Observations Regarding the Collatz Conjecture

Daniel Strecker, daniel.strecker@gmx.net

2022-07-22

Document version: 2022-07-23

1 Abstract

This is a loose collection of definitions and formulas useful when researching the Collatz conjecture. There is nothing new in this document. All of the ideas in here have been discovered by others earlier.

2 Collatz Functions and Sequences

Definition 2.1 (Original Collatz Function)**.** The original Collatz function (or original Collatz map, or simply "the Collatz map") is

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

Definition 2.2 (Original Collatz Sequence / OCS)**.** A Collatz sequence, or original Collatz sequence (OCS), is the sequence $(C^i(n_0))_{i \in \mathbb{N}_0}$, where $n_0 \in \mathbb{N}$ is the starting value and C^i is the i-times iterated function C.

Definition 2.3 (Shortcut Collatz Function)**.** The shortcut Collatz function is

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

Definition 2.4 (Shortcut Collatz Sequence / SCS)**.** A shortcut Collatz sequence (SCS) is the sequence $(T^i(n_0))_{i \in \mathbb{N}_0}$, where $n_0 \in \mathbb{N}$ is the starting value and T^i is the i-times iterated function T.

Remark 2.5. An SCS is similar to the OCS which starts at the same starting number n_0. The difference of the SCS is that it leaves out the even number that comes directly after every odd number in the OCS. This is because for odd n, the term $3n+1$ always evaluates to an even number, which is always divided by 2 in the next operation. The shortcut Collatz function integrates this division by 2 directly into the odd rule.

Remark 2.6. Note that $T(n)$ does a division by 2 exactly once in every application, so in the iterated function $T^i(n)$ there happen i many divisions by 2.

3 Cycles

Definition 3.1 (Cycle Member in an SCS)**.** The number $n \in \mathbb{N}$ is a member of a cycle in an SCS if, and only if, the sequence starting at n reaches this number again, i.e. iff $\exists m \in \mathbb{N} : T^m(n) = n$.

Remark 3.2. Note that $m = 0$ is excluded by $m \in \mathbb{N}$.

Remark 3.3. For example $T(1) = 2$ and $T(2) = 1$, so $T^2(1) = T(T(1)) = 1$ and $T^2(2) = 2$. Thus $\{1, 2\}$ are members of a cycle in an SCS. This is the only known positive cycle in all SCSes and it is easy to find, thus it is called the trivial Collatz cycle, or simply the trivial cycle.

4 Parity Vector

Definition 4.1 (Parity vector)**.** Let $n \in \mathbb{N}$. Each element v_i of the parity vector $v(n)$ is defined as

$$v_i(n) := \begin{cases} 1 & \text{if } T^i(n) \text{ is odd} \\ 0 & \text{if } T^i(n) \text{ is even} \end{cases} \quad \text{(starting at } i = 0\text{)}.$$

5 Iteration Speedups for Special $n \in \mathbb{N}$

Certain forms of $n \in \mathbb{N}$ allow to skip ahead many iterations at once. Some are shown below.

Lemma 5.1. *Let $n \in \mathbb{N}$ with $n = h \cdot 2^s + (2^s - 1)$, where $h \in \mathbb{N}_0$, $s \in \mathbb{N}$. Then $T(n) = (3h + 2)2^{s-1} + (2^{s-1} - 1)$.*

Proof. Let $n = h \cdot 2^s + (2^s - 1)$, where $h \in \mathbb{N}_0$, $s \in \mathbb{N}$.

Then $n = h \cdot 2^s + 2^s - 1 = 2^s(h + 1) - 1$. Note that n is odd, because 2^s is even and so is $2^s(h + 1)$, so subtracting 1 from this makes the result odd.

$$
\begin{aligned}
T(n) &= \frac{3n+1}{2} \\
&= \frac{3(2^s(h+1)-1)+1}{2} \\
&= \frac{3\cdot 2^s(h+1)-3+1}{2} \\
&= \frac{3\cdot 2^s(h+1)-2}{2} \\
&= \frac{2\cdot(3\cdot 2^{s-1}(h+1)-1)}{2} \\
&= 2^{s-1}3(h+1)-1 \\
&= 2^{s-1}(3h+3)-1 \\
&= 2^{s-1}(3h+2+1)-1 \\
&= 2^{s-1}(3h+2)+2^{s-1}-1 \\
&= (3h+2)2^{s-1}+(2^{s-1}-1)
\end{aligned}
$$

□

Proposition 5.2. Let $n \in \mathbb{N}$ with $n = h \cdot 2^s + (2^s - 1)$, where $h, s \in \mathbb{N}_0$. Then $T^s(n) = h \cdot 3^s + (3^s - 1)$.

