Periodicity and Repetition in Combinatorics on Words

by

Ming-wei Wang

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Computer Science

Waterloo, Ontario, Canada, 2004

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis concerns combinatorics on words. I present many results in this area, united by the common themes of periodicity and repetition. Most of these results have already appeared in journal or conference articles. Chapter 2 – Chapter 5 contain the most significant contribution of this thesis in the area of combinatorics on words. Below we give a brief synopsis of each chapter.

Chapter 1 introduces the subject area in general and some background information.

Chapter 2 and Chapter 3 grew out of attempts to prove the Decreasing Length Conjecture (DLC). The DLC states that if φ is a morphism over an alphabet of size n then for any word w, there exists $0 \leq i < j \leq n$ such that $|\varphi^i(w)| \leq |\varphi^j(w)|$. The DLC was proved by S. Cautis and S. Yazdani in *Periodicity, morphisms, and matrices* in *Theoret. Comput. Sci.* (**295**) 2003, 107-121.

More specifically, Chapter 2 gives two generalizations of the classical Fine and Wilf theorem which states that if $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ are two periodic sequences of real numbers, of period lengths h and k respectively, then

(a) If $f_n = g_n$ for $0 \leq n < h + k - \gcd(h, k)$, then $f_n = g_n$ for all $n \geq 0$.

(b) The conclusion in (a) would be false if $h + k - \gcd(h, k)$ were replaced by any smaller number.

We give similar results where equality in (a) is replaced by inequality and to where there are more than two sequences. These generalizations can be used to prove weak versions of the DLC.

Chapter 3 gives an essentially optimal bound to the following matrix problem. Let A be an $n \times n$ matrix with non-negative integer entries. Let $f(n)$ be the smallest integer such that for all A, there exist $i < j \leq f(n)$ such that $A^i \leq A^j$, where $A \leq B$ means each entry of A is less than or equal to the corresponding entry in B. The question is to find good upper bounds on $f(n)$. This problem has been attacked in two different ways. We give a method that proves an essentially optimal upper bound of $n + g(n)$ where $g(n)$ is the maximum order of an element of the symmetric group on n objects. A second approach yields a slightly worse upper bound. But this approach has a result of independent interest concerning *irreducible matrices*. A non-negative $n \times n$ matrix A is *irreducible* if $\sum_{i=0}^{n-1} A^i$ has all entries strictly positive. We show in Chapter 3 that if A is an irreducible $n \times n$ matrix, then there exists an integer $e > 0$ with $e = O(n \log n)$ such that the diagonal entries of A^e are all strictly positive. These results improve on results in my Master's thesis and are a version of the DLC in the matrix setting. They have direct applications to the growth rate of words in a D0L system.

Chapter 4 gives a complete characterization of two-sided fixed points of morphisms. A weak version of the DLC is used to prove a non-trivial case of the characterization. This characterization completes the previous work of Head and Lando on finite and one-sided fixed points of morphisms.

Chapter 5, 6 and 7 deal with avoiding different kinds of repetitions in infinite words.

Chapter 5 deals with problems about simultaneously avoiding cubes and large squares in infinite binary words. We use morphisms and fixed points to construct an infinite binary word that simultaneously avoid cubes and squares xx with $|x| \geq 4$. M. Dekking was the first to show such words exist. His construction used a non-uniform morphism. We use only uniform morphisms in Chapter 5. The construction in Chapter 5 is somewhat simpler than Dekking's.

Chapter 6 deals with problems of simultaneously avoiding several patterns at once. The patterns are generated by a simple arithmetic operation.

Chapter 7 proves a variant of a result of H. Friedman. We say a word y is a *subsequence* of a word z if y can be obtained by striking out zero or more symbols from z. Friedman proved that over any finite alphabet, there exists a longest finite word $x = x_1x_2\cdots x_n$ such that $x_ix_{i+1}\cdots x_{2i}$ is not a subsequence of $x_jx_{j+1}\cdots x_{2j}$ for $1 \leq i < j \leq n/2$. We call such words *self-avoiding*. We show that if "subsequence" is replaced by "subword" in defining self-avoiding, then there are infinite self-avoiding words over a 3-letter alphabet but not over binary or unary alphabets. This solves a question posed by Jean-Paul Allouche.

In Chapter 8 we give an application of the existence of infinitely many square-free words over a 3-letter alphabet. The *duplication language* generated by a word w is roughly speaking the set of words that can be obtained from w by repeatedly doubling the subwords of w. We use the existence of infinitely many square-free words over a 3-letter alphabet to prove that the duplication language generated by a word containing at least 3 distinct letters is not regular. This solves an open problem due to J. Dassow, V. Mitrana and Gh. Păun. It is known that the duplication language generated by a word over a binary alphabet is regular. It is not known whether such languages are context-free if the generator word contains at least 3 distinct letters. After the defence of my thesis I noticed that essentially the same argument was given by Ehrenfeucht and Rozenberg in *Regularity of languages generated by copying systems* in *Discrete Appl. Math.* (**8**) 1984, 313-317.

Chapter 9 defines a new "descriptive" measure of complexity of a word w by the minimal size of a deterministic finite automaton that accepts w (and possibly other words) but no other words of length $|w|$. Lower and upper bounds for various classes of words are proved in Chapter 9. Many of the proofs make essential use of repetitions in words.

Acknowledgements

I owe many thanks to my supervisor Jeff Shallit for his patience, encouragement and expertise. This thesis would not have been written without him. The Computer Science department at Waterloo also provided me with a friendly environment to work in. I am grateful for it. I think the most important thing I acquired during my time as a graduate student is a taste for research which will remain with me long after graduation. Without Jeff I probably would not have had this opportunity. I wish to thank him for it.

Contents

1 Introduction **1**

2 Variations on a Theorem of Fine and Wilf **7**

2.1 Introduction 7

2.2 First variation 8

2.3 Second variation: more than two periods 11

3 An Inequality for Non-Negative Matrices **16**

3.1 Introduction 16

3.2 The First Theorem 18

3.3 An Alternate approach 19

3.4 Lower bound 22

4 On Two-Sided Infinite Fixed Points of Morphisms **25**

4.1 Introduction and definitions 25

4.2 Finite and one-sided infinite fixed points 29

4.3 Two-sided infinite fixed points: the "pointed" case 30

4.4 Two-sided infinite fixed points: the "unpointed" case 32

4.5 Some examples 43

4.6 The equation $h(xy) = yx$ in finite words 44

5 Avoiding Large Squares in Infinite Binary Words **50**

5.1 Introduction . . . 50

5.2 A cubefree word without arbitrarily long squares . . . 51

6 New Problems of Pattern Avoidance **56**

6.1 Introduction . . . 56

6.2 Avoiding $x\sigma(x)$ . . . 57

6.3 Avoiding xx, $x\sigma(x)$, …, $x\sigma^j(x)$ simultaneously . . . 58

6.4 Even more results . . . 74

6.5 Avoiding $x\sigma^i(x)$ for all i . . . 75

7 Weakly Self-Avoiding Words and a Construction of Friedman **79**

7.1 Introduction . . . 79

7.2 Main Results . . . 80

7.3 Another construction . . . 83

8 On the irregularity of the duplication closure **84**

8.1 Open Problem . . . 86

9 Automatic Complexity of Strings **87**

9.1 Introduction . . . 87

9.2 Basic results . . . 89

9.3 Upper bounds . . . 92

9.4 Lower bounds . . . 94

9.5 Some specific examples . . . 100

9.6 Infinite words . . . 105

9.7 Open Problems . . . 111

10 Conclusion **112**

Bibliography **115**

Chapter 1

Introduction

The basic object of this thesis is a *word*, that is a sequence – finite or infinite – of symbols from a finite set. We present several results on periodicity and repetition in words and their applications. These results fall loosely under the subject of combinatorics on words [Lot83, Lot02, CK97, BK03].

A fundamental notion of words is periodicity. A word $w = w_1 w_2 \cdots$ is said to be *periodic with period* p if $w_i = w_{i+p}$ for $i = 1, 2, \ldots$. A fundamental and one of the oldest results about periodicity is the classical theorem of Fine and Wilf [FW65]. The Fine and Wilf theorem states that if $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ are two periodic sequences of real numbers, of period lengths h and k respectively, then

(a) If $f_n = g_n$ for $0 \leq n < h + k - \gcd(h,k)$, then $f_n = g_n$ for all $n \geq 0$.

(b) The conclusion in (a) would be false if $h + k - \gcd(h,k)$ were replaced by any smaller number.

In Chapter 2 we give several non-trivial generalizations of this theorem. We give results where equality in (a) is replaced by inequality and to where there are more than two sequences.

These generalizations can be used to prove weak versions of the Decreasing Length Conjecture (DLC) which was proved in [CMS$^+$03]. We can state DLC in terms of morphisms. A *morphism* is a map h from $\Sigma^* \to \Sigma^*$ such that $h(xy) = h(x)h(y)$ for all x, $y \in \Sigma^*$. The DLC states that if φ is a morphism over an alphabet of size n then for any word w, there exists $0 \leq i < j \leq n$ such that $|\varphi^i(w)| \leq |\varphi^j(w)|$.

There is an interesting connection between the Fine and Wilf theorem and infinite *sturmian words* [dLM94, Ber02]. An infinite *sturmian word* is an infinite binary word with exactly $n+1$ distinct subwords of length n for all $n \geq 1$. We consider the set $PER_{p,q}$ of finite words w of length $p+q-2$ with coprime periods p and q and is not a power of a single letter. Such w can be viewed as maximal common prefixes of sequences f and g for which the conclusion in part (a) of the Fine and Wilf theorem is false. It turns out that all such maximal common prefixes are binary words. In fact they are precisely the lenth $p+q-2$ prefixes of a subclass (*characteristic sturmian words*) of sturmian words. Also the set of subwords of sturmian words coincides with the set of subwords of $PER_{p,q}$ unioned over all coprime p and q.

Several other generalizations should be mentioned. Castelli, Mignosi and Restivo have generalized Fine and Wilf to three periods [CMR99]. They define a function $f(p_1, p_2, p_3) \sim \frac{1}{2}(p_1 + p_2 + p_3)$ such that if three periodic functions – with periods p_1, p_2 and p_3 respectively – have a common prefix of length at least $f(p_1, p_2, p_3)$, then they equal each other. Justin [Jus00] has extended this result to an arbitrary number of periods. Tijdeman and Zamboni [TZ03] give a fast algorithm to compute for a given set of periods a maximal common prefix w of periodic sequences for which the conclusion of the Fine and Wilf theorem is false. They showed that such a w is uniquely determined up to isomorphism and that it is a *palindrome*. A *palindrome* is a word w such that the reversal of w equals w. Simpson and Tijdeman [ST03] also give an multi-dimensional version of the Fine and Wilf theorem. Berstel and Boasson discussed Fine and Wilf theorem in the case of *partial words* [BB99, BH02]. A *partial words* is a word where there are some positions in the word that are undefined. Finally there is a connection to problems about covering the integers by residue classes, see [Sun04a, Sun04b].

In Chapter 3 we consider DLC in the matrix setting. The problem is as follows. Let A be an $n \times n$ matrix with non-negative integer entries. Let $f(n)$ be the smallest integer such that for all A, there exist $i < j \leq f(n)$ such that $A^i \leq A^j$, where $A \leq B$ means each entry of A is less than or equal to the corresponding entry in B. The question is to find good upper bounds on $f(n)$. This problem has been attacked in two different ways. We give a method that proves an essentially optimal upper bound of $n + g(n)$, where $g(n)$ is the maximum order of an element of the symmetric group on n objects. A second approach yields a slightly worse upper bound. But this approach has a result of independent interest concerning *irreducible matrices*. A non-negative $n \times n$ matrix A is *irreducible* if $\sum_{i=0}^{n-1} A^i$ has all entries strictly positive. We show in Chapter 3 that if A is an irreducible $n \times n$ matrix, then there exists an integer $e > 0$ with $e = O(n \log n)$ such that the diagonal entries of A^e are all strictly positive. These results improve on results in my Master's thesis [Wan99]. They have direct applications to the growth rate of words in a D0L system.

Morphisms and their fixed points are basic tools used in solving many problems on words. For example, nearly every explicit construction of an infinite word avoiding certain patterns involve the fixed point of a morphism [HM56, Lee57, Zec58, Ple70]. They are also worthy of study in their own right [CK97, HK97]. For example, define a morphism μ by $\mu(\mathtt{0}) = \mathtt{01}$ and $\mu(\mathtt{1}) = \mathtt{10}$. The Thue-Morse word

$$\mathbf{t} = \mathtt{0110100110010110} \cdots$$

is the unique one-sided infinite fixed point of μ which starts with $\mathtt{0}$. The word $\mathbf{t}$ is overlap-free, that is, contains no subword of the form $axaxa$, where $a \in \{\mathtt{0}, \mathtt{1}\}$, and $x \in (\mathtt{0} + \mathtt{1})^*$. The Thue-Morse word appears in many different contexts [AS99]. Morse [Mor21] introduced $\mathbf{t}$ in connection with a problem in differential geometry. It can be used to solve a problem in chess [Euw29, MH44]. It provides a solution to the Prouhet-Tarry-Escott problem in number theory [AL77, Pro51, Leh47]. It is used in solving the Burnside problem for groups:

Is every group with a finite number of generators and satisfying the identity $x^n = 1$ finite? [Adi79, NA68]. A simple transformation of $\mathbf{t}$ can be used to solve the Burnside problem for semigroups [MH44, RR85]. $\mathbf{t}$ also has many interesting extremal characterizations. For example, if an overlap-free binary sequence is a fixed point of a non-trivial morphism, then it is either equal to $\mathbf{t}$ or its complement $1001011001101001\cdots$ (1's and 0's are switched) [Séé85]. It is also the lexicographically largest overlap-free infinite binary word beginning with 0 [Ber95]. If we consider $\mathbf{t}$ as the binary expansion of a real number, then we have the following interesting extremal characterization. Let $\tau = \sum_{n\geq 1} t_n 2^{-n}$ where t_n is the $(n+1)$-th digit of $\mathbf{t}$. Define a set of real numbers Γ by

$$\Gamma = \{x \in [0,1] \mid \forall k \geq 0, 1-x \leq \{2^k x\} \leq x\},$$

where $\{x\}$ denotes the fractional part of x. Then τ is the least irrational element of Γ [AM01]. Hopefully these results convince the reader that fixed points of morphisms are interesting objects of study.

In Chapter 4 we give a complete classification of two-sided fixed points of morphisms. This classification effort led to the formulation of the Decreasing Length Conjecture. A weak version of DLC appears as Lemma 4.4.4 which is used to prove a non-trivial case of the characterization. This characterization completes the previous work of Head and Lando [Hea81, HL86] on finite and one-sided fixed points of morphisms. Foryś in [For04] studied a related problem.

The earliest systematic study of problems on words was initiated by Axel Thue at the beginning of last century. He published two long papers [Thu06, Thu12] in which he investigated among other things the structure of the set of *square-free* words over a three letter alphabet. A word is *square-free* if it does not contain two adjacent identical blocks of symbols. Thue also noted an interesting parallel between his method/results with those in diophantine equations in number theory. Chapter 5, 6 and 7 contain results on pattern avoidance which are descendants of this line of research. Chapter 8 contains a small application of the result that squares are avoidable over a 3-letter alphabet.

A *square* is a nonempty word of the form xx, as in the English word `murmur`. A *cube* is a nonempty word of the form xxx, as in the English sort-of-word `shshsh`. Chapter 5 deals with problems about simultaneously avoiding cubes and large squares in infinite binary words. We use morphisms and fixed points to construct an infinite binary word that simultaneously avoid cubes and squares xx with $|x| \geq 4$. M. Dekking was the first to show such words exist [Dek76]. His construction used a non-uniform morphism. We use only uniform morphisms in Chapter 5. The construction in Chapter 5 is somewhat simpler than Dekking's.

Chapter 6 deals with problems of simultaneously avoiding several patterns at once. Here we consider a generalization of Thue's problem. We define $\Sigma_k = \{0, 1, 2, \ldots, k-1\}$ for some integer $k \geq 2$, and we define the morphism $\sigma_k(a) = (a+1) \bmod k$. We omit the subscript k if it is clear from the context. In this chapter, we consider avoiding patterns of the form $x\sigma^i(x)$. The work here is a special case of more general pattern avoidance problems [BEM79, Cur93].

Chapter 7 proves a variant of a result of Friedman [Fri01, Fri00]. We say a word y is a *subsequence* of a word z if y can be obtained by striking out zero or more symbols from z. Friedman proved that over any finite alphabet, there exists a longest finite word $x = x_1x_2 \cdots x_n$ such that $x_ix_{i+1} \cdots x_{2i}$ is not a subsequence of $x_jx_{j+1} \cdots x_{2j}$ for $1 \leq i < j \leq n/2$. We call such words *self-avoiding*. We show that if "subsequence" is replaced by "subword" in defining self-avoiding, then there are infinite self-avoiding words over a 3-letter alphabet but not over binary or unary alphabets. This solves a question posed by Jean-Paul Allouche.

In Chapter 8 we give an application of the existence of infinitely many square-free words over a 3-letter alphabet. The *duplication language* generated by a word w is roughly speaking the set of words that can be obtained from w by repeatedly doubling the subwords of w. We used the existence of infinitely many square-free words over a 3-letter alphabet to prove that the duplication language generated by a word containing at least 3 distinct letters is not regular. This solves an open problem due to J. Dassow, V. Mitrana and Gh. Păun. It is known that the duplication language generated by a word over a binary alphabet is regular [DMP99]. It is not known whether such languages are context-free if the generator word contains at least 3 distinct letters. After the defence of my thesis I noticed that essentially the same argument was given by Ehrenfeucht and Rozenberg [ER84].

Descriptive complexity is an intrinsic property of words. The classical measure of Kolmogorov-Chaitin has proved to be a rich area of study with many applications. However, the Kolmogorov-Chaitin measure is not a computable measure. In Chapter 9, we propose a computable measure similar in spirit to the Kolmogorov-Chaitin measure and study some of its basic properties. We define for a word w, its *automatic complexity* $A(w)$ to be the minimal size of a deterministic finite automaton that accepts w (and possibly other words) but no other words of length $|w|$. It turns out that many aspects of this measure are closely related to repetitions in the word being measured. For example, a word w is said to be kth-power-free if w does not contain k adjacent identical blocks of symbols. Theorem 9.4.3 in Chapter 9 shows that if a word w is kth-power-free, then $A(w) \geq \frac{|w|+1}{k}$. These and other potential connections with combinatorics on words are the reason for Chapter 9's inclusion in this thesis.

Finally, we note that the most significant contributions to the area of combinatorics on words are in Chapter 2 through Chapter 5.

Chapter 2

Variations on a Theorem of Fine and Wilf

This chapter is based on the work of S. Cautis, S. Yazdani, F. Mignosi, J. Shallit and M.-w. Wang [MSW01, CMS$^+$03].

2.1 Introduction

In this chapter we explore several generalizations of a classical theorem of Fine and Wilf.

Periodicity is an important property of words that has applications in various domains. For instance, it has applications in string searching algorithms (cf. [CR94]), in formal languages (cf. for instance the pumping lemmas in Salomaa [Sal73]), and it is an important part of combinatorics on words (cf. [CK97, Ber02]).

We say a sequence $(f_n)_{n\geq 0}$ is periodic with period length $h \geq 1$ if $f_n = f_{n+h}$ for all $n \geq 0$. The following is a classical "folk theorem":

Theorem 2.1.1 *If $(f_n)_{n\geq 0}$ is a sequence of real numbers which is periodic with period lengths h and k, then it is periodic with period length* $\gcd(h, k)$.

Proof. By the extended Euclidean algorithm, there exist integers $r, s \geq 0$ such that $rh - sk = \gcd(h, k)$. Then we have

$$f_n = f_{n+rh} = f_{n+rh-sk} = f_{n+\gcd(h,k)}$$

for all $n \geq 0$. ■

The 1965 theorem of Fine and Wilf [FW65] is the following:

Theorem 2.1.2 *Let $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ be two periodic sequences of real numbers, of period lengths h and k respectively.*

(a) If $f_n = g_n$ for $0 \leq n < h + k - \gcd(h, k)$, then $f_n = g_n$ for all $n \geq 0$.

(b) The conclusion in (a) would be false if $h + k - \gcd(h, k)$ were replaced by any smaller number.

We first consider some variations on the theorem of Fine and Wilf in which equality is replaced by inequality.

2.2 First variation

We begin with a bit of notation and a lemma. Let $\mathbf{a} = (a_i)_{i\geq 0}$ be a sequence of real numbers, and let $p = (p_0, p_1, \ldots, p_{h-1})$ be a vector of real numbers of dimension $h \geq 1$. We will frequently need the new sequence $p \circ \mathbf{a}$ resulting from taking successive "'windows" of length h of $\mathbf{a}$ and forming their dot product with p. More formally, we define $p \circ \mathbf{a} := (\sum_{0\leq i<h} p_i a_{n+i})_{n\geq 0}$.

Lemma 2.2.1 *Let $p = (p_0, p_1, \ldots, p_{h-1})$ be a vector of $h \geq 1$ real numbers and $q = (q_0, q_1, \ldots, q_{k-1})$ be a vector of $k \geq 1$ real numbers. Then $q \circ (p \circ \mathbf{a}) = (qp) \circ \mathbf{a}$, where* *by qp we mean the vector $(r_0, r_1, \ldots, r_{h+k-2})$ defined by*

$$r_n = \sum_{\substack{0 \le i < h \\ 0 \le j < k \\ i+j=n}} p_i q_j.$$

Proof. Define $P(z) = \sum_{0 \le i < h} p_i z^i$, $Q(z) = \sum_{0 \le i < k} q_i z^i$, and $A(z) = \sum_{i \ge 0} a_i z^{-i}$. If $p \circ \mathbf{a} = (t_i)_{i \ge 0}$ then it is easy to see that $P(z)A(z) = \left(\sum_{i \ge 0} t_i z^{-i}\right) + W(z)$, where W is a polynomial of degree $\deg P$ such that $W(0) = 0$. If $q \circ (p \circ \mathbf{a}) = (u_i)_{i \ge 0}$ it follows that $Q(z)P(z)A(z) = \left(\sum_{i \ge 0} u_i z^{-i}\right) + S(z)$ where $S(0) = 0$. Hence $q \circ (p \circ \mathbf{a}) = (qp) \circ \mathbf{a}$. ■

For the rest of this chapter, we abuse notation slightly by writing $P \circ \mathbf{a}$ for $p \circ \mathbf{a}$, where $p = (p_0, p_1, \ldots, p_{h-1})$ and $P(z) = \sum_{0 \le i < h} p_i z^i$.

We are now ready to state and prove our first variation on the theorem of Fine and Wilf.

Theorem 2.2.2 *Let $\mathbf{f} = (f_n)_{n \ge 0}$, $\mathbf{g} = (g_n)_{n \ge 0}$ be two periodic sequences of real numbers, of period lengths h and k, respectively, such that*

$$\sum_{0 \le i < h} f_i \ge 0 \tag{2.1}$$

and

$$\sum_{0 \le j < k} g_j \le 0. \tag{2.2}$$

Let $d = \gcd(h, k)$. If

$$f_n \le g_n \quad \text{for } 0 \le n < h + k - d \tag{2.3}$$

then

(i) $f_n = g_n$ for all $n \ge 0$; and

(ii) $\sum_{j \le i < j+d} f_i = \sum_{j \le i < j+d} g_i = 0$ for all integers $j \ge 0$.

Proof. Let $d = \gcd(h, k)$, and define

$$\begin{aligned} P(z) &= 1 + z + \cdots + z^{h-1} = (z^h - 1)/(z - 1); \\ Q(z) &= 1 + z + \cdots + z^{k-1} = (z^k - 1)/(z - 1). \end{aligned}$$

Define

$$\begin{aligned} R(z) &= (z^k - 1)/(z^d - 1); \\ S(z) &= (z^h - 1)/(z^d - 1). \end{aligned}$$

Then none of P, Q, R, S is identically zero, but all have non-negative coefficients. By hypothesis (2.1) we have $P \circ \mathbf{f} \geq 0$. Hence $R \circ (P \circ \mathbf{f}) \geq 0$. But by Lemma 2.2.1 this means

$$RP \circ \mathbf{f} \geq 0. \tag{2.4}$$

Similarly by hypothesis (2.2) we have $Q \circ (-\mathbf{g}) \geq 0$; hence

$$SQ \circ (-\mathbf{g}) = S \circ (Q \circ (-\mathbf{g})) \geq 0. \tag{2.5}$$

Note that $RP = SQ$, and RP is a polynomial of degree $h + k - d - 1$. Define the coefficients e_i by $R(z)P(z) = \sum_{0 \leq i < h+k-d} e_i z^i$. By (2.4) and (2.5) we have

$$\sum_{0 \leq i < h+k-d} e_i (f_i - g_i) \geq 0. \tag{2.6}$$

Now we claim that all the coefficients e_i are strictly positive. To see this, note that

$$\begin{aligned} R(z)P(z) &= \frac{z^h - 1}{z^d - 1} \cdot \frac{z^k - 1}{z - 1} \\ &= (1 + z^d + z^{2d} + \cdots + z^{h-d})(1 + z + z^2 + \cdots + z^{k-1}). \end{aligned}$$

If $i < h$, write $i = qd + r$ where $0 \leq r < d$, and choose the term z^{qd} from the left factor and z^r from the right factor to see $e_i > 0$. If $h \leq i < h + k - d$, choose z^{h-d} from the left factor and z^{i-h+d} from the right factor to see $e_i > 0$.

Since the e_i are all strictly positive, combining the inequality (2.6) with the hypothesis (2.3) that $f_n \leq g_n$ for $0 \leq n < h + k - d$ gives $f_n = g_n$ for $0 \leq n < h + k - d$. But then, by the Fine and Wilf theorem, $f_n = g_n$ for all $n \geq 0$. This proves (a)(i).

Next we prove (a)(ii). Since $f_n = g_n$ for all $n \geq 0$, it follows that f is periodic of period length h and k, and hence by Theorem 2.1.1, of period d. The sum over the terms of this period must be 0, since if it were less than 0 this would contradict hypothesis (2.1), while if it were greater than 0 this would contradict hypothesis (2.2).

Then $f_j + f_{j+1} + \cdots + f_{j+d-1}$ is just a cyclic permutation of $f_0 + f_1 + \cdots + f_{d-1}$, which equals 0. A similar argument applies to g. ■

It is known that the bound $h + k - d$ is tight [MSW01].

Remarks. Theorem 2.2.2 is reminiscent of some classical theorems on trigonometric polynomials. For example, Fejér [Fej13] proved that a real trigonometric polynomial with 0 constant term

$$\lambda_1 \cos\theta + \mu_1 \sin\theta + \lambda_2 \cos(2\theta) + \mu_2 \sin(2\theta) + \cdots + \lambda_r \cos(r\theta) + \mu_r \sin(r\theta)$$

cannot have the same sign for all real θ unless it is identically zero. Also see Pólya and Szegö [PS76, pp. 80, 263] and Gilbert and Smyth [GS00].

2.3 Second variation: more than two periods

In this section we consider some variations on the Fine and Wilf theorem for more than two periods. For other generalizations of Fine and Wilf to more than two periods, see [CMR99, Jus00, TZ03, ST03].

For our first theorem, we need a little notation. For integers $p \geq 1$ let ω_p denote a primitive p'th root of unity, i.e., $\omega_p := e^{2\pi\sqrt{-1}/p}$. Define

$$R_p := \{\omega_p^i \ : \ 0 \leq i < p\} = \{\omega \in \mathbb{C} \ : \ \omega^p = 1\}.$$

Finally, for integers $h_1, h_2, \ldots, h_r \geq 1$ define

$$\gamma(h_1, h_2, \ldots, h_r) = |R_{h_1} \ \cup \ R_{h_2} \ \cup \ \cdots \ \cup \ R_{h_r}|,$$

the number of distinct roots of unity among the h_1'th, h_2'th, etc., roots of unity.

By the principle of inclusion-exclusion, it follows that

$$\gamma(h_1, h_2, \ldots, h_r) = \sum_{\substack{S \subseteq \{h_1, h_2, \ldots, h_r\} \\ S \neq \emptyset}} \gcd(S)(-1)^{|S|+1},$$

where by $\gcd(S)$ for S a nonempty set we mean the greatest common divisor of all elements of S. For example,

$$\gamma(6, 10, 15) = 6 + 10 + 15 - \gcd(6, 10) - \gcd(6, 15) - \gcd(10, 15) + \gcd(6, 10, 15) = 22.$$

Theorem 2.3.1 *Let $(f_i(n))_{n \geq 0}$, $1 \leq i \leq r$, be r periodic complex-valued sequences with period lengths $h_1, h_2, \ldots, h_r$, respectively. Suppose $\sum_{1 \leq i \leq r} f_i(n) = 0$ for $0 \leq n < \gamma(h_1, h_2, \ldots, h_r)$. Then $\sum_{1 \leq i \leq r} f_i(n) = 0$ for all $n \geq 0$.*

Proof. As Fine and Wilf observed [FW65], any periodic complex-valued sequence $(f(n))_{n \geq 0}$ of period length p can be written in the form

$$f(n) = \sum_{0 \leq i < p} c_i \omega_p^{in}$$

for some coefficients $c_0, c_1, \ldots, c_{p-1}$.

It follows that there exist coefficients $c_{i,j}$, $1 \le i \le r$ and $0 \le j < h_i$ such that

$$f_i(n) = \sum_{0 \le j < h_i} c_{i,j} \omega_{h_i}^{jn}.$$

Define

$$\begin{aligned} s &= [s_1, s_2, \ldots, s_m] \\ &= [1, \omega_{h_1}, \omega_{h_1}^2, \ldots, \omega_{h_1}^{h_1-1}, 1, \omega_{h_2}, \omega_{h_2}^2, \ldots, \omega_{h_2}^{h_2-1}, \ldots, 1, \omega_{h_r}, \omega_{h_r}^2, \ldots, \omega_{h_r}^{h_r-1}] \end{aligned}$$

where $m = h_1 + h_2 + \cdots + h_r$. Let $B := \gamma(h_1, h_2, \ldots, h_r)$ and define the $B \times m$ matrix $M = (t_{i,j})_{0 \le i < B, 1 \le j \le m}$ by $t_{i,j} := s_j^i$. Define the column vector

$$v := [c_{1,0}, c_{1,1}, \ldots, c_{1,h_1-1}, c_{2,0}, c_{2,1}, \ldots, c_{2,h_2-1}, \ldots, c_{r,0}, c_{r,1}, \ldots, c_{r,h_r-1}]^T.$$

Then the hypothesis of the theorem is $Mv = 0$. Some of the columns of M are identical because some of the entries in the vector s coincide. We may delete the repeated columns of M and sum the corresponding entries of v to get $M'v' = 0$, where M' is a $B \times B$ matrix and v' is a column vector with B entries. Now M' is a Vandermonde matrix and hence invertible, so $v' = 0$. It follows that $\sum_{1 \le i \le r} f_i(n) = 0$ for all n. ■

We next turn to another variation on Fine and Wilf for more than two periods. This generalization is more in the spirit of Theorem 2.2.2.

Theorem 2.3.2 *Let* $\mathbf{f}_1 = (f_1(n))_{n \ge 0}$, $\mathbf{f}_2 = (f_2(n))_{n \ge 0}$, $\ldots$, $\mathbf{f}_r = (f_r(n))_{n \ge 0}$ *be* r *periodic real-valued sequences of periods* $h_1, h_2, \ldots, h_r$*, respectively. Suppose that for all* i *with* $1 \le i \le r$*, we have*

$$\sum_{0 \le n < h_i} f_i(n) \ge 0.$$

If

$$\sum_{1\le i\le r} f_i(n) \le 0$$

for $0 \le n < h_1 + h_2 + \cdots + h_r - r + 1$, *then*

$$\sum_{1\le i\le r} f_i(n) = 0$$

for all $n \ge 0$.

