ_2, d) , where $\Psi_1: E \rightarrow \ell_\infty^d$ and $\Psi_2: \ell_\infty^d \rightarrow E$ are contractive linear maps and $d \in \mathbb{N}$, is δ -contractive for \mathcal{O} if

$$\|\Psi_2 \Psi_1 f - f\| < \delta$$

for all $f \in \mathcal{O}$. We define the δ -contractive information of \mathcal{O} as

$$H_{\text{ca}}(\mathcal{O}, \delta) := \min\{d \in \mathbb{N} : (\Psi_1, \Psi_2, d) \text{ is } \delta\text{-contractive for } \mathcal{O}\}.$$

Now we define similarly as before the *contractive approximation entropy* of T with respect to \mathcal{O} and δ as

$$h_{\text{ca}}(T, \mathcal{O}, \delta) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log \left(H_{\text{ca}} \left(\bigcup_{k=0}^{n-1} T^k \mathcal{O}, \delta \right) \right),$$

and the *contractive approximation entropy* of T as

$$h_{\text{ca}}(T) := \sup_{\delta > 0} \sup_{\mathcal{O} \subseteq E \text{ finite}} h_{\text{ca}}(T, \mathcal{O}, \delta).$$

We will use the following fact that is proved in [KL05, Proposition 3.1] as the main tool for the proof of Proposition 5.3.8.

Lemma 5.3.9. *Let K be a compact Hausdorff space and let $\varphi: K \rightarrow K$ be an invertible and continuous map. Denote by T_φ the associated Koopman operator on $C(K)$. Then*

$$h_{\text{ca}}(T_\varphi) = h_{\text{top}}(\varphi).$$

Proof of Proposition 5.3.8. Let $\mathcal{O} \subseteq C(K)$ be finite and let $\delta > 0$. Fix $n \in \mathbb{N}$. Let $P: C(K) \rightarrow C(K)$ be a Markov projection on minimal rank that satisfies

$$\|PT_\varphi^k f - T_\varphi^k f\| < \delta \tag{5.5}$$

for all $f \in \mathcal{O}$ and all $k \in \{0, \dots, n-1\}$. By Proposition 5.3.6 there exists a minimal partition of unity \mathcal{G} on K and a family of probability measures $\{\mu_g : g \in \mathcal{G}\}$ such that $\langle h, \mu_g \rangle = \delta_{g,h}$ for all $g, h \in \mathcal{G}$ and

$$Ph = \sum_{g \in \mathcal{G}} \langle h, \mu_g \rangle \cdot g.$$

Denote $d := |\mathcal{G}|$ and define the maps

$$\Psi_1: C(K) \rightarrow \ell_\infty(\mathcal{G}; \mathbb{K}), \quad h \mapsto (\langle h, \mu_g \rangle)_{g \in \mathcal{G}},$$

and

$$\Psi_2: \ell_\infty(\mathcal{G}; \mathbb{K}) \rightarrow C(K), \quad (c_g)_{g \in \mathcal{G}} \mapsto \sum_{g \in \mathcal{G}} c_g \cdot g.$$

Then by construction $\|\Psi_1\|, \|\Psi_2\| \leq 1$ and $P = \Psi_2\Psi_1$. Thus, by equation (5.5) the triple (Ψ_1, Ψ_2, d) is δ -contractive for $\bigcup_{k=0}^{n-1} T_\varphi^k \mathcal{O}$. It follows that

$$H_{\text{ca}} \left(\bigcup_{k=0}^{n-1} T_\varphi^k \mathcal{O}, \delta \right) \leq \dim(PC(K)) = |\mathcal{G}| = H_{\text{bl}} \left(\bigcup_{k=0}^{n-1} T_\varphi^k \mathcal{O}, \delta \right).$$

Taking logarithms, dividing by n and passing to the limit (superior) for $n \rightarrow \infty$ yields

$$h_{\text{ca}}(T_\varphi, \mathcal{O}, \delta) \leq h_{\text{bl}}(T_\varphi, \mathcal{O}, \delta).$$

Now taking the supremum over all $\delta > 0$ and all finite subsets $\mathcal{O} \subseteq C(K)$ together with Lemma 5.3.9 yields

$$h_{\text{top}}(\varphi) = h_{\text{ca}}(T_\varphi) \leq h_{\text{bl}}(T_\varphi). \quad \square$$

5.4 Notes

The results presented in this chapter aim to provide a new generalization of measure-theoretic and topological entropy via an entropy concept that is applicable to both Koopman operators on L^p of measure-preserving systems and Koopman operators on $C(K)$ of topological dynamical systems.

Other generalizations have been pursued by Günther Palm in his works [Pal76] and [Pal77]. This was further studied by Nikita Moriakov in his thesis [Mor16, Chapter 6]. Palm's main idea is different from our functional-analytic approach. The main object for him is a distributive lattice (V, \leq) with a greatest element 1 and a least element 0 (note the similarities and differences with the definition of a Boolean algebra in Section 2.1) that is endowed with a *measurement function* $m: V \rightarrow [0, \infty)$ satisfying $m(0) = 0$, $m(1) > 0$ and $m(a \vee b) = m(b)$ for all $a, b \in V$ with $m(a) = 0$. If $f: V \rightarrow V$ is a lattice homomorphism satisfying $m(f(a)) = 0$ whenever $m(a) = 0$, then (V, m, f) is called a *dynamical lattice*. A finite cover in this setting is any finite subset $\alpha \subseteq V$ with $\sup \alpha = 1$ and a cover α is a *refinement* ($\alpha \preceq \beta$) of a cover β if for every $a \in \alpha$ there exists $b \in \beta$ such that $a \leq b$. The Palm entropy is then given in multiple steps by first defining

$$H^*(\alpha) := - \sum_{a \in \alpha} \frac{m(a)}{M(\alpha)} \log \left(\frac{m(a)}{M(\alpha)} \right)$$

for a finite cover α of V , where $M(\alpha) := \sum_{a \in \alpha} m(a)$, and comparing this for all coarser covers with lower cardinality

$$\hat{H}(\alpha) := \sup \{H^*(\beta) : \beta \succeq \alpha, |\beta| \leq |\alpha|\}.$$

Then we set the *entropy of a cover* α to be

$$H(\alpha) := \inf \left\{ \sum_{k=1}^n \hat{H}(\beta_k) : n \in \mathbb{N}, \bigvee_{k=1}^n \beta_k \succeq \alpha \right\}.$$

The dynamic element comes into play by considering the iterated images of covers under the lattice homomorphism f and defining the entropy rate as

$$h(f) := \sup_{\alpha} \lim_{n \rightarrow \infty} \frac{1}{n} H \left(\bigvee_{k=0}^{n-1} f^k(\alpha) \right),$$

where the supremum is taken over all finite covers α of V .

If $\varphi: \Omega \rightarrow \Omega$ is a topological dynamical system on a compact Hausdorff space K , then take V to be the lattice of open subsets of K and $m(a) = 1$ if $a \neq \emptyset$ and $m(a) = 0$ otherwise. The lattice homomorphism is given by $f(a) = \varphi^{-1}(a)$ for all $a \in V$. In this case it is simple to see that $h(f) = h_{\text{top}}(\varphi)$, the classical topological entropy of φ . If (X, Σ, μ) is a probability space and $\varphi: X \rightarrow X$ is a measure-preserving dynamical system, then take V to be the lattice of measurable subsets of X and $m(a) = \mu(a)$ for all $a \in V$. Then let the lattice homomorphism be given by $f(a) = \varphi^{-1}(a)$ for all $a \in V$. In this case we can show that $h(f) = h_{\mu}(\varphi)$, the classical measure-theoretic entropy of φ . Palm's definition of lattice entropy is therefore much closer to the classical definitions and does not have an approximation element as our definition does.

Unsurprisingly, the result of Proposition 5.2.7 can be used to abbreviate the proof of Theorem 4.2.1, by following a similar chain of arguments to those used in the proof of Theorem 4.3.3. By virtue of Proposition 5.1.5 (iii) one might even be able to find a variant of Theorem 4.2.1 that works for delay subspaces F_n just as in Theorem 4.3.3. Note that there K_0 could be chosen independently of the delay n . This might be part of future research.

There should be few to no obstacles in generalizing the definition of Banach lattice entropy to group representations by Markov operators of an amenable group. The proof of Proposition 5.2.7 should be analogous, and the proof of Proposition 5.2.10 will require the use of Lindenstrauss' generalization of the Shannon–McMillan–Breiman theorem to amenable group actions [Lin01, Theorem 1.3] or [KL16, Theorem 9.25]. The remaining challenge will be to get rid of the ergodicity assumption. The proof of Proposition 5.3.7 should be adaptable using Følner sequences.

In order to generalize the proof of Proposition 5.3.8 we will need a generalization of the Lemma 5.3.9 to amenable group actions. The current proof of Lemma 5.3.9 given in [KL05, Proposition 3.1] relies on results in [Voi95], which in turn rely on results in [CNT87]. Following this chain of proofs is challenging and turned out to be beyond the scope of this work. However, it is possible that a generalization of Lemma 5.3.9 to amenable group actions exists.

Note that we have excluded the case $p = \infty$ in our treatment of Banach lattice entropy on the L^p -spaces in Section 5.2. The reason for this is that we conjecture that a similar characterization will not hold for the space $L^\infty(X)$.

Uniform Ergodicity for Koopman Operators

In Chapter 1 we introduced the concept of a topological dynamical system as a continuous mapping $\varphi: \Omega \rightarrow \Omega$ on a completely regular Hausdorff space Ω . Its Koopman operator (see Definition 1.1.6)

$$T_\varphi: C_b(\Omega) \rightarrow C_b(\Omega), \quad f \mapsto f \circ \varphi$$

was remarked to be a *unital lattice homomorphism*, i.e., it satisfies $T_\varphi \mathbb{1} = \mathbb{1}$ and $|T_\varphi f| = T_\varphi |f|$ for all $f \in C_b(\Omega)$. Similarly, in Chapter 2 we have extended the concept of Koopman operators to accommodate measure-preserving maps $\varphi: X \rightarrow X$ on a probability space $X = (X, \Sigma, \mu)$. For $p \in [1, \infty]$ it is defined to be the composition mapping (see Definition 2.1.3)

$$T_\varphi: L^p(X) \rightarrow L^p(X), \quad f \mapsto f \circ \varphi.$$

We noted in Remark 2.1.4 that in this case the Koopman operator is also a unital lattice homomorphism. It is exactly this structure of the Koopman operator that helps us further characterize the property of being uniformly ergodic. A linear operator $T: E \rightarrow E$ on a Banach space E is called *uniformly ergodic* if the sequence of Cesàro averages

$$A_n[T] := \frac{1}{n} \sum_{k=0}^{n-1} T^k$$

for $n \in \mathbb{N}$ converges in operator norm. Then the sequence of Cesàro averages converges to a projection $P: E \rightarrow E$ with $PE = \text{fix}(T)$ (see, e.g., [EFHN15, Theorem 8.5 on p. 138]).

In Section 6.1 we provide equivalent conditions for uniform ergodicity for lattice homomorphisms on Banach lattices by means of a simple spectral condition. We prove that this spectral condition, the uniform ergodicity, and the uniform convergence to a periodic operator are equivalent notions.

In Sections 6.2 and 6.3 we apply these results to Koopman operators of topological dynamical systems and measure-preserving dynamical systems, respectively. For Koopman operators of topological dynamical systems, we show that uniform convergence to a periodic operator can be strengthened to eventual periodicity of the underlying map. For measure-preserving Koopman operators, we show that uniform convergence is even equivalent to periodicity of the operator.

We close this chapter with Section 6.4, in which we discuss the stronger notions of quasi-compactness and compactness for lattice homomorphisms, and in particular, Koopman operators.

The Sections 6.1, 6.2 and 6.3 are based on the author's article [Höl25]. Section 6.4 has been newly developed for this thesis.

6.1 Uniform Ergodicity of Lattice Homomorphisms

A well-known characterization of uniform ergodicity of linear operators is the following from [Dun43, Theorem 3.16]:

Theorem 6.1.1. *Let E be a complex Banach space and let $T: E \rightarrow E$ be a bounded linear operator with $\lim_{n \rightarrow \infty} \frac{1}{n} \|T^n\| = 0$. Then T is uniformly ergodic if and only if one of the following conditions is true:*

- (i) *The point $\lambda = 1$ is in the resolvent set $\rho(T)$ of T .*
- (ii) *The point $\lambda = 1$ is a pole of the resolvent $R(\cdot, T)$.*

If the second condition is true, then $\lambda = 1$ is a pole of first order.

An operator $T: E \rightarrow E$ that satisfies $\lim_{n \rightarrow \infty} \frac{1}{n} \|T^n\| = 0$ is called *Cesàro bounded*. This condition implies immediately that the spectral radius is bounded by 1, i.e. $r(T) = \lim_{n \rightarrow \infty} \|T^n\|^{1/n} \leq 1$.

For both topological and measure-preserving systems, the fixed space $\text{fix}(T_\varphi)$ of the Koopman operator T_φ contains the constant functions. It follows that $1 \notin \rho(T_\varphi)$ and therefore the case (i) in Theorem 6.1.1 is never attained. Thus, a uniformly ergodic Koopman operator always has $\lambda = 1$ as a pole of the resolvent. It follows that 1 is a topologically isolated point in $\sigma(T_\varphi)$ if T_φ is uniformly ergodic. We will show in Theorem 6.1.2 that this condition is not only necessary but also sufficient for the uniform ergodicity of Koopman operators. We do so in the more general context of lattice homomorphisms with spectral radius 1, to which the Koopman operators belong.

