Journal of Integer Sequences, Vol. 21 (2018),
Article 18.5.4

A Generalization of Collatz Functions and Jacobsthal Numbers

Ji Young Choi
Department of Mathematics
Shippensburg University of Pennsylvania
1871 Old Main Drive
Shippensburg, PA 17257
USA
jychoi@ship.edu

Abstract

Let $b \geq 2$ be an integer and $g = b - 1$. We consider a generalization of the modified Collatz function: For any positive integer m, the g-Collatz function f_g divides m by g, if m is a multiple of g; otherwise, the g-Collatz function f_g is the least integer greater than or equal to $\frac{bm}{g}$. Using this g-Collatz function, we extend the Collatz problem, and we show that there are nontrivial cycles for some g. Then we show how the function f_g transforms the base-b representation of positive integers, and we study the sequence of the b-ary representation of integers generated by the function f_g, starting with a b-ary string representing b^N for an arbitrary large integer N. We show each b-ary string in the sequence has a repeating string, and the number of occurrences of each digit in each shortest repeating string generalizes Jacobsthal numbers.

1 Introduction

Definition 1. [2, 4] For any positive integer m, the *Collatz function* f_1 and the *modified Collatz function* f_2 on m are defined as follows:

$$f_1(m) := \begin{cases} \frac{m}{2}, & \text{if } m \text{ is even;} \\ 3m+1, & \text{if } m \text{ is odd,} \end{cases} \quad \text{and} \quad f_2(m) := \begin{cases} \frac{m}{2}, & \text{if } m \text{ is even;} \\ \frac{3m+1}{2} = \lceil \frac{3m}{2} \rceil, & \text{if } m \text{ is odd.} \end{cases} \tag{1}$$

Collatz [2] asked if every positive integer m is mapped to 1 by applying the Collatz function or the modified Collatz function repeatedly: $f_i^n(m) = 1$ for some integer n, where $i = 1$ or 2.

Conway [3] proved that a generalization of the Collatz problem is undecidable, and there are several ways to generalize the Collatz function. For example, Conway considered the following:

$$\bar{f}(m) = a_i m + b_i, m \equiv i \pmod{p}, \tag{2}$$

where a_i and b_i are rational numbers so that $\bar{f}(m)$ is an integer. Notice that the Collatz function (1) is the same as (2), when $p = 2$, $a_0 = \frac{1}{2}$, $b_0 = 0$, $a_1 = 3$, and $b_1 = 1$.

Throughout this paper, we let g be an integer greater than 1 and $b = g + 1$ (unless we specify otherwise), and we consider $p = g$, $a_0 = \frac{1}{g}$, $b_0 = 0$, $a_i = b$, $b_i = g - i$ for (2), as follows:

$$\bar{f}_g(m) := \begin{cases} \frac{m}{g}, & \text{if } m \equiv 0 \pmod{g}; \\ bm + g - i, & \text{if } m \equiv i \pmod{g} \text{ for } i = 1, 2, \ldots, g-1. \end{cases} \tag{3}$$

Then, as we modified f_1 to f_2, we modify $\bar{f}_g$ as shown in the following definition, by considering $\bar{f}_g(m)$, if g divides m; $\bar{f}_g^2(m)$, otherwise.

Definition 2. For any integer $g \geq 2$ and any positive integer m, the *g-Collatz function* f_g on m is defined as follows:

$$f_g(m) := \begin{cases} \frac{m}{g}, & \text{if } m \equiv 0 \pmod{g}; \\ \frac{bm+g-i}{g} = \lceil \frac{bm}{g} \rceil, & \text{if } m \equiv i \pmod{g} \text{ for } i = 1, 2, \ldots, g-1. \end{cases} \tag{4}$$

Now we extend the Collatz problem. We ask if every positive integer m is eventually mapped to 1 by repeatedly applying the g-Collatz function: if $f_g^n(m) = 1$ for some integer n, and we call it the *g-Collatz problem*. We can answer this g-Collatz problem immediately for some g by finding a nontrivial cycle. For example, when $g = 3$, we can find a nontrivial cycle in

$$5 \to 7 \to 10 \to 14 \to 19 \to 26 \to 35 \to 47 \to 63 \to 21 \to 7 \to \cdots.$$

Table 1 shows the minimum positive integer m such that $f_g^n(m) \neq 1$ for every n and the minimum integer $k > g$ such that $f_g^n(k) = k$ for some n, where $g = 3, 4, \ldots, 20$. When

g	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
m	5	11		7			31	34	588	767			49	35	19			63
k	7	23		23			35	42	642	1348			53	178	79			71

Table 1: Minimum m and $k > g$: $m \not\to 1$ and $k \to k$ by f_g repeatedly

$g = 5, 7, 8, 13, 14, 18$, and 19, every positive integer up to 2×10^9 is mapped to 1 by f_g^n for some n, but we do not know whether every integer beyond 2×10^9 also reaches 1. The

g-Collatz problem for $g = 5, 7, 8, 13, 14, 18$, and 19 seems hard, just as in the original Collatz problem.

Do we know which values of g provide a nontrivial cycle for the corresponding g-Collatz problem? That is, can we find a pattern for g which makes the g-Collatz problem different from the original Collatz problem? If we can answer this question, we can solve the original Collatz problem, since the original Collatz problem is the 2-Collatz problem. Hence, we want to work on a different property of the Collatz function to see if the property can be extended for all g.

The number of occurrences of each digit in each shortest repeating string in the ternary Collatz sequence starting with 3^N for an arbitrary large N is expressed with Jacobsthal numbers [1]. We wonder if we can extend this. That is, we want to know if we can generalize Jacobsthal numbers, to express the number of occurrences of each digit in each shortest repeating string in the b-ary g-Collatz sequence starting with b^N for an arbitrary large N, for all g. It is easy to do for some g, but not easy for all g.

In this paper, we provide two different generalizations of Jacobsthal numbers: one is defined for $g \not\equiv 2 \pmod 4$ and the other for $g \equiv 2 \pmod 8$, except $g = 2$. For $g \equiv 6 \pmod 8$, we may need to consider infinitely many cases, and we could provide infinitely many new types of generalizations of Jacobsthal numbers. This is desirable for future work.

Section 2 clarifies the notation in this paper. Section 3 shows how to apply the g-Collatz function to the base-b representation of positive integers. In Section 4, we study the shortest repeating string in the sequence of b-ary strings representing $f_g^n(b^N)$ for an arbitrary large integer N. In Section 5, we study the number of occurrences of each digit in each shortest repeating string in the b-ary g-Collatz sequence, when $g \not\equiv 6 \pmod 8$. Finally, in Section 6, we define two different types of generalizations of Jacobsthal numbers to express the number of each digit studied in Section 5.

2 Notation

Every base-b representation of an integer is a finite string in $\{0, 1, 2, \dots, g\}^*$, which is the set of all finite strings consisting of digits $0, 1, 2, \dots, g$. The set $\{0, 1, 2, \dots, g\}^*$ also includes the *empty string*, which contains no digits, denoted by ϵ [5].

The notation for the number of digits in a string is as follows:

Notation 3. [5] For any finite string x and a digit a, let $|x|$ denote the number of digits in x, and $|x|_a$ denote the number of occurrences of digit a's in x.

Lemma 4. *For any string x in $\{0, 1, 2, \dots, g\}^*$,*

$$|x| = \sum_{i=0}^{g} |x|_i.$$

For example, $|01011| = 5$, $|01011|_0 = 2$, $|01011|_1 = 3$, and $5 = 2 + 3$.

The following operation shows how to create a new string from given ones [5]:

Definition 5. For any strings x and y and any positive integer n, the *concatenation* of x and y, denoted by xy, is the string obtained by joining x and y end-to-end, and x^n denotes the concatenation of n x's. That is, if $x = a_1a_2 \cdots a_{|x|}$ and $y = b_1b_2 \cdots b_{|y|}$ for some $a_i, b_i \in \{0, 1, 2, \ldots, g\}$,

$$xy = a_1a_2 \cdots a_{|x|}b_1b_2 \cdots b_{|y|}, \text{ and } x^n = xx \cdots x \text{ (}n\text{ times).}$$

For a convention, x^0 is defined to be ϵ.

Lemma 6. *For any strings x and y and a nonnegative integer n, $|xy| = |x| + |y|$ and $|x^n| = n|x|$.*

For example, $101\ 00 = 10100$, $(10)^3 = 101010$, and $1 = 1\ (10)^0$. Then, $|101\ 00| = |101| + |00| = 3 + 2 = 5$, $|(10)^3| = 3|10| = 3 \cdot 2 = 5$, and $|\epsilon| = |(10)^0| = 0$.

Since the base-b representation of an integer is a string in $\{0, 1, 2, \ldots, g\}^*$, we call the base-b representation of an integer as a *b-ary string* throughout this paper. When we have to distinguish an integer and its b-ary string, we use the following notation.

Notation 7. For any integer m with its base-b representation x, we write $m = [x]_b$ or $(m)_b = x$.

For example, $5 = [12]_3$ and $(5)_3 = 12$. Then, $([x]_b)_b = x$ for any b-ary string x and $[(m)_b]_b = m$ for any integer m.

Throughout this paper, we use the convention that m is an integer and x is its b-ary string. When we apply the g-Collatz function f_g, we often phrase this in terms of how f_g transforms x to another b-ary string, and we do not mention m.

Notation 8. For a b-ary string x, we let $f_g(x)$ denote the *b-ary representation of* $f_g([x]_b)$. That is, $f_g(x) = (f_g([x]_b))_b$.

