

# SOME RESULTS ON RECURRENCE AND ENTROPY

DISSERTATION

Presented in Partial Fulfillment of the Requirements for  
the Degree Doctor of Philosophy in the Graduate  
School of the Ohio State University

By

Ronald Lee Pavlov Jr., B.S., M.S.

\* \* \* \* \*

The Ohio State University  
2007

**Dissertation Committee:**

Professor Vitaly Bergelson, Advisor

Professor Alexander Leibman

Professor Manfred Einsiedler

**Approved by**

---

Advisor

Graduate Program in Mathematics



## ABSTRACT

This thesis is comprised primarily of two separate portions. In the first portion, we exhibit, for any sparse enough increasing sequence  $\{p_n\}$  of integers, a totally minimal, totally uniquely ergodic, and topologically mixing system  $(X, T)$  and  $f \in C(X)$  for which the averages  $\frac{1}{N} \sum_{n=0}^{N-1} f(T^{p_n} x)$  fail to converge on a residual set in  $X$ , answering negatively an open question of Bergelson. We also construct here a totally minimal, totally uniquely ergodic, and topologically mixing system  $(X', T')$  and  $x' \in X'$  for which  $x' \notin \overline{\{T^{p_n} x'\}}$ .

In the second portion, we study perturbations of multidimensional shifts of finite type. Given any  $\mathbb{Z}^d$  shift of finite type  $X$  for  $d > 1$  and any word  $w$  in the language of  $X$ , denote by  $X_w$  the set of elements of  $X$  in which  $w$  does not appear. If  $X$  satisfies a uniform mixing condition called strong irreducibility, we obtain exponential upper and lower bounds on  $h^{top}(X) - h^{top}(X_w)$  dependent only on the size of  $w$ . This result generalizes a result of Lind about  $\mathbb{Z}$  shifts of finite type.

*To Dilip*

## ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor Vitaly Bergelson. This thesis could never have been completed without his help. His infectious enthusiasm for mathematics, along with a willingness to share his knowledge, were a constant inspiration.

I would also like to thank my other committee members, Professors Alexander Leibman and Manfred Einsiedler, for agreeing to serve on my committee and for many fruitful mathematical discussions.

Finally, I thank my friends and loved ones, without whom I would never have been able to get as far as I have. In no particular order:

Thank you to Mom and Dad, who have always been willing to do anything to support me and have always believed in me.

Thank you to Greg, without whose friendship and sense of humor graduate school would have been much less bearable.

Thank you to Jasmine, who was always there for me. I can never thank you enough for your constant love and support.

## VITA

2000-Present .....	Graduate Teaching Associate, The Ohio State University
1998-1999 .....	Student Instructional Associate, The Ohio State University
2003 .....	M.S. in Mathematics, The Ohio State University
2000 .....	B.S. in Mathematics, The Ohio State University

## **PUBLICATIONS**

### **Research Papers**

- *Some Counterexamples In Topological Dynamics* (Ergodic Theory and Dynamical Systems, to appear)
- *Perturbations of Multidimensional Shifts of Finite Type* (submitted)

## **FIELDS OF STUDY**

Major Field: Mathematics

Specialization: Ergodic Theory and Dynamical Systems

# TABLE OF CONTENTS

Abstract . . . . .	ii
Dedication . . . . .	iii
Acknowledgments . . . . .	iv
Vita . . . . .	v
List of Figures . . . . .	ix

CHAPTER	PAGE
1 Introduction . . . . .	1
2 Recurrence and convergence of ergodic averages along sparse sequences	8
2.1 Introduction . . . . .	8
2.2 Some general symbolic constructions . . . . .	16
2.3 Some symbolic counterexamples . . . . .	31
2.4 Some general constructions on connected manifolds . . . . .	42
2.5 Some counterexamples on connected manifolds . . . . .	55
2.6 A counterexample about simultaneous recurrence . . . . .	61
2.7 Questions . . . . .	65
3 Perturbations of multidimensional shifts of finite type . . . . .	67
3.1 Introduction . . . . .	67
3.2 Some measure-theoretic preliminaries . . . . .	84
3.3 A replacement theorem . . . . .	93
3.4 The proof of the main result . . . . .	114
3.5 A closer look at the main result . . . . .	146
3.6 An application to an undecidability question . . . . .	155
3.7 Questions . . . . .	159



Bibliography . . . . .	161
------------------------	-----

# LIST OF FIGURES

FIGURE	PAGE
3.1 $f_2$ 's action on a sample element of $Y$ . . . . .	73
3.2 A portion of a sample element of $Z$ . . . . .	76
3.3 $a_4$ . . . . .	77
3.4 $b_4$ . . . . .	77
3.5 $R_{k(j_1+R),k(j_2+R),\dots,k(j_d+R)}$ . . . . .	85
3.6 A standard replacement of $u$ . . . . .	95
3.7 A sample element $S$ of $R_j$ . . . . .	97
3.8 An element $S$ of $R_j$ associated to two standard replacements . . . . .	98
3.9 The suboctants of $B_j$ . . . . .	99
3.10 Elements $S, S'$ of $R_j$ whose difference is a multiple of $e_i$ . . . . .	100
3.11 An element $S$ of $R_j$ which contains $O$ . . . . .	106
3.12 $B_j$ . . . . .	107
3.13 $B'_j$ . . . . .	108
3.14 Intersecting occurrences of $w$ . . . . .	116
3.15 $S_i \setminus S_i^{(R)}$ . . . . .	120
3.16 Disallowed and allowed pairs of overlapping $w_{j,d}$ . . . . .	127
3.17 A point $x \in X_y$ . . . . .	138

3.18	The correspondence between copies of $\Gamma_W$ and points in $\Gamma_j^{(2)}$ . . . . .	139
3.19	How a copy of $\Gamma_W$ is filled if $b_{j,d}(p) = 1$ . . . . .	140
3.20	$f'$ . . . . .	141

# CHAPTER 1

## INTRODUCTION

This introduction will serve as a brief mathematical and historical overview of both of the main problems that we will examine in this thesis. Due to their somewhat disparate natures, the two portions of this thesis will each contain a more in-depth introduction as well. For this reason, we will for the most part relegate formal definitions to their pertinent introduction.

Ergodic theory is the study of average or long-term behavior of systems which evolve with time. For example, one can consider a particle bouncing in a box at fixed speed. Its position and velocity at any time can be represented by a vector in  $\mathbb{R}^6$ . One can then model this behavior by taking  $X$  to be the space of all possible positions and velocities for the particle, and  $T$  a self-map of  $X$  which, given a position and velocity, gives the position and velocity one second later. This pairing of a space  $X$  with a self-map  $T$  is called a **dynamical system**.

The more general setup is that of a group  $G$  acting on a space  $X$  with some sort of structure by maps  $\{T_g\}_{g \in G}$  preserving this structure. In this thesis,  $G$  is always  $\mathbb{Z}^d$  for some  $d \in \mathbb{N}$ . Measure-theoretic ergodic theory occurs when  $X$  is a measure space with a probability measure  $\mu$  invariant under each  $T_g$ , and so in this case we call

$(X, \mathcal{B}, \mu, \{T_g\}_{g \in G})$  a **measure-preserving dynamical system**. Topological dynamics occurs when  $X$  is a compact topological space and each  $T_g$  is a homeomorphism, and so in this case we call  $(X, \{T_g\}_{g \in G})$  a **topological dynamical system**. In either type of system, if  $G = \mathbb{Z}$ , then  $T_n = (T_1)^n$  for any integer  $n$ , and so we shorten  $(X, \mathcal{B}, \mu, \{T_n\}_{n \in \mathbb{Z}})$  to  $(X, \mathcal{B}, \mu, T)$  and  $(X, \{T_n\}_{n \in \mathbb{Z}})$  to  $(X, T)$ . (Here  $T = T_1$ .)

These cases are not as disjoint as they may first appear: the famed Bogoliouboff-Krylov theorem states that any topological dynamical system  $(X, \{T_g\}_{g \in G})$  has at least one invariant probability measure  $\mu$  as long as  $G$  is an amenable group. (An examination of amenability is beyond the scope of this thesis, but we mention that amenable groups are a class which contain all abelian groups. We only consider  $G = \mathbb{Z}^d$  in this thesis, so all topological dynamical systems examined here will possess invariant measures.) This sometimes allows one to examine topological properties of a system via properties of invariant measures. For instance, in Chapter 3, we will prove some purely combinatorial properties of a  $\mathbb{Z}^d$ -action by homeomorphisms of a Cantor set by means of studying the invariant measures of this action.

Chapter 2 deals primarily with ergodic averaging. One of the fundamental results of ergodic theory is Birkhoff's ergodic theorem:

**Theorem 1.0.1.** ([Bi]) *For any measure-preserving dynamical system  $(X, \mathcal{B}, \mu, T)$  and any  $f \in L^1(X)$ ,  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x)$  exists for  $\mu$ -almost every  $x \in X$ .*

Several results have been proven about the convergence of such averages when one averages not along all powers of  $T$ , but only along some distinguished subset of the

integers. ([Bou], [Bou2], [Wi]) In particular, when one averages along  $\{p(n)\}$  for a polynomial  $p(n)$  with integer coefficients, there is the following result of Bourgain:

**Theorem 2.1.12.** ([Bou], p. 7, Theorem 1) *For any measure-preserving dynamical system  $(X, \mathcal{B}, \mu, T)$ , for any polynomial  $q(t) \in \mathbb{Z}[t]$ , and for any  $f \in L^p(X, \mathcal{B}, \mu)$  with  $p > 1$ ,  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(T^{q(n)}x)$  exists for  $\mu$ -almost every  $x \in X$ .*

Theorem 2.1.12 can be interpreted as follows: for any polynomial  $q(t) \in \mathbb{Z}[t]$ , any measure-preserving system  $(X, \mathcal{B}, \mu, T)$ , and any measure-theoretically “nice” function  $f$ , the set of points  $x$  where  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(T^{q(n)}x)$  fails to converge is of measure zero, or negligible measure-theoretically. It is then natural to wonder whether or not there is a topological parallel to this result using topological notions of “niceness” (continuity) and negligibility (first category), and in fact such a question was posed by Bergelson:

**Question 2.1.13.** ([Be], p. 51, Question 5) *Assume that a topological dynamical system  $(X, T)$  is uniquely ergodic<sup>1</sup>, and let  $p \in \mathbb{Z}[t]$  and  $f \in C(X)$ . Is it true that for all but a first category set of points  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^{p(i)}x)$  exists?*

One of our results is a (quite negative) answer to Question 2.1.13:

**Theorem 2.1.15.** *For any increasing sequence  $\{p_n\}$  of integers with upper Banach density zero<sup>2</sup>, there exists a totally minimal, totally uniquely ergodic, and topologically*

---

<sup>1</sup>A topological dynamical system  $(X, \mathcal{B}, \mu, T)$  is uniquely ergodic if there exists exactly one  $T$ -invariant measure  $\mu$ .

<sup>2</sup>A set  $A \subseteq \mathbb{N}$  has upper Banach density zero if  $\limsup_{n \rightarrow \infty} \sup_{m \in \mathbb{N}} \frac{|\{m, m+1, \dots, m+n-1\} \cap A|}{n} = 0$ .

*mixing topological dynamical system  $(X, T)$  and a continuous function  $f$  on  $X$  with the property that  $\frac{1}{N} \sum_{n=0}^{N-1} f(T^{p_n}x)$  fails to converge for a residual set of  $x$ .*

We also examine recurrence along distinguished subsets of the integers, motivated primarily by the following result:

**Theorem 2.1.17.** ([BeL], p. 14, Corollary 1.8) *For any  $d \in \mathbb{N}$ , any minimal<sup>3</sup> topological dynamical system  $(X, \{T_v\}_{v \in \mathbb{Z}^d})$  and any polynomials  $q_1(t), q_2(t), \dots, q_d(t) \in \mathbb{Z}[t]$  with  $q_i(0) = 0$  for  $1 \leq i \leq d$ , for a residual set of  $x \in X$  there exists a sequence  $\{n_i\}$  of positive integers such that  $T_{q(n_i)e_i}x \rightarrow x$  for  $1 \leq i \leq d$ , where  $\{e_i\}_{i=1}^d$  is the standard orthonormal basis of  $\mathbb{R}^d$ .*

In particular, this implies that for any minimal topological dynamical system  $(X, T)$  and any polynomial  $q(t) \in \mathbb{Z}[t]$  with  $q(0) = 0$ , the set of  $x \in X$  with  $x \in \overline{\{T^{q(n)}x\}_{n \in \mathbb{N}}}$  is residual. The following result shows that for  $q$  with degree at least two, this residual set is not necessarily all of  $X$ .

**Theorem 2.1.18.** *For any increasing sequence  $\{p_n\}$  of integers with upper Banach density zero, there exists a totally minimal, totally uniquely ergodic, and topologically mixing topological dynamical system  $(X, T)$  and an uncountable set  $A \subset X$  such that for every  $x \in A$ , the sequence  $\{T^{p_n}x\}$  does not have  $x$  as a limit point, i.e. there is no sequence of positive integers  $\{n_i\}$  such that  $T^{p_{n_i}}x$  converges to  $x$ .*

Another consequence of Theorem 2.1.17 is that for any commuting minimal homeomorphisms  $T$  and  $S$  of a compact space  $X$ , there exists a residual set of  $x$  for which

---

<sup>3</sup>A topological dynamical system  $(X, \mathcal{B}, \mu, T)$  is minimal if there exist no nonempty proper closed  $T$ -invariant subsets of  $X$ .

there exists a sequence of positive integers  $\{n_i\}$  such that  $T^{n_i}x \rightarrow x$  and  $S^{n_i}x \rightarrow x$ . The following theorem shows that for some systems, it is the case that this residual set of  $x$  is not all of  $X$ .

**Theorem 2.1.21.** *There exists a totally minimal topological dynamical system  $(X, T)$  and a point  $x \in X$  such that for any positive integers  $r \neq s$ , any sequence  $\{n_i\}$  of integers satisfying  $T^{r n_i}x \rightarrow x$  and  $T^{s n_i}x \rightarrow x$  is eventually zero.*

The systems which we construct to prove Theorems 2.1.15 and 2.1.18 are **symbolic dynamical systems**. A symbolic dynamical system is defined by first choosing a finite set  $A$ , called the alphabet.  $\Omega = A^G$  endowed with the product topology is a compact space, and for any  $g \in G$ , we may define  $\sigma_g$  the shift homeomorphism by  $(\sigma_g \omega)(h) = \omega(hg)$  for  $\omega \in \Omega$ . Any closed set  $X \subseteq \Omega$  has a topology induced by  $\Omega$ , and if  $X$  is invariant under each  $T_g$ , then  $(X, \{\sigma_g\}_{g \in G})$  is a topological dynamical system which we call a symbolic dynamical system.

In Chapter 3, we examine symbolic dynamical systems exclusively. There, we consider a specific type of symbolic dynamical system called a **shift of finite type**. A  $\mathbb{Z}^d$ -shift of finite type is a symbolic dynamical system  $(X, \{\sigma_v\}_{v \in \mathbb{Z}^d})$  where  $X$  is defined by specifying a finite set  $\mathcal{F}$  of finite words (a word is a function from a finite subset of  $\mathbb{Z}^d$  to  $A$ ) and taking  $X$  to be the set of all elements of  $\Omega$  in which none of the words in  $\mathcal{F}$  appear. The shift of finite type  $X$  specified by a finite set  $\mathcal{F}$  of words in this way is denoted by  $\Omega_{\mathcal{F}}$ . For example, the set of all biinfinite sequences of zeroes and ones in which no two ones occur consecutively is a shift of finite type, as is the set of



all  $\mathbb{Z}^2$  arrays of zeroes and ones in which no three-by-four blocks of all zeroes or all ones appear.

We are particularly interested in the effects of forbidding a particular word from a shift of finite type  $X$ . Given any word  $w$ , define  $X_w$  to be the set of elements of  $X$  in which  $w$  does not appear. Then clearly  $X_w$  is a subset of  $X$ , and so the topological entropy  $h^{top}(X_w)$  of  $X_w$  is not greater than the topological entropy  $h^{top}(X)$  of  $X$ . It is natural to wonder how much the topological entropy drops by when  $w$  is removed, as it is a sort of measure of how important  $w$  is to the information-retaining capacity of  $X$ . In [L], Lind proved the following: (the condition  $w \in L_{\Gamma_n}(X)$  means that  $w \in A^{[1, \dots, n]^d}$  and that  $w$  appears in some element of  $X$ .)

**Theorem 3.1.16.** ([L], p. 360, Theorem 3) *For any topologically transitive  $\mathbb{Z}$ -shift of finite type  $X = \Omega_{\mathcal{F}}$  with positive topological entropy  $h^{top}(X)$ , there exist constants  $C_X$ ,  $D_X$ , and  $N_X$  such that for any  $n > N_X$  and any word  $w \in L_{\Gamma_n}(X)$ , if we denote by  $X_w$  the shift of finite type  $\Omega_{\mathcal{F} \sqcup \{w\}}$ , then*

$$\frac{C_X}{e^{h^{top}(X)n}} < h^{top}(X) - h^{top}(X_w) < \frac{D_X}{e^{h^{top}(X)n}}.$$

Our main result is a generalization of Theorem 3.1.16 for  $\mathbb{Z}^d$ -shifts of finite type which satisfy a mixing condition called strong irreducibility. (We formally define strong irreducibility in Definition 3.1.19.)

**Theorem 3.1.22.** *For any  $d > 1$  and any strongly irreducible  $\mathbb{Z}^d$ -shift  $X = \Omega_{\mathcal{F}}$  of finite type with uniform filling length  $R$  and positive topological entropy  $h^{top}(X)$ ,*

there exist constants  $N_X \in \mathbb{N}$  and  $D_X \in \mathbb{R}$  such that for any  $n > N_X$  and any word  $w \in L_{\Gamma_n}(X)$ , if we denote by  $X_w$  the shift of finite type  $\Omega_{\mathcal{F} \sqcup \{w\}}$ , then

$$\frac{1}{e^{h^{top}(X)(n+44R+70)^d}} < h^{top}(X) - h^{top}(X_w) < \frac{D_X}{e^{h^{top}(X)(n-2R)^d}}.$$

One way in which  $\mathbb{Z}^d$ -shifts of finite type are more difficult to deal with for  $d > 1$  is that given a finite collection  $\mathcal{F}$  of patterns, it is undecidable whether or not the shift of finite type  $X$  induced by  $\mathcal{F}$  is even nonempty. Our methods yield a situation in which this question can be answered:

**Theorem 3.6.1.** *For any alphabet  $A$ , there exist  $F, G \in \mathbb{N}$  such that for any  $m > 0$  and any finite set of words  $\mathcal{F}_m = \{w_k \in L_{\Gamma_{n_k}}(X) : 1 \leq k \leq m\}$  satisfying  $n_1 > G$  and  $n_k \geq F(n_{k-1})^{4d^2}$  for  $1 < k \leq m$ ,  $\Omega_{\mathcal{F}_m} \neq \emptyset$ .*

## CHAPTER 2

### RECURRENCE AND CONVERGENCE OF ERGODIC AVERAGES ALONG SPARSE SEQUENCES

#### 2.1 Introduction

In this chapter, we are concerned with the convergence of averages of the form  $\frac{1}{N} \sum_{n=0}^{N-1} f(T^{p_n}x)$  for an increasing sequence of integers  $\{p_n\}$ . We begin with some definitions.

**Definition 2.1.1.** A *measure-preserving dynamical system*  $(X, \mathcal{B}, \mu, \{T_g\}_{g \in G})$  consists of a measure space  $X$ , a probability measure  $\mu$  with  $\sigma$ -algebra  $\mathcal{B}$  of measurable sets, and a group action  $\{T_g\}_{g \in G}$  of transformations  $T_g : X \rightarrow X$  with  $\mu(T_g^{-1}A) = \mu(A)$  for all  $A \in \mathcal{B}$ .

**Definition 2.1.2.** A *measure-preserving dynamical system*  $(X, \mathcal{B}, \mu, \{T_g\}_{g \in G})$  is **ergodic** if any set  $A$  satisfying  $T_g A \subseteq A$  for all  $g \in G$  has  $\mu(A) \in \{0, 1\}$ .

**Definition 2.1.3.** A *topological dynamical system*  $(X, \{T_g\}_{g \in G})$  consists of a compact topological space  $X$  and a group action  $\{T_g\}_{g \in G}$  of homeomorphisms  $T_g : X \rightarrow X$ .

**Definition 2.1.4.** Given a topological dynamical system  $(X, \{T_g\}_{g \in G})$ , a Borel probability measure  $\mu$  on  $X$  is called **ergodic** if  $(X, \mathcal{B}(X), \mu, \{T_g\}_{g \in G})$  is an ergodic measure-preserving dynamical system, where  $\mathcal{B}(X)$  is the Borel  $\sigma$ -algebra of  $X$ .

In this chapter, all dynamical systems have  $G = \mathbb{Z}$ , so as already described we will use the notations  $(X, \mathcal{B}, \mu, T)$  and  $(X, T)$  for measure-preserving and topological dynamical systems respectively.

**Definition 2.1.5.** A topological dynamical system  $(X, T)$  is **minimal** if for any closed set  $K$  with  $T^{-1}K \subseteq K$ ,  $K = \emptyset$  or  $K = X$ .  $(X, T)$  is **totally minimal** if  $(X, T^n)$  is minimal for every  $n \in \mathbb{N}$ .

**Definition 2.1.6.** A topological dynamical system  $(X, T)$  is **uniquely ergodic** if there is only one Borel measure  $\mu$  on  $X$  such that  $\mu(A) = \mu(T^{-1}A)$  for every Borel set  $A \subseteq X$ .  $(X, T)$  is **totally uniquely ergodic** if  $(X, T^n)$  is uniquely ergodic for every  $n \in \mathbb{N}$ .

**Definition 2.1.7.** A topological dynamical system  $(X, T)$  is **topologically mixing** if for any open sets  $U, V \subseteq X$ , there exists  $N \in \mathbb{N}$  such that for any  $n > N$ ,  $U \cap T^n V \neq \emptyset$ .

**Definition 2.1.8.** For any set  $A \subseteq \mathbb{N}$ , the **upper Banach density** of  $A$  is defined by

$$d^*(A) = \limsup_{n \rightarrow \infty} \sup_{m \in \mathbb{N}} \frac{|\{m, m+1, \dots, m+n-1\} \cap A|}{n}.$$

**Definition 2.1.9.** For any set  $A \subseteq \mathbb{N}$ , the **upper density** of  $A$  is defined by

$$\bar{d}(A) = \limsup_{n \rightarrow \infty} \frac{|\{1, \dots, n\} \cap A|}{n}.$$

**Definition 2.1.10.** For a set  $A \subseteq \mathbb{N}$ , the **density** of  $A$  is defined by

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

if this limit exists.

**Definition 2.1.11.** Given a topological dynamical system  $(X, T)$  and a  $T$ -invariant Borel probability measure  $\mu$ , a point  $x \in X$  is  $(T, \mu)$ -**generic** if for every  $f \in C(X)$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \int f d\mu.$$

In the measure-preserving setup, there are several positive results about convergence of averages of the form  $\frac{1}{N} \sum_{n=0}^{N-1} f(T^{p_n} x)$ , including this theorem of Bourgain:

**Theorem 2.1.12.** ([Bou], p. 7, Theorem 1) For any measure-preserving dynamical system  $(X, \mathcal{B}, \mu, T)$ , for any polynomial  $q(t) \in \mathbb{Z}[t]$ , and for any  $f \in L^p(X, \mathcal{B}, \mu)$  with  $p > 1$ ,  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(T^{q(n)} x)$  exists for  $\mu$ -almost every  $x \in X$ .

The following question regarding a possible topological version of Theorem 2.1.12 was posed by Bergelson:

**Question 2.1.13.** ([Be], p. 51, Question 5) Assume that a topological dynamical system  $(X, T)$  is uniquely ergodic, and let  $p \in \mathbb{Z}[t]$  and  $f \in C(X)$ . Is it true that for all but a first category set of points  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^{p(i)} x)$  exists?

Bergelson added the hypothesis of unique ergodicity because it is a classical result that a system  $(X, T)$  is uniquely ergodic with unique  $T$ -invariant measure  $\mu$  if and only if for every  $x \in X$  and  $f \in C(X)$ ,  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \int_X f d\mu$ , and so in the topological setup this is a natural assumption to make about  $(X, T)$  to achieve the desired result.

However, Bergelson was particularly interested in the convergence of these averages to the “correct limit,” i.e.  $\int_X f d\mu$  where  $\mu$  is the unique  $T$ -invariant measure on  $X$ . To have any hope for such a result, it also becomes necessary to assume ergodicity for all powers of  $T$  in order to avoid some natural counterexamples related to distribution (mod  $k$ ) of  $p(n)$  for positive integers  $k$ . For example, if  $p(n) = n^2$ ,  $T$  is the permutation on  $X = \{0, 1, 2\}$  defined by  $Tx = x + 1 \pmod{3}$ ,  $\mu$  is normalized counting measure on  $X$ , and  $f = \chi_{\{0\}}$ , then  $(X, T)$  is obviously uniquely ergodic with unique invariant measure  $\mu = \frac{\delta_0 + \delta_1 + \delta_2}{3}$ , but

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^{p(i)} x) = \begin{cases} \frac{1}{3} & \text{if } x = 0, \\ 0 & \text{if } x = 1, \text{ and} \\ \frac{2}{3} & \text{if } x = 2. \end{cases}$$

To avoid such examples, we would need  $T$  to be totally ergodic as well as uniquely ergodic, and so it makes sense to assume total unique ergodicity to encompass both properties. Bergelson’s revised question then looks like this:

**Question 2.1.14.** *Assume that a topological dynamical system  $(X, T)$  is totally uniquely*

ergodic with unique  $T$ -invariant measure  $\mu$ , and let  $p \in \mathbb{Z}[t]$  and  $f \in C(X)$ . Is it true that for all but a first category set of points  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^{p(i)}x) = \int_X f d\mu$ ?

We answer Questions 2.1.13 and 2.1.14 negatively in the case where the degree of  $p$  is at least two, and in fact prove some slightly more general results. The level of generality depends on what hypotheses we place on the space  $X$ . In particular, we can exhibit more counterexamples in the case where  $X$  is a totally disconnected space than we can in the case where  $X$  is a connected space such as  $\mathbb{T}^k$ .

**Theorem 2.1.15.** *For any increasing sequence  $\{p_n\}$  of integers with upper Banach density zero, there exists a totally minimal, totally uniquely ergodic, and topologically mixing topological dynamical system  $(X, T)$  and a continuous function  $f$  on  $X$  with the property that  $\frac{1}{N} \sum_{n=0}^{N-1} f(T^{p_n}x)$  fails to converge for a residual set of  $x$ .*

**Theorem 2.1.16.** *For any increasing sequence  $\{p_n\}$  of integers with the property that for some integer  $d$ ,  $p_{n+1} < (p_{n+1} - p_n)^d$  for all sufficiently large  $n$ , there exists a totally minimal, totally uniquely ergodic, and topologically mixing topological dynamical system  $(X, T)$  and a continuous function  $f$  on  $X$  with the property that  $\frac{1}{N} \sum_{n=0}^{N-1} f(T^{p_n}x)$  fails to converge for a residual set of  $x$ . In addition, the space  $X$  is a connected  $2d + 9$ -manifold.*

We note that Theorems 2.1.15 and 2.1.16 answer Question 2.1.13 negatively for  $p$  with degree at least two, since the sequence  $p_n = p(n)$  for any nonlinear  $p(t) \in \mathbb{Z}[t]$  satisfies the hypotheses of both theorems.

Theorems 2.1.15 and 2.1.16 are about nonconvergence of ergodic averages along certain sequences of powers of  $x$ . We also prove two similar results about nonrecurrence

of points. As motivation, we note that a minimal system has the property that every point is recurrent. In other words, if  $(X, T)$  is minimal, then for all  $x \in X$ , it is the case that  $x \in \overline{\{T^n x\}_{n \in \mathbb{N}}}$ . If  $(X, T)$  is totally minimal, then all points are recurrent even along infinite arithmetic progressions: for any nonnegative integers  $a, b$ , and for all  $x \in X$ ,  $x \in \overline{\{T^{an+b} x\}_{n \in \mathbb{N}}}$ . It is then natural to wonder if the same is true for other sequences of powers of  $T$ , and in this vein there is the following result of Bergelson and Leibman, which is a corollary to their Polynomial van der Waerden theorem:

**Theorem 2.1.17.** ([BeL], p. 14, Corollary 1.8) *For any  $d \in \mathbb{N}$ , any minimal topological dynamical system  $(X, \{T_v\}_{v \in \mathbb{Z}^d})$  and any polynomials  $q_1(t), q_2(t), \dots, q_d(t) \in \mathbb{Z}[t]$  with  $q_i(0) = 0$  for  $1 \leq i \leq d$ , for a residual set of  $x \in X$  there exists a sequence  $\{n_i\}$  of positive integers such that  $T_{q(n_i)e_i} x \rightarrow x$  for  $1 \leq i \leq d$ , where  $\{e_i\}_{i=1}^d$  is the standard orthonormal basis of  $\mathbb{R}^d$ .*

In particular, Theorem 2.1.17 implies that for a minimal system  $(X, T)$  and any polynomial  $p(t)$  with  $p(0) = 0$ , the set of  $x$  such that  $x \in \overline{\{T^{p(n)} x\}_{n \in \mathbb{N}}}$  is residual. If  $(X, T)$  is assumed to be totally minimal, then by definition if the degree of  $p$  is one, then every point  $x \in X$  is in  $\overline{\{T^{p(n)} x\}_{n \in \mathbb{N}}}$ . The following two results imply that if the degree of  $p$  is greater than one, total minimality of  $(X, T)$  does not necessarily imply that every point  $x \in X$  is in  $\overline{\{T^{p(n)} x\}_{n \in \mathbb{N}}}$ .

**Theorem 2.1.18.** *For any increasing sequence  $\{p_n\}$  of integers with upper Banach density zero, there exists a totally minimal, totally uniquely ergodic, and topologically mixing topological dynamical system  $(X, T)$  and an uncountable set  $A \subset X$  such that*



for every  $x \in A$ , the sequence  $\{T^{p_n}x\}$  does not have  $x$  as a limit point, i.e. there is no sequence of positive integers  $\{n_i\}$  such that  $T^{p_{n_i}}x$  converges to  $x$ .

**Theorem 2.1.19.** *For any increasing sequence  $\{p_n\}$  of integers with the property that for some integer  $d$ ,  $p_{n+1} < (p_{n+1} - p_n)^d$  for all sufficiently large  $n$ , there exists a totally minimal, totally uniquely ergodic, and topologically mixing topological dynamical system  $(X, T)$  and a point  $x \in X$  such that the sequence  $\{T^{p_n}x\}$  does not have  $x$  as a limit point, i.e. there is no sequence of positive integers  $\{n_i\}$  such that  $T^{p_{n_i}}x$  converges to  $x$ . In addition, the space  $X$  is a connected  $2d + 7$ -manifold.*

The following simple lemma shows that if  $(X, T)$  is topologically mixing, then Theorems 2.1.18 and 2.1.19 cannot be improved too much, i.e. there is no increasing sequence  $\{p_n\}$  for which we can exhibit topologically mixing examples with a second category set of such nonrecurrent points.

**Lemma 2.1.20.** *If a topological dynamical system  $(X, T)$  is topologically mixing, then for any increasing sequence  $\{p_n\}$ , the set of  $x \in X$  for which  $x \notin \overline{\{T^{p_n}x\}}$  is of first category.*

**Proof:** Define  $C_\epsilon = \{x : d(x, T^{p_n}x) \geq \epsilon \ \forall n \in \mathbb{Z}\}$ . It is clear that  $C_\epsilon$  is closed. We claim that  $C_\epsilon$  contains no nonempty open set, which shows that it is nowhere dense, implying that  $C = \bigcup_{n=1}^{\infty} C_{\frac{1}{n}}$  the set of points  $x$  for which  $x$  is not a limit point of  $\{T^{p_n}x\}$  is of first category. Suppose, for a contradiction, that there is a nonempty open set  $U$  with  $U \subseteq C_\epsilon$  for some  $\epsilon$ . Then, there exists  $V$  with  $\text{diam}(V) < \epsilon$  such that  $V \subseteq U \subseteq C_\epsilon$ . By topological mixing, there exists  $n$  such that  $V \cap T^{-p_n}V \neq \emptyset$ . This

implies that there exists  $x \in V$  so that  $T^{p_n}x \in V$ . Since  $\text{diam}(V) < \epsilon$ ,  $d(x, T^{p_n}x) < \epsilon$ . However,  $x \in V \subseteq C_\epsilon$ , so we have a contradiction.

□

Theorem 2.1.18 shows that it is possible for this set of points nonrecurrent along  $p_n$  to be uncountable though.

We mention that some mixing condition is necessary for a statement like Lemma 2.1.20; as a simple example, consider an irrational circle rotation  $T : x \mapsto x + \alpha$  on the circle  $\mathbb{T}$ . There is clearly some increasing sequence of integers  $\{p_n\}$  such that  $p_n\alpha \pmod{1} \rightarrow \frac{1}{2}$ . Then, for any  $x \in \mathbb{T}$ ,  $T^{p_n}x \rightarrow x + \frac{1}{2}$ , and so for every  $x \in X$ ,  $\{T^{p_n}x\}$  does not have  $x$  as a limit point.

We also prove a result about simultaneous recurrence. Theorem 2.1.17 implies that for any commuting minimal homeomorphisms  $T$  and  $S$  of a compact space  $X$ , there exists a residual set of  $x$  for which there exists a sequence of positive integers  $\{n_i\}$  such that  $T^{n_i}x \rightarrow x$  and  $S^{n_i}x \rightarrow x$ . The following theorem shows that for some systems, it is the case that this residual set of  $x$  is not all of  $X$ . (In this particular example,  $T$  and  $S$  are powers of the same homeomorphism.)

**Theorem 2.1.21.** *There exists a totally minimal topological dynamical system  $(X, T)$  and a point  $x \in X$  such that for any positive integers  $r \neq s$ , any sequence  $\{n_i\}$  of integers satisfying  $T^{r n_i}x \rightarrow x$  and  $T^{s n_i}x \rightarrow x$  is eventually zero.*

Before proceeding with the proofs, we now give a brief description of the content of this chapter. In Section 2.2, we will describe some general symbolic constructions of

topological dynamical systems with particular mixing properties. At the end of this section, we will arrive at a construction of a system which is totally minimal, totally uniquely ergodic, and topologically mixing, and which has as a parameter a sequence of integers  $\{n_k\}$ .

In Section 2.3, by taking this sequence  $\{n_k\}$  to grow very quickly, we will show that the examples constructed in Section 2.2 are sufficient to prove Theorems 2.1.15 and 2.1.18. Some interesting questions also arise and are answered in Section 2.3 pertaining to the upper Banach density of countable unions of sets of upper Banach density zero.

In Section 2.4, we create a flow under a function with base transformation a skew product, which acts on a connected manifold, and which is totally minimal, totally uniquely ergodic, and topologically mixing. This transformation has as a parameter a function  $f \in C(\mathbb{T})$ . We use conditions of Fayad ([Fa]) on flows under functions to achieve topological mixing, and some conditions of Furstenberg ([Fu]) on skew products to prove total unique ergodicity.

In Section 2.5, by a judicious choice of  $f$ , we use the examples of Section 2.4 to prove Theorems 2.1.16 and 2.1.19.

In Section 2.6, we prove Theorem 2.1.21.

Finally, in Section 2.7 we give some open questions about strengthening our results.

## 2.2 Some general symbolic constructions

**Definition 2.2.1.** An ***alphabet*** is any finite set, whose elements are called ***letters***.

**Definition 2.2.2.** A **symbolic dynamical system** with alphabet  $A$  is a topological dynamical system  $(X, \{\sigma_g\}_{g \in G})$  where for any  $g \in G$ ,  $\sigma_g$  is the shift homeomorphism of  $\Omega = A^G$  defined by  $(\sigma_g \omega)(h) = \omega(hg)$  for  $\omega \in \Omega$ , and  $X$  is any closed subset of  $\Omega$  invariant under each  $\sigma_g$ .

**Definition 2.2.3.** A **word** on the alphabet  $A$  is any element of  $A^F$  for some finite set  $F \subset G$ .  $F$  is called the **shape** of  $w$ .

For any words  $v$  of length  $m$  and  $w$  of length  $n$ , we denote by  $vw$  their concatenation, i.e. the word  $v[1]v[2] \dots v[m]w[1]w[2] \dots w[n]$  of length  $m + n$ . We denote by  $w^k$  the word  $ww \dots w$  given by the concatenation of  $k$  copies of  $w$ .

**Definition 2.2.4.** Given any symbolic dynamical system  $(X, \{\sigma_g\}_{g \in G})$ , the **language** of  $X$ , denoted by  $L(X)$ , is the set of all words which appear as subwords of  $X$ .

All symbolic dynamical systems appearing in this chapter have alphabet  $\{0, 1\}$ ,  $G = \mathbb{Z}$ . We also restrict ourselves in this chapter to words with shape  $\{1, \dots, n\}$  for some  $n$ . In other words, a word can be thought of for now as a finite string of letters  $w = w(1)w(2) \dots w(n)$ . The number of letters in a word  $w$  is called its **length**.

**Definition 2.2.5.** A word  $w$  of length  $n$  is a **subword** of a sequence  $u \in \Omega$  if there exists  $k \in \mathbb{Z}$  such that  $u(i + k) = w(i)$  for  $1 \leq i \leq n$ . Analogously,  $w$  is a **subword** of a word  $v$  of length  $m$  if there exists  $0 < k < m - n$  such that  $v(i + k) = w(i)$  for  $1 \leq i \leq n$ .

We will outline three constructions which algorithmically create  $X$  such that  $(X, \sigma)$  have certain properties. (Here the “certain properties” in question depend on which

construction is used.) It should also be mentioned that many ideas from these constructions are taken from work of Hahn and Katznelson ([HaK]), where they also algorithmically constructed symbolic topological dynamical systems with certain ergodicity and mixing properties.

**Construction 1: (Minimal)** We define inductively  $n_k$  and  $A_k$ , which are, respectively, sequences of positive integers and sets of words on the alphabet  $\{0, 1\}$ . Each word in  $A_k$  is of length  $n_k$ . (We will use the term “ $A_k$ -word” to refer to a member of  $A_k$  from now on.) We define these as follows: always define  $n_1 = 1$  and  $A_1 = \{0, 1\}$ . Then, for any  $k \geq 1$ ,  $n_{k+1}$  is defined to be any integer greater than or equal to  $n_k|A_k|$  which is also a multiple of  $n_k$ , and then  $A_{k+1}$  is chosen to be the set of words of length  $n_{k+1}$  which are concatenations of  $A_k$ -words, containing each  $A_k$ -word in the concatenation at least once.

We then define  $X$  to be the set of all  $x \in \Omega$  which are limits of shifted  $A_k$ -words. In other words,  $x \in X$  if there exist  $w_k \in A_k$  and  $m_k \in \mathbb{Z}$  such that for all large enough  $i$ ,  $x(i) = w_k(i - m_k)$ . Since  $\Omega$  is compact, such  $x$  exist by a standard diagonalization argument. It is easily checkable that  $X$  is closed and  $\sigma$ -invariant. The claim is that regardless of the choice of the integers  $n_k$ , as long as  $n_k$  divides  $n_{k+1}$ , and  $n_{k+1} \geq n_k|A_k|$ ,  $(X, \sigma)$  is minimal. It suffices to show that for any  $y \in X$ ,  $\overline{\{\sigma^n y\}_{n \in \mathbb{Z}}} = X$ . Choose any  $y \in X$  and  $w \in L(X)$ . By the definition of  $X$ , there exists  $k$  and some  $A_k$ -word  $w_k$  such that  $w$  is a subword of  $w_k$ .  $w_k$  is an  $A_k$ -word, so by definition, every  $A_{k+1}$ -word contains  $w_k$ , and therefore  $w$ , as a subword. Finally, note that again by the definition of Construction 1,  $y$  is an biinfinite concatenation of  $A_{k+1}$ -words.

This implies that  $y(1) \dots y(2n_{k+1})$  contains some  $A_{k+1}$  word, and therefore  $w$ , as a subword, and so there exists  $n \in \mathbb{Z}$  so that  $\sigma^n y$  begins with  $w$ . Since  $w$  was an arbitrary word in  $L(X)$ , this implies that  $\overline{\{\sigma^n y\}_{n \in \mathbb{Z}}} = X$ , and so  $(X, \sigma)$  is minimal.

□

So, we have now demonstrated a way of constructing minimal  $(X, \sigma)$ . We will now make this construction a bit more complex in order to make  $(X, \sigma)$  totally minimal.

**Construction 2: (Totally minimal)** We define inductively  $n_k$  and  $A_k$ , which are, respectively, sequences of positive integers and sets of words on the alphabet  $\{0, 1\}$ . Each word in  $A_k$  is of length  $n_k$ . We define these as follows: always define  $n_1 = 1$  and  $A_1 = \{0, 1\}$ . Then, for any  $k \geq 1$ ,  $n_{k+1}$  is defined to be any integer greater than or equal to  $(k!)^2 n_k |A_k| + k! + n_k^2$ , and then  $A_{k+1}$  is chosen to be the set of words  $w$  of length  $n_{k+1}$  which are concatenations of  $A_k$ -words and the word 1 with the following properties: the word 1 does not appear at the beginning or end of  $w$ , only a single 1 can be concatenated between two  $A_k$ -words, and for every  $w' \in A_k$ , and for every  $0 \leq i < k!$ ,  $w$  appears in  $w'$  at an  $i \pmod{k!}$ -indexed place. That is, there exists  $m \equiv i \pmod{k!}$  with  $w(m)w(m+1) \dots w(m+n_k-1) = w'$ . From now on, to refer to this second condition, we say that every  $w' \in A_k$  occurs in  $w$  at places indexed by all residue classes modulo  $k!$ .

Since this construction is a bit complicated, a few quick examples may be in order. Suppose that  $n_2 = 6$ ,  $|A_2| = 4$ , and we choose  $n_3 = 134$ . Say that  $A_2 = \{a, b, c, d\}$ . Then  $w = abcd1abcd1dabcdabcabcdab$  is an  $A_3$ -word: each  $A_2$ -word appears at least once beginning with a letter of  $w$  with an odd index, and at least

once beginning with a letter of  $w$  with an even index. Examples of words which would not be  $A_3$ -words include  $abcd11abcdabcdababcabed$  (1 is concatenated twice between  $d$  and  $a$ ),  $abcda1bcd1dbcdabcbbbbbdc$  (occurrences of the word  $a$  begin only with even-indexed letters), or  $abcd1dcbabcdabcdabcdadb$  (wrong number of letters.)

We can then define  $X$  to again be the set of all  $x \in \Omega$  which are limits of shifted  $A_k$ -words. For this definition to make sense, it suffices to show that  $A_k$  is nonempty for all  $k$ , since if this is true then  $X \neq \emptyset$  by compactness of  $\Omega$  and a diagonalization argument just as in Construction 1.  $A_1$  is nonempty, so it suffices to show that  $A_k \neq \emptyset$  implies  $A_{k+1} \neq \emptyset$  for all  $k$ . For any  $k$ , assume that  $w_k \in A_k$ . Then, enumerate the elements of  $A_k$  by  $w_k = a_1, a_2, \dots, a_{|A_k|}$ , and define the words  $u_{k+1} = a_1^{k!} a_2^{k!} \dots a_{|A_k|}^{k!}$  and  $w' = (u_{k+1}1)^{k!} (a_11)^i a_1^{\frac{n_{k+1} - k!(k!n_k|A_k|+1) - i(n_{k+1})}{n_k}}$ , where  $i \equiv n_{k+1} - k! \pmod{n_k}$ . Since  $n_{k+1} > (k!)^2 n_k |A_k| + k! + n_k^2$ ,  $w'$  exists, and is a concatenation of  $A_k$ -words and the word 1 with length  $n_{k+1}$ . In  $w'$ , at most a single 1 is concatenated between any two  $A_k$ -words, and 1 does not appear at the beginning or end of  $w'$ . Also, since the length of  $u_{k+1}$  is divisible by  $k!$ , and since all  $A_k$ -words are subwords of  $u_{k+1}$ , all  $A_k$ -words appear in  $(u_{k+1}1)^{k!}$  at places indexed by all residue classes modulo  $k!$ , and so all  $A_k$ -words appear in  $w'$  at places indexed by all residue classes modulo  $k!$  as well. Therefore,  $w' \in A_{k+1}$ , and so  $A_{k+1}$  is nonempty.

We will show that  $(X, \sigma)$  is totally minimal. Fix any  $m > 0$ . We wish to show that for any  $y \in X$ ,  $\overline{\{\sigma^{mn}y\}_{n \in \mathbb{Z}}} = X$ . Choose any such  $y$ , and fix any word  $w \in L(X)$ . By definition of  $X$ , there exists  $k$  and an  $A_k$ -word  $w_k$  such that  $w$  is a subword of  $w_k$ . Without loss of generality, we assume that  $k > m$ . By the construction,  $w_k$

occurs in every  $A_{k+1}$ -word, and it occurs at places indexed by every residue class modulo  $k!$ . Since  $k > m$ , in particular this implies that  $w_k$ , and therefore  $w$ , occurs in every  $A_{k+1}$ -word at places indexed by every residue class modulo  $m$ . By definition of Construction 2,  $y$  is a biinfinite concatenation of  $A_{k+1}$ -words and the word 1. Therefore,  $y(1) \dots y(2n_{k+1} + 2)$  contains some  $A_{k+1}$ -word as a subword, and so  $w$  occurs in  $y(1) \dots y(2n_{k+1} + 2)$  at places indexed by every residue class modulo  $m$ . There then exists  $n$  so that  $\sigma^{mn}y$  begins with  $w$ . Since  $w$  was an arbitrary word of  $L(X)$ , this shows that  $\overline{\{\sigma^{mn}y\}_{n \in \mathbb{Z}}} = X$ , and since  $m$  was arbitrary, that  $(X, \sigma)$  is totally minimal.

□

We now define one more general type of construction, again more complex than the last, so that the system created is always totally uniquely ergodic and topologically mixing, in addition to being totally minimal. For this last construction, we first need a couple of definitions.

**Definition 2.2.6.** *For any integers  $0 \leq i < m$  and  $k$ , and  $w \in A_{k-1}$  and  $w' \in A_k$ , we define  $fr_{i,m}^*(w, w')$  to be the ratio of the number of occurrences of  $w$  as a concatenated  $A_{k-1}$ -word at  $i \pmod{m}$ -indexed places in  $w'$  to the total number of  $A_{k-1}$ -words concatenated in  $w'$ .*

We consider any positive integer to be equal to 0 (mod 1) for the purposes of this definition. An example is clearly in order: if  $A_1 = \{01, 10\}$ ,  $w = 01$ , (in Constructions 2 and 3,  $A_1$  is always taken to be  $\{0, 1\}$ , but here we deviate from this for



illustrative purposes) and  $w'$  is the  $A_2$ -word  $01|10|1|01|10$  (here vertical bars show where breaks in the concatenation occur), then  $w$  occurs twice out of four  $A_1$ -words, so  $fr_{0,1}^*(w, w') = \frac{1}{2}$ . Since one of these occurrences begins at  $w'(1)$  and one begins at  $w'(6)$ ,  $fr_{0,2}^*(w, w') = fr_{1,2}^*(w, w') = \frac{1}{4}$ . We make a quick note here that there could be some ambiguity here if an  $A_{k+1}$ -word could be decomposed as a concatenation of  $A_k$  words and ones in more than one way. For this reason, we just assume that when computing  $fr_{i,j}^*(w, w')$ , the definition of the  $A_{k+1}$  word  $w'$  includes its representation as a concatenation of  $A_k$ -words and ones. (i.e. in the example given,  $w'$  is defined as the concatenation  $01|10|1|01|10$  of  $A_1$ -words and ones, rather than the nine-letter word  $011010110$ .)

**Definition 2.2.7.** *Given any words  $w'$  of length  $n'$  and  $w$  of length  $n \leq n'$ , and any integers  $0 \leq i < m$ , define  $fr_{i,m}(w, w')$  to be the number of occurrences of  $w$  at  $i \pmod{m}$ -indexed places in  $w'$ , divided by  $n' - n + 1$ .*

Taking the previous example again,  $fr_{0,1}(w, w') = \frac{3}{8}$ , since  $01$  occurs three times as a subword of  $011010110$ . Since two of these occurrences begin at letters of  $w'$  with even indices and one begins at a letter of  $w'$  with odd index,  $fr_{0,2}(w, w') = \frac{2}{8}$  and  $fr_{1,2}(w, w') = \frac{1}{8}$ .

**Construction 3: (Totally minimal, totally uniquely ergodic, and topologically mixing)** We define inductively  $n_k$  and  $A_k$ , which are, respectively, sequences of positive integers and sets of words on the alphabet  $\{0, 1\}$ . Each word in  $A_k$  is of length  $n_k$ . We define these as follows: always define  $n_1 = 1$  and  $A_1 = \{0, 1\}$ . Then, we fix any sequence  $\{d_k\}$  of positive reals such that  $\sum_{k=1}^{\infty} d_k < \infty$  and define, for each

$k \geq 1$ , some  $n_{k+1} = C_k(k+1)!|A_k|n_k + p$  for any integer  $C_k > n_k > \frac{1}{d_k}$  and prime  $n_k < p \leq 2n_k$  (We may choose such a  $p$  by Bertrand's postulate. [HaW]) Note that this implies that  $(n_k, k!) = 1$  for all  $k \in \mathbb{N}$ . We then define  $A_{k+1}$  to be the set of words  $w'$  of length  $n_{k+1}$  with all of the same properties as in Construction 2, along with the property that, for any  $w \in A_k$ , and for any  $0 \leq i < k!$ ,  $fr_{i,k!}^*(w, w') \in [\frac{1-d_k}{k!|A_k|}, \frac{1+d_k}{k!|A_k|}]$ .

$X$  is again taken to be the set of  $x \in \Omega$  which are limits of shifted  $A_k$ -words. For this definition to make sense, it suffices to show that  $A_k$  is nonempty for all  $k$ , since if this is true then  $X \neq \emptyset$  by compactness of  $\Omega$  and a diagonalization argument just as in Construction 1.  $A_1$  is nonempty, so it suffices to show that  $A_k \neq \emptyset$  implies  $A_{k+1} \neq \emptyset$  for all  $k$ . For any  $k$ , assume that  $w_k \in A_k$ . Then, enumerate the elements of  $A_k$  by  $w_k = a_1, a_2, \dots, a_{|A_k|}$ , and define the words  $u_{k+1} = a_1^{k!} a_2^{k!} \dots a_{|A_k|}^{k!}$  and  $w' = (u_{k+1})^{C_k(k+1)-p} (u_{k+1} 1)^p$ . Clearly for large  $k$ ,  $w'$  exists, and is a concatenation of  $A_k$ -words and the word 1 with length  $n_{k+1}$ . In  $w'$ , at most a single 1 is concatenated between any two  $A_k$ -words, and 1 does not appear at the beginning or end of  $w'$ . Since  $(n_k, k!) = 1$ , for every  $0 \leq i < k!$  and  $x \in A_k$ ,  $x$  appears in  $u_{k+1}$  exactly once as a concatenated  $A_k$ -word at an  $i \pmod{k!}$ -indexed place. Therefore,  $x$  appears in  $w'$  exactly  $C_k(k+1)$  times as a concatenated  $A_k$ -word at  $i \pmod{k!}$ -indexed places, and so  $fr_{i,k!}^*(x, w') = \frac{1}{|A_k|n_k k!}$ . Since  $i$  and  $x$  were arbitrary,  $w' \in A_{k+1}$ , and so  $A_{k+1} \neq \emptyset$ .

$(X, \sigma)$  is totally minimal by the same proof used for Construction 2. We claim that  $(X, \sigma)$  is also totally uniquely ergodic. Take any word  $w \in L(X)$ , and any fixed integer  $j$ . We define two sequences  $\{m_k^{(j)}\}$  and  $\{M_k^{(j)}\}$  as follows:  $m_k^{(j)}$  is the minimum value

of  $fr_{i,j}(w, w')$ , where  $0 \leq i < j$  and  $w'$  ranges over all  $A_k$ -words, and  $M_k^{(j)}$  is the maximum value of  $fr_{i,j}(w, w')$ , where  $0 \leq i < j$  and  $w'$  ranges over all  $A_k$ -words.

