The Collatz Conjecture: A Conjugacy Approach

James T. Long III

Advisor: Dr. Michael J. Fraboni

Liaison: Dr. Benjamin J. Coleman
ABSTRACT. The Collatz Conjecture has perplexed mathematicians since its initial proposal over 70 years ago. Many different approaches have been suggested and formulated, but so far, all have fallen short of proving or disproving the conjecture. One of the more promising strategies involves topological conjugacy, which was initially applied to the conjecture by Bernstein and others as early as the 1990s. In this paper, we continue the work pioneered by Bernstein to make statements concerning the Non-trivial Cycles Conjecture. We also derive a necessary condition for an endomorphism of the shift map to induce a conjugacy, as well as classify the dynamics of a particularly interesting conjugacy that was initially discovered by Monks.
This thesis would not have been possible without the personal support of certain individuals who have embodied the past, present, and future of my life.

First and foremost, I give my respects to my mother. Though she was unable to see me through the writing of this thesis, my faith that she comforts me when I struggle and rejoices with me as I succeed will never wane.

Innumerable thanks go to my father and sister, who have been the best family I could have possibly asked for through thick and thin. Thanks also go to my stepmother Dawn, who has become an irreplaceable addition to my family.

Last and definitely not least, I express my gratitude over having Lindsey in my life. The ardor with which I pursued Honors is but a mere fraction of the love I pledge to our coming marriage.

I dedicate this to all of you.
Contents

Chapter 1. Background 3
  1. The 2-adic Integers 3
  2. Concepts from Dynamical Systems 6
  3. The Collatz Conjecture 7
  4. Conjugacy 9
  5. Summary of Results 14

Chapter 2. Cycles in $\mathcal{F}$ 15
  1. Restating the Collatz Conjecture 15
  2. Bernstein’s Inverse of $Q$ 18
  3. Generalizing the Inverse of $Q$ to $Q_{n,b,c,d}$ 19
  4. Analyzing $T$-cycles Using $\Phi$ 26

Chapter 3. Endomorphisms of the Shift Map 32
  1. Continuous Endomorphisms of $\sigma$ 32
  2. Solenoidal Parity Vector Functions 34
  3. In Pursuit of New Conjugacies of $\sigma$ 35

Chapter 4. The Dynamics of $D$ Revisited 40
  1. The Symmetry between $D$ and Its Parity Vector Function $\mathcal{P}$ 40
  2. Fixed Points of $D$ 43
  3. Eventually Periodic Points of $D$ with Period $2^n$ 47
4. Periodic Points of D with Period $2^n - 1$

Chapter 5. Conclusions and Future Research

Acknowledgments

Bibliography
CHAPTER 1

Background

In this chapter, we discuss the background material that the reader should be familiar with before proceeding. In particular, we define the 2-adic integers and summarize some of their properties, discuss some elementary concepts from dynamical systems, state the Collatz Conjecture, and then finally discuss the conjugacy-based methods that have been applied to the problem.

1. The 2-adic Integers

Throughout our study of the Collatz Conjecture, we will be working in the space of the 2-adic integers, so we first define them and some elementary definitions and concepts concerning them. We direct the reader desirous of a more thorough introduction to the 2-adic integers to [Gou00].

Definition. The 2-adic integers compose the set \( \mathbb{Z}_2 \) of infinite sequences \( a_0, a_1, \ldots \) such that \( a_i \in \{0, 1\} \) for every \( i \in \mathbb{N} \).

Remark. In this paper, we will simplify the presentation of 2-adic integers by omitting the commas between each element of \( \{0, 1\} \). For example, \( a = a_0, a_1, \ldots \) would be represented as \( a_0a_1 \cdots \). Also, we denote any repeating parts of 2-adic expressions with an overhead bar: \( a = 1101010 \cdots \) would be expressed as \( a = 1\overline{10} \), for instance. Finally, we will let \( \mathbb{N} \) be the set of nonnegative integers, \( \{0, 1, \ldots\} \).
A 2-adic integer $a_0a_1a_2\cdots$ can be thought of as a number in base 2 with an infinite binary expansion $\sum_{i=0}^{\infty} a_i2^i$. As a result, addition and multiplication on the 2-adics follows naturally from the operations of binary addition and multiplication. In addition, for every $x \in \mathbb{Z}_2$, the additive inverse $-x$ can be computed by taking the two’s complement of $x$. Thus, subtraction is also simple to perform. It is also well-known that $\mathbb{Z}_2$ is a ring under addition and multiplication, but is not a field. To illustrate, $\mathbb{Z}_2$ contains all members of $\mathbb{Q}$ with odd denominators (expressed as $\mathbb{Q}_{\text{odd}}$), but $2$ has no multiplicative inverse in $\mathbb{Z}_2$, and so some, but not all, of the nonzero members of $\mathbb{Z}_2$ have multiplicative inverses.

It is noteworthy that since $\mathbb{Z} \subset \mathbb{Q}_{\text{odd}}$ and $\mathbb{Q}_{\text{odd}} \subset \mathbb{Z}_2$, $\mathbb{Z} \subset \mathbb{Z}_2$. In particular, for any $a \in \mathbb{Z}$ with base 2 expansion $a_0a_1\cdots a_n$ for some $n \in \mathbb{N}$, $a$’s 2-adic expansion is $a_na_{n-1}\cdots a_00$. It is well-known that $\mathbb{Z}$ is a subring of $\mathbb{Z}_2$, and so $\mathbb{Z}_2$ is an extension of the regular integers.

**Example 1.1.** Since the base 2 representation of 3 is $11_2$, its 2-adic representation is $11\overline{0}$. Similarly, $13 = 1101_2$ has a 2-adic representation of $1011\overline{0}$.

**Example 1.2.** The 2-adic representation of $-3$ can be found by taking the two’s complement of $3 = 11\overline{0}$, which is $10\overline{1}$. This can be confirmed by showing that $10\overline{1}$ is the additive inverse of $3$: $10\overline{1} + 3 = 10\overline{1} + 11\overline{0} = \overline{0}$.

It should be noted that in $\mathbb{Z}_2$, we use the 2-adic metric instead of the standard Euclidean metric defined on the real numbers. This metric is expressed in terms of the 2-adic norm: for every $x \in \mathbb{Z}_2$, $|x|_2 = 2^{-n}$, where $n$ is the smallest $n \in \mathbb{N}$ such that $x_n = 1$. Then for every $x, y \in \mathbb{Z}_2$, the 2-adic metric $d_2(x, y) = |x - y|_2$. The following example illustrates that some
series which would not normally converge under the Euclidean metric will converge nicely under the 2-adic metric.

**Example 1.3.** The 2-adic representation of \(-\frac{1}{3}\) is \(\overline{10}\). This can be verified in two different ways. First, straightforward computation confirms that \(\overline{10}\) is the multiplicative inverse of \(-3\): \(\overline{10} \times \overline{3} = \overline{10} \times 10\overline{1} = 1\overline{0} = 1\). Alternatively, observe that \(\overline{10}\) can be expressed as a geometric series with common ratio 4: \(1 + 4 + 16 + \cdots\). Note that this series diverges under the Euclidean metric, as \(|4| = 4 > 1\). However, since the common ratio has 2-adic norm \(|4|_2 = |001\overline{0}|_2 = \frac{1}{2^2} = \frac{1}{4} < 1\), the series converges to the value \(\frac{1}{1-4} = \frac{1}{3}\) under the 2-adic metric. See \([Gou00]\) for a rigorous explanation of why the reasoning behind this example is valid.

Our study will make extensive usage of the notion of parity in \(\mathbb{Z}_2\). The following definition is a straightforward extension of the parity of binary numbers.

**Definition.** Let \(x \in \mathbb{Z}_2\) with \(x = x_0x_1 \cdots\). Such an \(x\) is called **odd** if \(x_0 = 1\). Otherwise, \(x\) is said to be **even**.

**Remark.** Although 2 does not have a multiplicative inverse in \(\mathbb{Z}_2\), it is well-known that for every \(x \in \mathbb{Z}_2\), \(x\) is divisible by 2 if and only if \(x\) is even. If \(x = 0x_1x_2 \cdots\), then \(\frac{x}{2} = x_1x_2 \cdots\).

Fortunately, it is not always necessary to derive the 2-adic expansion of an \(x \in \mathbb{Z}_2\) in order to determine its parity. For instance, if \(x \in \mathbb{Z}\), its parity in \(\mathbb{Z}_2\) is identical to its parity in \(\mathbb{Z}\). In general, if \(x \in \mathbb{Q}_{\text{odd}}\), its parity in \(\mathbb{Z}_2\) is identical to the parity of its numerator in \(\mathbb{Z}\).
2. Concepts from Dynamical Systems

Now that we have defined some basic ideas from the number system we will use in our work, we turn our attention to a few fundamental concepts from dynamical systems. For the curious, Devaney provides a more detailed exposition on the subject than needed here in [Dev92].

We first recall that for any \( f : S \to S \), \( x \in S \), and \( k \in \mathbb{N} \), \( f^k(x) \) denotes the \( k \)-fold composition of \( f \) with itself. In other words, \( f^0(x) = x \) and \( f^k(x) = f \circ f \circ \ldots \circ f(x) \) for \( k \geq 1 \). The reader should be warned that \( f^k(x) \) should not be mistaken for \( f(x)^k \), which is the typical notation used for functional exponentiation.

**Definition.** Let \( f : S \to S \). Then for any \( x \in S \), the **orbit of** \( x \) **under** \( f \), or the \( f \)-**orbit of** \( x \), is the infinite sequence \( x, f(x), f^2(x), \ldots \).

**Remark.** There are several alternative terms used in place of “orbit”; in [Lag85], for instance, Lagarias uses the word “trajectory” instead. For the sake of clarity, we will restrict our language to “orbit.”

**Example 1.4.** The orbit of \(-1\) under \( f(x) = x + 1 \) is \(-1, 0, 1, 2, \ldots \). Similarly, the orbit of \(-1\) under \( f(x) = x^2 \) is \(-1, 1, 1, 1, \ldots \).

The study of these orbits is crucial to our interests, and so we proceed to define several terms for expressing different types of orbits.

**Definition.** Let \( f : S \to S \). Then for any \( x \in S \), if there exists some \( m, n \in \mathbb{N} \) such that \( m < n \) and \( f^m(x) = f^n(x) \), then \( x \) is said to be **eventually periodic under** \( f \). In particular, when \( x^m = f^{m+1}(x) \), \( x \) is said to be an **eventually fixed point of** \( f \). In the special case where \( m = 0 \), these two types
of points are more precisely labeled as *periodic points* and *fixed points*, respectively. The smallest possible value of \( n - m \) is the *minimum period* of \( x \). Note that the minimum period of an eventually fixed point is always 1. When there is no risk of ambiguity, we sometimes refer to the minimum period of \( x \) simply as the *period* of \( x \).

**Example 1.5.** It is straightforward to verify that \( f(x) = x + 1 \) has no eventually periodic points, as \( f \) is strictly increasing.

**Example 1.6.** \( f(x) = x^2 \) has 0 and 1 as fixed points. In addition, the orbit of \(-1\) under \( f \) is eventually fixed after 1 iteration.

### 3. The Collatz Conjecture

Now that we have defined the important concepts from both the 2-adic integers and dynamical systems, we can at last turn our attention to our basic object of study.

**Definition.** The *Collatz function* is defined by \( T : \mathbb{Z}_2 \to \mathbb{Z}_2 \) such that for every \( x \in \mathbb{Z}_2 \),

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

**Remark.** Some articles, such as [Ber94], define the odd piece of \( T \) to be \( 3x + 1 \). In the context of the Collatz Conjecture, these two different definitions are essentially the same.

We now state the famous Collatz Conjecture, which concerns the behavior of \( T \)-orbits for positive integers. A classic history of the problem can be found in [Lag85].
CONJECTURE 1.7 (Collatz Conjecture). For every $x \in \mathbb{Z}^+$, the orbit of $x$ under $T$ contains $1$.

EXAMPLE 1.8. The orbit of $3$ under $T$ is $3, 5, 8, 4, 2, 1, 2, 1, \ldots$, which contains $1$.

Example 1.8 shows one of the most noteworthy dynamics of $T$: $1$ is a periodic point that cycles between itself and $2$. As a result, any $x \in \mathbb{Z}^+$ should eventually enter this cycle after iteration of $T$. Many attempts towards proving the conjecture have utilized this behavior. Chapter 2 will discuss some of these approaches in more detail.

The following example illustrates that the Collatz Conjecture does not hold for all of $\mathbb{Z}_2$.

EXAMPLE 1.9. The $T$-orbits of $-1$ and $0$ do not contain $1$ since they are both fixed points of $T$. Similarly, the orbit of $-\frac{1}{3}$ under $T$, $-\frac{1}{3}, 0, 0, \ldots$, does not contain $1$.

Naturally, it is interesting to study the number of iterations of $T$ that are necessary to obtain $1$ for a specific starting value of $x \in \mathbb{Z}_2$. We formally define a measure for such a value.

DEFINITION. Given $x \in \mathbb{Z}^+$, the total stopping time of $x$ under $T$ is the smallest value of $k \in \mathbb{Z}^+$ such that $T^k(x) = 1$. If no such value exists, the total stopping time is defined to be $\infty$.

EXAMPLE 1.10. The orbit of $26$ under $T$ is $26, 13, 20, 10, 5, 8, 4, 2, 1, \ldots$, and so the total stopping time of $26$ under $T$ is $8$. Interestingly enough, the total stopping time of $27$ under $T$ is $70$. 
We next provide formal descriptions of the three possible kinds of \( T \)-orbits of the positive integers.

