Jagiellonian University
Faculty of Mathematics and Computer Science
Institute of Mathematics

Topological Dynamics of Countable Amenable Groups

Marta Straszak

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Thesis Supervisor: Dominik Kwietniak

Kraków 2020

to my husband Damian

Acknowledgement

First of all, I would like to express my sincere gratitude to my advisor Dominik Kwietniak. I have been extremely lucky to have a supervisor who cared so much about my work. His friendly guidance and expert advice have been invaluable throughout the entire period of our collaboration. Without his support and encouragement, I would not have been able to complete this thesis.

I am also pleased to say thank you to Professor Klaudiusz Wójcik for the support he granted me during all years of my Ph.D. studies.

The completion of my dissertation would not have been possible without Martha Łącka, who has been my closest research partner and a true friend.

I thank Jakub Byszewski for his enthusiasm, many constructive discussions, and helpful advice.

I would also like to extend my gratitude to Łukasz Kosiński and Sylwester Zając, whose guidance and support I experienced in the very early stages of my research.

I would always remember my friends: Jakub Tarnawski, Andrzej Pezarski, Olena Karpel, Michał Świętek, Weronika Czerniawska, Paweł Potorski, and Maria Dudek. I am grateful for their encouragement and the fun-time we spent together. I also thank Pierre Guillon for his kindness and motivation.

I would not be where I am today without the support and encouragement of my parents, grandmother, and sister. They always believed in me like no one else ever had.

In the end, I would like to thank my husband Damian for his love and constant support. Not only he motivates me and keeps me going on, but he also serves me as an example of an excellent mathematician.

Abstract

Dynamical systems are ubiquitous in all of today's science. Virtually every natural process, whether it is in physics, biology, economics or sociology can be modeled as a dynamical system. Consequently, there has been a lot of research aimed to understand their mathematical properties and especially their asymptotic behavior. Although many of the nature-inspired dynamical systems are amenable to a detailed mathematical analysis and thus allow for reliable predictions of future behavior, it became clear quite early on that not all of them behave this way. The first glimpse of that phenomena comes from the work of Poincaré. More recently, in 1961, Lorenz discovered that his model of atmospheric convection is extremely sensitive to initial conditions, thus even deterministic dynamical systems derived from the physical laws of nature may generate chaos. One of the most important tools in the theory of dynamical systems, and also the main research subject of this thesis, is the notion of entropy, familiar also from physics, information theory and statistics.

In this dissertation, we are concerned with discrete dynamical systems that consist of a state-space X and an invertible map $T\colon X \to X$ that specifies how the states evolve in a unit of time. Another way of thinking about discrete dynamical systems is to identify them with the action of the additive group $\mathbb{Z}$ on X. This leads to a generalization, where we replace $\mathbb{Z}$ with an arbitrary countable discrete group G. Moreover, it is convenient to additionally assume that G is amenable. This assumption is necessary and natural to generalize basic concepts of ergodic theory such as entropy or Banach density, as well as to prove generalizations of fundamental results such as Krylov–Bogolyubov theorem or the pointwise ergodic theorem.

One of the main motivations for this thesis is the following fundamental question about the entropy in the archetypical example of a symbolic dynamical system, the Cantor space $X = \{0,1\}^G$. Which values between 0 and $\log 2$ can be realized as the entropy on a particular family of subsystems of X? One attempt of answering this question was made by F. Krieger, who proved via an intricate construction that each value can be realized as the entropy of some minimal subsystem of X, provided that G is residually finite. In this thesis, we present a novel approach to the proof of realizability theorem for minimal subsystems that allows us to extend it from residually finite groups to congruent monotileable groups. In our proof, the crucial realization is that if the space X is equipped with the Weyl metric, then the entropy (thought of as a function mapping $x \in X$ to the topological entropy of the closure of the G-orbit of x) becomes a continuous function (while the analogous property for the standard metric on $\{0,1\}^G$ obviously fails). In connection to that, we study the interplay between the Weyl metric, invariant measures on X and minimal subsystems of X, and obtain generalizations of the results of Downarowicz and Iwanik (who studied the case $G = \mathbb{Z}$).

As another contribution, we prove a new variant of the realizability theorem that proximal subsystems also achieve all values of entropy. This result we approach from a completely different angle. We develop the theory of subordinate shifts and show that we can control their entropy, manipulating the density of 1's in their primary shifts.

Contents

Chapter 1

Introduction

Dynamical systems are ubiquitous in all of today's science. Virtually every natural process, whether it is in physics, biology, economy or sociology can be modeled as a dynamical system and thus studied with the help of rich mathematical tools developed over a hundred years. To give a small sample of applications of dynamical systems let us here mention only: Maxwell's equations that describe electric and magnetic fields and how they interact with each other [46], in cosmology, exact solutions of Einstein's field equations allow to predict the formation of black holes and study the evolution of the Universe [3], in neuroscience dynamical systems are used to describe processes that occur in brain [34], in linguistics, they have applications in understanding how languages form [22] and in biology to study the evolution of species [45]. It is then natural to ask: what makes dynamical systems so useful in studying all these areas? The main reason, arguably, is that the mathematical models established for these physical phenomena are rather "tame". By that, we mean that the mathematical tools developed in dynamical systems to study these models allow for successful analysis of their properties and asymptotic behavior. In stark contrast to these stand the seminal discoveries of Poincaré [51] and Lorenz [44], (and later also Smale [57], May [47], Li and York [41]), who observed that the aforementioned tameness is by no means a given, even in physical systems. In particular, in his work on the 3-dimensional model of atmospheric convection Lorenz [44] states:

> "Two states differing by imperceptible amounts may eventually evolve into two considerably different states. If, then, there is any error whatever in observing the present state — and in any real system such errors seem inevitable — an acceptable prediction of an instantaneous state in the distant future may well be impossible.".

Through Lorenz's work, it became clear that certain dynamical systems, even though originating from nature, exhibit erratic behavior and remarkable sensitivity to initial conditions also known as chaos. This gave rise to the new area of study known as chaos theory. Lorenz summed up the theory in one sentence:

> "When the present determines the future, but the approximate present does not approximately determine the future.".

Since these discoveries researchers attempted to understand which dynamical systems are chaotic, what are the origins of chaos and how to measure it. The study of chaos in dynamical systems motivates further study of entropy, which, roughly speaking, describes the level of chaos in such a system, but also turns out to be a central quantity in thermodynamics, statistical mechanics or information theory (see [15], [13], [48]). It was introduced primarily as an invariant of conjugacy, but later it became an interesting subject of research on its own.

Mathematically, dynamical systems[1] are typically thought of as a map $T : X \to X$ where X is a universe, i.e., the set of all possible states of a given physical system, and T describes how the system transitions from a particular state to the next one. In applications, X is typically a subspace of $\mathbb{R}^n$ and T "implements a certain physical law", which in mathematics is generally abstracted out by assuming that X is a metric space and T is continuous. Such actions of T and the study of their properties are central to ergodic theory, which can be seen as an attempt to capture and understand chaos in a mathematically rigorous and precise manner. Entropy is one of the tools that allow to capture how much "disorder" does T inject via its action is the entropy $h(X) \in [0, +\infty]$ of such a system.

Thanks to the mathematical abstraction, we are not restricted to the study of, say "tame" subsets of $\mathbb{R}^n$, but we can work with arbitrary metric spaces. This importantly includes symbolic spaces such as the Baire space $\mathbb{N}^{\mathbb{Z}}$, or the Cantor space $\{0,1\}^{\mathbb{Z}}$, where the map T is now defined as simply shifting the corresponding (doubly-infinite) sequence of 0's and 1's one position to the left. The study of such spaces and the entropy of their subsets is crucial for information theory because they allow to digitalize dynamical systems occurring in nature so that no information is lost (see [15]).

An important generalization of the above setting becomes apparent after realizing that in the typical case when T is a homeomorphism of X, the dynamical system T defines, is a continuous action of $\mathbb{Z}$ onto X: indeed the action of $n \in \mathbb{Z}$ on $x \in X$ is to simply apply T to x n times, i.e., $n \cdot x \overset{\text{def}}{=} T^n(x)$. By replacing $\mathbb{Z}$ with arbitrary (countable, discrete) groups G we obtain structures that accurately describe physical systems in thermodynamics, quantum statistical mechanics or neural networks (see [54], [33], [23], [24]). To make progress in this generalized setting, one is required to formulate suitable generalizations of such fundamental notions as entropy or Banach density as well as ensure the existence of an invariant measure for any continuous group action. This does not seem to be straightforward when G is arbitrary. In this thesis, we work under the assumption that G is amenable, which seems to be the weakest assumption one can make for all these notions to still make sense. Intuitively, amenability guarantees that the averaging operation on bounded functions is invariant under translation by group elements.

One of the main motivations for this thesis comes from a question asked by Krieger in [37] about the entropy of subsets of the archetypical example of a symbolic space: the Cantor space[2] $\{0,1\}^G$. Since the entropy of every subspace of $\{0,1\}^G$ is between 0 and $\log(2)$ he asked whether every value in between these two can be realized as the entropy of a minimal (i.e., one that has no nontrivial invariant subspaces) subspace Y of $\{0,1\}^G$. He managed to answer this question positively in the case when G is residually finite. Krieger's proof proceeds via an intricate implicit combinatorial construction of such subspaces that seems to crucially rely on the assumption of G being residually finite. In this thesis, we provide an alternative approach to proving this realizability result for minimal subsystems which is arguably cleaner and simpler, and additionally allows us to extend it from residually finite groups to congruent monotileable groups[3] (for a definition see Section 2.3). Furthermore, we offer a generalization of the realizability question, i.e., whether all values of entropy can be realized by proximal subsystems (every two orbits of a proximal subsystem come arbitrarily close to each other during the group evolution) of $\{0,1\}^G$, and provide a partial (for residually finite groups) positive answer to it. On the way towards these generalizations of Krieger's realizability of entropy theorem we have encountered several other interesting mathematical objects and numerous questions that result from studying their properties. The short overview that we provide below

[1] Dynamical systems as formalized this way are known as discrete dynamical systems, as they describe the evolution of a state in discrete time steps: $x \mapsto T(x) \mapsto T^2(x) \mapsto T^3(x) \mapsto \ldots$, while continuous dynamical systems are typically described via differential equations and consequently the time is "continuous".

[2] The action of $g \in G$ onto $x \in \{0,1\}^G$ is defined as $(g \cdot x)(h) = x(hg)$ for $h \in G$.

[3] The notion of congruent monotileable group was introduced by P. Cecchi and M. I. Cortez in [8].

serves as a brief introduction to these questions and explains what progress does this thesis achieve in answering them.

Our approach to the realizability of entropy problem for minimal subsystems of $\{0,1\}^G$ can be, roughly, summarized in two steps: 1) prove that the entropy of a subsystem $\overline{Gx}$ varies continuously with $x \in \{0,1\}^G$, 2) find a subset $S \subseteq \{0,1\}^G$ with the following properties: a) S is connected, b) every subsystem $\overline{Gx}$ for $x \in S$ is minimal, c) among subsystems $\overline{Gx}$ for $x \in S$ there exist one with arbitrarily small (i.e., zero) and arbitrarily large (i.e., $\log(2)$) entropy. Note that having proved these two steps, the realizability of entropy for minimal systems follows easily: since S is connected, both small and large values of entropy are achieved within S, and entropy is continuous, then all possible values are achieved. Before the reader objects, it is important to clarify that such a plan in this form has no chance to succeed because the entropy function is actually not continuous with respect to the usual metric[4] in $\{0,1\}^G$. Yet interestingly, one can still carry out the steps with a small adjustment: consider a specially crafted metric D_W on $\{0,1\}^G$ instead of the one $\{0,1\}^G$ is equipped with. A choice of D_W under which we can make this plan work is the Weyl metric (see Definition 4.2) that was introduced in [35] and widely studied in [19], [6], [55] among others. We show, in particular, that entropy is indeed continuous with respect to D_W (see Theorem 5.2) and that by taking the set of all quasi-Toeplitz configurations[5] (see Definition 7.1) as S, the conditions a), b), c) are satisfied relative to D_W (see Lemma 7.3, Lemma 7.1 and Theorem 8.1). Recently, the author became aware of an unpublished manuscript[6] [52] that with some additional effort might likely yield a different proof of Theorem 8.1. For more details we refer to Chapter 8.

We perform a thorough study of properties of the Weyl metric in Chapters 4 and 5 of the thesis. In the former, we provide a definition along with a proof of several alternative formulations (see Definition 4.2, Lemma 4.9 and Theorem 4.1). While we mainly work in the most general setting with X being a compact metric space, we also obtain some results specific to the shift space $\{0,1\}^G$, such as a formula for the Weyl metric involving Banach density (see Theorem 4.2). In Chapter 5 we study further properties of the Weyl metric. We obtain the previously mentioned continuity of entropy result, which in the general case of arbitrary X becomes lower semi-continuity (see Theorem 5.1). In Section 5.2 of this chapter we also show that if X is equipped with the D_W metric, then the number of minimal components of $\overline{Gx}$ varies lower semicontinuously with $x \in X$, while in Section 5.3 we show a similar result: the function that assigns to $x \in X$ the set of invariant measures on $\overline{Gx}$ is uniformly continuous with respect to the Weyl metric. The results of this chapter generalize [19] from $G = \mathbb{Z}$ to arbitrary countable amenable group. Chapter 7 is devoted to the study of quasi-Toeplitz configurations. It shows in particular that every quasi-Toeplitz configuration generates a minimal subsystem (see also [8]) and that the family of quasi-Toeplitz configurations is path-connected with respect to the Weyl metric.

Even though the statement of the entropy realizability theorem for proximal systems is so similar to the version for minimal subsystems, the proof technique we employ is completely different. More precisely, for this variant of the theorem, we develop the theory of subordinate shifts (see Chapter 6 and for case $G = \mathbb{Z}$ see [39]) that turns out to be suitable for proximal subsystems achieving arbitrary values of entropy. What we show is that for an $x \in \{0,1\}^G$ (if constructed in a specific way) its subordinate, i.e., $Y \stackrel{\text{def}}{=} \{y \in \{0,1\}^G : y \leqslant x\}$ is proximal and has entropy equal, roughly, to the asymptotic density of 1's in x. This allows us to obtain a proof of entropy realizability by giving a construction of x's with any given asymptotic density of ones. An especially important role in the proof of proximality of such subordinate shifts plays a general formula for counting minimal subsystems that we develop in Chapter 3. The final chapter of the thesis, Chapter 8 serves as a summary and conclusion of the proofs for both versions of the entropy realizability theorems that

[4]Metric for the product topology on $\{0,1\}^G$ coming from the discrete topology on $\{0,1\}$.

[5]Quasi-Toeplitz configurations for residually finite groups were studied in [11], [10], [38].

[6]The author would like to thank Dominik Kwietniak and Benjamin Weiss for obtaining a copy of this manuscript.

we obtain.

Composition of the Thesis

The material of this thesis is partially based on the joint work [9] with Martha Łącka, the remaining results have not yet been published. A detailed breakdown of all the results chapter by chapter follows below. More historical comments and references are provided within each chapter separately.

In **Chapter 3** the main result is Theorem 3.1 that gives the full characterization of minimal subsystems of a transitive dynamical system. Its first part (i.e, formula (3.1)), that gives an upper bound on the number of minimal subsystems was published in [9]. Its second part (i.e., condition (ii)), that allows to find the minimal subsystem, was never published before.

The main results in **Chapter 4** are Theorems 4.1 and 4.2 that give equivalent formulas for the Weyl pseudometric in general and in the shift space respectively. Both theorems were published in [9]. However, Lemma 4.12 which is the main part of the proof of Theorem 4.2, was stated in [9] without the proof. Lemma 4.12 gives a general equivalent formula for the Weyl pseudometric that holds for any compact metric space X and any equicontinuous and uniformly bounded family of functions $\mathcal{K} \subseteq \mathbb{R}^{\mathbb{R}}$ such that the family $\mathcal{K}_G = \{x \mapsto \phi(gx) : \phi \in \mathcal{K}, g \in G\}$ separates the points of X (G is a countable amenable group). The detailed proof of Lemma 4.12 we present in this thesis was never published before. Alternative formulas for the Weyl psudometric that are presented in Lemmas 4.8 and 4.9 were publish in [9]. Moreover, the new result is Lemma 4.5, which gives an explicit formula for the upper Banach density of "periodic sets" in congruent monotileable groups.

In **Chapter 5** we study a quasi-uniform continuity in relation to topological entropy, minimal subsystems and invariant measures. All results of this chapter also appeared in [9], but here we give more detailed proofs. Subsection 5.1.1 of this chapter is devoted to the proof of Theorem 5.1, which states that the map that assigns to an element $x \in X$ the topological entropy of the closure of the orbit of x (denoted by $\overline{Gx}$) is lower semicontinuous with respect to the Weyl pseudometric. In Subsection 5.1.2 we show that if X is a shift space, then the same map as in Theorem 5.1 is quasi-uniform continuous (see Theorem 5.2). The aim of Section 5.2 is to prove Theorem 5.3, which states that the map that assigns to an element $x \in X$ the number of minimal components of $\overline{Gx}$ is lower semicontinuous with respect to the Weyl pseudometric. Finally, in Section 5.3 we show that the map that assigns to an element $x \in X$ the set of invariant probabilistic Borel measures supported in $\overline{Gx}$ is uniformly continuous with respect to the Weyl pseudometric (see Theorem 5.4). Moreover, in Lemma 5.9, that is crucial in our proof of Theorem 5.4, we justify that all the invariant measures supported in $\overline{Gx}$ are the distribution measures.

Chapter 6 deals with the topological entropy of subordinate shifts. In Theorem 6.1 we prove that for every subshift Z with zero entropy, entropy of its subordinate $\check{Z}$ can be expressed as the asymptotic density of ones appearing in Z. In Lemma 6.6 we prove that the entropy of $\check{Z}$ is also given by a maximum G-invariant measure of a cylinder $[1]_e$. Moreover, we prove that if G is reasidually finite and an element $z \in \{0,1\}^G$ is periodic, then the entropy of $\overline{Gz}$ is precisely an average number of ones appearing in z over the periodic set of indices (see Lemma 6.7). All results from this chapter are new and have never been published before in this generality.

Chapter 7 generalizes results from [9] from the case that G is amenable residually finite to all congruent monotileable groups. The aim of this chapter is to prove Lemma 7.3, which says that the family of quasi-Toeplitz configurations along a common Følner sequence is path-connected with respect to the Weyl pseudometric. Moreover, we define a regular quasi-Toeplitz configuration and in Lemma 7.4 we justify that every regular quasi-Toeplitz configuration is a quasi-uniform limit of periodic configurations.

In **Chapter 8** we present two main results of this thesis. The first one is Theorem 8.1, which states that if G is a congruent monotileable group acting on a finite alphabet $\mathscr{A}$ by shift, then for any number $\gamma \in [0, \log |\mathscr{A}|)$ there exists a minimal subshift with entropy γ. This generalizes [9], where an analogous theorem was proved for the case that G is an amenable residually finite group. The second result is Theorem 8.2, which says that if an amenable residually finite group acts on $\{0, 1\}$ by shits, then we can find a proximal subshift with any given entropy between 0 and $\log 2$. This theorem is new and was never published before.

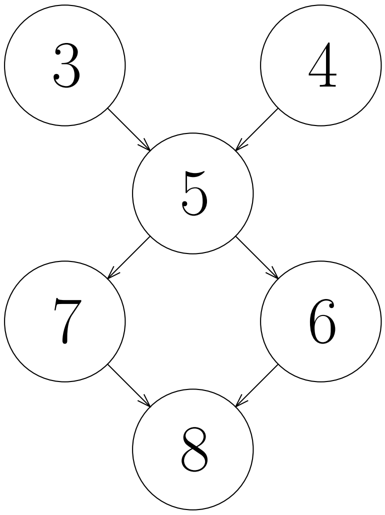

Figure 1.1: A diagram depicting interrelations between the chapters.

Chapter 2

Preliminaries

2.1 Background on Dynamical Systems

Throughout this thesis we study topological transformation groups, or in other words continuous group actions on topological spaces. While these notions are rather standard, for the sake of completeness we provide all the required definitions here. We follow the treatment of [12] and [30]. From now on G always denotes an infinite group and its identity element is denoted by e.

Definition 2.1 (Topological transformation group)**.** A **topological transformation group** is a triple (X, G, φ), where X is a compact metric space, G is a countable discrete group and φ is a continuous (left) group action of G on X, that is,

$$\varphi\colon G \times X \to X$$

is a map satisfying the following conditions:

1. $\varphi(e, x) = x$ for all $x \in X$,
2. $\varphi(g, \varphi(h, x)) = \varphi(gh, x)$ for any $g, h \in G$ and $x \in X$,
3. φ is continuous.

If such a group action φ of G on X is given, then we simply say that ***G* acts on *X*** or that φ is a ***G*-action** on X.

Note that since we restrict our attention to countable groups endowed with the discrete topology (countable discrete groups), the condition that φ is continuous is equivalent to the fact that the map $g \mapsto \varphi(g, x)$ is continuous for every $g \in G$.

In the literature, there is a commonly agreed upon convention to denote a group action by $\cdot$ (dot) and use the infix notation, i.e., $\varphi(g, x) \equiv g \cdot x$. Further, we often omit the dot, i.e., write gx instead of $g \cdot x$ whenever this does not lead to confusion. We also write $(\boldsymbol{X}, \boldsymbol{G})$ instead of (X, G, φ).

Observe that a topological transformation group from Definition 2.1 naturally generalizes the standard notion of a discrete-time dynamical system (X, T), where $T : X \to X$ is a homeomorphism (see for instance [36]). In this case the action $\cdot : \mathbb{Z} \times X \to X$ is defined as

$$n \cdot x \stackrel{\text{def}}{=} T^n(x) \qquad \text{for } n \in \mathbb{N} \text{ and } x \in X.$$

Notice that the above action $\cdot$ is an additive action of $\mathbb{Z}$ on X, which allows us to treat (X, T) as a dynamical system $(X, \mathbb{Z})$, i.e., a dynamical system generated by $\mathbb{Z}$.

In this thesis we use the name "dynamical system" as a synonym for "topological transformation group".

Let (X, G) be a dynamical system. For two subsets $S, T \subseteq G$ we often write ST to denote the set $\{st : s \in S, t \in T\} \subseteq G$ of pairwise products, similarly we define $SA := \{sx : s \in S,\, x \in A\}$ for $A \subseteq X$. We denote the **orbit** of an element $x \in X$ (with respect to a G-action) by

$$Gx \overset{\text{def}}{=} \{gx : g \in G\}$$

and by $\underline{x}_G \in X^G$, the **trajectory** of x, i.e., $\underline{x}_G(g) \overset{\text{def}}{=} gx$ for all $g \in G$. Note that trajectory is a function $G \to X$, while the orbit is the set of the values attained by the trajectory. We say that $x \in X$ is a **fixed point** if $Gx = \{x\}$. A subset $A \subseteq X$ is ***G*-invariant** if $GA \subseteq A$. If $Y \subseteq X$ is a non-empty closed G-ivariant set, we say that Y is a **subsystem** of X. What follows is a quite simple, yet useful property of invariant sets:

Lemma 2.1. *Let (X, G) be a dynamical system. If $A \subseteq X$ is invariant, then $\overline{A}$ (the closure of A) is invariant.*

Proof. Let $x \in \overline{A}$ and $g \in G$. There exist a sequence $\{x_n\}_{n \in \mathbb{N}} \subseteq A$ such that $x_n \to x$ $(n \to \infty)$. Since A is invariant, $gx_n \in A$ for every $n \in \mathbb{N}$. We also have $gx_n \to gx$ since G acts on X continuously. Therefore $gx \in \overline{A}$. $\square$

Observe that Lemma 2.1 implies that $\overline{Gx}$ is a subsystem of X for every $x \in X$.

When G acts on X then this automatically gives rise to G-actions (i.e. actions of G) on the space of functions Y^X (for any set Y) and on the space of Borel probability measures $\mathcal{M}(X)$ on X endowed with the weak-$\star$ topology. Formal definition of these actions follow

Definition 2.2 (Induced Group Actions)**.** Suppose that (X, G) is a dynamical system, then:

1. For an arbitrary set Y, the G-action induced on Y^X is defined as follows: for any $\tau : X \to Y$ and $g \in G$ the function $(g \cdot \tau) : X \to Y$ satisfies

$$(g \cdot \tau)(x) \overset{\text{def}}{=} \tau(gx) \quad \text{for } x \in X.$$

2. If $\mu \in \mathcal{M}(X)$ is any Borel probability measure on X then the result of the action of an element $g \in G$ on μ is denoted by $g_*(\mu)$ and is given by

$$g_*(\mu)(A) \overset{\text{def}}{=} \mu(g^{-1}A) \qquad \text{for any } A \in \mathcal{B}(X),$$

where $\mathcal{B}(X)$ denotes the collection of all Borel sets on X. This defines a dynamical system $(\mathcal{M}(X), G)$ on the compact space $\mathcal{M}(X)$.

In this thesis the following classes of dynamical systems often show up and play an important role

Definition 2.3 (Classes of dynamical stystems)**.** We say that a dynamical system (X, G) is

1. **transitive** if there exists an $x \in X$ such that $X = \overline{Gx}$;
2. **minimal** if X does not contain any non-empty, proper, closed G-invariant subset;
3. **proximal** if $\inf\limits_{g \in G} \rho(gx, gz) = 0$ for every $x, z \in X$.

As already mentioned, when it comes to X, we are mainly interested in compact metric spaces. The only assumption on G that we have made so far is that it is countable, now comes the second crucial assumption that allows us to study entropy and other familiar notions developed for dynamical systems generated by $\mathbb{Z}$, i.e., that G is an amenable group.

Definition 2.4 (Amenability)**.** We say that a countable group G is **amenable** if it admits a **Følner sequence** $\{F_n\}_{n\in\mathbb{N}}$ in G (where $\mathbb{N} = \{0, 1, 2, \ldots\}$), that is a sequence of non-empty finite subsets of G satisfying

$$\lim_{n\to\infty} \frac{|gF_n \Delta F_n|}{|F_n|} = 0 \qquad \text{for every } g \in G,$$

where Δ denotes a symmetric difference.

For a thorough discussion on amenable groups, including a historical perspective, basic properties and examples, we refer to [7].

From now on, without mentioning explicitly, we always assume that ***X* is a compact metric space equipped with a metric $\boldsymbol{\rho}$** satifying[1] $\boldsymbol{\rho(x, z) \leqslant 1}$ for $x, z \in X$ and ***G* is a discrete countable amenable group**. Amenable groups are especially important in the context of dynamical systems since they allow to naturally generalise the notion of topological entropy from classical dynamical systems (i.e., (X, T) where $T\colon X \to X$ is a continuous map) to dynamical systems generated by arbitrary (amenable) group actions. Entropy is the central invariant in the study of these actions and can be defined for an action of an amenable group G (with a Følner sequence $\{F_n\}_{n\in\mathbb{N}}$) on a space X as follows:

Definition 2.5 (Topological entropy)**.** For a finite open cover $\mathcal{U}$ of the space X let $\mathcal{N}(\mathcal{U})$ be the minimum cardinality of a subcover of $\mathcal{U}$. Define the join of two finite open covers $\mathcal{U}$ and $\mathcal{V}$ as

$$\mathcal{U} \vee \mathcal{V} \overset{\text{def}}{=} \{U \cap V : U \in \mathcal{U}, V \in \mathcal{V}\}.$$

Given $F = \{f_1, \ldots, f_s\} \subseteq G$ and $f \in F$ we write

$$f^{-1}\mathcal{U} \overset{\text{def}}{=} \{f^{-1}U : U \in \mathcal{U}\} \quad \text{and} \quad \mathcal{U}^F \overset{\text{def}}{=} (f_1^{-1}\mathcal{U}) \vee \ldots \vee (f_s^{-1}\mathcal{U}).$$

By $h(X,\mathcal{U})$ we denote the limit of the sequence $\left\{ \dfrac{\log \mathcal{N}\left(\mathcal{U}^{F_n}\right)}{|F_n|} \right\}_{n\in\mathbb{N}}$.
The **topological entropy** of (X, G) is defined as

$$h_{\text{top}}(X) \overset{\text{def}}{=} \sup\{h(X,\mathcal{U}) \; : \mathcal{U} \text{ is an open cover of } X\}.$$

Further, for any element $x \in X$ its entropy is defined as $h(x) \overset{\text{def}}{=} h_{\text{top}}(\overline{Gx})$.

Observe that $\vee$ is an associative operation on open covers thus $\mathcal{U}^F$ is well-defined. Moreover, note that for the above definition to make sense, one must show existence of $h(X,\mathcal{U})$. It is also not at all clear that $h(X,\mathcal{U})$ does not depend on the choice of a Følner sequence $\{F_n\}_{n\in\mathbb{N}}$. Indeed, the existence of $h(X,\mathcal{U})$ as well as its independence of the choice of the Følner sequence follows from [43, Theorem 6.1] (see Lemma 4.11).

2.2 Shift Spaces

Even though many of our results apply to the general case, when X is a compact, metric space, an important case that is of special focus is the Cantor space, i.e., when $X = \mathscr{A}^G$ (the space of all functions $G \to \mathscr{A}$), where $\mathscr{A}$ is a finite discrete metric space, called an **alphabet**. A standard way to define a metric ρ on $\mathscr{A}^G$ that induces the Tychonoff product topology on $\mathscr{A}^G$ is

$$\rho(z, y) \overset{\text{def}}{=} |\mathscr{A}|^{-k} \quad \text{with} \quad k \overset{\text{def}}{=} \min\{i \in \mathbb{N} : z(g_i) \neq y(g_i)\}, \tag{2.1}$$

[1] This is without loss of generality as a metric ρ can be always replaced by the equivalent metric $\min(\rho, 1)$.

where $(g_1, g_2, \ldots)$ is any bijective enumeration of elements of G. An especially important family of subsets of $\mathscr{A}^G$ are cylinders:

Definition 2.6 (Cylinder)**.** For a finite set $F \subseteq G$ and $w \in \mathscr{A}^F$, we define a **cylinder** in $\mathscr{A}^G$ by

$$[w]_F = \{x \in \mathscr{A}^G : x_F = w\},$$

where $x_F = x|_F$. If $a \in \mathscr{A}$ and $F = \{g\}$ for some $g \in G$, we denote for simplicity

$$[a]_g = \{x \in \mathscr{A}^G : x(g) = a\}.$$

We recall also that the cylinders are both open and closed in the topology induced by ρ and thus $\mathscr{A}^G$ endowed with ρ has a base consisting of clopen sets. A group action that is perhaps the most natural to consider in such symbolic spaces is *the shift*

Definition 2.7 (Shift action)**.** Given $\mathscr{A}$, we define the **shift action** $(\mathscr{A}^G, G)$ as the following action of G on $\mathscr{A}^G$:

$$\varphi\colon G \times \mathscr{A}^G \to \mathscr{A}^G \quad \text{(with the standard notation } \varphi(g, x) = g \cdot x\text{)},$$

where

$$(g \cdot x)(h) = x(hg) \quad \text{for } h \in G \text{ and } x \in \mathscr{A}^G.$$

Notice that such defined shift action is compatible with what one would think a shift is supposed to be in the case of $G = \mathbb{Z}$, i.e., it moves a sequence $x \in \mathscr{A}^{\mathbb{Z}}$ one place to the left, formally

$$\sigma\colon \mathscr{A}^{\mathbb{Z}} \to \mathscr{A}^{\mathbb{Z}} \qquad \text{satisfying} \qquad \sigma(x)(n) = x(n+1).$$

Closed, non-empty, invariant subspaces of a shift space are called **subshifts**.

2.3 Groups

We review a few notions related to groups as well as define certain classes of groups that are of particular interest in this thesis. We start by a definition of a *monotile* that allows us to study "periodic" sets that appear in connection with minimal and proximal dynamical systems. It was widely studied in [60].

Definition 2.8 (Monotile)**.** A (right) **monotile** of G is a finite set $F \subseteq G$ such that $e \in F$ and $\{Fc : c \in C\}$ is a partition of G for some $C \subseteq G$, i.e., the following two conditions are satisfied:

1. $\bigcup_{c \in C} Fc = G$;
2. $Fc_1 \cap Fc_2 = \varnothing$ for $c_1, c_2 \in C$ with $c_1 \neq c_2$.

We also say that C is a **set of centers** associated with F.

Note that the set $\{e\}$ is a monotile in every group. Moreover, it is important to realize, that a set of centers associated with a given monotile F does not need to be unique.

Example 2.1 (A monotile with uncountably many sets of centers)**.** Let $G = \mathbb{Z} \times \mathbb{Z}_2$. Then the set $F = \{(0,0), (0,1)\}$ is a monotile and for every $\{i_n\}_{n \in \mathbb{Z}} \in \{0,1\}^{\mathbb{Z}}$, the set $\{(n, i_n) : n \in \mathbb{Z}\}$ is a set of centers associated with F.

It is not known whether every amenable group admits a Følner sequence whose elements are monotiles, but it is clearly the case for $G = \mathbb{Z}$ or more generally $G = \mathbb{Z}^n$. This motivates the following definition, which was first introduced in [8] and generalizes the concept of amenable residually finite groups.

Definition 2.9 (Congruent monotileable group)**.** A countable amenable group G is **congruent monotileable** if it admits a Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ satisfying:

1. $F_0 = \{e\}$,
2. $\displaystyle\bigsqcup_{h\in J_n} F_n h = F_{n+1}$ for some $J_n \subseteq G$ with $e \in J_n$ for every $n \in \mathbb{N}$,
3. $\bigcup \mathcal{F} = G$.

If the above conditions hold, then we also say that $\mathcal{F}$ is a **congruent Følner sequence**.

We show that every congruent monotileable group G has an *elegant structure* (see Figure 2.1 and Lemma 2.3) that allows to define a periodic configuration over G.

Lemma 2.2 (Unique factorization)**.** *Let G be a congruent monotileable group and let $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ be a congruent Følner sequence in G. If sets $\{J_n\}_{n\in\mathbb{N}}$ in G are such that for every $n \in \mathbb{N}$ it holds*

$$e \in J_n \qquad \text{and} \qquad \bigsqcup_{h\in J_n} F_n h = F_{n+1},$$

then for every $g \in G$, there exists a unique element $(h_0, h_1, h_2, \ldots) \in J_0 \times J_1 \times J_2 \times \ldots$ such that $h_i = e$ for all but finitely many $i \in \mathbb{N}$ and $g = h_0 h_1 h_2 \ldots$.

