THE $\langle a, b\rangle$ SHADOW FUNCTIONS

ANDREW WOOD

1. Introduction

We introduce a class of dynamical systems, called $\langle a, b\rangle$ shadow functions, and develop a symbolic framework for analysing their behaviour.

Definition 1.1. Given positive integers $a < b$, the $\langle a, b\rangle$ *shadow function* $g\colon \mathbb{Z} \to \mathbb{Z}$ is defined by

$$n \mapsto \begin{cases} \lceil \frac{n}{a} \rceil & \text{if } a \text{ does not divide } n; \\ \frac{b}{a} n & \text{if } a \text{ divides } n. \end{cases}$$

2. Basic definitions

In this section we introduce some concepts that will allow us to work more easily with $\langle a, b\rangle$ shadow functions.

Definition 2.1. Let $a < b$ be positive integers and let g be the $\langle a, b\rangle$ shadow function. A non-empty sequence of integers $\boldsymbol{n} = \langle n_i \mid i \le m\rangle$ is a *chain* of g (of length m) if for all $i < m$,

$$n_{i+1} = g(n_i).$$

$\boldsymbol{n}$ is a *cycle* of g (of length m) if, in addition, $n_0 = n_m$.

Definition 2.2. A finite sequence of 0s and 1s is an *itinerary*.

Let $\boldsymbol{n} = \langle n_i \mid i \le m\rangle$ be a chain of the $\langle a, b\rangle$ shadow function. $\boldsymbol{\mu} = \langle \mu_i \mid i < m\rangle$ is the itinerary of $\boldsymbol{n}$ if for all $i < m$

$$\mu_i = \begin{cases} 0 & \text{if } a \text{ does not divide } n_i; \\ 1 & \text{if } a \text{ divides } n_i. \end{cases}$$

Definition 2.3. Let $\boldsymbol{\mu} = \langle \mu_i \mid i < m\rangle$ be an itinerary and let $\boldsymbol{k} = \langle k_i \mid i \le m\rangle$ be the sequence of integers such that for all $i \le m$,

$$k_i = \sum_{j<i} \mu_j.$$

Note that for all $i \le m$, $0 \le k_i \le i$.

We say $\boldsymbol{k}$ is the *itinerary counter* of $\boldsymbol{\mu}$ and k_m is the *itinerary count* of $\boldsymbol{\mu}$.

If $\boldsymbol{\mu}$ is the itinerary of a chain $\boldsymbol{n}$ of the $\langle a, b\rangle$ shadow function, we say $\boldsymbol{k}$ is the itinerary counter of $\boldsymbol{n}$ and k_m is the itinerary count of $\boldsymbol{n}$.

Definition 2.4. Given an itinerary $\boldsymbol{\mu}$ with itinerary count k, the *itinerary indexer* of $\boldsymbol{\mu}$ is the strictly increasing sequence of natural numbers $\boldsymbol{L} = \langle L_j \mid j < k\rangle$ such that for all $j < k$,

$$\mu_{L_j} = 1.$$

Note that $(j \mapsto L_j)$ is a bijection.

If $\boldsymbol{\mu}$ is the itinerary of a chain $\boldsymbol{n}$ of the $\langle a, b\rangle$ shadow function, we say $\boldsymbol{L}$ is the itinerary indexer of $\boldsymbol{n}$.

With thanks to Gareth.

Lemma 2.5. *Let $\boldsymbol{\mu}$ be an itinerary of length m. Let $\boldsymbol{k}$ be the itinerary counter of $\boldsymbol{\mu}$ and let $\boldsymbol{L}$ be the itinerary indexer of $\boldsymbol{\mu}$. Then, for all $i < m$ and all $j < k_m$, $L_j < i$ if and only if $j < k_i$.*

Proof. Let $i < m$ and $j < k_m$ be given.

First Suppose that $L_j < i$. Because $\boldsymbol{L}$ is increasing, $L_t < i$ for all $t \leq j$. Because $(j \mapsto L_j)$ is a bijection, $\mu_s = 1$ for at least $j+1$ distinct values of $s < i$. Then

$$k_i = \sum_{s<i} \mu_s \geq j+1$$

and therefore $j < k_i$.