Let E now be a complex Banach lattice and $T: E \rightarrow E$ be a lattice homomorphism that has spectral radius $r(T) = 1$. We investigate the long-term behavior of the powers $(T^n)_{n \in \mathbb{N}}$ in the case 1 is an isolated value of $\sigma(T)$. We recall that an eigenvalue $\lambda \in \mathbb{C}$ of a linear operator T is called *semi-simple* if $\ker(\lambda - T) = \ker((\lambda - T)^n)$ holds for $n = 2$, or equivalently, for all $n \in \mathbb{N}$. We say that T is *periodic* if there exists $k \in \mathbb{N}$ such that $T^k = I$.

Theorem 6.1.2. *Let E be a complex Banach lattice and $T: E \rightarrow E$ be a lattice homomorphism that has spectral radius $r(T) = 1$. Suppose that 1 is an isolated value in the spectrum $\sigma(T)$. Then there exists a closed lattice ideal I_{stab} and a closed sublattice E_{per} of E such that E decomposes as*

$$I_{\text{stab}} \oplus E_{\text{per}} = E$$

and

- (i) $I_{\text{stab}}, E_{\text{per}}$ are T -invariant,
- (ii) $T|_{E_{\text{per}}}$ is periodic and
- (iii) $\lim_{n \rightarrow \infty} \|(T|_{I_{\text{stab}}})^n\| = 0$.

Moreover, the peripheral spectrum $\sigma(T) \cap \mathbb{T}$ consists of a finite union of roots of unity, every $\lambda \in \sigma(T) \cap \mathbb{T}$ is a semi-simple eigenvalue and the equality

$$E_{\text{per}} = \bigoplus_{\lambda \in \sigma(T) \cap \mathbb{T}} \ker(\lambda - T) \tag{6.1}$$

holds, i.e., E_{per} is the direct sum of eigenspaces of T corresponding to unimodular eigenvalues of T .

Remark 6.1.3 (JdLG-decomposition). The decomposition $E = I_{\text{stab}} \oplus E_{\text{per}}$ in Theorem 6.1.2 is a special case of the Jacobs–de Leeuw–Glicksberg decomposition from Definition 2.2.6. In fact, we have that $E_{\text{per}} = E_{\text{rev}}$ and $I_{\text{stab}} = E_{\text{aws}}$. Note that in this case E_{per} is even a lattice ideal of E .

Our primary tool in establishing the proof of Theorem 6.1.2 is the observation that the spectrum of a lattice homomorphism T is cyclic (see [Sch74, Theorem V.4.4 on p. 325]), i.e., if $\lambda = |\lambda| e^{i\theta} \in \sigma(T)$ for some appropriate phase $\theta \in [0, 2\pi)$, then $|\lambda| e^{ik\theta} \in \sigma(T)$ for all $k \in \mathbb{Z}$. We employ this observation to describe the structure of the peripheral spectrum of T .

Lemma 6.1.4. *Let E be a complex Banach lattice and $T: E \rightarrow E$ be a Banach lattice homomorphism with $r(T) = 1$ and suppose that $\sigma(T) \cap \mathbb{T}$ is a proper subset of \mathbb{T} . Then the peripheral spectrum $\sigma(T) \cap \mathbb{T}$ is a finite union of finite subgroups of \mathbb{T} . In particular, the peripheral spectrum contains at most finitely many points.*

Proof. This follows from the cyclicity mentioned before the lemma and the density of $\{\alpha^n : n \in \mathbb{Z}\}$ in \mathbb{T} for each $\alpha \in \mathbb{T}$ that is not a root of unity [EFHN15, Theorem 2.36 and Example 2.37]. \square

We further require a version of Gelfand’s identity operator theorem. We first present its classical version. A bounded operator $T: E \rightarrow E$ on a Banach space E is called *doubly power-bounded* if T is invertible and $\sup_{k \in \mathbb{Z}} \|T^k\| < \infty$.

Proposition 6.1.5 (Gelfand’s $T = I$ theorem). *Let $T: E \rightarrow E$ be an invertible bounded linear operator on a Banach space E . If T is doubly power-bounded and $\sigma(T) = \{1\}$, then $T = I$.*

For the original proof of the result, we refer to [Gel41]. A more modern proof using functional calculus is given in [EN00, Theorem B.17]. A similar version of this result for a priori uniformly ergodic operators on Banach spaces is given in [LSS15, Lemma 2.1].

For lattice homomorphisms one can show the following special version of Proposition 6.1.5, which does not require the additional assumption of T being doubly power-bounded.

Proposition 6.1.6 ($T = I$ theorem for Lamperti operators). *Let $T: E \rightarrow E$ be a lattice homomorphism on the Banach lattice E . If $\sigma(T) = \{1\}$, then $T = I$.*

This proposition has been proved in [SWA78, Corollary 2.2] for the class of lattice homomorphisms and in [Are83, Corollary 3.6] for a more general class called Lamperti operators.

We are now positioned to prove our main theorem.

Proof of Theorem 6.1.2. It follows from the cyclicity of the spectrum, discussed before Lemma 6.1.4, and from the lemma itself that every point in the peripheral spectrum $\sigma(T) \cap \mathbb{T}$ is isolated in $\sigma(T)$.

Hence, there exists a constant $\rho \in (0, 1)$ such that every $\lambda \in \sigma(T) \cap \mathbb{D}$ satisfies $|\lambda| < \rho$. In other words, the spectral points outside of the peripheral spectrum can be bounded uniformly away from \mathbb{T} . In particular, we have $\sigma(T) \cap \rho \cdot \mathbb{T} = \emptyset$, and hence, we may define a bounded operator $P: E \rightarrow E$ by

$$P := \frac{1}{2\pi i} \oint_{|\lambda|=\rho} R(\lambda, T) d\lambda,$$

where $R(\cdot, T) = (\cdot - T)^{-1}$ denotes the resolvent of T and the path integral is integrated counterclockwise. This operator the spectral projection corresponding to the closed subset $\{\lambda \in \sigma(T) : |\lambda| < \rho\}$ of the spectrum of T (see [DS58, Section VII.3]) and decomposes the space E into two closed subspaces as

$$E = PE \oplus (I - P)E =: I_{\text{stab}} \oplus E_{\text{per}}. \quad (6.2)$$

Moreover, the image $I_{\text{stab}} = PE$ is a closed and T -invariant linear subspace of E with $r(T|_{I_{\text{stab}}}) < \rho < 1$ and $\sigma(T|_{E_{\text{per}}}) = \sigma(T) \cap \mathbb{T}$ (see [DS58, Theorem VII.3.20]).

Fix $\mu \in (r(T|_{I_{\text{stab}}}), \rho)$. To show that I_{stab} is an ideal we prove the equality

$$I_{\text{stab}} = \{x \in E : \lim_{n \rightarrow \infty} \mu^{-n} \|T^n x\| = 0\} =: I.$$

“ $I_{\text{stab}} \subseteq I$ ”: The inclusion holds, since we have

$$\lim_{n \rightarrow \infty} \mu^{-n} \|(T|_{I_{\text{stab}}})^n\| = 0$$

by the choice of μ .

“ $I \subseteq I_{\text{stab}}$ ”: Let $x \in I$. Then there exists a constant $C > 0$ such that $\|T^n x\| \leq C\mu^n$ for all $n \in \mathbb{N}$. Hence, the Neumann series

$$R(\lambda, T)x = \sum_{n=0}^{\infty} \lambda^{-(n+1)} T^n x$$

converges absolutely for all $\lambda \in \mathbb{C}$ with $|\lambda| > \mu$. Since $\mu < \rho < 1$, we obtain from Cauchy’s integral formula that

$$\begin{aligned} Px &= \frac{1}{2\pi i} \oint_{|\lambda|=\rho} R(\lambda, T)x d\lambda \\ &= \sum_{n=0}^{\infty} \frac{T^n x}{n!} \left(\frac{n!}{2\pi i} \oint_{|\lambda|=\rho} \frac{1}{\lambda^{n+1}} d\lambda \right) = x. \end{aligned}$$

Hence, $x = Px \in I_{\text{stab}}$ follows.

Observe that I is a lattice ideal, since T is a lattice homomorphism. As $I = I_{\text{stab}}$ is closed, it follows that the quotient space E/I_{stab} is a Banach lattice [Sch74, Proposition II.11.4 on p. 137]. Since I_{stab} is T -invariant we can define the induced operator of T on E/I_{stab} by

$$T_I: E/I_{\text{stab}} \rightarrow E/I_{\text{stab}}, \quad x + I_{\text{stab}} \mapsto Tx + I_{\text{stab}}.$$

Notice that T_γ is a lattice homomorphism on E/I_{stab} . From (6.2) we obtain that the mapping

$$S: E/I_{\text{stab}} \rightarrow E_{\text{per}}, \quad x + I_{\text{stab}} \mapsto (I - P)x$$

is an isomorphism of Banach spaces¹. Moreover, it follows that $T_\gamma = S^{-1}TS$, so T_γ is conjugate to $T|_{E_{\text{per}}}$. Thus, $\sigma(T_\gamma) = \sigma(T|_{E_{\text{per}}}) = \sigma(T) \cap \mathbb{T}$ and it follows that 1 is an isolated point in $\sigma(T_\gamma)$.

By Lemma 6.1.4, the peripheral spectrum of T consists of a finite union of roots of unity, and thus, there exists an $N \in \mathbb{N}$ such that $\sigma((T_\gamma)^N) = \{1\}$. From Proposition 6.1.6 we obtain that $(T_\gamma)^N$ is the identity operator on E/I_{stab} . From the conjugacy it follows that T^N acts as the identity on E_{per} . Thus, E_{per} is contained in the fixed space $\text{fix}(T^N)$. Since the powers of T converge strongly to 0 on I_{stab} , it follows from the decomposition (6.2) that we obtain the equality $E_{\text{per}} = \text{fix}(T^N)$. As fixed spaces of lattice homomorphisms are sublattices, it follows that E_{per} is a sublattice of E . Moreover, since $(T|_{E_{\text{per}}})^N$ is the identity on E_{per} , the operator $T|_{E_{\text{per}}}$ is periodic.

The decomposition (6.1) now follows from a standard result from linear algebra, see, e.g., [Glü17, Lemma 2.2.3], applied to the polynomial $p(X) = X^N - 1$. \square

Remark 6.1.7. (i) Note that the assumption $r(T) = 1$ automatically implies that $1 \in \sigma(T)$. This follows easily from the cyclicity of the spectrum discussed before Lemma 6.1.4 and is even true for the more general class of positive operators on Banach lattices [Mey91, Proposition 4.1.1 i)].

(ii) The fact that I_{stab} is a closed ideal also follows directly from [Are83, Theorem 4.1] and is proved there for the more general class of Lamperti operators.

Notice that, although I_{stab} is a closed ideal, it does not need to be a band, as the following example shows.

Example 6.1.8. There exists a Banach lattice E and a Banach lattice homomorphism $T: E \rightarrow E$ with $r(T) = 1$ that has 1 as an isolated point in its spectrum but the closed ideal I_{stab} is not a band.

Consider the Banach lattice $E := c(\mathbb{N})$ of all convergent sequences in \mathbb{C} endowed with the supremum norm. Define the operator

$$T: E \rightarrow E, \quad a = (a_n)_{n \in \mathbb{N}} \mapsto \lim_{n \rightarrow \infty} a_n \cdot \mathbb{1},$$

where $\mathbb{1} = (1)_{n \in \mathbb{N}}$. Since T is a projection it follows that $\sigma(T) = \{0, 1\}$ and that the conditions of Theorem 6.1.2 are satisfied. Since $E_{\text{per}} = \text{fix}(T) = \text{span}\{\mathbb{1}\}$, we have $I_{\text{stab}} = c_0(\mathbb{N})$, which is not a band in $c(\mathbb{N})$.

Let us state four corollaries of Theorem 6.1.2. In case T is isometric, we arrive at the following conclusion, which is already known from [DNP87, Proposition W.12].

¹At this stage of the proof we do not know whether E_{per} is a Banach lattice, so this isomorphism can only be one of Banach spaces.

Corollary 6.1.9. *Let E be a complex Banach lattice and $T: E \rightarrow E$ be an isometric lattice homomorphism with $r(T) = 1$ that has 1 as an isolated point in its spectrum. Then T is periodic.*

Proof. It follows from the isometry of T that $I_{\text{stab}} = \{0\}$. □

The second corollary connects the assumption that 1 is isolated in $\sigma(T)$ to two a priori distinct notions on the long-term behavior of $(T^n)_{n \in \mathbb{N}}$. We call a linear operator $T: E \rightarrow E$ *uniformly almost periodic* if there exists a periodic operator $S: E \rightarrow E$ such that

$$\lim_{n \rightarrow \infty} \|T^n - S^n\| = 0.$$

Corollary 6.1.10. *Let E be a complex Banach lattice and $T: E \rightarrow E$ be a lattice homomorphism with $r(T) = 1$. Then the following statements are equivalent.*

- (a) *The point 1 is isolated in the spectrum of T .*
- (b) *The operator T is uniformly almost periodic.*
- (c) *The operator T is uniformly ergodic.*

The uniform ergodicity for the more general classes of positive operators and Markov operators was studied by Michael Lin. We refer to [Lin78] and [Lin75], respectively.