Proof. Fix $g \in G$. First we justify the existence of a factorization of g. Since $\bigcup \mathcal{F} = G$, there exists $N \in \mathbb{N}$ such that $g \in F_N$. We set $h_n = e$ for $n \geqslant N$. We can also find $f_{N-1} \in F_{N-1}$ and $h_{N-1} \in J_{N-1}$ such that

$$g = f_{N-1} h_{N-1}.$$

Again, there are $f_{N-2} \in F_{N-2}$ and $h_{N-2} \in J_{N-2}$ such that

$$g = f_{N-2} h_{N-2} h_{N-1}.$$

Proceeding inductively, we obtain elements $h_i \in J_i$ for $i = 0, 1, \ldots, N-1$ such that (since $F_0 = \{e\}$)

$$g = h_0 h_1 \ldots h_{N-1}.$$

To see that this representation is unique, observe that the condition

$$\bigsqcup_{h\in J_n} F_n h = F_{n+1} \qquad \text{for every } n \in \mathbb{N}$$

implies that for every $n \in \mathbb{N}$ and for every $f_1, f_2 \in F_n$ and $h_1, h_2 \in J_n$ it holds

$$f_1 h_1 = f_2 h_2 \qquad \Longrightarrow \qquad f_1 = f_2 \text{ and } h_1 = h_2,$$

thus the sequence $(h_0, h_1, h_2, \ldots)$ is unique. ☐

As a consequence of a unique factorization of elements of G given by sets J_n (for $n \in \mathbb{N}$), we obtain sets of centers associated with $\mathcal{F}$ that has all nice properties required for the construction of a quasi-Toeplitz configuration (see Lemma 7.2).

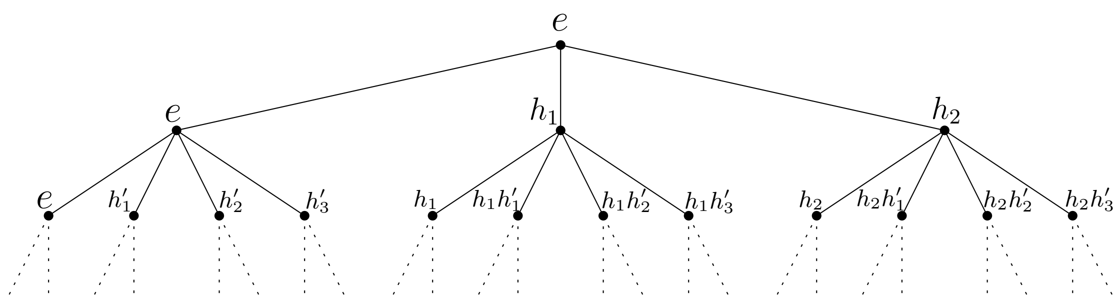

Figure 2.1: The proof of Lemma 2.3 shows us the structure of a congruent monotileable group G from a new perspective. We can think of G as a tree T such that every vertex is identified with some element of G according to the congruent Følner sequence $\{F_n\}_{n\in\mathbb{N}}$ and sets $\{J_n\}_{n\in\mathbb{N}}$ from Definition 2.9. The figure depicts a sample tree T for $J_0 = \{e, h_1, h_2\}$ and $J_1 = \{e, h'_1, h'_2, h'_3\}$. We construct T as follows: we set e to be the root of T. If $g \in G$ is a vertex of depth $n \in \mathbb{N}$, then we define its children as elements gh for $h \in J_n$. Then observe that F_n corresponds to the set of all vertices at depth n. For example $F_1 = \{e, h_1, h_2\}$. Moreover if $f \in F_n$, then fC_n is the set of all vertices of the full subtree of T rooted in f. For example h_1C_1 is the set $\{h_1, h_1, h_1h'_1, h_1h'_2, h_1h'_3, \ldots\}$.

Lemma 2.3. *Let G be a congruent monotileable group. Fix $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ a congruent Følner sequence in G with sets $\{J_n\}_{n\in\mathbb{N}}$ as in Definition 2.9. Then there exist a sequence $\mathcal{C} = \{C_n\}_{n\in\mathbb{N}}$ in G such that for every $n \in \mathbb{N}$ the following conditions hold*

1. *C_n is an associated with F_n set of centers,*
2. *$J_n \subseteq C_n$,*
3. *$C_n = \displaystyle\bigsqcup_{h\in J_n} hC_{n+1}$.*

We say that $\mathcal{C}$ is the **elegant sequence of centers** *associated with $\mathcal{F}$.*

Proof. Fix $n \in \mathbb{N}$. Observe that if $n \geqslant 1$, then

$$F_n = \{h_0h_1 \ldots h_{n-1} : h_i \in J_i \text{ for every } i = 0, 1, \ldots, n-1\}.$$

For $n \geqslant 1$ we define

$$C_n \stackrel{\text{def}}{=} \{h_nh_{n+1}h_{n+2} \ldots : h_i \in J_i \text{ for every } i \in \mathbb{N}_{\geqslant n} \text{ and } h_i = e \text{ for almost all } i \in \mathbb{N}_{\geqslant n}\}.$$

Notice that the condition 3 follows directly from the definition of C_n. It is also clear that $J_n \subseteq C_n$ since $e \in J_i$ for every $i \in \mathbb{N}$. Moreover, by Lemma 2.2 we have guaranteed that $F_nC_n = G$ and $F_nc \cap F_n\tilde{c} = \varnothing$ for every $c, \tilde{c} \in C_n$ such that $c \neq \tilde{c}$. ∎

It is worth emphasizing that from Lemma 2.3 we obtain that if $\{F_n\}_{n\in\mathbb{N}}$ is a congruent Følner sequence, then F_n is a monotile for every $n \in \mathbb{N}$.

As we have already mentioned, the notion of congruent monotileable group generalizes the concept of amenable residually finite group.

Definition 2.10 (Residually finite group)**.** A countable group G is *residually finite* if there exists a sequence $\{H_n\}_{n\in\mathbb{N}}$ of finite index normal subgroups satisfying:

1. $H_n \supseteq H_{n+1}$ for every $n \in \mathbb{N}$,

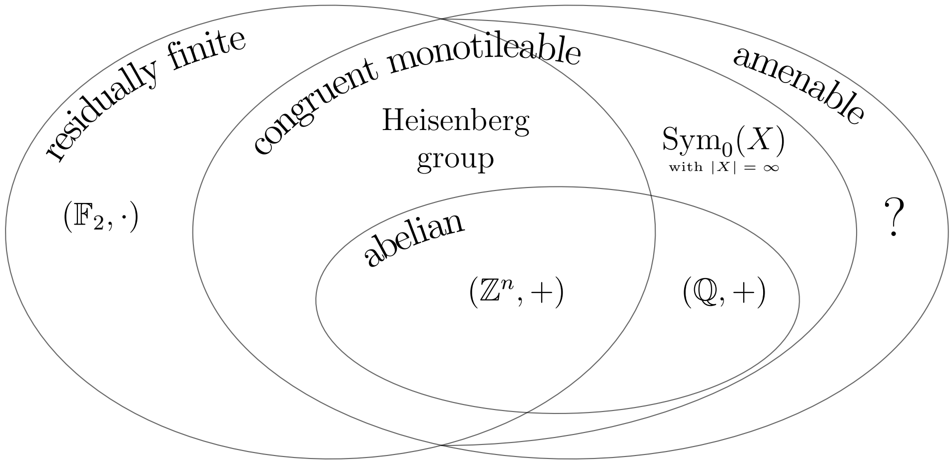

Figure 2.2: A diagram depicting inclusions between different classes of countable groups considered in this thesis. Congruent monotileable groups are a generalisation of residually finite amenable groups. It is not known whether every amenable group is congruent monotileable (this is denoted by a question mark in the diagram). The free group generated by two elements is denoted by $\mathbb{F}_2$. The group $\mathrm{Sym}_0(X)$ is the group of all permutations of X with finite support, where the support of permutation π of X is the set $\{x \in X : \pi(x) \neq x\}$. For the proof that every abelian group is coungruent monotileable see [8].

2. $\bigcap_{n\in\mathbb{N}} H_n = \{e\}$.

Recall that the index of a subgroup H in G is defined as the number of cosets of H in G, that is, the number $|\{gH : g \in G\}|$. We write $H \lhd_f G$ if H is a finite index normal subgroup of G.

It was proved in [11] that if a group is amenable and residually finite, then there exists a special Følner sequence associated to a sequence of finite index normal subgroups $\{H_n\}_{n\in\mathbb{N}}$.

Lemma 2.4. *[11, Lemma 4] If G is an amenable residually finite group then there exist a sequence $\{H_n\}_{n\in\mathbb{N}}$ with $H_n \lhd_f G$ and a Følner sequence $\{F_n\}_{n\in\mathbb{N}}$ satisfying:*

1. *$G = H_0 \supseteq H_1 \supseteq H_2 \supseteq \dots$ and $\bigcap_{n=0}^{\infty} H_n = \{e\}$,*
2. *$\{e\} = F_0 \subseteq F_1 \subseteq F_2 \subseteq \dots$ and $\bigcup_{n=0}^{\infty} F_n = G$,*
3. *$F_{i+1} = \bigsqcup_{h\in F_{i+1}\cap H_i} F_i h$ for every $i \in \mathbb{N}$,*
4. *for every $n \in \mathbb{N}$ the set F_n is a fundamental domain for G/H_n, that is, $|F_n| = |G/H_n|$ and $\{fH_n : f \in F_n\} = \{gH_n : g \in G\}$.*

Therefore if G is amenable and residually finite and $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ is given by Lemma 2.4, then $\mathcal{F}$ satisfies the conditions from Definition 2.9 hence G is a congruent monotileable group. Now, the relation between amenable residually finite groups and congruent monotileable groups is clear: a set of centers associated with a given element of a Følner sequence plays the role of a finite index normal subgroup in a group that is not necessarily residually finite. They give us some form of regularity in a group, but such a structure is not as nice as the one induced by a subgroup of finite index.

In Figure 2.2 we show how the classes of groups defined in this section relate to each other with respect to the inclusion.

Chapter 3

Minimal Subsystems

Since the main topic of this thesis is the entropy of some minimal or proximal dynamical systems, it is not surprising that the following questions are crucial in our investigations:

Is a given dynamical system minimal?

Is a given dynamical system proximal?

The purpose of this chapter is to develop tools allowing us to tackle questions of this type. More specifically, if (X, G) is a dynamical system, (where X is a compact metric space and G is an amenable group) Theorem 3.1 provides a formula that for a given $x \in X$ allows us to find the number of minimal components of $\overline{Gx}$, as well as a convenient method for finding these minimal subsystems.

This formula is then used to prove that a dynamical system generated by a quasi-Toeplitz configuration is minimal (Lemma 7.1). Additionally, as an almost straightforward consequence of Theorem 3.1, we obtain that the function assigning to an element $x \in X$, the number of minimal subsystems of $\overline{Gx}$ is lower-semicontiuous with respect to the Weyl pseudometric (Theorem 5.3). Finally, we use Theorem 3.1 to demonstrate that the dynamical system constructed in the proof of Theorem 8.2 is proximal, which we believe is especially tricky to derive without using this result.

To state Theorem 3.1 let us first introduce some useful notation. For an element $x \in X$ define

$$m(x) \stackrel{\text{def}}{=} \text{the number of minimal subsystems of } \overline{Gx}.$$

It is well known that $m(x)$ is always positive (for the proof see Lemma 3.4). In a discussion at the end of this section we prove that for the case $X = \{0,1\}^{\mathbb{Z}}$, all the values $1, 2, 3, \ldots, \infty$ are attainable by $m(x)$. The formula for computing $m(x)$ involves the notion of a **set of return times**, i.e., for a given $x \in X$ and $A \subseteq X$ we define

$$N(x, A) \stackrel{\text{def}}{=} \{g \in G \,:\, gx \in A\}.$$

Recall that for $Z \subseteq X$ and $\varepsilon > 0$ we let Z^ε to be the ε-neighborhood of Z. We are now ready to state the main result of this chapter.

Theorem 3.1. *If $x \in X$, then*

$$m(x) = \min \left\{ |Z| : Z \subseteq \overline{Gx} \textit{ and } N(x, Z^\varepsilon) \textit{ is syndetic for every } \varepsilon > 0 \right\}. \tag{3.1}$$

Moreover, the following hold:

i) If $N_1, N_2, \ldots, N_{m(x)}$ are all minimal subsystems of $\overline{Gx}$ and $Z = \{z_1, z_2, \ldots, z_{m(x)}\} \subseteq \overline{Gx}$ is such that $z_i \in N_i$ for every $i = 1, 2, \ldots, m(x)$, then Z realizes the minimum in (3.1).

ii) If $Z^\star \subseteq \overline{Gx}$ realizes the minimum in (3.1), then the family of the minimal subsystems of $\overline{Gx}$ is $\{\overline{Gz} : z \in Z^\star\}$.

The first part of the above theorem is a generalisation of [19, Lemma 1], where the case $G = \mathbb{Z}$ was studied. To make sense of the above theorem one needs to define what "syndetic" means. For a formal definition we refer to Section 3.1 (below) but for an intuitive understanding it is enough to think of syndetic subsets of G as very "large". For instance syndetic subsets of $\mathbb{Z}$ are exactly the ones that have uniformly bounded gaps between their elements. Thus, roughly, counting minimal components of $\overline{Gx}$ is equivalent to finding small sets of points $Z \subseteq \overline{Gx}$ such that the orbit of x visits the (arbitrarily) close vicinity of Z with high frequency.

3.1 Thick and Syndetic Sets

Before we proceed to the proof of Theorem 3.1 we need to introduce the concept of thick and syndetic sets and prove some of their properties.

Definition 3.1 (Thick and syndetic sets)**.** We define the following properties of subsets of G

1. **(Thick)** A set $T \subseteq G$ is **thick** if for every finite set $F \subseteq G$ there is $g \in G$ with $Fg \subseteq T$.

2. **(Syndetic)** A set $S \subseteq G$ is **syndetic** if there is a finite set $F \subseteq G$ such that $FS = G$.

More generally one can define left/right thick and left/right syndetic sets, yet in our setting only the above variants are relevant, hence we chose to stick to them and refer to them simply as *thick* and *syndetic*.

Example 3.1 (Thick and Syndetic sets in $\mathbb{Z}$)**.** A set $S \subseteq \mathbb{Z}$ is syndetic if and only if it has uniformly bounded gaps, i.e., there exists $k \in \mathbb{N}$ such that for every $n \in \mathbb{Z}$ one has

$$\{n, n+1, \ldots, n+k\} \cap S \neq \varnothing. \tag{3.2}$$

Indeed, to see that the latter condition implies that S is syndetic assume that there exists $k \in \mathbb{Z}$ such that (3.2) holds for every $n \in \mathbb{Z}$. That means that for every $n \in \mathbb{Z}$ there exist $s \in S$ and $j \in \{0, 1, \ldots, k\}$ such that $n + j = s$. Therefore $\mathbb{Z} = F + S$, where $F = \{-k, -k+1, \ldots, 0\}$.

On the other hand, let $F \in \mathrm{Fin}(\mathbb{Z})$ be such that $F + S = \mathbb{Z}$. Without loss of generality we can assume that $F = \{0, 1, \ldots, k\}$ for some $k \in \mathbb{N}$. Fix $n \in \mathbb{Z}$. We show that the condition (3.2) is satisfied. Observe that $n + k = j + s$ for some $s \in S$ and $j \in F$. Then $s = n + k - j$, which implies $s \in \{n, n+1, \ldots, n+k\}$ and the claim follows.

Similarly, one can show that a set $T \subseteq \mathbb{Z}$ is thick if and only if it contains arbitrarily long intervals of consecutive integers i.e., for every $k \in \mathbb{Z}$ there exists $n \in \mathbb{Z}$ such that $\{n, n+1, \ldots, n+k\} \subseteq T$.

Example 3.1 motivates the following general lemma that captures the duality between these two notions. (Note that this lemma holds rather trivially for $G = \mathbb{Z}$, given the characterisation of thick and syndetic sets provided in the Example 3.1.)

Lemma 3.1 (Duality between thick and syndetic sets)**.** *A set $S \subseteq G$ is syndetic if and only if $S \cap T \neq \varnothing$ for every thick set $T \subseteq G$.*

Proof. Assume that $S \subseteq G$ is syndetic. Let $F \subseteq G$ be a finite set such that $FS = G$. Fix a thick set $T \subseteq G$. Then there is $g \in G$ such that $F^{-1}g \subseteq T$. Since $FS = G$, there are $f \in F$ and $s \in S$ such that $g = fs$. Thus

$$T \supseteq F^{-1}g \ni f^{-1}g = f^{-1}fs = s,$$

which means that $s \in S \cap T$.
Now assume that $S \subseteq G$ is such that $S \cap T \neq \varnothing$ for every thick $T \subseteq G$. This implies that $G \backslash S$ is not thick. Therefore there exists $F \subseteq G$ finite such that for every $g \in G$ we have $Fg \cap S \neq \varnothing$. We claim that $F^{-1}S = G$. Take any $h \in G$. Let $f \in F$ be such that $fh \in S$. Then

$$h = (f^{-1}f)h = f^{-1}fh \in F^{-1}S. \qquad \square$$

Remark 3.1. An analogous property holds for thick sets: a set $T \subseteq G$ is thick if and only if $S \cap T \neq \varnothing$ for every syndetic set $S \subseteq G$.

The following lemma combined with Lemma 3.1 is useful in further investigations concerning syndetic sets (see also [19]).

Lemma 3.2. *If $Z \subseteq \overline{Gx}$ is non-empty and invariant*[1] *(not necessarily closed), then for any $\varepsilon > 0$ the set $N(x, Z^\varepsilon)$ is thick.*

Proof. Let $\varepsilon > 0$ and $F \in \mathrm{Fin}(G)$. We show that there exists $g \in G$ such that $Fg \subseteq N(x, Z^\varepsilon)$. Equivalently, we find $g \in G$ satifiying $fgx \in Z^\varepsilon$ for every $f \in F$. Let $\delta > 0$ be such that $\rho(a, b) < \delta$ implies $\rho(fa, fb) < \varepsilon$ for every $f \in F$ and $a, b \in X$. Now take arbitrary $z \in Z$ and choose $g \in G$ such that $\rho(gx, z) < \delta$. Then $\rho(fgx, fz) < \varepsilon$ for every $f \in F$. But $fz \in Z$ for every $f \in F$ since Z is G-invariant. Therefore $fgx \in Z^\varepsilon$ for every $f \in F$. $\square$

3.2 Counting Minimal Subsystems

This section is devoted to the proof of Theorem 3.1. The proof is an adaptation of a proof presented in [19].

Proof of Theorem 3.1. Let $x \in X$. It is well-known that $m(x) \geqslant 1$ (see also Lemma 3.4). It is most convenient to prove (3.1) by showing that the following three conditions are equivalent:

1. One has $m(x) \leqslant m$.
2. There exist closed G-invariant sets $N_1, N_2, \ldots, N_m \subseteq \overline{Gx}$ such that for every $z \in \overline{Gx}$ there exists $i = 1, \ldots, m$ satisfying $N_i \subseteq \overline{Gz}$.
3. There exists a set of m points $Z = \{z_1, \ldots, z_m\} \subseteq \overline{Gx}$ such that for every $\varepsilon > 0$ the set $N(x, Z^\varepsilon)$ is syndetic.

Note that the equivalence of conditions 1. and 3. proves (3.1) by applying the above for $m := m(x)$ and $m := m(x) - 1$.

To prove that (1) $\Rightarrow$ (2) let $N_1, \ldots, N_{m(x)}$ be all minimal subsystems of $\overline{Gx}$. If $m(x) < m$, put $N_k := N_1$ for $k = m(x)+1, m(x)+2, \ldots, m$. Note that for $z \in \overline{Gx}$ the set $\overline{Gz}$ is a subsystem of $\overline{Gx}$ and hence by Lemma 3.4 it contains a minimal subsystem N. Since N is also a minimal subsystem of $\overline{Gx}$, it must be equal to one of N_i for some $i = 1, \ldots, m$ (here we use the initial assumption that $n \geqslant 1$). Thus (2) holds.

[1]Recall that a set $A \subseteq X$ is invariant if $GA = \{gx : g \in G, x \in A\} \subseteq A$.

To show that (2) $\Rightarrow$ (3) take $z_i \in N_i$ for $i = 1, \dots, m$, where N_i's are as in (2). Denote $Z = \{z_1, z_2, \dots, z_m\}$. Suppose that $N(x, Z^\varepsilon)$ is not syndetic. Let $\{F_n\}_{n\in\mathbb{N}}$ be an increasing sequence of finite sets whose union is equal to G. Then for every $n \in \mathbb{N}$ one can find $g_n \in G$ such that for every $f \in F_n$ one has $fg_nx \notin Z^\varepsilon$. Let $z \in \overline{Gx}$ be a limit point of the sequence $\{g_nx\}_{n\in\mathbb{N}}$. We justify that $\overline{Gz} \cap Z^\varepsilon = \varnothing$, which contradicts (2). Fix $h \in G$. Then there exists $N \in \mathbb{N}$ such that $h \in F_n$ for every $n \geqslant N$. Therefore $hg_nx \notin Z^\varepsilon$ for $n \geqslant N$ and hence $hz \notin Z^\varepsilon$.

It remains to show that (3) $\Rightarrow$ (1). Suppose that $m(x) > m$. Let $Z_1, \dots Z_{m+1}$ be disjoint minimal subsystems of $\overline{Gx}$. Choose $\varepsilon > 0$ such that for any $i \neq j$ one has $\operatorname{dist}(Z_i, Z_j) > 2\varepsilon$. Notice that for any set $Z \subseteq \overline{Gx}$ consisting of m points $z_1, \dots, z_m$ there exists $i_0 = 1, \dots, m$ such that Z^ε is disjoint from $Z_{i_0}^\varepsilon$. Therefore $N(x, Z^\varepsilon) \cap N(x, Z_{i_0}^\varepsilon) = \varnothing$. By Lemma 3.2 the set $N(x, Z_{i_0}^\varepsilon)$ is thick. Hence, by Lemma 3.1, $N(x, Z^\varepsilon)$ is not syndetic contradicting (3). Therefore we must have $m(x) \leqslant m$. This concludes the proof of equivalence of conditions (1), (2) and (3).

Next, we justify (i). By following the steps (1) $\Rightarrow$ (2) and (2) $\Rightarrow$ (3) one can conclude that if $N_1, N_2, \dots, N_m$ are minimal subsystems of $\overline{Gx}$ and $Z = \{z_1, z_2, \dots, z_m\}$ is chosen such that $z_i \in N_i$ for every $i = 1, 2, \dots, m$, then Z indeed satisfies the condition (3).

Finally, we prove (ii). Since an orbit of every element in a minimal dynamical system is dense, it remains to show that if we have $Z = \{z_1, \dots, z_m\}$ as in condition 3. and $Z_1, Z_2, \dots, Z_m$ are all minimal subsystems of $\overline{Gx}$, then $Z \cap Z_i \neq \varnothing$ for every $i = 1, 2, \dots, m$. Fix $k \in \{1, 2, \dots, m\}$ and $\varepsilon > 0$. Since $N(x, Z_k^\varepsilon)$ is thick (Lemma 3.2) and $N(x, Z^\varepsilon)$ is syndetic, we have $N(x, Z_k^\varepsilon) \cap N(x, Z^\varepsilon) \neq \varnothing$ (by Lemma 3.1). Take $g \in N(x, Z_k^\varepsilon) \cap N(x, Z^\varepsilon)$. Then there exists $l \in \{2, \dots, m\}$ such that $\rho(gx, z_l) < \varepsilon$ and $\operatorname{dist}(gx, Z_k) < \varepsilon$. This implies $\operatorname{dist}(Z_k, z_l) < 2\varepsilon$. But ε can be arbitrarily small and Z_k is closed, hence $\operatorname{dist}(Z_k, Z) = 0$. $\square$

Example: application of Theorem 3.1.

Example 3.2 (Application of Theorem 3.1)**.** We consider an interesting example that shows how useful Theorem 3.1 might be. Consider the sequence $x_{\mathbb{P}} \in \{0,1\}^{\mathbb{Z}}$ given as

$$x_{\mathbb{P}}(k) = \begin{cases} 1 & \text{if } |k| \text{ is prime,} \\ 0 & \text{otherwise.} \end{cases}$$

We show that the subsystem $\overline{\mathbb{Z}x_{\mathbb{P}}}$ is proximal and has only one minimal subsystem: $\{0^\infty\}$.

Note that the subsystem $\overline{\mathbb{Z}x_{\mathbb{P}}}$ is rather complicated and tricky to "understand". Indeed, even the question of whether $0^\infty 1010^\infty$ belongs to $\overline{\mathbb{Z}x_{\mathbb{P}}}$ is a hard problem in Number Theory, intimately related to the Twin Prime Conjecture.

Using Theorem 3.1 our problem of showing that $\overline{\mathbb{Z}x_{\mathbb{P}}}$ is proximal boils down to proving that for every $\varepsilon > 0$ the set $N(x_{\mathbb{P}}, \{0^\infty\}^\varepsilon)$ is syndetic, which in turn reduces quite simply to proving the following number-theoretic lemma:

Lemma 3.3 (Large, frequent gaps between primes)**.** *For every number $k \in \mathbb{N}$ there exists $n \in \mathbb{N}$ such that every interval of natural numbers of length n contains k consecutive composite numbers.*

Proof. Fix $k \in \mathbb{N}$ and let $N := k + 1$. We claim that $n = 2N!$ is enough. To this end observe that each number of the form

$$lN! + j$$

for $l \in \mathbb{N}_{>0}$ and $j = 2, 3, \dots, N$ is composite. $\square$

Discussion: values attainable by $m(x)$

We first justify that for any $x \in X$ we have $m(x) > 0$.

Lemma 3.4. *For every dynamical system (X, G), there exists a minimal subsystem.*

Proof. We use Zorn's Lemma. Let

$$P = \{Y \subseteq X : (Y, G) \text{ is a subsystem of } (X, G)\} .$$

Then P is partially ordered by inclusion. Clearly, $P \neq \varnothing$ since $X \in P$. Now, take any non-empty descending chain $L \subseteq P$ and denote $Y_0 \stackrel{\text{def}}{=} \bigcap L$. Notice first, that L is a descending family of non-empty compact sets, hence by the Cantor's intersection theorem, we have $Y_0 \neq \varnothing$. Moreover, as the intersection of a family of closed sets, Y_0 is closed. We justify that it is also G-invariant. Let $y \in Y_0$ and $g \in G$. Then $y \in Y$ for every $Y \in L$, which implies $gy \in Y$ for every $Y \in L$. Hence $gy \in Y_0$. That means that the family $(P, \subseteq)$ satifies the assumptions of Zorn's Lemma. Therefore there exists a minimal element in P. ◻

Now we show that for every $m \in \mathbb{N}\backslash\{0\}$ there exists $x \in \{0,1\}^{\mathbb{Z}}$ such that $m(x) = m$. Denote $[n] = \{1, 2, \ldots, n\}$ for $n \in \mathbb{N}$ and let $[a, b]$ be the interval in $\mathbb{Z}$, that is, $[a, b] := [a, b] \cap \mathbb{Z}$ for $a, b \in \mathbb{Z}$.

Fix $m \in \mathbb{N}\backslash\{0\}$. For $i = 1, 2, \ldots, m$ let a word $w_i \in \{0,1\}^{[i]}$ be given by

$$w_i \stackrel{\text{def}}{=} 10^{i-1}.$$

Next define sequences $x, z_1, z_2, \ldots, z_m \in \{0,1\}^{\mathbb{Z}}$ as[2]

$$\begin{aligned} z_i &\stackrel{\text{def}}{=} \ldots w_i w_i . w_i w_i \ldots \qquad \text{for } i = 1, 2, \ldots, m, \\ x &\stackrel{\text{def}}{=} \ldots w_1^2 w_2^2 \ldots w_m^2 w_1 w_2 \ldots w_m . w_1 w_2 \ldots w_m w_1^2 w_2^2 \ldots w_m^2 w_1^3 w_2^3 \ldots w_m^3 \ldots \end{aligned}$$

Clearly, for every $i = 1, 2, \ldots, m$ one has $\overline{\mathbb{Z}z_i} \subseteq \overline{\mathbb{Z}x}$ and $\overline{\mathbb{Z}z_i}$ is minimal (since z_i is periodic). Moreover, if $i \neq j$ then $\overline{\mathbb{Z}z_i} \neq \overline{\mathbb{Z}z_j}$, simply because $\overline{\mathbb{Z}z_i}$ consists of i elements and $\overline{\mathbb{Z}z_j}$ consists of j elements. Hence $m(x) \geqslant m$.

To prove that $\overline{\mathbb{Z}z_i}$'s are the only minimal subshifts, we use Theorem 3.1 with $Z = \{z_1, z_2, \ldots, z_m\}$. Fix $\varepsilon > 0$ and $N \in \mathbb{N}$ satisfying $\frac{1}{N} < \varepsilon$. We justify that the set $N(x, Z^{\varepsilon})$ is syndetic. For this, we need to show that there exists $k \in \mathbb{N}$ such that for every $n \in \mathbb{Z}$ one can find $j \in [n, n+k]$ and $l = 1, 2, \ldots, m$ such that

$$x_{[j-N, j+N]} = (z_l)_{[-N,N]}.$$

In other words, we need to find $k \in \mathbb{N}$ such that in every interval $I \subseteq \mathbb{Z}$ of length k, there exists an interval $J \subseteq I$ of length $2N+1$ such that $x_J = (z_l)_{[-N,N]}$ for some $l = 1, 2, \ldots, m$. Note that for that purpose it is enough to guarantee that in every interval of length k there exists a subword consisting of $2N+1$ consecutive repetitions of some w_i (for any $i \in [m]$). To achieve this, it suffices to pick k to have length at least:

$$2 \cdot \left| w_1 w_2 \ldots w_m w_1^2 w_2^2 \ldots w_m^2 \ldots w_1^{2N+1} w_2^{2N+1} \ldots w_m^{2N+1} \right| ,$$

to be able to cover the whole "middle part" of x where there can be less than $2N+1$ consecutive repetitions of w_i's. Numerically,

$$k \stackrel{\text{def}}{=} 2 \cdot m \cdot (2N+1)^2$$

is clearly sufficient. Therefore $m(x) = m$ holds.

Finally, note that $\{0,1\}^{\mathbb{Z}}$ is transitive and it has infinitely many minimal subshifts, hence there exists $x \in \{0,1\}^{\mathbb{Z}}$ such that $m(x) = \infty$.

[2]For $x \in \{0,1\}^{\mathbb{Z}}$, we separate the $x(i)$ with $i \geqslant 0$ from those with $i < 0$ with a "decimal point".

3.3 Proximal Subsystems

In this section we use Theorem 3.1 to give a sufficient condition for a dynamical system to be proximal.

Lemma 3.5. *If a dynamical system (X, G) has a fixed point $z \in X$ that is the unique minimal subsystem, then (X, G) is proximal.*

Proof. Fix $\varepsilon > 0$. Observe first that if $x \in X \backslash \{z\}$, then $\overline{Gx}$ is a subsystem of X such that $\{z\}$ is its unique minimal subsystem. Moreover, by Theorem 3.1 part (i) the set $N(x, \{z\}^\varepsilon)$ is syndetic and by Lemma 3.2 it is thick.

We claim that for any given $x, y \in X$ there exists $g \in G$ such that $\rho(gx, gy) < 2\varepsilon$. If either x or y is equal to z then the claim holds trivially, because $z \in \overline{Gx}$ and z is a fixed point. We can thus assume $x, y \in X \backslash \{z\}$ and $x \neq y$. Since every syndetic set has a nonempty intersection with every thick set (see Lemma 3.1), there exists an element $g \in G$ such that

$$g \in N(x, \{z\}^\varepsilon) \cap N(y, \{z\}^\varepsilon).$$

Then $\rho(gx, z) \leqslant \varepsilon$ and $\rho(gy, z) \leqslant \varepsilon$, which implies $\rho(gx, gy) \leqslant 2\varepsilon$. □

Note that in the case when $G = \mathbb{Z}$, the converse of Lemma 3.5 is also true (see [1]). It is known that in general, if a dynamical system is proximal, then it has a unique minimal subsystem, but it does not need to be a singleton (see [12]). We present below an easy proof of this fact.

Lemma 3.6. *If (X, G) is proximal, then it has a unique minimal subsystem.*

Proof. Assume for the sake of contradiction that N_1 and N_2 are distinct minimal subsystems of X. Clearly, N_1 and N_2 are disjoint and closed, hence $\varepsilon := \operatorname{dist}(N_1, N_2) > 0$. Let $x \in N_1$ and $z \in N_2$. Then $\rho(gx, gz) \geqslant \varepsilon$ for every $g \in G$. This implies $\inf_{g \in G} \rho(gx, gz) \geqslant \varepsilon$ which is a contradiction with proximality of (X, G). □

Lemmas 3.5 and 3.6 lead to a natural question: Is it always true that the unique minimal subsystem for a proximal action is a singleton of a fixed point? This inspired a definition of strongly amenable group which was introduced by Glasner in [28]. A group is called strongly amenable if every its proximal action on a compact space has a fixed point. For example, all countable abelian groups are strongly amenable (see [28]). It is also known that if a group is not amenable, then it is not strongly amenable (see [28]). The only known example of a proximal dynamical system (X, G) without a fixed point and such that G is countable is an action of the Thompson's group presented in [32]. Unfortunately, it is not known whether the Thompson's group is amenable. One can find examples of actions of uncountable groups that are amenable but not strongly amenable in [27] and [29]. It is not clear whether all congruent monotileable groups are strongly amenable.

Chapter 4

The Weyl Pseudometric

In this chapter we introduce the Weyl pseudometric that is one of the crucial tools in our proof of a generalized Krieger's Theorem (Theorem 8.1). Given a dynamical system (X, G), the Weyl pseudometric D_W is defined on X^G and thus via a natural identification of points $x \in X$ with their trajcetories $\underline{x}_G \in X^G$ induces a pseudometric structure on X. A good example is when $X = \{0,1\}^G$ with the shift action, in which case D_W is equivalent (see Theorem 4.2) to a metric on X where the distance between $x, y \in \{0,1\}^G$ is the *upper Banach density* of the set of indices at which x and y differ, i.e., $D_W(x, y) = D^\star(\{g \in G : x(g) \neq y(g)\})$. Given that example, one might already expect that it interacts quite well with entropy, which indeed is the case, as demonstrated in Section 5.1.

Along with the Weyl pseudometric, in this chapter, we also study the related notion of Banach density as well as the Besicovitch pseudometric that, in a certain sense can be seen as a "simpler" variant of Weyl. Using these two notions we state and prove alternative formulas for the Weyl pseudometric (see Theorem 4.1, Lemma 4.8 and Lemma 4.9) that are generally useful in this thesis, but might also be of independent interest.

4.1 Densities in Amenable Groups

Before we proceed with the definition of the Weyl pseudometric let us start with a section that introduces the notion of Banach density in amenable groups. This is an adaptation of the familiar number-theoretic notion of density of a subset of $\mathbb{N}$ adjusted to our setting. This notion of density is then useful to define a handy, equivalent formula for the Weyl pseudometric that we actually work with. As a conclusion of this section we show a simple (yet surprisingly tricky to prove) formula for the Banach density of a special family of sets in G that can be thought of as "periodic".

We start by introducing the basic notation. Recall that by $\mathrm{Fin}(G)$ we denote the collection of all finite non-empty subsets of G and by $\mathcal{P}(A)$ we mean the power set of a set A, i.e., the collection of all its subsets. We use $\mathcal{F}$ to denote a Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ in G. Whenever a $\sup_{\mathcal{F}}$ appears in a formula, it is meant to be the supremum over all Følner seqences in G.