Now suppose that $L_j \geq i$. Because $\boldsymbol{L}$ is increasing, $L_t < i$ for at most j distinct values of t. Because $(j \mapsto L_j)$ is a bijection, $\mu_s = 1$ for at most j distinct values of $s < i$. Then

$$k_i = \sum_{s<i} \mu_s \leq j,$$

that is to say: $j \geq k_i$. □

Definition 2.6. Let $\boldsymbol{n}$ be a chain of the $\langle a, b\rangle$ shadow function, of length m. The *blueprint* of $\boldsymbol{n}$ is $\boldsymbol{R} = \langle R_i \mid i < m\rangle$ where, for all $i < m$,

$$R_i = a\left\lceil \frac{n_i}{a} \right\rceil - n_i.$$

Lemma 2.7. *Let $\boldsymbol{n}$ be a chain of the $\langle a, b\rangle$ shadow function, of length m, and let $\boldsymbol{R}$ be the blueprint of $\boldsymbol{n}$. Then, for all $i < m$, R_i is an integer, $0 \leq R_i < a$, and $R_i = 0$ if and only if $\mu_i = 1$.*

Proof. Fix $i < m$. From the definition, it is clear that R_i is an integer. Also,

$$\begin{aligned}
& \frac{n_i}{a} \leq \left\lceil \frac{n_i}{a} \right\rceil < \frac{n_i}{a} + 1 \\
\Longrightarrow \quad & n_i \leq a\left\lceil \frac{n_i}{a} \right\rceil < n_i + a \\
\Longrightarrow \quad & 0 \leq a\left\lceil \frac{n_i}{a} \right\rceil - n_i < a \\
\Longrightarrow \quad & 0 \leq R_i < a
\end{aligned}$$

and

$$\begin{aligned}
& R_i = 0 \\
\Longleftrightarrow \quad & a\left\lceil \frac{n_i}{a} \right\rceil - n_i = 0 \\
\Longleftrightarrow \quad & \left\lceil \frac{n_i}{a} \right\rceil = \frac{n_i}{a} \\
\Longleftrightarrow \quad & a \text{ divides } n_i \\
\Longleftrightarrow \quad & \mu_i = 1.
\end{aligned}$$

□

3. Formulae for chains and cycles

In this section we will establish a number of precise formulae relating the inital and final values of chains and cycles.

Proposition 3.1. *Let $\boldsymbol{n}$ be a chain of the $\langle a, b\rangle$ shadow function g, of length m. Let $\boldsymbol{k}$ be the itinerary counter of $\boldsymbol{n}$ and let $\boldsymbol{R}$ be the blueprint of $\boldsymbol{n}$. Then*

$$n_m = \frac{b^{k_m}}{a^m}\left(n_0 + \sum_{i<m} \frac{a^i}{b^{k_i}} R_i\right).$$

Proof. Fix positive integers $a < b$ and let g be the $\langle a, b \rangle$ shadow function. For each $m \in \mathbb{N}$, let $\phi(m)$ be the statement:

For all chains $\boldsymbol{n}$ of g of length m,

$$n_m = \frac{b^{k_m}}{a^m} \left(n_0 + \sum_{i<m} \frac{a^i}{b^{k_i}} R_i \right) \tag{1}$$

where $\boldsymbol{k}$ is the itinerary counter of $\boldsymbol{n}$ and $\boldsymbol{R}$ is the blueprint of $\boldsymbol{n}$.

We will prove that $\phi(m)$ is true for all $m \in \mathbb{N}$ by induction.

First, let $m = 0$ and let $\boldsymbol{n} = \langle n_0 \rangle$ be a chain of g of length 0. Then the itinerary counter of $\boldsymbol{n}$ is $\boldsymbol{k} = \langle k_0 \rangle = \langle 0 \rangle$ and the blueprint of $\boldsymbol{n}$ is $\boldsymbol{R} = \langle \rangle$. In this case, equation 1 becomes

$$\begin{aligned} n_0 &= \frac{b^0}{a^0} \left(n_0 + \sum_{i<0} \frac{a^i}{b^{k_i}} R_i \right) \\ &= (n_0 + 0) \\ &= n_0. \end{aligned}$$

Thus, $\phi(0)$ is true.