**Definition.** Let \( x \in \mathbb{Z}^+ \). The orbit of \( x \) under \( T \) can be classified in one of three different ways. If \( x \) has a finite total stopping time under \( T \), its orbit under \( T \) is said to be **convergent**. If, however, \( x \)'s total stopping time under \( T \) is \( \infty \) and \( x \) is an eventually periodic point of \( T \), the orbit of \( x \) under \( T \) is said to be **non-trivially cyclic**. In all other cases, the \( T \)-orbit of \( x \) is said to be **divergent**.

**Remark.** The Collatz Conjecture is equivalent to the statement that the orbits of all positive integers under \( T \) are convergent.

## 4. Conjugacy

Now that we have defined some of the basic terminology that has been used in studying the Collatz Conjecture, we can focus on defining the notion of conjugacy, the core concept of our research. In order to do so, we recall a definition from elementary topology.

**Definition.** Let \( A \) and \( B \) be topological spaces. Two mappings \( f : A \to B \) and \( g : B \to B \) are said to be **topologically conjugate** if there exists a homeomorphism \( h : B \to A \) such that \( f \circ h = h \circ g \).

**Example 1.11.** As Fraboni mentions in [Fra97], linear mappings on \( \mathbb{Z}_2 \) of the form \( ax+b \) for some \( a, b \in \mathbb{Z}_2 \) are simple examples of homeomorphisms.

We are now sufficiently versed to define conjugacy.

**Definition.** Let \( A \) and \( B \) be topological spaces. Two mappings \( f : A \to A \) and \( g : B \to B \) are said to be **topologically conjugate** if there exists a homeomorphism \( h : B \to A \) such that \( f \circ h = h \circ g \).
Remark. In the more general case where \( h \) is not necessarily continuous nor bijective, we say that \( h \) is a morphism from \( f \) to \( g \). Morphisms will be an important object of study in Chapter 3.

If two functions are topologically conjugate, they possess the same dynamics. To illustrate, suppose we have two mappings \( f : \mathbb{Z}_2 \rightarrow \mathbb{Z}_2 \) and \( g : \mathbb{Z}_2 \rightarrow \mathbb{Z}_2 \) that are topologically conjugate via a homeomorphism \( h : \mathbb{Z}_2 \rightarrow \mathbb{Z}_2 \). If \( x \in \mathbb{Z}_2 \) is a fixed point of \( f \), then \( h(x) \) is a fixed point of \( g \). A similar statement holds in the general case of eventually periodic points. Most importantly, it is well-known that if \( f \) is chaotic, then \( g \) is chaotic as well. (See [Dev92] for an in-depth discourse on chaotic maps) These facts result from the fact that iteration of a function is preserved under conjugacy: a straightforward proof by induction shows that for every \( n \in \mathbb{N} \), \( f^n = h \circ g^n \circ h^{-1} \).

See Monks’ [Mon08] for one such proof.

We now define a chaotic function that is conjugate to \( T \), and in doing so, show that \( T \) exhibits chaotic behavior as well.

Definition. The shift map is defined by the mapping \( \sigma : \mathbb{Z}_2 \rightarrow \mathbb{Z}_2 \) such that for every \( x \in \mathbb{Z}_2 \),

\[
\sigma(x) = \begin{cases} 
\frac{x}{2} & \text{if } x \text{ is even} \\
\frac{x - 1}{2} & \text{if } x \text{ is odd.}
\end{cases}
\]

Note that for any \( a \in \mathbb{Z}_2 \) such that \( a = a_0a_1a_2\cdots \), the shift map has the effect of mapping \( a \) to \( a_1a_2a_3\cdots \).

Example 1.12. To illustrate the “shift” effect of \( \sigma \), note that \( \sigma(11\bar{0}) = \sigma(3) = 1 = \bar{1}\bar{0} \). Similarly, \( \sigma(\bar{0}\bar{1}) = \sigma\left(-\frac{2}{3}\right) = -\frac{1}{3} = \bar{1}\bar{0} \).
Devaney proved in [Dev92] that the shift map is chaotic. To show that $T$ is chaotic as well, we provide a conjugacy between $\sigma$ and $T$.

**Definition.** The parity vector function associated with $T$ is defined by the map $Q : \mathbb{Z}_2 \to \mathbb{Z}_2$ such that for any $x, y \in \mathbb{Z}_2$ with $y = Q(x), y_i \equiv T^i(x) \mod 2$ for every $i \in \mathbb{N}$.

**Example 1.13.** Since the orbit of 3 under $T$ is 3, 5, 8, 4, 2, 1, 2, ..., $Q(3)$, 3’s parity vector under $T$, is 1100010.

The parity vector function has been applied to numerous aspects of the Collatz problem. In [Lag85], Lagarias showcases a few of these applications. We now state without proof the result that is most relevant to our study. The curious reader can find the proof in many sources, including Bernstein’s [Ber94].

**Theorem 1.14.** $T$ is conjugate to $\sigma$ by $Q$.

Since $T$ and $\sigma$ are conjugate, it follows that $T$ is chaotic. In theory, it would thus be possible to prove the Collatz Conjecture by studying the properties of the shift map and relating them back to $T$, but unfortunately, this is not a simple matter. In order for any properties of the shift map to be used to answer the conjecture, it would be necessary to compute the image of $\mathbb{Z}^+$ under the conjugacy $Q$, which is difficult.

In [Fra97], Fraboni constructed a family of functions that contained some members that were topologically conjugate to $T$ in simpler ways than $Q$. We define this family and present some important definitions before discussing these conjugacies.
DEFINITION. Let \( a, b, c, d \in \mathbb{Z}_2 \) with \( a, c, d \) odd and \( b \) even. The family of modular functions \( \mathcal{F} \) is the set of all functions \( f_{a,b,c,d} : \mathbb{Z}_2 \to \mathbb{Z}_2 \) such that

\[
f_{a,b,c,d}(x) = \begin{cases} 
\frac{ax + b}{2} & \text{if } x \text{ is even} \\
\frac{cx + d}{2} & \text{if } x \text{ is odd.}
\end{cases}
\]

EXAMPLE 1.15. Both \( T \) and \( \sigma \) are members of \( \mathcal{F} \), namely \( f_{1,0,3,1} \) and \( f_{1,0,1,-1} \) respectively.

The parity vector function has a straightforward generalization to \( \mathcal{F} \), which we define here.

DEFINITION. Let \( a, b, c, d \in \mathbb{Z}_2 \) such that \( a, c, d \) odd and \( b \) even. The parity vector function associated with \( f_{a,b,c,d} \) is defined by the map \( Q_{a,b,c,d} : \mathbb{Z}_2 \to \mathbb{Z}_2 \) such that for any \( x, y \in \mathbb{Z}_2 \) with \( y = Q_{a,b,c,d}(x) \), \( y_i \equiv f_{a,b,c,d}^i(x) \mod 2 \) for every \( i \in \mathbb{N} \).

EXAMPLE 1.16. The parity vector function associated with \( \sigma = f_{1,0,1,-1} \) is trivial since for every \( x \in \mathbb{Z}_2 \) and \( k \in \mathbb{N} \), \( x_k \equiv \sigma^k(x) \mod 2 \). Therefore, \( Q_{1,0,1,-1}(x) = x \).

Using this generalized parity vector function, it is possible to construct a conjugacy between the members of \( \mathcal{F} \) and \( T \). In [Fra97], Fraboni showed that every \( f \in \mathcal{F} \) is topologically conjugate to \( \sigma \) by its corresponding parity vector function \( Q_{a,b,c,d} \), just as with \( T \). Furthermore, conjugacy is transitive since it is an equivalence relation, and so it follows that \( f \) is conjugate to \( T \) as well. The proof is a straightforward formalization of the above argument, and so we omit it here.
Chapter 1: Background

Theorem 1.17 (Fraboni [Fra97]). Every $f \in \mathcal{F}$ is topologically conjugate to $T$.

In addition to stating and proving Theorem 1.17, Fraboni showed that certain members of $\mathcal{F}$ were related to $T$ by particularly nice conjugacies. The proof of the following result can be found in Fraboni’s [Fra97].

Theorem 1.18 (Fraboni [Fra97]). The set of all maps conjugate to $T$ by a linear homeomorphism $px + q$ consists of precisely those $f \in \mathcal{F}$ of the form $f_{1,q,3,p-q}$ with $p$ odd and $q$ even or $f_{3,p-q,1,q}$ with $p$ and $q$ odd.

Example 1.19. By Theorem 1.18, $f_{3,0,1,1}$ is topologically conjugate to $T$ by the homeomorphism $h(x) = x + 1$. Note that the orbit of 3 under $T$ is $3, 5, 8, 4, 2, 1, 2, \ldots$, while the orbit of $h(3) = 4$ under $f_{3,0,1,1}$ is $4, 6, 9, 5, 3, 2, 3, \ldots$, or $3 + 1, 5 + 1, 8 + 1, 4 + 1, 2 + 1, 1 + 1, 2 + 1, \ldots$. Clearly, 3 under iteration of $T$ exhibits the same behavior as 4 under iteration of $f_{3,0,1,1}$.

As we shall elaborate upon in Chapter 2, $f_{3,0,1,1}$, which we demonstrated in Example 1.19, will be useful to our study.

Fraboni showed in [Fra97] that by Theorem 1.18, the shift map is not conjugate to $T$ by a linear homeomorphism. It should be noted that $\mathcal{F}$ is by no means an exhaustive collection of the mappings which are topologically conjugate to $T$. For instance, in [Fra97], Fraboni constructed a mapping $f \notin \mathcal{F}$ that is conjugate to $T$ by a piecewise linear map. Fortunately, we only use members of $\mathcal{F}$ that are linearly conjugate to $T$ in our study, and so this stipulation is not a significant hindrance.
5. Summary of Results

In Chapter 2, we generalize Bernstein’s non-iterative inverse of the parity vector function $Q$ to $Q_{a,b,c,d}$; we then use this result to make decisive statements about the Non-trivial Cycles Conjecture, which concerns the periodic points of $T$.

Chapter 3 continues Monks’ efforts in [Mon08] to derive new conjugacies of $T$ from the continuous endomorphisms of $\sigma$, and states a necessary condition for these endomorphisms to be conjugate to $T$.

Finally, in Chapter 4, we turn our attention to one especially noteworthy endomorphism of $\sigma$ that was first discovered and studied by Monks. We classify some of the dynamics of this endomorphism in the hopes that it will be useful towards proving Monks’ reformulation of the Collatz Conjecture, which she states in [Mon08].
CHAPTER 2

Cycles in $\mathcal{F}$

As part of the ongoing effort to prove the Collatz Conjecture, mathematicians have often restated the conjecture in ways that will hopefully be more approachable. In [Ber94], Bernstein derives a conjugacy between $T$ and $\sigma$ that he uses to weaken the conditions on one such reformulation of the conjecture. In this chapter, we will brief the reader on the restatement Bernstein addressed, state Bernstein’s conjugacy and construct a generalization of it for $\mathcal{F}$, and finally apply our results back to this restatement.

1. Restating the Collatz Conjecture

Before we state the reformulation of the Collatz Conjecture that we will be focusing on throughout this chapter, we formalize the concept of a cycle to facilitate our discussion.

**Definition.** Let $x \in \mathbb{Z}_2$ and $a, b, c, d \in \mathbb{Z}_2$ with $a, c, d$ odd and $b$ even. If there exists some $n \in \mathbb{Z}^+$ such that $n$ is the smallest value for which $x = f_{a,b,c,d}^n(x)$, we say that $\{x, f_{a,b,c,d}(x), f_{a,b,c,d}^2(x), \ldots, f_{a,b,c,d}^{n-1}(x)\}$ is a cycle in $f_{a,b,c,d}$, or alternatively, an $f_{a,b,c,d}$-cycle. Additionally, we say that $n$ is the length of the cycle.

**Remark.** A similar definition exists in general for all mappings $f : S \rightarrow S$ for some set $S$.  

15
Chapter 2: Cycles in $\mathcal{F}$

**Example 2.1.** Since $\{1, 2\}$ is a $T$-cycle, the Collatz Conjecture is equivalent to the statement that for every $x \in \mathbb{Z}^+$, the orbit of $x$ under $T$ contains this cycle.

**Example 2.2.** It is straightforward to show that every $x \in \mathbb{Z}_2$ of the form $x = \overline{x_0x_1\cdots x_{n-1}}$ for some $n \in \mathbb{Z}^+$ is a $\sigma$-cycle, as $\sigma^n(x) = \sigma^n(\overline{x_0x_1\cdots x_{n-1}}) = \sigma^{n-1}(\overline{x_1\cdots x_{n-1}x_0}) = \cdots = \overline{x_0x_1\cdots x_{n-1}} = x$. Note that $\sigma^2(\overline{1010}) = \overline{1010}$, and so in general, $n$ is not necessarily the length of the cycle.

With this terminology, we can rephrase the Collatz Conjecture in terms of two conjectures pertaining to the cycles of $T$: the Non-trivial Cycles Conjecture and the Divergent Orbits Conjecture, both of which have been previously stated in numerous sources, including [MY04].

**Conjecture 2.3 (Non-trivial Cycles Conjecture).** The only $T$-cycle consisting of positive integers is $\{1, 2\}$.

Note that the Non-trivial Cycles Conjecture does not impose any restrictions on whether or not an arbitrary positive integer eventually enters a $T$-cycle. However, this stipulation is necessary since if any $x \in \mathbb{Z}_2$ were to have a divergent orbit under $T$, $x$ would be a counterexample for the Collatz Conjecture. We account for this possibility with the Divergent Orbits Conjecture.

