Symbolic Dynamics on Free Groups

Steven Piantadosi
Advisor: Dr. Karl Petersen

Contents

Introduction

Symbolic dynamics is the study of bi-infinite sequences of symbols under simple transformations. This subfield of dynamical systems theory has many deep connections to theoretical computer science, ergodic theory, information theory and number theory. Ideas from symbolic dynamics are applied in numerous scientific and engineering areas, ranging from studying the properties of crystals to constructing optimal codes for transmitting information.

This thesis attempts to generalize a few of the basic ideas from one-dimensional and two-dimensional symbolic dynamics to "colorings" of groups. This endeavor is interesting because many natural questions arise that have counter-intuitive answers, and by working towards understanding these more general systems, we may shed light on important questions that have already come up in symbolic dynamics.

Most of this thesis will study shifts of finite type on the free group. In Chapter 3, Theorem 3.5, we will determine when a shift of finite type on the free group admits a strongly periodic point. This turns out to be a rather subtle issue that involves difficult combinatorial properties of the shift of finite type. Next, we study the golden mean shift on the free group and derive a formula for its entropy in Chapter 4, Theorem 3.4. In doing so, we develop a new generalization of Fibonacci numbers and study their asymptotics. Finally, in Chapter 5, we present preliminary results related to "nonlinear continued fractions" which arise in studying the asymptotics of these generalized Fibonacci numbers.

Acknowledgments

I'd like to thank several people for helping me in various stages of this project. First, I'd like to thank my readers, Jane Hawkins and Sue Goodman. Klaus Schmidt provided helpful discussion about $\mathbb{Z}^d$ dynamics and the connection between periodic colorings and finite index subgroups. Discussions with Doug Lind about generalizing ideas from one-dimensional dynamics were also a huge help in the early stages of this project. In addition, both helped find flaws in earlier, incorrect proofs about the existence of strongly periodic colorings on the free group.

I'd also like to thank the Office of Undergraduate Research for providing me with funding to work on this project over the summer of 2005. The University Center for International Studies also provided me with a travel grant to pursue this research at the Schrödinger Institute in Vienna. I am especially grateful to Klaus Schmidt and Karl Petersen inviting me to visit Vienna, and to Sarah Bailey for being an excellent travel companion.

Most importantly, I'd like to thank Karl Petersen for being an amazing mentor and advisor. Dr. Petersen has patiently and expertly explained to me the intricacies of symbolic dynamics, ergodic theory, measure theory, and information theory. He has been instrumental in aiding my understanding, constructing and clarifying proofs, providing guidance, and editing numerous drafts to make my writing readable. I am especially appreciative of the effort and enthusiasm he has put into this project.

CHAPTER 1

Background

This chapter will recall basic facts about group theory, symbolic dynamics in one dimension, and symbolic dynamics in $\mathbb{Z}^d$. The discussion that follows is not meant to be comprehensive; rather, it is intended to mention the basic ideas in each area which will be the most useful for understanding symbolic dynamics on the free group. The discussion of group theory is closely based on [**9**] and [**6**].

1. Group theory

A group $\mathcal{G} = (G, \cdot)$ is a set of elements G and a binary operation $\cdot$ on G with the following properties:

(1) There exists $e \in G$ such that $e \cdot a = a \cdot e = a$ for all $a \in G$.

(2) For all $a \in G$ there exists $a^{-1} \in G$ such that $a \cdot a^{-1} = a^{-1} \cdot a = e$.

(3) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all $a, b, c \in G$.

Typically we drop the $\cdot$ notation and write ab instead of $a \cdot b$. In addition, we may call the set G a "group" so long as there is no ambiguity about the binary operation.

Definition 1.1. Two groups $(G_1, \cdot)$ and $(G_2, *)$ are *isomorphic* if there exists a bijection $\varphi : G_1 \to G_2$ such that $\varphi(x \cdot y) = \varphi(x) * \varphi(y)$ for $x, y \in G_1$. The function φ is called an *isomorphism*.

Four our purposes, two groups G_1 and G_2 will be considered "the same" if they are isomorphic.

1.1. Group presentations. In a group $\mathcal{G} = (G, \cdot)$, the operation $\cdot$ can be defined in several ways. For finite groups, $\cdot$ can be defined by explicitly writing a "multiplication" table. This will not work for infinite groups, but we can define $\cdot$ with a presentation.

A *presentation* of a group $\mathcal{G}$ is an expression of the form

$$\langle\, \sigma \mid R \,\rangle$$

where σ is a set of generators and R is a set of equivalences that describe the relationships among the generators. That is, we require that every element $g \in G$ can be written as some product $g = \sigma_{a_1}\sigma_{a_2}\ldots\sigma_{a_n}$, where each $\sigma_{a_i} \in \sigma$. The elements of R are equations such as

$$\sigma_{a_1}\sigma_{a_2}\ldots\sigma_{a_n} = \sigma_{b_1}\sigma_{b_2}\ldots\sigma_{b_m},$$

and R must satisfy the following two criteria: every equivalence in R holds in $\mathcal{G}$, and if

$$g = \sigma_{a_1} \dots \sigma_{a_n} = \sigma_{b_1} \dots \sigma_{b_m},$$

then it can be "deduced" from the relations that $\sigma_{a_1} \dots \sigma_{a_n} = \sigma_{b_1} \dots \sigma_{b_m}$. We avoid an in-depth discussion of what "deduced" means, as examples later in the chapter should make it intuitively clear.

To simplify notation, we will assume that R does not include the multiplicative properties necessary for all groups (e.g. that $aa^{-1} = e$), that $e \notin \sigma$, and that if $\sigma_i \in \sigma$ then $\sigma_i^{-1} \notin \sigma$. Also, when we wish to explicitly write out the sets σ and R in $\langle\, \sigma \mid R \,\rangle$, we leave out "{" and "}" for notational simplicity.

Example 1.2. The group of integers under addition, $\mathbb{Z}$, can be presented by $\langle\, \sigma \mid \emptyset \,\rangle$ where $\sigma = \{\sigma_1\}$. We will also write this presentation as $\langle\, \sigma_1 \mid \,\rangle$. Even though the binary operation for this group is addition, we will write elements as $\sigma_1\sigma_1 = \sigma_1^2$ rather than $\sigma_1 + \sigma_1$. Therefore, each element is given by σ_1^n for some $n \in \mathbb{Z}$.

Example 1.3. The cyclic group of order n, C_n, can be presented by $\langle\, \sigma_1 \mid \sigma_1^n = e \,\rangle$.

Example 1.4. The dihedral group D_n, which gives the symmetries of a regular n-gon, can be presented by $\langle\, r, s \mid r^2 = e, s^n = e, (rs)^n = e \,\rangle$.

Every group can also be represented by a *Cayley graph.* A Cayley graph $C = (V, E)$ of a group G is a set V of vertices and a set E of edges such that each vertex represents a unique element of G. Vertices representing $g_1, g_2 \in G$ are joined by a directed edge labeled σ_i from g_1 to g_2 if and only if $g_1\sigma_1 = g_2$. Therefore, each vertex has exactly one incoming and outgoing edge for each σ_i. Notice that these edges will depend on the presentation of G.

Example 1.5. Figure 1 shows part of the Cayley graph for $\mathbb{Z}$ presented by $\langle\, \sigma_1 \mid \,\rangle$. Group elements (vertices) are represented by black circles.

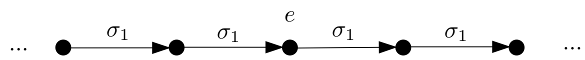

Figure 1. Part of the Cayley graph of $\mathbb{Z}$.

Example 1.6. Figure 2 shows part of the Cayley graph for $\mathbb{Z} \times \mathbb{Z}$ presented by $\langle\, \sigma_1, \sigma_2 \mid \sigma_1\sigma_2 = \sigma_2\sigma_1 \,\rangle$. In Figure 2, group elements (vertices) are represented by black circles. Assume solid arrows are labeled σ_1 and dotted lines are labeled σ_2.

1.2. Finitely generated groups. While there exists a presentation for every group, many, such as $\mathbb{R}\backslash\{0\}$ under multiplication, require the cardinality of σ or R to be infinite. Such groups can be difficult to analyze so here we restrict attention to finitely-generated groups:

Definition 1.7. A group G is *finitely generated* if there exists a presentation $\langle\, \sigma \mid R \,\rangle$ of G such that $|\sigma|$ and $|R|$ are finite.

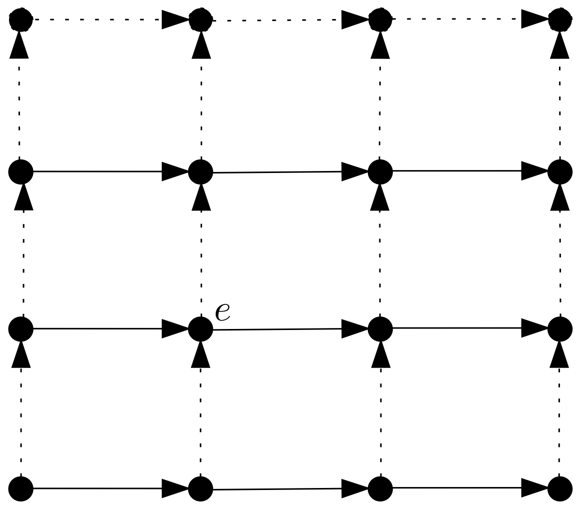

FIGURE 2. Part of the Cayley graph of $\mathbb{Z} \times \mathbb{Z}$.

In finitely-generated groups, we can define the "size", $|g|$, of a group element:

$$|g| = \min\{n : g = \sigma_{a_1}\sigma_{a_2}\ldots\sigma_{a_n}\}.$$

By convention, we say $|e| = 0$. Intuitively, $|g|$ corresponds to the number of edges one must travel on the Cayley graph of G to reach g from e. However, the value of $|g|$ may depend on the presentation of G.

1.3. The free group. The main objects of interest in this thesis will be free groups:

DEFINITION 1.8. A *free group* is any group G which can be presented by $\langle\, \sigma \mid \emptyset \,\rangle = \langle\, \sigma \mid \,\rangle$. If the cardinality of σ is q, we say that G is the free group *of order* q.

Note that our definition of the order of a group depends on our assumption that $\sigma_i \in \sigma$ implies $\sigma_i^{-1} \notin \sigma$. For example, $\mathbb{Z}$ is the finitely generated free group of order 1. In every free group, the set R of relations is empty, so each product $\sigma_{a_1}\sigma_{a_2}\ldots\sigma_{a_3}$ gives a unique group element. Part of the Cayley graph on two generators is shown in Figure 3.

1.4. Subgroups.

DEFINITION 1.9. A group $\mathcal{H} = (H, \cdot)$ is a *subgroup* of $\mathcal{G} = (G, \cdot)$ if $H \subseteq G$, $e \in H$, and if $x, y \in H$ then $xy \in H$ and $x^{-1} \in H$.

EXAMPLE 1.10. $\mathbb{Z}$ is a subgroup of $\mathbb{Z} \times \mathbb{Z}$. This should be clear from the fact that in $\langle\, \sigma_1, \sigma_2 \mid \sigma_1\sigma_2 = \sigma_2\sigma_1 \,\rangle$ we can consider elements of the form σ_1^n, which is simply the group $\mathbb{Z}$.

EXAMPLE 1.11. Suppose that $\mathcal{G} = C_{12}$, the cyclic group of order 12. Then C_2, C_3, C_4 and C_6 are subgroups of $\mathcal{G}$.

The *index* of a subgroup $H \subseteq G$ is roughly the number of different "copies" of H that there are in G. In $\mathbb{Z} \times \mathbb{Z}$ there are infinitely many "copies" of $\mathbb{Z}$

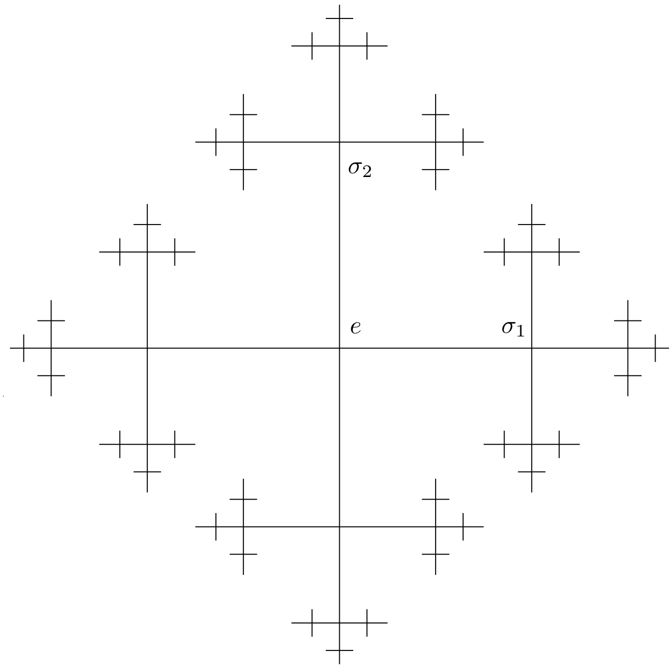

FIGURE 3. Part of the Cayley graph on two generators. Each intersection of lines represents a group element.

since we can generate a new copy of $\mathbb{Z}$ by considering the set $\{\sigma_2^k\sigma_1^n : n \in \mathbb{Z}\}$ for each k. Each choice of k produces a new set of group elements which is isomorphic to $\mathbb{Z}$, so the index of $\mathbb{Z}$ is infinite. This can be put more formally:

DEFINITION 1.12. The *index* of a subgroup $H \subseteq G$ is the cardinality of the set $\{aH : a \in G\}$. If the index of H is finite, H is called a *finite index subgroup*.

Thus, in a finite group, the index of the subgroup consisting of e is the cardinality of the group; in infinite groups, the index of the subgroup consisting of e is infinite. For each group G, G is a finite index subgroup of index 1. Here is a less trivial example:

EXAMPLE 1.13. Let G be presented by $\langle \sigma_1, \sigma_2 \mid \sigma_1\sigma_2 = \sigma_2\sigma_1, \sigma_2^k = e \rangle$ for some fixed k. In G, $\langle \sigma_1 \mid \ \rangle$ is a finite-index subgroup of index k and $\langle \sigma_2 \mid \sigma_2^k = e \rangle$ is subgroup of infinite index.

Given a subgroup H of G, each set (aH) where $a \in G$ is called a *coset* of H. The *group of left cosets* of H, denoted G/H, consists of all elements of the form (aH), where $a \in G$. The binary operation $\cdot$ defined on G/H is $(aH) \cdot (bH) = (abH)$. Similarly, the *group of right cosets* of H, denoted $G \setminus H$, consists of all elements of the form (Ha) where $a \in G$. The operation $\cdot$ on $G \setminus H$ is defined to be $(Ha) \cdot (Hb) = (Hab)$.

2. One-dimensional symbolic dynamics

2.1. Preliminaries. One-dimensional symbolic dynamics is the study of the behavior of bi-infinite sequences of symbols under simple transformations.

Typically, these symbols come from some finite alphabet $\mathcal{A}$, and an infinite sequence of symbols is called a *point.* A point $x \in \mathcal{A}^{\mathbb{Z}}$ may be written as

$$x = \ldots x_{-3}x_{-2}x_{-1}.x_0x_1x_2x_3\ldots$$

where each $x_i \in \mathcal{A}$. In this sequence, each x_i occupies a different "position," meaning that x_0 is the 0'th symbol, x_1 is the 1'st symbol, x_{-1} is the -1st symbol, and so on. This ordering is indicated by placing a decimal point to the left of the 0th symbol. For a point x, we may write $(x)_i$ to mean the symbol x_i.

A *block* is a finite sequence of symbols $b = b_1b_2\ldots b_n$, and we say that the block b has *length* $|b| = n$. A point x *contains* a block b if there exists some $k \in \mathbb{Z}$ such that $x_{k+j} = b_j$ for all $j \leq |b|$.

DEFINITION 2.1. Given any (infinite or finite) set of forbidden blocks F, define X_F to be the set of all points which do not contain any blocks in F. Such a set X_F is called a *shift space.*

EXAMPLE 2.2. $\mathsf{X}_\emptyset$ is called the *full shift* on the alphabet $\mathcal{A}$ and consists of all possible sequences in $\mathcal{A}$. Every shift space is a subset of the full shift.

EXAMPLE 2.3. Let $\mathcal{A} = \{0,1\}$ and $F = \{11, 00\}$. Then X_F consists of two points:

$$\ldots 0101.0101\ldots$$
$$\text{and}$$
$$\ldots 1010.1010\ldots.$$

EXAMPLE 2.4. Let $\mathcal{A} = \{0,1\}$ and

$$F = \{\underbrace{1\ldots1}_{n\text{ 1s}} : n \text{ is even }\}.$$

Then X_F consists of all bi-infinite sequences for which the maximal string of consecutive 1s always has odd length. For example,

$$\ldots 01110100.0111110\ldots \in \mathsf{X}_F$$

but

$$\ldots 0100101.101110\ldots \notin \mathsf{X}_F.$$

2.2. The shift transformation. Shift spaces are *shift invariant* sets, which informally means that given a point $x \in \mathsf{X}_F$, we can construct a new point in X_F by "shifting" the location of the decimal point in x. This is expressed formally by a *shift transformation* $\sigma : \mathsf{X}_F \to \mathsf{X}_F$, which acts by shifting the sequence one position to the left (or the decimal point one place to the right):

$$\begin{aligned}\sigma(x) &= \sigma(\ldots x_{-3}x_{-2}x_{-1}.x_0x_1x_2x_3\ldots)\\ &= \ldots x_{-3}x_{-2}x_{-1}x_0.x_1x_2x_3\ldots\end{aligned}$$

It should be obvious that $\sigma(\mathsf{X}_F) = \mathsf{X}_F$ and that σ^{-1} is also well-defined. In addition, we can assign a metric ρ to points in a shift space by setting $\rho(x,y) = \sup\{1/2^{|i|} : i \in \mathbb{Z}, x_i \neq y_i\}$. Under ρ, σ is continuous, and a shift space X_F is a *closed* set, meaning that X_F contains its limit points under ρ. In fact, we have the following theorem, proved in [**10**]:

THEOREM 2.5. *A set $X \subset \mathcal{A}^{\mathbb{Z}}$ is a shift space if and only if X is closed and shift-invariant.*

The shift transformation is important because its action on shift spaces creates a dynamical system. The formalism of one-dimensional symbolic dynamics describes dynamical systems that have a discrete finite space and discrete time. For example, suppose a system S can be in some finite set of states $\mathcal{A}$, and that transitions between states can be restricted by some set of blocks F. Then each point $x \in \mathsf{X}_F$ gives an entire history of the system. If we regard the current state of the system as x_0, then σ moves time forward one step since $(\sigma(x))_0 = x_1$. Many dynamical systems can be modeled exactly or almost exactly by symbolic dynamical systems, which, in general, are much easier to handle.

The shift transformation also defines one of the basic properties of points in a dynamical system, *periodicity*:

DEFINITION 2.6. A point $x \in \mathsf{X}_F$ is *periodic* if there exists $n \in \mathbb{Z}$ such that $\sigma^n(x) = x$.

Equivalently we could say that a point $x \in \mathsf{X}_F$ is periodic if the *orbit* $\{\sigma^n(x) : n \in \mathbb{Z}\}$ is finite.

