New and Existing Results on Circular Words

by

Jesse T. Johnson
B.Sc., Oregon State University, 2018

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Mathematics

Supervisory Committee

Dr. James Currie, Co-supervisor
(Department of Mathematics and Statistics, University of Winnipeg)

Dr. Garry MacGillivray, Co-supervisor
(Department of Mathematics)

ABSTRACT

Circular words, also known as necklaces, are combinatorial objects closely related to linear words. A brief history of circular words is given, from their early conception to present results. We introduce the concept of a *level* word, that being a word containing a equal or roughly equal amount of each letter. We characterize exactly the lengths for which level square free circular words on three letters exist. This is accomplished through a modification of Shur's construction of square-free circular words.

A word on two letters is called a *Frankel-Simpson* word if the only squares it contains are 00, 11, and 0101. Using the result mentioned above and several computer searches, we characterize exactly the lengths for which circular Frankel-Simpson words exist, and give an example or construction for each.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to extend my thanks to my advisors, Gary MacGillivray and James Currie, first and foremost. They have shaped my experience in graduate school, and have done an incredible amount to assist me and grow my abilities as a mathematician. Great thanks are also owed to Narad Rampersad and Lucas Mol, whose input was critical during the early phases of research. I would also like to extend my thanks to all the mathematicians who have taught and guided me until now, especially Yevgeniy Kovchegov, Clayton Petsche, Jason Seifkin, and Bob Burton. Finally, programming assistance came from from Braxton Cuneo and Alex Redfern, and writing advice came from Nicole Murker and Susan Cyganiak.

Dedication

To Evan. If I can make this, you can make anything at all. Just keep making stuff.

Chapter 1

Introduction

In the early 20th century, Axel Thue began the study of combinatorics on words with his research into **square-free words** [30]. Let Σ be a finite set. We refer to Σ as an **alphabet**, and its elements as **letters**. We will work in particular with the alphabets $\Sigma_m = \{0, 1, \ldots, m-1\}$, $A = \{a, b, c\}$, and $S = \{1, 2, 3\}$. We denote by Σ^* the free monoid over Σ, with identity ϵ. We call the elements of Σ^* **words**. As such, if u and v are words, with $u = u_1u_2\ldots u_n$ and $v = v_1v_2\ldots v_m$, then uv is a word with $uv = u_1u_2\ldots u_nv_1v_2\ldots v_m$. We call the binary operation on this monoid **concatenation**. In this case, we say that u is a **prefix** of uv and v is a **suffix**. We call u (resp., v) a **proper prefix** (resp., **proper suffix**) if $v \neq \epsilon$ (resp., $u \neq \epsilon$). More generally, if $w = uvz$, then v is a **factor** of w. In the case that $u, z \neq \epsilon$, we call v an **internal factor** of w. A word is said to **contain** its factors. A word of the form $s = uu$, $u \neq \epsilon$ is called a **square**. A word w which doesn't contain a square factor is said to be **square-free**.

In general, we use w_i to denote the ith letter of w, starting from w_1. If $u = u_1u_2\cdots u_n$, $u_i \in \Sigma$, then n is the **length** of u, and we write $|u| = n$. The set of words of length m over an alphabet Σ is denoted by Σ^m, and the set of words with length greater than 0 is Σ^+. For $a \in \Sigma$, we denote by $|u|_a$ the number of occurrences of a in u. We say that u_i has **index** i in u, and that u_i is the 'ith letter of u'. An infinite-length word on an alphabet Σ is a function from the natural numbers to Σ, where the letter at index i is the image of i under that function. Informally, one may think of an infinite-length word on an alphabet Σ as an infinitely long string of letters in Σ. An infinite-length word is typically called an ω-word.

A word using only two letters is called a **binary word**, while a word using exactly three letters is called a **ternary word**.

Phrased in modern language, Thue's first paper constructed a square-free ternary ω-word, and Thue's second paper constructed a binary ω-word containing no factor of the form $uvuvu$, with $u, v \in \Sigma_2$, and $u \neq \epsilon$. These early results allowed for the rest of the field to be built up from them [30, 31].

Circular Words

This thesis is focused primarily on **circular words**. If u, v, and w are words, with $w = uv$, we refer to vu as a **conjugate** of w. Conjugacy is an equivalence relation on Σ^*, and we refer to the equivalence classes of Σ^* under conjugacy as **circular words**, also called **necklaces**. For any word w, we let $[w]$ denote the circular word containing all conjugates of w. Conversely, the elements of $[w]$ are referred to as **linearizations** of $[w]$. For example, if $\Sigma = \{a, b, c, d, e\}$ and $w = beaded$, then $[w] = \{beaded, eadedb, adedbe, dedbea, edbead, dbeade\}$, and $edbead$ is one linearization of $[beaded]$. Equivalently, we may consider the indices i of the letters of a circular word $[u] = [u_1 u_2 \cdots u_n]$ to belong to $\mathbb{Z}_n$, the integers modulo n. Thus $u_{n+1} = u_1$, for example. It is natural to visualize the letters of $[w]$ being arranged in a circle, as shown in Figure 1.1. For this reason, circular words are also referred to as **necklaces**.

When speaking of w and circular word $[w]$, we may refer to w as a 'linear word' to emphasize that it is a single element of Σ^*. If $[w]$ is a circular word and $v \in \Sigma^*$, we say that v is a **factor** of $[w]$ if v is a factor of an element of $[w]$; equivalently, v is a factor of $[w]$ if v is a factor of a conjugate of w. A circular word $[w]$ is **square-free** if no factor of $[w]$ contains a square, i.e., if every conjugate of w is square-free.

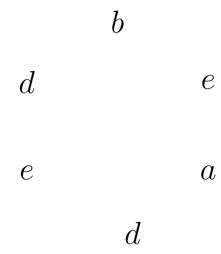

Figure 1.1: Circular word (*beaded*)

Morphisms

Both of Thue's results described so far were accomplished using the now ubiquitous tool of morphisms. Let Σ and T be alphabets. A map $\mu : \Sigma^* \to T^*$ is called a **morphism** if it is a monoid homomorphism, that is, if $\mu(uv) = \mu(u)\mu(v)$, for $u, v \in \Sigma^*$. The use of morphisms extends naturally to circular words, with $\mu([w]) = [\mu(w)]$, whenever $w \in \Sigma^*$ and μ a morphism.

A morphism is determined by its image on letters. If $u = a_1a_2 \cdots a_n$, $a_i \in \Sigma$, then $\mu(u) = \mu(a_1)\mu(a_2) \cdots \mu(a_n)$. We will also wish to allow **multi-valued morphisms** defined on the letters of Σ. A multi-valued morphism $\mu : \Sigma^* \to T^*$ maps each letter of Σ to a set of words of T^*. For arbitrary $u \in \Sigma^*$, $u = a_1a_2 \cdots a_n$, $a_i \in \Sigma$, we define $\mu(u)$ to be the set $\{v_1v_2 \cdots v_n : v_i \in \mu(a_i)\}$. If μ is a multi-valued morphism, we interpret 'u is a factor of $\mu(v)$' to mean that u is a factor of an element of $\mu(v)$.

Chapter 2

History

de Bruijn Sequences

The history of circular words can be traced back earlier than Thue's papers. *McMahon's* formula, found in 1892, gives the number of length n circular words on k letters as

$$\frac{1}{n}\sum_{d|n} k^d \phi(n/d),$$

where ϕ is Euler's totient function [19]. However, there were few developments into circular words until later in the 20th century. An early result closely related to McMahon's formula has to do with the number of primitive necklaces. A circular word $[w]$ is called a **primitive** circular word if none of its linearizations are equivalent. An example of a primitive circular word is given in figure 2.1, and an example of a circular word that is not primitive is given in figure 2.2.

Witt's Formula enumerates the number of primitive necklaces as

$$\frac{1}{n}\sum_{d|n} k^d \mu(n/d)$$

where μ is the Möbius function. This related result was not developed until 1937, signif-

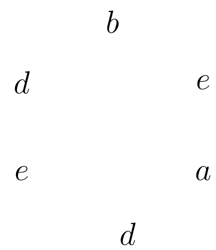

Figure 2.1: Primitive circular word (*beaded*)

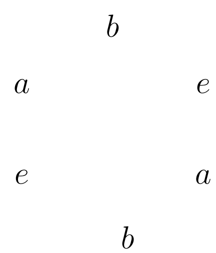

Figure 2.2: A circular word that is not primitive (*beabea*)

icantly later than its predecessor.[32]. Further development into the idea of circular words is that of a **de Bruijn sequence**. For a given length n, and an alphabet A of size k, a de Bruijn sequence of order n on A is a circular word containing as a factor every possible length n word on A exactly once. An example of a de Bruijn sequence of order 5 on Σ_2 is [4]:

$$00000100011001010011101011011111$$

We have the following theorem:

Theorem 2.1. *The number of distinct de Bruijn sequences of order n on Σ_k having length k^n is given by*

$$\frac{(k!)^{k^{n-1}}}{k^n}.$$

Although the existence of these sequences was shown in special cases in 1894 [24], the existence of a de Bruijn sequence for any n or k was not shown until 1934, by Martin [20]. Martin had only shown the existence of such words; the enumeration was found by de Bruijn in 1946, giving the sequences their name [11].

Palindromes and Reverses

To begin our look into more recent results, we examine **reverses** of a word. Given a word $u = u_1u_2 \dots u_{n-1}u_n$ we define the **reverse** of u to be $u^R = u_nu_{n-1} \cdots u_2u_1$. Building on this, a word u is called a **palindrome** if $u = u^R$. For instance, the word $u = 1001001001$ is a palindrome. For $c \in \Sigma_2$, let $\bar{c}$ be the element of Σ_2 such that $\bar{c} \neq c$. Then given $u = u_1u_2 \dots u_{n-1}u_n$ we define $\bar{u} = \overline{u_1}\overline{u_2} \dots \overline{u_{n-1}}u_n$. Then we say that a word u is an **antipalindrome** if $u = \bar{u}^R$. For instance, 10010110 is an antipalindrome.

We say a word u is a **subsequence** of a word v if there exists a function f from the indices of letters in u to indices of letters in v, such that $f(i) \geq i$. Put another way, one may form u by selecting letters in v such that each letter selected comes after the one before. To give an example, the word $u = business$ has the word $v = bins$ as a subsequence. In this case, the function used is $f(0) = 0$, $f(1) = 3$, $f(2) = 4$, and $f(3) = 6$

While studying protein folding, Lyngsø and Pedersen formed the following conjecture [18]:

Conjecture 2.1.1. *Given a circular word $[w]$ on Σ_2, if $|w|_0 = |w|_1$, then $[w]$ contains a subsequence w' such that w' is an antipalindrome, and $|w'| \geq \frac{2|w|}{3}$.*

Conjecture 2.1.1 has recently been given a palindromic counterpart [21]:

Conjecture 2.1.2. *Given a circular word* $[w]$ *on* Σ_2*,* $[w]$ *contains a subsequence* w' *such that* w' *is a palindrome, and* $|w'| \geq \frac{2|w|}{3}$*.*

Both conjectures remain unproven. Trivially, it may be shown that there exists w' with $|w'| \geq \frac{|w|}{2}$ in the case of Conjecture 2.1.2 by considering the maximally large subsequence containing only a single letter. It has been shown by Müllner and Rhyzikov that this is a tight bound, and they give examples of infinite classes of words for which the maximum size of any palindromic subsequence approaches $\frac{2|w|}{3}$ as $|w|$ increases.

Several related results have been found. For instance, consider the question of the number of distinct palindromes in a word. For any word w, Let $Pal(w)$ be the size of the set $\{p_1, p_2 \ldots\}$, where p_i is a palindromic factor of w, and $p_i \neq p_j$ for all i and j. In 2004, it was shown that $Pal(w) \geq |w|$ [6]. Prompted by this result, Simpson [28] has formed a similar theorem for circular words:

Theorem 2.2. *Given a circular word* $[w]$*,* $Pal([w]) < \frac{5|w|}{3}$*.*

It should be noted that this bound is nearly sharp. Simpson gives examples of circular words of length n containing $\frac{5n}{3} - 2$ distinct words.

Squares and powers

The first paper on combinatorics on words was by Thue, showing the existence of infinitely long square-free words on 3 letters, meaning that the concept of a word avoiding or encountering a repetitive pattern is approximately as old as the concept of words themselves. This is commonly generalized as follows:

Definition 1. *For a word* $u \in \Sigma^*$*, we define* u^2 *as* uu*,* u^3 *as* uuu*, and* u^n *as* n *consecutive repetitions of* u*. If* $i \in \mathbb{R}_{>1}$*, and* p *is the prefix of* u *of length* $\lceil (i - \lfloor i \rfloor) \cdot |u| \rceil$*, then* $u^i = u^{\lfloor i \rfloor} p$*.*

To give examples, if $u = lyrical$, then $u^{9/7} = lyrically$, and the word *ingraining* is the $\frac{10}{7}$-power of *ingrain*. This concept may also be understood in terms of **periodicity**. Given a word $w = w_1 w_1 \dots w_n$, we say a number p is a **period** of w if $w_i = w_{i+p}$ for all integers i between 1 and $n - p$. A word w is a k-**power** if it has a period p, with $p \leq \frac{|w|}{k}$; k is called the **exponent** of w. A word avoiding factors of the form x^3 is called cube-free, and a word avoiding factors of the form x^n is called n-power free. A slight alteration on this concept is that of an overlap:

Definition 2. *A word is called a $k+$ power if it is a j power for some $j > k$. In particular, a $2+$ power is called an* **overlap**.

Overlaps were studied by Thue as well as squares. In his 1912 paper, Thue famously constructed an ω-word on two letters that avoids all overlaps. Let $\mu : \Sigma_2 \to \Sigma_2$ be the morphism defined with $\mu(0) = 01$ and $\mu(1) = 10$. The **Thue-Morse word** is the ω-word formed by repeatedly iterating μ on the letter 0. This is denoted with $\mu^\omega(0) = \lim_{n\to\infty} \mu^n(0)$.