Now, suppose $\phi(m)$ is true for some $m \in \mathbb{N}$. Let $\boldsymbol{n}$ be a chain of g of length $m+1$. Let $\boldsymbol{\mu}$ be the itinerary of $\boldsymbol{n}$, let $\boldsymbol{k}$ be the itinerary counter of $\boldsymbol{n}$, and let $\boldsymbol{R}$ be the blueprint of $\boldsymbol{n}$. To show that $\phi(m+1)$ is true, we need to prove that

$$n_{m+1} = \frac{b^{k_{m+1}}}{a^{m+1}} \left(n_0 + \sum_{i<m+1} \frac{a^i}{b^{k_i}} R_i \right). \tag{2}$$

Observe that $\langle n_i \mid i \leq m \rangle$ is a chain of g of length m and that:

- $\langle \mu_i \mid i < m \rangle$ is the itinerary of $\langle n_i \mid i \leq m \rangle$;
- $\langle k_i \mid i \leq m \rangle$ is the itinerary counter of $\langle n_i \mid i \leq m \rangle$;
- $\langle R_i \mid i < m \rangle$ is the blueprint of $\langle n_i \mid i \leq m \rangle$.

Because $\phi(m)$ is true, we have equation 1.

Now, recall from the definition of $\boldsymbol{R}$ that,

$$\begin{aligned} & & R_m &= a \left\lceil \frac{n_m}{a} \right\rceil - n_m \\ \iff & & n_m + R_m &= a \left\lceil \frac{n_m}{a} \right\rceil \end{aligned}$$

When $\mu_m = 0$,

$$g(n_m) = \left\lceil \frac{n_m}{a} \right\rceil = \frac{1}{a}(n_m + R_m) = \frac{b^{\mu_m}}{a}(n_m + R_m).$$

When $\mu_m = 1$, a divides n_m so

$$g(n_m) = \frac{b}{a} n_m = \frac{b}{a} \left(a \left\lceil \frac{n_m}{a} \right\rceil \right) = \frac{b}{a}(n_m + R_m) = \frac{b^{\mu_m}}{a}(n_m + R_m).$$

In either case,

$$\begin{aligned}
n_{m+1} &= g(n_m) \\
&= \frac{b^{\mu_m}}{a}(n_m + R_m) \\
&= \frac{b^{\mu_m}}{a}\left(\frac{b^{k_m}}{a^m}\left(n_0 + \sum_{i<m}\frac{a^i}{b^{k_i}}R_i\right) + R_m\right) \\
&= \frac{b^{\mu_m}}{a}\left(\frac{b^{k_m}}{a^m}\left(n_0 + \sum_{i<m}\frac{a^i}{b^{k_i}}R_i + \frac{a^m}{b^{k_m}}R_m\right)\right) \\
&= \frac{b^{\mu_m}}{a}\left(\frac{b^{k_m}}{a^m}\left(n_0 + \sum_{i<m+1}\frac{a^i}{b^{k_i}}R_i\right)\right) \\
&= \frac{b^{k_m+\mu_m}}{a^{m+1}}\left(n_0 + \sum_{i<m+1}\frac{a^i}{b^{k_i}}R_i\right).
\end{aligned}$$

Finally, we can obtain equation 2 by observing that,

$$k_m + \mu_m = \left(\sum_{i<m}\mu_i\right) + \mu_m = \sum_{i<m+1}\mu_i = k_{m+1}.$$

Thus, $\phi(m+1)$ is true.

By induction, $\phi(m)$ is true for all $m \in \mathbb{N}$. □

Lemma 3.2. *Let* $\boldsymbol{n}$ *be a cycle of the* $\langle a, b\rangle$ *shadow function* g*, of length* $m > 0$*. Then*

$$a^m \neq b^k$$

where k *is the itinerary count of* $\boldsymbol{n}$*.*

Proof. First, note that because $a \neq b$ and $m \neq 0$, we must have $a^m \neq b^m$. Therefore, without loss of generality we may assume that $k < m$.

Let $\boldsymbol{k}$ be the itinerary counter of $\boldsymbol{n}$ and let $\boldsymbol{R}$ be the blueprint of $\boldsymbol{n}$. Proposition 3.1 tells us that

$$\begin{aligned}
n_m &= \frac{b^k}{a^m}\left(n_0 + \sum_{i<m}\frac{a^i}{b^{k_i}}R_i\right) \\
\implies \qquad a^m n_m &= b^k n_0 + \sum_{i<m} a^i b^{k-k_i} R_i.
\end{aligned}$$