EXAMPLE 2.7. The point $x = \ldots 010101.010101 \ldots$ is periodic since $\sigma^2(x) = x$. The point $x = \ldots 000000.100000 \ldots$ is not periodic.

2.3. Entropy and the golden mean shift. The set $\mathcal{L}(\mathsf{X}_F)$, also called the *language of* X_F, is defined to be the set of all finite blocks which appear in points in X_F. The language of X_F is important because it provides another way to characterize points in X_F and a connection to the "languages" studied in theory of computation. One of the most important characteristics of a shift space,*entropy*, is roughly the growth rate of the number of blocks in $\mathcal{L}(\mathsf{X}_F)$. Define

$$B_n(\mathsf{X}_F) = \{l \in \mathcal{L}(\mathsf{X}_F) : 0 < |l| \leq n\}.$$

Then we have:

DEFINITION 2.8. The *entropy* $h(\mathsf{X}_F)$ of a shift space X_F is defined to be

$$h(\mathsf{X}_F) = \lim_{n\to\infty} \frac{\log_2 |B_n(\mathsf{X}_F)|}{n}.$$

EXAMPLE 2.9. Let $\mathcal{A} = \{0, 1\}$ and $F = \{11\}$. X_F is called the *golden mean shift* because the entropy of X_F is equal to $\log \phi$, where ϕ is the golden mean (this example is taken from [**10**]).

To see this, note $B_1 = \{0, 1\}$ and $B_2 = \{00, 10, 01\}$. We can partition B_n into two sets, Z_n and W_n, consisting of blocks that end in 0 and 1 respectively. To construct W_{n+1}, we can put 1 on the end of each block in Z_n. To construct a block in Z_{n+1}, we can put 0 on the end of each block in Z_n or W_n. Thus,

$$|W_n| = |Z_{n-1}|$$
$$|Z_n| = |Z_{n-1}| + |W_{n-1}| = |Z_{n-1}| + |Z_{n-2}|.$$

This shows that the $|Z_n|$ are Fibonacci numbers F_i with $|Z_1| = 1 = F_2$, $|Z_2| = 2 = F_3$, and in general $|Z_i| = F_{i+1}$. Thus,

$$\begin{aligned} |B_n| &= |Z_n| + |W_n| \\ &= |Z_n| + |Z_{n-1}| \\ &= F_{n+1} + F_n \\ &= F_{n+2} \end{aligned}$$

Now, it is known by Binet's formula that

$$F_n = \frac{1}{\sqrt{5}}(\phi^n - \mu^n)$$

where

$$\phi = \frac{1+\sqrt{5}}{2} \quad \text{and} \quad \mu = \frac{1-\sqrt{5}}{2}.$$

We can find the entropy of X_F by

$$\begin{aligned} h(\mathsf{X}_F) &= \lim_{n\to\infty} \frac{\log |B_n(\mathsf{X}_F)|}{n} \\ &= \lim_{n\to\infty} \frac{\log F_{n+2}}{n} \\ &= \lim_{n\to\infty} \frac{\log \left[\frac{1}{\sqrt{5}}(\phi^{n+2} - \mu^{n+2})\right]}{n} \\ &= \lim_{n\to\infty} \frac{1}{n}\left[\log \frac{1}{\sqrt{5}} + (n+2)\log\phi + \log(1 - \frac{\mu^{n+2}}{\phi^{n+2}})\right] \\ &= \log\phi, \end{aligned}$$

since $\mu^n/\phi^n \to 0$.

The limit defining entropy always exists by subadditivity, and for certain classes of shift spaces it is easy to find. One such class of shift spaces is called the *shifts of finite type*:

DEFINITION 2.10. A *shift of finite type* is a shift space X_F for which the cardinality of F is finite.

Most of this thesis will be devoted to studying shifts of finite type because, while in one-dimension they are relatively easy to understand, on $\mathbb{Z} \times \mathbb{Z}$ their behavior is much more complex.

2.4. Graph representations of shifts of finite type. In general, we can represent any one-dimensional shift of finite type X_F with a directed graph Γ, in which either the edges or the vertices are labeled with symbols. In either case, Γ has a finite number of vertices and infinite walks on Γ correspond one-to-one with points in the shift space. A detailed discussion of such graphs can be found in [**10**], but for our purposes it will suffice to discuss an example and sketch the proof of a simple theorem about shifts of finite type.

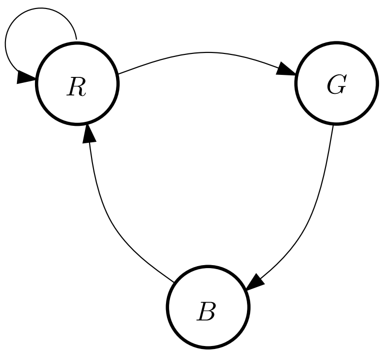

FIGURE 4. A graph representation of a one-dimensional shift of finite type.

EXAMPLE 2.11. Let $\mathcal{A} = \{R, G, B\}$ and $F = \{GR, GG, BG, BB, RB\}$. Figure 4 shows a graph representing X_F, where vertices correspond to symbols in $\mathcal{A}$. This graph shows that, for example, $BB \in F$ since there is no directed edge from B to B.

THEOREM 2.12. *A nonempty shift of finite type* X_F *always contains a periodic point.*

PROOF. If X_F is nonempty, then we can take an infinite walk on the graph Γ representing X_F. But Γ has a finite number of vertices, so we must visit some vertex v twice. That means that there exists a path from v back to v. Thus, we can construct a new walk on Γ which repeatedly goes from v to v. This new path will correspond to a periodic point in X_F. □

3. $\mathbb{Z}^d$ symbolic dynamics

Ideas from one-dimensional symbolic dynamics have also been generalized to multi-dimensional arrays of symbols. For example, in $\mathbb{Z} \times \mathbb{Z}$ a point is a two-dimensional array of symbols such as

$$\begin{array}{ccccc} & & \vdots & & \\ \cdots & a_{-1,1} & a_{0,1} & a_{1,1} & \cdots \\ \cdots & a_{-1,0} & a_{0,0} & a_{1,0} & \cdots \\ \cdots & a_{-1,-1} & a_{0,-1} & a_{1,-1} & \cdots \\ & & \vdots & & \end{array}$$

A point can be viewed as an assignment of a color in $\mathcal{A}$ to each element of $\mathbb{Z} \times \mathbb{Z}$. We will also call points *colorings*:

DEFINITION 3.1. Given a set or group S, a function $f : S \to \mathcal{A}$ is called a *coloring* of S.

In $\mathbb{Z}^d$, blocks consist of d-dimensional hypercubes, and shift spaces and shifts of finite type are defined analogously to the one-dimensional case. In $\mathbb{Z}^d$, there are d shift transformations $\sigma_1, \sigma_2, \ldots, \sigma_d$, where σ_i shifts a coloring along the i'th coordinate axis.

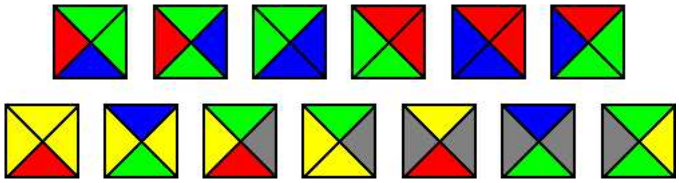

FIGURE 5. A set of 13 Wang tiles which will tile the plane only aperiodically.

The symbol at position $R \subset \mathbb{Z}^d$ in a point x is notated by x_R. With this notation, we can define the shift transformations by

$$(\sigma_i(x))_{(x_1,x_2,\ldots,x_d)} = x_{(x_1,\ldots,x_{i-1},x_i+1,x_{i+1},\ldots,x_d)}.$$

In addition, entropy is defined analogously to the one-dimensional case. If C_n is an d-dimensional hypercube of side length n, and B_n is the number of colorings of C_n that are allowed in X_F, then we define

$$h(\mathsf{X}_F) = \lim_{n\to\infty} \frac{\log B_n}{n^d}.$$

Entropy turns out to be difficult to compute for many simple $\mathbb{Z} \times \mathbb{Z}$ shifts of finite type; later we will return to discuss the golden mean shift on $\mathbb{Z} \times \mathbb{Z}$.

3.1. Wang tiles. $\mathbb{Z} \times \mathbb{Z}$ dynamics are much more difficult to analyze than $\mathbb{Z}$ because even simple systems such as shifts of finite type can have extraordinarily complicated behavior. For example, if a one-dimensional shift of finite type is nonempty, it must contain a periodic point. In two dimensions this is not true: there exist square tiles, called *Wang tiles*, which will only allow an aperiodic tiling of the plane. An example of Wang tiles is shown in Figure 5 [**1**]. One must place these in the plane without rotating them, so that colored edges match up and no gaps are left. The tiling of the plane by Wang tiles is "equivalent" to a two-dimensional shift of finite type, since we can give each Wang tile a unique symbol and use the colored edges to construct a set of forbidden blocks F.

It can be proved that the set of tiles in Figure 5 can tile the infinite plane, but that they will never produce a periodic coloring. This means that for all $T \in \mathbb{Z} \times \mathbb{Z}$ the orbit $\{T^n f : n \in \mathbb{Z}\}$ is infinite for all allowed colorings f by Wang tiles. This is perhaps surprising because the tiles are restricted only locally (by which can be adjacent), but they give rise to the global property that they never allow a periodic tiling. Interestingly, the existence of such shift spaces is closely tied to the fact that in two dimensions, determining whether a shift of finite type contains any points at all is formally undecidable.

3.2. The two-dimensional golden mean shift. $\mathbb{Z}^d$ dynamical systems can describe many complicated physical systems, allowing, for example, questions in statistical physics to be rephrased in $\mathbb{Z}^d$ dynamical terms. Perhaps the most well-known problem in this area is to determine the *hard square entropy constant*, which equals the entropy of a two-dimensional version of the golden mean shift. Consider $\mathbb{Z} \times \mathbb{Z}$ with the standard presentation

$$\langle\, \sigma_1, \sigma_2, \mid \sigma_1\sigma_2 = \sigma_2\sigma_1 \,\rangle.$$

Define X_F to be the set of colorings of $\mathbb{Z} \times \mathbb{Z}$ by $\mathcal{A} = \{0, 1\}$ such that no two 1's are adjacent horizontally (by σ_1) or vertically (by σ_2). That is, the set F of forbidden blocks consists of the following two blocks:

$$11 \qquad \begin{matrix} 1 \\ 1 \end{matrix}$$

If rooks could only attack adjacent squares, B_n would count the number of ways to place rooks on an $n \times n$ chessboard such that no two were attacking each other. The value of $h(\mathsf{X}_F)$ is important for understanding several physical systems and for many coding problems that arise in the study of two-dimensional run-length constrained channels [**14, 8, 12**]. Much effort has been put into simply achieving bounds for $h(\mathsf{X}_F)$, and $h(\mathsf{X}_F)$ has been computed to several decimal places [**4**]:

$$h(\mathsf{X}_F) = \log 1.5030480824\ldots$$

Aside from bounds and numerical computations, very little is known about this number, and even determining whether $h(\mathsf{X}_F)$ is algebraic is still an open problem.

CHAPTER 2

Symbolic dynamics on groups

The goal of this chapter is to generalize some of the basic ideas from one-dimensional dynamics to colorings of groups. This is motivated by the fact that we can view points in one-dimensional symbolic dynamics as assignments of symbols to each element of the group $\mathbb{Z}$. Similarly, points in $\mathbb{Z}^d$ are assignments of symbols to each element of $\mathbb{Z}^d$. In a sense, generalizing one-dimensional ideas to arbitrary groups is doomed from the start, since quantities such as entropy may no longer always exist by subadditivity. However, the next chapter will study the golden mean shift on the free group and show that the behavior of many systems can still be rich and interesting.

In this chapter, we will redefine much of the terminology from one-dimensional and $\mathbb{Z}^d$ dynamical systems. The definitions will be equivalent to the earlier definitions for the cases when the group under consideration is $\mathbb{Z}$ or $\mathbb{Z}^d$.

1. Basic definitions

1.1. The shift transformations. Let G be a finitely-generated group presented by $\langle\, \sigma \mid R \,\rangle$. Suppose G has generators $\sigma = \{\sigma_1, \ldots \sigma_n\}$, and let $\mathcal{A}$ be a finite alphabet of symbols. Recall a *coloring* of G is a function $f : G \to \mathcal{A}$. One can regard colorings as assigning to each vertex of the Cayley graph of G a symbol from $\mathcal{A}$.

Colorings will be considered points in the dynamical system $(\mathcal{A}^G, G)$ where the action is given by

$$gf(w) = f(gw)$$

for $g, w \in G$ and $f : G \to \mathcal{A}$. Since the σ_i are generators of G, we need only specify that $\sigma_i f(w) = f(\sigma_i w)$ for each $\sigma_i \in \sigma$. For simplicity, we will generally assume that groups are presented by a standard form: unless otherwise stated, $\mathbb{Z}^d$ will be presented by

$$\langle\, \sigma_1, \ldots, \sigma_d \mid \sigma_i \sigma_j = \sigma_j \sigma_i \text{ for all } i, j \leq d \,\rangle,$$

and the free group on $k+1$ generators will be presented by

$$\langle\, \sigma_1, \ldots, \sigma_{k+1} \mid \,\rangle.$$

However, it will occasionally be convenient to use a different presentation of a group.

1.2. Shift spaces. A *block* is a function $b : H \to \mathcal{A}$, where H is a finite subset of G. We make no stipulations that B is "connected" or of a certain shape. A coloring f *contains* a block $b : H \to \mathcal{A}$ if there exists $g \in G$ such that $f(gh) = b(h)$ for all $h \in H$. In other words, a coloring contains a block if the coloring can be shifted to agree with the block on all of H. We are now ready to define shift spaces:

DEFINITION 1.1. Given a group G and a set of forbidden blocks F, define X_F to be the set of all colorings of G that do not contain any block in F. X_F will be called a *shift space*.

In the case when $G = \mathbb{Z}^d$, we recover the previous definition of a $\mathbb{Z}^d$ symbolic dynamical system. Again, when the cardinality of F is finite, X_F is a called a *shift of finite type*.

If we define a metric $\rho : \mathsf{X}_F \to \mathbb{R}$ by

$$\rho(f_1, f_2) = \sup\{1/2^{|g|} : g \in G, f_1(g) \neq f_2(g)\}$$

then each shift transformation σ_i is continuous under ρ. Note that ρ is analogous to the one-dimensional case presented in Chapter 1.

EXAMPLE 1.2. On any group G, $X_\emptyset$ is a shift space consisting of all possible colorings of G by $\mathcal{A}$. This is called the *full shift on G*.

EXAMPLE 1.3. For each σ_i, define the block $b_{\sigma_i} : \{e, \sigma_i\} \to \{0, 1\}$ by $b_{\sigma_i}(e) = 1$ and $b_{\sigma_i}(\sigma_i) = 1$. If $F = \{b_{\sigma_1}, \ldots, b_{\sigma_n}\}$ then we call X_F the *golden mean shift on G*. When $G = \mathbb{Z}$ or $G = \mathbb{Z} \times \mathbb{Z}$, golden mean shift takes on its usual definition.

1.3. Entropy. We can also define entropy for shift spaces on groups. Define a "ball of radius n" by

$$C_n = \{g \in G : |g| \leq n\}.$$

DEFINITION 1.4. For $n = 0, 1, 2, \ldots$, let B_n be the number of colorings of C_n that appear in X_F. Then we define the *entropy* of X_F to be

$$h(\mathsf{X}_F) = \limsup_{n \to \infty} \frac{\log B_n}{|C_n|}.$$

Note that $h(\mathsf{X}_F)$ will always be defined; however, the lim sup may not necessarily equal the lim inf. On $\mathbb{Z}^d$, however, the lim sup equals the lim inf by a subadditivity argument [**13**]. The definition of entropy used here appears to be different from the definition of $\mathbb{Z}^d$ entropy given earlier because this definition considers "balls" of radius n and the earlier definition considered "cubes" of side length n. However, it turns out that the two give the same numerical value on $\mathbb{Z}^d$; in fact, one can consider any set of shapes that form a Følner sequence [**11**]. A *Følner sequence* is any sequence of subsets $F_1, F_2, \ldots$ of G for which $\lim_{n\to\infty} |gF_n \bigtriangleup F_n|/|F_n| = 0$ for any $g \in G$, where $\bigtriangleup$ is the symmetric difference operator. Informally, this says that the action of any group element does not move "move" the F_n too far. Følner sequences are also related to amenability in that a countable group is amenable if and only if it has a Følner sequence. In $\mathbb{Z}^d$ it is easy to check that both "balls" and "cubes" form Følner sequences and therefore give equal values for the entropy [**11**].

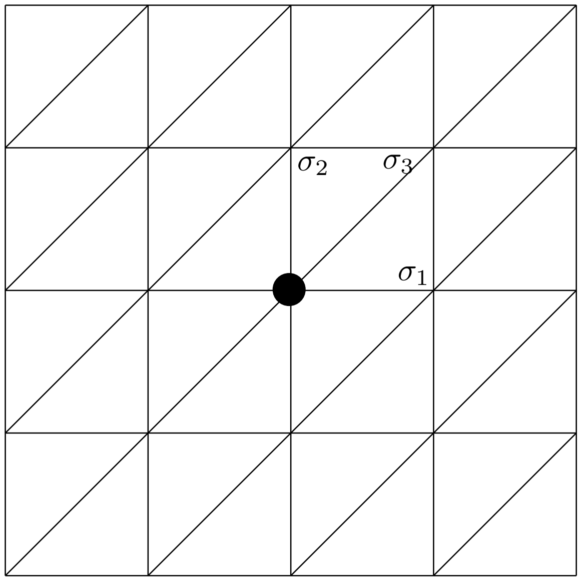

FIGURE 1. Part of the Cayley graph of $\mathbb{Z} \times \mathbb{Z}$ with a different presenation. Note $\sigma_1\sigma_2 = \sigma_2\sigma_1 = \sigma_3$.

1.4. Summary. Let us summarize what has been defined: given a group G, an alphabet $\mathcal{A}$, and a set F of forbidden blocks (functions from finite subsets of G to $\mathcal{A}$), we define the shift space X_F to be the collection of colorings which do not use any blocks in F. On this set, we define a metric ρ, such that f_1 and f_2 are "close" if they agree on a large block centered at $e \in G$. Under this metric, every shift action $T \in G$ is continuous under ρ. We also defined the entropy of X_F. All of these definitions were constructed so that they are equivalent to the one-dimensional definitions when $G = \mathbb{Z}$. Next, we use this framework to present a standard example from statistical physics. In the last section of this chapter, we will consider how to generalize the definition of periodicity.

2. The hard hexagon constant

We have already discussed the hard square entropy constant, which is the entropy of a two-dimensional version of the golden mean shift. A related constant is the *hard hexagon entropy constant*, which equals the entropy of the golden mean shift on a hexagonal grid. Consider $\mathbb{Z} \times \mathbb{Z}$ with the presentation

$$\langle\, \sigma_1, \sigma_2, \sigma_3 \mid \sigma_1\sigma_2 = \sigma_2\sigma_1, \sigma_1\sigma_2 = \sigma_3 \,\rangle,$$

and suppose that X_F consists of all colorings of $\mathbb{Z} \times \mathbb{Z}$ with this presentation such that no two adjacent elements are colored 1. Part of the Cayley graph of $\mathbb{Z} \times \mathbb{Z}$ with this presentation is shown in Figure 1. Note that this is the dual graph of a tiling by hexagons, which shows that the golden mean shift on this presentation is the same as the golden mean shift on a hexagonal grid. We can also rewrite these adjacency rules on the standard presentation of $\mathbb{Z} \times \mathbb{Z}$.