**Conjecture 2.4 (Divergent Orbits Conjecture).** For every $x \in \mathbb{Z}_2$, the orbit of $x$ under $T$ is not divergent. In other words, every $x \in \mathbb{Z}_2$ eventually enters a $T$-cycle.
EXAMPLE 2.5. Even the orbits of some negative integers under $T$ seem to exhibit cyclical behavior. The three cycles in the negative integers that are currently known are \{-1\}, \{-5, -7, -10\}, and \{-17, -25, -37, -55, -82, -41, -61, -91, -136, -68, -34\}.

It is straightforward to see that the Non-trivial Cycles Conjecture and the Divergent Orbits Conjecture together imply the Collatz Conjecture: for any $x \in \mathbb{Z}^+$, the Divergent Orbits Conjecture ensures that $x$ will eventually enter a $T$-cycle, while the Non-trivial Cycles Conjecture forces this cycle to be \{1, 2\}. Therefore, the orbit of $x$ under $T$ contains 1.

Several authors have mentioned an alternative to the Divergent Orbits Conjecture. In [BL96], Bernstein and Lagarias showed that the following conjecture implies the Divergent Orbits Conjecture.

**Conjecture 2.6 (Periodicity Conjecture).** $Q(Q_{\text{odd}}) \subseteq Q_{\text{odd}}$.

**Remark.** In earlier sources, such as Lagarias’ [Lag85], the Periodicity Conjecture was stated as “$Q(Q_{\text{odd}}) = Q_{\text{odd}}$,” but Bernstein showed in [Ber94] that $Q_{\text{odd}} \subseteq Q(Q_{\text{odd}})$, effectively weakening the hypothesis of the conjecture.

The Periodicity Conjecture implies the Divergent Orbits Conjecture since it is well-known that every $x \in Q_{\text{odd}}$ can be expressed as $x = x_0x_1 \cdots x_{m-1}x_mx_{m+1}x_{m+2} \cdots x_{m+n}$ for some $m,n \in \mathbb{Z}^+$. Thus, if for every $x \in Q_{\text{odd}}$, there exists a $y \in Q_{\text{odd}}$ such that $Q(x) = y$, $x$ would eventually enter a $T$-cycle due to the repeating nature of $y$, its parity vector under $T$. 
Chapter 2: Cycles in $\mathcal{F}$

As we will see, the Periodicity Conjecture is a promising rephrasing of the Divergent Orbits Conjecture, which makes it a viable possibility for helping to prove the Collatz Conjecture.

2. Bernstein’s Inverse of $Q$

Although $Q$ is simple to state, its iterative nature makes it difficult to analyze the orbits of arbitrary $x \in \mathbb{Z}_2$ under $T$, which would naturally help us in proving the Collatz Conjecture. In [Ber94], Bernstein managed to construct and express the inverse of $Q$ in the form of a non-iterative algebraic expression, which will prove useful in our endeavors.

**Definition.** Let $\Phi : \mathbb{Z}_2 \to \mathbb{Z}_2$ such that for every $x \in \mathbb{Z}_2$,

$$
\Phi(x) = \sum_{i=0}^{\infty} \frac{-x_i}{3^{S_i(x)}} 2^i,
$$

where $S_i(x) = x_0 + x_1 + \cdots + x_i$.

**Remark.** We can informally think of $S_i(x)$ as the number of 1s in the first $i + 1$ entries of $x = x_0 x_1 x_2 \cdots$.

Bernstein shows in [Ber94] that $\Phi$ is a bijection from $\mathbb{Z}_2$ to itself, and that both $\Phi$ and $\Phi^{-1}$ are continuous. In addition, he also proves that $\Phi$ is a morphism, and thus a conjugacy, from $\sigma$ to $T$. The following theorem, whose proof can be found in [Ber94], succinctly summarizes Bernstein’s results.

**Theorem 2.7 (Bernstein [Ber94]).** $\Phi$ is a bijection with $Q$ as its inverse.

**Example 2.8.** For $x = \overline{1}$, $\Phi(x) = -\frac{1}{3} + \frac{-2}{3^2} + \cdots = \frac{-1}{3} - \frac{2}{1 - \frac{2}{3}} = -1$. In other words, $-1$ is the unique 2-adic integer with parity vector $\overline{1}$ in $T$; this can be verified by computing $Q(-1)$. 

EXAMPLE 2.9. For \( x = \overline{10} \), \( \Phi(x) = -\frac{1}{3} + \frac{-2^2}{3^3} + \cdots = -\frac{1}{3} \cdot \frac{2^2}{1 - \frac{2}{3}} = 1 \). Similarly, it can be shown that for \( x = \overline{01} \), \( \Phi(x) = 2 \). Thus, any 2-adic integer with a parity vector that has the repeating stem \( \overline{10} \) or \( \overline{01} \) has a \( T \)-orbit that eventually enters the \( \{1, 2\} \) cycle.

The observant reader can see from the previous examples that parity vectors of the form \( x = x_0 x_1 \cdots x_{n-1} \) have images in \( \Phi \) that can be expressed as the finite sum of one or more geometric series. Our work in Section 4 will formalize this observation.

It is also straightforward to see that for every \( x \in \mathbb{Q}_{\text{odd}} \), there exists a \( y \in \mathbb{Q}_{\text{odd}} \) such that \( y = \Phi(x) \). Therefore, \( \mathbb{Q}_{\text{odd}} \subseteq \mathbb{Q}(\mathbb{Q}_{\text{odd}}) \). For one possible proof, see Bernstein’s [Ber94]. As stated earlier, showing that \( \mathbb{Q}(\mathbb{Q}_{\text{odd}}) \subseteq \mathbb{Q}_{\text{odd}} \) would be sufficient to prove the Periodicity Conjecture, and so the simplicity with which it can be shown that \( \mathbb{Q}_{\text{odd}} \subseteq \mathbb{Q}(\mathbb{Q}_{\text{odd}}) \) is promising.

3. Generalizing the Inverse of \( \mathbb{Q} \) to \( \mathbb{Q}_{a,b,c,d} \)

At the end of [Ber94], Bernstein mentions without proof that \( \Phi \) can be generalized to all functions \( f : \mathbb{Z}_2 \rightarrow \mathbb{Z}_2 \) such that for some \( c, d \in \mathbb{Z}_2 \) with \( c \) and \( d \) odd,

\[
f(x) = \begin{cases} 
\frac{x}{2} & \text{if } x \text{ is even} \\
\frac{cx + d}{2} & \text{if } x \text{ is odd.}
\end{cases}
\]
In our notation, this is the family of functions \( f_{1,0,c,d} \) for arbitrary odd \( c \) and \( d \). As it turns out, \( \Phi \) can actually be generalized to the entire family \( F \). We first define this generalization and then proceed to prove some important facts concerning it.

**Definition.** Given \( a, b, c, d \in \mathbb{Z}_2 \) with \( a, c, d \) odd and \( b \) even, let
\[
\Phi_{a,b,c,d}(x) = \sum_{i=0}^{\infty} \frac{-\left(b(1 - x_i) + dx_i\right)}{a^{i+1 - S_i(x)}c^{S_i(x)}} 2^i, \quad \text{where} \quad S_i(x) = x_0 + x_1 + \cdots + x_i.
\]

**Remark.** As with \( \Phi \), \( S_i(x) \) can be informally thought of as the number of 1s in the first \( i + 1 \) entries of \( x = x_0x_1\cdots \). Similarly, \( i + 1 - S_i(x) \) is the number of 0s in the first \( i + 1 \) entries of \( x = x_0x_1\cdots \).

With \( \Phi_{a,b,c,d} \) defined, we now show that for every \( f_{a,b,c,d} \in F \), \( \sigma \) and \( f_{a,b,c,d} \) are topologically conjugate via \( \Phi_{a,b,c,d} \).

**Theorem 2.10.** Given \( f_{a,b,c,d} \) with \( a, b, c, d \in \mathbb{Z}_2 \) and \( a, c, d \) odd and \( b \) even, \( \Phi_{a,b,c,d} \) is a topological conjugacy between \( f_{a,b,c,d} \) and \( \sigma \).

We need two lemmas in order to prove Theorem 2.10.

**Lemma 2.11.** For any \( x \in \mathbb{Z}_2 \), \( \Phi_{a,b,c,d}(x_0x_1\cdots) \equiv x_0 \mod 2 \).

**Proof.** The only term of \( \sum_{i=0}^{\infty} \frac{-\left(b(1 - x_i) + dx_i\right)}{a^{i+1 - S_i(x)}c^{S_i(x)}} 2^i \) that can affect the parity of \( \Phi_{a,b,c,d}(x) \) is the first one, as it is the only term that can possibly be odd (all the rest have a positive power of 2 as a factor). We can divide the remainder of the proof into two cases:

1. \( x_0 = 0 \)
   
   The first term of \( \sum_{i=0}^{\infty} \frac{-\left(b(1 - x_i) + dx_i\right)}{a^{i+1 - S_i(x)}c^{S_i(x)}} 2^i \) is \( -\frac{b}{a} \), which is even, and so \( \Phi_{a,b,c,d}(x) \equiv x_0 \mod 2 \).
Chapter 2: Cycles in \( \mathcal{F} \)

(2) \( x_0 = 1 \)

The first term of \( \sum_{i=0}^{\infty} \frac{-(b(1-x_i) + dx_i)}{a^{i+1-S_i(x)} c^S_i(x)} 2^i \) is \( -\frac{d}{c} \), which is odd, and so \( \Phi_{a,b,c,d}(x) \equiv x_0 \mod 2. \)

\[ \square \]

**Lemma 2.12.** \( \Phi_{a,b,c,d} \) is a morphism from \( f_{a,b,c,d} \) to \( \sigma \).

**Proof.** Let \( h = f_{a,b,c,d} \) and \( x \in \mathbb{Z}_2 \). We need to show that \( h \circ \Phi_{a,b,c,d} = \Phi_{a,b,c,d} \circ \sigma \).

The proof that \( h \circ \Phi_{a,b,c,d} = \Phi_{a,b,c,d} \circ \sigma \) can be divided into two cases:

(1) \( x_0 = 0 \)

\[ h \circ \Phi_{a,b,c,d}(x) = h(\Phi_{a,b,c,d}(x)) \]

\[ = h \left( \sum_{i=1}^{\infty} \frac{-(b(1-x_i) + dx_i)}{a^{i+1-S_i(x)} c^S_i(x)} 2^i \right) \]

\[ = h \left( -\frac{b}{a} + \sum_{i=1}^{\infty} \frac{-(b(1-x_i) + dx_i)}{a^{i+1-S_i(x)} c^S_i(x)} 2^i \right) \]

\[ = \frac{a}{2} \left( -\frac{b}{a} + \sum_{i=1}^{\infty} \frac{-(b(1-x_i) + dx_i)}{a^{i+1-S_i(x)} c^S_i(x)} 2^i \right) + b \]

\[ = \sum_{i=1}^{\infty} \frac{-(b(1-x_i) + dx_i)}{a^{i-S_i(x)} c^S_i(x)} 2^{i-1} \]

\[ = \sum_{i=0}^{\infty} \frac{-(b(1-x_{i+1}) + dx_{i+1})}{a^{i+1-S_{i+1}(x)} c^{S_{i+1}(x)}} 2^i \]

\[ = \Phi_{a,b,c,d}(\sigma(x)) \]

\[ = \Phi_{a,b,c,d} \circ \sigma(x) \]

(2) \( x_0 = 1 \)

\[ h \circ \Phi_{a,b,c,d}(x) = h(\Phi_{a,b,c,d}(x)) \]
Therefore, $\Phi_{a,b,c,d}$ is a morphism from $f_{a,b,c,d}$ to $\sigma$. \qed

We can now prove Theorem 2.10.

**Proof of Theorem 2.10.** From Lemma 2.12, we know that $\Phi_{a,b,c,d}$ is a morphism from $f_{a,b,c,d}$ to $\sigma$. It remains to be shown that $\Phi_{a,b,c,d}$ is a bijection and that both $\Phi_{a,b,c,d}$ and $\Phi_{a,b,c,d}^{-1}$ are continuous.

To prove that $\Phi_{a,b,c,d}$ is bijective, we show that for every $k \in \mathbb{Z}^+$, $x \equiv y \mod 2^k \Leftrightarrow \Phi_{a,b,c,d}(x) \equiv \Phi_{a,b,c,d}(y) \mod 2^k$. In other words, the first $k$ entries of $\Phi_{a,b,c,d}(x)$ and $\Phi_{a,b,c,d}(y)$ are equal. First, observe that for any $x, y \in \mathbb{Z}_2$ and $k \in \mathbb{Z}^+$ such that $x \equiv y \mod 2^k$ (in other words, the first $k$ entries of $x$ and $y$ are identical), the definition of $\Phi_{a,b,c,d}$ implies that $\Phi_{a,b,c,d}(x) \equiv \Phi_{a,b,c,d}(y) \mod 2^k$.

