Coincidences of Substitution Systems

Ian Hoffman
Advisor: Michael S. Keane

A Dissertation in Mathematics submitted to the faculty of Wesleyan University in partial fulfillment of the requirments for the degree of Master of Arts

Wesleyan University

Middletown, Connecticut

Spring, 2013

Acknowledgements

I would first like to thank Professor Michael Keane for being my graduate advisor and thesis tutor for the past two semesters. I had expressed an interest in Wesleyan's BA/MA program in Mathematics since the beginning of my junior year but had problems finding a professor to advise my research. Professor Keane took a chance on me to be his final graduate student prior to his retirement, and I took a chance on him advising me even though we never knew each other prior to my senior year. Please enjoy a long and healthy retirement.

Second, a thanks to department friend and past Visiting Professor Michel Dekking, a longtime collaborator with Professor Keane. I became acquainted with him during the fall 2011 semester, when he visited Wesleyan to teach and collaborate. Along with Professor Keane, he helped introduce me to some of his observations on substitution theory and I hope I have done a good job applying them to this thesis. *Dank je wel.*

Thanks as well to Professor Carol Wood, my undergraduate math advisor. She always kept me on the ball with suggesting various courses to take, and her decisions have helped me well all the way through to the end of this fifth year. As with Professor Keane, I wish her a long and healthy retirement.

Thanks to Wesleyan's diverse course programs, allowing me not only the opportunity to take a year of graduate school and work on a master's degree, but also offering a summer program back in 2010 in which I learned SuperCollider. To that, I thank Professors Ron Kuivila and James Lipton for taking their time out to teach during that summer and continuing to offer me programming help in the six semesters afterward.

Finally, thanks to all of the other math and music professors I've had in the last five years for getting me this far, and for my family and friends that have supported me and shown interest in my academic career at Wesleyan. It's been a great experience.

Contents

Introduction

Substitution is a branch of ergodic theory and dynamics where given a set of symbols, each symbol is mapped to an equal or preferably longer string of symbols within that set. Early theories were formulated in the early 20th century by Morse [8] and advanced by Dekking [1], Keane [5], Michel [6], et alii.

An important study of substitution is coincidence density, where various sequences, often under the same map of substitutions, are examined to determine how often equivalent symbols appear as the lengths of the sequences approach infinity. Michel's 1978 thesis [7] is the groundwork for these results, expanding upon his and others' theorems.

The symbol sets themselves can be of any finite size, though like many things in mathematics, it is easier to handle smaller sets first. In this case, the focus begins with the binary set $\{0, 1\}$, examining a sequence beginning with 0 and a sequence beginning with 1 under the same substitution map. The initial results are used to examine other types of coincidences of the $\{0, 1\}$ set between sequences of distinct substitutions. Finally, these results are applied towards q-ary sets where $q \geq 3$, and other hypotheses for general binary and q-ary cases are developed.

Chapter 1: Definitions of Substitution and Coincidence

Let I be a finite set. We call I an *alphabet* and the elements of I are called *symbols* or *letters*. The elements of $I^* = \bigcup_{n\geq 0} I^n$ are called *blocks* or *words*. If a block $b = b_1 b_2 \ldots b_n \in I^n$, then $|b| = n$ is the *length* of b. The *empty word* is notated as $\emptyset$ and has length 0. (Hence, $I^0 = \{\emptyset\}$.)

A *substitution* θ over the alphabet I is a map from I to $\bigcup_{n>1} I^n$. One basic substitution would be to send 0 to 01, and 1 to 110. A special substitution ϕ, called the *Fibonacci substitution*, sends 0 to 01 and 1 to 0: given $S_0 = 0$, its first several terms are $S_0 = 0$, $S_1 = 01$, $S_2 = 010 = S_1 S_0$, and $S_3 = 01001 = S_2 S_1$. We can see that for any $n \geq 2$, we will have $S_n = S_{n-1} S_{n-2}$ (hence the name). Unless stated otherwise, we will work in *binary substitutions* of the alphabet $I = \{0, 1\}$.

For each letter $i \in I$, $\theta(i)$ is a block $a_1^i a_2^i \ldots a_{L_i}^i \in I^{L_i}$, where $L_i = |\theta(i)|$. We say that θ is of *constant length* if $L_i = L_j$ for all $i, j \in I$; otherwise, θ is of *non-constant length*. Given a binary substitution θ, we denote the *substitution matrix* M_θ as a 2 x 2 matrix

$$\begin{bmatrix} x_0 & y_0 \\ x_1 & y_1 \end{bmatrix}$$

where $x_0 = card\{k : a_k^0 = 0\}$, $y_0 = card\{k : a_k^0 = 1\}$, $x_1 = card\{k : a_k^1 = 0\}$, and $y_1 = card\{k : a_k^1 = 1\}$, all of which are integers ≥ 0. Using our example above where $\theta(0) = 01$ and $\theta(1) = 110$, we have that $|\theta(0)| = 2$, $|\theta(1)| = 3$, and so θ is a non-constant length substitution with matrix

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

Clearly, for any $\theta(0) \in I^{L_0}$ and $\theta(1) \in I^{L_1}$, $x_0 + y_0 = L_0$ and $x_1 + y_1 = L_1$.

For an arbitrary block $b = b_1 b_2 \ldots b_n \in I^n$, we define $\theta(b) = \theta(b_1)\theta(b_2) \ldots \theta(b_n)$. Similarly, for any $i \in I$, we can use induction to define $\theta^2(i) = \theta(\theta(i)) = \theta(a_1^i) \ldots \theta(a_{L_i}^i)$, and so on. We may also define $\theta^0(i) = i$. Thus, for all $s \geq 1$, $\theta^{s+1}(i) = \theta(\theta^s(i))$.

Using our above θ, we have that $\theta^2(0) = \theta(0)\theta(1) = 01110$, $\theta^3(0) = \theta(\theta(01)) = 0111011001$, etc., and $\theta^2(1) = \theta(1)\theta(1)\theta(0) = 11011001$, $\theta^3(1) = 1101100111011001\ 01110$, etc. We can also shorten a string of n consecutive copies of the block $i_1 \ldots i_m$ as $(i_1 \ldots i_m)^n$, with parentheses optional if the string is a single symbol. For example, $010101 = (01)^3$ and $1111 = (1)^4$ or just 1^4.

If for all i, $\theta(i)$ begins with i, we define an infinite sequence $\theta^\infty(i)$ as the unique sequence that begins with $\theta^n(i)$ for all $n \geq 0$. We can also write $\theta^\infty(i)$ as $t^i = t_1^i t_2^i t_3^i \ldots$. For a given t^i, we denote the *density limit* of 0, if it exists, as

$$d_{t^i}(0) = \lim_{n\to\infty} \frac{card\{k \leq n : t_k^i = 0\}}{n},$$

with the respective change in the numerator made for 1.

In a binary alphabet, a pair of blocks $b = b_1 \dots b_m$, $c = c_1 \dots c_n$ is *balanced* if $m = n$ and $card\{k : b_k = 0\} = card\{k : c_k = 0\}$.

A balanced pair of blocks of length m is *minimal* if for all M such that $1 \leq M < m$, the blocks $b_1 \dots b_M$, $c_1 \dots c_M$ are not balanced. If b and c are balanced, we may write $b \sim c$. Working with our previous $\theta(0) = 01$, $\theta(1) = 110$ example, if we write out the first sixteen symbols of t^0 and t^1 we get

0111011011001110
1101100111011001

which is balanced; each block has ten 0's and six 1's. However, it's not minimal, for if we look at the first three entries in each string we see

011
110

which is a balanced pair of length three. If for sequences b and c there exists two minimal balanced pair of blocks

$$\begin{pmatrix} b_j \dots b_{j+m} \\ c_j \dots c_{j+m} \end{pmatrix} \text{and} \begin{pmatrix} b_k \dots b_{k+m} \\ c_k \dots c_{k+m} \end{pmatrix}$$

where $m \geq 2$, $(j+m) < k$, $b_j = c_k, \dots, b_{j+m} = c_{k+m}$, and $b_k = c_j, \dots, b_{k+m} = c_{j+m}$, such as

$$\begin{pmatrix} 011 \\ 110 \end{pmatrix} \text{and} \begin{pmatrix} 110 \\ 011 \end{pmatrix}$$

then we may let the two blocks be similar to each other. These pairs of blocks can be seen in the first three and last three indices of our two above sequences respectively, with the ten indices in between yielding a balance between the sequences as well.

Now suppose θ is a binary substitution whose matrix M_θ is positive. The Perron–Frobenius theorem tells us that M_θ has two distinct eigenvalues λ_1, λ_2. It also says that λ_1 is positive, $\lambda_1 > |\lambda_2|$ and that the right and left unit eigenvectors corresponding to λ_1 are positive. Suppose M_θ has eigenvalues $\lambda_1 > \lambda_2 = 1$. Let $w = [\alpha_0\, \alpha_1]^T$ be a right eigenvector corresponding to λ_1 (i.e., a Perron–Frobenius right eigenvector). We say that a pair of blocks $b = b_1 \dots b_m$, $c = c_1 \dots c_n$ is *weighted balanced* or *w-balanced* if

$$(card\{i : b_i = 0\} \ * \ \alpha_0) + (card\{j : b_j = 1\} \ * \ \alpha_1) =$$

$$(card\{i : c_i = 0\} \ * \ \alpha_0) + (card\{j : c_j = 1\} \ * \ \alpha_1).$$

While it's implicit that any pair of balanced blocks is also w-balanced, here it is not necessary that $m = n$. It is also not necessary that if b is a block of t^0, c is a block of t^1, and they are w-balanced, that b and c start at the same positions in their respective sequences. I shall use a new substitution, η, for a demonstration of weighted balance. Let $\eta(0) = 001$, $\eta(1) = 10110$; then

$$M_\eta = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$$

and $\lambda_1 = 4$, $\lambda_2 = 1$, $w = [1\,2]^T$. We could partition the beginning of $\eta^\infty(0)$ and $\eta^\infty(1)$ into the minimal balanced blocks

$$\begin{array}{lllll} 0010011 & 011 & 0001001101101 & 01 & 10 \\ 1011000 & 110 & 1101011000100 & 10 & 01 \end{array}.$$

The first pair of blocks is of length seven and the third is of length thirteen. A conjecture by P. Michel suggests that continuing this process farther will yield a subsequence of distinct pairs of minimal blocks whose lengths diverge [7]! Instead, by giving the 1's twice the weight of the 0's and looking for sums of equal weight from the left, as w suggests, our partitions of minimal w-balanced blocks will be consistently finer:

$$\begin{array}{cccccccccccccc} 00 & 10 & 01 & 1 & 011 & 00 & 01 & 00 & 110 & 1 & 10 & 1 & 01 & 10 \\ 1 & 01 & 10 & 00 & 110 & 1 & 10 & 1 & 011 & 00 & 01 & 00 & 10 & 01 \end{array}$$

If $b = w_j^0 w_{j+1}^0 \dots w_{j+|b|}^0$ and $c = w_k^1 w_{k+1}^1 \dots w_{k+|c|}^1$ are a pair of w-balanced blocks, then the *left slope* of the block is $j - k$; likewise, the *right slope* is $(j + |b|) - (k + |c|)$. If $|b| = |c|$ then we just call $j - k$ the *slope* of the block. We will explore w-balanced blocks and the substitutions of the above matrix M_η more in Chapters II and III.