Because $\boldsymbol{n}$ is a cycle, $n_0 = n_m$ and so

$$\left(a^m - b^k\right)n_0 = \sum_{i<m} a^i b^{k-k_i} R_i.$$

Since $k < m$, there must exist $s < m$ such that $\mu_s = 0$. From Lemma 2.7, we must have $R_s > 0$. Then,

$$\begin{aligned}
& a^s b^{k-k_s} R_s > 0 \\
\implies \quad & \sum_{i<m} a^i b^{k-k_i} R_i > 0 \\
\implies \quad & \left(a^m - b^k\right)n_0 > 0 \\
\implies \quad & a^m - b^k \neq 0 \\
\implies \quad & a^m \neq b^k.
\end{aligned}$$

□

Corollary 3.3. *Let $\boldsymbol{n}$ be a cycle of the $\langle a, b\rangle$ shadow function g, of length $m > 0$. Let $\boldsymbol{k}$ be the itinerary counter of $\boldsymbol{n}$ and let $\boldsymbol{R}$ be the blueprint of $\boldsymbol{n}$. Then*

$$n_0 = \frac{\sum_{i<m} a^i b^{k_m - k_i} R_i}{a^m - b^{k_m}}.$$

Proof. By Proposition 3.1,

$$\begin{aligned} && n_m &= \frac{b^{k_m}}{a^m}\left(n_0 + \sum_{i<m} \frac{a^i}{b^{k_i}} R_i\right) \\ \Longrightarrow && a^m n_m &= b^{k_m}\left(n_0 + \sum_{i<m} \frac{a^i}{b^{k_i}} R_i\right) \\ \Longrightarrow && a^m n_m - b^{k_m} n_0 &= \sum_{i<m} a^i b^{k_m - k_i} R_i. \end{aligned}$$

Because $\boldsymbol{n}$ is a cycle, $n_m = n_0$ and so

$$\left(a^m - b^{k_m}\right) n_0 = \sum_{i<m} a^i b^{k_m - k_i} R_i.$$

Finally, by Lemma 3.2, we can divide through by $\left(a^m - b^{k_m}\right)$, giving us

$$n_0 = \frac{\sum_{i<m} a^i b^{k_m - k_i} R_i}{a^m - b^{k_m}},$$

as required. □

4. Restrictions on cycles

In this section we'll use what we've built to establish bounds on parameters of cycles and conclude by observing for $\langle a, b\rangle$ shadow functions that admit any novel cycles then those cycles must be very long.

Corollary 4.1. *Let $\boldsymbol{n}$ be a chain of the $\langle a, b\rangle$ shadow function g, of length m. Then*

$$n_m \geq \frac{b^k}{a^m} n_0$$

where k is the itinerary count of $\boldsymbol{n}$.

Proof. Let $\boldsymbol{k}$ be the itinerary counter of $\boldsymbol{n}$ and let $\boldsymbol{R}$ be the blueprint of $\boldsymbol{n}$. Proposition 3.1 tells us that

$$n_m = \frac{b^k}{a^m}\left(n_0 + \sum_{i<m} \frac{a^i}{b^{k_i}} R_i\right).$$

By Lemma 2.7, For each $i < m$, $R_i \geq 0$ and therefore

$$n_m \geq \frac{b^k}{a^m} n_0.$$

□

Corollary 4.2. *Let $\boldsymbol{n}$ be a cycle of the $\langle a, b\rangle$ shadow function g, of length m, and suppose that $n_0 > 0$. Then*

$$b^k \leq a^m$$

where k is the itinerary count of $\boldsymbol{n}$.

Proof. By Corollary 4.1,

$$n_m \geq \frac{b^k}{a^m} n_0.$$

Because $\boldsymbol{n}$ is a cycle, $n_m = n_0$ and so

$$n_0 \geq \frac{b^k}{a^m} n_0.$$

Because $n_0 > 0$ we can divide through by n_0 yielding

$$a^m \geq b^k.$$

□

Proposition 4.3. *Let $\boldsymbol{n}$ be a chain of the $\langle a, b\rangle$ shadow function g, of length m. Suppose $s = \min \boldsymbol{n} > 1$. Then*

$$n_m \leq \frac{b^k}{a^m} \left(\frac{s}{s-1}\right)^{m-k} n_0$$

where k is the itinerary count of $\boldsymbol{n}$.

Proof. Let functions $h_0, h_1 \colon \mathbb{Q} \to \mathbb{Q}$ be such that for all $x \in \mathbb{Q}$,

$$h_0(x) = \frac{s}{a(s-1)} x$$
$$h_1(x) = \frac{b}{a} x$$

Let $\boldsymbol{x} = \langle x_i \mid i \leq m\rangle$ be such that $x_0 = n_0$ and, for all $i < m$,

$$x_{i+1} = h_{\mu_i}(x_i)$$

where $\boldsymbol{\mu}$ is the itinerary of $\boldsymbol{n}$.