Suppose that  $m_k^{(j)}$  and  $M_k^{(j)}$  are known, and that  $k > j$ . We wish to show that  $m_{k+1}^{(j)}$  and  $M_{k+1}^{(j)}$  are very close to each other. Let us consider any element  $w'$  of  $A_{k+1}$  and, for any fixed  $0 \leq i < j$ , see how few occurrences of  $w$  there could possibly be at  $i \pmod j$ -indexed places in  $w'$ . By the definition of Construction 3, for every  $w'' \in A_k$ , and  $0 \leq i' < k!$ , the ratio of the number of times  $w''$  occurs as a concatenated  $A_k$ -word in  $w'$  whose first letter is a letter of  $w'$  whose index is equal to  $i' \pmod{k!}$  to the total number of  $A_k$ -words concatenated in  $w'$  is at least  $\frac{1-d_k}{k!|A_k|}$ . Since  $j$  divides  $k!$ , then for any  $0 \leq i' < j$ , the ratio of the number of times that  $w''$  occurs as a concatenated  $A_k$ -word at  $i' \pmod j$ -indexed places in  $w'$  to the total number of  $A_k$ -words concatenated in  $w'$  is at least  $\frac{1-d_k}{j|A_k|}$ . Since the total number of  $A_k$ -words concatenated in  $w'$  is at least  $\frac{n_{k+1}}{n_k+1}$ , this implies that the number of such occurrences of  $w''$  in  $w'$  is at least  $\frac{1-d_k}{j|A_k|} \frac{n_{k+1}}{n_k+1}$  for any  $i'$  and  $w''$ . For any  $w''$  and  $i'$ , the number of times that  $w$  occurs at  $i \pmod j$ -indexed places in  $w'$  as a subword of an occurrence of  $w''$  that occurs at an  $i' \pmod j$ -indexed place in  $w$  is then at least  $\frac{1-d_k}{j|A_k|} \frac{n_{k+1}}{n_k+1} (n_k - |w| + 1) fr_{i-i' \pmod j, j}(w, w'')$ . Summing over all  $w'' \in A_k$  and  $0 \leq i' < j$ , the number of occurrences of  $w$  in  $w'$  at  $i \pmod j$ -indexed places is at least

$$(1 - d_k)n_{k+1} \frac{n_k - |w| + 1}{n_k + 1} \frac{\sum_{w'' \in A_k} \sum_{m=0}^{j-1} fr_{m,j}(w, w'')}{j|A_k|}.$$

Since  $w'$  was arbitrary in  $A_{k+1}$  and  $0 \leq i < j$  was arbitrary,

$$m_{k+1}^{(j)} \geq (1 - d_k) \frac{n_{k+1}}{n_{k+1} - |w| + 1} \frac{n_k - |w| + 1}{n_k + 1} \frac{\sum_{w'' \in A_k} \sum_{m=0}^{j-1} fr_{m,j}(w, w'')}{j|A_k|}.$$

Let us now bound from above the number of occurrences of  $w$  in  $w'$  at  $i \pmod j$ -indexed places. By precisely the same reasons as above, for any  $0 \leq i < j$ , the number of occurrences of  $w$  at  $i \pmod j$ -indexed places which lie entirely within a concatenated  $A_k$ -word in  $w'$  is not more than

$$(1 + d_k) n_{k+1} \frac{n_k - |w| + 1}{n_k} \frac{\sum_{w'' \in A_k} \sum_{m=0}^{j-1} fr_{m,j}(w, w'')}{j|A_k|}.$$

(The denominator of the first fraction changed because there are at most  $\frac{n_{k+1}}{n_k}$   $A_k$ -words concatenated in  $w'$ .) However, it is possible that there are occurrences of  $w$  in  $w'$  which do not lie entirely within a concatenated  $A_k$ -word in  $w'$ . The number of such occurrences of  $w$  is not more than  $|w| + 1$  times the number of concatenated  $A_k$ -words in  $w'$ , which in turn is less than or equal to  $(|w| + 1) \frac{n_{k+1}}{n_k}$ . This means that the number of occurrences of  $w$  at  $i \pmod j$ -indexed places in  $w'$  is bounded from above by

$$(1 + d_k) n_{k+1} \frac{n_k - |w| + 1}{n_k} \frac{\sum_{w'' \in A_k} \sum_{m=0}^{j-1} fr_{m,j}(w, w'')}{j|A_k|} + (|w| + 1) \frac{n_{k+1}}{n_k},$$

and since  $0 \leq i < j$  was arbitrary, this implies that

$$\begin{aligned} M_{k+1}^{(j)} &\leq (1 + d_k) \frac{n_{k+1}}{n_{k+1} - |w| + 1} \frac{n_k - |w| + 1}{n_k} \frac{\sum_{w'' \in A_k} \sum_{m=0}^{j-1} fr_{m,j}(w, w'')}{j|A_k|} \\ &\quad + \frac{n_{k+1}}{n_{k+1} - |w| + 1} \frac{|w| + 1}{n_k}. \end{aligned}$$

This implies that

$$M_{k+1}^{(j)} - m_{k+1}^{(j)} \leq 2d_k \frac{n_{k+1}}{n_{k+1} - |w| + 1} \frac{n_k - |w| + 1}{n_k} \frac{\sum_{w'' \in A_k} \sum_{m=0}^{j-1} fr_{m,j}(w, w'')}{j|A_k|} \\ + \frac{n_{k+1}}{n_{k+1} - |w| + 1} \frac{|w| + 1}{n_k}.$$

Since  $fr_{m,j}(w, w'') \leq 1$  for every  $0 \leq m < j$  and  $w'' \in A_k$ , for large  $k$  this shows that

$$M_{k+1}^{(j)} - m_{k+1}^{(j)} \leq 2d_k + \frac{2(|w|+1)}{n_k}, \text{ which clearly approaches zero as } k \rightarrow \infty.$$

We now note that since

$$m_{k+1}^{(j)} \geq (1 - d_k) \frac{n_{k+1}}{n_{k+1} - |w| + 1} \frac{n_k - |w| + 1}{n_k + 1} \frac{\sum_{w'' \in A_k} \sum_{m=0}^{j-1} fr_{m,j}(w, w'')}{j|A_k|},$$

and since by definition  $fr_{m,j}(w, w'') \geq m_k^{(j)}$  for all  $w'' \in A_k$ , we see that  $m_{k+1}^{(j)} \geq (1 - d_k) \frac{n_{k+1}}{n_{k+1} - |w| + 1} \frac{n_k - |w| + 1}{n_k + 1} m_k^{(j)}$ , and so

$$m_{k+1}^{(j)} - m_k^{(j)} \geq m_k^{(j)} [(1 - d_k) (1 + \frac{|w|}{n_{k+1} - |w| + 1}) (1 - \frac{|w| - 1}{n_k + 1}) - 1] \\ \geq -m_k^{(j)} (d_k + \frac{|w|}{n_k + 1}) \geq -(d_k + \frac{|w|}{n_k}).$$

By almost completely analogous reasoning, for large  $k$

$$M_{k+1}^{(j)} - M_k^{(j)} \leq M_k^{(j)} [(1 + d_k) (1 + \frac{|w| - 1}{n_{k+1} - |w| + 1}) (1 - \frac{|w| - 1}{n_k}) - 1] \\ + \frac{n_{k+1}}{n_{k+1} - |w| + 1} \frac{|w| + 1}{n_k} \leq M_k^{(j)} d_k + \frac{2(|w| + 1)}{n_k} \leq d_k + \frac{2|w| + 1}{n_k}.$$

Therefore,

$$m_{k+1}^{(j)} \leq M_{k+1}^{(j)} \leq M_k^{(j)} + d_k + \frac{2|w| + 1}{n_k} \leq m_k^{(j)} + 2d_{k-1} + \frac{2(|w| + 1)}{n_{k-1}} + d_k + \frac{2|w| + 1}{n_k}$$

$$\leq m_k^{(j)} + 2d_{k-1} + d_k + \frac{4|w| + 3}{n_{k-1}},$$

so  $|m_{k+1}^{(j)} - m_k^{(j)}| \leq d_k + 2d_{k-1} + \frac{4|w|+3}{n_{k-1}}$ . In a completely analogous fashion,  $|M_{k+1}^{(j)} - M_k^{(j)}| \leq d_k + 2d_{k-1} + \frac{3|w|+2}{n_{k-1}}$ . We know that  $\sum_{k=1}^{\infty} d_k$  converges, and since  $n_k \geq 2^k$  for all  $k$ ,  $\sum_{k=1}^{\infty} \frac{1}{n_k}$  converges as well. Therefore, we see that the sequences  $\{m_k^{(j)}\}$  and  $\{M_k^{(j)}\}$  are Cauchy, and converge. Since we also showed that  $M_k^{(j)} - m_k^{(j)} \rightarrow 0$ , we know that they have the same limit, call it  $\alpha$ .

This implies that for very large  $k$ ,  $|fr_{i,j}(w, w') - \alpha|$  is very small for every  $0 \leq i < j$  and  $w' \in A_k$ . We claim that this, in turn, implies that for very large  $N$ ,  $|fr_{i,j}(w, w'') - \alpha|$  is very small for every word  $w'' \in L(X)$  of length  $N$ : fix any  $\epsilon > 0$ , and take  $k$  such that  $|fr_{i,j}(w, w') - \alpha| < \frac{\epsilon}{2}$  for every  $0 \leq i < j$  and  $w' \in A_k$ , and such that  $\frac{1+|w|}{n_k} < \frac{\epsilon}{4}$ . Then for any word  $w'' \in L(X)$  of length at least  $\frac{8n_k}{\epsilon}$ ,  $w''$  is a subword of a concatenation of  $A_k$ -words and copies of the word 1. The number of full  $A_k$ -words appearing in the concatenation forming  $w''$  is at least  $\frac{|w''|}{n_k+1} - 2$ , and at most  $\frac{|w''|}{n_k}$ . So, the number of occurrences of  $w$  at  $i \pmod j$ -indexed places in  $w''$  which are contained entirely within a concatenated  $A_k$ -word is at least  $(\frac{(n_k-|w|+1)|w''|}{n_k+1} - 2(n_k - |w| + 1))(\alpha - \frac{\epsilon}{2}) \geq |w''|((1 - \frac{\epsilon}{4}) - \frac{\epsilon}{4})(\alpha - \frac{\epsilon}{2}) \geq |w''|(\alpha - \epsilon)$ , and at most  $|w''|\frac{n_k-|w|+1}{n_k}(\alpha + \frac{\epsilon}{2}) \leq |w''|(\alpha + \frac{\epsilon}{2})$ . Since there are at most  $\frac{(|w|+1)|w''|}{n_k} < |w''|\frac{\epsilon}{4}$  occurrences of  $w$  not contained entirely within a concatenated  $A_k$ -word, this implies that  $fr_{i,j}(w, w'')$  is at least  $\alpha - \epsilon$ , and at most  $\alpha + \epsilon$ . Since for any  $\epsilon > 0$ , this statement is true for any long enough word  $w'' \in L(X)$  and  $0 \leq i < j$ , we see that  $\frac{1}{n} \sum_{i=0}^{n-1} \chi_{[w]}(\sigma^{ij}y) \rightarrow \alpha$  uniformly for  $y \in X$ . Since  $w$  was arbitrary, and since characteristic functions of cylinder sets are dense in  $C(X)$ ,  $\frac{1}{n} \sum_{i=0}^{n-1} f(\sigma^{ij}y)$  approaches a uniform limit for all  $f \in C(X)$ , and so  $(X, \sigma^j)$

is uniquely ergodic for every  $j \in \mathbb{N}$ . Since an invariant measure for  $(X, \sigma)$  would be invariant for any  $(X, \sigma^j)$  as well, the unique invariant measure is the same for every  $j$ .

Finally, we claim that  $(X, \sigma)$  is also topologically mixing. Consider any words  $w, w' \in L(X)$ . By construction, there exists  $k$  so that there are  $A_k$ -words  $y, y'$  with  $w$  a subword of  $y$  and  $w'$  a subword of  $y'$ . We also claim that for any  $\frac{n_{k+1}}{6} < i < \frac{5n_{k+1}}{6}$ , there exists an  $A_{k+1}$ -word  $b_i$  where  $b_i(i+1)b_i(i+2) \dots b_i(i+n_k) = y$ , and similarly  $b'_i \in A_{k+1}$  with  $b'_i(i+1)b'_i(i+2) \dots b'_i(i+n_k) = y'$ . We show only the existence of  $b_i$ , as the proof for  $b'_i$  is trivially similar. Consider any  $\frac{n_{k+1}}{6} < i < \frac{5n_{k+1}}{6}$ , and take  $j = i \pmod{n_k}$ . Then, if we enumerate the elements of  $A_k$  by  $a_1, a_2, \dots, a_{|A_k|}$ , first define the word  $u_{k+1} = a_1^{k!} a_2^{k!} \dots a_{|A_k|}^{k!}$ , and then define the word  $y_i = (u_{k+1}1)^j (u_{k+1})^{C_k(k+1)-p} (u_{k+1}1)^{p-j}$ .  $y_i$  has the property that  $y_i(i+1)y_i(i+2) \dots y_i(i+n_k)$  is a concatenated  $A_k$ -word in  $y_i$  as long as  $y_i(i+1)y_i(i+2) \dots y_i(i+n_k)$  lies in the subword  $(u_{k+1})^{C_k(k+1)-p}$  of  $y_i$ , which is true for large  $k$  and  $i \in (\frac{n_{k+1}}{6}, \frac{5n_{k+1}}{6})$ . This means that if we reorder  $a_1, \dots, a_{|A_k|}$  in the definition of  $u_{k+1}$ , we may create a word  $b_i$  where  $b_i(i+1)b_i(i+2) \dots b_i(i+n_k) = y$ .  $b_i \in A_{k+1}$ , since for every  $0 \leq i < k!$  and  $x \in A_k$ ,  $x$  appears exactly  $C(k+1)$  times in  $b_i$  as a concatenated  $A_k$ -word at  $i \pmod{k!}$ -indexed places, implying that  $fr_{i,k!}^*(x, b_i) = \frac{1}{|A_k|n_k k!}$  (This uses the fact that  $(n_k, k!) = 1$ , which was already shown.) We create  $b'_i$  in the same way for each  $i$ . Since  $w$  is a subword of  $y$  and  $w'$  is a subword of  $y'$ , for every  $\frac{n_{k+1}}{6} + n_k \leq i \leq \frac{5n_{k+1}}{6} - n_k$ , it is easy to choose a word  $z_i$  to be  $b_j$  for properly chosen  $j$  so that  $z_i(i+1) \dots z_i(i+|w|) = w$ , and similarly  $z'_i$  so that  $z'_i(i+1) \dots z'_i(i+|w'|) = w'$ . For large  $k$ , this means that we can construct such  $z_i$  and  $z'_i$  for any  $i \in [\frac{n_{k+1}}{5}, \frac{4n_{k+1}}{5}]$ .

We will now use these  $z_i$  and  $z'_i$  to prove that for any  $n > |w| + n_{k+1}$ , there exists a word  $x \in L(X)$  of length  $n$  such that  $w x w' \in L(X)$ . We do this by proving a lemma:

**Lemma 2.2.8.** *For any  $t > k + 1$ , and for any  $0 \leq i, j < n_t$  such that there exists an  $A_t$ -word  $x$  where  $x(i+1)x(i+2) \dots x(i+n_{k+1})$  and  $x(j+1)x(j+2) \dots x(j+n_{k+1})$  are concatenated  $A_{k+1}$ -words in  $x$ , and for any two  $A_{k+1}$ -words  $z$  and  $z'$ , there exists an  $A_t$ -word  $x'$  where  $x'(i+1)x'(i+2) \dots x'(i+n_{k+1}) = z$  and  $x'(j+1)x'(j+2) \dots x'(j+n_{k+1}) = z'$ .*

**Proof:** We prove this by induction. First we prove the base case  $t = k + 2$ ; take an  $A_{k+2}$ -word  $x$  where  $x(i+1)x(i+2) \dots x(i+n_{k+1})$  and  $x(j+1)x(j+2) \dots x(j+n_{k+1})$  are concatenated  $A_{k+1}$ -words in  $x$ , call them  $a$  and  $b$  respectively. Since  $x$  is an  $A_{k+2}$ -word, there exists an occurrence of  $z$  at an  $(i \pmod{(k+1)!}) \pmod{(k+1)!}$ -indexed place, i.e. there exists  $i' \equiv i \pmod{(k+1)!}$  such that  $x(i'+1)x(i'+2) \dots x(i'+n_{k+1}) = z$ . Similarly, there exists  $j' \equiv j \pmod{(k+1)!}$  such that  $x(j'+1)x(j'+2) \dots x(j'+n_{k+1}) = z'$ . We now create  $x'$  by leaving almost all of  $x$  alone, but defining  $x'(i+1) \dots x'(i+n_{k+1}) = z$ ,  $x'(j+1) \dots x'(j+n_{k+1}) = z'$ ,  $x'(i'+1) \dots x'(i'+n_{k+1}) = a$ , and  $x'(j'+1) \dots x'(j'+n_{k+1}) = b$ . This new word  $x'$  is still a concatenation of  $A_{k+1}$ -words and ones, and since we switched two pairs of  $A_{k+1}$ -words which occurred at indices with the same residue class modulo  $(k+1)!$ ,  $fr_{i,(k+1)!}^*(w, x) = fr_{i,(k+1)!}^*(w, x')$  for all  $0 \leq i < (k+1)!$  and  $w \in A_{k+1}$ . Therefore,  $x'$  is an  $A_{k+2}$ -word, with  $z$  and  $z'$  occurring at the proper places, completing our proof of the base case.

Now, let us assume that the inductive hypothesis is true for a certain value of  $t$ , and prove it for  $t + 1$ . Consider an  $A_{t+1}$ -word  $x$  where  $x(i+1) \dots x(i+n_{k+1})$  and



$x(j+1) \dots x(j+n_{k+1})$  are concatenated  $A_{k+1}$ -words in  $x$ , call them  $a$  and  $b$ . Call the concatenated  $A_t$ -word that  $x(i+1) \dots x(i+n_{k+1})$  is a subword of  $a'$ , and denote by  $b'$  the corresponding  $A_t$ -word for  $x(j+1) \dots x(j+n_{k+1})$ . From now on, when we speak of these words  $a, b, a', b'$ , we are talking about the pertinent occurrences at the places within  $x$  already described. There are two cases; either  $a'$  and  $b'$  are the same; i.e. the same  $A_t$ -word in  $x$ , occurring at the same place, or they are not. If  $a'$  and  $b'$  do occur at the same place, then by the inductive hypothesis, there exists an  $A_t$ -word  $c$  with an occurrence of  $z$  at the same place as  $a$  occurs in  $a' = b'$ , and an occurrence of  $z'$  at the same place as  $b$  occurs in  $a' = b'$ . If we can replace  $a' = b'$  by  $c$  in  $x$ , then we will be done. If  $a'$  and  $b'$  do not occur at the same place, then since  $a$  is a concatenated  $A_{k+1}$ -word in  $a'$ , by the inductive hypothesis there exists an  $A_t$ -word  $a''$  such that  $a''$  has  $z$  occurring at the same place where  $a$  occurs in  $a'$ . Similarly, there exists an  $A_t$ -word  $b''$  such that  $b''$  has an occurrence of  $z'$  at the same place where  $b$  occurs in  $b'$ . If we replace  $a'$  by  $a''$  and  $b'$  by  $b''$  in  $x$ , then we will be done. So regardless of which case we are in, our goal is to replace one or two chosen  $A_t$ -words within  $x$  with one or two other  $A_t$ -words. We will show how to replace two, which clearly implies that replacing one is possible. We wish to replace  $a'$  by  $a''$  and  $b'$  by  $b''$ . We do this in exactly the same way as in the base case; say that  $a' = x(i'+1) \dots x(i'+n_t)$  and  $b' = x(j'+1) \dots x(j'+n_t)$ . Since  $a'' \in A_t$ , there exists  $i'' = i' \pmod{t!}$  and  $j'' = j' \pmod{t!}$  such that  $x(i''+1) \dots x(i''+n_t) = a''$  and  $x(j''+1) \dots x(j''+n_t) = b''$ . As in the base case, we create  $x'$  by making  $x'(i'+1) \dots x'(i'+n_t) = a''$  and  $x'(i''+1) \dots x'(i''+n_t) = a'$ ,  $x'(j'+1) \dots x'(j'+n_t) = b''$ ,

and  $x'(j'' + 1) \dots x'(j'' + n_t) = b'$ . Then  $x'$  is an  $A_{t+1}$ -word, and by construction  $x'(i + 1) \dots x'(i + n_{k+1}) = z$  and  $x'(j + 1) \dots x'(j + n_{k+1}) = z'$ .

□

Choose any sequence  $\{v_m\}$  of  $A_m$ -words for all  $m > k + 1$ . For any such  $m$ , take  $P_m = \{n : v_m(n + 1) \dots v_m(n + n_{k+1}) \text{ is a concatenated } A_{k+1} \text{ word in } v_m\}$ . Since  $v_m$  is a concatenation of  $A_{k+1}$ -words and ones, if we write the elements of  $P_m$  as  $p_1^{(m)} < p_2^{(m)} < \dots < p_t^{(m)}$ , then for any  $1 \leq l < t$ ,  $p_{l+1}^{(m)} - p_l^{(m)} \leq n_{k+1} + 1$ . For any  $1 < l < t$ , and  $i, j \in [\frac{n_{k+1}}{5}, \frac{4n_{k+1}}{5}]$ , by Lemma 2.2.8, there exists an  $A_m$ -word  $v$  with the property that  $v(p_1^{(m)} + 1) \dots v(p_1^{(m)} + n_{k+1}) = z_i$  and  $v(p_l^{(m)} + 1) \dots v(p_l^{(m)} + n_{k+1}) = z'_j$ . This implies that there is a subword of  $v$  of the form  $wxw'$  where the length of  $x$  is  $p_l^{(m)} - p_1^{(m)} + (j - i) - |w|$ . We note that  $j - i$  can take any integer value between  $-\frac{3n_{k+1}}{5}$  and  $\frac{3n_{k+1}}{5}$  inclusive. Therefore, the set of possible lengths of  $x$  for which  $wxw' \in L(X)$  contains

$$\bigcup_{l=2}^t \{(|w| + p_l^{(m)} - p_1^{(m)} + [-\frac{3n_{k+1}}{5}, \frac{3n_{k+1}}{5}])\}.$$

When  $l$  is increased by one,  $p_l^{(m)}$  is increased by at most  $n_{k+1} + 1$ . This, along with the fact that the intervals  $[-\frac{3n_{k+1}}{5}, \frac{3n_{k+1}}{5}]$  have length  $\frac{6n_{k+1}}{5}$ , which for large  $k$  exceeds  $n_{k+1} + 1$ , implies that this set of possible lengths of  $x$  contains  $[|w| + p_2^{(m)} - p_1^{(m)} - \frac{3n_{k+1}}{5}, |w| + p_t^{(m)} - p_1^{(m)} + \frac{3n_{k+1}}{5}] \supseteq [|w| + n_{k+1}, |w| + n_m - 2n_{k+1}]$ . Since this entire argument could be made for any  $m$ , we see that for any  $n > |w| + n_{k+1}$ , there exists  $x \in L(X)$  of length  $n$  so that  $wxw' \in L(X)$ . Then, for any nonempty open sets

$U, V \subseteq X$ , there exist  $w$  and  $w'$  such that  $[w] \subseteq U$  and  $[w'] \subseteq V$ . By the above arguments, there exists  $N$  so that for any  $n > N$ ,  $[w] \cap \sigma^n[w'] \neq \emptyset$ , implying that  $U \cap \sigma^n V \neq \emptyset$ . This shows that  $(X, \sigma)$  is topologically mixing.

□

## 2.3 Some symbolic counterexamples

**Proof of Theorem 2.1.15:** We take the continuous function  $f(y) = y(0)$  for all  $y \in X$ , and first note that

$$\begin{aligned} & \left\{ y \in X : \frac{1}{N} \sum_{n=0}^{N-1} f(\sigma^{p_n} y) \text{ does not converge} \right\} \\ & \supseteq \left( \bigcap_{n>0} \bigcup_{k>n} \left\{ y \in X : fr_{0,1}(0, y(p_1)y(p_2) \dots y(p_k)) < \frac{1}{4} \right\} \right) \cap \\ & \left( \bigcap_{n>0} \bigcup_{k>n} \left\{ y \in X : fr_{0,1}(0, y(p_1)y(p_2) \dots y(p_k)) > \frac{3}{4} \right\} \right), \end{aligned}$$

and that the latter set, call it  $B$ , is clearly a  $G_\delta$ . We will choose  $n_k$  so that  $B$  is dense in  $X$ . This will imply that  $B$  is a dense  $G_\delta$ , and since  $X$  is a complete metric space, by the Baire category theorem, that  $B$  is residual, which will prove Theorem 2.1.15. Now let us describe the construction of  $\{n_k\}$ .

Recall that we have assumed that the sequence  $\{p_n\}$  has upper Banach density zero. We define  $A := \{p_n : n \in \mathbb{N}\}$ . We also define the intervals of integers  $B_j = [2j!, (j+1)!] \cap \mathbb{N}$  for every  $j \in \mathbb{N}$ , and take any partition of  $\mathbb{N}$  into infinitely many disjoint infinite sets in  $\mathbb{N}$ , call them  $C_1, C_2, \dots$ . Define the set  $D_1 = \bigcup_{j \in C_1} B_j$ , and

then define the set  $A_1 = \{p_n : n \in D_1\} + 1$ . Next, choose some  $r_2$  large enough so that  $(\min C_{r_2})! > 2 \cdot 2$ , and define  $D_2 = \bigcup_{j \in C_{r_2}} B_j$ , and then define  $A_2 = \{p_n : n \in D_2\} + 2$ . Continuing in this way, we may inductively define  $A_k, D_k$  for all  $k \in \mathbb{N}$  so that for all  $k$ ,  $D_k = \bigcup_{j \in C_{r_k}} B_j$  for some  $r_k$  with the property that  $(\min C_{r_k})! > 2k$ , and  $A_k = \{p_n : n \in D_k\} + k$ . We will verify some properties of these sets. Most importantly, we denote by  $H$  the union  $\bigcup_{n=1}^{\infty} A_n$ , and claim that  $d^*(H) = 0$ . We show this by noting that  $H$  has a certain structure;  $H$  consists of shifted subintervals of  $A$ , separated by gaps which approach infinity. More rigorously:

**Lemma 2.3.1.** *There exist intervals  $I_k = [a_k, b_k] \cap \mathbb{N}$  and integers  $j_k$  such that  $H = \bigcup_{k=1}^{\infty} ((A \cap I_k) + j_k)$  and such that  $\lim_{k \rightarrow \infty} \left( \min((A \cap I_{k+1}) + j_{k+1}) - \max((A \cap I_k) + j_k) \right) = \infty$ .*

**Proof:** Take the set  $Q = \bigcup_{k=1}^{\infty} C_{r_k}$ , and denote its members by  $q_1 < q_2 < \dots$ . Then, for any  $k$ ,  $B_{q_k}$  is a subset of some  $D_s$ . The interval  $I_k$  is then defined to be  $[p_{2(q_k)!}, p_{(q_k+1)!}]$ , (which means  $a_k = p_{2(q_k)!}$  and  $b_k = p_{q_k(q_k)!}$ ) and  $j_k$  is defined to be  $s$ . It is just a rewriting of the definition of the  $A_k$  that  $H = \bigcup_{k=1}^{\infty} ((A \cap I_k) + j_k)$  with these notations. All that must be checked is that  $\lim_{k \rightarrow \infty} \left( \min((A \cap I_{k+1}) + j_{k+1}) - \max((A \cap I_k) + j_k) \right) = \infty$ . We will show that  $a_{k+1} + j_{k+1} - b_k - j_k \rightarrow \infty$ , which implies the desired result. Since  $q_{k+1} > q_k$ ,  $a_{k+1} - b_k = (q_{k+1})! > (q_k)!$ . So, we must simply show that  $(q_k)! - j_k \rightarrow \infty$ . Suppose that  $B_{q_k}$  is a subset of  $D_s$ . Then  $j_k = s$ . We also note that by construction,  $(\min C_{r_s})! > 2s$ . But, since  $B_{q_k} \subset D_s$ ,  $q_k \in C_{r_s}$ , and so  $\min C_{r_s} \leq q_k$ . Therefore,  $(q_k)! > 2s$ , and so  $(q_k)! - j_k = (q_k)! - s > \frac{(q_k)!}{2}$ , which

clearly shows that this quantity approaches  $\infty$ , since  $\{q_k\}$  is an increasing sequence of integers.

□

We now prove a general lemma that implies, in particular, that  $d^*(H) = 0$ .

**Lemma 2.3.2.** *If  $d^*(A) = 0$ , and if there exist intervals  $I_k = [a_k, b_k] \cap \mathbb{N}$  and integers  $j_k$  such that  $\lim_{k \rightarrow \infty} \left( \min((A \cap I_{k+1}) + j_{k+1}) - \max((A \cap I_k) + j_k) \right) = \infty$ , then the set  $B = \bigcup_{k=1}^{\infty} (A \cap I_k) + j_k$  has upper Banach density zero.*

**Proof:** Fix  $\epsilon > 0$ . By the fact that  $d^*(A) = 0$ , there exists  $N$  such that for any interval  $J$  of integers of length at least  $N$ ,  $\frac{|A \cap J|}{|J|} < \epsilon$ . Take  $J$  to be any interval of integers of length exactly  $N$ . Since  $\lim_{k \rightarrow \infty} \left( \min((A \cap I_{k+1}) + j_{k+1}) - \max((A \cap I_k) + j_k) \right) = \infty$ , there is some  $K$  such that if  $J$  has nonempty intersection with  $(A \cap I_k) + j_k$  for some  $k > K$ , it is disjoint from  $(A \cap I_{k'}) + j_{k'}$  for every  $k' \neq k$ . Therefore, for intervals  $J$  of integers of length  $N$  with large enough minimum element,  $J \cap B$  consists of a subset of a shifted copy of  $J \cap A$ , and so  $\frac{|B \cap J|}{|J|} \leq \frac{|A \cap J'|}{|J'|}$  for some interval  $J'$  of integers whose length is also  $N$ . This means that in this case,  $\frac{|B \cap J|}{|J|} < \epsilon$ . We have then shown that for every  $\epsilon$ , there exist  $N, M$  such that for any interval of integers  $J$  of length  $N$  with  $\min J > M$ ,  $\frac{|B \cap J|}{|J|} < \epsilon$ . We will show that this slightly modified definition still implies that  $d^*(B) = 0$ . Again fix  $\epsilon > 0$ , and define  $M$  and  $N$  as was just done. Now consider any interval of integers  $I$  with length at least  $\frac{N+M}{\epsilon}$ . Then, partition  $I$  into subintervals: define  $I_0 = I \cap \{1, \dots, M\}$ , and then break  $I \setminus I_0$  into consecutive subintervals of length  $N$ , called  $I_1, I_2, \dots, I_k$ . There may be one last subinterval left

over of length less than  $N$ ; call it  $I_{k+1}$  (which may be empty.) Note that  $|I| \geq Nk$ , or  $\frac{N}{|I|} \leq \frac{1}{k}$ . We see that

$$\begin{aligned} \frac{|B \cap I|}{|I|} &= \frac{|B \cap I_0|}{|I|} + \sum_{i=1}^k \frac{|B \cap I_i|}{|I|} + \frac{|B \cap I_{k+1}|}{|I|} \leq \frac{M}{|I|} + \frac{N}{|I|} \left( \sum_{i=1}^k \frac{|B \cap I_i|}{|I_i|} \right) + \frac{N}{|I|} \\ &\leq \frac{M+N}{|I|} + \frac{1}{k}(k\epsilon) < 2\epsilon. \end{aligned}$$

Since  $\epsilon$  was arbitrary,  $d^*(B) = 0$ .

□

By combining Lemmas 2.3.1 and 2.3.2,  $d^*(H) = 0$ . We will now create  $X$ . We note that this part of our construction uses only the fact that  $d^*(H) = 0$ , and no other properties. We take  $n_1 = 1$  and  $A_1 = \{0, 1\}$ . We recall that  $\{n_k\}$  must be a sequence of integers with the following properties: for all  $k > 1$ ,  $n_{k+1} = C_k(k+1)!|A_k|n_k + p$  for some positive integer  $C_k > n_k > \frac{1}{d_k}$  and prime  $n_k < p \leq 2n_k$ . We also require  $n_k$  to grow quickly enough so that for all  $k$ , and for any interval of integers  $I$  of length at least  $n_{k+1}$ ,  $\frac{|I \cap H|}{|I|} < \frac{d_k}{2k!|A_k|n_k}$ . That we may choose such  $n_k$  is a consequence of the fact that  $d^*(H) = 0$ . Using these  $n_k$ , we define  $A_k$  as in Construction 3. We now prove a lemma:

**Lemma 2.3.3.** *For any  $k \in \mathbb{N}$ ,  $m \in \mathbb{Z}$ , and for any  $u \in \{0, 1\}^{\mathbb{N}}$ , there exists an  $A_k$ -word  $v_{u,k,m}$  such that  $v_{u,k,m}(i - m) = u(i)$  for all  $i \in H \cap [m + 1, m + n_k]$ .*

**Proof:** This is proved by induction on  $k$ . Clearly the hypothesis is true for  $k = 1$  and for any  $u, m$ . Now suppose it to be true for a particular  $k$ . We will show

that it is true for  $k + 1$  and every  $u, m$ . We again construct an auxiliary word  $u_{k+1}$ : enumerate the elements of  $A_k$  by  $a_1, a_2, \dots, a_{|A_k|}$ . Then, we again define the word  $u_{k+1} = a_1^{k!} a_2^{k!} \dots a_{|A_k|}^{k!}$ . Define  $v'_{k+1} = (u_{k+1} 1)^p (u_{k+1})^{C_k(k+1)-p}$ , where  $n_{k+1} = C_k(k+1)!|A_k|n_k + p$  as above. We note that for any  $0 \leq i < k!$  and  $w \in A_k$ ,  $w$  occurs exactly  $C_k(k+1)$  times as a concatenated  $A_k$ -word at  $i \pmod{k!}$ -indexed places in  $v'_{k+1}$ . (This uses the fact that  $(n_k, k!) = 1$  for all  $k$ , which has already been shown.) Now, fix  $m \in \mathbb{Z}$ . We wish to construct an  $A_{k+1}$ -word  $v_{u,k+1,m}$  such that  $v_{u,k+1,m}(i - m) = u(i)$  for all  $i \in H \cap [m + 1, m + n_{k+1}]$ . We begin with the  $A_{k+1}$ -word  $v'_{k+1}$ . Clearly it is not necessarily true that  $v'_{k+1}(i - m) = u(i)$  for all  $i \in H \cap [m + 1, m + n_{k+1}]$ . We force this condition to be true by changing some of the  $A_k$ -words concatenated in  $v'_{k+1}$ . We show that this is possible; for any concatenated  $A_k$ -word in  $v'_{k+1}$ , say  $v'_{k+1}(j)v'_{k+1}(j+1) \dots v'_{k+1}(j+n_k-1)$ , the necessary condition is that ones or zeroes (depending on  $u$ ) be introduced at digits whose indices are of the form  $i - m$  for all  $i \in H \cap [m + j + 1, \dots, m + j + n_k - 1]$ . To do this, we replace this  $A_k$ -word by  $v_{u,k,m+j}$ , which by the inductive hypothesis has the correct digits of  $u$  at the desired places. So, we may change  $v'_{k+1}$  into a concatenation of  $A_k$ -words and ones, call it  $v_{u,k+1,m}$ , which has the proper digits of  $u$  in all desired places. This may be done by changing at most  $|H \cap [m + 1, \dots, m + n_{k+1}]| < \frac{d_k n_{k+1}}{2k!|A_k|n_k} \leq C_k(k+1)d_k$   $A_k$ -words. Therefore, since for any  $0 \leq i < k!$  and  $w \in A_k$ ,  $w$  occurred in  $v'_{k+1}$  as a concatenated  $A_k$ -word at  $i \pmod{k!}$ -indexed places exactly  $C_k(k+1)$  times,  $w$  occurs in  $v_{u,k+1,m}$  as a concatenated  $A_k$ -word at  $i \pmod{k!}$ -indexed places between  $C_k(k+1)(1 - d_k)$  and  $C_k(k+1)(1 + d_k)$  times. This implies that  $fr_{i,k!}^*(w, v_{u,k+1,m}) \in [\frac{1-d_k}{k!|A_k|}, \frac{1+d_k}{k!|A_k|}]$ , and since

$i$  and  $w$  were arbitrary, that  $v_{k+1,m}$  is an  $A_{k+1}$ -word. By induction, Lemma 2.3.3 is proved.

□

This implies in particular that for every  $u, k$  there exists an  $A_k$ -word  $v_{u,k,-n_{k-1}}$  with  $v_{u,k,-n_{k-1}}(i + n_{k-1}) = u(i)$  for all  $i \in H \cap (-n_{k-1}, \dots, n_k - n_{k-1}]$ . By a standard diagonalization argument, there exists a sequence  $\{k_j\}$  and  $x \in \Omega$  such that for all  $i \in \mathbb{Z}$ ,  $x(i) = v_{u,k_j,-n_{k_j-1}}(i + n_{k_j-1})$  for all large enough  $j$ . Since for every  $j$ ,  $v_{u,k_j,-n_{k_j-1}}(i + n_{k_j-1}) = u(i)$  for all  $i \in H \cap (-n_{k_j-1}, n_{k_j} - n_{k_j-1}]$ , clearly  $x(i) = u(i)$  for all  $i \in H$ . Since  $x$  is a limit of shifted  $A_k$ -words,  $x \in X$ . As mentioned above, this entire construction could be done with any set of zero upper Banach density in place of  $H$ , which lets us state the following corollary:

**Corollary 2.3.4.** *For any  $C \subset \mathbb{N}$  with  $d^*(C) = 0$ , there exists  $(X, \sigma)$  totally minimal, totally uniquely ergodic, topologically mixing, and with the property that for any sequence  $u \in \{0, 1\}^{\mathbb{N}}$ , there exists  $x_u \in X$  such that  $x_u(i) = u(i)$  for all  $i \in C$ .*

We use Corollary 2.3.4 to create  $(X, \sigma)$  which proves Theorem 2.1.15. Recall that  $H = \bigcup_{k=1}^{\infty} A_k$ , where  $A_k = \{p_n + k : n \in D_k\}$  for all  $k$ ,  $D_k = \bigcup_{j \in C_{r_k}} B_j$  for some  $r_k$ , and  $B_j = [2j!, (j+1)!] \cap \mathbb{Z}$  for all  $j$ . For each  $k$ , we write the elements of  $C_{r_k}$  in increasing order as  $c_{r_k}^{(1)}, c_{r_k}^{(2)}, \dots$ . We now decompose  $H$  into two disjoint subsets; define  $H_o = \{m \in H : m = p_n + k \text{ for some } n \in B_j \text{ where } j = c_{r_k}^{(i)} \text{ for odd } i\}$ , and  $H_e = \{m \in H : m = p_n + k \text{ for some } n \in B_j \text{ where } j = c_{r_k}^{(i)} \text{ for even } i\}$ . Since  $d^*(H) = 0$ , we use Corollary 2.3.4 to create a totally minimal, totally uniquely



ergodic, and topologically mixing  $(X, \sigma)$  and  $x \in X$  with  $x(n) = 0$  for  $n \in H_o$  and  $x(n) = 1$  for  $n \in H_e$ . Recall that we wish to show that for the continuous function  $f : y \mapsto y(0)$  from  $X$  to  $\{0, 1\}$ , the set of points  $y$  such that  $\frac{1}{N} \sum_{n=0}^{N-1} f(\sigma^{p_n} y)$  does not converge is residual. We showed earlier that it is sufficient to show that the set

$$B = \left( \bigcap_{n>0} \bigcup_{k>n} \left\{ y \in X : fr_{0,1}(0, y(p_1)y(p_2) \dots y(p_k)) < \frac{1}{4} \right\} \right) \cap \left( \bigcap_{n>0} \bigcup_{k>n} \left\{ y \in X : fr_{0,1}(0, y(p_1)y(p_2) \dots y(p_k)) > \frac{3}{4} \right\} \right)$$

is dense in  $X$ . Fix any  $j \in \mathbb{Z}$ . By the construction of  $H$ ,  $\{n : j + p_n \in H\} = D_k$  for some  $k$ ,  $x(i) = 0$  for all  $i \in H_o$ , and  $x(i) = 1$  for all  $i \in H_e$ . In particular,  $x(j + p_n) = 0$  for all  $n$  in  $B_{c_{r_k}^{(i)}}$  for odd  $i$  and  $x(j + p_n) = 1$  for all  $n$  in  $B_{c_{r_k}^{(i)}}$  for even  $i$ . But then for any odd integer  $i$ ,  $(\sigma^j x)(p_n) = 0$  for all integers  $n \in [2(c_{r_k}^{(i)})!, (c_{r_k}^{(i)} + 1)!]$ , and so

$$fr_{0,1}(0, (\sigma^j x)(p_1)(\sigma^j x)(p_2) \dots (\sigma^j x)(p_{(c_{r_k}^{(i)}+1)!})) \geq \frac{c_{r_k}^{(i)} - 1}{c_{r_k}^{(i)} + 1},$$

which is clearly larger than  $\frac{3}{4}$  for sufficiently large  $k$ . Similarly, for any even integer  $i$ ,  $(\sigma^j x)(p_n) = 1$  for all integers  $n \in [2(c_{r_k}^{(i)})!, (c_{r_k}^{(i)} + 1)!]$ , and so

$$fr_{0,1}(0, (\sigma^j x)(p_1)(\sigma^j x)(p_2) \dots (\sigma^j x)(p_{(c_{r_k}^{(i)}+1)!})) \leq \frac{2}{c_{r_k}^{(i)} + 1},$$

which is clearly less than  $\frac{1}{4}$  for sufficiently large  $k$ . Therefore,  $\sigma^j x \in B$ . Since  $j \in \mathbb{Z}$  was arbitrary, the orbit of  $x$  is a subset of  $B$ . Since  $X$  is minimal,  $B$  is dense, which completes the proof of Theorem 2.1.15.

□

We note that this proof in fact shows that the set

$$\left( \bigcap_{n>0} \bigcup_{k>n} \left\{ y \in X : fr_{0,1}(0, y(p_1)y(p_2) \dots y(p_k!)) < \frac{2}{k} \right\} \right) \cap \\ \left( \bigcap_{n>0} \bigcup_{k>n} \left\{ y \in X : fr_{0,1}(0, y(p_1)y(p_2) \dots y(p_k!)) > \frac{k-2}{k} \right\} \right)$$

is a residual set in  $X$ , and so we can also say that for a residual set of  $y \in X$ ,  $\liminf_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(\sigma^{p_n} y) = 0 = \inf_{x \in X} f(x)$  and  $\limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(\sigma^{p_n} y) = 1 = \sup_{x \in X} f(x)$ .

**Proof of Theorem 2.1.18:** We use Corollary 2.3.4. Consider any set  $C = \{p_n\}_{n \in \mathbb{N}}$  with  $d^*(C) = 0$ . Choose any set  $C' \subset \mathbb{N}$  with  $(C + 1) \subset C'$ ,  $1 \in C'$ ,  $d^*(C') = 0$ , and  $|C' \setminus (C + 1)| = \infty$ . Denote the elements of  $C' \setminus (C + 1)$  by  $b_1 < b_2 < \dots$ . By Corollary 2.3.4, we may construct a totally minimal, totally uniquely ergodic, topologically mixing  $(X, \sigma)$  with the property that for every  $u \in \{0, 1\}^{\mathbb{N}}$ , there exists  $x_u \in X$  with  $x_u(i) = u(i)$  for all  $i \in C'$ . For every  $v \in \{0, 1\}^{\mathbb{N}}$ , define some  $u_v \in \{0, 1\}^{\mathbb{N}}$  by  $u_v(1) = 0$ ,  $u_v(a) = 1$  for all  $a \in C + 1$ , and  $u_v(b_k) = v(k)$  for all  $k \in \mathbb{N}$ . Then, for any such  $v$ ,  $x_{u_v}(i) = 1$  for all  $i \in C + 1$ ,  $x_{u_v}(1) = 0$ , and  $x_{u_v}(b_k) = v(k)$  for all  $k \in \mathbb{Z}$ . Since for all  $n \in \mathbb{N}$ ,  $p_n + 1 \in C + 1$ ,  $(\sigma^{p_n} x_{u_v})(1) = x_{u_v}(p_n + 1) = 1$ , whereas  $x_{u_v}(1) = 0$ . It is then clear that  $x_{u_v}$  is not a limit point of  $\{\sigma^{p_n} x_{u_v}\}$ . Since  $x_{u_v}(b_k) = v(k)$  for all  $k \in \mathbb{N}$ ,  $x_{u_v} \neq x_{u_{v'}}$  for  $v \neq v'$ , and so the set  $\{x_{u_v}\}_{v \in \{0, 1\}^{\mathbb{N}}}$  is uncountable.

□

It is natural to wonder about one aspect of the proof; why is it that in our construction, we only force certain digits to occur along shifted subsets of  $A$ , rather than along entire shifted copies of  $A$ ? The reason comes from a combinatorial fact which is somewhat interesting in its own right:

**Example 2.3.5.** *There exists a set  $D \subseteq \mathbb{N}$  with  $d^*(D) = 0$  with the property that for any infinite set  $G$  of integers, the set  $D + G = \{d + g : d \in D, g \in G\}$  has upper Banach density one.*

**Proof:** We begin with the sequence  $d_n = 3^{\lfloor n/2 \rfloor}$ , where  $\lfloor n/2 \rfloor$  is the maximal integer  $k$  so that  $2^k | n$ . So,  $\{d_n\}$  begins  $1, 3, 1, 9, 1, 3, 1, 27, 1, 3, 1, 9, 1, 3, 1, \dots$ . Then, define  $c_n = \sum_{i=1}^n d_i$ .  $\{c_n\}$  is then an increasing sequence of integers, with the property that the  $n$ th gap  $c_{n+1} - c_n$  is  $d_{n+1}$  for all  $n$ . We first claim that  $d^*(\{c_n\}) = 0$ . Choose any positive integer  $k$ , and any  $2^k + 1$  consecutive elements  $c_i, \dots, c_{i+2^k}$  of the sequence  $\{c_n\}$ . There must be some integer  $j \in [0, 2^k - 1]$  such that  $2^k | i + j$ . This means that  $3^k | d_{i+j}$ , and so that  $c_{i+2^k} - c_i = \sum_{m=i}^{i+2^k-1} d_m > d_{i+j} \geq 3^k$ . This means that any interval of integers of length less than  $3^k$  can contain at most  $2^k$  elements of  $\{c_n\}$  for any  $k$ , which implies that  $d^*(\{c_n\}) = 0$  since  $\lim_{k \rightarrow \infty} \frac{2^k + 1}{3^k} = 0$ . Therefore, if we construct a new sequence by increasing the gaps  $\{d_n\}$ , it will still have upper Banach density zero. We will change  $\{d_n\}$  countably many times, never decreasing any element. In other words, we inductively construct, for every  $k \in \mathbb{N}$ , a sequence  $\{d_n^{(k)}\}$ , so that these sequences are nondecreasing in  $k$ , i.e.  $d_n^{(k)} \geq d_n^{(k-1)}$  for all  $n, k$ . We add the additional hypothesis that for every  $k$ ,  $d(\{n : d_n^{(k)} > d_n\}) = 0$ , in other

words that only a density zero subset of the elements of  $\{d_n\}$  have been changed after any step.

Step 1: We change  $d_n$  for some infinite, but density zero, set of  $n$ , so that for every positive integer  $m$ , there exists  $n$  such that  $d_n = m$ . This is clearly possible; for example, by increasing  $d_n$  for a density zero set of odd  $n$ . Call the resulting sequence  $\{d_n^{(1)}\}$ .

Step  $k$ : ( $k > 1$ ) Assume that we have already defined  $\{d_n^{(k-1)}\}$ , a sequence of integers with the property that  $d_n^{(k-1)} \geq d_n$  for all  $n$ , and that  $d(\{n : d_n^{(k-1)} > d_n\}) = 0$ . Define the set  $R_k \subset \mathbb{N}^k$  by  $R_k = \{(a_1, \dots, a_k) : a_i \in \mathbb{N}, a_i > 3^k\}$ . Since  $\{n : d_n^{(k-1)} > d_n\}$  has density zero, there exist infinitely many intervals of integers  $I_j$  such that  $|I_j| > 2^{k+1}$  for all  $j$ , and so that  $d_n^{(k-1)} = d_n$  for all  $n \in I_j$  for any  $j$ . Therefore, by the construction of the sequence  $\{d_n\}$ , each  $I_j$  contains a subinterval  $I'_j$  of integers of length  $2^k$  so that for every  $j$ , and for all  $n \in I'_j$ ,  $d_n^{(k-1)} \leq 3^k$ . We may also assume, by passing to a subset if necessary, that the union of all  $I'_j$  has density zero. We now take any bijection  $\phi$  from  $\mathbb{N}$  to  $R_{2^k}$ , and for every  $j$ , if  $I'_j = \{s, s+1, \dots, s+2^k-1\}$ , and if  $\phi(j) = (a_1, \dots, a_{2^k}) \in R_{2^k}$ , define  $d_s^{(k)} = a_1, d_{s+1}^{(k)} = a_2, \dots, d_{s+2^k-1}^{(k)} = a_{2^k}$ . After making these changes on each  $I'_j$ , for any  $m$  not in any  $I'_j$ , define  $d_m^{(k)} = d_m^{(k-1)}$ . This defines  $d_n^{(k)}$  for all  $n$ . Since for every  $n$ ,  $d_n^{(k)} \geq d_n^{(k-1)}$ , by the inductive hypothesis we see that  $d_n^{(k)} \geq d_n$  for all  $n$ . Also, since  $d(\{n : d_n^{(k)} > d_n^{(k-1)}\}) = 0$ , and since by the inductive hypothesis,  $d(\{n : d_n^{(k-1)} > d_n\}) = 0$ , we see that  $d(\{n : d_n^{(k)} > d_n\}) = 0$ , completing the inductive step.

It is a consequence of this construction that if we define  $H_k = \{n : d_n^{(k)} > d_n^{(k-1)}\}$

for every  $k$ , the sets  $H_k$  are pairwise disjoint. Therefore, the sequences  $\{d_n^{(k)}\}$  have a pointwise limit, call it  $\{e_n\}$ . By the construction, for any  $k$ , and for any  $k$ -tuple  $(a_1, \dots, a_k)$  of integers all greater than  $3^k$ , there exists  $m$  such that  $d_{m+i-1}^{(k)} = a_i$  for  $1 \leq i \leq k$ . And, since the  $H_k$  are disjoint, this means that  $e_{m+i-1} = a_i$  for  $1 \leq i \leq k$ . Now, define the sequence  $f_n = \sum_{k=1}^n e_n$  for all  $n$ . As already noted, since  $e_n \geq d_n$  for all  $n$ ,  $D = \{f_n\}$  has upper Banach density zero. We claim that for any infinite set  $G$  of integers,  $D + G$  contains arbitrarily long intervals of integers. Fix any infinite  $G = \{g_n\}$ , with  $g_1 < g_2 < \dots$ . For any  $k$ , there exist  $m_1, m_2, \dots, m_{2^k}, m_{2^k+1}$  so that  $g_{m_{j+1}} - g_{m_j} > 3^k$  for every  $1 \leq j \leq 2^k$ . Then, by construction, there exists  $m$  so that  $e_{m+j} = g_{m_{2^k-j+2}} - g_{m_{2^k-j+1}} + 1$  for  $1 \leq j \leq 2^k$ . This means that  $f_{m+i} = f_m + \sum_{j=1}^i e_{m+j} = f_m + \sum_{j=1}^i (g_{m_{2^k-j+2}} - g_{m_{2^k-j+1}} + 1) = (f_m + g_{m_{2^k+1}}) - g_{m_{2^k+1-i}} + i$  for  $1 \leq i \leq 2^k$ . But then for each such  $i$ ,  $f_{m+i} + g_{m_{2^k+1-i}} = f_m + g_{m_{2^k+1}} + i \in D + G$ . This implies that  $\{f_m + g_{m_{2^k+1}} + 1, \dots, f_m + g_{m_{2^k+1}} + 2^k\}$  is an interval of integers of length  $2^k$  which is a subset of  $D + G$ . Since  $k$  was arbitrary,  $D + G$  contains arbitrarily long intervals and so has upper Banach density one.

□

This answers our question: if we had, in the proof of Theorem 2.1.15, tried to force ones to occur along infinitely many shifted copies of our set  $A$  of upper Banach density zero, it's possible that no matter what set  $G$  of shifts we used, we would be attempting to force  $x(i) = 1$  for arbitrarily long intervals of integers  $i$  and some  $x \in X$ , which would imply that  $1^\infty \in X$ , yielding the closed invariant set  $\{1^\infty\} \subsetneq X$  and contradicting the minimality of  $X$ .