Proof. Let $n = h \cdot 2^s + (2^s - 1)$, where $h, s \in \mathbb{N}_0$.

Case by case analysis of the cases $s = 0$ and $s > 0$:

Case $s = 0$:

$$
\begin{aligned}
T^0(n) &= n \\
&= h\cdot 2^s + (2^s - 1) \\
&= h\cdot 2^0 + (2^0 - 1) \\
&= h\cdot 1 + (1 - 1) \\
&= h\cdot 3^0 + (3^0 - 1) \\
&= h\cdot 3^s + (3^s - 1)
\end{aligned}
$$

Case $s > 0$:

$s > 0 \implies s \in \mathbb{N}$, so lemma 5.1 can be applied:

$$
\begin{aligned}
T^s(n) &= T^{s-1}(T(n)) \\
&= T^{s-1}(T(h\cdot 2^s + 2^s - 1)) \\
&= T^{s-1}((3h+2)2^{s-1} + (2^{s-1} - 1))
\end{aligned}
$$

Now with $h' = 3h + 2$, you can apply lemma 5.1 another time and get

$$\begin{aligned} T^s(n) &= T^{s-1}((3h+2)2^{s-1} + (2^{s-1}-1)) \\ &= T^{s-2}((3(3h+2)+2)2^{s-2} + (2^{s-2}-1)) \\ &= T^{s-2}((3^2h + 3^1 \cdot 2 + 3^0 \cdot 2)2^{s-2} + (2^{s-2}-1)) \end{aligned}$$

When applying lemma 5.1 multiple times, a regularity shows as follows.

$$\begin{aligned} T^s(n) &= T^{s-2}((3^2h + 3^1 \cdot 2 + 3^0 \cdot 2)2^{s-2} + (2^{s-2}-1)) \\ &= T^{s-3}(3((3^2h + 3^1 \cdot 2 + 3^0 \cdot 2)+2)2^{s-3} + (2^{s-3}-1)) \\ &= T^{s-3}((3^3h + 3^2 \cdot 2 + 3^1 \cdot 2 + 3^0 \cdot 2)2^{s-3} + (2^{s-3}-1)) \\ &... \\ &= T^{s-s}((3^sh + 3^{s-1} \cdot 2 + 3^{s-2} \cdot 2 + ... + 3^0 \cdot 2)2^{s-s} + (2^{s-s}-1)) \\ &= T^0((3^sh + 2 \cdot (3^{s-1} + 3^{s-2} + ... + 3^0))2^0 + (2^0-1)) \\ &= h \cdot 3^s + 2 \cdot (3^{s-1} + 3^{s-2} + ... + 3^0) \\ &= h \cdot 3^s + 2\sum_{i=0}^{s-1} 3^i \\ &= h \cdot 3^s + 2\frac{3^{(s-1)+1}-1}{3-1} = 3^sh + 2\frac{3^s-1}{2} \\ &= h \cdot 3^s + 3^s - 1 \end{aligned}$$

□

Corollary 5.3. *Let $n \in \mathbb{N}$ with $n = 2^s - 1$, where $s \in \mathbb{N}_0$. Then $T^s(n) = 3^s - 1$.*

Proof. Let $n = 2^s - 1$, where $s \in \mathbb{N}_0$. Then with $h = 0$ we can rewrite this as $n = 2^s - 1 = h \cdot 2^s + (2^s - 1)$ and with proposition 5.2 it follows that $T^s(n) = h \cdot 3^s + (3^s - 1) = 3^s - 1$. □

Proposition 5.4. Let $n \in \mathbb{N}$ with $n = (h \cdot 2^s + (2^s - 1)) \cdot 2^z$, where $h, s, z \in \mathbb{N}_0$. Then $T^{s+z}(n) = h \cdot 3^s + (3^s - 1)$.

Proof. Let $n = (h \cdot 2^s + (2^s - 1)) \cdot 2^z$, where $h, s, z \in \mathbb{N}_0$.

Because n can be divided z times by 2, it follows that

$$T^z(n) = n \cdot \frac{1}{2^z} = (h \cdot 2^s + (2^s - 1)) \cdot 2^z \cdot \frac{1}{2^z} = h \cdot 2^s + (2^s - 1)$$

.

