Explorations of the Collatz Conjecture

Kelly S. LaTourette

Advisor: Dr. Michael Fraboni
Liaison: Dr. Benjamin Coleman

November 2007
Abstract

This paper presents results found from work on total stopping times of the Collatz Conjecture. We are especially interested in when the total stopping times of consecutive integers are equal, why they are equal, and finding runs of consecutive integers that have equal total stopping time. We use special parity sequences known as blocks, strings and stems to study the behaviors of the trajectories of these integers. Our work culminates in the conjecture that given any starting parity sequence, we can construct an arbitrarily long run of consecutive integers of the same total stopping time.
Dedication

This project is dedicated to my mother. She is my motivation to work hard, my role model for treating others with kindness, my inspiration to never lose hope and my guardian angel that is watching over me every day. I love and miss you dearly, Mom!
Acknowledgments

A HUGE thank you goes out to:

• Dr. Michael Fraboni, not just my academic advisor, but my friend as well. You believed in me when I had difficulty believing in myself. You have always cared not just about me as a student, but also about me as a person. Thank you for everything.

• My SOAR Team: Dr. Alicia Sevilla, Rebecca Angstadt and James Long. You were all a terrific help throughout our summer of research and your continued support as this became my Honors Project has been wonderful. I’m glad that we got to spend those countless hours discussing $3x + 1$ for a whole summer. Thank you!

• Patrick Sutton, my friend and Writing Center Tutor. Thank you for your many hours of reading drafts of this paper and for knowing just what to say to make me laugh when the going got tough.

• My wonderful family. Thanks for all of the support and encouragement in helping me fulfill my goals of becoming a successful college student. Knowing that I had your love with me always helped me through.

• My amazing friends, Andrew Abraham, Katie Bernstein, Jessica Fenske, Timothy Mills, Wesley Moser, Christopher Schilling, Rachael Todd, my Gamma Sigma Sigma brothers and sisters and all of the SOAR kids: You have all gone above and beyond the call of duty. You have encouraged me, helped me relax, and were there when I just needed someone to talk to. You are a truly fantastic group of friends and I am blessed to have you.

• My professors: Dr. Benjamin Coleman, Dr. Kevin Hartshorn, Dr. Robert Mayer and Dr. Nathan Shank: Thank you all so much for your kindness, understanding and support. Having professors like you is what makes me love learning here at Moravian so much!
Preface

It all began as a research project in the Student Opportunities for Academic Research (SOAR) Program in the summer of 2006. Along with students Rebecca Angstadt and James Long, as well as Drs. Michael Fraboni and Alicia Sevilla, I delved deep into the complexities of the Collatz Conjecture. For 10 weeks we met day after day to discuss our findings and our struggles as we tried to make some sense out of pieces of the conjecture. Rebecca turned her focus toward the negative integers and the three known cycles found through function iterations there. James, on the other hand, spent much of his time analyzing the work of Riho Terras and then eventually applying it to generalizations of our function. I set my sights on the total stopping times of the integers and constructing a massive graphical representation of the function. I enjoyed my work and was so intrigued by the conjecture that I have expanded my research into the Honors Project that is in front of you now. I hope that you enjoy reading it as much as I have enjoyed working on it.
Contents

1 Introduction ................................................. 6
  1.1 Definitions ........................................... 7

2 Total Stopping Time ....................................... 10

3 Consecutive Integers ..................................... 14
  3.1 Applying Garner’s Stems ............................... 19
  3.2 Consecutive Triples ................................... 21

4 Blocks and Strings ......................................... 23
  4.1 Corresponding Blocks and Strings ................. 23
  4.2 Block and String Structure ......................... 29
    4.2.1 Equal number of ones ......................... 29
    4.2.2 Difference Between Sums ...................... 29
    4.2.3 Switching 0’s and 1’s ......................... 31
    4.2.4 Proportion of Ones ............................ 36

5 Long Runs of Consecutive Integers ................. 41
Chapter 1

Introduction

The Collatz Conjecture is a well-known problem that has been puzzling mathematicians for years. The conjecture was first proposed in 1937 by Lothar Collatz, a German mathematician. Since the 1930’s it has been studied in a variety of ways by mathematicians all over the world. It has received other names, such as the Hailstone Numbers, the $3x+1$ Problem, the Ulam Conjecture and the Syracuse Problem, from the various mathematicians who have studied it. The past 70 years have led to the discovery of many different properties, extensions and representations of the basic idea of the conjecture. For a nice overview of work that has been done on the Collatz Conjecture, see *The $3x+1$ problem and its generalizations* by Jeffrey Lagarias [5].

Lagarias worked with David Applegate [1, 2] to develop their own ideas and theories on the problem. Their work on trees and total stopping time was of special interest to me, as we will see later on in this paper.

Another mathematician whose work was especially helpful, was Riho Terras [6]. He developed an algorithm that I used again and again throughout my re-
search. Most of my work, however, is an extension of work specifically done by Lynn Garner in *On heights in the Collatz 3n+1 problem* [4]. This work focuses on the heights (total stopping times) of integers. Guo-Gang Gao [3] produced work similar to Garner’s, but worked with a slightly different function than the one that we used.

For the purposes of this research, we will use the representation that is typically known as the $3x + 1$ function and is defined as follows.

For all positive integers $x$,

$$T(x) = \begin{cases} \frac{x}{2} & \text{if } x \text{ is even} \\ \frac{3x + 1}{2} & \text{if } x \text{ is odd} \end{cases}.$$

The $3x + 1$ conjecture states that all positive integers, when repeatedly iterated through this function, will enter the 2-1 cycle. We say that the integers one and two form a *cycle* because $T(1) = 2$ and $T(2) = 1$.

### 1.1 Definitions

In order to further this discussion of the conjecture, we must define a few terms that are commonly used when discussing this problem. The *trajectory* of an integer is the sequence that the integer follows through the $3x + 1$ function. The trajectory of $x$ can be represented as $\langle x, T^1(x), T^2(x), ... \rangle$. Then, another way to state the conjecture is that all positive integers have trajectories that contain two. Therefore, we usually look at the trajectory entries just up to the first occurrence of two. For example, the trajectory of 11 is $\langle 11, 17, 26, 13, 20, 10, 5, 8, 4, 2 \rangle$ and the trajectory of 27 is $\langle 27, 41, 62, 31, 47, 71, 107, 161, 242, 121, 182, 91, 137, \rangle$.
Using the trajectory, we can obtain the parity vector of an integer. Simply put, the entries of the parity vector are the elements of the trajectory reduced mod 2. We used the notation $P_n(x)$ to denote the first $n$ elements of the parity vector of $x$. We will call this the partial parity vector of length $n$. For example, $P_{10}(11) = \langle 1, 1, 0, 1, 0, 0, 1, 0, 0, 0 \rangle$.

The total stopping time $\sigma_\infty(x)$ of a positive integer $x$ is defined to be the number of iterations needed for that integer to reach one, or $\infty$ if the trajectory of $x$ does not contain one. In Chapter 2 we investigate the total stopping time through the construction of a graphical representation of the function, which is called a tree. This tree is based upon the inverse map of our function and helps us visualize the function. It also led us to identify several patterns found throughout the integer trajectories.

In Chapter 5, we define a run to be a set of consecutive integers that all have the same total stopping time. In order to study these runs we look at some special parity sequences known as blocks, strings and stems. We will define corresponding stems to be a pair of parity sequences of length $k$ such that if two consecutive integers have parity vectors that begin with this pair, then after $k$ iterations, their trajectories merge. See Definition 1 in Chapter 3 for a precise definition.

On the other hand, we will define corresponding blocks and corresponding
strings to be pairs of parity sequences of length $k$ such that if two consecutive integers have parity vectors that begin with this pair, then after $k$ iterations the next entries in their trajectories are consecutive. See Definitions 2 and 4 in Chapter 4 for further explanation. We use these parity sequences as a tool for investigating the behavior of the $3x + 1$ function, including finding long runs of consecutive integers.