## 2.4 Some general constructions on connected manifolds

We will now construct some totally minimal, totally uniquely ergodic, and topologically mixing topological dynamical systems  $(X, T)$  for which  $X$  is a connected manifold, and in Section 2.5 we will use such examples to prove Theorems 2.1.16 and 2.1.19. The constructions in question will use both skew products and flows under functions. We will be repeatedly using the topological space  $\mathbb{T}$ , which can most easily be considered as the half-open interval  $[0, 1)$  with 0 and 1 identified and the operation of addition (mod 1). (Whenever we refer to the addition of elements of  $\mathbb{T}$ , it should be understood to be addition (mod 1).) We also note that  $\mathbb{T}^n$  is a metric space for any  $n$  with metric  $d$  defined by  $d(x, y) = \min_{u, v \in \mathbb{Z}^n} \tilde{d}(x + u, y + v)$ , where  $\tilde{d}$  is the Euclidean metric in  $\mathbb{R}^n$ . We denote by  $\lambda_k$  Lebesgue measure on  $\mathbb{T}^k$  for any  $k > 0$ .

For any  $k > 1$ , irrational  $\alpha \in \mathbb{T}$  and continuous self-map  $f$  of  $\mathbb{T}$ , we define the homeomorphism  $S = S_{k, \alpha, f}$  on  $\mathbb{T}^k$  as follows: for any  $v = (v_1, \dots, v_k) \in \mathbb{T}^k$ ,  $(Sv)_1 = v_1 + \alpha$ ,  $(Sv)_2 = f(v_1) + v_2$ , and for every  $j > 2$ ,  $(Sv)_j = v_{j-1} + v_j$ .

**Theorem 2.4.1.** *For any  $f$  differentiable with  $\frac{1}{2} < f'(x) < \frac{3}{2}$  for all  $x \in \mathbb{T}$ ,  $S_{k, \alpha, f}$  is totally minimal and totally uniquely ergodic with respect to  $\lambda_k$ .*

**Proof:** During the proof, since  $k$ ,  $\alpha$ , and  $f$  are taken to be fixed, we suppress notational dependence and refer to  $S_{k, \alpha, f}$  simply as  $S$ . Fix any rectangles  $R = \prod_{i=1}^k R_i$  and  $R' = \prod_{i=1}^k R'_i$  in  $\mathbb{T}^k$ , where  $R_i$  and  $R'_i$  are intervals of length  $C$  for every  $1 \leq i \leq k$ . Now, fix any positive integer  $\ell$  such that  $\lambda_1((R_1 + \ell\alpha) \cap R'_1) > \frac{C}{2}$ . Fix any  $r_2 \in R_2, \dots, r_k \in R_k$ , and define the set

$$E_1 = \{x_1 \in R_1 : \pi_1(S^\ell(x_1, r_2, \dots, r_k)) \in R'_1\}.$$

(Here and elsewhere,  $\pi_i : \mathbb{T}^k \rightarrow \mathbb{T}$  is the usual projection map  $v \mapsto v_i$  for  $1 \leq i \leq k$ .) Since  $\pi_1(S^\ell(x_1, r_2, \dots, r_k)) = x_1 + \ell\alpha$ , by choice of  $\ell$  we know that  $E_1$  is an interval with  $\lambda_1(E_1) > \frac{C}{2}$ . Now, define the set

$$E_2 = \{x_1 \in R_1 : \pi_i(S^\ell(x_1, r_2, \dots, r_k)) \in R'_i, 1 \leq i \leq 2\}.$$

We will examine the structure of  $E_2$  by bounding  $\frac{\partial \pi_2(S^\ell(x_1, r_2, \dots, r_k))}{\partial x_1}$  from above and below. For this, we note that  $\pi_2(S^\ell(x_1, r_2, \dots, r_k)) = r_2 + \sum_{i=0}^{\ell-1} f(x_1 + i\alpha)$ , and make the observation that for any  $i$ ,  $f(x_1 + i\alpha)$  is equal modulo one to an increasing function in  $x_1$  whose slope is between  $\frac{1}{2}$  and  $\frac{3}{2}$ . Therefore, considered as a function of  $x_1$ ,  $\pi_2(S^\ell(x_1, r_2, \dots, r_k))$  is equal modulo one to an increasing function from  $[0, 1)$  to  $\mathbb{R}$  whose derivative is in  $(\frac{\ell}{2}, \frac{3\ell}{2})$  for all  $x_1$ . This implies that  $E_2$  is a union of many intervals separated by gaps of length less than  $\frac{1}{\ell}$ , where the length of all intervals but the first and last is greater than  $\frac{C}{2\ell}$ . For large  $\ell$ , this implies that  $E_2$  contains some set  $F_2$  a union of intervals of length  $\frac{D_2 C}{\ell}$  for some constant  $D_2$ , where  $\lambda_1(F_2) > B_2 C^2$  for some constant  $B_2$ .

We proceed inductively: for any  $2 \leq i < k$ , assume that we are given  $F_i$  a union of intervals whose lengths are  $D_i C \ell^{-(i-1)}$  for some constant  $D_i$ , where  $\lambda_1(F_i) > B_i C^i$  for some constant  $B_i$ , and such that for any  $x_1 \in F_i$ ,  $\pi_j(S^\ell(x_1, r_2, \dots, r_k)) \in R'_j$  for  $1 \leq j \leq i$ . We now wish to define  $F_{i+1}$ . Define the set

$$E_{i+1} = \{x_1 \in F_i : \pi_{i+1}(S^\ell(x_1, r_2, \dots, r_k)) \in R'_{i+1}\}.$$

For any interval  $I$  of length  $D_i C \ell^{-(i-1)}$  in  $F_i$ , let's examine  $I \cap E_{i+1}$ . We note that

$$\pi_{i+1}(S^\ell(x_1, r_2, \dots, r_k)) = \eta_{i+1, \ell}(r_2, \dots, r_{i+1}) + \sum_{j=0}^{\ell-i} \binom{\ell-j-1}{i-1} f(x_1 + j\alpha),$$

where  $\eta_{i+1, \ell}$  is some function of  $r_2, \dots, r_{i+1}$  which does not depend on  $x_1$ . So, as a function of  $x_1$ ,  $\pi_{i+1}(S^\ell(x_1, r_2, \dots, r_k))$  is equal modulo one to an increasing function from  $[0, 1)$  to  $\mathbb{R}$  whose derivative is between  $A\ell^i$  and  $B\ell^i$  for some constants  $A, B > 0$ . This implies that for large  $\ell$ ,  $I \cap E_{i+1}$  is a union of many intervals separated by gaps of length less than  $\frac{1}{B}\ell^{-i}$ , where the length of all but the first and last is greater than  $\frac{C}{A}\ell^{-i}$ . For large  $\ell$ , this implies that  $I \cap E_{i+1}$  contains  $F_{I, i+1}$  a union of intervals of length  $DC\ell^{-i}$  where  $\lambda_1(F_{I, i+1}) > EC\lambda_1(I)$  for some constants  $D, E$ . By taking  $F_{i+1}$  to be the union of all  $F_{I, i+1}$ , we see that  $F_{i+1}$  is a union of intervals of length  $D_{i+1}C\ell^{-i}$ , where  $\lambda_1(F_{i+1}) > B_{i+1}C^{i+1}$  for some constants  $D_{i+1}$  and  $B_{i+1}$ . Since  $F_{i+1} \subset E_{i+1}$ , for any  $x_1 \in F_{i+1}$ ,  $\pi_j(S^\ell(x_1, r_2, \dots, r_k)) \in R'_j$  for  $1 \leq j \leq i+1$ .

By inductively proceeding in this way, we eventually arrive at a set  $F_k$  where  $\lambda_1(F_k) > B_k C^k$  for some constant  $B_k$ , and where  $S^\ell(x_1, r_2, \dots, r_k) \in R'$  for every  $x_1 \in F_k$ . By integrating over all possible  $r_2, \dots, r_k$ , we see that  $\lambda_k(S^\ell R \cap R') > B_k C^{2k-1}$ . We have then shown that for any  $\ell$  with  $\lambda_1((R_1 + k\alpha) \cap R'_1) > \frac{C}{2}$ ,  $\lambda_k(S^\ell R \cap R') > B_k C^{2k-1}$ . Denote by  $L$  the set of such  $\ell$ . Then if we define  $R_\alpha$  to be the transformation  $x \mapsto x + \alpha$  on  $\mathbb{T}$ , then  $L = \{n : R_\alpha^n(0) \in J\}$ , where  $J = \{x \in \mathbb{T} : \lambda_1((R + x) \cap R') > \frac{C}{2}\}$ . It is easily checked that  $\lambda_1(J) = C$ . This means that for any  $M, N \in \mathbb{N}$ ,



$$\frac{1}{N}|L \cap \{0, M, \dots, M(N-1)\}| = \frac{1}{N} \sum_{i=0}^{N-1} \chi_J((R_\alpha^M)^i 0),$$

which approaches  $\lambda_1(J) = C$  as  $N \rightarrow \infty$  by total unique ergodicity of  $R_\alpha$  with respect to  $\lambda_1$ . Then, for large  $N$ ,

$$\frac{1}{N} \sum_{i=0}^{N-1} \lambda_k(S^{M\ell} R \cap R') \geq \frac{1}{N} |L \cap \{0, M, \dots, M(N-1)\}| B_k C^{2k-1} \rightarrow B_k C^{2k}.$$

Therefore,

$$\liminf_{N \rightarrow \infty} \frac{1}{N} \sum_{\ell=0}^{N-1} \lambda_k(S^{M\ell} R \cap R') \geq B_k \lambda_k(R) \lambda_k(R'). \quad (2.1)$$

We have shown that Equation 2.1 holds for  $R$  and  $R'$  arbitrary congruent cubes in  $\mathbb{T}^k$ . It is clear that it also holds for  $R$  and  $R'$  disjoint unions of congruent cubes. Suppose that for some  $M \in \mathbb{N}$ , there exists a Lebesgue measurable  $S^M$ -invariant set  $A \subseteq \mathbb{T}^k$  with  $\lambda_k(A) \in (0, 1)$ . By taking complements if necessary, without loss of generality we may assume that  $\lambda_k(A) \leq \frac{1}{2}$ . By regularity of Lebesgue measure, there exists  $\epsilon$  and  $A'$  a union of cubes with side length  $\epsilon$  such that  $(A')^c$  is also a union of cubes of side length  $\epsilon$ ,  $\lambda_k(A) \leq \lambda_k(A') \leq \frac{1}{2}$ , and  $\lambda_k(A \triangle A') < \frac{B_k}{2} \lambda_k(A)^2$ . Then, for any  $\ell \in \mathbb{N}$ , since  $S^{M\ell} A = A$  and  $\lambda_k(A \triangle A') < \frac{B_k}{2} \lambda_k(A)^2$ ,  $\lambda_k(A \triangle S^{M\ell} A') < \frac{B_k}{2} \lambda_k(A)^2$  as well. Similarly,  $\lambda_k(A^c \triangle S^{M\ell} (A')^c) < \frac{B_k}{2} \lambda_k(A)^2$ . Therefore,  $\lambda_k(S^{M\ell} A' \cap S^{M\ell} (A')^c) < B_k \lambda_k(A)^2$  for all  $\ell \in \mathbb{N}$ . Since  $\lambda_k(A) \leq \lambda_k(A') \leq \lambda_k((A')^c)$ , this contradicts Equation 2.1.

So,  $S^M$  is ergodic with respect to  $\lambda_k$  for every  $M > 0$ . We claim that this also implies unique ergodicity of  $S^M$  for every  $M > 0$ , which follows from an argument of

Furstenberg, and rests on the fact that  $S^M$  is a skew product over an irrational circle rotation. The following fact is shown in the proof of Lemma 2.1 in [Fu] on p. 578:

For any minimal system  $(X_0, T_0)$  which is uniquely ergodic with respect to a measure  $\mu_0$ , and any skew product  $T$  which acts on  $X = X_0 \times \mathbb{T}$  by  $T(x_0, y) = (T_0 x_0, y + h(x_0))$  where  $h : X_0 \rightarrow \mathbb{T}$  is a continuous function, if  $(X, T)$  is ergodic with respect to  $\mu_0 \times m$ , then  $(X, T)$  is minimal and uniquely ergodic with respect to  $\mu_0 \times m$ .

Denote by  $(S^M)^{(i)}$  the action of  $S$  on its first  $i$  coordinates for any  $1 \leq i \leq k$ . Since they are factors of  $S^M$ , each  $(S^M)^{(i)}$  is ergodic with respect to  $\lambda_i$ . Also, for each  $1 \leq i < k$ ,  $(S^M)^{(i+1)}$  is a skew product as described above with  $T_0 = (S^M)^{(i)}$ . We may then use Furstenberg's result and the fact that  $(S^M)^{(1)}$  is minimal and uniquely ergodic with respect to  $\lambda_1$  (since it is an irrational circle rotation) to see that  $(S^M)^{(2)}$  is minimal and uniquely ergodic with respect to  $\lambda_2$ . We can continue inductively in this fashion to arrive at the fact that  $S^M$  is minimal and uniquely ergodic with respect to  $\lambda_k$ . Since  $M$  was arbitrary,  $S$  is totally minimal and totally uniquely ergodic with respect to  $\lambda_k$ .

□

We will now use these skew products to define flows under functions which have all of the previous properties and are also topologically mixing. Define the continuous function  $g : \mathbb{T}^2 \rightarrow \mathbb{R}$  by

$$g(x, y) = 2 + \operatorname{Re} \left( \sum_{k=2}^{\infty} \frac{e^{2\pi i k x}}{e^k} + \sum_{l=2}^{\infty} \frac{e^{2\pi i l y}}{e^l} \right).$$

Note that  $1 < g(x, y) < 3$  for all  $x, y$ . We then define the space  $X = \{(v, x, y, t) : v \in \mathbb{T}^k, x, y \in \mathbb{T}, 0 \leq t \leq g(x, y)\}$  where  $(v, x, y, g(x, y))$  and  $(S_{k,\alpha,f}v, x + \gamma, y + \gamma', 0)$  are identified for all  $v, x, y$ .  $X$  is then homeomorphic to the mapping torus of  $\mathbb{T}^{k+2}$  and a continuous map, and so is a connected  $(k + 3)$ -manifold. For any irrational  $\gamma, \gamma' \in \mathbb{T}$ , we then define the continuous map  $T_{k+3,\alpha,\gamma,\gamma',f} : X \rightarrow X$  by

$$T_{k+3,\alpha,\gamma,\gamma',f}(v, x, y, t) = \begin{cases} (v, x, y, t + 1) & \text{if } t + 1 < g(x, y), \\ (S_{k,\alpha,f}v, x + \gamma, y + \gamma', t + 1 - g(x, y)) & \text{if } t + 1 \geq g(x, y). \end{cases}$$

Finally, we define  $\mu = (\int_{\mathbb{T}^{k+2}} g \, d\lambda_{k+2})^{-1} \lambda_{k+3}$  a  $T_{k+3,\alpha,\gamma,\gamma',f}$ -invariant Borel probability measure on  $X$ . We will prove the following:

**Theorem 2.4.2.** *For any  $f$  differentiable with  $\frac{1}{2} < f'(x) < \frac{3}{2}$  for all  $x \in \mathbb{T}$  and any irrational  $\alpha, \gamma, \gamma' \in \mathbb{T}$  which are linearly independent and which satisfy  $q'_n > e^{3q_n}$  and  $q_{n+1} > e^{3q'_n}$ , where  $\{q_n\}$  and  $\{q'_n\}$  are the digits in the continued fraction expansions of  $\gamma$  and  $\gamma'$  respectively,  $(X, T_{k+3,\alpha,\gamma,\gamma',f})$  is totally minimal, totally uniquely ergodic, and topologically mixing.*

Again, since  $\alpha, \gamma, \gamma'$ , and  $f$  are fixed, for now we suppress the dependence on these quantities in notation and denote the transformations  $T_{k+3,\alpha,\gamma,\gamma',f}$  and  $S_{k,\alpha,f}$  by  $T$  and  $S$  respectively. We also make the notation, for any integer  $\ell > 0$ ,  $g_\ell(x, y) = \sum_{i=0}^{\ell-1} g(x + i\gamma, y + i\gamma')$ , and define  $g_0(x, y) = 0$ . The proof of Theorem 2.4.2 rests mostly on the following lemma, which is essentially taken from [Fa].

**Lemma 2.4.3.** *For any sufficiently large integer  $n > 0$ ,  $y \in \mathbb{T}$ ,  $\ell \in [\frac{1}{2}e^{2q_n}, 2e^{2q'_n}]$ , and any  $x_0 \in \mathbb{T}$  with  $q_n x_0 \in [\frac{1}{6}, \frac{1}{3}]$ ,*

$$\frac{\ell q_n}{e^{q_n}} < \left| \frac{\partial g_\ell(x, y)}{\partial x}(x_0) \right| < \frac{7\ell q_n}{e^{q_n}}.$$

**Proof:**  $g_\ell(x, y) = 2\ell + \operatorname{Re} \left( \sum_{l=2}^{\infty} \frac{X(\ell, l)}{e^l} e^{2\pi i l x} + \sum_{m=2}^{\infty} \frac{Y(\ell, m)}{e^m} e^{2\pi i m y} \right)$ , where

$$X(\ell, l) = \frac{1 - e^{2\pi i l \ell \gamma}}{1 - e^{2\pi i l \gamma}},$$

$$Y(\ell, m) = \frac{1 - e^{2\pi i \ell m \gamma'}}{1 - e^{2\pi i m \gamma'}}.$$

The following facts are proved in [Fa], p. 454:

$$\text{For all } l \in \mathbb{N} \setminus \{0\}, \ell \in \mathbb{N}, |X(\ell, l)| \leq \ell. \quad (2.2)$$

$$\text{For all } n \in \mathbb{N}, l < q_n, \ell \in \mathbb{N}, |X(\ell, l)| \leq q_n. \quad (2.3)$$

$$\text{For all } n \in \mathbb{N}, l \in (q_n, 2q_n), \ell \in \mathbb{N}, |X(\ell, l)| \leq 2q_n. \quad (2.4)$$

$$\text{For any } \ell \leq \frac{q_{n+1}}{2}, |X(\ell, q_n)| \geq \frac{2\ell}{\pi}. \quad (2.5)$$

$$\text{For any } \ell \leq \frac{q_{n+1}}{2}, |\arg(X(\ell, q_n))| \leq \pi \frac{\ell - 1}{q_{n+1}}. \quad (2.6)$$

We will use these to prove Lemma 2.4.3. It is easy to check that

$$\frac{\partial g_\ell(x, y)}{\partial x}(x_0) = \operatorname{Re} \left( \sum_{l=2}^{\infty} 2\pi i l \frac{X(\ell, l)}{e^l} e^{2\pi i l x_0} \right)$$

$$\begin{aligned}
&= \operatorname{Re} \left( 2\pi i q_n \frac{|X(\ell, q_n)|}{e^{q_n}} e^{2\pi i q_n x_0} \right) + \operatorname{Re} \left( \sum_{l=2}^{q_n-1} 2\pi i l \frac{X(\ell, l)}{e^l} e^{2\pi i l x_0} \right) \\
&+ \operatorname{Re} \left( \sum_{l=q_n+1}^{2q_n-1} 2\pi i l \frac{X(\ell, l)}{e^l} e^{2\pi i l x_0} \right) + \operatorname{Re} \left( \sum_{l=2q_n}^{\infty} 2\pi i l \frac{X(\ell, l)}{e^l} e^{2\pi i l x_0} \right) \\
&+ \operatorname{Re} \left( 2\pi i q_n \frac{X(\ell, q_n) - |X(\ell, q_n)|}{e^{q_n}} e^{2\pi i q_n x_0} \right).
\end{aligned}$$

We bound the first term from above and below and the rest from above.

$$\begin{aligned}
&\left| \operatorname{Re} \left( 2\pi i q_n \frac{|X(\ell, q_n)|}{e^{q_n}} e^{2\pi i q_n x_0} \right) \right| = \frac{|X(\ell, q_n)|}{e^{q_n}} 2\pi q_n |\sin(2\pi q_n x_0)|, \text{ and since } \ell \in [\tfrac{1}{2}e^{2q_n}, 2e^{2q'_n}], \\
&\ell \leq \frac{q_{n+1}}{2}. \text{ By (2.2) and (2.5), } \tfrac{2}{\pi}\ell \leq |X(\ell, q_n)| \leq \ell. \text{ Since } q_n x_0 \in [\tfrac{1}{6}, \tfrac{1}{3}], \tfrac{1}{2} \leq \\
&|\sin(2\pi q_n x_0)| \leq 1. \text{ Therefore, } \tfrac{2\ell q_n}{e^{q_n}} \leq \left| \operatorname{Re} \left( 2\pi i q_n \frac{|X(\ell, q_n)|}{e^{q_n}} e^{2\pi i q_n x_0} \right) \right| \leq \tfrac{2\pi \ell q_n}{e^{q_n}}.
\end{aligned}$$

Next, by (2.3),

$$\left| \operatorname{Re} \left( \sum_{l=2}^{q_n-1} 2\pi i l \frac{X(\ell, l)}{e^l} e^{2\pi i l x_0} \right) \right| \leq 2\pi \sum_{l=2}^{q_n-1} l \frac{|X(\ell, l)|}{e^l} \leq 2\pi \sum_{l=2}^{q_n-1} \frac{q_n l}{e^l} \leq 2\pi q_n^2.$$

We similarly have

$$\left| \operatorname{Re} \left( \sum_{l=q_n+1}^{2q_n-1} 2\pi i l \frac{X(\ell, l)}{e^l} e^{2\pi i l x_0} \right) \right| \leq 4\pi q_n^2.$$

Also, from (2.2), we can conclude that for large  $n$

$$\left| \operatorname{Re} \left( \sum_{l=2q_n}^{\infty} 2\pi i l \frac{X(\ell, l)}{e^l} e^{2\pi i l x_0} \right) \right| \leq 2\pi \sum_{l=2q_n}^{\infty} l \frac{|X(\ell, l)|}{e^l} \leq 2\pi \ell \sum_{l=2q_n}^{\infty} \frac{l}{e^l} \leq \frac{2\pi \ell}{e^{1.5q_n}}.$$

Finally, from (2.6) and (2.2), we see that

$$|X(\ell, q_n) - |X(\ell, q_n)|| \leq 2|X(\ell, q_n)| \arg(X(\ell, q_n)) \leq 2\pi \frac{\ell - 1}{q_{n+1}} \ell,$$

and so since  $\ell \leq 2e^{2q'_n}$  and  $q_{n+1} \geq e^{3q'_n}$ , this implies that

$$\left| \operatorname{Re} \left( 2\pi i q_n \frac{X(\ell, q_n) - |X(\ell, q_n)|}{e^{q_n}} e^{2\pi i q_n x_0} \right) \right| \leq 8\pi^2 e^{-q'_n} q_n \frac{\ell}{e^{q_n}}.$$

By combining all of these bounds,

$$\begin{aligned} \frac{2\ell q_n}{e^{q_n}} - 2\pi q_n^2 - 4\pi q_n^2 - \frac{2\pi\ell}{e^{1.5q_n}} - 8\pi^2 e^{-q'_n} \frac{k\ell q_n}{e^{q_n}} &\leq \left| \frac{\partial g_\ell(x, y)}{\partial x}(x_0) \right| \\ &\leq \frac{2\pi\ell q_n}{e^{q_n}} + 2\pi q_n^2 + 4\pi q_n^2 + \frac{2\pi\ell}{e^{1.5q_n}} + 8\pi^2 e^{-q'_n} \frac{\ell q_n}{e^{q_n}}, \end{aligned}$$

and since  $\ell \geq \frac{1}{2}e^{2q_n}$ , for large  $n$  we have

$$\frac{\ell q_n}{e^{q_n}} \leq \left| \frac{\partial g_\ell(x, y)}{\partial x}(x_0) \right| \leq \frac{7\ell q_n}{e^{q_n}}.$$

□

The proof of the following fact is trivially similar:

**Lemma 2.4.4.** *For any sufficiently large integer  $n > 0$ ,  $x \in \mathbb{T}$ ,  $\ell \in [\frac{1}{2}e^{2q'_n}, 2e^{2q_{n+1}}]$ , and any  $y_0 \in \mathbb{T}$  with  $q'_n y_0 \in [\frac{1}{6}, \frac{1}{3}]$ ,*

$$\frac{\ell q'_n}{e^{q'_n}} < \left| \frac{\partial g_\ell(x, y)}{\partial y}(y_0) \right| < \frac{7\ell q'_n}{e^{q'_n}}.$$

**Proof of Theorem 2.4.2:** Consider any  $m \in [\frac{5}{3}e^{2q_n}, 2e^{2q'_n}]$ , and any cubes  $R = \prod_{i=1}^{k+3} R_i$  and  $R' = \prod_{i=1}^{k+3} R'_i$  where  $R_i$  and  $R'_i$  are intervals of length  $C$  for  $1 \leq i \leq k+3$ . Take intervals  $Q_{k+1}$  and  $Q_{k+2}$  of length  $\frac{C}{2}$  central in  $R_{k+1}$  and  $R_{k+2}$  respectively.

Define  $\tilde{R} = \prod_{i=1}^k R_i$ . We also make the following definition:  $\ell(m, x, y, t)$  is the integer  $\ell$  such that  $g_\ell(x, y) \leq m + t < g_{\ell+1}(x, y)$ . Alternately, for any  $v \in \mathbb{T}^k$ ,

$$T^m(v, x, y, t) = (S^{\ell(m, x, y, t)} v, x + \ell(m, x, y, t)\gamma, y + \ell(m, x, y, t)\gamma', m + t - g_{\ell(m, x, y, t)}(x, y)).$$

Now, fix any  $y \in Q_{k+2}$  and  $t \in Q_{k+3}$ . We first define the set

$$E'_1 = \{x_1 \in Q_{k+1} : q_n x_1 \in \left[\frac{1}{6}, \frac{1}{3}\right]\}.$$

For large  $m$ ,  $\lambda_1(E'_1) > D_1 C$  for some constant  $D_1 > 0$ , and  $E'_1$  is a union of many intervals, where all but the first and last have length  $\frac{1}{6q_n}$ . By removing those, we can define  $E''_1$  a union of intervals of length  $\frac{1}{6q_n}$  where  $\lambda_1(E''_1) > D_2 C$  for some constant  $D_2 > 0$ . We then define the set

$$E_2 = \{x_1 \in E''_1 : \lambda_1(R_1 + \ell(m, x_1, y, t)\alpha \cap R'_1) > \frac{C}{2}, Q_{k+1} + \ell(m, x_1, y, t)\gamma \subset R'_{k+1},$$

$$Q_{k+2} + \ell(m, x_1, y, t)\gamma' \subset R'_{k+2})\}.$$

We wish to use Lemma 2.4.3 to analyze the structure of  $E_2$ . First we note that since  $1 < g(x, y) < 3$  for all  $x, y$ , by definition of  $\ell(m, x, y, t)$ ,  $\ell(m, x, y, t) \leq m + t < 3(\ell(m, x, y, t) + 1)$ , and so for large enough  $n$ ,  $\ell(m, x, y, t) \leq m \leq \frac{10}{3}\ell(m, x, y, t)$ , or  $\ell(m, x, y, t) \in [3m, m]$ . By our assumption on  $m$ , this means that for any  $x, y, t$ ,  $\ell(m, x, y, t) \in [\frac{1}{2}e^{2q_n}, 2e^{2q'_n}]$ . Now, fix any interval  $I$  of length  $\frac{1}{6q_n}$  in  $E''_1$ , and let us examine  $E_2 \cap I$ . By the preceding remarks and Lemma 2.4.3,  $\frac{\ell(m, x_1, y, t)q_n}{e^{q_n}} < \left| \frac{\partial \pi_{k+3} T^m(v, x, y, t)}{\partial x}(x_1) \right| < \frac{7\ell(m, x_1, y, t)q_n}{e^{q_n}}$  for every  $x_1 \in I$ . Since  $\ell(m, x, y, t) \in [3m, m]$ ,

$$\frac{3mq_n}{e^{q_n}} < \left| \frac{\partial \pi_{k+3} T^m(v, x, y, t)}{\partial x}(x_1) \right| < \frac{7mq_n}{e^{q_n}} \quad (2.7)$$

for every  $x_1 \in I$ . This means that  $\frac{\partial \pi_{k+3} T^m(v, x, y, t)}{\partial x}$  has the same sign for all  $x_1 \in I$ , and without loss of generality we assume it to be positive. Let us define  $L$  to be the set of possible values for  $\ell(m, x_1, y, t)$  for  $x_1 \in I$ . (Since  $g_\ell$  is continuous for every  $\ell$ ,  $L$  is an interval of integers.) Then for any fixed  $\ell \in L$ , define  $I_\ell = \{x_1 \in I : \ell(m, x_1, y, t) = \ell\} = \{x_1 \in I : g_\ell(x_1, y) \leq m + t < g_{\ell+1}(x_1, y)\} = \{x_1 \in I : g_\ell(x_1, y) \in [m + t - g(x_1 + \ell\gamma, y + \ell\gamma'), m + t]\}$ . Since  $1 < g(x_1 + \ell\gamma, y + \ell\gamma') < 3$ , by (2.7)  $\frac{1}{\frac{7mq_n}{e^{qn}}} < m(I_\ell) < \frac{3}{\frac{3mq_n}{e^{qn}}}$ , or  $m(I_\ell) \in (\frac{1}{7} \frac{e^{qn}}{mq_n}, 10 \frac{e^{qn}}{mq_n})$  for all  $\ell \in L$  except the smallest and largest, for which  $m(I_\ell)$  could be smaller.

Since  $m > e^{2qn}$ , this means that the number of elements in  $L$  approaches infinity as  $m$  does. Now, we note that  $E_2 \cap I = \bigcup_{\ell \in L'} I_\ell$ , where  $L'$  is the set of  $\ell \in L$  where  $\lambda_1((R_1 + \ell\alpha) \cap R'_1) > \frac{C}{2}$ ,  $Q_{k+1} + \ell\gamma \subset R'_{k+1}$ , and  $Q_{k+2} + \ell\gamma' \subseteq R'_{k+2}$ . Then  $\frac{|L'|}{|L|} = \frac{1}{|L|} \sum_{\ell \in L} \chi_J(R_{\alpha, \gamma, \gamma'}^\ell(\mathbf{0}))$ , where  $R_{\alpha, \gamma, \gamma'}$  is the rotation on  $\mathbb{T}^3$  given by  $(a, b, c) \mapsto (a + \alpha, b + \gamma, c + \gamma')$  and  $J = \{(a, b, c) \in \mathbb{T}^3 : \lambda_1((R_1 + a) \cap R'_1) > \frac{C}{2}, Q_{k+1} + b \subset R'_{k+1}, Q_{k+2} + c \subset R'_{k+2}\}$ . Since  $\alpha$ ,  $\gamma$ , and  $\gamma'$  are rationally independent,  $R_{\alpha, \gamma, \gamma'}$  is uniquely ergodic, and so as  $m \rightarrow \infty$ ,  $\frac{|L'|}{|L|} \rightarrow \lambda_3(J) = \frac{C^3}{4}$ .

Due to the already established bounds on  $\lambda_1(I_\ell)$  for  $\ell \in L$ , this implies that there is a constant  $D_3 > 0$  so that  $\lambda_1(E_2 \cap I) > D_3 C^3 \lambda_1(I)$  for every interval  $I$  in  $E_1''$ . By removing the possibly shorter first and last subintervals of  $E_2 \cap I$ , we have  $F_I \subseteq E_2 \cap I$  a union of intervals of length greater than  $\frac{1}{7} \frac{e^{qn}}{mq_n}$  with  $\lambda_1(F_I) > D_4 C^3 \lambda_1(I)$  for some constant  $D_4 > 0$ . We take  $F_2$  to be the union of all  $F_I$ , and then  $\lambda_1(F_2) > D_5 C^4$  for some constant  $D_5 > 0$ . Finally, we define



$$E_3 = \{x_1 \in F_2 : \pi_{k+3}(T^m(v, x_1, y, t)) \in R'_{k+3} \forall v \in \mathbb{T}^k\}.$$

Note that  $F_2$  is a union of intervals  $I_\ell$ , and so fix any such interval  $I_\ell$  of length greater than  $\frac{1}{7} \frac{e^{q_n}}{mq_n}$ . By definition,  $\pi_{k+3}(T^m(v, x_1, y, t)) = m + t - g_\ell(x_1, y)$  for any  $x_1 \in I_\ell$ , and  $\pi_{k+3}(T^m(v, x_1, y, t))$  ranges monotonically from 0 to  $g(x_1 + \ell\gamma, y + \ell\gamma')$  as  $x_1$  increases over  $I_\ell$ . Since, by Lemma 2.4.3,  $\frac{\ell q_n}{e^{q_n}} < \left| \frac{\partial g_\ell(x, y)}{\partial x}(x_1) \right| < \frac{7\ell q_n}{e^{q_n}}$ ,  $\lambda_1(E_3 \cap I_\ell) > D_6 C \lambda_1(I_\ell)$  for some constant  $D_6 > 0$ . By taking the union over all  $I_\ell$ ,  $\lambda_1(E_3) > D_7 C^5$  for some  $D_7 > 0$ .

Consider any  $x_1 \in E_3$ . We know that  $\lambda_1((R_1 + \ell(m, x_1, y, t)\alpha) \cap R'_1) > \frac{C}{2}$ , and by the proof of Theorem 2.4.1, this implies that if we define the set  $A_{x_1} = \{v \in \tilde{R} : \pi_i(T^m(v, x_1, y, t)) \in R'_i, 1 \leq i \leq k\}$ , then  $\lambda_k(A_{x_1}) > B_k C^{2k-1}$  for some constant  $B_k > 0$ . But this means that for any  $v \in A_{x_1}$ ,

$$\begin{aligned} & T^m(v, x_1, y, t) \\ &= (S^{\ell(m, x_1, y, t)} v, x_1 + \ell(m, x_1, y, t)\gamma, y + \ell(m, x_1, y, t)\gamma', m + t - g_{\ell(m, x_1, y, t)}(x_1, y)) \end{aligned}$$

is in  $R'$  by definitions of  $E_3$  and  $A_{x_1}$ . So, there exists a constant  $D_8 > 0$  such that  $\lambda_{k+1}(\{(v, x_1) : T^m(v, x_1, y, t) \in R'\}) > D_8 C^{2k+4}$ . By integrating over all possible  $y \in Q_{k+2}$ ,  $t \in Q_{k+3}$ , we see that there is a constant  $D_9 > 0$  such that  $\mu((T^m R) \cap R') > \frac{D_9}{4} C^{2k+6} = D_9 \mu(R) \mu(R')$ .

This argument works for any large enough  $m \in [\frac{5}{3}e^{2q_n}, 2e^{2q'_n}]$  for some  $n$ . An analogous argument using Lemma 2.4.4 and which involves varying  $y$  instead of  $x$  shows that the same is true for any large enough  $m \in [\frac{5}{3}e^{2q'_n}, 2e^{q_{n+1}}]$ . However, this implies that

for all sufficiently large  $m$  and any congruent cubes  $R$  and  $R'$ ,  $\mu((T^m R) \cap R') > D_9 \mu(R) \mu(R')$ , which implies that  $(X, T)$  is totally ergodic with respect to  $\mu$  and topologically mixing. It remains to show that  $(X, T)$  is in fact totally uniquely ergodic.

Our proof is similar to the argument of Furstenberg used earlier, however since this is about a flow under a function and his argument was about skew products, we present the proof in its entirety here. Fix any  $M \in \mathbb{N}$ . Since  $T^M$  is ergodic with respect to  $\mu$ ,  $\mu$ -a.e. every point of  $X$  is  $(T^M, \mu)$ -generic. Since  $\mu$  is shift-invariant, and since shifts in the last coordinate commute with  $T^M$ , if a point  $(v, x, y, t) \in X$  is  $(T^M, \mu)$ -generic,  $(v, x, y, t')$  is as well for all  $0 \leq t' < g(x, y)$ . This implies that for  $\lambda_{k+2}$ -a.e.  $(v, x, y)$ , the fiber  $\{(v, x, y, t)\}_{0 \leq t < g(x, y)}$  consists of  $(T^M, \mu)$ -generic points. Denote by  $G$  this set of  $(v, x, y)$  which give rise to  $(T^M, \mu)$ -generic fibers. Choose any  $(v, x, y, t) \in X$ . We know that for large  $n$ ,  $\ell(nM, x, y, t) \in [.3nM, nM]$ , and that  $\ell(nM, x, y, t)$  is increasing in  $n$ . Therefore, the set  $\{\ell(nM, x, y, t) : n \in \mathbb{N}\}$  has positive density. The skew product  $U$  on  $\mathbb{T}^{k+2}$  defined by  $U(v, x, y) = (S_{k, \alpha, f} v, x + \gamma, y + \gamma')$  is totally uniquely ergodic with respect to  $\lambda_{k+2}$  for the same reason as in the proof of Theorem 2.4.1. (The only difference is that here the base case is the rotation on  $\mathbb{T}^3$  given by  $(x, y, z) \mapsto (x + \alpha, y + \gamma, z + \gamma')$  instead of an irrational rotation on  $\mathbb{T}$ . However, since  $\alpha$ ,  $\gamma$ , and  $\gamma'$  are rationally independent, this rotation is also totally uniquely ergodic and totally minimal.) This implies that the set  $\{\ell : U^\ell(v, x, y) \in G\}$  has density one. Together, these facts imply that there exists  $n$  such that  $U^{\ell(nM, x, y, t)}(v, x, y) \in G$ , or that  $T^{nM}(v, x, y, t) = (g, s)$  for some  $g \in G$  and  $s \in \mathbb{R}$ . This means that  $T^{nM}(v, x, y, t)$  is  $(T^M, \mu)$ -generic, and so  $(v, x, y, t)$  is as well. Since  $(v, x, y, t) \in X$  was arbitrary, every point in  $X$  is  $(T^M, \mu)$ -generic, and so  $T^M$  is uniquely ergodic with respect to

$\mu$ . Since  $\mu(U) > 0$  for every nonempty open set  $U$ ,  $T^M$  is minimal as well. Since  $M$  was arbitrary,  $(X, T)$  is totally uniquely ergodic, totally minimal, and topologically mixing.

□

## 2.5 Some counterexamples on connected manifolds

**Proof of Theorem 2.1.19:** Our transformation is  $T_{2d+7, \alpha, \gamma, \gamma', f}$  for properly chosen  $\alpha$ ,  $\gamma$ ,  $\gamma'$ , and  $f$ . We always take  $\alpha$  to be the golden ratio  $\frac{\sqrt{5}-1}{2}$  because of a classical fact from the theory of continued fractions:

**Lemma 2.5.1.** *For any  $n \in \mathbb{N}$ , the distance from  $n\alpha$  to the nearest integer is greater than  $\frac{1}{3n}$ .*

$\gamma$  and  $\gamma'$  can be any irrational elements of  $\mathbb{T}$  satisfying the hypotheses of Theorem 2.4.2. What remains is to define  $f$ . Before doing so, we use our sequence  $\{p_n\}$  to construct another increasing sequence of integers. For any  $n$ , take  $w_n = \ell(p_n, 0, 0, 0)$ . In other words, for any  $v \in \mathbb{T}^{2d+4}$ ,  $T^{p_n}(v, 0, 0, 0) = (S^{w_n}(v), w_n\gamma, w_n\gamma', p_n - g_{w_n}(0, 0))$ . (Here  $S = S_{2d+4, \alpha, f}$  and  $T = T_{2d+7, \alpha, \gamma, \gamma', f}$ .) As before, for large  $n$ ,  $w_n \in [.3p_n, p_n]$ . We claim that  $w_{n+1} < (w_{n+1} - w_n)^{d+1}$  for all large enough  $n$ . This is because  $w_{n+1} - w_n = \ell(p_{n+1} - p_n, w_n\gamma, w_n\gamma', p_n - g_{w_n}(0, 0))$ :

$$\begin{aligned} & \left( S^{w_{n+1}}v, w_{n+1}\gamma, w_{n+1}\gamma', p_{n+1} - g_{w_{n+1}}(0, 0) \right) = T^{p_{n+1}}(v, 0, 0, 0) \\ & = T^{p_{n+1}-p_n}(T^{p_n}(v, 0, 0, 0)) = T^{p_{n+1}-p_n}(S^{w_n}v, w_n\gamma, w_n\gamma', p_n - g_{w_n}(0, 0)) \end{aligned}$$

$$\begin{aligned}
&= \left( S^{w_n + \ell(p_{n+1} - p_n, w_n \gamma, w_n \gamma', p_n - g_{w_n}(0,0))} (S^{w_n}(v)), \right. \\
&\quad \left. (w_n + \ell(p_{n+1} - p_n, w_n \gamma, w_n \gamma', p_n - g_{w_n}(0,0))) \gamma, \right. \\
&\quad \left. (w_n + \ell(p_{n+1} - p_n, w_n \gamma, w_n \gamma', p_n - g_{w_n}(0,0))) \gamma', p_{n+1} - g_{w_{n+1}-1}(0,0) \right)
\end{aligned}$$

Therefore, since  $p_{n+1} - p_n \rightarrow \infty$ ,  $w_{n+1} - w_n \in [.3(p_{n+1} - p_n), p_{n+1} - p_n]$  for large  $n$  for the same reasons as before. This implies for large  $n$  that  $w_{n+1} \leq p_{n+1} < (p_{n+1} - p_n)^d \leq (\frac{10}{3}(w_{n+1} - w_n))^d \leq (\frac{10}{3})^d (w_{n+1} - w_n)^d < (w_{n+1} - w_n)^{d+1}$  since  $w_{n+1} - w_n \rightarrow \infty$ . We wish to choose  $f$  so that  $S^{w_n}(\mathbf{0})$  is bounded away from  $\mathbf{0}$ , where  $\mathbf{0} \in \mathbb{T}^{2d+4}$  is the zero vector.

We define  $f$  as an infinite sum:  $F(v_1) = v_1 + \sum_{i=1}^{\infty} c_i s_{x_i, \epsilon_i}(v_1)$  is a function from  $\mathbb{T}$  to  $\mathbb{R}$ , and  $f(v_1) = F(v_1) \pmod{1}$  is a self-map of  $\mathbb{T}$ . In this sum,  $c_i \in \mathbb{R}^+$  with  $\sum_{i=1}^{\infty} c_i < 1$ ,  $x_i \in \mathbb{T}$  and  $\epsilon_i > 0$  will be chosen later, and the function  $s_{x, \epsilon}(y)$  for any  $x \in \mathbb{T}$  and  $\epsilon > 0$  is a function defined by

$$s_{x, \epsilon}(y) = \begin{cases} \frac{\epsilon}{2\pi} \left[ \cos\left(\frac{\pi}{\epsilon}(y - x)\right) + 1 \right] & \text{if } x - \epsilon \leq y \leq x + \epsilon, \\ 0 & \text{otherwise.} \end{cases}$$

The pertinent properties of  $s_{x, \epsilon}$  are that it is nonzero only on the interval  $[x - \epsilon, x + \epsilon]$ , it attains a maximum of  $\frac{\epsilon}{\pi}$  at  $y = x$ , and that its derivative is bounded from above in absolute value by  $\frac{1}{2}$ . Since each term  $c_i s_{x_i, \epsilon_i}$  in the definition of  $F$  is a differentiable function with derivative bounded from above in absolute value by  $\frac{c_i}{2}$ , and since  $\sum_{i=1}^{\infty} c_i < 1$  and the identity function has derivative one everywhere,  $F$  is a differentiable function with  $F'(v_1) \in (\frac{1}{2}, \frac{3}{2})$  for all  $v_1 \in \mathbb{T}$ . This shows that for any

choice of  $c_i$  with  $\sum_{i=1}^{\infty} c_i < 1$ , and for any choice of  $x_i, \epsilon_i$ , by Theorem 2.4.2,  $(X, T)$  is totally minimal, totally uniquely ergodic, and topologically mixing.

We wish to choose  $f$  so that  $S^{w_n}(\mathbf{0})$  is bounded away from  $\mathbf{0}$ . The only quantities still to be chosen are  $c_i, x_i$  and  $\epsilon_i$ . We note that for any  $1 < k \leq 2d+4$  and any  $j \geq k-1$ ,  $\pi_k(S^j(\mathbf{0})) = \sum_{i=0}^{j-k+1} \binom{j-i-1}{k-2} f(i\alpha)$ . This can be proved by a quick induction, and is left to the reader. In particular, if we make the notation  $y_n = \pi_{2d+4}(S^{w_n}(\mathbf{0}))$  for any  $w_n \geq 2d+3$ , then  $y_n = \sum_{i=0}^{w_n-2d-3} \binom{w_n-i-1}{2d+2} f(i\alpha)$ . Our goal is to choose  $c_i, x_i$ , and  $\epsilon_i$  so that  $y_{w_n} = \frac{1}{3}$  for all sufficiently large  $n$ . To do this, we choose  $x_n = w_n\alpha$  for all  $n$ , and  $\epsilon_n = \inf_{0 \leq i < w_{n+1}, i \neq w_n} |x_n - i\alpha| > \frac{1}{3w_{n+1}}$  by Lemma 2.5.1. This guarantees that  $s_{x_n, \epsilon_n}(i\alpha) = 0$  for any  $0 \leq i < w_{n+1}, i \neq w_n$ . This means that each choice of  $c_i$  that we make changes the values of  $y_n$  for only  $n > i$ , and allows us to finally inductively define  $c_i$ .

Recall that our goal is to ensure that  $y_n = \sum_{i=0}^{w_n-2d-3} \binom{w_n-i-1}{2d+2} f(i\alpha) = \frac{1}{3}$  for all sufficiently large  $n$ . Note that since  $\{w_n\}$  is an increasing sequence of integers,  $w_n \geq n$  for all  $n$ . We have already shown that  $w_n < (w_n - w_{n-1})^{d+1}$  for all large  $n$ , and so  $w_n - w_{n-1} > n^{\frac{1}{d+1}}$ , and so  $w_n = w_1 + \sum_{i=2}^n (w_i - w_{i-1}) > \sum_{i=2}^n i^{\frac{1}{d+1}} > \frac{n}{2} \left(\frac{n}{2}\right)^{\frac{1}{d+1}} > \frac{1}{4} n^{1+\frac{1}{d+1}}$  for all large  $n$ , and so  $\sum_{n=1}^{\infty} w_n^{-1}$  converges, a fact which will be important momentarily. We choose  $N$  so that  $\sum_{n=N+1}^{\infty} w_n^{-1} < \frac{1}{6\pi(2d+2)!}$ . The procedure for defining the sequence  $c_n$  is then as follows:  $c_i = 0$  for  $1 \leq i \leq N$ . For any  $n > N$ , assume that  $y_i = \frac{1}{3}$  for  $N+1 < i \leq n$ . Then, we choose  $c_n$  so that  $y_{n+1} = \frac{1}{3}$ . Note that for any  $n > 1$ ,  $y_n = h_n(c_1, \dots, c_{n-2}) + c_{n-1} \frac{1}{\pi} \epsilon_{n-1} \binom{w_n - w_{n-1} - 1}{2d+2}$  for some  $h_n : \mathbb{T}^{n-2} \rightarrow \mathbb{T}$ . This means that taking  $c_n = \pi \frac{(\frac{1}{3} - h_{n+1}(c_1, \dots, c_{n-1})) \pmod{1}}{\epsilon_n \binom{w_{n+1} - w_n - 1}{2d+2}}$  gives  $y_{n+1} = \frac{1}{3}$ . Note that

$\binom{w_{n+1}-w_n-1}{2d+2} > \frac{1}{2(2d+2)!}(w_{n+1}-w_n)^{2d+2}$  for large  $n$ , and recall that  $\epsilon_n > \frac{1}{3w_{n+1}}$  for all  $n$  by Lemma 2.5.1. This means that  $c_n < 6\pi(2d+2)!w_{n+1}(w_{n+1}-w_n)^{-(2d+2)}$ , which, by the hypothesis on the sequence  $\{w_n\}$ , is less than  $6\pi(2d+2)!(w_{n+1})^{-1}$ , again for sufficiently large  $n$ . This means that  $\sum_{n=1}^{\infty} c_n < 6\pi(2d+2)!\sum_{n=N+1}^{\infty} w_{n+1}^{-1} < 1$  by definition of  $N$ . We have then chosen  $c_n$  so that  $d(S^{w_n}(\mathbf{0}), \mathbf{0}) \geq d(y_n, 0) = \frac{1}{3}$  for all  $n > N$ . For  $n \leq N$ , note that  $\pi_1(S^{w_n}(\mathbf{0})) = w_n\alpha \neq 0$ , since  $\alpha \notin \mathbb{Q}$ . This means that for all  $n$ ,  $d(S^{w_n}(\mathbf{0}), \mathbf{0}) \geq \min(\frac{1}{3}, \min_{1 \leq n \leq N} d(w_n\alpha, 0)) > 0$ .

However, by definition,  $\pi_{2d+4}(T^{p_n}(\mathbf{0}, 0, 0, 0)) = \pi_{2d+4}(S^{w_n}(\mathbf{0}))$  for all  $n$ . Therefore,  $T^{p_n}(\mathbf{0}, 0, 0, 0)$  is bounded away from  $(\mathbf{0}, 0, 0, 0)$ , and so since  $(X, T)$  is totally uniquely ergodic, totally minimal, and topologically mixing, we are done.

□

We note that there was nothing special about the number  $\frac{1}{3}$  in this proof, and so the proof of the following corollary is trivially similar:

**Corollary 2.5.2.** *For any increasing sequence  $\{w_n\}$  of integers with the property that for some integer  $d$ ,  $w_{n+1} < (w_{n+1} - w_n)^{d+1}$  for all sufficiently large  $n$ , and for any sequence  $\{z_n\} \subseteq \mathbb{T}$ , there exists  $f$  satisfying the hypotheses of Theorem 2.4.1 so that for all sufficiently large  $n$ ,  $\pi_{2d+4}((S_{2d+4, \alpha, f})^{w_n}(\mathbf{0})) = z_n$ .*

**Proof of Theorem 2.1.16:** Our transformation is  $T_{2d+9, \alpha, \gamma, \gamma', f}$  for the same  $\alpha$ ,  $\gamma$ , and  $\gamma'$  as before. We use the same strategy as we did to prove Theorem 2.1.15; in other words, we will be forcing certain types of nonrecurrence behavior along a set comprised of a union of infinitely many shifted subsequences of  $\{p_n\}$ .

We now proceed roughly as we did in proving Theorem 2.1.15. We define a sequence  $\{t_n\}$  by shifting different  $p_n$  by different amounts. First, define the intervals of integers  $B_j = [j! + 1, (j + 1)!] \cap \mathbb{N}$  for every  $j \in \mathbb{N}$ , and take any partition of  $\mathbb{N}$  into infinitely many disjoint infinite sets  $C_1, C_2, \dots$ . We denote the elements of  $C_i$ , written in increasing order, by  $c_i^{(1)}, c_i^{(2)}, \dots$ . Choose some  $s_1$  large enough so that  $p_{n+1} - p_n > 2 \cdot 1$  for  $n \geq (\inf C_{s_1})!$ , define the set  $D_1 = \bigcup_{j \in C_{s_1}} B_j$ , and then define  $t_n = p_n + 1$  for all  $n \in D_1$ . Next, choose some  $s_2$  large enough so that  $p_{n+1} - p_n > 2 \cdot 2$  for  $n \geq (\inf C_{s_2})!$ , and define  $D_2 = \bigcup_{j \in C_{s_2}} B_j$ , and then define  $t_n = p_n + 2$  for all  $n \in D_2$ . Continuing in this way, we may inductively define  $D_k$  for all  $k \in \mathbb{N}$  so that  $D_k = \bigcup_{j \in C_{s_k}} B_j$  for some  $s_k$  with the property that  $p_{n+1} - p_n > 2k$  for  $n \geq (\inf C_{s_k})!$ , and then define  $t_n = p_n + k$  for all  $n \in D_k$ . For any  $n \notin \bigcup_{k=1}^{\infty} D_k$ ,  $t_n = p_n$ . Note that by the construction, for any  $n, k$  where  $t_n = p_n + k$ , it must be the case that  $n \in D_k$ , and therefore that  $n > (\inf C_{s_k})!$ , and so that  $p_{n+1} - p_n > 2k$ . Therefore, since  $t_{n+1} \geq p_{n+1}$  for all  $n$ ,  $t_{n+1} - t_n \geq p_{n+1} - p_n - k > \frac{p_{n+1} - p_n}{2}$ , and so  $\{t_n\}$  is increasing. Since  $n - 1 \geq (\inf C_{s_k})!$ ,  $p_n - p_{n-1} > 2k$ , and so in particular  $p_n > 2k$ . This implies that  $t_n = p_n + k < 2p_n$  for all  $n$ . Finally, we see that  $t_{n+1} < 2p_{n+1} < 2(p_{n+1} - p_n)^d < 2^{d+1}(t_{n+1} - t_n)^d$  for all large enough  $n$ . Since  $t_{n+1} - t_n > \frac{p_{n+1} - p_n}{2} \rightarrow \infty$  as  $n \rightarrow \infty$ , this means that  $t_{n+1} < (t_{n+1} - t_n)^{d+1}$  for large enough  $n$ .