Proof. The proof is very similar to the proof of Theorem 2.2.2, and we indicate only what needs to be changed. First, we need the following easy generalization of Lemma 2.2.1.

Lemma 2.3.3 *If* $P_1, P_2, \ldots, P_r$ *are polynomials with real coefficients and* $\mathbf{a} = (a_n)_{n\ge 0}$ *then* $P_1 \circ (P_2 \circ \cdots \circ (P_r \circ \mathbf{a}) \cdots) = (P_1 \cdots P_r) \circ \mathbf{a}$.

For $1 \le i \le r$ define $P_i(z) = 1+z+\cdots+z^{h_i-1} = (z^{h_i}-1)/(z-1)$. Then by hypothesis $P_i \circ \mathbf{f}_i$ is a sequence of non-negative real numbers for each i, $1 \le i \le r$. It follows using Lemma 2.3.3 that if $P := P_1P_2\cdots P_r$, then $P \circ \mathbf{f}_i$ is a sequence of non-negative real numbers for $1 \le i \le r$. But P has degree $h_1 + h_2 + \cdots + h_r - r$ and hence has $h_1 + h_2 + \cdots + h_r - r + 1$ coefficients. Furthermore, all the coefficients of P are strictly positive. Hence if $\sum_{1\le i\le r} f_i(n) \le 0$ for $0 \le n < h_1+h_2+\cdots+h_r-r+1$, it follows that $\sum_{1\le i\le r} f_i(n) = 0$ for $0 \le n < h_1+h_2+\cdots+h_r-r+1$. Now $h_1 + h_2 + \cdots + h_r - r + 1 \ge \gamma(h_1, h_2, \ldots, h_r)$, since the left-hand side counts the total number of roots of unity among $R_{h_1}, \ldots, R_{h_r}$ without double-counting occurrences of 1, while the right-hand side counts the number of distinct roots of unity. But then $\sum_{1\le i\le r} f_i(n) = 0$ for all $n \ge 0$ by Theorem 2.3.1. ■

We note that the bound $h_1 + h_2 + \cdots + h_r - r + 1$ is not, in general, optimal, although the bound is optimal if the period lengths $h_1, h_2, \ldots, h_r$ are relatively prime.

One might be tempted to guess that the true bound, as in Theorem 2.3.1, is not $h_1 + h_2 + \cdots + h_r - r + 1$, but rather $\gamma(h_1, h_2, \ldots, h_r)$. This is not true, however. The following is an example of three periodic sequences of period lengths 6, 10, and 15, respectively, whose periods individually sum to 0 and such that $f_1(n) + f_2(n) + f_3(n) \leq 0$ for $0 \leq n < \gamma(6, 10, 15) = 22$, but not for $n = 22$.

$$\begin{aligned} f_1 &= (0,0,0,-1,1,0)^\omega; \\ f_2 &= (0,0,0,0,0,1,-1,0,-1,1)^\omega; \\ f_3 &= (0,0,0,1,-1,-1,1,0,1,0,-1,0,0,0,0)^\omega. \end{aligned}$$

The true bound is not known, although there is an algorithm to compute it. Theorem 2.3.2 can be used to prove a weak version of the Decreasing Length Conjecture. In the next chapter, we solve a problem similar in spirit to the Decreasing Length Conjecture.

Chapter 3

An Inequality for Non-Negative Matrices

This chapter is based on the work of J. Shallit and M.-w. Wang [SW99, Wan02].

3.1 Introduction

In [SW99, Wan99] we proved that for an $n \times n$ matrix A with non-negative integer entries there exist integers r, s with $0 \leq r < s \leq 2^n$ such that $A^r \leq A^s$. Bo improved the bound 2^n to $3^{n/2}$ [Bo00]. We give two results in this chapter. First, we improve the bound to $n + g(n)$ where g is the Landau function. Thus we are close to the known lower bound of $g(n)$ [SW99]. Second, we show that if A is an irreducible matrix then there is an integer i such that $A^i \geq I$ and $i = O(n \log n)$. We also give examples where $i = \Omega(n \log n / \log \log n)$. The second result can also be used to attack Theorem 3.2.1 below via a different method, though it gives a slightly worse bound. The results of this chapter has appeared in [Wan02].

Both Theorem 3.2.1 below and the results of Chapter 2 arose in connection with Lemma 4.4.4 of Chapter 4.

Remarks.

1. The significance of Theorem 3.2.1 below is that the upper bound on r and s depends only on n.

2. If $A^r \leq A^s$, then $A^s - A^r \geq 0$. Hence $A^t(A^s - A^r) \geq 0$ for all $t \geq 0$, and so it follows that $A^{r+t} \leq A^{s+t}$ for all $t \geq 0$.

While Theorem 3.2.1 appears to be new, there is some related work in the literature:

a. A non-negative matrix A is called *primitive* if there exists an integer $t \geq 1$ such that $A^t > 0$. The least such t is called the *exponent* of A and is denoted $\gamma(A)$. If A is an $n \times n$ primitive matrix, then Wielandt [Wie50] asserted $\gamma(A) \leq n^2 - 2n + 2$, and this bound is best possible. Wielandt's assertion was proved by Rosenblatt [Ros57], Holladay and Varga [HV58], Perkins [Per61], Dulmage and Mendelsohn [DM64], and Heap and Lynn [HL64]. In this case, evidently $I = A^0 \leq A^{\gamma(A)}$.

b. Rosenblatt [Ros57] studied the pattern of zero and nonzero entries in the powers of a non-negative matrix, and proved that there exist integers i, j such that the pattern of zeros and nonzeros in A^{t+i} is the same as that in A^t for all $t \geq j$. Also see [Ptá58, PS58, HL66b].

c. Marcus and May [MM62] studied the maximum number of zero entries in the powers of an irreducible matrix (see definition in Section 3.3). Pullman [Pul64] studied the maximum number of positive entries in the powers of a non-negative matrix, and the least power for which this maximum is assumed. Also see Heap and Lynn [HL66a].

d. Let $A = (a_{i,j})_{1 \leq i,j \leq n}$ be an $n \times n$ matrix with complex entries, and define $|A| = \max_{1 \leq i \leq n} \sum_{1 \leq j \leq n} |a_{ij}|$. Mařík and Pták [MP60] studied the least integer t such that $|A| = |A^2| = \cdots = |A^t| = 1$ implies $|A^r| = 1$ for all r.

3.2 The First Theorem

First we prove the following theorem. The technique used in the proof below is adapted from [CMS$^+$03] and is due to Sabin Cautis and Soroosh Yazdani.

Given matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ we say $A \geq B$ iff $a_{ij} \geq b_{ij}$ for all i, j. We recall that $g(n)$ is the maximum order of an element of the symmetric group on n objects. It is known that $g(n) = e^{O(\sqrt{n \log n})}$ [Lan03, Mil87].

Theorem 3.2.1 *Let $A \geq O$ be an $n \times n$ integer matrix. Then there exist integers $r < s$ with $1 \leq r \leq n$ and $s \leq n + g(n)$ such that $A^r \leq A^s$.*

Proof. Let $G = G(A)$ be the directed graph associated with A, i.e., we place a_{ij} edges from vertex v_i to vertex v_j (this may create multiple edges and self-loops). Then the entry (i, j) of A^t gives the number of distinct walks of length t from vertex i to vertex j in G. The length of a walk is the number of edges traversed.

Now consider some maximal set of vertices forming disjoint cycles $\{C_1, \ldots, C_k\}$ in G. Then the vertex set V of G can be written as the disjoint union

$$V = C_1 \cup \cdots \cup C_k \cup W$$

where W is the set of vertices which do not lie on any disjoint cycles. Note that W may be empty. Then any directed walk in G of length $|W|$ or greater must intersect some cycle C_i, for otherwise the walk would contain a cycle disjoint from $C_1, \ldots, C_k$, a contradiction. Now associate each walk of length $|W|$ or greater with the first cycle C_i it intersects. Define $P^t_{i,j,l}$ to be the number of directed walks of length t from vertex i to vertex j associated with cycle C_l. Let $A^t = [a^{(t)}_{ij}]$. Then for $t \geq |W|$, we have

$$a^{(t)}_{ij} = \sum_{l=1}^{k} P^t_{i,j,l}. \tag{3.1}$$

Observe that

$$P_{i,j,l}^{t} \leq P_{i,j,l}^{t+|C_l|} \tag{3.2}$$

because any walk of length t associated with C_l can be extended to a walk of length $t + |C_l|$ by traversing the cycle C_l once. This construction is 1-1 and it maps a walk associated with C_l to a walk associated with C_l, since C_l is the first cycle encountered by both walks.

Combining (3.1) and (3.2) we get for $t \geq |W|$

$$\begin{aligned} a_{ij}^{(t)} &= \sum_{l=1}^{k} P_{i,j,l}^{t} \\ &\leq \sum_{l=1}^{k} P_{i,j,l}^{t+\mathrm{lcm}(|C_1|,\ldots,|C_k|)} \\ &= a_{ij}^{(t+\mathrm{lcm}(|C_1|,\ldots,|C_k|))} \end{aligned}$$

Hence for $t \geq |W|$, $A^t \leq A^{t+\mathrm{lcm}(|C_1|,\ldots,|C_k|)}$. The theorem is proved. ∎

3.3 An Alternate approach

There is a different route to approach Theorem 3.2.1 using Theorem 3.3.1 below. A slightly worse bound can be obtained this way, though Theorem 3.3.1 is of independent interest. We give it below.

Let $A \geq O$ be an $n \times n$ matrix. A is *irreducible* if

$$\sum_{i=0}^{n-1} A^i > O.$$

Theorem 3.3.1 *If $A \geq O$ is an irreducible $n \times n$ matrix, then there exists an integer $e > 0$ with*

$$e = O(n \log n)$$

such that the diagonal entries of A^e are all strictly positive.

Theorem 3.3.1 is equivalent to the following theorem about strongly connected graphs which we prove.

Theorem 3.3.2 *Suppose G is a strongly connected graph on n vertices. Then there exists an integer $e > 0$ with*

$$e = O(n \log n)$$

such that for every vertex v there is a closed walk of length e containing v.

In the proof of Theorem 3.3.2 we will need the following fact.

A set of positive integers $S = \{a_1, \ldots, a_m\}$ is said to have the *distinct subset sum* property if no two distinct subsets of S sum to the same number. It is known that if $a_i \le n$ for $1 \le i \le m$, then $m = O(\log n)$ [Elk86].

Proof. We begin by picking an arbitrary vertex v and a set of closed walks $\mathcal{C} = \{C_1, \ldots, C_k\}$ with the following properties.

1. Each C_i contains v.

2. $\mathcal{C}$ *covers* G, i.e., every vertex of G is contained in at least one of the C_i's.

3. $|C_i| < 2n$ for $1 \le i \le k$ where $|C_i|$ denotes the length of C_i.

Since G is strongly connected, we can find such a set of closed walks $\mathcal{C}$. In fact we can find such a set with $k \le n$ but we will not need this fact below.

If a set of closed walks $\mathcal{C}$ contain at least one point in common then if we traverse any subset of $\mathcal{C}$ in any order and any number of times we get another closed walk. This is the reason for property 1. We use this fact implicitly below.

We need a bit of notation. If $\mathcal{C}$ is a set of closed walks, we denote by $\parallel \mathcal{C} \parallel$ the sum of the lengths of closed walks in $\mathcal{C}$.

Let $c_i = |C_i|$. Since each C_i contains v and $\mathcal{C}$ covers G, we see that every vertex of G is contained in a closed walk of length $c = \parallel \mathcal{C} \parallel = c_1 + \cdots + c_k$.

In general c may be as large as $\Omega(n^2)$. To get an $O(n \log n)$ upper bound we pick a subset $\mathcal{D} = \{D_1, \ldots, D_m\}$ of $\mathcal{C}$ with the following properties.

1. Let $d_i = |D_i|$. We require the set $S = \{d_1, \ldots, d_m\}$ to have the distinct subset sum property.

2. $\mathcal{D}$ is *maximal* in the sense that if we add any closed walk in $\mathcal{C} - \mathcal{D}$ to $\mathcal{D}$, then the set of lengths of closed walks in $\mathcal{D}$ no longer has the distinct subset sum property.

Now let $d = \parallel \mathcal{D} \parallel = d_1 + \cdots + d_m$. We claim that for every vertex w there is a closed walk of length d containing w. If the claim is true then we are done, since $m = O(\log n)$ and $d_i < 2n$. To see the claim, there are two cases.

Case 1: w lies on one of the closed walks in $\mathcal{D}$. Then w is contained in the closed walk T that traverses each D_i once and $|T| = d$.

Case 2: w does not lie on any closed walk in $\mathcal{D}$. Since $\mathcal{C}$ covers G, w lies on some closed walk in $\mathcal{C}$, say C_i. Since $\mathcal{D}$ is maximal, the set $S' = \{d_1, \ldots, d_m, c_i\}$ does not have the distinct subset sum property. Therefore there are two distinct subsets $U = \{u_1, \ldots, u_p\}$ and $H = \{h_1, \ldots, h_q\}$ of S' such that $u_1 + \cdots + u_p = h_1 + \cdots + h_q$. Since S has the distinct subset sum property, exactly one of U or H contains c_i, say U contains c_i. Now we can get a closed walk T of length d that contains w by traversing the closed walks that correspond to the u_i's once and then traversing the closed walks in $\mathcal{D} - \mathcal{H}$ once, where $\mathcal{H}$ consists of closed walks corresponding to the h_i's. T contains w because T traverses C_i which contains w. The length of T is d because

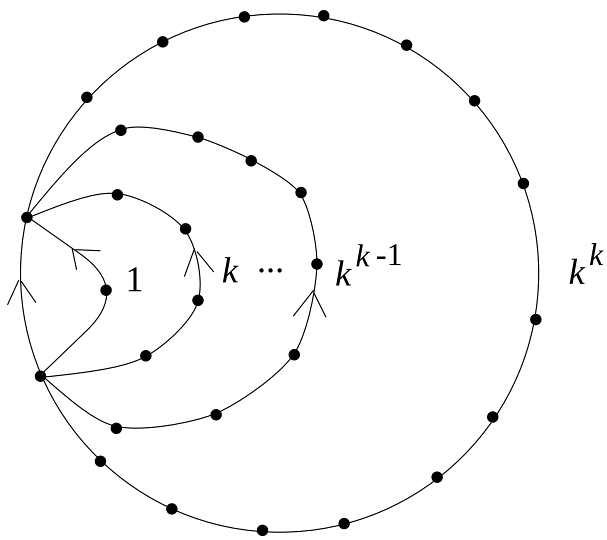
$$|T| = u_1 + \cdots + u_p + \parallel \mathcal{D} - \mathcal{H} \parallel = \parallel \mathcal{H} \parallel - \parallel \mathcal{D} - \mathcal{H} \parallel = \parallel \mathcal{D} \parallel = d.$$

This concludes the proof of the theorem. ∎

3.4 Lower bound

Given an irreducible matrix $A \geq O$, let e_A be the least integer for which the conclusion of Theorem 3.3.1 is true. We recall that in [SW99] we defined the function $\beta(n)$ to be the maximum of e_A over all $n \times n$ non-negative irreducible matrix A. Theorem 3.3.1 shows that $\beta(n) = O(n \log n)$. There is a lower bound of $\Omega(n \log n / \log \log n)$ for $\beta(n)$ due to J. Geelen. We sketch the construction below.

The lower bound is given by the following graph G.



Let $b_0 = k^k$ and $b_i = k^k + k^{i-1}$ for $1 \leq i \leq k$. Then we see that the length of any closed walk in G is a non-negative integer combination of numbers in $B = \{b_0, \ldots, b_k\}$.

We now define the *weight*, $W_B(t)$, of a number t with respect to B. If t cannot be written as a non-negative integer combination of elements of B then $W_B(t) = \infty$. Otherwise suppose

$$\begin{aligned} t &= \sum_{i=0}^{k} c_i b_i \\ &= ck^k + \sum_{i=1}^{k} d_i k^{i-1} \qquad (3.3) \end{aligned}$$

where $0 \leq d_i < k$. In this case we let

$$W_B(t) = \sum_{i=1}^{k} d_i.$$

Note that

$$c \geq W_B(t). \tag{3.4}$$

Let s be the least integer for which the conclusion of Theorem 3.3.2 is true. Since every vertex of G lies on a closed walk of length s we see that $W_B(s - b_i) < \infty$ for all i. By (3.3) and (3.4), we have

$$s \geq k^k \left(1 + \max_{b_i} W_B(s - b_i)\right).$$

We claim

$$\max_{b_i} W_B(s - b_i) \geq k - 1.$$

If the claim is true then we are done because

$$n = |G| = 1 + k + \cdots + k^k \leq 2k^k$$

while

$$s \geq k^k(1 + k - 1) = \Omega(n \log n / \log\log n).$$

To see the claim we write

$$\begin{aligned} s &= \sum_{i=0}^{k} c_i b_i \\ &= ck^k + \sum_{i=1}^{k} d_i k^{i-1} \end{aligned}$$

There are two cases.

Case 1: If $d_i \geq 1$ for $1 \leq i \leq k$, then the claim is true since $W_B(s - b_i) \geq k - 1$ for all i.

Case 2: If $d_i = 0$ for some i, then $W_B(s - b_i) \geq k - 1$.

So the claim is proved and we are done.

Chapter 4

On Two-Sided Infinite Fixed Points of Morphisms

This chapter is based on the work of J. Shallit and M.-w. Wang [SW02].

4.1 Introduction and definitions

Let Σ be a finite alphabet, and let $h : \Sigma^* \to \Sigma^*$ be a morphism on the free monoid, i.e., a map satisfying $h(xy) = h(x)h(y)$ for all $x, y \in \Sigma^*$. If a word w (finite or infinite) satisfies the equation $h(w) = w$, then we call w a *fixed point* of h. Both finite and infinite fixed points of morphisms have long been studied in formal languages. For example, in one of the earliest works on formal languages, Axel Thue [Thu12, Ber95] proved that the one-sided infinite word

$$\mathbf{t} = \mathtt{0110100110010110} \cdots$$

is overlap-free, that is, it contains no subword of the form $axaxa$, where $a \in \{\mathtt{0}, \mathtt{1}\}$, and $x \in (\mathtt{0} + \mathtt{1})^*$. Define a morphism μ by $\mu(\mathtt{0}) = \mathtt{01}$ and $\mu(\mathtt{1}) = \mathtt{10}$. The word $\mathbf{t}$, now called the Thue-Morse infinite word, is the unique one-sided infinite fixed point of μ which starts with $\mathtt{0}$. In fact, nearly every explicit construction of an infinite word avoiding certain patterns involves the fixed point of a morphism; for example, see [HM56, Lee57, Zec58, Ple70]. One-sided infinite fixed points of uniform morphisms also play a crucial role in the theory of automatic sequences; see, for example, [All87].

Because of their importance in formal languages, it is of great interest to characterize *all* the fixed points, both finite and infinite, of a morphism h. This problem was first studied by Head [Hea81], who characterized the finite fixed points of h. Later, Head and Lando [HL86] characterized the one-sided infinite fixed points of h. (For different proofs of these characterizations, see Hamm and Shallit [HS99].) In this chapter we complete the description of all fixed points of morphisms by characterizing the *two-sided* infinite fixed points of h. Two-sided infinite words (sometimes called *bi-infinite words* or *bi-infinite sequences*) play an important role in symbolic dynamics [LM95], and have also been studied in automata theory [NP82, NP86], cellular automata [Hur90], and the theory of codes [VTSS90, DT92]. See also the recent book by D. Perrin and J.-É. Pin [PP04].

We first introduce some notation, some of which is standard and can be found in [HU79]. For single letters, that is, elements of Σ, we use the lower case letters a, b, c, d. For finite words, we use the lower case letters t, u, v, w, x, y, z. For infinite words, we use bold-face letters $\mathbf{t}, \mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z}$. We let ϵ denote the empty word. If S is a set, then by Card S we mean the number of elements of S. We say $x \in \Sigma^*$ is a *subword* of $y \in \Sigma^*$ if there exist words $w, z \in \Sigma^*$ such that $y = wxz$.

If there exists an integer $j \geq 1$ such that $h^j(a) = \epsilon$, then the letter a is said to be *mortal*; otherwise a is *immortal*. The set of mortal letters associated with a morphism h is denoted by M_h. The *mortality exponent* of a morphism h is defined to be the least integer $t \geq 0$ such that $h^t(a) = \epsilon$ for all $a \in M_h$. (If $M_h = \emptyset$, we take $t = 0$.) We write the mortality exponent as $\exp(h) = t$. It is easy to prove that $\exp(h) \leq \text{Card } M_h$. If $h(a) \neq \epsilon$ for all $a \in \Sigma$, then h is *non-erasing*.

We let Σ^ω denote the set of all one-sided right-infinite words over the alphabet Σ. Most of the definitions above extend to Σ^ω in the obvious way. For example, if $\mathbf{w} = c_1c_2c_3\cdots$, then $h(\mathbf{w}) = h(c_1)h(c_2)h(c_3)\cdots$. If $L \subseteq \Sigma^*$ is a language, then we define

$$L^\omega := \{w_1 w_2 w_3 \cdots \ : \ w_i \in L - \{\epsilon\} \text{ for all } i \geq 1\}.$$

Perhaps slightly less obviously, we can also define the word $\overrightarrow{h^\omega}(a)$ for a letter a, provided $h(a) = wax$ and $w \in M_h^*$. In this case, there exists $t \geq 0$ such that $h^t(w) = \epsilon$. Then we define

$$\overrightarrow{h^\omega}(a) := h^{t-1}(w) \cdots h(w)\, w\, a\, x\, h(x)\, h^2(x) \cdots ,$$

which is infinite if and only if $x \notin M_h^*$. Note that the factorization of $h(a)$ as wax, with $w \in M_h^*$ and $x \notin M_h^*$, if it exists, is unique.

In a similar way, we let ${}^\omega\Sigma$ denote the set of all left-infinite words, which are of the form $\mathbf{w} = \cdots c_{-2}c_{-1}c_0$. We write $h(\mathbf{w}) = \cdots h(c_{-2})h(c_{-1})h(c_0)$. We define ${}^\omega L$ to be the set of left-infinite words formed by concatenating infinitely many words from L, that is,

$${}^\omega L := \{\cdots w_{-2}w_{-1}w_0 \ : \ w_i \in L - \{\epsilon\} \text{ for all } i \leq 0\}.$$

If $h(a) = wax$, and $w \notin M_h^*$, $x \in M_h^*$, then we define the left-infinite word

$$\overleftarrow{h^\omega}(a) := \cdots h^2(w)\, h(w)\, w\, a\, x\, h(x) \ \cdots h^{t-1}(x),$$

where $h^t(x) = \epsilon$. Again, if the factorization of $h(a)$ as wax exists, with $w \notin M_h^*$, $x \in M_h^*$, then it is unique.

We can convert left-infinite to right-infinite words (and vice versa) using the reverse operation, which is denoted $\mathbf{w}^R$. For example, if $w = c_0c_1c_2\cdots$, then $\mathbf{w}^R = \cdots c_2c_1c_0$.

We now turn to the notation for two-sided infinite words. These have been much less studied in the literature than one-sided words, and the notation has not been standardized. Some authors consider a pair of two-sided infinite words to be identical if they agree after applying a finite shift to one of the words. Other authors do not. (This distinction is sometimes called "unpointed" vs. "pointed" [Bea85].) In this chapter, we consider both the pointed and unpointed versions of the equation $h(\mathbf{w}) = \mathbf{w}$. As it turns out, the "pointed" version of this equation is quite easy to solve, based on known results, while the "unpointed" case is significantly more difficult. The latter is our main result, which appears as Theorem 4.4.1.

We let $\Sigma^{\mathbb{Z}}$ denote the set of all two-sided infinite words over the alphabet Σ, which are of the form $\cdots c_{-2}c_{-1}c_0.c_1c_2\cdots$. In displaying an infinite word as a concatenation of words, we use a decimal point to the left of the character c_1, to indicate how the word is indexed. Of course, the decimal point is not part of the word itself. We define the shift $\sigma(\mathbf{w})$ to be the infinite word obtained by shifting $\mathbf{w}$ to the left one position, so that

$$\sigma(\cdots c_{-2}c_{-1}c_0.c_1c_2c_3\cdots) = \cdots d_{-2}d_{-1}d_0.d_1d_2d_3\cdots,$$

where $d_i = c_{i+1}$ for all $i \in \mathbb{Z}$. By σ^k we mean the map σ iterated k times, so that

$$\sigma^k(\cdots c_{-2}c_{-1}c_0.c_1c_2c_3\cdots) = \cdots c_{k-1}c_k.c_{k+1}c_{k+2}\cdots.$$

If $\mathbf{w}, \mathbf{x}$ are both two-sided infinite words, and there exists an integer k such that $\mathbf{x} = \sigma^k(\mathbf{w})$, then we call $\mathbf{w}$ and $\mathbf{x}$ *conjugates*, and we write $\mathbf{w} \sim \mathbf{x}$. It is easy to see that $\sim$ is an equivalence relation. We extend this notation to languages as follows: if L is a set of two-sided infinite words, then by $\mathbf{w} \sim L$ we mean there exists $\mathbf{x} \in L$ such that $\mathbf{w} \sim \mathbf{x}$.

If w is a nonempty finite word, then by $w^{\mathbb{Z}}$ we mean the two-sided infinite word $\cdots www.www\cdots$. Using concatenation, we can join a left-infinite word $\mathbf{w} = \cdots c_{-2}c_{-1}c_0$ with a right-infinite word $\mathbf{x} = d_0d_1d_2\cdots$ to form a new two-sided infinite word, as follows:

$$\mathbf{w}.\mathbf{x} := \cdots c_{-2}c_{-1}c_0.d_0d_1d_2\cdots.$$

If $L \subseteq \Sigma^*$ is a set of words, then we define

$$L^{\mathbb{Z}} := \{\cdots w_{-2}w_{-1}w_0.w_1w_2\cdots \ : \ w_i \in L - \{\epsilon\} \text{ for all } i \in \mathbb{Z}\}.$$

If $\mathbf{w} = \cdots c_{-2}c_{-1}c_0.c_1c_2\cdots$, and h is a morphism, then we define

$$h(\mathbf{w}) := \cdots h(c_{-2})h(c_{-1})h(c_0).h(c_1)h(c_2)\cdots \tag{4.1}$$

Finally, if $i = |wa|$, $h(a) = wax$, and $w, x \notin M_h^*$, then we define

$$\overleftrightarrow{h^{\omega;i}}(a) := \cdots h^2(w)\, h(w)\, w\, .a\, x\, h(x)\, h^2(x)\cdots,$$

a two-sided infinite word. Note that in this case the factorization of $h(a)$ as wax is *not* necessarily unique, and we use the superscript i to indicate which a is being chosen.

We can produce one-sided infinite words from two-sided infinite words by ignoring the portion to the right or left of the decimal point. Suppose $\mathbf{w} = \cdots c_{-2}c_{-1}c_0.c_1c_2c_3\cdots$. We define

$$\mathrm{L}(\mathbf{w}) = \cdots c_{-2}c_{-1}c_0,$$

a left-infinite word, and

$$\mathrm{R}(\mathbf{w}) = c_1c_2c_3\cdots,$$

a right-infinite word.

4.2 Finite and one-sided infinite fixed points

In this section we recall the results of Head [Hea81] and Head and Lando [HL86].

Define

$$A_h = \{a \in \Sigma \ : \ \exists\, x, y \in \Sigma^* \text{ such that } h(a) = xay \text{ and } xy \in M_h^*\}$$

and

$$F_h = \{h^t(a) \ : \ a \in A_h \text{ and } t = \exp(h)\}.$$

Note that there is at most one way to write $h(a)$ in the form xay with $xy \in M_h^*$.

Theorem 4.2.1 *[Hea81] Let $h : \Sigma^* \to \Sigma^*$ be a morphism. Then a finite word $w \in \Sigma^*$ has the property that $w = h(w)$ if and only if $w \in F_h^*$.*

Theorem 4.2.2 *[HL86, HS99] The right-infinite word $\mathbf{w}$ is a fixed point of h if and only if at least one of the following two conditions holds:*

(a) $\mathbf{w} \in F_h^\omega$*; or*

(b) $\mathbf{w} \in F_h^* \ \overrightarrow{h^\omega}(a)$ *for some $a \in \Sigma$, and there exist $x \in M_h^*$ and $y \notin M_h^*$ such that $h(a) = xay$.*

Theorem 4.2.3 *[HL86, HS99] The left-infinite word $\mathbf{w}$ is a fixed point of h if and only if at least one of the following two conditions holds:*

(a) $\mathbf{w} \in {}^\omega F_h$*; or*

(b) $\mathbf{w} \in \overleftarrow{h^\omega}(a) F_h^*$ *for some $a \in \Sigma$, and there exist $x \notin M_h^*$ and $y \in M_h^*$ such that $h(a) = xay$.*

4.3 Two-sided infinite fixed points: the "pointed" case

Now we consider the equation $h(\mathbf{w}) = \mathbf{w}$ for two-sided infinite words. We have the following result:

Theorem 4.3.1 *The equation* $h(\mathbf{w}) = \mathbf{w}$ *has a solution if and only if at least one of the following conditions holds:*

(a) $\mathbf{w} \in F_h^{\mathbb{Z}}$*; or*

(b) $\mathbf{w} \in \overleftarrow{h^{\omega}}(a) F_h^*.F_h^{\omega}$ *for some* $a \in \Sigma$*, and there exist* $x \notin M_h^*$*,* $y \in M_h^*$ *such that* $h(a) = xay$*; or*

(c) $\mathbf{w} \in {}^{\omega}F_h.F_h^* \ \overrightarrow{h^{\omega}}(a)$ *for some* $a \in \Sigma$*, and there exist* $x \in M_h^*$*,* $y \notin M_h^*$ *such that* $h(a) = xay$*; or*

(d) $\mathbf{w} \in \overleftarrow{h^{\omega}}(a)\, F_h^*.F_h^* \ \overrightarrow{h^{\omega}}(b)$ *for some* $a, b \in \Sigma$ *and there exist* $x, z \notin M_h^*$*,* $y, w \in M_h^*$*, such that* $h(a) = xay$ *and* $h(b) = wbz$*.*

Proof. Let $\mathbf{w} = \cdots c_{-2}c_{-1}c_0.c_1c_2c_3\cdots$. By definition, we have

$$h(\mathbf{w}) = \cdots h(c_{-2})h(c_{-1})h(c_0).h(c_1)h(c_2)h(c_3)\cdots ,$$

so if $h(\mathbf{w}) = \mathbf{w}$, then we have $h(c_1c_2c_3\cdots) = c_1c_2c_3\cdots$ and $h(\cdots c_{-2}c_{-1}c_0) = \cdots c_{-2}c_{-1}c_0$.

We may now apply Theorem 4.2.2 (resp., Theorem 4.2.3) to $\mathrm{R}(\mathbf{w})$ (resp., $\mathrm{L}(\mathbf{w})$). There are 2 cases to consider for each side, giving $2 \times 2 = 4$ total cases. ■

Example. Let μ be the Thue-Morse morphism, which maps $\mathtt{0} \to \mathtt{01}$, and $\mathtt{1} \to \mathtt{10}$. Define $g = \mu^2$. Then $g(\mathtt{0}) = \mathtt{0110}$, $g(\mathtt{1}) = \mathtt{1001}$. Let $\mathbf{t} = \mathtt{01101001}\cdots$, the one-sided Thue-Morse infinite word. Then there are exactly 4 two-sided infinite fixed points of g, as follows:

$$\begin{aligned}
\mathbf{t}^R.\mathbf{t} &= \cdots\mathtt{10010110.01101001}\cdots \\
\overline{\mathbf{t}}^R.\mathbf{t} &= \cdots\mathtt{01101001.01101001}\cdots \\
\overline{\mathbf{t}}^R.\overline{\mathbf{t}} &= \cdots\mathtt{01101001.10010110}\cdots \\
\mathbf{t}^R.\overline{\mathbf{t}} &= \cdots\mathtt{10010110.10010110}\cdots .
\end{aligned}$$

All of these fall under case (d) of Theorem 4.3.1. Incidentally, all four of these words are overlap-free.