Proof of Corollary 6.1.10. “(a) \Rightarrow (b)”: This is Theorem 6.1.2.

“(b) \Rightarrow (c)”: Let $N \in \mathbb{N}$ be the period of S and define $P := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} S^k$, which exists since S is periodic. As noted in the introduction, the operator P is then a projection onto $\text{fix}(S)$. By (b), it follows that $\|A_n[T] - A_n[S]\| \rightarrow 0$, so $\|A_n[T] - P\| \rightarrow 0$.

“(c) \Rightarrow (a)”: If T is uniformly ergodic, then it follows that $\lim_{n \rightarrow \infty} \frac{1}{n} \|T^n\| = 0$, and thus, by Theorem 6.1.1 and the fact that $1 \in \sigma(T)$ (cf. Remark 6.1.7(i)), it follows that 1 is a pole of the resolvent $R(\cdot, T)$, so it is isolated in $\sigma(T)$. □

When we do not require that T has spectral radius 1, we obtain the following characterization of uniform ergodicity for lattice homomorphisms.

Corollary 6.1.11. *Let E be a complex Banach lattice. Then a Banach lattice homomorphism $T: E \rightarrow E$ is uniformly ergodic if and only if it is power-bounded and 1 is either isolated in the spectrum of T or in the resolvent set of T .*

Proof. “ \Rightarrow ”: The uniform ergodicity implies that $\lim_{n \rightarrow \infty} \frac{1}{n} \|T^n\| = 0$, which implies $r(T) \leq 1$. In case $r(T) < 1$, then the power-boundedness follows immediately and 1 is in the resolvent set of T . When $r(T) = 1$, then T is power-bounded and has 1 isolated in the spectrum of T by Corollary 6.1.10 and Theorem 6.1.2.

“ \Leftarrow ”: The power-boundedness implies that $r(T) \leq 1$. When $r(T) < 1$, then the uniform ergodicity is immediate. If $r(T) = 1$, then 1 must be in the spectrum of T by the cyclicity of the spectrum mentioned before Lemma 6.1.4. Thus, by assumption, 1 is isolated in $\sigma(T)$ and Corollary 6.1.10 implies the uniform ergodicity of T . □

We can now give an easy spectral characterization of all periodic lattice homomorphisms. Similar results for invertible positive operators can be found in [Zha92, Theorem 3.1] or [Zha93, Theorem 5.3]. For power-bounded operators on reflexive Banach spaces a similar result can be found in [Lin20, Corollary 2.7].

Corollary 6.1.12. *Let E be a complex Banach lattice and $T: E \rightarrow E$ be a lattice homomorphism. Then T is periodic if and only if its spectrum $\sigma(T)$ is a proper subset of \mathbb{T} .*

Proof. “ \Rightarrow ”: This is straightforward to see.

“ \Leftarrow ”: We may assume that $E \neq \{0\}$. If the spectrum of T is a proper subset of \mathbb{T} , then $r(T) = 1$ follows. Hence, by Lemma 6.1.4 the spectrum is a finite union of finite subgroups of \mathbb{T} , so 1 is isolated in the peripheral spectrum $\sigma(T) \cap \mathbb{T}$. As by assumption the peripheral spectrum coincides with the spectrum, it follows that 1 is isolated in $\sigma(T)$. Thus, by Theorem 6.1.2 it suffices to prove that $I_{\text{stab}} = \{0\}$.

Recall that in the proof of Theorem 6.1.2, the ideal I_{stab} was defined to be the image of the spectral projection P corresponding to the spectral values that are not peripheral. By the assumptions on the spectrum we have $P = 0$ from which the statement follows. \square

6.2 Uniform Ergodicity for Topological Dynamics

We recall that a Banach lattice E is an *AM-space* if it satisfies

$$\|x \vee y\| = \|x\| \vee \|y\|$$

for all $x, y \in E_+$. An element $e \in E_+$ is called a *unit* if for every $x \in E$ there exists a number $\lambda > 0$ such that $|x| \leq \lambda e$. An AM-space E with a unit e can always be equipped with an equivalent norm $\|\cdot\|_e$ satisfying $\|e\|_e = 1$, called the *gauge norm* of e (see [Sch74, Corollary of Proposition II.7.2 on p. 102]). The gauge norm of e is defined by

$$\|x\|_e := \inf\{\lambda > 0 : |x| \leq \lambda e\}$$

for all $x \in E$. Therefore, whenever we work with AM-spaces with a unit e we will tacitly assume that $\|e\| = 1$. Note that the renormed space $(E, \|\cdot\|_e)$ is again an AM-space with unit e .

Important examples of AM-spaces are the spaces of bounded and continuous functions on a topological space. For a topological space X we denote by $C_b(X)$ the space of all bounded complex-valued continuous functions endowed with the supremum norm. It is straightforward to see that $C_b(X)$ is an AM-space with the constant function $\mathbb{1}$ as a unit.

The next technical lemma brings [Küh21, Proposition 2.3 (I)] to the realm of lattice operators on AM-spaces that leave a unit fixed and is proved analogously.

Lemma 6.2.1. *Let E be a complex AM-space with unit e . Let $T: E \rightarrow E$ be a lattice homomorphism with $Te = e$. Let $I \subseteq E$ be a T -invariant and closed lattice ideal such that $\|(T|_I)^n\|$ converges to 0. Then $T|_I$ is nilpotent.*

Proof. By a representation theorem of Kakutani-Bohnenblust-Krein (see [AA02, Theorem 3.6]) we may assume that $E = C(K)$ for some compact Hausdorff space K and T maps the constant function $\mathbb{1}$ to itself.

By [Sch74, Theorem III.9.1] or Proposition 1.2.31 there exists a continuous function $\varphi: K \rightarrow K$ such that $T = T_\varphi$ with

$$T_\varphi: C(K) \rightarrow C(K), \quad f \mapsto f \circ \varphi.$$

Moreover, there exists a closed set $M \subseteq K$ such that

$$I = I_M := \{f \in C(K) : \forall x \in M : f(x) = 0\}$$

by [Sch74, Example 1 on p. 157] or Proposition 1.2.27.

We show that there exists an $n_0 \in \mathbb{N}$ such that $\varphi^{n_0}(K) \subseteq M$, since then for every $f \in I$ and $x \in K$ it follows that $T^{n_0}f(x) = f(\varphi^{n_0}(x)) = 0$. This implies the nilpotency. Assume, for a contradiction, that for every $n \in \mathbb{N}$ there exists $x \in K$ such that $\varphi^n(x) \in K \setminus M$. By Urysohn's lemma there exists a real-valued, non-negative $f \in I$ with $f(\varphi^n(x)) = 1$ and $\|f\| = 1$. Then

$$\|(T|_I)^n\| \geq \|(T|_I)^n f\| \geq f(\varphi^n(x)) = 1.$$

This contradicts the assumptions of the lemma. □

Combining the statement of Lemma 6.2.1 with that of Theorem 6.1.2 we obtain the following immediate corollary for AM-spaces. We will call a map ψ of some set into itself *eventually periodic* if there exist $k, p \in \mathbb{N}$ such that

$$\psi^{k+np} = \psi^k$$

for every $n \in \mathbb{N}$.

Corollary 6.2.2. *Let E be a complex AM-space with unit e . Let $T: E \rightarrow E$ be a Banach lattice homomorphism with $Te = e$. Suppose that 1 is isolated in $\sigma(T)$. Then T is eventually periodic.*

Proof. It follows from the fact that $\|T^n\| = \|T^n e\| = \|e\|$ for all $n \in \mathbb{N}$ (see [AA02, Lemma 3.2]) that $r(T) = 1$. Now the statement follows from Theorem 6.1.2 together with Lemma 6.2.1 applied to $T|_{I_{\text{stab}}}$. □

Notice that an eventually periodic operator is uniformly almost periodic, so by Corollary 6.1.10 the converse implication of Corollary 6.2.2 also holds.

Let us look at a simple example illustrating that the condition $Te = e$ for some unit $e \in E$ in Corollary 6.2.2 can not be dropped.

Example 6.2.3. There exists a complex AM-space E with unit e and a lattice homomorphism $T: E \rightarrow E$ with $r(T) = 1$, which is not eventually periodic.

Consider $E := \mathbb{R}^2$ endowed with the supremum norm and the lattice homomorphism $T: E \rightarrow E$ given by the matrix

$$T = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}.$$

Then $\sigma(T) = \{\frac{1}{2}, 1\}$. Moreover, $I_{\text{stab}} = \{f \in E : f(1) = 0\}$. However, T is not nilpotent on I_{stab} .

For Koopman operators of continuous dynamical systems on topological spaces Corollary 6.2.2 reads as follows. Recall from Definition 1.1.2 (i) that a topological space Ω is called *functionally Hausdorff* if for every pair of distinct points $\omega_1, \omega_2 \in \Omega$ there exists a continuous function $f \in C(\Omega)$ such that $f(\omega_1) \neq f(\omega_2)$. A moment's thought reveals that this condition is equivalent to the following: for every pair of distinct points $\omega_1, \omega_2 \in \Omega$ there exists a continuous function $f: \Omega \rightarrow [0, 1]$ with $f(\omega_1) = 1$ and $f(\omega_2) = 0$.

The following result generalizes [Sch71b, Satz 5] to functionally Hausdorff spaces. Also notice that a full description of the spectrum of lattice homomorphisms on $C(K)$ for compact and Hausdorff K can be found in [Sch71a, Theorem 2.7].

Theorem 6.2.4. *Let Ω be a functionally Hausdorff space and $\varphi: \Omega \rightarrow \Omega$ be continuous and consider the Koopman operator*

$$T_\varphi: C_b(\Omega) \rightarrow C_b(\Omega), \quad f \mapsto f \circ \varphi.$$

Then the following assertions are equivalent.

- (a) *The point 1 is isolated in $\sigma(T_\varphi)$.*
- (b) *The mapping φ is eventually periodic.*
- (c) *The operator T_φ is uniformly ergodic.*

Proof. The implications “(b) \Rightarrow (c)” and “(c) \Rightarrow (a)” immediately follow from Corollary 6.1.10.

“(a) \Rightarrow (b)”:
It follows from Corollary 6.2.2 that T_φ is eventually periodic. Let $p, k \in \mathbb{N}$ be such that $T_\varphi^{k+np} = T_\varphi^k$ for all $n \in \mathbb{N}$. Fix $\omega_1 \in \Omega$ and $n \in \mathbb{N}$. We show that $\varphi^{k+np}(\omega_1) = \varphi^k(\omega_1)$. Indeed, let $\omega_2 \in \Omega \setminus \{\varphi^k(\omega_1)\}$. Then, by the functionally Hausdorff property of Ω , there exists a continuous function $f: \Omega \rightarrow [0, 1]$ with $f(\varphi^k(\omega_1)) = 1$ and $f(\omega_2) = 0$. Hence, $f(\varphi^{k+np}(\omega_1)) = T_\varphi^{k+np}f(\omega_1) = T_\varphi^k f(\omega_1) = f(\varphi^k(\omega_1)) = 1$, and thus, $\varphi^{k+np}(\omega_1) \neq \omega_2$. Since ω_2 was arbitrary, it follows that $\varphi^{k+np}(\omega_1) = \varphi^k(\omega_1)$. \square

The next simple example shows that we cannot dismiss the functionally Hausdorff condition in Theorem 6.2.4. Recall that a topological space is called Urysohn if every two distinct points have open neighborhoods with disjoint closures. In particular, a Urysohn space is Hausdorff.

Example 6.2.5. There exists a Urysohn space Ω and a continuous map $\varphi: \Omega \rightarrow \Omega$ such that 1 is isolated in the spectrum of T_φ and φ is not eventually periodic.

There exists a (countable) Urysohn space $\tilde{\Omega}$ on which every continuous and real-valued function is constant [Hew46, Theorem 2]. Hence, it follows that $C_b(\tilde{\Omega}) = C_b(\tilde{\Omega}; \mathbb{R}) \oplus iC_b(\tilde{\Omega}; \mathbb{R})$ only consists of constant functions. Now set $\Omega := \tilde{\Omega}^{\mathbb{N}}$ and endow Ω with the product topology. It is straightforward to see that Ω is also Urysohn.

Moreover, $C_b(\Omega)$ also contains only the constant functions. Indeed, let $\omega = (\omega_n)_{n \in \mathbb{N}}$ and $\hat{\omega} = (\hat{\omega}_n)_{n \in \mathbb{N}}$ be elements in Ω and for each $k \in \mathbb{N}$ set $\omega^{(k)} = (\omega_n^{(k)})_{n \in \mathbb{N}} \in \Omega$ to be

$$\omega_n^{(k)} := \begin{cases} \hat{\omega}_n, & n \geq k, \\ \omega_n, & n < k. \end{cases}$$

Then for each $n \in \mathbb{N}$ we have $\omega_n^{(k)} \rightarrow \omega_n$ as $k \rightarrow \infty$, and thus, $\omega^{(k)} \rightarrow \omega$ in the product topology. Moreover, if $f \in C_b(\Omega)$ we have $f(\omega^{(k)}) = f(\hat{\omega})$ for all $k \in \mathbb{N}$, since the function f is constant in each single coordinate, and $\omega^{(k)}$ and $\hat{\omega}$ agree in all but finitely many coordinates. By the continuity of f it follows that $f(\omega) = \lim_{k \rightarrow \infty} f(\omega^{(k)}) = f(\hat{\omega})$. Since ω and $\hat{\omega}$ were arbitrary, we conclude that f is constant.