If we choose F to consist of the three blocks

$$\begin{array}{cc} 1 & 1 \end{array} \qquad \begin{array}{c} 1 \\ 1 \end{array} \qquad \begin{array}{cc} 1 & 0 \\ & 1 \end{array}$$

then the entropy $h(\mathsf{X}_F)$ is called the hard hexagon entropy constant, and it has been computed exactly [**7, 2**]. It is extraordinary that $\exp(h(\mathsf{X}_F))$ is algebraic and is given by

$$\exp(h(\mathsf{X}_F)) = \kappa_1\kappa_2\kappa_3\kappa_4 \approx 1.395485972\ldots,$$

where

$$\begin{aligned}
\kappa_1 &= 4^{-1}3^{5/4}11^{-5/12}a^{-2} \\
\kappa_2 &= \left[1 - \sqrt{1-a} + \sqrt{2+a+2\sqrt{1+a+a^2}}\right]^2 \\
\kappa_3 &= \left[-1 - \sqrt{1-a} + \sqrt{2+a+2\sqrt{1+a+a^2}}\right]^2 \\
\kappa_4 &= \left[\sqrt{1-b} + \sqrt{2+b+2\sqrt{1+b+b^2}}\right]^{-1/2},
\end{aligned}$$

and

$$\begin{aligned}
a &= \left[\frac{1}{4} + \frac{3}{8}b((c+1)^{1/3} - (c-1)^{1/3})\right]^{1/3} \\
b &= -\frac{124}{636}11^{1/3} \\
c &= \frac{2501}{11979}33^{1/2}.
\end{aligned}$$

The value of $h(\mathsf{X}_F)$ is important for determining the thermodynamic properties of a model in statistical physics that describes phase transitions of systems of rigid molecules. The above results were calculated by [**7, 2**] by using the Klein-Fricke theory of modular functions.

3. Periodic colorings

For the group $G = \mathbb{Z} = \langle\, \sigma_1 \mid \,\rangle$ that is used in one-dimensional symbolic dynamics, the definition of periodicity is straightforward: a coloring f is periodic if the orbit $\{\sigma_1^n f : n \in \mathbb{Z}\}$ is finite. However, the best way to generalize this to arbitrary groups is not clear. One possibility is to say that a point is periodic if there exists a group element whose orbit is periodic:

Definition 3.1 (Weakly Periodic). A coloring f of a group G is *weakly periodic* if there exists $T \in G$ with $T \neq e$ such that the orbit $\{T^n f : n \in \mathbb{Z}\}$ is finite.

Example 3.2. Suppose $G = \mathbb{Z} \times \mathbb{Z}$ and $\mathcal{A} = \{R, Y\}$. Let

$$f(\sigma_1^a\sigma_2^b) = \begin{cases} R \text{ if } b = 0 \\ Y \text{ otherwise} \end{cases}$$

Then the orbit $\sigma_1^2 f = f$ so f is weakly periodic.

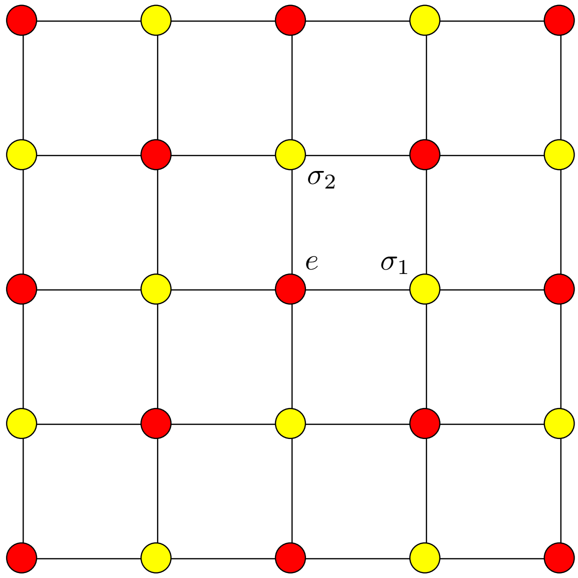

FIGURE 2. Part of the checkerboard coloring of $\mathbb{Z} \times \mathbb{Z}$.

Depending on the characteristics of G, a coloring f might be weakly periodic in a trivial sense: suppose that there is $\sigma_i \in \sigma$ such that $\sigma_i^r = e$ for some $r \in \mathbb{Z}$. Then clearly the orbit $\sigma_i^n f$ will be finite:

EXAMPLE 3.3. Fix $r \in \mathbb{N}$. Suppose that G is the group presented by $\langle \sigma_1, \sigma_2 \mid \sigma_1\sigma_2 = \sigma_2\sigma_1, \sigma_2^r = e \rangle$. Geometrically one can picture the Cayley graph of G as an infinitely long piece of graph paper of finite width that has been curled into an infinitely long cylinder. Every coloring f will be weakly periodic, since $\sigma_2^r f = f$.

One might also define "periodic" by considering the full orbit of a coloring under the entire group:

DEFINITION 3.4 (Strongly Periodic). A coloring f is *strongly periodic* if the orbit $\{Tf : T \in G\}$ is finite.

EXAMPLE 3.5. In Example 3.2 the coloring f is not strongly periodic since the orbit $\{\sigma_2^n f : n \in \mathbb{Z}\}$ is not finite.

EXAMPLE 3.6. Suppose $G = \mathbb{Z} \times \mathbb{Z}$ is presented by $\langle \sigma_1, \sigma_2 \mid \sigma_1\sigma_2 = \sigma_2\sigma_1 \rangle$ and $\mathcal{A} = \{R, Y\}$. Let

$$f(\sigma_1^a \sigma_2^b) = \begin{cases} R \text{ if } a \equiv b \mod 2 \\ Y \text{ otherwise .} \end{cases}$$

f might be called a *checkerboard coloring* of $\mathbb{Z} \times \mathbb{Z}$, part of which is shown in Figure 2. We have

$$\sigma_1 f = \sigma_2 f,$$

and

$$\sigma_1\sigma_2 f = \sigma_2\sigma_1 f = f.$$

Thus given any $T \in G$, $Tf = f$ or $Tf = \sigma_1 f$. In both cases, $T^2 f = f$, so f is strongly periodic.

Note that every strongly periodic point is also weakly periodic. We will also be concerned with aperiodic colorings:

Definition 3.7. A coloring f is *aperiodic* if it is not weakly periodic.

Example 3.8. Let $G = \mathbb{Z}$, $\mathcal{A} = \{R, Y\}$ and

$$f(\sigma_1^r) = \begin{cases} R \text{ if } r = 0 \\ Y \text{ otherwise.} \end{cases}$$

Then f is the one-dimensional point

$$\ldots YYYYYYR.YYYYYY\ldots,$$

and f is aperiodic.

4. An algebraic perspective on periodic colorings

Periodic colorings are defined in dynamical terms, but they can also be studied group-theoretically. Let

$$H = \{T \in G \mid Tf = f\}$$

be the *stabilizer* of a coloring f. It is easy to check that H is a normal subgroup of G. We can characterize the periodic properties of f by considering the group-theoretic properties of H:

Theorem 4.1. *H characterizes f in the following way:*

(*i*) *$H = \{e\}$ if and only if f is aperiodic (not weakly periodic).*

(*ii*) *H is of finite index if and only if f is strongly periodic.*

Proof. Note that f is weakly periodic if and only if there exists $T \in G$ such that $T \neq e$ and $Tf = f$. This implies (*i*).

For (*ii*), note that G can be written as a disjoint union of cosets $a_i H$ as

$$G = \bigcup_{a_i \in I} a_i H,$$

for some set I, where a_i is a representative of the i'th coset. This implies

$$Gf = \bigcup_{a_i \in I} a_i H f = \{a_i f : a_i \in I\}, \tag{1}$$

since H is the stabilizer of f. If H is finite index then I is finite so by (1) Gf must be finite. This shows f must be strongly periodic if H is finite index. Similarly, if H were not finite index, I would not be finite. Then (1) would show that Gf is not finite: $a_i f \neq a_j f$ for $i \neq j$, since the a_i are representatives of distinct cosets. This proves (*ii*). ☐

Therefore, finding periodic colorings turns out to be equivalent to finding finite index subgroups.

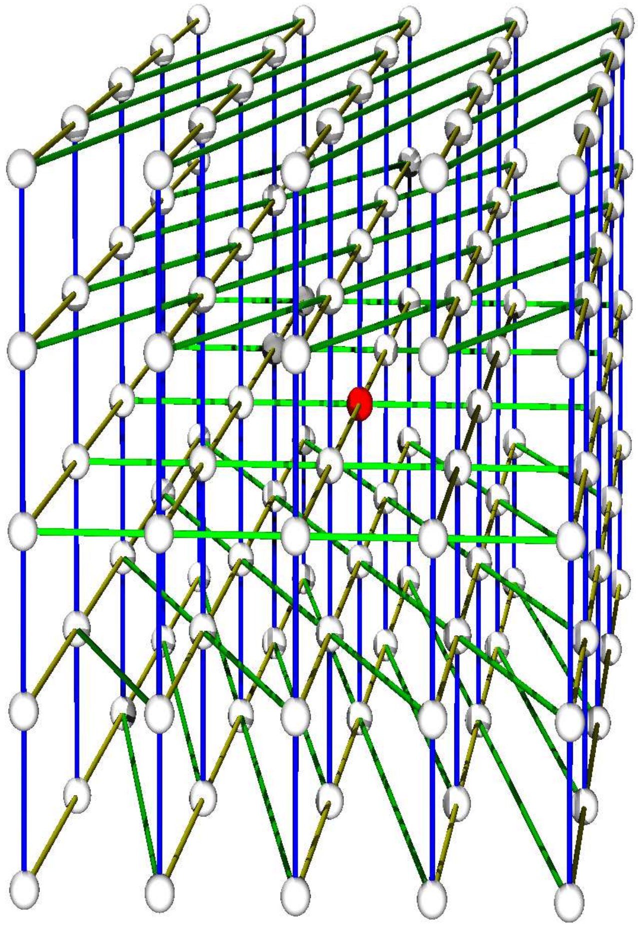

FIGURE 3. Part of the Cayley graph of $\mathbb{H}$. The red sphere corresponds to e.

5. An example: the Heisenberg group

The *discrete Heisenberg group*, $\mathbb{H}$, is a group of upper-triangular, three-by-three matrices. Elements of $\mathbb{H}$ are of the form

$$\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix},$$

where $a, b, c \in \mathbb{Z}$. $\mathbb{H}$ can be generated by elements x and y, where

$$x = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

With these two generators, $\mathbb{H}$ is presented by

$$\langle\, x, y \mid xxyx^{-1}y^{-1} = xyx^{-1}y^{-1}x, yxyx^{-1}y^{-1} = xyx^{-1}y^{-1}y \,\rangle.$$

We can simplify this notation by defining

$$z = xyx^{-1}y^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then $\mathbb{H}$ can be presented by

$$\langle\, x, y, z \mid xz = zx, yz = zy, z = xyx^{-1}y^{-1} \,\rangle. \tag{2}$$

The Cayley graph of $\mathbb{H}$ is shown in Figure 3.

In Figure 3, group elements are represented by white spheres. Generators x and y are represented by green and blue bars respectively, and z is represented by yellow bars. Note that $\mathbb{H}$ is similar to $\mathbb{Z}^3$, except that $\mathbb{H}$ has a "twist."

The presentation (2) makes it clear that $\mathbb{H}$ has $J = \mathbb{Z} \times \mathbb{Z}$ generated by $\langle\, x, z \mid xz = zx \,\rangle$ as a subgroup. Therefore, because there exist shifts of finite type on $\mathbb{Z} \times \mathbb{Z}$ that only admit aperiodic colorings, we can construct a shift of finite type rule on $\mathbb{H}$ such that J must be colored aperiodically (since J is isomorphic to $\mathbb{Z} \times \mathbb{Z}$). If we have some set of rules F that force J to be colored aperiodically but do not place any restrictions in the y direction, then X_F cannot contain a strongly periodic coloring, since the orbit $\{x^n f : n \in \mathbb{Z}\}$ will be finite for any $f \in \mathsf{X}_F$. Thus, we can construct a shift of finite type on the Heisenberg group that does not admit a strongly periodic point.

It should also be clear that if $f \in \mathsf{X}_F$ and $H = \{g \in \mathbb{H} : gf = f\}$ is the stabilizer of f, then $\mathbb{H}/H$ is isomorphic to $\mathbb{Z}$ if f is aperiodic and isomorphic to $\mathbb{Z} \mod r$ for some r if f is weakly periodic. It appears to be much more difficult to find a shift of finite type on $\mathbb{H}$ that does not allow a weakly periodic coloring, if such a shift of finite type even exists.

On $\mathbb{Z}^3$, there exist aperiodic sets of "Wang cubes" that will only tile $\mathbb{Z}^3$ aperiodically. An example of these can be found in [**5**], which constructs Wang cubes in $\mathbb{Z}^3$ by first constructing rules that color each $x-y$ plane aperiodically, using the regular Wang tile rules in the $x - y$ plane. The authors prove that such colorings will contain arbitrarily long stretches of a certain color in the x direction, and then cleverly use that property to restrict colorings in the z direction so that no coloring can be weakly periodic. The proof seems to hold potential for constructing an aperiodic shift of finite type on $\mathbb{H}$, since like $\mathbb{Z}^3$, $\mathbb{H}$ has $\mathbb{Z} \times \mathbb{Z}$ as a subgroup. However, direct application of the methods in [**5**] to our setting does not appear feasible.

CHAPTER 3

Shifts of finite type on the free group

1. Elementary properties

This section will study the periodic properties of shifts of finite type on the free group. We will study only *nearest-neighbor* shifts of finite type, which are shifts X_F for which each block in the set F of forbidden blocks consists of only two adjacent group elements; that is, F only restricts which elements can be adjacent by each $\sigma_i \in \sigma$. The restriction to nearest-neighbor shifts of finite type is mainly for notational simplicity, but it should be clear that given an arbitrary shift of finite type on the free group, we can "recode" it to a nearest-neighbor shift of finite type using a sliding block code scheme analogous to the one-dimensional case.

The following operation on one-dimensional shift spaces will be useful in this chapter:

Definition 1.1. Define $\mathsf{X}_F \mid_L = \mathsf{X}_{F \cup (\mathcal{A} \setminus L)}$ so that $\mathsf{X}_F \mid_L$ restricts the shift space X_F to the smaller alphabet $L \subset \mathcal{A}$.

Example 1.2. Let $\mathcal{A} = \{0, 1, 2\}$ and $F = \{11\}$. Then $\mathsf{X}_F \mid_{\{0,1\}}$ is the golden mean shift, $\mathsf{X}_F \mid_{\{0,2\}}$ is the full shift on $\{0, 2\}$, and $\mathsf{X}_F \mid_{\{1\}}$ is empty.

1.1. One-dimensional shifts of finite type in the free group. Since we are studying only nearest-neighbor shifts of finite type, we will assume that $F \subset \mathcal{A} \times \sigma \times \mathcal{A}$. Then, if $(a, \sigma_i, b) \in F$ with $a, b \in A$, the shift space X_F has the restriction that for any coloring $f \in \mathsf{X}_F$, $f(g) = a$ implies $f(g\sigma_i) \neq b$.

The free group is relatively easy to analyze because we can regard F as defining a one-dimensional shift of finite type for each of the generators. For example, we can consider only the shift transformation σ_i and only the adjacency rules in $\mathcal{A} \times \{\sigma_i\} \times \mathcal{A} \subset F$ to define a one-dimensional shift of finite type. This shift of finite type will be denoted X_{σ_i}.

Example 1.3. Let $\mathcal{A} = \{0, 1\}$, $\sigma = \{\sigma_1, \sigma_2\}$, and $F = \{\ (1, \sigma_1, 1)\ \}$. Then X_{σ_1} is the Golden Mean shift and X_{σ_2} is the full shift.

In the Cayley graph of G, there are infinitely many copies of X_{σ_i}. Given any $w \in G$ the shift of finite type denoted by $X_{\sigma_i}^{(w)}$ will consist of all allowed colorings of the vertices

$$\dots f(w\sigma_i^{-2}), f(w\sigma_i^{-1}), f(w), f(w\sigma_i^{1}), f(w\sigma_i^{2}), \dots.$$

A point $x \in X_{\sigma_i}^{(w)}$ is a function $f : \{w\sigma_i^n : n \in \mathbb{Z}\} \to \mathcal{A}$. In other words, $X_{\sigma_i}^{(w)}$ is a one-dimensional shift of finite type that colors all vertices in the σ_i direction starting at the group element w. Figure 1 shows the vertices in the

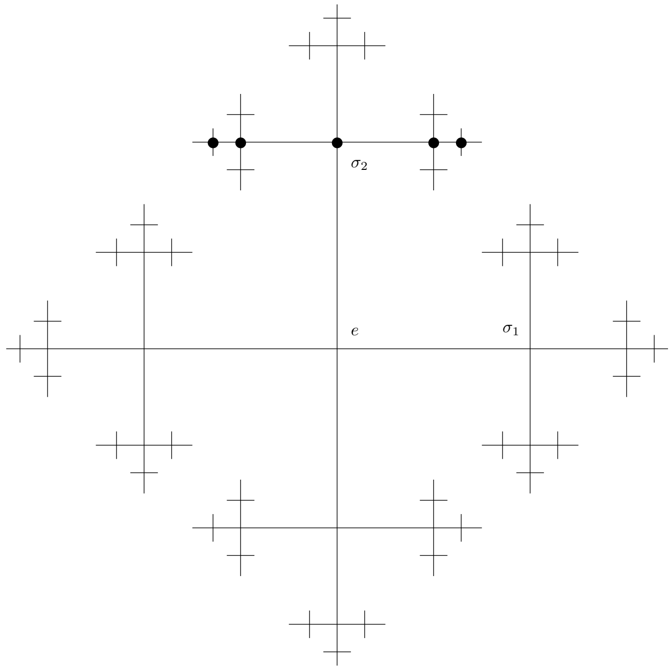

FIGURE 1. The black vertices are colored in $X_{\sigma_1}^{(\sigma_2)}$, for the free group on two generators.

Cayley graph of G that correspond to the one-dimensional shift of finite type $X_{\sigma_1}^{(\sigma_2)}$.

If the range of f is S, we say that f *uses* each color in S. Note that $X_{\sigma_i}^{(w)}$ and $X_{\sigma_i}^{(\sigma_i w)}$ color the same set of group elements, but $X_{\sigma_i}^{(w)}$ and $X_{\sigma_i}^{(\sigma_j w)}$ do not for $j \neq i$.

1.2. A way to specify colorings. Each one-dimensional shift of finite type X_{σ_i} sets restrictions about how G can be colored. However, it may be difficult to specify explicitly a given coloring $f \in \mathsf{X}_F$. Here we develop a way of presenting a coloring f of G which will be useful later in showing nonemptyness and constructing strongly periodic colorings.