Conversely, for any $x, y \in \mathbb{Z}_2$ such that $\Phi_{a,b,c,d}(x) \equiv \Phi_{a,b,c,d}(y) \mod 2^k$ for every $k \in \mathbb{Z}^+$, it follows that $x \equiv y \mod 2^k$. This can be verified
Chapter 2: Cycles in $\mathcal{F}$

by induction on $k$. The basis step follows directly from Lemma 2.11. For the inductive step, assume that $\Phi_{a,b,c,d}(x) \equiv \Phi_{a,b,c,d}(y) \mod 2^k$ implies that $x \equiv y \mod 2^k$. If we assume that $\Phi_{a,b,c,d}(x) \equiv \Phi_{a,b,c,d}(y) \mod 2^{k+1}$ and use the fact that $Q_{a,b,c,d}(z) \equiv Q_{a,b,c,d}(z + 2^{k+1}) \mod 2^{k+1}$ for every $z \in \mathbb{Z}_2$ (see [Fra97]), it follows from Lemma 2.12 that $\Phi_{a,b,c,d}(x_kx_{k+1}\cdots) \equiv \Phi_{a,b,c,d}(y_ky_{k+1}\cdots) \mod 2$:

$$\Phi_{a,b,c,d}(x_kx_{k+1}\cdots) = \Phi_{a,b,c,d}(\sigma^k(x_0x_1\cdots))$$

$$= f^k(\Phi_{a,b,c,d}(x_0x_1\cdots))$$

$$\equiv f^k(\Phi_{a,b,c,d}(y_0y_1\cdots)) \mod 2$$

$$\equiv \Phi_{a,b,c,d}(\sigma^k(y_0y_1\cdots)) \mod 2$$

$$\equiv \Phi_{a,b,c,d}(y_ky_{k+1}\cdots) \mod 2.$$  

By Lemma 2.11, $\Phi_{a,b,c,d}(x_kx_{k+1}\cdots) \equiv \Phi_{a,b,c,d}(y_ky_{k+1}\cdots) \mod 2$ implies that $x_k = y_k$. Furthermore, by the inductive hypothesis, $\Phi_{a,b,c,d}(x) \equiv \Phi_{a,b,c,d}(y) \mod 2^k$ implies that $x \equiv y \mod 2^k$. In other words, the first $k+1$ entries of $x$ and $y$ are identical, and so $\Phi_{a,b,c,d}(x) \equiv \Phi_{a,b,c,d}(y) \mod 2^{k+1}$ implies $x \equiv y \mod 2^{k+1}$. Therefore, for every $k \in \mathbb{Z}^+$, $x \equiv y \mod 2^k$.$\implies \Phi_{a,b,c,d}(x) \equiv \Phi_{a,b,c,d}(y) \mod 2^k$.

Since $x \equiv y \mod 2^k$.$\iff \Phi_{a,b,c,d}(x) \equiv \Phi_{a,b,c,d}(y) \mod 2^k$ for every $k \in \mathbb{Z}^+$, $\Phi_{a,b,c,d}$ is solenoidal (see Chapter 3 for a precise definition). Bernstein and Lagarias [BL96] showed that any solenoidal mapping is also a bijection, and so $\Phi_{a,b,c,d}$ is bijective.
We next show that $\Phi_{a,b,c,d}$ is continuous. Let $x_1, x_2 \in \mathbb{Z}_2$ and $\epsilon \in \mathbb{R}^+$. We then define $\delta = 2^{-n}$, where $n$ is the smallest nonnegative integer such that $2^{-n} < \epsilon$. If $|x_1 - x_2|_2 < \delta$, then the first $n$ entries of $x_1$ and $x_2$ are equal. It follows from the definition of $\Phi_{a,b,c,d}$ that the first $n$ entries of $\Phi_{a,b,c,d}(x_1)$ and $\Phi_{a,b,c,d}(x_2)$ are identical as well. Therefore, $|\Phi_{a,b,c,d}(x_1) - \Phi_{a,b,c,d}(x_2)|_2 < 2^{-n} < \epsilon$, and so $\Phi_{a,b,c,d}$ is continuous.

It remains to be shown that $\Phi^{-1}_{a,b,c,d}$ is continuous as well. Let $y_1, y_2 \in \mathbb{Z}_2$ and $\epsilon \in \mathbb{R}^+$. Since $\Phi_{a,b,c,d}$ is a bijection, there exists a unique $x_1 \in \mathbb{Z}_2$ such that $y_1 = \Phi(x_1)$; similarly, there exists a unique $x_2 \in \mathbb{Z}_2$ such that $y_2 = \Phi(x_2)$. We define $\delta = 2^{-n}$, where $n$ is the smallest nonnegative integer such that $2^{-n} < \epsilon$. If $|y_1 - y_2|_2 < \delta$, the first $n$ entries of $y_1$ and $y_2$, and thus $\Phi(x_1)$ and $\Phi(x_2)$, are equal. In other words, $\Phi(x_1) \equiv \Phi(x_2)$ mod $2^n$. As shown above, $\Phi(x_1) \equiv \Phi(x_2)$ mod $2^n$ implies that $x_1 \equiv x_2$ mod $2^n$, and so the first $n$ entries of $x_1$ and $x_2$ are also identical. Therefore, $|x_1 - x_2|_2 < 2^{-n} < \epsilon$, and so $\Phi^{-1}_{a,b,c,d}$ is continuous.

Therefore, $\Phi_{a,b,c,d}$ is a topological conjugacy from $f_{a,b,c,d}$ to $\sigma$. \qed

Since we have shown that $\Phi_{a,b,c,d}$ is a conjugacy between $f_{a,b,c,d}$ and $\sigma$, we can now show that $\Phi_{a,b,c,d}$ is in fact the inverse of $Q_{a,b,c,d}$.

**Corollary 2.13.** Given $f_{a,b,c,d}$ with $a, b, c, d \in \mathbb{Z}_2$ and $a, c, d$ odd and $b$ even, $\Phi_{a,b,c,d}$ is the inverse of the parity vector function associated with $f_{a,b,c,d}$, $Q_{a,b,c,d}$.

**Proof.** Let $q$ be the parity vector function associated with $f = f_{a,b,c,d}$. We need to show that $q \circ \Phi_{a,b,c,d}(x) = x$. This can be verified with computations involving Theorem 2.10 and Lemma 2.11: let $y = q \circ \Phi_{a,b,c,d}(x)$. By
the definition of the parity vector function, for every \( i \in \mathbb{N} \),

\[
y_i = f^i(\Phi_{a,b,c,d}(x)) \mod 2
= \Phi_{a,b,c,d}(\sigma^i(x)) \mod 2
= x_i.
\]

Therefore, \( y = x \). Since \( q \) is a bijection and \( q \circ \Phi_{a,b,c,d}(x) = x \), \( \Phi_{a,b,c,d} \) must be the unique inverse of \( q \). In other words, \( \Phi_{a,b,c,d} \) is the inverse of the parity vector function associated with \( f_{a,b,c,d} \), \( Q_{a,b,c,d} \). \( \square \)

**Example 2.14.** We showed in Example 1.16 that the parity vector function associated with \( \sigma = f_{1,0,1,-1} \) is the identity map on \( \mathbb{Z}_2 \). Since the inverse of the identity map is itself, \( \Phi_{1,0,1,-1} \) is the identity map as well. For example, note that \( \Phi_{1,0,1,-1}(1) = 1 + 2 + 4 + \cdots = \frac{1}{1-2} = -1 = \overline{1} \).

**Example 2.15.** Since the orbit of \( \frac{8}{3} \) under \( f_{1,2,1,3} \) is \( \frac{8}{3}, \frac{7}{3}, \frac{8}{3}, \ldots \), the parity vector of \( \frac{8}{3} \) under \( f_{1,2,1,3} \) is \( \overline{01} \). As we would expect, \( \Phi_{1,2,1,3}(\overline{01}) = -2 + -3 \cdot 2 + -2 \cdot 2^2 + -3 \cdot 2^3 + \cdots = (-2 + -2 \cdot 2^2 + \cdots) + (-3 \cdot 2 + -3 \cdot 2^3 + \cdots) = \frac{-2}{1-2^2} + \frac{-3 \cdot 2}{1-2^2} = \frac{8}{3} \). Note that it is well-known that every convergent series in the 2-adic integers converges absolutely, and so the rearrangement of terms we did as part of our computations is valid. See [Gou00] for an explanation of why the terms of an absolutely convergent series can be rearranged.

As we have shown, \( \Phi_{a,b,c,d} \) provides a non-iterative algebraic expression for the inverse of \( Q_{a,b,c,d} \), which is a form that we will be able to put to good use.
4. Analyzing T-cycles Using $\Phi$

Although we have a workable formula for the inverse of $Q_{a,b,c,d}$, it is apparent that computing $\Phi_{a,b,c,d}(x)$ for arbitrary $x \in \mathbb{Z}_2$ is not a straightforward task since it is an infinite summation. Luckily, we have seen that parity vectors of the form $x = x_0x_1\cdots x_{n-1}$ for some $n \in \mathbb{Z}^+$ can be expressed as a finite sum of geometric series, which are certainly much easier to compute.

If we think of cycles in terms of their corresponding parity vector, such as the $\{1, 2\}$ cycle in $T$ as $\overline{10}$ or $\overline{01}$, $\Phi_{a,b,c,d}$ gives us a method of computing the values in an arbitrary $f_{a,b,c,d}$-cycle if we know the parity vector of its orbit.

We next formalize this result to make it readily usable.

**Theorem 2.16.** For every finite cycle with parity vector $x = x_0x_1\cdots x_{n-1}$ for some $n \in \mathbb{Z}^+$, $\Phi_{a,b,c,d}(x) = \sum_{i=0}^{n-1} \frac{-(b(1-x_i) + dx_i)}{a^{i+1-S_i(x)}c^{S_i(x)}}2^i$.

**Proof.** Let $x = x_0x_1\cdots x_{n-1}$ be the parity vector of some cycle in $f_{a,b,c,d}$ and $u_i = \frac{-(b(1-x_i) + dx_i)}{a^{i+1-S_i(x)}c^{S_i(x)}}$ for every $i \in \mathbb{N}$. By Corollary 2.13, $\Phi_{a,b,c,d}(x) = \sum_{i=0}^{\infty} u_i2^i$. Since the terms of a convergent 2-adic series can be rearranged without changing the value the series converges to, $\Phi_{a,b,c,d}(x) = \sum_{i=0}^{\infty} u_{ni}2^{ni} + \sum_{i=0}^{\infty} u_{ni+1}2^{ni+1} + \cdots + \sum_{i=0}^{\infty} u_{ni+(n-1)}2^{ni+(n-1)}$, which is a sum of $n$ convergent geometric series with common ratio $\frac{2^n}{a^{n-S_n(x)}c^{S_n(x)}}$. Since the initial terms of these $n$ series are $u_0, u_1, \ldots$, and $u_{n-1}, \Phi_{a,b,c,d}(x) = \sum_{i=0}^{n-1} \frac{-(b(1-x_i) + dx_i)}{a^{i+1-S_i(x)}c^{S_i(x)}}2^i$.

$\square$
Example 2.17. We showed in Example 2.14 that \( \Phi_{1,0,1,-1}(\overline{1}) = \overline{1} \). To illustrate the usage of Theorem 2.16, we recalculate \( \Phi_{1,0,1,-1}(\overline{1}) \) by simply computing
\[
\Phi_{1,0,1,-1}(\overline{1}) = \frac{-1}{1 - \frac{10 \cdot 1^1}{2}} = -1 = \overline{1}.
\]

Example 2.18. Similarly, we arrive at our results from Example 2.15 by recalculating \( \Phi_{1,2,1,3}(\overline{01}) = \frac{8}{3} \) using Theorem 2.16:
\[
\Phi_{1,2,1,3}(\overline{01}) = \frac{-2}{11 \cdot 1^1} 
+ \frac{-3}{11 \cdot 1^2} = \frac{8}{3}.
\]

Example 2.19. Since \( \Phi_{1,0,1,-1}(x) = x \) for every \( x \in \mathbb{Z}_2 \), Theorem 2.16 provides a nice way of determining the value of any 2-adic integer \( a \) of the form \( a = a_0 a_1 \cdots a_{n-1} \) for some \( n \in \mathbb{Z}^+ \). For instance, we can determine that
\[
\overline{101} = \frac{-1}{11 \cdot 1^2} 
+ \frac{0}{11 \cdot 1^1} 
+ \frac{-1}{11 \cdot 1^2} = -\frac{5}{7}.
\]

We have seen that Bernstein’s \( \Phi \) mapping was useful in proving part of the Periodicity Conjecture. As we will soon show, our general mapping \( \Phi_{a,b,c,d} \) can be used to attack the Non-trivial Cycles Conjecture. Since Theorem 2.16 provides a nice expression for calculating \( \Phi_{a,b,c,d} \) for the parity vectors associated with cycles, we have a method of potentially verifying that the only \( T \)-cycle consisting of positive integers is the one with parity vector \( \overline{10} \) or \( \overline{01} \), both of which represent the same cycle. Contrariwise, if we could show that there exists some parity vector \( x = x_0 x_1 \cdots x_{n-1} \) for some \( n \in \mathbb{Z}^+ \) such that \( \Phi(x) \in \mathbb{Z}^+ \), we will have found a counterexample for the Collatz Conjecture.
We now show that a subset of cycles with a certain form of parity vector do not contain any positive integers.

**Theorem 2.20.** For any \( j \in \mathbb{Z}^+ \), there exists at most one \( k \in \mathbb{Z}^+ \) such that \( \underbrace{1 \cdots 1}_{j} \underbrace{0 \cdots 0}_{k} \) is the parity vector of a positive integer under \( T \).

The following lemma, which has a straightforward proof, is used to prove Theorem 2.20:

**Lemma 2.21.** For any \( x, y \in \mathbb{N} \), \( 2^x - 3^y \equiv 0 \mod 3 \) if and only if \( x \) is even and \( y = 0 \).