Definition 4.1 (Asymptotic densities and Banach densities)**.** The following formulas define the upper and lower asymptotic density of $A \subseteq G$ with respect to a Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$:

1. **(Upper asyptotic density)** $\bar{d}_{\mathcal{F}}(A) \stackrel{\text{def}}{=} \limsup_{n\to\infty} \dfrac{|A \cap F_n|}{|F_n|}$,
2. **(Lower asyptotic density)** $\underline{d}_{\mathcal{F}}(A) \stackrel{\text{def}}{=} \liminf_{n\to\infty} \dfrac{|A \cap F_n|}{|F_n|}$.

If $\bar{d}_{\mathcal{F}}(A) = \underline{d}_{\mathcal{F}}(A)$, then we say that A has the *natural density* with respect to $\mathcal{F}$ and we write $d_{\mathcal{F}}(A) = \bar{d}_{\mathcal{F}}(A)$.

The upper and lower Banach densities of $A \subseteq G$ are given by

1. **(Upper Banach density)** $\quad D^{\star}(A) \overset{\text{def}}{=} \inf_{F \in \text{Fin}(G)} \sup_{g \in G} \dfrac{|A \cap Fg|}{|F|},$

2. **(Lower Banach density)** $\quad D_{\star}(A) \overset{\text{def}}{=} 1 - D^{\star}(G \backslash A).$

If $D_{\star}(A) = D^{\star}(A)$, then we say that A has the *Banach density* $D(A) = D^{\star}(A)$.

Notice that the above definition of Banach density does not require the group G to be amenable. Unfortunately this particular definition is not convenient to work with densities, thus we state the following alternative formulations that were first obtained in [18] and [4].

Lemma 4.1. *For $A \subseteq G$, the upper Banach density is given by:*

1. *[18, Lemma 2.9]* $\quad D^{\star}(A) = \lim_{n \to \infty} \sup_{g \in G} \dfrac{|A \cap F_n g|}{|F_n|},$
   where $\mathcal{F} = \{F_n\}_{n \in \mathbb{N}}$ is an arbitrary Følner sequence. In particular, the above limit exists and does not depend on the choice of $\mathcal{F}$;

2. *[4, Lemma 3.3]* $\quad D^{\star}(A) = \sup_{\mathcal{F}} \bar{d}_{\mathcal{F}}(A).$
   (Recall that $\sup_{\mathcal{F}}$ means that the supremum is taken over all Følner sequences in G.)

In the following, we may use these formulas without further reference.

We now derive a rather simple and intuitive formula (in Lemma 4.5) for the upper Banach density of "periodic" sets. Such sets are crucial in the proof of Theorem 8.1. We first prove some preliminary results.

Observe the following relation between upper and lower asymptotic densities:

Lemma 4.2. *If $\mathcal{F}$ is a Følner sequence and $A \subseteq G$, then $\bar{d}_{\mathcal{F}}(A) = 1 - \underline{d}_{\mathcal{F}}(G \backslash A)$.*

Proof. Fix a Følner sequence $\mathcal{F} = \{F_n\}_{n \in \mathbb{N}}$ and let $A \subseteq G$. Then

$$\begin{aligned}
\underline{d}_{\mathcal{F}}(G \backslash A) &= \liminf_{n \to \infty} \frac{|(G \backslash A) \cap F_n|}{|F_n|} \\
&= \liminf_{n \to \infty} \frac{|F_n \backslash (A \cap F_n)|}{|F_n|} \\
&= \liminf_{n \to \infty} \left(1 - \frac{|A \cap F_n|}{|F_n|}\right) \\
&= 1 - \limsup_{n \to \infty} \frac{|A \cap F_n|}{|F_n|} \\
&= 1 - \bar{d}_{\mathcal{F}}(A).
\end{aligned}$$
$\square$

Next we prove a simple lemma that is a straightforward consequence of amenability.

Lemma 4.3. *If $\mathcal{F} = \{F_n\}_{n \in \mathbb{N}}$ is a Følner sequence in G and $F \in \text{Fin}(G)$, then for any $\varepsilon > 0$ and for all n large enough, one has $|FF_n \Delta F_n| < \varepsilon |F_n|$.*

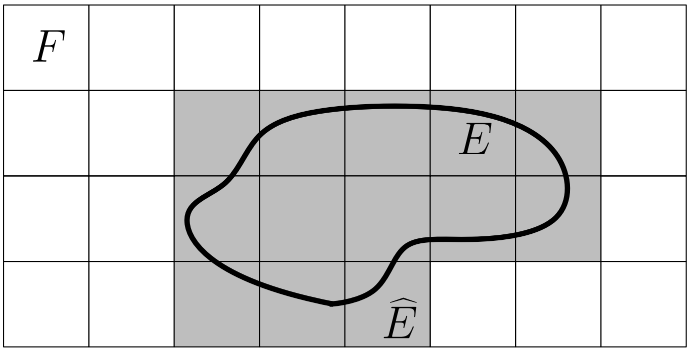

Figure 4.1: An illustration for Lemma 4.4. We can see a set $E \in \mathrm{Fin}(G)$ (the interior of the area marked with the fat line) and the set $\hat{E}$ defined in Lemma 4.4 (the grey area). The set $\hat{E}$ is the sum of shifted copies of F that intersect E (a copy of F is depicted as a single square).

Proof. Fix $\varepsilon > 0$. Since $\mathcal{F}$ is Følner, for every $f \in F$ there exists $N_f \in \mathbb{N}$ such that for every $n \geqslant N_f$ one has

$$\frac{|fF_n \Delta F_n|}{|F_n|} < \frac{\varepsilon}{|F|}.$$

Define $N \overset{\text{def}}{=} \max\{N_f : f \in F\}$. Then for every $n \geqslant N$ we have

$$\frac{1}{|F_n|}\left|\left(\bigcup_{f\in F} fF_n\right)\Delta F_n\right| \leqslant \frac{1}{|F_n|}\left|\bigcup_{f\in F}(fF_n \Delta F_n)\right| \leqslant \sum_{f\in F}\frac{|fF_n\Delta F_n|}{|F_n|} < \varepsilon. \qquad \square$$

The following technical lemma is the final preparation necessary to derive Lemma 4.5.

Lemma 4.4. *Let $F \subseteq G$ be a monotile (see Definition 2.8), $C \subseteq G$ be a set of centers for F, and let $\varepsilon > 0$. If a set $E \in \mathrm{Fin}(G)$ satisfies $|FF^{-1}E\Delta E| < \varepsilon|E|$, then for*

$$\hat{E} \overset{\text{def}}{=} \bigcup\{Fc : c \in C \textit{ is such that } Fc \cap E \neq \varnothing\}.$$

the following inequality holds

$$|\hat{E}| - |E| \leqslant \varepsilon|E|.$$

Proof. Notice that $E \subseteq FF^{-1}E$ since $e \in F$. This implies

$$|FF^{-1}E\Delta E| = |FF^{-1}E\backslash E| = |FF^{-1}E| - |E|.$$

Hence it is enough to prove that $|\hat{E}| \leqslant |FF^{-1}E|$ since $|FF^{-1}E\Delta E| < \varepsilon|E|$. We show that $\hat{E} \subseteq FF^{-1}E$. Choose $g \in \hat{E}$. Then there exists exactly one $c_g \in C$ and exactly one $f_g \in F$ such that $g = f_g c_g$. By the definition of $\hat{E}$, there exists $h \in E$ such that $h \in Fc_g$. That means that $h = f_h c_g$ for some $f_h \in F$. Combining these two observations, we obtain $g = f_g f_h^{-1} h \in FF^{-1}E$. $\square$

Finally, we are ready to prove the main technical result of this section.

Lemma 4.5. *If $F \subseteq G$ is a monotile with the associated set of centers C and $E \subseteq F$, then*

$$D^{\star}(EC) = \frac{|E|}{|F|}.$$

Proof. Let $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ be a Følner sequence in G and let $\varepsilon > 0$. For each $n \in \mathbb{N}$ we define (see Figure 4.1)

$$\widehat{F}_n \overset{\text{def}}{=} \bigcup\{Fc : c \in C \text{ satisfies } Fc \cap F_n \neq \varnothing\}.$$

By Lemma 4.3, there exists $N \in \mathbb{N}$ such that $|FF^{-1}F_n \Delta F_n| < \varepsilon|F_n|$ for every $n \geqslant N$. Then for any $n \geqslant N$ one has

$$\begin{aligned}
\frac{|EC \cap F_n|}{|F_n|} &= \frac{|EC \cap (\widehat{F}_n \backslash (\widehat{F}_n \backslash F_n))|}{|F_n|} \quad (\text{since } F_n \subseteq \widehat{F}_n) \\
&= \frac{|EC \cap \widehat{F}_n|}{|F_n|} - \frac{|EC \cap (\widehat{F}_n \backslash F_n)|}{|F_n|} \\
&\geqslant \frac{|EC \cap \widehat{F}_n|}{|F_n|} - \frac{|\widehat{F}_n| - |F_n|}{|F_n|} \\
&\geqslant \frac{|EC \cap \widehat{F}_n|}{|F_n|} - \frac{\varepsilon|F_n|}{|F_n|} \quad (\text{by Lemma 4.4}) \qquad (4.1) \\
&\geqslant \frac{|EC \cap \widehat{F}_n|}{|\widehat{F}_n|} - \varepsilon \\
&= \frac{|E||\{c \in C : Fc \cap F_n \neq \varnothing\}|}{|F||\{c \in C : Fc \cap F_n \neq \varnothing\}} - \varepsilon \\
&= \frac{|E|}{|F|} - \varepsilon.
\end{aligned}$$

Since ε was arbitrary and the above inequality holds for all sufficiently large n, we obtain

$$\bar{d}_{\mathcal{F}}(EC) \geqslant \underline{d}_{\mathcal{F}}(EC) \geqslant \frac{|E|}{|F|}.$$

Replacing E by $E' := F \backslash E$, we obtain

$$\underline{d}_{\mathcal{F}}(E'C) \geqslant \frac{|E'|}{|F|}.$$

Since $G \backslash E'C = EC$, using the relation between upper and lower asymptotic density (see Lemma 4.2), we obtain

$$\bar{d}_{\mathcal{F}}(EC) = 1 - \underline{d}_{\mathcal{F}}(E'C) \leqslant 1 - \frac{|E'|}{|F|} = \frac{|E|}{|F|}.$$

Hence

$$\bar{d}_{\mathcal{F}}(EC) = \frac{|E|}{|F|}.$$

Since $\mathcal{F}$ was arbitrary, taking the supremum over all Følner sequences, one has

$$D^\star(EC) = \frac{|E|}{|F|}. \qquad \square$$

As a corollary from the above lemma we obtain the following

Lemma 4.6. *Let $F \subseteq G$ be a monotile with associated set of centers $C \subseteq G$. If $A \subseteq G$ and $B \subseteq G$ are such that $A = E_1C$ and $B = E_2C$ for some $E_1, E_2 \subseteq F$, then*

1. $A \subseteq B$ *implies* $D^\star(A\backslash B) = D^\star(A) - D^\star(B)$,

2. $A \cap B = \varnothing$ *implies* $D^\star(A \cup B) = D^\star(A) + D^\star(B)$.

Proof. Clearly $A \subseteq B$ is equivalent to $E_1 \subseteq E_2$, and $A \cap B = \varnothing$ is equivalent to $E_1 \cap E_2 = \varnothing$, hence by Lemma 4.5 we have:

1. $A \subseteq B$ implies $D^\star(A\backslash B) = \dfrac{|E_1 \backslash E_2|}{|F|} = \dfrac{|E_1|}{|F|} - \dfrac{|E_2|}{|F|} = D^\star(A) - D^\star(B)$,

2. $A \cap B = \varnothing$ implies $D^\star(A \cup B) = \dfrac{|E_1 \cup E_2|}{|F|} = \dfrac{|E_1|}{|F|} + \dfrac{|E_2|}{|F|} = D^\star(A) + D^\star(B)$. □

4.2 The Weyl Pseudometric

In this section we introduce the Weyl pseudometric and establish its properties.

Definition 4.2 (Weyl pseudometric)**.** Given a Følner sequence $\mathcal{F} = \{F_n\}_{n=1}^{\infty}$, for $\underline{x}, \underline{z} \in X^G$ we define the (right) **Weyl pseudometric** as

$$D_W(\underline{x}, \underline{z}) \overset{\text{def}}{=} \limsup_{n\to\infty} \sup_{g\in G} \frac{1}{|F_n|} \sum_{f \in F_n g} \rho(\underline{x}(f), \underline{z}(f)).$$

The Weyl pseudometric induces a pseudometric on (X, G), namely given $x, z \in X$ we set

$$D_W(x, z) \overset{\text{def}}{=} D_W(\underline{x}_G, \underline{z}_G),$$

that is, $D_W(x, z)$ is determined by the Weyl pseudodistance between trajectories of x and z. We call the convergence in X induced by an action (X, G) and D_W the **quasi-uniform convergence**[1].

Note that by Lemma 4.9 the above formula does not depend on the choice of a Følner sequence $\mathcal{F}$.

Recall that a pseudometric satisfies the same axioms as a metric except that if the distance between x and y is zero, then may still be true $x \neq y$. Nevertheless, any pseudometric space naturally gives rise to a metric space on the quotient space (i.e., where all points which are at 0-pseudodistance from each other are identified). In our study, the fact that we work with pseudometrics and not metrics is never relevant. It is not hard to see that sequences $x, y \in \mathscr{A}^G$ that differ only on a finite set of indices are actually at pseudodistance 0 from each other, more examples appear later in this section.

We recall the notion of uniform continuity. Let Y and Z be spaces equipped with pseudometrics p_Y and p_Z respectively. For a function $\tau : (Y, p_Y) \to (Z, p_Z)$ we say that it is **uniformly continuous** if there exists a function $\Delta\colon \mathbb{R}_+ \to \mathbb{R}_+$ such that for every $\varepsilon > 0$ and for every $y_1, y_2 \in Y$ satisfying $p_Y(y_1, y_2) < \Delta(\varepsilon)$ one has $p_Z(\tau(y_1), \tau(y_2)) < \varepsilon$. We call such a function a **modulus of continuity** of τ. Having this, we can define the uniform equivalence of two (pseudo)metrics.

Definition 4.3 (Uniform equivalence)**.** Two pseudometrics D_1 and D_2 on X^G are **uniformly equivalent** if both identity functions $\mathrm{id}\colon (X^G, D_1) \to (X^G, D_2)$ and $\mathrm{id}\colon (X^G, D_2) \to (X^G, D_1)$ are both uniformly continuous.

We now state a technical, yet interesting result asserting that D_W can be expressed (up to a uniform equivalence) via Banach density. This form of D_W turns out especially convenient for proofs.

[1] The notion of *quasi-uniform convergence* for the case $G = \mathbb{Z}$ was introduced in [35], and then it was studied in [19], [17] and [40] among others.

Theorem 4.1. *The D_W pseudometric on X^G is uniformly equivalent to D'_W given by*

$$D'_W(\underline{x}, \underline{z}) \stackrel{\text{def}}{=} \inf\{\varepsilon > 0 \,:\, D^\star(\{f \in G \,:\, \rho(\underline{x}(f), \underline{z}(f)) > \varepsilon\}) < \varepsilon\}.$$

To prove this theorem it is helpful to introduce first the notion of the Besicovitch pseudometric. This way of measuring distance between trajectories is motivated by a notion used by Besicovitch in his study of almost periodic functions. The Besicovitch pseudometric appeared also in [2, 25, 26, 50].

Definition 4.4 (Besicovitch pseudometric)**.** The **Besicovitch pseudometric** for $\underline{x}, \underline{z} \in X^G$ along a Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ is defined as

$$D_{B,\mathcal{F}}(\underline{x}, \underline{z}) \stackrel{\text{def}}{=} \limsup_{N\to\infty} \frac{1}{|F_N|} \sum_{g\in F_N} \rho(\underline{x}(g), \underline{z}(g)).$$

Given a dynamical system (X, G) and $x, z \in X$ we define $D_{B,\mathcal{F}}(x, z)$ to be equal the $D_{B,\mathcal{F}}$-distance between trajectories of x and z, i.e.

$$D_{B,\mathcal{F}}(x, z) \stackrel{\text{def}}{=} D_{B,\mathcal{F}}(\underline{x}_G, \underline{z}_G).$$

One can notice the similarity between Besicovitch and Weyl pseudometrics, yet a formal relation between these two will be proved at a later point (see Lemma 4.8).

We introduce more terminology required to state the results of this section. Let Y and Z be spaces equipped with pseudometrics p_Y and p_Z respectively. If $\mathcal{K}$ is a family of functions $\tau : (Y, p_Y) \to (Z, p_Z)$ and there exists a common modulus of continuity Δ of them, then we call this family **uniformly equicontinuous**. In such a case Δ is called the **modulus of equicontinuity** of $\mathcal{K}$. Analogously we define a **modulus of uniform equivalence** of (pseudo)metrics.

The following lemma (along with its proof below) is adapted from [40] (Lemma 2 therein), where the case of $G = \mathbb{Z}$ was studied

Lemma 4.7. *Fix a Følner sequence $\mathcal{F}$ and for $\underline{x}, \underline{z} \in X^G$ define*

$$D'_{B,\mathcal{F}}(\underline{x}, \underline{z}) \stackrel{\text{def}}{=} \inf\{\delta > 0 : \bar{d}_{\mathcal{F}}(\{g \in G : \rho(\underline{x}(g), \underline{z}(g)) \geqslant \delta\}) < \delta\}. \tag{4.2}$$

Then $D_{B,\mathcal{F}}$ and $D'_{B,\mathcal{F}}$ are uniformly equivalent on X^G. Moreover, the modulus of uniform equivalence does not depend on the choice of a Følner sequence.

Proof. First we show that $\mathrm{id}\colon (X^G, D'_{B,\mathcal{F}}) \to (X^G, D_{B,\mathcal{F}})$ is uniformly continuous. For $\underline{x}, \underline{z} \in X^G$ and $\delta > 0$ define

$$J_\delta(\underline{x}, \underline{z}) \stackrel{\text{def}}{=} \{g \in G : \rho(\underline{x}(g), \underline{z}(g)) \geqslant \delta\}.$$

Notice that for every $g \in G$ one has[2], (recall that $\rho(x, z) \leqslant 1$ for $x, z \in X$)

$$\delta \mathbb{1}_{J_\delta}(g) \leqslant \rho(\underline{x}(g), \underline{z}(g)) \leqslant \mathbb{1}_{J_\delta}(g) + \delta,$$

which implies

$$\delta \bar{d}_{\mathcal{F}}(J_\delta(\underline{x}, \underline{z})) \leqslant D_{B,\mathcal{F}}(\underline{x}, \underline{z}) \leqslant \bar{d}_{\mathcal{F}}(J_\delta(\underline{x}, \underline{z})) + \delta. \tag{4.3}$$

Moreover,

$$D'_{B,\mathcal{F}}(\underline{x}, \underline{z}) < \delta \quad \text{if and only if} \quad \bar{d}_{\mathcal{F}}(J_\delta(\underline{x}, \underline{z})) < \delta. \tag{4.4}$$

[2]We denote by $\mathbb{1}_A$ the indicator function of a set A.

Fix $\varepsilon > 0$ and choose $\delta \in \left(0, \frac{\varepsilon}{2}\right)$. It follows from the second inequality in (4.3) and (4.4) that $D'_{B,\mathcal{F}}(\underline{x}, \underline{z}) < \delta$ implies $D_{B,\mathcal{F}}(\underline{x}, \underline{z}) < \varepsilon$. This yields uniform continuity of $\mathrm{id}\colon (X^G, D'_{B,\mathcal{F}}) \to (X^G, D_{B,\mathcal{F}})$. The modulus of uniform continuity is for example a function $\Delta\colon \mathbb{R}_+ \to \mathbb{R}_+$ given by $\Delta(\varepsilon) = \frac{\varepsilon}{2}$.

Now, we prove that $\mathrm{id}\colon (X^G, D_{B,\mathcal{F}}) \to (X^G, D'_{B,\mathcal{F}})$ is also uniformly continuous. To this end, fix $\varepsilon > 0$ and $\delta \in (0, \varepsilon^2)$. Take any pair $\underline{x}, \underline{z} \in X^G$ such that $D_{B,\mathcal{F}}(\underline{x}, \underline{z}) < \delta$. Use the first inequality in (4.3) to see that

$$\varepsilon \cdot \bar{d}_{\mathcal{F}}(J_\varepsilon(\underline{x}, \underline{z})) \leqslant D_{B,\mathcal{F}}(\underline{x}, \underline{z}).$$

Therefore, $D_{B,\mathcal{F}}(\underline{x}, \underline{z}) < \varepsilon^2$ implies $\bar{d}(J_\varepsilon(\underline{x}, \underline{z})) < \varepsilon$. Hence by (4.4) we obtain $D'_{B,\mathcal{F}}(\underline{x}, \underline{z}) < \varepsilon$ and $\Delta(\varepsilon) = \varepsilon^2$ as a modulus of uniform continuity. □

Corollary 4.1. *Let $\tilde{\rho}$ be a metric on X equivalent to ρ (i.e., ρ and $\tilde{\rho}$ generate the same topology on X) and $\tilde{D}_{B,\mathcal{F}}$ be defined as $D_{B,\mathcal{F}}$ above with $\tilde{\rho}$ in place of ρ. Then $\tilde{D}_{B,\mathcal{F}}$ and $D_{B,\mathcal{F}}$ are uniformly equivalent on X^G.*

The next lemma formally captures the close relation between Besicovitch and Weyl pseudometrics.

Lemma 4.8. *For every $\underline{x}, \underline{z} \in X^G$ one has*

$$D_W(\underline{x}, \underline{z}) = \sup_{\mathcal{F}} D_{B,\mathcal{F}}(\underline{x}, \underline{z}).$$

Proof. First notice that if $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ is a Følner sequence in G and $\{g_n\}_{n\in\mathbb{N}} \subseteq G$, then the sequence $\{F_n g_n\}_{n\in\mathbb{N}}$ is also Følner. Therefore

$$D_W(\underline{x}, \underline{z}) \leqslant \sup_{\mathcal{F}} D_{B,\mathcal{F}}(\underline{x}, \underline{z}) \quad \text{ for } \underline{x}, \underline{z} \in X^G.$$

Now we prove the reverse inequality. Let $\underline{x}, \underline{z} \in X^G$. Notice that it is enough to show

$$\inf_{K\in\mathrm{Fin}(G)} \sup_{g\in G} \frac{1}{|K|} \sum_{f\in Kg} \rho(\underline{x}(f), \underline{z}(f)) \geqslant \sup_{\mathcal{F}} D_{B,\mathcal{F}}(\underline{x}, \underline{z}).$$

Denote $\alpha \stackrel{\text{def}}{=} \sup_{\mathcal{F}} D_{B,\mathcal{F}}(\underline{x}, \underline{z})$ and assume $\alpha > 0$. Then we need to show that for every finite set $K \subseteq G$ there exists $g \in G$ such that for every $\beta \in (0,1)$ with $\alpha > \beta$ one has

$$\frac{1}{|K|} \sum_{f\in Kg} \rho(\underline{x}(f), \underline{z}(f)) > \beta. \tag{4.5}$$

Fix a finite set $K \subseteq G$. Choose a Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ in G and an increasing sequence of indices $\{k_n\}_{n\in\mathbb{N}} \subseteq \mathbb{N}$ such that for every $n \in \mathbb{N}$ one has

$$\frac{1}{|F_{k_n}|} \sum_{f\in F_{k_n}} \rho(\underline{x}(f), \underline{z}(f)) > \frac{\beta+\alpha}{2}. \tag{4.6}$$

Since K is finite, for $n \in \mathbb{N}$ large enough and for every $h \in K$ one has (see Lemma 4.3)

$$\frac{|F_n \backslash hF_n|}{|F_n|} \leqslant \frac{|hF_n \Delta F_n|}{|F_n|} < \frac{\alpha-\beta}{2}.$$

Consequently, for all $n \in \mathbb{N}$ large enough and any $h \in K$ we have (for the second inequality we use (4.6) and the fact that $\rho(x,z) \leqslant 1$ for $x, z \in X$)

$$\frac{1}{|F_{k_n}|} \sum_{f\in hF_{k_n}} \rho(\underline{x}(f), \underline{z}(f)) \geqslant \frac{1}{|F_{k_n}|} \sum_{f\in F_{k_n}} \rho(\underline{x}(f), \underline{z}(f)) - \frac{1}{|F_{k_n}|} \sum_{f\in F_{k_n}\backslash hF_{k_n}} \rho(\underline{x}(f), \underline{z}(f)) > \beta.$$

Therefore

$$\sum_{g\in F_{k_n}}\sum_{f\in Kg}\rho(\underline{x}(f),\underline{z}(f))=\sum_{h\in K}\sum_{f\in hF_{k_n}}\rho(\underline{x}(f),\underline{z}(f))>|F_{k_n}||K|\beta.$$

This means that average value of the sum $\sum_{f\in Kg}\rho(\underline{x}(f),\underline{z}(f))$ over $g\in F_{k_n}$ is greater than $|K|\beta$. Thus there exists $g\in F_{k_n}$ such that (4.5) holds. □

The proof of Theorem 4.1 is now a straightforward consequence of the above considerations.

Proof of Theorem 4.1. Combining Lemma 4.7 with Lemma 4.8 we obtain the claim. □

At this point we offer some examples that help in understanding the relation between the notions defined so far.

Example 4.1 (Distance in the Besicovitch pseudeometric depends on the choice of a Følner seqence.)**.** We construct $x\in\{0,1\}^{\mathbb{Z}}$ as follows: recall that for $n\in\mathbb{N}$ by s^n we denote the word of length n consisting of n symbols $s\in\{0,1\}$, that is:

$$s^n\overset{\text{def}}{=}\underbrace{sss\ldots sss}_{n}.$$

Define[3]

$$x=\ldots1^30^31^20^21.010^21^20^31^30^41^4\ldots$$

Now if we choose a Følner sequence such that its consequtive elements correspond to longer and longer segments of zeros, i.e.,

$$F_0=\{0\},\quad F_1=\{2,3\},\quad F_2=\{6,7,8\},\quad F_3=\{12,13,14,15\},\quad\ldots,$$

then $D_{B,\mathcal{F}}(x,0^{\mathbb{Z}})=0$. If we instead take a Følner sequence such that its consequtive elements correspond to longer and longer segments of ones, i.e.,

$$F_0=\{1\},\quad F_1=\{4,5\},\quad F_2=\{9,10,11\},\quad F_3=\{16,17,18,19\},\quad\ldots,$$

then we obtain $D_{B,\mathcal{F}}(x,0^{\mathbb{Z}})=1$. That is not the end of our options. Now take a Følner sequence such that the first half of its consequtive elements correspond to longer and longer segments of zeros and the second half corresponds to longer and longer segments of ones, i.e.,

$$F_0=\{0,1\},\quad F_1=\{2,3,4,5\},\quad F_2=\{6,7,8,\ldots,11\},\quad F_3=\{12,13,14,\ldots,19\},\quad\ldots,$$

then we have $D_{B,\mathcal{F}}(x,0^{\mathbb{Z}})=\frac{1}{2}$. Notice that for any $\alpha\in[0,1]$, we can choose a Følner sequence such that $D_{B,\mathcal{F}}(x,0^{\mathbb{Z}})=\alpha$.

Example 4.2 (Distinct elements with zero distance in the Weyl pseudometric)**.** Define $x\in\{0,1\}^{\mathbb{Z}}$ as follows: put zero on every position except 2^n, where we put one, that is

$$x(n)\overset{\text{def}}{=}\begin{cases}1 & \text{if } n=2^k \text{ for some } k\in\mathbb{N},\\ 0 & \text{otherwise.}\end{cases}$$

It is tedious to compute from the definition that $D_W(x,0^{\mathbb{Z}})=0$. If we use Theorem 4.2, then this holds trivially.

[3]For $x\in\{0,1\}^{\mathbb{Z}}$, we separate the $x(i)$ with $i\geqslant0$ from those with $i<0$ with a "decimal point".

Finally, at the end of this section we give two additional formulas for the Weyl Pseudometric. Notice that the formula b) makes it possible to define the Weyl pseudometric even if the group G is not necessarily amenable. Unfortunately, this extension does not seem to have good properties outside amenable setting (see [58]).

Lemma 4.9. *For any $\underline{x}, \underline{z} \in X^G$ one has*

a) $D_W(\underline{x}, \underline{z}) = \lim_{n\to\infty} \sup_{g\in G} \frac{1}{|F_n|} \sum_{f\in H_n g} \rho(\underline{x}(f), \underline{z}(f))$ *for any Følner sequence* $\{F_n\}_{n\in\mathbb{N}}$,

b) $D_W(\underline{x}, \underline{z}) = \inf_{F\in \mathrm{Fin}(G)} \frac{1}{|F|} \sup_{g\in G} \sum_{f\in F} \rho(\underline{x}(fg), \underline{z}(fg))$.

To prove Lemma 4.9 we introduce some machinery from [16].

A ***k*-cover** ($k \in \mathbb{N}$) of a set $F \in \mathrm{Fin}(G)$ is a tuple $(K_1, \ldots, K_r)$ of elements of $\mathrm{Fin}(G)$ such that every element $g \in F$ belongs to at least k sets among $K_1, \ldots, K_r$, i.e., for every $g \in F$ one has $|\{i \in \{1, 2, \ldots, r\} : g \in K_i\}| \geqslant k$. We say that a function $H\colon \mathrm{Fin}(G) \cup \{\varnothing\} \to [0, \infty)$

1. satisfies **Shearer's inequality** if for any $F \in \mathrm{Fin}(G)$ and any k-cover $(K_1, \ldots, K_r)$ of F, we have
$$H(F) \leqslant \frac{1}{k}(H(K_1) + \ldots + H(K_r)),$$
2. satisfies the **infimum rule** if for any Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ one has
$$\limsup_{n\to\infty} \frac{H(F_n)}{|F_n|} = \inf_{F\in\mathrm{Fin}(G)} \frac{H(F)}{|F|}.$$
3. is ***G*-invariant** if $H(Fg) = H(F)$ for every $g \in G$ and $F \in \mathrm{Fin}(G)$,
4. is **monotone** if for all $A, B \in \mathrm{Fin}(G)$ with $A \subseteq B$ one has $H(A) \leqslant H(B)$,
5. is **subadditive** if for all $A, B \in \mathrm{Fin}(G)$ with $A \cap B = \varnothing$ one has
$$H(A \cup B) \leqslant H(A) + H(B).$$

Lemma 4.10. *If $\underline{x}, \underline{z} \in X^G$ and the function $H\colon \mathrm{Fin}(G) \to [0, \infty)$ is given by*
$$H(F) \overset{\mathrm{def}}{=} \sup_{g\in G} \sum_{f\in Fg} \rho(\underline{x}(f), \underline{z}(f)),$$
then H is G-invariant and satisfies Shearer's inequality.

Proof. Notice that G-invariance follows directly from the definition of H. To show that it satisfies Shearer's inequality take $F \in \mathrm{Fin}(G)$ and let $(K_1, \ldots, K_r)$ be a k-cover of F. Every element of F belongs to K_i for at least k indices $i \in \{1, 2, \ldots, r\}$, hence
$$\begin{aligned}\frac{1}{k}\left(H(K_1) + \ldots + H(K_r)\right) &\geqslant \frac{1}{k} \sup_{g\in G} \left(\sum_{f\in K_1 g} \rho(\underline{x}(f), \underline{z}(f)) + \ldots + \sum_{f\in K_r g} \rho(\underline{x}(f), \underline{z}(f)) \right) \\ &\geqslant \frac{1}{k} \sup_{g\in G} \left(k \sum_{f\in Fg} \rho(\underline{x}(f), \underline{z}(f)) \right) = H(F). \qquad \square\end{aligned}$$

For the proof of Lemma 4.9 we also need the following fact, that was first established in [43]. We state it here without a proof.

Lemma 4.11 ([43, Theorem 6.1])**.** *If $H\colon \operatorname{Fin}(G) \to [0,\infty)$ is G-invariant, monotone, subadditive and such that $H(\varnothing) = 0$, then for any Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ a limit of the sequence*

$$\left\{\frac{H(F_n)}{|F_n|}\right\}_{n\in\mathbb{N}}$$

exists and does not depend on the choice of a Følner sequence $\mathcal{F}$.

Proof of Lemma 4.9. Part a) is a straigtforward consequence of Lemma 4.10 and Lemma 4.11. Part b) follows from the fact that every G-invariant, non-negative function on $\operatorname{Fin}(G)$ that satisfies Shearer's inequality, obeys the infimum rule (see [16, Proposition 3.3]). $\square$

4.3 The Weyl Pseudometric in a Shift Space

The focus of this section is the case $X = \mathscr{A}^G$. In particular, we state an alternative, simpler formula for D_W in this case.

Theorem 4.2. *If $X = \mathscr{A}^G$ (with the standard action of G on X by shifts), then D_W is uniformly equivalent to the pseudometric*

$$D^\star(\underline{x}, \underline{z}) = D^\star(\{g \in G : \underline{x}(g) \neq \underline{z}(g)\}).$$

Before we proceed to the proof, let us introduce some useful notation. Given a continuous function $\phi\colon X \to [0,1]$ we lift it to $\phi : X^G \to [0,1]^G$ in a straightforward way, for $y \in X^G$ define

$$\phi(y) \overset{\text{def}}{=} \{\phi(y_g)\}_{g\in G} \in [0,1]^G.$$

For a family of continuous functions $\mathcal{K} \subseteq [0,1]^X$ and a Følner sequence $\mathcal{F}$ in G, set

$$D_W^{\mathcal{K}}(x,z) \overset{\text{def}}{=} \sup_{\phi\in\mathcal{K}} \ddot{D}_W(\phi(\underline{x}_G), \phi(\underline{z}_G)), \quad D_{B,\mathcal{F}}^{\mathcal{K}}(x,z) \overset{\text{def}}{=} \sup_{\phi\in\mathcal{K}} \ddot{D}_{B,\mathcal{F}}(\phi(\underline{x}_G), \phi(\underline{z}_G)),$$

where $\ddot{D}_W$ ($\ddot{D}_{B,\mathcal{F}}$) denotes the Weyl (Besicovitch) pseudometric on $[0,1]^G$ corresponding to the Euclidean distance on $[0,1] \subseteq \mathbb{R}$.

The proof of Theorem 4.2 is a straightforward consequence of the following general lemma.

Lemma 4.12. *Let (X,G) be a dynamical system and let $\mathcal{K} \subseteq [0,1]^X$ be a uniformly equicontinuous family of functions such that*

$$\mathcal{K}_G = \{x \mapsto \phi(gx) : \phi \in \mathcal{K},\ g \in G\}$$

separates the points of X (i.e., for any $x,z \in X$ with $x \neq z$ there exists $\psi \in \mathcal{K}_G$ such that $\psi(x) \neq \psi(z)$). Then for any Følner sequence $\mathcal{F}$ in G, the pseudometrics $D_{B,\mathcal{F}}$ and $D_{B,\mathcal{F}}^{\mathcal{K}}$ are uniformly equivalent on X. Moreover, the modulus of uniform equivalence does not depend on the choice of $\mathcal{F}$ and consequently, the pseudometrics D_W and $D_W^{\mathcal{K}}$ are also uniformly equivalent.