Thue then found that this word can be used show the existence of a square free ω-word on three letters. Provided with an overlap free word $w \in \Sigma_2^*$, there is a square free word $x \in \Sigma_3^*$ with $f(x) = w$, where $f(a) = 0$, $f(b) = 01$, and $f(c) = 011$. This result can be extended to the circular case as follows:

Theorem 2.3. *[8] For any length $\ell \neq 5, 7, 9, 10, 15, 17$, there is a square-free ternary word of length ℓ.*

This was proven first by James Currie in 2002, and another method of constructing such words was found several years later, by Shur [27]. This result and the method used by Shur form the basis for the results shown in later chapters.

Many other significant results occur as a result of the Thue-Morse word. While there are square-free circular words on three letters for any length greater than 18, there are cube-free circular words on two letters for any length at all, and their elements are factors of the

Thue-Morse word [25]. It was afterwards found that the Thue-Morse word contains factors of every length above 209 with circular words that avoid $\frac{7}{3}+$ powers [1]. It was then shown that [2]:

Theorem 2.4. *For any length ℓ, the Thue-Morse word contains a factor v with $|v| = \ell$, and with $[v]$ avoiding all $\frac{5}{2}+$ powers.*

Note that Theorem 2.4 implies the two that came before it. Also note that this is a tight bound, as any binary circular word of length 5 must contain either a cube or a $\frac{5}{2}$ power as a factor.

Despite the tightness of this bound, a stronger theorem may be proved. Let the **critical exponent** of a word w be the greatest exponent of any nonempty factor of w. The **circular critical exponent** of $[w]$ is the greatest exponent of any nonempty factor of any conjugate of w. For instance, the circular critical exponent of the word *tomato* is 2, because *totoma* is a conjugate of *tomato*, and *toto* is a square. It was shown in 2018 that the circular critical exponent of any finite factor of the Thue-Morse word belongs to a finite list of rational numbers, specifically $\{1, 2, \frac{7}{3}, \frac{5}{2}, \frac{8}{3}, \frac{11}{4}, 3, \frac{10}{3}, \frac{7}{2}, 4, \frac{13}{3}, 5, 6\}$ [26]. This was shown by Jeffrey Shallit and Ramin Zarif, through use of a proving engine called Walnut.

How many distinct squares there can possibly occur as factors of a circular word of length n? Frankel and Simpson showed that in the linear case, a word of length n has at most $2n$ distinct squares [14]. Because of this, and because any square appearing in a circular word $[w]$ must appear in ww, we have the trivial bound of $4n$. A much sharper bound is given by Amit and Gawrychowski, who found a bound of $3.14n$ [3]. While this is a significant improvement, it is suspected from computer searches that the actual value is closer to $1.25n$.

Sturmian Words

An ω-word w is a **Sturmian word** if it contains exactly $n+1$ distinct factors of length n. Sturmian words have received a large amount of study in recent years because of their applications to a wide range of topics. We are interested in **Christoffel words**, which are regarded as a finite counterpart to Sturmian words. Given two coprime integers, q and p, consider the line segment on the latice $\mathbb{N} \times \mathbb{N}$ from $(0,0)$ to (q,p). Moving from $(0,0)$ to (q,p) along this line, code a 0 when this line segment passes an integer on the vertical axis, and code a 1 when this line segment passes an integer on the horizontal axis. This sequence of 0s and 1s gives a Christoffel word.

Given a linear Christoffel word w, the circular word $[w]$ is a **circular Sturmian word**. These circular Strumian words can be characterized in a variety of ways. For example: [7]

Theorem 2.5. *If $[w]$ is a circular Sturmian word, then $[w]$ has k factors of length $k+1$, for any whole number k less than $|w|$.*

To see another example, we define a word w to be **balanced** if for any letter a in w's alphabet, and for any two factors of w with equal length, called u and v, it is the case that $|u|_a - |v|_a \in -1, 0, 1$. Surprisingly, all circular Sturmian words are balanced [7]. Recall that a circular word is primitive if none of its elements are equal. In addition to the other properties mentioned, all circular Sturmian words are primitive.

A morphism μ is called a **Sturmian morphism** if, for any Sturmian word x, $\mu(x)$ is a Sturmian word as well. It is also the case that if $[x]$ is a circular Sturmian word, and μ is a Sturmian morphism, then $\mu([x])$ is Sturmian as well.

Chapter 3

Level Words

For a finite word w, we define the **density** of a letter a in w as $\frac{|w|_a}{|w|}$. If w is an ω-word, and p_i is the length i prefix of w, we define the density of a in w as $\lim_{i\to\infty} \frac{|p_i|_a}{|p_i|}$. It has been shown by Tarannikov [29] in 2002 and by Khalyavin [17] in 2007 that a ternary square-free ω-word must have a minimal density of 0.2746 or $\frac{883}{3215}$ for each of its letters. In this chapter, we explore an opposite concept: When is it possible for a ternary square-free word to have an equal, or nearly equal number of each letter in its alphabet? To formalize this concept, we introduce the following definition:

Definition 3. *A finite word w over Σ_3 is* **level** *if*

$$|w|_\alpha - 1 \leq |w|_\beta \leq |w|_\alpha + 1$$

for $\alpha, \beta \in \Sigma_3$.

If a linear word w avoids pattern p, then any factor of w avoids p as well. Therefore, to find a linear word avoiding pattern p of size ℓ, it suffices to find a linear word avoiding p of length greater than ℓ. For instance, to construct a overlap free word of length ℓ, take the length ℓ prefix of the Thue-Morse word. However, if $[w]$ is a circular word avoiding a pattern p, it is not necessarily the case that circularizations of factors of w avoid p. For example, length 5 ternary square-free circular words do not exist, so any circularization of a length 5 factor of a square-free circular word is not itself square-free.

In order to know the length of the image of a circular word under a particular morphism, it is necessary to know the composition of the preimage. Therefore, the precise theorem we will prove is:

Theorem 3.1. *For any length $l \neq 5, 7, 9, 10, 15, 17$, there is a level square-free ternary circular word of length ℓ.*

To outline our proof of this theorem, we first look at Shur's construction of square-free ternary words, and adjust it to find level square-free circular words of length $18n$, with $n \in \mathbb{N}$ and $n \neq 5, 7, 9, 10, 15, 17$. Following this, we find a set of factors that may be inserted into an encoding of a square-free word to find level square-free circular words of lengths besides $18n$.

To begin our proof of this theorem, we first look at Dejean's Conjecture:

Conjecture 3.1.1. *[12](Dejean, 1972) For any alphabet Σ_n, the infimum of k such that there is an ω-word on Σ_n^* avoiding all powers with exponents greater than k is $\frac{7}{4}$ if $n = 3$, $\frac{7}{5}$ if $n = 4$, and $\frac{n}{n-1}$ for other values of n.*

In proving this conjecture for $n = 4$, Pansiot [22] introduced **Pansiot encodings**, which have become a standard tool in the study of nonrepetitive words. Let $u = u_1 u_2 \cdots u_n, n \geq 2$ be a square-free word, with $u_i \in A$. Suppose we are given $u_1, u_2, \ldots, u_{i+1}$, $1 \leq i \leq n-2$. Since u is square-free, $u_{i+2} \neq u_{i+1}$. Since there are only 3 letters in A, u_{i+2} is therefore determined, once we know whether or not $u_{i+2} = u_i$. The **Pansiot encoding** of u is defined to be the word $\pi(u) = v_1 v_2 \cdots v_{n-2}$ where, for $1 \leq i \leq n-2$,

$$v_i = \begin{cases} 0, & u_i = u_{i+2} \\ 1, & u_i \neq u_{i+2} \end{cases}$$

Therefore u can be recovered from u_1, u_2 and $\pi(u)$.

Shur [27], introduces Pansiot encodings for circular words. The **Pansiot encoding** of $[u]$ is defined to be the circular word $\pi([u]) = [v_1 v_2 \cdots v_n]$ where, for $1 \leq i \leq n$,

$$v_i = \begin{cases} 0, & u_i = u_{i+2} \\ 1, & u_i \neq u_{i+2} \end{cases}$$

Here the arithmetic on indices is carried out in $\mathbb{Z}_n$.

Note that $\pi([u])$ is well-defined; a conjugate or 'rotation' of u yields a conjugate of the word $v_1 v_2 \cdots v_n$ defined above, by the same rotation.

As stated, our goal is to show that there is a level circular ternary square-free word for every length greater than 18, and to accomplish this, we modify the methods used to find square-free ternary circular words in general.

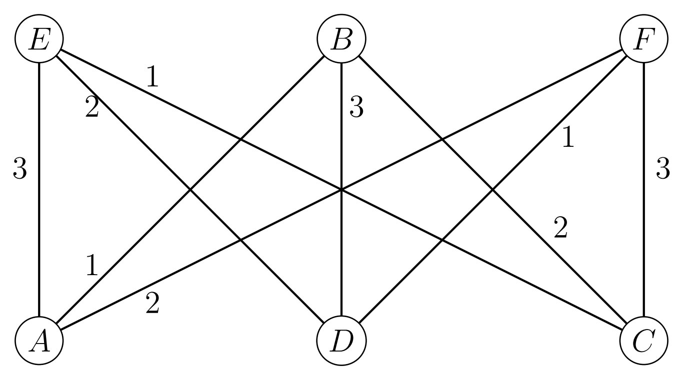

Figure 3.1: Shur's graph $\mathcal{G}$

Shur [27] defines a graph with labeled edges, isomorphic to graph $\mathcal{G}$ shown above. Because this graph is used only to discuss what strings label closed walks, the following definition and lemma provide a method by which we can bypass most consideration of this graph.

Definition 4. *Let $s \in S^*$. Let $\omega(s) = \sum_{i=0}^{|s|} (-1)^i s_i$, with each letter of s being considered as an integer.*

Lemma 3.1.1. *Let $s \in S^*$. If $|s|$ is even, and $\omega(s)$ is divisible by 3, then s is the sequence of edge labels of a closed walk on $\mathcal{G}$.*

Proof. Let $s = s_1 s_2 \dots s_n$ be an element of S^* with n even. The statement of Lemma 3.1.1 only requires that s is the sequence of edge labels of one closed walk, so let s' be a walk on $\mathcal{G}$ starting on the vertex A that is edge labeled by s. Consider the mapping $\mu(A) = \mu(E) = 0$, $\mu(D) = \mu(B) = 1$, $\mu(C) = \mu(F) = 2$. Let $s_{[i]}$ be the length i prefix of s, and let $s'_{[i]}$ be the walk labeled by $s_{[i]}$ starting at A. Let the last node of $s'_{[i]}$ be N_i. We will show by induction on i that $\mu(N_i) = \omega(s_{[i]}) \pmod 3$.

In the base case, let i equal 0. It is clear to see that $\mu(N_0) = 0 = \omega(s_0) \pmod 3$. So suppose $\mu(N_{i-1}) = \omega(s_{[i-1]}) \pmod 3$. Let ϕ be the last letter of $s_{[i]}$. Note that if $N_{i-1} \in \{A, D, C\}$, then $\mu(N_i) \pmod 3 = \mu(N_{i-1}) + \phi \pmod 3$, while if $N_{i-1} \in \{E, B, F\}$, then $\mu(N_i) \pmod 3 = \mu(N_{i-1}) - \phi \pmod 3$. Note also that if $i-1$ is even, then $N_{i-1} \in \{A, D, C\}$, while if $i-1$ is odd, then $N_{i-1} \in \{E, B, F\}$. Then

$$\begin{aligned}
\mu(N_i) &\equiv \mu(N_{i-1}) + (-1)^i \phi \\
&\equiv \omega(s_{[i-1]}) + (-1)^i \phi \\
&\equiv \sum_{j=0}^{i-1} (-1)^j s_j + (-1)^i \phi \\
&\equiv \sum_{j=0}^{i} (-1)^j s_j \\
&\equiv \omega(s_{[i]}) \pmod 3
\end{aligned}$$

If we let $i = n$, then $\mu(N_n) = 0$, implying that s' ends on either A or E. Note that $\mathcal{G}$ is bipartite. Because $|s|$ is even, s' must end on a node in the same disjoint set of nodes as the node that s' begins in, implying that $N_n = A$. Therefore, s' is a closed walk on $\mathcal{G}$ that is edge labeled by s.

□

Define the morphism $f : S \to \Sigma_2^*$ to be

$$f(1) = 01$$
$$f(2) = 011$$
$$f(3) = 0111$$

It was shown by Shur that the following requirements suffice for the sequence of edge labels of a walk on $\mathcal{G}$ to be a Pansiot encoding of a square free word [27]:

Lemma 3.1.2. *If $[w]$ is the sequence of edge labels of a closed walk of $\mathcal{G}$, then $f([w])$ is the Pansiot encoding of a square-free circular word if*

- *$[w]$ has no factor 11, 222, 223, 322, or 333*
- *$[w]$ has no factor $UxyU$ with Uxy a closed walk, where $|U| \geq 2$, and $x, y \in S$.*

Consider the substitution $h : A^* \to S^*$ given by

$$\begin{aligned} h(a) &= 123123 \\ h(b) &= 132132 \\ h(c) &= 131313 \end{aligned}$$

For $x \in A$ we refer to the word $h(x)$ as a **block**.

Remark 3.1.1. By, Lemma 3.1.1, Each block labels a closed walk on $\mathcal{G}$. Thus, for any word $w \in A^*$, $h(w)$ also labels a closed walk on $\mathcal{G}$.

Lemma 3.1.3. *The morphism h has the following properties:*

1. *Let $x, y \in A$. Let $s \in S^2$. If s is a suffix of both of $h(x)$ and $h(y)$, then $x = y$. Thus each letter $x \in A$ is determined by the length 2 suffix of $h(x)$.*

2. *Let $x, y \in A$. Let $p \in S^3$. If p is a prefix of both of $h(x)$ and $h(y)$, then $x = y$. Thus each letter $x \in A$ is determined by the length 3 prefix of $h(x)$.*

3. *Let $x, y, z \in A$, with $x \neq c$ and $y \neq z$. Let ϕ be a nonempty prefix of $h(y)$, and let ρ be a nonempty suffix of $h(z)$. Then $h(x) \neq \rho\phi$.*

Proof. The proof simply requires a finite amount of inspection:

1. The length 2 suffix of $h(a)$ is 23, the length 2 suffix of $h(b)$ is 32, and the length 2 suffix of $h(c)$ is 13. Therefore, the length 2 suffixes of blocks are distinct.