Suppose $n_i \leq x_i$ for some $i < m$.

If $\mu_i = 0$, then a does not divide n_i and

$$n_{i+1} = g(n_i) = \left\lceil \frac{n_i}{a} \right\rceil.$$

From this we know that

$$n_{i+1} - 1 < \frac{n_i}{a}.$$

Also, $n_{i+1} \geq s > 1$, so

$$\begin{aligned} & n_{i+1}(s-1) \leq s(n_{i+1}-1) \\ \implies & \left\lceil \frac{n_i}{a} \right\rceil (s-1) \leq s(n_{i+1}-1) \\ \implies & \left\lceil \frac{n_i}{a} \right\rceil (s-1) < s\frac{n_i}{a} \\ \implies & \left\lceil \frac{n_i}{a} \right\rceil < \frac{s}{a(s-1)} n_i \end{aligned}$$

Hence,

$$n_{i+1} = \left\lceil \frac{n_i}{a} \right\rceil < \frac{s}{a(s-1)} n_i \leq \frac{s}{a(s-1)} x_i = h_0(x_i) = x_{i+1}.$$

If instead $\mu_i = 1$, then a divides n_i and

$$n_{i+1} = g(n_i) = \frac{b}{a} n_i \leq \frac{b}{a} x_i = h_1(x_i) = x_{i+1}.$$

In both cases we have that $n_{i+1} \leq x_{i+1}$. By induction, we must have $n_m \leq x_m$.

Now, because $h_0 \circ h_1 = h_1 \circ h_0$,

$$x_m = \left(h_0^{m-k} \circ h_1^k\right)(x_0) = \left(\frac{s}{a(s-1)}\right)^{m-k} \left(\frac{b}{a}\right)^k x_0 = \frac{b^k}{a^m} \left(\frac{s}{s-1}\right)^{m-k} x_0.$$

Thus,

$$n_m \leq \frac{b^k}{a^m}\left(\frac{s}{s-1}\right)^{m-k} n_0.$$

□

Corollary 4.4. *Let* $\boldsymbol{n}$ *be a cycle of the* $\langle a, b\rangle$ *shadow function* g*, of length* $m > 0$*. Let* k *be the itinerary count of* $\boldsymbol{n}$ *and suppose* $k < m$*. Suppose* $s = \min \boldsymbol{n} > 1$*. Then*

$$s \leq \frac{a}{a - b^{\frac{k}{m}}}.$$

Proof. By Proposition 4.3,

$$n_m \leq \frac{b^k}{a^m}\left(\frac{s}{s-1}\right)^{m-k} n_0.$$

Because $\boldsymbol{n}$ is a cycle, $n_m = n_0$ and so

$$n_0 \leq \frac{b^k}{a^m}\left(\frac{s}{s-1}\right)^{m-k} n_0.$$

Because $n_0 > 0$ we can divide through by n_0. In what follows, recall that a, b, m, s, and $s-1$ are all positive here.

$$\begin{aligned}
& 1 \leq \frac{b^k}{a^m}\left(\frac{s}{s-1}\right)^{m-k} \\
\implies \quad & \frac{a^m}{b^k} \leq \left(\frac{s}{s-1}\right)^{m-k} \\
\implies \quad & \frac{a^m}{b^k} \leq \left(\frac{s}{s-1}\right)^{m} \\
\implies \quad & \frac{a}{b^{\frac{k}{m}}} \leq \frac{s}{s-1} \\
\implies \quad & \frac{a(s-1)}{s} \leq b^{\frac{k}{m}} \\
\implies \quad & a - \frac{a}{s} \leq b^{\frac{k}{m}} \\
\implies \quad & a - b^{\frac{k}{m}} \leq \frac{a}{s}
\end{aligned}$$

Together, Lemma 3.2 and Corollary 4.2 tell us that $b^k < a^m$. Therefore, $a - b^{\frac{k}{m}} > 0$ and so

$$s \leq \frac{a}{a - b^{\frac{k}{m}}}.$$

□

5. Collatz shadow function

5.1. **Introduction to Collatz.** The Collatz Conjecture is a famous, unsolved problem in mathematics posed by Lothar Collatz in 1937. In this section we hope to provide clarity on the connection between $\langle a, b\rangle$ shadow functions and the Collatz Conjecture by defining the Collatz shadow function and its basic properties.