DEFINITION 1.4. On the free group G on q generators, a set of $2q$ functions,

$$\{h_x : \mathcal{A} \to \mathcal{A} \text{ such that } x \in \sigma \text{ or } x^{-1} \in \sigma\}$$

is called a set of *coloring functions*.

Coloring functions can *specify* colorings in the following way. We begin by defining $f(e) = a$ for some $a \in \mathcal{A}$. We then repeat the following process: if $f(g)$ has been defined but $f(gx)$ has not yet been defined for $x \in \sigma$ or $x^{-1} \in \sigma$, we define $f(gx) = h_x(f(g))$. Therefore, once we have colored one vertex, the h_x inductively define the rest of the coloring.

It should be clear that if such coloring functions exist and obey the forbidden blocks so that we always have $(l, \sigma_i, h_{\sigma_i}(l)) \notin F$ and $(h_{\sigma_i^{-1}}(l), \sigma_i, l) \notin F$, then the shift space X_F must be nonempty.

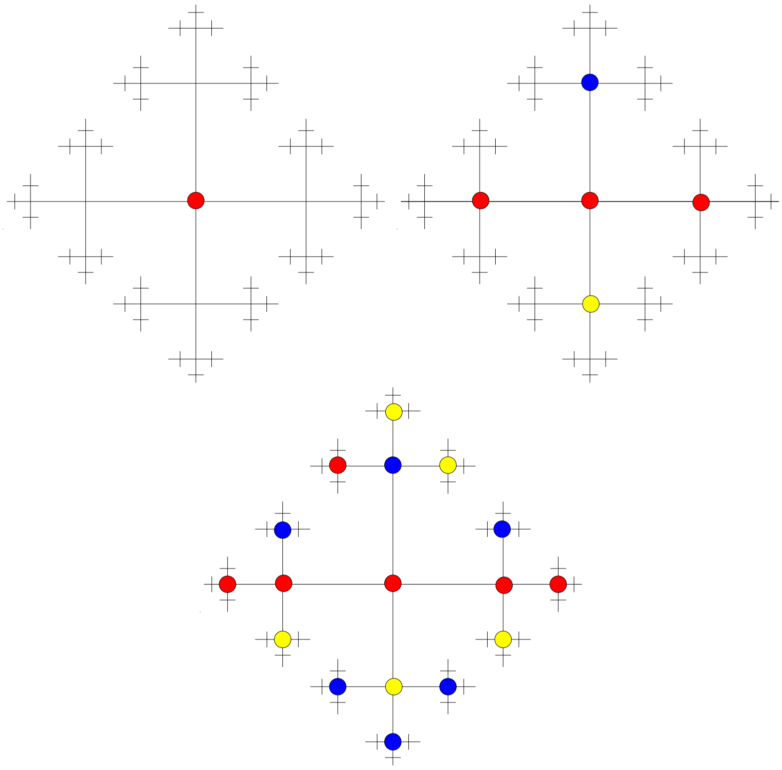

FIGURE 2. The first three steps in the coloring specified in Example 1.5.

EXAMPLE 1.5. Suppose G is the free group on two generators, $\mathcal{A} = \{R, Y, B\}$, and that the h_x are defined as follows:

$$\begin{aligned} h_{\sigma_1}(R) &= R & h_{\sigma_1^{-1}}(R) &= R \\ h_{\sigma_1}(B) &= Y & h_{\sigma_1^{-1}}(B) &= R \\ h_{\sigma_1}(Y) &= B & h_{\sigma_1^{-1}}(Y) &= B \end{aligned}$$

and

$$\begin{aligned} h_{\sigma_2}(R) &= B & h_{\sigma_2^{-1}}(R) &= Y \\ h_{\sigma_2}(B) &= Y & h_{\sigma_2^{-1}}(B) &= R \\ h_{\sigma_2}(Y) &= R & h_{\sigma_2^{-1}}(Y) &= B. \end{aligned}$$

If we begin by coloring e with R, then Figure 2 shows how the coloring of G proceeds. Throughout this thesis, R, G, and B in $\mathcal{A}$ will respectively stand for the colors red, green, and blue.

Also, if we require that the h_x are invertible, then the coloring specified by them will be strongly periodic:

THEOREM 1.6. *Suppose we color e by any $a \in \mathcal{A}$, and then produce a coloring $f : G \to \mathcal{A}$ using some set of coloring functions S. If $h_x^{-1} = h_{x^{-1}}$ for each $h_x \in S$, then f is strongly periodic.*

PROOF. We first note that a coloring f specified by S is *color-isotropic*, meaning that if $f(g_1) = f(g_2)$ then $f(g_1 w) = f(g_2 w)$ for all $w \in G$. This should be evident from the fact that each h_x is a bijection since $h_x^{-1} = h_{x^{-1}}$. But every color-isotropic coloring is necessarily strongly periodic, since any shift Tf of f will color e one of a finite number of colors. Each color that e can be will uniquely determine a coloring of the entire graph, so the orbit $\{Tf : T \in G\}$ must be finite. □

2. Nonemptyness and weakly periodic properties

First we will show that it is easy to determine if a shift of finite type on the free group contains any points. Then, we will show that the behavior of shifts of finite type on the free group is somewhere between the behavior of $\mathbb{Z}$ and $\mathbb{Z} \times \mathbb{Z}$: a nonempty shift of finite type on the free group may not contain a strongly periodic coloring (like $\mathbb{Z} \times \mathbb{Z}$), but it will always contain a weakly periodic coloring (like $\mathbb{Z}$).

THEOREM 2.1. *A shift of finite type X_F on the free group is nonempty if and only if there exists some $L \subseteq \mathcal{A}$ such that for every $l \in L$, each $X_{\sigma_i} |_L$ contains a point which uses the color l.*

PROOF. Suppose such an L exists. Fix $\sigma_i \in \sigma$. Then for each $l \in L$, each $X_{\sigma_i} |_L$ has a point which uses l, so there must exist $l' \in L$ such that $(l, \sigma_i, l') \notin F$. Set $h_{\sigma_i}(l) = l'$ for such an l'. Similarly, for each $l \in L$ there exists $l' \in L$ such that $(l', \sigma_i, l) \notin F$ so define $h_{\sigma_i^{-1}}(l) = l'$ for this l'. These h_x define a coloring function on the alphabet L which obeys the forbidden blocks, so G can be colored with L.

For the converse, suppose X_F is nonempty. Then X_F contains some coloring $f : G \to \mathcal{A}$. Let $S \subseteq \mathcal{A}$ be the range of f. Fix any $\sigma_i \in \sigma$ and $s \in S$. There exists $g \in G$ such that $f(g) = s$. As we let n range over $\mathbb{Z}$, $f(\sigma_i^n g)$ gives a point in $X_{\sigma_i} |_S$ which uses s. □

We next prove that, as in one-dimensional symbolic dynamics, any nonempty shift of finite type contains a weakly periodic coloring. While the notation of this proof is somewhat cumbersome, the idea is simple. To construct a weakly periodic coloring d, first color $X_{\sigma_1}^{(e)}$ periodically. To fill in the rest of the coloring, one must only make sure that if $d(\sigma_1^n) = d(\sigma_1^m)$ then $d(\sigma_1^n g) = d(\sigma_1^m g)$ for all $g \in G$.

THEOREM 2.2. *If X_F is nonempty, it contains a weakly periodic coloring.*

PROOF. Let $f : G \to \mathcal{A}$ be a coloring of G in X_F. Let $S = \{f(\sigma_1^n) : n \in \mathbb{Z}\}$ be the range of f on $X_{\sigma_1}^{(e)}$. We will need to keep track of one place that each color in S occurs in $X_{\sigma_1}^{(e)}$, so define $w : S \to \mathbb{Z}$ by choosing for each $l \in S$ an $n \in Z$ such that $f(\sigma_1^n) = l$, and define $w(l) = n$. There may be more than one choice for n, and it does not matter what n we choose so long as $f(\sigma_1^{w(l)}) = l$. Note that $X_{\sigma_1}^{(e)} |_S$ is a nonempty, one-dimensional shift of finite type, so it contains a periodic coloring, say $P : \{\sigma_1^n : n \in \mathbb{Z}\} \to S$, of some period p.

We will construct a weakly periodic coloring $d : G \to \mathcal{A}$. First, set $d(\sigma_1^n) = P(\sigma_1^n)$ for all $n \in \mathbb{Z}$ so that $X_{\sigma_1}^{(e)}$ is colored periodically. Note that every element in G is either of the form σ_1^n or $\sigma_1^n xg$ for some uniquely determined

$g \in \mathcal{G}$ and uniquely determined x, with $x \in \sigma$ or $x^{-1} \in \sigma$, $x \neq \sigma_1$, and $x \neq \sigma_1^{-1}$. So set

$$d(\sigma_1^n xg) = f(\sigma_1^{w(P(\sigma_1^n))} xg)$$

Then

$$\begin{aligned} \sigma_1^p d(\sigma_1^n xg) &= d(\sigma_1^{n+p} xg) \\ &= f(\sigma_1^{w(P(\sigma_1^{n+p}))} xg) \\ &= f(\sigma_1^{w(P(\sigma_1^n))} xg) \\ &= d(\sigma_1^n xg) \end{aligned}$$

since P has period p. This shows d is weakly periodic. Also, we must have $d \in \mathsf{X}_F$ since if $d(g_1) = a \in \mathcal{A}$ and $d(g_2) = b \in \mathcal{A}$, then there exists $g_1', g_2' \in G$ such that $f(g_1') = a$ and $f(g_2') = b$ by the construction of d. Since we are only considering nearest-neighbor shifts of finite type, this shows that d does not disobey any of the adjacency rules specified by F. □

3. Strongly periodic colorings of the free group

We now determine when a shift of finite type on the free group admits a strongly periodic coloring. This question turns out to be much more subtle than it may initially seem, and determining whether an arbitrary shift of finite type admits a strongly periodic coloring appears to be a difficult combinatorial problem.

3.1. Cycles as point representations. First, we will represent a periodic point in a one-dimensional shift of finite type with a *cycle*:

Definition 3.1. A *cycle* w is an expression $w = \overline{x_1 x_2 \dots x_n}$ where $x_i \in \mathcal{A}$. A cycle w *represents a periodic point* x if

$$x = \dots x_1 x_2 \dots x_2 . x_1 x_2 \dots x_n x_1 x_2 \dots x_n \dots .$$

A cycle $w = \overline{x_1 x_2 \dots x_n}$ has *length* $|w| = n$ and we will sometimes write $(w)_i = x_i$.

Note that a periodic point can be represented by infinitely many cycles and that the cycle length of a periodic point is a multiple of the least period of that point. In fact, the theorems of the next section could be restated in terms of multiples of the least period; however, doing so would be much more complicated notationally.

Example 3.2. If $x = \dots 1010101.101010 \dots$ then the following are all cycles that represent x: $\overline{10}$, $\overline{1010}$, $\overline{101010}$. The cycle $\overline{101}$ does not represent x.

We next define a function η_a on cycles which counts the number of times the color $a \in \mathcal{A}$ appears.

Definition 3.3. For a cycle w and $a \in \mathcal{A}$, define

$$\eta_a(w) = \#\{i : (w)_i = a, 1 \leq i \leq |w|\}.$$

If S is a set of cycles, define

$$\eta_a(S) = \sum_{w \in S} \eta_a(w).$$

EXAMPLE 3.4. If $w_1 = \overline{RGRB}$ then $\eta_R(w_1) = 2$, and $\eta_G(w_1) = \eta_B(w_1) = 1$. If $w_2 = \overline{RGB}$ then $\eta_R(\{w_1, w_2\}) = 3$, $\eta_B(\{w_1, w_2\}) = 2$ and $\eta_G(\{w_1, w_2\}) = 2$.

3.2. Existence of strongly periodic colorings. We now prove the main theorem of this section and chapter. The proof is notationally complex, so we follow each major step of the proof with an example.

THEOREM 3.5. *A shift of finite type on the free group on q generators contains a strongly periodic coloring if and only if there exist finite sets $S_{\sigma_1}, S_{\sigma_2}, \ldots, S_{\sigma_q}$ such that elements of S_{σ_i} are cycles that represent points in X_{σ_i}, and for all $a \in \mathcal{A}$ and $\sigma_i, \sigma_j \in \sigma$ we have $\eta_a(S_{\sigma_i}) = \eta_a(S_{\sigma_j})$.*

PROOF. We first prove the $\Leftarrow$ direction. Suppose such S_{σ_i} exist, and suppose $S_{\sigma_i} = \{w_1^{\sigma_i}, w_2^{\sigma_i}, \ldots, w_{p_{\sigma_i}}^{\sigma_i}\}$ for each $\sigma_i \in \sigma$. Define the alphabet $\mathcal{B}$ by

$$\mathcal{B} = \bigcup_{\sigma_i \in \sigma} \bigcup_{j=1}^{|S_{\sigma_i}|} \bigcup_{s=1}^{|w_j^{\sigma_i}|} \{x_{j,s}^{\sigma_i}\},$$

so that $\mathcal{B}$ consists of symbols such as $x_{1,1}^{\sigma_1}$, $x_{1,2}^{\sigma_1}$, $x_{2,1}^{\sigma_2}$, etc.. Each element $x_{j,s}^{\sigma_i}$ in $\mathcal{B}$ can be thought of as specifying a cycle $w_j^{\sigma_i}$ in S_{σ_i}, and a position s in that cycle.

EXAMPLE 3.6. Suppose we are attempting to color the free group with the shifts of finite type shown in Figure 3, and we have chosen

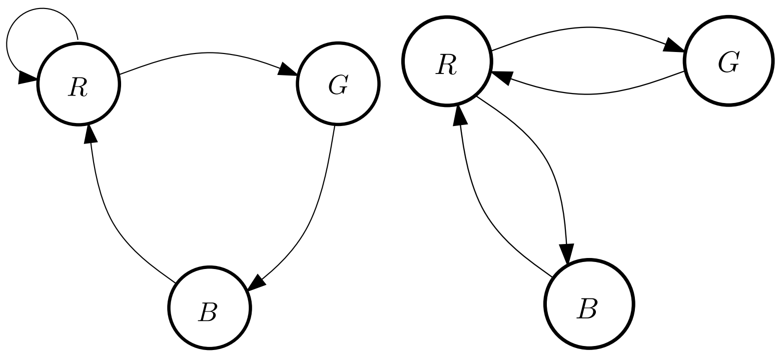

FIGURE 3. Shifts X_{σ_1} and X_{σ_2}.

$$S_{\sigma_1} = \{\overline{RRGB}\}$$

and

$$S_{\sigma_2} = \{\overline{RB}, \overline{RG}\},$$

with $w_1^{\sigma_2} = \overline{RB}$ and $w_2^{\sigma_2} = \overline{RG}$. Note that $\eta_R(S_{\sigma_1}) = \eta_R(S_{\sigma_2}) = 2$, $\eta_G(S_{\sigma_1}) = \eta_G(S_{\sigma_2}) = 1$ and $\eta_B(S_{\sigma_1}) = \eta_B(S_{\sigma_2}) = 1$. Then the alphabet

$$\mathcal{B} = \{x_{1,1}^{\sigma_1},\ x_{1,2}^{\sigma_1},\ x_{1,3}^{\sigma_1},\ x_{1,4}^{\sigma_1},\ x_{1,1}^{\sigma_2},\ x_{1,2}^{\sigma_2},\ x_{2,1}^{\sigma_2},\ x_{2,2}^{\sigma_2}\}.$$

Next, we define $\chi : \mathcal{B} \to \mathcal{A}$ by

$$\chi(x^{\sigma_i}_{j,s}) = (w^{\sigma_i}_j)_s$$

so that

$$w^{\sigma_i}_j = \overline{\chi(x^{\sigma_i}_{j,1})\chi(x^{\sigma_i}_{j,2})\dots\chi(x^{\sigma_i}_{j,|w^{\sigma_i}_j|})}.$$

That is, χ maps the symbol $x^{\sigma_i}_{j,s}$ to the color at the location (in a cycle) specified by $x^{\sigma_i}_{j,s}$.

EXAMPLE 3.7. In our example, we have,

$$\begin{aligned} \chi(x^{\sigma_1}_{1,1}) = \chi(x^{\sigma_1}_{1,2}) = \chi(x^{\sigma_2}_{1,1}) = \chi(x^{\sigma_2}_{2,1}) &= R \\ \chi(x^{\sigma_1}_{1,3}) = \chi(x^{\sigma_2}_{2,2}) &= G \\ \chi(x^{\sigma_1}_{1,4}) = \chi(x^{\sigma_2}_{1,2}) &= B. \end{aligned}$$

Since we assume $\eta_a(S_{\sigma_i}) = \eta_a(S_{\sigma_j})$ for all $a \in A$, $\sigma_i, \sigma_j \in \sigma$, we define η_a to be the common value of all $\eta_a(S_{\sigma_i})$. Define $N = \sum_{a\in\mathcal{A}} \eta_a$ and note we also have $N = \sum_{w\in S_{\sigma_i}} |w|$. Next we have the following lemma:

LEMMA 3.8. *Given $S_{\sigma_1}, S_{\sigma_2}, \dots, S_{\sigma_q}$ that satisfy the hypotheses of the theorem, then there exist sets $E_1, E_2, \dots E_N$ that partition $\mathcal{B}$ and meet the following conditions:*

(i) *If $x^{\sigma_i}_{j,s} \in E_r$ and $x^{\sigma_{i'}}_{j',s'} \in E_r$ then $\chi(x^{\sigma_i}_{j,s}) = \chi(x^{\sigma_{i'}}_{j',s'})$.*

(ii) *For every $\sigma_i \in \sigma$, each E_r contains exactly one element of the form $x^{\sigma_i}_{j,s}$ (where j and s may depend on σ_i).*