A *coincidence* between two finite or infinite sequences $v^1 = v_1^1 v_2^1 v_3^1 \dots$, $v^2 = v_1^2 v_2^2 v_3^2 \dots$ occurs when for some $n \geq 1$ we have $v_n^1 = v_n^2$. We may also define a coincidence in the same manner for a pair of blocks within the sequences. For infinite sequences we call

$$\delta = \lim_{N\to\infty} \frac{card\{n : n \leq N \text{ and } v_n^1 = v_n^2\}}{N},$$

whenever this limit exists, the *coincidence density* or *coincidence limit* of v_1 and v_2.

If θ is a substitution with sequences $\theta^\infty(0) = t^0$, $\theta^\infty(1) = t^1$, then θ is a *regular substitution* if the coincidence limit of t^0 and t^1 exists. We can notate this as $\delta(t^0, t^1)$ or just $\delta(\theta)$. If the limit does not exist, we shall call θ an *irregular substitution.*

Finally, let θ, ι be two distinct substitutions where $M_\theta = M_\iota$. We can say $\theta(i)$ and $\iota(i)$ are n *degrees apart*, notated $T(i) = n$, if we can transpose a mininum of n pairs of symbols in $\theta(i)$ to obtain $\iota(i)$. We can extend this definition to the whole of θ and ι by the sum $T(I) = \sum_{i\in I} T(i)$ and say θ and ι are *T(I) degrees apart.*

Example: Let $\theta : 0 \to 001, 1 \to 1101$ and $\iota : 0 \to 010, 1 \to 1101$. By transposing the second 0 and the 1 in $\theta(0)$ we get $\iota(0)$, so $T(0) = 1$. Since $\theta(1) = \iota(1)$ we have

$T(1) = 0$, thus $T(I) = 1$, and so θ and ι are one degree apart. If we changed $\iota(1)$ to 1011, then θ and ι would be two degrees apart.

Unless stated otherwise, we shall assume from here on:

a) All substitutions in this paper are of non-constant length.

b) For all $i \in I$, $|\theta(i)| \geq 2$.

c) For all $i \in I$ and substitutions θ, if $\theta(i) = a_1^i a_2^i \ldots a_{L_i}^i$ then $a_1^i = i$, and $a_m^i \neq i$ for at least one m such that $2 \leq m \leq L_i$.

We assume a) because P. Michel has proven that for binary substitutions θ of constant length, $\delta(\theta)$ always exists and is equal to 0 or 1, depending on if $\theta(0)$ and $\theta(1)$ have at least one coincidence or no coincidences respectively [7]. Therefore, the only "interesting" cases of binary substitutions will occur when the lengths are non-constant. We assume b) because if θ were a substitution such that, without loss of generality, $\theta(0) = 0$, then $\theta^n(0) = 0$ for all n, and therefore we cannot obtain a sequence of infinite length this way. The Fibonacci substitution is able to avert this because of how $\phi(1) \neq 1$ and $|\phi(0)| > 1$; thus $|\phi^{n+1}(1)| > |\phi^n(1)|$ for $n > 1$. Finally, we assume c) to ensure a "two-sided" quality to t^0 and t^1, where $\theta^n(0)$ begins with 0 and $\theta^n(1)$ begins with 1 for all n, and none of the sequences consist entirely of one symbol.

Much of the information about the behavior of substitutions and their coincidences originates from the substitution matrices. P. Michel proved that every non-constant length substitution whose matrix has eigenvalues $\lambda_1 > 1$, $\lambda_2 = 0$ is regular and has a rational coincidence limit [7]. In addition, for any such substitution that satisfies $\lambda_1 > 1$, $\lambda_2 = 0$, the set of minimal balanced pairs of blocks $\hat{I}$ is finite and nonempty. However, any other pair of feasible eigenvalues does not guarantee a result towards the regularity of substitutions whose matrices have such eigenvalues. The conjecture is that most of these substitutions are irregular, and for each such irregular substitution, its set of minimal balanced pairs of blocks is either empty or infinite.

One goal of this thesis is to explore such behaviors of substitutions where both eigenvalues are positive. We would like to know if there exist matrices of this type where every substitution sharing this matrix is irregular. We would also like to find out if every matrix of this type has at least one regular substitution (whether its coincidence limit is 0, 1, or somewhere in between).

Chapter 2: Some Previously Established Results

Let θ be a substitution with matrix M_θ and respective eigenvalues λ_1, λ_2. In the case of $\lambda_1 > 1$, $\lambda_2 = 0$, P. Michel shows that θ possesses a non-empty, finite set of minimal balanced pairs of blocks $\hat{I} = \{\hat{i}_1, \hat{i}_2, \ldots, \hat{i}_n\}$. If we order the $\hat{i}_j$'s in $\hat{I}$ by their first apppearance from the left, then

$$\begin{pmatrix} t^0 \\ t^1 \end{pmatrix} = \tilde{\theta}^\infty(\hat{i}_1) = \hat{t}_1\hat{t}_2\hat{t}_3\ldots.$$

We can also shorten $\tilde{\theta}^\infty(\hat{i}_1)$ to just $\hat{\theta}$. The coincidence density of t^0 and t^1 is given by the formula

$$\delta(\theta) = \frac{\sum_{j=1}^{n} c(\hat{i}_j)d(\hat{i}_j)}{\sum_{j=1}^{n} l(\hat{i}_j)d(\hat{i}_j)}$$

where $c(\hat{i}_j)$ is the number of coincidences of the pair $\hat{i}_j$, $l(\hat{i}_j)$ is the length of $\hat{i}_j$, and $d(\hat{i}_j)$ is the density limit of $\hat{i}_j$ in

$$\begin{pmatrix} t^0 \\ t^1 \end{pmatrix},$$

the rate at which $\hat{i}_j$ appears in the sequence $\hat{\theta}$, calculated by

$$d(\hat{i}_j) = \lim_{n\to\infty} \frac{card\{k \leq n : \hat{t}_k = \hat{i}_j\}}{n}.$$

If $\hat{i}_1$, $\hat{i}_2$, $\ldots$, $\hat{i}_n$ are the balanced pairs of θ, we may form an $n \times n$ substitution matrix $M_{\tilde{\theta}} = [\tilde{\theta}_{i,j}]$ where $\tilde{\theta}$ is the substitution θ applied to these pairs. In fact, if we let $w_L = [\beta_1 \ldots \beta_n]$ be a Perron–Frobenius left eigenvector for $M_{\tilde{\theta}}$, then $d(\hat{i}_k) = \beta_k/(\sum_{m=1}^{n} \beta_m)$ [7].

P. Michel also shows that if $\hat{I}$ is finite and the balanced pair

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \hat{0} \in \hat{I},$$

then $\delta(\theta) = 1$, and $\hat{0} \in \hat{I}$ iff the respective length one block $\hat{1} \in \hat{I}$ [7].

Example 1 (P. Michel). Let M_θ be the matrix

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix};$$

thus, there exist three possible substitutions for θ. Since its determinant is zero and eigenvalues are 3 and 0, we know all three substitutions are regular. The three are as follows:

$$\theta_1 : 0 \to 01, 1 \to 1100, \quad \theta_2 : 0 \to 01, 1 \to 1010, \quad \theta_3 : 0 \to 01, 1 \to 1001$$

In the case of θ_1, we have a set of four classes of balanced blocks

$$\begin{aligned} \hat{I} &= \left\{ \begin{pmatrix} 011 \\ 110 \end{pmatrix}, \begin{pmatrix} 10 \\ 01 \end{pmatrix}, \begin{pmatrix} 100 \\ 001 \end{pmatrix}, \begin{pmatrix} 1100 \\ 0101 \end{pmatrix} \right\} \\ &= \{\hat{a}, \hat{b}, \hat{c}, \hat{d}\} \end{aligned}$$

with substitutions $\hat{a} \to \hat{a}\hat{b}\hat{b}\hat{c}$, $\hat{b} \to \hat{a}\hat{c}$, $\hat{c} \to \hat{d}\hat{d}$, $\hat{d} \to \hat{a}\hat{b}\hat{c}\hat{d}$, giving us the 4 x 4 matrix

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

It turns out that $d(\hat{i}_j) = 1/4$ for all j. Therefore the coincidence density of θ_1 is

$$\frac{(1+0+1+1)(1/4)}{(3+2+3+4)(1/4)} = \frac{3}{9} = \frac{1}{3}.$$

For θ_2 and θ_3, we have three equivalence classes of balanced blocks

$$\hat{I} = \left\{ \begin{pmatrix} 01 \\ 10 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

but since we have $\hat{0} \in \hat{I}$, we conclude that $\delta(\theta_i) = 1$ for $i = 2, 3$. These three substitutions show us a very simple case of how different regular substitutions whose matrices are equal can still have different coincidence limits, even though each pair of substitutions is only one symbol transposition apart.

