A terminating, non-cyclic* path towards the Collatz conjecture

Jon Seymour <a_beautiful_k@wildducktheories.com>

14 October 2023

Abstract

"*Mathematics may not be ready for such problems*" - Paul Erdős on the Collatz conjecture.

It has long been known that the following identity applies to all paths of the Collatz sequence.

$$2^e x_n - 3^o x_0 = k$$

where x_0 is the initial term and x_n is the final term and e, o and k are path dependent parameters.

This paper derives a formula for k derived from 3 parameter sequences m_n, o_n, e_n.

$$m_n = x_n \pmod 2$$

$$o_n = \sum_{k=0}^{k=n} m_k$$

$$e_n = n + 1 - o_n$$

$$k_n = 2^{e_{n-1}} x_n - 3^{o_{n-1}} x_0 = \sum_{i=0}^{i=n-1} 2^{e_i} 3^{o_{n-1} - o_i} m_i$$

The resultant identity is not novel, for example see [1], but perhaps the technique by which it was derived may be interesting to some.

We also derive a recurrence relation that expresses each k_n in terms of k_{n-1}

$$k_n = 3^{m_{n-1}} k_{n-1} + 2^{e_{n-1}} m_{n-1}$$

and show that $k_0 = 0$

[1] First odd term of the sequence lower odd number n related to the $3 \cdot n + 1$ (https://mathoverflow.net/questions/448397/first-odd-term-of-the-sequence-lower-odd-number-n-related-to-the-3-cdot-n1?_gl=1*wy9ddp*_ga*MTAyNzkyMTgxNy4xNjYyNzY2MzQw*_ga_S812YQPLT2*MTY5NzI4ND on Maths Overflow Net (https://mathoverflow.net/)

A useful recurrence relation

This paper examines a recurrence relation of the form:

$$x_{n+1} = \frac{x_n + m_n(5x_n + 2)}{2}, m_n = x_n \pmod 2$$

This sequence is identical to the known Collatz sequence, more traditionally defined as:

$$x_{n+1} = 3x_n + 1, x \equiv 1 \pmod 2$$

$$x_{n+1} = \frac{x_n}{2}, x \equiv 0 \pmod 2$$

That the relation is equivalent to the Collatz sequence is evident by re-arranging and simplifying this equation:

$$x_{n+1} = m_n(3x_n + 1) + (1 - m_n)\frac{x}{2}$$

A single application of the recurrence relation can be represented as:

$$x_{n+1} = succ^{(1)}(x_n)$$

k-repeated applications of the recurrence relation can be represented thus:

$$x_{n+k} = succ^{(k)}(x_n)$$

Alternatively, x_n, can be expressed, for some n and x_0, thus:

$$x_n = succ^{(n)}(x_0)$$

Alternative forms of the recurrence relation

Using the identity:

$$5m + 1 = 6^m = 2^m 3^m \iff m \in \{0, 1\}$$

the same recurrence relation can also be expressed in all of the following ways:

$$x_{n+1} = \frac{(5m_n + 1)x_n + 2m_n}{2}$$

$$x_{n+1} = \frac{6^{m_n} x_n + 2m_n}{2}$$

$$x_{n+1} = \frac{2^{m_n} 3^{m_n} x_n + 2m_n}{2}$$

$$x_{n+1} = 2^{m_n - 1} 3^{m_n} x_n + m_n$$

and, finally, this:

$$x_{n+1} = 2^{m_n - 1}((2 + 1)^{m_n} x_n + m_n)$$

in which 3^{m_n} been replaced by $(2 + 1)^{m_n}$

In [1]:
```
from fractions import Fraction
import math

def m(x):
    return x%2

def succ(x):
    succ=(x+m(x)*(5*x+2))//2
    assert not(x%2 == 0) or succ == x//2
    assert not(x%2 == 1) or succ == 3*x+1
    assert succ == (6**(m(x))*x+2*m(x))//2
    assert succ == (2**(m(x))*3**(m(x))*x+2*m(x))//2
    assert succ == 2**(m(x)-1)*3**(m(x))*x+m(x)
    assert succ == 2**(m(x)-1)*(2+1)**(m(x))*x+m(x)
    return succ

def collatz(x0, n):
    x=x0
    for k in range(0, n):
        yield {"i": k, "x": x, "x0": x0, "m": m(x), "succ": succ(x)
        x=succ(x)

result=[e for e in collatz(27, 10)]
#result
```