**Proof.** Let \( x, y \in \mathbb{N} \). First, assume that \( x \) is even and \( y = 0 \). We need to show that \( 2^x - 3^y \equiv 0 \mod 3 \). Then \( 2^x - 3^y = 2^x - 1 \), and so \( 2^x - 1 \equiv 1 - 1 \equiv 0 \mod 3 \).

Conversely, assume that it is not the case that \( x \) is even and \( y = 0 \). We need to show that either \( 2^x - 3^y \equiv 1 \mod 3 \) or \( 2^x - 3^y \equiv 2 \mod 3 \). We have three cases:

1. \( x \) is odd and \( y = 0 \)
   
   \( 2^x - 3^y = 2^x - 1 \), and so \( 2^x - 1 \equiv 2 - 1 \equiv 1 \mod 3 \).

2. \( x \) is even and \( y \neq 0 \)
   
   \( 2^x - 3^y \equiv 1 - 0 \equiv 1 \mod 3 \)

3. \( x \) is odd and \( y \neq 0 \)
   
   \( 2^x - 3^y \equiv 2 - 0 \equiv 2 \mod 3 \)

\[ \square \]

An additional lemma is also needed; its proof follows directly from Theorem 2.16.
Chapter 2: Cycles in \( F \)

**Lemma 2.22.** For every \( y = \overbrace{1 \cdots 10 \cdots 0}^k \), \( \Phi_{3,0,1,1}(y) = 3^j \frac{2^k - 1}{2^{i+k} - 3^i} \).

The conjugacy \( \Phi_{3,0,1,1} \) will be useful in our proof since the structure of the parity vectors under \( f_{3,0,1,1} \) is closely related to those under \( T \).

Finally, the following inequality will also be used:

**Lemma 2.23.** For every \( j \in \mathbb{Z}^+ \), there is at most one \( k \in \mathbb{Z}^+ \) such that
\[
\frac{\ln 3}{\ln 2} j - \frac{\ln 3^j - 1}{\ln 2} < k \leq \frac{\ln 3^j - 1}{\ln 2}.
\]

**Proof.** Let \( g(x) = \frac{\ln 3}{\ln 2} x - x \) and \( h(x) = \frac{\ln 3^j - 1}{\ln 2} \). Note that \( g'(x) = \frac{\ln 3}{\ln 2} - 1 > 0 \) and \( \lim_{x \to \infty} h'(x) = \frac{\ln 3}{\ln 2} - 1 \), and so, since it is readily seen that \( h'(x) \) is increasing, \( h'(x) - g'(x) < 0 \) for all \( x > 0 \). Since \( h(1) - g(1) < 1 \) and \( h'(x) - g'(x) < 0 \), there can be at most one \( k \in \mathbb{Z}^+ \) such that \( g(j) \leq k \leq h(j) \). \( \square \)

We can now prove Theorem 2.20.

**Proof.** Given \( j, k \in \mathbb{Z}^+ \), let \( x \in \mathbb{Z}_2 \) have the parity vector \( \overbrace{1 \cdots 10 \cdots 0}^j \overbrace{1 \cdots 1}^k \) under \( T \). Consider \( f_{3,0,1,1}(x) \), which is topologically conjugate to \( T(x) \) by the homeomorphism \( h(x) = x + 1 \). The theorem is equivalent to showing that there is at most one cycle of \( f_{3,0,1,1} \) in \( h[\mathbb{Z}^+] = \{2, 3, 4, \ldots\} \) with a parity vector of the form \( \overbrace{0 \cdots 0}^{j} \overbrace{1 \cdots 1}^{k} \) and \( y = x + 1 \). Lemma 2.22 shows that \( y = 3^j \frac{2^k - 1}{2^{i+k} - 3^i} \).

We now have 3 cases:
Chapter 2: Cycles in $F$

(1) $k < \frac{\ln 3}{\ln 2} j - j$

Straightforward computation shows that the numerator and denominator of $\frac{2^k - 1}{2^{j+k} - 3^j}$ are positive and negative, respectively, and so $y \not\in \{2, 3, 4, \ldots\}$.

(2) $\frac{\ln 3}{\ln 2} j - j \leq k \leq \frac{\ln 3^j - 1}{\ln 2}$

Lemma 2.23 implies that there exists at most one value of $k$ such that $\frac{\ln 3}{\ln 2} j - j \leq k \leq \frac{\ln 3^j - 1}{\ln 2}$. For such a $k$, the numerator of $\frac{2^k - 1}{2^{j+k} - 3^j}$ is greater than the denominator, and so there is only one possible integer value of $y \in \{2, 3, \ldots\}$ on this interval.

(3) $k > \frac{\ln 2^j - 1}{\ln 2}$

By Lemma 2.21, $2^{j+k} - 3^j$ does not divide $3^j$. Furthermore, since the numerator of $\frac{2^k - 1}{2^{j+k} - 3^j}$ is less than the denominator on this interval, $y \not\in \{2, 3, 4, \ldots\}$

Therefore, there is at most one $k \in \mathbb{Z}^+$ such that $x \in \{2, 3, 4, \ldots\}$. □

In the proof of Theorem 2.20 above, the only value of $k$ that could possibly correspond to the parity vector of a cycle consisting of positive integers is one such that $\frac{\ln 3}{\ln 2} j - j \leq k \leq \frac{\ln 3^j - 1}{\ln 2}$. As we expect, $k = 1$ lies on this interval when $j = 1$: $\overline{10}$ corresponds to what we hope is the only cycle that contains positive integers. If we could prove that there is no value of $k$ that satisfies the above condition when $j > 1$, Theorem 2.20 could be strengthened to say that the only cycle with corresponding parity vector of the form
that consists of positive integers is \( \overline{10} \). For now, we leave this potential refinement as an open question.

We also note that in the process of proving Theorem 2.20, we were able to utilize a mapping conjugate to \( T \), namely \( f_{3,0,1,1} \), to achieve our desired result. Dealing with the images of \( \Phi \) alone would have been difficult, but strategic usage of conjugacy allowed us to circumvent this challenge. In doing so, we have made a potential step towards proving the Non-trivial Cycles Conjecture.
CHAPTER 3

Endomorphisms of the Shift Map

So far, we have seen that although the shift map $\sigma$ has behavior that has been well-documented (see [Hed69] for an especially thorough study), its dynamics are difficult to relate back to $T$ because of the complexity surrounding the conjugacy between them, namely the parity vector function $Q$. In [Mon08], Monks provides one possible solution to this dilemma: since conjugacy is an equivalence relation, any mapping $f$ that is conjugate to $\sigma$ will also be conjugate to $T$ as well. In this chapter, we provide the reader with the relevant studies that have been conducted on the shift map, summarize Monks’ discovery of a mapping with a particularly nice conjugacy to both $\sigma$ and $T$, and discuss methods which might prove useful for finding other conjugacies of $\sigma$.

1. Continuous Endomorphisms of $\sigma$

Endomorphisms are particularly important to our study of conjugacies of $\sigma$, so we will start by defining them.

**Definition.** Let $f : X \to X$. We say that $g : X \to X$ is an endomorphism of $f$ if $f \circ g = g \circ f$. In the special case where $g$ is bijective, $g$ is said to be an autoconjugacy of $f$.

**Remark.** Note that endomorphisms are simply morphisms from a function to itself.
EXAMPLE 3.1. Every function \( f : X \to X \) has two trivial endomorphisms, namely itself and the identity map \( I \) on \( X \) (since \( f \circ I = f = I \circ f \)). Since \( I \) is a bijection, it also happens to be an autoconjugacy of \( f \).

EXAMPLE 3.2. In [Hed69], Hedlund showed that \( \sigma \) has exactly two autoconjugacies: the identity map on \( \mathbb{Z}_2 \) and the “bit flip map” \( V(x) = -1 - x \), which has the effect of interchanging the 0s and 1s in \( x \). For instance, \( \sigma \circ V(1100) = \sigma(0011) = 011 = V(100) = V \circ \sigma(1100) \).

We are particularly interested in continuous endomorphisms \( h \) of \( \sigma \) since it is necessary that \( h \) be continuous if two mappings are to be topologically conjugate by \( h \). Fortunately, the classification of these particular endomorphisms is well-known. The curious reader may refer to either [Hed69] or [LM95] for a rigorous treatise on the topic.

DEFINITION. Let \( n \in \mathbb{Z}^+ \). We define \( B_n \) to be the set of all finite sequences \( a_0 a_1 \ldots a_{n-1} \) such that for every \( i \in \{0, 1, \ldots, n-1\} \), \( a_i \in \{0, 1\} \).

EXAMPLE 3.3. To illustrate, \( B_1 = \{0, 1\}, B_2 = \{00, 01, 10, 11\} \), and so on.

DEFINITION. Let \( n \in \mathbb{Z}^+ \) and \( f : B_n \to \{0, 1\} \). We define \( f_\infty : \mathbb{Z}_2 \to \mathbb{Z}_2 \) such that for \( x, y \in \mathbb{Z}_2 \), \( f_\infty(x) = y \), where \( y_i = f(x_i, x_{i+1}, \ldots, x_{i+n-1}) \) for every \( i \in \mathbb{N} \).

EXAMPLE 3.4. The shift map \( \sigma \) can be expressed in terms of this type of mapping. Let \( f : B_2 \to \{0, 1\} \) such that \( f(00) = f(10) = 0 \) and \( f(01) = f(11) = 1 \). It is easily verified that \( f_\infty = \sigma \).

EXAMPLE 3.5. For \( n \geq 2 \), some mappings with blocks in \( B_n \) are equivalent to mappings in \( B_{n-1} \). For instance, for \( f : B_3 \to \{0, 1\} \) such that
Chapter 3: Endomorphisms of the Shift Map

\[ f(000) = f(001) = 0, \ f(010) = f(011) = 1, \ f(100) = f(101) = 1, \text{ and} \]
\[ f(110) = f(111) = 0, \ f_\infty = \sigma \] since the image of \( f \) has no dependence on the third entry of the block.

In general, these “sliding block codes” (to borrow Lind’s terminology) compose the continuous endomorphisms of the shift map.

**Theorem 3.6 (Monks [Mon08]).** The continuous endomorphisms of the shift map \( \sigma \) are precisely those \( f_\infty \) for some \( n \in \mathbb{Z}^+ \) and \( f : \mathcal{B}_n \to \{0, 1\} \).

**Example 3.7.** Theorem 3.6 makes the continuous endomorphisms of \( \sigma \) easy to enumerate: for every \( n \in \mathbb{Z}^+ \), there are \( 2^{n+1} \) continuous endomorphisms formed from blocks of size \( n \) since \( |\mathcal{B}_n| = 2^n \) and there are 2 possible images for each of these \( 2^n \) blocks.

2. Solenoidal Parity Vector Functions

One additional challenge that arises from finding conjugacies of \( \sigma \) is the necessary stipulation that a conjugacy be bijective. Fortunately, there are well-known conditions that are sufficient for a map to be bijective. Much of Monks’ work was concerned with one such condition, as it is particularly simple to demonstrate for parity vector functions.

**Definition.** Let \( f : \mathbb{Z}_2 \to \mathbb{Z}_2 \). The parity vector function associated with \( f \) is the mapping \( p_f \) such that for \( x, y \in \mathbb{Z}_2 \), \( p_f(x) = y \) if \( y_i \equiv f_i(x) \mod 2 \) for every \( i \in \mathbb{N} \).

**Remark.** It is readily apparent that this definition extends our existing one for \( Q_{a,b,c,d} \).
DEFINITION. Let $g : \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$. We say that $g$ is **solenoidal** if for every $x, y \in \mathbb{Z}_2$ and $k \in \mathbb{Z}^+$, $x \equiv y \mod 2^k$ if and only if $g(x) \equiv g(y) \mod 2^k$.

Bernstein and Lagarias showed in [BL96] that $Q$ is bijective by demonstrating that it is solenoidal. Their argument generalizes to all maps, thus establishing a sufficient condition for bijectivity.

Monks showed that every continuous endomorphism $g$ of $\sigma$ with a solenoidal parity vector function $p_g$ is conjugate to $\sigma$ by $p_g$. We define one especially notable endomorphism which satisfies the given conditions.

DEFINITION. Let $f : \mathbb{B}_2 \rightarrow \{0, 1\}$ such that $f(00) = 0, f(01) = 1, f(10) = 1,$ and $f(11) = 0$. We say that $f_\infty$ is the **discrete derivative** $D$.

REMARK. Our motivation for the term “discrete derivative” can be seen by noting that for every $x \in \mathbb{B}_2$, $f(x) = |x_1 - x_0|$.

EXAMPLE 3.8. Monks showed that $D$ has precisely two fixed points, namely $0$ and $10$. Similarly, $\sigma$’s two fixed points are $0$ and $1$.

In addition, Monks showed that $\sigma$ and $D$, alongside their bit-flipped counterparts $V \circ \sigma$ and $V \circ D$, are the only such endomorphisms.

THEOREM 3.9 (Monks [Mon08]). The only continuous endomorphisms of the shift map $\sigma$ with solenoidal parity vector functions are $\sigma, V \circ \sigma, D,$ and $V \circ D$.