Before we move on towards cases of substitutions whose matrices have two positive eigenvalues, we shall introduce a pair of special w-balanced blocks and a theorem by Dekking. The blocks

$$\begin{pmatrix} & 0 \\ 0 & \end{pmatrix}, \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$$

with slope 1 are called the *skewed* or *sloped unit blocks* and are denoted by $\hat{0}_\natural$ and $\hat{1}_\natural$ respectively. Note that the offsetting of the 0's and 1's is intentional to emphasize inequality of index. This leads to:

Theorem 1 (Dekking) [2]. Let θ be a substitution with $\theta(0) = 0\ldots$, $\theta(1) = 1\ldots$, and $\lambda_1 > \lambda_2 = 1$ with respect to M_θ. If θ possesses a finite set of w-balanced blocks which contains both $\hat{0}$ and $\hat{0}_\natural$, then $\delta(\theta)$ does not exist.

This theorem induces a series of results concerning the matrix

$$\begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}.$$

Example 3. Given our assumptions from Chapter 1, there exist twelve feasible substitutions θ where the above matrix suffices. Ten of them are irregular. The two regular substitutions are $\theta_1 : 0 \to 010$, $1 \to 10101$ with limit 0, and $\theta_2 : 0 \to 001$, $1 \to 11001$ with limit $1/2$. The limit of θ_2 was originally calculated by Dekking and has been discussed in lecture [3], but as far as this author knows, he has not formally published a proof.

For θ_1, the result is rather obvious, as t^0 is the periodic sequence $01010101\ldots$ and t^1 is the periodic sequence $10101010\ldots$. Since the blocks 01 and 10 never coincide, we conclude t^0 and t^1 never coincide. Hence, $\delta(\theta_1) = 0$. $\delta(\theta_2)$ is a more subtle case and is best demonstrated by the following:

Theorem 2. The coincidence limit of t^0 and t^1, where $\theta(0) = 001$ and $\theta(1) = 11001$, is $1/2$.

Proof. For t^0 and t^1, $\hat{I}$ is empty; consider the beginnings of these sequences.

$$\begin{aligned} t^0 &= 0010011100100100111001110011100100100111001\ldots \\ t^1 &= 11001110010010011100111001110010010011100 1\ldots \end{aligned}$$

We observe that for $k \geq 4$, $t^0_k = t^1_k$ iff $t^0_k = t^0_{k+1}$ (and equivalently, $t^1_k = t^1_{k-1}$). Therefore, if $\delta(\theta_2)$ exists, we can express

$$\delta(\theta_2) = \lim_{n\to\infty} \frac{card\{k : k \leq n \text{ and } t^0_k = t^0_{k+1}\}}{n}.$$

Given the eigenvalues of M_θ we can use a Perron–Frobenius left eigenvector to find that $d_{t^0}(0) = 1/2 = d_{t^0}(1)$. We can concatenate strings of 0's and 1's to show that t^0 can be rewritten in terms of the blocks $x = 00$, $y = 1$, and $z = 111$. Thus we can rewrite t^0 as a ternary sequence $h^x = h^x_1 h^x_2 h^x_3 \ldots$, where $h^x_1 = x$ and $h^x = \eta^\infty(x)$, where η is the ternary substitution

$$x \to xyxy, \quad y \to yyxy, \quad z \to yyxzxzxy$$

Thus $h^x = xyxzxyxyxzxzxzxyxyxzxy\ldots$. We can rewrite h^x as a binary sequence of xy's and xz's via pairing the blocks $h^x_1 h^x_2$, $h^x_3 h^x_4$, $\ldots$, and we get $\eta(xy) = xyxzxy$, $\eta(xz) = xyxzxzxzxy$. This leads to the following

$$\textit{Claim.}\ \delta(\theta_2) = \frac{\sum_{i=x,y,z}(|i|-1)(d_{h^x}(i))}{\sum_{i=x,y,z}(|i|)(d_{h^x}(i))}$$

Clearly, $d_{h^x}(x) = 1/2$. Now we look at the half of h^x comprised of the y's and z's and form a binary substitution $\tilde{\eta}$, only dependent on the y's and z's. When we remove the x's from h^x we have that $\eta(xy) = xyxzxy$ "reduces" to the substitution $\tilde{\eta}(y) = yzy$, and $\eta(xz) = xyxzxzxzxy$ "reduces" to $\tilde{\eta}(z) = yzzzy$. Even though $\tilde{\eta}(z)$ does not begin with a z, we have that $M_{\tilde{\eta}} = M_\theta$ and so we can conclude $d_{h^x}(y) = 1/4 = d_{h^x}(z)$.

Proof of Claim. Because $t^0_k = t^0_{k+1}$ is our criterion for coincidence, and every 00 block is followed by a block of 1 or 111, we have that only the first 0 of each 00 in t^0 coincides with a 0 in t^1. Likewise, singleton 1s will never coincide because they are always followed by a 00, and for 111s, the first two 1s will coincide, but the third will not as the 111 will be followed a 00. Thus the equation proposed in the claim suffices, and

$$\delta(\theta_2) = \frac{(2-1)(1/2) + (1-1)(1/4) + (3-1)(1/4)}{(2)(1/2) + (1)(1/4) + (3)(1/4)} = \frac{1/2 + 0 + 1/2}{1 + 1/4 + 3/4} = \frac{1}{2}$$

which satisfies M. Dekking's calculation.

Corollary 2.1. Let θ be a substitution such that there exists M, $k \in \mathbf{N}$ such that for all $m \geq M$, $t^0_m = t^1_m$ iff $t^0_m = t^0_{m+k}$ or $t^1_m = t^1_{m+k}$. Suppose t^0 or t^1 can be rewritten in terms of distinct length blocks of 0's $\hat{0}_1, \hat{0}_2, \ldots, \hat{0}_x \in B_0$ and 1's $\hat{1}_1, \hat{1}_2, \ldots, \hat{1}_y \in B_1$ such that $|\hat{i}_j| \geq k$ for $i = 0, 1$ and $j = 1, \ldots, |B_i|$. If $t^0_m = t^0_{m+k}$ is the criterion for coincidence, then

$$\delta(\theta) = \frac{\sum_{j=1}^{|B_i|} \sum_{i=0}^{1} (|\hat{i}_j| - k)(d_{t^0}(\hat{i}_j))}{\sum_{j=1}^{|B_i|} \sum_{i=0}^{1} (|\hat{i}_j|)(d_{t^0}(\hat{i}_j))}$$

with the respective changes to d_{t^0} made if the criterion for coincidence is $t^1_m = t^1_{m+k}$.

Corollary 2.2. If θ is a binary substitution that satisfies the conditions of Corollary 2.1 and $|B_i| = 1$ for some i, then

$$\delta(\theta) = \frac{|\hat{i}_1| - k}{|\hat{i}_1|}$$

for the i such that $|B_i| = 1$.

Now, let θ_3 be the substitution $\theta_3(0) = 010$, $\theta_3(1) = 11010$. Note that the difference between θ_2 and θ_3 is a transposition of the last 0 and the last 1 for each substitution. Our partition of $\hat{\theta}$ into pairs of minimal w-balanced blocks begins as follows:

010	11	010	0	1	0	11	010	11	0	1	0
11	010	11	0	1	0	010	11	010	0	1	0

We see the blocks $\hat{0}$ (first appearance is the tenth block), $\hat{1}$ (eleventh), $\hat{0}_\natural$ (fourth), and $\hat{1}_\natural$ (fifth) all appear early in the partition. By Theorem 1, we conclude that θ_3 is an irregular substitution.

Even though $\delta(\theta_3)$ doesn't exist, there is something that we can conjecture about the behavior of coincidences based on the coincidence densities at particular points. A SuperCollider program [4] was used to count the coincidences of the first 10000[1] substitutions' worth of indices for $\hat{\theta}$, and then to calculate the coincidence densities at each of those indices and plot a line graph of the densities. (The accompanying code will be shown in Appendix A.) From the graph we observe that if a local maximum (minimum) for the coincidence density of θ_3 exists at the integer z, then the next successive local maximum (minimum) occurs approximately at the index $4z$. In the case of θ_3, several local maxima within the first 40000 indices occur at

Order	Index	Density
1	109	0.62727...
2	344	≈ 0.6261
3	1368	≈ 0.6253
4	5464	≈ 0.6251
5	21848	≈ 0.6250

Note that 344(4) - 8 = 1368, 1368(4) - 8 = 5464, and 5464(4) - 8 = 21848. Likewise, the table for the occurrences of several local minima in the same range (ignoring early deviations such as the density being 0 at the first index) is:

Order	Index	Density
1	169	0.5
2	681	0.5
3	2729	0.5
4	10921	0.5

Note that the first several minima remain constant at 1/2 for this substitution. We also see 169(4) + 5 = 681, 681(4) + 5 = 2729, and 2729(4) + 5 = 10921. There may be a connection between this observation and how the coincidence limit of θ_2 is 1/2, and we can certainly make a conjecture that $\liminf_{n\to\infty} c_n = 1/2$. However, if we try a substitution θ_4 where $0 \to 001$, $1 \to 11100$, our maxima and minima charts (within the first 40000 indices) are

Order	Index	Density
1	143	0.5625
2	677	≈ 0.5501
3	2703	≈ 0.5566
4	10917	≈ 0.5617

[1] 10000 substitutions was used as a sufficient round number for this substitution, which yields approximately 40000 indices for any substitution of equal matrix. By then several oscillations of coincidence density are present, and thus it is easy to grasp the behavior of the substitution. Recent SuperCollider builds have become faster in computation (see Appendix A), so 10000 may be substituted for a number many times larger if the programmer wishes.

and

Order	Index	Density
1	261	≈ 0.4809
2	1014	≈ 0.4887
3	4101	≈ 0.4993
4	16376	≈ 0.5072

respectively. We can see that although consecutive maxima and minima still occur at approximately 4:1 ratios, in general, the density values are slowly creeping up (consider the increase from the second to third maxima, then the third to the fourth). Perhaps after θ_4 is extended far enough we see very small differences between a local maximum and the next local minimum in sequence, but because of Theorem 1, $\liminf_{n\to\infty} \delta(\theta_4) \neq \limsup_{n\to\infty} \delta(\theta_4)$.