Now consider the left shift

$$\varphi: \Omega \rightarrow \Omega, \quad (\omega_n)_{n \in \mathbb{N}} \mapsto (\omega_{n+1})_{n \in \mathbb{N}}$$

which is continuous and not eventually periodic. However, since $C_b(\Omega)$ only contains constant functions, T_φ must be the identity operator.

6.3 Uniform Ergodicity of Measure-Preserving Dynamics

For Koopman operators of measure-preserving dynamical systems on L^p -spaces for $p \in [1, \infty)$ a similar result to Theorem 6.2.4 holds true, albeit one of its assertions is slightly stronger.

Theorem 6.3.1. *Let $\varphi: X \rightarrow X$ be a measure-preserving mapping on the probability space $X = (X, \Sigma, \mu)$ and consider for $p \in [1, \infty)$ the Koopman operator*

$$T_\varphi: L^p(X) \rightarrow L^p(X), \quad f \mapsto f \circ \varphi.$$

Then the following assertions are equivalent.

- (a) *The point 1 is isolated in $\sigma(T_\varphi)$.*
- (b) *The operator T_φ is periodic.*
- (c) *The operator T_φ is uniformly ergodic.*

Moreover, if the assertions (a) and (c) hold for some $p \in [1, \infty)$, then they hold for every $p \in [1, \infty)$.

Proof. The implications “(b) \Rightarrow (c)” and “(c) \Rightarrow (a)” immediately follow from Corollary 6.1.10.

“(a) \Rightarrow (b)”: Since φ is assumed to be measure-preserving, it follows that T_φ is an isometry. Hence, Corollary 6.1.9 implies that T_φ is periodic.

The last statement of the theorem follows from the observation that assertion (b) is independent of p . \square

Remark 6.3.2. An alternative proof of the implication “(a) \Rightarrow (b)” in Theorem 6.3.1 can be given as follows. By Theorem 6.1.2 there exists $p \in \mathbb{N}$ such that the operator sequence $(T_\varphi^{pn})_{n \in \mathbb{N}}$ converges in the operator norm topology to a projection Q . Since φ is measure-preserving, the constant function $\mathbb{1}$ is a fixed point of the dual operator T_φ' and we obtain $Q' \mathbb{1} = \mathbb{1}$. It follows that Q does not vanish on the set of non-zero positive functions in $L^p(X)$, i.e., $Qf \geq 0$ whenever $0 \leq f \in L^p(X)$. Thus, [GG18, Corollary 5.6] yields that T_φ^p is the identity mapping.

Theorem 6.3.1 admits a more abstract formulation, which holds for a larger class of Banach lattices than the L^p -spaces. Recall that L^p -spaces for $p \in [1, \infty)$ are Banach lattices with an order continuous norm (see Proposition 1.2.11 and Example (i) 1.2.12). Moreover, the L^p -norm is strictly monotone in the sense that it satisfies $\|f\|_p < \|g\|_p$ for every two $0 \leq f, g \in L^p(X)$ with $f \leq g$.

A function $\psi: A \rightarrow \mathbb{R}$ on a subset $A \subseteq E$ is called *strictly monotone* if $x \leq y$ implies $\psi(x) < \psi(y)$ for all $x, y \in A$. Moreover, an element $h \in E_+$ is called *quasi-interior* if the principal ideal I_h , defined by

$$I_h := \{x \in E : \exists C \geq 0 : |x| \leq Ch\},$$

is dense in E with respect to the norm topology.

If $X = (X, \Sigma, \mu)$ is a σ -finite measure space and $h \in L^p(X)$ with $p \in [1, \infty)$, then h is quasi-interior if and only if $h(x) > 0$ for μ -almost every $x \in X$ (see, e.g., [Sch74, Example 1 in Section II.6 on p. 98]).

Proposition 6.3.3. *Let E be a complex Banach lattice with order continuous norm and let $T: E \rightarrow E$ be a Banach lattice homomorphism with $r(T) = 1$. Assume there exists a strictly monotone function $\psi: E_+ \rightarrow \mathbb{R}$ such that $\psi(Tx) \leq \psi(x)$ for all $x \in E_+$ and that there exists a quasi-interior point $h \in E_+$ such that $Th = h$. If 1 is isolated in $\sigma(T)$, then T is periodic.*

Proof. By Theorem 6.1.2 the space E decomposes into a T -invariant and closed ideal I_{stab} , on which T^n converges uniformly to 0, and a T -invariant and closed sublattice E_{per} , on which T is periodic. So it suffices to show that $I_{\text{stab}} = \{0\}$.

Since E has order continuous norm the closed ideal I_{stab} is a projection band (see Proposition 1.2.11), and hence, there exists a band projection $P: E \rightarrow E$ onto I_{stab} . Since band projections satisfy $0 \leq P \leq I$, we have $Ph \leq h$, and thus, $TPh \leq Th \leq h$. As $TPh \in I_{\text{stab}}$, this implies $TPh = PTPh \leq Ph$, so Ph is a so-called *suffix point* of T .

Set $g := h - Ph \geq 0$. Then $0 \leq Ph - TPh = Tg - g$ shows that $Tg \geq g$. If $Tg \neq g$, then $\psi(Tg) > \psi(g)$, which contradicts the assumptions. Hence, we obtain that $Tg = g$, and thus, $TPh = Ph$. Therefore, $Ph \in I_{\text{stab}} \cap E_{\text{per}}$, which implies that $Ph = 0$. Since h is quasi-interior it follows that $P = 0$. This shows that $I_{\text{stab}} = \{0\}$, as claimed. \square

Remark 6.3.4. (i) In Proposition 6.3.3 we may replace the boundedness of T with respect to the strictly monotone function ψ by the following weaker condition: there exists a strictly monotone function $\psi : E_+ \rightarrow \mathbb{R}$ and an integer $n \in \mathbb{N}$ such that $\psi(T^n x) \leq \psi(x)$ for all $x \in E_+$.

Assuming this weaker condition, we can infer from Proposition 6.3.3 that T^n is periodic, which then implies that T must have been periodic in the first place.

(ii) The assumption $Th = h$ is necessary for the statement of Proposition 6.3.3 to be correct. Notice that Example 6.2.3, when E is endowed with the L^1 -norm, yields an lattice homomorphism $T : E \rightarrow E$ with $Th \leq h$ for every quasi-interior point $h \in E_+$, $\sigma(T) = \{\frac{1}{2}, 1\}$, $\|T\| \leq 1$. However, the assertion of Proposition 6.3.3 does not hold true.

The next two examples show that the existence of a strictly monotone function $\psi : E_+ \rightarrow \mathbb{R}$ with $\psi(Tx) \leq \psi(x)$ cannot be dropped in Proposition 6.3.3. First we use the simple matrix example from [Are83, Example after Theorem 4.1]. There the same example is used to demonstrate that the spectral projection P onto I_{stab} need not be positive.

Example 6.3.5. There exists a measure space X , a lattice homomorphism $T : L^1(X) \rightarrow L^1(X)$ with $Th = h$ and $r(T) = 1$ that has 1 isolated in $\sigma(T)$, satisfies $\|T\| > 1$ and is not periodic.

Indeed, consider the space $X = (X, \Sigma, \mu)$ with $X = \{0, 1\}$, $\Sigma = 2^X$ and μ being the counting measure, the constant map $\varphi = 0$. Notice that φ is not measure preserving, so Theorem 6.3.1 is not applicable. We consider the composition operator

$$T_\varphi : L^1(X) \rightarrow L^1(X), \quad f \mapsto f \circ \varphi.$$

Identifying $L^1(X) \cong (\mathbb{R}^2, \|\cdot\|_1)$ we can replace T_φ with a matrix $\hat{T}_\varphi \in \mathbb{R}^{2 \times 2}$ given by

$$\hat{T}_\varphi = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$

is a lattice homomorphism with $\sigma(T) = \{0, 1\}$. Moreover, notice that $\|T\| = 2$ and $h = (1, 1)$ is a quasi-interior point with $Th = h$. However, T is not periodic.

Notice that, in the above example, $T^n = T$ for every $n > 1$, so T is eventually periodic. This is an artifact of its finite-dimensionality. It is not true in general, as the next example demonstrates.

Example 6.3.6. There exists a measure space $X = (X, \Sigma, \mu)$, a lattice homomorphism $T : L^1(X) \rightarrow L^1(X)$ with $Th = h$ and $r(T) = 1$ for some quasi-interior point $h \in L(X)_+$ that has 1 isolated in $\sigma(T)$ and is not (eventually) periodic.

Let $X = [0, 2]$, Σ the Borel- σ -algebra on X and let μ be the Lebesgue measure on X . Consider the map

$$\varphi : X \rightarrow X, \quad x \mapsto \begin{cases} 2x, & x \leq 1, \\ x, & x \geq 1. \end{cases}$$

Notice that φ is not measure preserving, so Theorem 6.3.1 is not applicable.

The Banach lattice $E := L^1(X)$ has order continuous norm, contains the constant function $\mathbb{1}$ as a quasi-interior point and the map

$$T: E \rightarrow E, \quad f \mapsto f \circ \varphi.$$

is a lattice homomorphism with $T\mathbb{1} = \mathbb{1}$.

Consider the band

$$I := \{f \cdot \mathbb{1}_{[0,1]} : f \in L^1(X)\}$$

and notice that I is T -invariant, since for $f \in I$ we have $Tf(x) = 0$ for almost all $x \in [\frac{1}{2}, 2]$, so $T^n \mathbb{1}_{[0,1]} = \mathbb{1}_{[0,1/2^n]}$ for all $n \in \mathbb{N}$. Moreover, for every $f \in I$ we obtain

$$\|Tf\| = \int_0^1 |Tf(x)| \, dx = \int_0^1 |f(2x)| \, dx = \frac{1}{2} \int_0^2 |f(x)| \, dx = \frac{1}{2} \|f\|.$$

Hence, $r(T|_I) \leq \frac{1}{2}$.

Since I is closed and T -invariant, we can consider the quotient operator

$$T_I: E/I \rightarrow E/I, \quad f + I \mapsto Tf + I.$$

Then it is straightforward to see that T_I acts as the identity operator on E/I . Hence, $\sigma(T_I) = \{1\}$.

It is a standard result in spectral theory (see, e.g., [EN00, Proposition VI.2.15]) that $\sigma(T) \subseteq \sigma(T|_I) \cup \sigma(T_I)$. Hence, $r(T) = 1$ and 1 is isolated in $\sigma(T)$. In particular, Theorem 6.1.2 is applicable. It is straightforward to show that $I_{\text{stab}} = I$ and $E_{\text{per}} \cong E/I$.

Notice that $T^n \mathbb{1}_{[1,2]} = \mathbb{1}_{[1/2^n, 2]}$ for all $n \in \mathbb{N}$. This shows that T is not (eventually) periodic.

Observe that this does not contradict Proposition 6.3.3. The indicator function $f = \mathbb{1}_{[1,2]}$ also yields an example of a function satisfying $\psi(T^n f) > \psi(f)$ for every strictly monotone function $\psi: E_+ \rightarrow \mathbb{R}$ and every $n \in \mathbb{N}$.

6.4 Quasi-Compactness and Compactness

Unsurprisingly, there are conditions on operators that are even stronger than uniform ergodicity. To describe those, we need notions that are common in operator theory and especially in the theory of composition operators: compactness and the weaker quasi-compactness. Both have implications on how well an operator can be approximated.

Definition 6.4.1 (compact and quasi-compact operators). Let $T: E \rightarrow E$ be a bounded operator on a Banach space E .

- (i) We call T *compact* if $TB_{\leq 1}(0)$ is relatively compact in E , where $B_{\leq 1}(0)$ denotes the closed unit ball in E .
- (ii) We say that T is *quasi-compact* if there exists a compact operator $K: E \rightarrow E$ and an $n \in \mathbb{N}$ such that $\|T^n - K\| < 1$.

We denote by $\mathcal{K}(E)$ the set of all compact operators on a Banach space E . It is well-known that $\mathcal{K}(E)$ is a closed two-sided ideal in the Banach algebra $\mathcal{L}(E)$ of all bounded linear operators on E endowed with the operator norm. The *Calkin algebra* is defined as the quotient algebra $\mathcal{L}(E)/\mathcal{K}(E)$. The *essential norm* of an operator $T \in \mathcal{L}(E)$ is now defined to be the quotient norm of the equivalence class of T in $\mathcal{L}(E)/\mathcal{K}(E)$ and is given by

$$\|T\|_{\text{ess}} := \inf_{K \in \mathcal{K}(E)} \|T - K\|.$$

The essential norm is a seminorm on $\mathcal{L}(E)$ that vanishes precisely on the compact operators, so it defines a norm on the Calkin algebra. It follows from a result by Atkinson [Mur90, Theorem 1.4.16 on p. 28] that the spectrum of the equivalence class of T in the Calkin algebra coincides with the *essential spectrum*

$$\sigma_{\text{ess}}(T) := \{\lambda \in \mathbb{C} : \lambda I - T \text{ is not a Fredholm operator}\}.$$

We refer to [AA02, Definition 7.36 on p. 299] for a definition of Fredholm operators. We call $r_{\text{ess}}(T) := \sup\{|\lambda| : \lambda \in \sigma_{\text{ess}}(T)\}$ the *essential spectral radius* of T . It now follows from the result by Atkinson that

$$r_{\text{ess}}(T) = \lim_{n \rightarrow \infty} \|T^n\|_{\text{ess}}^{1/n}.$$

So the essential spectral radius coincides with the spectral radius of the equivalence class of T in the Calkin algebra.

