# Ergodic Theorems For Free Groups

by

**Roseanna Ferguson**

**M4K8 Scholarly Report**

Submitted to The University of Warwick

**Mathematics Institute**

April, 2016

# Contents

# 1 Introduction

This report is concerned with the generalisation of the ergodic theorems for actions of free groups, it will assume $f \in L^1(X)$ and discuss the convergence of a family of operators, $\mathcal{A}_n$.

**Definition 1.1.** *(Averaging operator). Let $F_m$ be the free non-abelian group on $m \geq 2$ generators, $c_1, c_2, \dots, c_m$. Then, for any measure preserving system $X = (X, \mathcal{X}, \mu)$, with $\mu(X) < \infty$, any $f \in L^1(X)$, any $x \in X$ and any $n \geq 1$, the averaging operators are defined as,*

$$\mathcal{A}_n f(x) \coloneqq \frac{1}{2m \times (2m-1)^{n-1}} \sum_{g \in F_m : |g| = n} f(T_g^{-1} x).$$

In particular, it will discuss the example, established by Terence Tao [17], of a measure preserving system $X$ and $f \in L^1(X)$ for which the sequence of averaging operators $\mathcal{A}_n f(x)$ is unbounded in $n$ for almost every $x$. For purposes of clarity, [17] restricted to the case of the free group on 2 generators and the main objective of this report will be to extend to the general case on $m \geq 2$ generators. The existence of such an example demonstrates that the pointwise and maximal ergodic theorems do not hold in $L^1$ for actions of $F_m$, $m \geq 2$.

Attempts will also be made to expand the justifications given in [17] in order to give a comprehensive and more accessible account. Finally, the concluding remarks will highlight the areas which required adaptation from the case on 2 generators.

## 1.1 Formulation of Definitions

This first section will introduce the key terminology of the report. Let $c_1, \dots, c_m$ be the generators of the free group $F_m$ and $W$ be the set

$$W = \{c_1, \dots, c_m, c_1^{-1}, \dots, c_m^{-1}\}.$$

**Definition 1.2.** *(Reduced word). Each non-identity element of $F_m$ can be represented by a unique reduced word. A reduced word is defined to be a word with letters in the set $W$ such that for every $i \in \{1, \dots, m\}$, $c_i$ and $c_i^{-1}$ are never adjacent. Furthermore, for $g \in F_m$ define the word length $|g|$ of $g$ to be the length of the unique reduced word that produces $g$.*

**Definition 1.3.** *(Action). Let $F_m$ be a group and $X$ a set. An (left) action of $F_m$ on $X$ is a map $T_g : X \to X$, defined for $g \in F_m$, $x \in X$ by $T_g(x)$, which satisfies the properties:*

*(i) $T_g T_h(x) = T_g(T_h(x))$ for all $g, h \in F_m$ and $x \in X$,*

*(ii) $T_1(x) = x$ for every $x \in X$.*

The action of the free group $F_m$, on the measure space $(X, \mathcal{X}, \mu)$, will be specified by the maps $T_{c_i}$ associated to each generator $c_i$.

**Definition 1.4.** *($F_m$-system). Define an $F_m$-system to be $(X, \mathcal{X}, \mu, (T_g)_{g \in F_m})$, where $(X, \mathcal{X}, \mu)$ is a measure space with $0 < \mu(X) < \infty$. $T_g : X \to X$ is a family of measure-preserving maps on $X$ for $g \in F_m$, where $T_1$ is the identity and $T_g T_h = T_{gh}$ $\forall g, h \in F_m$. In particular, $T_g$ are bi-measurable with $T_g^{-1} = T_{g^{-1}}$.*

- The system can be normalised by dividing $\mu$ by $\mu(X)$. However, in the last section of the report two of these systems will be combined into a third, where it will be convenient that an $F_m$-system is not required to have measure 1.
- Bi-measurable specifies that $T_g$ and its inverse are measurable with respect to the $\sigma$-algebra, $\mathcal{X}$.
- For a map $T : X \to X$ and a measure space $(X, \mathcal{X}, \mu)$, $T$ is measure preserving or equivalently $\mu$ is $T$-invariant if $\mu(T^{-1}(B)) = \mu(B)$ $\forall B \in \mathcal{X}$.

An $F_m$-system is fully specified by $m$ arbitrary invertible, bi-measurable, measure preserving maps, $T_{c_1}, \dots, T_{c_m} : X \to X$ corresponding to the generators of the free group, $F_m$. For $g = s_1 \cdots s_r \in F_m$, $T_g$ is defined by $T_g = T_{s_1} T_{s_2} \cdots T_{s_r}$.
Finally, ergodicity with respect to $F_m$ or $F_m^2$ is defined below, where $F_m^2$ denotes the index two subgroup of $F_m$ consisting of $g \in F_m$ with even word length. For comparison, the general definition for ergodicity can be found in the appendix, Section 9.1.
**Definition 1.5.** *($F_m$-invariant set). $B \in \mathcal{X}$ is an $F_m$-invariant set, if for all $g \in F_m$, $T_g^{-1}(B) = B$. Or equivalently, if for all $i \in \{1, \dots, m\}$, $T_{c_i}^{-1}(B) = B$. Similarly, $B \in \mathcal{X}$ is an $F_m^2$-invariant set if for all $g \in F_m$ with $|g| = 2n$, $n \in \mathbb{N}$, $T_g^{-1}(B) = B$.*
**Definition 1.6.** *($F_m$-ergodic). A system is $F_m$-ergodic if all $F_m$-invariant measurable sets either have zero or full measure i.e. if $B$ satisfies $T_g^{-1}(B) = B$ for all $g \in F_m$ then, $\mu(B) = 0$ or $\mu(X)$. Similarly, an $F_m$-system is $F_m^2$-ergodic if all $F_m^2$-invariant measurable sets either have zero or full measure.*

## 1.2 Ergodic Theory

The well known Birkhoff's ergodic theorem, the proof of which can obtained from [14], states
**Theorem 1.7.** *(Birkhoff's ergodic theorem). Let $(X, \mathcal{X}, \mu)$ be a probability space and let $T : X \to X$ be a measure preserving transformation. Let $\mathcal{I}$ denote the $\sigma$-algebra of $T$-invariant sets. Then, for every $f \in L^1(X)$*

$$\frac{1}{n} \sum_{j=0}^{n-1} f(T^j x) \to E(f|\mathcal{I}), \quad as \quad n \to \infty$$

*for $\mu$-almost every $x \in X$.*

This report proves there is no such analogous result for the case of actions of free groups on the space of $L^1$ functions for spherical averages. However, the example constructed in this report stands in contrast to other results for $L^1$ since, if slight adjustments are made to the assumptions, then a pointwise ergodic theorem does hold. For example, the pointwise ergodic theorem holds for finite measure preserving actions of the free *abelian* group on two generators, presented amongst other results in [8] and [6] established the case for the Cesáro means, $\frac{1}{N}\sum_{n\leq N}\mathcal{A}_n$, of spherical averages for $L^1$. Finally, the free group is not an amenable group but the pointwise ergodic theorem does hold for general locally compact amenable groups along Følner sequences that obey some restrictions, see [5].

**Definition 1.8.** ($L\log L$). *Let $(X,\mathcal{X},\mu)$ be a measure space. Then we say $\phi\in L\log L(X)$ if $|\phi|\log|\phi|\in L^1(X)$.*

The following theorem by Nevo and Stein [6] establishes the pointwise ergodic theorem for actions of free groups and functions in $L\log L$, they also prove the case for functions in $L^p$, $p>1$.

**Theorem 1.9.** *(Pointwise ergodic theorem). Let $(X,\mathcal{X},\mu,(T_g)_{g\in F_m})$ be an $F_m$-system. If $\int_X|f|\log(1+|f|)d\mu<\infty$, then $\mathcal{A}_{2n}f$ converges pointwise almost everywhere (and in $L^1(X)$ norm) to an $F_m^2$-invariant function. In particular, if $(X,\mathcal{X},\mu,(T_g)_{g\in F_m})$ is $F_m^2$-ergodic, then $\mathcal{A}_{2n}f$ converges pointwise almost everywhere in $L^1$ to the constant $\frac{1}{\mu(X)}\int_X fd\mu$.*

In fact, this will be used to prove the existence of a counterexample. The above also appears in Bufetov [1], from which the report will take other results, who used instead the 'Alternierende Verfahren' (alternating method) of Rota [12] to cover the $L\log L$ case. Bufetov [1] takes a similar perspective to this report by relating the averaging operators to a Markov operator.

The above theorem restricts to the case of even averages, $\mathcal{A}_{2n}$, and uses $F_m^2$ instead of $F_m$. This is necessary for functions satisfying $\int_X|f|\log(1+|f|)d\mu<\infty$ or $f\in L^p$, $p\geq 1$ and the following example, as suggested by Tao [17], demonstrates why.

**Example 1.10.** *Take $X=\{0,1\}$ with uniform measure and let*

$$T_{c_1}(x)=T_{c_2}(x)=\cdots=T_{c_m}(x)=\mathbb{1}_{\{x=0\}}.$$

*For any $f\in L^1(X)$ with $f(0)\neq f(1)$, $\mathcal{A}_nf$ neither converges pointwise nor converges in $L^1$, to any limit.*

*Proof.* Let $g\in F_m$, $T_g$ depends on whether the length of the word, $n=|g|$, is odd or even.

$$T_g(0)=\mathbb{1}_{\{n\text{ is odd}\}}\quad\text{ and }\quad T_g(1)=\mathbb{1}_{\{n\text{ is even}\}}.$$

It is required to show that $\exists\epsilon_1,\epsilon_2>0$ such that $\forall N\in\mathbb{N}$, $\exists n_1,m_1,n_2,m_2>N$ with,

$$|\mathcal{A}_{n_1}f(x)-\mathcal{A}_{m_1}f(x)|\geq\epsilon_1,\quad\text{for all }x\in X\quad\text{ and }\quad\|\mathcal{A}_{n_2}f-\mathcal{A}_{m_2}f\|_{L^1(X)}\geq\epsilon_2.$$

Let $f\in L^1$ be such that $|f(0)-f(1)|>0$. Take $n$ to be even then $T_g^{-1}(0)=0$ and the number of reduced words of length $n$ is $2m \times (2m-1)^{n-1}$, so

$$\begin{aligned}\mathcal{A}_n f(0) &:= \frac{1}{2m \times (2m-1)^{n-1}} \sum_{g \in F_m : |g| = n} f(0) \\ &= \frac{f(0)}{2m \times (2m-1)^{n-1}} \times \{\text{number of reduced words length } n\} \\ &= f(0).\end{aligned}$$

By similar calculations,

$$\mathcal{A}_n f(0) = \begin{cases} f(0) & \text{if} \quad n \text{ is even} \\ f(1) & \text{if} \quad n \text{ is odd} \end{cases} \quad \text{and} \quad \mathcal{A}_n f(1) = \begin{cases} f(1) & \text{if} \quad n \text{ is even} \\ f(0) & \text{if} \quad n \text{ is odd.} \end{cases}$$

Therefore, there exists $\epsilon_1 = |f(0) - f(1)| > 0$ such that for all $N \in \mathbb{N}$ there exists $n_1 = N+1, m_1 = N+2 > N$ (one is even and one is odd),

$$|\mathcal{A}_{n_1} f(1) - \mathcal{A}_{m_1} f(1)| = |\mathcal{A}_{n_1} f(0) - \mathcal{A}_{m_1} f(0)| = |f(1) - f(0)| = \epsilon_1.$$

So, $\mathcal{A}_n f(x)$ is not Cauchy for any $x \in X$ and cannot converge pointwise. Similarly, in $L^1$ there exists $\epsilon_2 = |f(0) - f(1)| > 0$ such that for all $N \in \mathbb{N}$ there exists $n_2 = N+1, m_2 = N+2 > N$ with,

$$\begin{aligned}\|\mathcal{A}_n f - \mathcal{A}_m f\|_{L^1(X)} &= |\mathcal{A}_n f(0) - \mathcal{A}_m f(0)| \int_{x=0} 1 d\mu(x) + |\mathcal{A}_n f(1) - \mathcal{A}_m f(1)| \int_{x=1} 1 d\mu(x) \\ &= |f(0) - f(1)| \mu(x=0) + |f(1) - f(0)| \mu(x=1) \\ &= |f(1) - f(0)| = \epsilon_2,\end{aligned}$$

since $\mu(1) = \mu(0) = 0.5$ by definition of uniform measure. Therefore, $\mathcal{A}_n f$ does not converge in $L^1$. Finally, take $f(0) = 1,\ f(1) = 2$ then

$$\int_X |f| \log(1 + |f|) = |f(0)| \log(1 + |f(0)|) \mu(0) + |f(1)| \log(1 + |f(1)|) \mu(1) < \infty.$$

This same function lies in $L^p$, for $p \geq 1$. □

## 1.3 Measure Theory

Finally, two key theorems from measure theory will be used in the report. The Kolmogorov extension theorem, the statement of which can be found in the appendix, Section 9.2, roughly speaking, states that if a measure is defined on a smaller set, which generates the $\sigma$-algebra, then the measure is fully specified on the whole $\sigma$-algebra. The second theorem is used several times to obtain a lower bound on a sequence of functions. The proof can be found in [13].

**Theorem 1.11.** *(Egorov's theorem). Assume $E$ has finite measure. Let $\{f_n\}$ be a sequence of measurable functions on $E$ that converges pointwise on $E$ to the real-valued function $f$. Then, for each $\epsilon > 0$, there is a closed set $F$ contained in $E$ for which $f_n \to f$, as $n \to \infty$, uniformly on $F$ and $\mu(E \setminus F) < \epsilon$.*

When $E$ is a finite measure space, the conclusion of Egorov's theorem can be restated to say, the set of functions converges *uniformly* to the limit outside of a set of measure $\epsilon$.

## 2 Summary

Firstly, recall the definition of the averaging operators.

**Definition 2.1.** *(Averaging operator). For any $f \in L^1(X) = L^1(X, \mathcal{X}, \mu)$, any $x \in X$ and any $n \geq 1$, we define the averaging operators as,*

$$\mathcal{A}_n f(x) := \frac{1}{2m \times (2m-1)^{n-1}} \sum_{g \in F_m : |g| = n} f(T_g^{-1} x).$$

**Remark 2.2.** *$2m \times (2m-1)^{n-1}$ is the number of reduced words of length $n$. Symmetry allows $T_g^{-1}$ to be replaced by $T_g$ in the definition since $T_g^{-1} = T_{g^{-1}}$ and the summation is over all words $g$ of length $n$. $\mathcal{A}_n f(x)$ can be interpreted as the $\mathbb{E}[f(T_g^{-1} x)]$ where $g$ is the random variable that assigns equal probability to each word in $F_m$ of length $n$.*

The objective of the report is to present the proof of the following statement, which is the generalisation of Tao [[17], Theorem 1.2].

**Theorem 2.3.** *(Counterexample). There exists an $F_m$-system, $(X, \mathcal{X}, \mu, (T_g)_{g \in F_m})$ and $f \in L^1(X)$ such that*

$$\sup_n |\mathcal{A}_{2n} f(x)| = \infty,$$

*for almost every $x \in X$. In particular, $\mathcal{A}_{2n} f(x)$ fails to converge to a limit as $n \to \infty$ for almost every $x \in X$.*

The proof of Theorem 2.3 will be structured as in [17]. Firstly, Theorem 2.3 will be reduced into a useful form. This will involve two reformulations. The first, involves reducing the problem to finding an $F_m$-system and an $L^\infty$ function for which the supremum of the averaging operators is bounded below but whose $L^1$ norm is bounded above by the size of the measure space. By combining an infinite number of copies of this $F_m$-system into a product system, it is possible to construct a counterexample. The second reformulation involves interpreting the averaging operators, $\mathcal{A}_{2n}$, as Markov operators.
Section 5 introduces two further theorems which together prove the second reformulation of Theorem 2.3. These theorems assert the existence of good systems, introduced in Section 4, on which a set of functions with certain technical properties can be defined. The main body of the proof is in proving these two theorems which take the form of the base case

and inductive step of an induction argument.

# 3 Reductions

This section will be centred around two reformulations of Theorem 2.3.

## 3.1 The First Reduction

The first reduction of Theorem 2.3, will be the generalisation of [[17], Theorem 2.1],

**Theorem 3.1.** *(Quantitative counterexample). Let $\alpha, \epsilon > 0$. Then there exists an $F_m$-system, $(X, \mathcal{X}, \mu, (T_g)_{g\in F_m})$, and a non-negative function $f \in L^\infty(X)$, with*

$$\|f\|_{L^1(X)} \leq \alpha\mu(X) \quad \textit{but such that} \quad \sup_n \mathcal{A}_{2n} f(x) \geq 1 - \epsilon,$$

*for all $x \in X$ outside a set of measure at most $\epsilon\mu(X)$.*

In order to prove Theorem 2.3 follows from Theorem 3.1, one of the Borel-Cantelli lemmas will be used, the statement of which can be found in the appendix, Section 9.3.

**Proposition 3.2.** *Theorem 3.1 implies Theorem 2.3.*

*Proof.* Proceed as in [17], in Theorem 3.1, take $\alpha = \epsilon = 2^{-k}$ and divide $\mu$ by $\mu(X)$ in order to assume, without loss of generality, that $\mu(X) = 1$. Therefore, for each natural number $k$ there exists an $F_m$-system, $(X_k, \mathcal{X}_k, \mu_k, (T_{g,k})_{g\in F_m})$, with $\mu_k(X_k) = 1$ and a non-negative function $f_k \in L^\infty(X_k)$ such that

$$\|f_k\|_{L^1(X_k)} \leq 2^{-k}\mu_k(X_k) = 2^{-k} \quad \text{and} \quad \sup_n \mathcal{A}_{2n} f_k(x)) \geq 1 - 2^{-k} \geq 1/2$$

for all $x \in X_k$ outside a set, $A_k$, of measure $2^{-k}\mu_k(X_k) = 2^{-k}$. Combine these $F_m$-systems by taking $X$ to be the Cartesian product $X := \prod_k X_k$ with product $\sigma$-algebra $\mathcal{X} := \prod_k \mathcal{X}_k$, product probability measure $\mu := \prod_k \mu_k$, so

$$\mu(X) = \prod_k \mu_k(X_k) = 1,$$

and product action $T_g := \biguplus_k T_{g,k}$. Altogether, this then forms the product system $(X, \mathcal{X}, \mu, (T_g)_{g\in F_m})$. Furthermore, each $f_k \in L^\infty(X_k)$ lifts to $\tilde{f}_k \in L^\infty(X)$ defined by,

$$\tilde{f}_k(x) = f_k(\pi_k(x))$$

for $x \in X$, where $\pi_k : X \to X_k$, defined by $\pi_k(x_1, x_2, \dots) = x_k$, is surjective. Thus, for $k \in \mathbb{N}$

$$\|\tilde{f}_k\|_{L^1(X)} = \|f_k \circ \pi_k\|_{L^1(X)} = \|f_k\|_{L^1(X_k)} \leq 2^{-k}$$

and

$$\sup_n \mathcal{A}_{2n}\tilde{f}_k(x) = \sup_n \mathcal{A}_{2n} f_k(\pi_k(x)) = \sup_n \mathcal{A}_{2n} f_k(y) \geq 1 - 2^{-k} \geq 1/2.$$

Where the last inequality holds for all $y = \pi_k(x) \in X_k$ outside of a set, $A_k$, of measure $2^{-k}$ inside $X_k$ and therefore, for all $x = \pi_k^{-1} y \in X$ outside of the set $\pi_k^{-1}(A_k)$. Noting that $\mu_j(X_j) = 1$ for all $j$,

$$\begin{aligned}\mu(\pi_k^{-1} A_k) &= \mu(X_1, \dots, X_{k-1}, A_k, X_{k+1}, \dots) \\ &= \mu_1(X_1) \cdots \mu_{k-1}(X_{k-1}) \mu_k(A_k) \mu_{k+1}(X_{k+1}) \cdots \\ &= \mu_k(A_k) = 2^{-k}.\end{aligned}$$

$\tilde{f}_k \in L^\infty(X)\ \forall k$ therefore, $\tilde{f}_k \in L^1(X)\ \forall k$ due to the nesting of $L^p$ spaces on a finite measure space. Set $f := \sum_k k\tilde{f}_k$, then

$$\|f\|_{L^1(X)} = \|\sum_{k=1}^{\infty} k\tilde{f}_k\|_{L^1(X)} \leq \sum_{k=1}^{\infty} k\|\tilde{f}_k\|_{L^1(X)} \leq \sum_{k=1}^{\infty} k2^{-k} < \infty,$$

so $f \in L^1(X)$.

$$\mu(x \in X | \sup_n \mathcal{A}_{2n}\tilde{f}_k(x) < 1/2) < \mu(\pi_k^{-1} A_k) = 2^{-k} \text{ and } \sum_{k=1}^{\infty} \mu(\pi_k^{-1} A_k) = \sum_{k=1}^{\infty} 2^{-k} = 1 < \infty.$$

Thus, by Borel-Cantelli for almost every $x$ there exists $k_0' \in \mathbb{N}$ such that for all $k \geq k_0'$,

$$\sup_n \mathcal{A}_{2n}\tilde{f}_k(x) \geq 1/2.$$

Consequently, $k_0 \sup_n \mathcal{A}_{2n}\tilde{f}_k(x) \geq k_0/2$ for any $k_0 \geq 0$. Using this, for any $k_0 \geq k_0'$,

$$\sup_n \mathcal{A}_{2n} f(x) = \sup_n \mathcal{A}_{2n} \sum_{k=1}^{\infty} k\tilde{f}_k(x) \geq \sup_n \mathcal{A}_{2n} \sum_{k=k_0}^{\infty} k\tilde{f}_k(x) \geq k_0 \sup_n \mathcal{A}_{2n}\tilde{f}_{k_0}(x) \geq k_0/2.$$

This gives the pointwise inequality,

$$\sup_n \mathcal{A}_{2n} f(x) \geq k_0/2, \quad \text{for all} \quad k_0 \geq k_0'.$$

Taking the limit as $k_0 \to \infty$ implies that for almost every $x \in X$,

$$|\sup_n A_{2n} f(x)| = \infty.$$

Therefore, a function $f \in L^1(X)$ exists where $X$ satisfies the conditions in Theorem 2.3 and $|\sup_n \mathcal{A}_{2n} f(x)| = \infty$ for almost every $x \in X$, completing the proof. □

The next section will establish a simple identity between $\mathcal{A}_n$ and the iterates of a particular Markov operator, $P$. This connection to Markov operators, which are defined below, allowed [17] to construct the counterexample using similar ideas to Ornstein's counterexample to an $L^1$ maximal inequality. Further details can be found in [16] where Ornstein's construction is presented in order to prove the following weaker statement.