3. In Pursuit of New Conjugacies of $\sigma$

Our first task here is to demonstrate that $D$ is a function conjugate to $\sigma$ that is distinct from the family of modular functions $\mathcal{F}$ that we studied in Chapter 2.
Chapter 3: Endomorphisms of the Shift Map

**Theorem 3.10.** D is not a member of \( \mathcal{F} \).

**Proof.** By way of contradiction, assume that D is a member of \( \mathcal{F} \). Then there exists an \( a, b, c, d \in \mathbb{Z}_2 \) with \( a, c, d \) odd and \( b \) even such that \( D = f_{a,b,c,d} \). Since \( D(\overline{0}) = 0, f_{a,b,c,d}(0) = \frac{b}{2} \) implies that \( b = 0 \). In addition, \( D(01\overline{0}) = 11\overline{0} = 3 \) and \( f_{a,0,c,d}(01\overline{0}) = f_{a,0,c,d}(2) = a \) together imply that \( a = 3 \). Finally, \( D(011\overline{0}) = 101\overline{0} = 5 \), but \( f_{3,0,c,d}(011\overline{0}) = f_{3,0,c,d}(6) = \frac{3 \cdot 6 + 0}{2} = 9 \neq 5 \). Therefore, D is not a member of \( \mathcal{F} \). \( \square \)

In addition, it is important to note that in our search for conjugacies, we need not restrict our attention to solenoidal parity vector functions. After all, a continuous endomorphism of \( \sigma \) could be conjugate to \( \sigma \) by a mapping other than a parity vector function (a trivial example is the identity map on \( \mathbb{Z}_2 \)). One potential way of looking for (or ruling out) endomorphisms that are also conjugate to \( \sigma \) is to compare their dynamics to those of \( \sigma \). The following theorem, which provides a restriction on the images of blocks of \( B_n \) that are shift-equivalent (such as 101, 011, and 110), demonstrates this approach in action.

**Theorem 3.11.** Let \( n \in \{3, 4, 5, \ldots \} \) and \( a \in B_n \) such that the cardinality of the set of elements of \( \mathbb{Z}_2 \) that are shift-equivalent to \( a_0 \cdots a_{n-1} \), \( S = \{ \sigma^k(a_0 \cdots a_{n-1}) | k \in \mathbb{N} \} \), is greater than 2. Then for any mapping \( f : B_n \to \{0, 1\} \) such that \( S \) is a subset of the preimage of either 0 or 1 under \( f \), \( f_\infty \) is not conjugate to \( \sigma \).

To prove Theorem 3.11, we will need to make use of a lemma concerning the dynamics of \( \sigma \). This lemma will also prove useful for our work in Chapter 4.
Lemma 3.12. Let $n \in \mathbb{Z}^+$. Then there are precisely $2^n$ points that reach a fixed point after exactly $n$ iterations of $\sigma$. In particular, $2^{n-1}$ points reach the fixed point $\bar{0}$, and the remaining $2^{n-1}$ points reach the fixed point $\bar{1}$.

Proof. Let $x \in \mathbb{Z}_2$ and $n \in \mathbb{Z}^+$ such that the orbit of $x$ under $\sigma$ reaches a fixed point after exactly $n$ iterations. It is easily seen that $x$ is of the form $x_0x_1 \cdots x_{n-2}1\bar{0}$ or $x_0x_1 \cdots x_{n-2}0\bar{1}$. Straightforward computation shows that there are exactly $2^{n-1}$ points that reach $\bar{0}$ after exactly $n$ iterations, and exactly $2^{n-1}$ points that reach $\bar{1}$ after exactly $n$ iterations. Therefore, there are $2^{n-1} + 2^{n-1} = 2^n$ points that reach a fixed point of $\sigma$ after exactly $n$ iterations. \Box

Example 3.13. By Lemma 3.12, there are precisely $2^3 = 8$ points that reach a fixed point of $\sigma$ after exactly 3 iterations. It is readily verifiable that these points are $001\bar{0}$, $011\bar{0}$, $101\bar{0}$, $111\bar{0}$, $000\bar{1}$, $010\bar{1}$, $100\bar{1}$, and $110\bar{1}$. Note that 4 of these eventually reach $\bar{0}$, and the remaining 4 reach $\bar{1}$.

We may now prove Theorem 3.11.

Proof of Theorem 3.11. Let $n \in \{3, 4, 5, \ldots\}$, $a \in \mathcal{B}_n$, $S = \{\sigma^k(a_0 \cdots a_{n-1}) | k \in \mathbb{N}\}$ such that $|S| > 2$, and $f : \mathcal{B}_n \to \{0, 1\}$ such that $S$ is a subset of the preimage of either 0 or 1 under $f$. To show that $f_\infty$ is not conjugate to $\sigma$, it suffices to demonstrate that $f_\infty$ exhibits different dynamics than $\sigma$.

We will divide our argument into three cases:

(1) $f(\underbrace{0 \cdots 0}_n) = f(\underbrace{1 \cdots 1}_n)$
Without loss of generality, assume that \( f(\underbrace{0 \cdots 0}_n) = f(\underbrace{1 \cdots 1}_n) = 0 \).

It is apparent that \( \overline{0} \) is a fixed point of \( f_\infty \). By assumption, for every \( x \in S \), either \( f_\infty(\underbrace{x_0 \cdots x_{n-1}}_n) = \overline{0} \) or \( \overline{x_0 \cdots x_{n-1}} = \overline{1} \). Therefore, every element of \( S \) reaches the fixed point \( \overline{0} \) after either exactly one or exactly two iterations. However, in the former case, it follows from Lemma 3.12 that there is only one point that reaches each fixed point of \( \sigma \) after exactly one iteration. Similarly, there are only two points that reach each fixed point of \( \sigma \) after exactly two iterations. Therefore, \( f_\infty \) has more eventually fixed points after either one or two iterations than \( \sigma \), and so the two mappings exhibit different dynamics.

(2) \( f(\underbrace{0 \cdots 0}_n) = 0 \) and \( f(\underbrace{1 \cdots 1}_n) = 1 \)

It is readily checked that \( \overline{0} \) and \( \overline{1} \) are both fixed points of \( f_\infty \). For every \( x \in S \), either \( f_\infty(\underbrace{x_0 \cdots x_{n-1}}_n) = \overline{0} \) or \( \overline{x_0 \cdots x_{n-1}} = \overline{1} \), and so \( \overline{x_0 \cdots x_{n-1}} \) reaches a fixed point in \( f_\infty \) after exactly one iteration. Since \( |S| > 2 \), \( f_\infty \) has more than two points that reach a fixed point after exactly one iteration. However, by Lemma 3.12, \( \sigma \) has only two such points. Therefore, \( f_\infty \) and \( \sigma \) have different dynamics.

(3) \( f(\underbrace{0 \cdots 0}_n) = 1 \) and \( f(\underbrace{1 \cdots 1}_n) = 0 \)

Since \( f_\infty(\overline{0}) = \overline{1} \) and \( f_\infty(\overline{1}) = \overline{0} \), it is apparent that \( \{\overline{0}, \overline{1}\} \) is a cycle of length 2 under iteration of \( f_\infty \). For every \( x \in S \), either \( f_\infty(\underbrace{x_0 \cdots x_{n-1}}_n) = \overline{0} \) or \( \overline{x_0 \cdots x_{n-1}} = \overline{1} \), and so each \( \overline{x_0 \cdots x_{n-1}} \) enters an \( f_\infty \)-cycle of length 2 after exactly one iteration. (Note that \( \overline{x_0 \cdots x_{n-1}} \) cannot start out in a 2-cycle since by our assumption that
Since \(|S| > 2\), \(x\) cannot equal \(0\) or \(1\), as \(|S|\) would be \(1\). Since \(|S| > 2\), there are more than two such points. However, there are only two points that enter a \(\sigma\)-cycle of length 2 after one iteration, namely 010 and 110, and so \(\sigma\) and \(f_{\infty}\) possess different dynamics.

Since the dynamics of \(f_{\infty}\) and \(\sigma\) are different, they cannot be conjugate to each other. \(\square\)

**Example 3.14.** Let \(n \in \{3, 4, 5, \ldots\}\). We can then define the “extended discrete derivative” \(D_n\) by \(D_n = f_{\infty}\) for \(f : \mathcal{B}_n \to \{0, 1\}\) such that \(f(x) = x_n - x_{n-1} - \cdots - x_0 \mod 2\) for every \(x \in \mathcal{B}_n\). Since \(\left\{ \sigma^k(\overbrace{\{0, \ldots, 0\}}^{n-1}) \mid k \in \mathbb{N} \right\} = n > 2\) and \(D_n(\sigma^k(\overbrace{\{0, \ldots, 0\}}^{n-1})) = 1\) for every \(k \in \mathbb{N}\), it follows directly from Theorem 3.11 that \(D_n\) is not conjugate to \(\sigma\).

In this section, we have continued the work started by Monks by investigating potential conjugacies of \(T\) and \(\sigma\) that might be simpler to study than the parity vector function \(Q\) by considering the continuous endomorphisms of \(\sigma\). By enhancing our understanding of the conjugacies of \(\sigma\), we stand to gain more tools for approaching the Collatz Conjecture. In addition, the counting arguments we developed here will be useful in Chapter 4, where we extend the existing classification of \(D\)’s dynamics.
CHAPTER 4

The Dynamics of D Revisited

In Chapter 3, we discussed the conjugacies of σ that arise when analyzing its continuous endomorphisms. In particular, the conjugacy of D possesses rich dynamics and beautiful symmetry that are not only fascinating to study, but relate nicely back to our goal of proving the Collatz Conjecture. To this end, we will devote this chapter to extending Monks’ analysis of D’s dynamics, which can be found in [Mon08].

1. The Symmetry between D and Its Parity Vector Function P

As mentioned before, the conjugacy between σ and T, namely Q, is complicated to work with, even when considering its non-iterative inverse Φ. In this section, we will show that the conjugacy between D and σ is in many ways much easier to handle thanks to the symmetry between D and the parity vector function associated with it. Due to the significance of this conjugacy, we will emphasize it with special notation.

DEFINITION. We let P denote the parity vector function associated with D.

The following theorem by Monks follows directly from the transitivity of conjugacy.
Theorem 4.1 (Monks [Mon08]). The map $R = \Phi \circ P$ is a conjugacy from $D$ to $T$.

Using $R$, Monks was able to reformulate the Collatz Conjecture, thus demonstrating the intrinsic connection between the mappings $D$ and $T$.

Theorem 4.2 (Monks [Mon08]). The following statements are equivalent:

1. The Collatz Conjecture is true.
2. For every $m \in \mathbb{Z}^+$, $R^{-1}(m)$ has reduced form $x_0x_1x_2x_3\cdots x_{2n+1}$ for some $n \in \mathbb{N}$.

Remark. We provide a precise definition of “reduced form” later in this chapter.

Example 4.3. Since $R(11\overline{0}) = \Phi \circ P(11\overline{0}) = \Phi(\overline{10}) = 1$, $R^{-1}(1) = 11\overline{0}$, as Theorem 4.2 suggests.

Note that understanding $P^{-1}$ allows us to better understand $R^{-1}$. Fortunately, the relationship between $P$ and $P^{-1}$ is much simpler than that of $Q$ and $\Phi$. The following theorem, which is the basis for $D$’s exquisite symmetry, compactly expresses this simplicity.

Theorem 4.4 (Monks [Mon08]). $P = P^{-1}$.

Example 4.5. By Theorem 4.4, $P = P^{-1}$, which implies that $P^2 = I$, where $I$ is the identity map on $\mathbb{Z}_2$. To illustrate, $P^2(11\overline{0}) = P(P(11\overline{0})) = P(\overline{10}) = 11\overline{0}$.

Example 4.6. The analogous statement for $Q$ and $\Phi$ is not true in general. Observe that $Q(3) = 1100\overline{11} = -\frac{23}{3}$, but $\Phi(3) = -\frac{5}{9} \neq -\frac{23}{3}$. 

Chapter 4: The Dynamics of D Revisited

42

Theorem 4.4 has two important consequences:

(1) Since $\mathcal{D}$ is its own inverse, we can understand $\mathcal{D}^{-1}$ by understanding $\mathcal{D}$.

(2) Results concerning the dynamics of $\mathcal{D}$ apply nicely to $\mathcal{D}$, and thus $\mathcal{D}^{-1}$.

Thus, in an effort to prove the second statement of Theorem 4.2, we will spend the remainder of this chapter studying the dynamics of $\mathcal{D}$. Before we can proceed, we must first define some simple terminology to facilitate our discussion.

**Definition.** Let $x \in \mathbb{Z}_2$. We say that $x$ is **eventually repeating** if $x = x_0x_1\cdots x_{t-1}x_t\cdots x_{t+m-1}$. Similarly, we say that $x$ is **repeating** when $x = x_0\cdots x_{m-1}$.

In other words, a 2-adic integer is eventually repeating if it consists of a finite stem of 0s and 1s followed by a repeating part. Note that the eventually periodic points of $\sigma$ are precisely those $x \in \mathbb{Z}_2$ such that $x$ is eventually repeating.

**Definition.** Let $x \in \mathbb{Z}_2$ such that $x = x_0x_1\cdots x_{t-1}x_t\cdots x_{t+m-1}$ (i.e. $x$ is eventually repeating). We say that $x$ is in **reduced form** if and only if $x_{t-1} \neq x_{t+m-1}$ and $m$ is the smallest integer such that $x$ can be expressed in this form. Furthermore, if $x$ is in reduced form, we say that $\|x\| = m$ and $x = t$. Similarly, if $x = x_0\cdots x_{m-1}$ (i.e. $x$ is repeating), $x$ is in **reduced form** if and only if $m$ is the smallest integer such that $x$ can be expressed in this form, and we define $\|x\| = m$ and $x = 0$. 
Remark. Informally, we can think of $\|x\|$ as a measure of the length of the repeating part of $x$ and $x$ as measure of the length of the finite stem of $x$.

Example 4.7. $\overline{1010}$ is not in reduced form because it violates both of the above conditions. However, we can simplify it into the appropriate form as follows: $1010 = \overline{10} = \overline{10}$. Note that $\|\overline{10}\| = 2$ and $\overline{10} = 0$.