Remark 6.4.2 (Disambiguation). It is simple to see that every compact operator is quasi-compact. Other definitions of quasi-compactness exist in the literature that do not have this property. Sometimes an operator $T: E \rightarrow E$ is said to be quasi-compact if

$$r_{\text{ess}}(T) < r(T),$$

where $r_{\text{ess}}(T)$ denotes the *essential spectral radius* of T . Note that the trivial operator $T = 0$ is not quasi-compact according to this definition, as $r_{\text{ess}}(T) = r(T) = 0$ in this case. However, it is compact. If we assume that $r(T) = 1$, then both definitions coincide, as the next lemma shows.

We note one important observation for the long-term behavior of quasi-compact operators.

Lemma 6.4.3. *Let $T: E \rightarrow E$ be a bounded operator on a Banach space E . Then the following assertions are equivalent.*

- (a) *The operator T is quasi-compact.*
- (b) *There exists a sequence $(K_n)_{n \in \mathbb{N}}$ of compact operators $K_n: E \rightarrow E$ such that $\lim_{n \rightarrow \infty} \|T^n - K_n\| = 0$.*
- (c) *The essential spectral radius of T satisfies $r_{\text{ess}}(T) < 1$.*

Proof. “(b) \Rightarrow (a)”: This is trivial.

“(a) \Rightarrow (c)”: Let $K: E \rightarrow E$ and $m \in \mathbb{N}$ such that $\|T^m - K\| < 1$. Then $\|\hat{T}^n\|_{\text{ess}} < 1$, where \hat{T} denotes the equivalence class of T in the Calkin algebra, and thus, $r_{\text{ess}}(T) < 1$.

“(c) \Rightarrow (b)”: Suppose that $r_{\text{ess}}(T) < 1$. Then, writing \hat{T} for the equivalence class of T in the Calkin algebra, we have $\|\hat{T}^n\|_{\text{ess}} \rightarrow 0$ as $n \rightarrow \infty$. Hence, the existence of a sequence $(K_n)_{n \in \mathbb{N}}$ of compact operators such that $\lim_{n \rightarrow \infty} \|T^n - K_n\| = 0$ follows directly. \square

Let us see in which sense compact and quasi-compact operators can be approximated by finite-rank operators.

For quasi-compact operators the spectral theorem tells us that in the statement of Lemma 6.4.3 (b) we may choose the compact operators K_n such that they have finite rank on the spectral subspace associated to the part of the spectrum outside the unit disk. So it follows that the iterates of a quasi-compact operator can be approximated by finite-rank operators in the long-term.

For compact operators more can be said, if the underlying Banach space has the approximation property. We say that a Banach space E has the *approximation property* if for every compact set $K \subseteq E$ and every $\varepsilon > 0$ there exists a finite-rank operator $S: E \rightarrow E$ such that $\|Sx - x\| \leq \varepsilon$, and E has the stronger *metric approximation property* if we can additionally assume that $\|S\| \leq 1$.

Note that a Banach space E has the approximation property if and only if for every Banach space F and any compact operator $K: F \rightarrow E$ there exists a sequence $(S_k)_{k \in \mathbb{N}}$ of finite-rank operators $S_k: F \rightarrow E$ such that $\lim_{k \rightarrow \infty} \|S_k - K\| = 0$ in operator norm [Jar81, Theorem 18.3.1 on p. 403]. If E has the metric approximation property, then we may choose the sequence $(S_k)_{k \in \mathbb{N}}$ such that $\|S_k\| \leq \|K\|$ for all $k \in \mathbb{N}$.

Hilbert spaces, L^p -spaces with $p \in [1, \infty]$ over arbitrary measure spaces and $C_b(\Omega)$ for completely regular Hausdorff spaces Ω all have the metric approximation property by [Jar81, Section 18.5].

As the space of compact operators is closed in the operator norm topology, it follows only compact operators can be approximated by finite-rank operators.

We will now use Theorem 6.1.2 in order to characterize quasi-compact lattice homomorphisms with spectral radius 1. For a lattice homomorphism T with $r(T) = 1$ and 1 isolated in $\sigma(T)$, recall from Theorem 6.1.2 the decomposition

$$E = I_{\text{stab}} \oplus E_{\text{per}}$$

into T -invariant subspaces on which T is uniformly stable and periodic, respectively. Note that in the next proposition we neither require the power-boundedness of T nor do we assume that $1 \in \sigma(T)$ as both conditions are automatic.

Proposition 6.4.4. *Let $T: E \rightarrow E$ be a lattice homomorphism on the Banach lattice E with $r(T) = 1$. Suppose that T is quasi-compact. Then 1 is isolated in the spectrum of T and E_{per} is finite-dimensional.*

Proof. Since T is a lattice homomorphism, it has cyclic spectrum by [Sch74, Theorem V.4.4 on p. 325]. As $r(T) = 1$ there exists $\lambda \in \sigma(T)$ with $|\lambda| = 1$. As then

$\{\lambda^k : k \in \mathbb{Z}\} \subseteq \sigma(T)$ is a subset of the spectrum of T and the spectrum is closed it follows that $1 \in \sigma(T)$.

Since Lemma 6.4.3 implies $r_{\text{ess}}(T) < 1$, it follows from [AA02, Lemma 7.43 on p. 300] that 1 is an isolated point in $\sigma(T)$. So Theorem 6.1.2 is applicable and yields the decomposition $E = I_{\text{stab}} \oplus E_{\text{per}}$. Let P be the projection onto E_{per} along I_{stab} . By periodicity, there exists a $p \in \mathbb{N}$ such that $(T|_{E_{\text{per}}})^p = I$. Since T is quasi-compact, there exists a sequence $(K_n)_{n \in \mathbb{N}}$ of compact operators such that $\lim_{n \rightarrow \infty} \|T^n - K_n\| = 0$ by Lemma 6.4.3. Thus,

$$\|P - K_{np}P\| = \|T^{np}P - K_{np}P\| \leq \|T^{np} - K_{np}\| \rightarrow 0.$$

Since each $K_{np}P$ is a compact operator for each $n \in \mathbb{N}$ and $\mathcal{K}(E)$ is closed in the operator norm topology, it follows that P is compact as well. Hence, $E_{\text{per}} = PE$ is finite-dimensional. \square

Let us note two simple corollaries.

Corollary 6.4.5. *Let $T: E \rightarrow E$ be a lattice homomorphism on the Banach lattice E with $r(T) = 1$. If T is quasi-compact, then it is uniformly ergodic.*

Proof. This follows easily from Proposition 6.4.4, which guarantees that 1 is isolated in the spectrum of T , and Theorem 6.1.2. \square

Note that this result is known for quasi-compact operators $T: E \rightarrow E$ on a Banach space E with $\lim_{n \rightarrow \infty} \frac{1}{n} \|T^n\| = 0$ (see [DNP87, Theorem W.9]).

We call an operator $T: E \rightarrow E$ *eventually finite-rank* if there exists $n \in \mathbb{N}$ such that T^n is a finite-rank operator.

Corollary 6.4.6. *Let $T: E \rightarrow E$ be a lattice homomorphism on the Banach lattice E with $r(T) = 1$. Suppose further that T is quasi-compact.*

- (i) *If E is a unital AM-space with unit $e \in E$ and $Te = e$, then T is eventually finite-rank.*
- (ii) *If T is an isometry, then E is finite-dimensional.*

Proof. “(i)”: This follows from the observation $T|_{I_{\text{stab}}}$ is nilpotent by Lemma 6.2.1 and the fact that E_{per} is finite-dimensional by the previous Proposition 6.4.4.

“(ii)”: Since T is an isometry, it follows from Corollary 6.1.9 that T is periodic, and thus, $E = E_{\text{per}}$. The claim now follows from Proposition 6.4.4. \square

Remark 6.4.7. Note that it follows from [Kre85, Theorem 2.8] that every quasi-compact isometry $T: E \rightarrow E$ on a Banach space E is actually periodic and that E is finite dimensional, so the conclusion of Corollary 6.4.6 (ii) also holds without the assumption that T is a lattice homomorphism.

For Koopman operators of measure-preserving dynamical systems on L^p -spaces for $p \in [1, \infty)$ we thus obtain the following characterization.

Proposition 6.4.8. *Let $\varphi: X \rightarrow X$ be a measure-preserving mapping on the probability space $X = (X, \Sigma, \mu)$ and consider for $p \in [1, \infty)$ its Koopman operator T_φ on $L^p(X)$. Then the following assertions are equivalent.*

- (a) *The operator T_φ is quasi-compact.*
- (b) *The operator T_φ is compact.*
- (c) *The space $L^p(X)$ is finite-dimensional.*

Proof. “(a) \Rightarrow (c)”: This follows from Corollary 6.4.6 (ii), since T_φ is an isometry.

“(c) \Rightarrow (b)”: If $L^p(X)$ is finite-dimensional, then every operator on $L^p(X)$ is compact.

“(b) \Rightarrow (a)”: Every compact operator is quasi-compact. \square

We may conclude that quasi-compact or compact Koopman operators on L^p -spaces for $p \in [1, \infty)$ are of the following form: the underlying measure space X is purely atomic with only finitely many atoms, and the mapping φ permutes those atoms.

For Koopman operators on $C_b(\Omega)$ for a completely regular Hausdorff space Ω a similar result holds.

Proposition 6.4.9. *Let $\varphi: \Omega \rightarrow \Omega$ be a continuous mapping on the normal Hausdorff space Ω and consider its Koopman operator T_φ on $C_b(\Omega)$. Then the following assertions are equivalent.*

- (a) *The operator T_φ is quasi-compact.*
- (b) *There exists an $n \in \mathbb{N}$ such that $\varphi^n(\Omega)$ is finite.*
- (c) *The operator T_φ is eventually finite-rank.*

Proof. “(a) \Rightarrow (c)”: This follows from Corollary 6.4.6 (i), since $C_b(\Omega)$ is a unital AM-space with unit $\mathbb{1}$ and $T_\varphi \mathbb{1} = \mathbb{1}$.

“(c) \Rightarrow (b)”: Suppose that T_φ^n has finite-rank for some $n \in \mathbb{N}$. Then by Tietze’s extension theorem it follows that $C_b(\overline{\varphi^n(\Omega)}) \subseteq \{f \circ \varphi^n : f \in C_b(\Omega)\} = T_\varphi^n C_b(\Omega)$, which is finite-dimensional. Hence, it follows that $\overline{\varphi^n(\Omega)}$ is finite, and thus, $\varphi^n(\Omega)$ is finite as well.

“(b) \Rightarrow (a)”: Since $\varphi^n(\Omega)$ is finite for some $n \in \mathbb{N}$, it follows that $T_\varphi^n C_b(\Omega) = C_b(\varphi^n(\Omega))$ is finite-dimensional. Hence, T_φ^n is a finite-rank operator, so T_φ is quasi-compact. \square

Compact Koopman operators on $C_b(\Omega)$ have a similar description.

Proposition 6.4.10. *Let $\varphi: \Omega \rightarrow \Omega$ be a continuous mapping on the normal Hausdorff space Ω and consider its Koopman operator T_φ on $C_b(\Omega)$. Then the following assertions are equivalent.*

- (a) *The operator T_φ is compact.*
- (b) *The space $\varphi(\Omega)$ is finite.*
- (c) *The operator T_φ is finite-rank.*

Proof. “(a) \Rightarrow (b)”: Suppose that T_φ is compact. We show that $C_b(\overline{\varphi(\Omega)})$ is finite-dimensional. Let $(f_n)_{n \in \mathbb{N}}$ be a sequence in $C_b(\overline{\varphi(\Omega)})$ with $\|f_n\|_\infty \leq 1$ for all $n \in \mathbb{N}$. Then by Tietze’s extension theorem there exists a sequence $(F_n)_{n \in \mathbb{N}}$ in $C_b(\Omega)$ with $\|F_n\|_\infty \leq 1$ and $F_n|_{\varphi(\Omega)} \circ \varphi = f_n$ for all $n \in \mathbb{N}$. Since T_φ is compact, the sequence $(T_\varphi F_n)_{n \in \mathbb{N}}$ has a convergent subsequence $(T_\varphi F_{n_k})_{k \in \mathbb{N}}$ in $C_b(\Omega)$. But $T_\varphi F_{n_k} = F_{n_k} \circ \varphi$ and thus, $(f_{n_k})_{k \in \mathbb{N}}$ is a convergent subsequence in $C_b(\overline{\varphi(\Omega)})$. Hence, the closed unit ball in $C_b(\overline{\varphi(\Omega)})$ is compact, which implies that $C_b(\overline{\varphi(\Omega)})$ is finite-dimensional. This implies that $\overline{\varphi(\Omega)}$ is finite, and thus, $\varphi(\Omega)$ is finite as well.