4.4 Two-sided infinite fixed points: the "unpointed" case

In this section, we characterize the two-sided infinite fixed points of a morphism in the "unpointed" case. That is, our goal is to characterize the solutions to $h(\mathbf{w}) \sim \mathbf{w}$. The following theorem is the main result of the chapter.

Theorem 4.4.1 *Let h be a morphism. Then the two-sided infinite word $\mathbf{w}$ satisfies the relation $h(\mathbf{w}) \sim \mathbf{w}$ if and only if at least one of the following conditions holds:*

(a) $\mathbf{w} \sim F_h^{\mathbb{Z}}$*; or*

(b) $\mathbf{w} \sim \overleftarrow{h^{\omega}}(a) \,.\, F_h^{\omega}$ *for some* $a \in \Sigma$*, and there exist* $x \notin M_h^*$ *and* $y \in M_h^*$ *such that* $h(a) = xay$*; or*

(c) $\mathbf{w} \sim {}^{\omega}F_h \,.\, \overrightarrow{h^{\omega}}(a)$ *for some* $a \in \Sigma$*, and there exist* $x \in M_h^*$ *and* $y \notin M_h^*$ *such that* $h(a) = xay$*; or*

(d) $\mathbf{w} \sim \overleftarrow{h^{\omega}}(a) \,.\, F_h^* \; \overrightarrow{h^{\omega}}(b)$ *for some* $a, b \in \Sigma$ *and there exist* $x, z \notin M_h^*$*,* $y, w \in M_h^*$*, such that* $h(a) = xay$ *and* $h(b) = wbz$*; or*

(e) $\mathbf{w} \sim \overleftrightarrow{h^{\omega;i}}(a)$ *for some* $a \in \Sigma$*, and there exist* $x, y \notin M_h^*$ *such that* $h(a) = xay$ *with* $|xa| = i$*; or*

(f) $\mathbf{w} = (xy)^{\mathbb{Z}}$ *for some* $x, y \in \Sigma^+$ *such that* $h(xy) = yx$*.*

Before we begin the proof of Theorem 4.4.1, we state and prove three useful lemmas.

Lemma 4.4.2 *Suppose* $\mathbf{w}$, $\mathbf{x}$ *are* 2 *two-sided infinite words with* $\mathbf{w} \sim \mathbf{x}$. *Then* $h(\mathbf{w}) \sim h(\mathbf{x})$.

Proof. Since $\mathbf{w} \sim \mathbf{x}$, there exists j such that $\mathbf{x} = \sigma^j(\mathbf{w})$. Then $h(\mathbf{x}) = \sigma^k(h(\mathbf{w}))$, where

$$k = \begin{cases} |h(c_1c_2\cdots c_j)|, & \text{if } j \geq 0; \\ -|h(c_{j+1}c_{j+2}\cdots c_{-1}c_0)|, & \text{if } j < 0. \end{cases} \tag{4.2}$$

■

Our second lemma concerns periodicity of infinite words. We say a two-sided infinite word

$$\mathbf{w} = \cdots c_{-2}c_{-1}c_0.c_1c_2\cdots$$

is *periodic* if there exists a nonempty word x such that $\mathbf{w} = x^{\mathbb{Z}}$, i.e., if there exists an integer $p \geq 1$ such that $c_k = c_{k+p}$ for all integers k. The integer p is called a *period* of $\mathbf{w}$.

Lemma 4.4.3 *Suppose* $\mathbf{w} = \cdots c_{-2}c_{-1}c_0.c_1c_2\cdots$ *is a two-sided infinite word such that there exists a one-sided right-infinite word* $\mathbf{x}$ *and infinitely many negative indices* $0 > i_1 > i_2 > \cdots$ *such that*

$$\mathbf{x} = c_{i_j}c_{i_j+1}c_{i_j+2}\cdots$$

for $j \geq 1$. *Then* $\mathbf{w}$ *is periodic.*

Proof. By assumption

$$\mathbf{x} = c_{i_j}c_{i_j+1}c_{i_j+2}\cdots = c_{i_{j+1}}c_{i_{j+1}+1}c_{i_{j+1}+2}\cdots$$

for $j \geq 1$. Hence $c_{i_j+k} = c_{i_{j+1}+k}$ for all $k \geq 0$, and so the right-infinite word $\mathbf{x}$ is periodic of period $i_j - i_{j+1}$. Since this is true for all $j \geq 1$, it follows that $\mathbf{x}$ is periodic of period $g = \gcd_{j\geq 1}(i_j - i_{j+1})$, i.e., $c_{i_j+k} = c_{i_j+g+k}$ for all $j \geq 1, k \geq 0$. Since $i_j \to -\infty$, it follows that $c_k = c_{k+g}$ for all k, and so $\mathbf{w}$ is periodic of period g. ■

Our third lemma is an application of Theorem 3.2.1 from Chapter 3 to the growth functions of iterated morphisms.

Lemma 4.4.4 *Let $h : \Sigma^* \to \Sigma^*$ be a morphism. Then*

(a) there exist integers i, j with $0 \le i < j$ and $|h^i(w)| \le |h^j(w)|$ for all $w \in \Sigma^*$; and

(b) there exists an integer M depending only on $k = \text{Card}\,\Sigma$ such that for all $h : \Sigma^* \to \Sigma^*$, we have $j \le M$.

We note that part (a) was asserted without proof by Cobham [Cob68]. The proof below connecting the lemma with Theorem 3.2.1 does not give the best upper bound M. For the best possible upper bound $M = |\Sigma|$, see [CMS$^+$03].

Proof. Let $|x|_a$ denote the number of occurrences of the letter a in the string x. Given a morphism $h : \Sigma^* \to \Sigma^*$ for $\Sigma = \{a_1, a_2, \ldots, a_d\}$, we define the *incidence matrix* $M = M(h)$ as follows:

$$M = (m_{i,j})_{1 \le i,j \le d}$$

where $m_{i,j} = |h(a_j)|_{a_i}$.

Note that

$$|h(w)|_{a_i} = \sum_{1 \le j \le d} |h(a_j)|_{a_i} \; |w|_{a_j}$$

and so

$$\begin{bmatrix} |h(w)|_{a_1} \\ |h(w)|_{a_2} \\ \vdots \\ |h(w)|_{a_d} \end{bmatrix} = M(h) \begin{bmatrix} |w|_{a_1} \\ |w|_{a_2} \\ \vdots \\ |w|_{a_d} \end{bmatrix}.$$

Hence

$$\begin{bmatrix} |h^n(w)|_{a_1} \\ |h^n(w)|_{a_2} \\ \vdots \\ |h^n(w)|_{a_d} \end{bmatrix} = (M(h))^n \begin{bmatrix} |w|_{a_1} \\ |w|_{a_2} \\ \vdots \\ |w|_{a_d} \end{bmatrix}$$

and finally

$$|h^n(w)| = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \end{bmatrix} M(h)^n \begin{bmatrix} |w|_{a_1} \\ |w|_{a_2} \\ \vdots \\ |w|_{a_d} \end{bmatrix}.$$

The lemma now follows from Theorem 3.2.1.

Now we can prove Theorem 4.4.1.

Proof. ($\Longleftarrow$): Suppose case (a) holds, and $\mathbf{w} \sim F_h^{\mathbb{Z}}$. Then there exists $\mathbf{x} \in F_h^{\mathbb{Z}}$ with $\mathbf{w} \sim \mathbf{x}$. Since $\mathbf{x} \in F_h^{\mathbb{Z}}$, we can write $\mathbf{x} = \cdots x_{-2}x_{-1}x_0.x_1x_2\cdots$, where $x_i \in F_h$ for all $i \in \mathbb{Z}$. Since $x_i \in F_h$, we have $h(x_i) = x_i$ for all $i \in \mathbb{Z}$. It follows that $h(\mathbf{x}) = \mathbf{x}$. Now, applying Lemma 4.4.2, we conclude that $h(\mathbf{w}) \sim h(\mathbf{x}) = \mathbf{x} \sim \mathbf{w}$.

Next, suppose case (b) holds, and $\mathbf{w} \sim \overleftarrow{h^{\omega}}(a).F_h^{\omega}$. Then $\mathbf{w} \sim \mathbf{x}$ for some $\mathbf{x}$ of the form

$$\mathbf{x} = \overleftarrow{h^{\omega}}(a).x_1x_2x_3\cdots,$$

where $x_i \in F_h$ for all $i \geq 1$, and $h(a) = xay$ with $x \notin M_h^*$ and $y \in M_h^*$. Then we have $h(\mathbf{x}) = \mathbf{x}$, and by Lemma 4.4.2, we conclude that $h(\mathbf{w}) \sim h(\mathbf{x}) = \mathbf{x} \sim \mathbf{w}$.

Cases (c), (d), and (e) are similar to case (b).

Finally, if case (f) holds, then

$$h(\mathbf{w}) = h(\cdots xyxy.xyxy\cdots) = \cdots yxyx.yxyx\cdots,$$

and so $h(\mathbf{w}) = \sigma^k(\mathbf{w})$ for $k = |x|$.

($\Longrightarrow$): First, we introduce some notation. We define $h_k(\mathbf{w}) = \sigma^{-k}(h(\mathbf{w}))$. More precisely, if

$$h_k(\cdots c_{-2}c_{-1}c_0.c_1c_2\cdots) = \cdots b_{-2}b_{-1}b_0.b_1b_2\cdots,$$

then we define

$$b_{s(i-1)+1}\cdots b_{s(i)} = h(c_i)$$

where

$$s(i) := \begin{cases} |h(c_1c_2\cdots c_i)| + k, & \text{if } i \geq 0; \\ k - |h(c_{i+1}c_{i+2}\cdots c_0)|, & \text{if } i < 0. \end{cases} \tag{4.3}$$

By hypothesis $\mathbf{w} = h_k(\mathbf{w})$ for some k. Then $h(c_i) = c_{s(i-1)+1}\cdots c_{s(i)}$ where s is defined as in Eq. (4.3). We define the set C as follows: $C = \{i \in \mathbb{Z} \ : \ s(i) = i\}$. Our argument is divided into two major cases, depending on whether or not C is empty.

Case 1: $C \neq \emptyset$. In this case there are four subcases, depending on the form of C.

Case 1a: $\inf C = -\infty$ and $\sup C = \infty$. Then there exists a two-sided infinite sequence $\ldots, e_{-2}, e_{-1}, e_0, e_1, e_2, \ldots$ such that $s(e_i) = e_i$ for all $i \in \mathbb{Z}$. But $s(e_i) = e_i$ and $s(e_{i+1}) = e_{i+1}$ together imply that

$$h(c_{e_i+1}\cdots c_{e_{i+1}}) = c_{s(e_i)+1}\cdots c_{s(e_{i+1})} = c_{e_i+1}\cdots c_{e_{i+1}}.$$

Hence, defining $x_i := c_{e_i+1}\cdots c_{e_{i+1}}$, we have $h(x_i) = x_i$ for all $i \in \mathbb{Z}$. By Theorem 4.2.1, then, $x_i \in F_h^*$ for all $i \in \mathbb{Z}$. Now define $\mathbf{x} = \cdots x_{-2}x_{-1}x_0.x_1x_2\cdots$. Since $\mathbf{w} \sim \mathbf{x}$, it follows that $\mathbf{w} \sim F_h^{\mathbb{Z}}$. This corresponds to case (a).

Case 1b: $\inf C = r > -\infty$ and $\sup C = \infty$. Then there exists an infinite sequence $e_0 = r, e_1, e_2, \ldots$ such that $s(e_i) = e_i$ for all integers $i \geq 0$. Define $\mathbf{z} := \cdots c_{r-2}c_{r-1}c_r$, a left-infinite word; then $h(\mathbf{z}) = \mathbf{z}$. It now follows from Theorem 4.2.3 and the form of $\mathbf{z}$ that $\mathbf{z} \in \overleftarrow{h^\omega}(a)$ for some a with $h(a) = xay$, $x \notin M_h^*$, $y \in M_h^*$.

For $i \geq 0$ define $y_i := c_{e_i+1} \cdots c_{e_{i+1}}$. As in Case 1a, we have $h(y_i) = y_i$ for all $i \geq 0$. Thus, letting $\mathbf{x} = \mathbf{z}.y_0 y_1 y_2 \cdots$, we have $\mathbf{x} \in \overleftarrow{h^\omega}(a).F_h^\omega$. Since $\mathbf{w} \sim \mathbf{x}$, it follows that $\mathbf{w} \sim \overleftarrow{h^\omega}(a).F_h^\omega$. This corresponds to case (b).

Case 1c: $\inf C = -\infty$ and $\sup C = t < \infty$. This is similar to Case 1b, and by the same reasoning we find $\mathbf{w} \sim {}^\omega F_h.\ \overrightarrow{h^\omega}(a)$ for some a with $h(a) = xay$ and $x \in M_h^*$, $y \notin M_h^*$. This corresponds to case (c).

Case 1d: $\inf C = r > -\infty$ and $\sup C = t < \infty$. Let $\mathbf{x} = \cdots c_{r-2} c_{r-1} c_r$ (a left-infinite word), $v = c_{r+1} \cdots c_t$ (a finite word), and $\mathbf{z} = c_{t+1} c_{t+2} \cdots$ (a right-infinite word). Then we have $h(\mathbf{x}) = \mathbf{x}$, $h(v) = v$, and $h(\mathbf{z}) = \mathbf{z}$. By Theorem 4.2.3 and the form of $\mathbf{x}$, there exists $a \in \Sigma$ such that $\mathbf{x} \in \overleftarrow{h^\omega}(a)$ with $h(a) = xay$, and $x \notin M_h^*$, $y \in M_h^*$. By Theorem 4.2.1 we know $v \in F_h^*$. By Theorem 4.2.2 and the form of $\mathbf{z}$, there exists $b \in \Sigma$ such that $\mathbf{z} \in \overrightarrow{h^\omega}(b)$ with $h(b) = wbz$, and $w \in M_h^*$, $z \notin M_h^*$. If we let $\mathbf{y} = \mathbf{x}.v\mathbf{z}$, then $\mathbf{w} \sim \mathbf{y}$, and so $\mathbf{w} \sim \overleftarrow{h^\omega}(a)\,.\,F_h^*\ \overrightarrow{h^\omega}(b)$. Thus case (d) holds.

Case 2: $C = \emptyset$. Once again there are several cases to consider.

Case 2a: There exist integers i, j with $i < j$ such that

$$s(i) > i \text{ but } s(j) < j. \tag{4.4}$$

Now consider the set

$$S = \{(i', j')\ :\ i' \geq i,\ j' \leq j,\ s(i') > i',\ \text{and } s(j') < j'\}.$$

By hypothesis, S is nonempty. Define

$$\begin{aligned} j_0 &:= \min\{j' : \exists i' \text{ such that } (i', j') \in S\}; \\ i_0 &:= \max\{i' : (i', j_0) \in S\}. \end{aligned}$$

Suppose there exists an integer k with $i_0 < k < j_0$. If $s(k) < k$, then $(i_0, k) \in S$ and $k < j_0$, contradicting the definition of j_0. If $s(k) > k$, then $(k, j_0) \in S$ and $k > i_0$, contradicting the definition of i_0. Hence $s(k) = k$. But this is impossible by our assumption. It follows that $j_0 = i_0 + 1$. Then $s(i_0) > i_0$, but $s(i_0 + 1) < i_0 + 1$, a contradiction, since $s(i_0) \leq s(i_0 + 1)$. Hence this case cannot occur.

Case 2b: There exists an integer r such that $s(i) < i$ for all $i < r$, and $s(i) > i$ for all $i \geq r$. Then $h(c_r) = c_{s(r-1)+1} \cdots c_{s(r)}$, which by the inequalities contains $c_{r-1}c_rc_{r+1}$ as a subword. Therefore, letting $a = c_r$, it follows that

$$\mathbf{w} \sim \mathbf{u}\, x \,.\, a\, y\, \mathbf{v},$$

where $\mathbf{u} = \cdots c_{s(r-1)-1}c_{s(r-1)}$ is a left-infinite word, $x = c_{s(r-1)+1} \cdots c_{r-1}$ and $y = c_{r+1} \cdots c_{s(r)}$ are finite words, and $\mathbf{v} = c_{s(r)+1}c_{s(r)+2} \cdots$ is a right-infinite word. Furthermore, we have $h(\mathbf{u}x) = \mathbf{u}$, $h(a) = xay$, and $h(y\mathbf{v}) = \mathbf{v}$.

Now the equation $h(y\mathbf{v}) = \mathbf{v}$ implies that $h(y)$ is a prefix of $\mathbf{v}$, and by an easy induction we have $h(y)h^2(y)h^3(y) \cdots$ is a prefix of $\mathbf{v}$. Suppose this prefix is finite. Then $y \in M_h^*$, and so $h(y)h^2(y)h^3(y) \cdots = h(y)h^2(y) \cdots h^t(y)$, where $t = \exp(h)$. Define $z = h(y)h^2(y) \cdots h^t(y)$. Then $s(r + |y| + |z|) = r + |y| + |z|$, a contradiction, since we have assumed $C = \emptyset$. It follows that $\mathbf{z} := h(y)h^2(y)h^3(y) \cdots$ is right-infinite and hence $y \notin M_h^*$.

By exactly the same reasoning, we find that $\cdots h^3(x)\, h^2(x)\, h(x)$ is a left-infinite suffix of $\mathbf{u}$. We conclude that $\mathbf{w} \sim \overleftrightarrow{h^{\omega;i}}(a)$, and hence case (e) holds.

Case 2c: $s(i) > i$ for all $i \in \mathbb{Z}$. Let $\mathbf{w} = \cdots c_{-2}c_{-1}c_0.c_1c_2 \cdots$.

Now consider the following factorization of certain conjugates of $\mathbf{w}$, as follows: for $i \leq 0$, we have $\mathbf{w} \sim \mathbf{x}_i\, y_i \,.\, \mathbf{z}_i$, where $\mathbf{x}_i = \cdots c_{i-2}c_{i-1}$ (a left-infinite word), $y_i = c_i \cdots c_{s(i-1)}$ (a finite word), and $\mathbf{z}_i = c_{s(i-1)+1}c_{s(i-1)+2}\cdots$ (a right-infinite word). Note that $i-1 < s(i-1)$ by assumption, so $i \leq s(i-1)$; hence y_i is nonempty. Evidently we have

$$\begin{aligned} h(\mathbf{x}_i) &= \mathbf{x}_i\, y_i; \quad \text{and} \\ h(y_i\, \mathbf{z}_i) &= \mathbf{z}_i. \end{aligned} \tag{4.5}$$

Now the equation $h(y_i\mathbf{z}_i) = \mathbf{z}_i$ implies that $h(y_i)$ is a prefix of $\mathbf{z}_i$. Now an easy induction, as in Case 2b, shows that $v := h(y_i)h^2(y_i)h^3(y_i)\cdots$ is a prefix of $\mathbf{z}_i$. If v were finite, then we would have $y_i \in M_h^*$, and so $s(j) = j$ for $j = s(i-1) + |v|$, a contradiction, since $C = \emptyset$. Hence v is right-infinite, and so $y_i \notin M_h^*$. There are now two cases to consider: (i) $\sup_{i\leq 0}(s(i) - i) < +\infty$, and (ii) $\sup_{i\leq 0}(s(i) - i) = +\infty$.

Case 2ci: Suppose $\sup_{i\leq 0}(s(i) - i) = d < +\infty$. It then follows that $|y_i| \leq d$. Hence there is a finite word u such that $y_i = u$ for infinitely many indices $i \leq 0$. From the above argument we see that the right-infinite word $h(u)h^2(u)h^3(u)\cdots$ is a suffix of $\mathbf{w}$, beginning at position $s(i-1)+1$, for infinitely many indices $i \leq 0$. We now use Lemma 4.4.3 to conclude that $\mathbf{w}$ is periodic.

Thus we can write $\mathbf{w} = \cdots c_{-2}c_{-1}c_0.c_1c_2\cdots$, and $\mathbf{w} = \cdots vvv.vvv\cdots$, where $v = c_1c_2\cdots c_p$ for some integer $p \geq 1$. Without loss of generality, we may assume p is minimal.

We claim $|h(v)| = p$. For if not we must have $|h(v)| = q$, for $q \neq p$, and then since $h(\mathbf{w}) \sim \mathbf{w}$, we would have $\mathbf{w}$ is periodic with periods p and q, hence periodic of period $\gcd(p, q)$. But since p was minimal we must have $p \,|\, q$. Hence $q \geq 2p$. Now let $s(p) = l$; since $s(i) > i$ for all i we must have $l > 0$. Then

$$h(c_1c_2\cdots c_p) = c_{s(-1)+1}\cdots c_{s(p)} = c_{l-q+1}\cdots c_l.$$

It now follows that

$$s(ip) = l - q + iq \tag{4.6}$$

for all integers i. Now $p < q$, so $p \le q - 1$, and hence $p < q - 1 + q/l$. Hence, multiplying by $-l$, we get $-lp > l - ql - q$. Now take $i = -l$ in Eq. (4.6), and we have

$$s(-lp) = l - q - lq < -lp,$$

a contradiction, since $s(i) > i$ for all i. It follows that $|h(v)| = p$.

There exists k such that $h(c_1 c_2 \cdots c_p) = c_{k+1} c_{k+2} \cdots c_{k+p}$. Using the division theorem, write $k = jp + r$, where $0 \le r < p$. Define

$$y = c_{k+1} \cdots c_{(j+1)p} = c_{r+1} \cdots c_p;$$

$$x = c_{(j+1)p+1} \cdots c_{k+p} = c_1 \cdots c_r.$$

We have $h(xy) = yx$, and $v = xy$. Then $\mathbf{w} = v^{\mathbb{Z}} = (xy)^{\mathbb{Z}}$.

We know that $|v| \ge 1$, so $xy \ne \epsilon$. Suppose $y = \epsilon$. Then $h(x) = x$, and so $x \in F_h^*$. It follows that $\mathbf{w} \in F_h^{\mathbb{Z}}$. A similar argument applies if $x = \epsilon$. However, if $\mathbf{w} \in F_h^{\mathbb{Z}}$, then $C \ne \emptyset$, a contradiction. Thus $x, y \ne \epsilon$, and case (f) holds.

Case 2cii: $\sup_{i \le 0}(s(i) - i) = +\infty$. Recall that $s(i) > i$ for all $i \in \mathbb{Z}$ and $\mathbf{w} = \cdots c_{-2} c_{-1} c_0 . c_1 c_2 \cdots$. Define

$$\begin{aligned} \mathbf{x} &:= \cdots c_{-2} c_{-1} c_0; \\ y &:= c_1 c_2 \cdots c_{s(0)}; \\ \mathbf{z} &:= c_{s(0)+1} c_{s(0)+2} \cdots . \end{aligned}$$

Then $\mathbf{w} = \mathbf{x}.y\mathbf{z}$ and $h(\mathbf{x}) = \mathbf{x}y$, $h(y\mathbf{z}) = \mathbf{z}$.

Define $B_j(k) = s^j(k) - s^{j-1}(k)$, where s^j denotes the j-fold composition of the function s with itself. First we prove the following technical lemma.

Lemma 4.4.5 *For all integer $r \geq 1$ there exists an integer $n \leq 0$ such that $B_j(n) > r$ for $1 \leq j \leq t$.*

Proof. By induction on t. For $t = 1$ the result follows since

$$\sup_{i \leq 0} B_1(i) = \sup_{i \leq 0}(s(i) - i) = +\infty.$$

Now assume the result is true for t; we prove it for $t+1$. Define $m := \max_{a \in \Sigma} |h(a)|$. By induction there exists an integer n_1 such that $B_j(n_1) > mr + m^{t+1}$ for $1 \leq j \leq t$. Then, by the definition of m there exist an integer $n_2 < n_1$ with $n_1 - n_2 < m$, and an integer n_3 such that $s(n_3) = n_2$.

Now $h(c_{n_3+1} \cdots c_{n_2}) = c_{s(n_3)+1} \cdots c_{s(n_2)}$, so $s(n_2) - s(n_3) \leq m(n_2 - n_3)$. Similarly, we have

$$s^j(n_2) - s^j(n_3) \leq m^j(n_2 - n_3) \tag{4.7}$$

for all $j \geq 0$. By the same reasoning, we have

$$s^j(n_1) - s^j(n_2) \leq m^j(n_1 - n_2) \leq m^j(m - 1) \tag{4.8}$$

for all $j \geq 0$. Thus we find

$$\begin{aligned}
B_1(n_3) &= s(n_3) - n_3 \\
&= n_2 - n_3 \\
&\geq \frac{s(n_2) - s(n_3)}{m} \quad \text{(by Eq. (4.7))} \\
&= \frac{s(n_2) - n_2}{m} \\
&= \frac{(s(n_1) - n_1) - ((s(n_1) - s(n_2)) - (n_1 - n_2))}{m} \\
&= \frac{B_1(n_1) - ((s(n_1) - s(n_2)) - (n_1 - n_2))}{m} \\
&> \frac{mr + m^{t+1} - m(m-1)}{m} \quad \text{(by induction and Eq. (4.8))} \\
&> r.
\end{aligned}$$

Similarly, for $2 \leq j \leq t+1$, we have

$$\begin{aligned}
B_j(n_3) &= s^j(n_3) - s^{j-1}(n_3) \\
&= s^{j-1}(n_2) - s^{j-2}(n_2) \\
&= (s^{j-1}(n_1) - s^{j-2}(n_1)) - ((s^{j-1}(n_1) - s^{j-1}(n_2)) - (s^{j-2}(n_1) - s^{j-2}(n_2))) \\
&= B_{j-1}(n_1) - ((s^{j-1}(n_1) - s^{j-1}(n_2)) - (s^{j-2}(n_1) - s^{j-2}(n_2))) \\
&> mr + m^{t+1} - m^{j-1}(m-1) \quad \text{(by Eq. (4.8))} \\
&\geq r.
\end{aligned}$$

It thus follows that we can take $n = n_3$. This completes the proof of Lemma 4.4.5. ■

Now let M be the integer specified in Lemma 4.4.4, and define $r := \sup_{1 \leq i \leq M} B_i(0)$. By Lemma 4.4.5 there exists an integer $n \leq 0$ such that $B_j(n) > r$ for $1 \leq j \leq M$. Define

$w := c_{n+1} \cdots c_0$. We have

$$\begin{aligned} |h^j(w)| &= s^j(0) - s^j(n); \quad \text{and} \\ |h^{j-1}(w)| &= s^{j-1}(0) - s^{j-1}(n). \end{aligned}$$

It follows that

$$\begin{aligned} |h^j(w)| &= (s^j(0) - s^{j-1}(0)) - (s^j(n) - s^{j-1}(n)) + |h^{j-1}(w)| \\ &= B_j(0) - B_j(n) + |h^{j-1}(w)| \\ &< B_j(0) - r + |h^{j-1}(w)| \\ &\leq |h^{j-1}(w)| \end{aligned}$$

for $1 \leq j \leq M$. But this contradicts Lemma 4.4.4. This contradiction shows that this case cannot occur.

Case 2d: $s(i) < i$ for all $i \in \mathbb{Z}$. This case is the mirror image of Case 2c[1], and the proof is identical. The proof of Theorem 4.4.1 is complete. ■

4.5 Some examples

In this section we consider some examples of Theorem 4.4.1.

Example 1. Consider the morphism f defined by $\mathtt{a} \to \mathtt{bb}$, $\mathtt{b} \to \epsilon$, $\mathtt{c} \to \mathtt{aad}$, $\mathtt{d} \to \mathtt{c}$. Let

$$\mathbf{w} = \cdots \mathtt{aadbbbbcaadbbbbc.aadbbbbcaadbbbbc} \cdots .$$

[1]Note that $s(i) > i$ for all i implies that $s(i-1) > i-1$. Therefore $s(i-1)+1 > i$, and hence Case 2d really *is* the mirror image of Case 2c.

Then

$$f(\mathbf{w}) = \cdots \mathtt{bbbbcaadbbbbcaad.bbbbcaadbbbbcaad} \cdots .$$

This falls under case (f) of Theorem 4.4.1.

Example 2. Consider the morphism φ defined by $\mathtt{0} \to \mathtt{201}$, $\mathtt{1} \to \mathtt{012}$, and $\mathtt{2} \to \mathtt{120}$. Then if

$$\mathbf{w} = \varphi^{\omega;2}(\mathtt{0}) = \cdots c_{-2}c_{-1}.c_0c_1c_2 \cdots = \cdots \mathtt{1202.01012} \cdots ,$$

we have $\varphi(\mathbf{w}) \sim \mathbf{w}$. This falls under case (d) of Theorem 4.4.1. Incidentally, c_i equals the sum of the digits, modulo 3, in the balanced ternary representation of i.

4.6 The equation $h(xy) = yx$ in finite words

It is not difficult to see that it is decidable whether any of conditions (a)–(e) of Theorem 4.4.1 hold. However, this is somewhat less obvious for condition (f), which demands that the equation $h(xy) = yx$ possess a nontrivial[2] solution. We conclude this chapter by discussing the solvability of this equation and give a characterization of the solution set.

To do so it is useful to extend the notation $\sim$, previously used for two-sided infinite words, to finite words. We say $w \sim z$ for $w, z \in \Sigma^*$ if w is a cyclic shift of z, i.e., if there exist $x, y \in \Sigma^*$ such that $w = xy$ and $z = yx$. It is now easy to verify that $\sim$ is an equivalence relation. Furthermore, if $w \sim z$, and h is a morphism, then $h(w) \sim h(z)$. Thus condition (f) can be restated as $h(z) \sim z$. The following theorem shows that the solvability of the equation $h(xy) = yx$ is decidable.

Theorem 4.6.1 *Let h be a morphism $h : \Sigma^* \to \Sigma^*$. Then $h(z) \sim z$ has a solution $z \neq \epsilon$ if and only if F_{h^d} is nonempty for some* $1 \leq d \leq \text{Card } \Sigma$.

[2]By nontrivial we mean $xy \neq \epsilon$.

Proof. $\Longleftarrow$: Suppose F_{h^d} is nonempty for some d, say $x \in F_{h^d}$. Then by definition of F_{h^d}, $h^d(x) = x$. Let $y = h(x) \cdots h^{d-1}(x)$ and $z = xy$. Then $h(xy) = yx$ and so $h(z) \sim z$.

$\Longrightarrow$: Suppose $h(z) \sim z$. Then $|h^n(z)| = |z|$ for all $n \geq 0$, and so there exist $0 \leq i < j$ such that $h^i(z) = h^j(z)$. In another word $h^i(z)$ is a finite fixed point of h^{j-i}. Hence $F_{h^{j-i}}$ is nonempty. This implies A_{h^d} is nonempty for some $1 \leq d \leq \text{Card } \Sigma$. Thus F_{h^d} is nonempty. ∎

Remarks.

1. Note that Theorem 4.6.1 does not characterize all the finite solutions of $h(z) \sim z$; it simply gives a necessary and sufficient condition for solutions to exist.