Finally, in Chapter 6, we discuss what the future might hold for further research on the total stopping times of the Collatz Conjecture.
Chapter 2

Total Stopping Time

Recall that the total stopping time $\sigma_\infty(x)$ of a positive integer $x$ is defined to be the number of iterations needed for that integer to reach one. In order to get a better visual representation of total stopping time, we can construct a tree to represent the function. Since we conjecture that all positive integers have trajectories that contain 1, the integer 1 will be the root of our tree. In order to generate the tree by starting with 1, we must use the inverse map of our function. We define the inverse map as shown below.

$$T^{-1}(x) = \begin{cases} 2x, & \frac{2x-1}{3} \\ \end{cases}$$

We will call $T^{-1}_0(x) = 2x$ the even piece of our inverse map and $T^{-1}_1(x) = \frac{2x-1}{3}$ the odd piece. On the tree, the even inverse is represented by vertical connections between positive integers. In other words, the integer $T^{-1}_0(x)$ is drawn directly above the positive integer $x$. We will refer to these vertical connections as trunks of the tree. New branches coming off these trunks are found using the odd piece of the inverse map. A branch only occurs when $T^{-1}_1(x)$ is equal to a positive integer $y$. For example, if $x = 5$, $T^{-1}_1(5) = \frac{2 \cdot 5 - 1}{3} = \frac{9}{3} = 3$, and 3 is the next node in the tree.
Figure 2.1: $3x + 1$ Tree

integer. On the tree, $T_1^{-1}(x)$ is located above and to the right of $x$ and connected to $x$ with a diagonal line. For example, see Figure 2.1, above. Note $T_0^{-1}(8) = 16$ and $T_1^{-1}(8) = 5$.

Applegate and Lagarias [1] made the following observations; every integer that has a branch coming off it that contains positive integers divisible by 3 is congruent to 5 mod 9 and $T_1^{-1}(x) \in \mathbb{Z} \iff x \equiv 2 \mod 3$. Then, the following two lemmas detail patterns that are related to these findings.

Lemma 1. If $x \equiv 0 \mod 3$, then for all $m \in \mathbb{Z}^+$, $T_1^{-1}(2^m x) \notin \mathbb{Z}^+$.

Proof. Let $x \equiv 0 \mod 3$. Then for all positive integers $m$, $2^m x \equiv 0 \mod 3$. Therefore, $T_1^{-1}(2^m x) \notin \mathbb{Z}^+$. □
Lemma 2. If \( x \equiv 2 \mod 3 \), then for all odd integers \( p \geq 0 \), \( T_1^{-1}(2^p x) \not\in \mathbb{Z}^+ \) and for all even integers \( q \geq 0 \), \( T_1^{-1}(2^q x) \in \mathbb{Z}^+ \).

Proof. Let \( x \equiv 2 \mod 3 \).

- For all odd integers \( p \geq 0 \), \( 2^p x \equiv 1 \mod 3 \). Therefore, \( T_1^{-1}(2^p x) \not\in \mathbb{Z}^+ \).

- For all even integers \( q \geq 0 \), \( 2^q x \equiv 2 \mod 3 \). Therefore, \( T_1^{-1}(2^q x) \in \mathbb{Z}^+ \).

\( \square \)

The tree in Figure 2.1 shows all integers that have a total stopping time between zero and eight. This tree clearly shows us the trajectory followed by each of the integers contained in it. The tree is also helpful to see which integers are at a specific total stopping time. All integers of the same total stopping time line up in a horizontal row. The above lemmas allow us to write an algorithm to count how many integers are at a given total stopping time. The code for this algorithm is found in Appendix A.

Using this algorithm, we found \( N_r \), the number of integers with a given total stopping time as shown in Table 2.1, on page 13.

In the next chapter, we will take a closer look at integers with the same total stopping time. Specifically, we will turn our focus toward integers that have the same total stopping time as an adjacent integer.
Table 2.1: Total Stopping Times

<table>
<thead>
<tr>
<th>$\sigma_{\infty}$</th>
<th>$N_{\sigma}$</th>
<th>$\sigma_{\infty}$</th>
<th>$N_{\sigma}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>13</td>
<td>31</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>14</td>
<td>39</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>15</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>16</td>
<td>68</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>17</td>
<td>91</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>18</td>
<td>120</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>19</td>
<td>159</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>20</td>
<td>211</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>21</td>
<td>282</td>
</tr>
</tbody>
</table>
Chapter 3

Consecutive Integers

One feature of the tree discussed in the previous chapter that has been widely noticed is the large number of occurrences of consecutive integers with equal total stopping time. Garner [4] stated that “about 70% of the integers less than 1024 have the same height [total stopping time] as an adjacent integer; we believe that a majority of the positive integers have that property.” We are curious about why these integers behave in this way, and in order to better understand the function and its patterns, we start by focusing on these consecutive integers that have the same total stopping time. If we can understand their behavior, we may gain insight into why other integers do not display this pattern.

From our tree it is obvious that for a pair of consecutive integers to have the same total stopping time, they must have trajectories that coincide after the same number of iterations. They may coincide for the first time at 8, or perhaps at an integer with a larger total stopping time. Using this information, Garner looked for patterns in the parity sequences of consecutive integers of the same total stopping time. As a result of this work, he defined parity sequences of the following form
as stems.

\[ s_i = \overline{01\ldots101} \text{ or } s'_i = \overline{11\ldots100} \text{ with } i \geq 0. \]

Garner conjectured that \( s_i \) and \( s'_i \) were the only possible form for corresponding stems. In addition, Garner conjectured that any consecutive integers must contain stems. Since we have not found any evidence that this is not the case, we formally state this conjecture below.

**Conjecture 1.** (Garner) Any pair of consecutive integers with equal total stopping time contain stems of the form \( s_i \) and \( s'_i \) in their parity vectors.

In order to explore the properties of consecutive integers, we must first consider what happens to any integer when it goes through a given sequence of odd and even pieces of our function regardless of the integer’s parity. We define \( T_0(x) = \frac{x}{2} \) and \( T_1(x) = \frac{3x+1}{2} \). Given the partial parity vector \( v = \langle v_1, v_2, \ldots, v_k \rangle \), the result of applying \( v \) to \( x \), denoted as \( T_v(x) \), is the result of the composition \( T_{v_k}(T_{v_{k-1}}(\ldots(T_{v_1}(x))\ldots)) \). For example,

\[ T_{\langle1,0,0\rangle}(13) = T_0(T_0(T_1(13))) = \frac{\frac{3(13)+1}{2}}{2} = 5. \]

We generalize Garner’s stems as follows:

**Definition 1.** A pair of parity sequences \( s \) and \( s' \) of length \( k \) are corresponding stems if for any integer \( x \), \( T_s(x) = T_{s'}(x+1) \) and for any initial subsequences \( v \) and \( v' \) of \( s \) and \( s' \) of equal length, \( |T_v(x) - T_{v'}(x+1)| \neq 1 \) and \( T_v(x) \neq T_{v'}(x+1) \).

Note that the second half of our definition ensures that our corresponding
stems are not made up of smaller corresponding stems or of blocks and strings. (See the next chapter for definitions of our blocks and strings.)

Clearly Garner’s Stems $s_i$ and $s'_i$ are corresponding stems in the above sense. To see this, given $s_i$ and $s'_i$, note $T_{s_i}(x) = T_{s'_i}(x + 1)$. Now proceed through $s_i$ and $s'_i$ backwards, applying appropriate inverse functions to see that their difference is never 1 and that they are never equal.

The result of applying a parity sequence $v = \langle v_1, v_2, ..., v_k \rangle$ to a number $x$ will be

$$\frac{3^n}{2^k}x + c + \left( \sum_{i=0}^{k-1} \frac{1}{2^{k-i}} \left( 3^{(v_k+v_{k-1}+...+v_{i+1})}v_{i+1} \right) \right)$$

(3.1)

where the sequence contains exactly $n$ ones and has length $k$. The proof of this fact involves basic distribution and simplification.

**Theorem 1.** Corresponding stems contain an equal number of ones.

**Proof.** Let $x$ and $x + 1$ be real numbers and let $s$ and $s'$ be corresponding stems with length $k$. Assume that $s$ has $n$ ones, and that $s'$ has $m$ ones. From Equation 3.1, we know that $T_{s_i}(x) = \frac{3^n}{2^k}x + c$, where $c$ does not depend on $x$ and that $T_{s'_i}(x + 1) = \frac{3^m}{2^k}(x + 1) + d$, where $d$ does not depend on $x$. Then, from the definition of corresponding stems, we know that $T_{s_i}(x) = T_{s'_i}(x + 1)$ for all $x \in \mathbb{R}$. Then we have,

$$\frac{3^n}{2^k}x + c = \frac{3^m}{2^k}(x + 1) + d$$

$$\frac{3^n}{2^k}x + c = \frac{3^m}{2^k}x + \frac{3^m}{2^k} + d.$$  

(3.2)
Now, if we look at the case where $x = 0$ then,

$$d = c - \frac{3^m}{2^k}.$$  

Next, we substitute this for $d$ in Equation 3.2 and obtain the following:

$$\frac{3^n}{2^k} x + c = \frac{3^m}{2^k} x + \frac{3^m}{2^k} + c - \frac{3^m}{2^k}.$$

Then, we simplify,

$$\frac{3^n}{2^k} x + c = \frac{3^m}{2^k} x + c
\frac{3^n}{2^k} x = \frac{3^m}{2^k} x
3^n = 3^m.$$

Therefore, $n = m$, and so $s$ and $s'$ must have an equal number of ones. □

From our definition of corresponding stems and Theorem 1, we can form an equation to test whether or not a pair of parity sequences is a set of corresponding stems.