“(b) \Rightarrow (c)”: If $\varphi(\Omega)$ is finite, then $T_\varphi C_b(\Omega) = C_b(\varphi(\Omega))$ is finite-dimensional. Hence, T_φ is a finite-rank operator.

“(c) \Rightarrow (a)”: Every finite-rank operator is compact. □

The author is currently unaware whether Proposition 6.4.9 and Proposition 6.4.10 remain true for Koopman operators on $C(\Omega)$, where Ω is only completely regular Hausdorff but not normal.

It is now instructive and rather simple to provide counterexamples of uniformly ergodic Koopman operators on $C_b(\Omega)$ that are not quasi-compact and of quasi-compact Koopman operators on $C_b(\Omega)$ that are not compact. Also note that the compactness of Koopman operators on $C_b(\Omega)$ does not imply the finite-dimensionality of $C_b(\Omega)$.

Example 6.4.11. (i) There exists a completely regular Hausdorff space Ω and a continuous mapping $\varphi: \Omega \rightarrow \Omega$ such that the Koopman operator T_φ on $C_b(\Omega)$ is uniformly ergodic, but not quasi-compact.

The identity mapping $\varphi = \text{id}_\Omega$ on any infinite completely regular Hausdorff space Ω provides such an example. Indeed, $T_\varphi = I$ is uniformly ergodic with mean ergodic projection $P = I$, but not quasi-compact by Proposition 6.4.9.

(ii) There exists a completely regular Hausdorff space Ω and a continuous mapping $\varphi: \Omega \rightarrow \Omega$ such that the Koopman operator T_φ on $C_b(\Omega)$ is quasi-compact, but not compact.

Let $n \geq 2$ and set $\Omega = \{0\} \cup (\{1, \dots, n\} \times \mathbb{N})$ endowed with the discrete topology. Consider the continuous mapping

$$\varphi: \Omega \rightarrow \Omega, \quad \omega \mapsto \begin{cases} 0, & \omega = 0, \\ 0, & \omega = (1, m) \text{ for some } m \in \mathbb{N}, \\ (k-1, m), & \omega = (k, m) \in \{2, \dots, n\} \times \mathbb{N}, \end{cases}$$

which is illustrated in Figure 6.1. Then $\varphi^n(\Omega) = \{0\}$ but $\varphi^k(\Omega)$ is infinite for all $k \in \{1, \dots, n-1\}$. Hence, by Proposition 6.4.9 the operator T_φ is quasi-compact, but by Proposition 6.4.10 it is not compact.

(iii) There exists a completely regular Hausdorff space Ω and a continuous mapping $\varphi: \Omega \rightarrow \Omega$ such that the Koopman operator T_φ on $C_b(\Omega)$ is compact, but $C_b(\Omega)$ is infinite-dimensional.

Setting $n = 1$ in the above example yields a space Ω of infinite cardinality such that T_φ is compact, as $\varphi(\Omega) = \{0\}$ is finite.

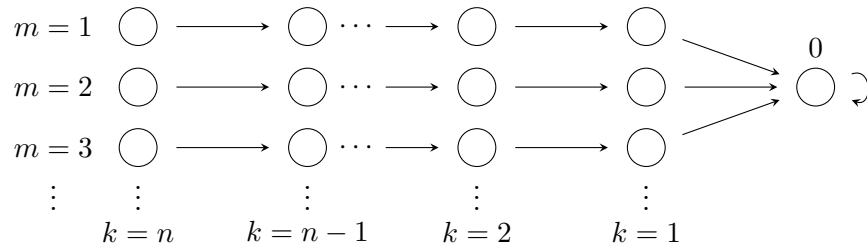


Figure 6.1: The space $\Omega = \{0\} \cup (\{1, \dots, n\} \times \mathbb{N})$ with the mapping φ from Example 6.4.11 (ii).

Remark 6.4.12. Note that Example 6.4.11 (ii) and (iii) rely heavily on the fact that φ is not surjective. In fact, if $\varphi(\Omega)$ is dense in Ω , then T_φ is an isometry by Proposition 1.3.5. Thus, if φ has a dense image, then T_φ is quasi-compact if and only if it is compact if and only if $C_b(\Omega)$ is finite-dimensional.

Note that the proof of Proposition 6.4.10 does not employ Corollary 6.4.6. In fact, one can prove Proposition 6.4.9 without using Proposition 6.4.9 and talking about essential spectral radii. To this end, notice that the equivalence between (a) and (b) in Lemma 6.4.3 is proved without the use of the essential spectral radius. Let us compile an alternative proof.

Alternative proof of Proposition 6.4.9 (a) \Rightarrow (b). By [DNP87, Theorem W.9], it follows that T_φ is uniformly ergodic, as $\lim_{n \rightarrow \infty} \frac{1}{n} \|T_\varphi^n\| = 0$. Thus, it follows from Theorem 6.2.4 that there exists $k_0, p \in \mathbb{N}$ such that $T_\varphi^{k_0+pk} = T_\varphi^{k_0}$ for all $k \in \mathbb{N}_0$. By Lemma 6.4.3 there exists a sequence $(K_n)_{n \in \mathbb{N}}$ such that $\lim_{n \rightarrow \infty} \|T_\varphi^n - K_n\| = 0$. In particular, this shows that $\lim_{k \rightarrow \infty} \|T_\varphi^{k_0} - K_{k_0+pk}\| = 0$. As $\mathcal{K}(E)$ is closed, it follows that $T_{\varphi^{k_0}} = T_\varphi^{k_0}$ is compact. Now Proposition 6.4.10 implies that $\varphi^{k_0}(\Omega)$ is finite. \square

Note that the complexity of the proof vanishes, as we mostly hide it in citing [DNP87, Theorem W.9].

6.5 Notes

Let us review some known results and how the contents of this chapter improve upon them:

In [DNP87, Proposition W.12] it is shown that an isometric lattice homomorphism T is uniformly ergodic if and only if it is periodic. We show that a non-isometric uniformly ergodic lattice homomorphism can be split into a periodic part and one that decays exponentially to 0 (see Theorem 6.1.2 and Corollary 6.1.10).

In [Ion63, Corollary 2] and [Foi64, Corollaire 1] it is shown that the spectrum of the composition operator on L^2 of an invertible and aperiodic measure-preserving system coincides with the complex unit circle. Hence, 1 is not isolated in the spectrum, and thus, the system cannot be uniformly ergodic. This observation was later strengthened to almost everywhere pointwise convergence in [Kre78] for transformations on the unit interval $(0, 1)$ endowed with the Lebesgue measure and composition operators on L^p and in [Pet83, Proposition 3.2.3] to ergodic systems on

non-atomic probability spaces. We show that every measure-preserving dynamical system for which the Cesàro averages have a uniform convergence rate must be periodic (see Theorem 6.3.1), thereby complementing the recent results in [CL23, Appendix].

In [Sch71b, Satz 5] it is proved that topological dynamical systems on compact Hausdorff spaces that are uniformly ergodic have uniformly bounded orbit lengths. In Theorem 6.2.4, we generalize this to non-compact spaces and show that, in this case, the dynamical system is eventually periodic.

The fact that quasi-compactness is a rare property for Koopman operators on L^p and $C_b(\Omega)$ alike does not mean that quasi-compactness is rare for all types of Koopman operators. In fact, if restricted to suitable subspaces of $C_b(\Omega)$, Koopman operators are often quasi-compact. We provide an example.

Example 6.5.1. Let K be a compact metric space with metric d . Let $\varphi: K \rightarrow K$ be a Lipschitz continuous mapping with Lipschitz constant $\text{Lip}(\varphi) < 1$. Denote by $\text{Lip}(K)$ the space of Lipschitz-continuous functions mapping to \mathbb{K} endowed with the norm

$$\|f\|_{\text{Lip}} := \|f\|_{\infty} + \sup_{\substack{x, y \in K \\ x \neq y}} \frac{|f(x) - f(y)|}{d(x, y)}$$

for all $f \in \text{Lip}(K)$. Then the Koopman operator $T_{\varphi}: \text{Lip}(K) \rightarrow \text{Lip}(K)$ defined by $T_{\varphi}f = f \circ \varphi$ for all $f \in \text{Lip}(K)$ is a quasi-compact contraction.

In order to prove the statement of the above example, we will use the following lemma, which is a generalization of a result by Cassius Ionescu Tulcea and Gheorghe Marinescu in [IM50]. It can be found as [HH01, Theorem II.5].

Lemma 6.5.2. *Let E be a Banach space with norm $\|\cdot\|$ and unit ball B . Denote by $T: E \rightarrow E$ a bounded linear operator. Suppose that there exists a semi-norm $\|\cdot\|_*$ on E and constants $k \in \mathbb{N}$, $0 \leq r < r(T)$, $R, M > 0$ such that $\|Tx\|_* \leq M\|x\|_*$ for all $x \in E$, the set TB is relatively compact with respect to the semi-norm $\|\cdot\|_*$ and*

$$\|T^k x\| \leq r^k \|\cdot\|_* + R\|x\| \tag{6.3}$$

holds for all $x \in E$. Then T is quasi-compact.

Note that the definition of quasi-compactness used in [HH01, Theorem II.5] is the one we have touched on in Remark 6.4.2 and coincides with our definition if $r(T) = 1$, which is the case in Example 6.5.1.

Proof of Example 6.5.1. The fact that T_{φ} is a contraction follows from the estimate

$$\begin{aligned} \|T_{\varphi}f\|_{\text{Lip}} &= \|T_{\varphi}f\|_{\infty} + \sup_{\substack{x, y \in K \\ x \neq y}} \frac{|(T_{\varphi}f)(x) - (T_{\varphi}f)(y)|}{d(x, y)} \\ &\leq \|f\|_{\infty} + \text{Lip}(\varphi) \cdot \sup_{\substack{x, y \in K \\ x \neq y}} \frac{|f(x) - f(y)|}{d(x, y)} \leq \|f\|_{\text{Lip}}. \end{aligned} \tag{6.4}$$

In particular, $r(T_{\varphi}) \leq 1$. Since $T_{\varphi} \mathbb{1} = \mathbb{1}$, it follows that $r(T_{\varphi}) = 1$.

We let $E = \text{Lip}(K)$ be endowed with $\|\cdot\|_{\text{Lip}}$ and let $\|\cdot\|_*$ be the supremum norm, which is a norm on $\text{Lip}(K)$. It is clear that $|T_\varphi f| \leq |f|$ for all $f \in \text{Lip}(K)$, so we may choose $M = 1$. Moreover, the set $T_\varphi B$ is relatively compact with respect to $\|\cdot\|_*$ by virtue of Arzelà–Ascoli’s theorem, since K is compact and the functions in $T_\varphi B$ are uniformly bounded and equicontinuous.

Further, it follows from (6.4) that (6.3) holds for $k = 1$, $r = \text{Lip}(\varphi) < 1 = r(T)$ and $R = 1$. Hence, we obtain from Lemma 6.5.2 that T_φ is quasi-compact. \square

Lemma 6.5.2 can be applied to a variety of spaces with greater regularity than $C_b(\Omega)$, such as spaces of Hölder continuous functions or spaces of differentiable functions. By using [Kre85, Theorem 2.8] we may conclude from the quasi-compactness of T_φ that its unimodular spectral points are eigenvalues with a finite-dimensional eigenspace. We also learn that much of the more complicated dynamics of T_φ cannot be observed this way by restricting to spaces of more regular functions.

We can further use an analogous reasoning as in the above results to find a class of sufficiently regular functions on which the Koopman operators is uniformly ergodic (cf. Section 4.1.3).

We should also note that the assumptions of Example 6.5.1 are strong enough to infer from Banach’s fixed point theorem that φ has a unique fixed point $x_0 \in K$ and that $\varphi^n(x) \rightarrow x_0$ for all $x \in K$ as $n \rightarrow \infty$. This implies that $T_\varphi^n f \rightarrow f(x_0) \cdot \mathbb{1}$ for all $f \in \text{Lip}(K)$ as $n \rightarrow \infty$. For more general dynamical systems one often needs to engineer the Banach space E and the semi-norm $\|\cdot\|_*$ to the needs of the given problem in order to apply Lemma 6.5.2. For Perron–Frobenius operators this is done in popular works such as [Bal00] for expanding instead of contracting systems and for hyperbolic systems. We refer to [Dem18] for an overview of different techniques.