2. As we have seen in Theorem 4.2.1, the set of finite solutions to $h(z) = z$ is finitely generated, in that the solution set can be written as S^* for some finite set T. However, the set of solutions to $h(z) \sim z$ need not even be context-free. For consider the morphism defined by $h(\mathtt{a}) = \mathtt{b}$, $h(\mathtt{b}) = \mathtt{c}$, $h(\mathtt{c}) = \mathtt{a}$, and let

$$T := \{z \in \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}^* \ : \ h(z) \sim z\}.$$

If T were context-free, then so would be $T \cap \mathtt{a}^*\mathtt{b}^*\mathtt{c}^*$. But

$$T \cap \mathtt{a}^*\mathtt{b}^*\mathtt{c}^* = \{\mathtt{a}^i\mathtt{b}^i\mathtt{c}^i \ : \ i \geq 0\}$$

which is not context-free.

We finish with a discussion of the set T of words z for which $h(z) \sim z$. From the proof of Theorem 4.6.1, we know that there exist $i < j$ such that $h^i(z)$ is a fixed point of h^{j-i}. Since $h^i(z) \sim z$, we may restrict our attention to the set $S = T \cap (\bigcup_{i \geq 1} F^*_{h^i})$. Our set T then is the set of all cyclic permutations of words in S.

To describe S we introduce an auxiliary morphism $h' : \Sigma' \to \Sigma'$, where $\Sigma' \subseteq \Sigma$. A letter a is a member of Σ' if and only if the following three conditions hold:

(1) a is an immortal letter of h;

(2) $h^i(a)$ contains exactly one immortal letter for all $i \geq 1$; and

(3) $h^i(a)$ contains a for some $i \geq 1$.

We define the morphism h' by $h'(a) = a'$ where a' is the unique immortal letter in $h(a)$.

The relation of h' to S is as follows. If $z \in S$, then $z \in F^*_{h^i}$ for some i. Hence $z = w_1 \cdots w_r$ where $w_j = x_j a_j y_j \in F_{h^i}$, and a_j is an immortal letter for $1 \leq j \leq r$. It follows easily that $a_j \in \Sigma'$. Since h cyclically permutes z, h' cyclically permutes $z' = a_1 \cdots a_r$. The words x_j and y_j are uniquely specified by i and a_j. Therefore, we now concentrate on the set T' of words z' that are cyclically permuted by h'.

Suppose $\Sigma' = \{a_1, \ldots, a_s\}$. Since h' acts as a permutation P on Σ', there exists a unique factorization of P into disjoint cycles $\{(a^i_0, \ldots, a^i_{t_i-1})\}_{i=1}^m$ where $h'(a^i_j) = a^i_{(j+1) \bmod t_i}$. The t_i's are the *length* of the cycles. We will construct a finite set R of regular languages from the set of cycles as follows.

Suppose $c = (b_0, \ldots, b_{t-1})$ is a cycle appearing in the factorization of P. Define

$$w_i = b_i b_{(i+1) \bmod t} \cdots b_{(i+t-1) \bmod t}$$

for $0 \leq i < t$. Let φ denote Euler's totient function, and let $k_1, \ldots, k_{\varphi(t)}$ be the integers in the range $[1, t]$ that are relatively prime to t. If $w_i = d_0 \cdots d_{t-1}$, define $u_{i,j} = d_0\ d_{k_j} \cdots d_{((t-1)k_j) \bmod t}$ for $1 \leq j \leq \varphi(t)$. Now for $1 \leq j \leq \varphi(t)$ we define

$$L_j(c) = \bigcup_{0 \leq i < t} u^*_{i,j}.$$

Thus from cycle c we constructed $\varphi(t)$ regular languages $L_j(c)$. We repeat this construction for each cycle and let R' be the set of the regular languages thus obtained from all cycles. Each regular language in R will be the union of several regular languages of R'. The union is defined as follows. Each regular language $L_j(c)$ in R' is associated with a pair (t, k_j) where t is the length of c and k_j is an integer relatively prime to t. A set of languages $L_{j_1}(c_1), \ldots, L_{j_n}(c_n)$ in R' belong to the same language of R if and only if the system of congruence equations $xk_{j_\alpha} = 1 \bmod t_\alpha$ for $1 \leq \alpha \leq n$ has a solution. Note that a language in R' may be contained in several languages of R.

We say a word w is the *perfect shuffle* of words $w_1, \ldots, w_r$ if $|w_1| = \cdots = |w_r|$ and the first r symbols of w are the first symbols of $w_1, \ldots, w_r$ in that order, the second r symbols of w are the second symbols of $w_1, \ldots, w_r$ in that order, and so on. We claim that z' is the perfect shuffle of words in a language in R if and only if h' cyclically permutes z'. We prove this below. For convenience we will refer to z' as z and h' as h in the remaining discussion.

Suppose z is a word in a language L in R. By our construction $z = z_0 \cdots z_{tn-1} = (d_0\ d_k \cdots d_{(t-1)k \bmod t})^n$ for some n where $d_0 \cdots d_{t-1}$ is a representation of some cycle $c = (b_0, \ldots, b_{t-1})$ (i.e. $d_0 = b_i$, $d_1 = b_{i+1 \bmod t}$, etc.) and k is relatively prime to t. By our definition of c, $h(z) = (d_{1 \bmod t}\ d_{k+1 \bmod t} \cdots d_{(t-1)k+1 \bmod t})^n$. And for each $1 \leq i \leq tn-1$ where $ik = 1 \bmod t$ we have $h(z) = (d_{ik \bmod t}\ d_{(i+1)k \bmod t} \cdots d_{(i+t-1)k \bmod t})^n = z_i\ z_{i+1 \bmod tn} \cdots z_{i-1}$. This shows, in particular that $h(z) \sim z$. Now suppose z is the perfect shuffle of two words w and w' in a language L in R. The general case of an arbitrary number of words is the same except it is notationally more cumbersome. We may assume with the same convention as above that $w = (d_0\ d_k \cdots d_{(t-1)k \bmod t})^n$ and $w' = (d'_0\ d'_{k'} \cdots d'_{(t'-1)k' \bmod t'})^{n'}$ and $z = z_0 \cdots z_{tn+t'n'-1} = d_0\ d'_0\ \cdots$. From the definitions we have 1) there exists i such that $ik = 1 \bmod t$ and $ik' = 1 \bmod t'$, and 2) $tn = t'n'$. Simple calculation shows that if such i exists, we may assume $1 \leq i < tn = t'n'$. As before we have

$$
\begin{aligned}
h(z) &= d_{1 \bmod t}\ d'_{1 \bmod t'}\ d_{k+1 \bmod t}\ d'_{k'+1 \bmod t'} \cdots \\
&= d_{ik \bmod t}\ d'_{ik' \bmod t'}\ d_{(i+1)k \bmod t}\ d'_{(i+1)k' \bmod t'} \cdots \\
&= z_{2i}\ z_{2i+1 \bmod 2tn} \cdots z_{2i-1}.
\end{aligned}
$$

This shows $h(z) \sim z$.

Now suppose $z = z_0 \cdots z_r$ is cyclically permuted by h. Then we have the following chain of indices: $0 = i_0, \ldots, i_l = 0$, where i_s are distinct for $0 \leq s < l$, and h acting on z as a cyclical permutation sends the index i_s to $i_{s+1 \bmod l}$. Note that there may be more than one way to define h's action on z. We fix an arbitrary one. We order the indices $i_0, \ldots, i_{l-1}$, say $i_0 = j_0 < \cdots < j_{l-1}$. Let $v_{j_s} = z_{j_s} \cdots z_{j_{s+1}-1}$ for $0 \leq s < l-1$ and $v_{j_{l-1}} = z_{j_{l-1}} \cdots z_r$. We claim that all the v_{j_s} are of the same length. To prove this it suffices to show $|v_{i_s}| = |v_{i_{s+1 \bmod l}}|$ for $0 \leq s \leq l-1$. Suppose this is false. Then there exists s such that $|v_{i_s}| > |v_{i_{s+1 \bmod l}}|$. Since h sends the index i_s to $i_{s+1 \bmod l}$ and $i_s + 1$ to $i_{s+1 \bmod l} + 1$ and so on, the difference in length implies that there is an index j with $i_s < j \leq i_s + |v_{i_s}| - 1$ such that h sends the index j to $j_{\alpha+1}$ where $j_\alpha = i_{s+1 \bmod l}$. This is impossible by our definition of v_{i_s}. Hence all v_{i_s} (or v_{j_s}) are of the same length. It follows that $h(v_{i_s}) = v_{i_{s+1 \bmod l}}$ since h sends the index i_s to $i_{s+1 \bmod l}$ and h is a one to one map.

Let $n = |v_{j_0}|$. For $0 \leq \alpha < n$ define

$$w_\alpha = z_{j_0+\alpha}\, z_{j_1+\alpha} \cdots z_{j_{l-1}+\alpha}.$$

Note that z is the perfect shuffle of $w_0, \ldots, w_{n-1}$. Suppose w is one of the w_{j_s}, say $w = y_0 \cdots y_{l-1}$. By our construction h cyclically permutes w and there exists a permutation $p_0, \ldots, p_{l-1}$ of $0, \ldots, l-1$ such that $p_0 = 0$ and h acting on w sends the index p_i to $p_{i+1 \bmod l}$. It follows that the set of distinct symbols of w must be the same as the set of symbols of one of the cycles, say $c = (b_0, \ldots, b_{t-1})$ of h. Let $b_{i_j} = y_j$ for $0 \leq j \leq l-1$. We claim that for each $0 \leq j \leq l-1$, $k_j = i_{j+1 \bmod l} - i_j \mod t$ has the same value, and $k = k_0$ is relatively prime to t. If the claim is true then we must have $w \in L_\gamma(c)$ where $k_\gamma = k$. And since we chose w to be an arbitrary w_α, the same applies to all w_α. First we show all the k_j have the same value. Let $\pi(j)$ be the permutation where h sends the index j to $\pi(j)$. It suffices to prove $\pi(j) - j \mod l$ is constant, say equal to d, for $0 \leq j \leq l-1$ because then 1) d must be relatively prime to l because $0, d, \ldots, (l-1)d$ cover all residue classes $(\bmod\, l)$, and 2) thus there exist unique $1 \leq k \leq l-1$ such that $kd = 1 \bmod l$. We note that $k_j = k \bmod t$ for $0 \leq j \leq l-1$. Now suppose not all $d_j = \pi(j) - j \bmod l$ have the same value. Then there must be a j such that $d_j > d_{j+1 \bmod l}$. Since h sends the index j to $\pi(j)$ and $j+1 \bmod l$ to $\pi(j)+1 \bmod l$ and so on, $d_j > d_{j+1 \bmod l}$ implies that there is a index β in the range $j+1 \bmod l \cdots \pi(j)-1 \bmod l$ such that h sends β to $\pi(\pi(j))$. But this is impossible because h permutes the indices and by definition h already sends $\pi(j)$ to $\pi(\pi(j))$. It remains to show that $k = k_0$ is relatively prime to t. Observe that $w = b_{i_0} \cdots b_{i_{l-1}} = b_{i_0} b_{i_0+k \bmod t} \cdots b_{i_0+(l-1)k \bmod t}$. Since w contains all distinct symbols in c, $0, k, \ldots, (l-1)k$ cover all residue classes $(\bmod\, t)$. Hence k is relatively prime to t.

We have shown above that $w_\alpha \in L_\alpha(c_\alpha)$ for $0 \leq \alpha < n$ where c_α is of length t_α and k_α is relatively prime to t_α. Suppose $c_\alpha = (b_0^\alpha, \cdots, b_{t_\alpha-1}^\alpha)$ for $0 \leq \alpha < n$. Since $L_\alpha(c_\alpha)$ contains all representations of c_α (i.e. $L_\alpha(c_\alpha)$ contains $b_i^\alpha \cdots b_{i+t_\alpha-1 \bmod t_\alpha}^\alpha$ for all $0 \leq i \leq t_\alpha - 1$, etc.) we may assume $w_\alpha = b_0^\alpha \cdots b_{k_\alpha(l-1) \bmod t_\alpha}^\alpha$ for $0 \leq \alpha < n$. We know $h(v_{i_s}) = v_{i_{s+1 \bmod l}}$. In particular we have $h(v_0) = h(v_{i_0}) = v_{i_1}$. Since z is a perfect shuffle of $w_0, \ldots, w_{n-1}$ we have

$$\begin{aligned} h(v_0) &= h(b_0^0 \cdots b_0^{n-1}) \\ &= b_1^0 \cdots b_1^{n-1} \\ &= b_{xk_0 \bmod t_0}^0 \cdots b_{xk_{n-1} \bmod t_{n-1}}^{n-1} \\ &= v_{i_1} \end{aligned}$$

for some $1 \leq x \leq l-1$. Thus the system of congruence equations $xk_\alpha = 1 \bmod t_\alpha$ for $0 \leq \alpha < n$ has a solution. Hence by the construction of R, $w_0, \ldots, w_{n-1}$ belong to a language in R. This concludes our discussion.

Chapter 5

Avoiding Large Squares in Infinite Binary Words

This chapter is based on the work of N. Rampersad, J. Shallit and M.-w. Wang [RSW03].

5.1 Introduction

A *square* is a nonempty word of the form xx, as in the English word `murmur`. It is easy to see that every word of length ≥ 4 constructed from the symbols 0 and 1 contains a square, so it is impossible to avoid squares in infinite binary words. However, in 1974, Entringer, Jackson, and Schatz [EJS74] proved the surprising fact that there exists an infinite binary word containing no squares xx with $|x| \geq 3$. Further, the bound 3 is best possible.

A *cube* is a nonempty word of the form xxx, as in the English sort-of-word `shshsh`. Dekking [Dek76] showed that there exists an infinite binary word that contains no cubes xxx and no squares yy with $|y| \geq 4$. Furthermore, the bound 4 is best possible.

Dekking's construction used iterated morphisms. If $h : \Sigma^* \to \Sigma^*$ and $h(a) = ax$ for some letter $a \in \Sigma$, then we say that h is *prolongable* on a, and we can then iterate h infinitely often to get the fixed point $h^\omega(a) := a\,x\,h(x)\,h^2(x)\,h^3(x)\cdots$.

A morphism is *k-uniform* if $|h(a)| = k$ for all $a \in \Sigma$; it is *uniform* if it is k-uniform for some k. Uniform morphisms have particularly nice properties. For example, the class of words generated by applying a coding to infinite iteration of k-uniform morphisms coincides with the class of k-automatic sequences, generated by finite automata [AS03].

Dekking's construction used a non-uniform morphism. In this chapter we show how to obtain, using the image of a uniform morphism, an infinite binary word that is cubefree and avoids squares yy with $|y| \geq 4$. Our construction is somewhat simpler than Dekking's.

5.2 A cubefree word without arbitrarily long squares

In this section we construct an infinite cubefree binary word avoiding squares yy with $|y| \geq 4$.

We introduce the following notation for alphabets: $\Sigma_k := \{0, 1, \ldots, k-1\}$.

Theorem 5.2.1 *There is a squarefree infinite word over Σ_4 with no occurrences of the subwords* 12*,* 13*,* 21*,* 32*,* 231*, or* 10302*.*

Proof. Let the morphism h be defined by

$$\begin{aligned} 0 &\rightarrow 0310201023 \\ 1 &\rightarrow 0310230102 \\ 2 &\rightarrow 0201031023 \\ 3 &\rightarrow 0203010201 \end{aligned}$$

Then we claim the fixed point $h^\omega(0)$ has the desired properties.

First, we claim that if $w \in \Sigma_4^*$ then $h(w)$ has no occurrences of 12, 13, 21, 32, 231, or 10302. For if any of these words occur as subwords of $h(w)$, they must occur within some $h(a)$ or straddling the boundary between $h(a)$ and $h(b)$, for some single letters a, b. They do not; this easy verification is left to the reader.

Next, we prove that if w is any squarefree word over Σ_4 having no occurrences of 12, 13, 21, or 32, then $h(w)$ is squarefree.

We argue by contradiction. Let $w = a_1 a_2 \cdots a_n$ be a squarefree string such that $h(w)$ contains a square, i.e., $h(w) = xyyz$ for some $x, z \in \Sigma_4^*$, $y \in \Sigma_4^+$. Without loss of generality, assume that w is a shortest such string, so that $0 \le |x|, |z| < 10$.

Case 1: $|y| \le 20$. In this case we can take $|w| \le 5$. To verify that $h(w)$ is squarefree, it therefore suffices to check each of the 49 possible words $w \in \Sigma_4^5$ to ensure that $h(w)$ is squarefree in each case.

Case 2: $|y| > 20$. First, we establish the following result.

Lemma 5.2.2 *(a) Suppose $h(ab) = th(c)u$ for some letters $a, b, c \in \Sigma_4$ and strings $t, u \in \Sigma_4^*$. Then this inclusion is trivial (that is, $t = \epsilon$ or $u = \epsilon$) or u is not a prefix of $h(d)$ for any $d \in \Sigma_4$.*

(b) Suppose there exist letters a, b, c and strings s, t, u, v such that $h(a) = st$, $h(b) = uv$, and $h(c) = sv$. Then either $a = c$ or $b = c$.

Proof.

(a) This can be verified with a short computation. In fact, the only a, b, c for which the equality $h(ab) = th(c)u$ holds nontrivially is $h(31) = th(2)u$, and in this case $t = 020301$, $u = 0102$, so u is not a prefix of any $h(d)$.

(b) This can also be verified with a short computation. If $|s| \ge 6$, then no two distinct letters have images under h that share a prefix of length 6. If $|s| \le 5$, then $|t| \ge 5$, and no two distinct letters have images under h that share a suffix of length 5.

■

Once Lemma 5.2.2 is established, the rest of the argument is fairly standard. It can be found, for example, in [KS03]; we omit the details here. ■

Theorem 5.2.3 *Let* $\mathbf{w}$ *be any infinite word satisfying the conditions of Theorem 5.2.1. Define a morphism* g *by*

$$\begin{aligned} 0 &\rightarrow 010011 \\ 1 &\rightarrow 010110 \\ 2 &\rightarrow 011001 \\ 3 &\rightarrow 011010 \end{aligned}$$

Then $g(\mathbf{w})$ *is a cubefree word containing no squares* xx *with* $|x| \geq 4$.

Before we begin the proof, we remark that all the words 12, 13, 21, 32, 231, 10302 must indeed be avoided, because

$$\begin{array}{rl} g(12) & \text{contains the squares } (0110)^2,\ (1100)^2,\ (1001)^2 \\ g(13) & \text{contains the square } (0110)^2 \\ g(21) & \text{contains the cube } (01)^3 \\ g(32) & \text{contains the square } (1001)^2 \\ g(231) & \text{contains the square } (10010110)^2 \\ g(10302) & \text{contains the square } (100100110110)^2. \end{array}$$

Proof. The proof parallels the proof of Theorem 5.2.1. Let $w = a_1 a_2 \cdots a_n$ be a squarefree string, with no occurrences of 12, 13, 21, 32, 231, or 10302. We first establish that if $g(w) = xyyz$ for some $x, z \in \Sigma_4^*$, $y \in \Sigma_4^+$, then $|y| \leq 3$. Without loss of generality, assume w is a shortest such string, so $0 \leq |x|, |z| < 6$.

Case 1: $|y| \leq 12$. In this case we can take $|w| \leq 5$. To verify that $g(w)$ contains no squares yy with $|y| \geq 4$, it suffices to check each of the 41 possible words $w \in \Sigma_4^5$.

Case 2: $|y| > 12$. First, we establish the analogue of Lemma 5.2.2.

Lemma 5.2.4 *(a) Suppose $g(ab) = tg(c)u$ for some letters $a, b, c \in \Sigma_4$ and strings $t, u \in \Sigma_4^*$. Then this inclusion is trivial (that is, $t = \epsilon$ or $u = \epsilon$) or u is not a prefix of $g(d)$ for any $d \in \Sigma_4$.*

(b) Suppose there exist letters a, b, c and strings s, t, u, v such that $g(a) = st$, $g(b) = uv$, and $g(c) = sv$. Then either $a = c$ or $b = c$, or $a = 2$, $b = 1$, $c = 3$, $s = 0110$, $t = 01$, $u = 0101$, $v = 10$.

Proof.

(a) This can be verified with a short computation. The only a, b, c for which $g(ab) = tg(c)u$ holds nontrivially are

$$\begin{aligned} g(01) &= 010\ g(3)\ 110 \\ g(10) &= 01\ g(2)\ 0011 \\ g(23) &= 0110\ g(1)\ 10. \end{aligned}$$

But none of 110, 0011, 10 are prefixes of any $g(d)$.

(b) If $|s| \geq 5$ then no two distinct letters have images under g that share a prefix of length 5. If $|s| \leq 3$ then $|t| \geq 3$, and no two distinct letters have images under g that share a suffix of length 3. Hence $|s| = 4$, $|t| = 2$. But only $g(2)$ and $g(3)$ share a prefix of length 4, and only $g(1)$ and $g(3)$ share a suffix of length 2.

■

The rest of the proof is exactly parallel to the proof of Theorem 5.2.1, with the following exception. When we get to the final case, where $|y|$ is divisible by 6, we can use Lemma 5.2.4 to rule out every case except where $x = 0101$, $z = 01$, $a_1 = 1$, $a_j = 3$, and $a_n = 2$. Thus $w = 1\alpha 3\alpha 2$ for some string $\alpha \in \Sigma_4^*$. This special case is ruled out by the following lemma:

Lemma 5.2.5 *Suppose $\alpha \in \Sigma_4^*$, and let $w = 1\alpha 3\alpha 2$. Then either w contains a square, or w contains an occurrence of one of the subwords* 12, 13, 21, 32, 231, *or* 10302.

Proof. This can be verified by checking (a) all strings w with $|w| \leq 4$, and (b) all strings of the form $w = abcw'de$, where $a, b, c, d, e \in \Sigma_4$ and $w' \in \Sigma_4^*$. (Here w' may be treated as an indeterminate.) ■

It now remains to show that if w is squarefree and contains no occurrence of 12, 13, 21, 32, 231, or 10302, then $g(w)$ is cubefree. If $g(w)$ contains a cube yyy, then it contains a square yy, and from what precedes we know $|y| \leq 3$. It therefore suffices to show that $g(w)$ contains no occurrence of 0^3, 1^3, $(01)^3$, $(10)^3$, $(001)^3$, $(010)^3$, $(011)^3$, $(100)^3$, $(101)^3$, $(110)^3$. The longest such string is of length 9, so it suffices to examine the 16 possibilities for $g(w)$ where $|w| = 3$. This is left to the reader.

The proof of Theorem 5.2.3 is now complete. ■

Corollary 5.2.6 *If g and h are defined as above, then*

$$g(h^\omega(0)) = 0100110110100101100100110110010100110101100100110110010111010\cdots$$

is cubefree, and avoids all squares xx with $|x| \geq 4$.

Chapter 6

New Problems of Pattern Avoidance

This chapter is based on the work of J. Loftus, J. Shallit and M.-w. Wang [LSW99].

6.1 Introduction

Pattern avoidance problems have long been studied in formal language theory, and have interesting applications to group theory, universal algebra, and other areas. For example, Axel Thue constructed an infinite squarefree word over $\{0, 1, 2\}$; i.e., a word that contains no subword of the form xx, where x is a nonempty word [Thu12, Ber95].

Eventually, generalizations of Thue's problem were considered. Erdős, for example, suggested trying to find infinite words containing no subword of the form xy, where y is a permutation of the letters of x. Such words are now sometimes called "abelian squarefree" [Bro71]. The reader can read other interesting papers on pattern avoidance [BEM79, Cur93, Cas93].

In this chapter, we consider some new generalizations of Thue's problem. We start with some notation. Let Σ, Γ be finite alphabets. We let Σ^ω denote the set of all one-sided infinite words over Σ, and we let $\Sigma^\infty = \Sigma^* \cup \Sigma^\omega$. If $x \in \Sigma^+$, then by x^ω we mean the one-sided infinite word $xxx\cdots$.

If there exist words $x, y \in \Sigma^*$, $w, z \in \Sigma^\infty$ such that $w = xyz$, then we say y is a finite *subword* of w. Suppose we are given a finite or infinite subset $P \subseteq \Sigma^*$. Then we say a word $w \in \Sigma^\infty$ *avoids* P if we cannot write $w = xyz$ such that $y \in P$. We say P is *avoidable* over Σ if it is possible to construct an infinite word $\mathbf{w} \in \Sigma^\omega$ which avoids P.

Sometimes we employ a common abuse of notation. For example, instead of saying that the infinite word $\mathbf{w}$ avoids $\{xx \ : \ x \in \Sigma^+\}$, we will instead simply say that $\mathbf{w}$ *avoids the pattern* xx. When we use this formulation, we always assume the strings in the pattern are nonempty.

We define $\Sigma_k = \{0, 1, 2, \ldots, k-1\}$ for some integer $k \geq 2$, and we define the morphism $\sigma_k(a) = (a + 1) \bmod k$. If the subscript k is clear from the context, we omit it. In this chapter, we consider avoiding patterns of the form $x\sigma^i(x)$.

We use two notational conventions that may be somewhat confusing. First, we think of the elements of Σ_k as residue class representatives so that, for example, -1 and 2 denote the same element of Σ_3. Second, since we allow negative numbers in words, we sometimes use the notation $(a_1, a_2, a_3, \ldots, a_n)$ to denote the word $a_1a_2a_3 \cdots a_n$. Thus, for example, 012 and $(0, -2, -1)$ denote the same element of Σ_3^*.

Some of the infinite words we construct arise from iterated morphisms. Call a morphism $h : \Gamma^* \to \Sigma^*$ *non-erasing* if $h(a) \neq \epsilon$ for all $a \in \Gamma$.

6.2 Avoiding $x\sigma(x)$

It is clear that over $\Sigma_2 = \{0, 1\}$, there are only two infinite words avoiding the pattern $x\sigma(x)$, namely 0^ω and 1^ω. However, we have the following result:

Theorem 6.2.1 *Over Σ_k for $k \geq 3$, there are uncountably many infinite words avoiding $x\sigma(x)$.*

Proof. Define $a_1 = 1$, and set $a_{i+1} = a_i + 1$ or $a_i + 2$, according to choice. Then

$$\mathbf{w} = \prod_{i \geq 1} ((-i) \bmod 3)^{a_i} = 2^{a_1}\, 1^{a_2}\, 0^{a_3}\, 2^{a_4}\, 1^{a_5}\, 0^{a_6} \cdots$$

avoids the pattern $x\sigma(x)$, and there are uncountably many such words. ■

6.3 Avoiding xx, $x\sigma(x)$, …, $x\sigma^j(x)$ simultaneously

The following theorem constitutes our main result. It characterizes, for each integer $j \geq 0$, the smallest integer k for which we can avoid the $j+1$ patterns xx, $x\sigma(x)$, …, $x\sigma^j(x)$ simultaneously over $\Sigma_k = \{0, 1, \ldots, k-1\}$.

Theorem 6.3.1 *(a) One can avoid the pattern xx over Σ_3, and 3 is best possible.*

(b) One can avoid the patterns xx and $x\sigma(x)$ simultaneously over Σ_5, and 5 is best possible.

(c) One can avoid the patterns xx, $x\sigma(x)$, $x\sigma^2(x)$ simultaneously over Σ_5, and 5 is best possible.

(d) One can avoid the patterns xx, $x\sigma(x)$, $x\sigma^2(x)$, $x\sigma^3(x)$ simultaneously over Σ_6, and 6 is best possible.

(e) For $j \geq 4$, one can avoid the $j+1$ patterns xx, $x\sigma(x)$, …, $x\sigma^j(x)$ simultaneously over Σ_{j+4}, and $j+4$ is best possible.

Remark. Our proofs of these facts are of two different types. First, in order to show that it is possible to avoid a certain set of patterns over Σ_k, we explicitly construct an infinite word over Σ_k having the desired property. Second, to show that k is optimal for a certain set of patterns, we use a classical breadth-first tree traversal technique, as follows:

Suppose we wish to avoid a given set of words P over Σ_k. We maintain a queue, Q, and initialize it with the empty word ϵ. If the queue is empty, we are done. Otherwise, we take the next element w from the queue, and form k new words by appending $0, 1, \ldots, k-1$ to it. For each new word wa, we check to see whether some suffix of wa occurs in P. If it does, we discard it; otherwise we add it to the queue.

If this algorithm terminates, we have proved that it is *not* possible to avoid P over Σ_k. The resulting proof may be represented in the form of a tree, with the leaves representing minimal length prefixes that contain an occurrence of one of the patterns as a suffix.

In the particular case of the patterns we discuss in this section, two additional efficiencies are possible. First, since a word w simultaneously avoids the patterns xx, $x\sigma(x)$, …, $x\sigma^j(x)$ iff $\sigma(w)$ does, we may without loss of generality consider only the words that begin with the letter 0. Second, if the last letter was a, then the next letter must be contained in the set $\{a + j + 1, \ldots, a + k - 1\}$, for otherwise our word would contain a length-2 subword of the form $x\sigma^i(x)$ for $0 \le i \le j$. This observation significantly cuts down on the branching factor of the trees we generate.

Proof of Theorem 6.3.1. Let us start with assertion (a). As already noted, a classical result due to Thue shows that one can avoid the pattern xx over $\Sigma_3 = \{0, 1, 2\}$ [Thu12, Ber95]. Furthermore, it is an old and easy observation that any word of length ≥ 4 over $\Sigma_2 = \{0, 1\}$ contains an occurrence of the pattern xx. More generally, we have

Proposition 6.3.2 *Let $k \ge 2$ be an integer, and let r be an integer with $1 \le r < k$. Then any word of length ≥ 4 over Σ_k contains an occurrence of the pattern $x\sigma^a(x)$ for some $a \not\equiv r$ (mod k).*

Proof. We use the tree traversal algorithm. Assume the first letter is 0; then if the next letter is $a \ne r$, we are done. Hence assume the next letter is r. Then, by a similar argument, the next letter must be $2r$, and the next $3r$. However, the word $(0, r, 2r, 3r)$ contains the pattern $x\sigma^{2r}(x)$ for $x = (0, r)$. Since $r \ne 0$, we have $2r \not\equiv r$ (mod k). ■

Now let us prove assertion (b) of Theorem 6.3.1. By Proposition 6.3.2 with $r = 2$, one cannot avoid the patterns xx and $x\sigma(x)$ simultaneously over Σ_3. We also have

Proposition 6.3.3 *Every word of length* ≥ 24 *over* Σ_4 *contains an occurrence of either* xx *or* $x\sigma(x)$.