PROOF. In this proof we temporarily change notation. Suppose $\mathcal{A} = \{1, 2, 3 \dots, m\}$ and

$$\mathcal{B} = \{b^1_1, b^1_2, \dots, b^1_{p_1}, b^2_1, b^2_2, \dots, b^2_{p_2}, \dots b^q_1, b^q_2, \dots, b^q_{p_q}\}$$

for some p_r, where $b^i_x = x^{\sigma_i}_{j,s}$ for some j, s. For each fixed $i < |\sigma|$, we can "sort" the $b^{\sigma_i}_x$ by color so that

$$\overbrace{b^{\sigma_i}_{y^i_1} b^{\sigma_i}_{y^i_2} \dots b^{\sigma_i}_{y^i_{\eta_1}}}^{\chi \text{ maps these to } 1\in\mathcal{A}} \; \overbrace{b^{\sigma_i}_{y^i_{(\eta_1+1)}} \; b^{\sigma_i}_{y_{(\eta_1+1)}v} \dots b^{\sigma_i}_{y^i_{(\eta_1+\eta_2)}}}^{\chi \text{ maps these to } 2\in\mathcal{A}} \dots\dots \overbrace{b^{\sigma_i}_{y^i_{(N-\eta_m+1)}} \; b^{\sigma_i}_{y^i_{(N-\eta_m+2)}} \dots b^{\sigma_i}_{y^i_{(N)}}}^{\chi \text{ maps these to } m\in\mathcal{A}},$$

for appropriate choices of the $y^i_a \in \mathbb{N}$ with $1 < a < p_i$. The conditions on the S_{σ_i} assure that

$$\#\{b^{\sigma_i}_x \in \mathcal{B} : \chi(b^{\sigma_i}_x) = a\} = \#\{b^{\sigma_j}_x \in \mathcal{B} : \chi(b^{\sigma_j}_x) = a\}$$

for all $a \in \mathcal{A}$ and $\sigma_i, \sigma_j \in \sigma$. Thus, when we sort by color, each row of $b^{\sigma_i}_x$ has the same number of elements (η_a) that χ maps to $a \in A$. That is, we can write

$$
\begin{array}{cccc}
\overbrace{\begin{array}{c} b^1_{y^1_1} b^1_{y_2} \dots b^1_{y^1_{\eta_1}} \\ b^2_{y^2_1} b^2_{y^2_2} \dots b^2_{y^2_{\eta_1}} \\ \vdots \\ \vdots \\ b^q_{y^q_1} b^q_{y^q_2} \dots b^q_{y^q_{\eta_1}} \end{array}}^{\chi \text{ maps these to } 1\in\mathcal{A}}
&
\overbrace{\begin{array}{c} b^1_{y^1_{(\eta_1+1)}} \; b^1_{y^1_{(\eta_1+1)}} \dots b^1_{y^1_{(\eta_1+\eta_2)}} \\ b^2_{y^2_{(\eta_1+1)}} \; b^2_{y^2_{(\eta_1+1)}} \dots b^2_{y^2_{(\eta_1+\eta_2)}} \\ \vdots \\ \vdots \\ b^q_{y^q_{(\eta_1+1)}} \; b^q_{y^q_{(\eta_1+1)}} \dots b^q_{y^q_{(\eta_1+\eta_2)}} \end{array}}^{\chi \text{ maps these to } 2\in\mathcal{A}}
&
\begin{array}{c} \dots \\ \dots \\ \vdots \\ \vdots \\ \dots \end{array}
&
\overbrace{\begin{array}{c} b^1_{y^1_{(N-\eta_m+1)}} \; b^1_{y^1_{(N-\eta_m+2)}} \dots b^1_{y^1_{(N)}} \\ b^2_{y^2_{(N-\eta_m+1)}} \; b^2_{y^2_{(N-\eta_m+2)}} \dots b^2_{y^2_{(N)}} \\ \vdots \\ \vdots \\ b^q_{y^q_{(N-\eta_m+1)}} \; b^q_{y^q_{(N-\eta_m+2)}} \dots b^q_{y^q_{(N)}} \end{array}}^{\chi \text{ maps these to } m\in\mathcal{A}}
\end{array}
$$

for appropriate choices of $y^i_a \in \mathbb{N}$ with $1 < a < p_i$. To get the E_r, simply read down the column of this array. That is, let

$$E_r = \bigcup_{i=1}^{q} \{b^i_{y^i_r}\}.$$

Then the fact that the rows are sorted by color and each row contains the same number of each color implies (i). The fact that each E_y contains exactly one element from each row implies (ii). $\square$

Continuing in the proof of the main theorem, we can apply Lemma 3.8 to choose $E_1, E_2, \dots E_N$ that satisfy (i) and (ii). Note that (i) assures that elements of a given E_r all specify locations in cycles that are colored the same color, while (ii) assures that colors from cycles in each S_{σ_i} are represented exactly once in each E_r, so that $|E_r| = |\sigma|$.

Example 3.9. In our example, there are several ways to choose the E_r. One is,

$$\begin{aligned} E_1 &= \{x^{\sigma_1}_{1,1}, x^{\sigma_2}_{1,1}\} \\ E_2 &= \{x^{\sigma_1}_{1,2}, x^{\sigma_2}_{2,1}\} \\ E_3 &= \{x^{\sigma_1}_{1,3}, x^{\sigma_2}_{2,2}\} \\ E_4 &= \{x^{\sigma_1}_{1,4}, x^{\sigma_2}_{1,2}\}. \end{aligned}$$

We will now use the E_r to construct a coloring of the free group with the alphabet $\{1, 2, \dots N\}$, which will later be projected down to a strongly periodic coloring by $\mathcal{A}$.

We produce a coloring by defining coloring functions

$$h_{\sigma_i} : \{1, 2, \dots N\} \to \{1, 2, \dots, N\}$$

for each $\sigma_i \in \sigma$. Given an $r \in \{1, \dots, N\}$ and $\sigma_i \in \sigma$, find the unique $j, s \in \mathbb{Z}$ such that $x^{\sigma_i}_{j,s} \in E_r$. Define $h_{\sigma_i}(r) = r'$, where r' is such that

$$x^{\sigma_i}_{j,(s+1 \bmod |w^{\sigma_i}_j|)} \in E_{r'}.$$

That is, given r, h_{σ_i} finds the element of the form $x^{\sigma_i}_{j,s} \in E_r$, and $h_{\sigma_i}(x^{\sigma_i}_{j,s})$ equals the number of the set containing the symbol adjacent to $x^{\sigma_i}_{j,s}$ in the cycle $w^{\sigma_i}_j$. Each h_{σ_i} is invertible, so we can define $h_{\sigma_i^{-1}} = h^{-1}_{\sigma_i}$.

EXAMPLE 3.10. In our example, $N = 4$ and the h_{σ_i} are given by

$$\begin{aligned} h_{\sigma_1}(1) &= 2 & h_{\sigma_2}(1) &= 4 \\ h_{\sigma_1}(2) &= 3 & h_{\sigma_2}(2) &= 3 \\ h_{\sigma_1}(3) &= 4 & h_{\sigma_2}(3) &= 2 \\ h_{\sigma_1}(4) &= 1 & h_{\sigma_2}(4) &= 1. \end{aligned}$$

Note that each h_{σ_i} is necessarily a bijection, and therefore it defines a strongly periodic coloring by Theorem 1.6. Suppose this strongly periodic coloring is $F : G \to \{1, 2, \ldots, N\}$. We project F down to a coloring of G by using χ. Define $f : G \to \mathcal{A}$ by $f(g) = \chi(x)$, where x is any arbitrary element of $E_{F(g)}$. Note the specific choice of x does not matter because of condition (i) above. The orbit $\{TF : T \in G\}$ is finite since F is strongly periodic, and χ maps each of these colorings to at most one distinct coloring, so the orbit

$$\{Tf : T \in G\} = \chi(\{TF : T \in G\})$$

must be finite. This shows f is strongly periodic.

EXAMPLE 3.11. Suppose we color e with 1. The coloring of G by $\{1, 2, 3, 4\}$ specified by the h_x above is strongly periodic. Figure 4 shows the coloring

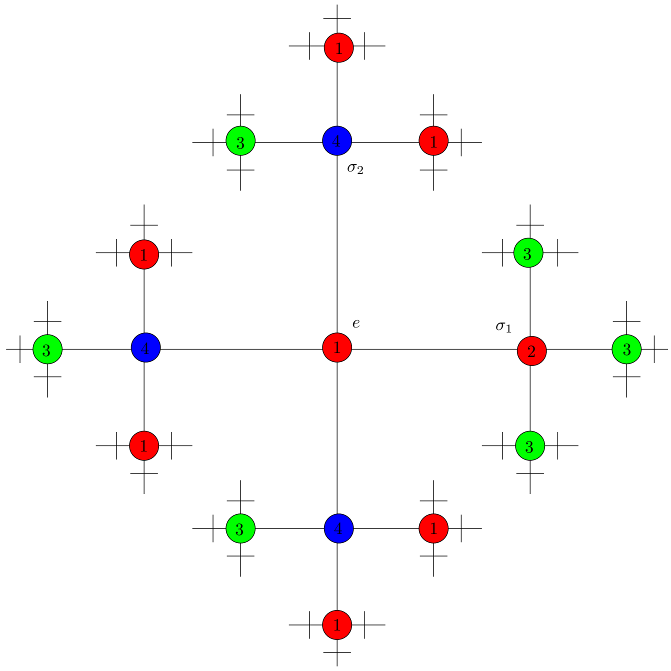

FIGURE 4. The coloring from Example 3.11.

the h_x specify over the alphabet $\{1, 2, 3, 4\}$, and the projection down onto $\{R, G, B\}$. Note that this coloring is color-isotropic over $\{1, 2, 3, 4\}$, but not $\{R, G, B\}$.

To prove the $\Rightarrow$ direction, we show that any strongly periodic coloring defines a set of cycles that meet the conditions of the theorem. Consider any strongly periodic coloring f which satisfies the shift of finite type constraints.

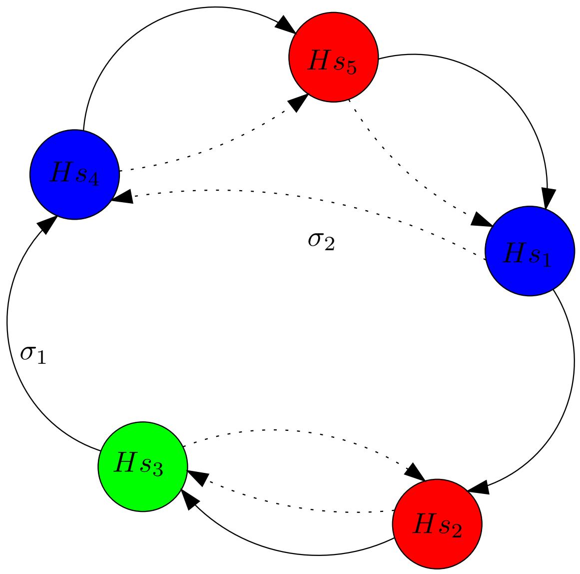

FIGURE 5. The Cayley graph of $G \setminus H$.

We can define the stabilizer of f, $H = \{T \in G : Tf = f\}$, which is a finite index subgroup since f is strongly periodic. H is like a fundamental region of the strongly periodic coloring.

Note that H is a normal subgroup, so consider the group of right cosets $G \setminus H$ with the binary operation $(Hr) \cdot (Hs) = (Hrs)$. Since H is a stabilizer of f, $hf(r) = f(r)$ for all $h \in H$ and $r \in G$. Thus, elements of the coset (aH) are all colored the same color $f(a)$, but these colors may not be distinct for different cosets.

EXAMPLE 3.12. Figure 5 shows a Cayley graph for $G \setminus H$, where H is a finite index subgroup. Each vertex in this graph is labeled with an element of $Hs_j \in G \setminus H$ and a color equal to $f(Hs_j)$, so that, for example, $f(Hs_1) = R$. There are two kinds of edges on this graph: solid correspond to σ_1, and dotted correspond to σ_2. Each vertex has exactly one incoming and outgoing edge of each type because $G \setminus H$ is a group. Given an arbitrary group element $g = \sigma_{a_1}\sigma_{a_2}\dots\sigma_{a_z}$, we can find $(Hg)(Hs_j)$ by following the appropriate sequence of dotted and solid edges from (Hs_j).

Now fix $\sigma_i \in \sigma$, and consider any coset (Hs). We define the *coset orbit*

$$O_{\sigma_i}(Hs) = \{(H\sigma_i)^n(Hs) : n \in \mathbb{Z}\}.$$

Since $G \setminus H$ is finite, each coset orbit $O_{\sigma_i}(Hs)$ must be finite. Note that two coset orbits $O_{\sigma_i}(Hs_1)$ and $O_{\sigma_i}(Hs_2)$ are either equal or disjoint. Therefore, there are some finite number of distinct coset orbits, $O_{\sigma_i}(Hs_1), O_{\sigma_i}(Hs_2), \dots, O_{\sigma_i}(Hs_{m_i})$, which together form a disjoint partition of $G \setminus H$ for each $\sigma_i \in \sigma$.

EXAMPLE 3.13. In Figure 5, we have the following disjoint coset orbits:

$$O_{\sigma_1}(Hs_1) = \{(Hs_1),\ (Hs_2),\ (Hs_3),\ (Hs_4),\ (Hs_5)\}$$
$$O_{\sigma_2}(Hs_1) = \{(Hs_1),\ (Hs_4),\ (Hs_5)\}$$
$$O_{\sigma_2}(Hs_2) = \{(Hs_2),\ (Hs_3)\}.$$

We alternatively could have chosen different Hs_j since, for example, $O_{\sigma_1}(Hs_1) = O_{\sigma_1}(Hs_5)$.

Each coset orbit $O_{\sigma_i}(Hs_j)$ is a collection of cosets given by

$$O_{\sigma_i}(Hs_j) = \{(Hs_j), (H\sigma_i s_j), (H\sigma_i^2 s_j), \ldots, (H\sigma_i^n s_j)\},$$

where n depends on i and j. We can therefore regard each coset orbit $O_{\sigma_i}(Hs_j)$ as specifying a cycle $C_{\sigma_i}(Hs_j)$ given by

$$C_{\sigma_i}(Hs_j) = \overline{f(s_j)f(\sigma_i^1 s_j)f(\sigma_i^2 s_j)\ldots f(\sigma_i^n s_j)}.$$

Since H is the stabilizer of a coloring in X_F, each $C_{\sigma_i}(Hs_1)$ represents a point in X_{σ_i}.

EXAMPLE 3.14. In the example from Figure 5, $C_{\sigma_1}(Hs_1) = \overline{RBRGB}$. Also, $C_{\sigma_2}(Hs_2) = \overline{BBR}$ and $C_{\sigma_2}(Hs_2) = \overline{RG}$. Note that we could have chosen different cycles (by shifting where we start), but these cycles still would have represented points in X_{σ_1} and X_{σ_2} since X_{σ_1} and X_{σ_2} are shift-invariant.

We then define the set

$$S_{\sigma_i} = \{C_{\sigma_i}(Hs_j) : 1 \leq j \leq m_i\},$$

where m_i is the number of distinct coset orbits. Note that

$$\begin{aligned}
\eta_a(S_{\sigma_i}) =& \sum_{j=1}^{m_i} \eta_a(C_{\sigma_i}(Hs_j)) \\
=& \sum_{j=1}^{m_i} \#\{i : (C_{\sigma_i}(Hs_j))_i = a, 1 \leq i \leq |C_{\sigma_i}(Hs_j)|\} \\
=& \sum_{j=1}^{m_i} \#\{x : x \in O_{\sigma_i}(Hs_j), \chi(x) = a\} \\
=& \#\{(Hs) \in G \setminus H : f(Hs) = a\},
\end{aligned}$$

since the $O_{\sigma_i}(Hs_j)$ form a disjoint partition of $G \setminus H$. However, the value

$$\#\{(Hs) \in G \setminus H : f(Hs) = a\}$$

is independent of σ_i, which shows that $\eta_a(S_{\sigma_i}) = \eta_a(S_{\sigma_j})$ for all $\sigma_i, \sigma_j \in \sigma$. This shows that the S_{σ_i} meet the hypotheses of the theorem.

EXAMPLE 3.15. In the example from Figure 5, we have that $S_{\sigma_1} = \{\overline{RBRGB}\}$ and $S_{\sigma_2} = \{\overline{BBR}, \overline{RG}\}$. Note that $\eta_a(S_{\sigma_1}) = \eta_a(S_{\sigma_2})$ for all $a \in \{R, G, B\}$.

□

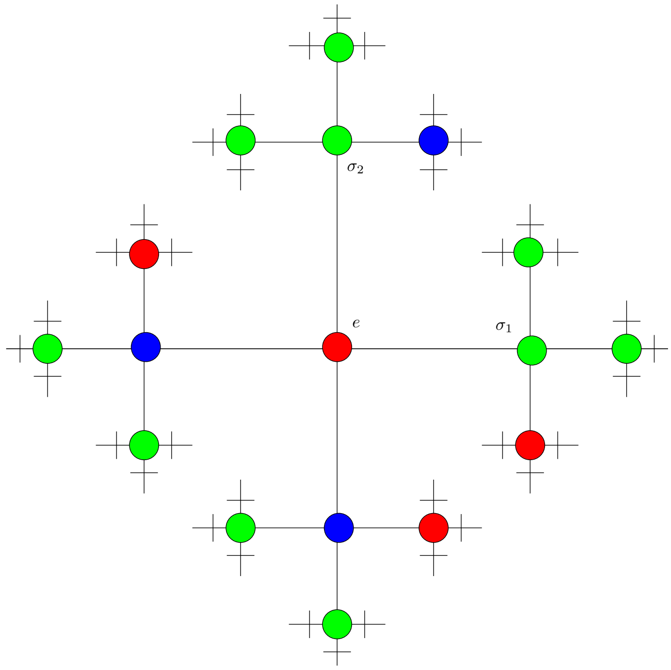

FIGURE 6. The coloring from Example 3.18.

This theorem is somewhat awkward to prove, and at first glance it may seem as though its hypotheses are difficult to work with. However, after stating a few obvious corollaries that justify proving the theorem, we will give an example of a case where the theorem can be used to show that a shift of finite type on the free group does not contain a strongly periodic point.

COROLLARY 3.16. *If there is a one-dimensional periodic point x which is an element of X_{σ_i} for all $\sigma_i \in \sigma$, then X_F contains a strongly periodic coloring.*

PROOF. The hypotheses of the theorem are satisfied by choosing $S_{\sigma_i} = \{y\}$ for all $\sigma_i \in \sigma$, where y is some cycle representing x. □

COROLLARY 3.17. *If $X_{\sigma_i} = X_{\sigma_j}$ for all $\sigma_i, \sigma_j \in \sigma$ and some X_{σ_i} is nonempty, then the shift of finite type on the free group contains a strongly periodic coloring.*

PROOF. Every nonempty one-dimensional shift of finite type contains a periodic point x, so apply the previous corollary. □

EXAMPLE 3.18. Consider the free group on two generators. Suppose X_{σ_1} and X_{σ_2} both contain the periodic point

$$x = \ldots RGGBRGGB.RGGBRGGB \ldots$$

which is represented by the cycle $w = \overline{RGGB}$. We can choose $S_{\sigma_1} = S_{\sigma_2} = \{w\}$ to satisfy the hypotheses of Theorem 3.5. If we follow the construction in Theorem 3.5, we can produce the periodic coloring of the free group shown in Figure 6.

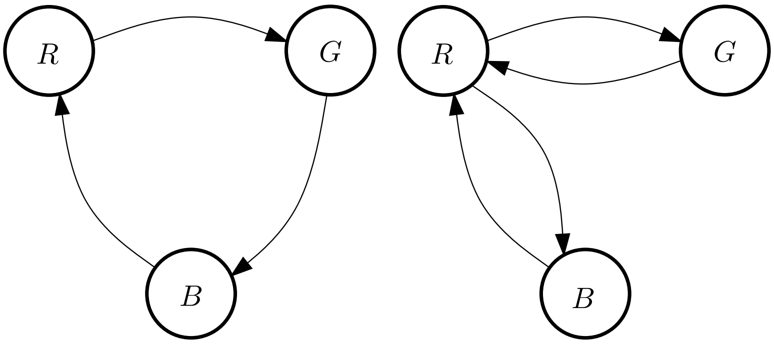

FIGURE 7. Shifts X_{σ_1} and X_{σ_2}.

4. An example that is not strongly periodic

Now we present an example of a shift of finite type on the free group where each one-dimensional shift of finite type is irreducible, but the free group still cannot be colored periodically. This, perhaps, is a counterintuitive result because irreducibility is a strong condition and implies the existence of many kinds of periodic points.