**Corollary 1.** Parity sequences $\langle v_1, v_2, ..., v_k \rangle$ and $\langle v'_1, v'_2, ..., v'_k \rangle$ are corresponding stems if and only if $v_1 \neq v'_1$ and

$$1 = \left| \sum_{i=1}^{k} (2^{i-1}) \left( \frac{v_i}{3^{(v_1+v_2+...+v_{i-1}+v'_i)}} - \frac{v'_i}{3^{(v'_1+v'_2+...+v'_{i-1}+v_i)}} \right) \right|.$$  

**Proof.** Recall, $v = \langle v_1, v_2, ..., v_k \rangle$ and $v' = \langle v'_1, v'_2, ..., v'_k \rangle$ are corresponding stems if and only if $T_v(x) = T'_{v'}(x + 1)$. So,

$$\frac{3^{(v_1+v_2+...+v_{k-1})}}{2^k}(x) + \sum_{i=0}^{k-1} \frac{1}{2^{k-i}} \left( 3^{(v_1+v_2+...+v_{i+1})}v_{i+1} \right) =$$
\[
\frac{3(v_1 + v_2 + \cdots + v_i)}{2^k} (x + 1) + \sum_{i=0}^{k-1} \frac{1}{2^{k-i}} \left(3^{(v_i + v_{i-1} + \cdots + v_{2i})} v_{i+1}^{'}\right).
\]

If \(v\) and \(v^{'}\) are stems then, from Theorem 1, we know that \(v\) and \(v^{'}\) must have the same number of ones, \(n\), so we can simplify our expression to the following:

\[
\sum_{i=0}^{k-1} \frac{1}{2^{k-i}} \left(3^{(v_i + v_{i-1} + \cdots + v_{2i})} v_{i+1}^{'}\right) = \frac{3^n}{2^k} + \sum_{i=0}^{k-1} \frac{1}{2^{k-i}} \left(3^{(v_i^{'} + v_{i-1} + \cdots + v_{2i})} v_{i+1}^{'}\right).
\]

So,

\[
\sum_{i=0}^{k-1} (2^i) \left(3^{(v_i + v_{i-1} + \cdots + v_{2i})} v_{i+1}^{'} - 3^{(v_i^{'} + v_{i-1} + \cdots + v_{2i})} v_{i+1}^{'}\right) = 3^n.
\]

Therefore, \(v\) and \(v^{'}\) are corresponding stems if and only if:

\[
1 = \sum_{i=0}^{k-1} (2^i) \left(\frac{3^{(v_i + v_{i-1} + \cdots + v_{2i})} v_{i+1}^{'} - 3^{(v_i^{'} + v_{i-1} + \cdots + v_{2i})} v_{i+1}^{'}}{3^n}\right)
\]

\[
= \sum_{i=0}^{k-1} (2^i) \left(\frac{v_{i+1}^{'} - v_{i+1}^{}}{3^{(v_i + v_{i-1} + \cdots + v_{2i})} - 3^{(v_i^{'} + v_{i-1} + \cdots + v_{2i})}}\right).
\]

If \(v\) and \(v^{'}\) satisfy the equation in the statement of the corollary, then we can work backwards through the steps shown above.

Since we do not know in what order \(v\) and \(v^{'}\) will be given to us, we must take the absolute value and obtain our result.

\[
\square
\]

It is easy to see that \(s_i\) and \(s_i^{'}\) satisfy the corollary. Now, we will look specifically at the use of Garner’s stems.
3.1 Applying Garner’s Stems

In this section we show that if an integer \( m \) has a parity vector beginning with \( s_i \), then the integer \( m + 1 \) must have a parity vector beginning with \( s'_i \). The following lemma and theorem combine to give this result.

**Lemma 3.** If \( P_{i+1}(m) = \langle 0, 1, 1, ..., 1 \rangle \), then \( P_{i+1}(m + 1) = \langle 1, 1, 1, ..., 1 \rangle \) and \( T^{i+1}(m + 1) = 3T^{i+1}(m) + 2 \).

**Proof.** We proceed by induction. Base case: \( i = 0 \)

Assume \( P_1(m) = \langle 0 \rangle \). Then we know that \( m \equiv 0 \mod 2 \). This means that \( m + 1 \equiv 1 \mod 2 \). So, \( P_1(m + 1) = \langle 1 \rangle \). Therefore the statement is true for \( i = 0 \).

Assume the statement is true for \( i = n - 1 \).

Next, suppose \( i = n \). Then, we are assuming \( T^n(m) \equiv 1 \mod 2 \). From Equation 3.1,

\[
T^n(m) = \frac{3^{n-1}}{2^n} (m) + \frac{1}{2} + \sum_{i=1}^{n-2} \left( \frac{1}{2} \right)^{n-i} 3^{n-i-1}
\]

\[
= \frac{1}{3} \left( \frac{3}{2} \right)^n m + \frac{1}{2} + \frac{1}{3} \sum_{i=1}^{n-2} \left( \frac{3}{2} \right)^{n-i}.
\]

Since we are assuming the statement is true for \( i = n - 1 \), \( P_{n}(m+1) = \langle 1, 1, 1, ..., 1 \rangle \).
\[
T^n(m + 1) = \frac{3^n}{2^n}(m + 1) + \frac{1}{2} + \sum_{i=0}^{n-2} \left( \frac{1}{2} \right)^{n-i} 3^{n-i-1}
\]
\[
= 3 \left( \frac{1}{3} \left( \frac{3}{2} \right)^n m + \frac{1}{2} + \frac{1}{3} \sum_{i=1}^{n-2} \left( \frac{3}{2} \right)^{n-i} \right) - 2 \frac{1}{3} \sum_{i=1}^{n-2} \left( \frac{3}{2} \right)^{n-i} + \frac{1}{3} \left( \frac{3}{2} \right)^n + \frac{1}{3} \left( \frac{3}{2} \right) - 1
\]
\[
= 3 \left( T^n(m) \right) - \frac{2}{3} \sum_{i=1}^{n-2} \left( \frac{3}{2} \right)^{n-i} + \frac{4}{3} \left( \frac{3}{2} \right)^n - 1
\]
\[
= 3 \left( T^n(m) \right) + 2.
\]

Then,
\[
T^n(m + 1) = 3 \left( T^n(m) \right) + 2 \equiv 1 \mod 2.
\]

Therefore, \( P_{n+1}(m + 1) = \langle 1 \ 1 \ldots \ 1 \rangle \) which concludes our proof. \( \square \)

**Theorem 2.** Given an integer \( m > 0 \), if \( P_{i+3}(m) = \langle 0 \ 1 \ 1 \ldots \ 1 \ 0 \ 1 \rangle \), then \( P_{i+3}(m + 1) = \langle 1 \ 1 \ 1 \ldots \ 1 \ 0 \ 0 \rangle \) and \( T^{i+3}(m) = T^{i+3}(m + 1) \).