Proof. We use the tree traversal algorithm. The resulting tree has depth 24 and contains 233 leaves. Figure 1 below lists these leaves in breadth-first order. ∎

0202	0203210202	031321310310	021020320210313	02102032021032020	0313213103132103131
0213	0210203203	031321031320	021020320210310	02102032021032021	0313213103132103132
0310	0210313210	031321031321	021020320210321	02103132131031320	0320210203202103203
02031	0210320213	032021032020	021031321310310	02103132131032132	0321310313213103202
03131	0313213102	032021032021	031321310321313	02103202102032020	0321310313213103203
03203	0313210202	032131031320	031321310321310	02103202102032131	0321310313213103210
021021	0313210203	032131032131	032021020320213	02103202102032132	02032102032021020320
031320	0313210310	032131032132	032021020321021	02103202102032103	02032102032021032020
032020	0320210202	032102032020	032131031321313	03132131031321020	02032102032021032021
032132	0320210313	032102032102	032131031321020	03132131031321021	02103132131031321020
032103	0320210310	032102032103	032131031321021	03132131031321032	02103132131031321021
0203203	0320210321	0203202102031	032131031321032	03213103132131031	02103132131031321032
0210202	0321310310	0203202103203	032102032021021	03213103132103131	03202102032021032020
0210310	0321310320	0203210203203	0203202102032020	03213103132103132	03202102032021032021
0210321	02032021021	0210203210202	0203202102032021	03210203202102031	03213103132131032132
0320213	02032102031	0210203210203	0203202102032131	03210203202103203	03210203202102032131
0321313	02102032020	0210313213102	0203202102032132	020320210203210202	03210203202102032132
0321021	02102032131	0210320210202	0203202102032103	020320210203210203	03210203202102032103
02032020	02102032132	0313213103131	0210203202102031	020321020320210202	021031321310313210310
02032132	02102032103	0313213103202	0210203202102032	020321020320210313	021032021020320210313
02032103	02103132132	0313213103203	0210203202103203	020321020320210310	021032021020320210310
02102031	03132103131	0313213103210	0210313213103131	020321020320210321	021032021020320210321
02103131	03202102031	0320210203203	0210313213103202	021031321310313213	032131031321310321313
02103203	03202103203	0321020320213	0210313213103203	021031321310321313	032131031321310321310
03132132	03213103131	02032021032020	0210313213103210	021031321310321310	032102032021020321021
03213102	03213103210	02032021032021	0210320210203203	021032021020320213	0203210203202102032131
03210202	03210203203	02032102032020	0313213103132131	021032021020321021	0203210203202102032132
020320213	03210203213	02102032021021	0313213103132132	031321310313210310	0203210203202102032103
020321313	020320210202	02103202102031	0320210203202102	032021020320210313	0210313213103132103131
020321021	020320210313	03132131031320	0320210203210202	032021020320210310	0210313213103132103132
021031320	020320210310	03132131032132	0320210203210203	032021020320210321	0210320210203202103203
021032020	020320210321	03202102032020	0321310313213102	032102032021020320	0321020320210203210202
031321313	020321020321	03202102032131	0321310313210310	032102032021032020	0321020320210203210203
031321021	021020320213	03202102032132	0321020320210202	032102032021032021	02032102032021020321021
031321032	021020321021	03202102032103	0321020320210313	0203210203202102031	02103202102032021032020
032021021	021031321313	03213103132132	0321020320210310	0203210203202103203	02103202102032021032021
032102031	021032021021	020320210203203	0321020320210321	0210320210203202102	020321020320210203210202
0203213102	021032021031	020321020320213	02032021020321021	0210320210203210202	020321020320210203210203
0203213103	021032021032	021020320210202	02032102032021021	0210320210203210203	

Figure 1: Leaves of the tree giving a proof of Proposition 6.3.3.

Thus we cannot avoid the patterns xx and $x\sigma(x)$ simultaneously over Σ_4. However, we can avoid the patterns xx and $x\sigma(x)$ simultaneously over Σ_5. This will follow from Theorem 6.3.4 below.

Next, let us prove assertion (c). As we have seen in Proposition 6.3.2 above, every word of length ≥ 4 over Σ_4 contains an occurrence of one of the patterns xx, $x\sigma(x)$, or $x\sigma^2(x)$.

We now show

Theorem 6.3.4 *It is possible to simultaneously avoid the patterns xx, $x\sigma(x)$, and $x\sigma^2(x)$ over Σ_5.*

Before starting the proof, we introduce some notation. If $\mathbf{w} = a_1a_2a_3\cdots$ is a word over Σ_k, then

$$\Delta(\mathbf{w}) := (a_2 - a_1, a_3 - a_2, a_4 - a_3, \ldots),$$

where the differences are, of course, taken mod k. Similarly, we write

$$S(\mathbf{w}) = (0, a_1, a_1 + a_2, a_1 + a_2 + a_3, \ldots),$$

where the sums are, of course, taken mod k. Note that $\Delta(S(\mathbf{w})) = \mathbf{w}$, and if $a_1 = 0$, then $S(\Delta(\mathbf{w})) = \mathbf{w}$. Finally, if $x = a_1 \cdots a_m \in \Sigma_k^*$, we define $s_k(x) = (\Sigma_{1\leq i\leq m} a_i) \bmod k$.

The following lemma relates occurrences of patterns of the form $x\sigma^a(x)$ in $\mathbf{w}$ to other, easier-to-study patterns in $\Delta(\mathbf{w})$.

Lemma 6.3.5 *Let $w \in \Sigma_k^\infty$, and let $a \in \Sigma_k$. Then w avoids the pattern $x\sigma^a(x)$ iff $\Delta(w)$ avoids $\{ycy \ : \ y \in \Sigma_k^*, \ c \in \Sigma_k, \ \text{and } s_k(yc) = a\}$.*

Proof. Suppose w contains an occurrence of the pattern $x\sigma^a(x)$. Write $x = b_1b_2\cdots b_i$. Then

$$w = w'b_1b_2\cdots b_i\sigma^a(b_1)\cdots\sigma^a(b_i)\cdots .$$

Thus

$$\Delta(w) = \Delta(w'b_1), (b_2 - b_1, \ldots, b_i - b_{i-1}, \sigma^a(b_1) - b_i, b_2 - b_1, \ldots, b_i - b_{i-1}, \ldots),$$

and hence contains ycy with

$$y = (b_2 - b_1, \ldots, b_i - b_{i-1}), \quad c = \sigma^a(b_1) - b_i.$$

Also

$$\begin{aligned} s_k(yc) &= (b_2 - b_1) + \cdots + (b_i - b_{i-1}) + \sigma^a(b_1) - b_i \\ &= (b_i - b_1) + (a + b_1 - b_i) \\ &= a. \end{aligned}$$

Now suppose $\Delta(w)$ contains a subword ycy with $y \in \Sigma_k^*$, $c \in \Sigma_k$, and $s_k(yc) = a$. Then $\Delta(w) = xycyz$ for some $x = b_1 b_2 \cdots b_j$ and $y = d_1 d_2 \cdots d_i$. Then

$$\Delta(w) = b_1 b_2 \cdots b_j d_1 d_2 \cdots d_i c d_1 d_2 \cdots d_i \cdots .$$

Then if e is the first letter of w, we have

$$\begin{aligned} w = {} & (e, e + b_1, e + b_1 + b_2, \ldots, e + b_1 + b_2 + \cdots + b_j, e + f + d_1, e + f + d_1 + d_2, \ldots, \\ & e + f + d_1 + d_2 + \cdots + d_i, e + f + g + c, e + f + g + c + d_1, e + f + g + c + d_1 + d_2, \ldots, \\ & e + f + g + c + d_1 + d_2 + \cdots + d_i, \ldots) \end{aligned}$$

where $f := b_1 + b_2 + \cdots + b_j$ and $g := d_1 + d_2 + \cdots + d_i$. It follows that w contains an occurrence of $x\sigma^a(x)$, where $x = (e + f, e + f + d_1, \ldots, e + f + d_1 + d_2 + \cdots + d_i)$ and $a = g + c$. But $g + c = s_k(d_1 d_2 \cdots d_i c) = s_k(yc)$. ■

Now to prove Theorem 6.3.4, it suffices to construct an infinite word $\mathbf{v}$ where $\mathbf{v}$ avoids

$$P_2 := \{ycy \ : \ y \in \Sigma_5^*, \ c \in \Sigma_5, \ \text{and} \ s_5(yc) \in \{0, 1, 2\}\}.$$

For then we could set $\mathbf{w} = S(\mathbf{v})$, and by Lemma 6.3.5, $\mathbf{w}$ avoids the patterns xx, $x\sigma(x)$, and $x\sigma^2(x)$ over Σ_5. We construct such a $\mathbf{v}$ using the following theorem.

Theorem 6.3.6 *Let h be the morphism over $\{3, 4\}$ defined by $h(4) = 4433$ and $h(3) =$*

44433. *Let w be a finite word. If w avoids P_2, then $h(w)$ avoids P_2.*

Proof. We prove the contrapositive.

Suppose $h(w)$ contains an occurrence of the pattern ycy with $y \in \Sigma_5^*$, $c \in \Sigma_5$, and $s_5(yc) \in \{0, 1, 2\}$. Write $h(w) = z_1 ycy z_2$. Without loss of generality, we may assume that $|z_1|$ is as small as possible, or, in other words, that the occurrence of ycy we are dealing with lies as far to the left as possible within $h(w)$.

Also note that $s_5(i) = s_5(h(i))$ for $i \in \{3, 4\}$, and so it follows that $s_5(w) = s_5(h(w))$ for all finite strings $w \in \{3, 4\}^*$.

We claim that if ycy is a subword of $h(w)$ for some w such that y, c obey the given conditions, then $|y| \geq 5$. Table 1 below suffices to prove this.

The explanation of the table is as follows. We examine all possible subwords yc of length ≤ 5 that occur in $\{4433, 44433\}^*$. For each such subword, it suffices to show that either $s_5(yc) \notin \{0, 1, 2\}$, or ycy cannot occur as a subword of $h(w)$ for any $w \in \{3, 4\}^*$. For this last check, it suffices to observe that if ycy contains any of the subwords $434, 343, 333$, or 4444, then it cannot occur as a subword of $h(w)$.

$\|y\|$	yc	$s_5(yc)$	ycy	ycy contains forbidden subword	if so, which one
0	3	3	3	no	
	4	4	4	no	
1	33	1	333	yes	333
	34	2	343	yes	343
	43	2	434	yes	434
	44	3	444	no	
2	334	0	33433	yes	343
	344	1	34434	yes	434
	433	0	43343	yes	343
	443	1	44344	yes	434
	444	2	44444	yes	4444
3	3344	4	3344334	no	
	3443	4	3443344	no	
	3444	0	3444344	yes	434
	4334	4	4334433	no	
	4433	4	4433443	no	
	4443	0	4443444	yes	434
4	33443	2	334433344	yes	333
	33444	3	334443344	no	
	34433	2	344333443	yes	333
	34443	3	344433444	no	
	43344	3	433444334	no	
	44334	3	443344433	no	
	44433	3	444334443	no	

Table 1: Proof that $|y| \geq 5$.

It follows that $|y| \geq 5$. There are now several cases to consider.

Case 1: y starts with 33. Then $ycy = 33\cdots\ c\ 33\cdots$. Since $h(w) \in \{4433, 44433\}^*$, we must have $c = 4$. Also, y must end with 4, and furthermore the letter immediately preceding the occurrence of ycy in $h(w)$ must be 4. We can therefore write $y = 33t4$ for some string t, and observe that $4\ 33t4\ 4\ 33t4 = 4y4y$ is a subword of $h(w)$. Now let $y' = 433t$, and note that $y'4y'$ is a subword of $h(w)$. But $s_5(y'4) = s_5(433t4) = s_5(33t44) = s_5(y4) \in \{0, 1, 2\}$, so $y'4y' \in P_2$, contradicting our assumption that ycy was the leftmost such occurrence in $h(w)$.

Case 2: y starts with 34. Then $ycy = 34\cdots\ c\ 34\cdots$, so $c = 3$. Thus $ycy = 34\cdots\ 3\ 34\cdots$, so y must end in 4, and further the letter immediately preceding the occurrence of ycy in $h(w)$ must be 3. We can therefore write $y = 34t3$ for some string t, and observe that $3\ 34t3\ 3\ 34t3 = 3y3y$ is a subword of $h(w)$. Now let $y' = 334t$ and note that $y'3y'$ is a subword of $h(w)$. But $s_5(y'3) = s_5(334t3) = s_5(34t33) = s_5(y3) \in \{0, 1, 2\}$, so $y'3y' \in P_2$, contradicting our assumption that ycy was the leftmost such occurrence in $h(w)$.

Case 3: y starts with 43. Then $ycy = 43\cdots\ c\ 43\cdots$, so $c = 4$, and further the letter immediately preceding the occurrence of ycy in $h(w)$ must be 4. Thus $y = 43t$. Write $t = t'b$, where $|b| = 1$. Then $y = 43t'b$. Then $4ycy = 4\ 43t'b\ 4\ 43t'b$ is a subword of $h(w)$. Let $y' = 443t'$. Then $y'by'$ is a subword of $h(w)$, and $s_5(y'b) = s_5(443t'b) = s_5(43t'b4) = s_5(y4) \in \{0, 1, 2\}$, so $y'by' \in P_2$, contradicting our assumption that ycy was the leftmost such occurrence in $h(w)$.

Case 4: y starts with 444. Then $ycy = 444\cdots\ c\ 444\cdots$, so $c = 3$, and further, y ends with 3. Since $|y| \geq 5$, we can write $y = 444t3$ for some string t. It follows that $y3y3 = 444t3\ 3\ 444t3\ 3$ is a subword of $h(w)$. Hence there exists a string u such that $h(3u) = y3$, and $3u3u$ is a subword of w. We have $s_5(u3) = s_5(3u) = s_5(y3) \in \{0, 1, 2\}$, so $u3u$ is an occurrence of a string of P_2 in w, as desired.

Case 5: y starts with 443. There are two subcases to consider:

Case 5a: $c = 3$. Then the last two characters of y must be 43. We have $ycy = 443\cdots 43\ 3\ 443\cdots 43$. Then $y3y3$ is a subword of $h(w)$, and there must exist u such that $h(4u) = y3$ and $u4u$ is a subword of w. Then $s_5(u4) = s_5(4u) = s_5(y3) \in \{0, 1, 2\}$, so $u4u$ is an occurrence of a string of P_2 in w, as desired.

Case 5b: $c = 4$. Then $ycy = 443\cdots\ 4\ 443\cdots$, so the last three characters of y must be 433. Since $|y| \geq 5$, we must have $y = 4433\cdots 433$. Write $y = 4433y'$. Then $ycy = 4433\ y'\ 44433\ y'$ is a subword of $h(w)$ and there exists u such that $h(u) = y'$. Then $h(u3u) = y'\ 44433\ y'$. Now $s_5(u3) = s_5(h(u3)) = s_5(y'44433) = s_5(4433y'4) = s_5(y4)$, so $u3u$ is an occurrence of a string of P_2 in w, as desired.

This completes the proof of Theorem 6.3.6. ■

Proof of Theorem 6.3.4. Define

$$\mathbf{v} = h^{\omega}(4) = 4433443344433444433\cdots .$$

We claim $\mathbf{v}$ avoids P_2. This follows because the word 4 avoids P_2, and by Theorem 6.3.6, if w avoids P_2 then so does $h(w)$. Now consider $S(\mathbf{v}) = 0431432032103104314\cdots$. From Lemma 6.3.5, it follows that $S(\mathbf{v})$ avoids the patterns xx, $x\sigma(x)$, and $x\sigma^2(x)$.

This completes the proof of Theorem 6.3.4, and hence assertion (c) of Theorem 6.3.1. ■

We now turn to assertion (d) of Theorem 6.3.1. From Proposition 6.3.3 with $r = 4$ we know any word of length ≥ 4 over Σ_5 contains an occurrence of one of the patterns xx, $x\sigma(x)$, $x\sigma^2(x)$, or $x\sigma^3(x)$. The methods of Theorem 6.3.6 and Lemma 6.3.5 lead immediately to

Theorem 6.3.7 *It is possible to simultaneously avoid the patterns xx, $x\sigma(x)$, $x\sigma^2(x)$, and $x\sigma^3(x)$ over Σ_6.*

Proof. We construct an infinite word $\mathbf{w}$ over Σ_6 such that $\mathbf{w}$ simultaneously avoids the patterns xx, $x\sigma(x)$, $x\sigma^2(x)$, and $x\sigma^3(x)$. Let g be the morphism over $\{4, 5\}$ defined by

$g(5) = 55544$ and $g(4) = 555544$. We claim that $\mathbf{w} = S(g^{\omega}(5))$ simultaneously avoids the patterns xx, $x\sigma(x)$, $x\sigma^2(x)$, and $x\sigma^3(x)$. The proof follows exactly the same plan as that of Theorem 6.3.6. We omit it here. ∎

Remark. We note that the morphisms used in the proof of Theorem 6.3.6 and Theorem 6.3.7 do not generalize to any other j. For example, if we were to define h analogously for $j = 4$, we would have $h(6) = 666655$ and $h(5) = 6666655$. By inspection, we see that $h(6) = 6666655$ contains ycy where $y = 66$ and $c = 6$. Hence $s_7(yc) = 4 = j$ and so $S(h(6))$ does not avoid the pattern $x\sigma^4(x)$.

Finally, we turn to assertion (e). First, we show it is not possible to avoid the patterns xx, $x\sigma(x)$, ..., $x\sigma^4(x)$ on 7 letters. Here the corresponding tree has 215 leaves, and the longest leaf has length 36. See Figure 2 below.

0531	0543205432	0653106542065432	06542065431654320543210	0654316543205432164321053210065
0542	0543216431	0654206543165431	06543165432054321643216	0532106421065310654206543165431
0643	0642106420	0654316543205431	053210642106531065420653	0543216432105321064210653106543
05320	0642106421	05321064210653105	054321643210532106421064	0642106531065420654316543205431
06420	0642106532	05432164321053216	064210653106542065431653	0653106542065431654320543216431
06532	0653106531	06421065310654205	065310654206543165432053	0654206543165432054321643210531
054310	0653106532	06531065420654310	065420654316543205432165	0654316543205432164321053210643
064216	0653106543	06542065431654321	065431654320543216432106	05321064210653106542065431654321
065316	0654206542	06543165432054320	0532106421065310654206542	05432164321053210642106531065421
065421	0654316542	053210642106531064	0543216432105321064210654	06421065310654206543165432054320
065432	05321064216	054321643210532105	0642106531065420654316542	06531065420654316543205432164320
0532165	05432164320	064210653106542064	0653106542065431654320542	06542065431654320543216432105320
0543164	06421065316	065310654206543164	0654206543165432054321642	06543165432054321643210532106420
0543206	06531065421	065420654316543206	0654316543205432164321054	053210642106531065420654316543206
0543210	06542065432	065431654320543210	05321064210653106542065432	054321643210532106421065310654205
0653105	06543165431	0532106421065310653	05432164321053210642106532	064210653106542065431654320543210
0654205	053210642105	0543216432105321065	06421065310654206543165431	065310654206543165432054321643216
0654310	054321643216	0642106531065420653	06531065420654316543205431	065420654316543205432164321053216
05321642	064210653105	0653106542065431653	06542065431654320543216431	065431654320543216432105321064216
05321643	065310654205	0654206543165432053	06543165432054321643210531	0532106421065310654206543165432053
05321054	065420654310	0654316543205432165	053210642106531065420654310	0543216432105321064210653106542064
05321065	065431654321	05321064210653106543	054321643210532106421065316	0642106531065420654316543205432165
05431653	0532106421064	05432164321053210643	064210653106542065431654321	0653106542065431654320543216432106
05431654	0543216432106	06421065310654206542	065310654206543165432054320	0654206543165432054321643210532105
05432053	0642106531064	06531065420654316542	065420654316543205432164320	0654316543205432164321053210642105
05432165	0653106542064	06542065431654320542	065431654320543216432105320	05321064210653106542065431654320542
06421053	0654206543164	06543165432054321642	0532106421065310654206543164	05432164321053210642106531065420653
06421054	0654316543206	053210642106531065421	0543216432105321064210653105	06421065310654206543165432054321642
06531064	05321064210654	054321643210532106420	0642106531065420654316543206	06531065420654316543205432164321054
06542064	05432164321054	064210653106542065432	0653106542065431654320543210	06542065431654320543216432105321065
06543164	06421065310653	065310654206543165431	0654206543165432054321643216	06543165432054321643210532106421064
053210531	06531065420653	065420654316543205431	0654316543205432164321053216	053210642106531065420654316543205431
053210643	06542065431653	065431654320543216431	05321064210653106542065431653	053210642106531065420654316543205432
054320542	06543165432053	0532106421065310654205	05432164321053210642106531064	054321643210532106421065310654206542
054321642	053210642106532	0543216432105321064216	06421065310654206543165432053	054321643210532106421065310654206543
064210643	054321643210531	0642106531065420654310	06531065420654316543205432165	064210653106542065431654320543216431
064210654	064210653106543	0653106542065431654321	06542065431654320543216432106	064210653106542065431654320543216432
065420653	065310654206542	0654206543165432054320	06543165432054321643210532105	065310654206543165432054321643210531
065431653	065420654316542	0654316543205432164320	053210642106531065420654316542	065310654206543165432054321643210532
0532105320	065431654320542	05321064210653106542064	054321643210532106421065310653	065420654316543205432164321053210642
0532105321	0532106421065316	05432164321053210642105	064210653106542065431654320542	065420654316543205432164321053210643
0532106420	0543216432105320	06421065310654206543164	065310654206543165432054321642	065431654320543216432105321064210653
0543205431	0642106531065421	06531065420654316543206	065420654316543205432164321054	065431654320543216432105321064210654

Figure 2: Leaves of the tree giving the proof of assertion (e)

Using the tree traversal algorithm, we can prove

Theorem 6.3.8 *One cannot avoid the patterns xx, $x\sigma(x)$, …, $x\sigma^j(x)$ on $j+3$ letters, for*

$j \geq 5$.

Proof. Consider trying to generate an infinite word $\mathbf{w}$ over $\mathbb{Z}$ starting with 0, subject to two conditions: (1) avoiding the pattern $x\sigma^i(x)$ for all i, where $|x| \geq 2$, and (2) avoiding all subwords of length 2 that are *not* of the form $(n, n-1)$ or $(n, n-2)$ for $n \in \mathbb{Z}$.

Let us now apply the tree traversal algorithm to this avoidance problem. The tree T so produced has 71 leaves and the longest leaf has length 12. All the occurrences of $x\sigma^i(x)$ found at the leaves of T, for $|x| \geq 2$, satisfy $i \in X = \{-3, -4, -6, -7, -8\}$.

Now consider the labels of this tree reduced modulo $j+3$. The patterns at the leaves are still of the form $x\sigma^i(x)$, except now i is reduced modulo $j+3$. In order for T to correctly represent a proof that the pattern $x\sigma^i(x)$ cannot be avoided for $0 \leq i \leq j$ we must check that $i \bmod (j+3) \in \{0, 1, \ldots, j\}$ for all $i \in X$. But this is clearly true for $j \geq 5$.

Figure 3 lists the leaves of T in coded form. We use the letters A, B, C, D, E, F, G to represent $10, 11, 12, 13, 14, 15, 16$ respectively, and the word $a_1 a_2 \cdots a_j$ represents the leaf $(-a_1, -a_2, \ldots, -a_j)$. ■

0246	024568AC	01356789A	0234568ABCE
0235	024568AB	01235789B	01356789BDF
0134	0245679A	01234689B	01246789ACD
02457	02456789	0245679BCE	01235789ABC
01357	0234689B	0245679BCD	01234689ABD
01245	0234689A	0245678ACE	0245678ACDEG
023467	0234579B	0234579ABD	0245678ACDEF
013568	02345689	0234579ABC	0234568ABCDF
012468	0135679B	0234568ABD	0234568ABCDE
012356	0135679A	0135678ACE	01356789BDEG
012345	0124678A	0135678ACD	01356789BDEF
0245689	0123578A	01356789BC	01246789ACEG

023468A	0123468A	01246789BD	01246789ACEF
0234578	0245679BD	01246789BC	01235789ABDF
0234567	0245678AB	01246789AB	01235789ABDE
0124679	0234579AC	01235789AC	01234689ABCE
0123579	0234568AC	01234689AC	01234689ABCD
0123467	0135678AB	0245678ACDF	

Figure 3: Leaves of the tree giving the proof of Theorem 6.3.8.

We now show it is possible to simultaneously avoid the patterns xx, $x\sigma(x)$, …, $x\sigma^j(x)$ on Σ_{j+4} for $j \geq 4$. Actually, we prove a more general result from which this result will follow.

Theorem 6.3.9 *Let $k \geq 4$ be an integer, and let $A \subset \Sigma_k$ such that* Card $A \leq k-3$. *Then it is possible to simultaneously avoid the patterns $\{x\sigma^a(x) \ : \ a \in A\}$ over Σ_k.*

Proof. Once again, the idea is to consider the first differences of words, modulo k. Suppose we can construct a word $\mathbf{w}$ over Σ_k such that $\mathbf{w}$ avoids both (i) the pattern ycy, where $|y| \geq 1$ and $|c| = 1$, and (ii) the letters $a \in A$. Then it follows from Lemma 6.3.5 that $S(\mathbf{w})$ avoids the pattern $x\sigma^a(x)$.

Lemma 6.3.10 *Let $\mathbf{w} = a_1a_2a_3\cdots$ be any squarefree word over Σ_3. Then the word $a_1a_1a_2a_2a_3a_3\cdots$ avoids the pattern ycy for $y \in \Sigma_3^+$ and $c \in \Sigma_3$.*

Remark. This lemma is due to J. Loftus.

Proof. Suppose $y = b_1b_2\cdots b_k$, and the pattern ycy occurs in $\mathbf{z} = a_1a_1a_2a_2a_3a_3\cdots$. There are three cases to consider, depending on $|y|$ and where y starts in $\mathbf{z}$.

Case 1: $|y|$ is even and y starts with $a_i a_i$. Let $k = 2j$. Then we have

$$\begin{array}{ccccccccc} b_1 & b_2 & \cdots & b_{2j} & c & b_1 & b_2 & \cdots & b_{2j} \\ & & & & = & & & & \\ a_i & a_i & \cdots & a_{i+j-1} & a_{i+j} & a_{i+j} & a_{i+j+1} & \cdots & a_{i+2j} \end{array}$$

and so $a_{i+j} = b_1 = a_i = b_2 = a_{i+j+1}$. It follows that $\mathbf{w}$ contains the square $a_{i+j}a_{i+j+1}$, a contradiction.

Case 2: $|y|$ is even and y starts with $a_i a_{i+1}$. Let $k = 2j$. Then we have

$$\begin{array}{ccccccccc} b_1 & b_2 & \cdots & b_{2j} & c & b_1 & b_2 & \cdots & b_{2j} \\ & & & & = & & & & \\ a_i & a_{i+1} & \cdots & a_{i+j} & a_{i+j} & a_{i+j+1} & a_{i+j+1} & \cdots & a_{i+2j} \end{array}$$

and so $a_i = b_1 = a_{i+j+1} = b_2 = a_{i+1}$. It follows that $\mathbf{w}$ contains the square $a_i a_{i+1}$, a contradiction.

Case 3: $|y|$ is odd. Let $k = 2j + 1$. Then either

$$\begin{array}{ccccccccccc} b_1 & b_2 & \cdots & b_{2j} & b_{2j+1} & c & b_1 & b_2 & \cdots & b_{2j} & b_{2j+1} \\ & & & & & = & & & & & \\ a_i & a_i & \cdots & a_{i+j-1} & a_{i+j} & a_{i+j} & a_{i+j+1} & a_{i+j+1} & \cdots & a_{i+2j} & a_{i+2j+1} \end{array}$$

or

$$\begin{array}{ccccccccccc} b_1 & b_2 & \cdots & b_{2j} & b_{2j+1} & c & b_1 & b_2 & \cdots & b_{2j} & b_{2j+1} \\ & & & & & = & & & & & \\ a_i & a_{i+1} & \cdots & a_{i+j} & a_{i+j} & a_{i+j+1} & a_{i+j+1} & a_{i+j+2} & \cdots & a_{i+2j+1} & a_{i+2j+1} \end{array}$$

In either case we find

$$\begin{aligned} a_i = &\quad b_1 &= a_{i+j+1} \\ a_{i+1} = &\quad b_3 &= a_{i+j+2} \\ &\quad \vdots & \\ a_{i+j} = &\quad b_{2j+1} &= a_{i+2j+1} \end{aligned}$$

It follows that $a_i a_{i+1} \cdots a_{i+j} = a_{i+j+1} a_{i+j+2} \cdots a_{i+2j+1}$ and so $\mathbf{w}$ contains the square $a_i a_{i+1} \cdots a_{i+2j+1}$, a contradiction. The proof of the Lemma is complete. ■

Remark. One cannot avoid the pattern ycy, with $|y| \geq 1$ and $|c| = 1$, over an alphabet of 2 letters. As the tree traversal algorithm shows, any word of length ≥ 7 over $\{0, 1\}$ contains an occurrence of ycy.

Now we can complete the proof of Theorem 6.3.9. Let $\mathbf{x}$ be any squarefree word over $\{0, 1, 2\}$. Since Card $A = k - 3$, we have $\Sigma_k - A = \{d, e, f\}$ for some distinct integers $0 \leq d, e, f < k$.

Consider the morphism $\varphi : \Sigma_{j+4}^* \to \Sigma_{j+4}^*$ defined as follows:

$$\begin{aligned} 0 &\to d\ d \\ 1 &\to e\ e \\ 2 &\to f\ f \end{aligned}$$

We claim $S(\varphi(\mathbf{x}))$ avoids the patterns xx, $x\sigma(x)$, ..., $x\sigma^j(x)$.

Let $\mathbf{v} = S(\varphi(\mathbf{x}))$. Then $\Delta(\mathbf{v}) = \varphi(\mathbf{x})$ clearly avoids ycy by Lemma 6.3.10, and it also avoids all the letters in A by construction. Then by Lemma 6.3.5, $\mathbf{v}$ avoids the patterns $x\sigma^a(x)$ for $a \in A$. ■

As a consequence we get

Corollary 6.3.11 *It is possible to simultaneously avoid the patterns xx, $x\sigma(x)$, …, $x\sigma^j(x)$ on Σ_{j+4} for $j \geq 4$.*

The proof of Theorem 6.3.1 is now complete. ■

6.4 Even more results

One may also consider the problem of avoiding other sets of patterns of the form $x\sigma^a(x)$. In this section, we let $j \geq 1$ be an integer, and consider avoiding the $2j+1$ patterns $x\sigma^{-j}(x)$, …, $x\sigma^{-1}(x)$, xx, $x\sigma(x)$, …, $x\sigma^j(x)$ simultaneously over the alphabet Σ_k.

Theorem 6.4.1 *For $j \geq 1$, one can simultaneously avoid the patterns $x\sigma^{-j}(x)$, …, $x\sigma^{-1}(x)$, xx, $x\sigma(x)$, …, $x\sigma^j(x)$ over Σ_{2j+4}, and this is best possible.*

Proof.

By Theorem 6.3.9 with $A = \{-j, 1-j, \ldots, -1, 0, 1, 2, \ldots j\}$, we see that we can simultaneously avoid the patterns $x\sigma^{-j}(x)$, …, $x\sigma^{-1}(x)$, xx, $x\sigma(x)$, …, $x\sigma^j(x)$ over Σ_{2j+4}.

It follows from Proposition 6.3.2 that one cannot avoid $x\sigma^{-j}(x)$, …, $x\sigma^{-1}(x)$, xx, $x\sigma(x)$, …, $x\sigma^j(x)$ over Σ_{2j+2} or smaller alphabet.

To prove that one cannot simultaneously avoid the patterns $x\sigma^{-j}(x)$, …, $x\sigma^{-1}(x)$, xx, $x\sigma(x)$, …, $x\sigma^j(x)$ over Σ_{2j+3}, we use the tree traversal algorithm. Then every word of length ≥ 8 over Σ_{2j+3} contains an occurrence of $x\sigma^l(x)$ for some l with $-j \leq l \leq j$. Figure 4 below gives the output of the tree traversal algorithm, showing that there are 24 leaves. Here $t = j+1$. ■

```
(0, -t, -2t, -3t)        (0, -t, -2t, -t, 0, t)        (0, -t, -2t, -t, 0, -t, -2t, -3t)
(0, -t, 0, -t)           (0, -t, 0, t, 0, t)           (0, -t, -2t, -t, 0, -t, -2t, -t)
(0, t, 0, t)             (0, t, 0, -t, 0, -t)          (0, -t, 0, t, 0, -t, 0, -t)
```

`(0, t, 2t, 3t)`	`(0, t, 2t, t, 0, -t)`	`(0, -t, 0, t, 0, -t, 0, t)`
`(0, -t, -2t, -t, -2t)`	`(0, -t, -2t, -t, 0, -t, 0)`	`(0, t, 0, -t, 0, t, 0, -t)`
`(0, -t, 0, t, 2t)`	`(0, -t, 0, t, 0, -t, -2t)`	`(0, t, 0, -t, 0, t, 0, t)`
`(0, t, 0, -t, -2t)`	`(0, t, 0, -t, 0, t, 2t)`	`(0, t, 2t, t, 0, t, 2t, t)`
`(0, t, 2t, t, 2t)`	`(0, t, 2t, t, 0, t, 0)`	`(0, t, 2t, t, 0, t, 2t, 3t)`

Figure 4: Leaves of the tree giving a proof of Theorem 6.4.1.