2. The length 3 prefix of $h(a)$ is 123, the length 3 prefix of $h(b)$ is 132, and the length 3 prefix of $h(c)$ is 131. Therefore, the length 3 prefixes of blocks are distinct.

3. This can be verified by exhaustively combining the possible prefixes and suffixes whose lengths add up to 6. Note that x cannot equal c, due to the case where $\rho = 1313$, the length 4 suffix of $h(c)$, and where $\phi = 13$, the length 2 prefix of $h(b)$.

□

Theorem 3.2. *Let $[v]$ be a square-free circular word over A. Let $[w] = h([v])$. Then $f([w])$ encodes a square-free circular word.*

Proof. By Shur's Lemma, it suffices to show that

- $[w]$ has no factor 11, 222, 223, 322, or 333

- $[w]$ has no factors $UxyU$ with $x, y \in S$, $U \in S^*$ where $|U| \geq 2$, and where Uxy is the label of a closed walk.

Let $[v]$ be a square-free circular word over A, and let $[w] = h([v])$. No element of $\{11, 222, 223, 322, 333\}$ appears as a factor of a concatenation of any two blocks, so no such element appears as a factor of $h(v)$.

Suppose for the purpose of finding a contradiction that $UxyU$ is a factor of $[w]$, where $x, y \in S$, $U \in S^*$, and Uxy is a closed walk on $\mathcal{G}$. Let $\nu = \nu_1\nu_2\dots$ be an element of $[v]$ such that $h(\nu\nu_1)$ contains $UxyU$. Let μ be the smallest factor of $\nu\nu_1$ such that $h(\mu)$ contains $UxyU$. It may be manually checked for any square-free ternary word t with $|t| \leq 6$, $h(t)$ does not contain any factor of the the form $UxyU$ as described. Therefore $|\mu| \geq 7$, and there is some factor of μ called $\hat{v}$ such that

$$U = sh(\hat{v})p$$

Where s is a proper suffix of a block and p is a proper prefix of a block. Trivially, this leads to:

$$UxyU = sh(\hat{v})pxysh(\hat{v})p.$$

Because $|\mu| \geq 7$, $h(\hat{v})pxysh(\hat{v})$ is the image of a word that is at least 5 letters long, so $|h(\hat{v})pxysh(\hat{v})| \geq 30$. Note that $|pxys| \leq 12$, so $2|h(\hat{v})| \geq 18$, and $|h(\hat{v})| \geq 9$. Conclude that $|\hat{v}| \geq 2$.

Because $|\hat{v}| \geq 2$, $\hat{v}$ contains at least one letter that is not c. Therefore, by the third item of Lemma 3.1.3, $h(\hat{v})$ appears in $h(\mu)$ only as the image of $\hat{v}$. Let $\varphi_2 \in A^*$ be such that $h(\varphi_2) = pxys$, and note that $1 \leq |\varphi_2| \leq 2$.

Case 1: Suppose $2|\hat{v}| + |\phi_2| < |v| - 1$. Then μ is of the form

$$\varphi_1\hat{v}\varphi_2\hat{v}\varphi_3,$$

where $\varphi_1, \varphi_3 \in A$ such that s is a suffix of $h(\varphi_1)$, and p is a prefix of $h(\varphi_3)$. Additionally, $|\mu| \leq |\nu|$, so that μ must be a factor of $[v]$, and so μ is square-free.

Suppose $|\varphi_2| = 2$, so that $|pxys| = |h(\varphi_2)| = 12$. Because $|x| = |y| = 1$, it must be that $|p| = |s| = 5$. As mentioned in Lemma 3.1.3 above, there is exactly one $\varphi_1 \in A$ such that $h(\varphi_1)$ has s as a suffix, and exactly one $\varphi_3 \in A$ such that $h(\varphi_3)$ with p as a prefix. φ_2 has p as a prefix, and so its first letter is φ_3. Similarly the second letter of φ_2 is φ_1. Then μ is

$$\varphi_1 \hat{v} \varphi_3 \varphi_1 \hat{v} \varphi_3,$$

which is a square, giving a contradiction. Suppose instead that $|\varphi_2| = 1$, which gives $|pxys| = |f(\varphi_2)| = 6$. Because $|xy| = 2$, $|p| + |s| = 4$. Suppose $|s| \geq 2$, so that there is only one block that s is a suffix of. Because s is a suffix of both φ_3 and φ_2, $\varphi_2 = \varphi_3$, and μ is

$$\varphi_1 \hat{v} \varphi_2 \hat{v} \varphi_2,$$

which contains a square. So suppose to the contrary that $|s| \leq 2$, so that $|p| \geq 3$ and there is only one block that p is a prefix of. Because p is a prefix of both φ_1 and φ_2, $\varphi_2 = \varphi_1$, and μ is

$$\varphi_2 \hat{v} \varphi_2 \hat{v} \varphi_3,$$

which also contains a square. This contradicts our assumption that v is square free, so conclude that $f([w])$ does not contain any $UxyU$, where U contains a block.

Case 2: Suppose $2|\hat{v}| + |\phi_2| = |v| - 1$, so that $\mu = \nu\nu_1$. Then we have

$$\nu = \varphi_1 \hat{v} \varphi_2 \hat{v},$$

where $\varphi_1 \in A$ such that s is a suffix of $h(\varphi_1)$, and p is a prefix of $h(\varphi_1)$. If $|\varphi_2| = 1$, then

$h(\varphi_2)$ is a single block, with p as a prefix and s as a suffix. Recall that $h(\varphi_2) = pxys$ so $p+2+s$. If $|p| \geq 3$, then φ_2 is uniquely determined, and so $\varphi_2 = \varphi_1$. On the other hand, if $|p| < 3$, then $|p| \leq 2$, and $|s| \geq 2$. In this case φ_2 is also uniquely determined, and $\varphi_2 = \varphi_1$. Then regardless of the size of p,

$$\nu = \varphi_1 \hat{v} \varphi_1 \hat{v},$$

which contradicts the square-freeness of ν.

So suppose $|\varphi_2| = 2$, and $h(\varphi_2) = pxys$. We have $|h(\varphi_2)| = 12$, so that $|pxys| = 12$, and $|p| + |s| = 10$. The factors p and s are both proper factors of blocks, and so $|p| < 6$, and $|s| < 6$. Combining these two facts gives that $|p| = |s| = 5$. Then p can only be a prefix of φ_1, and s can only be a suffix of φ_1. Therefore $h(\varphi_2) = h(\varphi_1)h(\varphi_1)$, implying that $\varphi_2 = \varphi_1\varphi_1$. This contradicts the square-freeness of ν.

□

A more complex proof that does not rely on direct verification of small cases is given in Appendix B.

To demonstrate the levelness of words constructed by h, we introduce the following definition and associated lemma.

Definition 5. *Given a Pansiot encoding μ, the decoding of μ is the word w defined such that*

1. $w_1 = a$
2. $w_2 = b$
3. *For $n \geq 3$, if $\mu_{n-2} = 0$, $w_n = w_{n-2}$*
4. *For $n \geq 3$, if $\mu_{n-2} = 1$, $w_n \neq w_{n-2}$ and $w_n \neq w_{n-1}$*

We may also notate the decoding of μ as $\Delta(\mu)$.

This definition may be given a circular counterpart:

Definition 6. *Let $[\mu]$ be a circular Pansiot encoding. Let w be the decoding of μ and let w^- be the length $|w|-2$ prefix of w. If w ends in ab, $[w^-]$ is the decoding of μ.*

We may also notate the decoding of $[\mu]$ with $[\Delta(\mu)]$.

Lemma 3.2.1. *Let $\mu, \nu \in \{0,1\}^*$ be Pansiot encodings where $\hat{w} = \hat{w}_1\hat{w}_2\cdots = \Delta(\mu)$, and where $\bar{w} = \bar{w}_1\bar{w}_2\cdots = Delta(\nu)$. If ab is a suffix of $\hat{w}$ and $\bar{w}$, then $\Delta([\mu\nu]) = [\hat{w}^-\bar{w}^-]$, where $\hat{w}^-$ is the length $|\hat{w}|-2$ prefix of $\hat{w}$ and $\bar{w}^-$ is the length $|\bar{w}|-2$ prefix of $\bar{w}$.*

Proof. Let w be $\Delta(\mu\nu)$. Trivially, $w_1 = a$ and $w_2 = b$. For $1 \le i \le |\mu|$, $(\mu\nu)_i = \mu_i$. If $w_j = \hat{w}_j$ for all $j < i$, then $w_{i+2} = \hat{w}_{i+2}$. By induction, $w_{i+2} = \hat{w}_{i+2}$ for all $1 \le i \le |\mu|$. Therefore, w has $\hat{w}$ as a prefix.

By assumption, $w_{|\mu|+1} = a$, and $w_{|\mu|+2} = b$. Note that $(\mu\nu)_{|\mu|+i} = \nu_i$. If $w_{j+|\mu|+2} = \bar{w}_{j+2}$ for all $j < i$, then, by the definition of decoding, $w_{i+|\mu|} = \bar{w}_i$. By induction, w has $\bar{w}$ as a suffix.

Because $|\hat{w}^-| + |\bar{w}| = |w|$, and because w has $\hat{w}^-$ as a prefix and $\bar{w}$ as a suffix, conclude that $w = \hat{w}^-\bar{w}$. Therefore the length $|\mu\nu|$ prefix of w is $\hat{w}^-\bar{w}^-$, and w ends with ab. By definition, the decoding of $\Delta([\mu\nu]) = [\hat{w}^-\bar{w}^-]$. □

Using this, we add to our previous results as follows:

Corollary 3.2.1. *Let $[v]$ be a square-free circular word over alphabet A. Let $[w] \in h([v])$. Then $f([w])$ encodes a level square-free circular word. In fact, $|\Delta(f([w]))|_a = |\Delta(f([w]))|_b = |\Delta(f([w]))|_c$.*

Proof. The decoding of each block is given in Table 3.1.

By inspection, each of these decodings ends in ab. Inductive use of Lemma 3.2.1 implies that the decoding of $f([w])$ is a circular concatenation of the words given in Table 3.2.

x	$h(x)$	$\Delta(f(h(x)))$
a	123123	abacabcbacbcacbabcab
b	132132	abacabcacbabcbacbcab
c	131313	abacabcacbcabcbabcab

Table 3.1: The decodings of the images of blocks under f

x	$\Delta(f(h(x)))^-$
a	abacabcbacbcacbabc
b	abacabcacbabcbacbc
c	abacabcacbcabcbabc

Table 3.2: The truncations of those decoding

Each of these words has exactly 6 instances of each letter. Therefore, the decoding of $f([w])$ is square-free, and has exactly $6|v|$ instances of each letter.

□

Theorem 3.2 may be used to construct a square-free level ternary word of any length of the form $18n$, with $n \notin \{5, 7, 9, 10, 15, 17\}$. To construct square-free circular words of other lengths, we will create a set of factors that maintain the desired properties when inserted into a word constructed with the morphism h. We establish the following lemma:

Lemma 3.2.2. *Let $v \in A^*$ be a word with b as a suffix and a as a prefix, such that $[v]$ is a square-free circular word, let $w = h(v)$, let $T \in S^*$, and let $s = 33T22$. If s is such that:*

1. *T has no suffix of the form $23123h(u)p$ or $123123h(u)p1$, where $u \in A^*$ and p is a prefix of a block.*

2. *T has no prefix of the form $qh(u)13213$ or $1qh(u)132132$, where $u \in A^*$ and q is a suffix of a block.*

3. *s has no factor $h(\mu)$, where $\mu \in A^*$ and $|\mu| \geq 2$, such that $[\mu]$ is square-free.*

4. *The word s labels a closed walk on $\mathcal{G}$*

5. *$2s1$ contains no factor $UxyU$ where $U \in S^*$, $x, y \in S$, and where Uxy is the label of a closed walk on $\mathcal{G}$*

6. *The word T contains no factor $\phi\phi$, where $\phi \in S$*

7. *The word T begins and ends with the letter 1*

Then we may conclude that $f([ws])$ encodes a square-free circular word.

Proof. By Lemma 3.2, it is sufficient to show that $[ws]$ has no factor from $\{11, 222, 223, 322, 333\}$, and no factor of the form $UxyU$ where $U \in S^*$, $x, y \in S$, and Uxy is the label of a closed walk on $\mathcal{G}$.

By conditions 6 and 7, The word s has no factor in $\{11, 222, 223, 322, 333\}$. Because w and s both do not contain any of the factors $\{11, 222, 223, 322, 333\}$, if any of these factors appear in $[ws]$, it must be a factor of either ws or sw. In fact, such a factor must have letters in both s and w. The last letter of s is 2, and the first letter of w is 1. The subfactor 21 does not appear in any factor on that list, so such a factor does not exist. The factor 32 is a suffix of w, and 33 is a prefix of s. Then any factor of ws of length 3 that contains a suffix of w and a prefix of s is a factor of 3233. Note that 3233 contains no factor from the list $\{11, 222, 223, 322, 333\}$, so these factors do not exist in ws or in sw.

Similarly, because $[w]$ and $[s]$ lack a factor $UxyU$, where $U \in S^*$, $x, y \in S$, and Uxy being the label of a closed walk on $\mathcal{G}$, if a factor $UxyU$ appears in $[ws]$, it must contain part of w, as well as part of s. There are four possibilities for the composition of $UxyU$:

1. $UxyU = s'w'$, where s' is a suffix of s, and w' is a prefix of w.

2. $UxyU = s'ws''$, where s' is a suffix of s, and s'' is a prefix of s.

3. $UxyU = w's'$, where w' is a suffix of w, and s' is a prefix of s.

4. $UxyU = w'sw''$, where w' is a suffix of w, and w'' is a prefix of w.

Remark 3.2.1. Note that 33 appears in $[ws]$ only once, as a prefix of s. Similarly, 22 appears in $[ws]$ only as a suffix of s. Therefore, in any of these cases, it is impossible for a length two prefix or suffix of s to appear entirely inside U. Any non-empty prefix or non-empty suffix of s appearing in $UxyU$ must either have length 1, or else one of its repeated letters must occur at x or y.

Case 1: *$UxyU = s'w'$, where s' is a suffix of s, and w' is a prefix of w.*

Because of the restrictions mentioned in Remark 3.2.1, Case 1 consists of subcases given in Figure 3.2.