Overall, we conjecture that if θ is an irregular substitution where $\lambda_1 > \lambda_2 = 1$ are the eigenvalues for M_θ, and n is the index of a local maximum (minimum) for the coincidence density of t^0 and t^1, then the next local maximum (minimum) will exist approximately at the index $n\lambda_1$.

Chapter 3: New Results and Alternative Approaches

In this chapter, I would like to introduce some new results and conjectures regarding various substitutions, and also present different ways of thinking about substitutions. We shall start with a basic theorem which has one case that was already covered in Chapter 2:

Theorem 3. Fix integers k, $m \geq 1$. Let Θ be the set of all substitutions θ such that

$$M_\theta = \begin{bmatrix} k+1 & k \\ m & m+1 \end{bmatrix}$$

for some k, $m \geq 1 \in \mathbf{Z}$. Then there exists at least one regular $\theta \in \Theta$.

Proof. Let $\theta(0) = (01)^k 0$, $\theta(1) = (10)^m 1$. Then $t^0 = 010101\ldots$, $t^1 = 101010\ldots$, and $\delta(\theta) = 0$.

The case from Chapter 2 in question was $\theta : 0 \to 010$, $1 \to 10101$. This gives way to a related theorem regarding periodic blocks of length 3; an accompanying conjecture/open problem will be discussed in Chapter 4.

Theorem 4. If θ is a binary substitution where $\theta(0)$ begins with 0 and $\theta(1)$ begins with 1, then neither t^0 nor t^1 can be a periodic sequence with a period of length three.

Proof. We shall prove these by contradiction. Since each periodic block of length 3 that begins with a 0 corresponds to a block beginning with a 1 by adding 1 to each symbol and reducing modulo 2 ($001 \to 110$, $010 \to 101$, $011 \to 100$), we only need to check three cases.

Case 1: Suppose θ is a substitution such that $t^0 = 001001\ldots$, with our assumptions from Chapter 1. The only way we can partition t^0 from the left and ensure the first two blocks are equal (corresponding to $\theta(00)$) is such that $|\theta(0)| = 3x$ for $x \geq 1$. By doing so, we force $\theta(1)$ to begin with a 0, which is a contradiction.

Case 2: Suppose θ is a substitution such that $t^0 = 010010\ldots$, thus $|\theta(0)| = 3y+1$, $|\theta(1)| = 3z+2$ for $y \geq 1$, $z \geq 0 \in \mathbf{Z}$, in order to preserve consistency of the $\theta(0)$ partitions. Thus we have $|\theta(010)| = (3y+1) + (3z+2) + (3y+1) = 3(2y+z) + 4$, which is not divisible by 3. In fact, the fourth partition (the third $\theta(0)$ block) begins with a 1, which immediately contradicts $\theta(0)$ in its own right. $\Rightarrow\Leftarrow$

Case 3: Finally, suppose θ is a substitution such that $t^0 = 011011\ldots$. Here, $|\theta(0)| = 3x+1$ or $3x+2$ for $x \geq 1 \in \mathbf{Z}$. So suppose $|\theta(0)| = 3x+1$. To preserve consistency of the $\theta(1)$ partitions, we know that 3 must divide $|\theta(1)|$. But this implies $|\theta(011)| = (3x+1) + 2(3y) = 3(x+2y) + 1$, which is not divisible by 3. The $|\theta(0)| = 3x+2$ case unfolds in the same manner. $\Rightarrow\Leftarrow$ Thus, no substitution sequence with a period of length 3 exists.

Theorem 5. Fix an integer $k \geq 1$ and let Θ be the set of all substitutions θ such that

$$M_\theta = \begin{bmatrix} k & 1 \\ k(k-1) & k+1 \end{bmatrix}.$$

Then there exists $\theta \in \Theta$ such that for all $n \geq (k+1)$, $t_n^0 = t_n^1$ iff $t_n^0 = t_{n+k-1}^0$.

Proof. Let θ be the substitution $0 \to 0^k 1$, $1 \to 11(\theta(0))^{k-1}$. By setting $a = 0^k$, $b = 1$, we find that θ behaves much like the θ_2 from Example 3 of Chapter 2, with $\eta(a) = (ab)^k$ and $\eta(b) = bb(ab)^{k-1}$, a constant length substitution of length $2k$ where the last $2k-1$ symbols of $\eta(a)$ and $\eta(b)$ match. As a result, every symbol of the $\{a,\, b\}$ substitution from the second index onward matches. Thus, the 1 at t_{k+1}^0 corresponds to the 1 at t_2^1, and so on. Therefore, the slope of this correspondence is $(k+1) - 2 = k - 1$. Thus if $t_n^0 = t_n^1$ when $n \geq k+1$, then $t_n^0 = t_{n+k-1}^0$; the converse is proved in a symmetric fashion.

Our next area of interest is *cross coincidences*, where, given substitutions $\theta \neq \eta$ such that $M_\theta = M_\eta$, we compare the coincidence limit of $\theta^\infty(i_1) = t^{i_1}$ to $\eta^\infty(i_2) = h^{i_2}$, where $i_1,\, i_2 \in I$. As we are comparing sequences derived from distinct substitutions, we shall use $\delta(t^{i_1},\, h^{i_2})$ to denote the coincidence limit of t^{i_1} and h^{i_2}, providing it exists. Since $\theta \neq \eta$, we may allow $i_1 = i_2$, and we will see that indeed, there exist pairs of θ, η where $i_1 = i_2$ and $\delta(t^{i_1},\, h^{i_2})$ does not exist.

Two notes should be observed regarding the balance of blocks in

$$\begin{pmatrix} t^{i_0} \\ h^{i_1} \end{pmatrix}:$$

1) As $M_\theta = M_\eta$, we have that if

$$\hat{a} = \begin{pmatrix} t_1 t_2 \ldots t_m \\ h_1 h_2 \ldots h_n \end{pmatrix} \in \hat{I}$$

is a pair of (w-)balanced blocks, then the blocks

$$\begin{pmatrix} \theta(t_1 t_2 \ldots t_m) \\ \eta(h_1 h_2 \ldots h_m) \end{pmatrix}$$

will also be (w-)balanced. This can be easily verified by taking the vector of cardinalities of 0 and 1 of the θ and η strings and multiplying it with $M_\theta = M_\eta$, and then applying the results to the w-balanced block equation if $\hat{a}$ is w-balanced but not balanced.

2) If two pairs of balanced blocks $\hat{a}$, $\hat{b}$ are row transpositions of each other as defined in Chapter 1, it is not necessarily true that $\hat{a}$ and $\hat{b}$ are similar once substitutions are applied. Consider the substitutions $\theta : 0 \to 001$, $1 \to 11010$, $\eta : 0 \to 001$, $1 \to 10011$ and the pairs

$$\hat{a} = \begin{pmatrix} 01 \\ 10 \end{pmatrix}, \ \hat{b} = \begin{pmatrix} 10 \\ 01 \end{pmatrix}.$$

Applying θ to the top row of each block and η to the bottom row, we find that the pairs of blocks

$$\begin{pmatrix} \theta(01) \\ \eta(10) \end{pmatrix} = \begin{pmatrix} 00111010 \\ 10011001 \end{pmatrix} \text{and} \begin{pmatrix} \theta(10) \\ \eta(01) \end{pmatrix} = \begin{pmatrix} 11010001 \\ 00110011 \end{pmatrix}$$

have the same length (eight) and same number of coincidences (four), but they partition into different sets of minimal balanced blocks. Therefore, they cannot be grouped together in terms of calculating the density limit of a particular type of balanced block.

A second example, $\theta : 0 \to 0101,\ 1 \to 10010,\ \eta : 0 \to 0110,\ 1 \to 10010$ shows that row transposition does not always preserve the number of coincidences. Using the same blocks for $\hat{a}$ and $\hat{b}$ as before, we see that the pairs of length nine blocks

$$\begin{pmatrix} \theta(01) \\ \eta(10) \end{pmatrix} = \begin{pmatrix} 010110010 \\ 100100110 \end{pmatrix} \text{and} \begin{pmatrix} \theta(10) \\ \eta(01) \end{pmatrix} = \begin{pmatrix} 100100101 \\ 011010010 \end{pmatrix}$$

have 5 and 1 coincidences respectively.

Now that we're familiar with the general behavior of balanced blocks originating from different substitutions with an equal matrix, some examples are in order. In all of these forthcoming examples we will assume that for all blocks, the sequence associated with θ is on the top row, and the sequence associated with η is on the bottom row.

Example 1. Let $\theta_1 : 0 \to 01,\ 1 \to 1100,\ \eta_1 : 0 \to 01,\ 1 \to 1001$. Here $T(I) = 1$ and we've seen before that $M_\theta = M_\eta$ has a zero eigenvalue. From the structure of balanced pairs observed in θ and η we might expect that for any crossing, $\hat{I}$ is nonempty and finite. If so, $\delta(t^{i_1},\ h^{i_2})$ exists for any pair $(i_1,\ i_2)$.

Let's cross t^0 with h^1. We find that

$$\begin{aligned} \hat{I} &= \left\{ \begin{pmatrix} 01 \\ 10 \end{pmatrix}, \begin{pmatrix} 1100 \\ 0101 \end{pmatrix}, \begin{pmatrix} 110 \\ 011 \end{pmatrix}, \begin{pmatrix} 100 \\ 001 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \\ &= \{\hat{a},\ \hat{b},\ \hat{c},\ \hat{d},\ \hat{0},\ \hat{1}\} \end{aligned}$$

which yields the substitutions

$$\begin{array}{lll} \hat{a} \to \hat{a}\hat{b}, & \hat{b} \to \hat{c}\hat{0}\hat{b}\hat{a}\hat{0}\hat{1}, & \hat{c} \to \hat{c}\hat{0}\hat{c}\hat{0}\hat{0}\hat{1}, \\ \hat{d} \to \hat{b}\hat{a}\hat{0}\hat{1}, & \hat{0} \to \hat{0}\hat{1}, & \hat{1} \to \hat{1}\hat{d} \end{array}.$$

When we take our 6×6 matrix of these substitutions and compute w_L, we see that $d(\hat{a}) = d(\hat{b}) = d(\hat{c}) = d(\hat{d}) = 0.1$ and $d(\hat{0}) = d(\hat{1}) = 0.3$. Working with our normal formula for coincidence density, we find that $\delta(t^0, h^1) = 5/9$. An alternate version of the coincidence density program on SuperCollider used in Chapter 2 confirms this result [4]. The other three cross coincidence limits can be computed in a similar fashion and are also equal to 5/9, as each pair of sequences winds up being partitioned into the same pairs of minimal balanced blocks.