We again define a sequence  $\{w_n\}$ : for any  $n$ , take  $w_n = \ell(t_n, 0, 0, 0)$ . In other words, for any  $v \in \mathbb{T}^{2d+6}$ ,  $T^{t_n}(v, 0, 0, 0) = (S^{w_n}(v), w_n\gamma, w_n\gamma', t_n - g_{w_n}(0, 0))$ . (Here  $S = S_{2d+6, \alpha, f}$  and  $T = T_{2d+9, \alpha, \gamma, \gamma', f}$ .) For exactly the same reasons as in the proof of Theorem 2.1.19,  $w_{n+1} < (w_{n+1} - w_n)^{d+2}$  for large  $n$ .

Therefore, by using Corollary 2.5.2, for any sequence  $\{z_n\} \subseteq \mathbb{T}$ , we may choose  $f$  such that  $\pi_{2d+6}(T^{t_n}(\mathbf{0}, 0, 0, 0)) = \pi_{2d+6}(S^{w_n}(\mathbf{0})) = z_n$  for all  $n$ . We define  $z_n = \frac{1}{3}$  for any  $n \in D_k$  where  $n \in B_j$  for  $j = c_{s_k}^{(i)}$  with odd  $i$ , and  $z_n = \frac{2}{3}$  for any  $n \in D_k$  where  $n \in B_j$  for  $j = c_{s_k}^{(i)}$  with even  $i$ . For any  $z_n$  not defined by these conditions,  $z_n$  may be anything. (We can take  $z_n = 0$  for such  $n$  if it is convenient.) Take  $h \in C(X)$  such that  $h(v, x, y, t) = 0$  if  $v_{2d+6} = \frac{2}{3}$ ,  $h(v, x, y, t) = 1$  if  $v_{2d+6} = \frac{1}{3}$ ,  $\inf_{x \in X} h(x) = 0$ , and  $\sup_{x \in X} h(x) = 1$ . Now, we note that

$$\begin{aligned} & \left\{ z \in X : \frac{1}{N} \sum_{n=0}^{N-1} h(T^{p_n} z) \text{ does not converge} \right\} \\ & \supseteq \left( \bigcap_{n>0} \bigcup_{k>n} \left\{ z \in X : \frac{|\{i : 1 \leq i \leq k, \pi_{2d+6}(T^{p_i}(z)) = \frac{1}{3}\}|}{k} > \frac{3}{4} \right\} \right) \cap \\ & \left( \bigcap_{n>0} \bigcup_{k>n} \left\{ z \in X : \frac{|\{i : 1 \leq i \leq k, \pi_{2d+6}(T^{p_i}(z)) = \frac{2}{3}\}|}{k} > \frac{3}{4} \right\} \right), \end{aligned}$$

and that the latter set, call it  $B$ , is clearly a  $G_\delta$ . We will show that  $B$  is dense in  $X$ . Choose any nonempty open set  $U \subset X$ . By minimality of  $T$ , there is some  $k$  so that  $T^k(\mathbf{0}, 0, 0, 0) \in U$ . By construction,  $t_n = p_n + k$  for all  $n \in D_k$ . Also by construction,  $D_k = \bigcup_{j \in C_{s_k}} B_j$ , and  $\pi_{2d+6}(T^{t_n}(\mathbf{0}, 0, 0, 0)) = \pi_{2d+6}(T^{p_n}(T^k(\mathbf{0}, 0, 0, 0))) = \frac{1}{3}$  for  $j = c_{s_k}^{(i)}$  and  $i$  odd. But then for any odd integer  $i$ ,  $\pi_{2d+6}(T^{p_n}(T^k(\mathbf{0}, 0, 0, 0))) = \frac{1}{3}$  for any  $n \in [(c_{s_k}^{(i)})! + 1, (c_{s_k}^{(i)} + 1)!]$ , and so

$$\frac{|\{i : 1 \leq i \leq (c_{s_k}^{(i)} + 1)!, \pi_{2d+6}(T^{p_i}(T^k(\mathbf{0}, 0, 0, 0))) = \frac{1}{3}\}|}{(c_{s_k}^{(i)} + 1)!} \geq \frac{c_{s_k}^{(i)}}{c_{s_k}^{(i)} + 1},$$

which is clearly larger than  $\frac{3}{4}$  for sufficiently large  $k$ . Similarly, for any even integer  $i$ ,  $\pi_{2d+6}(T^{p_n}(T^k(\mathbf{0}, 0, 0, 0))) = \frac{2}{3}$  for any  $n \in [(c_{s_k}^{(i)})! + 1, (c_{s_k}^{(i)} + 1)!]$ , and so



$$\frac{|\{i : 1 \leq i \leq (c_{s_k}^{(i)} + 1)!, \pi_{2d+6}(T^{p_i}(T^k(\mathbf{0}, 0, 0, 0))) = \frac{2}{3}\}|}{(c_{s_k}^{(i)} + 1)!} \geq \frac{c_{s_k}^{(i)}}{c_{s_k}^{(i)} + 1},$$

which is also greater than  $\frac{3}{4}$  for sufficiently large  $k$ . This implies that  $T^k(\mathbf{0}, 0, 0, 0) \in B$ , and so that  $B \cap U$  is nonempty. Since  $U$  was arbitrary, this shows that  $B$  is dense, and so a dense  $G_\delta$ . Therefore,  $B$  is a residual set by the Baire category theorem, and for every  $z \in B$ ,  $\frac{1}{N} \sum_{n=0}^{N-1} h(T^{p_n} z)$  does not converge.

□

We note that exactly as in the proof of Theorem 2.1.15, this in fact shows that for a residual set of  $z \in X$ ,  $\liminf_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} h(T^{p_n} z) = 0 = \inf_{x \in X} h(x)$  and  $\limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} h(T^{p_n} z) = 1 = \sup_{x \in X} h(x)$ .

## 2.6 A counterexample about simultaneous recurrence

Given a totally minimal system  $(X, T)$ , any point  $x \in X$ , and any two positive integers  $r$  and  $s$ , since  $(X, T^r)$  and  $(X, T^s)$  are minimal it must be true that  $x \in \overline{\{T^{rn}x\}_{n \in \mathbb{N}}}$  and  $x \in \overline{\{T^{sn}x\}_{n \in \mathbb{N}}}$ . It is then the case that there exist sequences of positive integers  $\{n_i\}$  and  $\{m_i\}$  such that  $T^{rn_i}x \rightarrow x$  and  $T^{sm_i}x \rightarrow x$ . By Theorem 2.1.17, for a residual set of  $x$ , it is possible to choose  $n_i = m_i$ . To prove Theorem 2.1.21, we must exhibit an example of a totally minimal system for which this residual set is not all of  $X$ .

**Proof of Theorem 2.1.21:** For any  $n$ , we define a symbolic dynamical system  $(X_n, \sigma)$  as follows: define  $x_n \in \Omega = \{0, 1\}^{\mathbb{Z}}$  by  $x_n(i) := \chi_{A_n}(i\alpha_n \pmod{1})$ , where

$\{\alpha_n\}_{n \in \mathbb{N}}$  is any sequence of rationally independent irrational reals, and  $A_n = \{0\} \cup \bigcup_{i=1}^{\infty} [\frac{n}{n^{2i}}, \frac{n+1}{n^{2i}})$ . Then, define  $\sigma$  to be the left shift on  $\Omega$ , and  $T$  to be the countable product of  $\sigma$  on  $\Omega^{\mathbb{N}}$ :  $T(\omega_1, \omega_2, \dots) := (\sigma\omega_1, \sigma\omega_2, \dots)$ .  $X$  is then defined to be the orbit closure of  $x = (x_1, x_2, \dots)$  under  $T$ :  $X = \overline{\{T^n x\}_{n \in \mathbb{Z}}}$ .

We claim that  $(X, T)$  and  $x$  prove Theorem 2.1.21. Consider any  $r > s > 0$ , and choose  $n > r$ . Consider a sequence  $\{n_i\}$  of integers where  $T^{rn_i}x \rightarrow x$  and  $T^{sn_i}x \rightarrow x$ . Then clearly  $\sigma^{rn_i}x_n \rightarrow x_n$  and  $\sigma^{sn_i}x_n \rightarrow x_n$ . Since  $x_n(0) = 1$ , this implies that  $x_n(rn_i) = x_n(sn_i) = 1$  for all large enough  $i$ . Therefore,  $rn_i\alpha_n \pmod{1}, sn_i\alpha_n \pmod{1} \in A_n$  for all large  $i$ . This implies that  $rsn_i\alpha_n \pmod{1} \in rA_n \pmod{1} \cap sA_n \pmod{1}$  for all large  $i$ . However,  $\sup rA_n = \frac{r(n+1)}{n^2} < 1$ , so  $rA_n \pmod{1} = rA_n$  and  $sA_n \pmod{1} = sA_n$ . We then see that  $rsn_i\alpha_n \pmod{1} \in rA_n \cap sA_n$ . If  $rsn_i\alpha_n \pmod{1} \neq 0$ , then  $rsn_i\alpha_n \in [\frac{rn}{n^{2j}}, \frac{r(n+1)}{n^{2j}}) \cap [\frac{sn}{n^{2k}}, \frac{s(n+1)}{n^{2k}})$  for some positive integers  $j, k$ . Since  $n > r, s$ , it must be the case that  $j = k$ , and then we have a contradiction since  $rn = (r-1)n + n \geq sn + n = s(n+1)$ . Therefore,  $rsn_i\alpha_n \pmod{1} = 0$ , and since  $\alpha_n \notin \mathbb{Q}$ ,  $n_i = 0$  for all large enough  $i$ .

It remains to be shown that  $(X, T)$  is totally minimal. Note that for any  $m \in \mathbb{Z}$ ,  $(X, T^m)$  is a system defined similarly to  $(X, T)$ , with the sequence of irrationals  $\{\alpha'_n\}_{n \in \mathbb{N}}$  given by  $\alpha'_n = m\alpha_n$  for  $n \in \mathbb{N}$ . Therefore, it suffices to show that  $(X, T)$  is minimal, since the definition of  $(X, T)$  requires only that the set of irrationals  $\{\alpha_n\}_{n \in \mathbb{N}}$  is rationally independent, and if this is true, it will be true of  $\{\alpha'_n\}_{n \in \mathbb{N}}$  as well. To show that  $(X, T)$  is minimal, we must show that the orbit of any  $y = (y_1, y_2, \dots) \in X$  under  $T$  is dense in  $X$ . It suffices to show that for any finite set of words  $\{w_i\}_{1 \leq i \leq k}$

with  $w_i$  a subword of  $x_i$  for  $1 \leq i \leq k$ , that there exists  $m > 0$  such that  $\sigma^m(y_i) \in [w_i]$  for  $1 \leq i \leq k$ .

Choose any  $1 \leq i \leq k$ . Since  $w_i$  is a subword of  $x_i$ , there exists  $p$  such that  $\sigma^p x_i \in [w_i]$ . This implies that if we denote by  $\ell_i$  the length of  $w_i$ , then the set  $\bigcap_{j=0}^{\ell_i-1} B_j^{(i)}$  contains  $p\alpha \pmod{1}$ , where  $B_j^{(i)} = (A_i - j\alpha_i) \pmod{1}$  if  $w_i(j) = 1$  and  $B_j^{(i)} = ((A_i - j\alpha_i) \pmod{1})^c$  if  $w_i(j) = 0$ . We now claim that the fact that  $\bigcap_{j=0}^{\ell_i-1} B_j^{(i)}$  is nonempty implies that it contains an open interval. Note that for each  $j$ ,  $B_j^{(i)}$  is either a union of half-open intervals or a union of half-open intervals with a singleton, and that due to the irrationality of  $\alpha_i$  and the rationality of all endpoints of intervals in  $A_i$ , the set of endpoints for intervals in  $B_j^{(i)}$  is disjoint from the set of endpoints of intervals in  $B_{j'}^{(i)}$  for any  $0 \leq j \neq j' < \ell_i$ . Since  $\bigcap_{j=0}^{\ell_i-1} B_j^{(i)} \neq \emptyset$ , there exist  $I_j \subset B_j^{(i)}$  such that  $\bigcap_{j=0}^{\ell_i-1} I_j \neq \emptyset$ , each  $I_j$  is either a half-open interval or a singleton, at most one  $I_j$  is a singleton, and the  $I_j$  have no endpoints in common. This implies that either  $\bigcap_{j=0}^{\ell_i-1} I_j$  contains an interval, implying that  $\bigcap_{j=0}^{\ell_i-1} B_j^{(i)}$  does as well, or that for some  $0 \leq j' < \ell_i$ ,  $I_{j'}$  is the singleton  $\{-j'\alpha_i \pmod{1}\}$ . In this case,  $I_j$  is not a singleton for  $j \neq j'$ , and so  $\bigcap_{0 \leq j < \ell_i, j \neq j'} I_j$  contains an interval  $I$ , and  $I_{j'} = \{-j'\alpha_i \pmod{1}\}$  lies in the interior. Choose  $\epsilon$  such that  $I_{j'} + \epsilon \subseteq I$ . Since  $I_{j'}$  is a singleton,  $B_{j'}^{(i)} = (A_i - j'\alpha_i) \pmod{1}$ . Choose  $k$  such that  $\frac{n+1}{n^{2k}} < \epsilon$ . Then  $[\frac{n}{n^{2k}}, \frac{n+1}{n^{2k}}) - j'\alpha_i \pmod{1} \subseteq B_{j'}^{(i)}$ , and by our choice of  $\epsilon$ ,  $[\frac{n}{n^{2k}}, \frac{n+1}{n^{2k}}) - j'\alpha_i \pmod{1} \subseteq I$  as well. Therefore,  $([\frac{n}{n^{2k}}, \frac{n+1}{n^{2k}}) - j'\alpha_i \pmod{1}) \cap (\bigcap_{0 \leq j < \ell_i, j \neq j'} I_j)$  contains an interval, and since  $I_j \subset B_j^{(i)}$  for  $j \neq j'$  and  $[\frac{n}{n^{2k}}, \frac{n+1}{n^{2k}}) - j'\alpha_i \pmod{1} \subseteq B_{j'}^{(i)}$ , we see that  $\bigcap_{j=0}^{\ell_i-1} B_j^{(i)}$  contains an interval as well.

For each  $1 \leq i \leq k$ , denote by  $J_i$  some interval contained in  $\bigcap_{j=0}^{\ell_i-1} B_j^{(i)}$ . Choose

some  $\delta < \min_{1 \leq i \leq k} |J_i|$ . We now note that since the sequence  $\{\alpha_n\}_{n \in \mathbb{N}}$  is rationally independent, the set  $\{(n\alpha_1 \pmod{1}, n\alpha_2 \pmod{1}, \dots, n\alpha_k \pmod{1})\}_{n \in \mathbb{N}}$  is dense in  $[0, 1]^k$ . Therefore, there exists  $r$  such that the set  $B = \{(n\alpha_1 \pmod{1}, n\alpha_2 \pmod{1}, \dots, n\alpha_k \pmod{1})\}_{0 \leq n < r}$  is  $\delta$ -dense, meaning that every point in  $[0, 1]^k$  is within a distance of less than  $\delta$  from some point in  $B$ . This means that  $\{(n\alpha_1 \pmod{1}, n\alpha_2 \pmod{1}, \dots, n\alpha_k \pmod{1})\}_{s \leq n < s+r} = B + (s\alpha_1, s\alpha_2, \dots, s\alpha_k) \pmod{1}$  is also  $\delta$ -dense for any  $s \in \mathbb{Z}$ . We note that then for any  $1 \leq i \leq k$  and any  $m \in \mathbb{Z}$ ,  $m\alpha_i \pmod{1} \in J_i$  implies that  $m\alpha_i \pmod{1} \in \bigcap_{j=0}^{\ell_i-1} B_j^{(i)}$ , and so that  $x_i(m)x_i(m+1) \dots x_i(m+\ell_i-1) = w_i$ , or that  $\sigma^m(x_i) \in [w_i]$ . Define  $\ell = \max_{1 \leq i \leq k} \ell_i$ . Now, choose any  $y = (y_1, y_2, \dots) \in X$ . Since  $y \in \overline{\{T^n x\}_{n \in \mathbb{Z}}}$ , there exists  $p$  such that  $y_i(0)y_i(1) \dots y_i(r+\ell) = x_i(p)x_i(p+1) \dots x_i(p+r+\ell)$  for  $1 \leq i \leq k$ . Then, by the above comments, there exists  $m \in [p, p+r] \cap \mathbb{Z}$  such that  $x_i(m)x_i(m+1) \dots x_i(m+\ell_i-1) = w_i$  for  $1 \leq i \leq k$ . Therefore,  $y_i(m-p)y_i(m-p+1) \dots y_i(m-p+\ell) = x_i(m)x_i(m+1) \dots x_i(m+\ell)$  for  $1 \leq i \leq k$ , and since  $\ell \geq \ell_i$  for  $1 \leq i \leq k$ ,  $\sigma^{m-p}(y_i) \in [w_i]$  for  $1 \leq i \leq k$ . Since  $\{w_i\}_{1 \leq i \leq k}$  were arbitrary, we have shown that the orbit of  $y$  under  $T$  is dense in  $X$ , and since  $y \in X$  was arbitrary, that  $(X, T)$  is minimal. We already showed that this in fact implies that  $(X, T)$  is totally minimal, and so are done.

□

## 2.7 Questions

There are some natural questions motivated by these results. For any totally minimal and totally uniquely ergodic system  $(X, T)$ , any nonempty open set  $U \subseteq X$  with

$\mu(U) \in (0, 1)$ , and any  $x \notin \overline{U}$ , take the set  $A = \{n \in \mathbb{N} : T^n x \in U\}$ . Since  $(X, T)$  is uniquely ergodic, the density  $d(A) := \lim_{n \rightarrow \infty} \frac{|\{1, 2, \dots, n\} \cap A|}{n}$  of  $A$  equals  $\mu(U) > 0$ , and so the sequence  $\{a_n\}$  of the elements of  $A$  written in increasing order does not satisfy the hypotheses of any of our theorems. However, since  $x \notin \overline{U}$ , and since  $T^{a_n} x \in U$  for all  $n$ ,  $T^{a_n} x$  is bounded away from  $x$ .

For a similar example, take  $T$  to be a totally minimal and totally uniquely ergodic isometry of a complete metric space  $(X, d)$  with diameter greater than 2. Take  $x, y \in X$  with  $d(x, y) > 2$ , and define  $f \in C(X)$  where  $f(z) = 0$  for all  $z \in B_1(x)$  and  $f(z) = 1$  for all  $z \in B_1(y)$ . Then, take the sets  $A = \{n \in \mathbb{N} : T^n x \in B_{\frac{1}{2}}(y)\}$  and  $B = \{n \in \mathbb{N} : T^n x \in B_{\frac{1}{2}}(x)\}$ . Since  $(X, T)$  is uniquely ergodic,  $d(A) = \mu(B_{\frac{1}{2}}(y)) > 0$  and  $d(B) = \mu(B_{\frac{1}{2}}(x)) > 0$ . For any  $z \in B_{\frac{1}{2}}(x)$ ,  $d(T^{a_n} z, y) \leq d(T^{a_n} z, T^{a_n} x) + d(T^{a_n} x, y) = d(z, x) + d(T^{a_n} x, y) < \frac{1}{2} + \frac{1}{2} = 1$ , and so  $f(T^{a_n} z) = 1$ . Also for any  $z \in B_{\frac{1}{2}}(x)$ ,  $d(T^{b_n} z, x) \leq d(T^{b_n} z, T^{b_n} x) + d(T^{b_n} x, x) = d(z, x) + d(T^{b_n} x, x) < \frac{1}{2} + \frac{1}{2} = 1$ , and so  $f(T^{b_n} z) = 0$ . This means that by making a sequence  $\{c_n\}$  by alternately choosing longer and longer subsequences of  $\{a_n\}$  and  $\{b_n\}$ , we could create such  $\{c_n\}$  with  $d(\{c_n\}) > 0$  (so  $\{c_n\}$  does not satisfy the hypotheses of any of our theorems) where  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(T^{c_n} z)$  fails to converge for any  $z$  in the second category set  $B_{\frac{1}{2}}(x)$ .

The point of these examples is to show that our hypotheses are certainly not the only ones under which examples of the types constructed in this paper exist. This brings up the following questions:

**Question 2.7.1.** *For what increasing sequences  $\{p_n\}$  of integers does there exist a*

*totally minimal and totally uniquely ergodic system  $(X, T)$  and a point  $x \in X$  such that  $T^{p_n}x$  is bounded away from  $x$ ?*

**Question 2.7.2.** *For what increasing sequences  $\{p_n\}$  of integers does there exist a totally minimal and totally uniquely ergodic system  $(X, T)$  and a function  $f \in C(X)$  such that  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(T^{p_n}x)$  fails to converge for a set of  $x$  of second category?*

Also, it is interesting that we could create examples for a wider class of sequences  $\{p_n\}$  when  $X$  was not connected. We would like to know whether or not this is necessary, i.e.

**Question 2.7.3.** *Given an increasing sequence  $\{p_n\}$  of integers and a totally minimal totally uniquely ergodic topological dynamical system  $(X, T)$  and  $x \in X$  such that  $T^{p_n}x$  is bounded away from  $x$ , must there exist a system with the same properties where  $X$  is a connected space?*

**Question 2.7.4.** *Given an increasing sequence  $\{p_n\}$  of integers and a totally minimal totally uniquely ergodic topological dynamical system  $(X, T)$  and  $f \in C(X)$  such that  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(T^{p_n}x)$  fails to converge for a set of  $x$  of second category, must there exist a system with the same properties where  $X$  is a connected space?*

# CHAPTER 3

## PERTURBATIONS OF MULTIDIMENSIONAL SHIFTS OF FINITE TYPE

### 3.1 Introduction

Consider a symbolic dynamical system with positive topological entropy  $h^{top}(X)$ . (A rigorous definition of topological entropy appears below, but informally it is the exponential growth rate of the number of different  $n$ -letter words appearing as subwords in elements of  $X$ .) Fix any word  $w \in A^{[1, \dots, n]}$  for  $n \in \mathbb{N}$  which appears as a subword of some element of  $X$ . Then, define  $X_w \subseteq \Omega$  to be the set of all elements of  $X$  in which  $w$  is not a subword. Clearly  $X_w \subset X$ , and so the topological entropy of  $X_w$  is not more than that of  $X$ . We wish to estimate the size of the drop in topological entropy  $h^{top}(X) - h^{top}(X_w)$ , and how this quantity behaves as  $n \rightarrow \infty$ . To more rigorously state and approach this problem, we begin with some definitions. Some of these terms have already been defined in Chapter 2, but we repeat the definitions for the more general setup of Chapter 3.

We denote by  $R_{j_1, j_2, \dots, j_d}$  the  $j_1 \times j_2 \times \dots \times j_d$  parallelepiped  $\prod_{i=1}^d \{1, 2, \dots, j_i\}$  in  $\mathbb{Z}^d$ , and use the shorthand notation  $\Gamma_j$  for the cube  $R_{j, j, \dots, j}$ .  $j_1, \dots, j_d$  are the **sizes** of  $R_{j_1, \dots, j_d}$ , and  $j$  is called the **size** of  $\Gamma_j$ . We use  $d$  to denote the  $l_\infty$  metric on  $\mathbb{R}^d$ : for

$p, q \in \mathbb{R}^d$ ,  $d(p, q) := |p - q|_\infty = \max_{1 \leq i \leq d} |p_i - q_i|$ . For each  $1 \leq i \leq d$ , we denote by  $e_i$  the  $i$ th element of the standard basis for  $\mathbb{Z}^d$ , and for  $p, q \in \mathbb{R}^d$ , the **distance in the  $e_i$ -direction** between  $p$  and  $q$  is defined to be  $|p_i - q_i|$ .

**Definition 3.1.1.** An **alphabet** is a finite set. For any alphabet  $A$ , we define the  $\mathbb{Z}^d$ -shift action  $\{\sigma_v\}_{v \in \mathbb{Z}^d}$  on  $\Omega = A^{\mathbb{Z}^d}$  as follows: for any  $v \in \mathbb{Z}^d$  and  $x \in \Omega$ ,  $(\sigma_v(x))(u) = x(v + u)$  for all  $u \in \mathbb{Z}^d$ .

**Definition 3.1.2.** A  $\mathbb{Z}^d$ -**subshift** on an alphabet  $A$  is a set  $X \subseteq \Omega$  with the following two properties:

- (i)  $X$  is shift-invariant, meaning that for any  $x \in X$  and  $p \in \mathbb{Z}^d$ ,  $\sigma_p(x) \in X$ .
- (ii)  $X$  is closed in the product topology on  $\Omega$ .

For any  $\mathbb{Z}^d$ -subshift  $X$ ,  $(X, \{\sigma_v\}_{v \in \mathbb{Z}^d})$  is a symbolic dynamical system.

**Definition 3.1.3.** A Borel probability measure of  $\Omega$  supported on a subshift  $X$  is **ergodic** if the measure-preserving dynamical system  $(X, \mathcal{B}(X), \mu, \{T_v\}_{v \in \mathbb{Z}^d})$  is ergodic, where  $\mathcal{B}(X)$  is the Borel  $\sigma$ -algebra of  $X$ .

**Definition 3.1.4.** A **word**  $u$  on the alphabet  $A$  is any mapping from a non-empty subset  $S$  of  $\mathbb{Z}^d$  to  $A$ , where  $S$  is called the **shape** of  $u$ . For words  $u$  and  $u'$ , where  $u$  has shape  $S$ , we say that  $u$  is a **subword** of  $u'$  if there exists  $p \in \mathbb{Z}^d$  such that  $u(q) = u'(q + p)$  for all  $q \in S$ . In this case, we say that  $u$  **occurs in**  $u'$  at  $S + p$ , or that  $u'(S + p) = u$ .

**Definition 3.1.5.** A word  $w$  with shape  $S$  is **periodic** with respect to  $v \in \mathbb{Z}^d$  if  $S \cap (S - v) \neq \emptyset$  and  $w(u) = w(u + v)$  for all  $u \in S \cap (S - v)$ . We also say that  $v$  is a **period** of  $w$ . A word is **aperiodic** if it has no periods.



**Definition 3.1.6.** For any  $T \subseteq S \subset \mathbb{Z}^d$ , we say that two words  $u$  and  $u'$  with shape  $S$  **agree on**  $T$  if  $u(T) = u'(T)$ , i.e.  $u$  and  $u'$  have the same letters on  $T$ .

**Definition 3.1.7.** The **boundary of thickness**  $k$  of a subset  $S$  of  $\mathbb{Z}^d$ , which is denoted by  $S^{(k)}$ , is the set of points  $p$  of  $S$  for which there exists a point  $q \in \mathbb{Z}^d \setminus S$  with  $d(p, q) \leq k$ . Whenever we refer to the **boundary** of a shape  $S$ , we mean the boundary of thickness one.

**Definition 3.1.8.**  $T \in \mathbb{Z}^d$  is called a **copy** of a shape  $S$  if  $T = S + v$  for some  $v \in \mathbb{Z}^d$ . This  $v$  is then called the **difference** of  $S$  and  $T$ , denoted by  $T - S$ . We say that a point  $p \in S$  **corresponds to** a point  $q \in T$  if  $p = q + (S - T)$ .

**Definition 3.1.9.** The **language** of a  $\mathbb{Z}^d$ -subshift  $X$ , denoted by  $L(X)$ , is the set of words which appear as subwords of elements of  $X$ . The set of words with a particular shape  $S$  which are in the language of  $X$  is denoted by  $L_S(X)$ .

**Definition 3.1.10.** For a  $\mathbb{Z}^d$ -subshift  $X$ , and for any finite shape  $S \subseteq \mathbb{Z}^d$ , the number of words in  $L_S(X)$  is denoted by  $H_S(X)$ , and the natural logarithm of this quantity is denoted by  $h_S(X)$ .

**Definition 3.1.11.** A  $\mathbb{Z}^d$ -**shift of finite type** is a  $\mathbb{Z}^d$ -subshift defined by specifying a finite collection of words on  $A$ , call this collection  $\mathcal{F}$ , and then taking  $X = \Omega_{\mathcal{F}}$  to be the set of elements of  $\Omega$  which do not contain any member of  $\mathcal{F}$  as subwords. Different collections  $\mathcal{F}$  could induce the same shift of finite type; for any fixed  $X$ , the **type** of  $X$  is the minimum nonnegative integer  $t$  such that for some  $\mathcal{F}$  consisting entirely of words with shape  $\Gamma_t$ ,  $X = \Omega_{\mathcal{F}}$ . A shift of type two is called a **Markov shift**.

**Definition 3.1.12.** A word  $v$  with shape  $S$  **forces** a word  $w$  with shape  $T$  in a shift of finite type  $X$  if any  $x \in X$  with  $x(S) = v$  also has  $x(T) = w$ .

**Definition 3.1.13.** A subshift  $X$  is **topologically transitive** if there exists  $x \in X$  for which  $\{\sigma_p(x)\}_{p \in \mathbb{Z}^d}$  is dense in  $X$ , i.e. if there is  $x \in X$  which contains every word in  $L(X)$  as a subword.

**Definition 3.1.14.** The **topological entropy** of a subshift  $X$ , denoted by  $h^{\text{top}}(X)$ , is defined by

$$h^{\text{top}}(X) = \lim_{j_1, j_2, \dots, j_d \rightarrow \infty} \frac{h_{R_{j_1, j_2, \dots, j_d}}(X)}{j_1 j_2 \cdots j_d}.$$

**Definition 3.1.15.** The **measure-theoretic entropy** of a subshift  $X$  with respect to a shift-invariant probability measure  $\mu$  on  $X$ , which is denoted by  $h_\mu(X)$ , is defined by

$$h_\mu(X) = \lim_{j_1, j_2, \dots, j_d \rightarrow \infty} \frac{1}{j_1 j_2 \cdots j_d} \sum_{w \in L_{R_{j_1, \dots, j_d}}(X)} f(\mu([w])),$$

where  $f(x) = -x \ln x$  for  $x > 0$  and  $f(0) = 0$ .

Both of these limits exist by subadditivity.

We now restate the question from the beginning of Chapter 3. Suppose that a  $\mathbb{Z}$ -subshift  $X$  on an alphabet  $A$  is given, and consider any  $n$ -letter word  $w \in L_{\Gamma_n}(X)$ . One can then define the subshift  $X_w$  as consisting of all elements of  $X$  which do not contain  $w$  as a subword. Since  $X_w \subseteq X$ ,  $h^{\text{top}}(X_w) \leq h^{\text{top}}(X)$ , but it is natural to wonder how much the entropy decreases by. For example, is it necessarily the case

that  $h^{top}(X) - h^{top}(X_w) \rightarrow 0$  as  $n \rightarrow \infty$ ? It is not hard to check that this is not always the case. For example, if  $X$  is minimal, then every  $x \in X$  contains every  $w \in L(X)$  as a subword. This means that for any word  $w \in L(X)$ ,  $X_w = \emptyset$ , and so  $h^{top}(X_w) = 0$ . This means that minimality is possibly too globally defined for our purposes. It is then natural to move to the completely local definition of a shift of finite type. In this setup, as was already mentioned in Chapter 1, there is the following result of Douglas Lind:

**Theorem 3.1.16.** ([L], p. 360, Theorem 3) *For any topologically transitive  $\mathbb{Z}$ -shift of finite type  $X = \Omega_{\mathcal{F}}$  with positive topological entropy  $h^{top}(X)$ , there exist constants  $C_X$ ,  $D_X$ , and  $N_X$  such that for any  $n > N_X$  and any word  $w \in L_{\Gamma_n}(X)$ , if we denote by  $X_w$  the shift of finite type  $\Omega_{\mathcal{F} \sqcup \{w\}}$ , then*

$$\frac{C_X}{e^{h^{top}(X)n}} < h^{top}(X) - h^{top}(X_w) < \frac{D_X}{e^{h^{top}(X)n}}.$$

Lind was apparently examining this question in order to prove some results about the automorphism group of a shift of finite type; he needed to know that  $h^{top}(X) - h^{top}(X_w) \rightarrow 0$  as the size  $n$  of the word  $w$  approaches infinity. ([BoyLR]). Some work on bounds of this type in the case where  $X = A^{\Omega}$  was also done by Guibas and Odylzko. ([GO1], [GO2])

We wish to prove a version of Theorem 3.1.16 for multidimensional shifts of finite type, i.e. shifts of finite type on  $\mathbb{Z}^d$  with  $d > 1$ . In the remainder of this section, we will discuss two relatively significant changes that we make in the statement of the multidimensional extension. This discussion will have the dual purposes of motivating

our result and illustrating some of the issues that arise in higher-dimensional symbolic dynamics. Our methods are similar in places to those used in [QT], where the effects on entropy of removing words from shifts of finite type are also examined.

The first problem in giving a multidimensional version of Theorem 3.1.16 comes in changing the exponents in the denominators. The most reasonable guess is that instead of using  $e^{h^{top}(X)n}$ , in the extension we should have  $e^{h^{top}(X)n^d}$ . However, we will show that this is a bit too much to hope for. Consider the following examples: take  $Y = \Omega$ .  $Y$  is clearly a shift of finite type, with the set of forbidden words  $\emptyset$ . For every  $j \in \mathbb{N}$ , we define a shift of finite type as follows: define the alphabet  $A^{(j)} = L_{\Gamma_j}(Y) = \{0, 1\}^{\Gamma_j}$ , and define  $f_j : Y \rightarrow (A^{(j)})^{\mathbb{Z}^d}$  by  $(f_j(x))(p) = x(p + \Gamma_j - (1, \dots, 1))$  for every  $x \in Y$  and  $p \in \mathbb{Z}^d$ . This map is usually called the higher block presentation of  $Y$ . (see [LM]) In other words, for every  $p \in \mathbb{Z}^d$ ,  $(f_j(x))(p)$  is defined to be the subword of  $x$  which lies in  $\Gamma_j + p - (1, \dots, 1)$ , or a copy of  $\Gamma_j$  whose least vertex in the usual lexicographic order on  $\mathbb{Z}^d$  is  $p$ . Figure 3.1 illustrates the action of  $f_2$  on a typical element of  $Y$  for  $d = 2$ : (the lower-leftmost entry of each array is at the point  $(1, \dots, 1) \in \mathbb{Z}^d$ )

In Figure 3.1, the horizontal and vertical lines separate letters of the alphabet. So, the left half shows a word with shape  $\Gamma_4$  whose letters are elements of  $A$ , namely zeroes and ones, and the right half shows a word with shape  $\Gamma_3$  on the alphabet  $A^{(2)}$ , which are themselves words with shape  $\Gamma_2$  whose letters are zeroes and ones.

With  $f_j$  so defined,  $f_j(Y)$  is a subshift with alphabet  $A^{(j)}$ . In fact it is easily checkable that  $f_j(Y)$  is a Markov shift;  $x \in (A^{(j)})^{\mathbb{Z}^d}$  is an element of  $f_j(Y)$  if and only if for

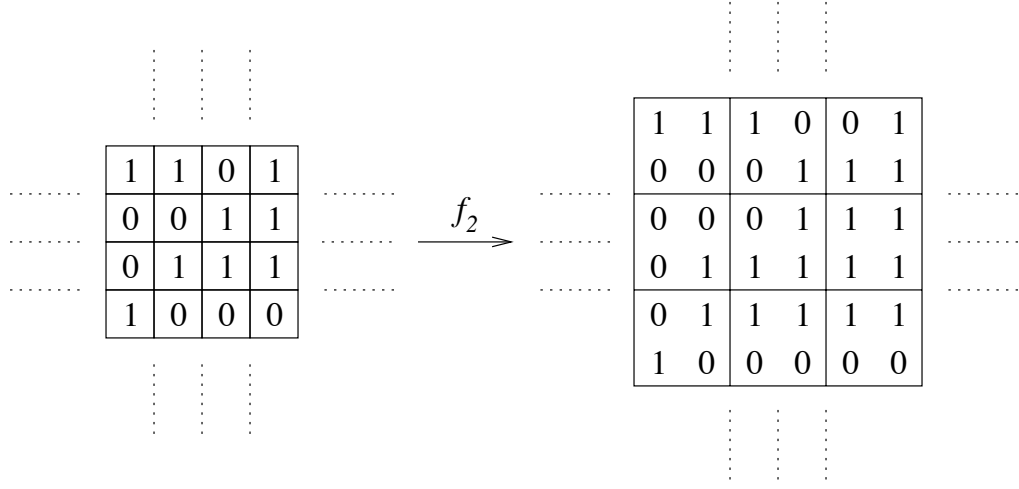


Figure 3.1:  $f_2$ 's action on a sample element of  $Y$

every  $p \in \mathbb{Z}^d$  and  $1 \leq i \leq d$ ,  $x(p)(R_{j,\dots,j,j-1,j,\dots,j} + e_i) = x(p + e_i)(R_{j,\dots,j,j-1,j,\dots,j})$ , where the  $j - 1$  is in the  $i$ th subindex in both cases. Since  $f_j$  is a shift-commuting bicontinuous bijection between  $Y$  and  $f_j(Y)$ ,  $f_j(Y)$  is topologically conjugate to  $Y$  for every  $j \in \mathbb{N}$ .

Since  $Y$  and  $f_j(Y)$  are such canonical examples of shifts of finite type (the full shift  $Y$  is in some sense the simplest shift of finite type, and every  $f_j(Y)$  is conjugate to it), any meaningful extension of Lind's result should certainly apply to them.  $f_j$  can be considered as a map between words just as easily as a function between subshifts: in the diagram above, if one disregards the dots surrounding the words, then  $f_2$  can be thought of as sending  $u \in L_{\Gamma_4}(Y)$  to  $f_2(u) \in L_{\Gamma_3}(f_2(Y))$ . For this reason, and since  $f_j$  is a topological conjugacy for all  $j$ , if we consider  $Y_w$  the shift of finite type resulting

from removing some word  $w$  with shape  $\Gamma_n$  from  $L(Y)$ , and  $(f_j(Y))_{f_j(w)}$  the shift of finite type resulting from removing  $f_j(w)$  from  $L(f_j(Y))$ , then  $Y_w$  and  $(f_j(Y))_{f_j(w)}$  are topologically conjugate as well, and therefore have the same topological entropy. In other words, the removal of a  $\Gamma_n$ -word from  $L(Y)$  results in the same drop in topological entropy as the removal of a  $\Gamma_{n-j+1}$ -word from  $L(f_j(Y))$ . Now, let us suppose that Lind's result can be extended by simply changing the  $n$  in the exponent to an  $n^d$ . This would result in something that looks like this:

**Possible Theorem 3.1.17.** *For any  $d > 1$  and any topologically transitive shift  $X = \Omega_{\mathcal{F}}$  of finite type with positive topological entropy  $h^{top}(X)$ , there exist constants  $C_X$ ,  $D_X$ , and  $N_X$  such that for any  $n > N_X$  and any word  $w \in L_{\Gamma_n}(X)$ , if we denote by  $X_w$  the shift of finite type  $\Omega_{\mathcal{F} \sqcup \{w\}}$ , then*

$$\frac{C_X}{e^{h^{top}(X)n^d}} < h^{top}(X) - h^{top}(X_w) < \frac{D_X}{e^{h^{top}(X)n^d}}.$$

We can now show that this possible extension cannot be true; suppose that there exist such constants  $C_Y$ ,  $D_Y$ , and  $N_Y$ , which satisfy Possible Theorem 3.1.17 for  $Y$ , and for any  $j > 1$ ,  $C_{f_j(Y)}$ ,  $D_{f_j(Y)}$ , and  $N_{f_j(Y)}$ , which satisfy Possible Theorem 3.1.17 for  $f_j(Y)$ . Then, for any word  $w \in L_{\Gamma_n}(Y)$ , where  $n > N_Y$ ,

$$\frac{C_Y}{e^{h^{top}(Y)n^d}} < h^{top}(Y) - h^{top}(Y_w) < \frac{D_Y}{e^{h^{top}(Y)n^d}}.$$

By the previous argument, this means that for any  $j$ ,

$$\frac{C_Y}{e^{h^{top}(Y)n^d}} < h^{top}(f_j(Y)) - h^{top}((f_j(Y))_{f_j(w)}) < \frac{D_Y}{e^{h^{top}(Y)n^d}}. \quad (3.1)$$

And by the definitions of  $C_{f_j(Y)}$ ,  $D_{f_j(Y)}$ , and  $N_{f_j(Y)}$ , and since  $f_j(w) \in L_{\Gamma_{n-j+1}}(f_j(Y))$ , as long as  $n - j + 1 > N_{f_j(Y)}$ , we know that

$$\frac{C_{f_j(Y)}}{e^{h^{top}(f_j(Y))(n-j+1)^d}} < h^{top}(f_j(Y)) - h^{top}((f_j(Y))_{f_j(w)}) < \frac{D_{f_j(Y)}}{e^{h^{top}(f_j(Y))(n-j+1)^d}}. \quad (3.2)$$

However, since  $d > 1$ , for large enough  $n$ , since  $h^{top}(f_j(Y)) = h^{top}(Y)$ ,  $e^{h^{top}(f_j(Y))n^d}$  grows much more quickly than  $e^{h^{top}(Y)(n-j+1)^d}$ . Therefore (3.1) and (3.2) contradict each other. The problem that must be addressed is that due to examples such as these, there must be some sort of a “tolerance range” around  $n$  in any meaningful multidimensional version of Theorem 3.1.16. In other words, we should amend our possible extension to the following:

**Possible Theorem 3.1.18.** *For any  $d > 1$  and any topologically transitive shift  $X = \Omega_{\mathcal{F}}$  of finite type with positive topological entropy  $h^{top}(X)$ , there exist constants  $C_X, D_X > 0$ ,  $A_X, B_X$ , and  $N_X \in \mathbb{N}$  such that for any  $n > N_X$  and any word  $w \in L_{\Gamma_n}(X)$ , if we denote by  $X_w$  the shift of finite type  $\Omega_{\mathcal{F} \sqcup \{w\}}$ , then*

$$\frac{C_X}{e^{h^{top}(X)(n+A_X)^d}} < h^{top}(X) - h^{top}(X_w) < \frac{D_X}{e^{h^{top}(X)(n+B_X)^d}}.$$

Since  $j$  could be arbitrarily large in the definition of  $f_j(Y)$ , it must be the case that  $A_X$  and  $B_X$  depend on  $X$  rather than being absolute constants. In fact, such constants  $A_X$  and  $B_X$  could be thought of as being present in Theorem 3.1.16 as well, but hidden inside the constants  $C_X$  and  $D_X$ . This new version is still not true, however.

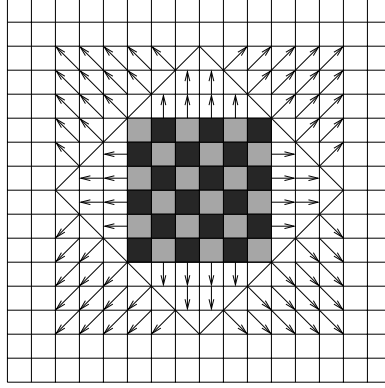


Figure 3.2: A portion of a sample element of  $Z$

This time, the problem is not with the conclusion, but the hypotheses. Topological transitivity is not nearly a strong enough hypothesis in several dimensions to get the type of result that we are after, and in fact we will see that topological mixing is not nearly good enough either. To show this, we consider a subshift called the *checkerboard island shift*, which we call  $Z$ , defined by Quas and Şahin in [QS].  $Z$  is a  $\mathbb{Z}^2$ -Markov shift, on an alphabet  $A$  with thirteen letters.  $Z$  is also topologically mixing; for a proof see [QS].

All letters of  $A$  and elements of  $L_{\Gamma_2}(X)$  are displayed in Figure 3.2. In other words, a two-by-two word of symbols is in  $L(Z)$  if and only if it appears as a subword of the word in Figure 3.2. Consider the word with shape  $\Gamma_8$  in Figure 3.3, which we call  $a_4$ .

It will not be shown here (details are in [QS]), but an occurrence of  $a_4$  in any element of  $Z$  forces the arrows and diagonal lines surrounding  $a_4$  in Figure 3.2, along with a boundary of thickness one filled with empty spaces. In other words, wherever  $a_4$ ,



which has shape  $\Gamma_8$ , appears, it forces a word, which we call  $b_4$ , of shape  $\Gamma_{14}$ .  $b_4$  appears in Figure 3.4.

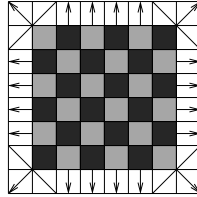


Figure 3.3:  $a_4$

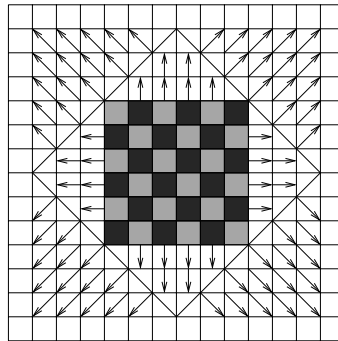


Figure 3.4:  $b_4$

In fact, for every  $n > 4$ , there is an analogous word  $a_n$  (a checkerboard of size  $\Gamma_{2n-2}$  with a boundary filled with arrows and diagonal lines in an analogous way to that

of  $a_4$ ) which forces a word  $b_n$  of size  $\Gamma_{4n-2}$  around it. We claim that for any  $n$ , the shift  $Z_{a_n}$  obtained by removing  $a_n$  from  $L(Z)$  and the shift  $Z_{b_n}$  obtained by removing  $b_n$  from  $L(Z)$  are equal. Since  $a_n$  is a subword of  $b_n$ , it is obvious that  $Z_{a_n} \subseteq Z_{b_n}$ . Suppose that  $Z_{b_n} \not\subseteq Z_{a_n}$ . Then there exists  $x \in Z$  which has an occurrence of  $a_n$ , but no occurrence of  $b_n$ , which, as has already been stated, is not possible. So,  $Z_{a_n} = Z_{b_n}$ . This implies that  $h^{top}(Z) - h^{top}(Z_{a_n}) = h^{top}(Z) - h^{top}(Z_{b_n})$ . However, since  $a_n \in L_{\Gamma_{2n}}(Z)$  and  $b_n \in L_{\Gamma_{4n-2}}(Z)$ , we have a contradiction to Possible Theorem 3.1.18 for the same reason as we did before. Since this is true for all  $n$ , if we wish to adjust the conditions of the conclusion again, we will have to change our “tolerance range” from  $[n + A_X, n + B_X]$  to something like  $[nA_X, nB_X]$ , which would weaken the result considerably. To obtain a conclusion of the strength we want, we need a stronger mixing property than topological mixing. The notion which we use is that of **strong irreducibility**.

**Definition 3.1.19.** A  $\mathbb{Z}^d$ -subshift  $X$  is **strongly irreducible with uniform filling length**  $R$  if for any two subsets  $S, T$  of  $\mathbb{Z}^d$ , and for all  $p \in \mathbb{Z}^d$  such that  $d(S, T + p) > R$  (that is, for all  $q \in S$  and  $r \in T + p$ ,  $d(q, r) > R$ ), and for any elements  $x, x' \in X$ , there exists  $y \in X$  such that  $y(S) = x(S)$  and  $y(T + p) = x'(T + p)$ .

For the sake of comparing, we also include a definition of topological mixing whose formulation looks similar to that of strong irreducibility:

**Definition 3.1.20.** A  $\mathbb{Z}^d$ -subshift  $X$  is **topologically mixing** if for any two subsets  $S$  and  $T$  of  $\mathbb{Z}^d$ , there exists  $R_{S,T}$  such that for all  $p \in \mathbb{Z}^d$  with the property that

$d(S, T + p) > R$ , and for any elements  $x, x' \in X$ , there exists  $y \in X$  such that  $y(S) = x(S)$  and  $y(T + p) = x'(T + p)$ .

The difference in the two definitions is subtle but important: for strong irreducibility, the distance  $R$  that must be present between two shapes in order to make interpolation possible is independent of the shapes involved, whereas in topological mixing, this distance can depend on the shapes. This distinction is irrelevant in one dimension, since strong irreducibility and topological mixing are the same concept for  $\mathbb{Z}$ -subshifts. (To see this, note that for a  $\mathbb{Z}$ -shift of finite type, a distance sufficient for interpolating between two letters of the alphabet is sufficient to interpolate between any two words.) However, strong irreducibility is a strictly stronger notion in more than one dimension. Pertinent for our purposes is the fact that if  $x(S) = w$  for some  $x$  in a strongly irreducible shift of finite type  $X$ , then the occurrence of  $w$  at  $S$  can only “force” letters to appear in the region of  $\mathbb{Z}^d$  within distance  $R$  of  $S$ . This is because for any  $v \in \mathbb{Z}^d$  with  $d(v, S) > R$  and any  $a \in A$ , by the definition of strong irreducibility there is some  $y \in X$  with  $y(S) = w$  and  $y(v) = a$ . For this reason, the checkerboard island system  $Z$  is topologically mixing but not strongly irreducible. Armed with this stronger hypothesis, we can now refine Possible Theorem 3.1.18 a bit more:

**Possible Theorem 3.1.21.** *For any  $d > 1$  and any strongly irreducible shift  $X = \Omega_{\mathcal{F}}$  of finite type with uniform filling length  $R$  and positive topological entropy  $h^{\text{top}}(X)$ , there exist constants  $C_X, D_X > 0$ ,  $A_X, B_X$ , and  $N_X \in \mathbb{N}$  such that for any  $n > N_X$  and any word  $w \in L_{\Gamma_n}(X)$ , if we denote by  $X_w$  the shift of finite type  $\Omega_{\mathcal{F} \sqcup \{w\}}$ , then*

$$\frac{C_X}{e^{h^{top}(X)(n+A_X)^d}} < h^{top}(X) - h^{top}(X_w) < \frac{D_X}{e^{h^{top}(X)(n+B_X)^d}}.$$

In fact the above is now provable with correct choices for the pertinent constants, and so we state our main result:

**Theorem 3.1.22.** *For any  $d > 1$  and any strongly irreducible shift  $X = \Omega_{\mathcal{F}}$  of finite type with uniform filling length  $R$  and positive topological entropy  $h^{top}(X)$ , there exist constants  $N_X \in \mathbb{N}$  and  $D_X \in \mathbb{R}$  such that for any  $n > N_X$  and any word  $w \in L_{\Gamma_n}(X)$ , if we denote by  $X_w$  the shift of finite type  $\Omega_{\mathcal{F} \sqcup \{w\}}$ , then*

$$\frac{1}{e^{h^{top}(X)(n+44R+70)^d}} < h^{top}(X) - h^{top}(X_w) < \frac{D_X}{e^{h^{top}(X)(n-2R)^d}}.$$

Before beginning with the preliminaries of a proof, we make an observation: our example shifts of finite type  $f_j(Y)$  used above to show the necessity of the constants  $A_X$  and  $B_X$  in this statement did not actually show that  $A_X$  and  $B_X$  are not the same; it is possible that  $A_{f_j(Y)} = B_{f_j(Y)} = j$  for all  $j$ . However, we may now introduce another example to show that in fact  $A_X$  and  $B_X$  must be distinct for certain  $X$ .

Consider, again for any  $j > 1$ , and any dimension  $d$ , the shift of finite type  $Z_j$  on the alphabet  $\{0, 1\}$  defined by the set of forbidden words  $\mathcal{F}_j = \{w \in \{0, 1\}^{\Gamma_j} : w \text{ has at least two ones}\}$ . In other words,  $Z_j$  consists of all infinite words of zeroes and ones on  $\mathbb{Z}^d$  such that any two ones are a  $d$ -distance of at least  $j$  from each other. By definition,  $Z_j$  is clearly a shift of finite type. We also claim that it is strongly irreducible with  $R = j - 1$ : consider any words  $x \in L_S(Z_j)$  and  $x' \in L_T(Z_j)$  such that  $d(S, T) > j - 1$ . Take  $y \in \{0, 1\}^{\mathbb{Z}^d}$  defined by  $y(S) = x$ ,  $y(T) = x'$ , and  $y(p) = 0$

for any  $p \notin S \cup T$ . We claim that  $y \in Z_j$ : suppose, on the contrary, that  $y \notin Z_j$ . Then there exist  $a, b \in \mathbb{Z}^d$  such that  $d(a, b) \leq j$ , and  $y(a) = y(b) = 1$ . If  $a, b \in S$ , then  $x$  contains two ones a  $d$ -distance less than  $j$  apart and cannot be in  $L(X)$ , a contradiction. We arrive at a similar contradiction if  $a, b \in T$ . Clearly we cannot have  $a \in S$  and  $b \in T$  since  $d(S, T) > j - 1$ , and similarly it cannot be the case that  $a \in T$  and  $b \in S$ . But since the only ones in  $y$  are in  $S$  or  $T$ , we have exhausted all possible cases. Therefore,  $y \in Z_j$ . By definition,  $Z_j$  is strongly irreducible with  $R = j - 1$ . We claim that in the language of Possible Theorem 3.1.21, it must be the case that  $A_{Z_j} \neq B_{Z_j}$ , and in fact that  $A_{Z_j} - B_{Z_j} \geq 2j - 2 = 2R$ .