**Theorem 3.3.** *Let $P : L^1(X) \to L^1(X)$ be a self adjoint Markov operator on a finite measure space $(X, \mathcal{X}, \mu)$, there does not exist a constant $C$ such that*

$$\| \sup_{n \geq 0} |P^n f| \|_{L^{1,\infty}(X)} \leq C \|f\|_{L^1(X)},$$

*for all self adjoint Markov operators on finite measure spaces.*

## 3.2 The Second Reduction

Before continuing to the second reduction of Theorem 2.3, a short discussion is given on Markov processes, adapted from [11], with a focus on Markov chains, which are discrete Markov processes.

Take a measure space $(W, \mathcal{X}, p)$, in particular $W$ is the state space and $p$ a probability measure. In the context of this report the state space will be the set of generators of $F_m$ and their inverses. A Markov process models a system displaying dependent random behaviour through time. The states of this system change according to a transition probability. The simplest case of a Markov process is the Markov chain which is defined to be a Markov process with a countable number of states. For ease of notation label the states by the integers.

**Definition 3.4.** *(Transition probability and transition matrix). For a Markov chain, the one-step transition probability from state $j$ to state $k$ is*

$$p_{j,k} = p_{j,k}^{(1)} \geq 0, \quad \text{where} \quad \sum_k p_{j,k} = 1.$$

*Further, the $(n+1)$-step transition probability from state $j$ to state $k$ is given recursively by $p_{j,k}^{(n+1)} = \sum_l p_{j,l} p_{l,k}^{(n)}$, for $n \in \mathbb{N}$. The transition function is often represented by the transition matrix, which is defined to be the matrix whose $i,j$-th entry is the $n$-step transition probability from state $i$ to state $j$. That is,*

$$(\Pi^{(n)})_{i,j} = p_{i,j}^{(n)},$$

*and note also that, $\Pi^{(n)} = \Pi^n$.*

Intuitively, $p_{i,j}$ represents the probability that the Markov process will be at state $j$ at time $n$, given the process was in state $i$ at time $n-1$. In this report, the transition probability of a state $x$ to another state $y$ is equally likely unless $y$ is the inverse of $x$, in which case it is zero.

**Definition 3.5.** *(Markov operator). The operator $P$ induced by a Markov transition* *function* $\Pi$,

$$(Pf)(x) = \int \Pi(x, dy) f(y),$$

*is called a Markov operator.*

The averaging operators do not display the semigroup property, $\mathcal{A}_n \mathcal{A}_m = \mathcal{A}_{n+m}$, therefore it is not possible to directly interpret them as Markov operators. Instead, the connection shall be demonstrated, following the reasoning of Bufetov [1], by lifting $X$ up to a $2m$-fold cover $\tilde{X}$.

**Definition 3.6.** *(The lifted measure space). Given an $F_m$-system, $(X, \mathcal{X}, \mu, (T_g)_{g \in F_m})$, let $(\tilde{X}, \tilde{\mathcal{X}}, \tilde{\mu})$ denote the lifted measure space which is defined to be the product of $(X, \mathcal{X}, \mu)$ and the $2m$ element space $W$, equipped with the uniform probability measure, $\nu$. Let $\pi : \tilde{X} \to X$ be the projection operator $\pi(x, s) := x$, for $x \in X$ and $s \in W$. This induces a pushforward operator, $\pi_* : L^1(\tilde{X}) \to L^1(X)$, and a pullback operator, $\pi^* : L^1(X) \to L^1(\tilde{X})$, defined by*

$$\pi_* \tilde{f}(x) := \frac{1}{2m} \sum_{s \in W} \tilde{f}(x, s) \quad \text{and} \quad \pi^* f(x, s) := f(x)$$

*for $f \in L^1(X)$ and $\tilde{f} \in L^1(\tilde{X})$.*

In particular, $\tilde{X} = X \times W$, $\tilde{\mu} = \mu \times \nu$ where $\nu$ is the uniform measure and $\tilde{\mu}(\tilde{X}) = \mu(X)$. The notation, $(\tilde{X}, \tilde{\mathcal{X}}, \tilde{\mu})$, will be used consistently to refer to a measure space lifted in this way. Similarly, $\tilde{f}$ is used to denote a function defined on the lifted measure space. Finally, note that, $\pi^* f = f \circ \pi$.

It is interesting to observe that $\pi_*$ is acting as an average. At a particular point $x$ it takes the average value of the function $\tilde{f}$ on each copy $X \times \{s\}$ for $s \in W$.

**Lemma 3.7.** *Where $Id$ denotes the identity function $Id(x) = x$,*

*(i) For any $f \in L^1(X)$, $\pi_* \pi^* f(x) = Id(x)$ for any $x \in X$, and*

*(ii) For any $\tilde{f} \in L^1(\tilde{X})$, $\pi^* \pi_* \tilde{f}(x, s_1) = \frac{1}{2m} \sum_{s \in W} \tilde{f}(x, s)$ for any $x \in X$, $s_1 \in W$.*

*Proof.* For the first identity, take $f \in L^1(X)$, fix $x \in X$

$$\pi_* \pi^* f(x) = \frac{1}{2m} \sum_{s \in W} \pi^* f(x, s) = \frac{1}{2m} \sum_{s \in W} f(x) = f(x).$$

For the second, take $\tilde{f} \in L^1(X \times W) = L^1(\tilde{X})$, fix $(x, s_1) \in \tilde{X}$

$$\pi^* \pi_* \tilde{f}(x, s_1) = \pi_* \tilde{f}(x) = \frac{1}{2m} \sum_{s \in W} \tilde{f}(x, s) \neq \tilde{f}(x, s_1).$$

□

The Markov operator, $P$, associated with the averaging operators can now be defined.

**Definition 3.8.** *(Markov operator, $P$). The Markov operator $P : L^1(\tilde{X}) \to L^1(\tilde{X})$ is defined by*

$$P\tilde{f}(x,s) := \frac{1}{2m-1} \sum_{s_1 \in W \setminus \{s^{-1}\}} \tilde{f}(T_{s_1}^{-1}x, s_1).$$

$P$ can be viewed as the Markov operator associated to a Markov chain that is defined according to the formation of possible words in $F_m$. Given the state is a particular point $(x,s)$ in $\tilde{X}$ then the process can move to one of $2m-1$ points, according to the uniform probability distribution. These points, $(T_{s_1}^{-1}x, s_1)$, correspond to the possible words $ss_1$, with $s$ fixed, of length 2. Before stating the identity relating $P$ to $\mathcal{A}_n$, the transition probabilities, transition matrix and an invariant measure associated to $P$, will be defined, as well as two other properties of $P$.

Firstly, the transition matrix is a $2m \times 2m$ matrix whose indices lie in $S = \{-m, \dots, -1, 1, \dots, m\}$,

$$\begin{bmatrix} \frac{1}{2m-1} & \frac{1}{2m-1} & \cdots & \frac{1}{2m-1} & 0 \\ \frac{1}{2m-1} & \iddots & \iddots & 0 & \frac{1}{2m-1} \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ \frac{1}{2m-1} & 0 & \iddots & \iddots & \frac{1}{2m-1} \\ 0 & \frac{1}{2m-1} & \cdots & \frac{1}{2m-1} & \frac{1}{2m-1} \end{bmatrix}.$$

This Markov chain has an invariant measure, $\pi_i = 1/2m$ for all $i \in S$, where an invariant measure is defined to be,

**Definition 3.9.** *(Invariant measure). For a Markov chain, with $2m$ possible states and transition matrix $\Pi = (p_{i,j})$, the $2m \times 1$ vector $\pi$ is an invariant measure if:*

$$\sum_{i \in S} \pi_i = 1, \ \textit{and} \ \pi_j = \sum_{i \in S} \pi_i p_{i,j}.$$

Clearly, for $\pi$ as defined above the entries add up to one, and for all $j \in S$,

$$\sum_{i \in S} \pi_i p_{i,j} = \sum_{i \in S \setminus \{-j\}} \frac{1}{2m} \frac{1}{2m-1} = (2m-1) \frac{1}{2m-1} \frac{1}{2m} = \frac{1}{2m} = \pi_j.$$

From [1], a Markov operator $P$ is self adjoint if it satisfies the symmetry condition,

$$\pi_i = \pi_{-i} \quad \text{and} \quad p_{-i,-j} = \frac{\pi_j p_{j,i}}{\pi_i}.$$

Therefore, $P$ is self adjoint, since $\pi_i = \pi_j$ for all $i, j \in S$ and for the non-trivial case $i \neq -j$,

$$\frac{\pi_j p_{j,i}}{\pi_i} = \frac{1}{2m} \frac{1}{2m-1} (2m) = \frac{1}{2m-1} = p_{-i,-j}.$$

Finally, the transition matrix of a Markov operator generates the free group if, $p_{i,j} = 0$ if and only if $i + j = 0$. Clearly, the transition matrix of $P$ satisfies this.

The following identity formulates the relationship between $P$ and $\mathcal{A}_n$.

**Proposition 3.10.**

$$\mathcal{A}_n f = \pi_* P^n \pi^* f \quad \textit{for any } f \in L^1(X) \textit{ and } n \geq 1.$$

*Proof.* Firstly, proof by induction shows,

$$P^n \tilde{f}(x,s) = \frac{1}{(2m-1)^n} \sum_{w=s_1 s_2 \cdots s_n \in F_m, |w|=n, s_1 \neq s^{-1}} \tilde{f}(T_w^{-1} x, s_n) \quad \text{for all } n \in \mathbb{N}. \tag{1}$$

The base case follows by definition of $P$,

$$P\tilde{f}(x,s) = \frac{1}{2m-1} \sum_{s_1 \in W \setminus \{s^{-1}\}} \tilde{f}(T_{s_1}^{-1} x, s_1) = \frac{1}{2m-1} \sum_{w=s_1 \in F_m, |w|=1, s_1 \neq s^{-1}} \tilde{f}(T_w^{-1} x, s_1)$$

Now suppose (1) is true for $n = k$ then,

$$\begin{aligned}
P^{k+1}\tilde{f}(x,s) &= P^{k-1}\left( \frac{1}{2m-1} \sum_{w=s_1 \in F_m, |w|=1, s_1 \neq s^{-1}} P\tilde{f}(T_w^{-1} x, s_1) \right) \\
&= P^{k-1}\left( \frac{1}{2m-1} \sum_{w=s_1 \in F_m, |w|=1, s_1 \neq s^{-1}} \frac{1}{2m-1} \sum_{w'=s_2 \in F_m, |w|=1, s_2 \neq s_1^{-1}} \tilde{f}(T_{w'}^{-1} T_w^{-1} x, s_2) \right) \\
&= \frac{1}{(2m-1)^2} \sum_{w=s_1 s_2 \in F_m, |w|=2, s_1 \neq s^{-1}} P^{k-1}\tilde{f}(T_w^{-1} x, s_2) \\
&= \frac{1}{(2m-1)^{k+1}} \sum_{w=s_1 s_2 \in F_m, |w|=2, s_1 \neq s^{-1}} \sum_{w'=s_3 s_4 \cdots s_{k+1} \in F_m, |w|=k-1, s_3 \neq s_2^{-1}} \tilde{f}(T_{w'}^{-1} T_w^{-1} x, s_{k+1}) \\
&= \frac{1}{(2m-1)^{k+1}} \sum_{w=s_1 s_2 \cdots s_{k+1} \in F_m, |w|=k+1, s_1 \neq s^{-1}} \tilde{f}(T_w^{-1} x, s_{k+1}).
\end{aligned}$$

Thus, by induction, (1) holds. Recall $\pi^* f(x,s) = f(x)$ and $\pi_* \tilde{f}(x) := \frac{1}{2m} \sum_{s \in W} \tilde{f}(x,s)$. So for any $n \in \mathbb{N}$ and $f \in L^1(X)$,

$$\begin{aligned}
\pi_*(P^n \pi^* f)(x) &= \frac{1}{2m} \sum_{s \in W} \frac{1}{(2m-1)^n} \sum_{w=s_1 s_2 \cdots s_n \in F_m, |w|=n, s_1 \neq s^{-1}} \pi^* f(T_w^{-1} x, s_n) \\
&= \frac{1}{2m \times (2m-1)^n} \sum_{s \in W} \sum_{w=s_1 s_2 \cdots s_n \in F_m, |w|=n, s_1 \neq s^{-1}} f(T_w^{-1} x).
\end{aligned}$$

Fix any $s_1 \in W$, summing over $s \in W \setminus \{s_1^{-1}\}$, $f(T_w^{-1})$ is counted $2m-1$ times. So, the summation above is equivalent to the summation of $(2m-1) \times f(T_w^{-1} x)$ over $w = s_1 \cdots s_n$ with $|w| = n$.

$$\pi_* P^n \pi^* f(x) = \frac{1}{2m \times (2m-1)^n} \sum_{w=s_1 s_2 \cdots s_n \in F_m, |w|=n} (2m-1) \times f(T_w^{-1} x)$$

$$= \frac{1}{2m \times (2m-1)^{n-1}} \sum_{w \in F_m, |w|=n} f(T_w^{-1}x) = \mathcal{A}_n f(x).$$

□

This identity allows another reformulation of Theorem 2.3, which corresponds to [[17], Theorem 2.2].

**Theorem 3.11.** *(Quantitative counterexample 2). Let $\alpha, \epsilon > 0$. Then there exists an $F_m$-system $(X, \mathcal{X}, \mu, (T_g)_{g \in F_m})$ and a non-negative function $\tilde{f} \in L^\infty(\tilde{X})$, such that*

$$\|\tilde{f}\|_{L^1(\tilde{X})} \le \alpha\mu(X)$$

*but such that*

$$\sup_n \pi_* P^{2n} \tilde{f}(x) \ge 1 - \epsilon$$

*for all $x \in X$ outside a set of measure at most $\epsilon\mu(X)$.*

**Proposition 3.12.** *Theorem 3.11 implies Theorem 3.1.*

*Proof.* Suppose Theorem 3.11 is true. As suggested by [17], set $f := 2m\pi_* \tilde{f}$ and take $\nu$ to be the uniform measure on $W$ then,

$$\|f\|_{L^1(X)} = 2m \int_X \frac{1}{2m} \sum_{s \in W} \tilde{f}(x,s) d\mu(x) = 2m \int_X \int_W \tilde{f}(x,s) d\nu \times \mu(x,s) = 2m \|\tilde{f}\|_{L^1(\tilde{X})}.$$

In Theorem 3.11 take $\alpha = \alpha/2m$, then

$$\frac{1}{2m}\|f\|_{L^1(X)} = \|\tilde{f}\|_{L^1(\tilde{X})} \le \frac{\alpha}{2m}\mu(X) \quad \text{and so} \quad \|f\|_{L^1(X)} \le \alpha\mu(X).$$

To show the second part of Theorem 3.1, first note that for any $s$ and any $x$

$$\pi^* f(x,s) = f(x) = 2m\pi_* \tilde{f}(x) = \sum_{s \in W} \tilde{f}(x,s) \ge \tilde{f}(x,s).$$

Combining this with the identity $\mathcal{A}_n f = \pi_* P^n \pi^* f$,

$$\sup_n \mathcal{A}_{2n} f(x) = \sup_n \pi_* P^{2n} \pi^* f(x) \ge \sup_n \pi_* P^{2n} \tilde{f}(x) \ge 1 - \epsilon$$

for all $x \in X$ outside a set of measure at most $\epsilon\mu(X)$, concluding the proof. □

# 4 Good Systems

This section will introduce the notion of a good system. These are $F_m$-systems with certain desirable properties associated to the generating maps $T_{c_1}, \dots, T_{c_m}$. It is a good system on which the function in Theorem 3.11 is constructed.

**Definition 4.1.** *(A good system). A good system, is an $F_m$-system, $(X, \mathcal{X}, \mu, (T_g)_{g\in F_m})$, which admits a decomposition $X = X_1 \cup X_2 \cup X_0$ into three disjoint sets $X_0, X_1, X_2$ with the properties:*

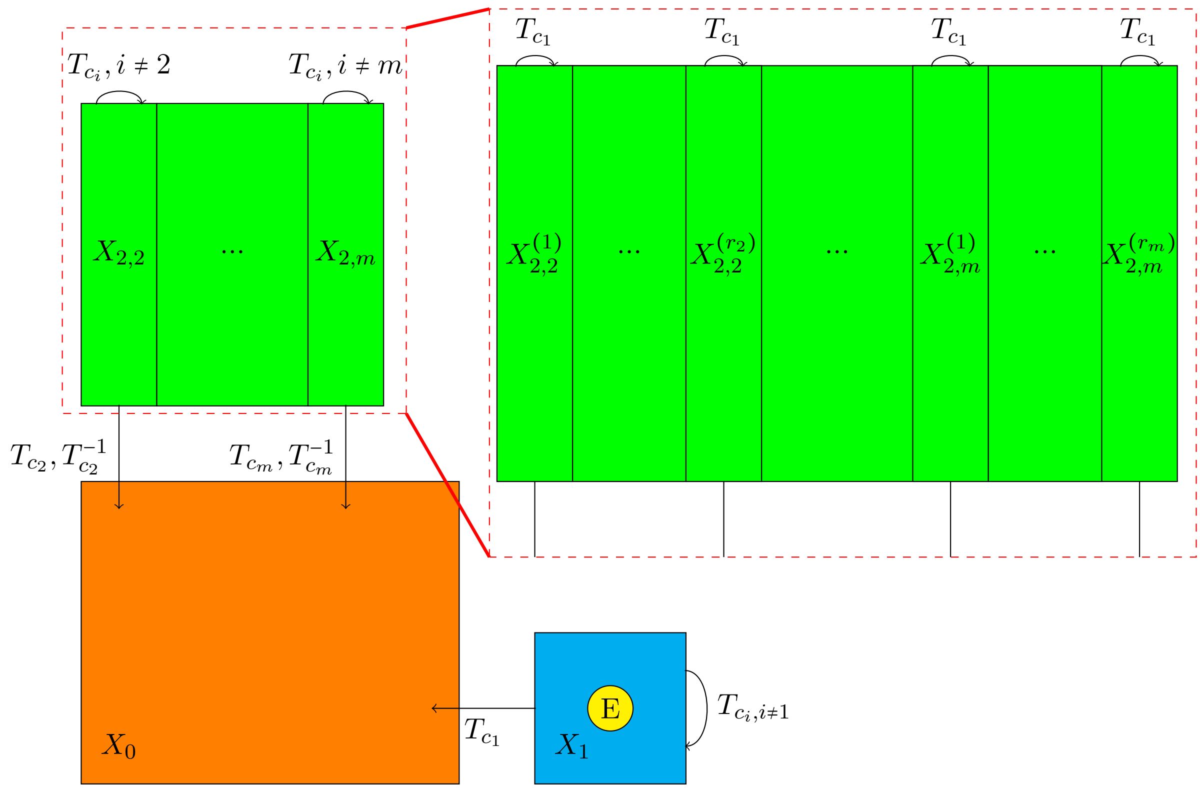
*(i) (Measure) $\mu(X_1) = \frac{m-1}{m^2}\mu(X)$, $\mu(X_2) = \frac{(m-1)^2}{m^2}\mu(X)$ and $\mu(X_0) = \frac{1}{m}\mu(X)$. Furthermore, for any $0 \le \kappa \le \mu(X_1)$, there exists a measurable subset of $X_1$, $B$, such that $\mu(B) = \kappa$.*

*(ii) (a)(Partition) $X_2$ can be partitioned into $m-1$ components $X_{2,2}, \dots, X_{2,m}$ of positive measure.*
*(b)(Ergodicity) For each $i = 2, \dots, m$, there exists $r_i < \infty$, such that $X_{2,i}$ can be partitioned into $T_{c_1}$-invariant components, $X_{2,i}^{(1)}, \dots, X_{2,i}^{(r_i)}$ and $T_{c_1}^2$ is ergodic on each of the components $X_{2,i}^{(k)}$; that is the only $T_{c_1}^2$-invariant measurable subsets of $X_{2,i}^{(k)}$ have either measure 0 or $\mu(X_{2,i}^{(k)})$ for $k = 1, \dots, r_i$.*

*(iii) (Invariance) $T_{c_1}X_2 = X_2$. For $i = 2, \dots, m$, $T_{c_i}X_1 = X_1$ and $T_{c_i}X_{2,j} = X_{2,j}$ for all $j \ne i$. The inclusions $T_{c_1}X_1 \subset T_{c_i}X_{2,i} \cup T_{c_i}^{-1}X_{2,i} \subset X_0$ hold for any $i = 1, \dots, m$.*

*(iv) (Generation) One has $X = \bigcup_{g\in F_m} T_g X_{2,i}^{(k)}$ up to null sets for each $k = 1, \dots, r_i$, and $i = 2, \dots, m$.*

Figure 1: This is a sketch of the structure of a good system along with part of the actions of $T_{c_i}, = 1, \dots, m$. A magnified view of $X_2$ is shown to indicate the second splitting of $X_2$ in Axiom (ii)(b). Further detail could be added to depict Axiom (iii).