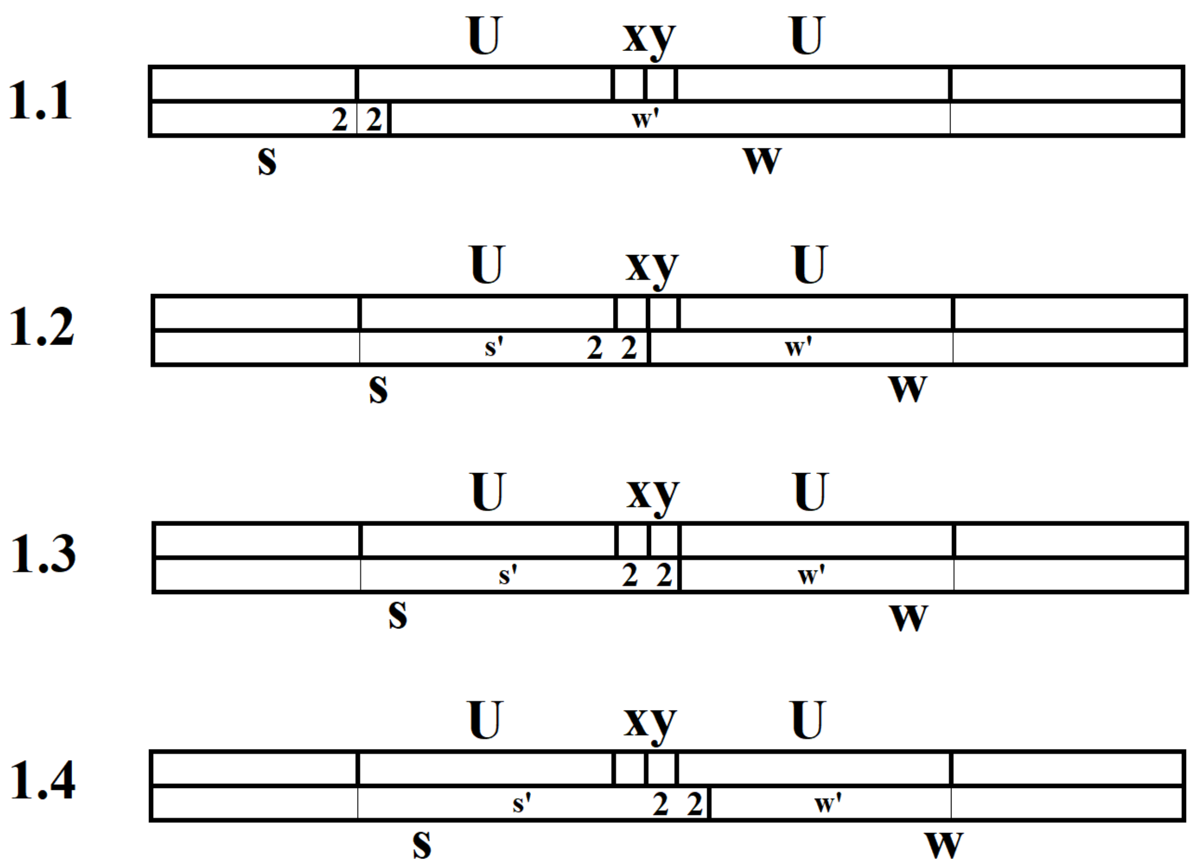

Figure 3.2: Subcases of Case 1

Specifically, in addition to the requirement that $UxyU = s'w'$, the subcases of Case 1 are:

- *Case 1.1*: $|s'| = 1$ and $|w'| = |UxyU| - 1$
- *Case 1.2*: $|s'| = |U| + 1$ and $|w'| = |U| + 1$.
- *Case 1.3*: $|s'| = |U| + 2$ and $|w'| = |U|$.
- *Case 1.4*: $|s'| = |U| + 3$ and $|w'| = |U| - 1$.

Case 1.1: First, suppose $w' \neq w$. Note that 2 is the last letter of w, so that $2w'$ is a factor of $[w]$. Because $[w]$ does not contain $UxyU$, This gives a contradiction. Suppose instead that $w' = w$, so that $UxyU = 2w$. The factor Uxy is a closed walk, so $|U|$ must be even. As an immediate result, $|UxyU|$ is even. However, $|w|$ is even, so $|2w|$ is odd, giving a contradiction.

Case 1.2: As mentioned, $|U|$ must be even. If $|U| = 2$, then $|w'| = 3$, and the length 3 prefix of w' is 123, so that $U = 23$. This is impossible, as the last letter of U must also be 2. If $|U| = 4$, then $w' = 12312$, so that $U = 2312$, and $Uxy = 231221$. However, by Lemma 3.1.1, 231221 is not a closed walk. If $|U| \geq 6$, then w' is of the form $123123h(u)p$, where $u \in A^*$ and p is a prefix of a block, so that $U = 23123h(u)p$. Let U^- be the prefix of U missing only the last letter. Then U^- is of the form $23123h(u)p$, and is a suffix of T, contradicting condition 1 of s.

Case 1.3: We have that U is a suffix of T, and so must end with the letter 1. Then $|U| \neq 2$, because the length 2 prefix of w is 12, which does not end in 1. Additionally, $|U| \neq 4$, as this would imply $U = 1231$, and $Uxy = 123122$ is not a closed walk by Lemma 3.1.1. If $|U| \geq 6$, then $U = w'$, and so $U = 123123h(u)p$ for some $u \in A^*$ and for some prefix of a block p. Therefore $123123h(u)p$ is a suffix of T, contradicting condition 1 of s.

Case 1.4: Here x is a suffix of T, so that $x = 1$. If $|U| = 2$, then $|w'| = 1$, and so $w' = 1$. Then $U = 21$, implying that T has 211 as a suffix, contradicting condition 6. Suppose instead that $|U| = 4$, so that $U = 2123$. Therefore, $Uxy = 212312$, which is not a closed walk by

Lemma 3.1.1. If $|U| = 6$, then $U = 212312$, and $Uxy = 21231212$, which is not a closed walk. If $|U| \geq 8$, so that $U = 2123123h(u)p$ for some $u \in A^*$ and for some prefix of a block p, then $123123h(u)p$ is a suffix of T, contradicting condition 1 of s.

Case 2: *$UxyU = s'ws''$, where s' is a suffix of s, and s'' is a prefix of s.*

Because of the restrictions mentioned in Remark 3.2.1, Case 2 consists of subcases given in Figure 3.3.

Specifically, in addition to the requirement that $UxyU = s'ws''$, the subcases of Case 2 are:

- *Case 2.1*: $|s'| = |s''| = 1$.
- *Case 2.2*: $|s'| = 1$, and $|s''| = |U| + 1$.
- *Case 2.3*: $|s'| = 1$, and $|s''| = |U| + 2$.
- *Case 2.4*: $|s'| = 1$, and $|s''| = |U| + 3$.
- *Case 2.2*: $|s'| = |U| + 1$, and $|s''| = 1$.
- *Case 2.3*: $|s'| = |U| + 2$, and $|s''| = 1$.
- *Case 2.4*: $|s'| = |U| + 3$, and $|s''| = 1$.

In Case 2.1, $UxyU = 2w3$. Note that $w = h(v)$, so $|w| \geq 12$. Therefore $|UxyU| \geq 14$, and $|U| \geq 6$. Because 132 is a suffix of U, 1323 is a suffix of w. However, 1323 does not appear as a factor of w, giving a contradiction.

Case 2.2 is impossible, as it implies that U begins with 2, but also that it begins with 3. Similarly, Case 2.3 requires that U begins with 2, and also that U begins with 1, by condition 7 of s. Case 2.5 implies that U ends with both 2 and 3, and Case 2.6 implies that U ends with both 1 and 3.

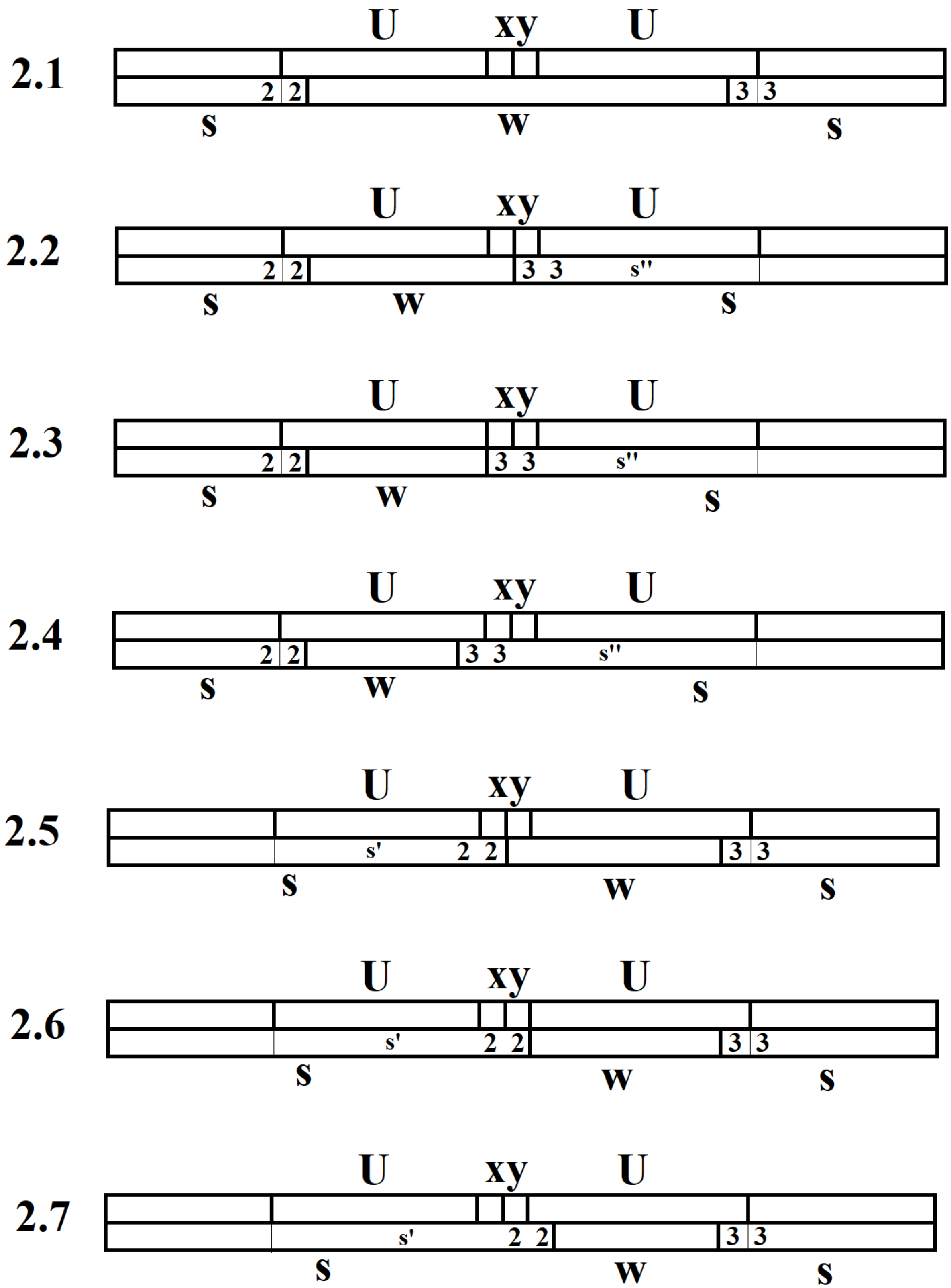

Figure 3.3: Subcases of Case 2

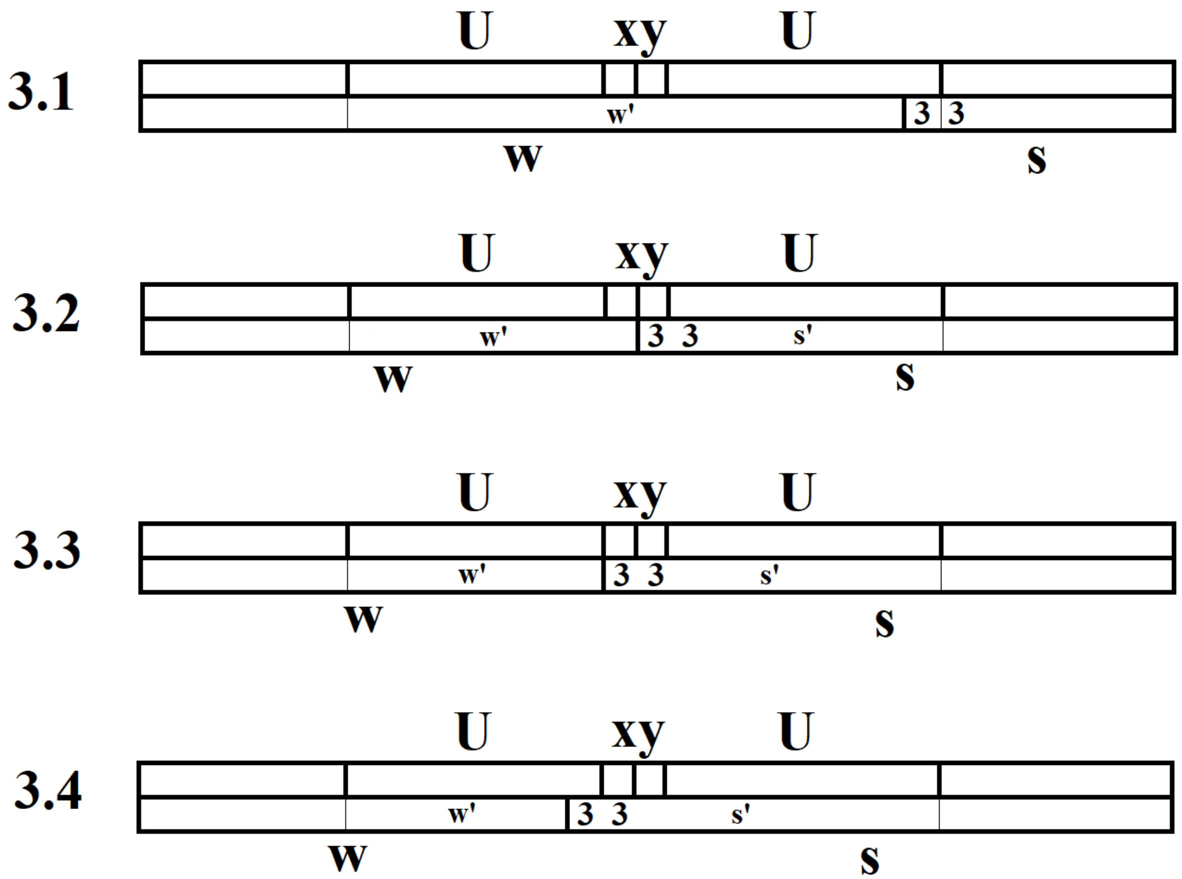

Figure 3.4: Subcases of Case 3

Cases 2.4 and 2.7 both imply that w is a factor of s. Because $w = h(v)$, and $|v| \geq 2$, this contradicts the condition 3 of s.