To prove this, we first make a definition: for any  $n > 0$ , we define the word  $a_n$  with shape  $\Gamma_{nj+1}$  by  $a_n(p) = 1$  if and only if all coordinates of  $p$  are equal to one modulo  $j$ .  $a_n \in L(Z_j)$  since it contains no two ones a  $d$ -distance of less than  $j$  apart. Let us also for any  $n > 0$  define the word  $b_n$  with shape  $\Gamma_{(n+2)j-1}$  where  $b_n$  has  $a_n$  as a subword occupying the central  $\Gamma_{nj+1}$ , and has zeroes on all of  $\Gamma_{(n+2)j-1}^{(j-1)}$ . By the definition of  $Z_j$ , any occurrence of  $a_n$  in an element of  $Z_j$  forces an occurrence of  $b_n$  containing it. Therefore, as argued for  $f_j(Y)$ ,  $(Z_j)_{a_n} = (Z_j)_{b_n}$  for any  $n$ . However, the size of the shape of  $b_n$  is  $2j - 2$  bigger than that of  $a_n$  for any  $n$ , and so for any  $A_{Z_j}$ ,  $B_{Z_j}$ ,  $C_{Z_j}$ , and  $D_{Z_j}$  which satisfy Possible Theorem 3.1.21, it must be the case that  $A_{Z_j} - B_{Z_j} \geq 2j - 2 = 2R$  for any  $j$ . Although this does not show that the bounds in Theorem 3.1.22 are optimal, it does show that  $A_X - B_X$  must be at least linear in  $R$ , and so the value of  $46R + 70$  for  $A_X - B_X$  attained in Theorem 3.1.22 can be improved by at most additive and multiplicative constants.

Before beginning to prove our main result, we now briefly outline the content of this paper. The main idea in the proof of Theorem 3.1.22 is to use strong irreducibility to find a word  $w' \neq w$  where  $w$  and  $w'$  agree on  $\Gamma_n^{(t)}$ , and then define a map which replaces all occurrences of  $w$  by  $w'$  in large words  $v \in L_{\Gamma_k}(X)$ . In order to know how many-to-one this map is, it is important to know how many times  $w$  and  $w'$  appear in a “typical”  $v \in L_{\Gamma_k}(X)$ . In Section 3.2, we use ergodic measures of maximal entropy to show that for fixed  $n$ , any  $w \in L_{\Gamma_n}(X)$ , and large  $k$ , there are “many” words  $v$  in  $L_{\Gamma_k}(X)$  for which we have good bounds from above and below on the number of occurrences of  $w$  in  $v$ . Namely, for a large set of  $v$ , the number of occurrences of  $w$  is roughly  $k^d \tilde{\mu}([w])$ , where  $\tilde{\mu}$  is an ergodic measure of maximal entropy on  $X$ . We then use a result of Burton and Steif ([BuS1], [BuS2]) to give upper and lower bounds on  $\tilde{\mu}([w])$  of the type  $e^{-h^{top}(X)(n+A)^d}$ .

There is a slight problem though; it is entirely possible that a replacement of  $w$  by  $w'$  could create a new occurrence of  $w$ , which could make this replacement map extremely unwieldy and hard to deal with. What we would like is a word  $w'$  such that replacing  $w$  by  $w'$  can never create a new occurrence of  $w$ . Unfortunately, this is impossible for some choices of  $w$ . In Section 3.3, we prove a slightly weaker fact that is still sufficient for our purposes; for any  $w \in L_{\Gamma_n}(X)$ , there exists a subword  $w' \in L_{\Gamma_m}(X)$  with  $m$  not much smaller than  $n$  for which there exists  $w'' \neq w'$  where  $w'$  and  $w''$  agree on  $\Gamma_m^{(t)}$  and where a replacement of  $w'$  by  $w''$  can never create a new occurrence of  $w'$ .

Section 3.4 contains the proof of Theorem 3.1.22. The proof is done by considering

maps between  $L(X)$  and  $L(X_w)$  which either introduce or destroy occurrences of  $w$  (which is done by using the result of Section 3.3 to perform various replacements), and estimating the sizes of images or preimages under these maps. We can make certain helpful assumptions about the words in which these replacements take place (such as the approximate number of occurrences of  $w$ ) by using the results of Section 3.2.

In Section 3.5, we deal with a simplified version of our main result where the word  $w$  in question is already chosen to have a nice property. (namely, that there exists  $w'$  such that replacing  $w$  by  $w'$  can neither accidentally create new  $w$  or  $w'$  or destroy existing  $w$  or  $w'$ ) In this case, we prove that  $(\ln 2)\tilde{\mu}_w([w']) \leq h^{top}(X) - h^{top}(X_w) \leq 2(\ln 2)\tilde{\mu}([w])$  for any  $\tilde{\mu}$  an ergodic measure of maximal entropy on  $X$  and  $\tilde{\mu}_w$  an ergodic measure of maximal entropy on  $X_w$ . We then show that it is not possible to prove a lower bound on  $h^{top}(X) - h^{top}(X_w)$  of the form  $C\tilde{\mu}([w])$  for constant  $C$ .

In Section 3.6, we prove a corollary using our methods: for any alphabet  $A$ , and given any set of words  $w_1, \dots, w_m$  where the size of  $w_{i+1}$  is much greater than that of  $w_i$  for  $1 \leq i < m$ ,  $Y_{w_1, \dots, w_m} \neq \emptyset$ , where  $Y = \Omega$  is the full shift. This can be seen to be related to some of the so-called undecidability questions in  $\mathbb{Z}^d$  shifts of finite type; given a set  $\mathcal{F}$  of words, it is algorithmically undecidable whether or not  $\Omega_{\mathcal{F}}$  is nonempty. The corollary described gives a condition under which this question is easily answered.

Finally, in Section 3.7, we give some open questions that arise in the course of proving our results.

### 3.2 Some measure-theoretic preliminaries

We will now move towards proving Theorem 3.1.22. We first need a lemma about strongly irreducible systems:

**Lemma 3.2.1.** *For any  $\mathbb{Z}^d$ -subshift  $X$  which is strongly irreducible with uniform filling length  $R$  and for any rectangular prism  $R_{j_1, j_2, \dots, j_d}$ ,  $e^{h^{top}(X)j_1 j_2 \dots j_d} \leq H_{R_{j_1, j_2, \dots, j_d}}(X) \leq e^{h^{top}(X)(j_1+R)(j_2+R) \dots (j_d+R)}$ .*

**Proof:** We first prove the first inequality: take any rectangular prism  $R_{j_1, \dots, j_d}$ . For any  $k \in \mathbb{N}$ , since  $R_{kj_1, \dots, kj_d}$  can be partitioned into  $t^d$  disjoint copies of  $R_{j_1, \dots, j_d}$ ,  $H_{R_{kj_1, \dots, kj_d}}(X) \leq H_{R_{j_1, \dots, j_d}}(X)^{k^d}$ . Take natural logarithms of both sides and divide by  $k^d j_1 \dots j_d$ :

$$\frac{h_{R_{kj_1, \dots, kj_d}}(X)}{k^d j_1 \dots j_d} \leq \frac{h_{R_{j_1, \dots, j_d}}(X)}{j_1 \dots j_d},$$

and by letting  $k$  approach infinity, we see that  $h^{top}(X) \leq \frac{h_{R_{j_1, \dots, j_d}}(X)}{j_1 \dots j_d}$ , which implies that  $e^{h^{top}(X)j_1 \dots j_d} \leq H_{R_{j_1, \dots, j_d}}(X)$ .

We prove the second inequality in almost the same way: again consider any rectangular prism  $R_{j_1, \dots, j_d}$ . For any positive integer  $k$ ,  $R_{k(j_1+R), k(j_2+R), \dots, k(j_d+R)}$  can be broken into  $k^d$  copies of  $R_{j_1+R, \dots, j_d+R}$ , and we can choose copies of  $R_{j_1, \dots, j_d}$  inside each  $R_{j_1+R, \dots, j_d+R}$  so that the  $d$ -distance between any pair of these  $R_{j_1, \dots, j_d}$  is greater than  $R$ . This is illustrated in Figure 3.5.

(A note on figures: all figures in this paper are drawn with  $d = 2$ , but this is in some sense without loss of generality; all of the ideas that we use can be illustrated via



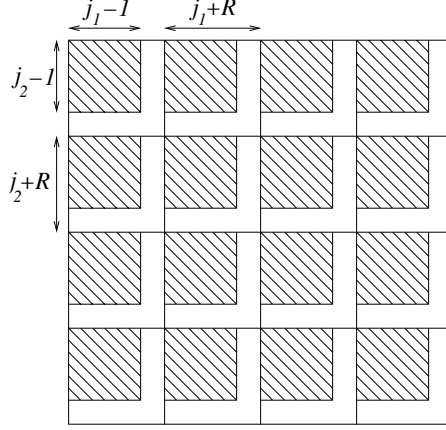


Figure 3.5:  $R_{k(j_1+R), k(j_2+R), \dots, k(j_d+R)}$

two-dimensional pictures. All constructions and descriptions are carried out in full generality.)

By strong irreducibility, for any way of filling these  $k^d$  copies of  $R_{j_1, \dots, j_d}$  with words in  $L(X)$ , the remainder can be filled to make a word in  $L_{R_{k(j_1+R), k(j_2+R), \dots, k(j_d+R)}}(X)$ . This implies that  $H_{R_{j_1, \dots, j_d}}(X)^{k^d} \leq H_{R_{k(j_1+R), k(j_2+R), \dots, k(j_d+R)}}(X)$ . Take natural logarithms of both sides and divide by  $k^d(j_1 + R)(j_2 + R) \cdots (j_d + R)$ :

$$\frac{h_{R_{j_1, \dots, j_d}}(X)}{(j_1 + R) \cdots (j_d + R)} \leq \frac{h_{R_{k(j_1+R), \dots, k(j_d+R)}}(X)}{k^d(j_1 + R) \cdots (j_d + R)},$$

and by letting  $k$  approach infinity, we see that  $\frac{h_{R_{j_1, \dots, j_d}}(X)}{(j_1 + R) \cdots (j_d + R)} \leq h^{top}(X)$ , or that  $H_{R_{j_1, \dots, j_d}}(X) \leq e^{h^{top}(X)(j_1+R)(j_2+R) \cdots (j_d+R)}$ .

□

To prove Theorem 3.1.22, we will be repeatedly using the concept of a measure of maximal entropy.

**Definition 3.2.2.** *A shift-invariant probability measure  $\mu$  on a subshift  $X$  is called a **measure of maximal entropy** if  $h_\mu(X) = h^{\text{top}}(X)$ .*

Such measures are said to have maximal entropy because of the following Variational Principle:

**Theorem 3.2.3.** *For any  $\mathbb{Z}^d$ -subshift  $X$ ,  $h^{\text{top}}(X) = \sup h_\mu(X)$ , where  $\mu$  ranges over all shift-invariant Borel probability measures on  $X$ . This supremum is achieved for some  $\mu$ .*

See [M] for a proof. A useful observation is that since the extreme points of the set of measures of maximal entropy are ergodic, and since the set of measures of maximal entropy is nonempty for subshifts, any subshift also has an ergodic measure of maximal entropy. (See [Wal] for details.) We need the following proposition:

**Proposition 3.2.4.** ([BuS2], p. 281, Proposition 1.20) *Let  $\mu$  be a measure of maximal entropy for a  $\mathbb{Z}^d$ -shift  $X$  of finite type  $t$ . Then for any shape  $U \subseteq \mathbb{Z}^d$ , the conditional distribution of  $\mu$  on  $U$  given a fixed word  $w \in L_{(U^c)^{(t)}}(X)$  is uniform over all words  $x \in L_U(X)$  such that the word  $y$  with shape  $U \cup (U^c)^{(t)}$  defined by  $y(U) = x$  and  $y((U^c)^{(t)}) = w$  is in  $L(X)$ .*

We have the following convenient restatement:

**Proposition 3.2.5.** *For any  $\mathbb{Z}^d$ -shift  $X$  of finite type  $t$ , and for any measure  $\mu$  of maximal entropy on  $X$ , and for any word  $w \in L_S(X)$ , and for any shape  $S^{(t)} \subseteq T \subseteq$*

$S$ ,  $\mu([w]) = \frac{\mu([w(T)])}{|L_S(X) \cap [w(T)]|}$ . In other words, any two words with the same shape  $S$  which agree on  $T$  have the same measure under  $\mu$ .

**Proof:** Consider any two words  $w, w' \in L_S(X)$  with  $w(T) = w'(T)$ . We claim that  $((S \setminus S^{(t)})^c)^{(t)} \subseteq S^{(t)}$ . Consider  $p \in ((S \setminus S^{(t)})^c)^{(t)}$ . By definition,  $p \in (S \setminus S^{(t)})^c$  and there exists  $q \in S \setminus S^{(t)}$  such that  $d(p, q) \leq t$ . Suppose that  $p \in S^c$ . Then, since  $d(p, q) \leq t$  and since  $q \in S$ , by definition,  $q \in S^{(t)}$ , a contradiction. Therefore,  $p \in S$ . Since  $p \in (S \setminus S^{(t)})^c$ , this implies that  $p \in S^{(t)}$ . Therefore,  $((S \setminus S^{(t)})^c)^{(t)} \subseteq S^{(t)}$ . We define  $x = w(S \setminus S^{(t)})$  and  $x' = w'(S \setminus S^{(t)})$ . Since  $S^{(t)} \subseteq T$ , and since  $((S \setminus S^{(t)})^c)^{(t)} \subseteq S^{(t)}$ ,  $w(((S \setminus S^{(t)})^c)^{(t)}) = w'(((S \setminus S^{(t)})^c)^{(t)})$ . Therefore, if we define  $z = w(((S \setminus S^{(t)})^c)^{(t)})$ , then since either  $x$  or  $x'$  could fill  $S \setminus S^{(t)}$  while  $z$  fills  $((S \setminus S^{(t)})^c)^{(t)}$ , we may take  $U = S \setminus S^{(t)}$  and use Proposition 3.2.4 to see that  $\mu(w) = \mu(w')$ .

□

**Lemma 3.2.6.** *For any strongly irreducible  $\mathbb{Z}^d$ -shift  $X$  of finite type  $t$  with uniform filling length  $R$ , and for any measure  $\mu$  of maximal entropy on  $X$ , and for any word  $u$  with shape  $S$  a subset of  $\Gamma_M$ ,*

$$\frac{1}{H_{(S \sqcup (S^c)^{(t+R)}) \setminus (S \sqcup (S^c)^{(t+R)})^{(t)}}(X)} \leq \mu([u]) \leq \frac{1}{H_{S \setminus S^{(R)}}(X)}.$$

**Proof:** Before we begin the proof, we recall that for any positive integer  $j$ ,  $S^{(j)}$  is the boundary of thickness  $j$  of  $S$ , and that  $(S^c)^{(j)}$  is the boundary of thickness  $j$  of

the complement of  $S$ , i.e. the set of all points  $p \notin S$  for which there exists  $q \in S$  with  $d(p, q) \leq j$ .

Fix  $u \in L_S(X)$ . We begin by bounding  $\mu([u])$  from below. Choose any word  $v \in L_{(S \sqcup (S^c)^{(t+R)})^{(t)}}(X)$ . We claim that  $d(S, (S \sqcup (S^c)^{(t+R)})^{(t)}) > R$ . Consider any  $p \in S$  and  $q \in (S \sqcup (S^c)^{(t+R)})^{(t)}$ . By definition, there exists  $r \notin S \sqcup (S^c)^{t+R}$  such that  $d(q, r) \leq t$ . Also by definition,  $d(p, r) > R + t$ . Therefore, by the triangle inequality,  $d(p, q) > R$ , as claimed. This means that by strong irreducibility, there exists a word  $w_v \in L_{S \sqcup (S^c)^{(t+R)}}(X)$  such that  $w_v(S) = u$  and  $w_v((S \sqcup (S^c)^{(t+R)})^{(t)}) = v$ . By Proposition 3.2.5,  $\mu([w_v]) = \frac{\mu([v])}{|L_{S \sqcup (S^c)^{(t+R)}}(X) \cap [v]|} \geq \frac{\mu([v])}{H_{(S \sqcup (S^c)^{(t+R)}) \setminus (S \sqcup (S^c)^{(t+R)})^{(t)}}(X)}}$ . The inequality comes from the fact that the number of words in  $L(X)$  with shape  $S \sqcup (S^c)^{(t+R)}$  with the fixed word  $v$  on  $(S \sqcup (S^c)^{(t+R)})^{(t)}$  is clearly less than or equal to the number of words in  $L(X)$  with shape  $(S \sqcup (S^c)^{(t+R)}) \setminus (S \sqcup (S^c)^{(t+R)})^{(t)}$ . For clarity, we rewrite:

$$\mu([w_v]) = \frac{\mu([v])}{H_{(S \sqcup (S^c)^{(t+R)}) \setminus (S \sqcup (S^c)^{(t+R)})^{(t)}}(X)}}. \quad (3.3)$$

If we sum (3.3) over all possible choices for  $v$ , then we get

$$\sum_{v \in L_{(S \sqcup (S^c)^{(t+R)})^{(t)}}(X)} \mu([w_v]) = \frac{\sum_{v \in L_{(S \sqcup (S^c)^{(t+R)})^{(t)}}(X)} \mu([v])}{H_{(S \sqcup (S^c)^{(t+R)}) \setminus (S \sqcup (S^c)^{(t+R)})^{(t)}}(X)}}. \quad (3.4)$$

Note that all  $w_v$  are distinct, and for all  $w_v$ ,  $w_v(S) = u$ . Therefore,

$$\bigcup_{v \in L_{(S \sqcup (S^c)^{(t+R)})^{(t)}}(X)} [w_v] \subseteq [u],$$

and so  $\sum_{v \in L_{(S \sqcup (S^c)^{(t+R))^{(t)}}(X)} \mu([w_v]) \leq \mu([u])$ . Since  $\bigcup_{v \in L_{(S \sqcup (S^c)^{(t+R))^{(t)}}(X)} [v] = X$ ,  $\sum_{v \in L_{(S \sqcup (S^c)^{(t+R))^{(t)}}(X)} \mu([v]) = 1$ . By combining these facts with (3.4), we see that

$$\mu([u]) \geq \frac{1}{H_{(S \sqcup (S^c)^{(t+R)) \setminus (S \sqcup (S^c)^{(t+R))^{(t)}}(X)}.$$

We now bound  $\mu([u])$  from above. Choose any word  $v \in L_{S \sqcup (S^c)^{(t)}}(X)$  with  $v(S) = u$ . We wish to use Proposition 3.2.5 with  $(S^c)^{(t)}$  as our  $T$ . To do so, we need to know that  $(S \sqcup (S^c)^{(t)})^{(t)} \subseteq (S^c)^{(t)} \subseteq S \sqcup (S^c)^{(t)}$ . The second containment is trivial. To prove the first, suppose that  $p \in (S \sqcup (S^c)^{(t)})^{(t)}$ . By definition, there then exists  $q \notin S \sqcup (S^c)^{(t)}$  such that  $d(p, q) \leq t$ . Since  $q \notin S \sqcup (S^c)^{(t)}$ , in particular  $q \notin S$ , or  $q \in S^c$ . Therefore, since  $d(p, q) \leq t$ ,  $p \in (S^c)^{(t)}$  by definition. We now apply Proposition 3.2.5:

$$\mu([v]) = \frac{\mu([v((S^c)^{(t)})])}{|L_{S \sqcup (S^c)^{(t)}}(X) \cap [v((S^c)^{(t)})]|} \quad (3.5)$$

We claim that  $d(S \setminus S^{(R)}, (S^c)^{(t)}) > R$ ; choose any  $p \in S \setminus S^{(R)}$  and  $q \in (S^c)^{(t)}$ . Clearly  $q \in S^c$ . If  $d(p, q) \leq R$ , then  $p \in S^{(R)}$ , a contradiction. Thus,  $d(p, q) > R$ . This implies that for any  $v$ , and for any  $v' \in L_{S \setminus S^{(R)}}(X)$ , there exists  $y \in L_{S \sqcup (S^c)^{(t)}}(X)$  with  $y((S^c)^{(t)}) = v((S^c)^{(t)})$  and  $y(S \setminus S^{(R)}) = v'$ . Therefore,  $|L_{S \sqcup (S^c)^{(t)}}(X) \cap [v((S^c)^{(t)})]| \geq |L_{S \setminus S^{(R)}}(X)|$ . By using this fact and summing (3.5) over all possible  $v$ , we get

$$\sum_{v \in L_{S \sqcup (S^c)^{(t)}}(X) \cap [u]} \mu([v]) \leq \frac{\sum_{v \in L_{S \sqcup (S^c)^{(t)}}(X) \cap [u]} \mu([v((S^c)^{(t)})])}{|L_{S \setminus S^{(R)}}(X)|}. \quad (3.6)$$

Note that since all  $v$  are distinct,  $\bigcup_{v \in L_{S \sqcup (S^c)^{(t)}}(X) \cap [u]} [v] = [u]$ , and so

$\sum_{v \in L_{S \sqcup (S^c)^{(t)}}(X) \cap [u]} \mu([v]) = \mu([u])$ . Also note that since all  $v$  are distinct, but all have  $v(S) = U$ , it must be the case that all  $v((S^c)^{(t)})$  are distinct, and so

$$\sum_{v \in L_{S \sqcup (S^c)^{(t)}}(X) \cap [u]} \mu([v((S^c)^{(t)})]) \leq 1.$$

Combining these facts with (3.6), we see that

$$\mu([u]) \leq \frac{1}{H_{S \setminus S^{(R)}}(X)},$$

which, along with the lower bound on  $\mu([u])$  already achieved, completes the proof. □

To avoid confusion, from now on we use  $\tilde{\mu}$  to denote an ergodic measure of maximal entropy on a subshift  $X$ , and  $\mu$  to denote a shift-invariant probability measure which may or may not be of maximal entropy.

**Lemma 3.2.7.** *For any  $\mathbb{Z}^d$ -subshift  $X$ , and for any shift-invariant ergodic probability measure  $\mu$  on  $X$ , and for any finite set of words  $u_1 \in L_{S_1}(X)$ ,  $u_2 \in L_{S_2}(X), \dots, u_j \in L_{S_j}(X)$  and  $\epsilon > 0$ , if we denote by  $A_{k,\epsilon,\mu,u_1,\dots,u_j}(X)$  the set of words in  $L_{\Gamma_k}(X)$  which have between  $k^d(\mu([u_i]) - \epsilon)$  and  $k^d(\mu([u_i]) + \epsilon)$  occurrences of  $u_i$  for all  $1 \leq i \leq j$ , then  $\liminf_{k \rightarrow \infty} \frac{\ln |A_{k,\epsilon,\mu,u_1,\dots,u_j}(X)|}{k^d} \geq h_\mu(X)$ .*

**Proof:** (For brevity, we use the shorthand notation  $\mu(A_{k,\epsilon,\mu,u_1,\dots,u_j}(X))$  for the measure of the union of all cylinder sets corresponding to words in  $A_{k,\epsilon,\mu,u_1,\dots,u_j}(X)$ .) Fix any  $\epsilon > 0$ . Firstly, we notice that it is a simple consequence of Birkhoff's ergodic theorem that  $\lim_{k \rightarrow \infty} \mu(A_{k,\epsilon,\mu,u_1,\dots,u_j}(X)) = 1$ : for each  $1 \leq i \leq j$ ,

$$\lim_{k \rightarrow \infty} \frac{1}{k^d} \sum_{p \in \Gamma_k} \chi_{[u_i]}(\sigma^p(x)) = \mu([u_i]) \text{ for almost every } x \in X,$$

where  $\sigma$  represents the  $\mathbb{Z}^d$  action by shifts on  $X$ . Convergence almost everywhere implies convergence in measure, meaning that

$$\lim_{k \rightarrow \infty} \mu \left( \left\{ x : \frac{1}{k^d} \sum_{p \in \Gamma_k} \chi_{[u_i]}(\sigma^p(x)) \in (\mu([u_i]) - \epsilon, \mu([u_i]) + \epsilon) \text{ for } 1 \leq i \leq j \right\} \right) = 1. \quad (3.7)$$

Denote by  $A'_{k,\epsilon,\mu,u_1,\dots,u_j}(X)$  the set whose measure is taken on the lefthand side of (3.7). Then whether or not  $x \in X$  lies in  $A'_{k,\epsilon,\mu,u_1,\dots,u_j}(X)$  depends only on its values on  $\Gamma_{k+M}$ , where  $M$  is the minimum integer such that all  $S_i$  are subsets of some copy of  $\Gamma_M$ . Therefore,  $A'_{k,\epsilon,\mu,u_1,\dots,u_j}(X)$  is a union of cylinder sets of words in  $L_{\Gamma_{k+M}}(X)$ . It is not hard to see that for large  $k$ , any word whose cylinder set is a subset of  $A'_{k,\frac{\epsilon}{2},\mu,u_1,\dots,u_j}(X)$  is an element of  $A_{k+M,\epsilon,\mu,u_1,\dots,u_j}(X)$ . Combining this with the fact that (3.7) is true for any  $\epsilon > 0$ , we see that  $\lim_{k \rightarrow \infty} \mu(A_{k,\epsilon,\mu,u_1,\dots,u_j}(X)) = 1$ .

We now write out the formula for the measure-theoretic entropy of  $X$ :

$$h_\mu(X, \xi) = \lim_{k \rightarrow \infty} \frac{1}{k^d} \sum_{u \in L_{\Gamma_n}(X)} f(\mu([u])),$$

where  $f(x) = -x \ln x$  for  $x > 0$  and  $f(0) = 0$ . Since  $\xi$  is a measure-theoretic generator for the  $\sigma$ -algebra of Borel sets in  $X$ , we know that  $h_\mu(X, \xi) = h_\mu(X)$ . (For more information on measure-theoretic and topological generators, see [Wal]) We now partition  $L_{\Gamma_k}(X)$  into the two pieces  $A_{k,\epsilon,\mu,u_1,\dots,u_j}(X)$  and  $A_{k,\epsilon,\mu,u_1,\dots,u_j}(X)^c$ :

$$\begin{aligned}
h_\mu(X) &= \lim_{k \rightarrow \infty} \left( \frac{1}{k^d} \sum_{u \in A_{k,\epsilon,\mu,u_1,\dots,u_j}(X)} f(\mu([u])) \right. \\
&\quad \left. + \frac{1}{k^d} \sum_{u' \in A_{k,\epsilon,\mu,u_1,\dots,u_j}(X)^c} f(\mu([u'])) \right).
\end{aligned}$$

To estimate each summand, we use the easily checkable fact that for any set of nonnegative reals  $\alpha_1, \alpha_2, \dots, \alpha_{j'}$  whose sum is  $\beta$ , the maximum value of  $\sum_{i=1}^{j'} f(\alpha_i)$  occurs when all terms are equal, and is  $\beta \ln \frac{j'}{\beta}$ . Using this, we see that the above sum is bounded from above by

$$\begin{aligned}
&\frac{1}{k^d} \left( \mu(A_{k,\epsilon,\mu,u_1,\dots,u_j}(X)) \ln \frac{|A_{k,\epsilon,\mu,u_1,\dots,u_j}(X)|}{\mu(A_{k,\epsilon,\mu,u_1,\dots,u_j}(X))} \right) \\
&+ \frac{1}{k^d} \left( (1 - \mu(A_{k,\epsilon,\mu,u_1,\dots,u_j}(X))) \ln \frac{H_{\Gamma_k}(X)}{1 - \mu(A_{k,\epsilon,\mu,u_1,\dots,u_j}(X))} \right).
\end{aligned}$$

Since, as  $k \rightarrow \infty$ ,  $\mu(A_{k,\epsilon,\mu,u_1,\dots,u_j}(X))$  approaches 1 and  $\frac{\ln |L_{\Gamma_k}(X)|}{k^d}$  approaches  $h^{\text{top}}(X) < \infty$ , the second term approaches zero as  $k \rightarrow \infty$ . By replacing  $\mu(A_{k,\epsilon,\mu,u_1,\dots,u_j}(X))$  by 1 in the limit in the first term, we see that

$$h_\mu(X) \leq \liminf_{k \rightarrow \infty} \frac{1}{k^d} \ln |A_{k,\epsilon,\mu,u_1,\dots,u_j}(X)|,$$

which was exactly what needed to be shown. □

**Corollary 3.2.8.** *For any strongly irreducible  $\mathbb{Z}^d$ -shift  $X$  of finite type  $t$  with uniform filling length  $R$ , for any  $\epsilon > 0$ , and for any positive integers  $k$  and  $M$ , denote by  $A_{k,\epsilon,M}(X)$  the set of words  $u$  in  $L_{\Gamma_k}(X)$  with between*



$k^d \left( \frac{1}{H_{\left( S_i \sqcup (S_i^c)^{(t+R)} \right) \setminus \left( S_i \sqcup (S_i^c)^{(t+R)} \right)^{(t)}(X)} - \epsilon \right)$  and  $k^d \left( \frac{1}{H_{S_i \setminus S_i^{(R)}}(X)} + \epsilon \right)$  occurrences of  $u_i \in L_{S_i}(X)$  for every word  $u_i$  with shape  $S_i \subseteq \Gamma_M$ . Then,

$$\lim_{k \rightarrow \infty} \frac{\ln |A_{k,\epsilon,M}(X)|}{k^d} = h^{top}(X).$$

**Proof:** Using a combination of Lemmas 3.2.6 and 3.2.7 for any fixed  $\tilde{\mu}$  an ergodic measure of maximal entropy on  $X$  shows that

$$\liminf_{k \rightarrow \infty} \frac{\ln |A_{k,\epsilon,M}(X)|}{k^d} \geq h_{\tilde{\mu}}(X) = h^{top}(X).$$

And by the definition of topological entropy,

$$\limsup_{k \rightarrow \infty} \frac{\ln |A_{k,\epsilon,M}(X)|}{k^d} \leq \limsup_{k \rightarrow \infty} \frac{h_{\Gamma_k}(X)}{k^d} = h^{top}(X).$$

□

### 3.3 A replacement theorem

We must prove an auxiliary theorem about replacements which is mostly combinatorial in nature, and which will be integral in the proof of the upper bound portion of Theorem 3.1.22.

**Theorem 3.3.1.** *For any strongly irreducible  $\mathbb{Z}^d$ -Markov shift  $X$  with uniform filling length  $R$ , there exists  $N$  depending on  $X$  such that for any choice of  $w \in L_{\Gamma_n}(X)$  with  $n > N$ , there exists  $w' \in L_{\Gamma_m}(X)$  a subword of  $w$ , where  $m > n - n^{1-\frac{1}{3d}}$ , and  $w'' \in L_{\Gamma_m}(X)$  so that  $w''$  and  $w'$  agree on the boundary, and with the property that*

replacing an occurrence of  $w'$  by  $w''$  in any element of  $X$  cannot possibly create a new occurrence of  $w'$ .

**Proof:** First let us make sure that the statement of Theorem 3.3.1 is totally clear. What this means is that we find  $w', w'' \in L_{\Gamma_m}(X)$  such that  $w'$  is a subword of  $w$ ,  $w'$  and  $w''$  agree on the boundary, and such that the following is true: for any  $x \in X$  with  $x(S) = w'$  for some  $S$  a copy of  $\Gamma_m$ , if one defines  $x' \in X$  by replacing  $x(S)$  by  $w''$ , then for any  $T$  a copy of  $\Gamma_m$  such that  $x'(T) = w'$ , it must be the case that  $x(T) = w'$ . We define  $l = \left\lceil \left( \frac{d \ln n}{h^{top}(X)} \right)^{\frac{1}{d}} \right\rceil + 1$  or  $l = \left\lceil \left( \frac{d \ln n}{h^{top}(X)} \right)^{\frac{1}{d}} \right\rceil + 2$ , whichever is odd. Then,  $H_{L_{\Gamma_l}}(X)$  is, by Lemma 3.2.1, at least  $e^{h^{top}(X)l^d} > e^{d \ln n} = n^d$ . This is greater than the number of words with shape  $\Gamma_l$  which occur as subwords of  $w$ , and so therefore there exists a word  $a \in L_{\Gamma_l}(X)$  which is not a subword of  $w$ . This allows us to make a definition:

**Definition 3.3.2.** *Given a subword  $u$  of  $w$ , choose a subword of  $u$  of shape  $\Gamma_{2R+l}$  which does not contain any portion of the boundary of  $u$ , and use strong irreducibility to replace this subword by any word of shape  $\Gamma_{2R+l}$  which has  $a$  as the subword occupying its central  $\Gamma_l$ . The word that  $u$  has become after this replacement is called a **standard replacement** of  $u$ , and so is any word created in this way.*

Choosing the  $\Gamma_{2R+l}$  to be replaced to be disjoint from the boundary of  $u$  ensures that any standard replacement of  $u$  agrees on the boundary with  $u$ . This means that since  $X$  is a Markov shift, any occurrence of  $u$  in a point  $x \in X$  can be replaced by a standard replacement  $u'$ , and the resulting element of  $\Omega$  is still in  $X$ .

**Definition 3.3.3.** *A subword  $u \in L_{\Gamma_k}(X)$  of  $w$  has **Property A** if for every standard*

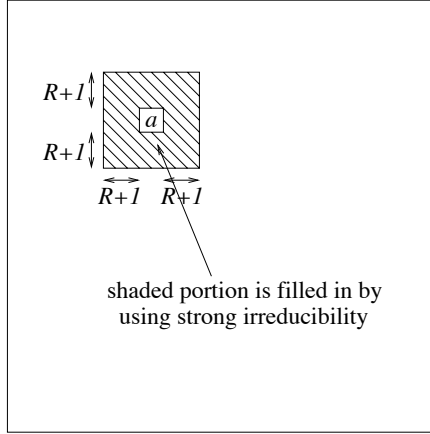


Figure 3.6: A standard replacement of  $u$

*replacement  $u'$  of  $u$ , replacing  $u$  by  $u'$  in some  $x \in X$  could possibly create a new occurrence of  $u$ . In other words, for every standard replacement  $u'$  of  $u$ , there exists  $x \in X$  and  $S, T$  copies of  $\Gamma_k$  such that  $x(S) = u$ ,  $x(T) \neq u$ , and if we define  $x' \in X$  by replacing  $x(S)$  by  $u'$ , then  $x'(T) = u$ .*

Clearly, if we can find  $u$  a subword of  $w$  which is large enough and does not have Property A, we will be done.

The organization of this induction is a bit odd, so we give a quick overview. We will construct a sequence  $w_j$  of subwords of  $w$ , where  $w_1 = w$  and each is a subword of the previous. Each  $w_j$  has shape a cube. We will be able to show that one of these  $w_j$  does not have Property A, and that it is large enough in the sense of Theorem 3.3.1. However, the construction of these  $w_j$  depends on a fact about periodicity of subwords of  $w$  which have Property A, which we must prove first. We will set up and prove

this periodicity fact for any  $w_j$  even though  $w_j$  has not yet been defined for  $j > 1$ . This is not a problem though; this proof depends only on  $w_j$  being a subword of  $w$ , which will be clearly true once the definitions of  $w_j$  are able to be made.

Let's assume that  $w_j$  has Property A for some  $j$ . For the purpose of the following constructions, it is helpful to picture  $w_j$  as fixed in space, so say that  $w_j$  has shape  $\Gamma_{n_j}$ , i.e. the cube with size  $n_j$  which lies entirely within the positive octant of  $\mathbb{Z}^d$  and has a vertex at  $(1, 1, \dots, 1)$ . To make sure to distinguish it from other copies of  $\Gamma_{n_j}$ , we call the shape that this fixed occurrence of  $w_j$  occupies  $B_j$ . We take a copy of  $\Gamma_{3(2R+l)\lceil n_j^{(1-\frac{1}{2d})} \rceil}$  or  $\Gamma_{3(2R+l)(\lceil n_j^{(1-\frac{1}{2d})} \rceil - 1)}$  (call it  $A_j$ , and its size  $k_j$ ), where the size is chosen to have the same parity as  $n_j$ , and which is central in  $B_j$ , and partition it into  $(3(\lceil n_j^{(1-\frac{1}{2d})} \rceil))^d$  or  $(3(\lceil n_j^{(1-\frac{1}{2d})} \rceil) - 1)^d$  disjoint copies of  $\Gamma_{2R+l}$ . Consider only the interior copies of  $\Gamma_{2R+l}$  in this partition, i.e. the ones which are disjoint from the boundary of  $A_j$ . For large  $n_j$ , there are more than  $2^d n_j^{d-\frac{1}{2}}$  of these, and to each one we may associate a standard replacement of  $w_j$ , and since  $w_j$  has Property A, also a copy of  $\Gamma_{n_j}$  which overlaps  $B_j$  which could somehow be filled with  $w_j$  at the same time that  $B_j$  is filled with the standard replacement in question. We use  $R_j$  to denote the set (possibly with repetitions) of these copies of  $\Gamma_{n_j}$ .

Recall that each of these represents a possible new occurrence of  $w_j$  created by performing a standard replacement. Since any of the standard replacements that we would perform is done by changing only letters in a particular  $\Gamma_{2R+l}$  in  $B_j$ , it must be the case that every element of  $R_j$  contains some portion of its associated  $\Gamma_{2R+l}$ . (Otherwise, the supposed “new” occurrence of  $w_j$  would have already existed before

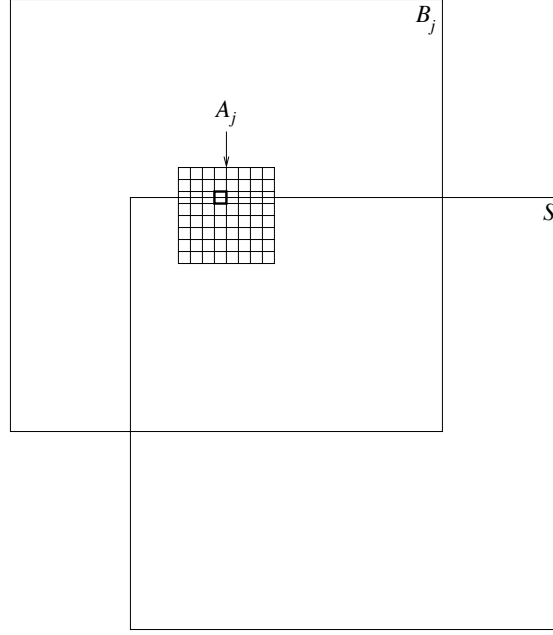


Figure 3.7: A sample element  $S$  of  $R_j$

the replacement.) Also, since every standard replacement results in some occurrence of  $a$  in  $A_j$ , and since  $a$  is not a subword of  $w_j$ , none of the elements in  $R_j$  contains all of its associated  $\Gamma_{2R+l}$ . In Figure 3.7, the element of  $R_j$  shown, which we call  $S$ , is associated to the standard replacement given by changing the letters in the highlighted  $\Gamma_{2R+l}$ .

We first make the claim that  $R_j$  contains no repeated elements, i.e. that two distinct copies of  $\Gamma_{2R+l}$  must be associated to distinct elements of  $R_j$ .

In Figure 3.8, suppose that  $S$  is associated to the standard replacements corresponding to each of  $U$  and  $V$ , the two highlighted copies of  $\Gamma_{2R+l}$ . This means that there

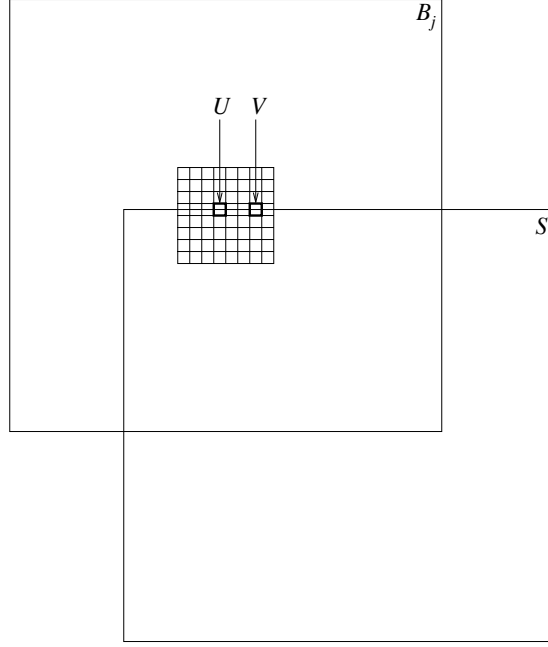


Figure 3.8: An element  $S$  of  $R_j$  associated to two standard replacements

exists a word  $u \in L_{B_j \cup S}(X)$  with  $u(S) = w_j$  and  $u(B_j) = w'_j$ , where  $w'_j$  is a standard replacement of  $w_j$  made by changing only the letters in  $U$ . The letters in  $U \cap S$  must be changed from their values in  $w_j$  in order for the occurrence of  $w_j$  in  $S$  to be “new.” Therefore,  $u(B_j \setminus U) = w_j(B_j \setminus U)$ , and  $u(U \cap S) \neq w_j(U \cap S)$ . Similarly, there exists a word  $u' \in L_{B_j \cup S}(X)$  with  $u'(S) = w_j$ ,  $u'(B_j \setminus V) = w_j(B_j \setminus V)$ , and  $u'(V \cap S) \neq w_j(V \cap S)$ . Since  $U$  and  $V$  are disjoint, this implies that  $u(U \cap S) \neq w_j(U \cap S)$  and  $u'(U \cap S) = w_j(U \cap S)$ , and so that  $u(U \cap S) \neq u'(U \cap S)$ . But  $u(S) = u'(S) = w_j$ , and since  $U \cap S \subseteq S$  we have a contradiction. Now we know that  $R_j$  consists of more than  $2^d n_j^{d-\frac{1}{2}}$  distinct copies of  $\Gamma_{n_j}$ .

**Definition 3.3.4.** A **suboctant** of  $B_j$  is a subcube of  $B_j$  which shares vertices with both  $B_j$  and  $A_j$ , is contained in  $B_j$ , and whose intersection with  $A_j$  contains only the shared vertex. (See Figure 3.9.)

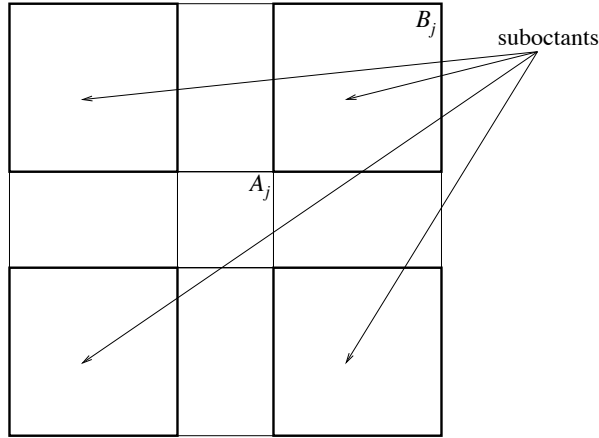


Figure 3.9: The suboctants of  $B_j$

It is fairly clear upon observation that each element of  $R_j$  must contain some suboctant of  $B_j$ , since each element of  $R_j$  contains some portion of  $A_j$ . Since  $R_j$  has more than  $2^d n_j^{d-\frac{1}{2}}$  distinct elements, there is some suboctant  $O_j$  which is contained in more than  $n_j^{d-\frac{1}{2}}$  of the copies of  $\Gamma_{n_j}$  in  $R_j$ .

We now prove a fact about the periodicity of a set which is contained in two elements of  $R_j$ . Suppose that two distinct elements  $S$  and  $S'$  of  $R_j$  are given such that  $S - S'$ , which we denote by  $v$ , is a multiple of  $e_i$  for some  $1 \leq i \leq d$ . We make the additional

The diagram illustrates the relativity of simultaneity in a spacetime context. It shows two frames,  $S$  (bottom) and  $S'$  (top), with their respective axes. A worldline  $r$  is shown. Points  $q$ ,  $t$ ,  $t+v$ ,  $p$ ,  $p+v$ , and  $r$  are marked. Lines of simultaneity are drawn for each frame, showing that events simultaneous in  $S$  are not simultaneous in  $S'$ . Vectors  $U$  and  $V$  are shown at points  $t$  and  $t+v$ . Velocities  $v$  and  $v'$  are indicated by arrows.

In Figure 3.10,  $t$  is any point of  $B_j$  such that  $t + v$  is also in  $B_j$ ,  $O'$  is the suboctant of  $B_j$  opposite  $O$ ,  $q$  is the vertex shared by  $O$  and  $B_j$ ,  $r$  is the vertex of  $B_j$  opposite



$q, p$  is the vertex of  $S$  corresponding to  $r$  in  $B_j$ , and  $v' = t - (p + v)$ .  $P$  and  $P'$  are subcubes of  $O$  and  $O'$ , respectively, which share vertices with  $A_j$  and whose size is  $\frac{n_j - k_j}{2} - k_j$ , or the size of  $O$  minus the size of  $A_j$ . (The dotted line portions are copies of  $A_j$  to remind us of the definition of  $P$  and  $P'$ , but are themselves irrelevant and are not named.) We define  $t'$  to be  $r + v'$ , giving us two new points  $t'$  and  $t' + v$ . Since  $S, S' \in R_j$ , they may each be filled with  $w_j$  when  $B_j$  is filled with the correct standard replacement of  $w_j$ . Call  $U$  the copy of  $\Gamma_{2R+l}$  on which  $w_j$  is altered to make the standard replacement which allows  $S$  to be filled with  $w_j$ , and  $V$  the copy of  $\Gamma_{2R+l}$  on which  $w_j$  is altered to make the standard replacement which allows  $S'$  to be filled with  $w_j$ . (A quick note: in several figures in this paper, we represent a point  $p \in \mathbb{Z}^d$  by the vector which points from  $\mathbf{0}$  to  $p$ .)

We first claim that if  $t, t + v \in P$ , then  $t', t' + v \in B_j \setminus (U \sqcup V)$ . Suppose that  $t, t + v \in P$ . Since  $S'$  contains some portion of  $V$ , but does not contain all of  $V$ , the same can be said of  $A_j$ ;  $S'$  contains some portion of  $A_j$ , but not all of  $A_j$ . Therefore, some coordinate of  $p + v$  is between  $\frac{n_j - k_j}{2} + 1$  and  $\frac{n_j + k_j}{2}$ . (We always use the word “between” in the inclusive sense.) The fact that  $S'$  has nonempty intersection with  $A_j$  also implies that all coordinates of  $p + v$  are between  $\frac{n_j - k_j}{2} + 1$  and  $n_j$ , since all coordinates of all elements of  $A_j$  are at least  $\frac{n_j - k_j}{2} + 1$ . Since  $t \in P$ , every coordinate of  $t$  is between  $k_j + 1$  and  $\frac{n_j - k_j}{2}$ . This implies that every coordinate of  $v'$  is nonpositive and at least  $-n_j + 1$ , and also that one coordinate of  $v'$  is at least  $-\frac{n_j - k_j}{2} + 1$ . Then,  $t' = r + v'$  is in  $B_j$ , and has one coordinate at least  $\frac{n_j + k_j}{2} + 1$ , and therefore does not lie in  $A_j$ , and so is not in  $U$  or  $V$ . Since  $t + v \in P$ , the same argument shows that

$t' + v$  is in  $B_j$ , but does not lie in  $U$  or  $V$ . Clearly, since  $U$  and  $V$  are disjoint from the boundary of  $A_j$  by their construction,  $t, t + v$  are not in  $U$  or  $V$  either.

Suppose that  $t, t + v, t', t' + v \in B_j \setminus (U \sqcup V)$ . Then by noting that  $S$  can be filled with  $w_j$  when  $B_j$  is filled with a standard replacement of  $w_j$  which agrees with  $w_j$  outside of  $U \sqcup V$ , we can infer that  $w_j(t) = w_j(t' + v)$  (this is because  $t \in S$  corresponds to  $t' + v \in B_j$ ), and by noting that  $S'$  can be filled with  $w_j$  when  $B_j$  is filled with a standard replacement of  $w_j$  which agrees with  $w_j$  outside of  $U \sqcup V$ , we infer that  $w_j(t) = w_j(t')$  and  $w_j(t + v) = w_j(t' + v)$ . (This is because  $t \in S'$  corresponds to  $t' \in B_j$ , and  $t + v \in S'$  corresponds to  $t' + v \in B_j$ ) This implies that  $w_j(t) = w_j(t + v)$ . This conclusion was dependent only on the fact that  $t, t + v, t', t' + v \in B_j \setminus (U \sqcup V)$ . We have already shown that this is true for any  $t, t + v \in P$ , and so  $w_j(P)$  is periodic with respect to  $v$ .

Note that in the course of this argument, we have also shown that  $w_j(t') = w_j(t' + v)$ . We claim that any pair of points in  $P'$  which are separated by  $v$  are  $t'$  and  $t' + v$  for some choice of  $t, t + v \in B_j \setminus (U \sqcup V)$ . To do this, we just determine  $t$  from  $t'$ : if  $t' \in P'$ , then every coordinate of  $v'$  is between  $-\frac{n_j - k_j}{2}$  and  $-k_j + 1$ . As already mentioned, all coordinates of  $p + v$  are between  $\frac{n_j - k_j}{2} + 1$  and  $n_j$ , and one coordinate is between  $\frac{n_j - k_j}{2} + 1$  and  $\frac{n_j + k_j}{2}$ . This implies that  $p + v + v' = t$  has all coordinates between 1 and  $n_j$ , and so is in  $B_j$ , and that one coordinate is between 1 and  $\frac{n_j - k_j}{2} + 1$ , which implies that  $t$  either does not lie in  $A_j$  or lies in the boundary of  $A_j$ , and, in either case, does not lie in  $U$  or  $V$ . Since  $t' + v \in P'$ , the same argument shows that  $t + v$  is in  $B_j$ , but does not lie in  $U$  or  $V$ . We have then shown that  $t'$  and  $t' + v$  could

be any pair of points in  $P'$  separated by  $v$ , and so the fact that  $w_j(t') = w_j(t' + v)$  in the above argument shows that  $w_j(P')$  is also periodic with respect to  $v$ .

We will use this fact momentarily; for now let's recall that we earlier showed that some suboctant  $O_j$  of  $B_j$  is contained in more than  $n_j^{d-\frac{1}{2}}$  distinct elements of  $R_j$ . As before, denote by  $q$  the vertex of  $B_j$  shared by  $B_j$  and  $O_j$ , and by  $r$  the vertex of  $B_j$  opposite  $q$ . It is clear upon observation that for any element  $S$  of  $R_j$  which contains  $O_j$ , the vertex of  $S$  which corresponds to  $r$  in  $B_j$  must be contained in  $B_j$ . Fix any  $e_i$  in  $\mathbb{Z}^d$ .  $B_j$  can be partitioned into  $n_j^{d-1}$  sets  $\{w + me_i\}_{0 \leq m < n_j}$  where  $w$  ranges over all points  $w$  in  $B_j$  with  $w_i = 1$ , and by the above comments, there are more than  $n_j^{d-\frac{1}{2}}$  distinct points in  $B_j$  which are vertices of elements of  $R_j$  which contain  $O_j$ . By the pigeonhole principle, this implies that one of the sets  $\{w + me_i\}_{0 \leq m < n_j}$  contains more than  $\sqrt{n_j}$  of these vertices. Again by the pigeonhole principle, this implies that there are two elements of  $R_j$ , call them  $S$  and  $S'$ , which contain  $O_j$  and where  $S' - S$  is a multiple of  $e_i$  whose length is less than  $\sqrt{n_j}$ . We make the notation  $v_i^{(j)} := S' - S$ . This in turn implies that  $w_j$  is periodic with respect to  $v_i^{(j)}$  on the regions above described as  $P$  and  $P'$ , and since these regions are dependent on  $O_j$  and  $O'_j$ , we call them  $P_j$  and  $P'_j$ . Since this can be done for all  $1 \leq i \leq d$ ,  $w_j(P_j)$  and  $w_j(P'_j)$  are periodic with respect to  $v_1^{(j)}, v_2^{(j)}, \dots, v_d^{(j)}$  where for each  $1 \leq i \leq d$ ,  $v_i^{(j)}$  is a multiple of  $e_i$  with length less than  $\sqrt{n_j}$ . This implies that their lengths are less than  $\sqrt{n}$  as well, since  $n_j \leq n$ .