To apply the g-Collatz function on a b-ary string, it is important to know whether a given b-ary string represents a multiple of g or not. For any integer m, we let $m \bmod g$ denote the least nonnegative residue of m modulo g, and $s_b(m)$ denote the digit sum of the base-b representation of m. That is, $m \bmod g$ is the remainder when m is divided by g, and if $(m)_b = a_1a_2 \cdots a_{k-1}a_k$,

$$s_b(m) = \sum_{i=1}^{k} a_i.$$

It is well-known that $s_b(m)$ is congruent to m modulo g. Hence,

$$m \bmod g \equiv s_b(m) \pmod{g}. \tag{5}$$

For a convention, we define the notation for the sum of digits in x, and the remainder when $[x]_b$ is divided by g, for any b-ary string x.

Definition 9. For any b-ary string x, we let $s(x)$ and $r(x)$ denote as follows:

$$s(x) = \begin{cases} s_b([x]_b), & \text{if } x \neq \epsilon; \\ 0, & \text{if } x = \epsilon, \end{cases} \text{ and } r(x) = \begin{cases} [x]_b \bmod g, & \text{if } x \neq \epsilon; \\ 0, & \text{if } x = \epsilon. \end{cases}$$

Lemma 10. *For any b-ary string x, $r(x) \equiv s(x) \pmod{g}$.*

To simplify arguments, we sometimes use the following notation.

Notation 11. For any integers a and b, let $a \equiv_g b$ denote $a \equiv b \pmod{g}$.

3 Generalized Collatz functions

For any nonzero digit $a \in \{0, 1, 2, \ldots, g\}$, the product $g \cdot a$ can be represented by a b-ary string of length 2, whose digit sum is g.

Lemma 12. *For any $a \in \{1, 2, \ldots, g\}$, the product $g \cdot a = [a-1, b-a]_b$ and the sum $s_b(g \cdot a) = g$.*

Proof. Since $g = b - 1$, $g \cdot a = (a-1) \cdot b + (b-a)$, so $s_b(g \cdot a) = a - 1 + b - a = g$. □

Lemma 13. *For any digits a_1 and a_0 in $\{0, 1, 2, \ldots, g\}$ with $a_1 < g$, if $[a_1 a_0]_b = g \cdot q + r$ for $0 \leq r < g$,*

$$q = \begin{cases} a_1, & \text{if } a_1 + a_0 < g; \\ a_1 + 1, & \text{if } a_1 + a_0 \geq g, \end{cases} \text{ and } r = \begin{cases} a_0 + q, & \text{if } a_0 + q < b; \\ a_0 + q - b, & \text{if } a_0 + q \geq b. \end{cases}$$

Proof. If $[a_1 a_0]_b < g$, the digit $a_1 = 0$ so it is obvious that $q = 0$ and $r = a_0$. Hence, assume $[a_1 a_0]_b \geq g$. Then, by Lemma 12, $g \cdot q = (q-1) \cdot b + b - q$, so

$$r = [a_1 a_0]_b - g \cdot q = (a_1 - q + 1) \cdot b + a_0 - b + q.$$

Since $a_1 < g$, the number $[a_1 a_0]_b < [g0]_b = g \cdot b$ so $q \leq b - 1$. Then, $a_0 - b + q \leq a_0 - 1 \leq g - 1$. Hence, if $a_0 - b + q \geq 0$, the remainder $r = a_0 - b + q$ and $a_1 - q + 1 = 0$, so $q = a_1 + 1$. Then, $a_0 - b + (a_1 + 1) \geq 0$, so $a_1 + a_0 \geq b - 1 = g$. If $a_0 + q - b < 0$, the remainder $r = a_0 - b + q + b = a_0 + q$ and $a_1 - q = 0$, so $q = a_1$. Hence, q and r are as desired. □

Lemma 14. *For any digit a_i's in $\{0, 1, 2, \ldots, g\}$, if $[a_1 a_2 \cdots a_k]_b \equiv 0 \pmod{g}$,*

$$\frac{[a_1 a_2 \cdots a_k]_b}{g} = [a_1' a_2' \cdots a_k']_b,$$

where $a_1' = 1$ if $a_1 = g$; 0 otherwise, and for $i > 1$,

$$a_i' = \begin{cases} r(a_1 a_2 \cdots a_{i-1}), & \text{if } a_i < g - r(a_1 a_2 \cdots a_{i-1}); \\ r(a_1 a_2 \cdots a_{i-1}) + 1, & \text{if } a_i \geq g - r(a_1 a_2 \cdots a_{i-1}). \end{cases}$$

Proof. It is obvious for a'_1. For any $i > 1$, let $r_i = r(a_1 a_2 \cdots a_{i-1})$. Then, r_i is a digit $< g$, and a'_i is the quotient when $[r_i a_i]_b$ is divided by g. Hence, $a'_i = r_i$ if $r_i + a_i < g$; $r_i + 1$ otherwise, by Lemma 13. That is, $a'_i = r_i$ if $a_i < g - r_i$; $r_i + 1$ otherwise. □

The following lemma shows how $\bar{f}_g$ defined in (3) transforms a b-ary string representing a non-multiple of g to a b-ary string representing a multiple of g.

Lemma 15. *For any b-ary string x, if $[x]_b \not\equiv 0 \pmod{g}$, $f_g(x) = f_g(xa)$, where $a = g - r(x)$.*

Proof. Since $[x]_b \not\equiv 0 \pmod{g}$, the number $\bar{f}_g([x]_b) = b[x]_b + g - r(x) = [x0]_b + a = [xa]_b$ by (3). Then, $[xa]_b \equiv_g s(xa) = s(x) + a \equiv_g r(x) + a \equiv 0 \pmod{g}$. Hence, by Definition 2, $f_g([x]_b) = \frac{[xa]_b}{g} = f_g([xa]_b)$. □

For example, when $b = 10$ and $x = 9107222$, $g = 9$ and $r(x) = 5$. Hence, $a = 9 - 5 = 4$ so $\bar{f}_9(x) = 9107222\ 4$ so $f_9(x) = f_9(91072224)$.

For a b-ary string x representing a multiple of g, the g-Collatz function f_g divides $[x]_b$ by g. Hence, by combining Lemma 14 and 15, we find how the g-Collatz function f_g transforms a b-ary string to a b-ary string.

Theorem 16. *For any digits a_i's in $\{0, 1, 2, \ldots, g\}$,*

$$f_g(a_1 a_2 \cdots a_k) = \begin{cases} a'_1 a'_2 \cdots a'_k, & \text{if } [a_1 a_2 \cdots a_k]_b \equiv 0 \pmod{g}; \\ a'_1 a'_2 \cdots a'_k a'_{k+1}, & \text{if } [a_1 a_2 \cdots a_k]_b \not\equiv 0 \pmod{g}, \end{cases}$$

where $a'_1 = 1$ if $a_1 = g$; 0 otherwise, $a'_{k+1} = r(a_1 a_2 \cdots a_k) + 1$, and for $i = 2, 3, \ldots, k$,

$$a'_i = \begin{cases} r(a_1 a_2 \cdots a_{i-1}), & \text{if } a_i < g - r(a_1 a_2 \cdots a_{i-1}); \\ r(a_1 a_2 \cdots a_{i-1}) + 1, & \text{if } a_i \geq g - r(a_1 a_2 \cdots a_{i-1}). \end{cases}$$

Proof. Let $x = a_1 a_2 \cdots a_k$. When $[x]_b \equiv 0 \pmod{g}$, it is obvious by Lemma 14. Hence, assume $[x]_b \not\equiv 0 \pmod{g}$. By Lemma 15, $f_g(x) = f_g(xa)$, where $a = g - r(x)$. Since $[xa]_b \equiv 0 \pmod{g}$, the number $f_g([x]_b) = \frac{[xa]_b}{g}$. Therefore, a'_i's are obtained by Lemma 14. Especially, $a'_{k+1} = r(x) + 1$, since $a = g - r(x)$. □

In this paper, when a head digit is transformed to digit 0, we **keep the new head digit** 0, so that $|f_g(x)| =$ either $|x|$ or $|x|+1$ for any string x. For example, when $b = 10$, $f_9(36099) = 04011$ and $f_9(36095) = 040106$ so that $|04011| = |36099|$ and $|040106| = |36095| + 1$.

Corollary 17. *For any digits a_i's in $\{0, 1, 2, \ldots, g\}$, if $f_g(a_1 a_2 \cdots a_k) = a'_1 a'_2 \cdots a'_k a'$ for some digits a'_i and a', allowing $a' = \epsilon$, the digit $a'_i \neq g - a_i$.*

Proof. If $a_i < g - r(a_1 a_2 \cdots a_{i-1})$, the digit $a'_i = r(a_1 a_2 \cdots a_{i-1}) < g - a_i$. If $a_i \geq g - r(a_1 a_2 \cdots a_{i-1})$, the number $r(a_1 a_2 \cdots a_{i-1}) \geq g - a_i$, so $a'_i = r(a_1 a_2 \cdots a_{i-1}) + 1 > r(a_1 a_2 \cdots a_{i-1}) \geq g - a_i$. □

a'	$r(x)=0$	1	2	3	$\cdots$	$g-3$	$g-2$	$g-1$
$a=0$	0	1	2	3	$\cdots$	$g-3$	$g-2$	$g-1$
1	0	1	2	3	$\cdots$	$g-3$	$g-2$	g
2	0	1	2	3	$\cdots$	$g-3$	$g-1$	g
3	0	1	2	3	$\cdots$	$g-2$	$g-1$	g
$\vdots$					$\vdots$			
$g-3$	0	1	2	4	$\cdots$	$g-2$	$g-1$	g
$g-2$	0	1	3	4	$\cdots$	$g-2$	$g-1$	g
$g-1$	0	2	3	4	$\cdots$	$g-2$	$g-1$	g
g	1	2	3	4	$\cdots$	$g-2$	$g-1$	g

Table 2: The image a' of a in $f_g(xay) = x'a'y'$

Table 2 shows how f_g transforms digit a to digit a' satisfying $f_g(xay) = x'a'y'$ for any b-ary strings x, x', y, and y' with $|x| = |x'|$. Notice that each image digit a' of digit a depends on $r(x)$, and there are g distinct images of each digit a.