A good system is split into $X_1$, $X_2$ and $X_0$, where in later constructions the boundary of

$X_0$ will consist of $X_1$ and $X_2$. $X_2$ is split up further into $m-1$ components and so the boundary of $X_0$ is made up of $m$ parts (that correspond to the $m$ generating maps, $T_{c_i}$). The second splitting of $X_{2,i}$ occurs due to the steps of the proof in Section 7 when two systems are glued together. If this second splitting is not specified then the new system is no longer a good system, which is required for the proof. As [17] notes, the relatively fewer conditions on $X_1$, in particular no assumption of ergodicity, will allow the modification of the maps $T_{c_1}, T_{c_2}, \ldots, T_{c_m}$ on $X_1$ in order to join two of these systems via $X_1$. Finally, observe $\mu(X_1 \cup X_2) = \frac{m-1}{m}\mu(X)$. A diagram of a good system is shown in Figure 1.
The following is from [[1], Lemma 1],

**Lemma 4.2.** *[1] Suppose the transition matrix of $P$, $\Pi$, generates the free group and satisfies the symmetry condition. Suppose the action of $F_m^2$ on the probability space $(X, \mathcal{X}, \mu)$ is ergodic. Then for any $\tilde{f} \in L\log L(\tilde{X})$, (where $\tilde{X} = X \times W$)*

$$P^n \tilde{f} \to \int_{\tilde{X}} \tilde{f} d\tilde{\mu},$$

*both $\tilde{\mu}$-almost everywhere and in $L^1(\tilde{X})$.*

Using Lemma 4.2 and Theorem 1.9, it can be shown good systems satisfy the following,

**Theorem 4.3.** *(Pointwise ergodic theorem for good systems). Every good system $(X, \mathcal{X}, \mu, (T_g)_{g \in F_m})$ is $F_m^2$-ergodic. In particular (by Theorem 1.9), for any $f \in L^\infty(X)$, the averages $\mathcal{A}_{2n} f$ converge pointwise almost everywhere and in $L^1$ norm to $\frac{1}{\mu(X)} \int_X f d\mu$. Furthermore, for any $\tilde{f} \in L^\infty(\tilde{X})$, $P^{2n}\tilde{f}$ converge pointwise almost everywhere and in $L^1$ norm to $\frac{1}{\mu(X)} \int_{\tilde{X}} \tilde{f} d\tilde{\mu}$.*

*Proof.* The proof uses the alternative characterisation of ergodicity, see the appendix, Proposition 9.1 and follows [[17], Lemma 2.3]. $F_m = F_m^2 \cup F_m^2 c_1$ since, for any $h \in F_m$, either $h \in F_m^2$ and $h$ has even length, or $h \in F_m^2 c_1$ and the word length is odd. Fix $i \in \{2, \ldots, m\}$ and $j \in \{1, \ldots, r_i\}$, by Axiom (iv) and the $T_{c_1}$-invariance of $X_{2,i}^{(j)}$, $X = \bigcup_{g \in F_m} T_g X_{2,i}^{(j)}$ up to null sets and $T_{c_1} X_{2,i}^{(j)} = X_{2,i}^{(j)}$. So,

$$X = \bigcup_{g \in F_m} T_g X_{2,i}^{(j)} = \bigcup_{g \in F_m^2 \cup F_m^2 c_1} T_g X_{2,i}^{(j)} = \bigcup_{g \in F_m^2} \left( T_g X_{2,i}^{(j)} \cup T_g T_{c_1} X_{2,i}^{(j)} \right) = \bigcup_{g \in F_m^2} T_g X_{2,i}^{(j)}.$$

Let $f \in L^\infty(X)$ ($\Rightarrow f \in L^1(X)$), be an $F_m^2$-invariant function i.e.

$$f \circ T_g = f, \quad \mu - a.e. \quad \text{for all } g \in F_m \text{ where } |g| \text{ is even, in particular } f \circ T_{c_1}^2 = f, \quad \mu - a.e.$$

Furthermore, by Axiom (ii)(b) of a good system, $T_{c_1}^2$ is ergodic on $X_{2,i}^{(j)}$ which implies $f$ is constant almost everywhere on $X_{2,i}^{(j)}$, say $c$. Moreover, for any $g \in F_m^2$, $f \circ T_g = f$ so $f$ is also equal to $c$ almost everywhere on $\bigcup_{g \in F_m^2} T_g X_{2,i}^{(j)} = X$ and every good system is $F_m^2$-ergodic. Now, using that a good system is $F_m^2$-ergodic, $f \in L^\infty(X)$ ($\Rightarrow f \in L\log L$) and Theorem 1.9, $\mathcal{A}_{2n} f$ converges pointwise almost everywhere and in $L^1$ to the constant $\frac{1}{\mu(X)} \int_X f d\mu$. The final statement follows from Lemma 4.2. Since $(X, \mathcal{X}, \mu)$ has finite measure, it is possible to normalise the measure by dividing $\mu$ by $\mu(X)$. $\Pi$ is the transition matrix associated with the Markov operator $P$ and it was shown earlier that $\Pi$ satisfies the symmetry condition and generates the free group. The previous part of the proof established that the action of $F_m^2$ on $(X, \mathcal{X}, \mu)$ is ergodic and $\tilde{f} \in L\log L(\tilde{X})$. Thus the assumptions of the Lemma 4.2 are satisfied and,

$$P^n \tilde{f} \to \int_{\tilde{X}} \tilde{f} d\frac{\tilde{\mu}}{\mu(X)} = \frac{1}{\mu(X)} \int_{\tilde{X}} \tilde{f} d\tilde{\mu}.$$

Thus,

$$P^{2n} \tilde{f} \to \frac{1}{\mu(X)} \int_{\tilde{X}} \tilde{f} d\tilde{\mu} \quad \text{both } \tilde{\mu}\text{-almost everywhere and in } L^1(\tilde{X}).$$

□

## 5 The Main Claims

This section will formulate two theorems, from which Theorem 3.11 can be deduced and the following two sections will then prove them. The theorems are based around constructing a sequence of functions on a good system satisfying certain properties. The following is the generalisation of [[17], Claim 2.4]. For any $\alpha > 0$, let $P(\alpha)$ denote the claim that:

**Claim 5.1.** $(P(\alpha))$. *For any $\epsilon > 0$, there exists a good system $(X, \mathcal{X}, \mu, (T_g)_{g \in F_m})$ with associated decomposition $X = X_1 \cup X_2 \cup X_0$, and a sequence of non-negative functions $\tilde{f}_n \in L^\infty(\tilde{X})$ for $n \in \mathbb{Z}$ with the following properties:*

- *(v) (Ancient Markov chain) $\tilde{f}_{n+1} = P\tilde{f}_n$ for all $n \in \mathbb{Z}$. Equivalently, $\tilde{f}_{n+m} = P^m \tilde{f}_n$ for all $n \in \mathbb{Z}$ and $m \in \mathbb{N}$. In particular, $\|\tilde{f}_n\|_{L^1(\tilde{X})}$ is independent of $n$.*
- *(vi) (Size) $\|\tilde{f}_n\|_{L^1(\tilde{X})} = \alpha\mu(X)$ for some $n \in \mathbb{Z}$ (and hence for all $n \in \mathbb{Z}$).*
- *(vii) (Early support) $\tilde{f}_n$ is supported in $\tilde{X}_0$ for all negative $n$. Furthermore, there exists a finite $A > 0$ such that $\tilde{f}_n$ is supported in a set of measure at most $A(2m-1)^n\mu(X)$, for all negative $n$.*
- *(viii) (Large maximum function)*
$$\sup_{n \in \mathbb{Z}} \pi_* \tilde{f}_{2n}(x) \geq 1 - \epsilon$$
*for all $x \in X$ outside of a set of measure at most $\epsilon\mu(X)$.*

The sequence $(P^n\tilde{f})_{n \geq 0}$ is extended to negative $n$, [17] describes this as $(P^n\tilde{f})_{n \geq 0}$ belonging to an 'ancient Markov chain' $(\tilde{f}_n)_{n \in \mathbb{Z}}$. This technical property allows a time delay to be put on the functions and in Section 7, when two systems are glued together using a suitable amount of coupling, the functions experience a delay on one measure space but not the other.

**Lemma 5.2.** *$\|\tilde{f}_n\|_{L^1(\tilde{X})}$ is independent of $n$.*

*Proof.* Recall, $\tilde{X}$ is defined to be the product of $(X, \mathcal{X}, \mu)$ and the $2m$ element space $W$ with the uniform probability measure $\nu$. Fix $n \in \mathbb{Z}$

$$\begin{aligned}
\|\tilde{f}_{n+1}\|_{L^1(\tilde{X})} &= \int_{\tilde{X}} P\tilde{f}_n(x,s)d\tilde{\mu}(x,s) \\
&= \int_{X\times W} \frac{1}{2m-1} \sum_{s_1 \in W \setminus s^{-1}} \tilde{f}_n(T_{s_1}^{-1}x, s_1)d(\mu\times\nu)(x,s) \\
&= \int_X \frac{1}{2m} \sum_{s\in W} \frac{1}{2m-1} \sum_{s_1 \in W\setminus s^{-1}} \tilde{f}_n(T_{s_1}^{-1}x, s_1)d\mu(x).
\end{aligned}$$

For every $s_1 \in W$, summing over $s \in W \setminus \{s_1^{-1}\}$, $\tilde{f}_n(T_{s_1}^{-1}x, s_1)$ is counted $2m-1$ times.

$$\begin{aligned}
\|\tilde{f}_{n+1}\|_{L^1(\tilde{X})} &= \int_X \frac{1}{2m}\frac{1}{2m-1} \sum_{s_1\in W} (2m-1)\times \tilde{f}_n(T_{s_1}^{-1}x, s_1)d\mu(x) \\
&= \frac{1}{2m} \sum_{s_1 \in W} \int_X \tilde{f}_n(T_{s_1}^{-1}x, s_1)d\mu(x) \\
&= \|\tilde{f}_n\|_{L^1(\tilde{X})}.
\end{aligned}$$

So, for all $n \in \mathbb{Z}$, $\|\tilde{f}_n\|_{L^1(\tilde{X})}$ is independent of $n$. □

By Axiom (vi), $\|\tilde{f}_n\|_{L^1(\tilde{X})} = \alpha\mu(X)$ for some $n \in \mathbb{Z}$. Therefore, since $\|\tilde{f}_n\|_{L^1(\tilde{X})}$ is independent of $n$, $\|\tilde{f}_n\|_{L^1(\tilde{X})} = \alpha\mu(X)$ for all $n \in \mathbb{Z}$.

**Theorem 5.3.** *If $P(\alpha)$ holds for an arbitrarily small set of $\alpha > 0$ then Theorem 3.11 holds.*

*Proof.* Assume that $P(\alpha)$ holds for an arbitrarily small set of $\alpha > 0$, and $\epsilon > 0$ is arbitrary. Claim: For any $N$,

$$\sup_{n\geq -2N} \pi_* \tilde{f}_{2n}(x) \geq 1-\epsilon$$

for all $x \in X$ outside of a set of measure at most $(\epsilon + \frac{1}{(2m-1)^2-1}A(2m-1)^{-2N})\mu(X)$.
Observe how similar this claim is to Axiom (viii), where $\sup_{n\in\mathbb{Z}} \pi_*\tilde{f}_{2n}(x) \geq 1-\epsilon$, for all $x \in X$ outside of a set of measure at most $\epsilon\mu(X)$.
However, if in Axiom (viii) the supremum is restricted to $n \geq -2N$, the inequality does not necessarily hold. So to use Axiom (viii), it is necessary to quantify what information was lost by discarding the functions, $\tilde{f}_{2n}$ for $n < -2N$.
Let $supp(\tilde{f}_{2n})$ denote the support of $\tilde{f}_{2n}$. If $x \notin supp(\tilde{f}_{2n})$, then $\tilde{f}_{2n}$ is 0 and removing $\tilde{f}_{2n}$ will not affect whether the supremum is larger than $1-\epsilon$. Fix $N$ and take $x \in supp(\tilde{f}_{2n})$ with $n < -2N$, then $n$ might have been the only value for which the inequality was satisfied. So taking the supremum over $n \geq -2N$ for any $x \in \bigcup_{n\leq -2N-1} supp(\tilde{f}_{2n})$,

$$\sup_{n\geq -2N} \pi_* \tilde{f}_{2n}(x) \geq 1-\epsilon$$

may not hold.

Axiom (vii) says, there exists a finite $A > 0$ such that $\tilde{f}_n$ is supported in a set of measure at most $A(2m-1)^n \mu(X)$ for all negative $n$. So,

$$\begin{aligned}\mu\Big(\bigcup_{n \leq -2N-1} supp(\tilde{f}_{2n})\Big) &\leq \sum_{n \leq -2N-1} \mu(supp(\tilde{f}_{2n})) \\ &\leq (2m-1)^{-2N-2} \sum_{n=0}^{\infty} A((2m-1)^2)^{-n} \mu(X) \\ &= \frac{1}{(2m-1)^2 - 1} A(2m-1)^{-2N} \mu(X).\end{aligned}$$

Moreover, Axiom (viii) holds for $x \in X$ outside a set of measure at most $\epsilon \mu(X)$. Therefore, for any $N$, $\sup_{n \geq -2N} \pi_* \tilde{f}_{2n}(x) \geq 1 - \epsilon$, for all $x \in X$ outside of a set of measure at most $(\epsilon + \frac{1}{(2m-1)^2-1} A(2m-1)^{-2N})\mu(X)$.
To obtain Theorem 3.11 take $\tilde{f} := \tilde{f}_{-2N}$ and recall the sequence of functions $\tilde{f}_n$ is part of an ancient Markov chain $(P^n \tilde{f})$, then

$$\sup_{n \in \mathbb{Z}} \pi_* P^{2n} \tilde{f}(x) = \sup_{n \in \mathbb{Z}} \pi_* P^{2n} \tilde{f}_{-2N}(x) = \sup_{n \geq -2N} \pi_* \tilde{f}_{2n}(x) \geq 1 - \epsilon.$$

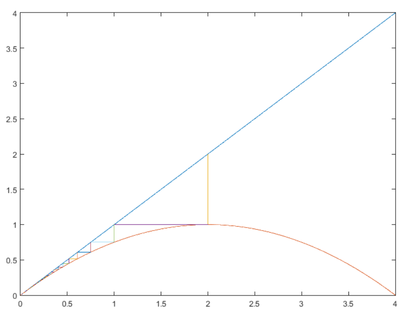
This is true for all $x \in X$ outside a set of measure at most $(\epsilon + \frac{1}{(2m-1)^2-1} A(2m-1)^{-2N})\mu(X)$. Now take $N$ large enough, depending on $\epsilon$ and $A$, such that $\epsilon$ in Theorem 3.11 is greater than $(\epsilon + \frac{1}{(2m-1)^2-1} A(2m-1)^{-2N})\mu(X)$. Thus for $\alpha, \epsilon > 0$ there exists an $F_m$-system $(X, \mathcal{X}, \mu, (T_g)_{g \in F_m})$ and a non-negative function $\tilde{f} := \tilde{f}_{-2N} \in L^\infty(\tilde{X})$, such that by Axiom (vi) satisfies $\|\tilde{f}\|_{L^1(\tilde{X})} \leq \alpha \mu(X)$ but such that $\sup_{n \in \mathbb{Z}} \pi_* P^{2n} \tilde{f}(x) \geq 1 - \epsilon$, for all $x \in X$ outside a set of measure at most $\epsilon \mu(X)$. □

**Theorem 5.4.** *$P(\alpha)$ holds for arbitrarily small $\alpha > 0$.*

Theorem 5.4 is the final reformulation of Theorem 2.3 and can be deduced from the following two theorems.
**Theorem 5.5.** *(Initial construction). The claim $P(\frac{2m-2}{m})$ is true.*

The following is an adaptation of [[7], Lemma 4], details of which can be found in [16].
**Theorem 5.6.** *(Iteration step). Suppose that $P(\alpha)$ holds for some $0 < \alpha \leq \frac{2m-2}{m}$. Then $P(\alpha(1-\alpha/4))$ is true.*
**Proposition 5.7.** *Theorem 5.5 and Theorem 5.6 are sufficient to show the existence of a counterexample.*

Figure 2: Cobweb diagram: $y = x(1 - x/4)$, $x_0 = 2$.




*Proof.* Theorems 5.5 and 5.6 imply the infimum of all $0 < \alpha \leq \frac{2m-2}{m} < 2$ for which $P(\alpha)$ holds is zero. This is because the iterative map $x_{n+1} = x_n(1 - x_n/4)$ has a fixed point at zero and iterating the map, $x_{n+1} \to 0$ if $x_1 \leq 2$. Therefore, assuming Theorem 5.5, applying

Theorem 5.6 iteratively, starting with $\frac{2m-2}{m}$, it is possible to obtain the statement $P(\alpha)$ for arbitrarily small $0 < \alpha < \frac{2m-2}{m}$. □

By looking at the associated cobweb diagram in Figure 2, it is clear that Proposition 5.7 holds.

# 6 The Initial Construction

This section will present the proof that $P(\frac{2m-2}{m})$ is true by detailing the construction of a good system $(X, \mathcal{X}, \mu, (T_g)_{g\in F_m})$ and functions $\tilde{f}_n$ which satisfy $P(\frac{2m-2}{m})$ *for every* $\epsilon > 0$. The measure space, $(X, \mathcal{X}, \mu)$, is built up from smaller pieces as follows:

- For each integer $n$, let $Y_n$ denote the space of half-infinite reduced words $(s_k)_{k\geq n} = s_n s_{n+1} s_{n+2} \cdots$, where $s_k \in W$, $\forall k$ and $c_i$ and $c_i^{-1}$ are never adjacent, for any $i$. In other words, $Y_n$ is the set of half-infinite reduced words that lie in $F_m$, starting in position $n$.
- This space will have the product $\sigma$-algebra $\mathcal{Y}_n$ defined to be the minimal $\sigma$-algebra for which the coordinate maps $(s_k)_{k\geq n} \to s_{k_0}$ are measurable for all $k_0 \geq n$.
- By the Kolmogorov extension theorem, construct the probability measure, $\mu_n$ on $\mathcal{Y}_n$, by defining the measure such that each finite reduced subword $s_n \cdots s_{n+r}$ for $r \geq 0$ occurs as an initial segment with measure $\frac{1}{2m\times(2m-1)^r}$.

Then take the disjoint union, $Y := \uplus_{n\in\mathbb{Z}} Y_n$, of the pieces $Y_n$. $Y$ admits an action $(S_g)_{g\in F_m}$ of $F_m$, with the action $S_s$ of a generator $s \in W$ defined for $s_n s_{n+1} s_{n+2} \cdots \in Y_n$, by setting

$$S_s(s_n s_{n+1} s_{n+2} \cdots) := s s_n s_{n+1} s_{n+2} \cdots \in Y_{n-1}, \quad \text{if } s \in W \setminus \{s_n^{-1}\}, \text{ and}$$

$$S_s(s_n s_{n+1} s_{n+2} \cdots) := s_{n+1} s_{n+2} \cdots \in Y_{n+1}, \quad \text{if } s = s_n^{-1}.$$

Similarly, the action of $S_g$ on $Y$ for a general reduced word, $g = w_1 \cdots w_n \in F_m$, of length $|g| = n$ can be defined. Thus $S_g$ is the operation of formal left-multiplication by $g$, after reducing any non-reduced words. Define the measure on $Y$ as $\mu_Y := \sum_{n\in\mathbb{Z}} (2m-1)^{-n} \mu_n$ then,

**Corollary 6.1.** *The action* $(S_g)_{g\in F_m}$ *is measure preserving with respect to the measure* $\mu_Y := \sum_{n\in\mathbb{Z}} (2m-1)^{-n} \mu_n$.