From this follows with proposition 5.2 that

$$\begin{aligned} T^{s+z}(n) &= T^s(T^z(n)) \\ &= T^s(h \cdot 2^s + (2^s - 1)) \\ &= h \cdot 3^s + 3^s - 1 \end{aligned}$$

. □

Corollary 5.5. *Let $n \in \mathbb{N}$ with $n = (2^s - 1) \cdot 2^z$, where $s, z \in \mathbb{N}_0$. Then $T^{s+z}(n) = 3^s - 1$.*

Proof. Let $n = (2^s - 1) \cdot 2^z$, where $s \in \mathbb{N}_0$. Then with $h = 0$ we can reformulate the expression for n as $n = (2^s - 1) \cdot 2^z = (h \cdot 2^s + (2^s - 1)) \cdot 2^z$ and with proposition 5.4 it follows that $T^{s+z}(n) = h \cdot 3^s + (3^s - 1) = 3^s - 1$. □

6 Decomposing T

Let $n \in \mathbb{N}$. Then n is the starting value of the SCS $(T^i(n))_{i \in \mathbb{N}_0}$.

Define

$$a_k := \begin{cases} \frac{3}{2} & \text{if } T^k(n) \text{ is odd} \\ \frac{1}{2} & \text{if } T^k(n) \text{ is even} \end{cases} \text{ (starting at } k = 0\text{)},$$

and

$$b_j := \begin{cases} \frac{1}{2} & \text{if } T^j(n) \text{ is odd} \\ 0 & \text{if } T^j(n) \text{ is even} \end{cases} \text{ (starting at } j = 0\text{)}.$$

Remark 6.1. a_k and b_j are actually functions of not only the indices k and j, but also of the starting value n. So technically it should be $a(n, k)$ and $b(n, j)$. However, for brevity of the following formulas, n is assumed to be fixed and not mentioned as arguments to a and b anymore. Moreover, k and j are used as indices rather than as a function arguments, so parentheses can be widely omitted to help readability.

The value n is givem, hence for $i = 0$, the value $T^0(n) = n$ is given. For all $i > 0$, the value $T^i(n)$ can be expressed as

$$\begin{aligned} T^i(n) &= b_{i\text{-}1} + a_{i\text{-}1} T^{i\text{-}1}(n) \\ &= b_{i\text{-}1} + a_{i\text{-}1}(b_{i\text{-}2} + a_{i\text{-}2} T^{i\text{-}2}(n)) \\ &= b_{i\text{-}1} + a_{i\text{-}1} b_{i\text{-}2} + a_{i\text{-}1} a_{i\text{-}2} T^{i\text{-}2}(n) \\ &= b_{i\text{-}1} + a_{i\text{-}1} b_{i\text{-}2} + a_{i\text{-}1} a_{i\text{-}2}(b_{i\text{-}3} + a_{i\text{-}3} T^{i\text{-}3}(n)) \\ &= b_{i\text{-}1} + a_{i\text{-}1} b_{i\text{-}2} + a_{i\text{-}1} a_{i\text{-}2} b_{i\text{-}3} + a_{i\text{-}1} a_{i\text{-}2} a_{i\text{-}3} T^{i\text{-}3}(n) \\ &= ... \\ &= b_{i\text{-}1} + a_{i\text{-}1} b_{i\text{-}2} + a_{i\text{-}1} a_{i\text{-}2} b_{i\text{-}3} + ... + a_{i\text{-}1} a_{i\text{-}2} ... a_1 b_0 + a_{i\text{-}1} a_{i\text{-}2} ... a_0 T^{i-i}(n) \\ &= b_{i\text{-}1} + a_{i\text{-}1} b_{i\text{-}2} + a_{i\text{-}1} a_{i\text{-}2} b_{i\text{-}3} + ... + a_{i\text{-}1} a_{i\text{-}2} ... a_1 b_0 + a_{i\text{-}1} a_{i\text{-}2} ... a_0 n \end{aligned}$$

.

And ultimately we write

$$T^i(n) = \underbrace{\sum_{j=0}^{i-1}(b_j \prod_{k=j+1}^{i-1} a_k)}_{\beta_i} + (\underbrace{\prod_{k=0}^{i-1} a_k}_{\alpha_i})n \tag{6.1}$$

.

With the definitions

$$\beta_i := \sum_{j=0}^{i-1}(b_j \prod_{k=j+1}^{i-1} a_k) \tag{6.2}$$

and

$$\alpha_i := \prod_{k=0}^{i-1} a_k \tag{6.3}$$

we simplify Eq. 6.1 to

$$T^i(n) = \beta_i + \alpha_i \cdot n \tag{6.4}$$

.

7 Notation

$\mathbb{N}$	natural numbers, not including 0
$\mathbb{N}_0$	natural numbers, including 0