Note 18. For any b-ary strings x_1, x_2, y_1, and y_2 and any digit a, let x'_1, x'_2, y'_1, and y'_2 be b-ary strings and a'_1 and a'_2 be digits satisfying $f_g(x_1ay_1) = x'_1a'_1y'_1$ and $f_g(x_2ay_2) = x'_2a'_2y'_2$ with $|x_1| = |x'_1|$ and $|x_2| = |x'_2|$. Then, $a'_1 = a'_2$ iff $r(x_1) = r(x_2)$.

Now consider the g-Collatz function f_g on a concatenation of b-ary strings.

Lemma 19. *For any b-ary strings y and z with $[y]_b \equiv 0 \pmod g$,*

$$f_g(yz) = f_g(y)f_g(z).$$

Proof. Since $s(y) \equiv_g [y]_b \equiv 0 \pmod g$, the number $[yz]_b \equiv_g s(yz) = s(y) + s(z) \equiv_g s(z) \equiv [z]_b \pmod g$. Hence, if $[z]_b \equiv 0 \pmod g$, $[f_g(yz)]_b = \frac{[yz]_b}{g}$ and

$$[f_g(y)f_g(z)]_b = \left[\left(\frac{[y]_b}{g}\right)_b \left(\frac{[z]_b}{g}\right)_b\right]_b = \left[\left(\frac{[y]_b}{g}\right)_b 0^{|z|}\right]_b + \frac{[z]_b}{g} = \frac{[y0^{|z|}]_b + [z]_b}{g} = \frac{[yz]_b}{g}.$$

If $[z]_b \not\equiv 0 \pmod g$, the string $f_g(z) = f_g(za)$, where $a = g - r(z)$ by Lemma 15. Since $[yz]_b \equiv [z]_b \pmod g$, the remainder $r(yz) = r(z)$. Hence, $a = g - r(yz)$, so $f_g(yz) = f_g(yza)$. Since $[za]_b \equiv 0 \pmod g$, the string $f_g(yz) = f_g(yza) = f_g(y)f_g(za) = f_g(y)f_g(z)$. □

Theorem 20. *For any b-ary strings y_i's and z, if $[y_i]_b \equiv 0 \pmod g$ for all i,*

$$f_g(y_1y_2\cdots y_kz) = f_g(y_1)f_g(y_2)\cdots f_g(y_k)f_g(z). \tag{6}$$

Proof. Since $[y_i]_b \equiv 0 \pmod g$ for all i, by Lemma 19, $f_g(y_1y_2\cdots y_kz) = f_g(y_1)f_g(y_2\cdots y_kz)$. Continuing this, (6) is obtained. □

In order to study a lengthy b-ary g-Collatz sequence with repeating digits in the following section, we present the following corollaries and theorem.

Corollary 21. *For any b-ary string x and z and for any positive integer k, if* $\gcd([x]_b, g) = d$,

$$f_g(x^k z) = \left(f_g(x^{\frac{g}{d}})\right)^{\left\lfloor \frac{dk}{g} \right\rfloor} f_g(x^{k \bmod \frac{g}{d}} z).$$

Proof. Since $k = \frac{g}{d}\left\lfloor \frac{dk}{g} \right\rfloor + k \bmod \frac{g}{d}$, the string $x^k = (x^{\frac{g}{d}})^{\left\lfloor \frac{dk}{g} \right\rfloor} x^{k \bmod \frac{g}{d}}$. Since $[x^{\frac{g}{d}}]_b \equiv_g s(x^{\frac{g}{d}}) = \frac{g}{d} \cdot s(x) \equiv_g \frac{g}{d}[x]_b \equiv_g g \cdot \frac{[x]_b}{d} \equiv 0 \pmod g$, we can apply Theorem 20. □

Theorem 22. *For any b-ary string x and digits a_i's, if $x = a_1 a_2 \cdots a_{|x|}$ and $\gcd([x]_b, g) = 1$, the b-ary string $f_g(x^g) = a'_1 a'_2 \cdots a'_{g|x|}$ for some digits a'_i's, where*

$$\{a'_k, a'_{|x|+k}, \ldots, a'_{(g-1)|x|+k}\} = \{a \in \{0, 1, 2, \ldots, g\} | a \neq g - a_k\}.$$

for any $k = 1, 2, \ldots, |x|$.

Proof. Since $[x^g]_b \equiv_g s(x^g) = g \cdot s(x) \equiv 0 \pmod g$, the number $|f_g(x^g)| = |x^g| = g|x|$. Hence, $f_g(x^g) = a'_1 a'_2 \cdots a'_{g|x|}$ for some digits a'_i's. Let $a_{i|x|+k}$ be the $(i|x| + k)$-th digit in the string x^g. Then, for any i and j with $0 \le i \neq j \le g - 1$, the digit $a_{i|x|+k} = a_{j|x|+k}$ but $r(a_1 a_2 \cdots a_{i|x|+k}) \neq r(a_1 a_2 \cdots a_{j|x|+k})$. (If so, $r(x^i) = r(x^j)$. Then, $i \cdot [x]_b \equiv_g i \cdot s(x) = s(x^i) \equiv_g s(x^j) = j \cdot [x]_b \equiv j \cdot r(x) \pmod g$, so $(i - j) \cdot [x]_b \equiv 0 \pmod g$, which is impossible, since $\gcd([x]_b, g) = 1$ and $0 \le i \neq j \le g - 1$.) Hence, $a'_{i|x|+k} \neq a'_{j|x|+k}$ by Note 18, so $|\{a'_k, a'_{|x|+k}, \ldots, a'_{(g-1)|x|+k}\}| = g$. Since $a'_{i|x|+k} \neq g - a_k$ by Corollary 17, the set $\{a'_k, a'_{|x|+k}, \ldots, a'_{(g-1)|x|+k}\}$ collects every possible image of a_k by f_g. That is, the set $\{a'_k, a'_{|x|+k}, \ldots, a'_{(g-1)|x|+k}\}$ contains every digit in $\{0, 1, 2, \ldots, g\}$ except $g - a_k$. □

Corollary 23. *For any b-ary string x and digit $a \in \{0, , 1, 2, \ldots, g\}$, if* $\gcd([x]_b, g) = 1$,

$$|f_g(x^g)|_a = |x| - |x|_{g-a}. \tag{7}$$

Proof. Theorem 22 shows that every digit a' in the string x is transformed to g distinct digits in $f_g(x^g)$, and each new digit a in $f_g(x^g)$ cannot be equal to $g - a'$. That is, every digit a in $f_g(x^g)$ is obtained by transforming digit a' in x, where $a' \neq g - a$, and there is no other way to obtain digit a in $f_g(x^g)$. Hence,

$$|f_g(x^g)|_a = \sum_{a' \neq g-a} |x|_{a'} = \sum_{a'=0}^{g} |x|_{a'} - |x|_{g-a} = |x| - |x|_{g-a}.$$

□

Corollary 24. *For any b-ary string x, if* $\gcd([x]_b, g) = 1$,

$$[f_g(x^g)]_b \equiv \begin{cases} [x]_b + \frac{g}{2} \pmod g, & \textit{if } g \textit{ is even and } |x| \textit{ is odd;} \\ [x]_b \pmod g, & \textit{otherwise.} \end{cases}$$

Proof. By Corollary 23,

$$
\begin{aligned}
[f_g(x^g)]_b \equiv_g s(f_g(x^g)) = \sum_{a=0}^{g} a|f_g(x^g)|_a &= \sum_{a=0}^{g} a(|x| - |x|_{g-a}) \\
&\equiv_g |x| \sum_{a=0}^{g} a + \sum_{a=0}^{g} (g-a)|x|_{g-a} \\
&= \frac{g(g+1)}{2}|x| + s(x) \\
&\equiv_g \frac{g(g+1)}{2}|x| + [x]_b.
\end{aligned}
$$

If $g+1$ or $|x|$ is even, the number $\frac{g(g+1)}{2}|x| \equiv 0 \pmod{g}$, so $[f_g(x^g)]_b \equiv [x]_b \pmod{g}$. If $g+1$ and $|x|$ are odd, g is even and $|x| = 2k+1$ for some integer k. Hence,

$$
\frac{g(g+1)}{2}|x| = \frac{g}{2}(g+1)(2k+1) \equiv_g \frac{g}{2} \cdot 2k + \frac{g}{2} = gk + \frac{g}{2} \equiv_g \frac{g}{2}.
$$

□

4 b-ary g-Collatz sequences

Consider a sequence of b-ary strings generated by the g-Collatz function f_g, starting with a b-ary string 10^N for any arbitrary large positive integer N. For example, when $g = 3$, $b = 4$. Then, the first few strings in the 4-ary 3-Collatz sequence starting with the string $1\ 0^{60}$ are as follows:

$$
\begin{aligned}
1\ 0^{60} &= 1000000000000000000000000000000000000000000000000000000000000; \\
f_3^1(1\ 0^{60}) &= 01111111111111111111111111111111111111111111111111111111111112; \\
f_3^2(1\ 0^{60}) &= 001301301301301301301301301301301301301301301301301301301301303; \\
f_3^3(1\ 0^{60}) &= 0002113231002113231002113231002113231002113231002113231002113233; \\
f_3^4(1\ 0^{60}) &= 0000302210112013321223131033000302210112013321223131033000302211.
\end{aligned}
$$

Since digit 0 repeats in the initial string 10^N, there exists a substring repeats in $f_g^n(10^N)$, ignoring the head digit and a tail string. For example, when $g = 3$, the substrings 013 and 002113231 repeat in the 4-ary strings $f_3^2(1\ 0^{60})$ and $f_3^3(1\ 0^{60})$, respectively.