Proof. Fix a Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ in G. First we show that the identity map

$$\operatorname{id}\colon (X, D_{B,\mathcal{F}}^{\mathcal{K}}) \to (X, D_{B,\mathcal{F}}).$$

is uniformly continuous with a modulus of uniform continuity independent of $\mathcal{F}$. We construct a metric $\tilde{\rho} : X \times X \to [0,1]$ on X that is equivalent to ρ but is more convenient to work with. Then, by Corollary 4.1, the pseudometric

$$\tilde{D}_{B,\mathcal{F}}(\underline{x}, \underline{z}) \stackrel{\text{def}}{=} \limsup_{N\to\infty} \frac{1}{|F_N|} \sum_{g\in F_N} \tilde{\rho}(\underline{x}(g), \underline{z}(g))$$

is uniformly equivalent to $D_{B,\mathcal{F}}$ and the modulus of uniform equicontinuity does not depend on the choice of the Følner sequence. Therefore, it will be enough to prove that the map

$$\mathrm{id}\colon (X, D^{\mathcal{K}}_{B,\mathcal{F}}) \to (X, \tilde{D}_{B,\mathcal{F}})$$

is uniformly continuous.

First, we construct the metric $\tilde{\rho}$. We claim that from $\mathcal{K}^G$ we can choose a countable subset $\{\psi_n : n \in \mathbb{N}\} \subseteq \mathcal{K}_G$ separating the points of X. Indeed, for every $n \in \mathbb{N}$, the set

$$Y_n \stackrel{\text{def}}{=} \left\{(x,z) \in X \times X : \rho(x,z) \geqslant \frac{1}{n}\right\}$$

is compact (as a closed subset of a compact metric space) and satisfies

$$Y_n \subseteq \bigcup_{\psi\in\mathcal{K}_G} U_\psi, \qquad \text{where} \qquad U_\psi = \{(x,z) \in X \times X : \psi(x) \neq \psi(z)\} \ \text{ for } \psi \in \mathcal{K}_G\,.$$

The family $\{U_\psi : \psi \in \mathcal{K}_G\}$ is clearly an open cover of Y_n. Hence for every $n \in \mathbb{N}$ we can choose a finite subcover $\mathcal{K}_n \subseteq \mathcal{K}_G$ of Y_n such that the collection $\{U_\psi : \psi \in \mathcal{K}_n\}$ is a cover of Y_n. Denote $\{\psi_n\}_{n\in\mathbb{N}} \stackrel{\text{def}}{=} \bigcup\{\mathcal{K}_n : n \in \mathbb{N}\}$. Then the collection

$$\mathcal{U} \stackrel{\text{def}}{=} \{U_{\psi_n} : n \in \mathbb{N}\} \tag{4.7}$$

is a countable cover of $\{(x,z) \in X \times X : x \neq z\}$. Therefore the set $\{\psi_n\}_{n\in\mathbb{N}} \subseteq \mathcal{K}_G$ separates the points of X. Also, since for every $n \in \mathbb{N}$ the function ψ_n belongs to $\mathcal{K}^G$, let us denote by $\phi_n \in \mathcal{K}$ and $g_n \in G$ a function and a group element such that $\psi_n(x) = \phi_n(g_n x)$ for every $x \in X$.

Take any sequence of weights $\{a_n\}_{n\in\mathbb{N}} \subseteq (0,1)$ such that $\sum_{n\in\mathbb{N}} a_n < 1$. For $x, z \in X$ define

$$\tilde{\rho}(x,z) \stackrel{\text{def}}{=} \sum_{n\in\mathbb{N}} a_n |\psi_n(x) - \psi_n(z)|.$$

Notice that $\tilde{\rho}$ is a metric on X which is equivalent to ρ. Indeed, first we show the continuity of the identity map $\mathrm{id}\colon (X,\rho) \to (X,\tilde{\rho})$. Fix $\varepsilon > 0$. Then there exists $m \in \mathbb{N}$ such that

$$\sum_{n=m}^{\infty} a_n \,|\psi_n(x) - \psi_n(z)| < \frac{\varepsilon}{2} \qquad \text{for } x, z \in X. \tag{4.8}$$

Since ψ_n's are uniformly continuous, for every $k = 0, 1, \ldots, m-1$ there exists $\delta_k > 0$ such that

$$\rho(x,z) < \delta_k \quad \text{ implies } \quad |\psi_k(x) - \psi_k(z)| < \frac{\varepsilon}{2m}.$$

Denote $\delta \stackrel{\text{def}}{=} \min\{\delta_k : k = 0, 1, \ldots, m-1\}$. Then for $x, z \in X$ satisfying $\rho(x,z) < \delta$, one has

$$\sum_{n=0}^{m-1} a_n \,|\psi_n(x) - \psi_n(z)| \leqslant \sum_{n=0}^{m-1} |\psi_n(x) - \psi_n(z)| < \frac{\varepsilon}{2}. \tag{4.9}$$

Combining (4.8) with (4.9) we obtain that whenever $\rho(x,z) < \delta$, then $\tilde{\rho}(x,z) < \varepsilon$ for $x, z \in X$, which shows that $\mathrm{id}\colon (X,\rho) \to (X,\tilde{\rho})$ is continuous. Note that our choice of δ does not depend on the choice of $\mathcal{F}$.

To show that $\mathrm{id}\colon (X,\tilde{\rho}) \to (X,\rho)$ is continuous, we fix $\varepsilon > 0$ and find $\delta > 0$ such that $\rho(x,z) \geqslant \varepsilon$ implies $\tilde{\rho}(x,z) \geqslant \delta$ for $x, z \in X$. Let

$$Y_\varepsilon \overset{\text{def}}{=} \{(x,z) \in X \times X : \rho(x,z) \geqslant \varepsilon\}\,.$$

Clearly, Y_ε is compact and $Y_\varepsilon \subseteq \bigcup \mathcal{U}$ (we defined $\mathcal{U}$ in (4.7)). Hence there exists in $\mathcal{U}$ a finite subcover of Y_ε, that is a finite set $I_\varepsilon \subseteq \mathbb{N}$ such that $Y_\varepsilon \subseteq \bigcup\{U_{\psi_n} : n \in I_\varepsilon\}$. Therefore

$$\max_{n\in I_\varepsilon} |\psi_n(x) - \psi_n(z)| > 0 \quad \text{ for } (x,z) \in Y_\varepsilon.$$

Then for $x, z \in Y_\varepsilon$ one has

$$\begin{aligned}
\sum_{n\in\mathbb{N}} a_n\, |\psi_n(x) - \psi_n(z)| &\geqslant \sum_{n\in I_\varepsilon} a_n\, |\psi_n(x) - \psi_n(z)| \\
&\geqslant \max_{n\in I_\varepsilon} a_n\, |\psi_n(x) - \psi_n(z)| \\
&\geqslant \left(\min_{n\in I_\varepsilon} a_n\right) \max_{n\in I_\varepsilon} |\psi_n(x) - \psi_n(z)| > 0.
\end{aligned}$$

Hence we can choose

$$\delta \overset{\text{def}}{=} \inf_{x,z\in Y_\varepsilon} \left(\min_{n\in I_\varepsilon} a_n\right) \max_{n\in I_\varepsilon} |\psi_n(x) - \psi_n(z)|\,,$$

with the guarantee that $\delta > 0$ because Y_ε is compact and I_ε is finite.

Next we prove that the identity map $\mathrm{id}\colon (X, D^{\mathcal{K}}_{B,\mathcal{F}}) \to (X, \tilde{D}_{B,\mathcal{F}})$ is continuous. Fix $\varepsilon > 0$ and choose $m \in \mathbb{N}$ such that

$$\sum_{n=m}^{\infty} a_n < \frac{\varepsilon}{4}.$$

Let $x, z \in X$ be such that $D^{\mathcal{K}}_{B,\mathcal{F}}(x,z) < \frac{\varepsilon}{2m}$. That is, we pick $\Delta(\varepsilon) = \frac{\varepsilon}{2m}$ independently of our choice of $\mathcal{F}$. We show that then $\tilde{D}_{B,\mathcal{F}}(x,z) \leqslant \varepsilon$. Indeed,

$$\begin{aligned}
\tilde{D}_{B,\mathcal{F}}(x,z) &= \limsup_{N\to\infty} \frac{1}{|F_N|} \sum_{g\in F_N} \tilde{\rho}(gx, gz) \\
&= \limsup_{N\to\infty} \frac{1}{|F_N|} \sum_{g\in F_N} \sum_{n\in\mathbb{N}} a_n\, |\psi_n(gx) - \psi_n(gz)| \\
&\leqslant \limsup_{N\to\infty} \frac{1}{|F_N|} \sum_{g\in F_N} \left(\sum_{n=0}^{m-1} a_n |\psi_n(gx) - \psi_n(gz)| + \frac{\varepsilon}{2} \right) \\
&= \frac{\varepsilon}{2} + \limsup_{N\to\infty} \frac{1}{|F_N|} \sum_{g\in F_N} \sum_{n=0}^{m-1} a_n |\psi_n(gx) - \psi_n(gz)| \\
&\leqslant \frac{\varepsilon}{2} + \sum_{n=0}^{m-1} a_n \left(\limsup_{N\to\infty} \frac{1}{|F_N|} \sum_{g\in|F_N|} |\psi_n(gx) - \psi_n(gz)| \right) \\
&= \frac{\varepsilon}{2} + \sum_{n=0}^{m-1} a_n \ddot{D}_{B,\mathcal{F}}(\psi(\underline{x}_G), \psi_n(\underline{z}_G)).
\end{aligned} \tag{4.10}$$

To bound (4.10) notice that G-invariance of $\ddot{D}_{B,\mathcal{F}}$ implies that for $n \in \mathbb{N}$ and $x, z \in X$ one has

$$\ddot{D}_{B,\mathcal{F}}(\psi_n(\underline{x}_G), \psi_n(\underline{z}_G)) = \ddot{D}_{B,\mathcal{F}}(\phi_n(g_n\underline{x}_G), \phi_n(g_n\underline{z}_G)) = \ddot{D}_{B,\mathcal{F}}(\phi_n(\underline{x}_G), \phi_n(\underline{z}_G)).$$

Therefore

$$\ddot{D}_{B,\mathcal{F}}(\psi_n(\underline{x}_G), \psi_n(\underline{z}_G)) \leqslant D^{\mathcal{K}}_{B,\mathcal{F}}(x,z)$$

and hence we obtain

$$\sum_{n=0}^{m-1} a_n \ddot{D}_{B,\mathcal{F}}(\psi(\underline{x}_G), \psi_n(\underline{z}_G)) < \sum_{n=0}^{m-1} a_n \frac{\varepsilon}{2m} < \frac{\varepsilon}{2}.$$

Finally, we prove that the map

$$\mathrm{id}\colon (X, D_{B,\mathcal{F}}) \to (X, D^{\mathcal{K}}_{B,\mathcal{F}})$$

is also uniformly continuous. Fix $\varepsilon > 0$. We need to show that there exists $\delta > 0$ such that for $x, z \in X$ one has

$$D_{B,\mathcal{F}}(x,z) \leqslant \delta \quad \Longrightarrow \quad \forall \phi \in \mathcal{K} \ \ \ddot{D}_{B,\mathcal{F}}(\phi(\underline{x}_G), \phi(\underline{z}_G)) \leqslant \varepsilon.$$

Moreover, our choice of δ should be independent of the choice of $\mathcal{F}$.

By Lemma 4.7 we know that $\ddot{D}_{B,\mathcal{F}}$ and $\ddot{D}'_{B,\mathcal{F}}$ are uniformly equivalent on $[0,1]^G$, where $\ddot{D}'_{B,\mathcal{F}}$ denotes the pseudometric on $[0,1]^G$ corresponding to the Euclidean distance on $[0,1]$ given by (4.2). Hence it is enough to prove that there exists $\delta > 0$ such that for $x, z \in X$ one has

$$D_{B,\mathcal{F}}(x,z) \leqslant \delta \quad \Longrightarrow \quad \forall \phi \in \mathcal{K} \ \ \ddot{D}'_{B,\mathcal{F}}(\phi(\underline{x}_G), \phi(\underline{z}_G)) \leqslant \varepsilon.$$

Rewriting it further we obtain that we need to show that there exists $\delta > 0$ such that for $x, z \in X$ and $\phi \in \mathcal{K}$ one has

$$\forall n \gg 0 \ \ \frac{1}{|F_n|} \sum_{g \in F_n} \rho(gx, gz) \leqslant \delta \quad \Longrightarrow \quad \forall n \gg 0 \ \ \frac{|\{g \in F_n : |\phi(gx) - \phi(gz)| \geqslant \varepsilon\}|}{|F_n|} \leqslant \varepsilon,$$

where "$\forall n \gg 0$" denotes "for all but finitely many n". Since $\mathcal{K}$ is uniformly equicontinuous, there exists $\delta \in (0, \varepsilon)$ such that for every $x, z \in X$ and for every $\phi \in \mathcal{K}$ we have that $\rho(x,y) < \delta^2$ implies $|\phi(x) - \phi(y)| < \varepsilon$. Now notice that the inequality

$$\frac{1}{|F_n|} \sum_{g \in F_n} \rho(gx, gz) \leqslant \delta^2 \tag{4.11}$$

means that the expected value of $\rho(gx, gz)$ when g is chosen uniformly at random from F_n is smaller than δ^2. Hence by the Markov inequality we obtain

$$\frac{|\{g \in F_n : \rho(gx, gz) \geqslant \delta\}|}{|F_n|} \leqslant \delta,$$

equivalently

$$\frac{|\{g \in F_n : \rho(gx, gz) < \delta\}|}{|F_n|} > 1 - \delta.$$

Hence, by the choice of δ, we obtain

$$\frac{|\{g \in F_n : |\phi(gx) - \phi(gz)| < \varepsilon\}|}{|F_n|} > 1 - \delta > 1 - \varepsilon.$$

Finally, we have

$$\frac{|\{g \in F_n : |\phi(gx) - \phi(gz)| \geqslant \varepsilon\}|}{|F_n|} \leqslant \varepsilon.$$

Since the modulus of uniform equivalence does not depend on the choice of $\mathcal{F}$, the second part of the theorem also holds. $\square$

Proof of Theorem 4.2. Observe that any finite family $\mathcal{K}$ is uniformly equicontinuous. Then taking $\mathcal{K} = \{\iota_e\}$, where $\iota_e(\underline{x}_G) = x_e$, and applying it to Lemma 4.12, we obtain the claim. $\square$

Chapter 5

Quasi-uniform Convergence

In this chapter we continue the study of the Weyl pseudometric defined in Chapter 4. More specifically, we investigate how the convergence in this pseudometric, i.e., quasi-uniform convergence, interacts with several important objects in ergodic theory such as topological entropy, minimal subsystems and invariant measures. These three are studied in the subsequent Sections 5.1, 5.2, 5.3. The result of Section 5.1 that the entropy of the subsystem $\overline{Gx}$ varies continuously with x (in the case of shift spaces) is a major component in our proof of Theorem 8.1. The results of the remaining two sections are not explicitly used in other chapters but still provide an insight into the topology that the Weyl pseudometric induces and thus are interesting on their own.

5.1 Topological Entropy

The purpose of this section is to study the interaction between the entropy function and the Weyl pseudometric. The results and proofs here presented are close adaptations of [19, Section 5]. We divide this section into two subsections 5.1.1 and 5.1.2 in which we study the general case of arbitrary dynamical systems and the special case of shift spaces, for which our result is stronger.

5.1.1 General Dynamical Systems

We start by stating the general result that we intend to prove in this subsection.

Theorem 5.1. *For a dynamical system (X, G), the function*

$$h\colon (X, D_W) \to \mathbb{R}_+ \cup \{\infty\} \qquad \textit{satisfying} \qquad h(x) \stackrel{\text{def}}{=} h_{\text{top}}(\overline{Gx})$$

is lower semicontinuous, that is, for every $x \in X$ and $\eta > 0$, there exists $\delta > 0$ such that for every $y \in X$ satisfying $D_W(x, y) < \delta$, it holds $h(y) \geqslant h(x) - \eta$.

The first question that arises given the above theorem is whether semicontinuity is the best one can hope for here. The example below shows that in general the function h is not continuous, on the other hand, in Subsection 5.1.2 we argue that in the case of shift spaces h is continuous.

Example 5.1. In [19, Example 3] it is shown that the function $(X, D_W) \ni x \mapsto h(x) \in \mathbb{R}_+ \cup \{\infty\}$ need not to be continuous even in case of $\mathbb{Z}$ action. We repeat this example here adjusted to the general case when G is a countable amenable discrete group.

Let $X = Y^G$ for some infinite compact metric space Y that has no isolated points. Let G act on X by the shift. Pick $y \in Y$ and $\{y_n\}_{n\in\mathbb{N}} \subseteq Y \backslash \{y\}$ such that $y_n \to y$ as $n \to \infty$. Let $x_n \in \{y, y_n\}^G$ be a point such that $\overline{Gx_n} = \{y, y_n\}^G$. Define $x = y^G$. Then $D_W(x, x_n) \to 0$ as $n \to \infty$, but $h(x) = 0$ while for every $n \in \mathbb{N}$ one has $h(x_n) = \log 2$.

Towards the proof of Theorem 5.1 we first establish an appropriate formula for the entropy that is simpler to work with than the classical definition (see Definition 2.5).

Definition 5.1 ((F, ε)-separated and (F, ε)-spanning sets)**.** Let $F \in \mathrm{Fin}(G)$ and $\delta, \varepsilon \in (0, 1]$. Let $\rho_F(x, y) = \max\{\rho(gx, gy) \colon g \in F\}$ for $x, y \in X$. A set $Z \subseteq X$ is called:

- $\boldsymbol{(F, \varepsilon)}$**-separated** if $\rho_F(x, z) > \varepsilon$ for $x, z \in Z$ with $x \neq z$. Let **sep** $\boldsymbol{(n, \varepsilon)}$ denote the maximum cardinality of an (F_n, ε)-separated set;

- $\boldsymbol{(F, \varepsilon)}$**-spanning** if for each $x \in X$ there is $z \in Z$ such that $\rho_F(x, z) \leqslant \varepsilon$. Let **span** $\boldsymbol{(n, \varepsilon)}$ denote the minimum cardinality of an (F_n, ε)-spanning set.

Note that compactness of X guarantees that both $\mathrm{sep}(n, \varepsilon)$ and $\mathrm{span}(n, \varepsilon)$ are finite for every $n \in \mathbb{N}$ and $\varepsilon > 0$. Moreover, for every $n \in \mathbb{N}$ both $\mathrm{sep}(n, \varepsilon)$ and $\mathrm{span}(n, \varepsilon)$ increase as ε decrease. Furthermore, it is noteworthy, that $\mathrm{span}(n, \varepsilon) \leqslant \mathrm{sep}(n, \varepsilon)$ for every $n \in \mathbb{N}$ and $\varepsilon > 0$. Indeed, if $Z \subseteq X$ is maximum (n, ε)-separated, then we cannot add any more point to Z such that it still has the separated property. Thus for every $x \in X$ we can find $z \in Z$ such that $\rho_{F_n}(x, z) \leqslant \varepsilon$, which means that Z is also (n, ε)-spanning.

Now, fix a Følner sequence $\mathcal{F} = \{F_n\}_{n \in \mathbb{N}}$ and denote:

$$h_{\mathrm{sep}}(X) \overset{\mathrm{def}}{=} \lim_{\varepsilon \searrow 0^+} \limsup_{n \to \infty} \frac{\log \mathrm{sep}(n, \varepsilon)}{|F_n|} \tag{5.1}$$

$$h_{\mathrm{span}}(X) \overset{\mathrm{def}}{=} \lim_{\varepsilon \searrow 0^+} \limsup_{n \to \infty} \frac{\log \mathrm{span}(n, \varepsilon)}{|F_n|}.$$

The following lemma is well known in case $G = \mathbb{Z}$.

Lemma 5.1. *For a dynamical system (X, G) it holds $h_{\mathrm{sep}}(X) = h_{\mathrm{span}}(X) = h_{\mathrm{top}}(X)$.*

Proof. The proof of this lemma is a straightforward adaptation of the proof in [15] for $G = \mathbb{Z}$. $\square$

Next, we introduce a generalization of (F, ε)-separated and (F, ε)-spanning sets.

Definition 5.2 ((F, ε, δ)-separated and (F, ε, δ)-spanning sets)**.** For $F \in \mathrm{Fin}(G)$ and $\delta, \varepsilon \in (0, 1]$ we say that $Z \subseteq X$ is:

- $\boldsymbol{(F, \varepsilon, \delta)}$**-separated** if for every $x, z \in Z$ either $x = z$ or

$$|\{g \in F \,:\, \rho(gx, gz) > \varepsilon\}| > \delta |F|. \tag{5.2}$$

  Let **sep** $\boldsymbol{(n, \varepsilon, \delta)}$ denote a minimum cardinality of an $(F_n, \varepsilon, \delta)$-separated set.

- $\boldsymbol{(F, \varepsilon, \delta)}$**-spanning** if for every $x \in X$ there exists $z \in Z$ such that

$$|\{g \in F \,:\, \rho(gx, gz) \leqslant \varepsilon\}| \geqslant (1 - \delta)|F|. \tag{5.3}$$

  Let **span** $\boldsymbol{(n, \varepsilon, \delta)}$ denote a maximum cardinality of an $(F_n, \varepsilon, \delta)$-spanning set.

Observe that if we put $\delta < \frac{1}{|F|}$ in formula (5.2), then we require that the set $\{g \in F : \rho(gx, gz) > \varepsilon\}$ has at least one element. This means that for $\delta < \frac{1}{|F|}$ we have $\operatorname{sep}(n, \varepsilon, \delta) = \operatorname{sep}(n, \varepsilon)$. Similarly, putting $\delta < \frac{1}{|F|}$ in formula (5.3), we require that the set $\{g \in F : \rho(gx, gz) \leqslant \varepsilon\}$ consists of all elements $g \in F$. Thus $\delta < \frac{1}{|F|}$ implies $\operatorname{span}(n, \varepsilon, \delta) = \operatorname{span}(n, \varepsilon)$. Moreover, for any $\delta > 0$ we have $\operatorname{sep}(n, \varepsilon, \delta) \leqslant \operatorname{sep}(n, \varepsilon)$ and $\operatorname{span}(n, \varepsilon, \delta) \leqslant \operatorname{span}(n, \varepsilon)$.

Analogously as before, for a fixed Følner sequence $\{F_n\}_{n\in\mathbb{N}}$ we denote

$$\tilde{h}_{\mathrm{sep}}(X) \stackrel{\mathrm{def}}{=} \lim_{\delta\searrow 0^+} \lim_{\varepsilon\searrow 0^+} \limsup_{n\to\infty} \frac{\log \operatorname{sep}(n, \varepsilon, \delta)}{|F_n|}, \tag{5.4}$$

$$\tilde{h}_{\mathrm{span}}(X) \stackrel{\mathrm{def}}{=} \lim_{\delta\searrow 0^+} \lim_{\varepsilon\searrow 0^+} \limsup_{n\to\infty} \frac{\log \operatorname{span}(n, \varepsilon, \delta)}{|F_n|}. \tag{5.5}$$

Below we show that $\tilde{h}_{\mathrm{sep}}(X) = \tilde{h}_{\mathrm{span}}(X) = h_{\mathrm{top}}(X)$. (Note that although we know that for small δ one has $\operatorname{sep}(n, \varepsilon, \delta) = \operatorname{sep}(n, \varepsilon)$ and $\operatorname{span}(n, \varepsilon, \delta) = \operatorname{span}(n, \varepsilon)$, these equalities do not hold trivially, because the limits as δ approaches 0 are the last one in both $\tilde{h}_{\mathrm{sep}}$ and $\tilde{h}_{\mathrm{span}}$ formulas.) This then allows us to use $\tilde{h}_{\mathrm{sep}}$ instead of h_{top} in the proof of Theorem 5.1. We start by some technical lemmas, the first one is from [56] and we state it without a proof.

Lemma 5.2 ([56], Lemma I.5.4)**.** *Let $H(\varepsilon) = -\varepsilon \log \varepsilon - (1 - \varepsilon) \log(1 - \varepsilon)$. If $\varepsilon \in (0, \frac{1}{2})$, then*

$$\sum_{j=0}^{\lfloor n\varepsilon \rfloor} \binom{n}{j} \leqslant 2^{nH(\varepsilon)} \quad \text{for } n \geqslant 1.$$

Lemma 5.3. *For a dynamical system (X, G) one has $h_{\mathrm{top}}(X) \leqslant \tilde{h}_{\mathrm{span}}(X)$.*

Proof. Notice first that since (see Definition 2.5)

$$h_{\mathrm{top}}(X) = \sup\{h(X, \mathcal{U}) : \mathcal{U} \text{ is an open cover of } X\},$$

it is enough to prove that for any finite open cover $\mathcal{U}$ of X one has

$$h(X, \mathcal{U}) \leqslant \tilde{h}_{\mathrm{span}}(X).$$

Fix an open cover $\mathcal{U}$ of X and let $2\varepsilon_0$ be its Lebesgue number. Choose $\delta \in \left(0, \frac{1}{2}\right)$ and $\varepsilon \in (0, \varepsilon_0)$. We show that

$$h(X, \mathcal{U}) \leqslant \limsup_{n\to\infty} \frac{\log \operatorname{span}(n, \varepsilon, \delta)}{|F_n|} + H(\delta) \log 2 + \delta. \tag{5.6}$$

Fix $n \in \mathbb{N}$. Let Z be a $(F_n, \varepsilon, \delta)$-spanning set such that $|Z| = \operatorname{span}(n, \varepsilon, \delta)$. For every $K \subseteq F_n$ and $y \in Z$ define

$$V(y, K) = \bigcap_{g\in K} g^{-1} \{a \in X : \rho(gy, a) < \varepsilon\},$$

Observe that $\operatorname{diam} gV(y, K) < 2\varepsilon < 2\varepsilon_0$ for each $g \in K$. Thus, by the choice of ε, for every $g \in K$ there exists $U_g \in \mathcal{U}$ such that $V(y, K) \subseteq g^{-1}U_g$. This in particular means that $V(y, K) \subseteq U$ for some $U \in \bigvee_{g\in K} g^{-1}\mathcal{U}$. Notice also that we can equivalently write

$$V(y, K) = \{a \in X : \forall g \in K,\ \rho(gy, ga) < \varepsilon\}.$$

which implies that $\{V(y,K) : y \in Z,\, K \subseteq F_n \text{ and } |K| > |F_n|(1-\delta)\}$ is an open cover of X. As a consequence of these two observations, we obtain that

$$\mathcal{W} = \bigcup_{K \subseteq F_n, |K| > |F_n|(1-\delta)} \left(\{V(y,K) \, : \, y \in Z\} \vee \bigvee_{g \in F_n \setminus K} g^{-1}\mathcal{U} \right)$$

is a refinement of a cover $\mathcal{U}^{F_n}$. Hence,

$$\begin{aligned} \mathcal{N}\left(\mathcal{U}^{F_n}\right) \leqslant \mathcal{N}(\mathcal{W}) &\leqslant \left|\{K \subseteq F_n \, : \, |K| > |F_n|(1-\delta)\}\right| |Z| |\mathcal{U}|^{|F_n|\delta} \\ &= \sum_{k=0}^{\lfloor |F_n|\delta \rfloor} \binom{|F_n|}{k} |Z||\mathcal{U}|^{|F_n|\delta} \\ &\leqslant 2^{|F_n|H(\delta)} \operatorname{span}(n,\varepsilon,\delta) |\mathcal{U}|^{|F_n|\delta}. \end{aligned}$$

This implies (5.6). Finally, observe that $H(\delta) \to 0$ as $\delta \to 0$ and thus

$$h(X,\mathcal{U}) \leqslant \lim_{\delta \searrow 0^+} \lim_{\varepsilon \searrow 0^+} \limsup_{n \to \infty} \frac{\log \operatorname{span}(n,\varepsilon,\delta)}{|F_n|}. \qquad \square$$

Lemma 5.4. *For a dynamical system* (X,G) *one has* $\tilde{h}_{\mathrm{span}}(X) = \tilde{h}_{\mathrm{sep}}(X) = h_{\mathrm{top}}(X)$.

Proof. Let $n \in \mathbb{N}$, $\varepsilon > 0$ and $\delta \in (0,1]$. Notice that

$$\operatorname{span}(n,\varepsilon,\delta) \leqslant \operatorname{sep}(n,\varepsilon,\delta). \tag{5.7}$$

Indeed, choose $Z \subseteq X$ to be a maximum (n,ε,δ)-separated set, that is $|Z| = \operatorname{sep}(n,\varepsilon,\delta)$. By the maximality of Z, for any $x \in X \setminus Z$ there exists $z \in Z$ such that condition (5.2) is not satisfied, but since (5.2) is a negation of (5.3), Z is also (n,ε,δ)-spanning.

Now, we show that

$$\operatorname{sep}(n,\varepsilon,\delta) \leqslant \operatorname{span}\left(n, \frac{\varepsilon}{2}, \frac{\delta}{2}\right). \tag{5.8}$$

Let Z be a minimum $(n, \frac{\varepsilon}{2}, \frac{\delta}{2})$-spanning set and V a maximum (n,ε,δ)-separated set. We construct an injection $\varphi \colon V \to Z$. Take $v \in V$, then there exists $z \in Z$ such that

$$\left| \left\{ g \in F \, : \, \rho(gv, gz) \leqslant \frac{\varepsilon}{2} \right\} \right| \geqslant \left(1 - \frac{\delta}{2}\right) |F|. \tag{5.9}$$

Put $\varphi(v) := z$. To check that φ is injective, assume that there are $v_1, v_2 \in V$ with $v_1 \neq v_2$ such that $\varphi(v_1) = \varphi(v_2) = z$ for some $z \in Z$. Denote

$$V_1 = \left\{ g \in F \, : \, \rho(gv_1, gz) \leqslant \frac{\varepsilon}{2} \right\}, \quad V_2 = \left\{ g \in F \, : \, \rho(gv_2, gz) \leqslant \frac{\varepsilon}{2} \right\}.$$

Observe that

$$V_1 \cap V_2 \subseteq \{ g \in F \, : \, \rho(gv_1, gv_2) \leqslant \varepsilon \},$$

thus

$$|V_1 \cap V_2| < (1-\delta)|F|. \tag{5.10}$$

On the other hand, using the triangle inequality and performing some elementary set operations, we obtain

$$\begin{aligned}|\{g \in F \,:\, \rho(gv_1, gv_2) \leqslant \varepsilon\}| &\geqslant |V_1 \cap V_2| \\ &\geqslant |F| - (|F| - |V_1|) - (|F| - |V_2|) \\ &\overset{(5.9)}{\geqslant} 2\left(1 - \frac{\delta}{2}\right)|F| - |F| \\ &= (1-\delta)\,|F|,\end{aligned}$$

which is a contradiction with (5.10). Combination of (5.7) and (5.8) gives us $\tilde{h}_{\mathrm{sep}}(X) = \tilde{h}_{\mathrm{span}}(X)$. Finally, using Lemma 5.3 we obtain

$$h_{\mathrm{top}}(X) \leqslant \tilde{h}_{\mathrm{span}}(X) = \tilde{h}_{\mathrm{sep}}(X) \leqslant h_{\mathrm{sep}}(X) = h_{\mathrm{top}}(X). \qquad \square$$

For $x \in X$, $n \in \mathbb{N}$, $\varepsilon > 0$ and $\delta \in (0,1]$ let $\mathrm{sep}(x,n,\varepsilon,\delta)$ be the maximum cardinality of $(F_n, \varepsilon, \delta)$-separated set for $(\overline{Gx}, G)$.

Lemma 5.5. *If $\varepsilon, \delta > 0$ and $x, y \in X$ are such that $D^*(\{g \in G : \rho(gx, gy) > \varepsilon\}) < \delta$, then for all $n \in \mathbb{N}$ large enough one has*

$$\mathrm{sep}(y,n,\varepsilon,\delta) \geqslant \mathrm{sep}(x,n,3\varepsilon,3\delta).$$

Proof. Fix $\varepsilon, \delta > 0$ and a Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$. Let $N_0 \in \mathbb{N}$ be such that for every $n \geqslant N_0$ and every $g \in G$ one has

$$|\{f \in F_n g \,:\, \rho(fx, fy) > \varepsilon\}| < \delta |F_n|.$$

Choose $n \geqslant N_0$. Denote $s := \mathrm{sep}(x,n,3\varepsilon,3\delta)$ and let $\{x_1, x_2, \ldots, x_s\} \subseteq \overline{Gx}$ be an $(x,n,3\varepsilon,3\delta)$-separated set. First, we justify that we can assume that for every $i \in \{1,2,\ldots,s\}$ one has $x_i = g_i x$ for some $g_i \in G$. Indeed, observe that since n is fixed and F_n is finite, we can choose $\eta > 0$ such that for every pair of distinct indices i,j it holds

$$\{g \in F_n : \rho(gx_i, gx_j) > 3\varepsilon\} = \{g \in F_n : \rho(gx_i, gx_j) > 3\varepsilon + 2\eta\}.$$

Thus, if $z_1, z_2, \ldots, z_s \in \overline{Gx}$ satisfy $\rho_{F_n}(z_i, x_i) < \eta$ (for $i = 1,2,\ldots,s$), then the set $\{z_1, z_2, \ldots, z_s\}$ is also $(x,n,3\varepsilon,3\delta)$-separated.

Next, define $Y = \{y_1, \ldots, y_s\}$ by $y_i = g_i y$ for every $i \in \{1,2,\ldots,s\}$. We show that Y is (y,n,ε,δ)-separated. Notice that for any pair of distinct indices i,j one has

$$\rho(fy_j, fy_i) \geqslant |\rho(fy_j, fx_i) - \rho(fx_i, fy_i)| \geqslant ||\rho(fx_i, fx_j) - \rho(fx_j, fy_j)| - \rho(fx_i, fy_i)|\,,$$

which implies

$$\begin{aligned}|\{f \in F_n \,:\, \rho(fy_i, fy_j) > \varepsilon\}| &\geqslant \left|\{f \in F_n : \rho(fy_i, fx_i) < \varepsilon,\ \rho(fx_i, fx_j) > 3\varepsilon,\ \rho(fx_j, fy_j) < \varepsilon\}\right| \\ &\geqslant (3\delta - \delta - \delta)|F_n| = \delta|F_n|. \qquad \square\end{aligned}$$

Now we have all the tools we need to prove Theorem 5.1.