Case 3: *$UxyU = w's'$, where w' is a suffix of w, and s' is a prefix of s.*

Because of the restrictions mentioned in Remark 3.2.1, Case 3 consists of subcases given in Figure 3.4.

Specifically, in addition to the requirement that $UxyU = w's'$, the subcases of Case 3 are:

- *Case 3.1*: $|s'| = 1$ and $|w'| = |UxyU| - 1$.

- *Case 3.2*: $|s'| = |U| + 1$ and $|w'| = |U| + 1$.

- *Case 3.3*: $|s'| = |U| + 2$ and $|w'| = |U| + 2$.

- *Case 3.4*: $|s'| = |U| + 3$ and $|w'| = |U| + 3$.

Case 3.1: Because w has 132132 as a suffix, U has 323 as a suffix. However, 323 does not appear in w, so Case 3.1 gives a contradiction.

Case 3.2: The length 5 suffix of w is 132 and 32132, so if $|U| = 2$, then $U = 13$. If $|U| = 4$, then $U = 3213$. Because 331 is a prefix of s, U must have 31 as a prefix, so $U \neq 2$ and $U \neq 4$. If $|U| \geq 6$, then $U = qh(u)13213$, where q is the suffix of some block, and $u \in A^*$. Because U begins with 3, q must be nonempty, so let q' be the suffix of q that includes everything but the first 3. Therefore $q'h(u)13213$ is a prefix of T, contradicting condition 2 of s.

Case 3.3: In this case, $U = w'$, and the first letter of U must be 1, as this is the third letter of s. However, the length 2 suffix of w is 32, and the length 4 suffix of w is 2132. Neither of these begin with 1, so $|U| \geq 6$. Therefore, because $s' = w'$, T has $qh(u)123123$ as a prefix, where q is the suffix of some block. This contradicts condition 2 of s.

Case 3.4: If $|U| = 2$, then $|w'| = 1$, so that $w' = 2$ and $Uxy = 2331$, which is not a closed walk by Lemma 3.1.1. Similarly, if $|U| = 4$, $w' = 132$, then $Uxy = 132331$, which is not a closed walk. If $|U| = 6$, $w' = 32132$, then $s' = 331w'3 = 331321323$, implying that t has 132132 as a prefix, contradicting the third condition of s. So, let $|U| \geq 8$, so that $w' = qh(u)132132$, where q is a suffix of a block, and where $u \in A*$. As mentioned, $s' = 331w'3$, meaning that that T has as a prefix $1qh(u)132132$, contradicting condition 2 of s.

Case 4: *$UxyU = w'sw''$, where w' is a suffix of w, and w'' is a prefix of w.*

In Case 4, it is inevitable that either the length 2 prefix or the length 2 suffix of s exist as a factor of U, because only one of these two factors of s coincide in part with xy. This gives a contradiction, as mentioned in Remark 3.2.1.

Because all cases are shown to be impossible, we conclude that $[ws]$ has no factor from $\{11, 222, 223, 322, 333\}$, and no factor of the form $UxyU$ where $U \in S^*$, $x, y \in S$, and Uxy being the label of a closed walk on $\mathcal{G}$. By Lemma 3.1.2, there is no square in $f([ws])$.

□

Table 3.3 gives a list of words that fulfill the requirements of s, found by computer search, along with the size of the images of these words under f.

As mentioned, we extend Lemma 3.2.2 to include words of other lengths not of the form $18n$, with $n \notin \{5, 7, 9, 10, 15, 17\}$.

Corollary 3.2.2. *Let $[v]$ be a square-free circular word over alphabet A with b as a suffix of v and a as a prefix. Let $w \in h(v)$, and let s be a word from Table 3.3. Then $f([ws])$ encodes a level square-free circular word.*

Proof. Let s be a word from Table 3.3, let $\psi = \Delta(f(s))$, and let ψ^- be the length $|\psi| - 2$ prefix of ψ. It may be checked by computer that ψ^- is level, and that ψ ends with ab. Let ω be the decoding of $f(w)$, and let ω^- be the length $|\omega| - 2$ prefix of ω. Recall that $|\omega|_a = |\omega|_b = |\omega|_c$ by Corollary 3.2.1, and that ω ends with ab, by Table 3.1. Let $\alpha, \beta \in A$. By the definition of a level word,

$$|\psi^-|_\alpha - 1 \leq |\psi^-|_\beta \leq |\psi^-|_\alpha + 1$$

$$|\psi^-|_\alpha + |\omega^-|_\alpha - 1 \leq |\psi^-|_\beta + |\omega^-|_\alpha \leq |\psi^-|_\alpha + |\omega^-|_\alpha + 1$$

$$|\psi^-|_\alpha + |\omega^-|_\alpha - 1 \leq |\psi^-|_\beta + |\omega^-|_\beta \leq |\psi^-|_\alpha + |\omega^-|_\alpha + 1$$

$$|\psi^-\omega^-|_\alpha - 1 \leq |\psi^-\omega^-|_\beta \leq |\psi^-\omega^-|_\alpha + 1$$

implying that $\psi^-\omega^-$ is level. The decoding of $f([ws])$, $[\psi^-\omega^-]$, must also be level. Therefore, $[\psi^-\omega^-]$ is a level square-free circular word.

|f(s)|	s
35	331212132122
36	331312312122
37	3312132122
38	331312323122
39	331323232122
40	33131212131122
41	33121213123122
42	33121231232122
43	33121231323122
44	33131323132122
45	33132323123122
46	3312121312132122
47	3312121321232122
48	3312121313232122
49	3312123232312122
50	3312131323232122
51	3312123232323122
52	3313132323232122
53	331212131212323122
54	331212131231323122
55	331212132312323122
56	331212132323232122
57	331213232312323122
58	331313232323123122
59	33121213121313232122
60	33121213123232312122
61	33121213132313232122

|f(s)|	s
62	33121232312323123122
63	33123123231232313122
64	3312123121232121313122
65	3312123121232312132122
66	3312123212131232313122
67	3312123123132132313122
68	3312123213231232313122
69	3312313213231231313122
70	3312312323213232313122
71	331212312132132312132122
72	331212312312132313123122
73	331212312123231232313122
74	331212312313232313123122
75	331212323123231232313122
76	331231231323231232313122
77	331231232323123232313122
78	33121231213213231232313122
79	33121231213213231232313122
80	33121231231232312323123122
81	33121232132323131232313122
82	33121232132323123232313122
83	33123132132323123232313122
84	3312121232132312313213132122
85	3312123121232132312323123122
86	3312123121232323131232313122
87	3312123123132323131232313122
88	3312123123132313123232313122

Table 3.3: Table of s values

□

By Corollary 3.2.2, if v is a square-free circular word beginning with c and ending with b, we may find a level square-free circular word of any length of the form

$$18|v| + i,$$

where $35 \leq i \leq 88$. There is no length such that there is no square-free ternary circular word v of length ℓ, $\ell + 1$, or $\ell + 2$. Therefore, because the interval $[35, 88]$ contains $18 \cdot 3$ consecutive integers, there exists a level square-free ternary circular word of any length at least $18 \cdot 2 + 35 = 71$. An exhaustive computer search shows that a level square-free word exists for each of these lengths besides 5, 7, 9, 10, 14, and 17, giving the required result. These words are given in Appendix A.

Chapter 4

Frankel-Simpson Words

As previously mentioned, there exist arbitrarily large words in Σ_3^* that avoid all squares. On the other hand, every word of Σ_2^* with length 4 or greater has a square. Nonetheless, a result on avoiding squares on the binary alphabet has been found.

Theorem 4.1. *[13] There exists an ω-word on Σ_2^* avoiding all factors of the form XX, with $X \in \Sigma_2^*$ and $X \neq 1, 0, 01$.*

To phrase this more naturally, one can form an arbitrarily long word on 2 letters avoiding all squares, if one allows the squares 00, 11, and 0101.

Definition 7. *A word $w \in \Sigma_2^*$ is called a **Frankel-Simpson** word, or an **FS** word, if w avoids all factors of the form XX, with $X \in \Sigma_{*2}$ and $X \neq 1, 0, 01$. A morphism that takes square-free words to Frankel-Simpson words is called a **Frankel-Simpson** morphism, or an **FS** morphism.*

This definition is extended naturally to the circular case:

Definition 8. *A word $[w]$ with $w \in \Sigma_2^*$ is called a **Frankel-Simpson** circular word, or a **FS** circular word, if every conjugate of w is an FS word.*

Theorem 4.1 extends to the circular case as follows, as will be shown:

Theorem 4.2. *For any length* $\ell \notin \{9, 10, 11, 13, 15, 16, 17, 18, 21, 22, 23, 25, 26, 27, 29, 31, 32, 33, 34, 35, 37, 40, 41, 42, 45, 47, 49, 53, 56, 59, 61, 64, 73, 83, 84\}$, *there exists a FS circular word* $[w]$ *with* $w \in \Sigma_2^\ell$.

To show this, we borrow a method used by Harju and Nowotka to find FS morphisms on four letters. We create square-free circular words on the alphabet $\{a, b, c, d\}$ based on level square-free words, and apply these FS morphisms to show the existence of arbitrarily large FS words. Finally, a computer search characterizes exactly the lengths for which circular FS words exist.

First, we prove a generalization of a method used by Harju and Nowotka. In their original work, the outline of this proof was used to demonstrate the square-freeness of a particular morphism acting on three letters [15]. Here, we use it to find FS morphisms acting on alphabets of any size:

Theorem 4.3. *[15] Let* $n \in \mathbb{N}$ *be such that* $n \geq 3$. *Suppose* $f : \Sigma_n^* \to \Sigma_2^*$ *is a morphism satisfying these properties:*

1. *For any square-free* $v \in \Sigma_n^3$, $f(v)$ *is an FS word.*
2. *There is a word* $p \in \Sigma_2^*$, $|p| \geq 3$, *such that:*
   (a) *For each* $a \in \Sigma_n$, p *is a prefix of* $f(a)$.
   (b) *If* $a_i \in \Sigma_n$, $1 \leq i \leq \ell$, *and* $f(a_1 a_2 \cdots a_\ell) = qpr$ *for some words* $q, r \in \Sigma_2^*$, *then* $q = \epsilon$ *or* $q = f(a_1 a_2 \cdots a_j)$, *some* $j \leq \ell$.

Then f *is an FS morphism.*

Proof. To begin with, note that the conditions imply that if $a, b \in \Sigma_n$ and $f(a)$ is a prefix of $f(b)$, then $a = b$. Otherwise, aba is a square-free word of length 3, with square prefix $f(a)f(a)$. However, $|f(a)| \geq |p| \geq 3$, so $f(a)f(a) \neq 00, 11$, or 0101. This contradicts condition 1.

For the sake of getting a contradiction, consider a square-free word $w_1w_2\cdots w_m$, with the $w_i \in \Sigma_n$, such that $f(w_1w_2\cdots w_m)$ contains a square xx, $x \neq \epsilon, 0, 1, 01$. Let m be as small as possible. By condition 1, $m \geq 4$. Since m is minimal, write

$$xx = W_1''W_2\cdots W_m',$$

Where, $W_i = f(w_i)$ for all $1 \leq i \leq m$, where W_1'' is a nonempty suffix of W_1, and where $W + m'$ is a nonempty prefix of W_m.

As per condition 2(a), write $W_2 = pW_2''$.

Case 1: $|x| < |W_1''|$ **or** $|x| < |W_m'|$

If $|x| < |W_1''|$, let W_1''' be the nonempty suffix of W_1'' so that $W_1'' = xW_1'''$. Then we find the second copy of x in xx can be written

$$x = W_1'''W_2\cdots W_m' = W_1'''pW_2''\cdots W_m'.$$

However, then

$$f(w_1) = W_1'W_1'' = W_1'xW_1''' = W_1'W_1'''pW_2''\cdots W_m'W_1'''$$

contains an instance of p at an index which contradicts condition 2(b).

Similarly, if $|x| < |W_m'|$, let W_m''' be such that $W_m' = W_m'''x$, and $W_m''' \neq \epsilon$. Then we find the first copy of x in xx can be written

$$x = W_1''W_2\cdots W_m''' = W_1'''pW_2''\cdots W_m'''.$$

However, then

$$f(w_m) = W_m'W_m'' = W_m'''xW_m'' = W_m'''W_1''pW_2''\cdots W_m'''W_m''$$

contains an instance of p at an index which contradicts condition 2(b).

Case 2: $|x| \geq |W_1''|, |W_m'|$

In this case we can write

$$\begin{aligned} x &= W_1'' \cdots W_j' \\ &= W_j'' \cdots W_m', \end{aligned}$$

for some j, $1 < j < m$, with $W_j = W_j' W_j''$.

If $j > 2$, then there is at least one instance of p in $x = W_1'' W_2 \cdots W_j'$, appearing as a prefix of W_2. On the other hand, if $j = 2$, then an instance of p appears as a prefix of W_{j+1} in $x = W_j'' W_{j+1} \cdots W_m$. In either case, there is at least one instance of p in x. For the sake of definiteness, adjusting notation if necessary, choose j so that $W_j' = \epsilon$ if x starts with p; that is, assume in all cases that $W_j'' \neq \epsilon$.

Case 2(a): Word x starts with p.