Definition 25. Let N be an arbitrary large integer. For any positive integer n and any g-Collatz function f_g, the nth repeating string u_{gn} is defined as the shortest string in the b-ary string $f_g^n(10^N)$ such that

$$
f_g^n(10^N) = 0(u_{gn})^{\left\lfloor \frac{N}{|u_{gn}|} \right\rfloor} t
$$

for some b-ary string t.

For example, u_{3n} for $n = 1, 2, 3, 4$ is as follows:

$$\begin{aligned} u_{31} &= 1; \\ u_{32} &= 013; \\ u_{33} &= 002113231; \\ u_{34} &= 000302210112013321223131033. \end{aligned}$$

Since $r(10^N) = 1$ for any $g \geq 2$, the string $f_g(10^N) = 01^N2$. Hence, $u_{g1} = 1$ for any $g \geq 2$. Then, by Theorem 16, $u_{g2} = f_g(1^g) = 012 \ldots (g-3)(g-2)g$. For example, u_{g2} for $g = 2, 3, \ldots, 9$ is as follows.

g	2	3	4	5	6	7	8	9
u_{g2}	02	013	0124	01235	012346	0123457	01234568	012345679

Then, $|u_{g2}| = g$, and $[u_{g2}]_b \equiv 1 \pmod{g}$ if g is odd; $1 + \frac{g}{2} \pmod{g}$ if g is even, by Corollary 24. Hence, $\gcd([u_{g2}]_b, g) = 1$ for odd g. For even g, we consider two cases: $g = 4k$ or $4k + 2$ for some integer k. If $g = 4k$, the number $[u_{g2}]_b \equiv_g 1 + \frac{g}{2} = 1 + 2k$. Since $\gcd(1 + 2k, k) = 1$ and $\gcd(1 + 2k, 4) = 1$, $\gcd(1 + 2k, 4k) = 1$. If $g = 4k + 2$, the number $[u_{g2}]_b \equiv_g 1 + \frac{g}{2} = 2(k + 1)$. Since $\gcd(k + 1, 2k + 1) = 1$, $\gcd(2(k + 1), 2(2k + 1)) = 2$. Hence,

$$\gcd([u_{g2}]_b, g) = \begin{cases} 1, & \text{if } g \not\equiv 2 \pmod{4}; \\ 2, & \text{if } g \equiv 2 \pmod{4}. \end{cases}$$

Therefore, by Corollary 21, the string $u_{g3} = f_g(u_{g2}^g)$ if $g \not\equiv 2 \pmod{4}$; $(u_{g2}^{\frac{g}{2}})$ otherwise. For example, u_{g3} for $g = 3, 4, 5, 6$ is as follows.

g	3	4	5	6
u_{g3}	002113231	0014340322421131	0014211253224043351545031	001405446153223631

Observation 26. For any integer $g \geq 2$,

$$u_{g1} = 1; \qquad |u_{g1}| = 1;$$

$$u_{g2} = 012 \cdots (g-3)(g-2)g; \qquad |u_{g2}| = g;$$

$$u_{g3} = \begin{cases} f_g(u_{g2}^g), & \text{if } g \not\equiv 2 \pmod{4}; \\ f_g(u_{g2}^{\frac{g}{2}}), & \text{if } g \equiv 2 \pmod{4}, \end{cases} \qquad |u_{g3}| = \begin{cases} g|u_{g2}| = g^2, & \text{if } g \not\equiv 2 \pmod{4}; \\ \frac{g}{2}|u_{g2}| = \frac{g^2}{2}, & \text{if } g \equiv 2 \pmod{4}. \end{cases}$$

To calculate the string u_{g4}, we need to know $\gcd([u_{g3}]_b, g)$. Since $|u_{g2}| = g$, the number $|u_{g2}|$ is even, if g is even. That is, there is no such case that g is even and $|u_{g2}|$ is odd. If $g \not\equiv 2 \pmod{4}$, $\gcd([u_{g2}]_b, g) = 1$. Hence, $[f_g(u_{g2}^g)]_b \equiv [u_{g2}]_b \pmod{g}$ by Corollary 24. Therefore,

$$\gcd([u_{g3}]_b, g) = 1, \text{ if } g \not\equiv 2 \pmod{4}. \tag{8}$$

Consider $g \equiv 2 \pmod 4$. Since $u_{g3} = f_g(u_{g2}^{\frac{g}{2}})$, let a_i''s be digits satisfying $u_{g3} = a_1'a_2'\cdots a_{\frac{g^2}{2}}'$, and Theorem 16 provides the following: For any $m = 0,1,2,\ldots,\frac{g}{2}-1$ and any digit $a = 0,1,2,\ldots,g-2$,

$$\begin{aligned} a'_{mg+a+1} &= \begin{cases} r(u_{g2}^m 012\cdots(a-2)(a-1)), & \text{if } a < g - r(u_{g2}^m 012\cdots(a-1)); \\ r(u_{g2}^m 012\cdots(a-2)(a-1))+1, & \text{if } a \geq g - r(u_{g2}^m 012\cdots(a-1)); \end{cases} \\ a'_{(m+1)g} &= r(u_{g2}^m 012\cdots(g-3)(g-2))+1, \end{aligned} \tag{9}$$

because $g \geq g - r(u_{g2}^m 012\cdots(g-3)(g-2))$. Since $r(u_{g2}) \equiv_g s(u_{g2}) = \frac{g(g+1)}{2} - g + 1 \equiv_g \frac{g}{2}+1$,

$$\begin{aligned} &r(u_{g2}^m 012\cdots(a-2)(a-1)) \equiv_g m\left(\frac{g}{2}+1\right) + \frac{a(a-1)}{2}; \\ &r(u_{g2}^m 012\cdots(g-3)(g-2)) \equiv_g m\left(\frac{g}{2}+1\right) + \frac{(g-2)(g-1)}{2} \equiv_g (m+1)\left(\frac{g}{2}+1\right). \end{aligned}$$

Since $\frac{g}{2}+1$ is even and $-g+2(\frac{g}{2}+1)=2$, $\gcd(g,\frac{g}{2}+1)=2$. Hence,

$$\left\{m\left(\frac{g}{2}+1\right) | m = 0,1,2,\ldots,\frac{g}{2}-1\right\} \equiv_g \{0,2,4,\ldots,g-4,g-2\}.$$

Since $\frac{a(a-1)}{2}$ is even if $a = 4q$ or $4q+1$; odd otherwise,

$$\left\{r\left(u_{g2}^m 01\cdots(a-1)\right) | m = 0,1,\ldots,\frac{g}{2}-1\right\} = \begin{cases} \{0,2,4,\ldots,g-2\}, & \text{if } a = 4q \text{ or } 4q+1; \\ \{1,3,5,\ldots,g-1\}, & \text{otherwise,} \end{cases}$$

$$\left\{r\left(u_{g2}^m 01\cdots(g-2)\right) | m = 0,1,\ldots,\frac{g}{2}-1\right\} = \{0,2,\ldots,g-4,g-2\}.$$

Therefore, by (9), for any digit $a = 0,1,\ldots,g-2$,

$$\begin{aligned} &\left\{a'_{mg+a+1} | m = 0,1,2,\ldots,\frac{g}{2}-1\right\} \\ &\qquad = \begin{cases} \{2k_1, 2k_2+1 | 0 \leq 2k_1 < g-a < 2k_2+1 < g\}, & \text{if } a = 4q \text{ or } 4q+1; \\ \{2k_1+1, 2k_2 | 0 < 2k_1+1 < g-a < 2k_2 \leq g\}, & \text{otherwise;} \end{cases} \\ &\left\{a'_{(m+1)g} | m = 0,1,2,\ldots,\frac{g}{2}-1\right\} = \{1,3,5,\ldots,g-3,g-1\}. \end{aligned}$$

Since $g-1$ is not a digit in u_{g2}, every even digit a' in u_{g3} is obtained by transforming the digit a's in $\{0,1,2,\ldots,g-2\}$, where $a = 4q$ or $4q+1$ with $a < g-a'$; $a = 4q+2$ or $4q+3$ with $a > g-a'$. Every odd digit a' in u_{g3} is obtained by transforming the digit a's in $\{0,1,2,\ldots,g-2,g\}$, where $a = 4q+2$ or $4q+3$ with $a < g-a'$; $a = 4q$ or $4q+1$ with $a > g-a'$; $a = g$. That is,

$$\begin{aligned} \text{if } a' \text{ is even,} \quad |u_{g3}|_{a'} &= |\{\ a|\ && a = 4q, 4q+1 \text{ with } 0 \leq a < g-a'; \\ & && a = 4q+2, 4q+3 \text{ with } g-a' < a \leq g-2 \quad \}|; \\ \text{if } a' \text{ is odd,} \quad |u_{g3}|_{a'} &= |\{\ a|\ && a = 4q+2, 4q+3 \text{ with } 0 \leq a < g-a'; \\ & && a = 4q, 4q+1 \text{ with } g-a' < a < g-1; \\ & && a = g \quad \}|. \end{aligned}$$