EXAMPLE 4.1. Consider the shifts of finite type shown in Figure 7. Cycles representing points in X_{σ_1} are always of the form

$$\overline{RGBRGB\ldots RGB}.$$

Cycles representing points in X_{σ_2} are always of the form

$$\overline{Ra_1Ra_2\ldots Ra_{n-1}Ra_n},$$

where each $a_j \in \{G, B\}$. This shows that for any set S_{σ_1} of cycles in X_{σ_1}, we must have

$$\eta_R(S_{\sigma_1}) = \eta_G(S_{\sigma_1}) = \eta_B(S_{\sigma_1}). \tag{3}$$

In addition, for any set S_{σ_2} of cycles in X_{σ_2} we must have

$$\eta_R(S_{\sigma_2}) = \eta_G(S_{\sigma_2}) + \eta_B(S_{\sigma_2}). \tag{4}$$

If $\eta_a(S_{\sigma_1}) = \eta_a(S_{\sigma_2}) = \eta_a$ for all $a \in \mathcal{A}$, then (3) and (4) become

$$\eta_R = \eta_G = \eta_B \tag{5}$$

and

$$\eta_R = \eta_G + \eta_B. \tag{6}$$

Clearly we cannot simultaneously solve (5) and (6) unless $\eta_G = 0$ or $\eta_B = 0$, which would imply that S_{σ_1} is empty, since every point in X_{σ_1} uses G and B. Therefore, even though both X_{σ_1} and X_{σ_2} are irreducible, X_F does not allow a strongly periodic coloring.

CHAPTER 4

The golden mean shift on the free group

In this section, we study the golden mean shift on the free group and determine an expression for its entropy in Theorem 3.4. In doing so, we develop a new generalization of Fibonacci numbers and analyze them using ideas from one-dimensional dynamical system theory.

1. Preliminaries

We will examine the golden mean shift on the free group with q generators, where $q \geq 2$. However, many results of this section will still hold true for $q = 1$ (ie $G = \mathbb{Z}$). We define $k = 2q - 1$ and assume that G has the standard presentation $\langle \sigma_1, \sigma_2, \ldots, \sigma_q \mid \rangle$. With this presentation, the Cayley graph of G is an infinite tree like the one shown in Figure 1. In Figure 1, each element of G is represented by an intersection of lines, with e at the center. The golden mean shift corresponds to coloring each vertex of this graph with the alphabet $\mathcal{A} = \{0, 1\}$ such that no two adjacent vertices are both colored 1. That is, the set of forbidden blocks $F = \{(1, \sigma_i, 1) : \sigma_i \in \sigma\}$, with the notation of the last chapter.

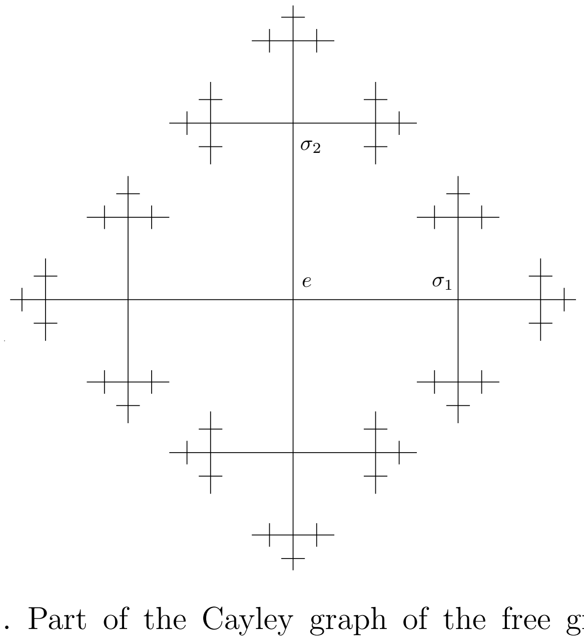

FIGURE 1. Part of the Cayley graph of the free group on 2 generators ($k = 3$).

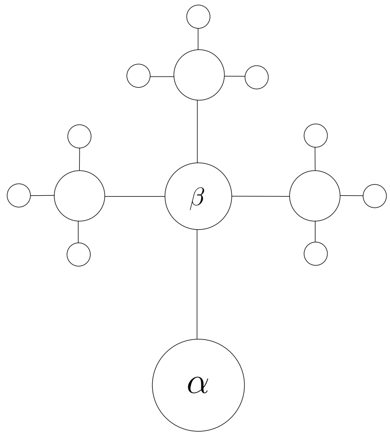

FIGURE 2. The tree T_3 for $k = 3$. Here group elements are represented as circles.

We first analyze finite blocks in X_F by using a combinatorial argument. Fix any $\sigma_i \in \sigma$ and define the set $T_n \subset G$ by

$$T_n = \{e\} \cup \{\sigma_i g : g \in G, |g| \leq n-1\}$$

for $n \geq 1$. T_n can be represented by a subset of the Cayley graph of G. Such a graph is shown in Figure 2. In the graph of T_n, we label the vertex corresponding to e by α and the vertex corresponding to σ_i by β. We call α the "root" of T_n and we say that T_n has *height* n.

DEFINITION 1.1. Let $c_0(n)$ and $c_1(n)$ be the numbers of possible colorings of T_n when α is colored 0 and 1, respectively.

THEOREM 1.2. *$c_0(n)$ satisfies the following recursion relationship: $c_0(1) = 2$, $c_0(2) = 2^k + 1$, and $c_0(n) = [c_0(n-1)]^k + [c_0(n-2)]^{k^2}$ for all $n \geq 2$.*

PROOF. It is easy to see by a simple combinatorial argument that $c_0(1) = 2$ and $c_0(2) = 2^k + 1$.

When α is colored 1, then β must be colored 0. But we can regard β as the "root" of k different trees each of height $n-1$. The colorings of each of the k trees for which β is the "root" can be chosen independently. Thus,

$$c_1(n) = [c_0(n-1)]^k \,. \tag{7}$$

Similarly, when α is colored 0, β can be colored either 0 or 1, so

$$c_0(n) = [c_0(n-1)]^k + [c_1(n-1)]^k \,.$$

By (7), this is the same as

$$c_0(n) = [c_0(n-1)]^k + [c_0(n-2)]^{k^2} \,,$$

which gives the theorem. □

Recall that $C_n = \{g \in G : |g| \leq n\}$ and that B_n is the number of allowed colorings of C_n in X_F. Then we have the following theorem:

THEOREM 1.3. *B_n satisfies $B_n = [c_0(n)]^{k+1} + [c_0(n-1)]^{k(k+1)}$.*

PROOF. Note that C_n consists of $k+1$ trees of height n that all share e as their root. When e is colored 0, there are $c_0(n)$ possible colorings of each of the $k+1$ trees of height n, each of which can be chosen independently, giving a total of $[c_0(n)]^{k+1}$ possible colorings.

Next, when e is colored 1, each of the $k+1$ trees has a root colored 1, and each tree can be colored independently. Therefore, there are $[c_1(n)]^{k+1} = [c_0(n-1)]^{k(k+1)}$ possible colorings of C_n when e is colored 1. Therefore,

$$B_n = [c_0(n)]^{k+1} + [c_0(n-1)]^{k(k+1)}.$$

□

DEFINITION 1.4. Two sequences (a_n) and (b_n) are *asymptotic* if

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 1.$$

We sometimes write this as $a_n \sim b_n$.

For notational simplicity, we define $a_n = c_0(n)$. So that

$$a_{n+1} = a_n^k + a_{n-1}^{k^2}.$$

Note that when $k = 1$, so that the free group under consideration is $\mathbb{Z}$, the a_n are Fibonacci numbers with the standard recursion formula. When $k \neq 1$, the nonlinear recursion sequence a_n will be central for understanding the golden mean shift on the free group. Namely, as we will show for small k, there are constants λ_1 and λ_2 such that $a_n \sim \lambda_1 \exp_2(\lambda_2 k^{n-1})$, where $\exp_2(x) = 2^x$. Theorem 3.4 will show λ_2 can be used to find the entropy of the golden mean shift. The next section will focus on the properties of the sequence (a_n) in preparation for determining its asymptotics.

2. Growth properties of (a_n)

Here we study the general growth properties of (a_n). Specifically we will determine when the limit

$$L = \lim_{n\to\infty} \frac{a_n}{a_{n-1}^k}$$

exists. We will show that the limit exists for sufficiently small k (recall $k = 2q-1$, where q is the number of generators) and equals the solution of $x^{k+1} = x^k + 1$ in the interval $I = [1, 2]$. The original motivation for considering this limit comes from an easy result about Fibonacci numbers: when $k = 1$, the a_n are Fibonacci numbers and the above limit is ϕ, the golden mean, which satisfies $\phi^2 = \phi + 1$. When k is sufficiently large, the limit may not exist but the values of the fraction will oscillate between two limits. In this case, the odd terms and even terms respectively converge to different limits.

We begin by defining the sequence (q_n) which will simplify the notation considerably.

DEFINITION 2.1. For each $n = 2, 3 \ldots$, let $q_n = a_n/a_{n-1}^k$.

We will next prove several basic properties of q_n.

PROPOSITION 2.2. *For each* $n = 2, 3, \ldots$, $q_{n+1} = 1 + 1/q_n^k$.

PROOF.

$$q_{n+1} = \frac{a_{n+1}}{a_n^k} = \frac{a_n^k + a_{n-1}^{k^2}}{a_n^k} = 1 + (\frac{a_{n-1}^k}{a_n})^k = 1 + \frac{1}{q_n^k}.$$

□

PROPOSITION 2.3. *For each $n = 3, 4, \ldots$, we have that $1 < q_n < 2$.*

PROOF. Note that $0 < q_2 = a_2/a_1^k = (2^k + 1)/2$. We will prove the proposition by induction. Suppose $0 < q_n$. Then $1/q_n^k > 0$, so $1 + 1/q_n^k > 1$, which is equivalent to $q_{n+1} > 1$, so inductively we have that $q_n > 1$ for all $n \geq 3$.

Next note $q_2 > 1$, so suppose $q_n > 1$. Then $q_n^k > 1$, so we have that $1/q_n^k < 1$. This implies $1 + 1/q_n^k < 2$, so $q_{n+1} < 2$. □

PROPOSITION 2.4. *For each $k = 2, 3, \ldots$ we have that $q_2 < q_3$ and $q_2 < q_4$.*

PROOF. For the first part, $q_2 < q_3$ only if $a_2^{k+1} < a_1^k a_3$. Now,

$$a_2^{k+1} = (2^k + 1)^{k+1} = (2^k + 1)^k(2^k + 1) = 2^k(2^k + 1)^k + (2^k + 1)^k,$$

and

$$a_1^k a_3 = a_1^k(a_2^k + a_1^{k^2}) = 2^k(2^k + 1)^k + 2^{k^2+k}.$$

This shows $a_2^{k+1} < a_1^k a_3$ only if $(2^k + 1)^k < 2^{k^2+k}$. But $(2^k + 1)^k < (2^{k+1})^k = 2^{k^2+k}$, so we must have $q_2 < q_3$.

Note $q_2 < q_4$ only if

$$a_2 a_3^k < a_4 a_1^k.$$

We have that

$$a_2 a_3^k = (2^k + 1)a_3^k = 2^k a_3^k + a_3^k$$

and

$$a_4 a_1^k = (a_3^k + a_2^{k^2})2^k = 2^k a_3^k + 2^k a_2^{k^2},$$

so $q_2 < q_4$ only if $a_3^k < 2^k a_2^{k^2}$. This is the same as requiring that

$$a_3 < 2a_2^k,$$

which is the same as

$$a_2^k + a_1^{k^2} < 2a_2^k.$$

This simplifies to $a_1^{k^2} < a_2^k$, or equivalently $2^{k^2} < (2^k + 1)^k$, which is obviously true. □

PROPOSITION 2.5. *For each $n = 2, 3, \ldots$ we have*

$$(i)\ q_{n+1} < q_{n-1} \quad \text{implies } q_n < q_{n+2}$$
$$(ii)\ q_{n+1} > q_{n-1} \quad \text{implies } q_n > q_{n+2}.$$

PROOF. We will prove only (i), as the proof for (ii) is analogous. Note that $q_{n+1} < q_{n-1}$ if and only if

$$\frac{a_{n+1}}{a_n^k} < \frac{a_{n-1}}{a_{n-2}^k},$$

which is equivalent to

$$a_{n-2}^k a_{n+1} < a_{n-1} a_n^k.$$

Raising each side to the k'th power gives

$$a_{n-2}^{k^2} a_{n+1}^k < a_{n-1}^k a_n^{k^2}.$$

Adding $a_{n-1}^k a_{n+1}^k$ to both sides and factoring gives

$$(a_{n-1}^k + a_{n-2}^{k^2}) a_{n+1}^k < a_{n-1}^k (a_{n+1}^k + a_n^{k^2}).$$

Since $a_{n-1}^k + a_{n-2}^{k^2} = a_n$ and $a_{n+1}^k + a_n^{k^2} = a_{n+2}$, this is equivalent to

$$a_n a_{n+1}^k < a_{n-1}^k a_{n+2},$$

which shows

$$\frac{a_{n+1}^k}{a_{n+2}} < \frac{a_{n-1}^k}{a_n}.$$

This is the same as

$$\frac{a_{n+2}}{a_{n+1}^k} > \frac{a_n}{a_{n-1}^k},$$

or equivalently $q_{n+2} > q_n$. □

We can use this property to show an unusual characteristic of (q_n). Let $E_n = q_{2n}$ and $O_n = q_{2n+1}$ for $n = 1, 2, \ldots$ define the subsequences of (q_n) which consist of the even- and odd-numbered terms of (q_n) respectively. Then we have,

PROPOSITION 2.6. *(E_n) increases monotonically to a limit $L_E \in I$ and (O_n) decreases monotonically to a limit $L_O \in I$.*

PROOF. We will prove this inductively. Proposition 2.4 shows that $q_2 < q_4$. So suppose that n is odd and $q_n < q_{n+2}$. Then (*ii*) of Proposition 2.5 implies $q_{n+1} > q_{n+3}$. But then (*i*) of Proposition 2.5 implies $q_{n+2} < q_{n+4}$, which shows that (E_n) is increasing. We can apply (*ii*) of Proposition 2.5 to show that (O_n) is decreasing. The fact that (E_n) and (O_n) converge is immediate from the fact that they are monotone and bounded in I. □

To summarize, we are interested in the existence of the limit

$$\lim_{n\to\infty} \frac{a_n}{a_{n-1}^k} = \lim_{n\to\infty} q_n.$$

We have shown that the even and odd terms of q_n always converge to some limits L_E and L_O respectively, each in the interval $I = [1, 2]$. The remainder of this section will be concerned with determining when $L_E = L_O$. This matter is not important to the entropy calculation, but is interesting because it turns out that $L_E = L_O$ for sufficiently small k, but not for large k. This result, perhaps, is counterintuitive because there does not exist an obvious reason why the limit should cease to exist for some large k, when it exists for small k.

To prove this, we study the iterates of a one-dimensional map f:

DEFINITION 2.7. Define $f : [1, 2] \to [1, 2]$ by $f(x) = 1 + x^{-k}$.

f is a useful function to define because $f(q_n) = q_{n+1}$ by Proposition 2.2. Note that $f(1) = 2 > 1$ and $f(2) = 1 + 2^{-k} < 2$, so there must be a fixed point $\alpha \in I = [1, 2]$. However, since $f'(x) < 0$ for all $x \in I$, α must be the unique fixed point in I.

Definition 2.8. Let α be the unique fixed point of f in I.

Note that α satisfies $\alpha^{k+1} = \alpha^k + 1$, since $f(\alpha) = 1 + \alpha^{-k} = \alpha$. Notice also that $f' < 0$, so f maps numbers less than α to numbers greater than α and vice versa.

Lemma 2.9. *For each $n = 2, 3, \ldots$, we have,*

$$(i)\ q_n > q_{n+1} \quad \text{implies } q_{n+1} < q_{n+2}$$
$$(ii)\ q_n < q_{n+1} \quad \text{implies } q_{n+1} > q_{n+2}.$$

Proof. We will prove only (i), as the proof for (ii) is analogous. Note that $q_n > q_{n+1}$ implies $q_n^k > q_{n+1}^k$, which shows $1/q_n^k < 1/q_{n+1}^k$. This implies $1 + 1/q_n^k < 1 + 1/q_{n+1}^k$, which is equivalent to $q_{n+1} < q_{n+2}$ □

Theorem 2.10. *$L_E \in [1, \alpha]$ and $L_O \in [\alpha, 2]$.*

Proof. We prove only the first part since the second is analogous. Because $f(x) > x$ if and only if $x < \alpha$, it suffices to prove that for n even, $q_n < f(q_n) = q_{n+1}$. We will prove this inductively. Proposition 2.4 shows $q_2 < q_3$. Suppose that $q_n < q_{n+1}$ with n odd. Then by (ii) of Lemma 2.9 we have that $q_{n+1} > q_{n+2}$. Then by (i) of Lemma 2.9 we have that $q_{n+2} < q_{n+3}$, proving the theorem. □

Next, to simplify notation, we also define a function g:

Definition 2.11. Define $g : [1, 2] \to [1, 2]$ by $g(x) = f^2(x) = f(f(x))$.

Note that $g(q_n) = q_{n+2}$, so that $g(O_n) = O_{n+1}$ and $g(E_n) = E_{n+1}$. This implies that the (possibly non-distinct) fixed points of g are L_E, L_O, and α.

2.1. The behavior of g for small k. In this section we will investigate the behavior of g' and determine for which $k \geq 2$ we have $\sup\{|g'(x)| : x \in [1, 2]\} = C < 1$. For these k, we will apply the contraction mapping theorem to conclude that g has a unique fixed point. Since α is a fixed point of g, we can conclude that α is the only fixed point of g and therefore we must have $L_E = L_O = \alpha$.

First note that

$$g'(x) = \frac{k^2}{x^{k+1}(1 + x^{-k})^{k+1}},$$

and also, after simplifying,

$$g''(x) = \frac{-k^2(k+1)x^{k^2-2}(1 - k + x^k)}{(1 + x^k)^{k+2}}.$$

We are interested in the maximum value that $g'(x)$ attains on $I = [1, 2]$, which will either be attained at a zero of $g''(x)$ or at one of the endpoints of

I. $g''(x) = 0$ only when $(1 - k + x^k) = 0$, or, equivalently, at $x = \sqrt[k]{k-1}$. Evaluating these possible maxima gives

$$g'(1) = \frac{k^2}{2^{k+1}},$$

$$g'(2) = \frac{k^2}{2^{k+1}(1 + 2^{-k})^{k+1}},$$

and

$$\begin{aligned} g'(\sqrt[k]{k-1}) &= \frac{k^2}{(k-1)^{1+1/k}(1 + (k-1)^{-1})^{k+1}} \\ &= \frac{k^2(k-1)^{k+1}}{(k-1)^{1+1/k}k^{k+1}} \\ &= \frac{(k-1)^{k-1/k}}{k^{k-1}}. \end{aligned}$$

Thus, for all $k \geq 2$ we have that $1 > |g'(1)| > |g'(2)|$. Note also that $g''(1) > 0$ and $g''(2) < 0$, so we know that $x = \sqrt[k]{k-1}$ is a maximum for g'. It is not immediately obvious for which values of k we have $g'(\sqrt[k]{k-1}) < 1$. This is equivalent to requiring

$$(k-1)^{k-1/k} < k^{k-1},$$

which can be simplified to

$$(k-1)^{k+1} < k^k$$

by raising each side to the $k/(k-1)$'th power. Still this inequality is not easy to solve analytically. One can numerically solve $(k-1)^{k+1} = k^k$ by using Mathematica's FindRoot, which implements Newton's method. This numerical approximation gives $k \approx 4.14104$.