*Proof.* For any $k \geq 0$, let $A_{r,k} = s_r s_{r+1} \cdots s_{r+k} \cdots \subset \mathcal{Y}_r$ be the set of reduced words lying in $Y_r$ where the first $k+1$ entries are fixed. Then, since $A_{r,k}$ only consists of elements in $Y_r$,

$$\mu_Y(A) = \sum_{n\in\mathbb{Z}} (2m-1)^{-n} \mu_n(A) = (2m-1)^{-r} \mu_r(A) = (2m-1)^{-r} \frac{1}{2m\times(2m-1)^k} = \frac{1}{2m\times(2m-1)^{r+k}}.$$

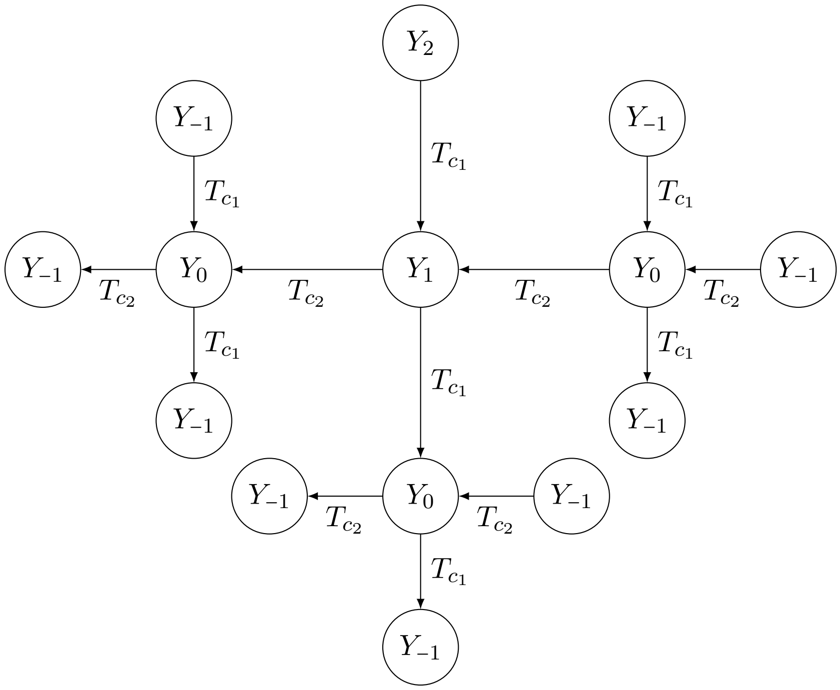



Figure 3: As displayed in [17], a fragment of the infinite measure space $Y$ in the case of 2 generators. The centre disk represents a portion of $Y_1$ consisting of reduced words $s_1 s_2 \cdots$ with initial letter $s_1 = c_1$. The remaining disks are images of this disk under shifts by various elements of $F_2$ and all have equal measure with respect to $\mu_Y$. This image should be compared with the infinite tree that is the Cayley graph of $F_2$, see Figure 4.

Consider the action $S_{c_1}$. Take $A_{r,k}$ as previously defined then either $s_r = c_1$ or $s_r \neq c_1$. In the latter case, $S_{c_1}^{-1}(s_r s_{r+1} \cdots s_{r+k} \cdots) = c_1^{-1} s_r s_{r+1} \cdots s_{r+k} \cdots$. Therefore,

$$\mu_Y(S_{c_1}^{-1}A) = \sum_{n \in \mathbb{Z}} (2m-1)^{-n} \mu_n(S_{c_1}^{-1}A) = (2m-1)^{-(r-1)} \mu_{r-1}(S_{c_1}^{-1}A) = \frac{1}{2m \times (2m-1)^{r+k}},$$

and in the case where $s_r = c_1$, $\mu_Y(S_{c_1}^{-1}A) = \frac{1}{2m \times (2m-1)^{r+k}}$. So $\mu_Y(S_{c_1}^{-1}(A)) = \mu_Y(A)$. Then, by the Kolmogorov extension theorem if all sets in the algebra that generates the $\sigma$-algebra exhibit the measure preserving property (as has just been shown) then the measure preserving property holds for any set in the $\sigma$-algebra. By similar calculations, this holds for all generators $c_i \in W$. Finally, for $g \in F_m$, $S_g$ are compositions of the maps $S_{c_i}$, so it follows they are also measure preserving. $\square$

As in [17] it is necessary to restrict $Y$ to the space $\uplus_{n \geq 0} Y_n = \uplus_{n \geq 1} Y_n \uplus Y_0$, otherwise the contribution from negative $n$ in $\mu_Y$ means it has infinite measure. Now, the shift maps $S_s$, $s \in W$ are partially undefined on the $Y_0$ boundary, however this is will be solved by redefining these maps on a quotient of $Y_0$. As [17] notes, this new space $\uplus_{n \geq 0} Y_n$ can be thought of as a suitably rescaled limit of an infinitely large ball in $F_m$ with $Y_0$ being the boundary of this ball and $Y_n$ lying increasingly deeper in the interior of the ball as $n$ increases, see Figure 3 for an illustration. Next, define a reflection operator $x \to \bar{x}$ on the

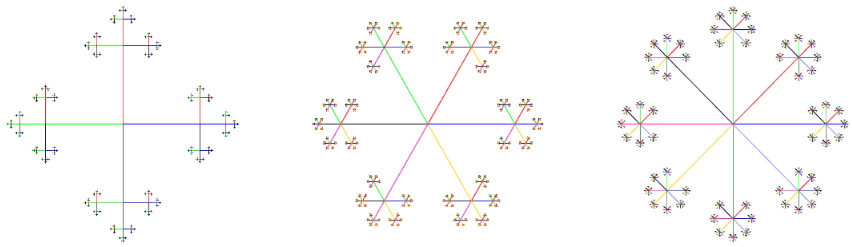


Figure 4: The Cayley graph on two, three and four generators. Comparing this to Figure 3, it gives an idea of what the illustration would look like for the general case on $m$ generators.

boundary $Y_0$ by mapping

$$\overline{s_0 s_1 s_2 \cdots} := s_0^{-1} s_1^{-1} s_2^{-1} \cdots.$$

**Corollary 6.2.** *The reflection map preserves the measure $\mu_0$.*

*Proof.* The measure $\mu_0$ is defined on the generating $\sigma$-algebra of words $x = s_0 s_1 \cdots s_k \cdots$ whose first $k+1$ letters are fixed by $\mu_0(x) = \frac{1}{2m \times (2m-1)^k}$, and so

$$\mu_0(\overline{s_0 s_1 \cdots s_k \cdots}) = \mu_0(s_0^{-1} s_1^{-1} \cdots s_k^{-1} \cdots) = \frac{1}{2m \times (2m-1)^k} = \mu_0(s_0 s_1 \cdots s_k \cdots).$$

Therefore, by the Kolmogorov extension theorem and the definition of measure preserving, the reflection map preserves the measure $\mu_0$. $\square$

Define the quotient space $Y_0/\sim := \{\{x, \bar{x}\} : x \in Y_0\}$ with corresponding probability measure, $m_0 := \mu_0/\sim$ on $Y_0/\sim$. The probability measure $m_0$ on $Y_0/\sim$, is defined by pushing forward the probability measure $\mu_0$ under the quotient map. The space $Y_0/\sim$ can be separated into the $m$ disjoint sets corresponding to the $m$ generators being in the 0-th position.

$$\begin{aligned} m_0(Y_0/\sim) &= m_0(c_1 \cdots \cup c_2 \cdots \cup \cdots \cup c_m \cdots) \\ &= \mu_0(c_1 \cdots \cup c_2 \cdots \cup \cdots \cup c_m \cdots \cup c_1^{-1} \cdots \cup c_2^{-1} \cdots \cup \cdots \cup c_m^{-1} \cdots) \\ &= \mu_0(c_1 \cdots) + \mu_0(c_1^{-1} \cdots) + \cdots + \mu_0(c_m \cdots) + \mu_0(c_m^{-1} \cdots) \\ &= \frac{1}{2m} + \frac{1}{2m} + \cdots + \frac{1}{2m} + \frac{1}{2m} = 1 \end{aligned}$$

**Lemma 6.3.** *$Y_0/\sim$ splits into 2 components*

$$((S_{c_1} Y_1 \cap Y_0) \cup (S_{c_1}^{-1} Y_1 \cap Y_0))/\sim \quad \textit{and} \quad \bigcup_{i=2,\ldots,m} ((S_{c_i} Y_1 \cap Y_0) \cup (S_{c_i}^{-1} Y_1 \cap Y_0))/\sim .$$

*with measure $\frac{1}{m}$ and $\frac{m-1}{m}$, respectively. The sets $S_{c_i} Y_1 \cap Y_0$, $S_{c_i}^{-1} Y_1 \cap Y_0$ are disjoint reflections of each other for $i = 1, \ldots, m$.*

*Proof.* Since both cases are analogous, take $((S_{c_1} Y_1 \cap Y_0) \cup (S_{c_1}^{-1} Y_1 \cap Y_0))/\sim$. $Y_1$ is the set of half infinite words starting at index 1. Let $x = s_1 s_2 \cdots \in Y_1$, $S_{c_1}$ corresponds to left multiplication by $c_1$, thus $S_{c_1} Y_1$ is the set of infinite words $c_1 s_1 s_2 \cdots$. Intersecting this with $Y_0$, $S_{c_1} Y_1 \cap Y_0$, it must be true that $s_1 \neq c_1^{-1}$ in order that when the expression $c_1 s_1 s_2 \cdots$ is reduced, it lies in $Y_0$. Similarly, $S_{c_1}^{-1} Y_1$ is the set of infinite words $c_1^{-1} s_1 s_2 \cdots \in Y_0$ if $s_1 \neq c_1$ or $s_2 \cdots \in Y_2$ if $s_1 = c_1$ and taking the intersection with $Y_0$ it follows $s_1 \neq c_1$ so that $S_{c_1}^{-1} s_1 s_2 \cdots$ lies in $Y_0$. The two sets $S_{c_1} Y_1 \cap Y_0$, $S_{c_1}^{-1} Y_1 \cap Y_0$ are disjoint since the first is all elements in $Y_1$ with starting element $c_1$ and the second is all elements in $Y_1$ with starting element $c_1^{-1}$. To find the measure of $((S_{c_1} Y_1 \cap Y_0) \cup (S_{c_1}^{-1} Y_1 \cap Y_0))/\sim$, consider the words in $Y_0$ beginning with $c_1$, which have measure $1/2m$ and the words in $Y_0$ beginning with $c_1^{-1}$, which also have measure $1/2m$. Under the equivalence relation $\sim$ these are the same set and therefore the measure under $m_0$ is $1/m$. Using the same reasoning, the measure of

$$\bigcup_{i=2,\dots,m} ((S_{c_i} Y_1 \cap Y_0) \cup (S_{c_i}^{-1} Y_1 \cap Y_0))/\sim \quad \text{is} \quad \tfrac{m-1}{m}.$$

Finally, $Y_0/\sim$ is the set of all possible words starting at index 0. It can be split according to the starting letter of the word. Since each element can occur in index 0 in only one of the sets defined above, it is true that they split $Y_0/\sim$ into two disjoint sets. □

Putting this altogether, $X$ can be defined as the quotient space $\uplus_{n\geq 1} Y_n \uplus (Y_0/\sim)$ with measure $\mu := \sum_{n\geq 1} (2m-1)^{-n} \mu_n + \frac{1}{2} m_0$. Since $\mu_n(Y_n) = 1$ for all $n$,

$$\mu(X) = \sum_{n\geq 1} (2m-1)^{-n} + 1/2 = \tfrac{m}{2(m-1)}.$$

Dividing by $\frac{m}{2(m-1)}$, it is possible to normalise $\mu$ however, this has no advantages and clutters the notation. Set $X_0 := \uplus_{n\geq 1} Y_n$, $X_1 := ((S_{c_1} Y_1 \cap Y_0) \cup (S_{c_1}^{-1} Y_1 \cap Y_0))/\sim$ and $X_2 = \bigcup_{i=2,\dots,m} ((S_{c_i} Y_1 \cap Y_0) \cup (S_{c_i}^{-1} Y_1 \cap Y_0))/\sim$ then letting $X = X_1 \cup X_2 \cup X_0$, it is required to show $X$ is a good system.

**Lemma 6.4.** *It is possible to construct measurable subsets of $X_1$ of arbitrary measure between 0 and $\mu(X_1) = \frac{1}{2m}$.*

*Proof.* An atom is a measurable set of positive measure that does not contain a set of smaller positive measure. Non-atomic measures have a continuum of values; that is if $\mu$ is an atomless measure and $A$ is a measurable set such that $\mu(A) > 0$ then for any $b$ such that $\mu(A) > b > 0$ there exists $B$ measurable such that $\mu(B) = b$.

Therefore, the objective is to prove $X_1$ is an atomless space. Take $s_1 s_2 \cdots s_k \cdots$ in $X_1$ which has measure $\frac{1}{2m\times(2m-1)^{k-1}}$ then it is possible to find a set of smaller measure, namely the set $s_1 \cdots s_{k+1} \cdots$ with measure $\frac{1}{2m\times(2m-1)^k}$. The fact that, $\bigcup_{k=0}^{\infty} \{\frac{r}{(2m-1)^k} : r = 1, \dots, (2m-1)^k\}$, is dense in the unit interval and the Kolmogorov extension theorem, concludes the proof. □

In [17], this property was deduced by observing that $X_1, X_2$ are Cantor spaces. In order to demonstrate Axiom (ii)(a), split $X_2$ into $X_{2,i}$ with $i = 2, \dots, m$ where $X_{2,i} = ((S_{c_i} Y_1 \cap$

$Y_0) \cup (S_{c_i}^{-1} Y_1 \cap Y_0))/\sim$. For Axiom (ii)(b), $r_i = 1$ for each $i = 2, \ldots, m$ therefore, $X_{2,i}^{(1)} = X_{2,i}$. Each $X_{2,i}$ can be identified as a measure space up to null sets to the unit circle with Haar measure. It is a standard result that the irrational translation map on the circle with respect to the Harr measure is ergodic. Thus, setting $T_{c_1}^2$ to be an irrational translation map gives a measure preserving invertible map, $(T_{c_1}^0)^2$, which is ergodic on $X_{2,i}$.

**Lemma 6.5.** *The set $X_{2,i}$ is homeomorphic to the unit circle.*

*Proof.* Fix $i$ then $X_{2,i}$ is the set of infinite words starting at index 0 beginning with $c_i$ or $c_i^{-1}$, $c_i s_1 s_2 \cdots$, with $s_j \in W$. Therefore, there is a bijection between these words and the base-$m$ expansions of the numbers between 0 and 1. Furthermore, the circle and interval are homeomorphic, so it is possible to identify the set $X_{2,i}$ with the unit circle up to null sets. □

**Definition 6.6.** *The maps $(T_{c_i})_{i=1,\ldots,m}$ are defined as,*

- *If $x \in X_0$, then for $i = 1, \ldots, m$, $T_{c_i} x$ is defined to be $S_{c_i} x$ projected onto $X$.*
- *If $x \in X_1$ then $T_{c_i} x := x$, for $i = 2, \ldots, m$ and $T_{c_1} x$ is defined to be $S_{c_1} x' \in Y_1$ where $x' \in S_{c_1}^{-1} Y_1 \cap Y_0$ is the lift of $x$ to $S_{c_1}^{-1} Y_1 \cap Y_0$.*
- *If $x \in X_2$, then $T_{c_1} x = T_{c_1}^0 x$. If $x \in X_{2,i}$, $i = 2, \ldots m$, $T_{c_j} x = x$ for $j \neq i$ and $T_{c_i} x$ is defined to be $S_{c_i} x' \in Y_1$ where $x' \in S_{c_i}^{-1} Y_1 \cap Y_0$ is the lift of $x$ to $S_{c_i}^{-1} Y_1 \cap Y_0$.*

*For the remaining $g \in F_m$, $T_g$ is then defined in the usual fashion.*

The shifts $T_{c_i} : X \to X$ for $i = 1, \ldots, m$ are essentially just $S_{c_i} : Y \to Y$ on $X_0$, where as on $X_0$, $X_1$ they are given by $T_0^{c_1}$ and the identity map. As in [17] the identity map is used for simplicity; an arbitrary measure preserving map could be substituted.

**Lemma 6.7.** *$T_{c_i}$ are invertible and measure preserving and hence $(X, \mathcal{X}, \mu, (T_g)_{g \in F_m})$ is an $F_m$-system.*

*Proof.* The invertibilty of the maps is clear checking the maps are measure preserving for each generator is completely analogous so just the case for $T_{c_1}$ is presented. It suffices to check for sets of half infinite words where the first $k+1$ elements have been specified. Take an arbitrary set of this form and denote it by $A_k$. Consider separately the cases for $X_1, X_2, X_0$. Clearly, $T_{c_1}$ is measure preserving on $X_2$ by definition of ergodicity. For $A_k \in X_1$ words start from index 0 and $s_0 = c_1$ or $s_0 = c_1^{-1}$ so

$$\mu(A_k) = \sum_{n \geq 1} (2m-1)^{-n} \mu_n(A_n) + \frac{1}{2} m_0(A_k) = \frac{1}{2} m_0(A_k) = \frac{1}{2m \times (2m-1)^k}.$$

$T_{c_1}^{-1}$ takes $X_1$ to $Y_1 \subset X_0$. Therefore,

$$\mu(T_{c_1}^{-1} A_k) = (2m-1)^{-1} \mu_1(T_{c_1}^{-1} A_k) = (2m-1)^{-1} \frac{1}{2m \times (2m-1)^{k-1}} = \frac{1}{2m \times (2m-1)^k},$$

and thus $T_{c_1}$ is measure preserving on $X_1$. Finally, on $X_0$ consider three cases. Take a set $A_{r,k}$ consisting of words of the form $s_r s_{r+1} s_{r+2} \cdots s_{r+k} \cdots \in Y_r$ where $r \geq 1$ and $k+1$ elements are fixed.

$$\mu(A_{r,k}) = (2m-1)^{-r} \mu_r(A_{r,k}) = (2m-1)^{-r} \frac{1}{2m \times (2m-1)^k} = \frac{1}{2m \times (2m-1)^{k+r}}.$$

Take $r = 1$ and consider $s_1 \neq c_1$, then

$$\mu(T_{c_1}^{-1} A_{1,k}) = \frac{1}{2} \frac{2}{2m \times (2m-1)^{k+1}} = \frac{1}{2m \times (2m-1)^{k+1}}.$$

For all other cases, $T_{c_1}^{-1}$ applied to $A_{r,k}$, gives elements in $A_{r-1,k}$, $r > 1$ or elements in $A_{r+1,k}$ $r \geq 1$. Either, $s_r = c_1$ then $T_{c_1}^{-1} A_{r,k} = A_{r+1,k} \subset Y_{r+1,k}$ and is the set whose first $k$ elements are fixed to be $s_{r+1} s_{r+2} \cdots s_{r+k}$. Or, $s_r \neq c_1$ and $T_{c_1}^{-1} A_{r,k} = A_{r-1,k} \subset Y_{r-1,k}$ is the set of elements $c_1^{-1} s_r s_{r+1} \cdots s_{r+k} \cdots$. So,

$$\mu(T_{c_1}^{-1} A_{r,k}) = (2m-1)^{-(r+1)} \mu_{r+1}(A_{r+1,k}) = (2m-1)^{-(r+1)} \frac{1}{2m \times (2m-1)^{k-1}} = \frac{1}{2m \times (2m-1)^{k+r}}.$$

$$\mu(T_{c_1}^{-1} A_{r,k}) = \mu(A_{r-1,k}) = (2m-1)^{-(r-1)} \frac{1}{2m \times (2m-1)^{k+1}} = \frac{1}{2m \times (2m-1)^{k+r}}.$$

□

**Lemma 6.8.** *Axiom (iv):* $X = \bigcup_{g \in F_m} T_g X_{2,i}^{(1)} = \bigcup_{g \in F_m} T_g X_{2,i}$ *up to null sets for each* $i = 2, \ldots, m$.