Bibliography

- [AA02] Yuriĭ A. Abramovich and Charalambos D. Aliprantis. *An invitation to operator theory*. Vol. 50. Grad. Stud. Math. Providence, RI: American Mathematical Society (AMS), 2002.
- [AB06] Charalambos D. Aliprantis and Kim C. Border. *Infinite dimensional analysis. A hitchhiker's guide*. 3rd ed. Berlin: Springer, 2006.
- [ACOS24] Moacir Aloisio, Silas L. Carvalho, César R. de Oliveira, and Edson Souza. „On spectral measures and convergence rates in von Neumann's ergodic theorem“. *Monatsh. Math.* 203.3 (2024), pp. 543–562.
- [And66] Tsuyoshi Ando. „Contractive projections in L_p -spaces“. *Pac. J. Math.* 17 (1966), pp. 391–405.
- [Are83] Wolfgang Arendt. „Spectral properties of Lamperti operators“. *Indiana Univ. Math. J.* 32 (1983), pp. 199–215.
- [Aus88] Joseph Auslander. *Minimal flows and their extensions*. Vol. 153. North-Holland Math. Stud. Amsterdam etc.: North-Holland, 1988.
- [Bal00] Viviane Baladi. *Positive transfer operators and decay of correlations*. Vol. 16. Adv. Ser. Nonlinear Dyn. Singapore: World Scientific, 2000.
- [Bil95] Patrick Billingsley. *Probability and measure*. 3rd ed. Chichester: John Wiley & Sons Ltd., 1995.
- [Bir67] Garrett Birkhoff. *Lattice theory*. 3rd edition. Vol. 25. Colloq. Publ., Am. Math. Soc. American Mathematical Society (AMS), Providence, RI, 1967.
- [Bir31] George D. Birkhoff. „Proof of the ergodic theorem“. *Proc. Natl. Acad. Sci. USA* 17 (1931), pp. 656–660.
- [BC24] Nicolas Boullé and Matthew J. Colbrook. *Multiplicative Dynamic Mode Decomposition*. 2024. arXiv: 2405.05334.
- [Cha67] Rafael V. Chacon. „A geometric construction of measure preserving transformations“ (1967).
- [Chr16] Ole Christensen. *An introduction to frames and Riesz bases*. 2nd edition. Appl. Numer. Harmon. Anal. Basel: Birkhäuser/Springer, 2016.
- [CL23] Guy Cohen and Michael Lin. *Uniform ergodicity and the one-sided ergodic Hilbert transform*. 2023. arXiv: 2310.15561.

- [Col23a] Matthew J. Colbrook. „The mpEDMD algorithm for data-driven computations of measure-preserving dynamical systems“. *SIAM J. Numer. Anal.* 61.3 (2023), pp. 1585–1608.
- [Col23b] Matthew J. Colbrook. *The Multiverse of Dynamic Mode Decomposition Algorithms*. 2023. arXiv: 2312.00137.
- [CAS23] Matthew J. Colbrook, Lorna J. Ayton, and Máté Szóke. „Residual dynamic mode decomposition: robust and verified Koopmanism“. *J. Fluid Mech.* 955 (2023). Id/No a21, p. 38.
- [CNT87] Alain Connes, Heide Narnhofer, and Walter E. Thirring. „Dynamical entropy of C^* -algebras and von Neumann algebras“. *Commun. Math. Phys.* 112 (1987), pp. 691–719.
- [Con81] John B. Conway. *Subnormal operators*. Vol. 51. Res. Notes Math., San Franc. Pitman Publishing, London, 1981.
- [Con85] John B. Conway. *A course in functional analysis*. Vol. 96. Grad. Texts Math. Springer, Cham, 1985.
- [Coo71] James B. Cooper. „The strict topology and spaces with mixed topologies“. *Proc. Amer. Math. Soc.* 30 (1971), pp. 583–592.
- [Coo87] James B. Cooper. *Saks spaces and applications to functional analysis. 2nd rev. ed.* Vol. 139. North-Holland Math. Stud. Elsevier, Amsterdam, 1987.
- [CFS82] Isaac P. Cornfeld, Sergei V. Fomin, and Yakov G. Sinai. *Ergodic theory*. Vol. 245. Grundlehren Math. Wiss. Springer, Cham, 1982.
- [Dem18] Mark F. Demers. „A gentle introduction to anisotropic Banach spaces“. *Chaos Solitons Fractals* 116 (2018), pp. 29–42.
- [DNP87] Roland Derndinger, Rainer Nagel, and Günther Palm. „Ergodic Theory in the Perspective of Functional Analysis“. Unfinished Book. 1987.
- [DL01] Yves Derriennic and Michael Lin. „Fractional Poisson equations and ergodic theorems for fractional coboundaries“. *Isr. J. Math.* 123 (2001), pp. 93–130.
- [DH24] Alexander Dobrick and Julian Hölz. „Uniform convergence of solutions to stochastic hybrid models of gene regulatory networks“. *J. Evol. Equ.* 24.3 (2024). Id/No 53, p. 17.
- [DHK24] Alexander Dobrick, Julian Hölz, and Markus Kunze. „Ultra Feller operators from a functional-analytic perspective“. *Stud. Math.* 279.3 (2024), pp. 243–271.
- [Dun43] Nelson Dunford. „Spectral theory. I. Convergence to projections“. *Trans. Am. Math. Soc.* 54 (1943), pp. 185–217.
- [DS58] Nelson Dunford and Jacob T. Schwartz. *Linear operators. I. General theory. (With the assistance of William G. Bade and Robert G. Bartle)*. Pure and Applied Mathematics. Vol. 7. New York and London: Interscience Publishers. xiv, 858 p. (1958). 1958.
- [Eis10] Tanja Eisner. *Stability of operators and operator semigroups*. Vol. 209. Oper. Theory: Adv. Appl. Basel: Birkhäuser, 2010.
- [EF25] Tanja Eisner and Bálint Farkas. *A journey through ergodic theorems*. Cham: Birkhäuser, 2025.

- [EFHN15] Tanja Eisner, Bálint Farkas, Markus Haase, and Rainer Nagel. *Operator theoretic aspects of ergodic theory*. Vol. 272. Grad. Texts Math. Cham: Springer, 2015.
- [EN00] Klaus-Jochen Engel and Rainer Nagel. *One-parameter semigroups for linear evolution equations*. Vol. 194. Graduate Texts in Mathematics. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafuno, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt. Springer-Verlag, New York, 2000.
- [Eng89] Ryszard Engelking. *General topology*. Rev. and compl. ed. Vol. 6. Sigma Ser. Pure Math. Berlin: Heldermann Verlag, 1989.
- [FK20] Bálint Farkas and Henrik Kreidler. „Towards a Koopman theory for dynamical systems on completely regular spaces“. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 378.2185 (2020), p. 20190617.
- [Foi64] Ciprian Foiaş. „Sur les mesures spectrales qui interviennent dans la théorie ergodique“. *J. Math. Mech.* 13 (1964), pp. 639–658.
- [Gel41] Isreal Gelfand. „Zur Theorie der Charaktere der Abelschen topologischen Gruppen.“ *Rec. Math. Moscou, n. Ser.* 9 (1941), pp. 49–50.
- [Ger18] Moritz Gerlach. „Convergence of dynamics and the Perron-Frobenius operator“. *Isr. J. Math.* 225.1 (2018), pp. 451–463.
- [GG18] Moritz Gerlach and Jochen Glück. „Lower bounds and the asymptotic behaviour of positive operator semigroups“. *Ergodic Theory Dynam. Systems* 38.8 (2018), pp. 3012–3041.
- [Gla03] Eli Glasner. *Ergodic theory via joinings*. Vol. 101. Math. Surv. Monogr. Providence, RI: American Mathematical Society (AMS), 2003.
- [Glü17] Jochen Glück. „Invariant Sets and Long Time Behaviour of Operator Semigroups“. Universität Ulm, 2017. doi: 10.18725/OPARU-4238.
- [GH23] Jochen Glück and Julian Hölz. „Eventual cone invariance revisited“. *Linear Algebra Appl.* 675 (2023), pp. 274–293.
- [GHT11] Alexander Gomilko, Markus Haase, and Yuri Tomilov. „On rates in mean ergodic theorems“. *Math. Res. Lett.* 18.2 (2011), pp. 201–213.
- [Gra05] Robert M. Gray. „Toeplitz and circulant matrices: a review.“ *Found. Trends Commun. Inf. Theory* 2.3 (2005), p. 93.
- [Hal44] Paul R. Halmos. „In general a measure preserving transformation is mixing“. *Ann. Math. (2)* 45 (1944), pp. 786–792.
- [Hal60] Paul R. Halmos. *Lectures on ergodic theory*. New York: Chelsea Publishing Company, VII, 99 p. (1960). 1960.
- [HK02] Boris Hasselblatt and Anatole Katok. „Principal structures“. In: *Handbook of dynamical systems. Volume 1A*. Amsterdam: North-Holland, 2002, pp. 1–203.
- [HH25] Till Hauser and Julian Hölz. *Entropy based lower dimension bounds for finite-time prediction of Dynamic Mode Decomposition algorithms*. 2025. arXiv: 2504.20269.

- [HH01] Hubert Hennion and Loïc Hervé. *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*. Vol. 1766. Lect. Notes Math. Berlin: Springer, 2001.
- [Hew46] Edwin Hewitt. „On two problems of Urysohn“. *Ann. of Math. (2)* 47 (1946), pp. 503–509.
- [Höl25] Julian Hölz. „A Note on the uniform ergodicity of dynamical systems“. *Israel J. Math.* (2025).
- [Ion63] Alexandra Ionescu Tulcea. „Random series and spectra of measure-preserving transformations“. In: *Ergodic theory. Proceedings of an international symposium, Tulane University, New Orleans, LA, USA, October, 1961*. New York, NY: Academic Press, 1963.
- [IM50] Cassius T. Ionescu Tulcea and Gheorghe I. Marinescu. „Théorie ergodique pour des classes d’opérations non complètement continues“. *Ann. Math. (2)* 52 (1950), pp. 140–147.
- [Jac56] Konrad Jacobs. „Ergodentheorie und fastperiodische Funktionen auf Halbgruppen“. *Math. Z.* 64 (1956), pp. 298–338.
- [JP83] Graham J. O. Jameson and Allan Pinkus. „Positive and minimal projections in function spaces“. *J. Approx. Theory* 37 (1983), pp. 182–195.
- [JK25] Asgar Jamneshan and Henrik Kreidler. „Ergodic Structure Theory and Applications“. ISEm script. 2025.
- [Jar81] Hans Jarchow. *Locally convex spaces*. Mathematische Leitfäden. [Mathematical Textbooks]. B. G. Teubner, Stuttgart, 1981.
- [Kac96] Aleksander G. Kachurovskii. „The rate of convergence in ergodic theorems“. *Russ. Math. Surv.* 51.4 (1996), pp. 653–703.
- [KT06] Anatole Katok and Jean-Paul Thouvenot. „Spectral properties and combinatorial constructions in ergodic theory“. In: *Handbook of dynamical systems. Volume 1B*. Amsterdam: Elsevier, 2006, pp. 649–743.
- [KL05] David Kerr and Hanfeng Li. „Dynamical entropy in Banach spaces“. *Invent. Math.* 162.3 (2005), pp. 649–686.
- [KL16] David Kerr and Hanfeng Li. *Ergodic theory. Independence and dichotomies*. Springer Monogr. Math. Cham: Springer, 2016.
- [KRW15] Ioannis G. Kevrekidis, Clarence W. Rowley, and Matthew O. Williams. „A kernel-based method for data-driven Koopman spectral analysis“. *J. Comput. Dyn.* 2.2 (2015), pp. 247–265.
- [Koo31] Bernard O. Koopman. „Hamiltonian systems and transformations in Hilbert space“. *Proc. Natl. Acad. Sci. USA* 17 (1931), pp. 315–318.
- [KN32] Bernard O. Koopman and John von Neumann. „Dynamical systems of continuous spectra“. *Proc. Natl. Acad. Sci. USA* 18 (1932), pp. 255–263.
- [KM18] Milan Korda and Igor Mezić. „On convergence of extended dynamic mode decomposition to the Koopman operator“. *J. Nonlinear Sci.* 28.2 (2018), pp. 687–710.

- [Köt69] Gottfried Köthe. *Topological vector spaces. I. (Translated by D. J. H. Garling)*. Vol. 159. Grundlehren Math. Wiss. Springer, Cham, 1969.
- [Kre78] Ulrich Krengel. „On the speed of convergence in the ergodic theorem“. *Monatsh. Math.* 86 (1978), pp. 3–6.
- [Kre85] Ulrich Krengel. *Ergodic theorems. With a supplement by Antoine Brunel*. Vol. 6. De Gruyter Stud. Math. De Gruyter, Berlin, 1985.
- [Kri86] Thomas L. Kriete. „An elementary approach to the multiplicity theory of multiplication operators“. *Rocky Mt. J. Math.* 16 (1986), pp. 23–32.
- [Küh21] Viktoria Kühner. „What can Koopmanism do for attractors in dynamical systems?“ *J. Anal.* 29.2 (2021), pp. 449–471.
- [Kun11] Markus Kunze. „A Pettis-type integral and applications to transition semigroups“. *Czechoslovak Math. J.* 61(136).2 (2011), pp. 437–459.
- [Küs21a] Kari V. Küster. „Decompositions of dynamical systems induced by the Koopman operator“. *Anal. Math.* 47.1 (2021), pp. 149–173.
- [Küs21b] Kari V. Küster. „Topological Dynamics via Structured Koopman Subsystems“. PhD thesis. 2021. doi: 10.15496/PUBLIKATION-63134.
- [LG59] Karel de Leeuw and Irving Glicksberg. „Almost periodic compactifications“. *Bull. Am. Math. Soc.* 65 (1959), pp. 134–139.
- [LG61] Karel de Leeuw and Irving Glicksberg. „Applications of almost periodic compactifications“. *Acta Math.* 105 (1961), pp. 63–97.
- [Lem24] Mariusz Lemańczyk. „On some spectral problems in ergodic theory“. In: *A vision for dynamics in the 21st century. The legacy of Anatole Katok. Based on a memorial conference, Będlewo, Poland, August 2019*. Cambridge: Cambridge University Press, 2024, pp. 318–336.
- [Lin75] Michael Lin. „Quasi-compactness and uniform ergodicity of Markov operators“. *Ann. Inst. Henri Poincaré, Nouv. Sér., Sect. B* 11 (1975), pp. 345–354.
- [Lin78] Michael Lin. „Quasi-compactness and uniform ergodicity of positive operators“. *Isr. J. Math.* 29 (1978), pp. 309–311.
- [Lin20] Michael Lin. *Reflexive Banach spaces with all power-bounded operators almost periodic*. 2020. arXiv: 2006.07850.
- [LSS15] Michael Lin, David Shoikhet, and Laurian Suciu. „Remarks on uniform ergodic theorems“. *Acta Sci. Math.* 81.1-2 (2015), pp. 251–283.
- [Lin01] Elon Lindenstrauss. „Pointwise theorems for amenable groups“. *Invent. Math.* 146.2 (2001), pp. 259–295.
- [LW00] Elon Lindenstrauss and Benjamin Weiss. „Mean topological dimension“. *Isr. J. Math.* 115 (2000), pp. 1–24.
- [MOA11] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: theory of majorization and its applications*. 2nd edition. Springer Ser. Stat. New York, NY: Springer, 2011.