**Definition 3.3.5.** A word  $w$  is **purely periodic** if for some positive integers  $n_1, \dots, n_d$ , it is periodic with respect to  $n_i e_i$  for  $1 \leq i \leq d$ .

We have then shown that if  $w_j$  has Property A, then there are regions  $P_j$  and  $P'_j$  as described above so that  $w_j(P_j)$  and  $w_j(P'_j)$  are purely periodic with all periods less than  $\sqrt{n}$ . In particular, the preceding is true for  $j = 1$  and  $w_1 = w$ . For each  $1 \leq i \leq d$ , we can then choose  $v_i^{(1)}$  a multiple of  $e_i$  which is a period for  $w_1(P_1)$  and  $w_1(P'_1)$  and which has length less than  $\sqrt{n}$ . We now choose  $w_2$  a subword of  $w_1$ . For sufficiently large  $n_1$ , we take  $w_2$  to be the subword of  $w_1$  with shape  $B_2 = \Gamma_{n_1 - ((10 \cdot 3^{2d})(2R+l) \lceil n_1^{1-\frac{1}{2d}} \rceil)}$  and which still has  $q$ , the vertex shared by  $O_1$  and  $B_1$ , as a vertex. The purpose of this is to cause the forced purely periodic portion  $P_1$  of  $w_1$  to contain a purely periodic central cube within  $w_2$ . We will choose subsequent  $w_j$  to be purely periodic on a central cube as well, which we will denote by  $C_j$ , and so  $w_j(C_j)$  will always be a purely periodic subword of  $w_j$ . In fact, for each  $j$ , we will choose  $w_{j+1}$  so that  $w_{j+1}(C_{j+1})$  is a subword of  $w_j(C_j)$ , and so each  $w_j(C_j)$  will be periodic with respect to  $v_i^{(1)}$  for  $1 \leq i \leq d$ . Due to the construction of  $w_2$ , we can take  $C_2$  to have shape  $\Gamma_{10 \cdot 3^{2d}(2R+l) \lceil n_1^{1-\frac{1}{2d}} \rceil}$  in  $w_2$ . We also move  $B_2$  in space to lie entirely within the positive octant of  $\mathbb{Z}^d$  with one vertex at  $(1, \dots, 1)$  (in other words,  $B_2 = \Gamma_{n_2}$ ), so that the phrasing of the earlier arguments still works.

This choice of  $w_2$  was done to create  $C_2$ : from now on we will follow a different inductive procedure. We need one more definition:

**Definition 3.3.6.** A ***superoctant*** of any  $B_j$  ( $j > 1$ ) is a subcube which shares vertices with both  $B_j$  and  $C_j$ , is contained in  $B_j$ , and contains  $C_j$ .

There is then a natural one-to-one correspondence between the superoctants and suboctants of  $B_j$ ; for every superoctant, there is exactly one suboctant which is a

subset of it. For each  $j \geq 2$  where  $w_j$  has Property A, we now describe how to construct  $w_{j+1}$ . For each such  $j$ , we define  $T_j$  to be the set of superoctants  $\overline{O}$  of  $B_j$  such that  $w_j(\overline{O})$  is periodic with respect to  $v_i^{(1)}$  for  $1 \leq i \leq d$ . We denote the size of  $C_j$  by  $m_j$ . Since  $w_1(P_1)$  is periodic with respect to each  $v_i^{(1)}$ , and since  $w_1(P_1) = w_2(\overline{O})$  for a superoctant  $\overline{O}$  of  $w_2$ , we have then already shown that  $T_2$  contains  $\overline{O}$ , and  $m_2 = 10 \cdot 3^{2d} (2R + l) \lceil n_1^{1-\frac{1}{2d}} \rceil$ . We assume that  $m_j \geq 10(2R + l)n_1^{1-\frac{1}{2d}}$  for all  $j$ , and in fact this will be clear once the construction is finished.

Let's now suppose that for some  $j \geq 2$ ,  $w_j$  has Property A. The idea is that we take  $w_{j+1}$  to be a suitable subword of  $w_j$  where  $C_{j+1}$  is smaller than  $C_j$ , but  $T_{j+1}$  is strictly larger than  $T_j$ . To show this, we have to prove a couple of claims. Firstly, we claim that any superoctant of  $B_j$  on which  $w_j$  is purely periodic with periods less than  $\sqrt{n}$  and which contains  $C_j$  must be periodic with respect to each  $v_i^{(1)}$ , i.e. must be in  $T_j$ . Since  $m_j \geq 10(2R + l)n^{1-\frac{1}{2d}}$ , clearly  $m_j \geq 4\sqrt{n}$  for all  $n$ .

We now claim that if  $T_j$  contains any pair of opposite superoctants  $\overline{O}$  and  $\overline{O}'$  of  $B_j$ , then neither of the suboctants  $O$  and  $O'$  associated to them can possibly be contained in any element of  $R_j$ .

We again assume without loss of generality that  $O$  is the least suboctant lexicographically of  $B_j$ . In Figure 3.11,  $w_j(\overline{O} \cup \overline{O}')$  (the portion of  $w_j$  shaded) is purely periodic with periods  $v_i^{(1)}$ , and  $U$  is the  $\Gamma_{2R+l}$  which is changed to make the standard replacement corresponding to  $S \in R_j$ . We assume that  $O \in S$  and derive a contradiction. We have denoted by  $C$  the region  $(U \cap S) + (B_k - S)$ , i.e.  $C$  in  $B_j$  corresponds to  $U \cap S$  in  $S$ .  $C \subset \overline{O}'$  because  $O \subset S$ , and for some  $1 \leq i \leq d$ , it is the case that

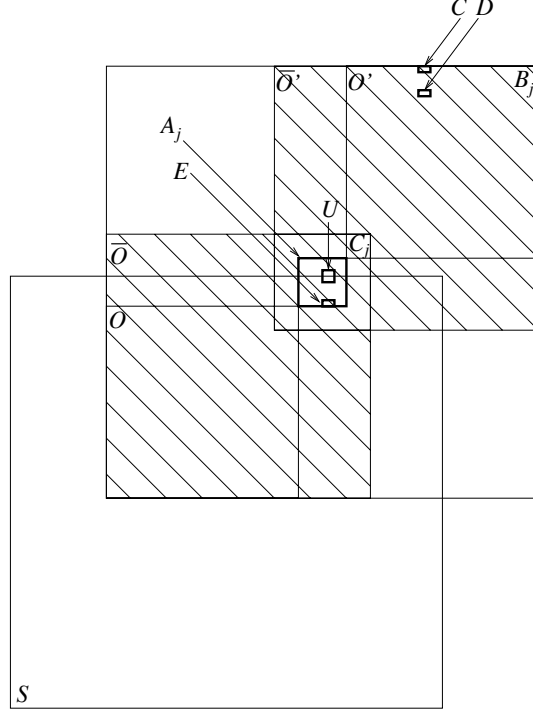


Figure 3.11: An element  $S$  of  $R_j$  which contains  $O$

$C$  consists of points whose  $i$ th coordinate is between  $n_j - k_j + 1$  and  $n_j$ . We then choose  $u_i$  which is a multiple of  $v_i^{(1)}$  such that  $C + u_i$  is a subset of  $\overline{O'}$  and  $U + u_i$  is a subset of  $\overline{O}$  which is disjoint from  $U$  (this can be done since  $v_i^{(1)}$  has length less than  $\sqrt{n}$ ; some multiple of it is greater than  $k_j$ , but still less than  $\frac{n_j - k_j}{2}$ ) and take  $D = C + u_i$  and  $E = (U \cap S) + u_i$ . It must be the case that filling  $B_j$  with a standard replacement for  $w_j$  in which no letters of  $B_j \setminus U$  are changed and at least some letters in  $U \cap S$  are changed can be done simultaneously with filling  $S$  with  $w_j$ . Call the word with shape  $B_j \cup S$  with this property  $u$ . We can restate our assumptions then

by saying  $u(B_j \setminus U) = w_j(B_j \setminus U)$ ,  $u(U \cap S) \neq w_j(U \cap S)$ , and  $u(S) = w_j$ . Since  $U \cap S$  in  $S$  corresponds to  $C$  in  $B_j$ , and since  $u(S) = w_j$ , it must be the case that  $u(U \cap S) = w_j(C)$ . By periodicity of  $w_j(\overline{O'})$  with respect to  $u_i$ ,  $w_j(C) = w_j(D)$ . Since  $D$  in  $B_j$  corresponds to  $E$  in  $S$ , and since  $u(S) = w_j$ ,  $w_j(D) = u(E)$ . Since  $u(B_j \setminus U) = w_j(B_j \setminus U)$  and  $E$  and  $U$  are disjoint,  $u(E) = w_j(E)$ . Finally, by periodicity of  $w_j(\overline{O})$  with respect to  $u_i$ ,  $w_j(E) = w_j(U \cap S)$ . But we have then shown that  $u(U \cap S) = w_j(U \cap S)$ , a contradiction. Therefore, if two opposite superoctants are contained in  $T_j$ , then neither of the suboctants associated to them may be contained in any element of  $R_j$ .

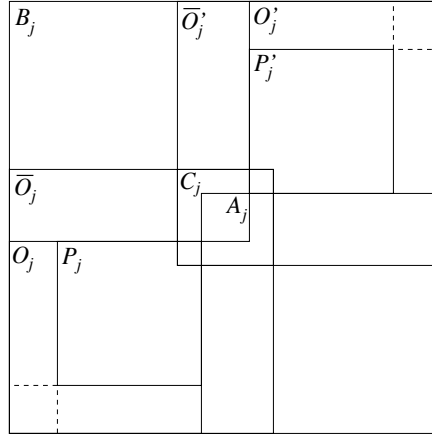


Figure 3.12:  $B_j$

Finally, we may make our inductive step. Consider any  $j \geq 2$  and  $w_j$  with Property A. By the earlier argument, we may then conclude that there are subcubes  $P_j$  and  $P'_j$

as defined above so that  $w_j(P_j)$  and  $w_j(P'_j)$  are purely periodic with periods at most  $\sqrt{n}$ . However, in the proof of this, we made use of the fact that the suboctant  $O_j$  of  $B_j$  corresponding to  $P_j$  is contained in many elements of  $R_j$ . Therefore, by the fact just shown, one of  $O_j$  or  $O'_j$  corresponds to a superoctant of  $B_j$  which is not in  $T_j$ .

First, we take a subword of  $w_j$ , call it  $w'_j$ , obtained by removing the boundary of thickness  $k_j$  from  $B_j$ . In other words, we take  $w'_j$  to be the subword whose shape is the central  $\Gamma_{n_j-2k_j}$  of  $B_j$ , which we call  $B'_j$ .

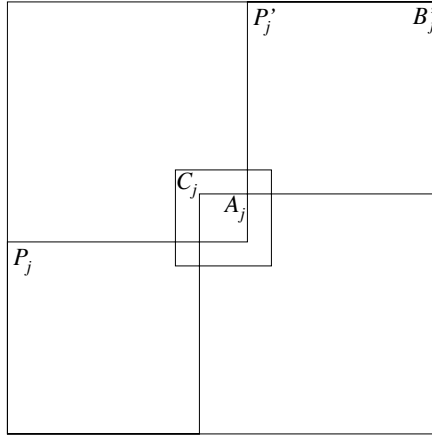


Figure 3.13:  $B'_j$

$w'_j(P_j)$  and  $w'_j(P'_j)$  are now subwords of  $w'_j$  occupying suboctants of  $B'_j$ . By construction,  $w'_j(P_j)$  and  $w'_j(P'_j)$  are also purely periodic with periods at most  $\sqrt{n}$ . However,



one of them is associated with a superoctant which is not in  $T_j$ , and we may assume that without loss of generality it is  $P_j$ . Also, since the size  $m_j$  of  $C_j$  is at least  $10(2R+l)n^{1-\frac{1}{2d}}$ , and since the size of  $A_j$  is at most  $3(2R+l)n^{1-\frac{1}{2d}}$ , the overlap between  $P_j$  and  $C_j$  is a cube whose size is at least  $\frac{m_j}{3}$ . Denote this cube by  $C_{j+1}$ .  $B_{j+1}$  is defined to be a subcube of  $B'_j$  which shares the same vertex with  $B'_j$  that  $B'_j$  shares with  $P_j$ , with the correct size so that the overlap between  $P_j$  and  $C_j$  is central. In doing this, we reduce the size of  $w'_j$  by at most  $m_j$ .  $w_{j+1}$  is defined to be  $w'_j(B_{j+1})$ . Then, since  $w_{j+1}(C_{j+1})$  is a subword of  $w_j(C_j)$ , it is also purely periodic with respect to each  $v_i^{(1)}$ , and by construction,  $C_{j+1}$  is central in  $B_{j+1}$ , so indeed  $C_{j+1}$  has the necessary properties. Each superoctant in  $T_j$  corresponds to a superoctant in  $T_{j+1}$ , and  $w_{j+1}(P_j)$  now occupies a superoctant of  $B_{j+1}$ , call it  $\overline{O}$ , on which  $w_{j+1}$  is purely periodic with periods at most  $\sqrt{n}$ . We claim that this means that  $\overline{O} \in T_{j+1}$  as well. For each  $1 \leq i \leq d$ , choose a period  $v_i^{(j)}$  of  $w_{j+1}(\overline{O})$  which is a multiple of  $e_i$  whose length is less than  $\sqrt{n}$ . For every  $j$ , as already mentioned,  $w_{j+1}(C_{j+1})$  is periodic with respect to  $v_i^{(1)}$  for  $1 \leq i \leq d$ . Now, consider any  $p \in \overline{O}$  such that  $p + v_k^{(1)} \in \overline{O}$  for some fixed  $1 \leq k \leq d$ . Since the size  $m_{j+1}$  of  $C_{j+1}$  is much greater than  $\sqrt{n}$ , there exists some linear combination  $\sum_{i=1}^d a_i v_i^{(j)}$  with  $a_i \in \mathbb{Z}$ , call it  $v$ , such that  $p + v \in C_{j+1}$  and  $(p + v) + v_k^{(1)} \in C_{j+1}$ . Then,  $w_{j+1}(p) = w_{j+1}(p + v)$  since  $v$  is a sum of periods for  $w_{j+1}(\overline{O})$ .  $w_{j+1}(p + v) = w_{j+1}(p + v + v_k^{(1)})$  since  $w_{j+1}(C_{j+1})$  is periodic with respect to  $v_k^{(1)}$ . Finally,  $w_{j+1}(p + v + v_k^{(1)}) = w_{j+1}(p + v_k^{(1)})$  since  $v$  is a sum of periods for  $w_{j+1}(\overline{O})$ , and therefore is itself a period of  $w_{j+1}(\overline{O})$ . Then, by definition,  $w_{j+1}(\overline{O})$  is periodic with respect to  $v_k^{(1)}$  as well, and since  $k$  was arbitrary, with respect to every  $v_i^{(1)}$  for  $1 \leq i \leq d$ . By definition, this shows that  $\overline{O} \in T_{j+1}$ .

Since  $\overline{O}$  corresponds to a superoctant of  $B_j$  that is not in  $T_j$ ,  $T_{j+1}$  contains at least one more element than  $T_j$ , and  $m_{j+1} \geq \frac{m_j}{3}$ . We again move  $B_j$  in space so that it lies entirely within the positive octant of  $\mathbb{Z}^d$  and has one vertex at  $(1, \dots, 1)$ . Since  $m_2 \geq 10 \cdot 3^{2^d} (2R + l) n^{1 - \frac{1}{2d}}$ , this implies that  $m_j \geq 10(2R + l) n^{1 - \frac{1}{2d}}$  for  $1 \leq j \leq 2^d$ . We will show that this induction never need proceed beyond  $j = 2^d$ , and so in the process will justify the claim already used that  $m_j \geq 10(2R + l) n^{1 - \frac{1}{2d}}$  for all  $j$  that we consider.

If each  $w_j$  for  $1 \leq j \leq 2^d$  has Property A, this means that  $T_{2^d+1}$  must have more than  $2^d$  elements. But, there are only  $2^d$  superoctants of  $B_{2^d+1}$ , and so we have a contradiction. This implies that some  $w_j$  must not have Property A, which we call  $w'$ .  $w'$  is created by at most  $2^d$  truncations of  $w$ , each of which reduces the size by at most  $k_2 + m_2$ , and so since  $k_2 < m_2$  for large enough  $n$ , the size  $m$  of  $w'$  is at least  $n - 10 \cdot 2^{d+1} 3^{2^d} (2R + l) \lceil n^{1 - \frac{1}{2d}} \rceil$ . Since for large  $n$ ,  $2R + l < c \ln n$  for some constant  $c$ , we can then say that there exists  $N'$  for which  $m > n - K \ln n \cdot n^{1 - \frac{1}{2d}}$  for  $n > N'$  and some constant  $K$ . Then, take  $N > N'$  so that  $K \ln n < n^{\frac{1}{6d}}$  for  $n > N$ . This implies that for  $n > N$ ,  $m > n - n^{1 - \frac{1}{3d}}$ . Since  $w'$  does not have Property A, there exists a standard replacement of  $w'$ , call it  $w''$ , which agrees on the boundary with  $w'$  by definition of standard replacement, and with the property that replacing  $w'$  by  $w''$  in some  $x \in X$  cannot possibly create a new occurrence of  $w'$ . Thus, Theorem 3.3.1 is proved.

□

We can prove as a corollary a version of Theorem 3.3.1 for any strongly irreducible shift of finite type:

**Corollary 3.3.7.** *For any strongly irreducible  $\mathbb{Z}^d$ -shift  $X$  of finite type  $t$  with uniform filling length  $R$ , there exists  $N$  depending on  $X$  such that for any choice of  $w \in L_{\Gamma_n}(X)$  with  $n > N$ , there exists a subword  $w'$  of  $w$ , with shape  $\Gamma_m$ , where  $m > n - n^{1-\frac{1}{3d}}$ , and  $w'' \in L_{\Gamma_m}(X)$  so that  $w''$  and  $w'$  agree on the boundary of thickness  $t$ , and with the property that replacing an occurrence of  $w'$  by  $w''$  in an element of  $X$  cannot possibly create a new occurrence of  $w'$ .*

**Proof:** Fix  $X$  a strongly irreducible  $\mathbb{Z}^d$ -shift of finite type  $t$  and uniform filling length  $R$ . Take the alphabet  $A^{(t)}$  whose letters are the elements of  $L_{\Gamma_t}(X)$ , i.e. the words in the language of  $X$  whose shape is  $\Gamma_t$ , and then define a map  $f$  from  $X$  to  $(A^{(t)})^{\mathbb{Z}^d}$  where for any  $p \in \mathbb{Z}^d$ ,  $(f(x))(p)$  is the subword of  $x$  which occupies  $\Gamma_t + p$ , i.e. the copy of  $\Gamma_t$  whose least vertex lexicographically is  $p$ . Then,  $f(X)$  is a Markov shift with alphabet  $A^{(t)}$ . The reader can check that  $f(X)$  is still strongly irreducible, but with a uniform filling length of  $R + t - 1$ . In short, this is because two shapes which are a distance of  $j$  apart in an element of  $x$  correspond to shapes a distance of  $j + t - 1$  apart in  $f(x)$  for  $j \in \mathbb{N}$ .

Since  $f(X)$  is a strongly irreducible Markov shift, by Theorem 3.3.1 there exists  $N'$  such that for any  $w \in L_{\Gamma_n}(f(X))$  with  $n > N'$ , there exist  $w' \in L_{\Gamma_m}$  a subword of  $w$ , with  $m > n - n^{1-\frac{1}{3d}}$ , and  $w''$  with the same shape as  $w'$  and which agrees on the boundary with  $w'$  such that replacing  $w'$  with  $w''$  in  $x' \in f(X)$  cannot result in the creation of a new occurrence of  $w'$ .

Now, take any word  $w \in L_{\Gamma_n}(X)$  with  $n > N + t - 1$ . Although  $f$  is defined as a map from one subshift to another, it should be fairly clear that as defined, it can also function as a map from  $L_{\Gamma_{k+t-1}}(X)$  to  $L_{\Gamma_k}(f(X))$ . So we can define  $f(w) \in L_{\Gamma_{n-t+1}}(f(X))$ . For ease of notation, we define  $\bar{n} = n - t + 1$ , so  $\bar{n} > N$ . Then, by Theorem 3.3.1, we can find  $v' \in L_{\Gamma_m}(f(X))$  a subword of  $f(w)$ , with  $m > \bar{n} - \bar{n}^{1-\frac{1}{3d}}$ , and  $v''$  which agrees with  $v'$  on the boundary so that replacing  $v'$  with  $v''$  in any element of  $f(X)$  cannot create a new occurrence of  $w'$ .  $f^{-1}(v')$  and  $f^{-1}(v'')$  are then subwords of  $w$  with shape  $\Gamma_{m+t-1}$  which agree on the boundary of thickness  $t$ . We wish to show that replacing  $f^{-1}(v')$  by  $f^{-1}(v'')$  in some  $x \in X$  cannot possibly result in a new occurrence of  $f^{-1}(v')$ . Suppose that this is not the case. Then there is  $x \in X$  and two copies of  $\Gamma_{m+t-1}$ , call them  $S$  and  $S'$ , such that  $x(S) = f^{-1}(v')$ ,  $x(S') \neq f^{-1}(v')$ , and if we define  $x' \in X$  by replacing  $x(S)$  by  $f^{-1}(v'')$ , then  $x'(S') = f^{-1}(v')$ . If we apply  $f$  to all of the objects in the above description, we arrive at a contradiction to the definition of  $v'$  and  $v''$ . It is therefore the case that replacing  $f^{-1}(v')$  with  $f^{-1}(v'')$  in any  $x \in X$  cannot create a new occurrence of  $f^{-1}(v')$ . Since  $v'$  is a subword of  $f(w)$ ,  $f^{-1}(v')$  is a subword of  $w$ .  $m > \bar{n} - \bar{n}^{1-\frac{1}{3d}}$ , and the size of the shape of  $f^{-1}(v')$  is  $m + t - 1 > n - \bar{n}^{1-\frac{1}{3d}} \geq n - n^{1-\frac{1}{3d}}$ , and so we are done.

□

Corollary 3.3.7 is the main tool in the proof of Theorem 3.1.22; very roughly, the idea is that for each occurrences of  $w$  in some large word in  $X$ , we destroy this occurrence of  $w$  by replacing the occurrence of  $w'$  which is a subword of it by  $w''$ .

The requirement that new occurrences of the destroyed words not appear during this process is to ensure that the process of getting rid of these occurrences of  $w$  takes as few steps as possible. It is natural to wonder why we go to the trouble of dealing with subwords in the statement of Corollary 3.3.7; i.e., why is it not possible, instead of dealing with the intermediate step of taking a subword of  $w$ , to just find a word  $w' \neq w$  such that replacing  $w$  by  $w'$  can never result in a new occurrence of  $w$  being created? The answer is that there are examples of strongly irreducible shifts of finite type  $X$  and words  $w \in L(X)$  for which this is impossible! Here is a quick example. Consider, for any dimension  $d$  and any  $n > 1$ , the full shift  $Y = \Omega$ , and the word  $w \in L_{\Gamma_n}(Y)$  defined by  $w(v) = 1$  if  $v_1 = v_2 = \dots = v_d = 1$ , and  $w(v) = 0$  if  $1 \leq v_i \leq n$  for  $1 \leq i \leq d$  and  $v_i > 1$  for some  $1 \leq i \leq d$ .  $w$  then has zeroes at every point of  $\Gamma_n$  except the least lexicographically. We claim that for any  $w' \neq w$ , there is some  $x \in Y$  such that  $x(\Gamma_n) = w$ , and replacing  $x(\Gamma_n)$  by  $w'$  creates a new occurrence of  $w$ . Consider any  $w' \neq w$ . Suppose that  $w'$  contains a one, i.e. there is some  $v \in \Gamma_n$  such that  $w'(v) = 1$ . Since  $w' \neq w$ , we can assume that  $v \neq (1, 1, \dots, 1)$ . Consider  $x \in Y$  defined by taking  $x(1, 1, \dots, 1) = 1$  and  $x(v) = 0$  for all other  $v \in \mathbb{Z}^d$ . Then,  $x(\Gamma_n) = w$ . Create a new  $x' \in Y$  by replacing  $x(\Gamma_n)$  by  $w'$ . Then, it is not hard to check that  $x'(\Gamma_n + (v - (1, 1, \dots, 1))) = w$ , and that  $x(\Gamma_n + (v - (1, 1, \dots, 1)))$  is the word consisting of all zeroes, and so not equal to  $w$ . This means that the replacement involved in changing  $x$  to  $x'$  created a new occurrence of  $w$ . The only other possibility for  $w'$  is that  $w'$  is the word consisting of all zeroes, i.e.  $w'(v) = 0$  for all  $v \in \Gamma_n$ . If this is the case, then define  $x \in Y$  by taking  $x(0, 1, \dots, 1) = x(1, 1, \dots, 1) = 1$  and  $x(v) = 0$  for all other  $v \in \mathbb{Z}^d$ . Again,  $x(\Gamma_n) = w$ . Create  $x' \in Y$  by replacing  $x(\Gamma_n)$

by  $w'$ . Then, it is not hard to check that  $x'(\Gamma_n - e_1) = w$ , and that  $x(\Gamma_n - e_1)$  had two ones, and is thus not equal to  $w$ . This means that in this case also, the replacement involved in changing  $x$  to  $x'$  created a new occurrence of  $w$ . We have then shown that for any  $w' \neq w$ , it is possible that a replacement of  $w$  by  $w'$  could create a new occurrence of  $w$ , and therefore the extra step of taking subwords in Theorem 3.3.1 and Corollary 3.3.7 is in fact necessary.

### 3.4 The proof of the main result

**Proof of Theorem 3.1.22:** We take  $X$  to be a strongly irreducible shift of finite type  $t$ . For any word  $w$  in  $L_{\Gamma_n}(X)$ , we call  $X_w$  the shift of finite type resulting from removing  $w$  from the language of  $X$ . We begin by proving an upper bound on  $h^{top}(X) - h^{top}(X_w)$  for sufficiently large  $n$ . By Corollary 3.3.7, as long as  $n$  is sufficiently large, there exists  $w' \in L_{\Gamma_m}(X)$  a subword of  $w$  with  $m > n - n^{1-\frac{1}{3d}}$  and  $w''$  with the same shape as  $w'$  and which agrees with  $w'$  on  $\Gamma_m^{(t)}$  with the property that replacing  $w'$  by  $w''$  cannot create new occurrences of  $w'$ .

For any  $k > m$ , we will define a mapping  $\phi_k : L_{\Gamma_{k+4m}}(X) \rightarrow L_{\Gamma_k}(X_w)$ .  $\phi_k$  will actually be defined as a composition of three maps:  $\alpha_k : L_{\Gamma_{k+4m}}(X) \rightarrow L_{\Gamma_{k+4m}}(X)$ ,  $\beta_k : \alpha_k(L_{\Gamma_{k+4m}}(X)) \rightarrow L_{\Gamma_{k+4m}}(X_w)$ , and  $\gamma_k : (\beta_k \circ \alpha_k)(L_{\Gamma_{k+4m}}(X)) \rightarrow L_{\Gamma_k}(X_w)$ . To define these, we need a definition:

**Definition 3.4.1.** *Two overlapping occurrences of the word  $w'$  which occur in copies  $S$  and  $S'$  of  $\Gamma_m$  are said to have **overlap of type B** if  $|S - S'|_\infty \geq \frac{m}{2} - 2n^{1-\frac{1}{3d}}$ .*

We point out an important property of this definition: given a fixed occurrence of  $w'$ , there are at most  $3^d m^d$  copies of  $\Gamma_m$  which could be filled with an overlapping  $w'$ , and so there are at most  $3^d m^d$  words which are made up of a pair of  $w'$  whose overlap is of type B. Call these words  $u_1, u_2, \dots, u_b$ , where  $b \leq 3^d m^d$ , and call  $S_i$  the shape of  $u_i$  for any  $1 \leq i \leq b$ . Given any  $u \in L_{\Gamma_{k+4m}}(X)$ , we define  $\alpha_k(v)$  by finding each occurrence of any of the  $u_i$  within  $u$ , and in each one, replacing the first lexicographically of the two occurrences of  $w'$  which make up  $u_i$  by  $w''$ . In order to make this operation well-defined, we must specify an order for these replacements to be done in; let us just choose the usual lexicographic order, and search for and replace each  $u_i$  in turn. (This means that we first replace the lexicographically first occurrence of  $u_1$ , then find the new lexicographically first occurrence of  $u_1$  and replace it, and continue until no  $u_1$  remain. We then perform the same procedure for  $u_2, u_3$ , etc.) Since replacing  $w'$  by  $w''$  can never create a new occurrence of  $w'$ , the resulting  $\Gamma_{k+4m}$ -word, which we call  $\alpha_k(u)$ , contains no  $u_i$ , i.e. contains no pair of occurrences of  $w'$  with overlap of type B.

For any  $\alpha_k(u) \in \alpha_k(\Gamma_{k+4m}(X))$ , we define  $(\beta_k \circ \alpha_k)(u)$  as follows: find the first occurrence of  $w$  lexicographically in  $\alpha_k(u)$ , and replace the  $w'$  which is a subword of this occurrence of  $w$  by  $w''$ . Denote by  $\tilde{w}$  the word that  $w$  becomes when its subword  $w'$  is replaced by  $w''$ . Repeat this procedure until there are no occurrences of  $w$  left, and call the resulting  $\Gamma_{k+4m}$ -word  $(\beta_k \circ \alpha_k)(u)$ . There is one fact that we must check; namely that when performing any of these replacements, we do not accidentally create a new occurrence of  $w$ . For a contradiction, assume that a particular replacement of  $w'$  by  $w''$  could create a new occurrence of  $w$ .

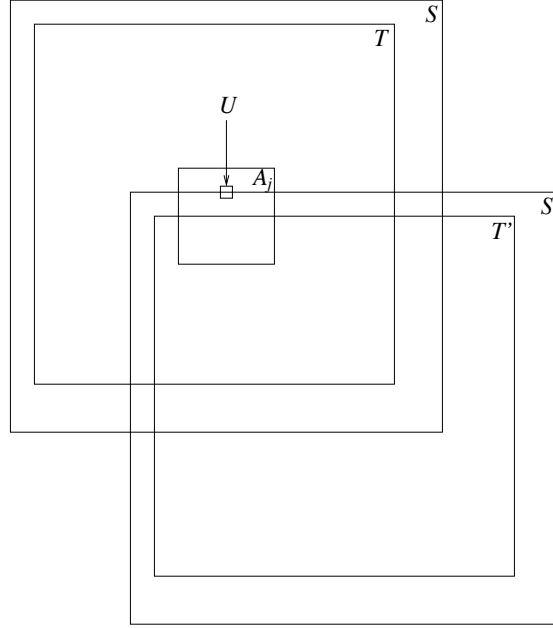


Figure 3.14: Intersecting occurrences of  $w$

In Figure 3.14,  $S$  is a copy of  $\Gamma_n$  which is filled with  $w$ ,  $T$  is a copy of  $\Gamma_m$  which is filled with  $w'$ ,  $U$  is the  $\Gamma_{2R+l}$  on which letters are changed to change the word on  $T$  from  $w'$  to  $w''$ ,  $S'$  is a copy of  $\Gamma_n$  which will be filled with  $w$  once the replacement is done,  $T'$  is the copy of  $\Gamma_m$  in  $S'$  corresponding to  $T$  in  $S$ , which will therefore be filled with  $w'$  once the replacement is done, and  $A_j$  is the copy of  $\Gamma_{3(2R+l)\left\lceil n_j^{1-\frac{1}{2d}} \right\rceil}$  or  $\Gamma_{3(2R+l)\left(\left\lceil n_j^{1-\frac{1}{2d}} \right\rceil - 1\right)}$  central in  $S$  in which  $U$  must lie, which comes from the proof of Theorem 3.3.1. For the same reasons as in the proof of Theorem 3.3.1,  $S'$  cannot contain all of  $U$ , but must contain some nonempty subset of  $U$ . Since the replacement cannot possibly create a new occurrence of  $w'$ ,  $T'$  must have already been filled with



$w'$  before the replacement occurred. This implies that before the replacement,  $T$  and  $T'$  were filled with overlapping occurrences of  $w'$ . We claim that these two occurrences had overlap of type  $B$ . Since  $U \subseteq A_j$ , and since the size of  $A_j$  is at most  $3(2R + l) \lceil n^{1-\frac{1}{2d}} \rceil$ , the distance in the  $e_i$ -direction between the center of  $U$  and the center of  $A_j$  is less than  $\frac{3}{2}(2R + l)n^{1-\frac{1}{2d}}$  for each  $1 \leq i \leq d$ .  $A_j$  is central in  $S$ , so the center of  $A_j$  is the same as the center of  $S$ . Since  $T$  is a subcube of  $S$  whose size is at most  $n^{1-\frac{1}{3d}}$  shorter, the distance in the  $e_i$ -direction between the center of  $S$  and the center of  $T$  is less than  $\frac{1}{2}n^{1-\frac{1}{3d}}$  for each  $1 \leq i \leq d$ . For the same reason, the distance in the  $e_i$ -direction between the center of  $S'$  and the center of  $T'$  is less than  $\frac{1}{2}n^{1-\frac{1}{3d}}$  for each  $1 \leq i \leq d$ . Finally, since  $U$  intersects the boundary of  $S'$ , and since  $U$  has size  $2R + l$ , there exists  $1 \leq i \leq d$  for which the distance in the  $e_i$ -direction between the center of  $S'$  and the center of  $U$  is between  $\frac{m}{2} - (R + \frac{l}{2})$  and  $\frac{m}{2} + (R + \frac{l}{2})$ . Putting all of these facts together, we see that there exists  $1 \leq i \leq d$  so that the distance in the  $e_i$ -direction between the center of  $T$  and the center of  $T'$  is between  $\frac{m}{2} - \frac{3}{2}(2R + l)n^{1-\frac{1}{2d}} - n^{1-\frac{1}{3d}} - (R + \frac{l}{2})$  and  $\frac{m}{2} + \frac{3}{2}(2R + l)n^{1-\frac{1}{2d}} + n^{1-\frac{1}{3d}} + (R + \frac{l}{2})$ , which implies that for large enough  $n$ , it is between  $\frac{m}{2} - 2n^{1-\frac{1}{3d}}$  and  $\frac{m}{2} + 2n^{1-\frac{1}{3d}}$ , which implies that before the replacement,  $T$  and  $T'$  were indeed filled with occurrences of  $w'$  with overlap of type  $B$ . However,  $\alpha_k(u)$  contained no pair of occurrences of  $w'$  with overlap of type  $B$ , and so since all replacements in the definition of  $\beta_k$  involve replacing some  $w'$  with  $w''$ , no new occurrences of  $w'$  can be created during these replacements. Therefore, there can never be a pair of occurrences of  $w'$  with overlap of type  $B$  during the application of  $\beta_k$  to a member of  $\alpha_k(\Gamma_{k+4m})$ . We therefore have a contradiction. Our original assumption was wrong, and none of these replacements

can create a new occurrence of  $w$ . So, when these replacements are finished, we have a word with shape  $\Gamma_{k+4m}$ , which we called  $(\beta_k \circ \alpha_k)(u)$ , which has no occurrences of  $w$ .

The definition of  $\gamma_k$  is much simpler: for any  $u \in \Gamma_{k+4m}(X)$ ,  $\gamma_k(u)$  is the subword of  $u$  occupying its central  $\Gamma_k$ . For any  $u \in \Gamma_{k+4m}(X)$ , we define  $\phi_k(u) = (\gamma_k \circ \beta_k \circ \alpha_k)(u)$ .

We now claim that  $\phi_k(u)$  is in  $L(X_w)$ , i.e. can be extended to a configuration on  $\mathbb{Z}^d$  which is in  $X$  and contains no occurrences of  $w$ . To do this, we inductively define a sequence of words  $d_j$ , where  $d_j \in L_{\Gamma_{k+4jm}}(X)$  and each  $d_j$  has no occurrences of  $w$ . Take  $d_0 = \phi_k(u)$  and  $d_1 = (\beta_k \circ \alpha_k)(u)$ . For any  $d_j$  for  $j > 0$ , define  $d_{j+1}$  as follows: since  $d_j \in L(X)$ , it can be extended to an infinite configuration in  $X$ . Use this fact to create a word  $d'_j \in L_{\Gamma_{k+4(j+1)m}}(X)$  which has  $d_j$  as the word occupying its central  $\Gamma_{k+4jm}$ . We then create  $d_{j+1}$  by performing the same replacements as in the definitions of  $\alpha_k$  and then  $\beta_k$ , in other words first finding and replacing any occurrences of  $u_i$  for any  $1 \leq i \leq b$ , and then finding and replacing any occurrences of  $w$ . For the same reasons as above,  $d_{j+1}$  has no occurrences of  $w$ . Since  $d_{j+1}$  agrees on the boundary of thickness  $t$  with  $d'_j$ ,  $d_{j+1} \in L(X)$ , completing the inductive step and allowing us to define  $d_j$  for all  $j$ . We also note that for any  $j$ , since  $d_j$  contained no occurrences of  $w$  or  $u_i$  for any  $1 \leq i \leq b$ , any occurrences of these words in  $d'_j$  must have had nonempty overlap with the newly created portion, i.e. must have been contained in the boundary of thickness  $4m$  of  $\Gamma_{k+4(j+1)m}$ . As a result,  $d_{j+1}$  and  $d_j$  must agree on their respective central  $\Gamma_{k+4(j-1)m}$ , since no replacements could have affected that portion. This means that we may define the “limit” of the  $d_j$ : for each

$j$ , since  $d_j \in L(X)$ , there exists  $z_j \in X$  with  $z_j(\Gamma_{k+4jm}) = d_j$ . Since  $d_{j+1}$  and  $d_j$  agree on their respective central  $\Gamma_{k+4(j-1)m}$ , the limit of the  $z_j$  exists, call it  $x$ . Since  $X$  is a subshift, it is closed and so  $x \in X$ . Since the central  $\Gamma_k$  of each  $d_j$  is equal to  $\phi_k(u)$ ,  $\phi_k(u)$  is a subword of  $x$ . Finally, since each  $d_j$  has no occurrences of  $w$  and  $x$  is a limit of the  $d_j$ ,  $x$  has no occurrences of  $w$ . This means that  $\phi_k(u) \in L_{\Gamma_k}(X_w)$  as claimed.

We have now defined  $\phi_k$  on all of  $L_{\Gamma_{k+4m}}(X)$ . For any  $\epsilon > 0$ , we can then deal with the restriction of  $\phi_k$  to  $A_{k+4m,\epsilon,3n}(X)$ . Since the shape of  $w$  is  $\Gamma_n \subseteq \Gamma_{3n}$ , any element of  $A_{k+4m,\epsilon,3n}(X)$  has fewer than  $(k+4m)^d(\frac{1}{H_{\Gamma_{n-2R}}(X)} + \epsilon)$  occurrences of  $w$ , and since each  $u_i$  has shape a subset of  $\Gamma_{3n}$ , any element of  $A_{k+4m,\epsilon,3n}(X)$  has fewer than  $(k+4m)^d(\frac{1}{H_{S_i \setminus S_i^{(R)}}(X)} + \epsilon)$  occurrences of  $u_i$  for each  $1 \leq i \leq b$ . By Lemma 3.2.1,  $H_{\Gamma_{n-2R}}(X) \geq e^{h^{top}(X)(n-2R)^d}$ . We now bound  $H_{S_i \setminus S_i^{(R)}}(X)$  in a similar fashion. First let's look at a picture of  $S_i \setminus S_i^{(R)}$  for any  $1 \leq j \leq b$ . Recall that  $S_i$  is the union of two overlapping copies of  $\Gamma_m$ , so  $S_i \setminus S_i^{(R)}$  is the union of two overlapping copies of  $\Gamma_{m-2R}$  concentric with the original copies of  $\Gamma_m$ . By the definition of type B overlap, there exists  $1 \leq i \leq d$  for which the distance in the  $e_i$ -direction between the centers of these two  $\Gamma_{m-2R}$  is between  $\frac{m}{2} - 2n^{1-\frac{1}{3d}}$  and  $\frac{m}{2} + 2n^{1-\frac{1}{3d}}$ .

In Figure 3.15, we denote the direction in question by  $e_i$ , the two copies of  $\Gamma_{m-2R}$  by  $S$  and  $S'$ , and the distance in the  $e_i$ -direction between the centers of  $S$  and  $S'$  by  $c$ . We then define  $T$  to be a rectangular prism which is a subset of  $S$  a distance of  $R+1$  away from  $T'$ . The sizes of  $T$  are  $m-2R$  in every direction but  $e_i$ , and  $c-R \geq m-2n^{1-\frac{1}{3d}}-R$  in the  $e_i$ -direction.

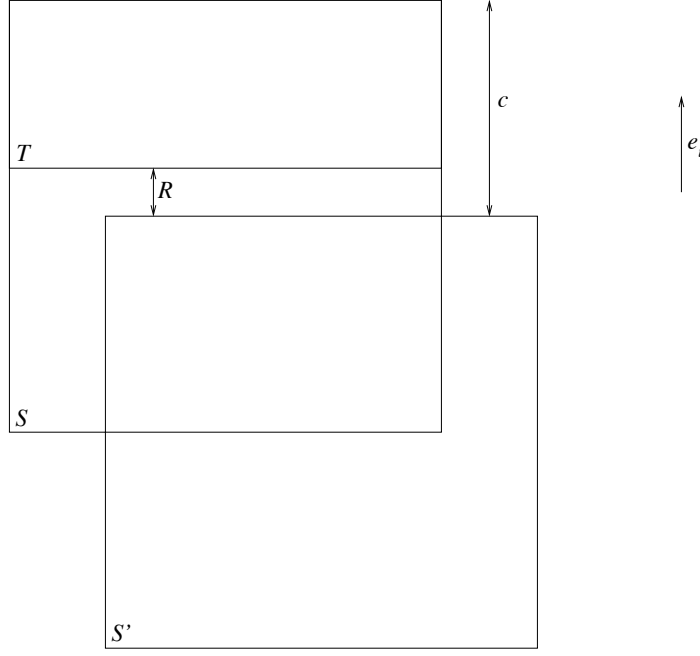


Figure 3.15:  $S_i \setminus S_i^{(R)}$

Clearly by just taking very large  $m$  we can assume that this dimension is greater than  $m - 3n^{1-\frac{1}{3d}}$ . Since  $d(T, S') = R + 1$  and  $T, S' \subseteq S_i \setminus S_i^{(R)}$ , by strong irreducibility  $H_{S_i \setminus S_i^{(R)}}(X) \geq H_T(X)H_{S'}(X) \geq e^{h^{top}(X)[(m-2R)^{d-1}(\frac{m}{2}-3n^{1-\frac{1}{3d}})+(m-2R)^d]}$ .  $R$ ,  $t$ , and  $n^{1-\frac{1}{3d}}$  are all small in relation to  $m$  for large enough  $m$ , and  $m$  asymptotically approaches  $n$  as  $n \rightarrow \infty$ , so for large  $n$ , this bound is greater than  $e^{h^{top}(X)1.4n^d}$ . Using these bounds, we can rewrite the above statement about  $A_{k+4m,\epsilon,3n}(X)$  a little more easily: every element of  $A_{k+4m,\epsilon,3n}(X)$  has fewer than  $(k+4m)^d(e^{-h^{top}(X)(n-2R)^d} + \epsilon)$  occurrences of  $w$  and fewer than  $(k+4m)^d(e^{-1.4h^{top}(X)n^d} + \epsilon)$  occurrences of any  $u_i$  for  $1 \leq i \leq b$ .

We now prove the upper bound on  $h^{top}(X) - h^{top}(X_w)$ . We do this by proving upper bounds on  $|\alpha_k^{-1}(v) \cap A_{k+4m,\epsilon,3n}(X)|$  for any  $v \in \alpha_k(L_{\Gamma_{k+4m}}(X))$ , on  $|\beta_k^{-1}(v') \cap \alpha_k(A_{k+4m,\epsilon,3n}(X))|$  for any  $v' \in L_{\Gamma_k}(X)$ , and on  $|\gamma_k^{-1}(v'')|$  for any  $v'' \in L_{\Gamma_k}(X_w)$ . The first upper bound is fairly straightforward: consider any  $u \in A_{k+4m,\epsilon,3n}(X)$ . Since  $b \leq 3^d m^d$ , the total number of letters in  $u$  which are part of an occurrence of any  $u_i$  is at most  $(2m^d)(3^d m^d)(k+4m)^d(e^{-1.4h^{top}(X)n^d} + \epsilon)$ , which for large  $m$  is less than  $(k+4m)^d(e^{-1.3h^{top}(X)n^d} + 3^{d+1}m^{2d}\epsilon)$ . Now, consider any  $v \in \alpha_k(L_{\Gamma_{k+4m}}(X))$ . By the previous reasoning, any  $u \in \alpha_k^{-1}(v) \cap A_{k+4m,\epsilon,3n}(X)$  differs from  $v$  on less than  $(k+4m)^d(e^{-1.3h^{top}(X)n^d} + 3^{d+1}m^{2d}\epsilon)$  letters, and so the number of possible  $u$  is at most

$$\begin{aligned} & |A|^{(k+4m)^d(e^{-1.3h^{top}(X)n^d} + 3^{d+1}m^{2d}\epsilon)} \sum_{i=0}^{\lfloor (k+4m)^d(e^{-1.3h^{top}(X)n^d} + 3^{d+1}m^{2d}\epsilon) \rfloor} \binom{(k+4m)^d}{i} \\ & \leq |A|^{(k+4m)^d(e^{-1.3h^{top}(X)n^d} + 3^{d+1}m^{2d}\epsilon)} (k+4m)^d \binom{(k+4m)^d}{\lfloor (k+4m)^d(e^{-1.3h^{top}(X)n^d} + 3^{d+1}m^{2d}\epsilon) \rfloor} \\ & \leq e^{(k+4m)^d e^{-h^{top}(X)(n-2R)^d + 3^{d+1}m^{2d}\epsilon} \ln |A|} \end{aligned}$$

for large values of  $n$ .

The second upper bound, the one on  $|\beta_k^{-1}(v') \cap \alpha_k(A_{k+4m,\epsilon,3n}(X))|$  for any  $v' \in L_{\Gamma_k}(X)$ , is slightly more difficult. Consider any  $u \in \alpha_k(A_{k+4m,\epsilon,3n}(X))$ . Since  $u$  is in the image of  $\alpha_k$ ,  $u$  has no occurrences of any  $u_i$ . Consider any pair of overlapping occurrences of  $w$  in  $u$ , say at  $S$  and  $S'$  copies of  $\Gamma_n$ . Then these occurrences of  $w$  contain occurrences of  $w'$  as subwords, say at  $T$  and  $T'$  copies of  $\Gamma_m$  with  $T \subset S$ ,  $T' \subset S'$ , and  $T - T' = S - S'$ . Also define  $U$  to be the copy of  $\Gamma_{2R+l}$  in  $T$  which would be changed to change  $u(T)$  to  $w''$ , and  $U'$  to be the corresponding copy of  $\Gamma_{2R+l}$  in  $T'$ . Since there are no

occurrences of  $u_i$ , it must be the case that  $|T - T'|_\infty < \frac{m}{2} - 2n^{1-\frac{1}{3d}}$ . This implies that  $T$  contains the central  $\Gamma_{3(2R+l)\lceil n^{1-\frac{1}{2d}} \rceil}$  of  $S$  for large  $n$ , and so that  $T$  contains  $U'$ . Similarly,  $T'$  contains  $U$ . Therefore, if  $u(T)$  is replaced by  $w''$  sometime during the replacements defining  $\beta_k$ , then  $u(S')$  will be changed to something other than  $w$ , and if  $u(T')$  is replaced by  $w''$  at some point in these replacements, then  $u(S)$  will be changed to something other than  $w$ . In other words, in any replacement involved in changing  $u$  to  $\beta_k(u)$ , if  $u(S)$  is changed from  $w$  to  $\tilde{w}$ , then all occurrences of  $w$  with nonempty overlap with  $S$  are also changed. Since new occurrences of  $w$  cannot be created during these replacements, this implies that  $(\beta_k(u))(S) = \tilde{w}$  for any such  $S$ . Note that we have also shown that the copies of  $\Gamma_n$  where replacements occur in the application of  $\beta_k$  must be disjoint. We now wish to bound from above the number of occurrences of  $\tilde{w}$  in  $\beta_k(u)$  for any  $u \in \alpha_k(A_{k+4m,\epsilon,3n}(X))$ . By the definition of  $A_{k+4m,\epsilon,3n}(X)$ , for any  $u' \in A_{k+4m,\epsilon,3n}(X)$ ,  $u'$  has at most  $(k+4m)^d(e^{-h^{top}(X)(n-2R)^d} + \epsilon)$  occurrences of  $\tilde{w}$ . We now claim that each of the replacements involved in  $\alpha_k$  and  $\beta_k$  could create at most  $9^d$  new occurrences of  $\tilde{w}$ , which rests on an aperiodicity property of  $\tilde{w}$ : namely, for large  $n$ , and for any  $x \in X$  with  $x(S) = x(S') = \tilde{w}$  for  $S$  and  $S'$  copies of  $\Gamma_n$ ,  $|S - S'|_\infty > \frac{n}{3}$ .

To show this, recall that in the definition of  $w''$ , we ensure that  $w''$  contains a subword  $a \in L_{\Gamma_{2R+l}}(X)$  which is not a subword of  $w$ . In addition,  $a$  lies in the central  $\Gamma_{3(2R+l)\lceil n^{1-\frac{1}{2d}} \rceil}$  in  $w'$ . Also recall that  $w'$  is a subword of  $w$  whose shape has size at least  $n - n^{1-\frac{1}{3d}}$ . Together, these imply that for large  $n$ ,  $\tilde{w}$  contains  $a$  as a subword of its central  $\Gamma_{\frac{n}{4}}$ . Since  $\tilde{w}$  and  $w$  agree outside this central  $\Gamma_{\frac{n}{4}}$ , and since  $a$  is not a subword of  $w$ , this implies that any occurrences of  $a$  in  $\tilde{w}$  have nonempty intersection with its

central  $\Gamma_{\frac{n}{4}}$ , and that there is at least one such occurrence of  $a$ . Now consider  $x \in X$  such that  $x(S) = x(S') = \tilde{w}$ , where  $S \neq S'$ . Also consider  $T$  a copy of  $\Gamma_{2R+l}$  such that  $x(T) = a$ . There are two possibilities: either  $T$  lies inside  $S'$  or  $T$  does not lie inside  $S'$ . If  $T$  lies inside  $S'$ , then since  $x(T) = a$ ,  $T$  must have nonempty intersection with the central  $\Gamma_{\frac{n}{4}}$  of  $S'$ . Again since  $x(T) = a$ ,  $T$  must have nonempty intersection with the central  $\Gamma_{\frac{n}{4}}$  of  $S$  as well. This implies that  $|S - S'|_{\infty} \leq \frac{n}{4} + (2R + l)$ . However, then  $S - S'$  is a period of  $\tilde{w}$ . Since  $0 \neq |S - S'|_{\infty} \leq \frac{n}{4} + (2R + l) < \frac{n}{3}$  for large  $n$ , there is some positive integer  $k$  so that  $T + k(S - S')$  lies in  $S$ , but has empty intersection with the empty central  $\Gamma_{\frac{n}{4}}$  of  $S$ . But by periodicity of  $\tilde{w}$  with respect to  $S - S'$ ,  $x(T + k(S - S')) = x(T) = a$ , which gives a contradiction. Therefore, it must be the case that  $T$  does not lie inside  $S'$ . Since  $T$  has nonempty intersection with the central  $\Gamma_{\frac{n}{4}}$  of  $S$ , this implies that  $|S - S'|_{\infty} \geq \frac{3n}{8} - (2R + l) > \frac{n}{3}$  for large  $n$ .