Definition 5.1. Let $f\colon \mathbb{Z} \to \mathbb{Z}$ be such that

$$n \mapsto \begin{cases} 3n+1 & \text{if } n \text{ is odd;} \\ \frac{n}{2} & \text{if } n \text{ is even.} \end{cases}$$

We call f the *Collatz function.*

Definition 5.2. Given $n \in \mathbb{Z}$, the *Collatz sequence* from n is the sequence

$$\langle f^i(n) \mid i \in \mathbb{N} \rangle$$

Conjecture 5.3 (The Collatz Conjecture)**.** *For all positive integers n, the Collatz sequence from n contains 1.*

Notice that when $n \in \mathbb{Z}$ is odd, $f(n) = 3n+1$ which is even and so $f^2(n) = \frac{3n+1}{2}$. Knowing this, we're naturally inclined to define a modified version of the Collatz function.

Definition 5.4. Let $f \colon \mathbb{Z} \to \mathbb{Z}$ be such that

$$n \mapsto \begin{cases} \frac{3n+1}{2} & \text{if } n \text{ is odd;} \\ \frac{n}{2} & \text{if } n \text{ is even.} \end{cases}$$

We call f the *modified Collatz function.*

Clearly, Conjecture 5.3 holds if and only if it holds for the modified Collatz function.

5.2. **A different perspective.** In studying the Collatz function, it is common to collapse adjacent even elements of a Collatz sequence together, dividing by 2 repeatedly until encountering an odd number. This technique shapes our view of the problem and influences the results we produce. We will define a Collatz-like function which allows us to more naturally collapse adjacent elements corresponding to the odd branch of the Collatz function and see what this different perspective yields.

Definition 5.5. Let f be the modified Collatz function and let $g \colon \mathbb{Z} \to \mathbb{Z}$ be defined by

$$n \mapsto f(n-1) + 1.$$

We call g the *Collatz shadow function.*

Lemma 5.6. *Let f be the modified Collatz function and g be the Collatz shadow function. Then, for all $n \in \mathbb{Z}$ and all $i \in \mathbb{N}$,*

$$g^i(n) = f^i(n-1) + 1.$$

Proof. We prove this by induction on i. First note that

$$g^0(n) = n = (n-1) + 1 = f^0(n-1) + 1.$$

Now, suppose that for some $i \in \mathbb{N}$, $g^i(n) = f^i(n-1) + 1$. Then

$$\begin{aligned} g^{i+1}(n) &= g\big(g^i(n)\big) \\ &= f\big(g^i(n) - 1\big) + 1 \\ &= f\big(f^i(n-1) + 1 - 1\big) + 1 \\ &= f^{i+1}(n-1) + 1. \end{aligned}$$

□

Conjecture 5.7. *Let g be the Collatz shadow function. Then for all integers $n \geq 2$, there exists $i \in \mathbb{N}$ such that $g^i(n) = 2$.*

By Lemma 5.6, we note that Conjecture 5.7 is equivalent to Conjecture 5.3.

Proposition 5.8. *Let g be the Collatz shadow function. Then for all $n \in \mathbb{Z}$,*

$$g(n) = \begin{cases} \frac{n+1}{2} & \textit{if } n \textit{ is odd;} \\ \frac{3}{2}n & \textit{if } n \textit{ is even.} \end{cases}$$

Proof. Fix $n \in \mathbb{Z}$.

If n is odd then $n-1$ is even and

$$g(n) = f(n-1) + 1 = \frac{n-1}{2} + 1 = \frac{n+1}{2}.$$

If n is even then $n-1$ is odd and

$$g(n) = f(n-1) + 1 = \frac{3(n-1)+1}{2} + 1 = \frac{3n-2}{2} + 1 = \frac{3}{2}n.$$

□

The formulation in Proposition 5.8 suggests a natural generalisation of Definition 1.1. Note that the $\langle 2, 3\rangle$ shadow function is precisely the Collatz shadow function.

5.3. **Collatz chains and cycles.** In this subsection we will give the precise formulae of chains and the parameters of cycles for the Collatz shadow function.