**Proof.** Suppose \( P_{i+3}(m) = \langle 0 \ 1 \ 1 \ldots \ 1 \ 0 \ 1 \rangle \). Then, \( T^{i+1}(m) \equiv 0 \mod 2 \) and \( T^{i+2}(m) \equiv 1 \mod 2 \). So, we have

\[
T^{i+2}(m) = \frac{T^{i+1}(m)}{2} \equiv 1 \mod 2
\]
\[
T^{i+3}(m) = \frac{3T^{i+1}(m)}{4} + \frac{1}{2}.
\]

Then from the previous lemma, \( P_{i+1}(m + 1) = \langle 1 \ 1 \ldots \ 1 \rangle \) and \( T^{i+1}(m + 1) = 3T^{i+1}(m) + 2 \). Now, since \( T^{i+1}(m) \equiv 0 \mod 2 \), \( T^{i+1}(m + 1) \equiv 0 \mod 2 \). Therefore, \( P_{i+2}(m + 1) = \langle 1 \ 1 \ldots \ 1 \ 0 \rangle \).
Then
\[ T_0(T^{i+1}(m + 1)) = \frac{3T^{i+1}(m) + 2}{2} = 3T^{i+2}(m) + 1 \]

Since \( T^{i+2}(m) \equiv 1 \mod 2 \),
\[ T_0(T^{i+1}(m + 1)) \equiv 0 \mod 2. \]

Therefore, \( P_{i+3}(m + 1) = \langle 1 \underbrace{1 \ldots 1}_{i 1's} 00 \rangle. \)

Also,
\[
T^{i+3}(m + 1) = T_0(T_0(T^{i+1}(m + 1)))
= \frac{T^{i+1}(m + 1)}{4}
= \frac{3T^{i+1}(m) + 2}{4}
= T^{i+3}(m).
\]

So, if \( P_{i+3}(m) = \langle 0 \underbrace{1 \ldots 1}_{i 1's} 01 \rangle \), then \( P_{i+3}(m + 1) = \langle 1 \underbrace{1 \ldots 1}_{i 1's} 00 \rangle \) and \( T^{i+3}(m) = T^{i+3}(m + 1) \).

\[
\square
\]

### 3.2 Consecutive Triples

Suppose that we are given two consecutive integers that have the same total stopping time whose parity vectors contain stems in the first \( i + 3 \) entries. If we apply the even inverse function to both of these integers, we will get two integers at
the same total stopping time that have a difference of two. Our theorem below states that the missing integer between them is also at the same total stopping time, forming a consecutive triple.

**Theorem 3.** Given consecutive positive integers, \( n \) and \( n + 1 \), with total stopping time \( k \), \( P_{i+3}(n) = s_i \), and \( P_{i+3}(n + 1) = s'_i \) where \( i \geq 1 \), there exists a triple of consecutive integers, \( m, m + 1, m + 2 \), of total stopping time \( k + 1 \) such that \( m = 2n \), \( m + 2 = 2n + 2 \) and \( P_3(m + 1) = \langle 100 \rangle \).

Proof. Suppose \( P_{i+3}(n) = s_i = \langle 0\overbrace{11\ldots1}^{i}\overbrace{01}^{i+1} \rangle \) and so if we let \( m = 2n \) then \( P_{i+4}(m) = \langle 0\overbrace{s_i}^{i+1} \rangle = \langle 0\overbrace{01\ldots1}^{i+1}01 \rangle = \langle \overbrace{11\ldots1}^{i+1}01 \rangle \). From Theorem 2, we know that \( P_3(m + 1) = \langle 100 \rangle = \langle \overbrace{s'_i}^{i+1} \rangle \) and so \( m \) and \( m + 1 \) have coinciding trajectories after 3 entries. Since \( T^{k+1}(m) = 1 \), \( m + 1 \) will have the same total stopping time as \( m \). We already know that \( m \) and \( m + 2 \) have the same total stopping time, therefore integers \( m, m + 1, m + 2 \) form a consecutive triple of the same total stopping time.

We now know that for every pair of consecutive integers of the same total stopping time that contains our stems as the first \( i + 3 \) entries of their parity vector, there exists a consecutive triple at the next level of the tree.

As a result of this chapter, we have some basic ideas of what makes a set of trajectories coincide for consecutive integers. In the next chapter, we will explore parity sequences that take a pair of consecutive integers to another pair of consecutive integers. These parity sequences will be used in conjunction with our stems in order to find more pairs of consecutive integers of the same total stopping time.

22
Chapter 4

Blocks and Strings

In the previous chapter, we discussed a special type of parity sequence called stems. Recall that when a set of these corresponding stems is applied to a pair of real numbers $x$ and $x + 1$, the result of both applications is the same real number $y$. We will now look at two more special types of parity sequences that we will call blocks and strings. These blocks and strings will take real numbers $x$ and $x + 1$ to some real numbers $y$ and $y + 1$ (not necessarily respectively). Blocks and strings are very helpful because we can append them to a parity sequence before a stem, and doing so will give us another set of consecutive integers of the same total stopping time.

4.1 Corresponding Blocks and Strings

Definition 2. A block is a pair of parity sequences of length $k$, $b$ and $b'$, such that for all positive integers $x$, $T_b(x) + 1 = T_{b'}(x + 1)$ and for any initial subsequences $v$ and $v'$ of $b$ and $b'$ of equal length, $|T_b(x) - T_{b'}(x + 1)| \neq 1$. 
Therefore, we can think of blocks as pairs of parity sequences that take two consecutive integers to two new consecutive integers, where the smaller of the first pair is mapped to the smaller of the second pair. Note that the second half of our above definition ensures that our blocks are not made up of smaller blocks. Those instances are defined below.

**Definition 3.** A block prefix is a pair of parity sequences of length $k$, $b$ and $b'$, such that for all positive integers $x$, $T_b(x) + 1 = T_{b'}(x + 1)$.

Garner’s Theorem 3 states that if $\langle v \rangle$ and $\langle v' \rangle$ are the parity vectors of a pair of consecutive integers of the same total stopping time, and $p$ and $p'$ are any prefix and its corresponding prefix, then $\langle pv \rangle$ and $\langle p'v' \rangle$ are parity vectors of a pair of consecutive integers of the same total stopping time. He provides no proof of this theorem, and we do not believe that we can state with certainty that the integers with the modified parity vectors will be of the same total stopping time. We know $\langle pv \rangle$ and $\langle p'v' \rangle$ are parity vectors of integers, but there is no guarantee that those integers enter the 2-1 cycle since the Collatz Conjecture has yet to be proven. Therefore, the following theorem is a weakened version of Garner’s Theorem 3 in that it does not guarantee the same total stopping time of the consecutive integers found. If $v$ and $v'$ contain stems, then the modified vectors would correspond to consecutive integers of the same total stopping time. However, as mentioned in the previous chapter, we do not know that having consecutive integers of the same total stopping time implies that their parity vectors contain stems.

**Lemma 4.** Let $v$ be a parity vector of length $k$ and $y$ a positive integer. Then $T_v(y)$
is an integer if and only if \( P_k(y) = v \).

**Proof.** Clearly if \( P_k(y) = v \), this implies that \( T_v(y) \in \mathbb{Z} \).

Note that \( T_1(x) \in \mathbb{Z} \) only when \( x \in \mathbb{Z} \) and \( T_0(x) \in \mathbb{Z} \) only when \( x \in \mathbb{Z} \). So if \( T_v(y) \in \mathbb{Z} \), then clearly \( P_k(y) = v \). □

**Theorem 4.** If \( \langle v \rangle \) and \( \langle v' \rangle \) are the parity vectors of length \( k \) of a pair of positive consecutive integers of the same total stopping time \( k \), \( \langle v \rangle \) corresponding to the smaller of the two, and \( p \) and \( p' \) are any block prefix of length \( s \), then there exist some positive integers \( y \) and \( y + 1 \) with \( P_{k+s}(y) = \langle pv \rangle \) and \( P_{k+s}(y + 1) = \langle p'v' \rangle \).

**Proof.** Given \( n \) and \( n + 1 \) of the same total stopping time, choose any positive integer \( y \) such that \( P_{k+s}(y) = \langle pv \rangle \). Then, \( P_s(y + 1) = \langle p' \rangle \) because we know that \( T_p(y) \in \mathbb{Z} \) and \( T_{p'}(y + 1) = T_p(y) + 1 \), by the definition of a block prefix, and so \( T_{p'}(y + 1) \in \mathbb{Z} \) which implies that \( p' \) is a parity vector of \( y + 1 \) by Lemma 4. Then, from the algorithm given by Terras [6], \( T_p(y) = n + a(2^k) \) for some positive integer \( a \). Using our definition of block prefix, \( T_{p'}(y + 1) = T_p(y) + 1 = n + 1 + a(2^k) \) for some positive integer \( a \). So, the parity vector of \( T_{p'}(y + 1) \) and the parity vector of \( n + 1 \) agree for \( k \) steps. Then the parity vector of \( T_{p'}(y + 1) \) is \( \langle v' \rangle \). This implies that the parity vector of \( y + 1 \) is \( \langle p'v' \rangle \). □

Using the definition of blocks and an extensive computer search, we can generate the blocks shown in Table 4.1 on page 26. This table contains the complete list of unique blocks (no prefixes) up to length 11. Those signified with asterisks are blocks that Garner [4] did not report.
Table 4.1: Corresponding Blocks