*Proof.* Recall, $X_{2,i}$ is the set of words in $X_2$ with $c_i$ or $c_i^{-1}$ as the first element. Fix $i \in \{2, \ldots, m\}$. Clearly, $\bigcup_{g \in F_m} T_g X_{2,i} \subset X$ and $X = X_1 \cup X_2 \cup X_0$, therefore just need to show that $X_1, X_2, X_0 \subset \bigcup_{g \in F_m} T_g X_{2,i}$ and clearly, $X_{2,i} \subset \bigcup_{g \in F_m} T_g X_{2,i}$. $Y_1 = T_{c_i} X_{2,i} \cup T_{c_i}^{-1} X_{2,i}$. Furthermore, $X_{2,j} \subset (T_{c_j} Y_1 \cup T_{c_j}^{-1} Y_1)$. So, for $i \neq j$,

$$X_{2,j} \subset T_{c_j} T_{c_i} X_{2,i} \cup T_{c_j} T_{c_i}^{-1} X_{2,i} \cup T_{c_j}^{-1} T_{c_i} X_{2,i} \cup T_{c_j}^{-1} T_{c_i}^{-1} X_{2,i}.$$

Therefore, for any $j = 2, \ldots, m$, $X_{2,j} \subset \bigcup_{g \in F_m} T_g X_{2,i}$. Again, using $Y_1 = T_{c_i} X_{2,i} \cup T_{c_i}^{-1} X_{2,i}$ and $X_1 \subset T_{c_1} Y_1 \cup T_{c_1}^{-1} Y_1$, it follows $X_1 \subset \bigcup_{g \in F_m} T_g X_{2,i}$.
The case for $X_0$ induction, starting with the base case, $Y_1 = T_{c_i} X_{2,i} \cup T_{c_i}^{-1} X_{2,i} \subset \bigcup_{g \in F_m} T_g X_{2,i}$. Assume $Y_{k-1} \subset \bigcup_{g \in F_m} T_g X_{2,i}$. $Y_k$ is the set of all half infinite words beginning at index $k$. Let $A_j$, $j = 1, \ldots, m$ be the set of words in $Y_k$ whose first element does not equal $c_j$ then $A_j \subset T_{c_j} Y_{k-1}$. Hence, $Y_k = \bigcup_{j=1}^m A_j \subset \bigcup_{j=1}^m T_{c_j} Y_{k-1}$ and by the induction hypothesis this is contained in $\bigcup_{g \in F_m} T_g X_{2,i}$. Therefore, by induction $Y_n \subset \bigcup_{g \in F_m} T_g X_{2,i}$ for all $n \geq 1$ and thus $X_0 = \bigcup_{n \geq 1} Y_n \subset \bigcup_{g \in F_m} T_g X_{2,i}$. □

**Proposition 6.9.** $(X, \mathcal{X}, \mu, (T_g)_{g \in F_m})$ *is a good system.*

*Proof.* $T_{c_i}$ are invertible and measure preserving by Lemma 6.7 and hence $(X, \mathcal{X}, \mu, (T_g)_{g\in F_m})$ is an $F_m$-system. Recall, $\mu(X) = \frac{m}{2(m-1)}$. Then,

$$\mu(X_0) = \sum_{n\geq 1} (2m-1)^{-n} = \tfrac{1}{2(m-1)} = \tfrac{1}{m}\mu(X)$$

and Lemma 6.3 implies

$$\mu(X_1) = \tfrac{1}{2m} = \tfrac{m-1}{m^2}\mu(X) \quad \text{and} \quad \mu(X_2) = \tfrac{m-1}{2m} = \tfrac{(m-1)^2}{m^2}\mu(X).$$

Axiom (i) follows from Lemma 6.4, Axiom (ii)(a), and (ii)$(b)$ have been verified and for Axiom (iii) note that

$$T_{c_1}X_1 \subset (S_{c_i}Y_0 \cap Y_1) \cup (S_{c_i}^{-1}Y_0 \cap Y_1) \subset T_{c_i}X_{2,i} \cup T_{c_i}^{-1}X_{2,i} = Y_1 \subset X_0$$

for all $i = 2, \ldots, m$, as required. Finally, Axiom (iv) was Lemma 6.8. □

Now, the aim is to construct a set of functions, on this good system, satisfying $P(\frac{2m-2}{2m})$.

**Proposition 6.10.** *It is possible to construct a sequence, $\tilde{f}_n$, of non-negative functions in $L^\infty(\tilde{X})$ for each $n \in \mathbb{Z}$ obeying Axiom (v)-(viii) with $\alpha = \frac{2m-2}{m}$.*

*Proof.* Recall that $\tilde{X}$ is the product of the space $X$ with the set $W$ equipped with the uniform measure, $\nu$. For negative $n$ define $\tilde{f}_n$ by setting

$$\tilde{f}_n(x,s) \coloneqq 2m \times (2m-1)^{-n-1},$$

whenever $x \in X$ and $s \in W$ are such that $x \in Y_{-n}$ and $S_s x \in Y_{-n-1}$ and $\tilde{f}_n(x,s) = 0$ otherwise. These are clearly non-negative functions in $L^\infty(\tilde{X})$.

Recall Axiom (vii): *$\tilde{f}_n$ is supported in $\tilde{X}_0$ for all negative $n$ and there exists a finite $A > 0$ such that $\tilde{f}_n$ is supported in a set of measure at most $A(2m-1)^n\mu(X)$ for all negative $n$.*

$X = X_1 \cup X_2 \cup X_0$ if $x \in X_1 \cup X_2$ then $x \in Y_0/\sim$ and in particular $x \notin Y_{-n}$ for $n < 0$, therefore $\tilde{f}_n(x,s) = 0$ and $\tilde{f}_n$ is supported in $\tilde{X}_0$ for all negative $n$.

Let $x = s_{-n}s_{-n+1}s_{-n+2}\cdots \in Y_{-n}$ then $\tilde{f}_n(x,s)$ is non-zero only if $S_s x = ss_{-n}s_{-n+1}s_{-n+2}\cdots \in Y_{-n-1}$, so $s_{-n} \neq s^{-1}$. The support is at most the set $\{(s_{-n}s_{-n+1}s_{-n+2}\cdots, s) \in \tilde{X} : s_{-n} \neq s^{-1}\}$.

$\mu \coloneqq \sum_{n\geq 1}(2m-1)^{-n}\mu_n + \frac{1}{2}m_0$ and this set only fixes the first element of the word, for which there are $2m-1$ possibilities, hence

$$\begin{aligned}\tilde{\mu}(\{(s_{-n}s_{-n+1}s_{-n+2}\cdots, s) \in \tilde{X} : s_{-n} \neq s^{-1}\}) &= \mu(\{s_{-n}s_{-n+1}s_{-n+2}\cdots : s_{-n} \neq s^{-1}\}) \times \nu(s) \\ &= (2m-1)^n \frac{2m-1}{2m} \times \frac{1}{2m}.\end{aligned}$$

Then summing each possible $s$, gives measure $(2m-1)^n \frac{2m-1}{2m}$. Since $\mu(X) = \frac{m}{2(m-1)}$, letting $A = \frac{(2m-1)(m-1)}{m^2}$ gives Axiom (vii).

Recall, Axiom (v): *$\tilde{f}_{n+m} = P^m \tilde{f}_n$ for all $n \in \mathbb{Z}$ and $m \in \mathbb{N}$.* Proceed by induction for $n < -2$.

For the base case

$$P\tilde{f}_{-2}(x,s) = \frac{1}{2m-1}\sum_{s'\in W\smallsetminus\{s^{-1}\}} \tilde{f}_{-2}(T_{s'}^{-1}x,s') = \frac{1}{2m-1}\sum_{s'\in W\smallsetminus\{s^{-1}\}} g(x,s').$$

Where,

$$\begin{aligned} g(x,s') &= \begin{cases} \tilde{f}_{-2}(T_{s'}^{-1}x,s') & \text{if } T_{s'}^{-1}x \in Y_2 \text{ and } S_{s'}T_{s'}^{-1}x \in Y_1 \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} 2m \times (2m-1) & \text{if } T_{s'}^{-1}x \in Y_2 \text{ and } T_{s'}T_{s'}^{-1}x = x \in Y_1 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

$g$ is non-zero when $T_{s'}^{-1}x \in Y_2$ and $x = s_1s_2\cdots \in Y_1$, which is if and only if $s' = s_1$. This happens precisely one time in the summation over $s' \in W\smallsetminus\{s^{-1}\}$ provided $s_1 \neq s^{-1}$. Finally, since $x \in Y_1$, $s_1 \neq s^{-1}$ is equivalent to $S_sx \in Y_0$. Therefore,

$$\begin{aligned} P\tilde{f}_{-2}(x,s) &= \begin{cases} \frac{1}{2m-1}2m \times (2m-1) & \text{if } S_sx \in Y_0 \text{ and } x \in Y_1 \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} 2m & \text{if } S_sx \in Y_0 \text{ and } x \in Y_1 \\ 0 & \text{otherwise} \end{cases} \\ &= \tilde{f}_{-1}(x,s). \end{aligned}$$

Now show $\tilde{f}_{-k} = P\tilde{f}_{-k-1}$ for $-k \leq -1$,

$$P\tilde{f}_{-k-1}(x,s) = \frac{1}{2m-1}\sum_{s'\in W\smallsetminus\{s^{-1}\}} \tilde{f}_{-k-1}(T_{s'}^{-1}x,s') = \frac{1}{2m-1}\sum_{s'\in W\smallsetminus\{s^{-1}\}} g(x,s')$$

$$g(x,s') = \begin{cases} 2m \times (2m-1)^k & \text{if } T_{s'}^{-1}x \in Y_{k+1} \text{ and } T_{s'}T_{s'}^{-1}x = x \in Y_k \\ 0 & \text{otherwise.} \end{cases}$$

$g$ is non-zero when $T_{s'}^{-1}x \in Y_{k+1}$ and $x = s_ks_{k+1}\cdots \in Y_k$ which is if and only if $s' = s_k$. This happens precisely one time in the summation, over $s \in W\smallsetminus\{s^{-1}\}$, provided $s_k \neq s^{-1}$. Finally, since $x \in Y_k$, $s_k \neq s^{-1}$ is equivalent to $T_sx \in Y_{k-1}$. Therefore,

$$\begin{aligned} P\tilde{f}_{-k-1}(x,s) &= \begin{cases} \frac{1}{2m-1}2m \times (2m-1)^k & \text{if } T_{s'}^{-1}x \in Y_{k+1} \text{ and } x \in Y_k \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} 2m \times (2m-1)^{k-1} & \text{if } S_sx \in Y_{k-1} \text{ and } x \in Y_k \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

$$= \tilde{f}_{-k}(x,s).$$

Axiom (v) then follows by defining $\tilde{f}_n$, for non-negative $n$, by the formula $\tilde{f}_n := P^{n+1}\tilde{f}_{-1}$.
Axiom (vi): $\|\tilde{f}_n\|_{L^1(\tilde{X})} = \alpha\mu(X)$ *for some* $n \in \mathbb{Z}$. Earlier, it was verified that the support for each function $f_n$ had measure $\frac{(2m-1)^{n+1}}{2m}$. Furthermore the function $f_n$ is constant on the support and equal to $(2m \times (2m-1)^{-n-1})$. Putting this together, for negative $n$,

$$\|\tilde{f}_n\|_{L^1(\tilde{X})} = \int_{supp(f_n)} \sum_{s\in W} \frac{1}{2m} 2m \times (2m-1)^{-n-1} d\mu(x) = \mu(supp(f_n)) 2m \times (2m-1)^{-n-1} = 1.$$

Take $\alpha = 1/\mu(X)$ with $\mu(X) = \frac{m}{2(m-1)}$ then Axiom (vi) follows (using Axiom (v) to extend to non-negative $n$).
Axiom (viii): $\sup_{n\in\mathbb{Z}} \pi_* \tilde{f}_{2n}(x) \geq 1-\epsilon$, *for all* $x \in X$ *outside of a set of measure at most* $\epsilon\mu(X)$. Theorem 4.3 states that for any good system $(X, \mathcal{X}, \mu, (T_g)_{g\in F_m})$ and $f \in L^\infty(\tilde{X})$, $P^{2n}\tilde{f}$ converges pointwise almost everywhere and in $L^1$ norm to $\frac{1}{\mu(X)} \int_{\tilde{X}} \tilde{f} d\tilde{\mu} = \alpha$ since, $\|\tilde{f}_n\|_{L^1(\tilde{X})} = \alpha\mu(X)$. Thus $\tilde{f}_{2n} = P^{2n}\tilde{f}_0$ converges pointwise almost everywhere to $\alpha$ as $n \to +\infty$. Using Egorov's theorem and noting that $\alpha = \frac{2m-2}{m} \geq 1$,

$$\sup_{n\in\mathbb{Z}} \pi_* \tilde{f}_{2n}(x) \geq \alpha - \epsilon \geq 1 - \epsilon$$

for all $x \in X$ outside of a set of measure at most $\epsilon\mu(X)$. □

# 7 The Inductive Step

The objective of this section is to prove Theorem 5.6. A similar construction to Ornstein's counterexample is used, this connection is clear when [[16], Lemma 2.1] is compared with Theorem 7.8. A sequence of functions, with different components will be defined on a space consisting of two identical good systems glued together. The components of the functions are defined according to which space they lie on. One of the spaces causes the functions to experience a time delay; that is, they are set equal to functions further back in the sequence. The Markov operator $P$ acts on these functions causing a small amount of leakage from one space to another. The time delay allows the other components to have achieved mixing in the portion where this occurs and finally, this allows the amplitude of the functions with time delay to be smaller than would have otherwise been necessary to make $\sup_n P^n f$ large. A definition of mixing is given in the appendix, Section 9.4. Now suppose there exists $\alpha$, $0 < \alpha \leq \frac{2m-2}{m}$, such that $P(\alpha)$ holds, (with $\epsilon$ replaced by $\epsilon/4$) and normalising $X$ to have measure 1 there exists a good system $(X, \mathcal{X}, \mu, (T_g)_{g\in F_m})$ with associated decomposition $X = X_1 \cup X_2 \cup X_0$, $\mu(X) = 1$ and a sequence of non-negative functions $\tilde{f}_n \in L^\infty(\tilde{X})$ for $n \in \mathbb{Z}$ with the following properties:

(v) (Ancient Markov chain) $\tilde{f}_{n+1} = P\tilde{f}_n$ for all $n \in \mathbb{Z}$.

(vi) (Size) $\|\tilde{f}_n\|_{L^1(\tilde{X})} = \alpha$ for all $n \in \mathbb{Z}$.

(vii) (Early support) $\tilde{f}_n$ is supported in $\tilde{X}_0$ for all negative $n$. Furthermore, there exists a finite $A > 0$ such that $\tilde{f}_n$ is supported in a set of measure at most $A(2m-1)^n$ for all negative $n$.

(viii) (Large maximum function)

$$\sup_{n \in \mathbb{Z}} \pi_* \tilde{f}_{2n}(x) \geq 1 - \epsilon/4$$

for all $x \in X$ outside a set of measure at most $\epsilon/4$.

The notation, $\tilde{X} = X \times W$, will continue to be used with the additional convention that $X' = X \times \{1, 2\}$ and $\tilde{X}' = X' \times W$. As before, $\tilde{f}$ refers to a function defined on $\tilde{X}$ and $\tilde{f}'$ will now denote a function defined on $\tilde{X}'$. This section will show the following theorem holds.

**Theorem 7.1.** *Given $P(\alpha)$ holds for some $0 < \alpha < \frac{2m-2}{m}$, $P(\alpha(1-\alpha/4))$ holds. More precisely, there exists a good system $(X', \mathcal{X}', \mu', (T'_g)_{g \in F_m})$ with associated decomposition $X' = X'_1 \cup X'_2 \cup X'_0$, Markov operator $P'$ and measure $\mu'(X') = 2$. Furthermore, it is possible to construct a sequence of non-negative functions $\tilde{f}'_n \in L^\infty(\tilde{X}')$ for $n \in \mathbb{Z}$ on this system with the following properties:*

*(v') (Ancient Markov Chain) $\tilde{f}'_{n+1} = P' \tilde{f}'_n$ for all $n \in \mathbb{Z}$.*

*(vi') (Size) $\|\tilde{f}'_n\|_{L^1(\tilde{X})} = \alpha(1-\alpha/4)\mu'(X') = \alpha(1-\alpha/4) \times 2 = \alpha(2-\alpha/2)$ for all $n \in \mathbb{Z}$.*

*(vii') (Early support) $\tilde{f}'_n$ is supported in $\tilde{X}'_0$ for all negative $n$. Furthermore there exists a finite $A' > 0$ such that $\tilde{f}'_n$ is supported in a set of measure at most $\mu'(X')A'(2m-1)^n = 2A'(2m-1)^n$ for all negative $n$.*

*(viii') (Large maximum function)*

$$\sup_{n \in \mathbb{Z}} \pi_* \tilde{f}'_{2n}(x') \geq 1 - \epsilon$$

*for all $x' \in X'$ outside a set of measure at most $2\epsilon = \epsilon\mu'(X')$.*

Firstly, an $F_m$-system, $(X', \mathcal{X}', \mu', (T'_g)_{g \in F_m})$, is constructed. This involves defining a measure space together with a set of bi-measurable maps. It is then shown that this system satisfies the four axioms of a good system. Finally, a set of functions is defined that satisfy the four axioms in Theorem 7.1. Section 5 demonstrated that Theorem 5.5 and Theorem 5.6 imply $P(\alpha)$ is true for arbitrarily small $\alpha$. This was the final reduced form of Theorem 2.3 and so proving Theorem 7.1 is the last stage of establishing the existence of a counterexample.

Firstly, construct the good system $(X', \mathcal{X}', \mu', (T'_g)_{g \in F_m})$ as in [17], by combining two copies of an arbitrary good system, $(X, \mathcal{X}, \mu, (T_g)_{g \in F_m})$, (satisfying Axioms (v)-(viii)) which are glued together by a small amount of coupling. This coupling is measured by the quantity $0 < \kappa < \frac{m-1}{m^2}$ which depends on $\epsilon, N$ and $\tilde{f}_n$. More precisely, define the measure space $(X', \mathcal{X}', \mu')$ to be the product of $(X, \mathcal{X}, \mu)$ with the two element set $\{1,2\}$ and counting measure, $\xi$. Using Axiom (i) there exists a subset $E$ of $X_1$ with measure $\kappa$. Next, define the shift maps $T'_{c_i} : X' \to X'$ for $i = 1, \dots, m$. The map $T'_{c_1}$ is a trivial lift of $T_{c_1}$, thus

$$T'_{c_1}(x, j) := (T_{c_1}x, j), \quad x \in X \quad \text{and} \quad j \in \{1,2\}.$$

The map $T'_{c_i}$, for $i = 2, \dots m$, controls the leakage between spaces and so it is defined differently for $E$ and $X \setminus E$.

$$T'_{c_i}(x, j) := \begin{cases} (T_{c_i}x, j) & \text{if } x \in X \setminus E \text{ and } j \in \{1,2\} \\ (T_{c_i}x, 3-j) & \text{if } x \in E \text{ and } j \in \{1,2\} \end{cases}$$

Then, define $T'_g$ for the remaining $g \in F_m$ in the usual fashion. This gives a family of measure preserving invertible maps on $X'$. Finally, partition $X' = X'_1 \cup X'_2 \cup X'_0$ where $X'_1 := X_1 \times \{1,2\}$, $X'_2 := X_2 \times \{1,2\}$ and $X'_0 := X_0 \times \{1,2\}$.

**Lemma 7.2.** *$(X', \mathcal{X}', \mu', (T'_g)_{g \in F_m})$ satisfies the following:*

*(i') (Measure) $\mu'(X'_1) = \frac{m-1}{m^2}\mu'(X')$, $\mu'(X'_1) = \frac{(m-1)^2}{m^2}\mu'(X')$ and $\mu'(X'_0) = \frac{1}{m}\mu'(X')$. Furthermore, for any $0 \le \kappa' \le \mu'(X'_1)$, there exists a measurable subset of $X'_1$ with measure $\kappa'$.*

*(ii') (a)(Partition) $X'_2$ can be partitioned into $m-1$ components $X'_{2,2}, \dots, X'_{2,m}$ of positive measure.*
*(b)(Ergodicity) For each $i = 2, \dots, m$, there exists $r_i < \infty$ such that $X'_{2,i}$ can be partitioned into $T'_{c_1}$-invariant components, $X'^{(1)}_{2,i}, \dots, X'^{(r_i)}_{2,i}$ and $T'^2_{c_1}$ is ergodic on each of the components $X'^{(k)}_{2,i}$; that is the only $T'^2_{c_1}$-invariant measurable subsets of $X'^{(k)}_{2,i}$ have either measure 0 or $\mu'(X'^{(k)}_{2,i})$ for $k = 1, \dots, r_i$.*

*(iii') (Invariance) $T'_{c_1}X'_2 = X'_2$. For $i = 2, \dots, m$, $T'_{c_i}X'_1 = X'_1$ and $T'_{c_i}X'_{2,j} = X'_{2,j}$ for all $j \ne i$. The inclusions $T'_{c_1}X'_1 \subset T'_{c_i}X'_{2,i} \cup T'^{-1}_{c_i}X'_{2,i} \subset X'_0$ hold for any $i = 2, \dots, m$.*

*Proof.* $\mu(X) = 1$ so, $\mu'(X') = \mu(X) \times \xi(\{1,2\}) = 2$. Using that $X$ satisfies Axiom (i),

$$\begin{aligned} \mu'(X'_1)) &= \mu(X_1) \times \xi(\{1,2\}) = \frac{m-1}{m^2}\mu(X) \times 2 = \frac{m-1}{m^2}\mu'(X') \\ \mu'(X'_2) &= \mu(X_2) \times \xi(\{1,2\}) = \frac{(m-1)^2}{m^2}\mu(X) \times 2 = \frac{(m-1)^2}{m^2}\mu'(X') \\ \mu'(X'_0) &= \mu(X_0) \times \xi(\{1,2\}) = \frac{1}{m}\mu(X) \times 2 = \frac{1}{m}\mu'(X'). \end{aligned}$$

For any $0 \le \kappa \le \mu(X_1)$, there is a measurable subset of $X_1$ of measure equal to $\kappa$. Fix $\kappa'$, with $0 \le \kappa' \le \mu'(X'_1)$ then there exists measurable $A \in X_1$, with $\mu(A) = \kappa'/2$. Furthermore,

$A\times\{1,2\}$ is measurable in $X_1'$ with $\mu'(A\times\{1,2\}) = \mu(A)\times\xi(\{1,2\}) = \kappa\times 2 = \kappa'$. Hence Axiom (i') is satisfied.