6.5 Avoiding $x\sigma^i(x)$ for all i

Generalizing the results of the previous section, we may ask if it is possible to avoid the patterns $x\sigma^i(x)$ *for all* i. Unfortunately, this is clearly impossible, for if a word z begins with $i\,j$, then it contains a subword of the form $i\,\sigma^{j-i}(i)$.

However, we can relax our conditions for avoidance, as follows: we say an infinite word *weakly avoids* the patterns $x\sigma^i(x)$ if it contains no subwords of the form $x\sigma^i(x)$ with $|x| \geq 2$. (In contrast, our previous notion of avoidability we will call *strong*.)

Proposition 6.5.1 *Over Σ_2, every word of length ≥ 8 contains a subword of the form $x\sigma^i(x)$ for some $i \geq 0$, with $|x| \geq 2$.*

Proof. Our simple tree traversal algorithm proves this. The tree generated has 24 leaves, and the leaves are given in Figure 5. ■

0000	000101	00010000
0011	001001	00010001
0101	010000	00100010
0110	011100	00100011
00011	0001001	01000100

00101	0010000	01000101
01001	0100011	01110110
01111	0111010	01110111

Figure 5: Leaves of the tree giving a proof of Proposition 6.5.1.

However, it is possible to weakly avoid the patterns $x\sigma^i(x)$ for all $i \geq 0$ over Σ_3. Let $\mathbf{w}$ be any squarefree word over $\{0, 1, 2\}$, and consider the morphism f which maps

$$\begin{aligned} 0 &\rightarrow 00 \\ 1 &\rightarrow 10 \\ 2 &\rightarrow 20. \end{aligned}$$

Theorem 6.5.2 *The infinite word $f(\mathbf{w})$ weakly avoids the patterns $x\sigma^i(x)$ for all $i \geq 0$.*

Proof. Let $\mathbf{w} = c_1c_2c_3\cdots$, and $f(\mathbf{w})$ contains a subword of the form $z = x\sigma^i(x)$ for some i and $|x| \geq 2$. There are two cases, depending on $|x|$ mod 2.

Case 1: $|x| \equiv 0 \pmod 2$. In this case, there are two possibilities, depending where x starts in $f(\mathbf{w})$:

$$\begin{aligned} z &= \overbrace{d_1 0 d_2 0 \cdots d_j 0}^{x} \mid \overbrace{d_{j+1} 0 \cdots d_{2j} 0}^{\sigma^i(x)} \\ z &= 0 d_1 0 d_2 \cdots 0 d_j \mid 0 d_{j+1} \cdots 0 d_{2j} \end{aligned}$$

where $d_t = c_{t+k}$ for some integer $k \geq 0$. Comparing the second symbol in the first case, or the first symbol in the second case, we see that if $z = x\sigma^i(x)$, then $i = 0$. Hence $d_t = d_{j+t}$ for $1 \leq t \leq j$, and so $c_{k+t} = c_{k+j+t}$ for $1 \leq t \leq j$, contradicting the assumption that $\mathbf{w}$ was squarefree.

Case 2: $|x| \equiv 1 \pmod 2$.

$$
\begin{aligned}
z &= \overbrace{d_1 0 d_2 \cdots}^{x} \mid \overbrace{0 d_j 0 \cdots}^{\sigma^i(x)} \\
z &= 0 d_1 0 \cdots \mid d_j 0 d_{j+1} \cdots
\end{aligned}
$$

If $z = x\sigma^i(x)$, then, in the first case, we must have $d_1 = d_2$, and in the second $d_j = d_{j+1}$. Both correspond to a square in $\mathbf{w}$, a contradiction. ■

We might also try weakly avoiding $x\sigma^i(x)$ for $0 < i < k$ over Σ_k, while simultaneously (strongly) avoiding xx.

Theorem 6.5.3 *If $k = 4$, one can, over Σ_k, simultaneously weakly avoid $x\sigma^i(x)$ for $0 < i < k$ and strongly avoid xx. Here k is best possible.*

Proof. We can weakly avoid $x\sigma^i(x)$ for $0 < i < k$ and strongly avoid xx over Σ_4 as follows: let $\mathbf{w}$ be any squarefree word over $\{1, 2, 3\}$, and consider the morphism f which maps

$$
\begin{aligned}
1 &\rightarrow 10 \\
2 &\rightarrow 20 \\
3 &\rightarrow 30.
\end{aligned}
$$

Then it follows from the same method of the proof of Theorem 6.5.2 that $f(\mathbf{w})$ weakly avoids $x\sigma^i(x)$ for all i. However, it is clear from the construction that $f(\mathbf{w})$ has no subword of the form cc for $c \in \Sigma_4$, so $f(\mathbf{w})$ also strongly avoids xx.

On the other hand, the tree traversal algorithm shows that over Σ_3, any word of length ≥ 8 has a (weak) occurrence of $x\sigma^i(x)$ with $0 < i < 3$, or a strong occurrence of xx. The tree generated has 24 leaves, and the leaves are given in Figure 6. ■

0101	010202	01020101

0120	012102	01020102
0202	020101	01210120
0210	021201	01210121
01021	0102012	02010201
01212	0121010	02010202
02012	0201021	02120210
02121	0212020	02120212

Figure 6: Leaves of the tree giving a proof of Theorem 6.5.3.

Chapter 7

Weakly Self-Avoiding Words and a Construction of Friedman

This chapter is based on the work of J. Shallit and M.-w. Wang [SW01b].

7.1 Introduction

We say a word y is a *subsequence* of a word z if y can be obtained by striking out 0 or more symbols from z. For example, "iron" is a subsequence of "introduction". We say a word y is a *subword* of a word z if there exist words w, x such that $z = wyx$. For example, "duct" is a subword of "introduction".

We use the notation $x[k]$ to denote the k'th letter chosen from the string x. (The first letter of a string is $x[1]$.) We write $x[a..b]$ to denote the subword of x of length $b - a + 1$ starting at position a and ending at position b.

Recently H. Friedman has found a remarkable construction that generates extremely large numbers [Fri01, Fri00]. Namely, consider words over a finite alphabet Σ of cardinality k. If an infinite word $\mathbf{x}$ has the property that for all i, j with $0 < i < j$ the subword $\mathbf{x}[i..2i]$ is not a subsequence of $\mathbf{x}[j..2j]$, we call it *self-avoiding*. We apply the same definition for a finite word x of length n, imposing the additional restriction that $j \le n/2$.

Friedman shows there are no infinite self-avoiding words over a finite alphabet. Furthermore, he shows that for each k there exists a longest finite self-avoiding word x over an alphabet of size k. Call $n(k)$ the length of such a word. Then clearly $n(1) = 3$ and a simple argument shows that $n(2) = 11$. Friedman shows that $n(3)$ is greater than the incomprehensibly large number $A_{7198}(158386)$, where A is the Ackermann function.

Jean-Paul Allouche asked what happens when "subsequence" is replaced by "subword". A priori we do not expect results as strange as Friedman's, since there are no infinite anti-chains for the partial order defined by "x is a subsequence of y", while there *are* infinite anti-chains for the partial order defined by "x is a subword of y".

7.2 Main Results

If an infinite word $\mathbf{x}$ has the property that for all i, j with $0 \le i < j$ the subword $\mathbf{x}[i..2i]$ is not a subword of $\mathbf{x}[j..2j]$, we call it *weakly self-avoiding*. If x is a finite word of length n, we apply the same definition with the additional restriction that $j \le n/2$.

Theorem 7.2.1 *Let* $\Sigma = \{0, 1, \ldots, k-1\}$.

(a) If $k = 1$*, the longest weakly self-avoiding word is of length* 3*, namely* 000.

(b) If $k = 2$*, there are no weakly self-avoiding words of length* > 13*. There are* 8 *longest weakly self-avoiding words, namely* 0010111111010*,* 0010111111011*,* 0011110101010*,* 0011110101011 *and the four words obtained by changing* 0 *to* 1 *and* 1 *to* 0.

(c) If $k \ge 3$*, there exists an infinite weakly self-avoiding word.*

Proof.

(a) If a word x over $\Sigma = \{0\}$ is of length ≥ 4, then it must contain 0000 as a prefix. Then $x[1..2] = 00$ is a subword of $x[2..4] = 000$.

(b) To prove this result, we create a tree whose root is labeled with ϵ, the empty word. If a node's label x is weakly self-avoiding, then it has two children labeled $x0$ and $x1$. This tree is finite if and only if there is a longest weakly self-avoiding word. In this case, the leaves of the tree represent non-weakly-self-avoiding words that are minimal in the sense that any proper prefix is weakly self-avoiding.

Now we use the breadth-first tree traversal technique of Section 6.3 in the previous chapter.

If this algorithm terminates, we have proved that there is a longest weakly self-avoiding word. The proof may be concisely represented by listing the leaves in breadth-first order. We may shorten the tree by assuming, without loss of generality, that the root is labeled 0.

When we perform this procedure, we obtain a tree with 92 leaves, whose longest label is of length 14. The following list describes this tree:

0000	00111100	0011010101	001011111011
0001	00111110	0011010110	001011111100
0101	00111111	0011010111	001011111110
001000	01000000	0011101000	001011111111
001001	01000001	0011101001	001110101000
001010	01000010	0011101011	001110101001
001100	01000011	0011110100	001110101010
010001	01100001	0011110110	001110101011
010010	01100010	0011110111	001111010100
010011	01100011	0110000000	001111010110
011001	01110001	0110000001	001111010111
011010	01110010	0110000010	011100000000
011011	01110011	0110000011	011100000001
011101	0010110100	0111000001	011100000010
011110	0010110101	0111000010	011100000011

011111	0010110110	0111000011	00101111110100
00101100	0010110111	001011110100	00101111110101
00110100	0010111000	001011110101	00101111110110
00110110	0010111001	001011110110	00101111110111
00110111	0010111010	001011110111	00111101010100
00111000	0010111011	001011111000	00111101010101
00111001	0010111100	001011111001	00111101010110
00111011	0011010100	001011111010	00111101010111

Figure 1: Leaves of the tree giving a proof of Theorem 7.2.1 (b)

(c) Consider the word

$$\begin{aligned} \mathbf{x} &= 22010110111011111011111110111111111110\cdots \\ &= 2\,2\,0\,1\,0\,1^2\,0\,1^3\,0\,1^5\,0\,1^7\,0\,1^{11}\,0\,1^{15}\,0\,1^{23}\,0\,1^{31}\,0\,1^{47}\,0\,\cdots \end{aligned}$$

where there are 0's in positions $3, 5, 8, 12, 18, 26, 38, 54, 78, 110, 158, \ldots$. More precisely, define $f_{2n+1} = 5 \cdot 2^n - 2$ for $n \geq 0$, and $f_{2n} = 7 \cdot 2^{n-1} - 2$ for $n \geq 1$. Then $\mathbf{x}$ has 0's only in the positions given by f_i for $i \geq 1$.

First we claim that if $i \geq 3$, then any subword of the form $\mathbf{x}[i..2i]$ contains exactly two 0's. This is easily verified for $i = 3$. If $5 \cdot 2^n - 1 \leq i < 7 \cdot 2^n - 1$ and $n \geq 0$, then there are 0's at positions $7 \cdot 2^n - 2$ and $5 \cdot 2^{n+1} - 2$. (The next 0 is at position $7 \cdot 2^{n+1} - 2$, which is $> 2(7 \cdot 2^n - 2)$.) On the other hand, if $7 \cdot 2^{n-1} - 1 \leq i < 5 \cdot 2^n - 1$ for $n \geq 1$, then there are 0's at positions $5 \cdot 2^n - 2$ and $7 \cdot 2^n - 2$. (The next 0 is at position $5 \cdot 2^{n+1} - 2$, which is $> 2 \cdot (5 \cdot 2^n - 2)$.)

Now we prove that $\mathbf{x}$ is weakly self-avoiding. Clearly $\mathbf{x}[1..2] = 22$ is not a subword of any subword of the form $\mathbf{x}[j..2j]$ for any $j \geq 2$. Similarly, $\mathbf{x}[2..4] = 201$ is not a subword of any subword of the form $\mathbf{x}[j..2j]$ for any $j \geq 3$. Now consider subwords of the form

$t := \mathbf{x}[i..2i]$ and $t' := \mathbf{x}[j..2j]$ for $i, j \geq 3$ and $i < j$. From above we know $t = 1^u01^v01^w$, and $t' = 1^{u'}01^{v'}01^{w'}$. For t to be a subword of t' we must have $u \leq u'$, $v = v'$, and $w \leq w'$. But since the blocks of 1's in $\mathbf{x}$ are distinct in size, this means that the middle block of 1's in t and t' must occur in the same positions of $\mathbf{x}$. Then $u \leq u'$ implies $i \geq j$, a contradiction. ■

7.3 Another construction

Friedman has also considered variations on his construction, such as the following: let $M_2(n)$ denote the length of the longest finite word $\mathbf{x}$ over $\{0, 1\}$ such that $\mathbf{x}[i..2i]$ is not a subsequence of $\mathbf{x}[j..2j]$ for $n \leq i < j$. We can again consider this where "subsequence" is replaced by "subword".

Theorem 7.3.1 *There exists an infinite word $\mathbf{x}$ over $\{0, 1\}$ such that $\mathbf{x}[i..2i]$ is not a subword of $\mathbf{x}[j..2j]$ for all i, j with $2 \leq i < j$.*

Proof. Let

$$\begin{aligned} \mathbf{x} &= 0\,0\,1\,0\,0\,1^3\,0\,1^2\,0\,1^7\,0\,1^5\,0\,1^{15}\,0\,1^{11}\,0\,1^{31}\,0\,1^{23}\,\cdots \\ &= 0\,0\,1\,0\,0\,1^{g_1}\,0\,1^{g_2}\,0\,1^{g_3}\,0\,\cdots \end{aligned}$$

where $g_1 = 3$, $g_2 = 2$, and $g_n = 2g_{n-2} + 1$ for $n \geq 3$. Then a proof similar to that above shows that every subword of the form $\mathbf{x}[i..2i]$ contains exactly two 0's, and hence, since the g_i are all distinct, we have $\mathbf{x}[i..2i]$ is not a subword of $\mathbf{x}[j..2j]$ for $j > i > 1$. ■

Chapter 8

On the irregularity of the duplication closure

This chapter is based on [Wan00]. After the defence of my thesis I noticed that essentially the same argument was given by Ehrenfeucht and Rozenberg [ER84].

In this chapter we solve a problem proposed in "On the regularity of the duplication closure" by J. Dassow, V. Mitrana and Gh. Păun [DMP99].

First we need some notation. Let $w \in \Sigma^*$ be a word. Define $D(w)$ to be the set of words u for which there exists x, y, $z \in \Sigma^*$ with $w = xyz$ and $u = xyyz$. We extend the definition of D to set of words in the natural way. We also define $D^2 = D(D)$. So powers of D correspond to compositions. Finally let $D^* = \bigcup_{i=1}^{\infty} D^i$.

We prove the following.

Theorem 8.0.2 *Suppose w is a word containing at least 3 distinct letters. Then $D^*(w)$ is not regular.*

We assume $w = abc$ below. The general case follows easily from this. Let Σ be an alphabet.

Lemma 8.0.3 *Suppose $u = abc\,u'$ where $u' \in \Sigma^*$, then there exists $v \in \Sigma^*$ such that $uv \in D^*(w)$.*

Proof: We show how to construct $z = uv$ iteratively. Initially we set $z = abc$. Suppose $u = a_1 a_2 ... a_k$ and $z = b_1 b_2 ... b_l$. Initially we have that $a_i = b_i$ for $1 \leq i \leq 3$. Then for each $4 \leq i \leq k$, we do the following: we find the largest index $j < i$ such that $b_j = a_i$. Then we perform a subword copy on z using indices $j...i-1$. This means that now $z = b_1 b_2 ... b_j ... b_{i-1} b_j ... b_{i-1} b_i ... b_l$. The effect of this is to make the prefixes of u and z agree on all indices up to and including i. For example, suppose $u = abcbacca$. We construct z iteratively as follows, where the underlined portion displays the subword to be repeated:

$a\underline{bc} \to \underline{abcb}c \to ab\underline{cba}bcbc \to abcba\underline{c}babcbc \to abcb\underline{acc}babcbc \to abcbaccaccbabcbc.$ ■

Now observe that each word $x \in D^*(w)$ is obtained from the word w by a sequence of subword doubling operations. Let $t(x)$ be the minimal number of doubling operations to obtain x from w. We have

Lemma 8.0.4

$$t(x) \geq \log_2(|x|/3).$$

Proof: Each doubling operation at most doubles the length of the previous word and the starting word $w = abc$ is of length 3. The lower bound follows. ■

Lemma 8.0.5 *Suppose $u = abc\,u' \in \Sigma^*$ is square free. Let v be a shortest word such that $uv \in D^*(w)$. Then*

$$|v| \geq \log_2(|u|/3). \tag{8.1}$$

Proof: By the definition of t, uv is obtained from w by a sequence of at least $t(uv)$ subword doubling operations. Since u is square free, each of these doubling operations must result in at least one additional letter outside u, i.e., in v. It follows that

$$|v| \geq t(uv) \geq \log_2(|uv|/3) \geq \log_2(|u|/3).$$

■

Now we are ready to prove the theorem using Myhill-Nerode's characterization of regular languages. We construct an infinite sequence of pairwise inequivalent words as follows.

We start by defining $W_1 = abc$. For $i \geq 1$, we define W_{i+1} inductively as follows: let V_i be such that $W_i V_i \in D^*(w)$. Then we choose W_{i+1} to be a square free word, starting with abc, such that $\log_2(|W_{i+1}|/3) > |V_i|$. Such a word exists because there are infinitely many square free words over a 3 letter alphabet. This length condition ensures (by Lemma 8.0.5) that $W_{i+1} V_j \notin D^*(w)$ for all $j \leq i$. It follows that the W_i are pairwise inequivalent. Since there are infinitely many W_i, by Myhill-Nerode $D^*(w)$ is not regular.

8.1 Open Problem

The most obvious question is still open. Is $D^*(w)$ context-free?

Chapter 9

Automatic Complexity of Strings

This chapter is based on the work of J. Shallit and M.-w. Wang [SW01a].

9.1 Introduction

We are interested in a computable measure of complexity for finite strings x over a finite alphabet, typically $\{0, 1\}$. Any such measure should reflect, in some sense, how "complicated" the string x is.

Of course, any such discussion must start with Kolmogorov-Chaitin complexity [LV97] $C(x)$, which (roughly speaking) measures the complexity of a string x as the size of the shortest pair

$$(T, y) = (\text{Turing machine description, input})$$

such that T on input y outputs x. Not only does $C(x)$ measure the complexity of x, but also the pair (T, y) can be viewed as the optimal way to compress the string x.

However it has three major deficiencies (the first two are equivalent):

1. It is uncomputable! It is known that "$C(x) < n$" is computably enumerable, but "$C(x) \geq n$" is not computably enumerable.

2. There is no effective procedure for finding a compression pair (T, y).

3. K depends somewhat on the particular model of universal Turing machine chosen, and is defined in a machine-independent way only up to an additive constant.

One consequence of deficiency (3) above is that since $C(xx) = C(x) + O(1)$, with the constant depending on the particular model of universal Turing machine chosen, it doesn't make sense to ask if $C(xx) > C(x)$ for any, most, or all strings x. We will see below, however, that in the measure of complexity proposed in this chapter, we have $A(xx) \geq A(x)$ for all strings x, and in fact there are infinitely many strings x for which this inequality is strict; see Theorems 9.2.3 and 9.5.4.

It would be nice to find a measure without these deficiencies. Turing machines are extremely powerful, and this suggests that we could replace the Turing machine with a less powerful model and hope to find a computable measure.

For example, we could consider replacing the Turing machine with a context-free grammar (CFG). We choose, perhaps arbitrarily, some measure of the complexity of a context-free grammar, and then ask for the smallest grammar G such that $L(G) = \{x\}$.

If we demand that the context-free grammar be in Chomsky normal form (i.e., all productions are of the form $A \rightarrow BC$ or $A \rightarrow a$ where A, B, C are variables and a is a terminal), and use the number of variables as the measure of a grammar's size, then then we get a well-known measure of complexity associated with "word chains". Diwan [Diw86] was apparently the first to study this measure; for other papers see [BB88, Rot89, AB89, Alt90, Bou92].

In this chapter we consider replacing the Turing machine with a deterministic finite automaton, or DFA.

Given a string x, in analogy with the word chain problem mentioned above, we might seek to find a smallest DFA M such that $L(M) = \{x\}$. But this is clearly uninteresting, since if $|x| = n$, a smallest such DFA always has exactly $n + 2$ states. Hence we consider relaxing the requirement somewhat.

If a DFA M has the property that it accepts a string x, but no other strings of length $|x|$, we say *M accepts x uniquely*. In this chapter, we examine the consequences of the following definition. We define $A(x)$, the *automatic complexity* of x, to be the smallest number of states in any DFA M that accepts x uniquely. Of course, there may be many such DFA's with the smallest number of states. We do not care how M behaves on strings that are shorter or longer than x. More formally,

Definition 9.1.1 Let $\Sigma = \{0, 1\}$ and $x \in \Sigma^*$ with $|x| = n$. Define $A(x)$ to be the smallest number of states in any DFA M such that $L(M) \cap \Sigma^n = \{x\}$.

J. Shallit and Y. Breitbart [SB96] explored a similar notion of descriptional complexity for languages. However, that measure turns out to be uninteresting for the case of a single string.

There is a connection between the measure studied in this chapter and the so-called "separating words" problem, which, given two strings w and x, both of length $\leq n$, asks for the number $B(w, x)$ of states in the smallest DFA M such that M *separates* w from x, i.e., M accepts exactly one of $\{w, x\}$. It is known that $B(w, x) = O(\log n)$ if $|w| \neq |x|$, and $B(w, x) = O(n^{2/5}(\log n)^{3/5})$ if $|w| = |x|$; see, for example, [GK86, Rob89, Rob96]. Given w, the function $A(w)$ can be viewed as measuring the size of the smallest DFA M such that M separates w from $\Sigma^{|w|} - \{w\}$.

9.2 Basic results

Clearly $A(x) \leq |x| + 1$, since we can uniquely accept any string of length $|x|$ with a chain of $|x|$ states that loops back to the start state, plus one additional "dead" state. It follows that A is computable, since we can simply examine all finite automata with $|x| + 1$ or fewer states, and test each DFA by brute force to see if x is accepted uniquely. As we will see below, it is possible to improve this algorithm somewhat, but we still do not know if $A(x)$ is computable in time polynomial in $|x|$.

It is possible, however, that $A(x)$ is significantly smaller than $|x|+1$. Roughly speaking, there are two ways to save states. The first is to use a loop. For example, the DFA in Figure 9.1 shows that $A(0^9 1^8) \leq 8$. (Unspecified transitions go to a "dead state" which is not shown.)

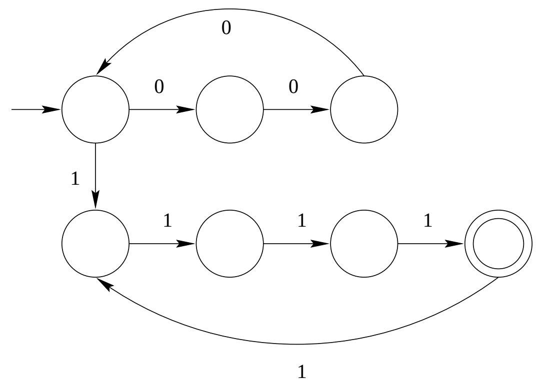


Figure 9.1: Automaton uniquely accepting $0^9 1^8$.

The second way to save states is through reuse. For example, you can reuse states, if the string is of the form $x\,y\,z\,\overline{y}^R\,w$, as shown in Figure 9.2. (By $\overline{y}$ we mean the string obtained by changing 0 to 1 and vice versa.)

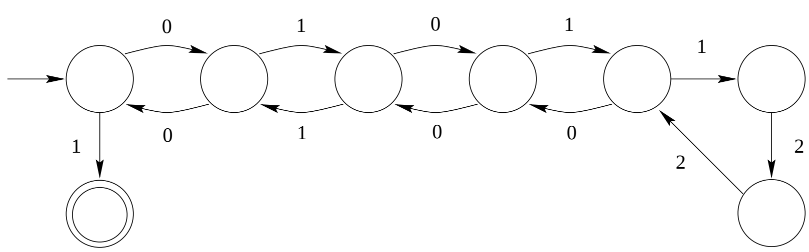


Figure 9.2: Automaton uniquely accepting 010112200101

Hopefully the reader is already convinced that this definition is somewhat natural and worthy of study.[1] Let us first see if the definition is useful; for example, can we use the measure as a data compression technique?

The answer is yes, in the following sense.

[1]But if not, there are some alternatives that also may be of interest. For example, we could define $B(x)$ to be the smallest number of states in any DFA M such that x is the lexicographically least string of length $|x|$ accepted by M.

Theorem 9.2.1 *Given a description of a DFA M which uniquely accepts x, and the length $n = |x|$, we can efficiently recover x.*

Proof. By "efficiently", we mean polynomial in the description size of M and n, the length of x.

Let $M = (Q, \Sigma, \delta, q_1, F)$ be a DFA uniquely accepting x. Let $Q = \{q_1, q_2, \ldots, q_r\}$, with $r = |Q|$. Create a directed graph $G = (V, E)$ with vertex set V defined as follows:

$$V = \{p_{i,j} \ : \ 1 \leq i \leq r, \ 0 \leq j \leq n\}.$$

Create a directed edge $(p_{i,j}, p_{k,l})$ labeled a if $\delta(q_i, a) = q_k$ and $l = j + 1$. Note that G is acyclic.

Since M uniquely accepts x, there exists a single index t with $q_t \in F$ such that there is exactly one path from $p_{1,0}$ to $p_{t,n+1}$, and for all $u \neq t$, there is no path from $p_{1,0}$ to $p_{u,n+1}$. We can now find this path using, for example, depth-first search, in $O(|V| + |E|) = O((n+1)|Q||\Sigma|)$ time. ■

Given a DFA, we can also efficiently decide if it uniquely accepts a given x.

Theorem 9.2.2 *Given a DFA M with r states and a string x of length $n \geq 1$, we can determine in $O(n + r^3 \log n)$ steps whether M uniquely accepts x.*

Proof. Let $M = (Q, \Sigma, \delta, q_1, F)$, where $Q = \{q_1, q_2, \ldots, q_r\}$. We can determine if M accepts x by simply simulating it on x, which can be done in $O(n)$ time.

Now create a matrix $M = (a_{i,j})_{1 \leq i,j \leq n}$ where $a_{i,j} = \text{Card}\{b \in \Sigma \ : \ \delta(q_i, b) = q_j\}$. Then an easy induction gives that if $M^k = (c_{i,j,k})_{1 \leq i,j \leq n}$, then $c_{i,j,k} = \text{Card}\{x \in \Sigma^k \ : \ \delta(q_i, x) = q_j\}$. Now compute $\sum_{j : q_j \in F} c_{1,j,|x|}$. This sum is 1 iff M uniquely accepts x.

Thus it suffices to compute M^k efficiently. To do so we can use the familiar "binary method" of exponentiation; see, for example, [BS96]. Furthermore, during the computation of M^k, we can always reduce an entry that is ≥ 2 to 2. The result is a matrix M' with entries in $\{0, 1, 2\}$ with the property that if an entry of M^k is 0 or 1, so is the corresponding entry of M', and if an entry of M^k is 2 or more, the corresponding entry in M' is 2. Since the sizes of the entries of M' are bounded by 2, it follows that this computation can be done in $O(r^3 \log n)$ bit operations. ■

Our last theorem of this section is the following:

Theorem 9.2.3 *We have $A(xx) \geq A(x)$ for all strings x.*

The following simple proof was shown to us at the DCAGRS 2000 workshop in London, Ontario, by Kai Salomaa:

Proof. Consider the DFA $M = (Q, \Sigma, \delta, q_0, F)$ minimizing $A(xx)$. Then we know there is only one path of length $|xx|$ from q_0 to a state of F, and this path is labeled xx. Let $q = \delta(q_0, x)$. Construct a new DFA $M' = (Q, \Sigma, \delta, q_0, \{q\})$. Then we claim M' uniquely accepts x. For if not, there exists another string $w \neq x$, $|w| = |x|$, such that $\delta(q_0, w) = q$. Then $\delta(q_0, wx) \in F$, and so M accepts wx, another string of length $|xx|$, and $wx \neq xx$. This contradiction proves that $A(x) \leq A(xx)$, as desired. ■

In Theorem 9.5.4 below we show that in fact $A(xx) > A(x)$ for infinitely many strings x.

9.3 Upper bounds

In this section we prove some upper bounds on $A(x)$.

Theorem 9.3.1 *Let $x \in \Sigma^*$ with $|\Sigma| = k \geq 2$ and $|x| = n$. Suppose $n > k^t + t - 1$. Then $A(x) \leq n + 2 - t$.*

Proof. If $n > k^t + t - 1$, then $x = a_1 a_2 \cdots a_n$ has at least $k^t + 1$ subwords of length t. Hence some subword of length t appears at least twice in x. Let y be a longest repeated subword, and let the first two occurrences of y be denoted y' and y'' (they may overlap).

Then we have the two factorizations shown in Figure 9.3.

$x =$

<table>
<tr><td>u</td><td>y'</td><td colspan="2">v</td><td>w</td></tr>
<tr><td>u</td><td colspan="2">v'</td><td>y''</td><td>w</td></tr>
</table>

Figure 9.3: Two factorizations of x.

where $y = y' = y''$. Furthermore, since y is a longest repeated subword, we know that either $w = \epsilon$ or the first letter of v differs from the first letter of w.

By a classic theorem of Lyndon & Schützenberger [LS62], the equality $yv = v'y$ implies that there exist strings r, s and an integer $e \geq 0$ such that

$$\begin{aligned} y &= (rs)^e r \\ v &= sr \\ v' &= rs. \end{aligned}$$

Thus $x = u(rs)^{e+1}rw$. It follows that the first letter of s differs from the first letter of w, so we can accept x uniquely with a DFA as in Figure 9.4.

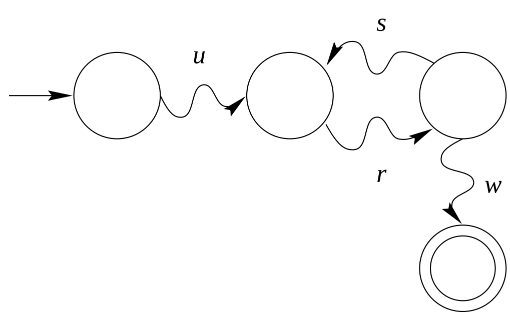


Figure 9.4: DFA uniquely accepting $x = u(rs)^{e+1}rw$.

The total number of states is $|ursw| + 2 = n + 2 - |y|$. ■

Theorem 9.3.2 *Let $x \in \{0,1\}^n$. Then*

$$A(x) \leq \frac{3}{4}n + (\log n)\sqrt{\frac{n}{8}}$$

for almost all strings x.

Proof. (Sketch.) The idea is to write $x = x'ax''$ where $|x'| = |x''| = \lfloor \frac{n}{2} \rfloor$ and $a \in \{\epsilon, 0, 1\}$. Then the expected number of mismatches between x' and x''^R is $\frac{n}{4} + O(1)$, with standard deviation $\sqrt{\frac{n}{8}} + O(1)$. We can now build a DFA for x', and attempt to reuse states corresponding to the mismatches between x' and x''^R, as in Figure 9.2. ■

9.4 Lower bounds

First, we show by a simple counting argument the existence of a constant C such that almost all strings x of length n satisfy $A(x) > C\frac{n}{\log n}$.