<table>
<thead>
<tr>
<th>$b$</th>
<th>$b'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>01</td>
</tr>
<tr>
<td>11000</td>
<td>00101</td>
</tr>
<tr>
<td>1101000</td>
<td>0011001</td>
</tr>
<tr>
<td>11100000</td>
<td>00011001</td>
</tr>
<tr>
<td>11001000</td>
<td>00100101</td>
</tr>
<tr>
<td>11011000</td>
<td>00111100</td>
</tr>
<tr>
<td>111001000</td>
<td>000111100</td>
</tr>
<tr>
<td>110101000</td>
<td>001101001</td>
</tr>
<tr>
<td>*1110001000</td>
<td>*0001101001</td>
</tr>
<tr>
<td>1101001000</td>
<td>0011000101</td>
</tr>
<tr>
<td>*1100101000</td>
<td>*0010011001</td>
</tr>
<tr>
<td>1101011000</td>
<td>001101100</td>
</tr>
<tr>
<td>*11001100000</td>
<td>*00100011001</td>
</tr>
<tr>
<td>11011100000</td>
<td>00111001001</td>
</tr>
<tr>
<td>1110000001000</td>
<td>00011000101</td>
</tr>
<tr>
<td>*11001001000</td>
<td>*00100100101</td>
</tr>
<tr>
<td>11010101000</td>
<td>00110101001</td>
</tr>
<tr>
<td>*11100011000</td>
<td>*00011011100</td>
</tr>
<tr>
<td>*11001011000</td>
<td>*00100111100</td>
</tr>
<tr>
<td>*11011010010</td>
<td>*00111101100</td>
</tr>
<tr>
<td>*11011001010</td>
<td>*00111101100</td>
</tr>
</tbody>
</table>

26
**Definition 4.** A string is a pair of parity sequences of length $k$, $q$ and $q'$, such that for all positive integers $x$, $T_q(x) = T_{q'}(x + 1) + 1$ and for any initial subsequences $v$ and $v'$ of $q$ and $q'$ of equal length, $|T_q(x) - T_{q'}(x + 1)| \neq 1$.

Then, because of the specificity of the second half of our definition, we must also define the following, more general case.

**Definition 5.** A string prefix is a pair of parity sequences of length $k$, $q$ and $q'$, such that for all positive integers $x$, $T_q(x) = T_{q'}(x + 1) + 1$.

Theorem 5 below is similar to Garner’s Theorem 4, as our previous Theorem was similar to Garner’s Theorem 3, in that we cannot say that the new set of consecutive integers found will have the same total stopping time.

**Theorem 5.** If $\langle v \rangle$ and $\langle v' \rangle$ are the parity vectors of length $k$ of a pair of consecutive integers of the same total stopping time $k$, $v$ corresponding to the smaller of the two, and $q$ and $q'$ are a string prefix of length $s$, then there exist some positive integers $m$ and $m + 1$ with $P_{k+s}(m) = \langle qv' \rangle$ and $P_{k+s}(m + 1) = \langle q'v \rangle$.

The proof of this theorem is analogous to the proof of Theorem 4.

Using the definition of strings, we can construct a table of strings. Table 4.2, found on page 28, shows all strings up to length 11, with newly found strings signified by an asterisk.

Garner [4] has conjectured that the list of all blocks and strings is infinite. We have not found any evidence to suggest that this is not the case.
Table 4.2: Corresponding Strings

<table>
<thead>
<tr>
<th>q</th>
<th>q'</th>
</tr>
</thead>
<tbody>
<tr>
<td>000011</td>
<td>101000</td>
</tr>
<tr>
<td>000110</td>
<td>101100</td>
</tr>
<tr>
<td>010001</td>
<td>110100</td>
</tr>
<tr>
<td>010011</td>
<td>110110</td>
</tr>
<tr>
<td>011000</td>
<td>111010</td>
</tr>
<tr>
<td>*000111</td>
<td>*101100</td>
</tr>
<tr>
<td>000101</td>
<td>101001</td>
</tr>
<tr>
<td>011001</td>
<td>111010</td>
</tr>
<tr>
<td>011100</td>
<td>111010</td>
</tr>
<tr>
<td>*0001111</td>
<td>*1011011</td>
</tr>
<tr>
<td>0001011</td>
<td>10111000</td>
</tr>
<tr>
<td>0100111</td>
<td>11011010</td>
</tr>
<tr>
<td>*00011001</td>
<td>*10110010</td>
</tr>
<tr>
<td>01000101</td>
<td>11010010</td>
</tr>
<tr>
<td>*011100111</td>
<td>*1111011</td>
</tr>
<tr>
<td>000101111</td>
<td>101110100</td>
</tr>
<tr>
<td>*010011111</td>
<td>*1101111</td>
</tr>
<tr>
<td>*000111101</td>
<td>*10110101</td>
</tr>
<tr>
<td>*0001011101</td>
<td>*101101010</td>
</tr>
<tr>
<td>0100101101</td>
<td>1101101000</td>
</tr>
<tr>
<td>0110011101</td>
<td>1101101000</td>
</tr>
<tr>
<td>*01110011101</td>
<td>*111101100</td>
</tr>
<tr>
<td>*01111000111</td>
<td>*1111101</td>
</tr>
<tr>
<td>00010111100</td>
<td>1011100100</td>
</tr>
<tr>
<td>*011100111100</td>
<td>*1101100100</td>
</tr>
<tr>
<td>*01001111101</td>
<td>*101101010</td>
</tr>
<tr>
<td>*0001111001</td>
<td>*101101001</td>
</tr>
<tr>
<td>01100010110</td>
<td>1110100100</td>
</tr>
<tr>
<td>*01110011011</td>
<td>*111110110</td>
</tr>
<tr>
<td>*01111100011</td>
<td>*111110100</td>
</tr>
<tr>
<td>*011111100011</td>
<td>*1111110100</td>
</tr>
</tbody>
</table>
4.2 Block and String Structure

Since block and string prefixes have proven to be important to discovering new sets of consecutive integers of the same total stopping time, we want to study their structure in order to help us understand how they are formed and why they work. The following sections showcase our results in these efforts.

4.2.1 Equal number of ones

Initial looks at Tables 4.1 and 4.2 reveal that all sets of blocks and strings are simply permutations of one another. That is, each parity sequence in the pair that compose a block or string have an equal number of ones and zeros. This is illustrated in the following theorem.

**Theorem 6.** Block and string prefixes contain an equal number of ones.

The proof of this theorem is analogous to the proof of Theorem 1, found in Chapter 3.

4.2.2 Difference Between Sums

From our definitions of block and string prefixes, Theorem 6 and Equation 3.1, we can form an equation that will tell us whether or not a pair of parity sequences is a block or string prefix, as shown in the following corollary.

**Corollary 2.** Parity sequences \(v = \langle v_1, v_2, ..., v_k \rangle\) and \(v' = \langle v'_1, v'_2, ..., v'_k \rangle\) are block or string prefixes if and only if their parity sequences have the same number of 1’s.
and

\[ 1 = \left| \sum_{i=0}^{k-1} \frac{1}{2k-i} \left( 3^{(v'_k + v'_{k-1} + \ldots + v'_2 + v'_1)} v'_{i+1} - 3^{(v_k + v_{k-1} + \ldots + v_2 + v_1)} v_{i+1} \right) + \frac{3^{(v'_k + v'_{k-1} + \ldots + v'_2 + v'_1)}}{2^k} \right|. \]

**Proof.** Recall, \( v = \langle v_1, v_2, \ldots, v_k \rangle \) and \( v' = \langle v'_1, v'_2, \ldots, v'_k \rangle \) are block prefixes if and only if \( T_v(x) + 1 = T_{v'}(x + 1) \). From Equation 3.1, \( v \) and \( v' \) are block prefixes if and only if

\[
\left( \frac{3^{(v_k + v_{k-1} + \ldots + v_2 + v_1)}}{2^k} (x) + \sum_{i=0}^{k-1} \frac{1}{2k-i} \left( 3^{(v_k + v_{k-1} + \ldots + v_2 + v_1)} v_{i+1} \right) \right) + 1 = \frac{3^{(v'_k + v'_{k-1} + \ldots + v'_2 + v'_1)}}{2^k} (x + 1) + \sum_{i=0}^{k-2} \frac{1}{2k-i} \left( 3^{(v'_k + v'_{k-1} + \ldots + v'_2 + v'_1)} v'_{i+1} \right).
\]