Proof of Theorem 5.1. Since D_W and D'_W are uniformly equivalent (see Theorem 4.1), it is enough to show that $h(x)$ is lower semicontinuous with respect to D'_W. We use the $\tilde{h}_{\mathrm{sep}}$ formula (see (5.4)). Fix $\eta > 0$ and $x \in X$ such that $h(x) < \infty$. There exists $\delta, \varepsilon > 0$ such that

$$\limsup_{n\to\infty} \frac{\log \mathrm{sep}(x,n,\varepsilon,\delta)}{|F_n|} > h(x) - \eta$$

It follows from Lemma 5.5 that for every $y \in X$ such that $D'_W(x, y) < \delta$ one has

$$\limsup_{n\to\infty} \frac{\log \operatorname{sep}(y, n, \frac{\varepsilon}{3}, \frac{\delta}{3})}{|F_n|} \geqslant \limsup_{n\to\infty} \frac{\log \operatorname{sep}(x, n, \varepsilon, \delta)}{|F_n|} > h(x) - \eta,$$

which implies $h(y) > h(x) - \eta$. If $h(x) = \infty$ then the proof is similar. $\square$

5.1.2 Shift Spaces

In this section we prove a strengthening of Theorem 5.1 that claims that the entropy function is not only lower semicontinuous but even continuous, but under the assumption that $X = \mathscr{A}^G$, where $\mathscr{A}$ is a finite set, and G acts on $\mathscr{A}^G$ by shifts.

Theorem 5.2. *For every shift space $(\mathscr{A}^G, G)$ the entropy function*

$$h : (\mathscr{A}^G, D_W) \to \mathbb{R}_+ \cup \{\infty\} \qquad \text{given by} \qquad h(x) \stackrel{\text{def}}{=} h_{\text{top}}(\overline{Gx})$$

is continuous.

We note that the techniques used in the proof of the above theorem, when compared to Theorem 5.1, are different. Before we proceed to the proof, we first introduce notation that allows us to state a formula for entropy that is specific for the case of $X = \mathscr{A}^G$.

Definition 5.3. For $Y \subseteq \mathscr{A}^G$ and $F \in \operatorname{Fin}(G)$, by $\mathcal{B}_F(Y)$ we denote a collection of all patterns that appear in elements of Y over all the translations of F. Formally, we say that $p \in \mathcal{B}_F(Y) \subseteq \mathscr{A}^F$ if there exists $x \in Y$ and $g \in G$ such that for any $f \in F$, we have $p(f) = x(fg)$. For $x \in \mathscr{A}^G$ we denote $\mathcal{B}_F(x) \stackrel{\text{def}}{=} \mathcal{B}_F(\{x\})$.

The lemma below establishes a convenient formula for entropy in shift spaces. This Lemma is well known for the $\mathbb{Z}$ case, see also [7].

Lemma 5.6. *For any subshift $Y \subseteq \mathscr{A}^G$ and a Følner sequence $\{F_n\}_{n\in\mathbb{N}}$, one has*

$$h_{\text{top}}(Y) = \lim_{n\to\infty} \frac{\log |\mathcal{B}_{F_n}(Y)|}{|F_n|}. \tag{5.11}$$

Proof. The proof is divided into two steps. The first step is to show that for any shift space Y, the limit

$$\lim_{n\to\infty} \frac{\log |\mathcal{B}_{F_n}(Y)|}{|F_n|} \tag{5.12}$$

exists and does not depend on the choice of the Følner sequence. In the second step we choose a special Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ (that is two-sided) and show that for such $\mathcal{F}$ (5.11) holds. We remark that in the case when G is abelian, the first step is not necessary because each Følner sequence is two-sided.

Step 1. Fix Y. To prove the existence of limit (5.12) and its independence of the choice of a Følner sequence we use Lemma 4.11. Therefore, we need to show that the function $H : \operatorname{Fin}(G) \to [0, \infty)$ defined as

$$H(F) = \log |\mathcal{B}_F(Y)|$$

is G-invariant, monotone and subadditive. Indeed, G-invariance of H follows from the fact that Y is a shift space. Moreover, if $F \subseteq F'$, then clearly $|\mathcal{B}_F(Y)| \leqslant |\mathcal{B}_{F'}(Y)|$, thus H is monotone. Finally, observe that for any disjoint finite subsets $F, F' \subseteq G$, the condition

$$H(F \cup F') \leqslant H(F) + H(F'),$$

can be equivalently written as

$$|\mathcal{B}_{F\cup F'}(Y)| \leqslant |\mathcal{B}_F(Y)||\mathcal{B}_{F'}(Y)|.$$

But $|\mathcal{B}_F(Y)||\mathcal{B}_{F'}(Y)|$ is the maximum possible number of patterns that can appear in Y over $F \cup F'$, thus H is also subadditive.

Step 2. Let $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ be a Følner sequence such that $F_n = F_n^{-1}$ for every $n \in \mathbb{N}$ (it exists by [49]). First, observe that the G-invariance of Y implies that $\mathcal{B}_F(Y) = \{y_F : y \in Y\}$ for every $F \in \mathrm{Fin}(G)$. Since G is countable, we can write $G = \{g_0, g_1, g_2, \ldots\}$ and assume that $g_0 = e$, we also use a metric on $\mathscr{A}^G$ compatible with this indexing (see (2.1)). Denote $I_l := \{g_0, \ldots, g_l\}$ for $l \in \mathbb{N}$ and $S_{l,n} := F_n I_l$ for $l, n \in \mathbb{N}$. Now, define an equivalence relation on Y as follows, for $F \subseteq G$ one has

$$y \sim_F z \quad \text{ if and only if } \quad y_F = z_F.$$

Elements of $Y/\sim_F$ denote by $z_{[F]}$. Finally, choose $A_{l,n}$ $(, n \in \mathbb{N})$ to be a set of representatives of $Y/\sim_{S_{l,n}}$. Thus, in other words, $A_{l,n}$ $(l, n \in \mathbb{N})$ is a set satisfying the following two properties:

- for every $z_{[S_{l,n}]} \in Y/\sim_{S_{l,n}}$ there is $y \in A_{l,n}$ such that $y \in z_{[S_{l,n}]}$,
- if $y, z \in A_{l,n}$, then $y_{[S_{l,n}]} \neq z_{[S_{l,n}]}$.

Notice that $A_{l,n}$ is a maximum $(n, |\mathscr{A}|^{-l})$-separated set, hence $s(n, |\mathscr{A}|^{-l}) = |A_{l,n}|$. We also have

$$\mathcal{B}_{F_n}(Y) \subseteq \{y_{F_n} : y \in A_{l,n}\} \qquad \text{for every } l, n \in \mathbb{N}, \tag{5.13}$$

since $F_n \subseteq S_{l,n}$. Moreover, observe that since $F_n = F_n^{-1}$ for every $n \in \mathbb{N}$, it holds

$$|S_{l,n} \triangle F_n| = |(F_n I_l)^{-1} \triangle F_n^{-1}| = |I_l^{-1} F_n \triangle F_n| \qquad \text{for every } l, n \in \mathbb{N}, \tag{5.14}$$

Now, fix $l \in \mathbb{N}$ and $\eta > 0$, then since $\{F_n\}_{n\in\mathbb{N}}$ is Følner (see Lemma 4.3) there exists $N \in \mathbb{N}$ such that for $n > N$, one has

$$|I_l^{-1} F_n \triangle F_n| \leqslant \eta |F_n|, \tag{5.15}$$

Then, combining (5.13), (5.14) and (5.15), we obtain that for every $n > N$ it holds

$$|\mathcal{B}_{F_n}(Y)| \leqslant |A_{l,n}| \leqslant |\mathcal{B}_{F_n}(Y)||\mathscr{A}|^{|S_{l,n}\triangle F_n|} \leqslant |\mathcal{B}_{F_n}(Y)||\mathscr{A}|^{\eta|F_n|},$$

which implies

$$\limsup_{n\to\infty} \frac{\log|\mathcal{B}_{F_n}(Y)|}{|F_n|} \leqslant \lim_{l\to\infty} \limsup_{n\to\infty} \frac{\log|A_{l,n}|}{|F_n|} \leqslant \limsup_{n\to\infty} \frac{\log|\mathcal{B}_{F_n}(Y)|}{|F_n|} + \eta \log|\mathscr{A}|.$$

Therefore, since η was arbitrary, we obtain

$$h_{\mathrm{top}}(Y) = \limsup_{n\to\infty} \frac{\log|\mathcal{B}_{F_n}(Y)|}{|F_n|}. \qquad \square$$

As a corollary we obtain the following

Lemma 5.7. *For any $x \in \mathscr{A}^G$ and a Følner sequence $\{F_n\}_{n\in\mathbb{N}}$,*

$$h(x) = \lim_{n\to\infty} \frac{\log|\mathcal{B}_{F_n}(x)|}{|F_n|}. \tag{5.16}$$

Notice that the above lemma implies that $h(x) \leqslant \log|\mathscr{A}|$ for any $x \in \mathscr{A}^G$.

Proof of Theorem 5.2. We consider two points from $\mathscr{A}^G$ that are D^*-close and estimate their entropy using formula (5.16).

Let $\mathcal{F} = \{F_n\}_{n=1}^{\infty}$ be a Følner sequence and $\varepsilon > 0$. Pick $\delta \in (0, \frac{1}{2})$ so that $t < \delta$ implies $H(t) < \varepsilon$. Notice that if $x, z \in \mathscr{A}^G$ satisfy $D^*(x, z) < \delta$, then for all $g \in G$ and all $n \in \mathbb{N}$ large enough we have

$$|\{f \in F_n : x_{fg} \neq z_{fg}\}| < \delta |F_n|.$$

By Lemma 5.2 we have that for all $n \in \mathbb{N}$ large enough

$$|\mathcal{B}_{F_n}(z)| \leqslant |\mathcal{B}_{F_n}(x)||\mathscr{A}|^{\delta|F_n|} \sum_{k=0}^{\lfloor \delta n \rfloor} \binom{|F_n|}{k} \leqslant |\mathscr{A}|^{\delta|F_n|} 2^{H(\delta)|F_n|} |\mathcal{B}_{F_n}(x)|.$$

Thus, using formula (5.16), we obtain

$$h(z) \leqslant \delta \log |\mathscr{A}| + H(\delta) \log 2 + h(x).$$

Interchanging the roles of x and z finishes the proof. $\square$

5.2 The Number of Minimal Components of $\overline{Gx}$

Recall that by $m(x) \in \mathbb{N} \cup \{\infty\}$ we denote the number of minimal components of $\overline{Gx}$. Our aim is to prove that m is lower semicontinuous with respect to D_W, which generalizes [19, Theorem 1].

Theorem 5.3. *The function $(X, D_W) \ni x \mapsto m(x) \in \mathbb{N} \cup \{\infty\}$ is lower semicontinuous.*

Proof. Fix $x \in X$. Choose minimal sets $Z_1, \ldots, Z_{m(x)} \subseteq \overline{Gx}$ and $\varepsilon > 0$ such that $\operatorname{dist}(Z_i, Z_j) > 2\varepsilon$ whenever $i \neq j$. Fix $y \in X$ with $D'_W(x, y) < \frac{\varepsilon}{6}$, we show that $m(y) \geqslant m(x)$. Let $\{F_n\}_{n \in \mathbb{N}}$ be a Følner sequence and let $Z(y) = \{z_1, \ldots, z_{m(y)}\} \subseteq \overline{Gy}$ be the set of points generating minimal subsystems in $\overline{Gy}$, that is, $\overline{Gz_i}$ is a minimal subsystem in $\overline{Gy}$ for every $i = 1, 2, \ldots, m(y)$ (see Theorem 3.1). Now, suppose that the following claim holds.
<u>Claim:</u> There exists $N \in \mathbb{N}$ such that for every $k \in \{1, 2, \ldots, m(x)\}$, there exists $r(k) \in \{1, 2, \ldots, m(y)\}$ such that more than half of elements $f \in F_N$ satisfy

$$\operatorname{dist}(f z_{r(k)}, Z_k) < \varepsilon.$$

From this claim, it follows that for any distinct $k, l \in \{1, 2, \ldots, m(x)\}$ we can find $f \in F_N$ such that

$$\operatorname{dist}(f z_{r(k)}, Z_k) < \varepsilon \quad \text{and} \quad \operatorname{dist}(f z_{r(l)}, Z_l) < \varepsilon.$$

Therefore, since $\operatorname{dist}(Z_k, Z_l) > 2\varepsilon$, we have $r(k) \neq r(l)$ and hence $m(y) \geqslant m(x)$.

It remains to prove the Claim. The inequality $D'_W(x, y) < \frac{\varepsilon}{6}$ implies that there exists $N \in \mathbb{N}$ such that for every $g \in G$ one has

$$\left| \left\{ f \in F_N g \;:\; \rho(fx, fy) > \frac{\varepsilon}{6} \right\} \right| < \frac{\varepsilon}{6} |F_N| < \frac{1}{2} |F_N|. \tag{5.17}$$

Next, we choose $\delta \in (0, \frac{\varepsilon}{3})$ such that if $a, b \in X$ satisfy $\rho(a, b) < 2\delta$, then $\rho(fa, fb) < \varepsilon/3$ for every $f \in F_N$. Fix $k = 1, 2, \ldots, m(x)$. Then $N(x, Z_k^\delta) \cap N(y, Z(y)^\delta) \neq \varnothing$ (Lemma 3.2 and Lemma 3.1). Hence there exist $g(k) \in G$ and $r(k) = 1, 2, \ldots, m(y)$ such that

$$g(k)x \in Z_k^\delta \quad \text{and} \quad \rho(g(k)y, z_{r(k)}) < \delta. \tag{5.18}$$

Finally, keeping in mind that $gZ_k \subseteq Z_k$ for every $g \in G$ and combining (5.17) and (5.18), we get that for more than the half of elements $f \in F_N$ one has

$$\operatorname{dist}(fz_{r(k)}, Z_k) \leqslant \rho(fz_{r(k)}, fg(k)y) + \rho(fg(k)x, fg(k)y) + \operatorname{dist}(fg(k)x, Z_k) < \varepsilon. \qquad \square$$

Remark 5.1. Note that the example presented in Remark 5.1 shows that there is a G action such that m is not D_W-continuous. Such a system exists for every countable amenable group G (cf. [19, Example 2]).

5.3 Invariant Measures

Before we state Theorem 5.4 that is the main result of this section let us first introduce some useful notation and basic facts about measures. First recall that the space $\mathcal{M}(X)$ (of all Borel probability measures on X) equipped with the weak-$\star$ topology is metrizable with the **Prokhorov metric** (see [5]) given by

$$D_P(\mu,\nu) = \inf\{\varepsilon > 0 \,:\, \forall B \in \mathcal{B}(X) \quad \mu(B) \leqslant \nu(B^\varepsilon) + \varepsilon \text{ and } \nu(B) \leqslant \mu(B^\varepsilon) + \varepsilon\},$$

where as usual B^ε denotes the ε-hull of a set $B \subseteq X$, that is,

$$B^\varepsilon = \{y \in X \,:\, \exists x \in B \quad \rho(x,y) < \varepsilon\}.$$

Therefore $(\mathcal{M}(X), D_P)$ is a compact metric space. This allows us to define the Hausdorff metric $\mathrm{d_H}$ on nonempty compact subsets of $\mathcal{M}(X)$, i.e., for any nonempty compact sets $A, B \subseteq \mathcal{M}(x)$ we have

$$\mathrm{d_H}(A,B) = \inf\{\varepsilon > 0 : A \subseteq B^\varepsilon \text{ and } B \subseteq A^\varepsilon\},$$

or equivalently

$$\mathrm{d_H}(A,B) = \max\left\{\sup_{\mu\in A}\inf_{\nu\in B} D_P(\mu,\nu), \sup_{\nu\in B}\inf_{\mu\in A} D_P(\mu,\nu)\right\}.$$

Next we define the main subject of study in this section, i.e., G-invariant measures.

Definition 5.4 (G-invariant measure)**.** A measure $\mu \in \mathcal{M}(X)$ is ***G*-invariant** if for every $g \in G$ it holds

$$\mu(B) = \mu(gB) \qquad \text{for every } B \in \mathcal{B}(X).$$

By $\mathcal{M}_G(X)$ we denote the set of all G-invariant measures on X.

Below we prove an important lemma that serves as the main tool for constructing invariant measures. In particular, it implies the generalized Kryloff-Bogoliouboff theorem which states that the set $\mathcal{M}_G(X)$ is non-empty. We adapt the proof from [20] to the case of arbitrary groups.

Lemma 5.8. *Let G be a countable amenable group that acts on a compact metric space X continuously. Let $\{\nu_n\}_{n\in\mathbb{N}}$ be any sequence in $\mathcal{M}(X)$ and let $\{F_n\}_{n\in\mathbb{N}}$ be a Følner sequence in G. Then any weak*-limit point of the sequence $\{\mu_n\}_{n\in\mathbb{N}}$ defined by*

$$\mu_n = \frac{1}{|F_n|}\sum_{g\in F_n} g_*\nu_n$$

is a member of $\mathcal{M}_G(X)$.

Proof. Let μ be a weak*-limit point of $\{\mu_n\}_{n\in\mathbb{N}}$. Fix $\varphi \in \mathcal{C}(X)$ and $g \in G$. We need to prove that

$$\int g \cdot \varphi d\mu = \int \varphi d\mu,$$

where $(g \cdot \varphi)(x) \stackrel{\text{def}}{=} \varphi(gx)$ for $x \in X$ (see Definition 2.2). Since for any $n \in \mathbb{N}$ we have

$$\left|\int g \cdot \varphi d\mu - \int \varphi d\mu\right| \leqslant \left|\int g \cdot \varphi d\mu - \int g \cdot \varphi d\mu_n\right| + \left|\int g \cdot \varphi d\mu_n - \int \varphi d\mu_n\right| + \left|\int \varphi d\mu_n - \int \varphi d\mu\right|,$$

it is enough to show that

$$\left|\int g \cdot \varphi d\mu_n - \int \varphi d\mu_n\right| \to 0 \quad (n \to \infty).$$

Indeed,

$$\begin{aligned}
\left|\int g \cdot \varphi d\mu_n - \int \varphi d\mu_n\right| &= \frac{1}{|F_n|}\left|\sum_{h\in F_n}\int g \cdot \varphi d h_* \nu_n - \sum_{h\in F_n}\int \varphi d h_*\nu_n\right| \\
&= \frac{1}{|F_n|}\left|\sum_{h\in F_n}\int h \cdot (g \cdot \varphi) d\nu_n - \sum_{h\in F_n}\int h \cdot \varphi d\nu_n\right| \\
&= \frac{1}{|F_n|}\left|\sum_{h\in F_n}\int (gh) \cdot \varphi d\nu_n - \sum_{h\in F_n}\int h \cdot \varphi d\nu_n\right| \\
&= \frac{1}{|F_n|}\left|\sum_{h\in gF_n}\int h \cdot \varphi d\nu_n - \sum_{h\in F_n}\int h \cdot \varphi d\nu_n\right| \\
&\leqslant \frac{1}{|F_n|}\sum_{h\in gF_n\Delta F_n}\left|\int h \cdot \varphi d\nu_n\right| \\
&\leqslant \frac{|gF_n \Delta F_n|}{|F_n|}\|\varphi\|_\infty.
\end{aligned}$$

Since $\{F_n\}_{n\in\mathbb{N}}$ is a Følner sequence, we have

$$\lim_{n\to\infty}\frac{|gF_n\Delta F_n|}{|F_n|} = 0.$$

Further $\|\varphi\|_\infty < \infty$ since φ is a continuous function on a compact space. The lemma follows. □

The fact that $\mathcal{M}_G(X)$ is non-empty follows now from the compactness of $\mathcal{M}(X)$. We are now ready to state the main result of this section.

Theorem 5.4. *For a dynamical system (X, G), the function*

$$\mathcal{M}_{\mathrm{G}} \colon (X, D_W) \to (2^{\mathcal{M}(X)}, \mathrm{d_H}) \qquad \textit{given by} \qquad \mathcal{M}_{\mathrm{G}}(x) = \mathcal{M}_G(\overline{Gx})$$

is uniformly continuous.

Before we present the proof of Theorem 5.4, we discuss its corollary.

Corollary 5.1. *The number of ergodic*[1] *invariant measures supported at $\overline{Gx}$ is lower semicontinuous with respect to D_W.*

[1] Recall that $\mu \in \mathcal{M}_G(X)$ is ergodic if $gB = B$ implies $\mu(B) = 1$ or $\mu(B) = 0$ for every $B \in \mathcal{B}(X)$ and $g \in G$.

Proof. Note first that since ergodic invariant measures are extreme points of $\mathcal{M}_G(X)$, it is just enough to show that the number of extreme points of a compact convex subset of a compact space varies lower semicontinuously with respect to $\mathrm{d_H}$. To see this assume that $K_n \subseteq \mathcal{M}_G(X)$ for $n \in \mathbb{N}$ and $K \subseteq \mathcal{M}_G(X)$ are compact, convex sets such that $K_n \to K$ in the Hausdorff metric. For every $n \in \mathbb{N}$ denote by s_n the number of extreme points of K_n and let s denote the number of extreme points K. We can assume that $\liminf\limits_{n\to\infty} s_n = m$ for some $m \in \mathbb{N}$ as the case $m = \infty$ is easy to deal with. Let $(l_n)_{n\in N}$ be a sequence of indices such that $l_n \to \infty$ and $s_{l_n} = m$ for all $n \in \mathbb{N}$.

For every $n \in \mathbb{N}$ let $V_n := \{v_1^n, \ldots, v_m^n\} \subseteq K_{l_n}$ be the set of extreme points of K_{l_n}, i.e., $\mathrm{conv}(V_n) = K_{l_n}$, where $\mathrm{conv}(V_n)$ denotes the convex hull of V_n. We can assume (restricting to a subsequence if necessary) that for every $1 \leqslant i \leqslant m$ one has $v_i^n \to v_i$ for some $v_i \in K$. We show that $\mathrm{conv}\{v_1, \ldots, v_m\} = K$. To this end fix $x \in K$. For $1 \leqslant i \leqslant m$ and $n \in \mathbb{N}$ pick $\alpha_m^{(n)} \in [0,1]$ such that:

- for every $n \in \mathbb{N}$ one has $\alpha_1^{(n)} + \ldots + \alpha_m^{(n)} = 1$,
- $\alpha_1^{(n)} v_1^n + \ldots + \alpha_m^{(n)} v_m^n \to x$ as $n \to \infty$.

By restricting to a subsequence, we can assume for every $i = 1, 2, \ldots, m$ we have $\alpha_i^{(n)} \to \alpha_i$ for some $\alpha_i \in [0,1]$. Then $\alpha_1 + \ldots + \alpha_m = 1$ and $\alpha_1 s_1 + \ldots + \alpha_m s_m = x$. This finishes the proof. □

Before we proceed with the proof of Theorem 5.4, we define a family of measures that allows us to characterize the collection of G-invariant measures.

Definition 5.5. The **empirical measure** $\mathfrak{m}(\underline{x}, F) \in \mathcal{M}(X)$ of an element $\underline{x} \in X^G$ with respect to a finite set $F \subseteq G$ is given by

$$\mathfrak{m}(\underline{x}, F) \stackrel{\text{def}}{=} \frac{1}{|F|} \sum_{g\in F} \delta_{\underline{x}(g)}.$$

A measure $\mu \in \mathcal{M}(X)$ is a **distribution measure** for $\underline{x} \in X^G$ and a Følner seqence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ if μ is a limit point of the sequence $\{\mathfrak{m}(\underline{x}, F_n)\}_{n\in\mathbb{N}}$. The set of all distribution measures of $\underline{x}$ along $\mathcal{F}$ is denoted by $\mathrm{Dis}_{\mathcal{F}}(\underline{x})$ and the set of all distribution measures of $\underline{x}$ we define as

$$\mathrm{Dis}(\underline{x}) \stackrel{\text{def}}{=} \bigcup_{\mathcal{F}} \mathrm{Dis}_{\mathcal{F}}(\underline{x}).$$

For $x \in X$ we put $\mathrm{Dis}_{\mathcal{F}}(x) \stackrel{\text{def}}{=} \mathrm{Dis}_{\mathcal{F}}(\underline{x}_G)$ and $\mathrm{Dis}(x) \stackrel{\text{def}}{=} \mathrm{Dis}(\underline{x}_G)$.

Note that $\mathrm{Dis}_{\mathcal{F}}(\underline{x})$ is a closed and nonempty subset of $\mathcal{M}(X)$ for every Følner sequence $\mathcal{F}$. Now, observe that Theorem 5.4 is a straightforward consequence of the following two lemmas.

Lemma 5.9. *For every $x \in X$, the set* $\mathrm{Dis}(x)$ *is compact and one has* $\mathcal{M}_G(x) = \mathrm{Dis}(x)$.

Lemma 5.10. *The function* $\mathrm{Dis}\colon (X, D_W) \to (2^{\mathcal{M}(X)}, \mathrm{d_H})$ *is uniformly continuous.*

The proofs of the above lemmas are completely independent of each other and are established in the two subsequent sections.

Proof of Lemma 5.9

We start with the proof of the first part of Lemma 5.9.

Lemma 5.11. *For any $x \in X$ the set* $\mathrm{Dis}(x)$ *is compact.*

Proof. Fix $x \in X$, we just need to prove that $\mathcal{M}_G(x)$ is closed. Pick $\{\mu_k\}_{k\in\mathbb{N}} \subseteq \mathrm{Dis}(x)$ such that $\mu_k \to \mu$ as $k \to \infty$ for some $\mu \in \mathcal{M}(X)$. We show that $\mu \in \mathrm{Dis}(x)$. For every $k \in \mathbb{N}$ let $\mathcal{F}^{(k)} = \{F_n^{(k)}\}_{n\in\mathbb{N}}$ be a Følner sequence such that $\mathfrak{m}(x, F_n^{(k)}) \to \mu_k$ as $n \to \infty$. Enumerate elements of G as $g_1, g_2, \ldots$. For every $k \in \mathbb{N}$ choose $n_k \in \mathbb{N}$ such that

$$D_P\left(\mathfrak{m}(x, F_{n_k}^{(k)}), \mu_k\right) < \frac{1}{k}$$

and for every $i \leqslant k$ one has

$$\frac{\left|g_i F_{n_k}^{(k)} \triangle F_{n_k}^{(k)}\right|}{\left|F_{n_k}^{(k)}\right|} < \frac{1}{k}.$$

Define $\mathcal{F} = \{F_{n_k}^{(k)}\}_{k\in\mathbb{N}}$. Then $\mathcal{F}$ is a Følner sequence and $\mathrm{Dis}_{\mathcal{F}}(x) = \{\mu\}$. □

In the proof of the second part of Lemma 5.9 we use Lemma 5.12 that is a straightforward consequence of the Pointwise Ergodic Theorem for amenable groups (for a proof of the latter result we refer to [42]). To state it we need to define a special class of Følner sequences.

Definition 5.6 (Tempered Følner sequence)**.** A Følner sequence $\{F_n\}_{n\in\mathbb{N}}$ in G is **tempered** if there exists a constant $C > 0$ such that for every $n \in \mathbb{N}$ it holds

$$\left|\bigcup_{k<n} F_k^{-1} F_n\right| \leqslant C\,|F_n|\,.$$

It is proved in [42] that every Følner sequence has a tempered subsequence, which in particular means that every amenable group has a tempered Følner sequence.

Theorem 5.5 (Pointwise Ergodic Theorem [42])**.** *Let $\{F_n\}_{n\in\mathbb{N}}$ be a tempered Følner sequence in G and $\mu \in \mathcal{M}_G(X)$. Then for every function $\phi \in L^1(\mu)$, there exists a G-invariant function $\phi^\star \in L^1(\mu)$ satisfying*

$$\lim_{n\to\infty} \frac{1}{|F_n|} \sum_{g\in F_n} \phi(gx) = \phi^\star(x) \qquad \mu\text{-almost everywhere.}$$

Moreover, if μ is ergodic, then

$$\phi^\star(x) = \int_X \phi d\mu \qquad \mu\text{-almost everywhere.}$$

As a corollary we will obtain Lemma 5.12, which claims existence of generic points. To state it, we need the following definition:

Definition 5.7 (Generic point)**.** Let $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ be a Følner sequence and let $\mu \in \mathcal{M}_G(X)$ be an ergodic measure. A point $x \in X$ is **generic** for μ along $\mathcal{F}$ if for every continuous function $\phi \in \mathbb{R}^X$ one has

$$\lim_{n\to\infty} \frac{1}{|F_n|} \sum_{g\in F_n} \phi(gx) = \int_X \phi d\mu,$$

in other words

$$\lim_{n\to\infty} \mathfrak{m}(x, F_n) = \mu.$$

Lemma 5.12. *If $\mathcal{F}$ is a tempered Følner sequence in G and $\mu \in \mathcal{M}_G(X)$ is ergodic, then μ-almost every point $x \in X$ is generic for μ along $\mathcal{F}$.*

Proof. This is a straightforward consequence of Pointwise Ergodic Theorem (see Theorem 5.5). □

Note that by Lemma 5.12, for every ergodic measure $\mu \in \mathcal{M}_G(X)$, there exists a generic point $x \in X$ such that $x \in \operatorname{supp}(\mu)$ (where $\operatorname{supp}(\mu)$ denotes the support of μ, that is, the set of all points $x \in X$ such that for every open set $U \subseteq X$ with $x \in U$, one has $\mu(U) > 0$). Indeed, this follows from the fact that both those sets, $\operatorname{supp}(\mu)$ and the set of generic points for μ, are of measure 1.

To proceed we need the following three simple lemmas, we provide proofs for the sake of completeness.

Lemma 5.13. *Let $x \in X$ and $Z = \overline{Gx}$. Take an ergodic measure $\mu \in \mathcal{M}_G(Z)$ and choose its generic point $z \in \operatorname{supp}(\mu)$. Then for every open neighborhood $U \subseteq Z$ of z there exists infinitely many elements $g \in G$ such that $gx \in U$.*

Proof. Fix an open neighborhood $U \subseteq Z$ of z, without loss of generality U is an open ball of radius $\varepsilon > 0$ around z. Assume first that $z = g_0 x$ for some $g_0 \in G$. Since $z \in \operatorname{supp}(\mu)$ and U is open, we have $\mu(U) > 0$. But z is generic for μ, hence there must be infinitely many $g \in G$ such that $gz \in U$.

If $z \notin Gx$, then there exists a sequence $\{g_n\}_{n \in \mathbb{N}} \subseteq G$ such that $\lim_{n \to \infty} g_n x = z$ and $g_n x \neq z$ for every $n \in \mathbb{N}$. By taking a subsequence, we might assume that $\rho(g_n x, z) < \varepsilon$ for every $n \in \mathbb{N}$ and $\rho(g_n x, z)$ is strictly decreasing with n. It follows that all g_n's are pairwise distinct and $g_n x \in U$ for every $n \in \mathbb{N}$. □

Lemma 5.14. *Let $F \in \operatorname{Fin}(G)$ and $x, z \in X$. If for every $f \in F$ one has $\rho(fx, fz) \leqslant \varepsilon$, then $D_P(\mathfrak{m}(x, F), \mathfrak{m}(z, F)) \leqslant \varepsilon$.*

Proof. Take any Borel set $B \subseteq X$ and any $\varepsilon > 0$. By the assumption we have

$$|\{f \in F : fx \in B\}| \leqslant |\{f \in F : fz \in B^\varepsilon\}|,$$

which implies

$$\frac{1}{|F|} |\{f \in F : fx \in B\}| \leqslant \frac{1}{|F|} |\{f \in F : fz \in B^\varepsilon\}| + \varepsilon.$$

Therefore $\mathfrak{m}(x, F)(B) \leqslant \mathfrak{m}(y, F)(B^\varepsilon) + \varepsilon$. Interchanging the role of x and z we obtain the claim. □

Lemma 5.15. *Let $\alpha_1, \ldots, \alpha_k \in [0,1]$ be such that $\sum_{i=1}^k \alpha_i = 1$. Then for all $\mu_1, \ldots, \mu_k, \nu_1, \ldots, \nu_k \in \mathcal{M}(X)$ one has*

$$D_P\left(\sum_{i=1}^k \alpha_i \mu_i, \sum_{i=1}^k \alpha_i \nu_i\right) \leqslant \max\left\{D_P(\mu_i, \nu_i) \,:\, 1 \leqslant i \leqslant k\right\}.$$

Proof. Suppose that $D_P(\mu_i, \nu_i) \leqslant \varepsilon$ for some $\varepsilon \geqslant 0$ and every $i = 1, 2, \ldots, k$. Then, for every Borel set $B \subseteq X$, we have

$$\begin{aligned}\left(\sum_{i=1}^k \alpha_i \mu_i\right)(B) &= \sum_{i=1}^k \alpha_i \mu_i(B) \\ &\leqslant \sum_{i=1}^k \alpha_i (\nu_i(B^\varepsilon) + \varepsilon) \\ &= \left(\sum_{i=1}^k \alpha_i \nu_i\right)(B^\varepsilon) + \varepsilon.\end{aligned}$$

By exchanging the roles of μ's and ν's, we arrive at the claim. □

Proof of Lemma 5.9. It is obvious that $\mathrm{Dis}(x) \subseteq \mathcal{M}_G(x)$. Fix $\mu \in \mathcal{M}_G(x)$ and a tempered Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$. From the Krein-Milman theorem we know that μ belongs to the closure of the convex hull of all extreme points of $\mathcal{M}_G(x)$. Since the extreme points of $\mathcal{M}_G(x)$ are exactly the ergodic measures $\mathcal{M}_G^e(\overline{Gx})$, we can express μ as a limit of the form

$$\mu = \lim_{n\to\infty} \frac{1}{L_n} \sum_{i=1}^{L_n} \nu_i^{(n)},$$

where for each $n \in \mathbb{N}$, L_n is a natural number and for $i \in \{1, 2, \ldots, L_n\}$ one has $\nu_i^{(n)} \in \mathcal{M}_G^e(\overline{Gx})$ (the measures $\nu_i^{(n)}$ need not to be pairwise different).

Since the set $\mathrm{Dis}(x)$ is closed (by Lemma 5.11), it is enough to prove the following general statement: for every sequence of ergodic measures $\nu_1, \nu_2, \ldots, \nu_N$ and for every $\varepsilon > 0$ there exists a Følner sequence $\mathcal{F}_\varepsilon = \{F_{k,\varepsilon}\}_{k\in\mathbb{N}}$ such that for every $k \in \mathbb{N}$ large enough one has

$$D_P\left(\frac{1}{N}\sum_{i=1}^{N} \nu_i,\, \mathfrak{m}\left(x, F_{k,\varepsilon}\right)\right) < \varepsilon.$$

For $i = 1, 2, \ldots, N$ let $x_i \in \mathrm{supp}(\nu_i)$ be a generic point for ν_i along $\mathcal{F}$ (see Lemma 5.12), i.e., $\lim_{k\to\infty} \mathfrak{m}(x_i, F_k) = \nu_i$. From now on, we assume that k is large enough so that for every $i \in \{1, 2, \ldots, N\}$ it holds

$$D_P\left(\nu_i, \mathfrak{m}(x_i, F_k)\right) < \frac{\varepsilon}{2}.$$

Let $\delta > 0$ be such that if $f \in F_k$, $a, b \in X$ and $\rho(a, b) < \delta$ then $\rho(fa, fb) < \frac{\varepsilon}{2}$.