By Axiom (ii)(a) for $X$, $X_2$ can be partitioned into $m-1$ components $X_{2,2}, X_{2,3}, \ldots, X_{2,m}$ of positive measure, which induces a partition of $X_2'$ into the $m-1$ components $X_{2,i}' = X_{2,i}\times\{1,2\}$ for $i=2,\ldots,m$. By Axiom (ii)(b), for each $i=2,\ldots,m$ these can be partitioned into finitely many, $T_{c_1}$-invariant components, $X_{2,i}^{(1)},\ldots,X_{2,i}^{(r_i)}$ such that $T_{c_1}^2$ is ergodic on each of the components $X_{2,i}^{(k)}$. This induces a partition of $X_{2,i}'$ into the $2r_i$ components $X_{2,i}^{(1)}\times\{1\},\ldots,X_{2,i}^{(r_i)}\times\{1\},X_{2,i}^{(1)}\times\{2\},\ldots,X_{2,i}^{(r_i)}\times\{2\}$. Each of these components are clearly $T_{c_1}'^2$-ergodic.

Now, $T_{c_1}'X_2' = X_2'$ follows from $T_{c_1}'X_2\times\{j\} = X_2\times\{j\}$ for $j\in\{1,2\}$. Similarly, for $i=2,\ldots,m$ and $j\neq i$, $T_{c_i}'X_{2,i}' = X_{2,i}'$. $T_{c_i}'X_1' = X_1'$ for $i=2,\ldots,m$ follows from the definition of $T_{c_i}'$ for $i=2,\ldots,m$, which gives

$$\begin{aligned} T_{c_i}'X_1' &= T_{c_i}'(E\times\{1\})\cup T_{c_i}'(E\times\{2\})\cup T_{c_i}'(X_1\setminus E\times\{1\})\cup T_{c_i}'(X_1\setminus E\times\{2\})\\ &= (E\times\{2\})\cup(E\times\{1\})\cup(X_1\setminus E\times\{1\})\cup(X_1\setminus E\times\{2\}) = X_1'. \end{aligned}$$

Finally, observe $T_{c_1}'$ is a trivial lift of $T_{c_1}$ and $T_{c_i}'$ is a trivial lift of $T_{c_i}$ on $X_2'$. Then using that $X$ satisfies Axiom (iii), $T_{c_1}X_1\subset T_{c_i}X_{2,i}\cup T_{c_i}^{-1}X_{2,i}\subset X_0$ for any $i=2,\ldots,m$, it follows

$$T_{c_1}'X_1' = ((T_{c_1}X_1)\times\{1,2\})\subset(T_{c_i}X_{2,i}\cup T_{c_i}^{-1}X_{2,i})\times\{1,2\} = T_{c_i}'X_{2,i}'\cup T_{c_i}'^{-1}X_{2,i}'\subset X_0\times\{1,2\} = X_0'$$

□

**Proposition 7.3.** *If $\kappa$ is sufficiently small then $(X',\mathcal{X}',\mu',(T_g')_{g\in F_m})$ is a good system.*

*Proof.* Axioms (i')-(iii') are verified in Lemma 7.2 so using the same argument as [[17], Proposition 4.1] it now remains to verify Axiom (iv'); $X' = \bigcup_{g\in F_m} T_g'(X_{2,i}^{(k)}\times\{j\})$ up to null sets for each $k=1,\ldots,r_i$, for $i=2,\ldots,m$, and $j=1,2$. Fix $i,j$ and $k$. Denote the right-hand side by $Y$ which is $F_m$-invariant and contains $X_{2,i}^{(k)}\times\{j\}$. Using Axiom (iv) for $X$, $X=\bigcup_{g\in F_m} T_g X_{2,i}^{(k)}$, and the pigeon hole principle there exists $g\in F_m$ such that $T_gX_{2,i}^{(k)}$ intersects $E$ in a set of positive measure. Assume the word length $|g|$ of $g$ is minimal among all $g$ with this property, thus $T_hX_{2,i}^{(k)}\cap E$ is null whenever $|h|<|g|$ and since $T_h'$ is a trivial lift outside of $E\times\{1,2\}$, $T_h'$, for $|h|\le|g|$ will not map $j$ to $3-j$. Hence, $T_g'(X_{2,i}^{(k)}\times\{j\})$ intersects $E\times\{j\}$ in a set of positive measure. Recall, $T_g'$, $g\in F_m$ are measure preserving, so using definition of $T_{c_i}'$, $i\neq 1$, and Axiom(iii') it follows,

$$\begin{aligned} 0 &< \mu'([T_g'(X_{2,i}^{(k)}\times\{j\})]\cap[E\times\{j\}])\\ &\le \mu'([T_{c_ig}'(X_{2,i}^{(k)}\times\{j\})]\cap[T_{c_i}E\times\{3-j\}])\\ &\le \mu'([T_{c_ig}'(X_{2,i}^{(k)}\times\{j\})]\cap[X_1\times\{3-j\}])\\ &\le \mu'([T_{c_1c_ig}'(X_{2,i}^{(k)}\times\{j\})]\cap[T_{c_1}'(X_1\times\{3-j\})]) \end{aligned}$$

$$\leq \mu'([T'_{c_1c_ig}(X_{2,i}^{(k)} \times \{j\})] \cap [T'_{c_i}(X_{2,i} \times \{3-j\}) \cup T'_{c_i^{-1}}(X_{2,i} \times \{3-j\})])$$
$$\leq \mu'([T'_{c_1c_ig}(X_{2,i}^{(k)} \times \{j\})] \cap [T'_{c_i}(X_{2,i} \times \{3-j\})]) + \mu'([T'_{c_1c_ig}(X_{2,i}^{(k)} \times \{j\})] \cap [T'_{c_i^{-1}}(X_{2,i} \times \{3-j\})])$$
$$\leq \mu'([T'_{c_i^{-1}c_1c_ig}(X_{2,i}^{(k)} \times \{j\})] \cap [(X_{2,i} \times \{3-j\})]) + \mu'([T'_{c_ic_1c_ig}(X_{2,i}^{(k)} \times \{j\})] \cap [(X_{2,i} \times \{3-j\})]).$$

In particular, $Y$ intersects $X_{2,i} \times \{3-j\}$ in a set of positive measure and thus also intersects $X_{2,i}^{(k')} \times \{3-j\}$ for some $k' = 1, \ldots, r_i$. Now since $Y \cap [(X_{2,i}^{(k')} \times \{3-j\})] \subset (X_{2,i}^{(k')} \times \{3-j\})$ is a positive measurable set and both $Y$ and $(X_{2,i}^{(k')} \times \{3-j\})$ are $T_{c_1}'^2$-invariant, by Axiom (ii')(b) it follows that this set must have measure $\mu(X_{2,i}^{(k')})$ and thus $Y$ contains $(X_{2,i}^{(k')} \times \{3-j\})$ up to null sets, for some $k' = 1, \ldots, r_i$.
By another appeal to Axiom (iv) and the pigeon-hole principle it is possible to find $g_{k'} \in F_m$ such that $T_{g_{k'}} X_{2,i}^{(k')}$ and $X_{2,i}^{(k)}$ intersect in a set of positive measure. There are only $r_i$ choices for $k'$, so the word length $g_{k'}$ can be bounded above and the measure of $T_{g_{k'}} X_{2,i}^{(k')} \cap X_{2,i}^{(k)}$ bounded below by quantities independent of $\kappa$. Start in $X_{2,i}^{(k')}$ and apply the maps $T'_{c_i}$. Suppose that in order to end up in $X_{2,i}^{(k)}$ it is necessary to go via $X_1$. If not then $T'_{g_{k'}}(X_{2,i}^{(k')} \times \{3-j\})$ and $X_{2,i}^{(k)} \times \{3-j\}$ intersect in a set of positive measure since $T'_{c_i}$ is a trivial lift of $T_{c_i}$ for $i = 1, \ldots, m$. Using that $|g_{k'}|$ is bounded above there exists $L < \infty$ and $s_1, \ldots, s_L \in W$ such that $g_{k'} = s_1s_2 \ldots s_L$. Furthermore, there exist positive integers $l_1 < l_2 < \ldots < l_p$ with $p < L$ such that $h_{l_1} = s_1 \ldots s_{l_1}$, is the first sub-word of $g_{k'}$, (i.e $g_{k'} = h_{l_1} s_{l_1+1} \ldots s_L$) such that $A_1 = T_{h_{l_1}} X_{2,i}^{(k')} \cap X_1$ has positive measure. Similarly, $l_r$ is defined to be the $r-th$ time that $A_r = T_{h_{l_r}} X_{2,i}^{(k')} \cap X_1$ has positive measure. Take $\mu(E) < \min_{1 \leq r \leq p} \mu(A_r)$, then $T'_{g_{k'}}(X_{2,i}^{(k')} \times \{3-j\})$ and $X_{2,i}^{(k)} \times \{3-j\}$ intersect in a set of positive.
Hence, if $\kappa$ and hence $E$ is small enough then $T'_{g_{k'}}(X_{2,i}^{(k')} \times \{3-j\})$ and $X_{2,i}^{(k)} \times \{3-j\}$ also intersect in a set of positive measure thus $Y$ must intersect $X_{2,i}^{(k)} \times \{3-j\}$ in a set of positive measure. So, by $T_{c_1}'^2$-ergodicity of $X_{2,i}'^{(k)}$, $Y$ contains $X_{2,i}^{(k)} \times \{3-j\}$ up to null sets. Since $Y$ already contained $X_{2,i}^{(k)} \times \{j\}$, it now follows $X_{2,i}^{(k)} \times \{1,2\}$ is contained in $Y$ up to null sets.
Now for any $(x, j') \in X'$, from Axiom (iv) it follows that $x = T_g y$ for some $y \in X_{2,i}^k$ and $g \in F_m$. This implies that $(x, j') = T'_g(y, j'')$ for some $j'' \in \{1,2\}$ and hence $(x, j') \in Y$ for almost every $(x, j') \in X$ which gives Axiom (iv') for $X'$ as required. □

With a good system defined, it remains to find a set of functions satisfying the axioms of Theorem 7.1. Let $M \in \mathbb{N}$ be large, depending on all previous quantities, (in particular $\kappa$), to be chosen later. The functions $\tilde{f}'_n \in L^1(\tilde{X}')$ will be defined for negative $n$ by the formulae

$$\tilde{f}'_n(x, 1, s) := \tilde{f}_n(x, s) \text{ and } \tilde{f}'_n(x, 2, s) := (1 - \alpha/2)\tilde{f}_{n-2M}(x, s)$$

for any $x \in X$ and $s \in W$. As Tao [17] remarks, informally $\tilde{f}'_n$ has two components, the one on $X \times \{2\}$ experiences a significant time delay, $n \to n - 2M$ and also a slight reduction in amplitude by factor $(1-\alpha/2)$. This delay causes the dynamics of $X \times \{1\}$ to have mixed almost completely, so that half the mass of the $X \times \{1\}$ component is spread out almost uniformly over $X \times \{2\}$ allowing the amplitude reduction for the $X \times \{2\}$ component.

**Lemma 7.4.** *Axiom (vii') (Early support) $\tilde{f}'_n$ is supported in $\tilde{X}'_0$ for all negative $n$. Furthermore there exists a finite $A' > 0$ such that $\tilde{f}'_n$ is supported in a set of measure at most $\mu'(X')A'(2m-1)^n = 2A'(2m-1)^n$ for all negative $n$. (Note the constant $A'$ can depend on $M$).*

*Proof.* Let $n < 0$, $\tilde{f}'_n(x,1,s) := \tilde{f}_n(x,s)$ and $\tilde{f}'_n(x,2,s) := (1-\alpha/2)\tilde{f}_{n-2M}(x,s)$ and by Axiom (vii) $\tilde{f}_n$ and $\tilde{f}_{n-2M}$ are supported in $\tilde{X}_0$. Therefore, $\tilde{f}'_n$ is supported on $\tilde{X}'_0$. Furthermore, there exists $A > 0$ such that $\tilde{f}'_n(x,1,s) := \tilde{f}_n(x,s)$ is supported in a set of measure at most $A(2m-1)^n \mu(X)$ and $\tilde{f}'_n(x,2,s) := (1-\alpha/2)\tilde{f}_{n-2M}(x,s)$ is supported in a set of measure at most $A(2m-1)^{n-2M}\mu(X)$. Thus $\tilde{f}'_n$ is supported in a set of measure at most $\frac{A}{2}(1+(2m-1)^{-2M}) \times (2m-1)^n \times 2\mu(X) = A'(2m-1)^n \mu'(X')$, where $A' = \frac{A}{2}(1+(2m-1)^{-2M})$. ☐

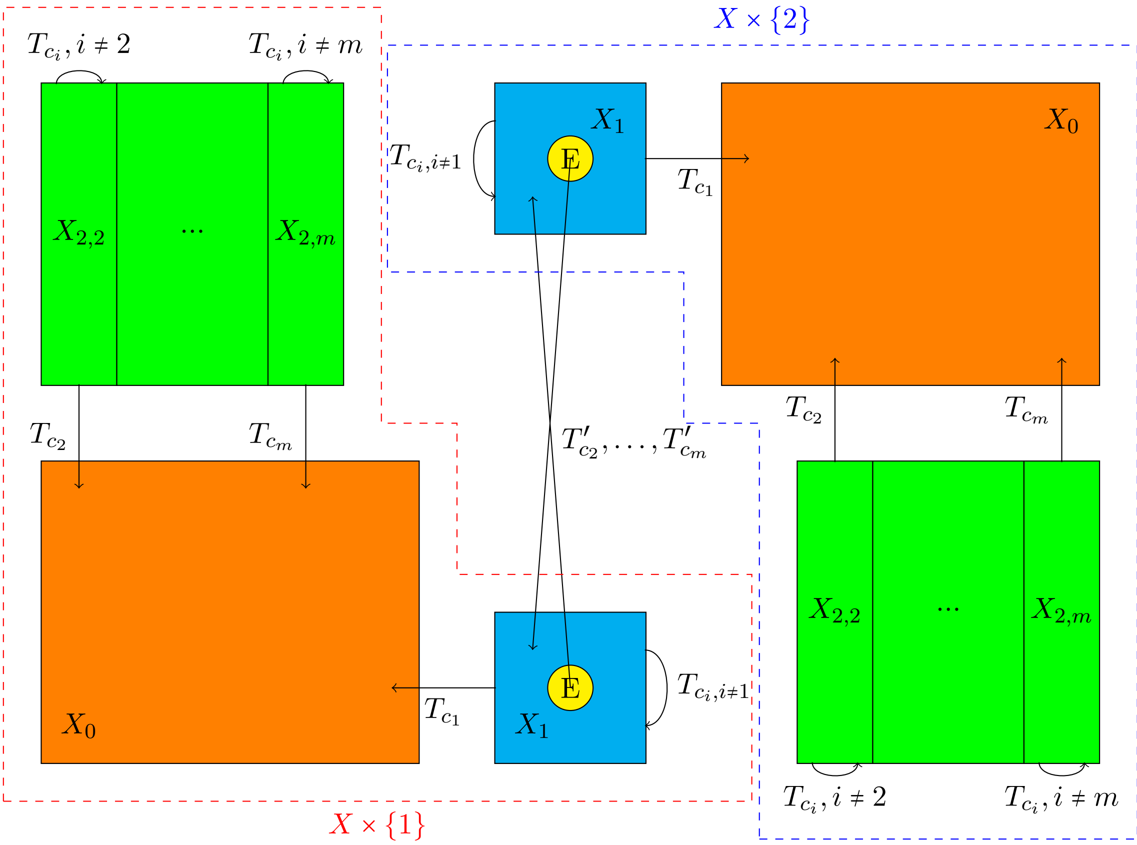



Figure 5: This illustrates the $F_m$-system constructed from two good systems glued together, interacting via the small set $E$. Only part of the action of $T_{c_i}$ for $i = 1, \ldots, m$ and the structure of $X_2$, are displayed.

Let $P'$ be defined as,

$$P'\tilde{f}'(x,i,s) := \frac{1}{2m-1} \sum_{s_1 \in W \setminus \{s^{-1}\}} \tilde{f}(T'^{-1}_{s_1}(x,i), s_1).$$

**Lemma 7.5.** *Axiom (v')(Ancient Markov chain) $\tilde{f}'_{n+1} = P'\tilde{f}'_n$ for all $n \in \mathbb{Z}$.*

*Proof.* If $x \notin E$ then it is clear that,

$$\begin{aligned}
P'\tilde{f}'_{-2}(x,i,s) &= \frac{1}{2m-1}\sum_{s_1 \in W \setminus \{s^{-1}\}} \tilde{f}'_{-2}(T'^{-1}_{s_1}(x,i), s_1) \\
&= \frac{1}{2m-1}\sum_{s_1 \in W \setminus \{s^{-1}\}} \tilde{f}'_{-2}(T^{-1}_{s_1}x, i, s_1) \\
&= \begin{cases} \frac{1}{2m-1}\sum_{s_1 \in W \setminus \{s^{-1}\}} \tilde{f}_{-2}(T^{-1}_{s_1}x, s_1) & \text{if } i = 1 \\ \frac{1}{2m-1}\sum_{s_1 \in W \setminus \{s^{-1}\}} (1-\alpha/2)\tilde{f}_{-2-2M}(T^{-1}_{s_1}x, s_1) & \text{if } i = 2 \end{cases} \\
&= \begin{cases} P\tilde{f}_{-2}(x,s) & \text{if } i = 1 \\ (1-\alpha/2)P\tilde{f}_{-2-2M}(x,s) & \text{if } i = 2 \end{cases} \\
&= \begin{cases} \tilde{f}_{-1}(x,s) & \text{if } i = 1 \\ (1-\alpha/2)\tilde{f}_{-1-2M}(x,s) & \text{if } i = 2 \end{cases} \\
&= \tilde{f}'_{-1}(x,i,s).
\end{aligned}$$

For $x \in E$, $n < 0$, $\tilde{f}'_n$ and $\tilde{f}_n$ are zero on $X_1$ and $X'_1$ by Axiom (vii) and (vii'), respectively. Suppose $i = 1$,

$$\begin{aligned}
P'\tilde{f}'_{-2}(x,1,s) &= \frac{1}{2m-1}\left(\sum_{s_1 \in \{c_1, c_1^{-1}\} \setminus \{s^{-1}\}} \tilde{f}'_{-2}(T'^{-1}_{s_1}(x,1), s_1) + \sum_{s_1 \in W \setminus \{s^{-1}, c_1, c_1^{-1}\}} \tilde{f}'_{-2}(T'^{-1}_{s_1}(x,1), s_1)\right) \\
&= \frac{1}{2m-1}\left(\sum_{s_1 \in \{c_1, c_1^{-1}\} \setminus \{s^{-1}\}} \tilde{f}'_{-2}(T^{-1}_{s_1}x, 1, s_1) + \sum_{s_1 \in W \setminus \{s^{-1}, c_1, c_1^{-1}\}} \tilde{f}'_{-2}(T^{-1}_{s_1}x, 2, s_1)\right) \\
&= \frac{1}{2m-1}\left(\sum_{s_1 \in \{c_1, c_1^{-1}\} \setminus \{s^{-1}\}} \tilde{f}_{-2}(T^{-1}_{s_1}x, s_1) + \sum_{s_1 \in W \setminus \{s^{-1}, c_1, c_1^{-1}\}} (1-\alpha/2)\tilde{f}_{-2-2M}(T^{-1}_{s_1}x, s_1)\right) \\
&= 0 = \tilde{f}'_{-1}(x,1,s).
\end{aligned}$$

The cases for i=2 and $n \le -2$ are completely analogous. So $\tilde{f}'_{n+1} = P'\tilde{f}'_n$ for all $n \le -2$. Axiom (v') then follows by defining $\tilde{f}'_n$ for $n \ge 0$ by,

$$\tilde{f}'_n = (P')^{n+1}\tilde{f}'_{-1}.$$

□

Clearly, the $\tilde{f}'_n$ are non-negative and in $L^\infty_{\tilde{X}'}$.

**Lemma 7.6.** *Axiom (vi') (Size) $\|\tilde{f}'_n\|_{L^1(\tilde{X})} = \alpha(1-\alpha/4)\mu'(X') = \alpha(2-\alpha/2)$ for all $n \in \mathbb{Z}$.*

*Proof.*

$$\begin{aligned}\|\tilde{f}'_n\|_{L^1(\tilde{X}')} &= \int_{\{1\}}\int_{\tilde{X}} \tilde{f}'_n(x,1,s)d\tilde{\mu}d\xi + \int_{\{2\}}\int_{\tilde{X}} \tilde{f}'_n(x,2,s)d\tilde{\mu}d\xi \\ &= \xi(\{1\})\int_{\tilde{X}} \tilde{f}_n(x,s)d\tilde{\mu} + \xi(\{2\})\int_{\tilde{X}}(1-\alpha/2)\tilde{f}_{n-2M}(x,s)d\tilde{\mu} \\ &= \|\tilde{f}_n\|_{L^1(\tilde{X})} + (1-\alpha/2)\|\tilde{f}_{n-2M}\|_{L^1(\tilde{X})} \\ &= \alpha + (1-\alpha/2)\alpha = (2-\alpha/2)\alpha = \alpha(1-\alpha/4)\mu'(X')\end{aligned}$$

using $\|\tilde{f}_n\|_{L^1(\tilde{X})} = \alpha$ by Axiom (vi). From Axiom (v'), it follows $\|\tilde{f}'_n\|_{L^1(\tilde{X})} = \alpha(2-\alpha/2)$ for all $n$. □

The final two propositions establish separate bounds on $X \times \{1\}$ and $X \times \{2\}$ and Axiom (viii') then follows by a union bound. The first of these propositions is [[17], Proposition 4.2].