Definition 2.12. Let $\Omega \approx 4.14104$ be defined by

$$\Omega = \inf\{k : (k-1)^{k+1} < k^k\}.$$

Proposition 2.13. *For $k > \Omega$ we have $(k-1)^{k+1} > k^k$ and for $k < \Omega$ we have $(k-1)^{k+1} < k^k$.*

Proof. Note $(2-1)^{2+1} < 2^2$, so we can prove the theorem by showing that the equation $(x-1)^{x+1} = x^x$ has one solution. Let

$$h(x) = \frac{(x-1)^{x+1}}{x^x} = \left(1 - \frac{1}{x}\right)^x (x-1)$$

for $x > 1$. Then $h(x) \to 0$ as $x \to 1^+$ and $h(x) \to \infty$ as $x \to \infty$ since $(1 - 1/x)^x \to 1/e$ as $x \to \infty$. However, it is easy to check that

$$\frac{dh(x)}{dx} > 0$$

for all $x > 1$, so h must have a unique point for which $h(x) = 1$. This implies that there is exactly one value of k which satisfies $(k-1)^{k+1} = k^k$, and that this value of k must equal Ω. □

Thus, for $k < \Omega$ we know that $\sup\{g'(x) : x \in [1,2]\} = C < 1$ for all x. Therefore, for $k < \Omega$ the Mean Value Theorem implies that for all $x, y \in I$ we have

$$|g(x) - g(y)| = |g'(c)|\,|x - y| \le C\,|x - y|$$

for some $x \in I$. Thus, when $k < \Omega$ the Contraction Mapping Theorem implies that there will be a unique fixed point in I:

THEOREM 2.14 (Contraction mapping theorem). *Suppose X is a complete metric space and $f : X \to X$ has the property that there exists $C \in \mathbb{R}$ with $0 \le C < 1$ such that*

$$|f(x) - f(y)| \le C\,|x - y|$$

for all $x, y \in X$. Then there exists a unique fixed point $\alpha \in X$ such that $f(\alpha) = \alpha$, and $f^n(x_0) \to \alpha$ for all $x_0 \in X$.

Since for $k < \Omega$ there is only one fixed point in I and the sequences (E_n) and (O_n) both converge, we must have $L_E = L_O = \alpha$. For $k < \Omega$, we know that the limit

$$\lim_{n\to\infty} q_n = \lim_{n\to\infty} \frac{a_n}{a_{n-1}^k} = \alpha,$$

where α is the root of

$$\alpha^{k+1} - \alpha - 1 = 0 \tag{8}$$

in $I = [1,2]$. The fact that there is a single fixed point for $k < \Omega$ implies that there is exactly one α in I which satisfies (8). It is not clear what happens when $k = \Omega$ since in that case, $g'(x) = 1$ for exactly one point and we can no longer apply the contraction mapping theorem.

2.2. The behavior of g for large k. Here we will show that for $k > \Omega$, there are two distinct fixed points besides α, one strictly less than α and one strictly greater than α. Therefore, L_E and L_O are different, and neither equals α, which implies the limit $\lim_{n\to\infty} a_n/a_{n-1}^k = \lim_{n\to\infty} q_n$ does not exist.

LEMMA 2.15. *$k > \Omega$ implies $\alpha < \sqrt[k]{k-1}$.*

PROOF. Note that $\alpha < x$ implies $f(x) < x$, since f is decreasing and α is the only fixed point of f. Therefore, to check the proposition we must check whether $f(\sqrt[k]{k-1}) < \sqrt[k]{k-1}$. Note

$$f(\sqrt[k]{k-1}) = 1 + \frac{1}{k-1} = \frac{k}{k-1}.$$

Now,

$$\frac{k}{k-1} < \sqrt[k]{k-1}$$

if and only if

$$k^k < (k-1)^{k+1},$$

which is true only when $k > \Omega$ by Proposition 2.13. ☐

PROPOSITION 2.16. *$k > \Omega$ implies $g'(\alpha) > 1$.*

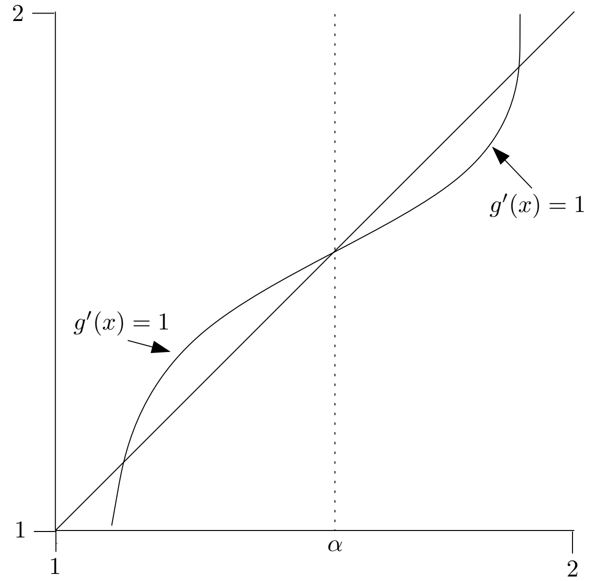

FIGURE 3. Plot of $g(x)$ for the proof of Theorem 2.17.

PROOF. $g'(\alpha) = f'(f(\alpha))f'(\alpha) = [f'(\alpha)]^2 = k^2(\alpha^2)^{-k-1} > 1$ if and only if $k^2 > (\alpha^2)^{k+1}$, or, equivalently, $k > \alpha^{k+1}$. But $\alpha^{k+1} = \alpha^k + 1$, so $k > \alpha^{k+1}$ only if $\sqrt[k]{k-1} > \alpha$. By Lemma 2.15 this is true since $k > \Omega$. $\square$

This proposition implies that α is a repelling fixed point for g when $k > \Omega$.

THEOREM 2.17. *For $k > \Omega$ there exist exactly three values in $I = [1, 2]$ for which $g(x) = x$. One of these fixed points is α, one is strictly less than α, and one is strictly greater than α.*

PROOF. We have shown that g must have fixed points L_O and L_E such that $E_n \uparrow L_E$ and $O_n \downarrow L_O$. For $k > \Omega$, Proposition 2.16 shows that α is a repelling fixed point, which implies that we cannot have $L_E = \alpha$ or $L_O = \alpha$ because L_E and L_O are the limit points of $g^n(q_2)$ and $g^n(q_3)$ respectively.

To see that there cannot be more than three fixed points, note $g'(1) < 1$ and $g'(2) < 1$. Since g'' has only one zero in I, this implies that $g'(x) = 1$ for at most two values in I. Therefore, g can cross the line $y = x$ at most three times (see Figure 3). $\square$

Thus, when $k > \Omega$ the limit

$$\lim_{n\to\infty} \frac{a_n}{a_{n-1}^k}$$

does not exist and values of $a_n/a_{n-1}^k = q_n$ oscillate between two limits points, L_E and L_O.

Figure 4 shows graphs of x, $g(x)$ and $f^{50}(x)$ for $k = 3$ and $k = 5$. Note that for $k = 3 < \Omega \approx 4.141$, $f^{50}(x)$ seem to take one value, which equals α. However, for $k = 5 > \Omega$, $f^{50}(x)$ seems to take two values, which equal L_E and L_O.

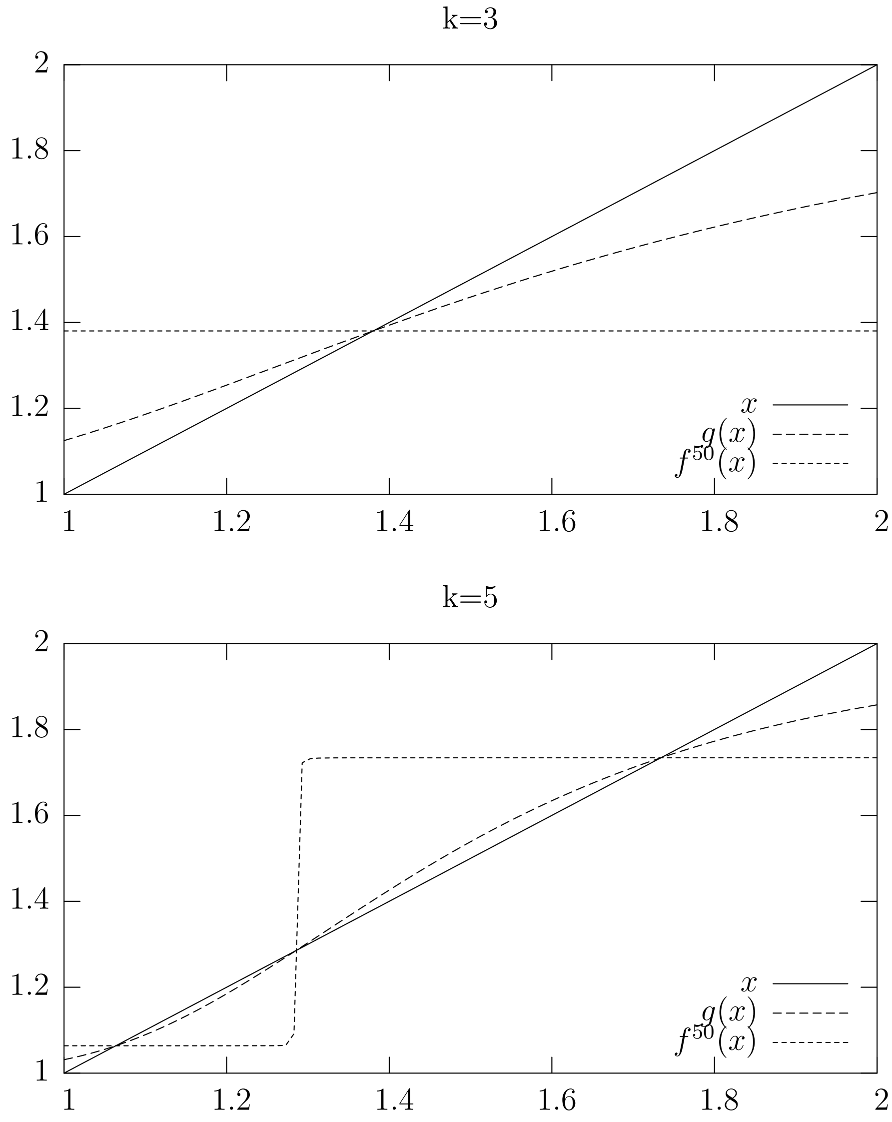

FIGURE 4. Figures showing x, $g(x)$ and $f^{50}(x)$.

We can give another characterization of α, L_E and L_O by noting that $g(x) = x$ only when

$$x = 1 + (1 + x^{-k})^{-k},$$

or, equivalently,

$$(x - 1)(1 + x^{-k})^k - 1 = 0,$$

which can be simplified to

$$(x - 1)(x^k + 1)^k - x^{k^2} = 0.$$

This is a polynomial with integral coefficients, and the fixed points of g correspond to the zeros of this polynomial. The above discussion shows that when $k < \Omega$ this polynomial has exactly one root in I. Similarly, when $k > \Omega$ this polynomial has exactly 3 roots in I.

3. The asymptotics of (a_n)

We are now ready to determine the asymptotics of (a_n), in which the limits L_E and L_O will arise naturally.

PROPOSITION 3.1. *For each* $n = 2, 3, \ldots$ *we have*

$$a_n = 2^{k^{n-1}} \prod_{j=2}^{n} q_j^{k^{n-j}}.$$

PROOF. We prove this proposition inductively. Recall that $q_n = a_n / a_{n-1}^k$, so $a_n = q_n a_{n-1}^k$. Note that the theorem is satisfied for $n = 2$, since $a_2 = q_2 a_1^k = 2^k q_2$. Now we suppose the proposition is true for n and prove it for $n+1$:

$$\begin{aligned} a_{n+1} = q_{n+1} a_n^k = q_{n+1} \left[2^{k^{n-1}} \prod_{j=2}^{n} q_j^{k^{n-j}} \right]^k &= q_{n+1} 2^{k^n} \prod_{j=2}^{n} q_j^{k^{n-j+1}} \\ &= 2^{k^n} \prod_{j=2}^{n+1} q_j^{k^{n+1-j}}. \end{aligned}$$

□

Note that Proposition 3.1 implies

$$\begin{aligned} a_n &= 2^{k^{n-1}} \prod_{j=2}^{n} q_j^{k^{n-j}} \\ &= 2^{k^{n-1}} \exp_2 \left(\sum_{j=2}^{n} \log_2 q_j^{k^{n-j}} \right) \\ &= 2^{k^{n-1}} \exp_2 \left(\sum_{j=2}^{n} k^{n-j} \log_2 q_j \right) \\ &= 2^{k^{n-1}} \exp_2 \left(k^n \sum_{j=2}^{n} \frac{\log_2 q_j}{k^j} \right) \\ &= \exp_2 \left(k^{n-1} + k^n \sum_{j=2}^{n} \frac{\log_2 q_j}{k^j} \right) \\ &= \exp_2 \left[k^{n-1} \left(1 + k \sum_{j=2}^{n} \frac{\log_2 q_j}{k^j} \right) \right]. \end{aligned}$$

Next we define the sum

$$A_n = 1 + k \sum_{j=2}^{n} \frac{\log_2 q_j}{k^j},$$

so that

$$a_n = \exp_2(k^{n-1}A_n).$$

Note that

$$A_n = 1 + k\frac{\log_2 q_2}{k^2} + k\sum_{j=3}^{n}\frac{\log_2 q_j}{k^j}$$

so that

$$A_n \leq 1 + k\frac{\log_2 q_2}{k^2} + k\sum_{j=3}^{n}\frac{1}{k^j} < \infty,$$

since $1 < q_n < 2$ for $n = 3, 4, \ldots$. This implies A_n must converge to some value A given by

$$A = \lim_{n\to\infty} A_n = 1 + k\sum_{j=2}^{\infty} k^{-j}\log_2 q_j.$$

Thus, it is reasonable to guess that a_n is asymptotic to $\lambda_1 \exp_2(k^{n-1}A)$ for some λ_1.

THEOREM 3.2. *For $k = 2, 3, \ldots$ and n even, we have $a_n \sim \lambda_1 \exp_2(k^{n-1}\lambda_2)$ with*

$$\lambda_1 = \exp_2\left(-\frac{\log_2 L_E + k\log_2 L_O}{k^2-1}\right) \qquad \textit{and} \qquad \lambda_2 = A,$$

where A is as above and L_O and L_E are the limits of (O_n) and (E_n) respectively.

PROOF. We will prove this by showing that

$$\lim_{n\to\infty}\frac{\exp_2(k^{n-1}A)}{a_n} = \lim_{n\to\infty}\frac{\exp_2(k^{n-1}A)}{\exp_2(k^{n-1}A_n)} = \exp\left(\frac{\log_2 L_E + k\log_2 L_O}{k^2-1}\right).$$

Note that

$$\lim_{n\to\infty}\frac{\exp_2(k^{n-1}A)}{\exp_2(k^{n-1}A_n)} = \exp_2\left(k^{n-1}(A-A_n)\right).$$

We define

$$\begin{aligned}
W_n &= k^{n-1}(A - A_n)\\
&= k^{n-1}\left(k\sum_{j=2}^{\infty}k^{-j}\log_2 q_j - k\sum_{j=2}^{n}k^{-j}\log_2 q_j\right)\\
&= k^n\sum_{j=n+1}^{\infty}k^{-j}\log_2 q_j\\
&= \sum_{j=n+1}^{\infty}k^{n-j}\log_2 q_j\\
&= \sum_{j=1}^{\infty}k^{-j}\log_2 q_{j+n}.
\end{aligned}$$

For all $r \geq 3$, we have $1 < q_r < 2$ so $\log_2 q_r > 0$. This implies that we can rearrange the sum in W_n, splitting it into even and odd terms to give

$$W_n = \left[\sum_{j=1}^{\infty} k^{-2j} \log_2 q_{2j+n} + \sum_{j=1}^{\infty} k^{-2j+1} \log_2 q_{(2j-1)+n}\right]. \tag{9}$$

We will show that $W_n \to (\log_2 L_E + k \log_2 L_O)/(k^2 - 1)$. Fix $\epsilon > 0$. We have already shown that $(O_n) \to L_O$ and $(E_n) \to L_E$. The log function is continuous, so this implies $(\log_2 O_n) \to \log_2 L_O$ and $(\log_2 E_n) \to \log_2 L_E$. So choose N such that for all $n > N$ we have

$$|\log_2 O_n - \log_2 L_O| < \epsilon \text{ and } |\log_2 E_n - \log_2 L_E| < \epsilon.$$

Then we know that for n even,

$$\begin{aligned}
&\left|W_n - \left((\log_2 L_E)\sum_{j=1}^{\infty} k^{-2j} + (\log_2 L_O)\sum_{j=1}^{\infty} k^{-2j+1}\right)\right| \\
&\quad = \left|\sum_{j=1}^{\infty} k^{-2j}(\log_2 q_{2j+n} - \log_2 L_E) + \sum_{j=1}^{\infty} k^{-2j+1}(\log_2 q_{2j-1} - \log_2 L_O)\right| \\
&\quad < \epsilon \sum_{j=1}^{\infty} k^{-j} \leq \epsilon.
\end{aligned} \tag{10}$$

Therefore for n even,

$$\begin{aligned}
W_n &\to (\log_2 L_E)\sum_{j=1}^{\infty} k^{-2j} + (\log_2 L_O)\sum_{j=1}^{\infty} k^{-2j+1} \\
&= (\log_2 L_E)\frac{1}{k^2 - 1} + (\log_2 L_O)\frac{k}{k^2 - 1} \\
&= \frac{\log_2 L_E + k \log_2 L_O}{k^2 - 1}.
\end{aligned}$$

Since the exponential function is continuous, this implies

$$\begin{aligned}
\frac{\exp_2(k^{n-1}A)}{\exp_2(k^{n-1}A_n)} &= \exp_2(k^{n-1}(A - A_n)) \\
&= \exp_2(W_n) \to \exp_2\left(\frac{\log L_E + k \log L_O}{k^2 - 1}\right),
\end{aligned}$$

proving the theorem. ☐

THEOREM 3.3. *For $k = 2, 3, \ldots$ and n odd, we have $a_n \sim \lambda_1 \exp_2(k^{n-1}\lambda_2)$ with*

$$\lambda_1 = \exp_2\left(-\frac{\log_2 L_O + k \log_2 L_E}{k^2 - 1}\right) \quad \text{and} \quad \lambda_2 = A,$$

where A is as above and L_E and L_O are the limits of (O_n) and (E_n) respectively.