Now, assume \( v \) and \( v' \) are block prefixes. Then, from Theorem 6, we know that \( v \) and \( v' \) must have the same number of ones so we can simplify our equation from above to the following:

\[
1 + \sum_{i=0}^{k-1} \frac{1}{2k-i} \left( 3^{(v'_k + v'_{k-1} + \ldots + v'_2 + v'_1)} v'_{i+1} \right) = \frac{3^{(v'_k + v'_{k-1} + \ldots + v'_2 + v'_1)}}{2^k} + \sum_{i=0}^{k-1} \frac{1}{2k-i} \left( 3^{(v'_k + v'_{k-1} + \ldots + v'_2 + v'_1)} v'_{i+1} \right).
\]

So,

\[
1 = \sum_{i=0}^{k-1} \frac{1}{2k-i} \left( 3^{(v'_k + v'_{k-1} + \ldots + v'_2 + v'_1)} v'_{i+1} - 3^{(v_k + v_{k-1} + \ldots + v_2 + v_1)} v_{i+1} \right) + \frac{3^{(v'_k + v'_{k-1} + \ldots + v'_2 + v'_1)}}{2^k}.
\]

On the other hand, if \( v \) and \( v' \) have an equal number of 1’s and the equality in our corollary holds, then do the steps above backwards and we have our result.

Similarly, \( v \) and \( v' \) are string prefixes if and only if:

\[
-1 = \sum_{i=0}^{k-1} \frac{1}{2k-i} \left( 3^{(v'_k + v'_{k-1} + \ldots + v'_2 + v'_1)} v'_{i+1} - 3^{(v_k + v_{k-1} + \ldots + v_2 + v_1)} v_{i+1} \right) + \frac{3^{(v'_k + v'_{k-1} + \ldots + v'_2 + v'_1)}}{2^k}.
\]

30
Therefore, \( v \) and \( v' \) are block or string prefixes if and only if:

\[
1 = \left| \frac{\sum_{i=0}^{k-1} \frac{1}{2^{k-i}} \left( 3^{(v'_k + v'_{k-1} + \ldots + v'_2)}v'_{i+1} - 3^{(v_k + v_{k-1} + \ldots + v_2)}v_{i+1} \right) + \frac{3^{(v'_k + v_{k-1} + \ldots + v'_2)}}{2^k}} \right|.
\]

\( \Box \)

### 4.2.3 Switching 0’s and 1’s

Since we discovered earlier that corresponding blocks and strings are just permutations of each other, we began to look at what would happen if we took a parity sequence and moved the ones around to form a new permutation of the original. If we take a single 1 and have it switch positions with an adjacent 0, we have the following result.

**Lemma 5.** Given a parity sequence \( v = \langle v_1, v_2, \ldots, v_{k-1}, v_k \rangle \) such that \( v_a = 1 \) and \( v_{a+1} = 0 \) for some \( a < k \), if we form a new parity sequence \( v' \) by switching \( v_a \) and \( v_{a+1} \), then \( T_{v'}(x) - T_v(x) = \Delta \) where

\[
\Delta = \frac{3^{v_a + v_{a+3} + \ldots + v_k}}{2^{k-a+1}}.
\]

**Proof.** Given \( v \) as stated above, we will use

\[
v' = \langle v_1, v_2, \ldots v_{a-1}, v_{a+1}, v_a, v_{a+2}, \ldots, v_{k-1}, v_k \rangle
\]

and we will let \( n \) be the number of ones in \( v \) and \( v' \). From Equation 3.1 and Theorem 6 we know that when \( v \) is applied to a number \( x \)

\[
T_v(x) = \frac{3^n}{2^k}x + \sum_{i=0}^{k-1} \frac{3^{(v_k + v_{k-1} + \ldots + v_2)}}{2^{k-i}}v_{i+1}.
\]

31
Similarly, the result of $\nu'$ being applied to a number $x$, is

$$T_{\nu'}(x) = \frac{3^n}{2^k}x + \sum_{i=0}^{k-1} \frac{3^{(v_k+v_{k-1}+\ldots+v_{a+2})}}{2^k}v_{i+1}.$$  

- Case 1: If $a < k - 1$, then

$$T_{\nu}(x) = \frac{3^n}{2^k}x + \sum_{i=0}^{a-2} \frac{3^{(v_k+v_{k-1}+\ldots+v_{i+2})}}{2^k}v_{i+1} + \frac{3^{(v_k+v_{k-1}+\ldots+v_{a+1})}}{2^{k-a+1}}v_a$$

$$+ \frac{3^{(v_k+v_{k-1}+\ldots+v_{a+2})}}{2^{k-a}}v_{a+1} + \sum_{i=a+1}^{k-1} \frac{3^{(v_k+v_{k-1}+\ldots+v_{i+2})}}{2^k}v_{i+1}$$

and

$$T_{\nu'}(x) = \frac{3^n}{2^k}x + \sum_{i=0}^{a-2} \frac{3^{(v_k+v_{k-1}+\ldots+v_{i+2})}}{2^k}v_{i+1} + \frac{3^{(v_k+v_{k-1}+\ldots+v_{a+1})}}{2^{k-a+1}}v_a$$

$$+ \frac{3^{(v_k+v_{k-1}+\ldots+v_{a+2})}}{2^{k-a}}v_a + \sum_{i=a+1}^{k-1} \frac{3^{(v_k+v_{k-1}+\ldots+v_{i+2})}}{2^k}v_{i+1}$$

The difference between $T_{\nu'}(x)$ and $T_{\nu}(x)$ can be found as follows:

$$\Delta = \frac{3^{(v_k+v_{k-1}+\ldots+v_{a+2})}}{2^{k-a+1}}v_{a+1} + \frac{3^{(v_k+v_{k-1}+\ldots+v_{a+2})}}{2^{k-a}}v_a$$

$$- \frac{3^{(v_k+v_{k-1}+\ldots+v_{a+2})}}{2^{k-a+1}}v_a - \frac{3^{(v_k+v_{k-1}+\ldots+v_{a+2})}}{2^{k-a}}v_{a+1}.$$  

Then since $v_a = 1$ and $v_{a+1} = 0$, we have

$$\Delta = \frac{3^{(v_k+v_{k-1}+\ldots+v_{a+2})}}{2^{k-a+1}}.$$  

- Case 2: If $a = k - 1$, then $T_{\nu}(x)$ can be written as follows since $k = a + 1$:

$$\frac{3^n}{2^{a+1}}x + \sum_{i=0}^{a-2} \frac{3^{(v_{a+1}+v_a+v_{a-1}+\ldots+v_{i+2})}}{2^{a+1-i}}v_{i+1} + \frac{3^{v_{a+1}}}{2^a}v_a + \frac{1}{2}v_{a+1}.$$
Then,
\[ v' = \langle v_1, v_2, \ldots, v_{a-1}, v_{a+1}, v_a \rangle \]

When \( v' \) is applied to \( x \), then the result is
\[ T_{v'}(x) = \frac{3^a}{2^{a+1}}x + \sum_{i=0}^{a-2} \frac{3^{v_a+v_{a+1}+v_{a-1}+\ldots+v_i+2}}{2^{a+1-i}}v_{i+1} + \frac{3^{v_a}}{2^a}v_{a+1} + \frac{1}{2}v_a \]

Then the difference between the results simplifies to
\[ \Delta = \frac{3^{v_a}}{4}v_{a+1} + \frac{1}{2}v_a - \frac{3^{v_{a+1}}}{2^2}v_a - \frac{1}{2}v_{a+1} \]
\[ = v_{a+1}\left(\frac{3^{v_a}}{4} - \frac{1}{2}\right) + v_a\left(\frac{1}{2} - \frac{3^{v_{a+1}}}{2^2}\right) \]

Then, since \( v_a = 1 \) and \( v_{a+1} = 0 \) we have,
\[ \Delta = \frac{3^{v_{a+2}+v_{a+3}+\ldots+v_k}}{2^{k-a+1}}. \]

\[ \square \]

Note that in our above lemma, \( \Delta \) is positive. If we were to move our 1 to the left instead of to the right, our \( \Delta \) would be negative.

We generalize the result from the previous lemma by moving the \( t^{th} \) 1 in \( v \) to the right such that it is still the \( t^{th} \) 1 in \( v' \) as well.