**Proposition 7.7.** *If $\kappa$ is sufficiently small (depending on $\epsilon, N$ and the $\tilde{f}_n$ but without dependence on $M$), then*

$$\sup_{n\in\mathbb{Z}} \pi_* \tilde{f}'_{2n}(x,1) \geq 1-\epsilon$$

*for all $x \in X$ outside a set of measure at most $\epsilon$.*

*Proof.* $\tilde{f}'_n(x,1,s) = \tilde{f}_n(x,s)$, and hence $\max(\tilde{f}_n(x,s) - \tilde{f}'_n(x,1,s), 0) = 0$, for all negative $n$, $x \in X$ and $s \in W$. $P$ is a contraction on $L^\infty$ i.e. $\|P\| \leq 1$ since,

$$|P\tilde{f}(x,s)| \leq \frac{1}{2m-1}\sum_{s_1 \in W\smallsetminus\{s^{-1}\}} |\tilde{f}(T_{s_1}^{-1}x, s_1)| \leq \frac{1}{2m-1}\sum_{s_1\in W\smallsetminus\{s^{-1}\}} \|\tilde{f}\|_{L^\infty} = \|\tilde{f}\|_{L^\infty}.$$

So for non-negative $n$, since $\tilde{f}'_n := (P')^{n+1}\tilde{f}'_{-1}$, $\tilde{f}'_n$ are uniformly bounded in $L^\infty$ by some quantity $B$ independent of $\kappa$. Recall, $\tilde{f}_n := (P)^{n+1}\tilde{f}_{-1}$ for all negative $n$. Crucially, when $i = 1$ by definition of $T'_g$ for $g \in F_m$, $P$ applied to $\tilde{f}_{-1}$ differs from $P'$ applied to $\tilde{f}'_{-1}$ only when $x \in E$.

$$\text{Claim:} \quad \int_{\tilde{X}} \max(\tilde{f}_n(x,s) - \tilde{f}'_n(x,1,s), 0)d\tilde{\mu}(x,s) \leq C_{B,n}\kappa, \quad \text{for all } n \in \mathbb{Z}.$$

The proof of this claim follows by induction. Starting with the base case for $n = 0$,

$$\begin{aligned}&\int_{\tilde{X}} \max(\tilde{f}_0(x,s) - \tilde{f}'_0(x,1,s), 0)d\tilde{\mu}(x,s) \\ &= \int_{\tilde{X}} \max\Big(\frac{1}{2m-1}\sum_{s_1\in W\smallsetminus\{s^{-1}\}} \big(\tilde{f}_{-1}(T_{s_1}^{-1}x, s_1) - \tilde{f}'_{-1}(T'^{-1}_{s_1}(x,1), s_1)\big), 0\Big)d\tilde{\mu}(x,s) \\ &= \int_{\tilde{X}} \max\Big(\frac{1}{2m-1}\Big[\sum_{s_1\in W\smallsetminus\{s^{-1},c_1,c_1^{-1}\}} \big(\tilde{f}_{-1}(T_{s_1}^{-1}x, s_1) - \tilde{f}'_{-1}(T'^{-1}_{s_1}(x,1), s_1)\big) \\ &+ \sum_{s_1\in\{c_1,c_1^{-1}\}\smallsetminus\{s^{-1}\}} \big(\tilde{f}_{-1}(T_{s_1}^{-1}x, s_1) - \tilde{f}_{-1}(T_{s_1}^{-1}x, s_1)\big)\Big], 0\Big)d\tilde{\mu}(x,s)\end{aligned}$$

$$
\begin{aligned}
&= \int_{E\times W} \max\left(\frac{1}{2m-1}\sum_{s_1\in W\smallsetminus\{s^{-1},c_1,c_1^{-1}\}}\left(\tilde{f}_{-1}(T_{s_1}^{-1}x,s_1)-\tilde{f}'_{-1}(T'^{-1}_{s_1}(x,1),s_1)\right),0\right)d\tilde{\mu}(x,s)\\
&\le \frac{1}{2m\times(2m-1)}\int_E\sum_{s\in W}\sum_{s_1\in W\smallsetminus\{s^{-1},c_1,c_1^{-1}\}}\max\left(\tilde{f}_{-1}(T_{s_1}^{-1}x,s_1)-\tilde{f}'_{-1}(T'^{-1}_{s_1}(x,1),s_1),0\right)d\mu(x)\\
&\le \frac{1}{2m\times(2m-1)}\int_E(2m-1)\sum_{s_1\in W\smallsetminus\{c_1,c_1^{-1}\}}\max\left(\tilde{f}_{-1}(T_{s_1}^{-1}x,s_1)-\tilde{f}'_{-1}(T'^{-1}_{s_1}(x,1),s_1),0\right)d\mu(x)\\
&\le \frac{1}{2m}\int_E\sum_{s_1\in W\smallsetminus\{c_1,c_1^{-1}\}}\max\left(\tilde{f}_{-1}(T_{s_1}^{-1}x,s_1)-\tilde{f}'_{-1}(T'^{-1}_{s_1}(x,1),s_1),0\right)d\mu(x)\\
&\le \frac{1}{2m}\int_E\sum_{s_1\in W\smallsetminus\{c_1,c_1^{-1}\}}2Bd\mu(x)\\
&\le \frac{2B(m-1)}{m}\kappa = C_{B,0}\kappa
\end{aligned}
$$

In order to identify a pattern, look at the case for $n=1$,

$$
\begin{aligned}
&\int_{\tilde{X}}\max(\tilde{f}_1(x,s)-\tilde{f}'_1(x,1,s),0)d\tilde{\mu}(x,s)\\
&= \int_{E\times W}\max(\tilde{f}_1(x,s)-\tilde{f}'_1(x,1,s),0)d\tilde{\mu}(x,s)\\
&+ \int_{\cup_{s'\in W}T_{s'}E\cap E^c\times W}\max(\tilde{f}_1(x,s)-\tilde{f}'_1(x,1,s),0)d\tilde{\mu}(x,s)\\
&\le 2B\kappa+\int_{\cup_{s'\in W}T_{s'}E\cap E^c\times W}\max(\tilde{f}_1(x,s)-\tilde{f}'_1(x,1,s),0)d\tilde{\mu}(x,s)\\
&= 2B\kappa+\sum_{s'\in W}\frac{1}{2m}\sum_{s\in W}\int_{T_{s'}E\cap E^c}\max(\frac{1}{2m-1}\sum_{s_1\in W\smallsetminus\{s^{-1}\}}\tilde{f}_0(T_{s_1}^{-1}x,s_1)-\tilde{f}'_0(T'^{-1}_{s_1}(x,1),s_1),0)d\mu(x)\\
&\le 2B\kappa+\sum_{s'\in W}\frac{1}{2m}\int_{T_{s'}E\cap E^c}(2m-1)\sum_{s_1\in W}\max(\frac{1}{2m-1}\tilde{f}_0(T_{s_1}^{-1}x,s_1)-\tilde{f}'_0(T'^{-1}_{s_1}(x,1),s_1),0)d\mu(x)\\
&= 2B\kappa+\sum_{s'\in W}\frac{1}{2m}\int_{T_{s'}E\cap E^c}\sum_{s_1\in W}\max(\tilde{f}_0(T_{s_1}^{-1}x,s_1)-\tilde{f}'_0(T_{s_1}^{-1}x,1,s_1),0)d\mu(x)
\end{aligned}
$$

Since $x\notin E$ and $x\in T_{s'}E$, using the previous case, the sum is only non-zero if $s_1=s'$. Otherwise $\tilde{f}_0(T_{s_1}^{-1}x,s_1)=\tilde{f}'_0(T'^{-1}_{s_1}(x,1),s_1)$. So,

$$
\begin{aligned}
&= 2B\kappa+\sum_{s'\in W}\frac{1}{2m}\int_{T_{s'}E\cap E^c}\max(\tilde{f}_0(T_{s'}^{-1}x,s')-\tilde{f}'_0(T_{s'}^{-1}x,1,s'),0)d\mu(x)\\
&\le 2B\kappa+\int_{\tilde{X}}\max(\tilde{f}_0(x,s')-\tilde{f}'_0(x,1,s'),0)d\tilde{\mu}(x,s')\\
&= \left(1+\frac{m-1}{m}\right)2B\kappa = C_{B,1}\kappa.
\end{aligned}
$$

Proceed by induction, take $A=\cup_{g:|g|=n}T_gE\cap(\cup_{g:|g|<n}T_gE)^c$ and $A'=\cup_{g:|s'g|=n}T_{s'}T_gE\cap$

$(\cup_{g:|g|<n} T_g E)^c$

$$\begin{aligned}
&\int_{\tilde{X}} \max(\tilde{f}_n(x,s) - \tilde{f}'_n(x,1,s),0) d\tilde{\mu}(x,s) \\
&= \int_{\cup_{g:|g|<n} T_g E \times W} \max(\tilde{f}_n(x,s) - \tilde{f}'_n(x,1,s),0) d\tilde{\mu}(x,s) \\
&+ \int_{A\times W} \max(\tilde{f}_n(x,s) - \tilde{f}'_n(x,1,s),0) d\tilde{\mu}(x,s) \\
&= 2DB\kappa + \int_{A\times W} \max(\tilde{f}_n(x,s) - \tilde{f}'_n(x,1,s),0) d\tilde{\mu}(x,s) \\
&= 2DB\kappa + \int_A \frac{1}{2m} \sum_{s\in W} \max(\frac{1}{2m-1} \sum_{s_1 \in W \setminus \{s^{-1}\}} \left(\tilde{f}_{n-1}(T_{s_1}^{-1}x, s_1) - \tilde{f}'_{n-1}(T_{s_1}^{-1}x,1,s_1)\right), 0) d\tilde{\mu}(x,s) \\
&\leq 2DB\kappa + \sum_{s'\in W} \int_{A'} \frac{1}{2m} \sum_{s_1\in W} \max(\tilde{f}_{n-1}(T_{s_1}^{-1}x, s_1) - \tilde{f}'_{n-1}(T_{s_1}^{-1}x,1,s_1),0) d\tilde{\mu}(x,s)
\end{aligned}$$

Since, $x \in A'$ the term above is non-zero only when $s_1 = s'$. Otherwise,

$$\tilde{f}_{n-1}(T_{s_1}^{-1}x, s_1) = \tilde{f}'_{n-1}(T'^{-1}_{s_1}(x,1), s_1).$$

$$\begin{aligned}
&\leq 2DB\kappa + \frac{1}{2m} \sum_{s'\in W} \int_{A'} \max(\tilde{f}_{n-1}(T_{s'}^{-1}x, s') - \tilde{f}'_{n-1}(T_{s'}^{-1}x,1,s'),0) d\mu(x) \\
&\leq 2DB\kappa + \int_{\tilde{X}} \max(\tilde{f}_{n-1}(x,s') - \tilde{f}'_{n-1}(x,1,s'),0) d\mu(x,s') \\
&\leq 2DB\kappa + C_{B,n-1}\kappa.
\end{aligned}$$

Where $D = D(m,n) = \sum_{k=1}^{n-1} 2m \times (2m-1)^{k-1}$. Thus the claim follows by induction. Recall, $\pi_* \tilde{f}(x) = \frac{1}{2m} \sum_{s\in W} \tilde{f}(x,s)$, and

$$\begin{aligned}
C_{B,n}\kappa &\geq \int_{\tilde{X}} \max(\tilde{f}_n(x,s) - \tilde{f}'_n(x,1,s),0) d\tilde{\mu}(x,s) \\
&\geq \int_X \max(\frac{1}{2m} \sum_{s\in W} \tilde{f}_n(x,s) - \frac{1}{2m} \sum_{s\in W} \tilde{f}'_n(x,1,s), 0) d\mu(x) \\
&= \int_X \max(\pi_* \tilde{f}_n(x) - \pi_* \tilde{f}'_n(x,1), 0) d\mu(x).
\end{aligned}$$

Furthermore, for every $-N < n < N$ and every $x$

$$\pi_* \tilde{f}_{2n}(x) \leq \pi_* \tilde{f}_{2n}(x) - \pi_* \tilde{f}'_{2n}(x,1) + \sup_{-N\leq n\leq N} \pi_* \tilde{f}'_{2n}(x,1)$$

Thus, taking suprema

$$\begin{aligned}
\sup_{-N<n<N} \pi_* \tilde{f}_{2n}(x) &\leq \sup_{-N<n<N} \left(\pi_* \tilde{f}_{2n}(x) - \pi_* \tilde{f}'_{2n}(x,1) + \sup_{-N\leq n\leq N} \pi_* \tilde{f}'_{2n}(x,1)\right) \\
&= \sup_{-N<n<N} \left(\pi_* \tilde{f}_{2n}(x) - \pi_* \tilde{f}'_{2n}(x,1)\right) + \sup_{-N\leq n\leq N} \pi_* \tilde{f}'_{2n}(x,1),
\end{aligned}$$

and rearranging,

$$\sup_{-N<n<N} \pi_* \tilde{f}_{2n}(x) - \sup_{-N\le n\le N} \pi_* \tilde{f}'_{2n}(x,1) \le \sup_{-N<n<N} \left(\pi_* \tilde{f}_{2n}(x) - \pi_* \tilde{f}'_{2n}(x,1)\right).$$

In particular,

$$\begin{aligned}\max\left(\sup_{-N<n<N} \pi_* \tilde{f}_{2n}(x) - \sup_{-N\le n\le N} \pi_* \tilde{f}'_{2n}(x,1), 0\right) &\le \max\left(\sup_{-N<n<N} \left(\pi_* \tilde{f}_{2n}(x) - \pi_* \tilde{f}'_{2n}(x,1)\right), 0\right) \\ &= \sup_{-N<n<-N} \max\left(\pi_* \tilde{f}_{2n}(x) - \pi_* \tilde{f}'_{2n}(x,1), 0\right).\end{aligned}$$

So,

$$\begin{aligned}&\int_X \max\left(\sup_{-N\le n\le N} \pi_* \tilde{f}_{2n}(x) - \sup_{-N\le n\le N} \pi_* \tilde{f}'_{2n}(x,1), 0\right) d\mu(x) \\ &\le \int_X \max\left(\sup_{-N\le n\le N} \left(\pi_* \tilde{f}_{2n}(x) - \pi_* \tilde{f}'_{2n}(x,1)\right), 0\right) d\mu(x) \\ &= \sup_{-N\le n\le N} \int_X \max\left(\pi_* \tilde{f}_{2n}(x) - \pi_* \tilde{f}'_{2n}(x,1), 0\right) d\mu(x) \\ &\le \sup_{-N\le n\le N} C_{B,2n}\kappa = C'_{B,N}\kappa.\end{aligned}$$

Thus, $\int_X \max\left(\sup_{-N\le n\le N} \pi_* \tilde{f}_{2n}(x) - \sup_{-N\le n\le N} \pi_* \tilde{f}'_{2n}(x,1), 0\right) d\mu(x) \le C'_{B,N}\kappa$, for some $C'_{B,N}$ independent of $\kappa$. Markov's inequality states that

$$\mu(\{x \in X : |f(x)| \ge \delta\}) \le \frac{1}{\delta}\int_X |f(x)|$$

Since, $\frac{1}{\delta}\int_X \max\left(\sup_{-N\le n\le N} \pi_* \tilde{f}_{2n}(x) - \sup_{-N\le n\le N} \pi_* \tilde{f}'_{2n}(x,1), 0\right) d\mu(x) \le \frac{1}{\delta} C'_{B,N}\kappa$, taking $\delta = \frac{\epsilon}{3}$ and $\kappa$ small enough say $\kappa < \dfrac{\epsilon^2}{3^2 C'_{B,N}}$ then

$$\max\left(\sup_{-N\le n\le N} \pi_* \tilde{f}_{2n}(x) - \sup_{-N\le n\le N} \pi_* \tilde{f}'_{2n}(x,1), 0\right) > \epsilon/3$$

for $x \in X$ of measure at most $\dfrac{1}{\delta} C'_{B,N}\kappa < \epsilon/3$. Therefore reversing the inequality, and taking the complement of the set,

$$\sup_{-N\le n\le N} \pi_* \tilde{f}_{2n}(x) - \sup_{-N\le n\le N} \pi_* \tilde{f}'_{2n}(x,1) \le \max\left(\sup_{-N\le n\le N} \pi_* \tilde{f}_{2n}(x) - \sup_{-N\le n\le N} \pi_* \tilde{f}'_{2n}(x,1), 0\right) < \epsilon/3, \tag{2}$$

for all $x \in X$ outside of a set of measure at most $\epsilon/3$. Now by Egorov's theorem and Axiom (viii), there exists a natural number $N$ such that

$$\sup_{-N\le n\le N} \pi_* \tilde{f}_{2n}(x) \ge 1 - \epsilon/4 \ge 1 - \epsilon/3$$

for all $x \in X$ outside a set of measure at most $\epsilon/3$. Thus using this

$$1 - \epsilon/3 - \sup_{-N \le n \le N} \pi_* \tilde{f}'_{2n}(x,1) \le \sup_{-N \le n \le N} \pi_* \tilde{f}_{2n}(x) - \sup_{-N \le n \le N} \pi_* \tilde{f}'_{2n}(x,1) < \epsilon/3.$$

Hence, rearranging

$$1 - \frac{2\epsilon}{3} < \sup_{-N \le n \le N} \pi_* \tilde{f}'_{2n}(x,1)$$

and

$$\sup_{n \in \mathbb{Z}} \pi_* \tilde{f}'_{2n}(x,1) \ge \sup_{-N \le n \le N} \pi_* \tilde{f}'_{2n}(x,1) \ge 1 - \frac{2\epsilon}{3} > 1 - \epsilon$$

for all $x \in X$ outside of a set of measure at most $2/3\epsilon$, concluding the proof. $\square$

The following is [[17], Proposition 4.3].

**Proposition 7.8.** *If $\kappa$ is sufficiently small (depending on $\epsilon, N$ and the $\tilde{f}_n$ but without dependence on $M$),*

$$\sup_{n \in \mathbb{Z}} \pi_* \tilde{f}'_{2n}(x,2) \ge 1 - \epsilon,$$

*for all $x \in X$ outside a set of measure at most $\epsilon$.*

*Proof.* Split $\tilde{f}'_n$ into,

$$\tilde{f}'_n = \tilde{f}'_{n,1} + \tilde{f}'_{n,2}$$

where for negative $n$,

$$\tilde{f}'_{n,i}(x,j,s) = \begin{cases} \tilde{f}'_n(x,j,s) & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

For non-negative $n$, $\tilde{f}'_{n,i}$ is propagated by $P'$, $\tilde{f}'_{n,i} := (P')^{n+1} \tilde{f}'_{-1,i}$. More specifically,

$$\tilde{f}'_{n,i}(x,j,s) = \begin{cases} \frac{1}{(2m-1)^n} \sum_{w = s_1 \cdots s_n \in F_m, s_1 \neq s^{-1}, |w| = n} \tilde{f}'_{-1}(T'^{-1}_w(x,j), s_n) & \text{if } T'^{-1}_w(x,j) = (T^{-1}_w x, i) \\ 0 & \text{otherwise.} \end{cases}$$

Take $i = 2$, for all $n < 2M$ $\tilde{f}'_{n,2}$ is supported on $X_0 \times \{2\} \times W$, and $\tilde{f}'_{n,2}(x,2,s) = (1 - \alpha/2)\tilde{f}_{n-2M}(x,s)$, for all $x \in X$ and $s \in W$. For negative $n$, this follows from Axiom (vii'). Take $0 \le n < 2M$ arbitrary,

$$\begin{aligned} \tilde{f}'_{n,2}(x,j,s) &= \begin{cases} \frac{1}{(2m-1)^n} \sum_{w = s_1 \cdots s_n \in F_m, s_1 \neq s^{-1}, |w| = n} \tilde{f}'_{-1}(T'^{-1}_w(x,j), s_n) & \text{if } T'^{-1}_w(x,j) = (T^{-1}_w x, 2) \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} \frac{(1-\alpha/2)}{(2m-1)^n} \sum_{w = s_1 \cdots s_n \in F_m, s_1 \neq s^{-1}, |w| = n} \tilde{f}_{-1-2M}(T^{-1}_w x, s_n) & \text{if } T'^{-1}_w(x,j) = (T^{-1}_w x, 2) \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

$$= \begin{cases} (1-\alpha/2)\tilde{f}_{-1-2M+n}(x,s) & \text{if } T_w'^{-1}(x,j) = (T_w^{-1}x,2) \\ 0 & \text{otherwise.} \end{cases}$$