PROOF. The proof is identical to the proof of the previous theorem, except that for odd n, L_E and L_O are switched in (10) and every step after. ☐

Note for n even we have,

$$\lambda_1 = \exp_2\left(-\frac{\log_2 L_E + k\log_2 L_O}{k^2-1}\right)$$

and for n odd we have,

$$\lambda_1 = \exp_2\left(-\frac{\log_2 L_O + k\log_2 L_E}{k^2-1}\right).$$

When $k < \Omega$, we know that $L_E = L_O = \alpha$, so these expressions both simplify to

$$\lambda_1 = \exp_2\left(-\frac{(k+1)\log_2\alpha}{k^2-1}\right) = \exp_2\left(-\frac{\log_2\alpha}{k-1}\right) = \alpha^{\frac{1}{1-k}}.$$

Thus, for $k < \Omega$ we know $a_n \sim \lambda_1 \exp_2(\lambda_2 k^{n-1})$ for all n, where $\lambda_1 = \alpha^{1/(1-k)}$.

THEOREM 3.4. *The entropy of the golden mean shift is given by $h(\mathsf{X}_F) = \lambda_2(k-1)/k$.*

PROOF. To prove the theorem we show that if $a_n \sim \lambda_1 \exp_2(\lambda_2 k^{n-1})$, then

$$\lim_{n\to\infty}\frac{\log_2 B_n}{|C_n|} = \lambda_2\frac{k-1}{k}.$$

Since λ_1 does not appear in this expression, the entropy of the golden mean shift is independent of the value of λ_1. Therefore, it is irrelevant to the entropy calculation that λ_1 may depend on whether we look at the even or the odd terms of a_n.

In the free group, the number of elements $g \in G$ such that $|g| = n$ is given by $(k+1)k^{n-1}$ for $n \geq 1$. Therefore,

$$|C_n| = 1 + \sum_{j=1}^{n}(k+1)k^{j-1} = 1 + (k+1)\frac{k^n-1}{k-1}.$$

Next,

$$\lim_{n\to\infty}\frac{\log_2 B_n}{|C_n|} = \lim_{n\to\infty}\frac{\log_2\left[a_n^{k+1} + a_{n-1}^{k(k+1)}\right]}{|C_n|}.$$

Substituting for a_n and simplifying gives

$$\lim_{n\to\infty}\frac{\log_2\left[\lambda_1^{k+1}\exp_2((k+1)\lambda_2 k^{n-1})u_n + \lambda_1^{k(k+1)}\exp_2(\lambda_2(k+1)k^{n-1})u_{n-1}\right]}{|C_n|},$$

where $u_n \to 1$. Factoring and simplifying gives,

$$\begin{aligned}
\lim_{n\to\infty} \frac{\log_2 B_n}{|C_n|} &= \lim_{n\to\infty} \frac{(k+1)\lambda_2 k^{n-1} + \log_2\left[\lambda_1^{k+1} u_n + \lambda_1^{k(k+1)} u_{n-1}\right]}{|C_n|} \\
&= \lim_{n\to\infty} \frac{(k+1)k^{n-1}\lambda_2}{1+(k+1)\frac{k^n-1}{k-1}} \\
&= \lim_{n\to\infty} \frac{(k+1)k^{n-1}(k-1)\lambda_2}{(k-1)+(k+1)(k^n-1)} \\
&= \lambda_2 \frac{k-1}{k}.
\end{aligned}$$

□

Theorem 3.4 shows that the entropy of the golden mean shift on the free group is therefore given by

$$h(\mathsf{X}_F) = \lambda_2 \frac{k-1}{k} = A\frac{k-1}{k} = \frac{k-1}{k}\left(1 + k\sum_{j=2}^{\infty} k^{-j}\log_3 q_j\right). \tag{11}$$

The next section will determine this value numerically for various k. Finding a closed form of $h(\mathsf{X}_F)$ is still an open problem. However, we did find a closed form for λ_1, which implies that perhaps it will be easier in other problems, such as the square ice problem, to find the constant in front of the exponential, rather than the actual entropy.

4. Numerical computations related to the golden mean shift

In this section we present several tables of numerical results related to the golden mean shift on the free group. All the following results were computed using a Mathematica script written by the author.

The following shows values of a_n for $k = 2, 3, 4$.

	$k=2$	$k=3$	$k=4$
a_1	2.000000000	2.000000000	2.00000000
a_2	5.000000000	9.000000000	17.00000000
a_3	41.00000000	1241.000000	149057.0000
a_4	2306.000000	$2.298661010 \times 10^{9}$	$5.423002381 \times 10^{20}$
a_5	$8.143397000 \times 10^{6}$	$1.912721905 \times 10^{28}$	$1.458683650 \times 10^{83}$
a_6	$9.459216733 \times 10^{13}$	$8.789440240 \times 10^{84}$	$5.086901266 \times 10^{332}$
a_7	$1.334534603 \times 10^{28}$	$1.021684124 \times 10^{255}$	$1.089720930 \times 10^{1331}$
a_8	$2.581592044 \times 10^{56}$	$1.379550025 \times 10^{765}$	$1.611162872 \times 10^{5324}$
a_9	$9.836516530 \times 10^{112}$	$3.838471331 \times 10^{2295}$	$1.069248835 \times 10^{21297}$
a_{10}	$1.411741836 \times 10^{226}$	$7.465377822 \times 10^{6886}$	$1.513291731 \times 10^{85188}$
a_{11}	$2.929207829 \times 10^{452}$	$5.969536653 \times 10^{20660}$	$8.163515502 \times 10^{340752}$
a_{12}	$1.255236735 \times 10^{905}$	$2.847487925 \times 10^{61982}$	$5.197706700 \times 10^{1363011}$
a_{13}	$2.311827621 \times 10^{1810}$	$3.271440374 \times 10^{185947}$	$1.118949619 \times 10^{5452047}$

The following shows values of $h(\mathsf{X}_F)$ for $k = 2, \ldots, 10$. These values were computed by computing the first 8 terms of (11). The error was computed by finding an upper bound for the remaining terms in (11).

	$h(\mathsf{X}_F)$	error
$k=2$	0.7320165404	± 0.00492914
$k=3$	0.7748023747	± 0.000144245
$k=4$	0.8096298193	± 0.0000128363
$k=5$	0.8372456350	$\pm 2.01898 \times 10^{-6}$
$k=6$	0.8587918886	$\pm 4.50767 \times 10^{-7}$
$k=7$	0.8757141659	$\pm 1.27686 \times 10^{-7}$
$k=8$	0.8892044026	$\pm 4.29787 \times 10^{-8}$
$k=9$	0.9001414916	$\pm 1.6489 \times 10^{-8}$
$k=10$	0.9091550972	$\pm 7.01033 \times 10^{-9}$

These data suggest the following conjecture, which we make no attempt to prove here:

CONJECTURE 4.1. *As $q \to \infty$, $h(\mathsf{X}_F) \to 1$, where X_F is the golden mean shift on the free group on q generators.*

The following shows the real roots in $I = [1,2]$ of $f(x) = 1 + x^{-k}$ for $k = 2, \ldots, 10$. Note that where there is more than one root in I, the lesser root is L_E and the greater is L_O.

	Root(s) of f
$k=2$	1.465571232
$k=3$	1.380277569
$k=4$	1.324717957
$k=5$	1.063770006, 1.734110265
$k=6$	1.023326214, 1.870793951
$k=7$	1.009904994, 1.933332438
$k=8$	1.004504650, 1.964682475
$k=9$	1.002127834, 1.981051658
$k=10$	1.001027902, 1.989778850

These data suggest the following conjecture, which we make no attempt to prove here:

CONJECTURE 4.2. *As $q \to \infty$, $L_E \to 1$ and $L_O \to 2$.*

CHAPTER 5

Further topics

1. Nonlinear continued fractions

We showed in Section 2 of Chapter 4 that iterated map

$$f(x) = 1 + x^{-k}$$

on $[1, 2]$ has a single fixed point for small enough k, but three fixed points for large k. This is interesting because iterations $f(x), f^2(x), f^3(x), \ldots$ can be expanded into a *nonlinear continued fraction.* For example,

$$f^3(x) = 1 + \cfrac{1}{\left(1 + \cfrac{1}{\left(1 + \cfrac{1}{\left(1 + \cfrac{1}{x^k}\right)^k}\right)^k}\right)^k}$$

which appears similar to a standard continued fraction, except that each denominator has a k'th power. We showed that for small k, the value of the infinite continued fraction given by

$$\lim_{n\to\infty} f^n(x)$$

is equal to a root α of $y^{k+1} - y^k = 1$, where the value of α is independent of x. However, for large enough k, the value of this limit will oscillate between two limit points, indicating that the continued fraction is not well-defined. This suggests that perhaps such "nonlinear continued fractions" behave in suprising and interesting ways. Here we present a preliminary analysis of nonlinear continued fractions.

2. The nonlinear Gauss map

We begin by defining a nonlinear version of the Gauss map:

Definition 2.1. For a fixed $k \geq 2$, define the k-Gauss map $\phi_k : [0, 1] \to [0, 1]$ by

$$\phi_k(x) = \frac{1}{\sqrt[k]{x}} - \left\lfloor \frac{1}{\sqrt[k]{x}} \right\rfloor .$$

The k-Gauss map partitions the unit interval into an infinite collection of intervals $I_1, I_2, \ldots$ where

$$I_n = \left(\frac{1}{(n-1)^k}, \frac{1}{n^k} \right].$$

Each I_n is mapped by ϕ_k bijectively onto the unit interval. As in the typical continued fraction case where $k = 1$, given any $x \in [0,1]$ we can record which intervals iterates of ϕ_k map x into:

Definition 2.2. For each $x \in [0,1]$, the *sequence of convergents* is the sequence of integers (a_m) such that $\phi_k^m(x) \in I_{a_m}$ for all $m \in \mathbb{N}$.

In analogy to the typical definition of continued fraction we may write

$$x = [a_1, a_2, a_3, a_4, \ldots]_k = \frac{1}{\left(a_1 + \left(\frac{1}{a_2 + \left(\frac{1}{a_3 + \left(\frac{1}{a_4 + \ldots} \right)} \right)^k} \right)^k \right)^k}.$$

For the case where $k = 1$, it can easily be proved that each x gives a unique sequence (a_m); however, for $k > \Omega$, different values $x, y \in [0,1]$ will give the same sequence (a_m).

3. When continued fractions are not well-defined

Note that on each I_n there will be a single fixed point. We will consider only the fixed point in I_1, which will be designated β. β is given by

$$\beta = \frac{1}{\sqrt[k]{\beta}} - 1,$$

or equivalently

$$(\beta + 1)^k = \frac{1}{\beta}.$$

Then we have the following:

Lemma 3.1. *When $k > \Omega$ we have $\beta > 1/(k-1)$.*

Proof. It should be clear that $1/(k-1) \in I_1$ since $1/2^k \leq 1/(k-1) \leq 1$ for $k \geq 2$. Since ϕ_k is monotonically decreasing on I_1, we can check this lemma by proving that $\phi_k(1/(k-1)) > 1/(k-1)$. Note that since $1/(k-1) \in I_1$,

$$\phi_k \left(\frac{1}{k-1} \right) > \frac{1}{k-1}$$

when

$$\begin{aligned} \sqrt[k]{k-1}-1 &> \frac{1}{k-1} \\ \sqrt[k]{k-1} &> \frac{k}{k-1} \\ k-1 &> \frac{k^k}{(k-1)^k} \\ (k-1)^{k+1} &> k^k, \end{aligned}$$

which is true since $k > \Omega$. ☐

THEOREM 3.2. *For $k > \Omega$, there exist $x, y \in [0,1]$ such that $x \neq y$ but $\phi_k^m(x) \in I_1$ and $\phi_k^m(y) \in I_1$ for all $m \in \mathbb{N}$.*

PROOF. We show this by proving that for $k > \Omega$, β is an attracting fixed point of ϕ_k. When β is an attracting fixed point, there will exist an interval $J = (\beta - \delta, \beta + \delta) \subset I_1$ such that $\phi_k(J) \subset J$. This implies for all $x \in J$, we have that $\phi_k^m(x) \in I_1$ for all $m \in \mathbb{N}$.

To see that for $k > \Omega$, β is an attracting fixed point, note,

$$\phi_k'(x) = \frac{-1}{k} x^{-1/k-1} = \frac{-1}{k} x^{-(k+1)/k}.$$

Next, $|\phi_k'(\beta)| < 1$ when

$$\begin{aligned} \frac{1}{k}\beta^{-(k+1)/k} &< 1 \\ \beta^{-(k+1)} &< k^k \\ \left(\frac{1}{\beta}\right)^{k+1} &< k^k. \end{aligned}$$

Since $1/\beta = (\beta+1)^k$, this is equivalent to requiring

$$(\beta+1)^{k+1} < k$$

or

$$(\beta+1)^k + \beta(\beta+1)^k < k. \tag{12}$$

We know by Lemma 3.1 that for $k > \Omega$ we have $\beta > 1/(k-1)$, so

$$1/\beta = (\beta+1)^k < k-1.$$

Also,

$$k - 1 + \beta(\beta+1)^k = k$$

since

$$(\beta+1)^k = 1/\beta.$$

By (12), these imply,

$$(\beta+1)^k + \beta(\beta+1)^k < k - 1 + \beta(\beta+1)^k = k,$$

which gives the theorem. ☐

Therefore, we have shown that for $k > \Omega$, the representation as nonlinear continued fractions is not unique: two different values may have the same fraction representation. Next, we offer the following conjecture, but make no attempt here to prove it:

CONJECTURE 3.3. *Suppose $k < \Omega$. Then given $x, y \in [0,1]$ with $x \neq y$, there exist $N, a, b \in \mathbb{N}$ such that $\phi_k^N(x) \in I_a$ and $\phi_k^N(y) \in I_b$ but $a \neq b$.*

This conjecture is equivalent to saying that numbers are uniquely represented by the sequence (a_m) of convergents, and therefore the nonlinear continued fraction representation is well-defined. The next sections will explore nonlinear continued fractions and assume that Conjecture 3.3 is true.

4. Conjugacy between (I, ϕ_k) and (I, ϕ_1)

Note that ϕ_k acts as a shift on $[a_1, a_2, a_3, \ldots]_k$ since

$$\phi_k([a_1, a_2, a_3, a_4 \ldots]_k) = [a_2, a_3, a_4 \ldots]_k. \tag{13}$$

This shows that for any k, ϕ_k behaves similarly to the $k = 1$ Gauss map since both act as a shift on $\mathbb{N}^{\mathbb{N}}$.

DEFINITION 4.1. Define $\theta_k : [0,1] \to [0,1]$ by

$$\theta_k([a_1, a_2, a_3, \ldots]_k) = [a_1, a_2, a_3, \ldots]_1.$$

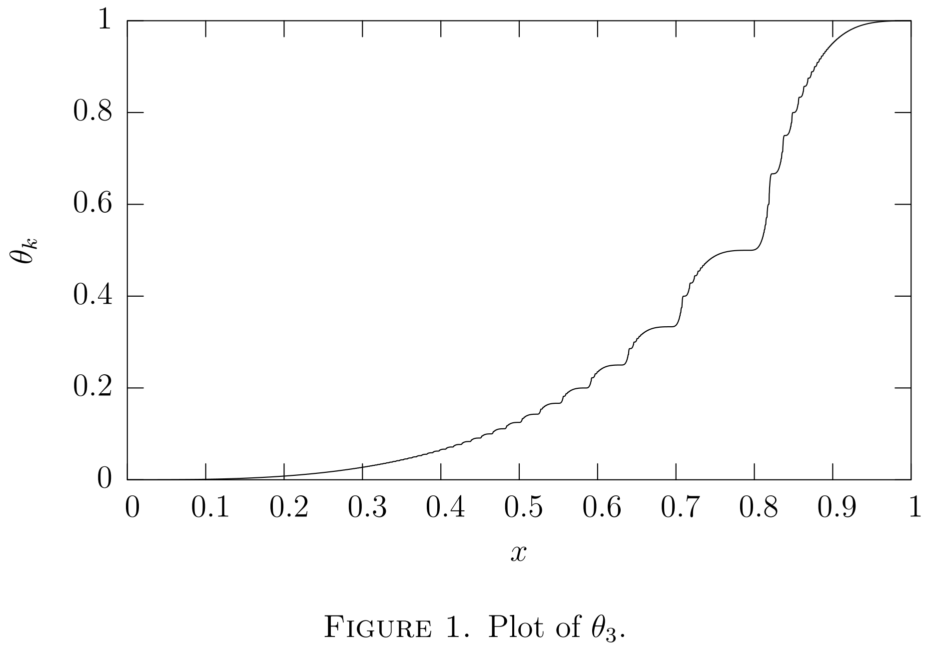

FIGURE 1. Plot of θ_3.

Thus, θ_k removes the k'th powers from the denominators of a nonlinear continued fraction. Figure 1 shows a plot of θ_3. When $k < \Omega$, this appears to be a well-defined function. However, Theorem 3.2 shows that θ is not a function for $k > \Omega$. Figure 1 leads to the following conjecture:

CONJECTURE 4.2. *For $k < \Omega$, θ_k is monotone.*

If θ_k is monotone, then it is continuous and differentiable almost everywhere, so θ_k would act as a topological conjugacy between (I, ϕ_1) and (I, ϕ_k). That is, by (13), we can then observe that

$$\theta_k \circ \phi_k = \phi_1 \circ \theta_k.$$

In addition, it is well-known (see [**3**]) that an invariant measure for the Gauss map ϕ_1 is given by

$$G(X) = \frac{1}{\log 2} \int_X \frac{1}{1+x} dx.$$

If θ_k is a conjugacy between (I, ϕ_k) and (I, ϕ_1), then we know that an invariant measure for ϕ_k is given by

$$G_k(X) = \frac{1}{\log 2} \int_{\theta_k(X)} \frac{1}{1+x} dx.$$

Further exploration of this would require a better grasp of how cylinder sets behave under θ_k.

Bibliography

1. Anon., *Wang tile*, http://en.wikipedia.org/wiki/Wang_tile.
2. R. J. Baxter, *Hard hexagons: Exact solution*, J. Physics A **13** (1980), 1023–1030.
3. P. Billingsley, *Ergodic Theory and Information*, Wiley, 1965.
4. S. Finch, *Mathematical Constants*, Cambridge University Press, 2003.
5. Karel Culik II and Jarkko Kari, *An aperiodic set of wang cubes*, Symposium on Theoretical Aspects of Computer Science, 1996, pp. 137–146.
6. D.L. Johnson, *Symmetries*, Springer, 2001.
7. G. S. Joyce, *On the hard hexagon model and the theory of modular functions*, Phil. Trans. Royal Soc. London A **325** (1988), 643–702.
8. A. Kato and K. Zeger, *On the capacity of two-dimensional run-length constrained channels*, IEEE Trans. Inform. Theory **45** (1999), 1527–1540.
9. S. Lang, *Undergraduate Algebra*, Springer, 1990.
10. D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, 1995.
11. D. Ornstein and B. Weiss, *Entropy and isomorphism theorems for actions of amenable groups*, J. Analyse Math **48** (1987), 1–141.
12. P.H. Siegel R.M. Roth and J. J. Wolf, *Efficient coding schemes for the hard square model*, IEEE Trans. Inform. Theory **47** (2001), 1166–1176.
13. K. Schmidt, *Dynamical Systems of Algebraic Origin*, Birkhauser, 1995.
14. W. Weeks and R.E. Blahut, *The capacity and coding gain of certain checkerboard codes*, IEEE Trans. Inform. Theory **44** (1998), 1193–1203.