We claim that there exist $g_1^{(k)}, g_2^{(k)}, \ldots, g_N^{(k)} \in G$ such that:

1. for every $i \in \{1, 2, \ldots, N\}$ it holds $\rho(g_i^{(k)}x, x_i) < \delta$,
2. the sets $F_k g_1^{(k)}, F_k g_2^{(k)}, \ldots, F_k g_N^{(k)}$ are pairwise disjoint.

The claim simply follows from the fact that for every $i \in \{1, 2, \ldots, N\}$ there are infinitely many $g \in G$ such that $\rho(gx, x_i) < \delta$ (see Lemma 5.13).

Now, Lemma 5.15 combined with Lemma 5.14 imply

$$D_P\left(\frac{1}{N}\sum_{i=1}^{N} \nu_i, \frac{1}{N}\sum_{i=1}^{N} \mathfrak{m}(x, F_k g_i^{(k)})\right) \leqslant \max_{1\leqslant i\leqslant N} D_P(\nu_i, \mathfrak{m}(x, F_k g_i^{(k)})) \leqslant \varepsilon. \tag{5.19}$$

It remains to observe that the sequence $\mathcal{F}_\varepsilon = \{F_{\varepsilon,k}\}_{k=1}^{\infty}$ given by

$$F_{\varepsilon,k} \overset{\text{def}}{=} \bigsqcup_{i=1}^{N} F_k g_i^{(k)},$$

is a Følner sequence and we have

$$\frac{1}{N}\sum_{i=1}^{N} \mathfrak{m}(x, F_k g_i^{(k)}) = \mathfrak{m}(x, F_{\varepsilon,k}),$$

which combined with (5.19) concludes the proof. □

Proof of Lemma 5.10

The following lemma is the basic component of the proof of Lemma 5.10.

Lemma 5.16. *For every* $\underline{x}, \underline{z} \in X^G$ *and for every* $\mu \in \mathrm{Dis}(x)$ *there exists* $\nu \in \mathrm{Dis}(z)$ *with* $D_P(\mu, \nu) \leqslant D_W(\underline{x}, \underline{z})$.

Proof. Denote $\varepsilon := D_W(\underline{x}, \underline{z})$. Take any $\mu \in \mathrm{Dis}(x)$. There exists a Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ such that

$$\mu = \lim_{n\to\infty} \mathfrak{m}(\underline{x}, F_n).$$

Let $\nu \in \mathrm{Dis}_{\mathcal{F}}(z)$ be arbitrary, we a can assume, by passing to a subsequence if necessary that

$$\nu = \lim_{n\to\infty} \mathfrak{m}(\underline{z}, F_n).$$

Recall that by Lemma 4.8 we have $D_W = \sup_{\mathcal{F}} D_{B,\mathcal{F}}$, hence in particular $D_{B,\mathcal{F}}(x, z) \leqslant \varepsilon$. This means that there exists $n \in \mathbb{N}$ such that for $n > N$, one has

$$|\{g \in F_n : \rho(\underline{x}(g), \underline{z}(g)) \geqslant \varepsilon\}| < \varepsilon |F_n|. \tag{5.20}$$

Let $B \in \mathcal{B}(X)$. We show that $\mu(B) \leqslant \nu(B^\varepsilon) + \varepsilon$. Notice that it is enough to prove that for $n \in \mathbb{N}$ large enough, we have

$$\mathfrak{m}(\underline{x}, F_n)(B) \leqslant \mathfrak{m}(\underline{z}, F_n)(B^\varepsilon) + \varepsilon.$$

For $n > N$, by (5.20) we obtain

$$\begin{aligned}
&\mathfrak{m}(\underline{x}, F_n)(B) = \frac{|\{g \in F_n : \underline{x}(g) \in B\}|}{|F_n|} \\
&= \frac{|\{g \in F_n : \underline{x}(g) \in B \text{ and } \rho(\underline{x}(g), \underline{z}(g)) < \varepsilon\}|}{|F_n|} + \frac{|\{g \in F_n : \underline{x}(g) \in B \text{ and } \rho(\underline{x}(g), \underline{z}(g)) \geqslant \varepsilon\}|}{|F_n|} \\
&\leqslant \frac{|\{g \in F_n : \underline{z}(g) \in B^\varepsilon\}|}{|F_n|} + \varepsilon = \mathfrak{m}(\underline{z}, F_n)(B^\varepsilon) + \varepsilon.
\end{aligned}$$

Interchanging the roles of μ and ν we obtain $\nu(B) \leqslant \mu(B^\varepsilon) + \varepsilon$, and hence $D_P(\mu, \nu) < \varepsilon$. □

The proof of Lemma 5.10 now follows easily.

Proof of Lemma 5.10. It follows straight from the definition of Hausdorff distance and Lemma 5.16 that the Dis function is Lipschitz-continuous with the Lipschitz constant 1, and thus uniformly continuous. □

Chapter 6

Subordinate Shifts

Given a shift space $Z \subseteq \{0,1\}^G$ one can define its so-called *subordinate shift* $\check{Z} \subseteq \{0,1\}^G$ by saying that $x \in \check{Z}$ if and only if there exists a point z in Z such that x is obtained by replacing some 1's with 0's in z (or in other words $x \leqslant z$ coordinate-wise). This notion was defined and studied for the shift space $\{0,1,\ldots,n\}^{\mathbb{N}}$ in [39].

It is worth noting that even if Z is simple, its subordinate shift may be complicated, for example take $Z = \{1^G\}$, then $\check{Z} = \{0,1\}^G$. However, what turns out, is that computing the entropy of a subordinate shift $\check{Z}$ under the assumption that $h_{\text{top}}(Z) = 0$ is much easier than computing the entropy in general. In Theorem 6.1 we prove that if $h_{\text{top}}(Z) = 0$, the entropy of $\check{Z}$ coincides with the "asymptotic density of ones" in Z. This means that by constructing a space Z with zero entropy and a particular density of 1's in Z, we can obtain a space $\check{Z}$ having that given entropy.

Theorem 6.1 plays a crucial role in the proof of Theorem 8.2, where the existence of a proximal dynamical system with a given entropy, boils down to a construction of a zero-entropy configuration with a given density of 1's. Additionally, Theorem 6.1 paves a new way for computing the entropy of subordinate shifts: in Lemma 6.6 we prove that for Z with $h_{\text{top}}(Z) = 0$, we have that $h_{\text{top}}(\check{Z})$ is equal to the maximum (over all G-invariant measures) of the measure of the cylinder[1] $[1]_e$.

6.1 Transitivity of Subordinate Shifts

In this section we give a characterization of a family of shifts such that their subordinate shifts are transitive. We also introduce some basic properties of the operation of taking subordinate set .

Definition 6.1 (Subordinate shift)**.** For a set $U \subseteq \{0,1\}^G$ we define its **subordinate set** as

$$\check{U} \stackrel{\text{def}}{=} \left\{z \in \{0,1\}^G : \exists x \in U \quad z \leqslant x\right\},$$

where $x \leqslant z$ means that $x(g) \leqslant z(g)$ for every $g \in G$. In particular, if $Z \subseteq \{0,1\}^G$ is a subshift, then we say that $\check{Z}$ is its **subordinate shift**.

The aim of this section is to prove the following.

Lemma 6.1 (Transitivity of subordinate shifts)**.** *Let $x \in X$ and let $\{F_n\}_{n\in\mathbb{N}}$ be a Følner sequence in G. Denote $Z = \overline{Gx}$. If for every $n \in \mathbb{N}$ there exist infinitely many elements $g \in G$ such that $x_{F_n} \leqslant x_{F_n g}$, then there exists $z \in \check{Z}$ such that $z \leqslant x$ and $\overline{Gz} = \check{Z}$.*

[1] Recall that $[1]_e = \{x \in \{0,1\}^G : x(e) = 1\}$ (see Definition 2.6).

Before we proceed with the proof, we explain that the word "shift" in the name "subordinate shift" is indeed justified.

Lemma 6.2. *If $Z \subseteq \{0,1\}^G$ is a subshift, then its subordinate shift $\check{Z}$ is also a subshift in $\{0,1\}^G$ (i.e., it is closed and invariant).*

Proof. Clearly, $\check{Z}$ is G-invariant. We show that it is closed. Let $\{x^{(n)}\}_{n\in\mathbb{N}} \subseteq \check{Z}$ converge to some $x \in \{0,1\}^G$. Then there exists $\{z^{(n)}\}_{n\in\mathbb{N}} \subseteq Z$ such that $x^{(n)}(g) \leqslant z^{(n)}(g)$ for every $n \in \mathbb{N}$ and $g \in G$. Since Z is compact, there exists a subsequence $\{z^{(k_n)}\}_{n\in\mathbb{N}}$ that converges to some element $z \in Z$. For simplicity assume that $k_n = n$ for every $n \in \mathbb{N}$. We show that $x(g) \leqslant z(g)$ for every $g \in G$. Fix $g \in G$. If $z(g) = 0$, then there exists $N \in \mathbb{N}$ such that for $n > N$ we have $z^{(n)}(g) = 0$. Therefore $x^{(n)}(g) = 0$ for $n > N$, which implies $x(g) = 0$. $\square$

Next, we show how operations of taking a subordinate set and a closure of a set interact with each other.

Lemma 6.3. *The closure of a set commutes with the operation of taking its subordinate set, that is, for every $U \subseteq \{0,1\}^G$ we have $\check{\overline{U}} = \overline{\check{U}}$.*

Proof. Clearly, $\check{U} \subseteq \check{\overline{U}}$. Thus, since $\check{\overline{U}}$ is closed (by Lemma 6.2), we have $\overline{\check{U}} \subseteq \check{\overline{U}}$. Next, we justify the converse inclusion. Fix a Følner sequence $\{F_n\}_{n\in\mathbb{N}}$. Let $x \in \check{\overline{U}}$. Then there exists $y \in \overline{U}$ such that $x \leqslant y$. We can find $\{y_n\}_{n\in\mathbb{N}} \subseteq U$ such that $y_n \to y$ as $n \to \infty$. Passing to a subsequence if necessary, we can assume that $x_{F_n} \leqslant (y_n)_{F_n}$ for every $n \in \mathbb{N}$. Now, for every $n \in \mathbb{N}$ define $x_n \in \{0,1\}^G$ as follows:

$$(x_n)_{F_n} = x_{F_n} \qquad \text{and} \qquad (x_n)_{G\setminus F_n} \equiv 0.$$

Then, certainly, $x_n \to x$ as $n \to \infty$. Moreover, for every $n \in \mathbb{N}$ we have $x_n \leqslant y_n$, thus $x_n \in \check{U}$. Hence $x \in \overline{\check{U}}$. $\square$

We need one more technical lemma to prove Lemma 6.1.

Lemma 6.4. *Let $x \in X$ and let $\{F_n\}_{n\in\mathbb{N}}$ be a Følner sequence in G. If $z \in \widecheck{\overline{Gx}}$ is such that for every $n \in \mathbb{N}$ and for every $w \in \{0,1\}^{F_n}$ satisfying $w \leqslant x_{F_n}$ there exists $g \in G$ such that $z_{F_n g} = w$, then $\overline{Gz} = \widecheck{\overline{Gx}}$.*

Proof. Denote $Z = \overline{Gx}$. Since $z \in \check{Z}$, it is clear that $Gz \subseteq \check{Z}$. But $\check{Z}$ is closed (by Lemma6.2), hence we also have $\overline{Gz} \subseteq \check{Z}$. Now, we prove the converse inclusion. By Lemma 6.3 we have $\check{Z} = \overline{\widecheck{Gx}}$. Thus, since $\overline{Gz}$ is closed, it is enough to show that $\widecheck{Gx} \subseteq \overline{Gz}$. Fix $y \in \widecheck{Gx}$. Then there exists $g \in G$ such that $y \leqslant gx$, which implies $g^{-1}y \leqslant x$. This in particular means that $(g^{-1}y)_{F_n} \leqslant x_{F_n}$ for every $n \in \mathbb{N}$. Therefore $g^{-1}y \in \overline{Gz}$, and thus $y \in \overline{Gz}$. $\square$

Proof of Lemma 6.1. Denote $Z = \overline{Gx}$ and recall that

$$\check{Z} = \{y \in \{0,1\}^G : \exists z \in Z \ \ \forall g \in G \ \ y(g) \leqslant z(g)\}.$$

Moreover, for every $n \in \mathbb{N}$ define

$$W_n \stackrel{\text{def}}{=} \{w \in \{0,1\}^{F_n} : w \leqslant x_{F_n}\}$$

and denote

$$W \stackrel{\text{def}}{=} \bigcup_{n\in\mathbb{N}} W_n.$$

Now, let $w_1, w_2, w_3, \ldots$ be an arbitrary enumeration of W, further for every $k \in \mathbb{N}$ denote by n_k the unique index such that $w_k \in W_{n_k}$.

If we construct an element $z \in \check{Z}$ such that for every $k \in \mathbb{N}$ the following holds:

$$\text{Exh(k): there exists } g_k \in G \text{ such that } z_{F_{n_k} g_k} = w_k,$$

then by Lemma 6.4 we obtain the claim.

We inductively construct non-icreasing sequence of elements $z_k \in \check{Z}$ and subsequently define z as its limit:

$$z \stackrel{\text{def}}{=} \lim_{k \to \infty} z_k.$$

Note that the existence of this limit follows from monotonicity of $\{z_k\}_{k \in \mathbb{N}}$. In the k-th step of the construction we define $z_k \in \check{Z}$ modifying z_{k-1} such that Exh(k) is satisfied for z_k. To assure that this does not break $Exh(i)$ for any $i < k$, we need to be careful choosing indices in z_k that we modify. Formally, in the k-th step we choose $g_k \in G$ and construct $z_k \in \check{Z}$ such that:

(1) $z_k \leqslant z_{k-1}$,

(2) $(z_k)_{F_{n_k} g_k} = w_k$,

(3) $(z_k)_{G \setminus F_{n_k} g_k} = (z_{k-1})_{G \setminus F_{n_k} g_k}$

(4) $F_{n_k} g_k$ is disjoint from $F_{n_i} g_i$ for every $i < k$.

In the base step of the construction for $k = 0$, we set $z_0 = x$. In the k-th step, having $z_{k-1} \in \check{Z}$ constructed, we define $z_k \in \check{Z}$ as follows: as g_k we choose an element $g \in G$ such that

(i) $x_{F_{n_k}} \leqslant x_{F_{n_k} g}$,

(ii) $F_{n_k} g \cap F_{n_i} g_i = \varnothing$ for every $i < k$.

Note that by the assumption about x, condition (i) is satisfied for infinitely many elements $g \in G$. Further, condition (ii) holds for all but finitely many $g \in G$. Consequently, there exists $g_k \in G$ that satisfies both these conditions.

Next we define $z_k \in \check{Z}$ as follows:

$$(z_k)_{F_{n_k} g_k} \stackrel{\text{def}}{=} w_k \qquad \text{and} \qquad (z_k)_{G \setminus F_{n_k} g_k} \stackrel{\text{def}}{=} (z_{k-1})_{G \setminus F_{n_k} g_k}.$$

Now, observe that conditions (2) and (3) follow directly from the definition of z_k. Further, condition (4) follows from (ii). Finally, condition (1) holds since $w_k \leqslant x_{F_{n_k}} \leqslant x_{F_{n_k} g_k}$. $\square$

In the example below we show that it is not true in general that a subordinate of a transitive shift space is transitive.

Example 6.1. Let $x \in \{0,1\}^{\mathbb{Z}}$ be defined as

$$x \stackrel{\text{def}}{=} \ldots 0000.110000 \ldots.$$

Then $\overline{Gx} = Gx \cup \{0^{\mathbb{Z}}\}$ and $\breve{\overline{Gx}} = Gx \cup Gy \cup \{0^{\mathbb{Z}}\}$, where

$$y \stackrel{\text{def}}{=} \ldots 0000.100000 \ldots.$$

Now, it is clear that $\breve{\overline{Gx}}$ is not transitive, because up to shifts there are only three distinct elements in this space: x, y and $0^{\mathbb{Z}}$, none of which has a dense orbit.

6.2 Entropy

The main result of this section is to prove that the maximum asymptotic density of ones in a subshift Z coincides with $h_{\text{top}}(\check{Z})$ provided $h_{\text{top}}(Z) = 0$.

Theorem 6.1. *Let G be a countable amenable group with a Følner sequence $\{F_n\}_{n\in\mathbb{N}}$. If $Z \subseteq \{0,1\}^G$ is a subshift such that $h_{\text{top}}(Z) = 0$, then the entropy of its subordinate shift $\check{Z}$ is given by*

$$h_{\text{top}}(\check{Z}) = \log(2) \lim_{n\to\infty} \frac{1}{|F_n|} \max\left\{ \sum_{g\in F_n} w(g) : w \in \mathcal{B}_{F_n}(Z) \right\}.$$

Proof. Let $\{F_n\}_{n\in\mathbb{N}}$ be a Følner sequence in G. Since $\check{Z} \subseteq \{0,1\}^G$ is a subshift, we can use the formula from Lemma 5.6, that is,

$$h_{\text{top}}(\check{Z}) = \lim_{n\to\infty} \frac{\log|\mathcal{B}_{F_n}(\check{Z})|}{|F_n|},$$

where $\mathcal{B}_{F_n}(\check{Z})$ denotes the collection of all patterns that appear in elements of $\check{Z}$ over all the translations of F_n (Definition 5.3). By the assumption we have

$$0 = \lim_{n\to\infty} \frac{\log|\mathcal{B}_{F_n}(Z)|}{|F_n|} = \lim_{n\to\infty} \log|\mathcal{B}_{F_n}(Z)|^{|F_n|^{-1}},$$

thus

$$\lim_{n\to\infty} |\mathcal{B}_{F_n}(Z)|^{|F_n|^{-1}} = 1.$$

This gives us a bound on the number of patterns appearing in Z over F_n for sufficiently large n. Formally, for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for every $n \geqslant N$, we have

$$|\mathcal{B}_{F_n}(Z)| \leqslant (1+\varepsilon)^{|F_n|}.$$

Therefore

$$|\mathcal{B}_{F_n}(\check{Z})| \leqslant (1+\varepsilon)^{|F_n|} 2^{d_n}, \tag{6.1}$$

where

$$d_n \stackrel{\text{def}}{=} \max\left\{ \sum_{g\in F_n} w(g) : w \in \mathcal{B}_{F_n}(Z) \right\}.$$

Clearly, we also have

$$|\mathcal{B}_{F_n}(\check{Z})| \geqslant 2^{d_n}, \tag{6.2}$$

Combining inequalities (6.1) and (6.2), taking logarithm and dividing by $|F_n|$, we obtain

$$\frac{d_n}{|F_n|}\log 2 \leqslant \frac{\log|\mathcal{B}_{F_n}(\check{Z})|}{|F_n|} \leqslant \frac{d_n}{|F_n|}\log 2 + \log(1+\varepsilon).$$

Then, passing to the limit, one has

$$\limsup_{n\to\infty} \frac{d_n}{|F_n|}\log 2 \leqslant \lim_{n\to\infty} \frac{\log|\mathcal{B}_{F_n}(\check{Z})|}{|F_n|} \leqslant \liminf_{n\to\infty} \frac{d_n}{|F_n|}\log 2 + \log(1+\varepsilon).$$

Since ε was arbitrary, we obtain

$$\lim_{n\to\infty} \frac{\log|\mathcal{B}_{F_n}(\check{Z})|}{|F_n|} = \lim_{n\to\infty} \frac{d_n}{|F_n|}\log 2. \qquad \square$$

As our next result we prove that the maximum measure of the identity cylinder is equal, roughly, to the asymptotic maximum density of ones in x. (Recall that by $[1]_e$ we denote the cylinder $\{x \in \{0,1\}^G : x(e) = 1\}$)

Lemma 6.5. *Let G be a countable amenable group. If $Z \subseteq \{0,1\}^G$ is a subshift, then*

$$\max_{\mu \in \mathcal{M}_G(Z)} \mu([1]_e) = \lim_{n\to\infty} \frac{1}{|F_n|} \max\left\{ \sum_{g\in F_n} w(g) : w \in \mathcal{B}_{F_n}(Z) \right\}.$$

Proof. First, note that the limit in the right-hand side exists by Lemma 4.11.

Next, fix a subshift $Z \subseteq \{0,1\}^G$ and denote

$$d_n \stackrel{\text{def}}{=} \max\left\{ \sum_{g\in F_n} w(g) : w \in \mathcal{B}_{F_n}(Z) \right\}.$$

First we show that

$$\lim_{n\to\infty} \frac{d_n}{|F_n|} \geqslant \max_{\mu \in \mathcal{M}_G(Z)} \mu([1]_e).$$

In other words, we prove that for every $\mu \in \mathcal{M}_G(Z)$, we have

$$\alpha \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{d_n}{|F_n|} \geqslant \mu([1]_e).$$

Assume to the contrary that there exists $\mu \in \mathcal{M}_G(Z)$ and $\varepsilon > 0$ such that $\mu([1]_e) = \alpha + \varepsilon$. Then, since μ is G-invariant, for every $n \in \mathbb{N}$ we obtain

$$\sum_{g\in F_n} \mu([1]_g) = |F_n|\mu([1]_e) = |F_n|(\alpha + \varepsilon). \tag{6.3}$$

On the other hand, keepeing in mind that for any $w \in \{0,1\}^{F_n}$, the sum $\sum_{g\in F_n} w(g)$ is just a number of ones occuring in w, we compute

$$\sum_{g\in F_n} \mu([1]_g) = \sum_{g\in F_n} \sum_{\substack{w\in\{0,1\}^{F_n} \\ w(g)=1}} \mu([w]_{F_n}) = \sum_{w\in\{0,1\}^{F_n}} \left(\sum_{g\in F_n} w(g) \right) \mu([w]_{F_n}). \tag{6.4}$$

Combining (6.3) and (6.4) we obtain

$$\sum_{w\in\{0,1\}^{F_n}} \left(\sum_{g\in F_n} w(g) \right) \mu([w]_{F_n}) = |F_n|(\alpha + \varepsilon).$$

Therefore, since

$$\sum_{w\in\{0,1\}^{F_n}} \mu([w]_{F_n}) = 1,$$

there must be $w \in \{0,1\}^{F_n}$ such that

$$\sum_{g\in F_n} w(g) \geqslant |F_n|(\alpha + \varepsilon) \quad \text{and} \quad \mu([w]_{F_n}) > 0,$$

which is a contradiction with the definition of α.

To prove the opposite inequality we construct a measure $\mu \in \mathcal{M}_G(Z)$ satisfying

$$\lim_{n\to\infty} \frac{d_n}{|F_n|} \leqslant \mu([1]_e)$$

as follows: for every $n \in \mathbb{N}$ we choose $w^{(n)} \in \mathcal{B}_{F_n}(Z)$ such that

$$w^{(n)} \in \operatorname{argmax} \left\{ \sum_{g\in F_n} w(g) : w \in \mathcal{B}_{F_n}(Z) \right\}.$$

We also take $x^{(n)} \in Z$ such that $x^{(n)}_{F_n} = w^{(n)}$. Now, we define a sequence $\{\mu_n\}_{n\in\mathbb{N}}$ of measures in $\mathcal{M}(Z)$ by

$$\mu_n = \frac{1}{|F_n|} \sum_{g\in F_n} g_* \delta_{x^{(n)}} = \frac{1}{|F_n|} \sum_{g\in F_n} \delta_{gx^{(n)}}.$$

Since $\mathcal{M}(Z)$ is weak*-compact, the sequence $\{\mu_n\}_{n\in\mathbb{N}}$ has a convergent subsequence $\{\mu_{k_n}\}_{n\in\mathbb{N}}$; call $\mu \in \mathcal{M}(Z)$ its limit. Moreover, as we show in Lemma 5.8, μ is necessarily invariant. For brevity we simply assume $k_n = n$. Since μ is the weak*-limit of $\{\mu_n\}_n$ it follows that for every closed subset $B \subseteq \overline{Gx}$ we have

$$\limsup_{n\to\infty} \mu_n(B) \leqslant \mu(B).$$

Hence, (as $[1]_e$ is closed in $\overline{Gx}$) it is enough to prove that for every $n \in \mathbb{N}$ we have

$$\frac{d_n}{|F_n|} \leqslant \mu_n([1]_e).$$

Notice that for every $g \in G$ and $n \in \mathbb{N}$, we have

$$\delta_{gx^{(n)}}([1]_e) = \begin{cases} 1 & \text{if } gx^{(n)} \in [1]_e \\ 0 & \text{otherwise} \end{cases} = \begin{cases} 1 & \text{if } x^{(n)}(g) = 1 \\ 0 & \text{otherwise} \end{cases}.$$

Therefore

$$\mu_n([1]_e) = \frac{1}{|F_n|} \sum_{g\in F_n} \delta_{gx^{(n)}}([1]_e) = \frac{1}{|F_n|} \sum_{g\in F_n} x^{(n)}(g) = \frac{1}{|F_n|} \sum_{g\in F_n} w^{(n)}(g) = \frac{d_n}{|F_n|},$$

where the last equality holds by the choice of $w^{(n)}$. This finishes the proof. □

The above lemma leads us to a new formula for the entropy of a subordinate shift.

Lemma 6.6. *Let G be a countable amenable group. If $Z \subseteq \{0,1\}^G$ is a subshift such that $h_{\mathrm{top}}(Z) = 0$, then the entropy of its subordinate shift $\check{Z}$ is given by*

$$h_{\mathrm{top}}(\check{Z}) = \max_{\mu\in\mathcal{M}_G(Z)} \mu([1]_e) \log 2.$$

Proof. This is a straigtforward consequence of Theorem 6.1 combined with Lemma 6.5. □

6.3 Periodic Shifts

In this section we assume that G is an amenable residually finite group (Definition 2.10) with a Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ associated to a sequence of normal subgroups $\{H_n\}_{n\in\mathbb{N}} \subseteq \mathcal{P}(G)$ provided by Lemma 2.4.

Definition 6.2 (Periodic configuration)**.** Let H be a subgroup of G and let $F \subseteq G$ be a fundamental domain for G/H. A configuration $x \in \{0,1\}^G$ is **periodic** with period H, if

$$x(g) = x(gh) \quad \text{for every } g \in F \text{ and } h \in H.$$

The aim of this section is to prove the following

Lemma 6.7. *For a periodic configuration $z \in \{0,1\}^G$ with period H_N, the entropy of the subordinate shift $\check{Z}$ of $Z := \overline{Gz}$ is given by*

$$h_{\text{top}}(\check{Z}) = \frac{\log 2}{|F_N|} \sum_{g\in F_N} z(g).$$

To prove this lemma, we want to use Theorem 6.1, hence first we need to show that its assumptions are satisfied.

Lemma 6.8. *Let $N \in \mathbb{N}$. If $x \in \{0,1\}^G$ is periodic with period H_N, then $h_{\text{top}}(\overline{Gx}) = 0$.*

Proof. We prove that (see Lemma 5.7)

$$\lim_{n\to\infty} \frac{\log|\mathcal{B}_{F_n}(x)|}{|F_n|} = 0.$$

Notice that it is enough to show that for arbitrary $g \in G$ we have

$$x_{F_N\tilde{g}} = x_{F_N g} \quad \text{for every } \tilde{g} \in gH_N. \tag{6.5}$$

Indeed, this imples that the number of different configurations appearing in x over shifts of F_N is bounded from above by the number of equivalence classes of G/H_N, which is equal to the cardinality of F_N. By Condition 3. in Lemma 2.4 the same is true for every $n > N$. Then we obtain

$$\lim_{n\to\infty} \frac{\log|\mathcal{B}_{F_n}(x)|}{|F_n|} \leqslant \lim_{n\to\infty} \frac{\log|F_n|}{|F_n|} = 0.$$

To prove (6.5), fix $g \in G$ and $\tilde{g} \in gH_N$. Then there exists $h \in H_N$ such that $\tilde{g} = gh$. Let $f \in F_N$, then

$$x(f\tilde{g}) = x(fgh) = x(fg).$$

□

Next, we prove that shifting a periodic monotile, does not change the number of ones in it.

Lemma 6.9. *Let G be a residually finite group, H its normal subgroup and let F be a fundamental domain for G/H. Define $\Phi\colon G \to F$ by $\Phi(g) = gh^{-1}$, where $h \in H$ is the unique element such that $g \in Fh$. Then $\Phi|_{Fg}$ is bijective for every $g \in G$.*

We refer to Figure 6.1 for an illustration of what Φ is.

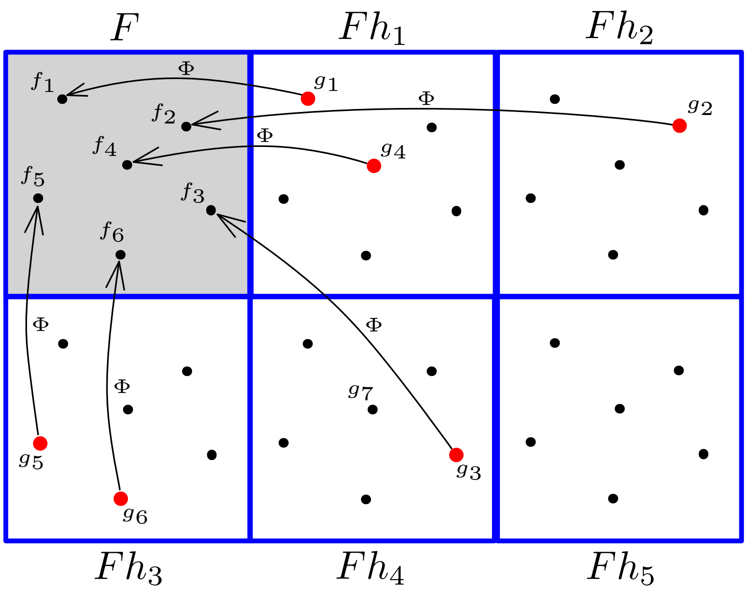

Figure 6.1: We can see a sample set $F = \{f_1, \dots, f_6\}$ and its shifts by some elements $h_1, \dots, h_5 \in H$. An element $g \in G \backslash H$ gives us a set $Fg = \{g_1, \dots, g_6\}$ (marked in red). Arrows show how the map Φ acts on elements of Fg. Hence for instance, we have $\Phi(g_1) = g_1 h_1^{-1} = f_1$ and $\Phi(g_3) = g_3 h_4^{-1} = f_3$.

Proof of Lemma 6.9. Let $f_1, f_2 \in F$ and assume that $\Phi(f_1 g) = \Phi(f_2 g)$. Then there exist $h_1, h_2 \in H$ such that $f_1 g h_1^{-1}$, $f_2 g h_2^{-1} \in F$ and

$$f_1 g h_1^{-1} = f_2 g h_2^{-1}.$$

This implies that

$$f_1 = f_2 g h_2^{-1} h_1 g^{-1}.$$

Clearly, $g h_2^{-1} h_1 g^{-1} \in H$ since H is normal. But $F \cap Fh = \varnothing$ for every $h \in H \backslash \{e\}$. Therefore $g h_2^{-1} h_1 g^{-1} = e$ and thus $f_1 = f_2$. $\square$

Now the proof of Lemma 6.7 is an almost straightforward consequence of the above considerations.

Proof of Lemma 6.7. Since $z \in \{0,1\}^G$ is periodic, it satisfies $h_{\text{top}}(\overline{Gz}) = 0$ (Lemma 6.8.), hence we can use Theorem 6.1. Notice that it is enough to prove that

$$\frac{1}{|F_N|} \max \left\{ \sum_{g \in F_N} w(g) : w \in \mathcal{B}_{F_N}(\overline{Gz}) \right\} = \frac{1}{|F_N|} \sum_{g \in F_N} z(g)$$

because then, by periodicity of z, for every $n > N$ we obtain

$$\frac{1}{|F_n|} \max \left\{ \sum_{g \in F_n} w(g) : w \in \mathcal{B}_{F_n}(\overline{Gz}) \right\} = \frac{1}{|F_N|} \sum_{g \in F_N} z(g).$$

Clearly we have

$$\max \left\{ \sum_{g \in F_N} w(g) : w \in \mathcal{B}_{F_N}(\overline{Gz}) \right\} = \max \left\{ \sum_{g \in F_N h} z(g) : h \in G \right\}.$$

Let $\Phi\colon G \to F_N$ satisfy $\Phi(g) = gh^{-1}$, where $h \in H_N$ is the unique element such that $g \in F_N h$. By Lemma 6.9. we know that $\Phi|_{F_N g}$ is bijective for every $g \in G$ (see Figure 6.1) which means that

$$\sum_{g \in F_N h} z(g) = \sum_{g \in F_N} z(g) \quad \text{for every } h \in G. \qquad \square$$

Chapter 7

Quasi-Toeplitz Configurations

In this chapter we develop tools necessary in our approach of proving Theorem 8.1. More specifically, we recall the notion of a Toeplitz sequence (see [35], [19]) and its generalization[1]: a quasi-Toeplitz configuration (see [8]) and study their properties. Most importantly, we show that the orbit closure of a quasi-Toeplitz configuration forms a minimal subsystem (see also [8]), and further that the set of all quasi-Toeplitz configurations is path-connected with respect to the Weyl pseudometric (defined in Section 4.2). These are the two main observations that allow us later to derive Theorem 8.1.

7.1 Definition

In this chapter we assume that G is a congruent monotileable, hence amenable, group. We work with the shift space $\mathscr{A}^G$ where $\mathscr{A}$ is a finite alphabet. What follows is a definition of a quasi-Toeplitz configuration.

Definition 7.1 (Quasi-Toeplitz Configuration)**.** An element $x \in \mathscr{A}^G$ is called a **quasi-Toeplitz configuration** if there exists a congruent Følner sequence $\{F_n\}_{n\in\mathbb{N}}$ with an associated elegant sequence of centers $\{C_n\}_{n\in\mathbb{N}}$ (see Lemma 2.3) such that for every $g \in G$ there exists $m \in \mathbb{N}$ such that $x(g) = x(gc)$ for any $c \in C_m$.

It is also convenient to introduce the following notation: given $C \subseteq G$ and $x \in \mathscr{A}^G$ we define

$$\mathrm{Per}_C(x) \stackrel{\text{def}}{=} \big\{g \in G \,:\, x(g) = x(gh) \text{ for every } h \in C\big\}.$$

With this notation it is not hard to see that $x \in \mathscr{A}^G$ is a quasi-Toeplitz configuration if and only if there exists a congruent Følner sequence $\{F_n\}_{n\in\mathbb{N}}$ with an elegant sequence of centers $\{C_n\}_{n\in\mathbb{N}}$ satisfying

$$\bigcup_{n=0}^{\infty} \mathrm{Per}_{C_n}(x) = G.$$

Quasi-periodic configuration is a special case of quasi-Toeplitz configurations.

[1]Toeplitz sequence is a two-sided sequence, hence a function defined over $\mathbb{Z}$, such that its every element occurs periodically. Its first generalization is a Toeplitz configuration, which is defined over an amenable residually finite group G instead of $\mathbb{Z}$. Toeplitz configurations were widely studied in [11], [10], [38].