Since, $-1-2M+n<0$ then $\tilde{f}'_{n,2} = (1-\alpha/2)\tilde{f}_{-1-2M+n}$ or zero, therefore $\tilde{f}'_{n,2}$ is supported on $X_0 \times \{2\} \times W$.
Observe that the $\tilde{f}'_{n,1}$ component of $\tilde{f}'_n$ does not depend on $M$. From Theorem 4.3, for any $\tilde{f} \in L^\infty(\tilde{X})$, $(P^{2n}\tilde{f})$ converges pointwise almost everywhere and in $L^1$ norm to $\frac{1}{\mu(X)}\int_{\tilde{X}} \tilde{f} d\tilde{\mu}$.
So applying this to $\tilde{f}'_{n,1}$, $\tilde{f}'_{n,1}$ converges pointwise and almost everywhere as $n$ goes to infinity to the constant $\alpha/2$. Furthermore, $\pi_*\tilde{f}'_{n,1}$ converges pointwise almost everywhere to the same constant,

$$\begin{aligned} \frac{1}{\mu'(\tilde{X}')}\int_{X'} \pi_*\tilde{f}'_{-1,1}(x,i)d\mu'(x,i) &= \frac{1}{\mu'(\tilde{X}')}\int_X \pi_*\tilde{f}_{-1}(x)d\mu(x) \\ &= \frac{1}{\mu'(\tilde{X}')}\int_X \frac{1}{2m}\sum_{s\in W}\tilde{f}_{-1}(x,s)d\mu(x) \\ &= \frac{1}{2}\int_{\tilde{X}} \tilde{f}_{-1}(x,s)d\tilde{\mu}(x,s) \\ &= \alpha/2. \end{aligned}$$

Thus by Egorov's theorem and assuming $M$ is sufficiently large (depending on previous quantities such as $\epsilon$, $\kappa$ and the $\tilde{f}_n$ but without any circular dependency on $M$ on itself),

$$\inf_{n\ge 2M-2N} \pi_*\tilde{f}'_{n,1}(x,2) \ge \alpha/2 - \epsilon/3 \tag{3}$$

for all $x \in X$ outside of a set of measure at most $\epsilon/3$.
By the same method as was used in the previous proposition, to obtain equation (2), as well as the facts shown previously about $\tilde{f}'_{2n,2}$ it is possible to show,

$$(1-\alpha/2)\sup_{M-N\le n\le M+N} \pi_*\tilde{f}_{2n-2M}(x) - \sup_{M-N\le n\le M+N} \pi_*\tilde{f}'_{2n,2}(x,2) \le \frac{\epsilon}{3}.$$

for all $x \in X$ outside of a set of measure at most $\epsilon/3$. (See the appendix, Section 9.5 for details). Rearranging,

$$\sup_{M-N\le n\le M+N} \pi_*\tilde{f}'_{2n,2}(x,2) \ge (1-\alpha/2)\sup_{M-N\le n\le M+N} \pi_*\tilde{f}_{2n-2M}(x) - \frac{\epsilon}{3},$$

and using (3),

$$\sup_{M-N\le n\le M+N} \pi_*\tilde{f}'_{2n,2}(x,2) + \inf_{n\ge 2M-2N} \pi_*\tilde{f}'_{n,1}(x,2) \ge (1-\alpha/2)\sup_{M-N\le n\le M+N} \pi_*\tilde{f}_{2n-2M}(x) - \frac{\epsilon}{3} + \alpha/2 - \epsilon/3.$$

Noting,

$$\begin{aligned}
\sup_{n\in\mathbb{Z}} \pi_* \tilde{f}'_{2n}(x,2) &\geq \sup_{n\geq M-N} \pi_* \tilde{f}'_{2n}(x,2) \\
&= \sup_{n\geq M-N} \left(\pi_* \tilde{f}'_{2n,2}(x,2) + \pi_* \tilde{f}'_{2n,1}(x,2)\right) \\
&= \sup_{n\geq M-N} \pi_* \tilde{f}'_{2n,2}(x,2) + \sup_{n\geq M-N} \pi_* \tilde{f}'_{2n,1}(x,2) \\
&\geq \sup_{n\geq M-N} \pi_* \tilde{f}'_{2n,2}(x,2) + \inf_{n\geq M-N} \pi_* \tilde{f}'_{2n,1}(x,2) \\
&\geq \sup_{M-N\leq n\leq M+N} \pi_* \tilde{f}'_{2n,2}(x,2) + \inf_{n\geq 2M-2N} \pi_* \tilde{f}'_{n,1}(x,2),
\end{aligned}$$

therefore,

$$\sup_{n\in\mathbb{Z}} \pi_* \tilde{f}'_{2n}(x,2) \geq (1-\alpha/2) \sup_{M-N\leq n\leq M+N} \pi_* \tilde{f}_{2n-2M}(x) + \alpha/2 - \frac{2\epsilon}{3}.$$

Finally, with a change of variables,

$$\sup_{n\in\mathbb{Z}} \pi_* \tilde{f}'_{2n}(x,2) \geq (1-\alpha/2) \sup_{-N\leq n\leq N} \pi_* \tilde{f}_{2n}(x) + \alpha/2 - \frac{2\epsilon}{3}.$$

for all $x \in X$ outside a set of measure at most $2\epsilon/3$. Now by Egorov's theorem and Axiom (viii), there exists a natural number $N$ such that $\sup_{-N\leq n\leq N} \pi_* \tilde{f}_{2n}(x) \geq 1-\epsilon/3$, for all $x \in X$ outside a set of measure at most $\epsilon/3$. Combining this with the above,

$$\begin{aligned}
\sup_{n\in\mathbb{Z}} \pi_* \tilde{f}'_{n,2} &\geq \alpha/2 + (1-\alpha/2) \sup_{-N\leq n\leq N} \pi_* \tilde{f}_{2n}(x) - 2\epsilon/3 \\
&\geq \alpha/2 + (1-\alpha/2)(1-\epsilon/3) - 2\epsilon/3 \geq 1-\epsilon,
\end{aligned}$$

outside of a set of measure a most $\epsilon$, concluding the proof. □

# 8 Concluding Remarks

The details of the case of the free group on 2 generators were outlined in [17] and this report attempted to extend to the case on $m$ generators. The structure of the proof remained the same and required no adjustment. The details in most of the construction of a counterexample only required minimal alterations. The biggest difference arose when defining the good system in Section 4, where the structure of $X_2$ and the dynamics of the maps $T_{c_i}$ for $i = 1, \ldots, m$ had to be adapted. Finally, the measure and the definition of the maps in the initial construction in Section 6 also required changes in order to extend to the general case.

# 9 Appendix

## 9.1 Definition of Ergodicity

**Definition 9.1.** *(Ergodic). Let $(X, \mathcal{X}, \mu)$ be a probability space and $T : X \to X$ a measure preserving transformation, then $T$ is an ergodic transformation (or $\mu$ is an ergodic measure) if, for every $B \in \mathcal{X}$ such that*

$$T^{-1}(B) = B.$$

*Then $\mu(B) = 0$ or $1$.*

The following is used as an alternative criterion for establishing ergodicity throughout the report.

**Proposition 9.2.** *(Alternative characterisation of ergodicity). Let $T$ be a measure preserving transformation of $(X, \mathcal{X}, \mu)$. The following are equivalent,*

*(i) $T$ is ergodic.*

*(ii) Whenever $f \in L^1(X)$ satisfies $f = f \circ T$ $\mu$−almost everywhere, then $f$ is constant $\mu$-almost everywhere.*

A proof can be found in [14].

## 9.2 Measure Theory

As mentioned in the introduction the report uses the following obtained from [14].

**Theorem.** *(The Kolmogorov extension theorem). Let $\mathcal{A}$ be an algebra of subsets of $X$. Suppose that $\mu : \mathcal{A} \to \mathbb{R} \cup \{+\infty\}$ satisfies:*

*(i) $\mu(\varnothing) = 0$,*

*(ii) There exists finitely or countably many sets $X_n \in \mathcal{A}$ such that $X = \bigcup_n X_n$ and $\mu(X_n) < \infty$, and*

*(iii) If $E_n \in \mathcal{A}$, $n \geq 1$ are pairwise disjoint and if $\bigcup_{n=1}^{\infty} E_n \in \mathcal{A}$ then*

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n).$$

*Then there is a unique measure, $\mu : \mathcal{B}(\mathcal{A}) \to \mathbb{R}^+$, which is an extension of $\mu : \mathcal{A} \to \mathbb{R}$.*

## 9.3 Borel Cantelli

**Lemma 9.3.** *(Borel-Cantelli). Let* $(En : n \in \mathbb{N})$ *be a sequence of measurable sets from a fixed probability space* $(X, \mathcal{X}, \mu)$, $\mu(X) = 1$. *Then*

$$\sum_n \mu(E_n) < \infty \quad \Rightarrow \quad \mu(\limsup_n E_n) = \mu(E_n \quad i.o.) = 0,$$

*where* $E_n$ *i.o is shorthand for* $E_n$ *occurs infinitely often in* $n$.

A proof can be found in [4]. $\mu(E_n \quad i.o.) = 0$ means that almost surely a sequence of events $E_n$ does not occur infinitely often. If this is true then almost surely the sequence of events occurs for only finitely many $n$ and thus, almost surely, there exists a random constant $N_0 < \infty$ such that $\mu(E_n) = 0$ for all $n \geq N_0$. Borel-Cantelli is used in Proposition 3.2 and $\mu(E_n \quad i.o.) = 0$ is interpreted as described here.

## 9.4 Mixing

The following was adapted from [18]. If $T$ is a measure preserving transformation of a probability space then from the ergodic theorem it can be deduced that $T$ is ergodic if and only if $\forall A, B \in \mathcal{X}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} m(T^{-i}A \cap B) = m(A)m(B)$$

The concepts of strong and weak mixing can be obtained by making slight adjustments to the above statement.

**Definition 9.4.** *(Mixing). Let* $T$ *be a measure preserving transformation of a probability space* $(X, \mathcal{X}, \mu)$

- $T$ *is weak mixing if* $\forall A, B \in \mathcal{X}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \left| m(T^{-i}A \cap B) - m(A)m(B) \right| = 0$$

- $T$ *is strong-mixing if* $\forall A, B \in \mathcal{X}$

$$\lim_{n\to\infty} m(T^{-i}A \cap B) = m(A)m(B)$$

**Remark 9.5.** *Strong* $\Rightarrow$ *Weak* $\Rightarrow$ *Ergodic.*

### 9.5 Calculation in Proposition 7.8

For all $n < 2M$. The objective is to find a bound for $\tilde{f}'_{n,2}$ by repeating the arguments used to prove Proposition 7.7. If $\kappa$ is sufficiently small depending on $\epsilon, N$ but without any dependence on $M$, then the claim is

$$(1-\alpha/2) \sup_{M-N\leq n\leq M+N} \pi_* \tilde{f}_{2n-2M}(x) - \sup_{M-N\leq n\leq M+N} \pi_* \tilde{f}'_{2n,2}(x,2) \leq \frac{\epsilon}{3}.$$

for all $x \in X$ outside of a set of measure at most $\epsilon/3$.

*Proof.* By Construction,

$$\tilde{f}'_{n,2}(x,2,s) = (1-\alpha/2)\tilde{f}_{n-2M}(x,s)$$

for all non-negative $n$ and for $0 \leq n < 2M$, $(1-\alpha/2)\tilde{f}_{n-2M}(x,s) = 0$ if $x \in X \setminus X_0$ and furthermore, either $\tilde{f}'_{n,2}(x,2,s) = (1-\alpha/2)\tilde{f}_{n-2M}(x,s)$ or $x$ is such that $T'^{-1}_g(x,2) = (T_g x, 1)$ in which case $\tilde{f}'_{n,2}(x,2,s) = 0$. Hence $\max((1-\alpha/2)\tilde{f}_n(x,s) - \tilde{f}'_{n,2}(x,2,s), 0) = 0$ for all $n < 2M$. Observe $P$ is a contraction on $L^\infty$. So for non-negative $n$, $\tilde{f}'_n$ are uniformly bounded in $L^\infty$ by some quantity $B$ independent of $\kappa$. Furthermore, for all non-negative $n$ $\tilde{f}'_{n,2} := (P')^{n+1}\tilde{f}'_{-1,2}$ and $\tilde{f}_n := (P)^{n+1}\tilde{f}_{-1}$. Recall the definition of $P$ as

$$Pf(x,s) = \frac{1}{2m-1} \sum_{s'\in W\setminus\{s^{-1}\}} \tilde{f}(T_{s'}^{-1}x, s')$$

and similarly,

$$P'\tilde{f}(x,j,s) = \frac{1}{2m-1} \sum_{s'\in W\setminus\{s^{-1}\}} \tilde{f}(T'^{-1}_{s'}x, j, s').$$

Where $T'$ was defined previously. The claim is as follows,

$$\int_{\tilde{X}} \max(\tilde{f}_n(x,s) - \tilde{f}'_{n,2}(x,2,s), 0)\, d\tilde{\mu}(x,s) \leq C_{B,n}\kappa.$$

The proof is similar to previous. Start with $n = 2M$, for $n = 2M-1$ the integral is 0 as previous arguments demonstrate and the proof follows by induction, using a method similar to that used for equation (2). So

$$\int_X \max((1-\alpha/2)\pi_*\tilde{f}_n(x) - \pi_*\tilde{f}'_{n,2}(x,2), 0)\, d\tilde{\mu}(x,s) \leq C_{B,n}\kappa.$$

This is true for every $n$. For every $M-N < n < M+N$ and every $x$,

$$(1-\alpha/2)\pi_*\tilde{f}_{2n-2M}(x) \leq (1-\alpha/2)\pi_*\tilde{f}_{2n-2M}(x) - \pi_*\tilde{f}'_{2n,2}(x,2) + \sup_{M-N\leq n\leq M+N} \pi_*\tilde{f}'_{2n,2}(x,2).$$

Thus, taking suprema,

$$
\begin{aligned}
&(1-\alpha/2) \sup_{M-N<n<M+N} \pi_* \tilde{f}_{2n-2M}(x) \\
&\le \sup_{M-N<n<M+N} \left((1-\alpha/2)\pi_* \tilde{f}_{2n-2M}(x) - \pi_* \tilde{f}'_{2n,2}(x,2)\right) + \sup_{M-N<n<M+N} \left( \sup_{M-N\le n\le M+N} \pi_* \tilde{f}'_{2n,2}(x,2) \right) \\
&= \sup_{M-N<n<M+N} \left((1-\alpha/2)\pi_* \tilde{f}_{2n-2M}(x) - \pi_* \tilde{f}'_{2n,2}(x,2)\right) + \sup_{M-N\le n\le M+N} \pi_* \tilde{f}'_{2n,2}(x,2),
\end{aligned}
$$

and rearranging,

$$
\begin{aligned}
&(1-\alpha/2) \sup_{M-N<n<M+N} \pi_* \tilde{f}_{2n-2M}(x) - \sup_{M-N\le n\le M+N} \pi_* \tilde{f}'_{2n,2}(x,2) \\
&\le \sup_{M-N<n<M+N} \left((1-\alpha/2)\pi_* \tilde{f}_{2n-2M}(x) - \pi_* \tilde{f}'_{2n,2}(x,2)\right)
\end{aligned}
$$

and in particular,

$$
\begin{aligned}
&\max\left((1-\alpha/2) \sup_{M-N<n<M+N} \pi_* \tilde{f}_{2n-2M}(x) - \sup_{M-N\le n\le M+N} \pi_* \tilde{f}'_{2n,2}(x,2), 0\right) \\
&\le \max\left( \sup_{M-N<n<M+N} \left((1-\alpha/2)\pi_* \tilde{f}_{2n-2M}(x) - \pi_* \tilde{f}'_{2n,2}(x,2)\right), 0\right) \\
&= \sup_{M-N<n<M+N} \max\left((1-\alpha/2)\pi_* \tilde{f}_{2n-2M}(x) - \pi_* \tilde{f}'_{2n,2}(x,2), 0\right).
\end{aligned}
$$

So,

$$
\begin{aligned}
&\int_X \max\left((1-\alpha/2) \sup_{M-N\le n\le M+N} \pi_* \tilde{f}_{2n-2M}(x) - \sup_{M-N\le n\le M+N} \pi_* \tilde{f}'_{2n,2}(x,2), 0\right) d\mu(x) \\
&\le \int_X \max\left( \sup_{M-N\le n\le M+N} \left((1-\alpha/2)\pi_* \tilde{f}_{2n-2M}(x) - \pi_* \tilde{f}'_{2n,2}(x,2)\right), 0\right) d\mu(x) \\
&= \sup_{M-N\le n\le M+N} \int_X \max\left((1-\alpha/2)\pi_* \tilde{f}_{2n-2M}(x) - \pi_* \tilde{f}'_{2n,2}(x,2), 0\right) d\mu(x) \\
&\le \sup_{M-N\le n\le M+N} C_{B,n}\kappa = C'_{B,N}\kappa.
\end{aligned}
$$

Thus,

$$
\int_X \max\left((1-\alpha/2) \sup_{M-N\le n\le M+N} \pi_* \tilde{f}_{2n-2M}(x) - \sup_{M-N\le n\le M+N} \pi_* \tilde{f}'_{2n,2}(x,2), 0\right) d\mu(x) \le C'_{B,N}\kappa
$$

for some $C'_{B,N}$ independent of $\kappa$. Markov's inequality states that

$$
\mu(\{x \in X \| f(x)| \ge \delta\}) \le \frac{1}{\delta} \int_X |f(x)|
$$

and since

$$\frac{1}{\delta}\int_X \max\left((1-\alpha/2)\sup_{M-N\le n\le M+N}\pi_*\tilde{f}_{2n-2M}(x) - \sup_{M-N\le n\le M+N}\pi_*\tilde{f}'_{2n,2}(x,2), 0\right) d\mu(x) \le \frac{1}{\delta}C'_{B,N}\kappa$$

taking $\delta = \dfrac{\epsilon}{3}$ and $\kappa$ small enough say $\kappa < \dfrac{\epsilon^2}{3^2 C'_{B,N}}$,

$$\max\left((1-\alpha/2)\sup_{M-N\le n\le M+N}\pi_*\tilde{f}_{2n-2M}(x) - \sup_{M-N\le n\le M+N}\pi_*\tilde{f}'_{2n,2}(x,2), 0\right) > \epsilon/3,$$

for $x \in X$ of measure less than or equal to $\frac{1}{\delta}C'_{B,N}\kappa < \epsilon/3$. Furthermore, taking the reverse inequality with the complement of the set the previous inequality held for,

$$\begin{aligned}
&(1-\alpha/2)\sup_{M-N\le n\le M+N}\pi_*\tilde{f}_{2n-2M}(x) - \sup_{M-N\le n\le M+N}\pi_*\tilde{f}'_{2n,2}(x,2)\\
&\le \max\left((1-\alpha/2)\sup_{M-N\le n\le M+N}\pi_*\tilde{f}_{2n-2M}(x) - \sup_{M-N\le n\le M+N}\pi_*\tilde{f}'_{2n,2}(x,2), 0\right)\\
&\le \max\left((1-\alpha/2)\sup_{M-N\le n\le M+N}\pi_*\tilde{f}_{2n-2M}(x) - \sup_{M-N\le n\le M+N}\pi_*\tilde{f}'_{2n,2}(x,2), 0\right)
\end{aligned}$$

therefore,

$$(1-\alpha/2)\sup_{M-N\le n\le M+N}\pi_*\tilde{f}_{2n-2M}(x) - \sup_{M-N\le n\le M+N}\pi_*\tilde{f}'_{2n,2}(x,2) < \epsilon/3.$$

for $x \in X$ outside a set of measure at most $\epsilon/3$. □

# References


[1] Alexander I Bufetov. Convergence of spherical averages for actions of free groups. *Annals of Mathematics*, 155(3):929–944, 2002.

[2] Allen Hatcher. *Algebraic topology*. Cambridge University Press, 2002.

[3] Derek Holt. (MA249) Algebra II course notes.

[4] Wilfrid Kendall. ST318 Probability theory course notes.

[5] Elon Lindenstrauss. Pointwise theorems for amenable groups. *Inventiones Mathematicae*, 146(2):259–295, 2001.

[6] Amos Nevo and Elias M Stein. A generalization of Birkhoff's pointwise ergodic theorem. *Acta Mathematica*, 173(1):135–154, 1994.

[7] Donald S Ornstein. On pointwise behavior of iterates of a self-adjoint operator. *Journal of Mathematics and Mechanics*, 18(5):473, 1968.

[8] Donald S Ornstein and Benjamin Weiss. Entropy and isomorphism theorems for actions of amenable groups. *Journal d'Analyse Mathématique*, 48(1):1–141, 1987.

[9] Karl E Petersen. *Ergodic theory*, volume 2. Cambridge University Press, 1989.

[10] Mark Pollicott and Richard Sharp. Ergodic theorems for actions of hyperbolic groups. *Proceedings of the American Mathematical Society*, 141(5):1749–1757, 2013.

[11] Murray Rosenblatt. *Markov Processes, Structure and Asymptotic Behavior: Structure and Asymptotic Behavior*, volume 184. Springer Science & Business Media, 2012.

[12] Gian-Carlo Rota. An alternierende verfahren for general positive operators. *Bulletin of the American Mathematical Society*, 68(2):95–102, 1962.

[13] Halsey L Royden and Patrick Fitzpatrick. *Real analysis*. Macmillan New York, 1988.

[14] Richard Sharp. MA427 Ergodic theory course notes.

[15] Elias M Stein. On the maximal ergodic theorem. *Proceedings of the National Academy of Sciences*, 47(12):1894–1897, 1961.

[16] Terence Tao. Ornstein's counterexample to an $L^1$ maximal inequality. `https://terrytao.wordpress.com/2015/05/18/failure-of-the-l1-pointwise-and-maximal-ergodic-theorems-for-the-free-group/`. Accessed: 25-01-2016.

[17] Terence Tao. Failure of the $L^1$ pointwise and maximal ergodic theorems for the free group. *arXiv preprint arXiv:1505.04725*, 2015.

[18] Peter Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer Science & Business Media, 2000.