Let $g = 4k+2$ for some integer k. Then,

$$|u_{g3}|_0 = |\{4q, 4q+1 | q = 0, 1, 2, \ldots, k\} - \{4k+1\}| = 2k+1;$$
$$|u_{g3}|_1 = |\{4q+2, 4q+3 | q = 0, 1, 2, \ldots, k-1\} \cup \{4k+2\}| = 2k+1;$$
$$|u_{g3}|_2 = |\{4q, 4q+1 | q = 0, 1, 2, \ldots, k-1\}| = 2k,$$

and for any $p = 1, 2, \ldots, k$,

$$\begin{aligned}
|u_{g3}|_{4p-1} =& |\{4q+2 | q = 0, \ldots, k-p\} \cup \{4q+3 | q = 0, \ldots, k-p-1\} \\
&\cup \{4q | q = k-p+1, \ldots, k\} \cup \{4q+1 | q = k-p+1, \ldots, k-1\} \cup \{4k+2\}|; \\
|u_{g3}|_{4p} =& |\{4q | q = 0, \ldots, k-p\} \cup \{4q+1 | q = 0, \ldots, k-p\} \\
&\cup \{4q+2 | q = k-p+1, \ldots, k-1\} \cup \{4q+3 | q = k-p, \ldots, k-1\}; \\
|u_{g3}|_{4p+1} =& |\{4q+2 | q = 0, \ldots, k-p-1\} \cup \{4q+3 | q = 0, \ldots, k-p-1\} \\
&\cup \{4q | q = k-p+1, \ldots, k\} \cup \{4q+1 | q = k-p+1, \ldots, k-1\} \cup \{4k+2\}|; \\
|u_{g3}|_{4p+2} =& |\{4q | q = 0, \ldots, k-p-1\} \cup \{4q+1 | q = 0, \ldots, k-p-1\} \\
&\cup \{4q+2 | q = k-p, \ldots, k-1\} \cup \{4q+3 | q = k-p, \ldots, k-1\}.
\end{aligned}$$

Hence, $|u_{g3}|_{4p-1} = |u_{g3}|_{4p} = 2k+1$ and $|u_{g3}|_{4p+1} = |u_{g3}|_{4p+2} = 2k$.

Observation 27. Let $g = 4k+2$ for some integer $k \geq 0$. For any digit $a \in \{0, 1, 2, \ldots, g-1, g\}$,

$$|u_{g3}|_a = \begin{cases} \frac{g}{2}, & \text{if } a = 0, 1, 4p-1, 4p; \\ \frac{g}{2} - 1, & \text{if } a = 2, 4p+1, 4p+2, \end{cases} \quad \text{for } p = 1, 2, \ldots, k.$$

Lemma 28. *For any $g \geq 2$,*

$$\gcd([u_{g3}]_b, g) = \begin{cases} 1, & \text{if } g \not\equiv 6 \pmod 8 \\ 2, & \text{if } g \equiv 6 \pmod 8. \end{cases}$$

Proof. The case when $g \not\equiv 2 \pmod 4$ is shown in (8). Hence, assume $g \equiv 2 \pmod 4$. Let $g = 4k+2$ for some k. Then, by Observation 27,

$$\begin{aligned}
\sum_{a=0}^{g} a|u_{g3}|_b &= \Big(0 + 1 + \sum_{p=1}^{k}(4p - 1 + 4p)\Big) \cdot \frac{g}{2} + \Big(2 + \sum_{p=1}^{k}(4p+1+4p+2)\Big) \cdot \Big(\frac{g}{2} - 1\Big) \\
&= \frac{g}{2} + g \cdot \Big(1 + \sum_{p=1}^{k}(8p+1)\Big) - 2 - \sum_{p=1}^{k}(8p+3) \\
&\equiv_g (2k+1) - 2 - 4k(k+1) - 3k = -4k^2 - 5k - 1 \equiv_g -3k - 1 \equiv_g k+1.
\end{aligned}$$

Since $[u_{g3}]_b \equiv_g s(u_{g3})$, the number $[u_{g3}]_b \equiv_g k+1$. Then, $\gcd([u_{g3}]_b, 2(2k+1))) = 1$ if k is even; 2 if k is odd, since $\gcd(k+1, 2k+1) = 1$. That is,

$$\gcd([u_{g3}]_b, g) = \begin{cases} 1, & \text{if } g \equiv 2 \pmod 8; \\ 2, & \text{if } g \equiv 6 \pmod 8. \end{cases}$$

□

If $\gcd([u_{g3}]_b, g) = 1$, f_g transforms the string u_{g3}^g to u_{g4}. Hence,

$$u_{g4} = f_g(u_{g3}^g), \text{ if } g \not\equiv 6 \pmod 8. \tag{10}$$

We will focus on the case when $g \not\equiv 6 \pmod 8$ from now on, because it is relatively easy to generate u_{gn} for all $n \geq 4$.

Theorem 29. *For any positive integer $n \geq 3$, if $g \not\equiv 6 \pmod 8$,*

(1) $\gcd([u_{gn}]_b, g) = 1$*;*

(2) $u_{g,n+1} = f_g(u_{gn}^g)$*;*

(3) $|u_{gn}| = g^{n-1}$ *if* $g \not\equiv 2 \pmod 4$*;* $g^{n-1}/2$ *if* $g \equiv 2 \pmod 8$*.*

Proof. The proof is done by mathematical induction on n. The base case is shown in Lemma 28, (10), and Observation 26. Assume (1), (2), and (3) are true for all $n-1 \geq 3$:

$$\gcd([u_{g,n-1}]_b, g) = 1; \qquad u_{gn} = f_g(u_{g,n-1}^g); \qquad |u_{g,n-1}| = \begin{cases} g^{n-2}, & \text{if } g \not\equiv 2 \pmod 4; \\ g^{n-2}/2, & \text{if } g \equiv 2 \pmod 8. \end{cases}$$

Since g is even, $|u_{g,n-1}|$ is even. Hence, there is no such case that g is even and $|u_{g,n-1}|$ is odd. Since $\gcd([u_{g,n-1}]_b, g) = 1$, the number $[f_g(u_{g,n-1}^g)]_b \equiv [u_{g,n-1}]_b \pmod g$ by Corollary 24. Since $u_{gn} = f_g(u_{g,n-1}^g)$, (1) holds.

Since u_{gn} is the shortest repeating string in the string $f_g^n(10^N)$, the shortest repeating string in $f_g^{n+1}(10^N)$ should be $f_g(u_{gn}^h)$ for some integer h. Since $\gcd([u_{gn}]_b, g) = 1$ by (1), the string $f_g(u_{gn}^g)$ repeats in $f_g^{n+1}(10^N)$ by Corollary 21: for some string t,

$$f_g^{n+1}(10^N) = 0 \left(f_g(u_{gn}^g)\right)^{\left\lfloor \frac{N}{g|u_{gn}|} \right\rfloor} t.$$

For any $0 < h < g$, the number $[u_{gn}^h]_b \equiv_g s(u_{gn}^h) = h \cdot s(u_{gn}) \equiv_g h \cdot [u_{gn}]_b \not\equiv 0 \pmod g$, since $\gcd([u_{gn}]_b, g) = 1$. Hence, the string $f_g(u_{gn}^h)$ cannot repeat in $f_g^{n+1}(10^N)$ by Theorem 16, if $h < g$. Therefore, $f_g(u_{gn}^g)$ is the shortest repeating string in $f_g^{n+1}(10^N)$. Hence, (2) holds.

By Induction hypothesis, $|u_{gn}| = |f_g(u_{g,n-1}^g)| = |u_{g,n-1}^g| = g|u_{g,n-1}|$, so (3) holds. □

Note 30. For any $g \not\equiv 2 \pmod 4$, Theorem 29 holds for all $n \geq 1$.

5 The number of digits in b-ary g-Collatz sequences

Let's count the number of occurrences of each digit in the string u_{gn} for $g \not\equiv 6 \pmod 8$. First, we simplify the notation.