This shows that even though $\hat{I}$ is finite, $\hat{0}$ and $\hat{1} \in \hat{I}$ does not imply $\delta(t^{i_1}, h^{i_2}) = 1$. In fact, for any matrix with a zero eigenvalue, the finiteness of $\hat{I}$ and the fact that for at least one of $i = 0, 1$ we have that $\theta(i) \neq \eta(i)$ is enough to confirm the negative implication.

Example 2. Let $\theta_2 : 0 \to 001, 1 \to 11001, \eta_2 : 0 \to 010, 1 \to 10101$; recall these are the two regular substitutions corresponding to the matrix

$$\begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}.$$

If we try to cross t^{i_1} with h^{i_2} for any pair $(i_1, i_2) \in I \times I$, SuperCollider suggests that $\delta(t^{i_1}, h^{i_2})$ exists and equals 1/2. Unfortunately, we will also see that trying to partition any crossing into minimal (w-)balanced blocks is futile. The balanced block partition yields a set of minimal pairs with a subsequence of distinct blocks of diverging length, and the w-balanced block partition contains both $\hat{0}$ and $\hat{0}_{\natural}$. However, there are several things we do know about the behaviors of θ and η that will allow us to use a system of linear equations to obtain our projected coincidence limit:

1) $d(0) = d(1) = 1/2$ for both θ and η.

2) η is a periodic substitution with period 01 for h^0, 10 for h^1.

3) The 0's in t^0 and t^1 are arranged such that there is a limit of half the 0's occuring at odd indices and the other half occuring at even indices.

4) For $i \in \{0, 1\}$ and all $n \in \mathbf{Z}^+, t^i_n = h^0_n$ iff $t^i_n \neq h^1_n$ and vice versa (this is a consequence of the periodicity of η). Thus, if $\delta(t^i, h^0)$ exists, then $\delta(t^i, h^1)$ exists and is equal to $1 - \delta(t^i, h^0)$.

Let's first consider the crossing of t^0 and h^0. Our criteria for coincidence between these sequences is, given $n \in \mathbf{Z}^+$, that $t^0_n = h^0_n$ iff "n is odd and $t^0_n = 0$" or "n is even and $t^0_n = 1$". Let $P(odd\ (even),\ 0\ (1))$ denote the probability that, given a random positive integer n, that n is odd (even) *and* $t^n_0 = 0$ (1). Thus we obtain the following system:

$$\begin{aligned}
\frac{1}{2} &= P(odd,\ 0) + P(even,\ 0) \\
\frac{1}{2} &= P(odd,\ 0) + P(odd,\ 1) \\
\frac{1}{2} &= P(odd,\ 1) + P(even,\ 1) \\
\delta(t^0,\ h^0) &= P(odd,\ 0) + P(even,\ 1)
\end{aligned}$$

From note 3) above, we deduce $P(odd,\ 0) = P(even,\ 0) = 1/4$, therefore $P(odd,\ 1) = P(even,\ 1) = 1/4$, and so $\delta(t^0,\ h^0) = 1/2$. Note 4) gives us $\delta(t^0,\ h^1) = 1/2$. Finally, we use notes 1) and 4) again to confirm that $\delta(t^1,\ h^0) = c(t^1,\ h^1) = 1/2$ as well.

Example 3. Let $\theta_3 : 0 \to 001,\ 1 \to 11010$, $\eta_3 : 0 \to 001,\ 1 \to 11100$. Neither substitution yields a periodic sequence, and partitioning each combination of $(t^{i_1},\ h^{i_2})$ into w-balanced blocks shows that both $\hat{0}$ and $\hat{0}_\natural$ will appear in the partition, implying that there is no cross coincidence limit for any of the four pairs, which is verified by graphing on SuperCollider. Thus, for pairs of substitutions where the set of balanced blocks may be empty or infinite, there are no guarantees as to whether or not coincidence limits between sequences of the different substitutions exists.

Another way of comparing behaviors of substitutions $\theta \neq \eta$ with $M_\theta = M_\eta$ is by *sum substitutions.* For a binary substitution θ, the sum substitution sequence $t^\Sigma = t_1^\Sigma t_2^\Sigma t_3^\Sigma \ldots$ (with t^Σ pronounced "t-sum" or "t-sigma") is defined by either of these equivalent formulas as $n \to \infty$:

$$\begin{aligned}
t_n^\Sigma &= t_n^0 + t_n^1 \bmod 2 \\
t_n^\Sigma &= \begin{cases} 0 \text{ if } t_n^0 = t_n^1 \\ 1 \text{ if } t_n^0 \neq t_n^1 \end{cases}
\end{aligned}$$

For example, take the substitution $\theta : 0 \to 001, 1 \to 11001$. We will show the first 20 entries of the substitution sequences and make the respective entries for t^Σ out of those:

$$\begin{aligned}
t^0 &= 00100111001001001110\ldots \\
t^1 &= 11001110010010011100\ldots \\
t^\Sigma &= 11101001011011010010\ldots
\end{aligned}$$

Theorem 6. Given a substitution θ, $d_{t^\Sigma}(0)$ and $d_{t^\Sigma}(1)$ exist iff $\delta(\theta)$ exists.

Proof. This simply draws from the second formula for the construction of t_n^Σ, as $t_n^\Sigma = 0$ iff $t_n^0 = t_n^1$, and so $d_{t^\Sigma}(1) = 1 - d_{t^\Sigma}(0)$ if $d_{t^\Sigma}(0)$ exists.

Using the above example, we have $d_{t^\Sigma}(0) = d_{t^\Sigma}(1) = 1/2$. The reasonable coincidence limits to examine when introducing sequences like these is that between θ^Σ and η^Σ whenever $\theta \neq \eta$. If the substitutions θ, η behave in similar fashions (i.e., the indices at which t^0, t^1 match/mismatch compared to such indices for h^0, h^1) at a consistent rate, then $\delta(\theta^\Sigma, \eta^\Sigma)$ should exist. There are two immediate observations taken from the result of Theorem 6:

Theorem 7. Let θ, η be two subsitutions whose respective coincidence limits are both 0 or both 1. Then $\delta(t^\Sigma, h^\Sigma) = 1$.

Proof. If the coincidence limits are both 0, then $d_{t^\Sigma}(1) = d_{h^\Sigma}(1) = 1$ as the lengths approach infinity. Thus each string will have large strings of 1's that will always coincide, and $\delta(w^\Sigma,\ h^\Sigma) = 1$. The proof is symmetric if the limits are both 1, as w^Σ, h^Σ will consist of large strings of 0's behaving in the same fashion.

Theorem 8. Let θ, η be regular substitutions such that $\delta(\theta) = 1$ and $\delta(\eta) = p$. Then $\delta(t^\Sigma,\ h^\Sigma)$ exists and is equal to p. Similarly, if $\delta(\theta) = 0$ and $\delta(\eta) = p$ then $\delta(t^\Sigma,\ h^\Sigma) = 1 - p$.

Proof. Suppose $\delta(t^0,\ t^1) = 1$. $d_{t^\Sigma}(0) = 1$ and $d_{h^\Sigma}(0) = p$ as the indices approach infinity, thus a large string of 0's in t^Σ will line up with a string in h^Σ where the density of 0 is p and 1 is $1 - p$, thus the sum coincidence limit exists and will approach p. Likewise, $\delta(\theta) = 0$ implies $d_{t^\Sigma}(1) = 1$ and so the sum coincidence limit between t^Σ and h^Σ will approach $1 - p$.

The challenge with trying to find the coincidence limits of a pair of sequences such as t^Σ and h^Σ is that while we can easily compute each sequence separately, it is difficult to construct substitutions corresponding to the sum sequences themselves, and therefore it is difficult to predict anything about whether or not a limit exists. An alternate method is to define $x_n^\Sigma = 2x_n^0 + x_n^1$ for $x = t$ or h as necessary, which gives us a bijection from the ordered pairs $\{(0,\ 0),\ (0,\ 1),\ (1,\ 0),\ (1,\ 1)\}$ of $(t_n^\Sigma,\ h_n^\Sigma)$ to a quaternary alphabet $\{0,\ 1,\ 2,\ 3\}$, with $(0,\ 0) \to 0$ and so on respectively. Here we will need to tweak our definition of a balanced pair of blocks. In this case, a pair of blocks

$$\begin{pmatrix} t_1^\Sigma t_2^\Sigma \dots t_{L_i}^\Sigma \\ h_1^\Sigma h_2^\Sigma \dots h_{L_i}^\Sigma \end{pmatrix}$$

is balanced if

$$card\{n : 1 \le n \le L_i,\ t_n^\Sigma = 1\} = card\{n : 1 \le n \le L_i,\ t_n^\Sigma = 2\}$$

and

$$card\{n : 1 \le n \le L_i,\ h_n^\Sigma = 1\} = card\{n : 1 \le n \le L_i,\ h_n^\Sigma = 2\}.$$

That is, θ and η contain separate sequences of pairs of balanced blocks whose lengths sum to L_i. (0's and 3's in the quaternary alphabet correspond to a coincidence of 0's and 1's in the binary alphabet respectively, so we only need to worry about counting the indices n where $\theta_n^0 \neq \theta_n^1$, likewise with η.) Using this approach, we have that

$$\delta(\theta^\Sigma,\ \eta^\Sigma) = \frac{card\{k \leq n : \theta_k^\Sigma = \eta_k^\Sigma \text{ or } \theta_k^\Sigma + \eta_k^\Sigma = 0 \bmod 3\}}{n}.$$

These challenges and other ideas related to sum coincidences will be referenced again in Chapter 4.