- [MMS20] Alexandre Mauroy, Igor Mezić, and Yoshihiko Susuki, eds. *The Koopman operator in systems and control. Concepts, methodologies and applications*. Vol. 484. Lect. Notes Control Inf. Sci. Cham: Springer, 2020.
- [Mey91] Peter Meyer-Nieberg. *Banach lattices*. Universitext. Berlin etc.: Springer-Verlag, 1991.
- [Mor16] Nikita V. Moriakov. „Entropy and Kolmogorov Complexity“. PhD thesis. 2016. doi: 10.4233/UUID:565EB3E7-E0EA-4A88-ABBB-EB0AE2C6C36F.
- [MNR12] Thomas Müller-Gronbach, Erich Novak, and Klaus Ritter. *Monte Carlo Algorithmen*. Springer-Lehrb. Berlin: Springer, 2012.
- [Mur90] Gerard J. Murphy. *C*-algebras and operator theory*. Boston, MA etc.: Academic Press, Inc., 1990.
- [Nad20] Mahendra G. Nadkarni. *Spectral theory of dynamical systems*. 2nd expanded edition. Vol. 15. Texts Read. Math. Singapore: Springer; New Delhi: Hindustan Book Agency, 2020.
- [Neu32] John von Neumann. „Proof of the quasi-ergodic hypothesis“. *Proc. Natl. Acad. Sci. USA* 18 (1932), pp. 70–82.
- [Pal76] Günther Palm. „Entropie und Erzeuger in dynamischen Verbänden“. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* 36 (1976), pp. 27–45.
- [Pal77] Günther Palm. „A common generalization of topological and measure-theoretic entropy“. In: *International conference on dynamical systems in mathematical physics, Rennes, France, September 14–21, 1975*. Paris: Société Mathématique de France (SMF), 1977.
- [Par69] William Parry. *Entropy and generators in ergodic theory*. Math. Lect. Note Ser. The Benjamin/Cummings Publishing Company, Reading, MA, 1969.
- [Ped89] Gert K. Pedersen. *Analysis now*. Vol. 118. Grad. Texts Math. New York etc.: Springer-Verlag, 1989.
- [Pet83] Karl Petersen. *Ergodic theory*. Vol. 2. Camb. Stud. Adv. Math. Cambridge University Press, Cambridge, 1983.
- [Phe01] Robert R. Phelps. *Lectures on Choquet’s theorem*. Vol. 1757. Lect. Notes Math. Berlin: Springer, 2001.
- [Pie74] Albrecht Pietsch. „s-numbers of operators in Banach spaces“. *Stud. Math.* 51 (1974), pp. 201–223.
- [Pie07] Albrecht Pietsch. *History of Banach spaces and linear operators*. Basel: Birkhäuser, 2007.
- [Que10] Martine Queffélec. *Substitution dynamical systems. Spectral analysis*. 2nd ed. Vol. 1294. Lect. Notes Math. Dordrecht: Springer, 2010.
- [Rak97] Vladimir Rakočević. „On continuity of the Moore-Penrose and Drazin inverses“. *Mat. Vesn.* 49.3-4 (1997), pp. 163–172.
- [RS80] Michael Reed and Barry Simon. *Methods of modern mathematical physics. I: Functional analysis. Rev. and enl. ed.* New York etc.: Academic Press. 1980.

- [Rok48] Vladimir A. Rokhlin. „A ”general” measure-preserving transformation is not mixing“. *Dokl. Akad. Nauk SSSR, n. Ser.* 60 (1948), pp. 349–351.
- [Rok67] Vladimir A. Rokhlin. „Lectures on the entropy theory of measure-preserving transformations“. *Russ. Math. Surv.* 22.5 (1967), pp. 1–52.
- [Rud87] Walter Rudin. *Real and complex analysis*. 3rd ed. New York, NY: McGraw-Hill, 1987.
- [Sch74] Helmut H. Schaefer. *Banach lattices and positive operators*. Die Grundlehren der mathematischen Wissenschaften, Band 215. Springer-Verlag, New York-Heidelberg, 1974.
- [SWA78] Helmut H. Schaefer, Manfred Wolff, and Wolfgang Arendt. „On lattice isomorphisms with positive real spectrum and groups of positive operators“. *Math. Z.* 164 (1978), pp. 115–123.
- [SW99] Helmut H. Schaefer and Manfred P. Wolff. *Topological vector spaces*. Second. Vol. 3. Graduate Texts in Mathematics. Springer-Verlag, New York, 1999.
- [Sch71a] Egon Scheffold. „Das Spektrum von Verbandsoperatoren in Banachverbänden“. *Math. Z.* 123 (1971), pp. 177–190.
- [Sch71b] Egon Scheffold. „Die erzeugte Algebra eines Markov-Verbandsoperators in $C(X)$ “. *Math. Ann.* 193 (1971), pp. 76–82.
- [Sch10] Peter J. Schmid. „Dynamic mode decomposition of numerical and experimental data“. *J. Fluid Mech.* 656 (2010), pp. 5–28.
- [SKK16] Christof Schütte, Péter Koltai, and Stefan Klus. „On the numerical approximation of the Perron-Frobenius and Koopman operator“. *J. Comput. Dyn.* 3.1 (2016), pp. 51–79.
- [Sen72] F. Dennis Sentiilles. „Bounded continuous functions on a completely regular space“. *Trans. Amer. Math. Soc.* 168 (1972), pp. 311–336.
- [SS95] Lynn A. Steen and J. Arthur Seebach Jr. *Counterexamples in topology*. Reprint of the second (1978) edition. Dover Publications, Inc., Mineola, NY, 1995.
- [Ste69] Gilbert W. Stewart. „On the continuity of the generalized inverse“. *SIAM J. Appl. Math.* 17 (1969), pp. 33–45.
- [Voi95] Dan Voiculescu. „Dynamical approximation entropies and topological entropy in operator algebras“. *Commun. Math. Phys.* 170.2 (1995), pp. 249–281.
- [Wal82] Peter Walters. *An introduction to ergodic theory*. Vol. 79. Grad. Texts Math. Springer, Cham, 1982.
- [Wei89] George Weiss. „Weakly ℓ^p -stable linear operators are power stable“. *Int. J. Syst. Sci.* 20.11 (1989), pp. 2323–2328.
- [WKR15] Matthew O. Williams, Ioannis G. Kevrekidis, and Clarence W. Rowley. „A data-driven approximation of the koopman operator: extending dynamic mode decomposition“. *J. Nonlinear Sci.* 25.6 (2015), pp. 1307–1346.
- [Yan15] Kesong Yan. „Conditional entropy and fiber entropy for amenable group actions“. *J. Differ. Equations* 259.7 (2015), pp. 3004–3031.

- [Yos38] Kosaku Yosida. „Mean ergodic theorem in Banach spaces“. *Proc. Imp. Acad. Japan* 14 (1938), pp. 292–294.
- [YK38] Kosaku Yosida and Shizuo Kakutani. „Application of mean ergodic theorem to the problems of Markoff’s process“. *Proc. Imp. Acad. Japan* 14 (1938), pp. 333–339.
- [Zha92] Xiao-Dong Zhang. „Some aspects of the spectral theory of positive operators“. *Acta Appl. Math.* 27.1-2 (1992), pp. 135–142.
- [Zha93] Xiao-Dong Zhang. „On spectral properties of positive operators“. *Indag. Math., New Ser.* 4.1 (1993), pp. 111–127.

Index

A			
adjoint operator	12, 26	composition operator	16
AL-space	22	conditional expectation	128
algebra		conjugacy	19
Boolean	54	consistent topology	25
algebra ideal	35	continuous functions	
almost weakly stable	60	pointwise order	17
AM-space	17, 163	contractive approx. entropy	152
analysis operator	77, 112	convergence in density	61
approximation entropy		covering number	102
fractional	90		
approximation property	171	D	
asymptotic density	61	decay of correlation	81
convergence	61	delay observable	111
atom (measure space)	99	delay subspace	111
		disjoint complement	31
B		disjoint elements	31
Banach lattice		distributive lattice	54
AM-space	17	doubly power-bounded	159
continuous functions	17	duality pairing	11
quasi-interior point	138, 167	dynamic refinement	71
real part	29	dynamical lattice	153
band	31		
principal	31	E	
projection band	31	eigenvalue	
band projection	28	semi-simple	158
Bernoulli shift	52	entropy	
two-sided	52	measure-theoretic	72
Boolean algebra	54	ergodic sampling	108
		ergodicity	59
		mean ergodic projection	58
		mean ergodicity	58
		uniform	157
		essential norm	170
		essential spectral radius	170
		eventually finite-rank operator	172
		F	
		Fejér approximation	85
		fractional approximation entropy	90
		frame	115
C			
C^* -algebra			
continuous functions	18		
Calkin algebra	170		
Cesàro averages	58		
Cesàro bounded	158		
common refinement	71		
of open covers	102		
compact operator	12, 169		
completely regular	14		

analysis operator	77, 112	operator	54, 138
frame operator	77	projection	128, 138
Gram matrix	77, 113	Markov chain	124
synthesis operator	113	time reversal	126
tight	115	mean ergodic decomposition	60
frame operator	77	mean ergodic projection	58
functionally Hausdorff	14, 165	mean ergodicity	58
		measure	21
G		locally finite	21
Gram matrix	77, 113	outer regular	21
		positive	21
H		pushforward	52, 64, 110
Hölder adjoint	53	Radon	21
		regular	57
I		support	65
ideal	35	tight	21, 57
vanishing ideal	35	total variation	21
identity operator	12	measure-preserving map	51
iid sampling	108	invertibility	51
information hull	110	measure-preserving systems	
inner product	11	ergodicity	59
interval preserving	20	weak topology	64
invariant measure	12	measure-theoretic entropy	72
irrational rotation	52	measurement function	153
		mixing	
J		strong	63
Jacob–de Leeuw–Glicksberg		weak	61
decomposition	60	multiplicity function	94
almost weakly stable	60		
reversible	60	N	
join of open covers	102	norm-preserving	34
		normal space	14
K		norming dual pair	24
Koopman operator	16, 53	consistent topology	25
adjoint	53	Mackey topology	25
contravariance	16	of Banach lattices	27
		weak topology	25
L		O	
lattice		observable	
distributive	54	delay	111
lattice homomorphism	157	full state	134
unital	18, 157	open cover	102
lattice ideal	35	operator	
linear operator	12	compact	169
linear span	11	eventually finite-rank	172
		quasi-compact	169
M		order interval	20, 28, 29
Mackey–Arens theorem	26		
Markov		P	
automorphism	139	partition	
embedding	54		
isomorphism	54		

common refinement	71	spectrum	12
dynamic refinement	71	standard Lebesgue	56
partition of unity	149	strict topology	50
periodicity	158	strongly connected	14
Perron–Frobenius operator	22, 53	strongly mixing	63
covariance	23	subcover	102
pointwise convergence	47	subfix point	167
Polish space	56	support of a measure	65
principal band	31	synthesis operator	113
probability measure	12		
projection band	31	T	
pseudoinverse	115	Toeplitz matrix	82
pushforward measure	12, 52, 64,	symbol	82
110		Wiener class	82
		topological entropy	103
Q		topology	
quasi-compact operator	169	completely regular	14
quasi-completeness	50	consistent	25
quasi-interior point	138, 167	functionally Hausdorff	14, 165
		normal	14
R		of pointwise convergence	47
Radon measure	21	strongly connected	14
refinement of open covers	102	total variation measure	12, 21
resolvent	12		
resolvent set	12	U	
reversible	60	uniform ergodicity	157
Riesz–Kantorovich property	30		
rotation system	52	V	
		vanishing ideal	35
S			
semi-simple eigenvalue	158	W	
spectral measure	85, 94	weakly mixing	61