2. Fixed Points of $D$

Monks classified the eventually fixed points of $D$ in [Mon08]. Thanks to the symmetry of $P$, the classification for certain periodic points of $D$ follows nicely, as we will see.

Theorem 4.8 (Monks [Mon08]). Let $x \in \mathbb{Z}_2$. The orbit of $x$ under $D$ reaches a fixed point after at most $2^n$ iterations if and only if for some $n \in \mathbb{N}$ the reduced form of $x$ is either $\overline{x_0x_1\cdots x_{2^n-1}}$ (in which case it reaches the fixed point $\overline{0}$) or $x_0\overline{x_1x_2\cdots x_{2^n}}$ (in which case it reaches the fixed point $1\overline{0}$).

Example 4.9. By Theorem 4.8, $\overline{1011}$ is an eventually fixed point of $D$, as we easily verify: the orbit of $\overline{1011}$ under $D$ is $\overline{1011, 1100, 01, 1, 0, 0, \ldots}$. Similarly, the $D$-orbit of $\overline{1001}$ is $\overline{1001, 10, 0, 0, 0, \ldots}$. Note that $\overline{1011}$ and $\overline{1001}$ reach a fixed point after a different number of iterations; Theorem 4.8 only provides an upper bound for the number of iterations of $D$ it takes to reach a fixed point.

Since the parity vector of a point that eventually reaches the fixed point $\overline{0}$ under iteration of $D$ is of the form $x_0x_1\cdots x_{k-1}0$ for some $k \in \mathbb{Z}^+$, and the parity vector of some periodic point of $D$ with period $2^n$ for some $n \in \mathbb{N}$ is
of the form $x_0x_1\cdots x_{2^n-1}$, the following theorem and Theorem 4.8 illustrate the symmetry that often emerges in the dynamics of $D$.

**Theorem 4.10** (Monks [Mon08]). Let $x \in \mathbb{N}$. Then $x$ is a periodic point of $D$ with period $2^n$, where $n$ is the smallest nonnegative integer such that $2^n \geq x$.

**Remark.** Theorem 4.10 relies on the fact that every $x \in \mathbb{N}$ has an eventually repeating 2-adic representation, which we know to be true from Chapter 1.

**Example 4.11.** Since $\overline{1010} = 3$, by Theorem 4.10, $101\overline{0}$ is a periodic point of $D$ with period $2^2 = 4 > 3$. Indeed, the orbit of $101\overline{0}$ under $D$ is easily seen to be $101\overline{0}, 111\overline{0}, 001\overline{0}, 011\overline{0}, 101\overline{0}, \ldots$. Note that $\mathcal{P}(101\overline{0}) = \overline{1100}$, which by Theorem 4.8 is an eventually fixed point of $D$. In general, every eventually fixed point of $D$ is related to a unique period point of $D$ of period $2^k$ for some $k \in \mathbb{N}$.

As Example 4.9 alluded to, although Theorem 4.8 classifies all of the eventually fixed points of $D$, it does not specify the exact number of iterations it takes for a given point to reach either $\overline{0}$ or $1\overline{0}$, the fixed points of $D$. To understand the parity vectors of $D$ better, it is important that we understand exactly how many iterations of $D$ are necessary before a fixed point is reached. The following theorem shows an instance where determining the number of iterations is simple.

**Theorem 4.12.** Let $x \in \mathbb{Z}_2$ and $n \in \mathbb{N}$. Then $x$ has reduced form $\overline{x_0x_1\cdots x_{2^n-1}}$ or $x_0x_1x_2\cdots x_{2^n}$ and has an odd number of 1s in its repeating part if and only if the orbit of $x$ under $D$ reaches a fixed point after exactly $2^n$ iterations.
To prove Theorem 4.12, we will need a few lemmas. We first construct a mapping \( d \) that operates on finite blocks in \( B_n \) for some \( n \in \mathbb{Z}^+ \) (as opposed to \( D \), whose domain consists of 2-adic integers, which are infinite sequences).

**Definition.** For any \( n \in \{2, 3, 4, \ldots\} \) and \( x \in B_n \), define \( d : B_n \to B_{n-1} \) by 
\[
d(x_0 x_1 \cdots x_{n-1}) = y_0 y_1 \cdots y_{n-2},
\]
where for every \( i \in \{0, 1, \ldots, n-2\} \), 
\[
y_i = |x_i - x_{i+1}|.
\]

The following lemma by Monks shows that \( d \) and \( D \) have a simple, albeit important, relationship.

**Lemma 4.13 (Monks [Mon08]).** Let \( x \in \mathbb{Z}_2, n \in \mathbb{Z}^+, \) and \( y = D^n(x) \). For every \( i \in \mathbb{N} \), 
\[
y_i = d^n(x_i x_{i+1} \cdots x_{i+n}).
\]

Monks used Lemma 4.13 to prove the following important result concerning blocks that satisfy certain conditions.

**Lemma 4.14 (Monks [Mon08]).** Let \( n \in \mathbb{N} \) and \( a \in B_{2^n} \). Then 
\[
d^{2^n - 1}(a) = \begin{cases} 
0 & \text{if } a \text{ contains an even number of 1s} \\
1 & \text{otherwise}.
\end{cases}
\]

We can now prove Theorem 4.12.

**Proof of Theorem 4.12.** Let \( x \in \mathbb{Z}_2 \) and \( n \in \mathbb{N} \). First, assume that \( x \) has reduced form \( x_0 x_1 \cdots x_{2^n - 1} \) and has an odd number of 1s in its repeating part. By Lemma 4.14, 
\[
d^{2^n - 1}(x_i x_{i+1} \cdots x_{i+2^n - 1}) = 1
\]
for every \( i \in \mathbb{N} \). It follows from Lemma 4.13 that \( D^{2^n - 1}(x) = \overline{1} \), and so \( D^{2^n}(x) = \overline{0} \). Therefore, the orbit
of $x$ under $D$ reaches the fixed point $\bar{0}$ after exactly $2^n$ iterations. A similar argument shows that when $x$ has reduced form $x_0x_1x_2\cdots x_{2^n}$ and has an odd number of 1s in its repeating part, the orbit of $x$ under $D$ reaches the fixed point $1\bar{0}$ after exactly $2^n$ iterations.

Conversely, assume that the orbit of $x$ under $D$ reaches the fixed point $\bar{0}$ after exactly $2^n$ iterations. By Lemma 3.12, there are precisely $2^{2n-1}$ points whose orbits under $\sigma$ reach the fixed point $\bar{0}$ after exactly $2^n$ iterations. Since $D$ is conjugate to $\sigma$ by $P$ and $P$ maps $\bar{0}$ to $\bar{0}$, there are precisely $2^{2n-1}$ points whose orbits under $D$ reach the fixed point $\bar{0}$ after exactly $2^n$ iterations. However, there are exactly $2^{2n-1}$ points with reduced form $\bar{x_0x_1\cdots x_{2^n-1}}$ and an odd number of 1s under their repeating part. Therefore, we have enumerated every point whose orbit under $D$ reaches the fixed point $\bar{0}$ after $2^n$ iterations, and so $x$ must have reduced form $\bar{x_0x_1\cdots x_{2^n-1}}$ and an odd number of 1s under its repeating part. A similar argument shows that if $x$ reaches $1\bar{0}$ after exactly $2^n$ iterations, it must have reduced form $x_0\bar{x_1x_2\cdots x_{2^n}}$ and an odd number of 1s under its repeating part.

Therefore, $x$ has reduced form $\bar{x_0x_1\cdots x_{2^n-1}}$ or $x_0\bar{x_1x_2\cdots x_{2^n}}$ and has an odd number of 1s in its repeating part if and only if the orbit of $x$ under $D$ reaches a fixed point after exactly $2^n$ iterations. \hfill $\square$

**Example 4.15.** As in Example 4.9, we consider the points $\overline{1001}$ and $\overline{1011}$. Since $\overline{1011}$ has an odd number of 1s in its repeating part, we can immediately apply Theorem 4.12 to conclude that the orbit of $\overline{1011}$ under $D$ reaches a fixed point after exactly 4 iterations. Therefore, without any computation, we know the parity vector of $\overline{1011}$ under $D$ must have reduced form
Chapter 4: The Dynamics of $D$ Revisited

$x_0x_1x_2x_3\bar{0}$. Note that with regards to $100\bar{1}$, Theorem 4.12 only lets us conclude that the orbit of $100\bar{1}$ under $D$ reaches $\bar{0}$ in strictly less than 4 iterations, as $100\bar{1}$ has an even number of 1s in its repeating part.

3. Eventually Periodic Points of $D$ with Period $2^n$

In this section, we extend Theorem 4.10 from all periodic points of $D$ with period $2^n$ for some $n \in \mathbb{N}$ to all eventually periodic points with period $2^n$. Similar to Theorem 4.12, we will also provide conditions where we can easily determine the parity vector of such a point under $D$. We first state an important result concerning the stem of an eventually repeating 2-adic integer.

**Lemma 4.16 (Monks [Mon08]).** Let $x \in \mathbb{Z}_2$ such that $x$ is eventually repeating. Then for every $k \in \mathbb{N}$, $D^k(x) = x$.

**Remark.** Informally, we say the length of the finite stem of $x$ is preserved under iteration on $D$.

Using Lemma 4.16, we can now state our desired classification.

**Theorem 4.17.** Let $x \in \mathbb{Z}_2$ such that for some $k, l \in \mathbb{Z}^+$, $x$ has reduced form $x_0x_1 \cdots x_{k-1}x_kx_{k+1} \cdots x_{k+2l-1}$. Then $x$ is an eventually periodic point of $D$ of period $2^n$ after exactly $m$ iterations, where $n$ is the smallest nonnegative integer such that $2^n \geq k$ and $m$ is the smallest integer such that $D^m(x_kx_{k+1} \cdots x_{k+2l-1}) = \bar{0}$.

**Proof.** Let $x \in \mathbb{Z}_2$ such that for some $k, l \in \mathbb{Z}^+$, $x$ has reduced form $x_0x_1 \cdots x_{k-1}x_kx_{k+1} \cdots x_{k+2l-1}$. Furthermore, let $n$ be the smallest nonnegative integer such that $2^n \geq k$ and $m$ be the smallest integer such that
Chapter 4: The Dynamics of D Revisited

\[ D^m(x_kx_{k+1}\cdots x_{k+2^l-1}) = \bar{0} \] (such an integer \( m \) is guaranteed to exist by Theorem 4.8). Then \( D^m(x) = D^m(x_0x_1\cdots x_{k-1}x_kx_{k+1}\cdots x_{k+2^l-1}) = y_0y_1\cdots y_{k-1}\bar{0} \). Note that \( y_0y_1\cdots y_{k-1}\bar{0} \) must be in reduced form since by Lemma 4.16, \( x = k = y_0y_1\cdots y_{k-1}\bar{0} \). Since Theorem 4.10 shows that \( y_0y_1\cdots y_{k-1}\bar{0} \) is a periodic point of period \( 2^n \), \( x \) is an eventually periodic point of \( D \) of period \( 2^n \) after exactly \( m \) iterations. \( \square \)

**Remark.** Recall that the dynamics of \( x \) when it is repeating (as opposed to eventually repeating) were previously classified by Theorem 4.8.

Note that in our proof of Theorem 4.17, we used Theorem 4.8 to show that there existed a smallest integer \( m \) such that \( D^m(x_kx_{k+1}\cdots x_{k+2^l-1}) = 0 \). However, Theorem 4.8 only provides an upper bound for \( m \), namely \( 2^l \), and so in general, we cannot use it to determine exactly how many iterations it takes the \( D \)-orbit of \( x \) to reach a periodic point.

**Example 4.18.** Since \( \|10101\| = 4 \) and \( 10101 = 2 \), Theorem 4.17 lets us conclude that the \( D \)-orbit of 10101 enters a 2-cycle after at most 4 iterations. Computing the orbit of 10101 under \( D \) confirms this; the orbit is 10101, 111100, 0001, 010, 110, 010, \ldots.

Note that in Example 4.18, Theorem 4.8’s bound on the number of iterations it took the orbit of 10101 under \( D \) to reach a periodic point was tight. The following corollary, which we prove using Theorem 4.12 instead of Theorem 4.8, allows us to state in general when this occurs.

**Corollary 4.19.** Let \( x \in \mathbb{Z}_2 \) such that for some \( k, l \in \mathbb{Z}^+ \), \( x \) has reduced form \( x_0x_1\cdots x_{2^k-1}x_{2^k+1}\cdots x_{2^k+2^l-1} \) and the repeating part of \( x \) contains an odd
number of 1s. Then \( x \) is an eventually periodic point of \( D \) of period \( 2^k \) after exactly \( 2^l \) iterations.

**Proof.** Let \( x \in \mathbb{Z}_2 \) such that for some \( k, l \in \mathbb{Z}^+ \), \( x \) has reduced form \( x_0x_1 \cdots x_{2^k-1}x_{2^k}x_{2^k+1} \cdots x_{2^k+2^l-1} \). By Theorem 4.12, the smallest integer \( n \) such that \( D^n(x_{2^k}x_{2^k+1} \cdots x_{2^k+2^l-1}) = 0 \) is \( 2^l \). Also note that \( x = 2^k \). By Theorem 4.17, the orbit of \( x \) under \( D \) reaches a periodic point of period \( 2^k \) after exactly \( 2^l \) iterations. □

**Remark.** Recall that the dynamics of \( x \) when it is repeating (as opposed to eventually repeating) and has an odd number of 1s in its repeating part were previously classified by Theorem 4.12.