The sequences o_n, e_n

Let $o_0 = m_0, o_{n+1} = o_n + m_{n+1}, n > 0.$

Let $e_n = n + 1 - o_n$

Clearly:

$o_n = \sum_{k=0}^{k=n} m_k$

$e_n = \sum_{k=0}^{k=n} (1 - m_k)$

Interpretation

The series o_n and e_n is that they represent the number of odd and even terms in the path between x_0 and x_n.

In [2]:
```
def o(seq):
    for e in seq:
        if e["i"] == 0:
            e["o"]=e["m"]
        else:
            e["o"]=pred["o"]+e["m"]
        assert e["o"] <= e["i"]+1
        yield e
        pred=e

def e(seq):
    for e in o(seq):
        e["e"]=e["i"]+1-e["o"]
        assert e["o"]+e["e"]==e["i"]+1
        yield e

result=[e for e in e(collatz(27, 10))]
#result
```

An expression for x_n in terms of x_0.

the sequences a_n, c_n and d_n

In this section we rewrite each term x_n as combination of terms from three simpler monotonically incresing sequences:

- a_n (*the accumulator sequence*)
- c_n (*the coefficient sequence*)
- d_n (*the divisor sequence*).

Let the first term of the sequence, x_n be named x_0 with subsequent terms being x_1, x_2 et cetera.

Then we can define:

$x_n = \frac{c_n x_0 + a_n}{d_n}$

with $a_0 = 0$, $c_0 = 1$ and $d_0 = 1$.

Now, let's derive an expression for x_{n+1} in terms of x_0, a_n, c_n and d_n

Recall that:

$x_{n+1} = \frac{x_n + m_n(5x_n + 2)}{2}$

Substituting $\frac{c_n x_0 + a_n}{d_n}$ for x_n yields:

$$x_{n+1} = \frac{\frac{cn\,x_0+an}{dn} + m_n(5(\frac{cn\,x_0+an}{dn})+2)}{2}$$

Simplifying:

$$x_{n+1} = \frac{c_n x_0 + a_n + m_n(5(c_n x_0 + a_n) + 2d_n)}{2d_n}$$

Redistributing terms:

$$x_{n+1} = \frac{c_n(5m_n+1)x_0 + a_n(5m_n+1) + m_n 2d_n}{2d_n}$$

Noting the identity:

$$5m_n + 1 = 6^{m_n}$$

a_n can be further simplified to:

$$x_{n+1} = \frac{6^{m_n} c_n x_0 + 6^{m_n} a_n + m_n 2d_n}{2d_n}$$

This is of the form:

$$x_{n+1} = \frac{c_{n+1} x_0 + a_{n+1}}{d_{n+1}}$$

with:

$$a_{n+1} = 6^{m_n} a_n + m_n 2d_n$$

$$c_{n+1} = 6^{m_n} c_n$$

$$d_{n+1} = 2d_n$$

In [3]:
```
def d(seq):
    for e in seq:
        if e["i"] == 0:
            e["d"]=1
        else:
            e["d"]=(2*pred["d"])
        yield e
        pred=e

def c(seq):
    for e in seq:
        if e["i"] == 0:
            e["c"]=1
        else:
            e["c"]=(6**pred["m"])*pred["c"]
        yield e
        pred=e

def a(seq):
    for e in d(seq):
        if e["i"] == 0:
            e["a"]=0
        else:
            e["a"]=(6**pred["m"])*pred["a"]+(2*pred["m"]*pred["d"])
        yield e
        pred=e

def a_c_d(seq):
    for e in c(a(seq)):
        if e["i"] > 0:
            assert e["x"] == (e["c"]*e["x0"]+e["a"])/e["d"]
        yield e
        pred=e


results=[e for e in a_c_d(collatz(27, 10))]
#results
```