The chapter concludes with an introduction of concepts concerning substitution sequences in q-ary alphabets, where $q \geq 3$. Without loss of generality we may assume such a q-ary alphabet to consist of the symbols $\{0,\ 1,\ \ldots,\ q-1\}$. If $q > 10$, we could add Roman letters, Greek letters, or other suitable symbols after 9 if multiple digit numbers may become confusing. For a q-ary substitution θ and its associated sequences $t^0,\ t^1,\ \ldots,\ t^{q-1}$ we define a coincidence iff $t_n^0 = t_n^1 = \ldots = t_n^{q-1}$, and for $2 \leq k < n$, a *k-coincidence* iff there exists a subset $\{i_1, \ldots, i_k\}$ of the alphabet such that $t_n^{i_1} = \ldots = t_n^{i_k}$ at the index n.

Our notation for coincidence limits will retain familiarity: δ for coincidences and δ^k for k-coincidences. If we choose to examine (k-)coincidences within a proper subset of sequences, we list them in parentheses after the $\delta^{(k)}$, much like our cross and sum coincidence notations.

The definition of a constant and nonconstant substitution remains the same, but we may also introduce a substitution as *semiconstant* if the substitutions for at least two $i \in I$ have equal length, but not all of them. Similarly, a substitution can be defined as *semiregular* if $\delta(t^{i_1},\ \ldots,\ t^{i_k})$ exists for the proper subset $\{i_1,\ \ldots,\ i_k\}$, $k \geq 2$, but $\delta(\theta)$ does not exist.

Example. Let θ be the ternary substitution $0 \to 0121$, $1 \to 102$, $2 \to 2210$. This substitution is semiconstant as $|\theta(0)| = |\theta(2)| \neq |\theta(1)|$. θ corresponds to the 3×3 matrix

$$M_\theta = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

and these sequences:

$$\begin{aligned} t^0 &= 01211022210102102012122 10\ldots \\ t^1 &= 10201212210012110222101 02\ldots \\ t^2 &= 22102210102012122102210 10\ldots \end{aligned}$$

In these opening entries, there are only two coincidences (the 2's in the fourteenth index and the 1's at the fifteenth) but several 2-coincidences exist, starting with the

2's in the third index, the 0's in the fourth, and so on. The first minimal balanced block for the above substitution is of length nine and consists of four 2's, three 1's and two 0's. We could also search for minimal blocks restricted to a subset of the t^i's. For example, if we restrict our search to t^0 and t^1 we find balanced blocks of lengths two, five, one, and one within the length nine block covering the beginning of each sequence.

Other general definitions from the first chapter, such as the density of a symbol or block, and coincidence limit, retain their meaning in an arbitrary q-ary alphabet, keeping in mind the extra conditions needed to satisfy every symbol in the alphabet. We can expand some of the same principles developed in Chapter 2 to determine if there exists a coincidence limit for $\{t^0, \ldots, t^{q-1}\}$ or a k-coincidence limit among the subset $\{t^{i_1}, \ldots, t^{i_k}\}$, particularly that if there is a finite, nonempty set of balanced blocks for a set of (k) sequences, then those sequences have a (k-)coincidence limit.

Periodic substitutions in larger alphabets are also easy to construct; a simple ternary substitution where each sequence has a period of length three is $\pi_1 : 0 \to 0120,\ 1 \to 1201,\ 2 \to 2012$. Every period has a pairwise coincidence density of zero, and thus we can conclude the (2-)coincidence density of any subset of the sequences is 0. However, if we consider $\pi_2 : 0 \to 010,\ 1 \to 101,\ 2 \to 202$ we see that the sequences starting with 0 and 1 are periodic, but the sequence starting with 2 is not (although there are some nice patterns involving sets of blocks in powers of three in t^2 nonetheless).

We may conclude by proving two theorems about the behaviors of q-ary substitutions of constant and semiconstant length:

Theorem 9. Let θ be a q-ary substitution of constant length, and suppose there exists a subset $\{i_1, \ldots, i_j\}$, $2 \leq j \leq q$, of $\{0, \ldots, q-1\}$ such that the $\theta(i_j)$'s have a (k-)coincidence, with $2 \leq k < j$ if necessary. Then $\delta^{(k)}(t^{i_1}, \ldots, t^{i_k})$ exists and is equal to 1. If the substitutions never coincide, then $\delta^{(k)}(t^{i_1}, \ldots, t^{i_j})$ exists and is equal to 0.

Proof. This is an "inductive" corollary to P. Michel's proof [7] that constant binary substitutions have coincidence limits of 1 or 0; Michel proves the "base" case of $q = 2$. Suppose the substitutions all coincide at some point. Then $\delta(t^{i_n}, t^{i_{n+1}}) = 1$ for all $n \leq j-1$, and so we conclude $\delta(\theta) = 1$. Since having a coincidence implies a k-coincidence for all $2 \leq k < n$, $\delta^k(t^{i_1}, \ldots, t^{i_j}) = 1$ for any subset of size $2 \leq k < j \leq q$ as well.

Now suppose the substitutions never coincide. Then for any $n \in \mathbf{Z}^+$, the symbols $t^0_n, \ldots, t^{q-1}_n$ are all distinct, and so $\delta^{(k)}(t^{i_1}, \ldots, t^{i_j}) = 0$.

Theorem 10. Let θ be a q-ary semiconstant substitution, and let $p \neq r \in I$ such that $|\theta(p)| = |\theta(r)| = n$. Denote $\theta(p) = a^1_p \ldots a^n_p$ and $\theta(r) = a^1_r \ldots a^n_r$. Suppose there exists $k \leq n$ such that $a^k_p = a^k_r$, and the blocks

$$\hat{a} = \left(\begin{array}{c} a_p^1 \dots a_p^{k-1} \\ a_r^1 \dots a_r^{k-1} \end{array} \right), \ \hat{b} = \left(\begin{array}{c} a_p^{k+1} \dots a_p^n \\ a_r^{k+1} \dots a_r^n \end{array} \right)$$

are balanced. Then $\delta(t^p, t^r) = 1$.

Proof. Again, this theorem draws from P. Michel's original proof for binary constant substitutions that coincide, and we can even use one of his original formulas. First of all, we can write

$$\left(\begin{array}{c} \theta(p) \\ \theta(r) \end{array} \right) = \left(\begin{array}{c} a_p^1 \dots a_p^{k-1} a_p^k a_p^{k+1} \dots a_p^n \\ a_r^1 \dots a_r^{k-1} a_r^k a_r^{k+1} \dots a_r^n \end{array} \right) = \left(\begin{array}{ccc} (\hat{a}) & \begin{array}{c} a_p^k \\ a_p^k \end{array} & (\hat{b}) \end{array} \right)$$

Since $|\theta(p)| = |\theta(r)|$ and $\hat{a}$, $\hat{b}$ are balanced, we have that $\theta^x(\hat{a})$ and $\theta^x(\hat{b})$ are balanced for all x, and so

$$\left(\begin{array}{c} \theta^x(p) \\ \theta^x(r) \end{array} \right) = \left(\begin{array}{ccc} \theta^x(\hat{a}) & \begin{array}{c} \theta^x(a_p^k) \\ \theta^x(a_p^k) \end{array} & \theta^x(\hat{b}) \end{array} \right)$$

is balanced for all x. Therefore,

$$card\{m \leq |\theta^x(p)| : t_p^m \neq t_r^m\} \leq (n-1)^x$$

and as $x \to \infty$, this implies $\delta(t^i, t^j) = 1$.

Chapter 4: Open Problems and Some Conjectures

We conclude with some general questions and hypotheses that arise as a result of the results explored in Chapter 3. We shall work roughly in the order the results were presented in the chapter.

Problem 1. Let θ be a q-ary substitution with alphabet I such that there exists $i \in I$ with w^i a periodic sequence. If p is the length of the period in w^i, what values can p take relative to q? Does it matter if θ is constant, semiconstant, or non-constant?

Hypotheses. It's easy to see that if $1 \leq p \leq q$, then there exists a q-ary substitution with a period of length p. Let $J \subset I$ with $J = \{j_1, \ldots, j_p\}$, and without loss of generality, suppose $\theta(j_1) = j_1 j_2 \ldots j_p j_1$. If we set $\theta(j_2) = j_2 \ldots j_p j_1 j_2, \ldots,$ $\theta(j_p) = j_p j_1 \ldots j_{p-1} j_p$ then we have t^j periodic for all $j \in J$ with period of length p. θ does not have to be constant among these p substitutions either; $\theta(j_k)$ could be equal to $(j_k \ldots j_{k-1})^{x_k} j_k$ for distinct x_k among the j_k's but the periodicity will still remain.

We've proven that it's impossible for the case of $q = 2$, $p = 3$, regardless of if θ is constant or non-constant. Perhaps there exist substitutions which have periods where 2 is a proper divisor or even an n^{th} root of p, or maybe 1 and 2 are the only feasible values of p. We can adjust these hypotheses for greater q as necessary. However, I am not confident we will have a case where $p > q$; the proof likely lies in more modular arithmetic as seen in Chapter 3.

Problem 2. Is every binary substitution θ of the form $0 \to 0^k 1$, $1 \to 11(\theta^{k-1}(0))$ regular?

Hypothesis. I believe that every substitution of this form is regular. Theorem 5 shows that for each $k \geq 1$ we have $n \geq k+1$ implies $t^0_n = t^1_n$ iff $t^0_n = t^0_{n+k-1}$. Therefore, we should be able to concatenate the blocks of 0^k, 1, and 111 such that we can partition them into a sequence of a finite set of concatenated blocks, each of which has a calculatable density and a length longer than the slope. Then we can apply a related substitution η in terms of those blocks (much like we've seen in Chapter 2) to determine how many of the symbols will coincide with the symbol $k-1$ indices greater, and thus determine a coincidence limit for θ.

Problem 3. Let θ be a binary substitution such that there exist $m, N \in \mathbf{Z}^+$ with for all $n \geq N$ we have $t^0_n = t^1_n$ iff $t^0_n = t^0_{n+m}$ or $t^1_n = t^1_{n+m}$. Can M_θ have a form not equal to

$$\begin{bmatrix} k & 1 \\ k(k-1) & k+1 \end{bmatrix}$$

(or its row transposition) for some positive integer k? What if θ is q-ary for $q \geq 3$?