Now consider any replacement of  $w'$  by  $w''$  in the replacements defining  $\alpha_k$  and  $\beta_k$ , and suppose that this replacement occurs at  $S$  a copy of  $\Gamma_m$ . Specifically, suppose that  $v, v' \in L_{\Gamma_{k+4m}}(X)$  where  $v(S) = w'$ ,  $v'(S) = w''$ , and  $v$  and  $v'$  agree outside of  $S$ . Also assume that  $T$  a copy of  $\Gamma_n$  is a location for an occurrence of  $\tilde{w}$  newly created in this replacement, i.e.  $v'(T) = \tilde{w}$ ,  $v(T) \neq \tilde{w}$ . Since  $v$  and  $v'$  agree outside  $S$ ,  $S$  and  $T$  must have nonempty intersection, meaning that any possible location for  $T$  given a fixed  $S$  lies in a copy of  $\Gamma_{m+2n}$  concentric with  $S$ . However, by the aperiodicity fact about  $\tilde{w}$  above, any  $T, T'$  both of which contain newly created occurrences of  $\tilde{w}$  must satisfy  $|T - T'|_{\infty} > \frac{n}{3}$ . This means that at each replacement, there are no more than  $9^d$  newly created occurrences of  $\tilde{w}$ . We now bound from above the total number of replacements performed in applying  $\alpha_k$  to  $u \in A_{k+4m, \epsilon, 3n}(X)$ . As already

shown, for large  $n$  the total number of occurrences of any  $u_i$  in  $u \in A_{k+4m,\epsilon,3n}(X)$  is smaller than  $3^d m^d (k+4m)^d (e^{-1.4h^{top}(X)n^d} + \epsilon)$ , and so the total number of occurrences of  $\tilde{w}$  created during these replacements is less than  $(27m)^d (k+4m)^d (e^{-1.4h^{top}(X)n^d} + \epsilon)$ , which is less than  $(k+4m)^d (e^{-h^{top}(X)(n-2R)^d} + (27m)^d \epsilon)$  for large  $n$ . We now bound from above the total number of replacements performed in applying  $\beta_k$  to  $\alpha_k(u)$ . For this, we need to bound from above the total number of occurrences of  $w$  in  $\alpha_k(u)$ .  $u$  itself had at most  $(k+4m)^d (e^{-h^{top}(X)(n-2R)^d} + \epsilon)$  occurrences of  $w$  by definition of  $A_{k+4m,\epsilon,3n}(X)$ . While changing  $u$  to  $\alpha_k(u)$ , less than  $3^d m^d (k+4m)^d (e^{-1.4h^{top}(X)n^d} + \epsilon)$  replacements are made. Each one of these could create at most  $(m+2n)^d$  new occurrences of  $w$  by the same reasoning we used for  $\tilde{w}$  (we just don't have any aperiodicity facts about  $w$  to use.) Therefore,  $\alpha_k(u)$  has less than  $(k+4m)^d (e^{-h^{top}(X)(n-2R)^d} + \epsilon) + (m+2n)^d 3^d m^d (k+4m)^d (e^{-1.4h^{top}(X)n^d} + \epsilon)$  occurrences of  $w$ , which is less than  $(k+4m)^d (2e^{-h^{top}(X)(n-2R)^d} + (12m)^d \epsilon)$  for large  $n$ . This means that the number of replacements involved in changing  $\alpha_k(u)$  to  $(\beta_k \circ \alpha_k)(u)$  is less than  $(k+4m)^d (2e^{-h^{top}(X)(n-2R)^d} + (12m)^d \epsilon)$  as well. Since each of these can create at most  $9^d$  new occurrences of  $\tilde{w}$ , the number of occurrences of  $\tilde{w}$  created during these replacements is less than  $(k+4m)^d (2 \cdot 9^d e^{-h^{top}(X)(n-2R)^d} + (108m)^d \epsilon)$ . So, the total number of occurrences of  $\tilde{w}$  created in the process of changing  $u$  to  $(\beta_k \circ \alpha_k)(u)$  is less than  $(k+4m)^d ((2 \cdot 9^d + 1)e^{-h^{top}(X)(n-2R)^d} + ((108m)^d + (27m)^d \epsilon))$ . Since  $u$  itself contained at most  $(k+4m)^d (e^{-h^{top}(X)(n-2R)^d} + \epsilon)$  occurrences of  $\tilde{w}$  by definition of  $A_{k+4m,\epsilon,3n}(X)$ , this means that  $(\beta_k \circ \alpha_k)(u)$  contains less than  $(k+4m)^d (20^d e^{-h^{top}(X)(n-2R)^d} + (216m)^d \epsilon)$  occurrences of  $\tilde{w}$ .

We now collect the three facts we need to bound  $|\beta_k^{-1}(v') \cap \alpha_k(A_{k+4m,\epsilon,3n}(X))|$  from



above for any  $v' \in L_{\Gamma_k}(X)$ . Firstly, for any  $u \in A_{k+4m,\epsilon,3n}(X)$ , we know that for any  $S$  a copy of  $\Gamma_n$  which is the location of a replacement during the process of changing  $\alpha_k(u)$  to  $(\beta_k \circ \alpha_k)(u)$ ,  $((\beta_k \circ \alpha_k)(u))(S) = \tilde{w}$ . In other words, once  $w$  is changed to  $\tilde{w}$  during the application of  $\beta_k$  to  $\alpha_k(u)$ , that occurrence of  $\tilde{w}$  will not be changed during future replacements. Secondly, all replacements performed in the application of  $\beta_k$  to  $\alpha_k(u)$  are disjoint. This means that knowing  $(\beta_k \circ \alpha_k)(u)$ , along with knowing the locations of all copies of  $\Gamma_n$  where replacements occur in changing  $\alpha_k(u)$  to  $(\beta_k \circ \alpha_k)(u)$ , is enough to uniquely determine  $\alpha_k(u)$ . Finally,  $(\beta_k \circ \alpha_k)(u)$  contains fewer than  $(k+4m)^d(20^d e^{-h^{top}(X)(n-2R)^d} + (216m)^d \epsilon)$  occurrences of  $\tilde{w}$ . This implies that for any  $v' \in L_{\Gamma_k}(X)$ ,  $|\beta_k^{-1}(v') \cap \alpha_k(A_{k+4m,\epsilon,3n}(X))| \leq 2^{(k+4m)^d(20^d e^{-h^{top}(X)(n-2R)^d} + (216m)^d \epsilon)}$ , the total number of subsets of the locations of occurrences of  $\tilde{w}$  in  $v'$  which could have been the locations of replacements in the application of  $\beta_k$ .

Finally, we give an upper bound on  $|\gamma_k^{-1}(v'')|$  for any  $v'' \in L_{\Gamma_k}(X)$ , which is straightforward: any  $u \in \gamma_k^{-1}(v'')$  must have its central  $\Gamma_k$  filled with  $v''$ . This means that  $|\gamma_k^{-1}(v'')| \leq |A|^{(k+4m)^d - k^d} < |A|^{5dmk^{d-1}}$  for large  $k$ .

Putting all of these upper bounds together, we see that  $|\phi_k^{-1}(v) \cap A_{k+4m,\epsilon,3n}(X)| <$

$$\begin{aligned} & e^{(k+4m)^d e^{-h^{top}(X)(n-2R)^d} + \ln |A| 3^{d+1} m^{2d\epsilon}} \cdot 2^{(k+4m)^d(20^d e^{-h^{top}(X)(n-2R)^d} + (216m)^d \epsilon)} \cdot |A|^{5dmk^{d-1}} \\ & \leq e^{(k+4m)^d C e^{-h^{top}(X)(n-2R)^d} + D\epsilon} \end{aligned}$$

for constants  $C$  independent of  $n$  and  $k$  and  $D$  independent of  $k$ . This implies that

$$H_{\Gamma_k}(X_w) \geq \frac{|A_{k+4m,\epsilon,3n}(X)|}{e^{(k+4m)^d C e^{-h^{top}(X)(n-2R)^d} + D\epsilon}}.$$

We then take natural logarithms of both sides, divide by  $(k + 4m)^d$ , and let  $k \rightarrow \infty$  to get

$$\lim_{k \rightarrow \infty} \frac{k^d}{(k + 4m)^d} \frac{h_{\Gamma_k}(X_w)}{k^d} \geq \lim_{k \rightarrow \infty} \frac{\ln |A_{k+4m, \epsilon, 3n}(X)|}{(k + 4m)^d} - C e^{-h^{\text{top}}(X)(n-2R)^d} - D\epsilon,$$

and by using the definition of entropy and Corollary 3.2.8, we are left with

$$h^{\text{top}}(X_w) \geq h^{\text{top}}(X) - C e^{-h^{\text{top}}(X)(n-2R)^d} - D\epsilon.$$

$\epsilon$  was arbitrary, so we may let it approach zero, leaving

$$h^{\text{top}}(X) - h^{\text{top}}(X_w) < \frac{C}{e^{h^{\text{top}}(X)(n-2R)^d}}.$$

A different tactic must be used to prove a lower bound for  $h^{\text{top}}(X) - h^{\text{top}}(X_w)$ . We in a sense proceed in the opposite way from our upper bound: we will define a map  $\psi_k$  which sends any word in  $L_{\Gamma_k}(X_w)$  to a subset of  $L_{\Gamma_k}(X)$  such that for any  $u \neq u' \in L_{\Gamma_k}(X_w)$ ,  $\psi_k(u)$  and  $\psi_k(u')$  are disjoint. This map will create occurrences of  $w$  instead of eliminating them.

We first need another auxiliary combinatorial theorem, again about a sort of lack of periodicity.

**Theorem 3.4.2.** *For any  $d > 1$  and any strongly irreducible  $\mathbb{Z}^d$ -shift  $X$  of finite type  $t$  with uniform filling length  $R$ , there exists a constant  $W$  dependent on  $X$  such that for all sufficiently large  $j$ , there is a word  $w_{j,d} \in L_{\Gamma_{jW}}^{(2W)}(X)$  with the following aperiodicity condition: there cannot exist two overlapping occurrences of  $w_{j,d}$  such that one has nonempty intersection with the empty central  $\Gamma_{(j-4)W}$  of the other.*

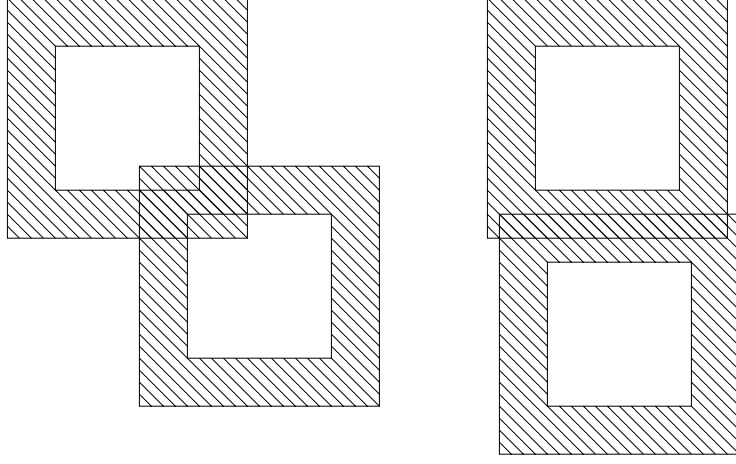


Figure 3.16: Disallowed and allowed pairs of overlapping  $w_{j,d}$

**Proof:** To make the statement of Theorem 3.4.2 more clear, we refer to Figure 3.16, which gives an example of a pair of overlapping occurrences of  $w_{j,d}$  which is ruled out by the given aperiodicity condition (left) and a pair which is not. (right)

We begin by giving a concrete example, for any fixed  $d$  and  $j \geq 3d+5$ , of an aperiodic word  $a_{j,d}$  in  $\{0, 1\}^{\Gamma_j}$ . For any  $p \in \Gamma_j$ , we define  $a_{j,d}(p) = 1$  if any  $p_i$  for  $1 \leq i \leq d-1$  is not equal to 1 or  $j$ . This leaves only  $a_{j,d}(p)$  where each of  $p_1, p_2, \dots, p_{d-1}$  are 1 or  $j$  to be defined. We think of this undefined portion as a set of  $2^{d-1}$  one-dimensional  $j$ -letter words: for every  $(p_1, p_2, \dots, p_{d-1}) \in \{1, j\}^{d-1}$ , we think of the yet-to-be-defined  $a_{j,d}(p_1, p_2, \dots, p_{d-1}, 1)$ ,  $a_{j,d}(p_1, p_2, \dots, p_{d-1}, 2), \dots$ ,  $a_{j,d}(p_1, p_2, \dots, p_{d-1}, j)$  as the letters of a  $j$ -letter word. Now, to fill these portions in, we define  $2^{d-1}$   $j$ -letter words, which we will call  $h_0, h_1, \dots, h_{2^{d-1}-1}$ , such that no two may overlap each other, i.e. if the final  $k$  letters of  $h_i$  are equal to the initial  $k$  letters

of  $h_{i'}$  for some  $k > 0$ , then  $k = j$  and  $i = i'$ . We do this in the following way: the first four letters of any  $h_i$  are defined to be 0001, and the final four letters are defined to be 0111. The  $3d - 3$  letters following the initial 0001 in any  $h_i$  are defined as follows: concatenate  $d - 1$  three-letter words, determined by  $i$ 's  $d - 1$ -digit binary expansion: each 0 in the binary expansion corresponds to the word 001, and each 1 in the binary expansion corresponds to 011. The remaining  $j - (3d + 5)$  letters of  $h_i$  which precede the final 0111 are alternating zeroes and ones, beginning with a zero.

An example should be helpful: suppose that  $d = 3$  and  $j = 18$ . Then we create four 18-letter words  $h_0, h_1, h_2$ , and  $h_3$ . The initial four letters of  $h_0$  are 0001. The next  $3d - 3 = 6$  letters are dependent on the two-digit binary expansion of the subscript 0: since it is 00, the next six letters are 001001. The next  $j - (3d + 5) = 4$  letters are alternating zeroes and ones beginning with a zero, i.e. 0101, and the final four letters are 0111. This gives  $h_0 = 000100100101010111$ . Using similar reasoning, we see that  $h_1 = 000100101101010111$ ,  $h_2 = 000101100101010111$ , and  $h_3 = 000101101101010111$ .

We claim that this set of  $2^{d-1}$  words has the nonoverlapping property described earlier. Suppose that for some  $k > 0$  and  $0 \leq i, i' < 2^{d-1}$ , the initial  $k$  letters of  $h_i$  are the same as the final  $k$  letters of  $h_{i'}$ .  $k$  cannot equal 1 or 2, because the first  $k$  letters of  $h_i$  would then be 0 or 00, and  $h_{i'}$  does not end with either of those words. So,  $k \geq 3$ . Then the initial  $k$  letters of  $h_i$  begin with 000, and since the only place that 000 occurs in  $h_{i'}$  is at the beginning, this implies that  $k = j$ , and since  $\{h_i\}_{i=0}^{2^{d-1}-1}$  contains  $2^{d-1}$  distinct words, that  $i = i'$ .

We now, for every choice of  $(p_1, p_2, \dots, p_{d-1}) \in \{1, j\}^{d-1}$ , choose an  $h_i$ , and define  $a_{j,d}(p_1, p_2, \dots, p_{d-1}, i') = h_i(i')$  for  $1 \leq i' \leq j$ . Assigning the  $2^{d-1}$   $h_i$  to the  $2^{d-1}$  choices for  $(p_1, p_2, \dots, p_{d-1})$  in a one-to-one way completes the definition of  $a_{j,d}$ . We now make the claim that  $a_{j,d}$  is aperiodic. Suppose not; then there exists  $v = (v_1, v_2, \dots, v_d) \neq 0$  with  $|v_i| < j$  for  $1 \leq i \leq d$  such that  $a_{j,d}(r) = a_{j,d}(r + v)$  for all  $r \in \Gamma_j$  such that  $r + v \in \Gamma_j$ . Since  $-v$  could serve as a period as well, we may assume without loss of generality that  $v_d \geq 0$ . We choose a vertex  $q$  of  $\Gamma_{j,d}$  based on  $v$ : for each  $1 \leq i \leq d$ , if  $v_i \geq 0$ ,  $q_i = 1$ . If  $v_i < 0$ ,  $q_i = j$ . In this way, we ensure that  $q + v \in \Gamma_j$ , and therefore that  $a_{j,d}(q) = a_{j,d}(q + v)$ . Since  $q_i \in \{1, j\}$  for  $1 \leq i \leq d-1$  and  $q_d = 1$ , by construction  $a_{j,d}(q) = 0$ . Therefore,  $a_{j,d}(q + v) = 0$ . However, again due to construction, the only zeroes in  $a_{j,d}$  lie at points whose first  $d-1$  coordinates are either 1 or  $j$ . This implies that the first  $d-1$  coordinates of  $q + v$  are either 1 or  $j$ . Let's denote by  $h_i$  the  $j$ -letter word where  $h_i(k) = a_{j,d}(q + (k-1)e_d)$  for  $1 \leq k \leq j$ , and by  $h_{i'}$  the  $j$ -letter word where  $h_{i'}(k) = a_{j,d}(q + v + (k-1-v_d)e_d)$ . Due to the supposed periodicity of  $a_{j,d}$ , the first  $j - v_d$  letters of  $h_i$  are the same as the final  $j - v_d$  letters of  $h_{i'}$ . Since no two distinct words in  $\{h_n\}_{n=1}^{2^{d-1}}$  may overlap,  $v_d = 0$  and  $h_i = h_{i'}$ . This implies that the first  $d-1$  coordinates of  $q$  are the same as the first  $d-1$  coordinates of  $q + v$ , and therefore that the first  $d-1$  coordinates of  $v$  are zero, implying that  $v = 0$ , a contradiction. Thus,  $a_{j,d}$  is aperiodic.

We now use  $a_{j,d-1}$  as a tool to create, for any  $j \geq 3d+2$ , a word  $b_{j,d} \in \{0, 1\}^{\Gamma_j^{(2)}}$  with the aperiodicity property outlined in the conclusion of Theorem 3.4.2: there cannot exist two overlapping occurrences of  $b_{j,d}$  such that one has nonempty intersection with the empty central  $\Gamma_{j-4}$  of the other.  $b_{j,d}(p)$  is defined to be 0 if  $p_d = 1$  or 2,

and defined to be 1 if  $3 \leq p_d \leq j - 1$ . If  $p_d = j$ , then  $b_{j,d}(v) = a_{j,d-1}(p_1, \dots, p_{d-1})$ . Suppose that  $b_{j,d}$  is periodic with respect to  $v \neq 0$ . Since our supposed overlap is of the form described, it must be the case that  $|v_i| \leq j - 2$  for  $1 \leq i \leq d$ . We can again assume without loss of generality that  $v_d \geq 0$ . Choose a point  $q$  of  $\Gamma_j^{(2)}$  as follows: for each  $1 \leq i \leq d - 2$ , take  $q_i = 1$  if  $v_i \geq 0$ , and  $q_i = j - 1$  if  $v_i < 0$ . Choose  $q_{d-1} = j - 1 - v_{d-1}$  if  $v_{d-1} \geq 0$ , and  $q_{d-1} = 1 - v_{d-1}$  if  $v_{d-1} < 0$ . Finally, take  $q_d = 1$ . For any  $r \in \mathbb{Z}^d$  all of whose coordinates are zero or one, the first  $d - 1$  coordinates of  $q + r$  are between 1 and  $j$ , and the  $d$ th coordinate is 1 or 2, implying that  $q + r \in \Gamma_j^{(2)}$ . This means that a copy of  $\Gamma_2$  with its least vertex lexicographically at  $q$  is a subset of  $\Gamma_j^{(2)}$ , call it  $S$ . Since  $S$  is composed entirely of points whose  $d$ th coordinate is 1 or 2,  $b_{j,d}(S)$  is filled with zeroes. Again, for any point  $r$  all of whose coordinates are zero or one, the first  $d - 2$  coordinates of  $q + v + r$  are between 1 and  $j$ , the  $d - 1$ th coordinate is 1, 2,  $j - 1$ , or  $j$ , and the  $d$ th coordinate is between 1 and  $j$ . Therefore, any such  $q + v + r$  is in  $\Gamma_j^{(2)}$ , implying that a copy of  $\Gamma_2$  with its least vertex lexicographically at  $q + v$  is a subset of  $\Gamma_j^{(2)}$  as well, which is then  $S + v$ . By periodicity,  $b_{j,d}(S + v)$  must be filled with zeroes as well. However, by construction, for any copy of  $\Gamma_2$  which is filled with zeroes in  $b_{j,d}$ , the  $d$ th coordinate of its least vertex lexicographically is 1. Since  $q_d = 1$ , it must be the case that  $v_d = 0$ . This implies that  $a_{j,d-1}$  is periodic with period  $(v_1, v_2, \dots, v_{d-1})$ , which implies that  $v_i = 0$  for  $1 \leq i \leq d - 1$  as well, and so  $v = 0$ , a contradiction. Therefore,  $b_{j,d}$  has the desired aperiodicity property. Note, however, that this does not prove Theorem 3.4.2, since  $b_{j,d}$  is a word in the full shift, and there is no reason for it to be in  $L(X)$ .

We now turn to our shift of finite type  $X$ . First, we will be constructing a word

$y \in L_{\Gamma_{R+4}}(X)$  for which  $X_y$  is nonempty. Obviously, finding a word  $y$  such that  $X_y$  is nonempty could be done by using the already proven upper bound for  $h^{top}(X) - h^{top}(X_y)$ , but this may require  $y$  to have shape with very large size. We need the shape of  $y$  to have a small size to prove as tight a lower bound as possible, and for our proof, we need a lower bound on  $H_{\Gamma_k}(X)$  for relatively small  $k$ . For large  $k$ , Lemma 3.2.1 gives an exponential lower bound for  $H_{\Gamma_k}(X)$ , but doesn't really tell us anything about small  $k$ . For this reason, we prove the following lemma:

**Lemma 3.4.3.** *For any  $\mathbb{Z}^d$ -subshift  $X$  with  $h^{top}(X) > 0$ , and finite  $S \subset \mathbb{Z}^d$ ,  $H_S(X) > |S|$ .*

**Proof:** We write  $S = \{s_1, s_2, \dots, s_{|S|}\}$  where  $s_i$  comes before  $s_{i+1}$  in the usual lexicographical order for  $1 \leq i < |S|$ , and make the notation  $S_i = \{s_1, \dots, s_i\}$  for all  $1 \leq i \leq |S|$ . Suppose that  $H_S(X) \leq |S|$ . Then, since  $H_{S_1}(X) = |A| > 1$ , it must be the case that for some  $1 \leq j < |S|$ ,  $H_{S_{j+1}}(X) \leq H_{S_j}(X)$ . Since  $S_j \subset S_{j+1}$ , this means that for every word  $w \in L_{S_j}(X)$ , there is a unique way to extend it to a word in  $L_{S_{j+1}}(X)$ . In other words, for any  $x \in X$ , given  $x(S_j)$ ,  $x(s_{j+1})$  is forced. By shift-invariance of  $X$ , for any  $x \in X$ , given  $x(S_j - s_{j+1})$ ,  $x(\mathbf{0})$  is forced. Since  $S$  is finite, take  $N > \text{diam}(S)$ . Then,  $S_j - s_{j+1}$  consists of elements of  $\mathbb{Z}^d$  within a  $d$ -distance of less than  $N$  from  $\mathbf{0}$  and lexicographically less than  $\mathbf{0}$ . For this reason, we make the notation  $H_{n,d} = \{v \in (\Gamma_{2n-1} - \mathbf{n}) : v \text{ is less than } \mathbf{0} \text{ lexicographically}\}$ , where  $\mathbf{n} = (n, n, \dots, n)$ . It is then clear that  $S_j - s_{j+1} \subseteq H_{N,d}$ , and so we note that  $x(H_{N,d})$  forces  $x(\mathbf{0})$ . We need a few more pieces of notation;  $I_{n,d} = \{v \in (\Gamma_{2n-1} - \mathbf{n}) : v_d < 0\}$ ,  $G_{n,k,d} = \{v \in \mathbb{Z}^d : v_i > 0 \text{ for } 1 \leq i \leq d, \sum_{i=1}^d n^{i-1}v_i \leq k\}$ , and  $L_{n,k,d} = (\Gamma_{k+n} -$

$\mathbf{n}) \setminus \Gamma_k$ . We quickly note a few useful facts about these sets which can be checked by the reader:  $H_{n,d+1} = I_{n,d} \sqcup (H_{n,d} \times \{0\})$ , and  $G_{n,k,d+1} = \bigsqcup_{i=1}^{\lfloor \frac{k}{n^d} \rfloor} (G_{n,k-in^d,d} \times \{i\})$ . We now claim that if  $x(H_{n,d})$  forces  $x(\mathbf{0})$  for all  $x \in X$ , then  $x(L_{n,k,d})$  forces  $x(G_{n,k,d})$  for any  $k$ . We prove this claim by induction on  $d$ .

For  $d = 1$ , the claim is fairly easy to check (and is in fact a classical theorem due to Hedlund and Morse ([HeM])); it amounts to showing that if any  $n$  consecutive letters of any  $x$  force the next, then any  $n$  consecutive letters of  $x$  force the next  $k$  for any  $k$ . But this is clear; if  $x(1), x(2), \dots, x(n)$  force  $x(n+1)$ , then  $x(2), \dots, x(n+1)$  force  $x(n+2)$ , and we can proceed like this indefinitely. Thus, the claim is proven for  $d = 1$ . Now suppose that it is true for a fixed  $d$ , and we will prove it for  $d + 1$ . This proof will also be by induction; since we wish to show that  $x(L_{n,k,d+1})$  forces  $x(G_{n,k,d+1})$ , and since  $G_{n,k,d+1} = \bigsqcup_{i=1}^{\lfloor \frac{k}{n^d} \rfloor} (G_{n,k-in^d,d} \times \{i\})$ , it suffices to show that  $x(L_{n,k,d+1} \cup \bigsqcup_{i=1}^j (G_{n,k-in^d,d} \times \{i\}))$  forces  $x(G_{n,k-(j+1)n^d,d} \times \{j+1\})$  for every  $0 \leq j < \lfloor \frac{k}{n^d} \rfloor$ . Fix any such  $j$ , and suppose that  $x(L_{n,k,d+1} \cup \bigsqcup_{i=1}^j (G_{n,k-in^d,d} \times \{i\}))$  is given. We will show that  $x(G_{n,k-(j+1)n^d,d} \times \{j+1\})$  is forced. Consider any  $v = (v', j+1) \in G_{n,k-(j+1)n^d,d} \times \{j+1\}$ . We claim that  $v + I_{n,d+1} \subseteq L_{n,k,d+1} \cup \bigsqcup_{i=1}^j (G_{n,k-in^d,d} \times \{j\})$ . Since  $I_{n,d+1} = (\Gamma_{2n-1} - \mathbf{n}) \times \{-n+1, \dots, 0\}$ , it suffices to show that  $v' + (\Gamma_{2n-1} - \mathbf{n}) \subseteq (L_{n,k,d+1} \cup \bigsqcup_{i=1}^j (G_{n,k-in^d,d} \times \{j\})) \cap (\mathbb{Z}^d \times \{i\})$  for  $j-n < i \leq j$ . Since  $(L_{n,k,d+1} \cup \bigsqcup_{i=1}^j (G_{n,k-in^d,d} \times \{j\})) \cap (\mathbb{Z}^d \times \{i\}) = (L_{n,k,d} \cup \Gamma_k) \times \{i\}$  for  $-n+1 \leq i \leq 0$ , and  $(L_{n,k,d+1} \cup \bigsqcup_{i=1}^j (G_{n,k-in^d,d} \times \{j\})) \cap (\mathbb{Z}^d \times \{i\}) = (L_{n,k,d} \cup G_{n,k-in^d,d}) \times \{i\}$  for  $i > 0$ , and since  $G_{n,k-(i+1)n^d,d} \subseteq G_{n,k-in^d,d} \subseteq \Gamma_k$  for  $i > 0$ , it suffices to show that  $G_{n,k-(i+1)n^d,d} + (\Gamma_{2n-1} - \mathbf{n}) \subseteq L_{n,k,d} \cup G_{n,k-in^d,d}$  for  $i > 0$ . Consider any  $v \in G_{n,k-(i+1)n^d,d}$  and any  $v' \in (\Gamma_{2n-1} - \mathbf{n})$ . By definition of



$G_{n,k-(i+1)n^d,d}$ ,  $\sum_{i=1}^d v_i n^{i-1} \leq k - (i+1)n^d$ , and  $v_i > 0$  for  $1 \leq i \leq d$ . By choice of  $v'$ ,  $-n+1 \leq v'_i \leq n-1$  for  $1 \leq i \leq d$ . Therefore,  $-n+1 < (v+v')_i$  for  $1 \leq i \leq d$ , and  $\sum_{i=1}^d (v+v')_i n^{i-1} \leq (k-(i+1)n^d) + (\sum_{i=0}^{d-1} n^i) < k - in^d$ . There are then two cases; if any coordinate of  $v+v'$  is nonpositive, then  $v+v' \in L_{n,k,d}$ , and if all coordinates are positive, then by definition,  $v+v' \in G_{n,k-in^d,d}$ . So, indeed  $G_{n,k-(i+1)n^d,d} + (\Gamma_{2n-1} - \mathbf{n}) \subseteq L_{n,k,d} \sqcup G_{n,k-in^d,d}$ , and so for any  $v \in G_{n,k-(j+1)n^d,d} \times \{j+1\}$ ,  $v + I_{n,d+1} \subseteq L_{n,k,d+1} \cup \bigsqcup_{i=1}^j (G_{n,k-in^d,d} \times \{i\})$ . This means that since  $H_{n,d+1} = I_{n,d+1} \sqcup (H_{n,d} \times \{0\})$ , for any  $v \in (G_{n,k-(j+1)n^d,d} \times \{j+1\})$ ,  $x(v + (H_{n,d} \times \{0\}))$  forces  $x(v)$ . But now since  $L_{n,k-(j+1)n^d,d} \times \{j+1\} \subset L_{n,k,d+1}$ , by the inductive hypothesis  $x(G_{n,k-(j+1)n^d,d})$  is in fact forced by  $x(L_{n,k,d+1} \cup \bigsqcup_{i=1}^j (G_{n,k-in^d,d} \times \{i\}))$ . Since  $j$  was arbitrary here, as described above this shows that  $x(L_{n,k,d+1})$  forces  $x(G_{n,k,d+1})$  as claimed, and so by induction we see that for any  $d$ , if  $x(H_{n,d})$  forces  $x(\mathbf{0})$  for all  $x \in X$ , then  $x(L_{n,k,d})$  forces  $x(G_{n,k,d})$  for any  $k$ .

We finally return to our original example of  $X$  for which  $H_S(X) \leq S$ . Recall that we showed that there is then some  $N$  such that  $x(H_{N,d})$  forces  $x(\mathbf{0})$  for all  $x$ . By the claim just shown,  $x(L_{N,k,d})$  then forces  $x(G_{N,k,d})$  for all  $k$ . This then shows that  $H_{G_{N,k,d}}(X) \leq H_{L_{N,k,d}}(X)$  for any  $k$ . Note that  $\Gamma_{\lfloor \frac{k}{dN^d} \rfloor} \subseteq G_{N,k,d}$  for all  $k$ . Therefore,  $H_{\Gamma_n}(X) \leq H_{G_{N,ndN^d,d}}(X) \leq H_{L_{N,ndN^d,d}}(X)$  for all  $n$ . Since  $|L_{N,ndN^d,d}| = (ndN^d + N)^d - (ndN^d)^d \leq 2Nd(ndN^d)^{d-1} = Cn^{d-1}$  for large  $n$  and a constant  $C$  independent of  $n$ , we see that  $H_{\Gamma_n}(X) \leq |A|^{Cn^{d-1}}$  for all  $n$ , and so by definition of entropy,  $h^{top}(X) = \lim_{n \rightarrow \infty} \frac{\ln H_{\Gamma_n}(X)}{n^d} \leq \limsup_{n \rightarrow \infty} \frac{Cn^{d-1} \ln |A|}{n^d} = 0$ . Therefore,  $h^{top}(X) = 0$ , a contradiction to the hypotheses. Our initial assumption was therefore wrong, and so  $H_S(X) > |S|$  for all finite shapes  $S$ .

□

Before moving on with our proof, we must comment on Lemma 3.4.3. In the course of the proof, we show that if for some total order  $\leq$  on  $\mathbb{Z}^d$  which is preserved under addition, some finite set  $S \subset \mathbb{Z}^d$ , and some  $t \notin S$  with  $t \geq s$  for all  $s \in S$ ,  $x(S)$  forces  $x(t)$ , then the topological entropy of  $X$  is zero. It is natural to wonder whether the assumption about ordering is necessary. In fact it is:

**Lemma 3.4.4.** *For any finite  $S \subset \mathbb{Z}^d$  and  $t \notin S$  such that there is no addition-preserved total order on  $\mathbb{Z}^d$  with respect to which  $t$  is greater than all elements of  $S$ , there exists a subshift  $X$  such that  $x(S)$  forces  $x(t)$  for all  $x \in X$ , but  $h^{\text{top}}(X) > 0$ .*

**Proof:** Suppose that for some  $S \subset \mathbb{Z}^d$  finite and  $t \notin S$ , it is the case that for no addition-preserved total order on  $\mathbb{Z}^d$  is  $t$  greater than all elements of  $S$ . We claim that this implies that  $t$  lies in the convex hull of  $S$ . To see this, suppose that  $t$  does not lie in the convex hull of  $S$ . Then clearly there is a  $d - 1$ -dimensional hyperplane  $L$  of  $\mathbb{R}^d$  such that  $t$  and  $S$  lie on opposite sides of  $L$ . Define the subspace  $L'$  of  $\mathbb{R}^d$  which equals  $L - \ell$  for any  $\ell \in L$ . Without loss of generality, we may assume that  $L \cap \mathbb{Z}^d = \emptyset$ , since such an assumption requires only a small rotation of  $L$ . Then, define an addition-preserved total order  $\leq$  on  $\mathbb{Z}^d$  by arbitrarily fixing one of the open half-spaces  $H$  that  $L'$  splits  $\mathbb{R}^d$  into and defining  $v \leq v'$  if  $v' - v \in H \sqcup \{\mathbf{0}\}$ . Then, since  $t$  and  $S$  are on opposite sides of  $L$ , for all  $s \in S$ ,  $t - s$  is on the same side of  $L'$ , and so either  $t \leq s$  for all  $s \in S$  or  $t \geq s$  for all  $s \in S$ . By reversing  $\leq$  if necessary, we may assume without loss of generality that  $t$  is greater than all elements of  $S$  under  $\leq$ . This is a contradiction, and so  $t$  lies in the convex hull of  $S$ .

To make things easier, by shift-invariance, we can without loss of generality assume that  $t = \mathbf{0} := (0, \dots, 0)$ . We will, for any  $S$  such that  $\mathbf{0}$  lies in the convex hull of  $S$ , construct a subshift  $X$  with positive topological entropy such that  $x(S)$  forces  $x(\mathbf{0})$  for any  $x \in X$ . By decomposing the convex hull of  $S$  into simplices, we can, for some  $1 \leq d' \leq d$ , find  $S'$  a linearly independent subset of  $S$  of cardinality  $d' + 1$  such that the convex hull of  $S'$  contains  $\mathbf{0}$  and is contained in a  $d'$ -dimensional subspace of  $\mathbb{R}^d$ , call it  $L$ . Denote the elements of  $S'$  by  $s_1, s_2, \dots, s_{d'+1}$ . Then there exist positive reals  $n_1, \dots, n_{d'+1}$  such that  $\mathbf{0} = \sum_{i=1}^{d'+1} n_i s_i$ . Since all elements of  $S$  are in  $\mathbb{Z}^d$ , we can take all  $n_i$  to be rational, and so without loss of generality integers. This implies that  $s_{d'+1} = \sum_{i=1}^{d'} -\frac{n_i}{n_{d'+1}} s_i$ . Take  $N > \frac{n_i}{n_{d'+1}} + 1$  for all  $1 \leq i \leq d'$ . Take the  $d'$ -dimensional parallelepiped  $R = \{\sum_{i=1}^{d'} r_i s_i : 0 \leq r_i < N\}$  contained in  $L$ . One may then tessellate  $L$  with copies of  $R$ :  $L = \bigsqcup_{v \in \mathbb{Z}^d} (R + \sum_{i=1}^{d'} N v_i s_i)$ . (Intersecting with  $\mathbb{Z}^d$  on both sides leads to a similar tessellation of  $L \cap \mathbb{Z}^d$  by copies of  $R \cap \mathbb{Z}^d$ .) If  $L \subsetneq \mathbb{R}^d$ , then choose linearly independent  $u_1, \dots, u_{d-d'}$  in  $\mathbb{Z}^d$  such that  $(L \cap \mathbb{Z}^d) + \sum_{j=1}^{d-d'} (\mathbb{Z} u_j) = \mathbb{Z}^d$ . We will define  $X$ , a subshift with alphabet  $A = \{0, 1\} \times (R \cap \mathbb{Z}^d)$ , such that  $x(S')$  forces  $x(\mathbf{0})$  for any  $x \in X$ , which clearly implies that  $x(S)$  forces  $x(\mathbf{0})$  for any  $x \in X$  as well. Choose any  $y \in \Omega$ . We define  $x \in X$  as follows: for any  $v \in \mathbb{Z}^d$  and  $r \in (R \cap \mathbb{Z}^d)$ , define  $x(r + \sum_{i=1}^{d'} N v_i s_i + \sum_{j=1}^{d-d'} v_{j+d'} u_j) = (y(v), r)$ . Due to the already mentioned tessellation of  $L \cap \mathbb{Z}^d$  by copies of  $R \cap \mathbb{Z}^d$  and tessellation of  $\mathbb{Z}^d$  by copies of  $L \cap \mathbb{Z}^d$ , this defines  $x$  on all of  $\mathbb{Z}^d$ . Take  $X'$  to be the set of all elements of  $\Omega$  constructed in this way.  $X'$  is not necessarily shift-invariant, but by the construction, it is invariant under shifts by  $N s_i$  for  $1 \leq i \leq d'$  and  $u_j$  for  $1 \leq j \leq d - d'$ . Since the collection  $\{N s_1, N s_2, \dots, N s_{d'}, u_1, \dots, u_{d-d'}\}$  is linearly independent, this means that a finite

union of shifts of  $X'$ , call it  $X$ , is a shift-invariant set. It is also clear that  $X$  is closed, and thus it is a subshift.

We first show that  $x(S')$  forces  $x(\mathbf{0})$  for any  $x \in X$ . By the definition of  $X$ , it is sufficient to show that  $x'(S' + p)$  forces  $x'(p)$  for any  $x' \in X'$  and  $p \in \mathbb{Z}^d$ . Since  $p \in \mathbb{Z}^d$ , there exists some  $r \in (R \cap \mathbb{Z}^d)$  and  $v \in \mathbb{Z}^d$  such that  $p = r + \sum_{i=1}^{d'} N v_i s_i + \sum_{j=1}^{d-d'} v_{j+d'} u_j$ . Since  $r \in R$ ,  $r = \sum_{i=1}^{d'} r_i s_i$  for some positive reals  $r_i$ . For any  $1 \leq i \leq d'$ ,  $r + s_i$  is in  $R$  unless the  $s_i$  coefficient of  $r + s_i$  is greater than or equal to  $N$ . Therefore, if  $r + s_i \notin R$  for all  $1 \leq i \leq d'$ , then it must be the case that  $r_i \geq N - 1$  for all  $1 \leq i \leq d'$ . But then since  $s_{d'+1} = \sum_{i=1}^{d'} -\frac{n_i}{n_{d'+1}} s_i$  as noted earlier, and since  $\frac{n_i}{n_{d'+1}} < N - 1$  for all  $1 \leq i \leq d'$ ,  $r + s_{d'+1} = \sum_{i=1}^{d'} (r_i - \frac{n_i}{n_{d'+1}}) s_i$ , with all coefficients in  $[0, N)$ . This implies that  $r + s_{d'+1} \in R$ . We then showed that if  $r + s_i \notin R$  for all  $1 \leq i \leq d'$ , then  $r + s_{d'+1} \in R$ . In other words, there is some  $1 \leq j \leq d' + 1$  such that  $r + s_j \in R$ . Fix this  $j$ . Then,  $p + s_j = (r + s_j) + \sum_{i=1}^{d'} N v_i s_i + \sum_{j=1}^{d-d'} v_{j+d'} u_j$ . Say that  $x'(p + s_j) = (r + s_j, q)$  for some  $m \in \{0, 1\}$ . Then by definition of  $X'$ , since  $p$  and  $p + s_j$  lie in the same copy of  $R$ , it must be the case that  $x'(p) = (r, q)$ . This means that  $x'(p + s_j)$  forced  $x'(p)$ . Since  $1 \leq j \leq d' + 1$ , this shows that  $x'(S' + p)$  forces  $x'(p)$  for any  $p \in \mathbb{Z}^d$ , and so that  $x(S')$  forces  $x(\mathbf{0})$  for any  $x \in X$ .

It remains to show that  $h^{top}(X) > 0$ . Choose  $M > |N s_i|_\infty, |u_j|_\infty$  for  $1 \leq i \leq d'$ ,  $1 \leq j \leq d - d'$ . Then,  $R + \{\sum_{i=1}^{d'} N k_i s_i + \sum_{j=1}^{d-d'} k_{j+d'} u_j : |k_i| \leq K\} \subseteq [-(K + 1)Md, (K + 1)Md]^d$  for any  $K$ . By the definition of  $X$ , there are at least  $2^{|[-K, K]^d|}$  ways to fill letters of  $x$  on  $R + \{\sum_{i=1}^{d'} N k_i s_i + \sum_{j=1}^{d-d'} k_{j+d'} u_j : |k_i| \leq K\}$ , and so we see that  $H_{[-(K+1)Md, (K+1)Md]^d}(X) = H_{\Gamma_{2(K+1)Md+1}}(X) \geq 2^{(2K+1)^d}$ . This means that

$$\frac{h_{\Gamma_{2(K+1)Md+1}}(X)}{(2(K+1)Md+1)^d} \geq \frac{(\ln 2)(2K+1)^d}{(2(K+1)Md+1)^d} \geq \frac{1}{(4Md)^d}$$

for all  $K$ , and therefore, by letting  $K$  approach infinity, we see that  $h^{top}(X) \geq \frac{1}{(4Md)^d} > 0$ .

□

We claim that for a strongly irreducible shift of finite type  $X$ ,  $H_{\Gamma_{R+4}}(X) > (2R+4)^d$ . We define  $S_1 = \{1, 2\} \times \{1, 2, \dots, R+4\}^{d-1} \subset \Gamma_{R+4}$  and  $S_2 = \{R+3, R+4\} \times \{1, 2, \dots, R+4\}^{d-1} \subset \Gamma_{R+4}$ . Then  $d(S_1, S_2) > R$ , and so by strong irreducibility  $H_{\Gamma_{R+4}}(X) \geq H_{S_1}(X)H_{S_2}(X)$ , which is greater than  $4(R+4)^{2d-2}$  by Lemma 3.4.3. Since  $R+4 \geq 4$ ,

$$4(R+4)^{2d-2} \geq 4^{d-1}(R+4)^d \geq 2^d(R+4)^d = (2R+8)^d > (2R+4)^d.$$

Therefore,  $H_{\Gamma_{R+4}}(X) > (2R+4)^d$ . Consider any  $y \in L_{\Gamma_{3R+7}}(X)$ . Then, since there are  $(2R+4)^d$  copies of  $\Gamma_{R+4}$  contained in  $\Gamma_{3R+7}$ , there are at most  $(2R+4)^d$  different words in  $L_{\Gamma_{R+4}}(X)$  which are subwords of  $y$ . Since  $H_{\Gamma_{R+4}}(X) > (2R+4)^d$ , this means that there is  $z \in L_{\Gamma_{R+4}}(X)$  such that  $z$  is not a subword of  $y$ . Then, again using strong irreducibility, we create  $x \in X$  such that for any  $v \in \mathbb{Z}^d$ ,  $x(\Gamma_{R+4} + (2R+4)v) = z$ . (See Figure 3.17.) Then, for any  $S$  a copy of  $\Gamma_{3R+7}$ ,  $x(S)$  contains a  $z$ , and therefore  $x(S) \neq y$ . Thus,  $y$  is not a subword of  $x$ , and so  $X_y$  contains at least one point and is nonempty.

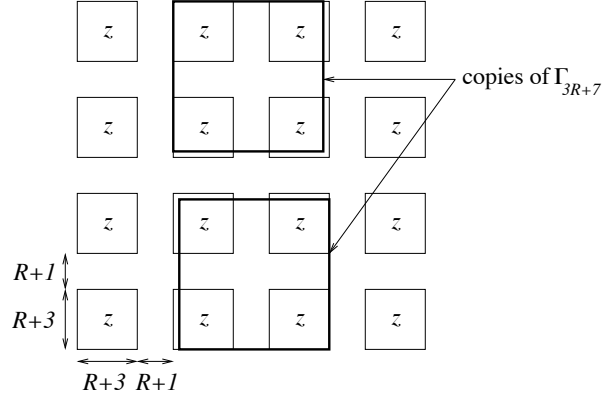


Figure 3.17: A point  $x \in X_y$

We will use  $y$  and  $b_{j,d}$  to create, for  $W = 8R + 14$ , and for any  $j \geq 3d + 2$ , a word  $w_{j,d} \in L_{\Gamma_{jW}^{(2W)}}(X)$  such that there cannot exist two overlapping occurrences of  $w_{j,d}$  where one has nonempty intersection with the empty central  $\Gamma_{(j-4)W}$  of the other. To do this, we first partition  $\Gamma_{jW}^{(2W)}$  into disjoint copies of  $\Gamma_W$ . The disjoint copies of  $\Gamma_W$  then have an obvious bijective correspondence to the points of  $\Gamma_j^{(2)}$ , illustrated in Figure 3.18.

We then use each entry of  $b_{j,d}$  to assign entries of  $w_{j,d}$  to the corresponding  $\Gamma_W$  in  $\Gamma_{jW}^{(2W)}$ . For  $p \in \Gamma_j^{(2)}$ , if  $b_{j,d}(p) = 0$ , then the least copy lexicographically of  $\Gamma_{W-R}$  in the  $\Gamma_W$  corresponding to  $p$  is filled with any word in  $L_{\Gamma_{W-R}}(X_y)$ , i.e. a word without any occurrences of  $y$ . If  $b_{j,d}(p) = 1$ , then the least copy lexicographically of  $\Gamma_{W-R}$  in the  $\Gamma_W$  corresponding to  $p$  has  $2^d$  occurrences of  $y$  placed inside it, each one sharing a vertex with it.

The remainder of  $\Gamma_{jW}^{(2W)}$  is then filled to make a word in  $L(X)$  by using strong irreducibility, since all of the filled portions are a distance of at least  $R + 1$  from each other. We claim that this word  $w_{j,d} \in L_{\Gamma_{jW}^{(2W)}}(X)$  has the desired aperiodicity condition. Suppose not; then there exist two overlapping occurrences of  $w_{j,d}$  such that one has nonempty overlap with the empty central  $\Gamma_{(j-4)W}$  of the other, which implies that  $w_{j,d}$  is periodic with respect to some  $v \neq 0$  with  $|v_i| < (j - 2)W$  for  $1 \leq i \leq d$ . We then define  $v'$  by defining  $v'_i$  to be the closest multiple of  $W$  to  $v_i$  for  $1 \leq i \leq d$ . If two are equally close, choose either. In this way, we ensure that  $|v_i - v'_i| \leq \frac{W}{2}$  for each  $1 \leq i \leq d$ . We make the notation  $v'' := v - v'$ . Since each coordinate of  $v'$  is divisible by  $W$ ,  $\frac{v'}{W}$  has integer coordinates.

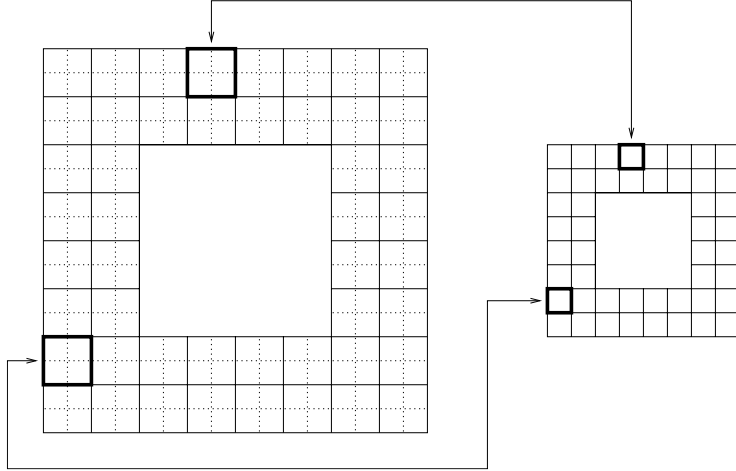


Figure 3.18: The correspondence between copies of  $\Gamma_W$  and points in  $\Gamma_j^{(2)}$

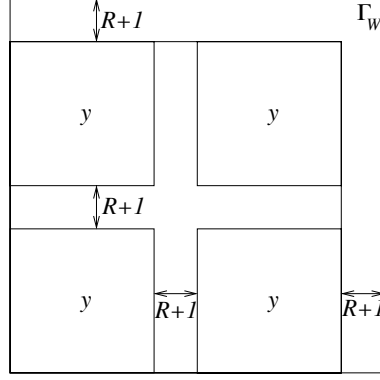


Figure 3.19: How a copy of  $\Gamma_W$  is filled if  $b_{j,d}(p) = 1$

Since  $|v'_i| \leq (j-2)W$  for all  $i$ , every coordinate of  $\frac{v'}{W}$  has absolute value at most  $j-2$ . This implies that either  $v' = 0$  or  $\frac{v'}{W}$  is the difference between two overlapping occurrences of  $\Gamma_j^{(2)}$ , one of which has nonempty intersection with the empty central  $\Gamma_{j-4}$  of the other. Assume that the latter is the case. Then, by the already proven aperiodicity condition on  $b_{n,d}$ , this implies that there exist  $q, q + \frac{v'}{W} \in \Gamma_j^{(2)}$  such that  $b_{j,d}(q) \neq b_{j,d}(q + \frac{v'}{W})$ . Without loss of generality, we assume that  $b_{j,d}(q) = 1$  and  $b_{j,d}(q + \frac{v'}{W}) = 0$ .

Let's call  $S$  the central  $\Gamma_{W-2R}$  in the  $\Gamma_W$  in  $\Gamma_{jW}^{(2W)}$  corresponding to  $q$  and  $T$  the central  $\Gamma_{W-2R}$  in the  $\Gamma_W$  in  $\Gamma_{jW}^{(2W)}$  corresponding to  $q + \frac{v'}{W}$ . Then  $T - S = v'$ , and  $w_{j,d}$  is periodic with respect to  $v$ ,  $w_{j,d}(S \cap (T - v)) = w_{j,d}((S + v) \cap T)$ .  $T - v = (T - v') - v'' = S - v''$ . Since every coordinate of  $v''$  is at most  $\frac{W}{2}$ ,  $S \cap (S - v'')$ , and hence  $S \cap (T - v)$ , is nonempty. In fact, it is a rectangular prism  $R_{s_1, \dots, s_d}$  where  $s_i \geq (W - 2R) - \frac{W}{2} = 3R + 7$ ,  $1 \leq i \leq d$ . Therefore,  $S \cap (T - v)$  contains one of



the  $2^d$  copies of  $\Gamma_{3R+7}$  in  $S$  which share a vertex with  $S$ . Since  $b_{j,d}(q) = 1$ , and since  $S$  is the central  $\Gamma_{W-2R}$  in the  $\Gamma_W$  corresponding to  $q$ , every one of these subcubes of  $S$  contains an occurrence of  $y$  in  $w_{j,d}$ . Therefore,  $w_{j,d}(S \cap (T - v))$  has  $y$  as a subword. However, since  $(S + v) \cap T \subseteq T$ ,  $w_{j,d}((S + v) \cap T)$  is a subword of  $w_{j,d}(T)$ . Since  $b_{j,d}(q + \frac{v'}{R}) = 0$ , and since  $T$  is the central  $\Gamma_{W-2R}$  of the  $\Gamma_W$  corresponding to  $q + \frac{v'}{R}$ ,  $w_{j,d}(T)$  contains no occurrences of  $y$ . Therefore,  $w_{j,d}((S + v) \cap T)$  contains no occurrences of  $y$  either. Since  $w_{j,d}(S \cap (T - v)) = w_{j,d}((S + v) \cap T)$ , we have a contradiction.

The only remaining case is when  $v' = 0$ , i.e.  $|v_i| \leq \frac{W}{2}$  for  $1 \leq i \leq d$ . We then simply take an integer multiple  $nv'$  of  $v'$  such that some coordinate of  $nv'$  is greater than  $\frac{W}{2}$ , but at most  $W$ . Then,  $nv'$  is also a period of  $w_{j,d}$ , and we can repeat the argument just made for the same contradiction.