Definition 31. For any positive integer n and any digit $a \in \{0, 1, 2, \ldots, g\}$,

$$a_{gn} := \begin{cases} |u_{g,n+1}|_a, & \text{if } g \not\equiv 2 \pmod 4; \\ |u_{g,n+2}|_a, & \text{if } g \equiv 2 \pmod 8. \end{cases}$$

For example, a_{gn} for $g = 2, 3, 4, 10$ is as follows.

n	0_{2n}	1_{2n}	2_{2n}
1	1	1	0
2	2	1	1
3	3	3	2
4	6	5	5
5	11	11	10

n	0_{3n}	1_{3n}	2_{3n}	3_{3n}
1	1	1	0	1
2	2	3	2	2
3	7	7	6	7
4	20	21	20	20
5	61	61	60	61

n	0_{4n}	1_{4n}	2_{4n}	3_{4n}	4_{4n}
1	1	1	1	0	1
2	3	4	3	3	3
3	13	13	13	12	13
4	51	52	51	51	51
5	205	205	205	204	205

n	$0_{10,n}$	$1_{10,n}$	$2_{10,n}$	$3_{10,n}$	$4_{10,n}$	$5_{10,n}$	$6_{10,n}$	$7_{10,n}$	$8_{10,n}$	$9_{10,n}$	$10_{10,n}$
1	5	5	4	5	5	4	4	5	5	4	4
2	46	46	45	45	46	46	45	45	46	45	45
3	455	455	454	455	455	454	454	455	455	454	454
4	4546	4546	4545	4545	4546	4546	4545	4545	4546	4545	4545
5	45455	45455	45454	45455	45455	45454	45454	45455	45455	45454	45454

Lemma 32. *For any positive integer n,*

(1) $\sum_{a=0}^{g} a_{gn} = g^n$ *if* $g \not\equiv 2 \pmod 4$*;* $\frac{g^{n+1}}{2}$ *if* $g \equiv 2 \pmod 8$*;*

(2) $a_{g,n+1} = \sum_{a' \neq g-a} a'_{gn}$ *for any* $g \not\equiv 6 \pmod 8$*.*

Proof. Theorem 29 (3) and Note 30 provide (1), and Corollary 23 provides (2). □

Lemma 33. *Suppose $g \not\equiv 2 \pmod 4$. Then,*

$$\textit{if } n \textit{ is odd, } a_{gn} = \begin{cases} 0_{gn}, & \textit{if } a \neq g-1; \\ 0_{gn} - 1, & \textit{if } a = g-1, \end{cases} \textit{ and if } n \textit{ is even, } a_{gn} = \begin{cases} 0_{gn}, & \textit{if } a \neq 1; \\ 0_{gn} + 1, & \textit{if } a = 1. \end{cases}$$

Proof. The proof is done by mathematical induction. The base case is obvious by Observation 26: $a_{g1} = 1$ if $a \neq g-1$; 0 if $a = g-1$, since $u_{g2} = 012\cdots(g-2)g$. Induction hypothesis: suppose it is true for all $n-1$. If n is even, $n-1$ is odd. By induction hypothesis, $a_{g,n-1} = 0_{g,n-1}$ for all $a \neq g-1$ and $(g-1)_{gn} = 0_{g,n-1} - 1$. Since $a_{gn} = \sum_{a' \neq g-a} a'_{g,n-1}$ by Lemma 32 (2), $a_{gn} = (g-1)0_{g,n-1} + (g-1)_{g,n-1} = g \cdot 0_{g,n-1} - 1$ for all $a \neq 1$ and $1_{gn} = g \cdot 0_{g,n-1}$. Hence, $a_{gn} = 0_{gn}$ for all $a \neq 1$ and $1_{gn} = 0_{gn} + 1$. Similarly, we can prove the case when n is odd. □

Lemma 34. *For any $g = 8k+2$ for some integer $k \geq 1$,*

$$\textit{if } n \textit{ is odd, } a_{gn} = \begin{cases} 0_{gn}, & \textit{if } a = 0, 1, 4p-1, 4p; \\ 0_{gn} - 1, & \textit{if } a = 2, 4p+1, 4p+2, \end{cases} \textit{ for } p = 1, 2, \ldots, 2k;$$

$$\textit{if } n \textit{ is even, } a_{gn} = \begin{cases} g_{gn} + 1, & \textit{if } a = 4q, 4q+1, g-2; \\ g_{gn}, & \textit{if } a = 4q+2, 4q+3, g-1, g, \end{cases} \textit{ for } q = 0, 1, \ldots, 2k-1.$$

Proof. The proof is done by mathematical induction. The base case is shown in Observation 27. Induction Hypothesis: suppose it is true for all $n-1$. If n is even, $n-1$ is odd. By induction hypothesis, $a_{g,n-1} = 0_{g,n-1}$ if $a = 0, 1, 4p-1, 4p$ and $a_{g,n-1} = 0_{g,n-1} - 1$ if $a = 2, 4p+1, 4p+2$ for $p = 1, 2, \ldots, 2k$. By Lemma 32 (2), $a_{gn} = (\frac{g}{2}+1)\cdot 0_{g,n-1} + (\frac{g}{2}-1)\cdot(0_{g,n-1}-1) = g\cdot 0_{g,n-1} - \frac{g}{2} + 1$ if $g - a = 2, 4p+1, 4p+2$; $\frac{g}{2}\cdot 0_{g,n-1} + \frac{g}{2}\cdot(0_{g,n-1}-1) = g\cdot 0_{g,n-1} - \frac{g}{2}$ otherwise. Since $g = 8k+2$, the number $g - a = 2, 4p+1, 4p+2$ implies $a = g-2, 4(2k-p)+1, 4(2k-p)$. Since $p = 1, 2, \ldots, 2k$, the number $2k-p = 0, 1, \ldots, 2k-1$. Hence, $a_{gn} = 0_{gn}$ if $a = 4q, 4q+1, g-2$; $0_{gn}-1$ otherwise, for $q = 0, 1, 2, \ldots, 2k-1$. Similarly, we can prove the case when n is odd. □

When $g = 2$, we have a different arrangement for a_{gn}[1].

Note 35.

$$\text{If } n \text{ is odd, } a_{2n} = \begin{cases} 0_{2n}, & \text{if } a = 0, 1; \\ 0_{2n} - 1, & \text{if } a = 2, \end{cases} \text{ and if } n \text{ is even, } a_{2n} = \begin{cases} g_{2n}, & \text{if } a = 1, 2; \\ g_{2n} + 1, & \text{if } a = 0. \end{cases}$$

Corollary 36. *For any positive integer n, if $g \not\equiv 2 \pmod 4$,*

(1) *if n is odd, $0_n = g\cdot 0_{g,n-1} + 1$; $(g-1)_{gn} = g\cdot 0_{g,n-1}$;*

(2) *if n is even, $0_n = g\cdot 0_{g,n-1} - 1$; $1_{gn} = g\cdot 0_{g,n-1}$.*

Proof. If n is odd, $n-1$ is even. By Lemma 33, $a_{g,n-1} = 0_{g,n-1}$ for all $a \neq 1$ and $1_{g,n-1} = 0_{g,n-1} + 1$. Then, by Lemma 32 (2),

$$\begin{aligned} 0_{gn} &= 0_{g,n-1} \quad + 1_{g,n-1} + 2_{g,n-1} + \cdots + (g-2)_{g,n-1} + (g-1)_{g,n-1}; \\ (g-1)_{gn} &= 0_{g,n-1} \qquad\qquad\quad + 2_{g,n-1} + \cdots + (g-2)_{g,n-1} + (g-1)_{g,n-1} + g_{g,n-1}. \end{aligned}$$

Hence, (1) is obtained. Similarly, we can prove (2). □

Corollary 37. *For any positive integer n, if $g \equiv 2 \pmod 8$,*

(1) *if n is odd, $0_n = \frac{g}{2}0_{n-1} + \frac{g}{2}g_{n-1}$; $g_n = (\frac{g}{2}-1)0_{n-1} + (\frac{g}{2}+1)g_{n-1}$;*

(2) *if n is even, $g_n = \frac{g}{2}0_{n-1} + \frac{g}{2}g_{n-1}$; $0_n = (\frac{g}{2}-1)g_{n-1} + (\frac{g}{2}+1)0_{n-1}$.*

Proof. If n is odd, $n-1$ is even. By Lemma 34 and Note 35, $a_{g,n-1} = 0_{g,n-1}$ if $a = 4q, 4q+1, 2$ and $a_{g,n-1} = g_{g,n-1}$ if $a = 4q+2, 4q+3, g-1, g$ for any $q = 0, 1, \ldots, , 2k-1$. By Lemma 32 (2),

$$0_{gn} = 0_{g,n-1} + 1_{g,n-1} + \cdots + (g-1)_{g,n-1} \text{ and } g_{gn} = 1_{g,n-1} + 2_{g,n-1} + \cdots + g_{g,n-1}.$$

Hence, (1) is obtained. Similarly, we can prove (2). □

Corollary 38. *For any positive integer n,*

$$\begin{array}{rll} & \text{if } g \not\equiv 2 \pmod 4, & |\{a \in \{0,1,2,\ldots,g\} | a_{gn} = 0_{gn} = g_{gn}\}| = g; \\ & & |\{a \in \{0,1,2,\ldots,g\} | a_{gn} = 0_{gn} + (-1)^n\}| = 1, \\ \text{and} & \text{if } g \equiv 2 \pmod 8, & |\{a \in \{0,1,2,\ldots,g\} | a_{gn} = 0_{gn}\}| = \frac{g}{2} + \frac{1-(-1)^n}{2}; \\ & & |\{a \in \{0,1,2,\ldots,g\} | a_{gn} = g_{gn}\}| = \frac{g}{2} + \frac{1+(-1)^n}{2}. \end{array}$$

Proof. It is obtained by sorting and counting a_{gn}'s using Lemma 33, 34 and Note 35. □

Even if there are b distinct digits in a b-ary string, there are only two distinct values to identify a_{gn}, for any g and n. If $g \not\equiv 2 \pmod 4$, there are g many a_{gn}'s with $a_{gn} = 0_{gn} = g_{gn}$ and only one a_{gn} with $a_{gn} \neq 0_{gn}$. If $g \equiv 2 \pmod 8$, there are $\frac{g}{2}+1$ many a_{gn}'s with $a_{gn} = 0_{gn}$ (g_{gn} for even n) and $\frac{g}{2}$ many a_{gn}'s with $a_{gn} \neq 0_{gn}$ (g_{gn} for even n) for odd n. Hence, we consider the majority of a_{gn}'s and the minority of a_{gn}'s as follows:

Definition 39. For any $g \not\equiv 6 \pmod 8$ and any positive integer n,

$$M_{gn} = \begin{cases} 0_{gn}, & \text{if } n \text{ is odd;} \\ g_{gn}, & \text{if } n \text{ is even,} \end{cases} \quad \text{and } m_{gn} = M_{gn} + (-1)^n.$$

For example, the following shows the first few M_{gn} and m_{gn} for $g = 2, 3, 4, 5, 7, 10$.

n	1	2	3	4	5	6
M_{2n}	1	1	3	5	11	21
M_{3n}	1	2	7	20	61	182
M_{4n}	1	3	13	51	205	819
M_{5n}	1	4	21	104	521	2604
M_{7n}	1	6	43	300	2101	14706
$M_{10,n}$	5	45	455	4545	45455	454545

n	1	2	3	4	5	6
m_{2n}	0	2	2	6	10	22
m_{3n}	0	3	6	21	60	183
m_{4n}	0	4	12	52	204	820
m_{5n}	0	5	20	105	520	2605
m_{7n}	0	7	42	301	2100	14707
$m_{10,n}$	4	46	454	4546	45454	454546

Then, we have a recurrence relation as follows:

Theorem 40. *For any $g \not\equiv 6 \pmod 8$ and any integer $n \geq 3$,*

$$M_{gn} = (g-1)M_{g,n-1} + gM_{g,n-2}.$$

Proof. By Lemma 36 and 37, for any positive integer n, whether n is even or odd,

if $g \not\equiv 2 \pmod 4$, $M_{gn} = gM_{g,n-1} + (-1)^{n-1}$ and $m_{gn} = gM_{g,n-1}$;
if $g \equiv 2 \pmod 8$, $M_{gn} = \dfrac{g}{2}M_{g,n-1} + \dfrac{g}{2}m_{g,n-1}$ and $m_{gn} = \left(\dfrac{g}{2}+1\right)M_{g,n-1} + \left(\dfrac{g}{2}-1\right)m_{g,n-1}$.

Hence, if $g \not\equiv 2 \pmod 4$,

$$M_{gn} = (g-1)M_{g,n-1} + M_{g,n-1} + (-1)^{n-1} = (g-1)M_{g,n-1} + m_{g,n-1} = (g-1)M_{g,n-1} + gM_{g,n-2},$$

and if $g \equiv 2 \pmod 8$,

$$\begin{aligned}
M_{gn} &= \frac{g}{2}M_{g,n-1} + \frac{g}{2}m_{g,n-1} \\
&= \frac{g}{2}\left(\frac{g}{2}M_{g,n-2} + \frac{g}{2}m_{g,n-2}\right) + \frac{g}{2}\Bigg(\left(\frac{g}{2}+1\right)M_{g,n-2} + \left(\frac{g}{2}-1\right)m_{g,n-2}\Bigg) \\
&= \frac{g^2+g}{2}\cdot M_{g,n-2} + \frac{g^2-g}{2}\cdot m_{g,n-2} \\
&= \frac{g^2-g}{2}\left(M_{g,n-2} + m_{g,n-2}\right) + gM_{g,n-2} \\
&= (g-1)\frac{g}{2}\left(M_{g,n-2} + m_{g,n-2}\right) + gM_{g,n-2} = (g-1)M_{g,n-1} + gM_{g,n-2}.
\end{aligned}$$

□

Furthermore, the explicit formulae are calculated as follows:

Theorem 41. *For any nonnegative integer n,*

$$\text{if } g \not\equiv 2 \pmod 4,\ M_{gn} = \frac{g^n - (-1)^n}{b},\ \text{and if } g \equiv 2 \pmod 8,\ M_{gn} = \frac{g}{2}\cdot\frac{g^n-(-1)^n}{b}.$$

Proof. By Lemma 32 (1), Corollary 38, and Definition 39,

if $g \not\equiv 2 \pmod 4$, $gM_{gn} + m_{gn} = gM_{gn} + M_{gn} + (-1)^n = g^n$;

if $g \equiv 2 \pmod 8$, $\left(\frac{g}{2}+1\right)M_{gn} + \frac{g}{2}m_{gn} = \left(\frac{g}{2}+1\right)M_{gn} + \frac{g}{2}\Big(M_{gn} + (-1)^n\Big) = \frac{g^{n+1}}{2}$.

Since $b = g+1$, the explicit formulae are calculated as desired. □

Notice that $M_{g1} = 1$ for all $g \not\equiv 2 \pmod 4$ and $M_{g1} = \frac{g}{2}$ for all $g \equiv 2 \pmod 8$.

6 Generalized Jacobsthal numbers

For any nonnegative integer n, the n-th Jacobsthal number, J_n (A001045), and the nth almost Jacobsthal numbers, A_n (A005578) and B_n (A000975), [6] are defined as follows:

$$J_n = \frac{2^n - (-1)^n}{3}; \qquad A_n = \left\lceil \frac{2^n}{3} \right\rceil; \qquad B_n = \left\lceil \frac{2^n-2}{3} \right\rceil, \tag{11}$$

and they satisfy the following [1]: for any positive number n,

$$J_n = 1_{2n}, \qquad A_n = 0_{2n}, \qquad B_n = 2_{2n}, \qquad J_n + A_n + B_n = 2^n. \tag{12}$$

In order to generalize (12), we first generalize the sequences in (11).

Definition 42. For any integer $g \geq 2$ and any nonnegative integer n, the nth g-Jacobsthal number, J_{gn}, and the nth almost g-Jacobsthal numbers, A_{gn} and B_{gn}, are defined as

$$J_{gn} = \frac{g^n - (-1)^n}{g+1}; \qquad A_{gn} := \left\lceil \frac{g^n}{g+1} \right\rceil; \qquad B_n := \left\lceil \frac{g^n - g}{g+1} \right\rceil.$$

The g-Jacobsthal numbers $(J_{gn})_{g \geq n}$ can be generated by the following recurrence as well.

Lemma 43. *The g-Jacobsthal numbers $(J_{gn})_{n\geq 0}$ satisfy the following:*

(1) $J_{g0} = 0, J_{g1} = 1, J_{gn} = (g-1)J_{g,n-1} + gJ_{g,n-2}$ *for* $n \geq 2$*;*

(2) $J_{g0} = 0, J_{g,n-1} + J_{g,n} = g^{n-1}$ *for* $n \geq 1$.

Obviously, $J_{2n} = J_n$, $A_{2n} = A_n$, and $B_{2n} = B_n$. The sequence $(A_{3n})_{n\geq 0}$ is identified as A122983, and $(J_{gn})_{n\geq 0}$ for $g = 2, 3, 4, 5, 6$ is identified as follows:

g	2	3	4	5	6
$(J_{gn})_{n\geq 0}$	A001045	A015518	A015521	A015331	A015540

Furthermore, consider $Y_{gn} := J_{gn} + (-1)^n = \frac{g^n + g(-1)^n}{g+1}$. Then, $(Y_{gn})_{n\geq 0}$ for $g = 2, 3, 4, 5, 6$ is identified as follows:

g	2	3	4	5	6
$(Y_{gn})_{n\geq 0}$	A078008	A0054878	A109499	A109500	A109501

The following shows the first few J_{gn}'s.

n	0	1	2	3	4	5	6	7
J_{2n}	0	1	1	3	5	11	21	43
J_{3n}	0	1	2	7	20	61	182	547
J_{4n}	0	1	3	13	51	205	819	3277
J_{5n}	0	1	4	21	104	521	2604	13021
J_{6n}	0	1	5	31	185	1111	6665	39991
J_{7n}	0	1	6	43	300	2101	14706	102943

Now, we can express M_{gn} and m_{gn} defined in Section 5 in terms of J_{gn}.

Theorem 44. *For any positive integer n,*

(1) *if* $g \not\equiv 2 \pmod 4$, $M_{gn} = J_{gn}$ *and* $m_{gn} = J_{gn} + (-1)^n$*;*

(2) *if* $g \equiv 2 \pmod 8$, $M_{gn} = \frac{g}{2}J_{gn}$ *and* $m_{gn} = \frac{g}{2}J_{gn} + (-1)^n$.

Proof. It is obtained by Definition 39, Theorem 41, and Definition 42. □

Hence, we can express all a_{gn}'s defined in Section 5 in terms of J_{gn}.