More precisely, we prove

Theorem 9.4.1 *Suppose $|\Sigma| = k \geq 2$, and let $0 < \varepsilon, \delta < 1$ be fixed. If n is sufficiently large, then*

$$A(x) \geq (1 - \delta)\varepsilon \frac{\log k}{k} \frac{n}{\log n}$$

for all strings $x \in \Sigma^n$, with at most $k^{\varepsilon n}$ exceptions.

Proof. It is easy to see there are at most q^{qk+1} essentially distinct automata with $\leq q$ states and exactly one final state. (The factor q^{qk} comes from the transition function, and the factor q comes from the assignment of final states. Note we can simulate a DFA with $\leq q$ states by one with exactly q states, by simply adding non-connected states, if necessary.) Each of these automata uniquely accepts at most one string of length n. Thus if

$$q^{qk+1} < k^{\varepsilon n}, \tag{9.1}$$

then at most $k^{\varepsilon n}$ different strings of length n can be represented. Now a routine calculation shows that if $q < (1 - \delta)\varepsilon \frac{\log k}{k} \frac{n}{\log n}$, then the inequality (9.1) holds. ■

It is possible to improve this bound as follows:

Theorem 9.4.2 *We have $A(x) \geq n/13$ for almost all strings $x \in \{0,1\}^n$.*

Proof. Suppose M is a DFA with $A(x)$ states that uniquely accepts x. Let $n = |x|$. Consider the transition diagram D of M, which is a labeled directed graph whose vertices are the states of M and whose (labeled) edges correspond to transitions. We define the *accepting path* P for x to be the sequence of $n+1$ edges traversed in this graph. Note that the first element of P is an edge labeled ϵ that enters the initial state q_0 of M. We define the *abbreviated accepting path* P' to be the sequence of edges obtained from P by considering each edge in order and deleting it if it has previously been traversed. The idea is to encode P' in a space-efficient manner so that x can be recovered.

The outdegree of each vertex encountered along P' is at most 2, since M is a DFA. We claim the indegree of each vertex is at most 2. If not, then let v be a vertex with indegree at least 3. Then there are at least three distinct edges entering v, say g_1, g_2, g_3. Let x_1 be a prefix of x such that the edge g_1 is used in the last transition when the DFA processes x_1. (If $v = q_0$, the initial state, we may have $x_1 = \epsilon$.) Let x_1x_2 be a prefix of x such that g_2 is used in the last transition when processing x_1x_2, $x_2 \neq \epsilon$. Let $x_1x_2x_3$ be a prefix of x such that g_3 is used in the last transition when processing $x_1x_2x_3$, $x_3 \neq \epsilon$. Finally, let x_4 be such that $x = x_1x_2x_3x_4$. Then $x' := x_1x_3x_2x_4$ is also accepted by M, and $|x| = |x'|$. If $x = x'$, then $x_2x_3 = x_3x_2$. Then, by a theorem of Lyndon & Schützenberger [LS62], there exist a string $z \neq \epsilon$ and integers $i, j \geq 1$ such that $x_2 = z^i$, $x_3 = z^j$. Now if there is a path labeled z^i going from v to v, and a path labeled z^j from v to v, then there is a path labeled $z^{\gcd(i,j)}$ from v to v. But then $g_2 = g_3$, a contradiction. Hence $x \neq x'$, contradicting the hypothesis that x is accepted uniquely.

Now consider the vertices visited by $P' = (e_0, e_1, \ldots, e_t)$, the abbreviated accepting path for x. Each vertex v is of exactly one of the following types:

Type 1. There is exactly one edge e_i of P' entering v and there is exactly one edge e_{i+1} leaving v.

Type 2. There are exactly two edges, e_i and e_j, $i < j$, entering v, and exactly one edge

e_{i+1} leaving v.

Type 3. There is one edge, e_i, entering v, and exactly two edges, e_{i+1} and e_j, $i < j$, leaving v.

Type 4. There are exactly two edges, e_i and e_j, entering v, with $i < j$, and there are exactly two edges, e_{i+1} and e_{j+1}, leaving v.

We now describe a space-efficient encoding E of P' which will avoid recording the state numbers. Instead, we record the labels of the edges along with some additional information that tells us what type each vertex is, and allows us to recover how these vertices are connected.

If $P' = (e_0, e_1, \ldots, e_t)$, then we define $E(i, n)$ to be a certain encoding, over the alphabet $\{0, 1, [,]_0,]_1, *, +\}$ of the edges $(e_i, \ldots, e_n)$. We also define a_i to be the label of the edge e_i corresponding to the symbol causing the transition. The meaning of the symbols is as follows: 0 and 1 represent the labels on the edges of P'. A left bracket [represents a vertex that is the target of a backedge. A right bracket ($]_0$ or $]_1$) represents a backedge labeled with its subscript. The symbol $+$ represents a vertex of outdegree 2, and the symbol $*$ (introduced later) represents a final state.

The base case is when $i > n$, in which case we define $E(i, n) = \epsilon$. For the inductive definition there are four cases, depending on the type of the vertex reached by the directed edge e_i, given in Figures 9.5–9.8.

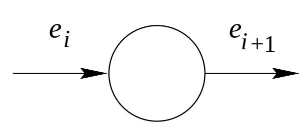


Figure 9.5: Vertex of type 1: $E(i, n) := a_i \ E(i + 1, n)$

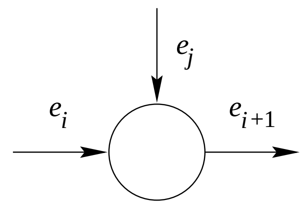


Figure 9.6: Vertex of type 2: $E(i,n) = a_i\,[\,E(i+1,j-1)\,]_{a_j}\ E(j+1,n)$

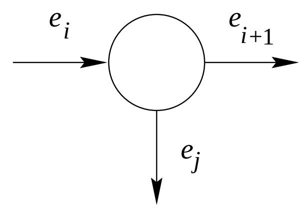


Figure 9.7: Vertex of type 3: $E(i,n) = a_i + E(i+1,n)$

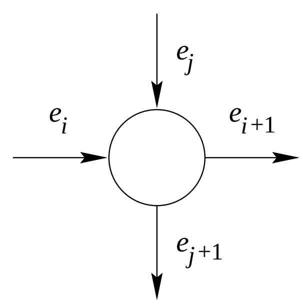


Figure 9.8: Vertex of type 4: $E(i,n) = a_i\,[\,+E(i+1,j-1)\,]_{a_j}E(j+1,n)$

Finally, if $P' = (e_0, e_1, \ldots, e_t)$, we define $E(x)$ to be $E(0,t)$ with a symbol $*$ inserted after the symbol leading to the (unique) accepting state, followed by the symbol #, followed by the base-2 representation of $n = |x|$, followed by ##. Thus $E(x)$ is a self-delimiting encoding of x over the 8-symbol alphabet $\{0, 1, +, *, [,]_0,]_1, \#\}$. We consider some examples of this encoding.

Figure	String	Encoding
Figure 9.1	$0^9 1^8$	$[+00]_0 1[111*]_1 \#10001\#\#$
Figure 9.10	0110100110	$[+0[+1[+10]_1]_0]_0 110 * \#1010\#\#$
Figure 9.11	01101001100101	$0[1 * 101[001+]_1]_0 \#1110\#\#$

We leave it to the reader to verify that P' can be reconstructed from $E(0,t)$ and x can be reconstructed from $E(x)$. It is easy to prove by induction that $|E(a,b)| \leq 2(b-a+1)$. Now P' has at most $2A(x)$ edges with nonempty labels, so we find $|E(0,t)| \leq 4A(x)+2$. It follows that $|E(x)| \leq 4A(x)+6+\log_2|x|$. Since E is over an 8-letter alphabet, it can be recoded over $\{0,1\}$ using three bits for each symbol. It follows that $C(x) \leq 12A(x) + 18 + 3\log_2|x|$. On the other hand, it is known that $C(x) \geq |x| - \log_2|x|$ for almost all x. Hence $A(x) \geq |x|/13$ for almost all x. ∎

Remark. We have not tried to optimize the constant 13 in Theorem 9.4.2. H. Petersen informs us (personal communication) that 13 can be reduced to 7.

We can improve the lower bound for certain kinds of strings, as follows:

Theorem 9.4.3 *Suppose $w \in \Sigma^*$ is kth-power-free for some integer $k \geq 2$, i.e., w contains no subword of the form x^k with $x \neq \epsilon$. Then $A(w) \geq \frac{|w|+1}{k}$.*

Proof. Let $w = a_1 a_2 \cdots a_n$ be uniquely accepted by some DFA $M = (Q, \Sigma, \delta, q_0, A)$, and define $p_i := \delta(q_0, a_1 a_2 \cdots a_i)$ for $0 \leq i \leq n$.

Suppose some state is visited at least $k+1$ times on the acceptance path for w. Then there exist indices $i_1, i_2, \ldots, i_{k+1}$ such that

$$p_{i_1} = p_{i_2} = \cdots = p_{i_{k+1}}.$$

Define

$$\begin{aligned}
w_0 &= a_1 a_2 \cdots a_{i_1} \\
w_1 &= a_{i_1+1} \cdots a_{i_2} \\
w_2 &= a_{i_2+1} \cdots a_{i_3} \\
&\vdots \\
w_k &= a_{i_k+1} \cdots a_{i_{k+1}} \\
w_{k+1} &= a_{i_{k+1}} \cdots a_n.
\end{aligned}$$

Then M uniquely accepts $w = w_0 w_1 w_2 w_3 \cdots w_{k+1}$. However, it also accepts, for example, $w' = w_0 w_2 w_1 w_3 \cdots w_{k+1}$. But $|w'| = |w|$. If $w_1 \neq w_2$, this gives a contradiction. Hence $w_1 = w_2$. By a similar argument we find $w_i = w_j$ for $1 \leq i, j \leq k$. It follows that $w = w_0 w_1^k w_{k+1}$, and so w contains a k'th power, a contradiction.

Thus we have shown that no state can be visited $k+1$ times on the acceptance path for w. Now for $0 \leq i < |Q|$ let b_i be the number of times state q_i is visited on the acceptance path for w. Then we have

$$\sum_{0 \leq i < |Q|} b_i q_i = n + 1.$$

But by the argument above $0 \leq b_i \leq k$. Thus

$$n + 1 = \sum_{0 \leq i < |Q|} b_i q_i \leq \sum_{0 \leq i < |Q|} k q_i = k|Q|.$$

It follows that $|Q| \geq (n+1)/k$, and so $A(w) = |Q| \geq (n+1)/k$, as desired. ∎

9.5 Some specific examples

In this section we determine the automatic complexity for some particular examples. There are interesting connections to number theory.

Theorem 9.5.1 *We have* $A(0^n 1^n) = O(\sqrt{n})$.

Proof. Assume $n \geq 1$. Let $r = \lfloor \sqrt{n} \rfloor$, so $r^2 \leq n < (r+1)^2$. Write $n = r^2 + a$. Then $0 \leq a \leq 2r$ and $r \geq 1$. Then we can accept $0^n 1^n$ with a DFA of the form given in Figure 9.9. (Unspecified transitions go to a "dead state" which is not shown.)

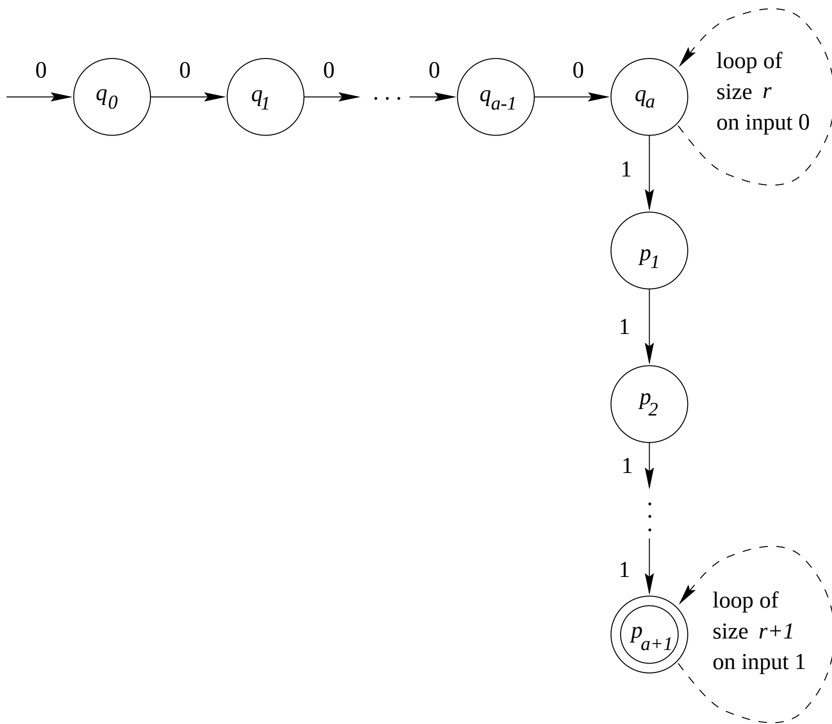


Figure 9.9: Automaton uniquely accepting $0^n 1^n$, where $n = r^2 + a$.

This DFA does indeed accept $0^n 1^n$ because

1. We go from state q_0 to state q_a on 0^a;

2. We then go around the loop at q_a r times;

3. Next on 1^a we go from q_a to p_{a+1};

4. Finally, we go around the loop at p_{a+1} $r-1$ times.

This path accepts $0^a(0^r)^r 1 1^a (1^{r+1})^{r-1} = 0^{r^2+a}1^{r^2+a}$.

On the other hand, we claim that this DFA accepts no other string of length $2n$. Suppose it did. Then any accepting path must go around the loop on q_a b times and the loop on p_{a+1} c times. Then

$$2n = a + br + a + 1 + c(r+1).$$

Since $n = r^2 + a$, it follows that $2r^2 - 1 = br + c(r+1)$. Reducing modulo r, we get $c \equiv -1 \pmod{r}$. Thus $c \in \{r-1, 2r-1, \ldots\}$. But if $c \geq 2r-1$ then the string would be of length $\geq (2r-1)(r+1) + 2a + 1 = 2r^2 + r - 1 + 2a + 1 \geq 2n + r > 2n$, a contradiction.

Finally, our DFA uses $a + 1 + r - 1 + a + 1 + r = 2r + 2a + 1 \leq 6r + 1 \leq 6\sqrt{n} + 1$ states.

■

We now show that the bound of $O(\sqrt{n})$ is tight. First we state the following lemma:

Lemma 9.5.2 *Let c, d be integers ≥ 1. Suppose the linear diophantine equation $N = xc + yd$ is solvable in integers, i.e., suppose $\gcd(c,d) \,|\, N$. If $N > 2cd - c - d$, then the linear diophantine equation $N = xc + yd$ has at least two solutions in non-negative integers x, y.*

The proof is easy and left to the reader. This result (in a more general form) has recently been proved independently by Beck & Robins [BR00].

We now prove

Theorem 9.5.3 *Any DFA that uniquely accepts $0^n 1^n$ must have at least $\sqrt{n} - 1$ states.*

Proof. Suppose M is a DFA with $< \sqrt{n} - 1$ states that uniquely accepts $0^n 1^n$. Define $p_i = \delta(q_0, 0^i)$ for $0 \leq i \leq n$. Since M has $< n$ states, some state must be repeated, and thus there must be a "loop" of $r \geq 1$ states that is repeated j times, for some integer $j \geq 0$. There may also be a "tail" at the beginning, and at the end we may not go around the "loop" an integral number of times. Let $s = n - rj$. Then $r, s < \sqrt{n} - 1$.

Similarly, define $r_i = \delta(p_n, 1^i)$ for $0 \le i \le n$. By the same argument there must be a "loop" of $u \ge 1$ states that is repeated k times, for some integer $k \ge 0$. Let $t = n - ku$. Then $t, u < \sqrt{n} - 1$.

Since M accepts $0^n 1^n$ uniquely, it must be the case that the equation $ra + ub = 2n - s - t$ has exactly one solution $(a, b) = (j, k)$. Then, by Lemma 9.5.2, we have $2n - s - t \le 2ru - r - u$. Thus $2n - 2(\sqrt{n} - 1) \le 2n - s - t \le 2ru - r - u \le 2(\sqrt{n} - 1)(\sqrt{n} - 1) - 2$. But then $2\sqrt{n} + 2 \le 0$, a contradiction. ∎

We can now exhibit infinitely many strings for which $A(xx) > A(x)$.

Theorem 9.5.4 *Let $x = 0^n 1$. Then $A(xx) = \Omega(\sqrt{n})$, but $A(x) = O(1)$.*

Proof. It is clear that $A(x) = O(1)$, since we can accept $0^n 1$ uniquely with a 3-state DFA. However, mimicking the lower bound proof of Theorem 9.5.3 above, it is easy to see that $A(0^n 1 0^n 1) = \Omega(\sqrt{n})$. ∎

It is possible to generalize Theorem 9.5.1. We need some technical lemmas. The first concerns solvability of certain linear diophantine equations.

Lemma 9.5.5 *Let $k \ge 1$, and let $n_1, n_2, \ldots, n_k$ be positive integers, relatively prime in pairs. Let $r \ge 0$ be an integer. Define $P := n_1 n_2 \cdots n_k$. If $r = 0$, the linear diophantine equation*

$$a_1 \frac{P}{n_1} + a_2 \frac{P}{n_2} + \cdots + a_k \frac{P}{n_k} = (n_1 - 1)\frac{P}{n_1} + (n_2 - 1)\frac{P}{n_2} + \cdots + (n_k - 1)\frac{P}{n_k} - rP \qquad (9.2)$$

has a unique solution

$$(a_1, a_2, \ldots, a_k) = (n_1 - 1, n_2 - 1, \ldots, n_k - 1)$$

in non-negative integers. If $r \ge 1$, then (9.2) has no solutions in non-negative integers.

Proof. By induction on k. If $k = 1$ Eq. (9.2) becomes $a_1 = n_1 - 1 - rn_1$. If $r = 0$ this equation has the unique solution $a_1 = n_1 - 1$, but if $r \geq 1$ then clearly there are no solutions in non-negative integers.

Now assume the result is true for $1, 2, \ldots, k-1$. We prove it for k. Consider Eq. (9.2) mod n_k. We get

$$a_k \frac{P}{n_k} \equiv -\frac{P}{n_k} \pmod{n_k}.$$

Since the n_i are pairwise relatively prime, it follows that $a_k \equiv -1 \pmod{n_k}$. Since a_k is a non-negative integer, we can therefore write $a_k = jn_k - 1$ for some integer $j \geq 1$.

Now substitute $a_k = jn_k - 1$ in Eq. (9.2). After a little easy algebra, we get

$$\begin{aligned} a_1 \frac{P}{n_1} + a_2 \frac{P}{n_2} + \cdots + a_{k-1} \frac{P}{n_{k-1}} = \\ (n_1 - 1)\frac{P}{n_1} + (n_2 - 1)\frac{P}{n_2} + \cdots + (n_{k-1} - 1)\frac{P}{n_{k-1}} - (j + r - 1)P \end{aligned} \tag{9.3}$$

By induction Eq. (9.3) has a solution iff $j + r - 1 = 0$. But $j \geq 1$. Hence $j = 1$ and $a_k = n_k - 1$, and hence Eq. (9.3) has a solution iff $r = 0$. If $r = 0$, by induction the solution is $(a_1, a_2, \ldots, a_{k-1}) = (n_1 - 1, n_2 - 1, \ldots, n_{k-1} - 1)$.

Lemma 9.5.6 *(a) If $M^{1/k} > 2B$, then*

$$\frac{M}{M^{1/k} - B} < M^{\frac{k-1}{k}} + 2BM^{\frac{k-2}{k}}.$$

(b) If $0 < B < A$ and $k \geq 1$, then

$$(A - B)^k \geq A^k - kA^{k-1}B.$$

Proof.

(a) We have

$$(M^{1/k} - B)(M^{\frac{k-1}{k}} + 2BM^{\frac{k-2}{k}}) = M + BM^{\frac{k-2}{k}}(M^{1/k} - 2B) > M.$$

(b) An easy induction on k proves that if $0 < x < 1$ and $k \geq 1$ then $(1-x)^k \geq 1 - kx$. Now let $x = B/A$ and multiply by A^k.

For our last lemma, we will need a certain number-theoretic function. For $t \geq 1$, define $f(t)$ to be the least integer n such that every set of n consecutive positive integers contains a subset of size t that is pairwise relatively prime. Then, for example, $f(4) = 6$, since the set $\{2, 3, 4, 5, 6\}$ contains no subset of 4 relatively prime integers, while it is easy to check that every set of 6 consecutive positive integers does.

It seems quite difficult to estimate f precisely. However, the following lemma follows easily from results of Erdős and Selfridge [ES71]:

Lemma 9.5.7 *For all $\delta > 0$ and t sufficiently large we have $f(t) < t^{2+\delta}$.*

Proof. Erdős and Selfridge defined $F(n, k)$ to be the largest subset of pairwise relatively prime integers in $\{n+1, n+2, \ldots, n+k\}$, and proved that $\min_{n \geq 0} F(n, k) > k^{1/2-\epsilon}$. Now let $k = t^{2+5\epsilon}$ for some $\epsilon < 1/10$. We find

$$\min_{n \geq 0} F(n, t^{2+5\epsilon}) > (t^{2+5\epsilon})^{1/2-\epsilon} > t^{1+\epsilon/2-5\epsilon^2} > t,$$

since $\epsilon < 1/10$. Hence for all $0 < \epsilon < 1/10$ and all t sufficiently large, any $t^{2+5\epsilon}$ consecutive integers contains a pairwise relatively prime subset of cardinality $> t$. In other words, $f(t) < t^{2+\delta}$ where $\delta = 5\epsilon$. ■

We are now ready to prove

Theorem 9.5.8 *Let $a_1, a_2, \ldots, a_k$ be k distinct symbols. Then $A(a_1^n a_2^n \cdots a_k^n) = O(n^{1-1/k})$, where the constant in the big-O may depend on k.*

Proof. The idea is as follows: we choose k pairwise relatively prime integers, each $\leq n^{1/k}$, say $n_1, n_2, \ldots, n_k$. Let $P = n_1 n_2 \cdots n_k$. We then form a DFA similar to that in Figure 9.9, with k loops, one on each a_i, $1 \leq i \leq k$, of size P/n_i. Each loop is preceded by a "tail" of length $n - (P/n_i)(n_i - 1) = n - P + P/n_i$. By Lemma 9.5.5, this DFA uniquely accepts $a_1^n a_2^n \cdots a_k^n$.

The total number of states is $\leq N$, where

$$N := 1 + k(n - P) + 2 \sum_{1 \leq i \leq k} \frac{P}{n_i}. \tag{9.4}$$

By Lemma 9.5 we can choose the pairwise relatively prime numbers n_i such that $n^{1/k} - k^{2+\delta} < n_i < n^{1/k}$. Setting $A = n^{1/k}$ and $B = k^{2+\delta}$ in Lemma 9.5.6 (b) we obtain

$$P = n_1 n_2 \cdots n_k \geq n - k^{3+\delta} n^{\frac{k-1}{k}}.$$

Hence

$$k(n - P) < k^{4+\delta} n^{\frac{k-1}{k}}. \tag{9.5}$$

On the other hand, setting $M = P$ and $B = k^{2+\delta}$ in Lemma 9.5.6 (a) we obtain

$$\frac{P}{n_i} < P^{\frac{k-1}{k}} + 2k^{2+\delta} P^{\frac{k-2}{k}} \tag{9.6}$$

for all n sufficiently large. Combining Eqs. (9.4)–(9.6), we obtain $N = O(k^{4+\delta} n^{\frac{k-1}{k}})$, as desired. ■

9.6 Infinite words

Up to now we have been dealing with finite words. However, it is also interesting to consider the case of infinite words. In this chapter, by an infinite word we will mean a one-sided, right-infinite word, i.e., a map from $\mathbb{N}$ to Σ. For an infinite word $\mathbf{x}$ we are interested in computing

$$I(\mathbf{x}) = \liminf_{x \text{ is a prefix of } \mathbf{x}} \frac{A(x)}{|x|}$$

and

$$S(\mathbf{x}) = \limsup_{x \text{ is a prefix of } \mathbf{x}} \frac{A(x)}{|x|}.$$

for "interesting" infinite words $\mathbf{x}$.

We start with the Thue-Morse word $\mathbf{t}$. Let μ be a morphism defined by $\mu(0) = 01$, $\mu(1) = 10$. Then $\mathbf{t} = t_0 t_1 t_2 \cdots = \lim_{n\to\infty} \mu^n(0)$. We define $T(r) = t_0 t_1 \cdots t_{r-1}$, the prefix of $\mathbf{t}$ of length r.

Theorem 9.6.1 *We have*

$$I(\mathbf{t}) \geq \frac{1}{3}$$

and

$$S(\mathbf{t}) \leq \frac{2}{3}.$$

Proof. The lower bound for $I(\mathbf{t})$ follows immediately from Theorem 9.4.3, since, as is well-known, the Thue-Morse word is cube-free.

For the upper bound, we break the argument up as follows. We claim that we can accept $T(m)$ using $h(m)$ states, where h is given in the table below.

m	$h(m)$
$2 \cdot 2^{2n} \leq m \leq 3 \cdot 2^{2n}$	$m + 3 - 2^{2n}$
$3 \cdot 2^{2n} < m < 4 \cdot 2^{2n}$	$2 \cdot 2^{2n} + 2$
$4 \cdot 2^{2n} \leq m < 5 \cdot 2^{2n}$	$m + 2 - 2 \cdot 2^{2n}$
$5 \cdot 2^{2n} \leq m \leq 6 \cdot 2^{2n}$	$m + 1 - 2 \cdot 2^{2n}$
$6 \cdot 2^{2n} < m < 8 \cdot 2^{2n}$	$4 \cdot 2^{2n} + 2$

For $2 \cdot 2^{2n} \leq m \leq 3 \cdot 2^{2n}$, we use the fact that $T(2 \cdot 2^{2n}) = T(2^{2n})\overline{T(2^{2n})^R}$, which allows us to reuse $2^{2n} - 1$ states, as illustrated in Figure 9.10.

For $3 \cdot 2^{2n} \leq m < 4 \cdot 2^{2n}$, we use the fact that $T(4 \cdot 2^{2n}) = T(2^{2n})(\overline{T(2^{2n})})^2 T(2^{2n})$, which allows us to reuse 2^{2n} states in an inner loop and $m - 3 \cdot 2^{2n}$ states in an outer loop, as illustrated in Figure 9.11.

For $4 \cdot 2^{2n} \leq m < 5 \cdot 2^{2n}$, it is easiest to give the encoding of the corresponding machine, as introduced in Section 9.4:

$$t_0[t_1 \cdots t_{m-3\cdot 2^{2n}-1} * t_{m-3\cdot 2^{2n}}[t_{m-3\cdot 2^{2n}+1} \cdots t_{2^{2n+1}-1} + t_{2^{2n+1}} \cdots t_{m-2^{2n+1}-1}]_{t_{m-2^{2n+1}}}]_{t_{3\cdot 2^{2n}}}$$

This is illustrated in Figure 9.12.

For $5 \cdot 2^{2n} \leq m \leq 6 \cdot 2^{2n}$, we use the fact that $T(5 \cdot 2^{2n}) = (T(2^{2n})\overline{T(2^{2n})T(2^{2n})})^{5/3}$, as illustrated in Figure 9.13.

For $6 \cdot 2^{2n} \leq m \leq 8 \cdot 2^{2n}$, we use the fact that $T(m) = T(2^{2n+3} - m)x\overline{x}^R$, where $x = t_{2^{2n+3}-m} \cdots t_{2^{2n+2}}$, as illustrated in Figure 9.14.

In the figures that follow, unspecified transitions go to a "dead state" which is not shown.

■

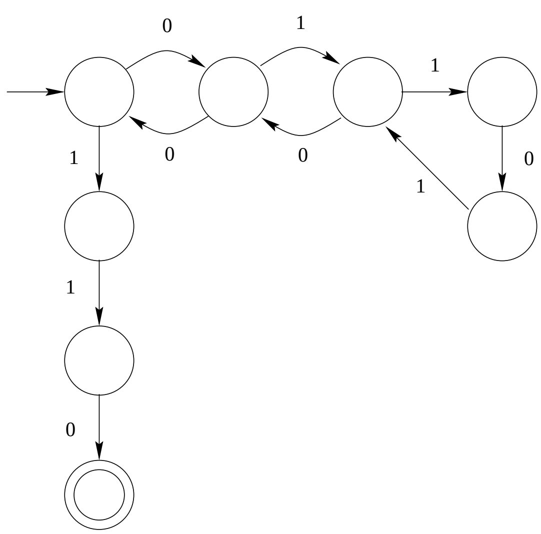


Figure 9.10: Automaton uniquely accepting t_{10}

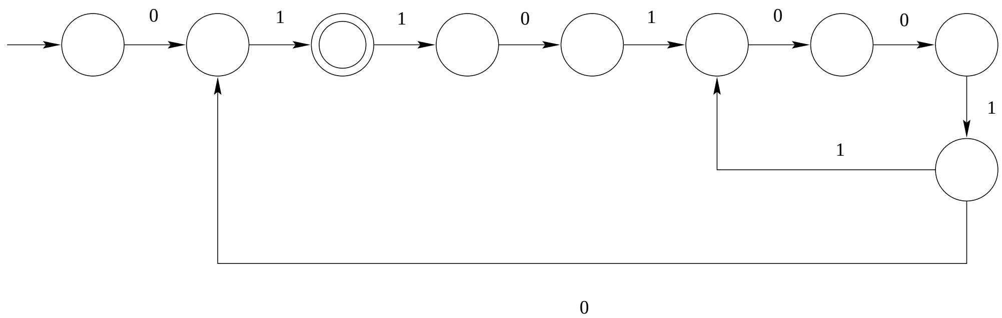


Figure 9.11: Automaton uniquely accepting t_{14}

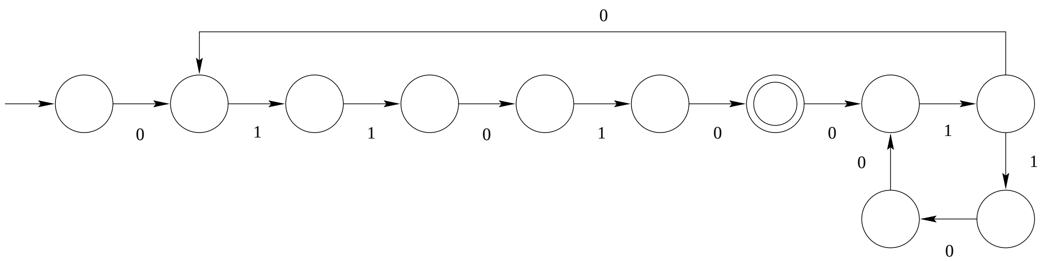


Figure 9.12: Automaton uniquely accepting t_{18}

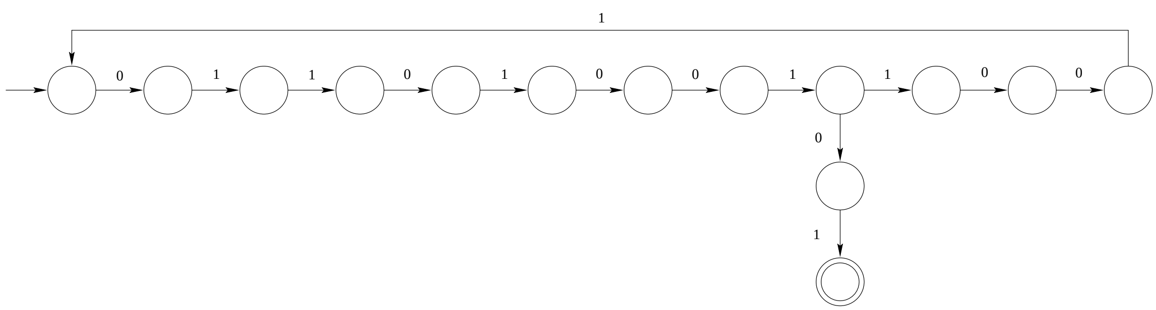


Figure 9.13: Automaton uniquely accepting t_{22}

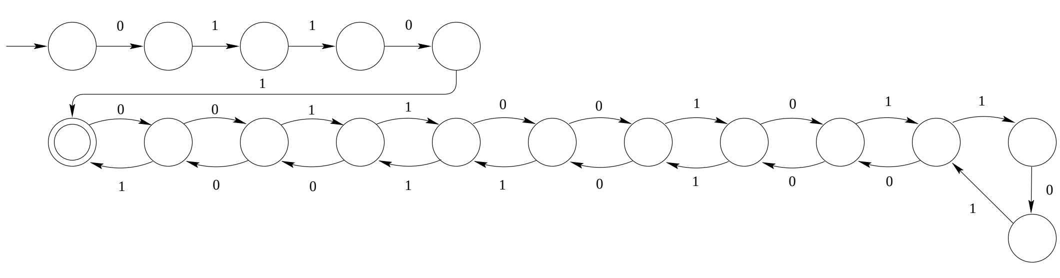


Figure 9.14: Automaton uniquely accepting t_{26}

If we consider the Thue-Morse word on three symbols [Ber79], we can get a sharper result.