Hypotheses. I think there are infinitely many binary substitutions that imply the "slope" criterion for coincidence. The best way to find such a substitution is to create one where one of the symbols must always appear in a string of its own symbol of constant length (such as the 0^k strings as seen in Theorem 5). Once we find a case where the substitutions of 0 and 1 have the smallest possible length, we should be able to construct an infinite family of similar substitutions by adding powers of strings to both sequences as necessary. Compare this strategy with the form we already have, where $\theta(1)$ ends with a number of $\theta(0)$ strings equal to the number of 0's at the beginning of $\theta(0)$.

If $q \geq 3$ the situation becomes more challenging, but I think it could be manageable. Perhaps we need to force $q-1$ of the symbols to appear in a constant length string of its own symbol and let the final symbol be flexible in length. I also think the order of such strings needs to be cyclic: for example, in a ternary sequence we would have to have a string of 0's followed by a string of 1's, then a string of 2s, then back to 0's, and so on.

Problem 4. Suppose θ, η are irregular binary substitutions with $M_\theta = M_\eta$. Can there exist a pair of sequences t^{i_1}, h^{i_2} such that $\delta(t^{i_1}, h^{i_2})$ exists? What if θ and η are regular but there exist irregular substitutions with an equal matrix?

Hypotheses. If it is to be believed that pairs of balanced blocks of cross sequences mimic the behaviors (emptyness and/or finiteness) of the balanced blocks of the two substitutions they're derived from, then I think that if θ and η are irregular, then no crossing of the substitutions will yield a coincidence limit. As well, M. Dekking's theorem on w-balanced blocks will probably apply in appropriate situations (which is only advantageous whenever the entries of the Perron–Frobenius right eigenvector have a rational ratio).

If θ and η are regular, then the regularity of their crossing may be dependent on which pairs $(i_1, i_2) \in I \times I$ are chosen. If the substitutions are a small number of degrees apart and for $i = 0$ or 1 we have $\theta(i) = \eta(i)$ then balanced blocks for $\delta(t^i, h^i)$ could come fairly easily, or perhaps as futilely as before. (Remember, $\theta : 0 \to 001, 1 \to 11001$ is regular but has an empty set of balanced blocks.) If the set of blocks is indeed empty or infinite, we should then test for w-balanced blocks, and proceed as usual. However, if the crossed sequences begin rather similarly, that might imply that a cross coincidence limit exists.

Problem 5. Suppose θ, η are regular binary substitutions with $M_\theta \neq M_\eta$. Can we cross t^{i_1}, h^{i_2} such that $\delta(t^{i_1}, h^{i_2})$ exists? Will this happen for every pair t^{i_1}, h^{i_2}?

Hypotheses. The best way to handle this might be to make a 4×4 matrix

$$M_{\theta\eta} = \begin{bmatrix} M_\theta & 0 \\ 0 & M_\eta \end{bmatrix}$$

and compute its eigenvalues. The purpose of making it a 4×4 matrix with the distinct quadrants is to emphsize the differences in behavior for 0 and/or 1 in the two substitutions. If we have $\lambda_1 > 1$ and $\lambda_2,\ \lambda_3,\ \lambda_4 = 0$ then I feel very confident that the crossing is regular, since that would be the logical 4×4 analogue to the 2×2 case where $\lambda_1 > 1 > \lambda_2 = 0$ implies all substitutions of that matrix are regular.

There is also one immediate family of cases where possibly irregular substitutions can be crossed such that their coincidence limit is 0. Let η equal the substitution for θ where 1 is added to every symbol in θ then reduced mod 2. For example, if $\theta : 0 \to 0110, 1 \to 10011$ then $\eta : 0 \to 01100, 1 \to 1001$, and we have $\delta(t^0,\ h^1) = 0$ (and the same for $\delta(t^1,\ h^0)$). Note that in these such cases M_η is the row transposition of M_θ.

Problem 6. Suppose θ and η are regular binary substitutions, with coincidence limits $0 < \delta(\theta),\ \delta(\eta) < 1$. Can we find a relationship or formula between $\delta(t_0,\ t_1)$, $\delta(h_0,\ h_1)$, and $\delta(t^\Sigma,\ h^\Sigma)$? Does it matter if $M_\theta = M_\eta$?

Hypotheses. The situation in which a coincidence limit was 0 or 1 was covered in the prior chapter, so it's natural to consider what would happen in these intermediate cases. Note that the only cases of regular substitutions that have been studied result in rational coincidence limits, so if we let $\delta(t_0,\ t_1) = \frac{p_1}{q_1}$, $c_n(h_0,\ h_1) = \frac{p_2}{q_2}$ where the p_i and $q_i \in \mathbf{Z}^+$ and relatively prime, then my initial hypothesis was that $\delta(t^\Sigma,\ h^\Sigma)$ exists, is rational, and has a denominator that divides the lcm of q_1 and q_2. If these limits do exist, then the challenge becomes determining if we can always partition t^Σ and h^Σ into balanced blocks to get that result, and if so, if can we do it in terms of 0's and 1's, or if we have to use the $\{0,\ 1,\ 2,\ 3\}$ alphabet to do so.

Having computed an example on SuperCollider [4], my hypothesis appears to be incorrect. I used $\theta : 0 \to 001,\ 1 \to 11001$ and $\eta : 0 \to 01,\ 1 \to 1100$ as my testing substitutions. As seen before, both have coincidence limits of $1/2$. I used the $\{0,\ 1,\ 2,\ 3\}$ alphabet and mod 3 criterion for coincidence, testing the sum sequences over the first 16000 substitutions. Without any initial regard for balanced blocks between t^Σ and h^Σ, which might be empty as it is because θ has an empty set of balanced blocks, the sum coincidence limit does not appear to exist. Even still, as n increases, the coincidence ratios oscillate around $5/9$ as opposed to some $p/2^i$ for p odd. So, the best way to treat this problem might be to examine pairs of substitutions whose balanced block sets are both nonempty and finite, and go from there.

Conclusion

I believe that exploring the cases related to these problems will give us a better understanding of how substitutions among various sizes of alphabets will work. There are certainly enough special circumstances presented in these problems, particularly in the binary alphabet, such that enough proofs of these theorems in the binary case can induce an intuition about similar behaviors in other q-ary cases. From there, we could be able to subdivide these cases into other interesting situations. They might not have to be quite in depth as "Does a cross k-coincidence limit exist between the sequences $(\theta_1)^0$, $(\theta_2)^1$, $\ldots$, $(\theta_q)^{q-1}$ for any particular $k < q$?" but in a mathematical realm as recently developed and relatively unexplored as the dynamics or substitutions, the possibilities are still wide open.

Bibliography

[1] Dekking, F. Michel. The spectrum of dynamical systems arising from substitutions of constant length. Z. Wahrschleinlichkeitsttheorie verw. Gebiete **41**, 221–239 (1978).

[2] Dekking, F. Michel. Diagonals in 2D-tilings and coincidence densities of substitutions. arXiv:1302.3759 [math.DS] (2013).

[3] Dekking, F. Michel. (2012, December 10). Coincidence densities of substitutions, "DNA-matching" of sibling streams. Lecture conducted from Instituut-Lorentz for theoretical physics, Universiteit Leiden, Netherlands.

[4] Hoffman, Ian. February 27, 2013. Computer program source code. "Substitution Generator and Coincidence Calculator Samples."

[5] Keane, Michael S. Generalized Morse sequences. Z. Wahrschleinlichkeitsttheorie verw. Gebiete **10**, 335–353 (1968).

[6] Michel, Pierre. Sur les ensembles minimaux engendrés par les substitutions de longeur non constante. Thèse. Université de Rennes (1975).

[7] Michel, Pierre. Coincidence values and Spectra of Substitutions. Z. Wahrschleinlichkeitsttheorie verw. Gebiete **42**, 205–227 (1978).

[8] Morse, Marston. Recurrent geodesics on a surface of negative curvature. Trans. Amer. Math. Soc. **22**, 84–100 (1921).

Appendix A: SuperCollider Source Codes

These are samples of the various codes I developed throughout the course of my research. While SuperCollider is most widely used for and known as an audio synthesis language, it is capable of performing a variety of mathematical and Boolean operations as well, many kinds of which are present in the following codes. The three samples here are basic codes for computing and plotting ratios of coincidences, sum coincidences, and cross coincidences for binary substitutions; the comments in the codes should also prove a handy walkthrough as well. Those familiar with programming in the Ruby or C family languages should recognize many of the commands and operations seen below.

The most recent build used for this thesis is SuperCollider IDE version 3.6.3 with the Qt version 4.8.4 toolkit. On the computer I used (Windows 7 with 1.83 GHz processing speed), it can construct the first 50000 substitutions[2] of both sequences for $\theta : 0 \to 001,\ 1 \to 11100$ (t^0 and t^1 have 200,000 and 200,008 indices respectively), compute their coincidence ratios for those first 200,000 indices, and plot a line graph of the ratios in under 45 seconds.

```
Substitution Generator and Coincidence Calculator Samples by Ian Hoffman
ihoffman (AT) wesleyan.edu Updated February 27, 2013

These samples are freeware. This source can be used or modified for non-
commercial purpose, but proper credit must be given. Please ask the cre-
ator if this code is to be used for other purposes.

The code is designed for use in SuperCollider and other related languages.
```

```
// This program was created by Ian Hoffman for his thesis on Coincidences
// in Substitution Sequences.
//
// The program accepts the substitutions of the individual symbols of an arb-
// itrary alphabet, as well as concatenating the substitutions of sequences
// of symbols into a complete sequence. Once the sequences have been ful-
// ly generated, the user may request the density of coincidences between
// two sequences, by asking the program to check for equality of entries at
// every feasible index (i.e. each index where an entry is present in both
// sequences). Once the ratios after each index have been calculated, the
```