**Lemma 6.** Given parity sequences of length \( k \), \( v \) and \( v' \) each having \( n \) 1’s, if we obtain \( v' \) from \( v \) by moving the \( t^{th} \) 1 in \( v \), from the \( a^{th} \) position \( r \geq 0 \) places to the right, such that it is still the \( t^{th} \) 1 in \( v' \), then \( T_{v'}(x) - T_v(x) = \Delta \) where
\[ \Delta = \frac{3^{a-t}}{2^{k-a+1}}(2^r - 1). \]
Proof. By Lemma 5, if \( v' \) is obtained from \( v \) by moving the \( t \)th 1 one place to the right we have
\[
T_{v'}(x) - T_v(x) = \frac{3^{n-t}}{2^{k-a+1}} = \frac{3^{n-t}}{2^{k+1}} 2^a.
\]
where \( n = v_k + ... + v_1 \) and \( a \) is the position of the \( t \)th one in \( v \). Thus if \( v' \) is obtained from \( v \) by moving the \( t \)th 1 \( r \) places to the right we have
\[
T_{v'}(x) - T_v(x) = \frac{3^{n-t}}{2^{k+1}} (2^a + 2^{a+1} + ... + 2^{a+r-1})
\]
\[
= \frac{3^{n-t}}{2^{k+1}} 2^a (1 + 2 + 2^2 + ... + 2^{r-1})
\]
\[
= \frac{3^{n-t}}{2^{k+1}} 2^a (2^r - 1)
\]
\[
= \frac{3^{n-t}}{2^{k-a+1}} (2^r - 1)
\]
\]
\[
\square
\]
Let \( s(t) \) be the starting position of our \( t \)th 1 in \( v \) and let \( e(t) \) be the ending position of the \( t \)th 1 in \( v' \). Then we define \( \delta_t = \text{sgn}(s(t) - e(t)) \). If we assume that the \( t \)th 1 in \( v \) is not the only 1 that is moved between \( v \) and \( v' \), we have the following lemma, which follows directly from above.

**Lemma 7.** Given parity sequences of length \( k \), \( v \) and \( v' \), each having \( n \) ones, if we interchange 0’s and 1’s in \( v \) to form \( v' \), then \( T_{v'}(x) - T_v(x) = \Delta \), where
\[
\Delta = \sum_{t=1}^{n} \delta_t \frac{3^{n-t}}{2^{k-a_t+1}} (2^{r_t} - 1)
\]
and where \( \delta_t = \text{sgn}(s(t) - e(t)) \), \( r_t = |s(t) - e(t)| \) and \( a_t = s(t) \).
We can use Lemma 7, Equation 3.1, Theorem 6 and our definitions of blocks and strings to come up with a new formula for checking to see if a set of parity sequences is in fact a block or string prefix.

**Theorem 7.** If \( v \) and \( v' \) are parity sequences with \( n \) ones and \( k \) entries, they are block or string prefixes if and only if
\[
2^k - 3^n = \sum_{i=1}^{n} \delta_i \frac{3^{n-i}}{2^{k-a+1}} (2^r - 1).
\]

**Proof.** First, assume \( v = \langle v_1, v_2, ..., v_{k-1}, v_k \rangle \) and \( v' = \langle v'_1, v'_2, ..., v'_{k-1}, v'_k \rangle \) are block prefixes. Then, by definition, we know that \( T_v(x) + 1 = T_{v'}(x + 1) \). From Equation 3.1 and Theorem 6 we know that when \( v' \) is applied to \( x + 1 \) the result is
\[
T_{v'}(x + 1) = \frac{3^n}{2^k} (x + 1) + \sum_{i=0}^{k-1} \frac{3^{i+1} v'_{i+1}}{2^k} = T_{v'}(x) + \frac{3^n}{2^k}.
\]
Therefore, the following should be true: \( T_v(x) + 1 = T_{v'}(x) + \frac{3^n}{2^k} \). Using the substitution of \( T_{v'}(x) = T_v(x) + \Delta \) from Lemma 7,
\[
T_v(x) + 1 = T_v(x) + \Delta + \frac{3^n}{2^k}
\]
\[
\Delta = 1 - \frac{3^n}{2^k}
\]
\[
2^k \Delta = 2^k - 3^n.
\]
From Lemma 7, \( 2^k \Delta = \sum_{i=1}^{n} \delta_i \frac{3^{n-i}}{2^{k-a+1}} (2^r - 1) \). Therefore,
\[
2^k - 3^n = \sum_{i=1}^{n} \delta_i \frac{3^{n-i}}{2^{k-a+1}} (2^r - 1).
\]
A similar proof can be written for the case where \( v \) and \( v' \) are string prefixes. \( \square \)
4.2.4 Proportion of Ones

Next, we set out to determine the ratio between the number of ones in a block or string and the length of the block or string. We started by investigating our extreme values, that is we looked at the case of having all zeros and the case of having all ones in our blocks and strings. It is obvious that neither of these values work, since corresponding blocks and strings must have an equal number of ones by Theorem 6. We also know that since the numbers that the sequences are representing are consecutive, they must have different starting entries. This means that each must have at least a single zero and a single one.

Now let’s look at the case of just a single zero in the parity sequence.

**Theorem 8.** Given a block or string prefix of length $k > 2$ with $n$ ones, $n \neq k - 1$.

**Proof.** Suppose we have a block prefix $b$, of length $k$ where the number of ones in $b$ is $n = k - 1$. Since $b$ and $b'$ are the parity vectors of consecutive integers, either $b_1 = 0$ or $b'_1 = 0$.

Case 1: $b_1 = 0$. Assume that the 0 is at position $m$ in $b'$. So for $1 \leq t < m$, each 1 moves one position to the left and so $r_t = 1$, $\delta_t = -1$ and $a_t = t + 1$. For $m \leq t \leq k - 1$, $\delta_t = 0$.

$$\Delta = \sum_{t=1}^{m-1} (-1) \frac{3^{n-t}}{2^{k-t}} (2^t \cdot 1) + 0$$

$$= 3^{n+1-m} 2^{m-k} - 3^n 2^{1-k}$$

$$= 3^{k-m} 2^{m-k} - 3^{k-1} 2^{1-k}$$

36
Since, $2^k \Delta = 2^k - 3^n$,

$$3^{k-m}2^m - 3^{k-1}2^1 = 2^k - 3^{k-1}$$

$$3^{k-m}2^m - 3^{k-1} = 2^k$$

Note that the right hand side is even which implies that $m = 0$, which contradicts our initial assumptions.

Case 2: $b'_1 = 0$. Assume that the 0 is at position $m$ in $b$. So for $1 \leq t < m$, each 1 moves one position to the right and so $r_t = 1$, $\delta_t = 1$ and $a_t = t$. For $m \leq t \leq k - 1$, $\delta_t = 0$.

$$\Delta = \sum_{t=1}^{m-1} (1) \frac{3^{n-t}}{2^{k-t+1}} (2^1 - 1) + 0$$

$$= -3^{n+1-m}2^{m-k-1} + 3^n2^{-k}$$

$$= -3^{k-m}2^{m-k-1} + 3^{k-1}2^{-k}$$

Since, $2^k \Delta = 2^k - 3^n$,

$$-3^{k-m}2^{m-1} + 3^{k-1} = 2^k - 3^{k-1}$$

$$-3^{k-m}2^{m-2} + 3^{k-1} = 2^{k-1}$$

The right hand side is even so $m = 2$.

$$-3^{k-2}2^{2-2} + 3^{k-1} = 2^{k-1}$$

$$(2)3^{k-2} = 2^{k-1}$$
\[ 3^{k-2} = 2^{k-2} \]

\[ k = 2 \]

Therefore, the only time that our block or string prefixes may have a single zero is the case where our parity sequences are of the form \((10)\) and \((01)\).

\[ \square \]

Next, we will look at the case with just a single one in the sequence.

**Theorem 9.** Given a block or string prefix of length \(k > 2\), with \(n\) ones, \(n \neq 1\).

**Proof.** Suppose we have a block prefix \(b\), of length \(k\) where the number of ones in \(b\) is \(n = 1\). Then,

\[ \Delta = \delta_1 \frac{3^{1-1}}{2^{k-a_1+1}} (2^{r_1} - 1) = \delta_1 \frac{2^{r_1} - 1}{2^{k-a_1+1}} \]

Then, \(2^k \Delta\) should equal \(2^k - 3^n\).

\[ \delta_1 \frac{2^{r_1} - 1}{2^{k-a_1+1}} = 2^k - 3 \]

\[ \delta_1 \left( 2^{r_1+a_1-1} - 2^{a_1-1} \right) = 2^k - 3 \]

We know that one of our two parity sequences must start with a one, so we have two cases.