Theorem 9.6.2 *Let* $\mathbf{u} = 102120102012\cdots$ *be the infinite Thue-Morse word on three symbols, generated by* $1 \to 102$, $0 \to 12$, $2 \to 0$*. Then*

$$I(\mathbf{u}) = \frac{1}{2}.$$

Proof. The lower bound comes from Theorem 9.4.3. For the upper bound, we claim that if we let rs be a prefix of $\mathbf{u}$ with $|r| = |s| = 2^{2k}$ for some $k \geq 0$, then r differs in every single position from s^R. Hence we may reuse states in analogy with Figure 9.2.

To prove the claim, let $\mathbf{u} = u_0u_1u_2\cdots$. Let $r = u_0u_1\cdots u_{2^{2k}-1}$ and $s = u_{2^{2k}}\cdots u_{2^{2k+1}-1}$. It is known that $u_i = (2t_i + t_{i+1}) \bmod 3$, where $\mathbf{t} = t_0t_1t_2\cdots$ is the Thue-Morse word. For $0 \leq i < 2^{2k+1}$ we have $t_i = 1 - t_{2^{2k+1}-i-1}$. Now $\mathbf{t}$ is cube-free, so $t_{i-1}t_it_{i+1} \not\in \{000, 111\}$. Hence $t_i + t_{i+1} \neq 2t_{i-1}$, and $t_j \in \{0, 1\}$ for $j \geq 0$, so $t_i + t_{i+1} \not\equiv 2t_{i-1} \pmod 3$. Adding t_i to this last incongruence, we have

$$\begin{aligned} u_i &\equiv 2t_i + t_{i+1} \not\equiv t_i + 2t_{i-1} \\ &\equiv 2(1 - t_i) + (1 - t_{i-1}) \\ &\equiv 2t_{2^{2k+1}-i-1} + t_{2^{2k+1}-i} \\ &\equiv u_{2^{2k+1}-i-1} \pmod 3. \end{aligned}$$

This proves that r differs in every position from s^R. ■

One might wonder if $S(\mathbf{x}) = 0$ implies that $\mathbf{x}$ is ultimately periodic. The answer is no, as the following example shows:

Theorem 9.6.3 *Let* $\mathbf{v} = 0^2\,1\,0^4\,1\,0^{16}\,1\,0^{256}\,10^{65536}\,1\cdots$*. Then* $S(\mathbf{v}) = 0$*.*

Proof. First, we need the following lemma.

Lemma 9.6.4 *Let* $y \in \{0, 1\}^*$*. Then for all* $r, s \geq 1$ *we have* $A(y10^r10^s) \leq |y| + 6\sqrt{m}$*, where* $m = \max(r, s)$*.*

Proof. Suppose $r \geq s$. (The proof for $r < s$ is similar and is left to the reader.) Define $t := \lfloor \sqrt{r} \rfloor$ and $u = \min(\lfloor s/t \rfloor, t)$. Now construct the following DFA:

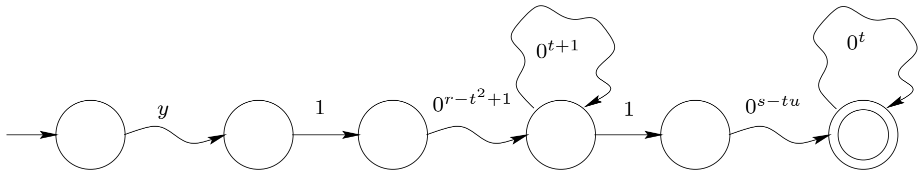


Figure 9.15: DFA uniquely accepting $y10^r10^s$

We claim this DFA uniquely accepts $y10^r10^s$. It clearly accepts this string, by going around the first loop $t-1$ times and the second loop u times. To see that acceptance is unique, consider going around the first loop $k \geq 0$ times and the second loop $j \geq 0$ times. This gives a string of length $|y| + 1 + r - t^2 + 1 + k(t+1) + 1 + s - tu + jt$. Setting this equal to the desired length of $|y| + 1 + r + 1 + s$, we get the linear diophantine equation

$$r - t^2 + 1 + k(t+1) + s - tu + jt = r + s;$$

in other words, $k(t+1) + jt = t^2 - 1 + tu$. Now consider this equation modulo t. We find $k \equiv -1 \pmod{t}$. Suppose $k \geq 2t - 1$. Then $t^2 - 1 + tu = k(t+1) + jt \geq (2t-1)(t+1) + jt$. Simplifying, we obtain $u \geq t + 1 + j$. But $j \geq 0$, so $u \geq t + 1$, contradicting the definition of u. It follows that $k = t - 1$, and hence $j = u$, as desired.

Our DFA has $N := |y| + 1 + r - t^2 + 1 + t + 1 + s - tu + t - 1 + 1$ states. Since $\sqrt{r} - 1 \leq t \leq \sqrt{r}$, it follows that $r - t^2 \leq 2\sqrt{r} - 1$ and

$$\begin{aligned} s - tu &\leq s - t\min(\lfloor s/t \rfloor, t) = \max(s - t\lfloor s/t \rfloor, s - t^2) = \max(s \bmod t, s - t^2) \\ &\leq \max(t, r - t^2) \leq \max(\sqrt{r}, 2\sqrt{r} - 1) \leq 2\sqrt{r} - 1. \end{aligned}$$

Hence $N < |y| + 6\sqrt{r}$. ■

Now we can complete the proof of Theorem 9.6.3. Every sufficiently long prefix of $\mathbf{v}$ is of the form

$$x = 0^2\,1\,0^{2^2}\,1\,0^{2^{2^2}}\,1\,0^{2^{2^3}}\,1\cdots 1\,0^{2^{2^n}}\,1\,0^a$$

where $0 \le a \le 2^{2^{n+1}}$. Let $y = 0^2\,1\,0^{2^2}\,1\,0^{2^{2^2}}\,1\,0^{2^{2^3}}\,1\cdots 1\,0^{2^{2^{n-1}}}$, $r = 2^{2^n}$, and $s = a$. Then $|x| = 2^{2^n} + a + O(2^{2^{n-1}})$, while Lemma 9.6.4 states that $A(x) \le 6\sqrt{2^{2^n}+a} + O(2^{2^{n-1}})$. It follows that $A(x)/|x| = O(1/\sqrt{x})$, and so $S(\mathbf{v}) = 0$. ■

9.7 Open Problems

There are many open problems related to this work. For example, is $A(x)$ computable in polynomial time? Does the inequality $A(x) < A(xx)$ hold for almost all x ?

Chapter 10

Conclusion

In this thesis, we presented several results on periodicity and repetition in words. The most significant contributions are the results in Chapter 2 – Chapter 5.

In Chapter 2 we gave several non-trivial generalizations of the Fine and Wilf theorem. In particular, we generalize the Fine and Wilf theorem to inequalities and to more than two periodic sequences.

In Chapter 3 we gave two results on non-negative matrices that have direct applications to the growth rate of words in a D0L system.

In Chapter 4 we gave a complete classification of two-sided fixed points of morphisms. This characterization completes the previous work of Head and Lando [Hea81, HL86] on finite and one-sided fixed points of morphisms. It also led to the formulation of the Decreasing Length Conjecture.

In Chapter 5 we gave a method (following Thue) to construct an infinite binary word that simultaneously avoid cubes and squares xx with $|x| \geq 4$. This method can be applied to other pattern avoidance problems.

There are many potential areas for further research. We list some of them below.

1. There is a version of the Fine and Wilf theorem for continuous functions. It is as follows. Let $f(x)$ and $g(x)$ be continuous periodic real functions of periods α, β, respectively.

(a) $\alpha/\beta = p/q$ where p, q are coprime positive integers. Then if $f(x) = g(x)$ on an interval of length $\alpha + \beta - \beta/q$, then $f(x) = g(x)$ everywhere. Furthermore, the result would be false if $\alpha + \beta - \beta/q$ were replaced by anything smaller.

(b) α/β is irrational. Then if $f(x) = g(x)$ on an interval of length $\alpha + \beta$, then $f(x) = g(x)$ everywhere. Furthermore, the result would be false if $\alpha + \beta$ were replaced by anything smaller.

None of the generalizations mentioned in this thesis have been proved/adapted for continuous functions. It would be interesting to extend these results to continuous functions. The proof of the upper bounds in both cases above can probably be adapted to these generalizations. To prove matching lower bounds seems to require something new.

2. Let A be a non-negative $n \times n$ irreducible matrix. In Chapter 3 we showed that there is an integer $e = O(n \log n)$ such that A^e has strictly positive diagonal entries. We also showed a lower bound of $\Omega(n \log n / \log \log n)$ on e. It would be nice to have matching upper and lower bounds here.

3. L. Ilie, P. Ochem and J. Shallit [IOS04] considered the following type of pattern avoidance problem.

   Let $\alpha > 1$ be a rational number, and let $\ell \geq 1$ be an integer. A word w is a *repetition of order α and length ℓ* if we can write it as $w = x^n x'$ where x' is a prefix of x, $|x| = l$, and $|w| = \alpha|x|$. For brevity, we also call w a (α, ℓ)-*repetition.* Notice that an α-power is an (α, ℓ)-repetition for some ℓ. We say a word is (α, ℓ)-*free* if it contains no factor that is a (α', ℓ')-repetition for $\alpha' \geq \alpha$ and $\ell' \geq \ell$. We say a word is (α^+, ℓ)-*free* if it is (α', ℓ)-free for all $\alpha' > \alpha$.

   For integers $k \geq 2$ and $\ell \geq 1$, we define the *generalized repetition threshold* $R(k, \ell)$ as the real number α such that either

(a) over Σ_k there exists an (α^+, ℓ)-free infinite word, but all (α, ℓ)-free words are finite; or

(b) over Σ_k there exists a (α, ℓ)-free infinite word, but for all $\epsilon > 0$, all $(\alpha - \epsilon, \ell)$-free words are finite.

$R(k, \ell)$ generalizes the repetition threshold of Brandenburg [Bra83]. Apparently Pansiot also suggested looking at this generalization at the end of his paper [Pan84]. [IOS04] gives several results on $R(k, \ell)$ and contains many open problems. We mention two below.

Conjecture 10.0.1 *$R(3, \ell) = 1 + \frac{1}{\ell}$ for $\ell \geq 2$.*

Conjecture 10.0.2 *$R(4, \ell) = 1 + \frac{1}{\ell+2}$ for $\ell \geq 2$.*

These conjectures are weakly supported by numeric evidence. A lot of related problems could be investigated.

4. Is the language $D^*(w)$ defined in Chapter 8 context-free?

5. Is the automatic complexity function $A(x)$ defined in Chapter 9 computable in polynomial time?

6. Does the inequality $A(xx) > A(x)$ hold for almost all x?

7. In Chapter 9, we showed that given an integer $k \geq 1$ and a "large" N, there exists k positive integers $c_1, \ldots, c_k$ each approximately $N^{(k-1)/k}$ and a positive integer C with the following two properties:

   1. C can be written as a unique positive linear combination of c_i.

   2. Let $C = \sum_i x_i c_i$. Then $x_i c_i \sim N$.

   Is it possible to find c_i's approximately $\sqrt{N}$? This would improve the bound in Theorem 9.5.8 and is of some independent interest.

Bibliography

[AB89] A. Arnold and S. Brlek. Optimal word chains for the Thue-Morse word. *Inform. Comput.*, 83:140–151, 1989.

[Adi79] S. I. Adian. *The Burnside problem and identities in groups*, volume 95 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, 1979.

[AL77] A. Alder and S.-Y. R. Li. Magic cubes and Prouhet sequences. *Amer. Math. Monthly*, 84:618–627, 1977.

[All87] J.-P. Allouche. Automates finis en théorie des nombres. *Exposition. Math.*, 5:239–266, 1987.

[Alt90] I. Althöfer. Tight lower bounds for the length of word chains. *Inform. Process. Lett.*, 34:275–276, 1990.

[AM01] J.-P. Allouche and Cosnard M. Non-integer bases, iteration of continuous real maps, and an arithmetic self-similar set. *Acta Math. Hungar.*, 91:325–332, 2001.

[AS99] J.-P. Allouche and J. Shallit. The ubiquitous Prouhet-Thue-Morse sequence. In C. Ding, T. Helleseth, and H. Niederreiter, editors, *Sequences and Their Applications: Proceedings of SETA '98*, pages 1–16. Springer-Verlag, 1999.

[AS03] J.-P. Allouche and J. Shallit. *Automatic Sequences: Theory, Applications, Generalizations*. Cambridge University Press, 2003.

[BB99] J. Berstel and L. Boasson. Partial words and a theorem of Fine and Wilf. *Theoret. Comput. Sci.*, 218:135–141, 1999.

[BB88] J. Berstel and S. Brlek. On the length of word chains. *Inform. Process. Lett.*, 26:23–28, 1987/88.

[Bea85] D. Beaquier. Ensembles reconnaissables de mots bi-infinis. In M. Nivat and D. Perrin, editors, *Automata on Infinite Words*, volume 192 of *Lecture Notes in Computer Science*, pages 28–46. Springer-Verlag, 1985.

[BEM79] D. A. Bean, G. Ehrenfeucht, and G. McNulty. Avoidable patterns in strings of symbols. *Pacific J. Math.*, 85:261–294, 1979.

[Ber79] J. Berstel. Sur la construction de mots sans carré. *Séminaire de Théorie des Nombres*, pages 18.01–18.15, 1978–1979.

[Ber95] J. Berstel. *Axel Thue's Papers on Repetitions in Words: a Translation.* Number 20 in Publications du Laboratoire de Combinatoire et d'Informatique Mathématique. Université du Québec à Montréal, February 1995.

[Ber02] J. Berstel. Sturmian words. In M. Lothaire, editor, *Algebraic Combinatorics on Words*, pages 40–97. Cambridge University Press, 2002. Preliminary version available at `http://www-igm.univ-mlv.fr/~berstel/Lothaire/`.

[BH02] F. Blanchet-Sadri and R. A. Hegstrom. Partial words and a theorem of Fine and Wilf revisited. *Theoret. Comput. Sci.*, 270:401–419, 2002.

[BK03] J. Berstel and J. Karhumäki. Combinatorics on words – a tutorial. Technical Report 530, Turku Centre for Computer Science, June 2003.

[Bo00] Z. Bo. Improvements on inequalities for non-negative matrices. *Australasian J. Combinat.*, 21:251–255, 2000.

[Bou92] M. Bousquet-Mélou. The number of minimal word chains computing the Thue-Morse word. *Inform. Process. Lett.*, 44:57–64, 1992.

[BR00] M. Beck and S. Robins. A formula related to the Frobenius problem in two dimensions. Unpublished manuscript, January, 2000.

[Bra83] F.-J. Brandenburg. Uniformly growing k-th power-free homomorphisms. *Theoret. Comput. Sci.*, 23:69–82, 1983.

[Bro71] T. C. Brown. Is there a sequence on four symbols in which no two adjacent segments are permutations of one another? *Amer. Math. Monthly*, 78:886–888, 1971.

[BS96] E. Bach and J. Shallit. *Algorithmic Number Theory.* The MIT Press, 1996.

[Cas93] J. Cassaigne. Unavoidable binary patterns. *Acta Informatica*, 30:385–395, 1993.

[CK97] C. Choffrut and J. Karhumäki. Combinatorics of words. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 1, pages 329–438. Springer-Verlag, 1997.

[CMR99] M. G. Castelli, F. Mignosi, and A. Restivo. Fine and Wilf's theorem for three periods and a generalization of sturmian words. *Theoret. Comput. Sci.*, 218:83–94, 1999.

[CMS+03] S. Cautis, F. Mignosi, J. Shallit, M.-w. Wang, and S. Yazdani. Periodicity, morphisms, and matrices. *Theoret. Comput. Sci.*, 295:107–121, 2003.

[Cob68] A. Cobham. On the Hartmanis-Stearns problem for a class of tag machines. In *IEEE Conference Record of 1968 Ninth Annual Symposium on Switching and Automata Theory*, pages 51–60, 1968. Also appeared as IBM Research Technical Report RC-2178, August 23 1968.

[CR94] M. Crochemore and W. Rytter. *Text Algorithms*. Oxford University Press, 1994.

[Cur93] J. D. Currie. Open problems in pattern avoidance. *Amer. Math. Monthly*, 100:790–793, 1993.

[Dek76] F. M. Dekking. On repetitions of blocks in binary sequences. *J. Combin. Theory. Ser. A*, 20:292–299, 1976.

[Diw86] A. A. Diwan. A new combinatorial complexity measure for languages. Technical report, Computer Science Group, Tata Institute, Bombay, 1986.

[dLM94] A. de Luca and F. Mignosi. some combinatorial properties of Sturmian words. *Theoret. Comput. Sci.*, 136:361–385, 1994.

[DM64] A. L. Dulmage and N. S. Mendelsohn. Gaps in the exponent set of primitive matrices. *Illinois J. Math.*, 8:642–656, 1964.

[DMP99] J. Dassow, V. Mitrana, and Gh. Păun. On the regularity of duplication closure. *Bull. European Assoc. Theor. Comput. Sci.*, 68:133–136, 1999.

[DT92] J. Devolder and E. Timmerman. Finitary codes for biinfinite words. *RAIRO Inform. Théor. App.*, 26:363–386, 1992.

[EJS74] R. C. Entringer, D. E. Jackson, and J. A. Schatz. On nonrepetitive sequences. *J. Combin. Theory. Ser. A*, 16:159–164, 1974.

[Elk86] N. Elkies. An improved lower bound on the greatest element of a sum-distinct set of fixed order. *J. Combin. Theory. Ser. A*, 41:89–94, 1986.

[ER84] A. Ehrenfeucht and G. Rozenberg. On regularity of languages generated by copying systems. *Disc. Appl. Math.*, 8:313–317, 1984.

[ES71] P. Erdős and J. L. Selfridge. Complete prime subsets of consecutive integers. In R. S. D. Thomas and H. C. Williams, editors, *Proceedings of the Manitoba Conference on Numerical Mathematics*, pages 1–14. Univ. Manitoba Press, 1971.

[Euw29] M. Euwe. Mengentheoretische betrachtungen über das schachspiel. *Proc. Konin. Akad. Wetenschappen, Amsterdam*, 32:633–642, 1929.

[Fej13] L. Fejér. Sur les polynomes harmoniques quelconques. *C. R. Acad. Sci. Paris*, 157:506–509, 1913.

[For04] W. Foryś. Asymptotic behaviour of bi-infinite words. *Theoret. Inform. Appl.*, 38:27–48, 2004.

[Fri00] H. Friedman. Enormous integers in real life. available at `http://www.math.ohio-state.edu/foundations/manuscripts.html`, 2000.

[Fri01] H. Friedman. Long finite sequences. *J. Combin. Theory. Ser. A*, 95:102–144, 2001.

[FW65] N. J. Fine and H. S. Wilf. Uniqueness theorems for periodic functions. *Proc. Amer. Math. Soc.*, 16:109–114, 1965.

[GK86] P. Goralčík and V. Koubek. On discerning words by automata. In L. Kott, editor, *Proc. 13th Int'l Conf. on Automata, Languages, and Programming (ICALP)*, volume 226 of *Lecture Notes in Computer Science*, pages 116–122. Springer-Verlag, 1986.

[GS00] A. D. Gilbert and C. J. Smyth. Zero-mean cosine polynomials which are non-negative for as long as possible. *J. London Math. Soc.*, 62:489–504, 200.

[Hea81] T. Head. Fixed languages and the adult languages of $0L$ schemes. *Internat. J. Comput. Math.*, 10:103–107, 1981.

[HK97] T. Harju and J. Karhumäki. Morphisms. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 1, pages 439–510. Springer-Verlag, 1997.

[HL64] B. R. Heap and M. S. Lynn. The index of primitivity of a non-negative matrix. *Numer. Math.*, 6:120–141, 1964.

[HL66a] B. R. Heap and M. S. Lynn. The structure of powers of nonnegative matrices II. The index of maximum density. *SIAM J. Appl. Math.*, 14:762–777, 1966.

[HL66b] B. R. Heap and M. S. Lynn. The structure of powers of nonnegative matrices I. The index of convergence. *SIAM J. Appl. Math.*, 14:610–639, 1966.

[HL86] T. Head and B. Lando. Fixed and stationary ω-words and ω-languages. In G. Rozenberg and A. Salomaa, editors, *The Book of L*, pages 147–156. Springer-Verlag, 1986.

[HM56] D. Hawkins and W. E. Mientka. On sequences which contain no repetitions. *Math. Student*, 24:185–187, 1956.

[HS99] D. Hamm and J. Shallit. Characterization of finite and one-sided infinite fixed points of morphisms on free monoids. Technical Report CS-99-17, University of Waterloo, July 1999.

[HU79] J. E. Hopcroft and J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, 1979.

[Hur90] L. P. Hurd. Recursive cellular automata invariant sets. *Complex Systems*, 4:119–129, 1990.

[HV58] J. C. Holladay and R. S. Varga. On powers of non-negative matrices. *Proc. Amer. Math. Soc.*, 9:631–634, 1958.

[IOS04] L. Ilie, P. Ochem, and J. Shallit. A generalization of repetition threshold. *submitted*, 2004. Preprint available at `http://arxiv.org/abs/math.CO/0310144`.

[Jus00] J. Justin. On a paper by Castelli, Mignosi, Restivo. *Theoret. Inform. Appl.*, 34:373–377, 2000.

[KS03] J. Karhumäki and J. Shallit. Polynomial versus exponential growth in repetition-free binary words. *To appear, J. Combinatorial Theory Ser. A*, 2003. Preprint available at `http://www.arxiv.org/abs/math.CO/0304095`.

[Lan03] E. Landau. Über die Maximalordnung der Permutationen gegebenen Grades. *Archiv der Mathematik und Physik*, 5:92–103, 1903.

[Lee57] J. Leech. A problem on strings of beads. *Math. Gazette*, 41:277–278, 1957.

[Leh47] D. H. Lehmer. The Tarry-Escott problem. *Scripta Math.*, 13:37–41, 1947.

[LM95] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, 1995.

[Lot83] M. Lothaire. *Combinatorics on Words*, volume 17 of *Encyclopedia of Mathematics and Its Applications*. Addison-Wesley, 1983.

[Lot02] M. Lothaire. *Algebraic Combinatorics on Words*. Cambridge University Press, 2002.

[LS62] R. C. Lyndon and M. P. Schützenberger. The equation $a^M = b^N c^P$ in a free group. *Michigan Math. J.*, 9:289–298, 1962.

[LSW99] J. Loftus, J. Shallit, and M.-w. Wang. New problems of pattern avoidance. In G. Rozenberg and W. Thomas, editors, *DLT '99, Developments in Language Theory*, pages 185–199. World Scientific Press, 1999.

[LV97] M. Li and P. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer-Verlag, 1997.

[MH44] M. Morse and G. A. Hedlund. Unending chess, symbolic dynamics and a problem in semigroups. *Duke Math. J.*, 11:1–7, 1944.

[Mil87] W. Miller. The maximum order of an element of a finite symmetric group. *Amer. Math. Monthly*, 94:497–506, 1987.

[MM62] M. Marcus and F. May. The maximum number of zeros in the powers of an indecomposable matrix. *Duke Math. J.*, 29:581–588, 1962.

[Mor21] M. Morse. Recurrent geodesics on a surface of negative curvature. *Trans. Amer. Math. Soc.*, 22:84–100, 1921.

[MP60] J. Mařík and V. Pták. Norms, spectra, and combinatorial properties of matrices. *Czech. Math. J.*, 10:181–196, 1960.

[MSW01] F. Mignosi, J. Shallit, and M.-w. Wang. Variations on a theorem of Fine & Wilf. In J. Sgall, A. Pultr, and P. Kolman, editors, *Proceedings of MFCS 2001*, volume 2136 of *Lecture Notes in Computer Science*, pages 512–523. Springer-Verlag, 2001.

[NA68] P. S. Novikov and S. I. Adian. Infinite periodic groups, I, II, III. *Izv. Akad. Nauk SSSR. Ser. Mat.*, 32:212–244, 251–524, 709–731, 1968.

[NP82] M. Nivat and D. Perrin. Ensembles reconnaissables de mots biinfinis. In *Proc. Fourteenth Ann. ACM Symp. Theor. Comput.*, pages 47–59. ACM, 1982.

[NP86] M. Nivat and D. Perrin. Ensembles reconnaissables de mots biinfinis. *Canad. J. Math.*, 38:513–537, 1986.

[Pan84] J.-J. Pansiot. A propos d'une conjecture de F. Dejean sur les répétitions dans les mots. *Disc. Appl. Math.*, 7:297–311, 1984.

[Per61] P. Perkins. A theorem on regular matrices. *Pacific J. Math.*, 11:1529–1533, 1961.

[Ple70] P. A. B. Pleasants. Non-repetitive sequences. *Proc. Cambridge Phil. Soc.*, 68:267–274, 1970.

[PP04] D. Perrin and J.-É. Pin. *Infinite Words: Automata, Semigroups, Logic and Games*, volume 141 of *Pure and Applied Mathematics*. Academic Press, 2004.

[Pro51] E. Prouhet. Mémoire sur quelques relations entres les puissances des nombres. *C. R. Acad. Sci. Paris Sér. I*, 33:225, 1851.

[PS58] V. Pták and J. Sedláček. On the index of imprimitivity of nonnegative matrices. *Czech. Math. J.*, 8:496–501, 1958. In Russian. English summary.

[PS76] G. Pólya and G. Szegö. *Problems and Theorems in Analysis II*. Springer-Verlag, 1976.

[Ptá58] V. Pták. On a combinatorial theorem and its application to nonnegative matrices. *Czech. Math. J.*, 8:487–495, 1958. In Russian. English summary.

[Pul64] N. Pullman. On the number of positive entries in the powers of a non-negative matrix. *Canad. Math. Bull.*, 7:525–537, 1964.

[Rob89] J. M. Robson. Separating strings with small automata. *Inform. Process. Lett.*, 30:209–214, 1989.

[Rob96] J. M. Robson. Separating words with machines and groups. *RAIRO Inform. Théor. App.*, 30:81–86, 1996.

[Ros57] D. Rosenblatt. On the graphs and asymptotic forms of finite Boolean relation matrices and stochastic matrices. *Naval Res. Logist. Quart.*, 4:151–167, 1957.

[Rot89] P. Roth. A note on word chains and regular languages. *Inform. Process. Lett.*, 30:15–18, 1989.

[RR85] A. Restivo and C. Reutenauer. Rational languages and the Burnside problem. *Theoret. Comput. Sci.*, 40:13–30, 1985.

[RSW03] N. Rampersad, J. Shallit, and M.-.w Wang. Avoiding large squares in infinite binary words. In Tero Harju and Juhani Karhumäki, editors, *WORDS'03, Proc. 4th International Conference on Combinatorics on Words*, volume 27 of *TUCS General Publication*, pages 185–197. Turku Centre for Computer Science, 2003.

[Sal73] A. Salomaa. *Formal Languages*. Academic Press, 1973.

[SB96] J. Shallit and Y. Breitbart. Automaticity I: Properties of a measure of descriptional complexity. *J. Comput. System Sci.*, 53:10–25, 1996.

[Séé85] P. Séébold. Sequences generated by infinitely iterated morphisms. *Disc. Appl. Math.*, 11:255–264, 1985.

[ST03] R. J. Simpson and R. Tijdeman. Multi-dimensional versions of a theorem of Fine and Wilf and a formula of Sylvester. *Proc. Amer. Math. Soc.*, 131:1661–1671, 2003.

[Sun04a] Z. W. Sun. Arithmetic properties of periodic maps. *Math. Res. Lett.*, 11:187–196, 2004.

[Sun04b] Z. W. Sun. A local-global theorem on periodic maps. *submitted*, 2004. Preprint available at `http://front.math.ucdavis.edu/math.NT/0404137`.

[SW99] J. Shallit and M.-w. Wang. An inequality for non-negative matrices. *Linear Algebra and Its Applications*, 290:135–144, 1999.

[SW01a] J. Shallit and M.-.w Wang. Automatic complexity of strings. *J. Automata, Languages, and Combinatorics*, 6:537–554, 2001.

[SW01b] J. Shallit and M.-w. Wang. Weakly self-avoiding words and a construction of Friedman. *Electronic J. Combinatorics*, 8, 2001. available at `http://www.combinatorics.org/Volume_8/Abstracts/v8i1n2.html`.

[SW02] J. Shallit and M.-w. Wang. On infinite two-sided fixed points of morphisms. *Theoret. Comput. Sci.*, 270:659–675, 2002.

[Thu06] A. Thue. Über unendliche zeichenreihen. *Norske vid. Selsk. Skr. I. Mat. Nat. Kl.*, 7:1–22, 1906.

[Thu12] A. Thue. Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. *Norske vid. Selsk. Skr. I. Mat. Nat. Kl.*, 10:1–67, 1912. Reprinted in *Selected Mathematical Papers of Axel Thue*, T. Nagell, editor, Universitetsforlaget, Oslo, 1977, pp. 413–478.

[TZ03] R. Tijdeman and L. Zamboni. Fine and Wilf words for any periods. *Indag. Math.*, 14:135–147, 2003.

[VTSS90] D. L. Van, D. G. Thomas, K. G. Subramanian, and R. Siromoney. Bi-infinitary codes. *RAIRO Inform. Théor. App.*, 24:67–87, 1990.

[Wan99] M.-w. Wang. Subword complexity and a matrix inequality. Master's thesis, University of Waterloo, 1999.

[Wan00] M.-w. Wang. On the irregularity of the duplication closure. *Bull. European Assoc. Theor. Comput. Sci.*, 70:162–163, 2000.

[Wan02] M.-w. Wang. An inequality for non-negative matrices II. *Linear Algebra and Its Applications*, 348:259–264, 2002.

[Wie50] H. Wielandt. Unzerlegbare, nicht negative Matrizen. *Math. Zeitschift*, 52:642–648, 1950.

[Zec58] T. Zech. Wiederholungsfreie Folgen. *Z. Angew. Math. Mech.*, 38:206–209, 1958.