[2] Replacing the "9999" in the code seen on the following page with "49999"

```
// user may also request a line graph to be drawn which approximates the
// contour of the ratios as the number of checked indices increases.
//
// Other optional functions that are demonstrated among the various samples
// include the positions and sizes of balanced blocks (sections of equal length
// and position in the two sequences such that any given symbol appears the
// same amount of times in each sequence), cross coincidences (in which two
// different substitutions are presented and one sequence is taken from each
// substitution), and sum coincidences (where addition is performed on substi-
// tution sequences on an index-by-index basis and then examined modulo 3).
```

```
// Sample 1: A Basic Test for Counting Coincidences

// This code examines the coincidences of the substitution
// 0 -> 001, 1 -> 11100

x = Array.with(0, 0, 1).postln;
y = Array.with(1, 1, 1, 0, 0).postln;

// x is an array that starts with a single 0 and substitutes
// 0 for 001 and 1 for 11100
// y is an array that starts with a single 1 and substitutes the same as x

m = 1;

for (m, 9999, {case
     {x[m] == 0} {case
          {y[m] == 0} {x = x.add(0).add(0).add(1);
     y = y.add(0).add(0).add(1); m = m+1}
          {y[m] == 1} {x = x.add(0).add(0).add(1);
     y = y.add(1).add(1).add(1).add(0).add(0); m = m+1}}
     {x[m] == 1} {case
          {y[m] == 0} {x = x.add(1).add(1).add(1).add(0).add(0);
     y = y.add(0).add(0).add(1); m = m+1}
          {y[m] == 1} {x = x.add(1).add(1).add(1).add(0).add(0);
     y = y.add(1).add(1).add(1).add(0).add(0); m = m+1}}
     });

x.postln;
y.postln;
m.postln;
```

```
// The initial m shows one substitution has already been performed.
// The second m shows 10000 substitutions have been performed.
// x now corresponds to the first 10000 substitutions of the infinite
// sequence beginning with 0; y is respectively the same with the
// sequence beginning with 1.

c = 0;
d = 1;
~ratio = Array.new;
for (d, List[x.size, y.size].minItem, {if (x[d-1] == y[d-1],
     {c = c+1; (c/d).postln; ~ratio = ~ratio.add((c/d));
          c.postln; d.postln; d = d+1;},
     {(c/d).postln; ~ratio = ~ratio.add((c/d)); c = c;
          c.postln; d.postln; d = d+1;}
     )
     }
);

// The fraction (c/d) tells us the number of coincidences c found
// after examining the first d indices of x and y, until the minimum
// length (size) of x and y has been reached. Because these rati-
// os are all kept in their own Array, we can use them to plot the
// line graph as seen below:

~ratio.plot("theta^inf(0) vs. theta^inf(1)", minval: 1/3, maxval: 2/3);

// Because some ratios early on take values outside of the inter-
// val [1/3, 2/3], SuperCollider shows them at the top or bottom
// of the graph window as opposed to not showing them at all
// (like a graphing calculator would do). However, since the se-
// quence of ratios quickly become bounded by [1/3, 2/3], it's use-
// ful to use this as our range since it's easier to see the contours
// of the graph. Note that the coincidence limit of the substitution
// does not exist.

// Below I will sample some zoom-ins near some of the local minima
// and maxima seen on the graph.

~ratio[250..300].plot("First good local minimum",
     minval: 0.45, maxval: 0.5);
~ratio[600..700].plot("First good local maximum",
```

```
     minval: 0.54, maxval: 0.5505);
~ratio[1000..1100].plot("Second local minimum",
     minval: 0.485, maxval: 0.4975);
~ratio[2700..2800].plot("Second local maximum",
     minval: 0.55, maxval: 0.56);

// If we wish to find the exact local minima and maxima, and their
// index numbers, we can use these lines as well. If we only want
// to view these one at a time, we can delete the .postln as neces-
// sary before performing the operation.

~ratio[3000..5000].minItem.postln;
     ((~ratio[3000..5000].minIndex)+3000).postln;
~ratio[10000..11500].maxItem.postln;
     ((~ratio[10000..11500].maxIndex)+10000).postln;
~ratio[16000..17000].minItem.postln;
     ((~ratio[16000..17000].minIndex)+16000).postln;
~ratio[43000..44000].maxItem.postln;
     ((~ratio[43000..44000].maxIndex)+43000).postln;
~ratio[65000..66000].minItem.postln;
     ((~ratio[65000..66000].minIndex)+65000);

// In addition, the eigenvalues for the matrix associated with this
// substitution suggest that the set of balanced blocks is either
// empty or infinite. We can run the following block of code to de-
// termine how to partition our pair of sequences into balanced
// blocks.

~sum = 0;
for (~sum, List[x.size, y.size].minItem,
     {if (x[0..~sum].sum == y[0..~sum].sum,
     {~sum = ~sum+1;
     ("Balance found after"+~sum.asString+"entries;
     the sum of the digits of either sequence is"+x[0..(~sum-1)].sum).postln},
     {~sum = ~sum+1})
     }
     ); (List[x.size, y.size].minItem.asString+"total entries")
```

```
// Sample 2: Cross Coincidences Between Two Substitutions

// Here we will consider two substitutions "theta" and "eta".
```

```
// "Theta" maps 0 to 001 and 1 to 11010,
// and "eta" maps 0 to 010 and 1 to 11010. We will use the
// code to show the first 10000 substitutions of the infinite
// sequences of theta beginning with 1, and eta beginning
// with 1, then analyze the coincidences in those sequences.

// The code is essentially the same as before, but note the dif-
// ferences that do appear.

x = Array.with(1, 1, 0, 1, 0).postln;
y = Array.with(1, 1, 0, 1, 0).postln;

m = 1;

for (m, 7999, {case
     {x[m] == 0} {case
          {y[m] == 0} {x = x.add(0).add(0).add(1);
     y = y.add(0).add(1).add(0); m = m+1}
          {y[m] == 1} {x = x.add(0).add(0).add(1);
     y = y.add(1).add(1).add(0).add(1).add(0); m = m+1}}
     {x[m] == 1} {case
          {y[m] == 0} {x = x.add(1).add(1).add(0).add(1).add(0);
     y = y.add(0).add(1).add(0); m = m+1}
          {y[m] == 1} {x = x.add(1).add(1).add(0).add(1).add(0);
     y = y.add(1).add(1).add(0).add(1).add(0); m = m+1}}
});

x.postln;
y.postln;
m.postln;

c = 0;
d = 1;
~ratio = Array.new;
for (d, List[x.size, y.size].minItem, {if (x[d-1] == y[d-1],
     {c = c+1; (c/d).postln; ~ratio = ~ratio.add((c/d));
          c.postln; d.postln; d = d+1;},
     {(c/d).postln; ~ratio = ~ratio.add((c/d)); c = c;
          c.postln; d.postln; d = d+1;}
     )
     }
);
```

```
~ratio.plot("theta and eta", minval: 0.575, maxval: 0.625);
```

```
// Sample 3: Sum Coincidences Between Two Substitutions

// Our last sample uses the same substitutions as seen in Sample 2;
// "theta" mapping 0 to 001 and 1 to 11010,
// and "eta" mapping 0 to 010 and 1 to 11010.

x = Array.with(0, 0, 1).postln;
y = Array.with(1, 1, 0, 1, 0).postln;

m = 1;

for (m, 9999, {case
     {x[m] == 0} {case
          {y[m] == 0} {x = x.add(0).add(0).add(1);
     y = y.add(0).add(0).add(1); m = m+1}
          {y[m] == 1} {x = x.add(0).add(0).add(1);
     y = y.add(1).add(1).add(0).add(1).add(0); m = m+1}}
     {x[m] == 1} {case
          {y[m] == 0} {x = x.add(1).add(1).add(0).add(1).add(0);
     y = y.add(0).add(0).add(1); m = m+1}
          {y[m] == 1} {x = x.add(1).add(1).add(0).add(1).add(0);
     y = y.add(1).add(1).add(0).add(1).add(0); m = m+1}}
});

x.postln;
y.postln;
m;

z = ((2*x)+y).postln;

// z is the sum sequence of theta, where 0 corresponds to a 0 in
// the theta(0) sequence and a 0 in the theta(1) sequence, 1 cor-
// responds to 0 in theta(0) and 1 in theta(1), and so on, over
// the course of the first 10000 substitutions in each sequence.

p = Array.with(0, 1, 0).postln;
q = Array.with(1, 1, 0, 1, 0).postln;

n = 1;
```

```
for (n, 9999, {case
     {p[n] == 0} {case
          {q[n] == 0} {p = p.add(0).add(1).add(0);
     q = q.add(0).add(1).add(0); n = n+1}
          {q[n] == 1} {p = p.add(0).add(1).add(0);
     q = q.add(1).add(1).add(0).add(1).add(0); n = n+1}}
     {p[n] == 1} {case
          {q[n] == 0} {p = p.add(1).add(1).add(0).add(1).add(0);
     q = q.add(0).add(1).add(0); n = n+1}
          {q[n] == 1} {p = p.add(1).add(1).add(0).add(1).add(0);
     q = q.add(1).add(1).add(0).add(1).add(0); n = n+1}}
});

p.postln;
q.postln;
n;

r = ((2*p)+q);

// r is the eta analogue of z...

c = 0;
d = 1;
~ratio = Array.new;
for (d, List[z.size, r.size].minItem, {if (((z[d-1] + r[d-1])%3 == 0) ||
     (z[d-1] == r[d-1]),
     {c = c+1; (c/d).postln; ~ratio = ~ratio.add((c/d));
          c.postln; d.postln; d = d+1;},
     {(c/d).postln; ~ratio = ~ratio.add((c/d)); c = c;
          c.postln; d.postln; d = d+1;}
     )
     }
);
~ratio.plot("theta sum and eta sum", minval: (5/9)-0.1,
     maxval: (5/9)+0.1)

// "(z[d-1] + r[d-1])%3 == 0" is the Boolean command for checking
// whether or not the sum in question is congruent to 0 modulo 3.

// What we're examining between z and r is whether or not their
// coincidences and non-coincidences coincide at a regular rate.
```

```
// Since 0s and 3s correspond to the coincidences in z or r, and 1s
// and 2s do so with the non-coincidences, we know that coinciden-
// ces occur iff the sum *or* the difference is 0 mod 3.
// Afterwards, we can plot a graph as usual, and we will see
// that a sum coincidence limit will likely not exist due to the con-
// tours of the graph.
```