Case 1: Suppose that the first in the pair starts with 1. That makes \(a_1 = 1\) and \(\delta_1 = 1\), which gives us

\[ 2^{r_1} - 1 = 2^k - 3 \]
\[ 2 = 2^k - 2^{r_1} \]
\[ 1 = 2^{k-1} - 2^{r_1-1} \]

Which implies that \( k = 2 \) and \( r_1 = 1 \) which is the case where our two parity
sequences are \( 10 \) and \( 01 \).

Case 2: Suppose that the second in the pair starts with 1. That makes \( r_1 = a_1 - 1 \)
and \( \delta_1 = -1 \), which gives us
\[ (-1) \left( 2^{a_1-2} - 2^{a_1-1} \right) = 2^k - 3 \]

The left side of this equation will always be even, whereas the right side will
always be odd. This is a contradiction.

Therefore, a block or string prefix will only have a single one if it is of length
two and of the form specified above.

\[ \square \]

We can continue our argument above by writing similar proofs that will show
us all of the blocks that contain two ones, all of the blocks that contain three ones,
and so on. As we increase the number of ones that the sequence has, the number
of cases that the proof has grows rapidly, to the point that it is unreasonable to
continue using this method much beyond the case of a block or string prefix with
three ones.

From our block and string data in Tables 4.1 and 4.2, we have found values for
the minimum, maximum and average value of our \( \frac{n}{k} \) ratio. We believe that these
values will approximate the actual values for all blocks and strings. Therefore, we
have the following conjecture.
Conjecture 2. For all blocks and strings of length $k$ with $n$ ones,

$$\frac{1}{3} \leq \frac{n}{k} \leq \frac{7}{11}$$

and, on average

$$\frac{n}{k} \approx \frac{5}{11}.$$

We have just seen the importance of blocks and strings in the parity sequences of consecutive integers of the same total stopping time. The next chapter will expand on these ideas as we explore long runs of consecutive integers with the same total stopping time.
Chapter 5

Long Runs of Consecutive Integers

Recall from Chapter 1 that we define a run to be a set of consecutive integers that all have the same total stopping time. Garner [4] pointed out that runs are simply many sets of overlapping pairs of consecutive integers of the same total stopping time. This suggests a way to build long runs. First, start with a run, then try to make the last integer of the run the first part of a pair. If this can be done, then we have found a new run with at least one more consecutive integer than our starting run. Using this idea, we feel that given any starting parity sequence of length \( k \), we can construct an arbitrarily long run where that parity sequence is the first \( k \) terms of the parity vector of the first integer in the run. The following conjecture states this idea more formally.

**Conjecture 3.** Consider any parity vector of length \( k \), which corresponds to the set of integers of the form \( 2^k m + b_k \), where \( b_k \) is found using Terras’[6] algorithm, \( m \) is some positive integer, and \( \sigma_\infty(2^k m + b_k) = \sigma_\infty(2^k m + b_k + 1) = \ldots = \sigma_\infty(2^k m + b_k + (g-1)) \) (a run of \( g \) consecutive integers of the same total stopping time.) There exists some positive integer \( m' \) such that \( \sigma_\infty(2^k m' + b_k) = \sigma_\infty(2^k m' + b_k + g) \), a run
of \( g + 1 \) integers.

Using this conjecture, we can construct the flow chart found in Figure 5.1 above. The flow chart shows how to search a parity sequence \( v \) to determine whether or not it is the first part of a pair of consecutive integers of the same total stopping time.

In our search, we first check the initial entry of our given parity sequence, \( \langle v \rangle \). If it is a one, we begin searching for blocks that fit the parity sequence. If the initial entry is a zero, we begin searching for strings or stems that fit it. After finding one of our defined special sequences (blocks, strings or stems) that fits the start of the sequence, we check the next entry and continue to move along through the flow chart, searching for the designated special sequences. From our definition of stems, we know to stop once a stem is found because we will now have the first part of a pair. If we run out of entries before reaching a stem, entries
can be appended to the sequence in order to make the sequence match the desired special sequence. In these cases, make a stem whenever possible. In order to actually make a new pair, we must then form the parity sequence of the next consecutive integer by using the corresponding sequences in the specified order to the sequence we have just searched. We can then use the work of Terras [6] to find actual integer values to which this sequence corresponds.

The basic ideas from Conjecture 3 and Figure 5.1 led us to write a computer program that allowed us to find long runs of consecutive integers of the same total stopping time. In this program, we start with an integer and find how many consecutive integers there are in a row that have the same total stopping time. We do this by calculating and comparing the trajectories of each. While doing the comparison, we keep track of how many steps it takes for the trajectories to coincide. Next, because of our algorithm from Terras [6], we know that if we add an integer multiple of 2 raised to the power of how many steps it takes for the trajectory to coincide to our given integer, we will get another run that will be at least as long as the first. Note that this is assuming that every coinciding pair has to contain a stem. In practice, we have always found this to be true. We can take our new integer and calculate how many consecutive integers of the same total stopping time are less than and greater than our integer, to find the length of our run. We believe that we can continue to do these operations indefinitely. Our only real constraint at the moment is the time it takes the computer to run the calculations. So far, the longest run that we have found is a 39,116-tuple where the first integer in that run is \(47,223,664,828,697,525,864,383 + 5 \times 2^{274}\).
An obvious restriction on finding these long runs are the powers of two. Note that any power of two has parity vector \((0, 0, ..., 0, 0)\). A parity vector of this type is the fastest way to get to 1, as can be seen on our tree from Chapter 2. Any non-power of two cannot possibly reach 1 that quickly, therefore \(2^n\) cannot be part of a pair. Since the distance between \(2^n\) and \(2^{n+1}\) is \(2^n\), a run of length \(2^n\) must contain only integers larger than \(2^{n+1}\).
Chapter 6

Future Work

Throughout the last 18 months, I have been asked many times, “Did you prove it yet?” Now that I am at the end of this research my answer is, unfortunately, still a sad “no.” I do feel, though, that my work has left the doors wide open for much more related work on this problem and even helped in the ultimate quest to prove the conjecture. I have discovered many interesting results that allowed me to prove new theorems, formulate new conjectures and strengthened my faith that someday we will be able to prove that the Collatz Conjecture is true. Therefore, I feel as though this has been a very worthwhile endeavor for me to pursue.

One place where there is room for much more investigation is our stems, blocks and strings. Are Garner’s stems really the only form of corresponding stems that exist? What is it about that string of 1’s in those stems that makes them work and bring those trajectories together? Are there infinitely many blocks and strings? Are our corresponding pairs that we have presented unique? What are the bounds for the proportion of ones found in these blocks and strings? These are essential questions that could provide a great opportunity for further research.
Our work with the long runs of consecutive integers in Chapter 5 also has room for research expansion. Do arbitrarily long runs actually exist? Can we make them from any starting parity sequence? Are the majority of the positive integers in a run? Is there any way for us to predict where a run will occur? These ideas, again, also provide areas for exploration.

As you can see, there are still so many questions left to answer in regards to this problem. We have only just touched the surface of the work to be done and the results that will come from our attempts to solve this intriguing conjecture.
Appendix A

Branch Count Procedure

branchcount:=proc(b, k, c)
# counts the number of integers at height k above an integer b
# assuming that b is at height c
local count,w,s,v,r;

if b=1 or b=2 then # this is taking care of the base case.
    if k>2 then return(branchcount(4, k, 2));
    else return(1);
    fi;
elif b=2ˆc then count:=1; # count the integer 2ˆk
else count:=0; # don't double-count anything else
fi;
ellif b=2ˆc then count:=1; # count the integer 2ˆk
else count:=0;
    fi;

if b mod 3 = 0 then count:=count+0;

elif b mod 3 = 1 then
w:=floor((k-c)/2);  # this is how many branchings will occur above b
count:=count+w;
if w>0 then
for s from 1 to w do
    count:=count+branchcount((b*(2**(2*s))-1)/3,k,c+2*s);
od;
fi;
elif b mod 3 = 2 then
v:=ceil((k-c)/2);
count:=count+v;  # this is how many branchings will occur above b
if v>0 then
    for r from 1 to v do
        count:=count+branchcount((b*(2**(2*r-1))-1)/3,k,c+2*r-1);
od;
fi;
fi;
return(count);
end:
Bibliography