If x starts with p, then condition 2(b) forces $W_1'' = W_1$. Our choice of notation gives $W_j'' = W_j$. Since W_1 and W_j are prefixes of x, one must be a prefix of the other, and, as noted at the beginning of this proof, this forces $w_1 = w_j$. Therefore $W_1 = W_j$.

We prove by induction that for $1 \leq i \leq j-2$, $w_1 \cdots w_i = w_j \cdots w_{j+i-1}$, and $W_{i+1} \cdots W_{j-1} = W_{j+i} \cdots W_m'$. The base case of this induction, when $i = 1$, has just been shown.

Suppose that for some k, $1 \leq k < j-2$, we have $w_1 \cdots w_k = w_j \cdots w_{j+k-1}$, and $W_{k+1} \cdots W_{j-1} = W_{j+k} \cdots W_m'$. Then one of W_{k+1} and W_{j+k} is a prefix of the other, giving $w_{k+1} = w_{j+k}$, yielding the induction step.

Setting $i = j-1$, we see that $w_1 \cdots w_{j-1} = w_j \cdots w_{2j-2}$. However, now w contains the square $(w_1 \cdots w_{j-1})^2$. Since $|w_1| \geq |p| = 3$, this is a contradiction.

Case 2(b): Word x doesn't start with p.

The first p in x is at the beginning of W_2: If $x = W_1''W_2 \cdots W_j'$ has an instance of p of index i, $1 < i < |W_1'| + 1$, then $f(w_1w_2)$ contains an instance of p of index properly between 1 and $|f(w_1)| + 1$, violating property 2(b). Thus the least index of p in x is $|W_1'| + 1$. However, an analogous argument observing that $x = W_j''W_{j+1} \cdots W_m'$ yields least index of $p = |W_j''| + 1$. Thus W_1'' and W_j'' are prefixes of x with the same length, forcing $W_1'' = W_j''$. Now, $W_2 \cdots W_j' = W_{j+1} \cdots W_m'$, so that one of W_2 and W_{j+1} is a prefix of the other, forcing $w_2 = w_{j+1}$.

We prove by induction that for $2 \leq i \leq j-2$, $w_2 \cdots w_i = w_{j+1} \cdots w_{j+i-1}$, and $W_{i+1} \cdots W_{j-1}W_j' = W_{j+i} \cdots W_{m-1}W_m'$. We have just established the base case of this induction, when $i = 2$.

Suppose that for some k, $1 \leq k < j-1$, we have $w_2 \cdots w_k = w_{j+1} \cdots w_{j+k-1}$, and $W_{k+1} \cdots W_{j-1}W_j' = W_{j+k} \cdots W_{m-1}W_m'$. Then one of W_{k+1} and W_{j+k} is a prefix of the other, giving $w_{k+1} = w_{j+k}$, yielding the induction step.

When $i = j-1$, we find $W_j' = W_m'$. Since one of W_j and W_m must be a prefix of the other, $w_j = w_m$. Then w contains the square $w_2 \cdots w_j w_{j+1} \cdots w_m = (w_2 \cdots w_j)^2$. Since $|w_2| \geq |p| = 3$, this is a contradiction.

□

This theorem can be used to find linear FS words. In general, theorems used to find linear words possessing a certain property can be used to find circular words with that property by the following lemma, due to Narad Rampersad [23]:

Lemma 4.3.1. *If f is a square-free morphism from Σ_n to Σ_m, and $[w]$ is a square-free circular word with $|w| \geq 2$, then $f([w])$ is a square-free circular word.*

Proof. We prove this by contradiction. Write $w = w_1w_2 \cdots w_\ell$, $w_\ell \in \Sigma_n$. Let $f(w_i) = W_i$, $1 \leq i \leq \ell$. Suppose there is a square in $f(w)$. Replacing w with one of its conjugates if necessary, we can assume that $W_1''W_2 \cdots W_\ell W_1'$ is a linearization of $[f(w)]$ containing a square, where $W_1 = W_1'W_1''$. Then $W_1W_2 \cdots W_\ell W_1 = f(w_1w_2 \cdots w_\ell w_1)$ also contains this square. Since f is square-free, this implies that $w_1w_2 \cdots w_\ell w_1$ contains some square xx. Both $w_1w_2 \cdots w_\ell$ and $w_2 \cdots w_\ell w_1$ are linearizations of w, and are thus square-free. It follows that $xx = w_1w_2 \cdots w_\ell w_1$. However, x then begins and ends with letter w_1, so that w_1w_1 appears at the center of xx, whence w contains the square w_1w_1.This is a contradiction. □

This may be expanded to account for FS morphisms

Corollary 4.3.1. *If f is a FS morphism from Σ_n to Σ_2, and $[w]$ is a square-free circular word with $|w| \geq 2$, then $f([w])$ is an FS circular word.*

Proof. Repeat the previous proof, replacing 'containing a square' by 'containing a square other than 00, 11, 0101', and 'square-free' by 'an FS morphism'. □

As a result of these two observations, we form a light corollary:

Corollary 4.3.2. *If f is a square-free morphism from Σ_n to Σ_m, and $[w]$ is square free circular word with $|w| \geq 2$, then $f([w])$ is a square free circular word.*

Combining Corollary 4.3.2 with Theorem 4.3 trivially gives the following corollary:

Corollary 4.3.3. *Let Σ_n be an alphabet on n letters. Suppose $f : \Sigma_n^+ \to \Sigma_2^+$ is a morphism satisfying the properties described in Theorem 4.3. If $[w]$ is a square free circular word, then $f([w])$ is a FS circular word.*

This result allows for easy computer searches of FS morphisms. We introduce the following result to find the lengths of circular FS words generated by a FS morphism.

Lemma 4.3.2. *Let $f : \Sigma_4 \to \Sigma_2^*$ be a Frankel-Simpson morphism with $|f(a)| = \alpha$, $|f(b)| = \beta$, $|f(c)| = \gamma$, and $|f(d)| = \delta$. Let $\phi \in \{-1, 0, 1\}$, let $r \in \{a, b, c\}$, let $n \in \mathbb{N}^+$, and let $i \in \mathbb{N}$. Suppose the following hold:*

- $\phi + 3n \notin \{5, 7, 9, 10, 14, 1\}$.
- *If* $r \neq a$, *then* $i \leq n$.
- *If* $r = a$, *then* $i \leq n + \phi$.

Then there exists a circular Frankel-Simpson word of length

$$n(\alpha + \beta + \gamma) + \phi\alpha + i(\delta - |f(r)|)$$

There also exist circular Frankel-Simpson words of lengths described by this expression on any permutation of $\{\alpha, \beta, \gamma, \delta\}$.

Proof. Let $f : \Sigma_4 \to \Sigma_2^*$ be an FS morphism with $|f(a)| = \alpha$, $|f(b)| = \beta$, $|f(c)| = \gamma$, and $|f(d)| = \delta$. By Theorem 3.1, there is some $w \in A^*$ such that $[w]$ is a level circular word, with $|w|_a = n + \phi$, while $|w|_b, |w|_c = n$. Let $r \in \{a, b, c\}$. Note that $|w|_r \geq i$. Let w' be formed by replacing any i instances of r with d. Then

$$|f(w')| = \alpha|w'|_a + \beta|w'|_b + \gamma|w'|_c + \delta|w'|_d$$

Suppose $r = a$, so that $\alpha = |f(r)|$. Then

$$\begin{aligned}
|f(w')| &= \alpha|w'|_a + \beta|w'|_b + \gamma|w'|_c + \delta|w'|_d \\
&= \alpha(n + \phi - i) + \beta n + \gamma n + \delta i \\
&= \alpha n + \phi\alpha - i\alpha + \beta n + \gamma n + \delta i \\
&= n(\alpha + \beta + \gamma) + \phi\alpha - i|f(r)| + \delta i \\
&= n(\alpha + \beta + \gamma) + \phi\alpha + i(\delta - |f(r)|).
\end{aligned}$$

Suppose instead that $r = b$ and $\beta = |f(r)|$. Then

$$\begin{aligned} |f(w')| &= \alpha|w'|_a + \beta|w'|_b + \gamma|w'|_c + \delta|w'|_d \\ &= \alpha(n+\phi) + \beta(n-i) + \gamma n + \delta i \\ &= \alpha n + \phi\alpha - i\beta + \beta n + \gamma n + \delta i \\ &= n(\alpha+\beta+\gamma) + \phi\alpha - i|f(r)| + \delta i \\ &= n(\alpha+\beta+\gamma) + \phi\alpha + i(\delta - |f(r)|). \end{aligned}$$

The analogous argument shows that $f(w') = n(\alpha+\beta+\gamma) + \phi\alpha + i(\delta - |f(r)|)$ if $r = c$.

Let p be a permutation on $\{\alpha, \beta, \gamma, \delta\}$, and let f_p be the FS function formed from $f_p(a) = p(\alpha)$, $f_p(b) = p(\beta)$, $f_p(c) = p(\gamma)$, and $f_p(d) = p(\delta)$. As a result,

$$|f_p(w')| = n(p(\alpha) + p(\beta) + p(\gamma)) + \phi p(\alpha) + i(p(\delta) - p(\rho))$$

satisfying the initially required condition. □

We performed a computer search for morphisms that satisfy the requirements outlined in Corollary 4.3.3. Selected results of this search are given in Table 4.1.

Finally, we restate our central theorem with proof:

Theorem 4.4. *For any length* $\ell \notin \{9, 10, 11, 13, 15, 16, 17, 18, 21, 22, 23, 25, 26, 27, 29, 31, 32, 33, 34, 35, 37, 40, 41, 42, 45, 47, 49, 53, 56, 59, 61, 64, 73, 83, 84\}$*, there exists a FS circular word* $[w]$ *with* $w \in \Sigma_2^\ell$.

Proof. Let ℓ be a length with $\ell \geq 7350$. Then let $\lambda = \left\lfloor \frac{\ell}{50} \right\rfloor$. If λ is a multiple of 3, let $\phi = 0$, while if λ is one greater than multiple of 3, let $\phi = 1$, and if λ is one less than a multiple of 3, let $\phi = -1$. Then $\lambda - \phi$ is a multiple of 3, so let $n = \frac{\lambda-\phi}{3}$. Because $\ell \geq 7350$, $\left\lfloor \frac{\ell}{50} \right\rfloor = \lambda \geq 147$, and so $\frac{\lambda-\phi}{3} = n \geq 49$. Then let $i = \ell - 50\left(\left\lfloor \frac{\ell}{50} \right\rfloor\right)$. We have that $i \leq 49$, so that $i \leq n$.

$f_1(a) = 011001110001100101110001$ $f_1(b) = 011001110001100101100010111001$ $f_1(c) = 01100111000101110010110001011100011001011000111001$ $f_1(d) = 011001110001011100101100011100101110001011000111001$	$f_8(a) = 01100111000101110010110001011100011001011000111001$ $f_8(b) = 01100111000110010110001011100101100011100101110001$ $f_8(c) = 01100111000110010111000101100011100101100010111001$ $f_8(d) = 011001110001011100101100011100101110001011000111001$
$f_2(a) = 0110011100010111001$ $f_2(b) = 011001110001011000111001$ $f_2(c) = 0110011100011001011000111001$ $f_2(d) = 011001110001100101100010111001$	$f_9(a) = 0110011100010111001$ $f_9(b) = 011001110001100101110001$ $f_9(c) = 011001110001100101100011100101110001$ $f_9(d) = 0110011100011001011000101110001100101110001$
$f_3(a) = 0110011100011001011000111001$ $f_3(b) = 011001110001100101100010111001$ $f_3(c) = 01100111000101100011100101110001$ $f_3(d) = 011001110001100101110001011000111001$	$f_{10}(a) = 011001110001100101110001$ $f_{10}(b) = 011001110001011100101100011100101110001$ $f_{10}(c) = 01100111000101110010110001011100011001011000111001$ $f_{10}(d) = 01100111000110010110001011100101100011100101110001$
$f_4(a) = 0110011100010111001$ $f_4(b) = 011001110001100101100010111001$ $f_4(c) = 011001110001100101100011100101110001$ $f_4(d) = 011001110001100101110001011000111001$	$f_{11}(a) = 0110011100010111001$ $f_{11}(b) = 01100111000101100011100101100010111001$ $f_{11}(c) = 01100111000101100011100101110001100101100010111001$ $f_{11}(d) = 01100111000110010111000101100011100101100010111001$
$f_5(a) = 0110011100010111001$ $f_5(b) = 011001110001100101100010111001$ $f_5(c) = 01100111000101100011100101110001$ $f_5(d) = 011001110001100101110001011000111001$	$f_{12}(a) = 0110011100010111001$ $f_{12}(b) = 01100111000101100011100101100010111001$ $f_{12}(c) = 01100111000101100011100101110001100101100010111001$ $f_{12}(d) = 01100111000110010111000101100011100101100010111001$
$f_6(a) = 0110011100010111001$ $f_6(b) = 011001110001100101100010111001$ $f_6(c) = 011001110001011000111001011100011001011000111001$ $f_6(d) = 011001110001100101100011100101110001011000111001$	$f_{13}(a) = 011001110001100101110001$ $f_{13}(b) = 011001110001100101100010111001$ $f_{13}(c) = 011001110001011100101100011100101110001$ $f_{13}(d) = 01100111000101110010110001011100011001011000111001$
$f_7(a) = 0110011100010111001$ $f_7(b) = 011001110001011000111001011100011001011000111001$ $f_7(c) = 01100111000101100011100101110001100101100010111001$ $f_7(d) = 01100111000110010111000101100011100101100010111001$	$f_{14}(a) = 011001110001100101110001$ $f_{14}(b) = 011001110001100101100011100101110001$ $f_{14}(c) = 01100111000101110010110001011100011001011000111001$ $f_{14}(d) = 01100111000110010110001011100101100011100101110001$

Table 4.1: A selection of FS morphisms

Morphism f_8 in Table 4.1 has $|f_8(a)| = 50$, $|f_8(b)| = 50$, $|f_8(c)| = 50$, and $|f_8(d)| = 51$. By Lemma 4.3.2, there is a FS circular word of length

$$n(50 + 50 + 50) + \phi 50 + i(51 - 50) = n(150) + \phi 50 + i.$$

Plugging in the values for n, ϕ, and i described above shows that f_8 gives a FS circular word of size

$$\begin{aligned} n(150) + \phi 50 + i &= 50(3n + \phi) + (\ell - 50\lambda) \\ &= 50\left(3\frac{\lambda - \phi}{3} + \phi\right) + (\ell - 50\lambda) \\ &= 50\left(\lambda - \phi + \phi\right) + (\ell - 50\lambda) \\ &= 50\lambda + \ell - 50\lambda \\ &= \ell \end{aligned}$$

In a similar way, the morphisms in Table 4.1 (and the morphisms obtained from them by permuting a, b, c, d) give FS circular words of every length less than 7350 except:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 72, 74, 75, 76, 78, 80, 83, 84, 88, 89, 108, 114, 173

Table 4.2: Lengths for which no morphism given provides a FS circular word

We performed a computer search where every FS word with a length in Table 4.2 was checked to see whether it is a circular FS word. This found an example for all lengths besides

9, 10, 11, 13, 15, 16, 17, 18, 21, 22, 23, 25, 26, 27, 29, 31, 32, 33, 34, 35, 37, 40, 41, 42, 45, 47, 49, 53, 56, 59, 61, 64, 73, 83, 84

Table 4.3: Lengths for which there exists no FS circular word

The specific examples of each length of FS word found by this search are given in Appendix C. The remaining lengths are shown not to have any circular Frankel-Simpson words. Therefore, we have proven Theorem 4.2.