Definition 7.2 (Quasi-periodic configuration)**.** An element $x \in \mathscr{A}^G$ is a **quasi-periodic configuration** if there exists a congruent Følner sequence $\{F_n\}_{n \in \mathbb{N}}$ with an elegant sequence of centers $\{C_n\}_{n \in \mathbb{N}}$ and there exists $N \in \mathbb{N}$ such that for every $g \in G$ it holds $x(g) = x(gc)$ for any $c \in C_N$. Then we say that x is quasi-periodic with period C_N.

7.2 Examples

Before we proceed with the study of properties of quasi-Toeplitz configurations we give some examples to provide intuitions and help in visualizing this important notion. We discuss two special cases: when $G = \mathbb{Z}$ and when $G = \mathbb{Q}$.

7.2.1 Toeplitz Sequences over $\mathbb{Z}$

In the case of $G = \mathbb{Z}$, quasi-Toeplitz configurations coincide with the notion of Toeplitz sequences and are well studied in the literature [35], [19], [21], [14]. In fact, the definition of quasi-Toeplitz configurations comes as a generalization of Toeplitz sequences from $\mathbb{Z}$ to more general groups. For clarity we repeat the definition for this special case.

Definition 7.3 (Toeplitz Sequence)**.** We call $x \in \mathscr{A}^{\mathbb{Z}}$ a Toeplitz seqence if each of its elements occurs periodically, i.e., for every $n \in \mathbb{Z}$ there exists $p \in \mathbb{N}$ such that we have $x(n + ip) = x(n)$ for every $i \in \mathbb{Z}$.

Remark 7.1. Notice that if $x \in \mathscr{A}^{Z}$ satisfies the above definition then it also satisfies Definition 7.1 in case $G = \mathbb{Z}$. Indeed, let $F_n = \{0, 1, \ldots, n! - 1\}$ and $C_n = n!\mathbb{Z}$ for every $n \in \mathbb{N}$. If $x(n) \in \mathscr{A}$ (for some $n \in \mathbb{N}$) occurs with some period $p \in \mathbb{N}$, then $x(n) = x(n + c)$ for every $c \in C_p$.

We now proceed with a construction of Toeplitz sequences that is in a sense "complete", i.e., each sequence constructed this way is Toeplitz, and also each Toeplitz sequence can be constructed this way (by an appropriate choice of parameters of the construction).

Construction of a Toeplitz sequence over $\mathbb{Z}$.

We construct Toeplitz sequences inductively. For this, it is convenient to temporarily "add" to the alphabet $\mathscr{A}$ a symbol "$\star$" denoting "empty place" and obtain a new alphabet $\mathscr{A}_\star = \mathscr{A} \cup \{\star\}$.
Base step. We start with a sequence of "empty places" $x_0 \in \mathscr{A}_\star^{\mathbb{Z}}$ satisfying $x_0 \equiv \star$. Now, the idea is to "exhaust" all the empty places one by one.
Induction step. In the n-th step (with $n \geqslant 1$) we construct $x_n \in \mathscr{A}_\star^{\mathbb{Z}}$ as follows: if there is no star left in x_{n-1} then we end the construction by setting $x_n = x_{n+1} = x_{n+2} = \ldots = x_{n-1}$. Otherwise we pick an index $k_n \in \mathbb{Z}$ to be the k such that $x_{n-1}(k) = \star$ and k has the smallest absolute value out of all such (with ties resolved arbitrarily).

Next, we choose an element $a_n \in \mathscr{A}$ and assign it to position k_n. Finally, we need to decide on the period $p_n \in \mathbb{N}$ that determines how often the newly assigned element will repeat. Note that p_n needs to satisfy the condition that all the positions among $k_n + p_n\mathbb{Z}$ are occupied by $\star$'s in x_{n-1}, otherwise we would override some previously set value. It is not hard to see that indeed such a period always exists. At the end, we modify x_{n-1} by "filling empty places" with a_n at appropriate positions to obtain x_n, thus formally

$$x_n(k) = \begin{cases} a_n & \text{if } k \in k_n + p_n\mathbb{Z}, \\ x_{n-1}(k) & \text{otherwise.} \end{cases}$$

$$
\begin{array}{llcccccccccccccccc}
 & x_0: & \dots & \star & \star & \star & \star & \star & \star & .\star & \star & \star & \star & \star & \star & \star & \star & \dots \\
p_1 = 3, & x_1: & \dots & \boldsymbol{a_1} & \star & \star & \boldsymbol{a_1} & \star & \star & .\boldsymbol{a_1} & \star & \star & \boldsymbol{a_1} & \star & \star & \boldsymbol{a_1} & \star & \dots \\
p_2 = 6, & x_2: & \dots & a_1 & \boldsymbol{a_2} & \star & a_1 & \star & \star & .a_1 & \boldsymbol{a_2} & \star & a_1 & \star & \star & a_1 & \boldsymbol{a_2} & \dots \\
p_3 = 6, & x_3: & \dots & a_1 & a_2 & \star & a_1 & \star & \boldsymbol{a_3} & .a_1 & a_2 & \star & a_1 & \star & \boldsymbol{a_3} & a_1 & a_2 & \dots \\
p_4 = 12, & x_4: & \dots & a_1 & a_2 & \star & a_1 & \star & a_3 & .a_1 & a_2 & \boldsymbol{a_4} & a_1 & \star & a_3 & a_1 & a_2 & \dots \\
p_5 = 6, & x_5: & \dots & a_1 & a_2 & \star & a_1 & \boldsymbol{a_5} & a_3 & .a_1 & a_2 & a_4 & a_1 & \boldsymbol{a_5} & a_3 & a_1 & a_2 & \dots
\end{array}
$$

Figure 7.1: An example of a construction of a Toeplitz sequence over $\mathbb{Z}$.

Limit step. Having $\{x_n\}_{n\in\mathbb{N}}$ constructed, we set

$$x = \lim_{n\to\infty} x_n.$$

We refer to Figure 7.1 for an example of how such a construction might look like.

7.2.2 Quasi-Toeplitz Configurations over $\mathbb{Q}$

The second example we would like to demonstrate is when $G = (\mathbb{Q}, +)$ (for simplicity we also choose $\mathscr{A} = \{0, 1\}$). We note that this case is fundamentally different to $\mathbb{Z}$, as the additive group of rationals is not residually finite (since it is abelian and not finitely generated). Still, $\mathbb{Q}$ is congruent monotileable, which, as it turns out, is enough to prove interesting properties of quasi-Toeplitz configurations in this setting.

We would like to assign 0's and 1's to rational numbers such that they appear with some "period" given by a set of centers associated with a certain monotile. Therefore we first need to construct a congruent Følner sequence. Before we proceed with the construction, we introduce some useful notation. In this section we denote intervals of rational numbers by (a, b), instead of $(a, b) \cap \mathbb{Q}$. We start with a definition of a set that serves as a fundamental building block for creating monotiles

$$J(a, b, k) \stackrel{\text{def}}{=} [a, b) \cap \left\{ \frac{j}{k} : j \in \mathbb{Z} \right\} \qquad \text{for } a, b \in \mathbb{Z} \quad k \in \mathbb{N}\backslash\{0\}.$$

Notice that $J(a, b, k)$ is a monotile since

$$C(a, b, k) \stackrel{\text{def}}{=} \left[0, \frac{1}{k}\right) + (b - a)\mathbb{Z}$$

is the set of centers associated with $J(a, b, k)$. We define the following two operations on $J(a, b, k)$ (see also Figure 7.2 for a graphical explanation):

1. Extend $[J(a, b, k)] \stackrel{\text{def}}{=} (J(a, b, k) - (b - a)) \sqcup J(a, b, k) \sqcup (J(a, b, k) + (b - a))$,
2. Condense $[J(a, b, k), p] \stackrel{\text{def}}{=} \displaystyle\bigsqcup_{j=0}^{p-1} \left(J(a, b, k) + \frac{j}{kp} \right)$, for any positive integer p.

Observe that the operation Extend triples the length of $J(a, b, k)$ by attaching both to the left and to the right of it a shifted copy of it. The Condense operation makes p copies of the set $J(a, b, k)$ and puts them one after another with spacing $\frac{1}{kp}$. Moreover, the following relations hold

$$\begin{aligned}
&\text{Extend}\,[J(a, b, k)] = J(a - (b - a), b + (b - a), k),\\
&\text{Condense}\,[J(a, b, k), p] = J(a, b, kp).
\end{aligned}$$

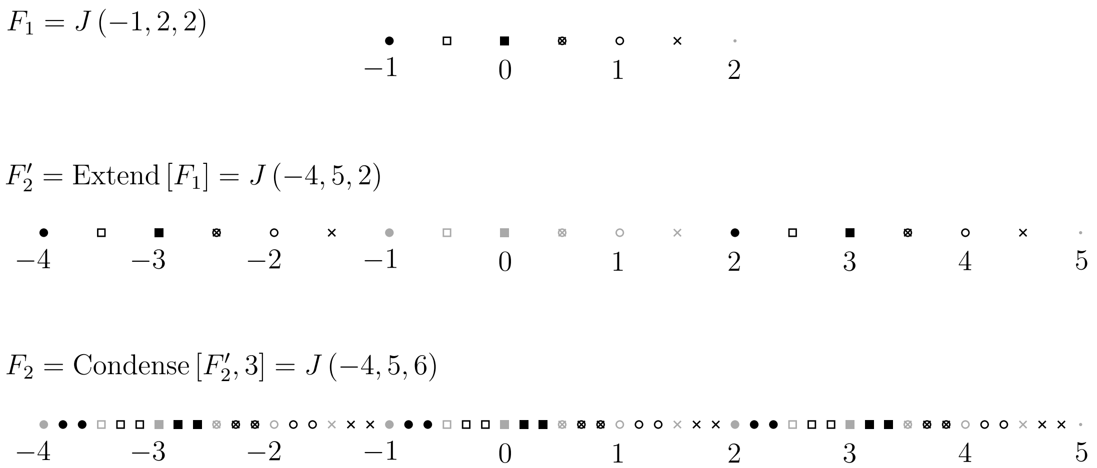

Figure 7.2: An illustration of the second step of the construction of a congruent Følner seqence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ in $(\mathbb{Q}, +)$ with $\{q_i\}_{i\in\mathbb{N}}$ such that $q_0 = 2$ and $q_1 = 3$. We start with F_1 already constructed. Its elements are marked with different symbols to better illustrate how does F_2 arise by copying and shifting F_1. New copies obtained as a result of the Extend and Condense operations respectively are marked with black, and their arguments are grey.

Note that both these operations given a monotile create a new monotile by copying and shifting a certain number of copies of the initial monotile. It is therefore not a surprise that we can use these operations to construct a congruent Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$. Essentially, the only non-trivial constraint to satisfy is that $\bigcup \mathcal{F} = \mathbb{Q}$. Towards this end, we need to keep extending our monotile and carefully use the condense operation. For this purpose define a sequence of primes $\{q_i\}_{i\in\mathbb{N}} \subseteq \mathbb{N}$ such that every prime number occurs infinitely many times, that is, for every $p \in \mathbb{P}$ the set $\{i : q_i = p\}$ is infinite (where $\mathbb{P}$ denotes the set of primes).

Now we are ready to construct $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$. The construction is inductive. In the base step $n = 0$ we put $F_0 \stackrel{\text{def}}{=} J(0,1,1)$, that is, $F_0 = \{0\}$. Next assume that we have already constructed F_n for some $n \in \mathbb{N}$. We define F_{n+1} as follows (see Figure 7.2)

$$F'_{n+1} \stackrel{\text{def}}{=} \text{Extend}[F_n],$$
$$F_{n+1} \stackrel{\text{def}}{=} \text{Condense}\left[F'_{n+1}, q_n\right].$$

We now justify that such a sequence $\mathcal{F}$ indeed satisfies $\bigcup \mathcal{F} = \mathbb{Q}$. Let $p, q \in \mathbb{Z}$. We show that there exists $N \in \mathbb{N}$ such that $\frac{p}{q} \in F_N$. For $n \in \mathbb{N}$ define α_n, a_n, b_n such that

$$F_n = J(a_n, b_n, \alpha_n).$$

It is not hard to verify by induction that $a_n = -\frac{1}{2}(3^n - 1)$, $b_n = \frac{1}{2}(3^n + 1)$ and $\alpha_n = \prod_{i=0}^{n-1} q_i$. Since, $\{q_i\}_{i\in\mathbb{N}}$ contains every prime infinitely many times, there exists $n_1 \in \mathbb{N}$ such that q divides α_{n_1}. We can also find $n_2 \in \mathbb{N}$ satisfying

$$a_{n_2} < \frac{p}{q} < b_{n_2},$$

since clearly $a_n \to -\infty$ and $b_n \to \infty$. By taking $N \stackrel{\text{def}}{=} \max\{n_1, n_2\}$ we obtain that, by definition, $\frac{p}{q} \in F_N$.

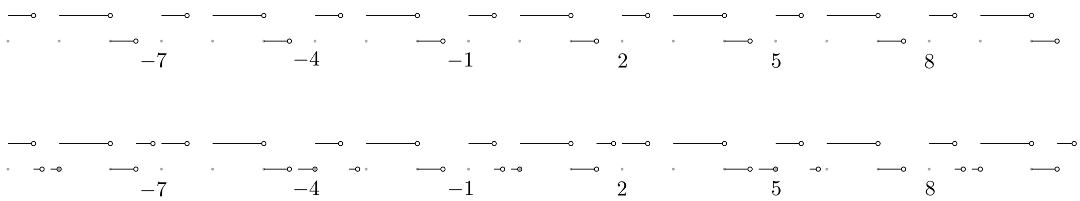

Figure 7.3: An illustration of the initial two steps of the construction of a quasi-Toeplitz configuration in $\mathbb{Q}$.

It is easy to see from the construction of $\mathcal{F}$ that remaining conditions from Definition 2.9 and the Følner condition are satisfied. Therefore $\{F_n\}_{n\in\mathbb{N}}$ is a congruent Følner sequence.

Now we construct a quasi-Toeplitz configuration $x \in \{0,1\}^{\mathbb{Q}}$ with respect to the congruent Følner sequence $\mathcal{F}$ in a similar way to how we built a Toeplitz sequence in $\{0,1\}^{\mathbb{Z}}$. Fix an arbitrary enumeration of elements of $\mathbb{Q}$. Denote by $\{C_n\}_{n\in\mathbb{N}}$ the sequence of centers associated with the sets in $\mathcal{F}$.

In the base step $n = 0$ we arbitrarily define $x(g)$ for g in some arbitrarily chosen set $G_0 \subseteq F_0$ (in particular, we might choose $G_0 = \varnothing$). Then copy this "pattern to the set $G_0 + C_0 \subseteq \mathbb{Q}$, that is, put $x(g+h) \overset{\text{def}}{=} x(g)$ for $g \in G_0$ and $h \in C_0$. Therefore in the base step we define x on the set $G_0 + C_0$, i.e., set $x(g+c) = x(g)$ for all $g \in G_0$ and $c \in C_0$. At the start of the $(n+1)$-th step of the construction we have x defined on the set of indices

$$T_n \overset{\text{def}}{=} \bigsqcup_{i=0}^{n} (G_i + C_i).$$

We simply choose a new set $G_{n+1} \subseteq F_{n+1} \backslash T_n$ such that in the first place we choose available elements from F_{n+1} with the smallest index (according to the enumeration of $\mathbb{Q}$). Next we define x on G_{n+1} and extend this "pattern" to the set $G_{n+1} + C_{n+1}$, that is, we set $x(g+h) \overset{\text{def}}{=} x(g)$ for $g \in G_{n+1}$ and $h \in C_{n+1}$.

Note that choosing T_n according to the enumeration of $\mathbb{Q}$ for every $n \in \mathbb{N}$, guarantees that $\bigcup_{n\in\mathbb{N}} T_n = \mathbb{Q}$, thus x is defined over all elements in $\mathbb{Q}$.

As an example, we show two steps of the above construction with $\mathcal{F}$ defined for $\{q_i\}_{i\in\mathbb{N}}$ satisfying $q_0 = 2$ and $q_1 = 3$ (see Figure 7.3). For F_1 and F_2 in this setting, see Figure 7.2. Choose $G_0 \overset{\text{def}}{=} \varnothing$. Subsequently, we let $G_1 \overset{\text{def}}{=} \{-1, 0, \frac{1}{2}, 1\}$. We define x on G_1 and extend this to $G_1 + C_1$ (for the formula for C_1 see (7.2.2))

$$x(r) \overset{\text{def}}{=} \begin{cases} 1 & \text{if } r \in \{-1, 0, \frac{1}{2}\} + [0, \frac{1}{2}) + 3\mathbb{Z}, \\ 0 & \text{if } r \in 1 + [0, \frac{1}{2}) + 3\mathbb{Z}. \end{cases}$$

Next, we define $G_2 \subseteq F_2$. We work here with an enumeration of elements in G that is consistent with $\mathcal{F}$ and thus we choose to G_2 elements from $F_1 \backslash G_1$, we set $G_2 \overset{\text{def}}{=} \{-3\frac{2}{6}, -\frac{1}{2}, -\frac{1}{6}, 1\frac{1}{2}, 1\frac{4}{6}, 4\frac{4}{6}, 4\frac{5}{6}\}$. Then we define x on G_2 and extend it to $G_2 + C_2$ as follows

$$x(r) \overset{\text{def}}{=} \begin{cases} 1 & \text{if } r \in \{1\frac{1}{2}, 1\frac{4}{6}\} + [0, \frac{1}{6}) + 9\mathbb{Z}, \\ 0 & \text{if } r \in \{-3\frac{2}{6}, -\frac{1}{2}, -\frac{1}{6}, 4\frac{4}{6}, 4\frac{5}{6}\} + [0, \frac{1}{6}) + 9\mathbb{Z}. \end{cases}$$

7.3 Properties

We are ready to prove some interesting properties of quasi-Toeplitz configurations that are instrumental in our proof of Theorem 8.1. The first property is that orbits of quasi-Toeplitz configurations yield minimal subsystems.

Lemma 7.1. *If $x \in \mathscr{A}^G$ is a quasi-Toeplitz configuration with respect to a congruent Følner sequence $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$, then $\overline{Gx}$ is a minimal set.*

Proof. By Theorem 3.1 it is enough to show that $N(x, \{x\}^\varepsilon)$ is syndetic for every $\varepsilon > 0$. (Recall that $N(x, A) = \{g \in G : gx \in A\}$, where $x \in X$ and $A \subseteq X$; and B^ε is an ε-hull of a set $B \subseteq X$.)

Fix $\varepsilon > 0$ and take $F \subseteq G$ such that $|F|^{-1} < \varepsilon$. It suffices to prove that the set

$$\{g \in G : gx_F = x_F\}$$

is syndetic. To this end, it is enough to show that

$$C_n \subseteq \{g \in G : gx_F = x_F\} \tag{7.1}$$

for some $n \in \mathbb{N}$ since C_n is syndetic ($F_nC_n = G$). Choose $N \in \mathbb{N}$ such that $F \subseteq F_N$. By the definition of x, there exists $m \geqslant N$ such that $x_{F_N} = x_{F_Nc}$ for every $c \in C_m$. We show that C_m satisfies (7.1) (i.e., one can take $n := m$). Take any $c \in C_m$, then indeed $F \subseteq F_N$ implies

$$cx_F = x_{Fc} = x_F. \qquad \square$$

Now we show that the family of quasi-Toeplitz configurations is path-connected with respect to the Weyl pseudometric, which is a consequence of the following lemma.

Lemma 7.2. *If G is a congruent monotileable group then there exists a function $\Psi\colon [0,1] \to \{0,1\}^G$ such that $\Psi(0) = 0^G$, $\Psi(1) = 1^G$ and for every $s, t \in [0,1]$, $\Psi(s), \Psi(t)$ are quasi-Toeplitz configurations satisfying $D^\star(\Psi(s), \Psi(t)) \leqslant |s - t|$.*

Proof. Let $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ be a congruent Følner sequence in G with sets $\{J_n\}_{n\in\mathbb{N}}$ as in Definition 2.9 and associated elegant sequence of centers $\{C_n\}_{n\in\mathbb{N}}$. Since G is countable, we can enumerate its elements as $g_1, g_2, g_3, \ldots$ such that if $g_i \in F_n$ and $g_j \in F_{n+1}\backslash F_n$, then $i < j$. Fix $t \in [0,1]$. We define $\Psi(t) = x^{(t)} \in \{0,1\}^G$ inductively.

In the base step we put $x_0^{(t)} \stackrel{\text{def}}{=} \star^G \in \{0,1,\star\}^G$ (by the symbol "$\star$" we understand an "empty place"). In the $(n+1)$-th step we define a quasi-periodic configuration $x_{n+1}^{(t)} \in \{0,1,\star\}^G$ such the number of 1's in F_{n+1} divided by $|F_{n+1}|$ approximates t from below with accuracy $\frac{1}{|F_n|}$ (see Figure 7.4).

Assume that we have already constructed some $x_n^{(t)}$. Denote by k_{n+1} the number of $\star$'s appearing in $x_n^{(t)}$ over F_{n+1}. We construct $x_{n+1}^{(t)}$ by substituting all (but one) of these $\star$'s with 0's and 1's and then copying this new pattern we obtained over F_{n+1} to the whole group G. We proceed as follows, choose $l_{n+1} \in \{0, 1, \ldots, k_{n+1}\}$ such that the number $\frac{l_{n+1}}{k_{n+1}}$ approximates t from below. Then we substitute l_{n+1} stars in F_{n+1} with 1's, $k_{n+1} - l_{n+1} - 1$ stars we substitute with 0's and one remaining star we leave without change. Next we copy the pattern we defined over F_{n+1} to the whole group G. Finally, we define $x^{(t)}$ as a quasi-uniform limit of the sequence $\left\{x_n^{(t)}\right\}_{n\in\mathbb{N}}$.

Formally, let $D_0(t) = \varnothing$ and $E_0(t) = \varnothing$. In the $(n+1)$-th step we construct sets $D_{n+1}(t) \subseteq G$ and $E_{n+1}(t) \subseteq G$ which determine positions of 1's and 0's in $x_{n+1}^{(t)}$ respectively. Assume that we

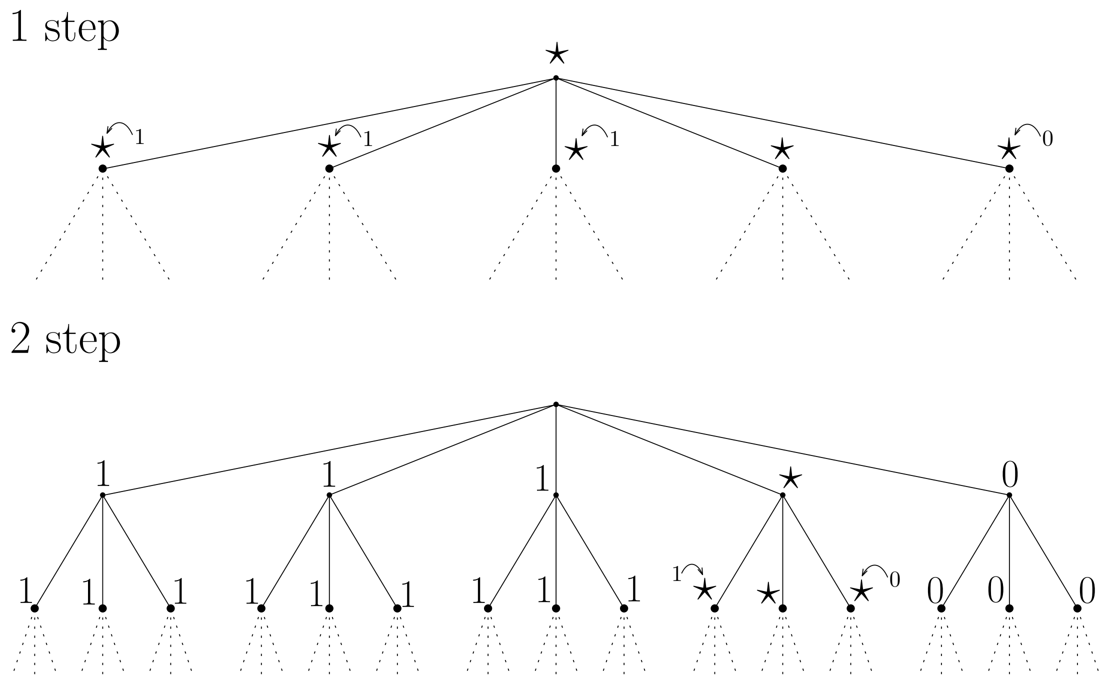

Figure 7.4: An illustration of the first two steps of the construction from the proof of Lemma 7.2 for $t = 0.7$. (For the construction of the tree graph of G see Figure 2.1.) Sample set F_1 consists of 5 elements (all vertices of depth 1). The initial state in the first step is F_1 filled with $\star$'s. We need to decide how many $\star$'s we replace with 1's and how many with 0's. Hence we need to find $k = 0, 1, \ldots, 5$ satisfying $\frac{k}{5} < 0.7 \leqslant \frac{k+1}{5}$. Clearly, $k = 3$. Therefore the first three $\star$'s we replace with 1's, the next star we leave without change, and the remaining star we replace by zero (according to the global ordering of G). Next we copy this pattern using C_1 to the all G. In the second step we are interested in vertices of depth 2. There are three $\star$'s in F_2 hence we need to choose $k = 0, 1, 2, 3$ such that $\frac{9+k}{15} < 0.7 \leqslant \frac{9+k+1}{15}$. Thus $k = 1$, which means that the first star we replace by 1, the next star we leave without change, and the remaining star we replace by 0. Then we copy this pattern using C_2 to the all G.

have already defined sets $D_n(t)$ and $E_n(t)$ for some $n \in \mathbb{N}$. Let $f \in G$ be the position occupied by the "star" in F_n after the n-th step of the construction, that is,

$$\{f\} \stackrel{\text{def}}{=} F_n \backslash (D_n(t) \cup E_n(t)).$$

Then fJ_n is the set of positions occupied by $\star$'s in F_{n+1} after the n-th step of the construction. Denote $s := |J_n|$ and let $f_1, f_2, \ldots, f_s$ be the elements of fJ_n indexed as given by the global ordering of G. Denote by p the number of 1's in F_{n+1} after the n-th step of the construction, i.e.,

$$p \stackrel{\text{def}}{=} |D_n(t) \cap F_{n+1}|.$$

Next we choose the number of $\star$'s in F_{n+1} that we replace with 1's:

$$q \stackrel{\text{def}}{=} \max\left\{\tilde{q} \,:\, \tilde{q} \in \{0, 1, \ldots, s\} \text{ and } \frac{p + \tilde{q}}{|F_{n+1}|} < t\right\}, \tag{7.2}$$

and define

$$D_{n+1}(t) \stackrel{\text{def}}{=} D_n(t) \cup \{f_1, f_2, \ldots, f_q\}\, C_{n+1},$$
$$E_{n+1}(t) \stackrel{\text{def}}{=} E_n(t) \cup \{f_{q+2}, f_{q+3}, \ldots, f_s\}\, C_{n+1}.$$

At the end we have constructed sequences

$$D_0(t) \subseteq D_1(t) \subseteq D_2(t) \subseteq \ldots \qquad \text{and} \qquad E_0(t) \subseteq E_1(t) \subseteq E_2(t) \subseteq \ldots.$$

Denote their unions by:

$$D(t) \stackrel{\text{def}}{=} \bigcup_{n \in \mathbb{N}} D_n(t) \quad \text{and} \quad E(t) \stackrel{\text{def}}{=} \bigcup_{n \in \mathbb{N}} E_n(t).$$

Notice that $D(t) \cup E(t) = G$ and $D(t) \cap E(t) = \varnothing$. Then $x^{(t)} \in \{0, 1\}^G$ given by

$$x^{(t)}(g) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } g \in D(t), \\ 0 & \text{if } g \in E(t) \end{cases}$$

is a quasi-Toeplitz configuration. Hence the function $\Psi(t) = x^{(t)}$ satisfies the first two claims. To justify that it is quasi-continuous (i.e., continuous with respect to D_W), fix $s, t \in [0, 1]$ with $s < t$. Note that if $D_n(s)$ and $D_n(t)$ are as in the construction above, then $D_n(s) \subseteq D_n(t)$ for every $n \in \mathbb{N}$. To prove this, one has to consider the first step of the construction $N \in \mathbb{N}$ such that $D_N(t) \neq D_N(s)$. Next, observe that $D_N(s) \subseteq D_N(t)$. Then proceed inductively. This fact implies that

$$D^\star\left(x^{(s)}, x^{(t)}\right) = D^\star\left(\left\{g \in G \,:\, x^{(s)}(g) \neq x^{(t)}(g)\right\}\right) = D^\star(D(t) \backslash D(s)).$$

Now we show that $D^\star(D(t) \backslash D(s)) = t - s$. To this end, observe that Lemma 4.5 applied to $E := D_n(t) \cap F_n$ and $F := F_n$ yields

$$D^\star(D_n(t)) = \frac{|D_n(t) \cap F_n|}{|F_n|}.$$

Hence, by the construction (see (7.2)) and again applying Lemma 4.5, one has

$$D^\star(D_n(t)) < t \leqslant D^\star(D_n(t)) + \frac{1}{|F_n|} = D^\star(D_n'(t)), \tag{7.3}$$

where $D'_n(t) := G\backslash E_n(t)$. Since $D_n(t)\backslash D'_n(s) \subseteq D(t)\backslash D(s)$ for every $n \in \mathbb{N}$ and $D'_n(s) \subseteq D_n(t)$ for n large enough, combining (7.3) with Lemma 4.6 we obtain

$$t - s - \frac{2}{|F_n|} \leqslant D^\star(D(t)\backslash D(s)) \tag{7.4}$$

for n large enough. Analogously we show that $D(t)\backslash D(s) \subseteq D'_n(t)\backslash D_n(s)$ implies

$$D^\star(D(t)\backslash D(s)) \leqslant t - s + \frac{2}{|F_n|}. \tag{7.5}$$

Combining (7.4), (7.5) and passing with n to ∞ yields $D^\star(D(t)\backslash D(s)) = t - s$. □

Lemma 7.3. *Let $\mathscr{A}$ be a finite alphabet and let $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}}$ be a congruent Følner sequence with associated elegant sequence of centers $\mathcal{C} = \{C_n\}_{n\in\mathbb{N}}$. Then the family of quasi-Toeplitz configurations with respect to $\mathcal{F}$ and $\mathcal{C}$ over $\mathscr{A}^G$ is D_W-path-connected.*

Proof. We can realize any finite alphabet $\mathscr{A}$ as a discrete subset of $\mathbb{R}$, i.e., assume that $\mathscr{A} \subseteq \mathbb{R}$. Pick two quasi-Toeplitz configurations $z, z' \in \mathscr{A}^G$ with respect to $\mathcal{F}$ and $\mathcal{C}$. We construct a path

$$\{u^{(t)} : t \in [0,1]\} \subseteq \mathscr{A}^G$$

connecting z with z'. Let $\Psi\colon [0,1] \to \{0,1\}^G$ and $x^{(t)} \in \{0,1\}^G$ for $t \in [0,1]$ be defined as in Lemma 7.2 with $\mathcal{F}$ and $\mathcal{C}$ in the construction. For $t \in [0,1]$ define $u^{(t)} \in \mathscr{A}^G$ by

$$u^{(t)}(g) \stackrel{\text{def}}{=} x^{(t)}(g)z(g) + (1 - x^{(t)}(g))z'(g).$$

Note that $D^\star(u^{(s)}, u^{(t)}) \leqslant D^\star(x^{(s)}, x^{(t)})$ and $D^\star(x^{(s)}, x^{(t)}) \to 0$ as $s \to t$. Now we show that $u^{(t)}$ is a quasi-Toeplitz configuration for every $t \in [0,1]$. Fix $t \in [0,1]$ and $g \in G$. We will find $N \in \mathbb{N}$ such that $u^{(t)}(gh) = u^{(t)}(g)$ for every $h \in C_N$. Since $x^{(t)}$, z and z' are quasi-Toeplitz configurations with respect to $\mathcal{F}$, there exist $n_1, n_2, n_3 \in \mathbb{N}$ such that $g \in \mathrm{Per}_{C_{n_1}}(x^{(t)})$, $g \in \mathrm{Per}_{C_{n_2}}(z)$ and $g \in \mathrm{Per}_{C_{n_3}}(z')$. Choose $N = \max\{n_1, n_2, n_3\}$. Then $C_N \subseteq C_{n_i}$ for $i = 1, 2, 3$. Therefore for every $h \in C$ one has

$$u^{(t)}(gh) = x^{(t)}(gh)z(gh) + (1 - x^{(t)}(gh))z'(gh) = x^{(t)}(g)z(g) + (1 - x^{(t)}(g))z'(g) = u^{(t)}(g). \quad \square$$

We conclude this section with an interesting, auxiliary result on regular quasi-Toeplitz configurations. Let us start by a definition.

Definition 7.4. We call a quasi-Toeplitz configuration $x \in \mathscr{A}^G$ **regular** if there exists a congruent sequence $\{F_n\}_{n\in\mathbb{N}}$ with an elegant sequence of centers $\{C_n\}_{n\in\mathbb{N}}$ such that

$$\sup_{n\in\mathbb{N}} D^\star\big(\mathrm{Per}_{C_n}(x)\big) = 1.$$

Lemma 7.4. *Every regular quasi-Toeplitz configuration is a quasi uniform limit of a sequence of quasi-periodic configurations.*

Proof. Let $x \in \mathscr{A}^G$ be a regular quasi-Toeplitz configuration with a congruent sequence $\{F_n\}_{n\in\mathbb{N}}$ and an elegant sequence of centers $\{C_n\}_{n\in\mathbb{N}}$. Notice that if $m \leqslant n$, then $\mathrm{Per}_{C_m}(x) \subseteq \mathrm{Per}_{C_n}(x)$ for $m, n \in \mathbb{N}$. Therefore

$$1 = \sup_{n\in\mathbb{N}} D^\star(\mathrm{Per}_{C_n}(x)) = \lim_{n\to\infty} D^\star(\mathrm{Per}_{C_n}(x)).$$

Define $x^{(n)} \in \mathscr{A}^G$ as $x^{(n)}(fc) = x(f)$ for every $f \in F_n$ and $c \in C_n$. Then

$$\{g \in G : x^{(n)}(g) \neq x(g)\} \subseteq G\backslash\mathrm{Per}_{C_n}(x),$$

which implies

$$D^\star(x^{(n)}, x) \leqslant 1 - D^\star(\mathrm{Per}_{C_n}(x)) \to 0 \quad \text{as} \quad n \to \infty. \quad \square$$

Chapter 8

Realizability of Entropy

In this chapter we finally combine various tools developed throughout this thesis to deduce its two main results: the realizability of entropy for minimal subsystems and proximal subsystems. We hope that the theory developed around these tools in the previous chapters is already interesting by itself. It is this particular application that indicates that this theory is also useful and allows to both obtain new, simpler proofs of old results and to prove new theorems.