Proposition 5.9. *Let $\boldsymbol{n}$ be a chain of the Collatz shadow function, of length m. Let k be the itinerary count of $\boldsymbol{n}$ and $\boldsymbol{L}$ be the itinerary indexer of $\boldsymbol{n}$. Then*

$$n_0 - 1 = \frac{2^m}{3^k}(n_m - 1) - \sum_{j<k} \frac{2^{L_j}}{3^{j+1}}.$$

Proof. Let $k = k_m$ and let $\boldsymbol{R}$ be the blueprint of $\boldsymbol{n}$. Because the Collatz shadow function is precisely the $\langle 2, 3\rangle$ shadow function, Proposition 3.1 gives us that

$$n_m = \frac{3^k}{2^m}\left(n_0 + \sum_{i<m} \frac{2^i}{3^{k_i}} R_i\right). \tag{3}$$

We will focus our attention on the summation within this equation.

$$\begin{aligned}\sum_{i<m} \frac{2^i}{3^{k_i}} R_i &= \sum_{j=0}^{k} \sum_{\substack{i<m\\ k_i=j}} \frac{2^i}{3^j} R_i \\ &= \sum_{\substack{i<m\\ k_i=0}} \frac{2^i}{3^0} R_i + \sum_{j=1}^{k-1} \sum_{\substack{i<m\\ k_i=j}} \frac{2^i}{3^j} R_i + \sum_{\substack{i<m\\ k_i=k}} \frac{2^i}{3^k} R_i.\end{aligned}$$

By Lemma 2.5, for all $i < m$ and all $1 \le j \le k$,

$$k_i \ge j \iff j - 1 < k_i \iff L_{j-1} < i \iff i \ge L_{j-1} + 1.$$

Also, for all $i < m$ and all $0 \le j \le k-1$,

$$k_i \le j \iff j \not< k_i \iff L_j \not< i \iff i \le L_j.$$

With this, we see that

$$\begin{aligned}&\sum_{\substack{i<m\\ k_i=0}} \frac{2^i}{3^0} R_i + \sum_{j=1}^{k-1} \sum_{\substack{i<m\\ k_i=j}} \frac{2^i}{3^j} R_i + \sum_{\substack{i<m\\ k_i=k}} \frac{2^i}{3^k} R_i \\ &= \sum_{i=0}^{L_0} \frac{2^i}{3^0} R_i + \sum_{j=1}^{k-1} \sum_{i=L_{j-1}+1}^{L_j} \frac{2^i}{3^j} R_i + \sum_{i=L_{k-1}+1}^{m-1} \frac{2^i}{3^k} R_i.\end{aligned}$$

Now, by Lemma 2.7, for all $i < m$, $R_i = 0$ or $R_i = 1$ and $R_i = 0$ if and only if $\mu_i = 1$. Because $(j \mapsto L_j)$ is a bijection, for all $i < m$, $R_i = 0$ if and only if there exists $j < k$ such that $L_j = i$. Therefore

$$\begin{aligned}&\sum_{i=0}^{L_0} \frac{2^i}{3^0} R_i + \sum_{j=1}^{k-1} \sum_{i=L_{j-1}+1}^{L_j} \frac{2^i}{3^j} R_i + \sum_{i=L_{k-1}+1}^{m-1} \frac{2^i}{3^k} R_i \\ &= \sum_{i=0}^{L_0-1} \frac{2^i}{3^0} + \sum_{j=1}^{k-1} \sum_{i=L_{j-1}+1}^{L_j-1} \frac{2^i}{3^j} + \sum_{i=L_{k-1}+1}^{m-1} \frac{2^i}{3^k}.\end{aligned}$$

Note the elementary identity: for all integers $s \leq t$,

$$\begin{aligned}\sum_{i=s}^{t-1} 2^i &= \sum_{i=s}^{t-1} 2^{i+1} - \sum_{i=s}^{t-1} 2^i \\ &= \sum_{i=s+1}^{t} 2^i - \sum_{i=s}^{t-1} 2^i \\ &= 2^t - 2^s.\end{aligned}$$