□

Chapter 5

Open Problems

Let p and w be words. We say that p is a **pattern**, and call the image of p under some non-erasing morphism an **instance** of p. The word w **avoids** p if there is no factor of w that is an instance of p. The pattern p is k-avoidable if there are arbitrarily long words of Σ_k^* which avoid p.

Let t be the Thue-Morse word. For any odd index i, we note that $t_i t_{i+1} \in \{10, 01\}$. If v is a factor of t, and $|v|$ is odd, then there is some word $v' \in \{10, 01\}^*$ and some $x \in \Sigma_2$ such that either $v = v'x$ or $v = xv'$. In either case, v is level. As mentioned previously, for every length ℓ, there is a factor v of t such that $|v| = l$, and $[v]$ avoids $\frac{5}{2}+$ powers. Combining these two observations gives

Remark 5.0.1. For any odd length ℓ, there is a level binary circular word $[v]$ of length ℓ avoiding $\frac{5}{2}+$ powers.

As demonstrated, if there exists a square-free ternary circular word of length ℓ, there is a length ℓ level square-free circular ternary word. Based on this sparse evidence, we form the following conjecture

Conjecture 5.0.1. *Let p be a pattern, and let $[w]$ be a circular word on k letters that avoids p. Then there is some level circular word $[w']$ on k letters of length $|w|$ that avoids p.*

We are also interested in the opposite question: how far can a square-free circular words be from level? This may be formalized as follows:

Question 5.0.1. *Given a length ℓ, what is the largest ratio r such that there is a length ℓ square-free circular word w with $\frac{|w|_a}{|w|} \geq r$?*

Along similar lines, one may ask:

Question 5.0.2. *Given a length ℓ, what is the smallest ratio r such that there is a length ℓ square-free circular word w with $\frac{|w|_a}{|w|} \leq r$?*

We also note that if a circular word of length ℓ on alphabet Σ avoids a pattern p, then p is also avoided on circular words of length ℓ on larger alphabets. This is not trivial if the requirement that the word be level is added. Nonetheless, we conjecture:

Conjecture 5.0.2. *For any integer $k \geq 3$, there is a level square-free circular word on Σ_k of any length greater than 18.*

As mentioned, there are ternary circular words avoiding squares of any length greater than 17. Additionally, there is a binary circular word in the Thue-Morse word avoiding $\frac{5}{2}+$ powers. Inspired by these examples, Currie has formed the following conjecture:

Conjecture 5.0.3. *[9] if a pattern p is k-avoidable, there are arbitrarily long circular words on k letters avoiding p.*

Bibliography

[1] A. Aberkane, J. D. Currie, The Thue–Morse word contains circular $5/2^{+}$ power free words of every length *Theoretical Computer Science*, **11** (2004) 573–581

[2] A. Aberkane, J. D. Currie, Attainable lengths for circular binary words avoiding k powers *Bulletin of the Belgian Mathematical Society*, **12(4)** (2005) 525–534

[3] M. Amit, P. Gawrychowski, Distinct Squares in Circular Words, In: Fici G., Sciortino M., Venturini R. (eds) String Processing and Information Retrieval. SPIRE 2017. Lecture Notes in Computer Science, vol 10508. Springer, Cham

[4] J. Berstel, D. Perrin, The origins of combinatorics on words *European Journal of Combinatorics* **28(3)** (2005) 996–1022

[5] A. Blondin Massé, S. Brlek, A. Garon, S. Labbé, Equations on palindromes and circular words *Theoretical Computer Science*, **412** (2011) 2922–2930

[6] S. Brlek, S. Hamel, M. Nivat, C. Reutenauer On the palindromic complexity of infinite words *Internat. J. Found. Comput. Sci.*, **15** (2004) 293–306

[7] G. Castiglione, A.Restivo, M.Sciortino, Circular sturmian words and Hopcroft's algorithm *Theoretical Computer Science*, **410(10)** (2009) 4372–4381

[8] J. Currie, There Are Ternary Circular Square-Free Words of Length n for $n \geq 18$, *Electr. J. Comb.* **9** (2002), 10 (electronic)

[9] J. D. Currie, Pattern avoidance: themes and variations, *Theoretical Computer Science* **339(1)**, 7–18 (2005)

[10] J. Currie, N. Rampersad, T. Harju & P. Ochem, Some further results on squarefree arithmetic progressions in infinite words, *Theoretical Computer Science* **799(12)** (2019) 140–148

[11] N. G. de Bruijn, A combinatorial problem *Nederl. Akad. Wetensch. Proc* **49(7)** (1946) 758–764

[12] F. Dejean, Sur un théorème de Thue, *Journal of Combinatorial Theory*, **13(1)** (1972) 90-99

[13] A. Frankel, J. Simpson, How Many Squares Must a Binary Sequence Contain?, *Electr. J. Comb.* 2(R2):9pp, (1995)

[14] A. Fraenkel and J. Simpson, How many squares can a string contain? *J. Comb. Theory A*, **82(1)** (1998) 112–120

[15] T. Harju, D. Nowotka, Binary Words with Few Squares. *Bulletin of the EATCS* 89 (2006)

[16] T. Harju, Disposability in Square-Free Words, (2019) arXiv:1911.09973v1

[17] A. Khalyavin, The minimal density of a letter in an infinite ternary square-free word is $\frac{883}{3215}\ldots$ *Journal of Integer Sequences* **10(6)**(2007) Article 07.6.5

[18] Lyngsø, R.B., Pedersen, C.N. Protein folding in the 2D HP model *Tech. rep., University of Aarhu*, **16(6)** (1999) 15–99

[19] P. A. MacMahon, Application of a theory of permutations in circular procession to the theory of numbers *Proc. London Math. Soc.* **s1-23(1)** (1892) 305-–318

[20] M. H. Martin, A problem in arrangements *Bull. Amer. Math. Soc.* **58(4)** (1934) 727–734

[21] C. Müllner, A. Rhyzikov, Palindromic Subsequences in Finite Words, arXiv:1901.07502 (2019)

[22] J.-J. Pansiot, A propos d'une conjecture de F. Dejean sur les répétitions dans les mots, *Discr. Appl. Math.* **7** (1984), 297–311.

[23] N. Rampersad, Personal communication.

[24] C. F. Sainte-Marie, Solution to problem number 58 *L'intermédiaire des mathématiciens* (1894)

[25] D. Sean Fitzpatrick, There are binary cube-free circular words of length n contained within the Thue Morse word for all positive integers n *Ars Combinatoria*, **74(1)**(2005)

[26] J. Shallit, R. Zarifi Circular critical exponents for Thue–Morse factors, *RAIRO-Theor. Inf. Appl.*, **53** (2019) 37–49

[27] A. M. Shur, On ternary square-free circular words, *The electronic Journal of Combinatorics* **17(1)** (2010) R140

[28] J. Simpson, Palindromes in circular words *Theoretical Computer Science*, **550** (2014) 66–78

[29] Y. Tarannikov, The minimal density of a letter in an infinite ternary square-free word is 0.2746... *Journal of Integer Sequences* 5(2): Article 02.2.2 (2002) Article 02.2.2

[30] A. Thue, Über unendliche Zeichenreihen *Norske Vid. Selsk. Skr. Kl.* **7** (1906) 1–22

[31] A. Thue, Über die gegenseitige Loge gleicher Teile gewisser Zeichenreihen *Norske Vid. Selsk. Skr.* **1** (1912) 1–67

[32] E. Witt, Treue Darstellung Lieschen Ringe *J. Reine Angew. Math.* **177(1)** (1937) 152–160

Appendix A

A List of Level Words

$4 = |f(11)|$

$6 = |f(22)|$

$8 = |f(33)|$

$11 = |f(1213)|$

$12 = |f(1232)|$

$13 = |f(1323)|$

$15 = |f(121212)|$

$16 = |f(122122)|$

$19 = |f(133213)|$

$20 = |f(133232)|$

$21 = |f(232323)|$

$22 = |f(12312113)|$

$23 = |f(13121232)|$

$24 = |f(13121323)|$

$25 = |f(13132133)|$

$26 = |f(13132323)|$

$27 = |f(12332323))|$

$28 = |f(13323233)|$

$29 = |f(1212132133)|$

$30 = |f(1212132323)|$

$31 = |f(1212133233)|$

$32 = |f(1212323233)|$

$34 = |f(1232323323)|$

$35 = |f(1323233233)|$

$37 = |f(121212323323)|$

$38 = |f(121231233233)|$

$39 = |f(121233133233)|$

$40 = |f(131313233233)|$

$41 = |f(131233233233)|$

$42 = |f(123233233233)|$

$43 = |f(12212123313233)|$

$45 = |f(12121233233233)|$

$$53 = |f(1212323313323323)|$$

$$54 = |f(131313123123132132)|$$

$$55 = |f(1231323233233233)|$$

$$56 = |f(1332323233133233)|$$

$$57 = |f(122121213233233233)|$$

$$58 = |f(122121323313323323)|$$

$$59 = |f(122121323233233233)|$$

$$60 = |f(123131232323323323)|$$

$$61 = |f(123132323313323233)|$$

$$62 = |f(133123233133233233)|$$

$$63 = |f(1332323231332332 33)|$$

$$64 = |f(1332332331332332 33)|$$

$$65 = |f(12123121323233233233)|$$

$$66 = |f(12123123233133233233)|$$

$$67 = |f(12123313233233233133)|$$

$$68 = |f(12313133232323233233)|$$

$$69 = |f(12312213123122131231 2213)|$$

$$70 = |f(12312213123131231231 2213)|$$

Appendix B

Alternate Proof of Theorem 3.2

First, we create a counterpart to Lemma 3.1.3:

Lemma B.0.1. *The morphism h has the following properties:*

1. *Let $x, y \in A$. Let $s \in S^2$. If s is a suffix of both of $h(x)$ and $h(y)$, then $x = y$. Thus each letter $x \in A$ is determined by the length 2 suffix of $h(x)$.*
2. *Let $x, y \in A$. Let $p \in S^3$. If p is a prefix of both of $h(x)$ and $h(y)$, then $x = y$. Thus each letter $x \in A$ is determined by the length 3 prefix of $h(x)$.*
3. *Let $x, y \in A$, and suppose $s \in S^2$ is a suffix of $h(x)$, $p \in S^3$ is a prefix of $h(y)$. Suppose $u, v \in A$, $u \neq v$, and $h(uv) = \rho s p \sigma$, some $\rho, \sigma \in S^*$. Then $u = x$, $v = y$, and $h(u) = \rho s$, $h(v) = p\sigma$.*

Proof. The first two of these properties were shown in Lemma 3.1.3

This can be verified by exhaustively listing the possible length 2 suffixes s and length 3 prefixes p, and noting that sp only ever occurs in $h(w)$ with s as suffix of a block and p as prefix of a block is w is a square-free word of A^2. For example, if $s = 23$ (the length 2 suffix of $h(a)$) and $p = 131$ (a length 3 prefix of $h(c)$), then the first 1 in $sp = 23131$ can never be one of the underlined 1's in $h(a) = \underline{1}23\underline{1}23$, $h(b) = \underline{1}32\underline{1}32$, $\underline{1}3\underline{1}3\underline{1}3 \in h(c)$, or $12\underline{1}2\underline{1}2 \in h(c)$, so that in any occurrence of sp in $h(w)$, the first 1 must occur as a prefix of $h(c)$, and the result holds. The requirement that the letters u and v of w are distinct rules out cases such as $h(aa) = 1231\underline{23123}123$.

□

Theorem 3.2 is:

Theorem B.1. *Let $[v]$ be a square-free circular word over A. Let $[w] = h([v])$. Then $f([w])$ encodes a square-free circular word.*

Proof. By Lemma 3.1.2, it suffices to show that

- $[w]$ has no factor 11, 222, 223, 322, or 333
- $[w]$ has no factors $UxyU$ with $x, y \in S$, $U \in S^*$ where $|U| \geq 2$, and where Uxy is the label of a closed walk.

Let $[v]$ be a square-free circular word over A, and let $[w] = h([v])$. No element of $\{11, 222, 223, 322, 333\}$ appears as a factor of a concatenation of any two blocks, so no such element appears as a factor of $h(v)$.

Suppose for the purpose of finding a contradiction that $UxyU$ is a factor of $[w]$, where $x, y \in S$, $U \in S^*$, and Uxy is a closed walk on $\mathcal{G}$. Note that closed walks on $\mathcal{G}$ have even length, so $|Uxy|$ is even, and hence $|U|$ is even.

Suppose $|U| = 2$, so that $|Uxy| = 4$. Because Uxy is closed, the possible values of $|Uxy$ are:

$$\{1323, 1312, 1213, 1232, 3231, 3212, 3132, 3121, 2321, 2313, 2131, 2123\}$$

Of these, only 1312, 3212, 3132, 2313, 2131, and 2123 appear at all inside $[w]$. Therefore, the possible values of $UxyU$ are:

$$\{131213, 321232, 313231, 231323, 213121, 212321\}$$

Of these, all contain either 121, 323, or 232. None of these factors appear inside $[w]$, so $|U| \neq 2$.