8.1 Minimal Subsystems

In this section, we show that each possible finite value of entropy is realizable by some minimal dynamical subsystem of $\mathscr{A}^G$ assuming only that G is a congruent monotilable group. Recall that this result was initially proved by Krieger [37] and later reproved by [9] assuming that the group G is residually finite. Our generalization includes in particular all virtually nilpotent and all abelian groups; perhaps the simplest example that is captured now, but was not previously is $(\mathbb{Q}, +)$ (the additive group of rationals).

Recently, it was brought to author's attention that a result by Rosenthal in an unpublished manuscript[1] [52] might likely be used to give a different proof of the Theorem 8.1. Specifically, Rosenthal's main theorem asserts that if (Y, Σ, ν) is a standard probability space with a free ergodic measure preserving action of an amenable group G, then there exists a compact metric space X and an action of G on X preserving a Borel probability measure μ such that (X, G) is a minimal dynamical system and μ is unique measure invariant for the action of G on X. Furthermore, the ergodic measure preserving action of G on X is isomorphic with the action of G on Y. While the theorem does not specify what the space X is, one might try to repeat the proof (thus also go through the argument of [31], [53] and [59]) associating with Y a minimal subshift of $\{0, 1, \ldots, d-1\}^G$ whenever the Kolmogorov-Sinai entropy of (Y, ν, G) is strictly smaller than $\log d$.

Theorem 8.1 (Realizability of entropy for minimal systems)**.** *Let G be a congruent monotliable group and $\mathscr{A}$ be a finite alphabet. Then for every number $\gamma \in [0, 1)$ there exists a minimal subsystem Y of the dynamical system $(\mathscr{A}^G, G)$ such that $h_{\text{top}}(Y) = \gamma \log |\mathscr{A}|$.*

Proof. We show that for every number $\gamma \in [0, 1)$ there exists an element $x \in \mathscr{A}^G$ being a quasi-Toeplitz configuration such that

$$h(x) = h_{\text{top}}(\overline{Gx}) = \gamma \log |\mathscr{A}|. \tag{8.1}$$

[1] The author would like to thank Dominik Kwietniak and Benjamin Weiss for obtaining a copy of this manuscript.

Having established that, to conclude the Theorem it is enough to apply Lemma 7.1 which says that each such subsystem $\overline{Gx}$ is minimal.

To prove existence of an x satisfying (8.1) for every $\gamma \in [0,1)$ we use the properties of quasi-Toeplitz sequences that we have established in previous chapters. First of all, by Lemma 7.3 the space of quasi-Toeplitz configurations is path-connected with respect to the Weyl pseudometric. Secondly, as proved in Theorem 5.2, the entropy function $h : \mathscr{A}^G \to [0, \log|\mathscr{A}|)$ is continuous on $\mathscr{A}^G$ (when equipped with the Weyl pseudometric).

Therefore, it is enough to find a quasi-Toeplitz configuration that has entropy 0 and then to find another quasi-Toeplitz configurations that have entropy arbitrarily close to $\log|\mathscr{A}|$. The former is simple: the constant configuration a^G for any $a \in \mathscr{A}$ is clearly quasi-Toeplitz and has entropy 0. To show the latter, let us fix any $\gamma \in (0,1)$, we show that there exists a quasi-Toeplitz configuration $x \in \mathscr{A}^G$ such that $h(x) \geqslant \gamma \log|\mathscr{A}|$. To this end, let $\{F_n\}_{n\in\mathbb{N}}$ be a congruent Følner sequence with the associated elegant sequence $\mathcal{C} = \{C_n\}_{n\in\mathbb{N}}$. Passing to a subsequence if necessary, we may assume that

$$|F_{n+1}| \geqslant |F_n||\mathscr{A}|^{|F_n|}.$$

For any $n \in \mathbb{N}$ we let

$$r_n \overset{\text{def}}{=} \left\lfloor (1-\gamma)\frac{|F_n|}{2^{n+1}} \right\rfloor.$$

We construct a sequence $\{G_n\}_{n\in\mathbb{N}}$ of subsets of G satisfying:

1. $G_n \subseteq F_n$ and $|G_n| = r_n$ for every $n \in \mathbb{N}$,
2. $G_i \cap G_jC_j = \varnothing$ for $i \neq j$,
3. $\displaystyle\bigcup_{n\in\mathbb{N}} G_nC_n = G$.

Since G is countable, we can enumerate its elements as $g_1, g_2, g_3, \ldots$. Moreover, since $F_0 \subseteq F_1 \subseteq F_2 \subseteq \ldots$ and $\bigcup_{n\in\mathbb{N}} F_n = G$, we can make sure that the enumeration is "consistent" with the Følner sequence $\{F_n\}_{n\in\mathbb{N}}$, i.e., whenever $g_i \in F_n$ and $g_j \in F_k \backslash F_n$ for some $n < k$ then $i < j$.

The construction of $G_0, G_1, \ldots$ is inductive. Intuitively, G_n is taken to be r_n elements with smallest indices that do not appear in any "shift" of a previously selected subset G_i (for $i < n$), i.e., not in G_iC_i (for any $i < n$). One can imagine this process of constructing $G_0, G_1, \ldots$ as starting with the entire sequence $g_1, g_2, g_3, \ldots$ and then whenever G_n is constructed, all its elements along with all their shifts by C_n (i.e. G_nC_n) are erased from the sequence.

To formalize the above, define

$$G_0 \overset{\text{def}}{=} \{g_1, \ldots, g_{r_0}\} \subseteq F_0.$$

Now, assuming that G_n is already constructed, G_{n+1} is chosen to be

$$G_{n+1} \overset{\text{def}}{=} \{g_{k_1}, \ldots, g_{k_{r_{n+1}}}\},$$

where

$$k_j \overset{\text{def}}{=} \min\left\{ i : g_i \notin \bigcup_{k=0}^{n} G_kC_k \ \text{ and } \ g_i \neq g_{k_l} \text{ for every } l = 1, 2, \ldots, j-1 \right\}.$$

Now we construct $x \in \mathscr{A}^G$ inductively as follows. We start by setting x_{F_0} arbitrarily. Subsequently, we copy the "pattern" at $G_0 \subseteq F_0$ on all shifts of G_0 by C_0, that is for every $g \in G_0$ and $c \in C_0$ we set $x(gc) := x(g)$.

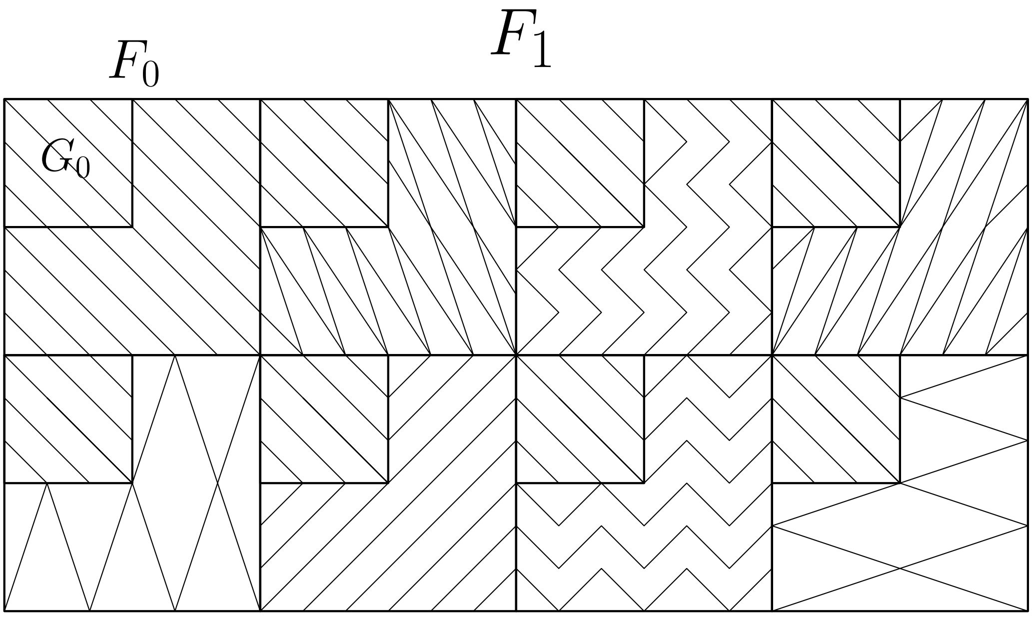

Figure 8.1: One step of the construction of a Toeplitz sequence $x \in \mathscr{A}^G$ from the proof of Theorem 8.1. In the 0-th step we fill F_0 with arbitrary pattern, let's say diagonal lines, and copy this pattern from G_0 to the all group shifting it by elements from C_0. Next we fill the remaining place in F_1 as follows: assume that $|\mathscr{A}|^{|F_0 \setminus G_0|} = 8$. That means that there are eight possibilities to fill the remaining place over the copies of F_0 in F_1 and we have just so many copies of F_0 in F_1. Therefore we fill each copy of $F_0 \setminus G_0$ in F_1 with a different pattern.

We now describe the induction step: fix any $n \in \mathbb{N}$, after steps $0, 1, \ldots, n$ we are left with x defined on the set $F_n \cup T_n \subseteq G$, where

$$T_n \stackrel{\text{def}}{=} \bigcup_{j=0}^{n} G_j C_j.$$

The goal is now to define x on the remaining part of F_{n+1}. Let J_n be the subset of C_n such that

$$F_{n+1} = \bigcup_{c \in J_n} F_n c.$$

Each such "translated" monotile $F_n c$ for $c \in J_n$, $c \neq e$, has a small part of it already filled (that is, x is defined for g in $F_n c \cap T_n$), but the remaining $|F_n \setminus T_n|$ elements of $F_n c$ are still to be set. The idea is now to exhaust all $|\mathscr{A}|^{|F_n \setminus T_n|}$ possible patterns (see Figure 8.1) over copies of F_n when filling F_{n+1}. Since by the choice of the Følner sequence, $|J_n| \geqslant |\mathscr{A}|^{|F_n|}$, this is indeed possible. Having x defined over F_{n+1}, as previously, we copy the part over G_{n+1} over the whole group, formally for every $g \in G_{n+1}$ and $c \in C_{n+1}$ we put $x(gc) := x(g)$. Now x is defined over $F_{n+1} \cup T_{n+1}$.

Having described the construction we are ready to prove the desired properties of x. First of all, note that x is a well-defined quasi-Toeplitz configuration; condition 3 ensures that x is defined at every index $g \in G$ and "all its elements are quasi-periodic".

The next step is to bound the entropy of x from below. To this end, recall that in the construction of x restricted to F_{n+1} we guarantee that

$$|\mathcal{B}_{F_n}(x)| \geqslant |\mathscr{A}|^{|F_n \setminus T_n|}.$$

Moreover, we have

$$
\begin{aligned}
|F_n \backslash T_n| &= |F_n| - \sum_{i=0}^{n} r_i \frac{|F_n|}{|F_i|} \\
&\geqslant |F_n| - \sum_{i=0}^{n} (1-\gamma)\frac{|F_i|}{2^{i+1}} \frac{|F_n|}{|F_i|} \\
&\geqslant |F_n| - \sum_{i=0}^{\infty} (1-\gamma)|F_n| 2^{-i-1} \\
&\geqslant |F_n| - (1-\gamma)|F_n| \\
&= \gamma |F_n|
\end{aligned}
$$

Thus

$$
h(x) \geqslant \limsup_{n\to\infty} \frac{\log|\mathcal{B}_{F_n}(x)|}{|F_n|} \geqslant \gamma \log|\mathscr{A}|. \qquad \square
$$

Remark 8.1. Note that if G is a congruent monotileable group, then there does not exist a minimal subsystem Y of $(\mathscr{A}^G, G)$ such that $h_{\text{top}}(Y) = \log|\mathscr{A}|$. We justify this fact in case $|\mathscr{A}| = 2$ (if $|\mathscr{A}| > 2$, the proof is the same).

Assume that $Y \subseteq \mathscr{A}^G$ is a minimal subshift. Let $\{F_n\}_{n\in\mathbb{N}}$ be a congruent Følner sequence. Then, since $Y \neq \mathscr{A}^G$, there exists $N \in \mathbb{N}$ such that

$$
|\mathcal{B}_{F_N}(Y)| \leqslant 2^{|F_N|} - 1.
$$

Moreover, for every $m > N$, there exists $K_m \subseteq G$ such that $F_m = F_N K_m$ and $F_N g \cap F_N h = \varnothing$ for every $g, h \in K_m$ (since G is coungruent monotileable). This implies that for every $m > N$, one has

$$
|\mathcal{B}_{F_m}(Y)| \leqslant \left(2^{|F_N|} - 1\right)^{|K_m|}.
$$

Therefore (note that $|F_m| = |K_m||F_N|$)

$$
\frac{\log|\mathcal{B}_{F_m}(Y)|}{|F_m|} \leqslant \frac{|K_m|\log\left(2^{|F_N|}-1\right)}{|F_m|} = \frac{\log\left(2^{|F_N|}-1\right)}{|F_N|} < \log 2,
$$

which implies

$$
h_{\text{top}}(Y) = \lim_{m\to\infty} \frac{\log|\mathcal{B}_{F_m}(Y)|}{|F_m|} < \log 2.
$$

8.2 Proximal Subsystems

In this section we prove that all possible values of entropy are realizable by proximal subsystems of $\{0,1\}^G$ whenever G is a residually finite group. While we state and prove it for the case of binary alphabet, extending it to arbitrary finite alphabets is straightforward, see Remark 8.2.

Theorem 8.2 (Realizability of entropy for proximal systems)**.** *Let G be a residually finite amenable group. Then, for every number $\gamma \in [0, 1)$ there exists a proximal transitive dynamical subsystem (Z, G) of the dynamical system $(\{0,1\}^G, G)$ such that $h(Z) = \gamma \log 2$.*

$F_{n_{k+1}}$

F_{n_k}	F_{n_k}	0	0
F_{n_k}	F_{n_k}	F_{n_k}	0
F_{n_k}	F_{n_k}	F_{n_k}	F_{n_k}

Figure 8.2: The $(k+1)$-th step of the construction of a pattern over $F_{n_{k+1}}$ in $x^{(k+1)}$ from the proof of Theorem 8.2. We copy the pattern from F_{n_k} to $m_k - r_k$ shifts of F_{n_k} by some elements from H_{n_k} and we fill the remaining places with zeros.

Proof. Let $\mathcal{F} = \{F_n\}_{n\in\mathbb{N}} \subseteq \mathcal{P}(G)$ be a Følner sequence associated to a sequence of normal subgroups $\{H_n\}_{n\in\mathbb{N}} \subseteq \mathcal{P}(G)$ as provided by Lemma 2.4. Fix $\gamma \in [0, 1)$. We define an element $x \in \{0,1\}^G$ such that the density of the set $\{g \in G : x(g) = 1\}$ along $\mathcal{F}$ is γ. Then we consider the subordinate $\check{Z}$ of $Z := \overline{Gx}$, i.e., the set

$$\check{Z} \stackrel{\text{def}}{=} \{y \in \{0,1\}^G : \exists z \in Z \ \ \forall g \in G \ \ y(g) \leqslant z(g)\}.$$

We prove that $\check{Z}$ is proximal and its entropy is $h_{\text{top}}(\check{Z}) = \gamma \log 2$.

We construct x inductively. We begin with $x^{(0)} \equiv 1$ and $n_0 = 0$. In the $(k+1)$-th step, given integers n_i and configurations $x^{(i)} \in \{0,1\}^G$ for $i = 0, 1, \ldots, k$ already constructed, we define

$$d_i \stackrel{\text{def}}{=} \frac{1}{|F_{n_i}|} \sum_{g\in F_{n_i}} x^{(i)}(g) \qquad \text{for } i = 0, 1, \ldots, k.$$

Then we construct a configuration $x^{(k+1)} \in \{0,1\}^G$ and choose an integer n_{k+1} satisfying:

(1) $n_{k+1} > n_k$,

(2) $x^{(k+1)}_{F_{n_{k+1}}h} = x^{(k+1)}_{F_{n_{k+1}}}$ for every $h \in H_{n_{k+1}}$ (i.e., $x^{(k+1)}$ is periodic with period $H_{n_{k+1}}$),

(3) $x^{(k+1)}_{F_{n_k}} = x^{(k)}_{F_{n_k}}$,

(4) $1 = d_0 > d_1 > d_2 > \ldots > d_{k+1} > \gamma$,

(5) $|d_{k+1} - \gamma| \leqslant \dfrac{1}{2}|d_k - \gamma|$.

Notice that the condition (3) implies that the sequence $\{x^{(n)}\}_{n\in\mathbb{N}}$ is convergent and we can define $x \in \{0,1\}^G$ as

$$x \stackrel{\text{def}}{=} \lim_{n\to\infty} x^{(n)}.$$

Further, the condition (5) implies that $d_k \to \gamma$, which will be useful later on.

Let us now go back to the construction and justify that we can perform the step from k to $(k+1)$. Assuming that we have already constructed $x^{(k)}$ as a periodic configuration with period H_{n_k}, we construct $x^{(k+1)}$ as follows. Intuitively: after choosing n_{k+1}, we copy the pattern from F_{n_k} in $x^{(k)}$ over some (not all) shifts of F_{n_k} contained in $F_{n_{k+1}}$ and then we fill with zeros the remaining places. Finally, we copy the pattern from $F_{n_{k+1}}$ to the whole group G using $H_{n_{k+1}}$. When choosing n_{k+1} the goal is for the density of ones in $F_{n_{k+1}}$ to satisfy

$$\gamma < d_{k+1} \leqslant \frac{1}{2}(d_k + \gamma) < d_k.$$

Hence two problems may appear, if we fill $F_{n_{k+1}}$ with too many zeros, then $d_{k+1} < \gamma$, and if we fill it with too little zeros, then d_k might never converge to γ. Therefore we need to choose n_{k+1} such that $F_{n_{k+1}}$ fits with a lot of copies of F_{n_k} so that we can "balance out" the number of zeros just right.

Let us now write down the above idea formally. By Lemma 2.4 we have that

$$F_{n_{k+1}} = \bigsqcup \{F_{n_k} g : g \in F_{n_{k+1}} \cap H_{n_k}\}.$$

We divide $F_{n_{k+1}} \cap H_{n_k}$ into two disjoint parts S_k and Z_k (to be specified later), and assign:

$$x^{(k+1)}(gh) \overset{\text{def}}{=} \begin{cases} x(g) & \text{if } h \in S_k, \\ 0 & \text{if } h \in Z_k, \end{cases} \qquad \text{for } g \in F_k.$$

Having $x^{(k+1)}$ defined over $F_{n_{k+1}}$, we copy this pattern to G:

$$x^{(k+1)}(gh) \overset{\text{def}}{=} x^{(k+1)}(g) \quad \text{ for } g \in F_{n_{k+1}},\ h \in H_{n_{k+1}}.$$

To finish the construction, it is enough to specify S_k and Z_k. Denote

$$m_k \overset{\text{def}}{=} |F_{n_{k+1}} \ \cap \ H_{n_k}| \quad \text{ and } \quad r_k \overset{\text{def}}{=} |Z_k|.$$

We will determine r_k. Let $\varepsilon_k := d_k - \gamma$ for every $k \in \mathbb{N}$. Notice that by conditions (4) and (5) in the induction thesis we have $0 < \varepsilon_{k+1} < \frac{1}{2}\varepsilon_k$. Since

$$d_{k+1} = \frac{m_k - r_k}{m_k} d_k,$$

we obtain

$$\varepsilon_{k+1} = d_k - \gamma - \frac{r_k}{m_k} d_k = \varepsilon_k - \frac{r_k}{m_k} d_k.$$

Hence in conditions (4) and (5) we require

$$0 < \varepsilon_k - \frac{r_k}{m_k} d_k < \frac{1}{2}\varepsilon_k,$$

which is equivalent to

$$\frac{\varepsilon_k}{2d_k} m_k < r_k < \frac{\varepsilon_k}{d_k} m_k.$$

Therefore it is enought to choose n_{k+1} such that $F_{n_{k+1}}$ is so large that there exists an integer between $\frac{\varepsilon_k}{2d_k} m_k$ and $\frac{\varepsilon_k}{d_k} m_k$. This is possible, since ε_k and d_k are fixed positive numbers and $m_k \to \infty$ as we go with n_{k+1} to infinity. We take this integer to be r_k, i.e., the number of copies of F_{n_k} in $F_{n_{k+1}}$ that we fill with zeros. We refer to Figure 8.2 for a graphical explanation of this construction.

From now on, for simplicity of notation we assume that $n_k = k$ for all $k \in \mathbb{N}$. The next step is to prove that $h_{\text{top}}(\breve{Z}) = \gamma \log 2$. To this end, fix $n \in \mathbb{N}$ and denote

$$\breve{Z}^{(n)} \stackrel{\text{def}}{=} \{y \in \{0,1\}^G : \exists z \in \overline{Gx^{(n)}} \ \ \forall g \in G \ \ y(g) \leqslant z(g)\}.$$

Notice that

$$\breve{Z} \subseteq \breve{Z}^{(n)}.$$

Indeed, $x(g) \leqslant x^{(n)}(g)$ for every $g \in G$. Thus $x \in \breve{Z}^{(n)}$. This and Lemma 6.7 imply that

$$h_{\text{top}}(\breve{Z}) \leqslant h_{\text{top}}(\breve{Z}^{(n)}) = d_n \log 2.$$

Therefore, since n was arbitrary, we have

$$h_{\text{top}}(\breve{Z}) \leqslant \gamma \log 2.$$

On the other hand for every $n \in \mathbb{N}$ it holds

$$\left|\mathcal{B}_{F_n}(\breve{Z})\right| \geqslant 2^{|F_n| d_n}.$$

Hence

$$\frac{1}{|F_n|} \log \left(\left|\mathcal{B}_{F_n}(\breve{Z})\right|\right) \geqslant d_n \log 2 \geqslant \gamma \log 2,$$

and passing with n to ∞ we consequently have $h_{\text{top}}(\breve{Z}) \geqslant \gamma \log 2$.

Finally we are left with the task of proving that $(\breve{Z}, G)$ is proximal. Before we proceed with the proof, recall that the metric ρ on $\{0,1\}^G$ is defined as:

$$\rho(z, y) = 2^{-k} \quad \text{with} \quad k = \min\{i \in \mathbb{N} : z(g_i) \neq y(g_i)\},$$

where $(g_1, g_2, \ldots)$ is any bijective enumeration of elements of G. It will be convenient for us to assume that this enumeration is consistent with the Følner sequence in the following sense:

$$\exists n \in \mathbb{N} \quad g_l \in F_n \quad \text{and} \quad g_k \notin F_n \quad \Longrightarrow \quad l < k.$$

In other words, elements in F_0 go first, next elements in $F_1 \backslash F_0$, next in $F_2 \backslash F_1$ and so on.

Now, observe that by the construction of x, it is clear that for every $n \in \mathbb{N}$ we have $x_{F_n} = x_{F_n g}$ for infinitely many $g \in G$. Hence, applying Lemma 6.1 we obtain that $\breve{Z}$ is transitive. Let $z \in \breve{Z}$ be such that $\overline{Gz} = \breve{Z}$. We now apply Theorem 3.1 to show that $\{0^G\}$ is the only minimal subsystem of $\breve{Z} = \overline{Gz}$, i.e., $m(z) = 1$. For this, we need to show that for every $\varepsilon > 0$ the set

$$S \stackrel{\text{def}}{=} \{g \in G : \rho(gz, 0) < \varepsilon\}$$

is syndetic. Fix $\varepsilon > 0$. Then there exists an $n \in \mathbb{N}$ such that $|F_n| > \frac{1}{\varepsilon}$. By the construction of x, there exists $h_0 \in H_n$ such that

$$x_{F_n h_0} \equiv 0.$$

Then for every $h \in H_{n+1}$ we have

$$x_{F_n h_0 h} \equiv 0.$$

But $z \leqslant x$, hence for every $h \in H_{n+1}$ we also have

$$z_{F_n h_0 h} \equiv 0.$$

Thus

$$\rho(h_0 h z, 0) < \varepsilon \quad \text{for } h \in H_{n+1},$$

which means that $h_0 H_{n+1} \subseteq S$. Therefore taking $F = F_{n+1} h_0^{-1}$ we obtain

$$FS \supseteq F h_0 H_{n+1} = F_{n+1} h_0^{-1} h_0 H_{n+1} = F_{n+1} H_{n+1} = G.$$

Consequently, the claim holds and $\check{Z}$ indeed is a proximal system with 0^G being its only fixed point. □

Remark 8.2. The general version of Theorem 8.2 (i.e., for arbitrary alphabet $\mathscr{A} = \{0, 1, \ldots, k-1\}$) can be obtained by repeating the proof of Theorem 8.2 with a small modification: in the construction of x we replace 1's by $(k-1)$'s.

For this proof to go through we need a new version of Theorem 6.1, yet a complete generalization is not necessary. It is sufficient to prove that if a subshift $Z \subseteq \{0, k-1\}^G \subseteq \mathscr{A}^G$ is such that $h_{\mathrm{top}}(Z) = 0$, then the entropy of the subordinate shift $\check{Z} \subseteq \mathscr{A}^G$ is given by

$$h_{\mathrm{top}}(\check{Z}) = \log(k) \lim_{n\to\infty} \frac{1}{|F_n|} \max\left\{ \sum_{g\in F_n} \frac{w(g)}{k-1} : w \in \mathcal{B}_{F_n}(Z) \right\}.$$

This, however, can be argued in the same way as for the binary alphabet, since only two characters: 0 and $k-1$ play a role.

Bibliography

[1] Ethan Akin and Sergii Kolyada. Li–yorke sensitivity. *Nonlinearity*, 16(4):1421–1433, may 2003.

[2] Joseph Auslander. Mean-L-stable systems. *Illinois J. Math.*, 3:566–579, 1959.

[3] Sebastian Bahamonde, Christian G. Böhmer, Sante Carloni, Edmund J. Copeland, Wei Fang, and Nicola Tamanini. Dynamical systems applied to cosmology: Dark energy and modified gravity. *Physics Reports*, 775-777:1–122, Nov 2018.

[4] Mathias Beiglböck, Vitaly Bergelson, and Alexander Fish. Sumset phenomenon in countable amenable groups. *Adv. Math.*, 223(2):416–432, 2010.

[5] Patrick Billingsley. *Convergence of Probability Measures.* Wiley Series in Probability and Mathematical Statistics. Wiley, 1968.

[6] Francois Blanchard, Enrico Formenti, and Petr Kurka. Cellular automata in the Cantor, Besicovitch, and Weyl topological spaces. *Complex Systems*, 11(2):107–123, 1997.

[7] Tullio Ceccherini-Silberstein and Michel Coornaert. *Cellular Automata and Groups.* Springer Monographs in Mathematics. Springer Berlin Heidelberg, 2010.

[8] Paulina Cecchi and María Isabel Cortez. Invariant measures for actions of congruent monotileable amenable groups. *Groups, Geometry, and Dynamics*, 13(3):821–839, May 2019.

[9] Martha Łącka and Marta Straszak. Quasi-uniform convergence in dynamical systems generated by an amenable group action. *Journal of the London Mathematical Society*, 98(3):687–707, 2018.

[10] María Isabel Cortez and Samuel Petite. G-odometers and their almost one-to-one extensions. *J. Lond. Math. Soc. (2)*, 78(1):1–20, 2008.

[11] María Isabel Cortez and Samuel Petite. Invariant measures and orbit equivalence for generalized toeplitz subshifts. *Groups Geom. Dyn.*, 8(4):1007–1045, 2014.

[12] J. de Vries. *Elements of Topological Dynamics.* Mathematics and Its Applications. Springer Netherlands, 1993.

[13] Tomasz Downarowicz. Entropia. 2001.

[14] Tomasz Downarowicz. Survey of odometers and toeplitz flows. *Contemporary Mathematics*, 385:7–38, 2005.

[15] Tomasz Downarowicz. *Entropy in Dynamical Systems*. New Mathematical Monographs. Cambridge University Press, 2011.

[16] Tomasz Downarowicz, Bartosz Frej, and Pierre-Paul Romagnoli. Shearer's inequality and infimum rule for shannon entropy and topological entropy. In *Dynamics and Numbers, Contemporary Mathematics, vol. 669*, pages 76–89. AMS, 2016.

[17] Tomasz Downarowicz and Eli Glasner. Isomorphic extensions and applications. *Topol. Methods Nonlinear Anal.*, 48(1):321–338, 2016.

[18] Tomasz Downarowicz, Dawid Huczek, and Guohua Zhang. Tilings of amenable groups. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2019, 02 2015.

[19] Tomasz Downarowicz and Anzelm Iwanik. Quasi-uniform convergence in compact dynamical systems. *Studia Math.*, 89(1):11–25, 1988.

[20] Manfred Einsiedler and Thomas Ward. *Ergodic Theory: with a view towards Number Theory*. Graduate Texts in Mathematics. Springer London, 2010.

[21] El Houcein El Abdalaoui, Stanisław Kasjan, and Mariusz Lemańczyk. 0-1 sequences of the thue-morse type and sarnak's conjecture. *Proceedings of the American Mathematical Society*, 144(1):161–176, 2016.

[22] Jeffrey L. Elman. Language as a dynamical system. *Mind as motion: Explorations in the dynamics of cognition*, pages 195–225, 1995.

[23] Gerard G. Emch. *Algebraic Methods in Statistical Mechanics and Quantum Field Theory*. Dover Publications, 2014.

[24] Gerard G. Emch and Chuang Liu. *The Logic of Thermostatistical Physics*. Physics and astronomy online library. Springer, 2002.

[25] Sergei Vasilyevich Fomin. On dynamical systems with a purely point spectrum. *Doklady Akad. Nauk SSSR (N.S.)*, 77:29–32, 1951.

[26] Gabriel Fuhrmann, Maik Gröger, and Tobias Jäger. Amorphic complexity. *Nonlinearity*, 29(2):528–565, jan 2016.

[27] Eli Glasner and Benjamin Weiss. Minimal actions of the group $\mathbb{S}(Z)$ of permutations of the integers. *Geometric and Functional Analysis GAFA*, 12:964–988, 2002.

[28] Shmuel Glasner. *Proximal flows*. Lecture Notes in Mathematics. Springer, Berlin, Heidelberg, 1976.

[29] Shmuel Glasner. Proximal flows of lie groups. *Israel Journal of Mathematics*, 45(2):97–99, Jun 1983.

[30] Walter Helbig Gottschalk and Gustav Arnold Hedlund. *Topological Dynamics*. American Mathematical Society: Colloquium publications. American Mathematical Society, 1955.

[31] G. Hansel and J. P. Raoult. Ergodicity, uniformity and unique ergodicity. *Indiana University Mathematics Journal*, 23(3):221–237, 1973.

[32] Yair Hartman, Kate Juschenko, Omer Tamuz, and Pooya Vahidi Ferdowsi. Thompson's group f is not strongly amenable. *Ergodic Theory and Dynamical Systems*, 39(4):925–929, 2019.

[33] Michael Hochman and Tom Meyerovitch. A characterization of the entropies of multidimensional shifts of finite type. *Annals of Mathematics*, 171(3):2011–2038, May 2010.

[34] E.M. Izhikevich. *Dynamical Systems in Neuroscience.* Computational neuroscience Dynamical systems in neuroscience. MIT Press, 2007.

[35] Konrad Jacobs and Michael Keane. 0-1-sequences of toeplitz type. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 13(2):123–131, Jun 1969.

[36] Anatole Katok and Boris Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1995.

[37] Fabrice Krieger. Sous-décalages de Toeplitz sur les groupes moyennables résiduellement finis. *J. Lond. Math. Soc. (2)*, 75(2):447–462, 2007.

[38] Fabrice Krieger. Toeplitz subshifts and odometers for residually finite groups. In *École de Théorie Ergodique*, volume 20 of *Sémin. Congr.*, pages 147–161. Soc. Math. France, Paris, 2010.

[39] Marcin Kulczycki, Dominik Kwietniak, and Jian Li. Entropy of subordinate shift spaces. *The American Mathematical Monthly*, 125:141–148, 02 2018.

[40] Dominik Kwietniak, Martha Łącka, and Piotr Oprocha. Generic points for dynamical systems with average shadowing. *ArXiv e-prints*, March 2016.

[41] Tien-Yien Li and James A. Yorke. Period three implies chaos. *The American Mathematical Monthly*, 82(10):985–992, 1975.

[42] Elon Lindenstrauss. Pointwise theorems for amenable groups. *Invent. Math.*, 146(2):259–295, 2001.

[43] Elon Lindenstrauss and Benjamin Weiss. Mean topological dimension. *Israel J. Math.*, 115:1–24, 2000.

[44] Edward Norton Lorenz. Deterministic nonperiodic flow. *Journal of Atmospheric Sciences*, 20(2):130–148, Mar 1963.

[45] Paul Marrow, Ulf Dieckmann, and Richard Law. Evolutionary dynamics of predator-prey systems: an ecological perspective. *Journal of mathematical biology*, 34(5-6):556–578, 1996.

[46] James Clerk Maxwell. A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, 155:459–512, 1865.

[47] Robert May. Simple mathematical models with very complicated dynamics. *Nature*, 261:459–467, 1976.

[48] Marc Mézard and Andrea Montanari. *Information, physics, and computation.* Oxford Univ. Press, 2009.

[49] Isaac Namioka. Folner's conditions for amenable semi-groups. *Mathematica Scandinavica*, 15:18–28, Jun. 1964.

[50] John C. Oxtoby. Ergodic sets. *Bull. Amer. Math. Soc.*, 58:116–136, 1952.

[51] Henri Poincaré. *New Methods of Celestial Mechanics.* History of Modern Physics. Springer Verlag, 1993.

[52] Alan Rosenthal. Strictly ergodic models and amenable group actions. Unpublished.

[53] Alan Rosenthal. On strictly ergodic models for commuting ergodic transformations. *Annales de l'I.H.P. Probabilités et statistiques*, 25(1):73–92, 1989.

[54] David Ruelle. *Thermodynamic Formalism: The Mathematical Structure of Equilibrium Statistical Mechanics.* Cambridge Mathematical Library. Cambridge University Press, 2 edition, 2004.

[55] Ville Salo and Ilkka Törmä. Geometry and dynamics of the Besicovitch and Weyl spaces. In *Developments in language theory*, volume 7410 of *Lecture Notes in Comput. Sci.*, pages 465–470. Springer, Heidelberg, 2012.

[56] Paul C. Shields. *The ergodic theory of discrete sample paths*, volume 13 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 1996.

[57] Stephen Smale. Differentiable dynamical systems. *Bull. Amer. Math. Soc.*, 73(6):747–817, 11 1967.

[58] Slawomir Solecki. Size of subsets of groups and haar null sets. *Geometric & Functional Analysis GAFA*, 15(1):246–273, Feb 2005.

[59] Benjamin Weiss. Strictly ergodic models for dynamical systems. *Bull. Amer. Math. Soc. (N.S.)*, 13(2):143–146, 10 1985.

[60] Benjamin Weiss. *Topology, ergodic theory, real algebraic geometry*, chapter Monotileable amenable groups, pages 257–262. American Mathematical Society, Providence, RI, 2001.