□

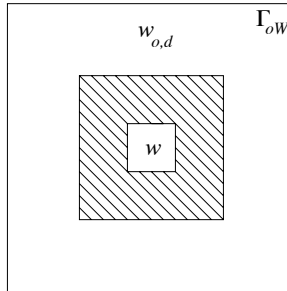


Figure 3.20:  $f'$

We now return to our proof of the lower bound on  $h^{top}(X) - h^{top}(X_w)$ . We suppose that our word  $w \in L_{\Gamma_n}(X)$  which we are removing from  $L(X)$  has  $n > (3d + 5)W$ . Take  $o$  to be the smallest integer such that  $(o - 4)W > n + 2R$ . Then,  $(o - 4)W > (3d + 5)W$ , and so  $o \geq 3d + 2$ . We also define  $n' = oW + 2R$ . By the definition of  $o$ ,  $n' < n + 4R + 5W \leq n + 44R + 70$ . Take  $\tilde{\mu}_w$  to be any ergodic measure of maximal entropy on  $X_w$ . Since

$$\sum_{w' \in L_{\Gamma_{n'}}(X_w)} \tilde{\mu}_w([w']) = \tilde{\mu}_w\left(\bigcup_{w' \in L_{\Gamma_{n'}}(X_w)} [w']\right) = \tilde{\mu}_w(X_w) = 1,$$

there must exist some word  $f \in L_{\Gamma_{n'}}(X_w)$  such that  $\tilde{\mu}_w([f]) \geq \frac{1}{H_{\Gamma_{n'}}(X_w)}$ . We note that since  $X_w \subset X$ ,  $H_{\Gamma_{n'}}(X_w) \leq H_{\Gamma_{n'}}(X)$ . By Lemma 3.2.1,  $H_{\Gamma_{n'}}(X) \leq e^{h^{top}(X)(n'+R)^d}$ . Therefore,  $\tilde{\mu}_w([f]) \geq e^{-h^{top}(X)(n'+R)^d}$ .

This means that by using Lemma 3.2.7, we see that if, for any  $\epsilon > 0$ , we denote by  $B_{k,\epsilon}(X_w)$  the set of words in  $L_{\Gamma_k}(X_w)$  which have at least  $k^d(e^{-h^{top}(X)(n'+R)^d} - \epsilon)$  occurrences of  $f$ , then  $\lim_{k \rightarrow \infty} \frac{\ln |B_{k,\epsilon}(X_w)|}{k^d} = h^{top}(X_w)$ .

In Figure 3.20, we show a word  $f' \in L_{\Gamma_{n'-2R}}(X)$  constructed as follows: the copy of  $\Gamma_{oW}^{(2W)}$  is filled with  $w_{o,d}$  as constructed in Theorem 3.4.2, and the central  $\Gamma_n$  is filled with  $w$ . The remaining shaded portion is filled using strong irreducibility to create a word  $f' \in L(X)$ .

We may now finally describe our map  $\psi_k$ . Consider any  $u \in B_{k,\epsilon}(X_w)$ . By definition,  $u$  has at least  $k^d(e^{-h^{top}(X)(n'+R)^d} - \epsilon)$  occurrences of  $f$ . We choose a set of disjoint occurrences of  $f$  using a simple algorithm: choose any occurrence of  $f$  to begin, call it  $f^{(1)}$ . At most  $3^d n'^d$  occurrences of  $f$  overlap  $f^{(1)}$ , and so we choose any occurrence

of  $f$  which does not overlap  $f^{(1)}$ , and call it  $f^{(2)}$ .  $f^{(3)}$  is any occurrence of  $f$  which does not overlap  $f^{(1)}$  or  $f^{(2)}$ , and we continue in this fashion. We are able to choose at least  $k^d 3^{-d} n'^{-d} (e^{-h^{top}(X)(n'+R)^d} - \epsilon)$  disjoint occurrences of  $f$  in  $u$  in this way. We then choose any nonempty subset of this chosen set of occurrences of  $f$ , and for each chosen occurrence of  $f$ , if its shape is  $U$  a copy of  $\Gamma_{n'}$ , we use strong irreducibility to replace the central  $\Gamma_{n'-2R}$  of  $U$  by  $f'$ .  $\psi_k(u)$  is defined to be the set of words which could be created by performing such replacements on  $u$ . The cardinality of  $\psi_k(u)$  is then at least  $2^{k^d 3^{-d} n'^{-d} (e^{-h^{top}(X)(n'+R)^d} - \epsilon)} - 1$  for any  $u \in B_{k,\epsilon}(X_w)$ .

We claim that  $\psi_k(u) \cap \psi_k(u') = \emptyset$  for any  $u \neq u' \in B_{k,\epsilon}(X_w)$ . To show this, we first show an auxiliary claim: that no occurrence of  $f'$  is ever incidentally created in the replacement process of  $\psi_k$ . In other words, when some occurrences of  $f$  in  $u \in B_{k,\epsilon}(X_w)$  are replaced by words containing  $f'$  to create an element of  $\psi_k(u)$ , the only occurrences of  $f'$  in the result are the ones which are subwords of each replaced occurrence of  $f$ . Suppose that this is not the case; that for some  $v \in B_{k,\epsilon}(X_w)$  and  $v' \in \psi_k(v)$ ,  $v'$  contains an occurrence of  $f'$  which occupies a copy of  $\Gamma_{n'-2R}$ , which we call  $B'$ , which was not the central  $\Gamma_{n'-2R}$  of a copy of  $\Gamma_n$  which was occupied by one of the replaced occurrences of  $f$  in  $v$ . Denote by  $B''$  the central  $\Gamma_n$  of  $B'$  and by  $B$  the copy of  $\Gamma_{n'}$  in which  $B'$  is central. Then  $v'(B'') = w$ . Since  $v \in L(X_w)$ ,  $v(B'') \neq w$ . This implies that one of the replacements made had nonempty intersection with  $B''$ , otherwise  $v(B'') = v'(B'')$ , a contradiction. Call the  $\Gamma_{n'}$  where this replacement was made  $C$ , and call its central  $\Gamma_{n'-2R}$   $C'$ . By our hypothesis,  $B$  was not one of the replaced copies of  $\Gamma_{n'}$ , and so  $B \neq C$ . Since all of the replacements made were disjoint,  $v'(C') = f'$ . We know that  $v'(B') = f'$  as well. Since  $C \cap B'' \neq \emptyset$ ,

$|B - C|_\infty \leq n + \frac{n'-n}{2} = \frac{n'+n}{2}$ . But,  $\frac{n'+n}{2} = \frac{n'-oW}{2} + \frac{oW+n}{2}$ . Since  $n' = oW + 2R + 2t$ ,  $\frac{n'-oW}{2} + \frac{oW+n}{2} = R + t + \frac{oW+n}{2}$ . It was part of the definition of  $o$  that  $(o-4)W > n + 2R + 2t$ , so  $R + t < \frac{(o-4)W-n}{2}$ . Therefore,  $R + t + \frac{oW+n}{2} < \frac{(o-4)W-n}{2} + \frac{oW+n}{2} = (o-2)W$ . We have then shown that  $|B - C|_\infty < (o-2)W$ . Since  $B'$  is central in  $B$  and  $C'$  is central in  $C$ ,  $B' - C' = B - C$  and so  $|B' - C'|_\infty < (o-2)W$  as well. We make one more notation: call  $E'$  the boundary of thickness  $2W$  of  $B'$  and  $F'$  the boundary of thickness  $2W$  of  $C'$ . Since  $B \neq C$ ,  $E' \neq F'$ . Then, since  $v'(B') = v'(C') = f'$ ,  $v'(E') = v'(F') = w_{o,d}$ . Also,  $E' - F' = B' - C'$ , so  $|E' - F'|_\infty < (o-2)W$ , which implies that  $E'$  has nonempty intersection with the central  $\Gamma_{(o-4)W}$  of  $F'$ , a contradiction to the aperiodicity property of  $w_{o,d}$ .

We have then shown that the only occurrences of  $f'$  in any  $v' \in \psi_k(v)$  for some  $v \in B_{k,\epsilon}(X_w)$  are those which occupy the central  $\Gamma_{n'-2R}$  of replaced occurrences of  $f$ . We also have a sort of converse result: for any such  $v, v'$ , and for every  $U$  a copy  $\Gamma_{n'}$  which is filled with an occurrence of  $f$  in  $v$  which is replaced, the central  $\Gamma_{n'-2R}$  of  $U$  is filled with  $f'$  in  $v'$ . This fact rests on the disjointness of the replacements made in the definition of  $\psi_k$ ; since these replacements are disjoint, each letter is changed at most once.

Now, consider any  $v' \in \psi_k(v)$  for  $v \in B_{k,\epsilon}(X_w)$ . Since  $v' \in \psi_k(v)$ ,  $v'$  has at least one occurrence of  $f'$ . Since the only occurrences of  $f'$  in  $v'$  come from replacements during the application of  $\psi_k$ , we know that for every  $U$  a copy of  $\Gamma_{n'-2R}$  with  $v'(U) = f'$ , if we call  $U'$  the copy of  $\Gamma_{n'}$  in which  $U$  is central, it must be the case that  $v(U') = f$ . We also know that these are the only replacements performed in turning  $v$  into  $v'$ ,

since any others would have resulted in more occurrences of  $f'$ . So,  $v'$  determines the set of replacements which were made in  $v$ . However, trivially  $v = v'$  outside of the regions where replacements occurred, and so  $v'$  determines the letters of  $v$  outside regions where replacements occurred, as well. Therefore,  $v$  is uniquely determined by  $v'$ , and so we know that for  $v \neq v' \in B_{k,\epsilon}(X_w)$ ,  $\psi_k(v) \cap \psi_k(v') = \emptyset$ .

Since  $|\psi_k(v)| \geq 2^{k^d 3^{-d} n'^{-d} (e^{-h^{top}(X)(n'+R)^d} - \epsilon)} - 1 > 2^{k^d 3^{-d+1} n^{-d} (e^{-h^{top}(X)(n'+R)^d} - \epsilon)}$  for  $v \in B_{k,\epsilon}(X_w)$ , we have shown that

$$H_{\Gamma_k}(X) > 2^{k^d 3^{-d+1} n^{-d} (e^{-h^{top}(X)(n'+R)^d} - \epsilon)} |B_{k,\epsilon}(X_w)|$$

for large  $k$ . Take natural logarithms of both sides, divide by  $k^d$ , and let  $k \rightarrow \infty$  to see that

$$h^{top}(X) \geq (\ln 2)(3^{-d+1} n^{-d} (e^{-h^{top}(X)(n'+R)^d} - \epsilon)) + h^{top}(X_w).$$

Since  $\epsilon$  was arbitrary, we allow it to approach zero, and so

$$h^{top}(X) - h^{top}(X_w) \geq \ln 2 \cdot 3^{-d+1} n^{-d} e^{-h^{top}(X)(n'+R)^d},$$

which for large enough  $n$ , gives

$$h^{top}(X) - h^{top}(X_w) > e^{-h^{top}(X)(n'+R+1)^d},$$

or, by replacing  $n'$  by its maximum possible value  $n + 44R + 69$ ,

$$h^{top}(X) - h^{top}(X_w) > \frac{1}{e^{h^{top}(X)(n+44R+70)^d}}.$$

Combining this with the earlier upper bound on  $h^{top}(X) - h^{top}(X_w)$ , we have

$$\frac{1}{e^{h^{top}(X)(n+44R+70)^d}} < h^{top}(X) - h^{top}(X_w) < \frac{D_X}{e^{h^{top}(X)(n-2R)^d}}.$$

□

### 3.5 A closer look at the main result

We here wish to briefly review the proof of Theorem 3.1.22, and point out that most of the complications encountered in the proof were due to possible problems with periodicity of  $w$ . If some assumptions about  $w$ 's aperiodicity are made, the proof can be simplified significantly. We will skip some of the details of these proofs, as they are just simplified versions of arguments already made in the main text.

**Theorem 3.5.1.** *For any strongly irreducible shift  $X = \Omega_{\mathcal{F}}$  of finite type with uniform filling length  $R$  and positive topological entropy  $h^{top}(X)$ , and for any words  $w, w' \in L_{\Gamma_n}(X)$  such that  $w$  and  $w'$  agree on the boundary of thickness  $t$  and any replacement of  $w$  by  $w'$  or vice versa cannot create or destroy any other occurrences of  $w$  and  $w'$ , if we denote by  $X_w$  the shift of finite type  $\Omega_{\mathcal{F} \sqcup \{w\}}$ , then for any  $\tilde{\mu}_w$  an ergodic measure of maximal entropy on  $X_w$  and any  $\tilde{\mu}$  an ergodic measure of maximal entropy on  $X$ ,*

$$(\ln 2)\tilde{\mu}_w([w']) \leq h^{top}(X) - h^{top}(X_w) \leq 2(\ln 2)\tilde{\mu}([w]).$$

**Proof:** To prove an upper bound on  $h^{top}(X) - h^{top}(X_w)$ , define a mapping  $\theta_k$  (similar to  $\phi_k$  from the upper bound portion of the proof of Theorem 3.1.22) on  $L_{\Gamma_{k+2n}}(X)$ :

given  $v \in L_{\Gamma_{k+2n}}(X)$ , one replaces the occurrences of  $w$  by  $w'$ , going in order lexicographically. Since none of these replacements can make new occurrences of  $w$ , this process terminates. The word occupying the central  $\Gamma_k$  of the resulting word is then in  $L_{\Gamma_k}(X_w)$  for the same reasons as before, and we call this word  $\theta_k(v)$ . We then wish to bound from above the size of  $\theta_k^{-1}(v) \cap A_{k+2n, \epsilon, \tilde{\mu}, w, w'}(X)$  for any  $v \in L_{\Gamma_{k+2n}}(X)$  and any ergodic measure of maximal entropy  $\tilde{\mu}$  on  $X$ . Recall that  $A_{k+2n, \epsilon, \tilde{\mu}, w, w'}(X)$  is the set of words in  $L_{\Gamma_{k+2n}}(X)$  which have between  $(k+2n)^d(\tilde{\mu}([x]) - \epsilon)$  and  $(k+2n)^d(\tilde{\mu}([x]) + \epsilon)$  occurrences of  $x$  for  $x = w, w'$ .

Consider any  $x \in \theta_k^{-1}(v) \cap A_{k+2n, \epsilon, \tilde{\mu}, w, w'}(X)$ . Because of the fact that occurrences of  $w$  and  $w'$  are not incidentally created nor destroyed during the replacements made in the process of changing  $x$  to  $v$ , the word occupying a shape  $S$  a copy of  $\Gamma_n$  will be changed from  $w$  to  $w'$  at some point during these replacements if and only if  $x(S) = w$ . If  $x(S) = w$ , then  $v(S) = w'$ , since once an occurrence of  $w'$  is created, it cannot be eliminated in any of the other replacements made in the process of changing  $x$  to  $\theta_k(x) = v$ . Therefore, any replacements made are at copies  $S$  of  $\Gamma_n$  such that  $v(S) = w'$ . We have then shown that if we denote by  $S_1, \dots, S_p$  the set of copies  $S$  of  $\Gamma_n$  such that  $v(S) = w'$ , then  $x$  agrees with  $v$  on  $\Gamma_k \setminus \left( \bigcup_{i=1}^p S_i \right)$ , and for each  $i$ ,  $x(S_i) = w$  or  $x(S_i) = w'$ . This shows that there are at most  $2^p$  choices for  $x(\Gamma_k)$ , and so since there are at most  $|A|^{(k+2n)^d - k^d}$  ways to fill  $\Gamma_{k+2n}^{(n)}$  with letters, that  $|\theta_k^{-1}(v)| \leq 2^p |A|^{(k+2n)^d - k^d}$ . But,  $v = \theta_k(x)$ , and  $x \in A_{k+2n, \epsilon, \tilde{\mu}, w, w'}(X)$ . Therefore,  $p$ , the number of occurrences of  $w'$  in  $v$ , is equal to  $q + r$ , where  $q$  is the number of occurrences of  $w$  in  $x$  and  $r$  is the number of occurrences of  $w'$  in  $x$ . By definition of  $A_{k+2n, \epsilon, n}(X)$ ,  $q \leq (k+2n)^d(\tilde{\mu}([w]) + \epsilon)$  and  $r \leq (k+2n)^d(\tilde{\mu}([w']) + \epsilon)$ . By Proposition 3.2.5,

since  $w$  and  $w'$  agree on their boundary of thickness  $t$ ,  $\tilde{\mu}([w]) = \tilde{\mu}([w'])$ . Therefore,  $p \leq 2(k+2n)^d(\tilde{\mu}([w]) + \epsilon)$ , and so we see that if  $\theta_k^{-1}(v) \cap A_{k+2n, \epsilon, \tilde{\mu}, w, w'}(X)$  is nonempty, then it has cardinality at most  $2^{2(k+2n)^d(\tilde{\mu}([w]) + \epsilon)} |A|^{(k+2n)^d - k^d}$ . This implies that

$$H_{\Gamma_k}(X_w) \geq \frac{|A_{k+2n, \epsilon, \tilde{\mu}, w, w'}(X)|}{2^{2(k+2n)^d(\tilde{\mu}([w]) + \epsilon)} |A|^{(k+2n)^d - k^d}}.$$

Take natural logarithms of both sides, divide by  $(k+2n)^d$ , and let  $k \rightarrow \infty$  to get

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{k^d}{(k+2n)^d} \frac{h_{\Gamma_k}(X_w)}{k^d} &\geq \lim_{k \rightarrow \infty} \frac{\ln |A_{k+2n, \epsilon, \tilde{\mu}, w, w'}(X)|}{(k+2n)^d} - 2 \ln 2(\tilde{\mu}([w]) + \epsilon) \\ &\quad - \frac{\ln |A|((k+2n)^d - k^d)}{(k+2n)^d}, \end{aligned}$$

and by using the definition of entropy and Lemma 3.2.7, we see that

$$h^{top}(X_w) \geq h_{\tilde{\mu}}(X) - 2(\ln 2)(\tilde{\mu}([w]) + \epsilon).$$

Since  $\tilde{\mu}$  is a measure of maximal entropy, and since  $\epsilon$  can be arbitrarily small, we may rearrange to arrive at

$$h^{top}(X) - h^{top}(X_w) \leq 2(\ln 2)\tilde{\mu}([w]).$$

We could then use Lemma 3.2.6 to further bound this from above, but we leave the bound as it is for now, to emphasize the fact that our upper bound comes from  $\tilde{\mu}([w])$ .

We can also give a similarly shortened proof of a lower bound on  $h^{top}(X) - h^{top}(X_w)$ . Define a map  $\delta_k$  (similar to  $\psi_k$  from the proof of the lower bound portion of Theorem 3.1.22) from  $L_{\Gamma_k}(X_w)$  to  $\mathcal{P}(L_{\Gamma_k}(X))$  as follows: for any  $v \in L_{\Gamma_k}(X_w)$  which has at least one occurrence of  $w'$ , take any nonempty subset of the occurrences of  $w'$  in  $v$



and replace them all by  $w$ . It is possible to perform all of these replacements simultaneously because the hypotheses of Theorem 3.5.1 imply that the portions of any two occurrences of  $w'$  which would need to be changed to turn them into  $w$  are disjoint. (If not, then a replacement of one of them by  $w$  would destroy the other.) Therefore, for any ergodic measure of maximal entropy  $\tilde{\mu}_w$  on  $X_w$  and for any  $v \in A_{k,\epsilon,\tilde{\mu}_w,w'}(X)$ ,  $|\delta_k(v)| \geq 2^{k^d(\tilde{\mu}_w([w'])-\epsilon)} - 1$ . For exactly the same reasons as in the earlier lower bound proof,  $\delta_k(v) \cap \delta_k(v') = \emptyset$  for  $v \neq v' \in A_{k,\epsilon,\tilde{\mu}_w,w'}(X_w)$ . Therefore,

$$H_{\Gamma_k}(X) > (2^{k^d(\tilde{\mu}_w([w'])-\epsilon)} - 1)|A_{k,\epsilon,\tilde{\mu}_w,w'}(X_w)|.$$

Take natural logarithms of both sides, divide by  $k^d$ , and let  $k \rightarrow \infty$  to see that

$$h^{top}(X) \geq (\ln 2)(\tilde{\mu}_w([w']) - \epsilon) + \limsup_{k \rightarrow \infty} \frac{\ln |A_{k,\epsilon,\tilde{\mu}_w,w'}(X_w)|}{k^d},$$

and by using Lemma 3.2.7 and letting  $\epsilon$  approach zero since it is arbitrary,

$$h^{top}(X) \geq (\ln 2)\tilde{\mu}_w([w']) + h_{\tilde{\mu}_w}(X_w).$$

Finally, we use the fact that  $\tilde{\mu}_w$  is a measure of maximal entropy on  $X_w$  to rewrite as

$$h^{top}(X) - h^{top}(X_w) \geq (\ln 2)\tilde{\mu}_w([w']).$$

□

We recall that the upper and lower bounds in Theorem 3.1.16 (the one-dimensional case) have a ratio bounded away from zero and infinity, but that we showed that it is

impossible to have such close bounds in the multidimensional case which depend only on the size of  $w$ . It is then natural to wonder whether or not it is possible, at least under the hypotheses on  $w$  from Theorem 3.5.1 and for large  $n$ , to have a lower bound of the form  $C\tilde{\mu}([w])$  for  $h^{top}(X) - h^{top}(X_w)$ . It turns out that this cannot be true if there are multiple measures of maximal entropy. In one dimension, the uniqueness of a measure of maximal entropy of any mixing shift of finite type is well-known, but Burton and Steif ([BuS1], [BuS2]) have shown that there are multidimensional strongly irreducible shifts of finite type which have more than one measure of maximal entropy. In order to show that this type of lower bound is too good to be true, we need a lemma:

**Lemma 3.5.2.** *If  $X$  is a strongly irreducible  $\mathbb{Z}^d$ -shift of finite type  $t$  and if we denote by  $E_n(X)$  the set of words  $w \in L_{\Gamma_n}(X)$  for which there exists  $w' \in L_{\Gamma_n}(X)$  where  $w$  and  $w'$  agree on the boundary of thickness  $t$  and such that replacing  $w$  by  $w'$  or vice versa can never incidentally create nor destroy an occurrence of  $w$  or  $w'$ , then for any ergodic measure  $\tilde{\mu}$  of maximal entropy on  $X$ ,  $\lim_{n \rightarrow \infty} \tilde{\mu}(E_n(X)) = 1$ .*

**Proof:** We claim that as long as  $n$  is sufficiently large, and  $w$  has the property that given any two disjoint copies  $S$  and  $T$  of  $\Gamma_{\lfloor \sqrt{n} \rfloor}$  contained in  $\Gamma_n$ ,  $w(S) \neq w(T)$ , then  $w \in E_n(X)$ . Consider any  $w$  with the property just described. Then create a standard replacement of  $w$  in the same way as in the proof of Theorem 3.3.1, by changing  $w$  only on a central copy of  $\Gamma_{2R+l}$ . (Recall that  $l = \left\lceil \left( \frac{d \ln n}{h^{top}(X)} \right)^{\frac{1}{d}} \right\rceil + 1$  or  $l = \left\lceil \left( \frac{d \ln n}{h^{top}(X)} \right)^{\frac{1}{d}} \right\rceil + 2$  and that  $l$  is chosen so that there exists  $a \in L_{\Gamma_l}(X)$  so that  $a$  is not a subword of  $w$ . We before chose  $l$  to be odd, but here we take it to have the

same parity as  $n$  so that there is a central  $\Gamma_l$  within  $\Gamma_n$ .) We denote by  $K$  the central copy of  $\Gamma_{2R+l}$  on which  $w$  is changed to create  $w'$ , and so  $w(\Gamma_n \setminus K) = w'(\Gamma_n \setminus K)$ . We have four cases that we would like to show are impossible:

**Case 1:** Suppose that a replacement of  $w$  by  $w'$  could possibly create a new occurrence of  $w$ . More rigorously, that there exists  $x \in X$  and copies  $T$  and  $T'$  of  $\Gamma_n$  such that  $x(T) = w$ , and if we define by  $x'$  the element of  $X$  created by replacing  $x(T)$  by  $w'$ , then  $x'(T') = w$ , but  $x(T') \neq w$ . This means that  $x'(T) = w'$  and  $x'(T') = w$ . Since the central  $\Gamma_l$  in  $w'$  is occupied by  $a$ , this implies that  $|T' - T|_\infty > \frac{n-l}{2}$ , or else  $x'(T') = w$  would have to contain this occurrence of  $a$ . However, in order for  $x'(T')$  to be a newly created occurrence of  $w$ , it must be the case that  $T'$  has nonempty overlap with  $K + (T - \Gamma_n)$ , and so  $|T' - T|_\infty < \frac{n+(2R+l)}{2}$ . By this upper bound on  $|T' - T|_\infty$ ,  $T \cap T'$  contains a cube of size at least  $\frac{n-(2R+l)}{2}$ , which must contain a copy of  $\Gamma_{\lfloor \sqrt{n} \rfloor}$  disjoint from the very small sets  $K + (T - \Gamma_n)$  and  $K + (T' - \Gamma_n)$ , call it  $U$ . Since  $x'(T) = w'$ ,  $x'(U) = w'(U - (T - \Gamma_n))$ . Since  $U$  is disjoint from  $K + (T - \Gamma_n)$ ,  $w'(U - (T - \Gamma_n)) = w(U - (T - \Gamma_n))$ . Since  $x'(T') = w$ ,  $x'(U) = w(U - (T' - \Gamma_n))$ . Therefore,  $w(U - (T - \Gamma_n)) = w(U - (T' - \Gamma_n))$ . But, since  $|T' - T|_\infty > \frac{n-l}{2}$ , which is greater than  $\sqrt{n}$  for large  $n$ ,  $U - (T - \Gamma_n)$  and  $U - (T' - \Gamma_n)$  are disjoint copies of  $\Gamma_{\lfloor \sqrt{n} \rfloor}$ , and so we have a contradiction.

**Case 2:** Suppose that a replacement of  $w$  by  $w'$  could possibly destroy an existing occurrence of  $w'$ . More rigorously, that there exists  $x \in X$  and copies  $T$  and  $T'$  of  $\Gamma_n$  such that  $x(T) = w$ , and if we define by  $x'$  the element of  $X$  created by replacing

$x(T)$  by  $w'$ , then  $x'(T') \neq w'$ , but  $x(T') = w'$ . This means that  $x(T') = w'$  and  $x(T) = w$ . Everything from here proceeds to a contradiction exactly as in Case 1.

**Case 3:** Suppose that a replacement of  $w$  by  $w'$  could possibly destroy an existing occurrence of  $w$ . More rigorously, that there exists  $x \in X$  and copies  $T$  and  $T'$  of  $\Gamma_n$  such that  $x(T) = w$ , and if we define by  $x'$  the element of  $X$  created by replacing  $x(T)$  by  $w'$ , then  $x'(T') \neq w$ , but  $x(T') = w$ . This means that  $x(T) = w$  and  $x(T') = w$ . However, since  $x'(T') \neq w$ , it must be the case that  $K + (T - \Gamma_n)$  has nonempty intersection with  $T'$ . This again implies that  $|T' - T|_\infty < \frac{n+(2R+l)}{2}$ . It is then possible to choose a positive integer  $k$  so that  $\sqrt{n} < |k(T' - T)|_\infty < n - \sqrt{n}$ . Then choose any  $U$  a copy of  $\Gamma_{\lfloor \sqrt{n} \rfloor}$  so that  $U, U + k(T - T') \subseteq T$ . Since  $x(T) = w$ ,  $x(U) = w(U)$ . Since  $x(T') = w$ ,  $x(U) = w(U + (T - T'))$ . So,  $w(U) = w(U + (T - T'))$ . In the same way, we can see that  $w(U) = w(U + k(T - T'))$ , and since  $|k(T - T')|_\infty > \sqrt{n}$ ,  $U$  and  $U + k(T - T')$  are disjoint and so we have a contradiction.

**Case 4:** Suppose that a replacement of  $w$  by  $w'$  could possibly create a new occurrence of  $w'$ . More rigorously, that there exists  $x \in X$  and copies  $T$  and  $T'$  of  $\Gamma_n$  such that  $x(T) = w$ , and if we define by  $x'$  the element of  $X$  created by replacing  $x(T)$  by  $w'$ , then  $x'(T') = w'$ , but  $x(T') \neq w'$ . This means that  $x'(T) = w'$  and  $x'(T') = w'$ . However, since  $x(T') \neq w'$ , it must be the case that  $K + (T - \Gamma_n)$  has nonempty intersection with  $T'$ . This again implies that  $|T' - T|_\infty < \frac{n+(2R+l)}{2}$ . It is then possible to choose a positive integer  $k$  so that  $\sqrt{n} < |k(T' - T)|_\infty < n - \sqrt{n}$ . Then, choose any  $U$  a copy of  $\Gamma_{\lfloor \sqrt{n} \rfloor}$  so that  $U, U + k(T - T') \subseteq T$ , and both are disjoint from  $K + (T - \Gamma_n)$ . Then, since  $x'(T) = w'$ ,  $x'(U) = w'(U)$ . Since  $x'(T') = w'$ ,

$x'(U) = w'(U + (T - T'))$ . So,  $w'(U) = w'(U + (T - T'))$ . In the same way, we can see that  $w'(U) = w'(U + k(T - T'))$ . Since  $U$  and  $U + k(T - T')$  are disjoint from  $K + (T - \Gamma_n)$ ,  $w(U) = w'(U)$  and  $w(U + k(T - T')) = w'(U + k(T - T'))$ . Therefore,  $w(U) = w(U + k(T - T'))$ . Since  $|k(T - T')|_\infty > \sqrt{n}$ ,  $U$  and  $U + k(T - T')$  are disjoint and so we have a contradiction.

Therefore, as long as  $w$  has  $w(S) \neq w(T)$  for  $S$  and  $T$  disjoint copies of  $\Gamma_{\lfloor \sqrt{n} \rfloor}$  in  $\Gamma_n$ ,  $w \in E_n(X)$ . This implies that

$$L_{\Gamma_n}(X) \setminus E_n(X) \subseteq \bigcup_{v \in L_{\Gamma_{\lfloor \sqrt{n} \rfloor}}(X)} \left( \bigcup_{u' \in [-n, n]^d, |u'|_\infty > \sqrt{n}} \left( \bigcup_{u \in [-n, n]^d} ([v] \cap ([v] + u')) + u \right) \right). \quad (3.8)$$

We can now use Lemma 3.2.6 to see that for any  $v \in L_{\Gamma_{\lfloor \sqrt{n} \rfloor}}(X)$ , any  $u'$  with  $|u'|_\infty > \sqrt{n}$ , and any ergodic measure of maximal entropy  $\tilde{\mu}$  on  $X$ ,

$$\tilde{\mu}([v] \cap ([v] + u')) \leq \frac{1}{H_{\left( \Gamma_{\lfloor \sqrt{n} \rfloor} \cup (\Gamma_{\lfloor \sqrt{n} \rfloor} + u') \right) \setminus \left( \Gamma_{\lfloor \sqrt{n} \rfloor} \cup (\Gamma_{\lfloor \sqrt{n} \rfloor} + u') \right)^{(R)}(X)}}.$$

Since  $|u'|_\infty > \sqrt{n}$ ,

$$\begin{aligned} & \left( \Gamma_{\lfloor \sqrt{n} \rfloor} \cup (\Gamma_{\lfloor \sqrt{n} \rfloor} + u') \right) \setminus \left( \Gamma_{\lfloor \sqrt{n} \rfloor} \cup (\Gamma_{\lfloor \sqrt{n} \rfloor} + u') \right)^{(R)} \\ &= \left( \Gamma_{\lfloor \sqrt{n} \rfloor} \setminus \Gamma_{\lfloor \sqrt{n} \rfloor}^{(R)} \right) \cup \left( \left( \Gamma_{\lfloor \sqrt{n} \rfloor} \setminus \Gamma_{\lfloor \sqrt{n} \rfloor}^{(R)} \right) + u' \right). \end{aligned}$$

Since  $|u'|_\infty > \sqrt{n}$ ,  $d\left( \left( \Gamma_{\lfloor \sqrt{n} \rfloor} \setminus \Gamma_{\lfloor \sqrt{n} \rfloor}^{(R)} \right), \left( \Gamma_{\lfloor \sqrt{n} \rfloor} \setminus \Gamma_{\lfloor \sqrt{n} \rfloor}^{(R)} \right) + u' \right) > \sqrt{n} > R$  for large  $n$ , and so by strong irreducibility,

$$H_{\left( \Gamma_{\lfloor \sqrt{n} \rfloor} \setminus \Gamma_{\lfloor \sqrt{n} \rfloor}^{(R)} \right) \cup \left( \left( \Gamma_{\lfloor \sqrt{n} \rfloor} \setminus \Gamma_{\lfloor \sqrt{n} \rfloor}^{(R)} \right) + u' \right)}(X) = H_{\left( \Gamma_{\lfloor \sqrt{n} \rfloor} \setminus \Gamma_{\lfloor \sqrt{n} \rfloor}^{(R)} \right)}(X)^2,$$

which is at least  $e^{h^{top}(X)2(\sqrt{n}-2R)^d}$  by Lemma 3.2.1. Combining this with shift-invariance of  $\tilde{\mu}$  and (3.8), we see that

$$\tilde{\mu}(L_{\Gamma_n}(X) \setminus E_n(X)) \leq H_{\Gamma_{\lfloor \sqrt{n} \rfloor}}(X)(2n)^{2d} e^{-h^{top}(X)2(\sqrt{n}-2R)^d}.$$

Again by Lemma 3.2.1,  $H_{\Gamma_{\lfloor \sqrt{n} \rfloor}}(X) \leq e^{h^{top}(X)(\sqrt{n}+R)^d}$ . Therefore,

$$\tilde{\mu}(L_{\Gamma_n}(X) \setminus E_n(X)) \leq (2n)^{2d} e^{-h^{top}(X)(2(\sqrt{n}-2R)^d - (\sqrt{n}+R)^d)},$$

which clearly approaches zero as  $n \rightarrow \infty$ .

□

Now, suppose that for any strongly irreducible  $\mathbb{Z}^d$ -shift of finite type  $X$ , there exist constants  $C, N > 0$  and an ergodic measure of maximal entropy  $\tilde{\mu}$  on  $X$  such that for any  $w \in E_n(X)$  with  $n > N$ ,  $C\tilde{\mu}([w]) < h^{top}(X) - h^{top}(X_w)$ . For a contradiction, choose any  $X$  with more than one measure of maximal entropy. Since the extreme points of the set of measures of maximal entropy are the ergodic measures of maximal entropy, there must exist some ergodic measure of maximal entropy  $\tilde{\mu}' \neq \tilde{\mu}$  on  $X$ . Then we have

$$C\tilde{\mu}([w]) < h^{top}(X) - h^{top}(X_w) < (2 \ln 2)\tilde{\mu}'([w]) \quad (3.9)$$

for all  $w \in E_n$  with  $n > N$ . However, since  $\tilde{\mu}$  and  $\tilde{\mu}'$  are ergodic, they must be mutually singular. Therefore, for any  $\epsilon > 0$ , there exists some open set  $U_\epsilon$  such that  $\tilde{\mu}(U_\epsilon) < \epsilon$  and  $\tilde{\mu}'(U_\epsilon) > 1 - \epsilon$ . By Lemma 3.5.2, for large enough  $n$   $\tilde{\mu}'(U_\epsilon \cap E_n) > 1 - 2\epsilon$ ,

and obviously  $\tilde{\mu}(U_\epsilon \cap E_n) < \epsilon$ . Since  $U_\epsilon$  is an open set in  $X$ , it is a union of cylinder sets. This means that there exists  $N' > N$  so that for any  $n > N'$ , there exists some  $V_\epsilon \subseteq U_\epsilon$  a union of cylinder sets of words in  $L_{\Gamma_n}(X)$  with  $\tilde{\mu}'(V_\epsilon \cap E_n) > 1 - 3\epsilon$ , and again  $\tilde{\mu}(V_\epsilon \cap E_n) < \epsilon$ . However, we assumed that  $C\tilde{\mu}([w]) < h^{top}(X) - h^{top}(X_w) < (2 \ln n)\tilde{\mu}'([w])$  for all  $w \in E_n$  and  $n > N$ , and so since  $V_\epsilon \cap E_n$  is a union of cylinder sets associated to  $E_n$ -words, for any  $n > N'$  we may sum (3.9) over  $w \in V_\epsilon \cap E_n$  to see that  $C\tilde{\mu}(V_\epsilon \cap E_n) < (2 \ln 2)\tilde{\mu}'(V_\epsilon \cap E_n)$ , and so  $C\epsilon < (2 \ln 2)(1 - 3\epsilon)$ , a contradiction for small enough  $\epsilon$ . This implies that if  $d > 1$ , then there does not exist a lower bound on  $h^{top}(X) - h^{top}(X_w)$  of the form  $C\tilde{\mu}([w])$  which holds for all words  $w \in E_n$  for sufficiently large  $n$ .

### 3.6 An application to an undecidability question

One of the complexities of multidimensional symbolic dynamics is that it is algorithmically undecidable, given just the set  $\mathcal{F}$  of forbidden words, whether or not a shift of finite type in  $\mathbb{Z}^d$  is nonempty. (See [B], [R], [Wan] for details.) One application of Corollary 3.3.7 is a condition under which  $\Omega_{\mathcal{F}}$  is nonempty:

**Theorem 3.6.1.** *For any alphabet  $A$ , there exist  $F, G \in \mathbb{N}$  such that for any  $m > 0$  and any finite set of words  $\mathcal{F}_m = \{w_k \in L_{\Gamma_{n_k}}(X) : 1 \leq k \leq m\}$  satisfying  $n_1 > G$  and  $n_k \geq F(n_{k-1})^{4d^2}$  for  $1 < k \leq m$ ,  $\Omega_{\mathcal{F}_m} \neq \emptyset$ .*

To prove this, we need the following lemma:

**Lemma 3.6.2.** *For any strongly irreducible  $\mathbb{Z}^d$ -shift  $X = \Omega_{\mathcal{F}}$  of finite type  $t$  with uniform filling length  $R$ , there exist  $C, E \in \mathbb{N}$  dependent only on  $d$  such that for*

any  $w \in L_{\Gamma_n}(X)$  with  $n > \max(C(R+1)^{4d^2}, t^2, E)$ , if we denote by  $X_w$  the shift of finite type  $\Omega_{\mathcal{F} \sqcup \{w\}}$ , then there is some  $w' \in L_{\Gamma_m}(X)$  a subword of  $w$  such that  $X_{w'}$  is nonempty and strongly irreducible with uniform filling length at most  $2n + R$ .

**Proof:** Suppose that such a shift  $X$  is given. By Corollary 3.3.7, there exists  $N$  such that for any  $w \in L_{\Gamma_n}(X)$  with  $n > N$ , there exist  $w', w'' \in L_{\Gamma_m}(X)$  such that  $w'$  is a subword of  $w$ ,  $w'$  and  $w''$  agree on  $\Gamma_n^{(t)}$ , and replacing an occurrence of  $w'$  by  $w''$  in an element of  $X$  cannot possibly create a new occurrence of  $w'$ . (We are not using the full force of Corollary 3.3.7 here, in that we do not care about the size  $m$  of  $w'$ , only that  $w'$  exists.) By reviewing the proof of Corollary 3.3.7, we see that a sufficient condition for the existence of  $w'$  is that

$$(n - t + 1) - \left( C' \left( \frac{\ln(n - t + 1)}{h^{top}(X)} + R \right) \right) (n - t + 1)^{1 - \frac{1}{2d}} > 0$$

for some constant  $C'$  depending only on  $d$ . Some algebraic manipulation shows that the above is implied if  $n > C'^{2d} \left( \frac{\ln(n - t + 1)}{h^{top}(X)} + R \right)^{2d} + t$ . By Lemma 3.2.1,  $e^{h^{top}(X)(R+1)^d} \geq H_{\Gamma_1}(X) \geq 2$ , so  $h^{top}(X) \geq \frac{\ln 2}{(R+1)^d}$ . Therefore, it is sufficient to have  $n > C'^{2d} \left( \frac{1}{\ln 2} (R+1)^d \ln n + R \right)^{2d} + t$ . Since  $(R+1)^d > R$  and  $n > t^2$ , for any  $n > 1$  a sufficient condition is  $n > 2C'^{2d} (2 \ln n (R+1)^d)^{2d} + \sqrt{n}$ . There is a constant  $E$  dependent only on  $d$  so that if  $n > E$ ,  $(\ln n)^{2d} < \sqrt{n}$ . For such  $n$ , it would be enough to have  $n > 2C'^{2d} \sqrt{n} (2(R+1)^d)^{2d} + \sqrt{n}$ , or by dividing both sides by  $\sqrt{n}$  and then squaring both sides,  $n > C(R+1)^{4d^2}$  for some constant  $C$  dependent only on  $d$ . So, as long as  $n > \max(C(R+1)^{4d^2}, t^2, E)$ , such a  $w'$  exists.

Consider any two shapes  $S, T \subseteq \mathbb{Z}^d$  such that  $d(S, T) > 2n + R$ , and any words  $y \in$



$L_S(X_{w'})$  and  $z \in L_T(X_{w'})$ . Since  $y \in L_S(X_{w'})$ , there exists  $y' \in L_{S \sqcup (S^c)^{(n)}}(X_{w'})$  such that  $y'(S) = y$ . Similarly, there exists  $z' \in L_{T \sqcup (T^c)^{(n)}}(X_{w'})$  such that  $z'(T) = z$ . For any  $p \in S \sqcup (S^c)^{(n)}$ , by definition there exists  $p' \in S$  such that  $d(p, p') \leq n$ . Similarly, for any  $q \in T \sqcup (T^c)^{(n)}$ , there exists  $q' \in T$  such that  $d(q, q') \leq n$ . But  $d(S, T) > 2n + R$ , so  $d(p', q') > 2n + R$ . By the triangle inequality,  $d(p, q) > R$ . Therefore, since  $p, q$  were arbitrary,  $d(S \sqcup (S^c)^{(n)}, T \sqcup (T^c)^{(n)}) > R$ , and so by strong irreducibility of  $X$  there exists  $x \in X$  such that  $x(S \sqcup (S^c)^{(n)}) = y'$  and  $x(T \sqcup (T^c)^{(n)}) = z'$ . We now fix any ordering of the elements of  $\mathbb{Z}^d$  with a least element (say lexicographically with respect to polar coordinates) and replace each element of  $w'$  by  $w''$  in turn with respect to this order. In this way, we eventually arrive at  $x' \in X$  which has no occurrences of  $w'$  and is thus an element of  $X_{w'}$ . Note that since  $x(S \sqcup (S^c)^{(n)}) = y'$  and  $x(T \sqcup (T^c)^{(n)})$  are words in  $L(X_{w'})$ , they had no occurrences of  $w'$ . Therefore, any of the replaced occurrences of  $w'$  had nonempty intersection with  $((S \sqcup (S^c)^{(n)}) \sqcup (T \sqcup (T^c)^{(n)}))^c$ . Consider such a replaced occurrence which occupies  $U$  a copy of  $\Gamma_m$ . From the fact just noted, there exists  $p \in U$  such that  $p \notin S \sqcup (S^c)^{(n)} \sqcup T \sqcup (T^c)^{(n)}$ . This implies that  $d(p, S) > n$  and  $d(p, T) > n$ . Therefore, since the size of  $U$  is  $m \leq n$ ,  $U$  is disjoint from both  $S$  and  $T$ . Since  $U$  is an arbitrary replaced occurrence of  $w'$ , this implies that  $x$  remained unchanged on  $S$  and  $T$  throughout the process of changing it to  $x'$ , and so  $x'(S) = x(S) = y$  and  $x'(T) = x(T) = z$ . By definition, we have shown that  $X_{w'}$  is nonempty and strongly irreducible with uniform filling length at most  $2n + R$ .

□

**Proof of Theorem 3.6.1:** We prove Theorem 3.6.1 by induction. Our inductive

hypothesis is that for any  $\mathcal{F}_m$  as described in the hypotheses,  $\Omega_{\mathcal{F}_m}$  contains a strongly irreducible shift of finite type with uniform filling length less than  $4n_m$ . First note that by Theorem 3.1.22 and Lemma 3.6.2, there certainly exists  $G$  such that for any word  $w_1 \in A^{\Gamma_{n_1}}$  with  $n_1 > G$ , there is a subword  $w'_1$  of  $w_1$  such that  $\Omega_{\{w'_1\}}$  is nonempty and strongly irreducible with uniform filling length  $R_1 < 4n_1$ . We take this  $G$  to be greater than the  $E$  from Lemma 3.6.2, and take our  $F$  to be  $5^{4d^2}C$ , where  $C$  is from Lemma 3.6.2. Now suppose the hypothesis to be true for  $m$ , and consider any  $\mathcal{F}_{m+1} = \{w_1, w_2, \dots, w_{m+1}\}$  satisfying the hypotheses of Theorem 3.6.1. By the inductive hypothesis,  $\Omega_{\{w_1, w_2, \dots, w_m\}}$  contains a strongly irreducible shift  $X_m$  of finite type  $t_m$  with uniform filling length  $R_m < 4n_m$ . There are two cases: either  $w_{m+1} \in L(X_m)$  or not. If  $w_{m+1} \notin L(X_m)$ , then  $(X_m)_{w_{m+1}} = X_m$ , and so since  $X_m = (X_m)_{w_{m+1}} \subseteq \Omega_{\mathcal{F}_{m+1}}$ , in this case  $\Omega_{\mathcal{F}_{m+1}}$  contains a strongly irreducible shift of finite type with uniform filling length  $R_m < 4n_m < 4n_{m+1}$  and the inductive hypothesis is verified for  $m+1$ . So, we suppose that  $w_{m+1} \in L(X_m)$ . We will show that  $n_{m+1} > \max(C(R_m + 1)^{4d^2}, t_m^2, E)$ . By the hypotheses of Theorem 3.6.1, we see that  $n_{m+1} > F(n_m)^{4d^2} = 5^{4d^2}C(n_m)^{4d^2} > C(5n_m)^{4d^2} > C(R_m + 1)^{4d^2}$ . Since the largest size of a word in  $\mathcal{F}$  is  $n_m$ ,  $t_m \leq n_m$ , and so  $n_{m+1} > F(n_m)^{4d^2} > t_m^2$ . Finally,  $n_{m+1} > n_1 > G > E$ . Therefore,  $w_{m+1}$  satisfies the hypotheses of Lemma 3.6.2 and so there is  $w'_{m+1}$  a subword of  $w_{m+1}$  such that  $(X_m)_{w'_{m+1}}$  is nonempty and strongly irreducible with uniform filling length less than  $4n_{m+1}$ . Also,  $(X_m)_{w'_{m+1}} \subseteq (X_m)_{w_{m+1}} \subseteq \Omega_{\mathcal{F}_{m+1}}$ , and so the induction is complete. We have then shown that for any  $m$ ,  $\Omega_{\mathcal{F}_m}$  contains a nonempty shift and is therefore nonempty itself.

□

### 3.7 Questions

A few questions suggest themselves based on our results. The first, and perhaps most obvious, is simply how the bounds in Theorem 3.1.22 could be improved. (We suspect that at least the lower bound is far from optimal.)

**Question 3.7.1.** *Can the constants  $A_X$  and  $B_X$  in the statement of Possible Theorem 3.1.21 be improved from their values in Theorem 3.1.22?*

We know at least that the gap between  $A_X$  and  $B_X$  can be improved only up to possible additive and multiplicative constants; we showed earlier that there are examples which force  $A_X - B_X \geq 2R$  for strongly irreducible shifts of finite type with arbitrarily large  $R$ , and our Theorem 3.1.22 has  $A_X - B_X = 46R + 70$ . However, it would be nice to achieve optimal values.

Next, recall that we showed in the introduction that no result nearly as strong as Theorem 3.1.22 can be true for shifts of finite type which are only assumed to be topologically mixing. However, it is probable that something can be said. So, we can ask

**Question 3.7.2.** *If  $X$  is a topologically mixing  $\mathbb{Z}^d$ -shift  $X = \Omega_{\mathcal{F}}$  of finite type with positive topological entropy  $h^{\text{top}}(X)$ , and if we denote by  $X_w$  the shift of finite type  $\Omega_{\mathcal{F} \sqcup \{w\}}$ , then is it true that for every  $\epsilon > 0$ , there exists  $N$  such that for all  $w \in L_{\Gamma_n}(X)$  with  $n > N$ ,  $h^{\text{top}}(X) - h^{\text{top}}(X_w) < \epsilon$ ? If so, can anything be said about the rate at which  $h^{\text{top}}(X) - h^{\text{top}}(X_w)$  approaches zero as the size  $n$  of  $w \in L_{\Gamma_n}(X)$  approaches infinity?*

If it is the case that  $h^{top}(X) - h^{top}(X_w)$  approaches zero as  $n \rightarrow \infty$ , then it seems likely that some new techniques would be necessary for a proof, since there exist non-strongly irreducible shifts of finite type whose languages contain words  $w$  for which there exists no word  $w' \neq w$  agreeing with  $w$  on its border. For example, in the checkerboard island shift  $Z$ , any square word which is a finite portion of a checkerboard is forced by its boundary.

## BIBLIOGRAPHY

- [B] ROBERT BERGER, *The undecidability of the domino problem*, Mem. Amer. Math. Soc. **66** (1966)
- [Be] VITALY BERGELSON, *Ergodic Ramsey Theory - an Update*, Ergodic Theory of  $\mathbb{Z}^d$ -actions, M. Pollicott and K. Schmidt, editors, London Math. Soc. Lecture Notes Series **228**, Cambridge University Press (1996), 1–61.
- [BeL] VITALY BERGELSON AND ALEXANDER LEIBMAN, *Polynomial extensions of van der Waerden’s and Szemerédi’s theorems*, Journal of AMS **9** (1996), no. 3, 725–753.
- [Bi] G.D. BIRKHOFF, *Proof of the ergodic theorem*, Proc. Nat. Acad. Sci. USA, **17** (1931) 656–660.
- [Bou] JEAN BOURGAIN, *Pointwise ergodic theorems for arithmetic sets*, Publ. Math. IHES **69** (1989), no. 3, 5–45.
- [Bou2] JEAN BOURGAIN, *On pointwise ergodic theorems for arithmetic sets*, C. R. Acad. Sci. Paris Sér. I Math. **305** (1987), no. 10, 397–402.
- [BoyLR] MICHAEL BOYLE, DOUGLAS LIND, AND DANIEL RUDOLPH, *The automorphism group of a shift of finite type*, Trans. Amer. Math. Soc., **306** (1988), no. 1, 71–114.
- [BuS1] ROBERT BURTON AND JEFFREY STEIF, *Non-uniqueness of measures of maximal entropy*, Ergodic Theory Dynam. Systems, **14** (1994), no. 2, 213–235.
- [BuS2] ROBERT BURTON AND JEFFREY STEIF, *New results on measures of maximal entropy*, Israel J. Math. **89** (1995), nos. 1-3, 275–300.
- [Fa] BASSAM FAYAD, *Analytic mixing reparametrizations of irrational flows*, Ergodic Theory Dynam. Systems **22** (2002), no. 2, 437–468.

- [Fu] HILLEL FURSTENBERG, *Strict ergodicity and transformation of the torus*, Amer. J. Math. **83** (1961), 573–601.
- [GO1] LEO GUIBAS AND ANDREW ODLYZKO, *Maximal prefix-synchronized codes*, SIAM J. Appl. Math. **35** (1978), 401–418.
- [GO2] LEO GUIBAS AND ANDREW ODLYZKO, *Periods in strings*, J. Combin. Theory Ser. A, **30** (1981), 19–42.
- [HaK] FRANK HAHN AND YITZHAK KATZNELSON, *On the entropy of uniquely ergodic transformations*, Trans. Amer. Math. Soc. **126** (1967), 335–360.
- [HaW] G.H. HARDY AND E.M. WRIGHT, *An introduction to the theory of numbers*, Clarendon Press, 1979.
- [HeM] GUSTAV HEDLUND AND MARSTON MORSE, *Symbolic dynamics*, Amer. J. Math., **60** (1938), no. 4, 815–866.
- [L] DOUGLAS LIND, *Perturbations of shifts of finite type*, Siam J. Discrete Math. **2** (1989), no. 3, 350–365.
- [LM] DOUGLAS LIND AND BRIAN MARCUS, *An Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, 1995.
- [M] MICHAEL MISIUREWICZ, *A short proof of the variational principle for a  $\mathbb{Z}_+^n$ -action on a compact space*, Asterisque **40** (1975), 147–157.
- [QS] ANTHONY QUAS AND A.A. ŞAHİN, *Entropy gaps and locally maximal entropy in  $\mathbb{Z}^d$  subshifts*, Ergodic Theory Dynam. Systems **23** (2003), 1227–1245.
- [QT] ANTHONY QUAS AND PAUL TROW, *Subshifts of multi-dimensional shifts of finite type*, Ergodic Theory Dynam. Systems **20** (2000), no. 3, 859–874.
- [R] RAPHAEL M. ROBINSON, *Undecidability and nonperiodicity for tilings of the plane*, Invent. Math **12** (1971), 177–209.
- [Wal] PETER WALTERS, *An Introduction to Ergodic Theory*, Springer-Verlag, 1982.
- [Wan] HAO WANG, *Proving theorems by pattern recognition II*, AT&T Bell Labs. Tech. J. **40** (1961), 1–41.

- [Wi] MÁTÉ WIERDL, *Pointwise ergodic theorem along the prime numbers*, Israel J. Math. **64** (1988), no. 3, 315–336.