Suppose instead that $|U| = 4$, so that $|Uxy| = 6$. Up to conjugation, the closed walks of length 6 are:

$$\{131313, 121212, 123123, 132132, 232323\}$$

However, every conjugate of 121212 contains 121, and every conjugate of 232323 contains 232, so Uxy is not a conjugate of either of these. Therefore, the possible values of Uxy are:

$$\{131313, 313131, 123123, 231231, 312312, 132132, 321321, 213213\}$$

and the coresponding values of $UxyU$ are:

$$\{1313131313, 3131313131, 1231231231, 2312312312,$$

$$3123123123, 1321321321, 3213213213, 2132132132\}$$

Because each block is uniquely determined by the length 2 suffix or length 3 prefix, we can determine that the first two of these values implies that cc is a factor of $[w]$, while the next 3 imply a factor of aa, and the final 3 imply a factor of bb. Each of these cases contradict the square-freeness of $[w]$, so $|U| \neq 4$.

We conclude that $|U| \neq 2, 4$, so that $|U| \geq 6$. It follows that U is not an internal factor of a block.

Write $UxyU = sh(u)pp'h(v)t$ where for some letters $\sigma, \rho, \tau \in A$, we have $h(\sigma) = s's$, $h(\rho) = pp'$, $h(\tau) = tt'$, $s', p', t' \neq \epsilon$; $Uxy = sh(\nu)p$, $U = p'h(\omega)t$; and both ρu and $\nu\rho\upsilon\tau$ are factors of $[w]$.

Claim B.1.1. *Either $s = p = \epsilon$ or $s = p'$. It follows that $|Uxy| \equiv 0 \pmod 6$.*

Proof of claim: By Remark 3.1.1, $h(u)$ is the label of a closed walk. Since $Uxy = sh(u)p$ is also the label of a closed walk, we conclude that sp is also the label of a closed walk. By the symmetry of $\mathcal{G}$, we can assume that sp is the label of a closed walk which arrives at vertex A after crossing edges labeled by s. This means that walking the edges labeled by either s^R or p, starting at A, we arrive at the same vertex.

We form cases based on the vertex we arrive at walking the labels of s^R, starting at A:

Case: We arrive at F.

Suppose that, walking the labels of s^R, starting at A, takes us to F. The possible values of s are then 123, 32, 313, and 2. Recall that s^R indicates the reverse of s. Since sp labels a closed walk, the walk p also takes us from A to F. Considering the possible walks from A to F, we see that the candidate values for p are 123, 131, 313, 321, 313, 13213, 31231, 32323 and 12121. Of these, only 123, 131, and 13213, appear as prefixes of a block.

An additional consideration allows us to rule out several of these possibilities; since $sh(u)p = Uxy$, and $p'h(\omega)t = U$, we conclude that one of s and p' is a prefix of the other.

- In the case $s = 123$, we conclude that the first letter of p' is 1. Recall that the candidate values for p are 13213, 131, and 123. Then $h(\rho) = pp'$ has prefix 132131, 1311, or 1231; however, neither 132131 nor 1311, is a prefix of the image of a letter, so that if $s = 123$, the only one of the four possibilities for p which is allowable is that $p = 123$, so that $p' = 123 = s$.

- If $s = 32$, then pp' must have prefix 1233, 132133, or 13132, none of which is possible.

- If $s = 313$, then pp' must have prefix 1233, 132133, or 13131; the only possibility is that $p = 131, p' = 313 = s$.

- If $s = 2$, then pp' must have prefix 1232, 132132, or 1312; the only possibility is that $p = 13213$, with $p' = s = 2$.

In every case, we are forced to conclude that $s = p'$.

Case: We arrive at E.

Suppose that, walking the labels of s^R, starting at A, takes us to E. The possible values of s are then 3 and 132. Similarly, the candidate values for p are 12312, 132, and 13131.

- In the case $s = 3$, the first letter of p' is 3; however, neither of 1323 and 1213 is a prefix of the image of a letter, so that p is one of 12312 and 13131; again $p' = 3 = s$.
- If $s = 132$, the first letter of p' is 1; since none of 123121, 131311 and is a prefix of the image of a letter, p is 132, forcing one of 12312 and 13131, and again $p' = 132 = s$.
- If $s = 212$, then pp' must have prefix 123122, 1322, orr 131312; None of these are possible.

In every case, we are forced to conclude that $s = p'$.

Case: We arrive at D.

The Similar case analysis shows that p is 1231 and $s = p'$, namely, 23.

Case: We arrive at C.

Case analysis gives that p is 1321 or 1313, and $s = p'$.

Case: We arrive at B.

In this case, the only possibility for p is $p = 1$, and the candidates for s are the words 31313, 23123, and 32132.

Case: We arrive at A.

Considering the four possible vertex sequences, and recalling that $s' \neq \epsilon$, we conclude that $s = \epsilon$. If $|U| \geq 9$, then U contains a factor concatenating a length 2 suffix and a length 3 prefix as in Remark 3.1.3, forcing $|Uxy| \equiv 0 \pmod 6$. However, $|Uxy| = |sh(u)p|$, and $|s| = 0$, forcing $|p| \equiv 0 \pmod 6$, so that $|p| = 0$, since $p' \neq \epsilon$. Thus $p = \epsilon$, as desired.

Suppose $|U| < 9$. Closed walks on $\mathcal{G}$ have even length, so $|Uxy|$ is even, forcing $|U|$ to be even. Thus $|U| \leq 8$, and $|Uxy| \leq 10$. Given $s = \epsilon$, we find that Uxy is a prefix of $h(u\rho)$. However, Uxy is a closed walk, so that the vertex sequences at the beginning of the claim show that $|Uxy|$ must be 6, whence $|U| = 4$. Thus $Uxy = h(\alpha)$, some $\alpha \in A$, and the second U in $UxyU$ is a length 4 prefix of $h(\beta)$, some $\beta \in A$, where $\alpha\beta$ is a prefix of $u\rho$. By Lemma 3.1.3 $\alpha = \beta$, since the length 4 images of their prefixes agree. This contradicts the square-freeness of w.

We have established in all cases that that either $s = p = \epsilon$ or $s = p'$. In both cases, $|s| \equiv |p'| \pmod 6$, so that $|sp| \equiv |p'| + |p| \equiv 0 \pmod 6$. Thus, $|Uxy| = |sh(u)p| \equiv 0 \pmod 6$.

(End proof of Claim)

Since , $|Uxy| \equiv 0 \pmod 6$, it follows that $|UxyU| \equiv 4 \pmod 6$. Certainly $|[w]| \equiv 0 \pmod 6$.

Recall that $UxyU = sh(u)pp'h(\omega)t$, where for some letters $\sigma, \rho, \tau \in A$ we have $h(\sigma) = s's$, $h(\rho) = pp'$, $h(\tau) = tt'$, $s', p', t' \neq \epsilon$; $Uxy = sh(u)p$, $U = p'h(\omega)t$; and both $\sigma u\rho\omega$ and $u\rho\omega\tau$ are factors of $[v]$.

Case: We have $s = p = \epsilon$.

Suppose $s = p = \epsilon$. If $|UxyU| \equiv 4 \pmod 6$ then $|h(v)pp'h(\omega)t| = 0$, and $|t| = 4$. Let $\alpha \in A$ be the last letter of u. Then the length 4 prefix of $h(\alpha)$ is the length 4 suffix of

$U = h(u)(xy)^{-1} = p'h(\omega)t$, and is therefore t. By Lemma 3.1.3, this forces $t' = xy$ and $\alpha = \tau$. Then $Uxy = h(u) = h(\rho\omega\tau)$. We now reach a contradiction, because $u\rho\omega\tau$ is a factor of $[w]$, but $u\rho\omega\tau$ is a square.

Case: We have $s = p' \neq \epsilon$.

Suppose we have $s = p' \neq \epsilon$.

Suppose $|h(w)| = |UxyU| + 2$.

In this case, the length requirement forces $\sigma = \tau$, and $t' = zs$ for some $z \in S^2$. Both s and p' are prefixes of U.

Recall that $|UxyU| \equiv 4 \pmod 6$ and $|h(w)| \equiv 0 \pmod 6$.

Claim B.1.2. *We have $z = xy$.*

Proof of claim: If $|s| = 1$, then $|p| = 5 = |tz|$. Since p is a suffix of Uxy, xy is a suffix of p, and $|t| = |p(xy)^{-1}| = 3$. This implies that $z = xy$. Similarly, if $|s| = |p'| \geq 2$, by Lemma 3.1.3, $tz = p$ and $\rho = \tau$. Then z is a suffix of p, which is a suffix of Uxy, and again $z = xy$.

(End proof of Claim)

From the claim, we have $UxyUxy = sh(u)pp'h(\omega)tz$, so that $p'h(\omega)tz = Uxy$. We thus reach a contradiction, because $u\rho\omega\tau$ is a factor of $[w]$, but $h(u\rho\omega\tau) = (s^{-1}Uxys)(s^{-1}Uxys)$ is a square.

Suppose $|h(w)| > |UxyU| + 2$.

Here $|h(w)| \geq |UxyU| + 8$, and we can write $UxyU = sh(u)pp'h(\omega)t$, where for some letters $\sigma, \rho, \tau \in A$ we have $h(\sigma) = s's$, $h(\rho) = pp'$, $h(\tau) = tt'$, $s', p', t' \neq \epsilon$; $Uxy = sh(u)p$, $U = p'h(\omega)t$; and $\sigma u\rho\omega\tau$ is a factor of $[w]$.

Both s and p' are prefixes of U. We consider $|s|$:

If $2 \leq |s| \leq 4$, then, because $|s| \geq 2$, Lemma 3.1.3, forces $s' = p$ and $\sigma = \rho$. Because $|p'| = |s| \leq 4$, we conclude $|s'| = |p| \geq 2$. Write $s' = p = p"xy$. It follows that $p"$ is a suffix of U, and $t = p"$. We conclude that $h(\sigma u \rho \omega) = (p"xyU(p")^{-1})(p"xyU(p")^{-1})$, which is a square. This is a contradiction.

If $|s| = |p'| = 1$, then, because $|p| = 6 - |p'| = 5$, xy is again suffix of p. Write $p = p"xy$. Then $p"$ is a suffix of U, and (because $|xyU| \equiv 0 \pmod 6$), $t = p"$. From $|p"| = |t| = 3$, Lemma 3.1.3, forces $\rho = \tau$, and $t' = xyp'$ so that $h(u\rho\omega\tau) = ((p')^{-1}Uxyp')((p'^{-1})Uxyp')$, which is a square, giving a contradiction.

In the remaining possibility, $|s| = |p'| = 5$. We thus have $|s'| = |p| = 1$. Let α be the last letter of u. From $sh(\alpha)p = Uxy$, we can write $h(\alpha) = \alpha' x$, some $\alpha' \in S^*$, and $y = p = s'$. Because $|U| \equiv 4 \pmod 6$, we find that $t = \alpha'$, and Lemma 3.1.3, forces $\alpha = \tau$. Then $h(\sigma u \rho \omega \tau) = yUxyUx$, which is a square, giving a contradiction.

$\square$

Appendix C

A list of FS Words

0

00

000

0001

00011

000111

0001011

00010111

000101100111

00010111001011

0001011100011001011

00010110001110010111

0001011000111001011100111

000110010110001110010110011 1

00010111001011001110001100101 1

0001011000111001011001110001100101 11

000101100011100101100010111001011001 11

0001011000111001011000101110001100101 11

0001011001110001011100101100111000110010111

0001011000111001011100010110011100011001011 1

000101100111000101110010110001011100011001011 1

000101100011100101100111000110010110001110010111

00010110001110010110001011100101100111000110010111

000101100011100101100010111000110010110001110010111

0001011001110001100101100011100101100111000110010111

000101100011100101100111000110010110001011100101100111

0001011000111001011000101110001100101100011100101100111

00010110001110010110001011100011001011000101110010110 0111

00010110001110010110011100010111001011000101110001100 10111

000101100011100101110001011001110001100101100011100101100111

0001011000111001011000101110010110011100011001011000111001 0111

00010110001110010111000101100111000101110010110011100011001 0111

000101100011100101100010111000110010110001011100101100011100 10111

0001011000111001011000101110010110011100011001011000111001011 00111

000101100011100101100010111000110010110001110010110011100011001 0111

0001011000111001011001110001011000111001011100010110011100011001 0111

00010110001110010110001011100011001011000101110010110011100011001 0111

000101100011100101100010111000110010111000101100111000101110010110 0111

0001011001110001011100101100111000110010110001110010110011100011001 0111

000101100011100101100010111001011001110001100101100010111001011000111001 0111

0001011000111001011000101110001100101110001011000111001011000101110010110 0111

00010110001110010110001011100011001011100010110001110010111000101100111000110010111

000101100011100101100010111000110010110001011100101100111000101110010110001110010111

00010110001110010110001011100011001011000111001011100010110001110010110011100011 0010

111

00010110001110010110001011100011001011000101110010110011100011001011000101110001100 1

0111

000101100011100101100010111000110010110001011100101100011100101110001011001110001100

10111

0001011000111001011000101110001100101100011100101110001011001110001011100101100011 10

010111

000101100011100101100010111000110010110001011100101100111000101110010110001011100011

0010111

000101100011100101100010111001011001110001100101100011100101110001100101100010111001

01100111

000101100011100101100010111000110010110001011100101100011100101110001100101100011100

101100111

000101100011100101100010111000110010110001110010111000101100111000101110010110011100

0110010111

000101100011100101100010111000110010110001011100101100011100101110001100101100010111

00101100111

000101100011100101100010111000110010110001011100101100011100101110001011001110001011

100101100111

000101100011100101100010111000110010110001011100101100111000110010110001110010110011

1000110010111

000101100011100101100010111000110010110001110010110011100011001011100010110011100010

11100101100111

000101100011100101100010111000110010110001110010110011100010110001110010111000101100

111000110010111

000110010110001110010111000101100011100101100010111000110010110001110010110011100010

110001110010111000101100111000110010110001011100101100011100101110001011000111001011

00111

Appendix D

Code

All code used is available upon request.