Corollary 45. *For any positive integer n, if $g \not\equiv 2 \pmod 4$,*

$$a_{g,2n-1} = \begin{cases} J_{g,2n-1}, & \text{if } a \neq g-1; \\ J_{g,2n-1} - 1, & \text{if } a = g-1 \end{cases} \quad \text{and} \quad a_{g,2n} = \begin{cases} J_{g,2n}, & \text{if } a \neq 1; \\ J_{g,2n} + 1, & \text{if } a = 1, \end{cases}$$

and if $g = 8k+2$ with $k \geq 1$,

$$a_{g,2n-1} = \begin{cases} \frac{g}{2} J_{g,2n-1}, & \text{if } a = 0, 1, 4p-1, 4p; \\ \frac{g}{2} J_{g,2n-1} - 1, & \text{if } a = 2, 4p+1, 4p+2, \end{cases} \quad \text{for } p = 1, 2, \ldots, 2k;$$

$$a_{g,2n} = \begin{cases} \frac{g}{2} J_{g,2n}, & \text{if } a = 4q+2, 4q+3, g-1, g; \\ \frac{g}{2} J_{g,2n} + 1, & \text{if } a = 4q, 4q+1, g-2, \end{cases} \quad \text{for } q = 0, 1, \ldots, 2k-1.$$

Proof. It is obtained by Theorem 44, Definition 39, Lemma 33, and Lemma 34. □

Now, let us compare A_{gn} and B_{gn} with J_{gn}. The following shows the first few A_{gn}, J_{gn}, and B_{gn}, when $g = 3$ and 4.

n	0	1	2	3	4	5	6
A_{3n}	1	1	3	7	21	61	183
J_{3n}	0	1	2	7	20	61	182
B_{3n}	0	0	2	6	20	60	182

n	0	1	2	3	4	5	6
A_{4n}	1	1	4	13	52	205	820
J_{4n}	0	1	3	13	51	205	819
B_{4n}	0	0	3	12	51	204	819

We can identify both A_{gn} and B_{gn} as J_{gn} or $J_{gn} \pm 1$.

Lemma 46. *For any nonnegative integer n,*

$$A_{gn} = \begin{cases} J_{gn}, & \text{if } n \text{ is odd;} \\ J_{gn} + 1, & \text{if } n \text{ is even.} \end{cases} \qquad B_{gn} = \begin{cases} J_{gn} - 1, & \text{if } n \text{ is odd;} \\ J_{gn}, & \text{if } n \text{ is even.} \end{cases}$$

Proof. Since $g^n = ((g+1)-1)^n = \sum_{i=0}^{n-1} \binom{n}{i} (g+1)^{n-i}(-1)^i + (-1)^n$, the number $g^n - (-1)^n$ is divisible by $g+1$. Hence,

$$\left\lceil \frac{g^n}{g+1} \right\rceil = \frac{g^n - (-1)^n}{g+1} + \left\lceil \frac{(-1)^n}{g+1} \right\rceil = J_{gn} + \begin{cases} 0, & \text{if } n \text{ is odd;} \\ 1, & \text{if } n \text{ is even.} \end{cases}$$

and

$$\left\lceil \frac{g^n - g}{g+1} \right\rceil = \frac{g^n - (-1)^n}{g+1} + \left\lceil \frac{(-1)^n - g}{g+1} \right\rceil = J_{gn} + \begin{cases} -1, & \text{if } n \text{ is odd;} \\ 0, & \text{if } n \text{ is even.} \end{cases}$$

□

Therefore, we can generalize (12), when $g \not\equiv 2 \pmod 4$.

Theorem 47. *For any positive integer n, if $g \not\equiv 2 \pmod 4$,*

$$a_{gn} = \begin{cases} J_{gn}, & \text{if } a \neq 1, g-1; \\ A_{gn}, & \text{if } a = 1; \\ B_{gn}, & \text{if } a = g-1, \end{cases} \quad \text{and } (g-1)J_{gn} + A_{gn} + B_{gn} = g^n.$$

Proof. Corollary 45 and Lemma 46 identify a_{gn}, and Lemma 32 (1) provides the relation. □

To work on the case when $g \equiv 2 \pmod 8$, we express A_{gn} and B_{gn} more explicitly.

Corollary 48. *For any nonnegative integer n,*

$$A_{gn} = \frac{g^n + (\frac{g+1}{2} - 1)(-1)^n + \frac{g+1}{2}}{g+1} \text{ and } B_{gn} = \frac{g^n + (\frac{g+1}{2} - 1)(-1)^n - \frac{g+1}{2}}{g+1}. \tag{13}$$

Proof. Since $\frac{(-1)^n+1}{2} = 0$ for odd n; 1 for even n, and $\frac{(-1)^n-1}{2} = -1$ for odd n; 0 for even n,

$$A_{gn} = \frac{g^n - (-1)^n}{g+1} + \frac{(-1)^n + 1}{2} \text{ and } B_{gn} = \frac{g^n - (-1)^n}{g+1} + \frac{(-1)^n - 1}{2},$$

which is modified to (13). □

Notice that A_n and B_n are as follows:

$$A_n = \frac{g^n + (\frac{3}{2} - 1)(-1)^n + \frac{3}{2}}{3} \text{ and } B_n = \frac{g^n + (\frac{3}{2} - 1)(-1)^n - \frac{3}{2}}{3}. \tag{14}$$

We can obtain (13) from (14) by replacing $\frac{3}{2}$ with $\frac{g+1}{2}$. However, we can consider replacing $\frac{3}{2}$ with $\frac{g+1}{g}$. If we consider $J_{2n} = \frac{2}{2} \cdot J_n$, $A_{2n} = \frac{2}{2} \cdot A_n$, and $B_{2n} = \frac{2}{2} \cdot B_n$, we can also consider replacing $\frac{2}{2}$ with $\frac{g}{2}$. Taking all of these new considerations, we now have the following:

$$\frac{g}{2} \cdot \frac{g^n - (-1)^n}{g+1}; \quad \frac{g}{2} \cdot \frac{g^n + (\frac{g+1}{g} - 1)(-1)^n + \frac{g+1}{g}}{g+1}; \quad \frac{g}{2} \cdot \frac{g^n + (\frac{g+1}{g} - 1)(-1)^n - \frac{g+1}{g}}{g+1}. \tag{15}$$

By simplifying (15), we have another generalization of (11).

Definition 49. For any integer $g \geq 2$ any nonnegative integer n, we define J'_{gn}, A'_{gn}, and B'_{gn} as follows:

$$J'_{gn} = \frac{g^{n+1} - g(-1)^n}{2(g+1)}; \qquad A'_{gn} = \left\lceil \frac{g^{n+1}}{2(g+1)} \right\rceil; \qquad B'_{gn} = \left\lceil \frac{g^{n+1} - g - 2}{2(g+1)} \right\rceil.$$

Lemma 50. *For any nonnegative integer n,*

$$J'_{gn} = \frac{g}{2} J_{gn}; \quad A'_{gn} = J'_{gn} + \begin{cases} 0, & \text{if } n \text{ is odd;} \\ 1, & \text{otherwise} \end{cases}; \quad B'_{gn} = J'_{gn} + \begin{cases} -1, & \text{if } n \text{ is odd;} \\ 0, & \text{otherwise} \end{cases}$$

Proof. By Definition 42 and 49, it is obvious that $J'_{gn} = \frac{g}{2} J_{gn}$. Since

$$A'_{gn} = \frac{g^{n+1} - g(-1)^n}{2(g+1)} + \left\lceil \frac{g(-1)^n}{2(g+1)} \right\rceil \text{ and } B'_{gn} = \frac{g^{n+1} - g(-1)^n}{2(g+1)} + \left\lceil \frac{g(-1)^n - g - 2}{2(g+1)} \right\rceil,$$

A'_{gn} and B'_{gn} are shown as desired. □

Theorem 51. *For any positive integer n, if $g = 8k+2$ for some integer $k \geq 1$,*

$$a_{gn} = \begin{cases} J'_{gn}, & \text{if } a = 4p-1 \text{ for } p = 1, 2, \dots, 2k; \\ A'_{gn}, & \text{if } a = 0, 1, 4p \text{ for } p = 1, 2, \dots, 2k; \\ B'_{gn}, & \text{if } a = 4p-2, g-1, g \text{ for } p = 1, 2, \dots, 2k; \\ J'_{gn} + (-1)^n, & \text{if } a = 4p+1 \text{ for } p = 1, 2, \dots, 2k-1, \end{cases}$$

and

$$(4k-1)J'_{gn} + (2k+2)A'_{gn} + (2k+2)B'_{gn} = \frac{g^{n+1}}{2} - (2k-1)(-1)^n.$$

Proof. Lemma 50 and Corollary 45 identify a_{gn}, and Lemma 32 (1) provides

$$2k \cdot J'_{gn} + (2k+2)A'_{gn} + (2k+2)B'_{gn} + (2k-1)\left(J'_{gn} + (-1)^n\right) = \frac{g^{n+1}}{2}.$$

□

7 Acknowledgments

I would like to thank the reviewer for his/her suggestions, which helped to improve this paper.

References

[1] J. Choi, Ternary modified Collatz sequences and Jacobsthal numbers, *J. Integer Seq.* **19** (2016) 16.75.

[2] L. Collatz, On the motivation and origin of the $(3n+1)$ problem, *J. Qufu Normal University, Natural Science Edition* **3** (1986) 9–11.

[3] J. Conway, Unpredictable iterations, *Conf. Proc. Number Theory, Univ. of Colorado, Boulder*(1972) 49-52.

[4] R. Terras, A stopping time problem on the positive integers, *Acta Arith.* **30** (1976) 241–252.

[5] J. Shallit, *A Second Course in Formal Languages and Automata Theory*, Cambridge University Press, 2009.

[6] N. J. A. Sloane, Online Encyclopedia of Integer Sequences, `http://oeis.org`.

2010 *Mathematics Subject Classification*: Primary 11A67; Secondary 11B75.
Keywords: Collatz problem, Jacobsthal number, base-b representation, $3n + 1$ problem.

(Concerned with sequences A000975, A001045, A005578, A015331, A015518, A015521, A015540, A054878, A078008, A109499, A109500, A109501, and A122983.)

Received November 1 2017; revised version received March 29 2018; May 11 2018. Published in *Journal of Integer Sequences*, May 25 2018.

Return to Journal of Integer Sequences home page.