**Example 4.20.** Since 101011 has an odd number of 1s in its repeating part, Corollary 4.19 shows us that 101011 enters a 2-cycle after exactly 4 iterations. Therefore, we can immediately conclude that 101011’s parity vector under \( D \) has reduced form \( x_0x_1x_2x_3x_4x_5 \). Interestingly enough, we can apply Corollary 4.17 to this parity vector to conclude that it is itself an eventually periodic point of period 4 after exactly 2 iterations. In general, we could not determine the exact number of iterations of \( D \) that would have been necessary to reach a periodic point, but in this case, if \( x_0x_1x_2x_3x_4x_5 \) did not have an odd number of 1s in its repeating part, it would not be in reduced form.

### 4. Periodic Points of \( D \) with Period \( 2^n - 1 \)

In this section, we attempt to classify the periodic points of period \( 2^n - 1 \) for some \( n \in \mathbb{Z}^+ \), which for reasons we will elaborate upon proves to be much more challenging than classifying the ones of period \( 2^n \).
The following theorem classifies some of the periodic points of $D$ with (not necessarily minimum) period $2^n - 1$. In other words, these points could have a smaller period that is some factor of $2^n - 1$.

**Theorem 4.21.** Let $x \in B_{2^n - 1}$ for some $n \in \mathbb{Z}^+$ such that $x$ contains an even number of 1s in its repeating part. Then $x_0 x_1 \cdots x_{2^n - 2}$ is a periodic point of $D$ with (not necessarily minimum) period $2^n - 1$.

**Proof.** Let $x \in B_{2^n - 1}$ for some $n \in \mathbb{Z}^+$ such that $x$ contains an even number of 1s in its repeating part. Furthermore, let $y = x_0 x_1 \cdots x_{2^n - 2}$. For every $i \in \mathbb{N}$, $y_i = y_{i + 2^n - 1}$, and so by Lemma 4.14, $d^{2^n - 1}(y_i y_{i+1} \cdots y_{i+2^n-1}) = 0$ if and only if $y_1 = 0$. It follows from Lemma 4.13 that $D^{2^n - 1}(y) = y$. □

**Example 4.22.** Since $110$ has an even number of 1s in its repeating part, Theorem 4.21 lets us conclude that $110$ is a periodic point of $D$ with (not necessarily minimum) period 3. However, the only divisors of 3 are 1 and 3, and since we know by Theorem 4.8 that $110$ is not an eventually fixed point of $D$, $110$ must have period 3. The orbit of $110$ under $D$ is $110, 011, 101, 110, \ldots$, which verifies our conclusion. We will generalize the reasoning of this example in Corollary 4.27.

To apply Theorem 4.21 to eventually periodic points with period $2^n - 1$, we will need the following theorem.

**Theorem 4.23 (Monks [Mon08]).** For every $x, y \in \mathbb{Z}_2$ such that $D(x) = y$, $x$ and $V(x)$ are the only preimages of $y$ under $D$.

**Example 4.24.** Since $D(1\overline{0}) = 1\overline{0}$, Theorem 4.23 lets us conclude that the only preimages of $1\overline{0}$ under $D$ are $1\overline{0}$ and $V(1\overline{0}) = 0\overline{1}$. 
We may now use Theorems 4.21 and 4.23 to classify some of the eventually periodic points of \( D \) with (not necessarily minimum) period \( 2^n - 1 \).

**Corollary 4.25.** Let \( x \in B_{2^n - 1} \) for some \( n \in \mathbb{Z}^+ \) such that \( x \) contains an odd number of 1s in its repeating part. Then \( x_0 x_1 \cdots x_{2^n - 2} \) is an eventually periodic point of \( D \) of (not necessarily minimum) period \( 2^n - 1 \) after exactly 1 iteration.

**Proof.** Let \( x \in B_{2^n - 1} \) for some \( n \in \mathbb{Z}^+ \) such that \( x \) contains an odd number of 1s in its repeating part. Furthermore, let \( y = x_0 x_1 \cdots x_{2^n - 2} \). Note that \( y \) is not a periodic point of \( D \) of period \( 2^n - 1 \) because by reasoning similar to the proof of Theorem 4.21, \( d^{2^n - 1}(y_1 y_{i+1} \cdots y_{i+2^n - 1}) = 0 \) if and only if \( y_i = 1 \). Therefore, by Lemma 4.13, \( D^{2^n - 1}(y) = V(y) \neq y \).

We need to show that although \( y \) is not a periodic point of \( D \) of period \( 2^n - 1 \), it is nevertheless an eventually periodic point. Let \( z = D(y) \). By Theorem 4.23, the preimages of \( z \) under \( D \) are precisely \( y \) and \( V(y) \). Note that \( V(y) = v_0 v_1 \cdots v_{2^n - 2} \), where \( v \in B_{2^n - 1} \) such that for every \( i \in \{0, 1, \ldots, 2^n - 2\} \), \( v_i = 1 - x_i \). Also observe that \( v \) has an even number of 1s in its repeating part. Therefore, by Theorem 4.21, \( V(y) \), and thus \( z \), are periodic points of \( D \) with period \( 2^n - 1 \), and so \( y \) is an eventually periodic point of \( D \) of period \( 2^n - 1 \) after exactly 1 iteration. \( \square \)

**Example 4.26.** Since \( 010 \) has an odd number of 1s in its repeating part, we can use Corollary 4.25 to conclude that \( 010 \) is an eventually periodic point of \( D \) with (not necessarily minimum) period 3 after exactly 1 iteration. Calculating the orbit of \( 010 \) under \( D \) shows us that \( 010 \) in fact enters the same \( D \)-cycle as the one in Example 4.22: \( 010, 110, 011, 101, 110, \ldots \).
The next two corollaries address the special case where $2^n - 1$ is a Mersenne prime, in which case we can conclude that $2^n - 1$ is in fact the minimum period. Furthermore, these corollaries provide a complete classification of all such periodic points.

**Corollary 4.27.** Let $p$ be a Mersenne prime of the form $2^n - 1$ for some $n \in \mathbb{Z}^+$. Then the periodic points of $D$ of period $p$ are precisely those $x \in \mathbb{Z}_2$ with reduced form $\overline{x_0 x_1 \cdots x_{2^n-2}}$ or $\overline{x_0 x_1 x_2 \cdots x_{2^n-1}}$ and repeating part consisting of an even number of 1s.

**Proof.** Let $p$ be a Mersenne prime of the form $2^n - 1$ for some $n \in \mathbb{Z}^+$. Note that there are precisely $2^n - 2$ periodic points of $\sigma$ of period $2^n - 1$ since the only two repeating 2-adic integers $s$ of the form $\overline{s_0 s_1 \cdots s_{2^n-2}}$ that are not in reduced form are $\overline{0}$ and $\overline{1}$. Since $D$ is conjugate to $\sigma$, if we can find $2^n - 2$ periodic points of $D$ of period $2^n - 1$, we will have enumerated all such periodic points.

Let $x \in \mathbb{Z}_2$ with reduced form $\overline{x_0 x_1 \cdots x_{2^n-2}}$ and repeating part consisting of an even number of 1s. It follows from Theorem 4.21 that $x$ is a periodic point of $D$ of period $2^n - 1$. Further, this period must be the minimum period since $2^n - 1$ is prime and $x$ is neither of the two fixed points of $D$. Since there are $\frac{2^n - 2}{2}$ such points, we have enumerated exactly half of the points we need.

Let $y \in \mathbb{Z}_2$ with reduced form $\overline{y_0 y_1 y_2 \cdots y_{2^n-1}}$ and repeating part consisting of an even number of 1s. It follows from Theorem 4.21 that $\overline{y_1 y_2 \cdots y_{2^n-1}}$ is a periodic point of $D$ of period $2^n - 1$. Therefore, $D^{2^n-1}(y) = \overline{z_0 y_1 y_2 \cdots y_{2^n-1}}$. By Lemma 4.16, $z_0 = y_0$, and so $D^{2^n-1}(y) = y$. Thus, $y$ is
a periodic point of \( D \) of period \( 2^n - 1 \). As before, this period must be the minimum period since \( 2^n - 1 \) is prime and \( x \) is not a fixed point of \( D \). Since there are \( \frac{2^n - 2}{2} \) such points, we have enumerated all of the periodic points of \( D \) of period \( p \). \( \square \)

**Corollary 4.28.** Let \( p \) be a Mersenne prime of the form \( 2^n - 1 \) for some \( n \in \mathbb{Z}^+ \). Then the eventually periodic points of \( D \) of period \( p \) after exactly 1 iteration are precisely those \( x \in \mathbb{Z}_2 \) with reduced form \( x_0x_1\cdots x_{2^n-2} \) or \( x_0x_1x_2\cdots x_{2^n} \) and repeating part containing an odd number of 1s.

**Proof.** Let \( p \) be a Mersenne prime of the form \( 2^n - 1 \) for some \( n \in \mathbb{Z}^+ \). The corollary follows directly from adapting the logic of Corollaries 4.25 and 4.27 to the fact that there are precisely \( 2^n - 2 \) eventually periodic points of \( \sigma \) of period \( 2^n - 1 \) after exactly 1 iteration. \( \square \)

In order to fully classify the eventually periodic points of \( D \) with period \( 2^n - 1 \) when \( 2^n - 1 \) is not necessarily prime, it will be necessary to address some of the challenges we encountered in this section. For instance, Theorem 4.21 only provides an upper bound for the minimum period of 2-adic integers \( x \) with the form \( x_0x_1\cdots x_{2^n-2} \). To illustrate, \( \|100011110101100\| = 2^4 - 1 = 15 \), but \( 100011110101100 \) is a periodic point of \( D \) with period 5.

In addition, we need some mechanism for classifying eventually periodic points of \( D \) with period \( 2^n - 1 \) after more than 1 iteration. Theorem 4.17 follows nicely from the fact that by Theorem 4.8, \( x_0x_1\cdots x_k \) has easily understood behavior provided that \( k = 2^m - 1 \) for some \( m \in \mathbb{N} \). On the other hand, when \( k \neq 2^m - 1 \) for any \( m \in \mathbb{N} \), the behavior of \( x_0x_1\cdots x_k \)
is much harder to deduce. For instance, every periodic point of $D$ with reduced form $x_0x_1\cdots x_5$ has period $6 = 2 \cdot 3$, but every periodic point of $D$ with reduced form $x_0x_1\cdots x_{13}$ has period $819$. Interestingly enough, $819 = 63 \cdot 13$, and $63$ is a Mersenne number, leading us to conjecture that some peculiar relationship between Mersenne numbers and the periodic points of $D$ exists.

**Conjecture 4.29.** Let $n \in \mathbb{Z}^+$ and $x \in \mathbb{Z}_2$ such that $x = x_0x_1\cdots x_{n-1}$. If $x$ is a periodic point of $D$, then the period of $x$ is a factor of $2^d \cdot m$, where $d$ is the largest nonnegative integer such that $2^d$ divides $n$ and $m$ is the smallest Mersenne number that has $n/2^d$ as a factor.

**Remark.** Note that for every $n \in \mathbb{Z}^+$ such that $n$ is odd, there exists a Mersenne number $m$ such that $n$ divides $m$. By Euler’s Theorem, $n$ always divides $m = 2^{\phi(n)} - 1$, where $\phi(n)$ denotes the number of positive integers less than and relatively prime to $n$.

We leave this fascinating question open for future research.
CHAPTER 5

Conclusions and Future Research

Throughout the course of our research, we have derived many new results about the Collatz Conjecture, but much more work will be necessary to definitively prove or disprove this tenacious problem. We summarize our results and open questions here.

In Chapter 2, we generalized Bernstein’s $\Phi$ mapping to construct $\Phi_{a,b,c,d}$, a non-iterative inverse of $f_{a,b,c,d}$’s parity vector function, $Q_{a,b,c,d}$. Using $\Phi_{a,b,c,d}$ and Fraboni’s results in [Fra97], we were able to develop an expression for computing arbitrary $f_{a,b,c,d}$-cycles, which enabled us to conclusively rule out cycles that could have been counterexamples for the Non-trivial Cycles Conjecture. As mentioned earlier, Theorem 2.20 currently accounts for most, but not all, $T$-cycles of the form $1\cdots10\cdots0$. Resolving this issue would be of use towards proving the Non-trivial Cycles Conjecture.

In Chapter 3, we continued Monks’ work in [Mon08] by searching for new conjugacies of $T$ among the continuous endomorphisms of $\sigma$, the shift map. We ultimately concluded that although there are continuous endomorphisms that are conjugate to $\sigma$ by a mapping other than its parity vector, finding them would prove difficult. Fortunately, some of the techniques that we applied to search for new conjugacies turned out to be helpful in our later analysis on the dynamics of $D$ in Chapter 4.
Lastly, in Chapter 4, in an attempt to prove Monks’ reformulation of the Collatz Conjecture, we proved useful results concerning the dynamics of $D$, and thus $P$, its parity vector function. In particular, we used Monks’ characterization of the eventually fixed points of $D$ and the periodic points of $D$ with period $2^k$ to derive a complete classification of all of the eventually periodic points of $D$ with period $2^k$. In addition, we also discovered special cases where the form of certain parity vectors under $D$ can be derived without any computation. The dynamics of $D$, however, have yet to be fully classified. Most notably, our attempts to classify the periodic points of period $2^k - 1$ have yielded some peculiar connections to number theory, and our intuition suggests that this is a rich area of future research.
Acknowledgments

The author would like to thank Dr. Alicia Sevilla for her assistance in the formulation of Conjecture 4.29.
Bibliography