With this, we see that

$$\begin{aligned}&\sum_{i=0}^{L_0-1} \frac{2^i}{3^0} + \sum_{j=1}^{k-1} \sum_{i=L_{j-1}+1}^{L_j-1} \frac{2^i}{3^j} + \sum_{i=L_{k-1}+1}^{m-1} \frac{2^i}{3^k} \\ &= \frac{2^{L_0} - 2^0}{3^0} + \sum_{j=1}^{k-1} \frac{2^{L_j} - 2^{L_{j-1}+1}}{3^j} + \frac{2^m - 2^{L_{k-1}+1}}{3^k} \\ &= \left(\frac{2^m}{3^k} - \frac{2^0}{3^0}\right) + \frac{2^{L_0}}{3^0} + \sum_{j=1}^{k-1} \frac{2^{L_j}}{3^j} - \sum_{j=1}^{k-1} \frac{2^{L_{j-1}+1}}{3^j} - \frac{2^{L_{k-1}+1}}{3^k} \\ &= \left(\frac{2^m}{3^k} - \frac{2^0}{3^0}\right) + \left(\frac{2^{L_0}}{3^0} + \sum_{j=1}^{k-1} \frac{2^{L_j}}{3^j}\right) - \left(\sum_{j=0}^{k-2} \frac{2^{L_j+1}}{3^{j+1}} + \frac{2^{L_{k-1}+1}}{3^{(k-1)+1}}\right) \\ &= \left(\frac{2^m}{3^k} - \frac{2^0}{3^0}\right) + \sum_{j=0}^{k-1} \frac{2^{L_j}}{3^j} - \sum_{j=0}^{k-1} \frac{2^{L_j+1}}{3^{j+1}} \\ &= \left(\frac{2^m}{3^k} - \frac{2^0}{3^0}\right) + \sum_{j=0}^{k-1} \left(1 - \frac{2}{3}\right) \frac{2^{L_j}}{3^j} \\ &= \frac{2^m}{3^k} - 1 + \sum_{j<k} \frac{2^{L_j}}{3^{j+1}}.\end{aligned}$$

Substituting this into equation 3 yields

$$\begin{aligned}&& n_m &= \frac{3^k}{2^m} \left(n_0 + \frac{2^m}{3^k} - 1 + \sum_{j<k} \frac{2^{L_j}}{3^{j+1}}\right) \\ \implies && \frac{2^m}{3^k} n_m - \frac{2^m}{3^k} &= n_0 - 1 + \sum_{j<k} \frac{2^{L_j}}{3^{j+1}} \\ \implies && \frac{2^m}{3^k}(n_m - 1) - \sum_{j<k} \frac{2^{L_j}}{3^{j+1}} &= n_0 - 1,\end{aligned}$$

as required. □

As of this writing, the Collatz conjecture has been verified for all positive integers up to and including 87×2^{60}. This corresponds to verifying Conjecture 5.7 for all integers from 2 up to and including $87 \times 2^{60} + 1$.

In light of Corollary 4.4, we're encouraged to search for natural numbers m and k such that

$$\frac{2}{2 - 3^{\frac{k}{m}}} > 87 \times 2^{60} + 1.$$

A computer search returns 10 439 860 591 as the least possible value of m. It follows that there are no novel cycles of the Collatz shadow function of length less than ten billion.

5.4. **Thoughts for further work.** If Conjecture 5.7 is true then for every integer $n \geq 2$ there exists a chain $\boldsymbol{n}$ of the Collatz shadow function such that $n_0 = n$ and $n_m = 2$. By Proposition 5.9, we would have

$$n - 1 = \frac{2^m}{3^k}(2 - 1) - \sum_{j<k} \frac{2^{L_j}}{3^{j+1}}$$

where k is the itinerary count of $\boldsymbol{n}$ and $\boldsymbol{L}$ is the itinerary indexer of $\boldsymbol{n}$. Additionally, it seems intuitive that for the Collatz shadow function, m, k, and $\boldsymbol{L}$ would fully specify the chain $\boldsymbol{n}$ with $n_m = 2$. Accounting for the cycle $\langle 2, 3, 2\rangle$, we make the following conjecture.

Conjecture 5.10. *Every integer greater than* 2 *can be expressed in the form*

$$\frac{2^m}{3^k} - \sum_{j<k} \frac{2^{L_j}}{3^{j+1}}$$

where $m \in \mathbb{N}$ *and* $\langle L_j \mid j < k\rangle$ *is a strictly increasing sequence of natural numbers below* m*. Moreover, the expression is unique up to appending copies of the cycle* $\langle 2, 3, 2\rangle$.

Some examples:

$$3 = \frac{32}{9} - \left(\frac{1}{3} + \frac{2}{9}\right)$$
$$4 = \frac{4}{1}$$
$$5 = \frac{16}{3} - \left(\frac{1}{3}\right)$$
$$6 = \frac{64}{9} - \left(\frac{2}{3} + \frac{4}{9}\right)$$
$$7 = \frac{2048}{243} - \left(\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \frac{16}{81} + \frac{128}{243}\right)$$