Rewriting a_{n+1}, c_{n+1}, d_{n+1}, x_{n+1} in terms of x_0

Case: $x_0 \rightarrow x_1$

We can check this for n=0 as follows:

$$a_1 = 6^{m_0} a_0 + 2m_0 d_0 = 2m_0 d_0 = 2m_0$$

$$c_1 = 6^{m_0} c_0 = 6^{m_0}$$

$$d_1 = 2d_0 = 2$$

$$x_1 = \frac{6^{m_0} x_0 + 2m_0}{2} = \frac{(5m_0+2)x_0 + m_0}{2} = \frac{x_0 + m_0(5x_0+2)}{2} = succ^{(1)}(x_0)$$

Case: $x_0 \rightarrow x_1 \rightarrow x_2$

$$a_2 = 6^{m_1} a_1 + 2d_1 m_1 = 6^{m_1} m_0 2 + m_1 4$$

$$c_2 = 6^{m_1} c_1 = 6^{m_1} 6^{m_0}$$

$$d_2 = 2d_1 = 4$$

$$x_2 = \frac{6^{m_1} 6^{m_0} x_0 + 6^{m_1} m_0 2 + m_1 4}{4} = succ^{(2)}(x_0)$$

Case: $x_0 \rightarrow x_1 \rightarrow x_2 \rightarrow x_3$

$$a_3 = 6^{m_2} a_2 + 2d_2 m_2 = 6^{m_2}(6^{m_1} 2m_0 + 4m_1) + m_2 8 = 6^{m_2} 6^{m_1} m_0 2 + 6^{m_2} m_1 4 + m_2 8$$

$$a_3 = 6^{o_2}\left(\sum_{i=0}^{i=2} \frac{2^{i+1} m_i}{6^{o_i}}\right)$$

$$c_3 = 6^{m_2} 6^{m_1} 6^{m_0} = 6^{o_2}$$

$$d_3 = 8 = 2^3$$

$$x_3 = \frac{6^{o_2} x_0 + 6^{o_2}\left(\sum_{i=0}^{i=2} \frac{2^{i+1} m_i}{6^{o_i}}\right)}{2^3}$$

$$x_3 = 6^{o_2}\left(\frac{x_0 + \sum_{i=0}^{i=2} \frac{2^{i+1} m_i}{6^{o_i}})}{2^3}\right) = succ^{(3)}(x_0)$$

Case: $x_0 \rightarrow \ldots \rightarrow x_n \rightarrow x_{n+1}$

It can be shown that this generalises to:

$$x_{n+1} = 6^{o_n}\left(\frac{x_0 + \sum_{i=0}^{i=n} \frac{2^{i+1} m_i}{o_i}}{2^{n+1}}\right)$$

Replacing n with $n-1$ everywhere yields:

$$x_n = 6^{o_{n-1}}\left(\frac{x_0 + \sum_{i=0}^{i=n-1} \frac{2^{i+1} m_i}{6^{o_i}}}{2^n}\right)$$

$$x_n = 6^{o_{n-1}}\left(\frac{x_0 + \sum_{i=0}^{i=n-1} \frac{2^{i+1} m_i}{6^{o_i}}}{2^n}\right)$$

Recognising that $n - o_{n-1} = e_{n-1}$ we get:

$$x_n = 2^{o_{n-1}} 3^{o_{n-1}}\left(\frac{x_0 + \sum_{i=0}^{i=n-1} \frac{2^{i+1} m_i}{2^{o_i} 3^{o_i}}}{2^{o_{n-1}+e_{n-1}}}\right)$$

$$x_n = \frac{3^{o_{n-1}}(x_0+\sum_{i=0}^{i=n-1} \frac{2^{e_i} m_i}{3^{o_i}})}{2^{e_{n-1}}}$$

$$x_n = \frac{3^{o_{n-1}}}{2^{e_{n-1}}}(x_0 + \sum_{i=0}^{i=n-1} \frac{2^{e_i} m_i}{3^{o_i}}) = succ^{(n)}(x_0)$$

A Beautiful k

Re-arranged, the expression for x_n yields an expression for k_n the value of $2^{e_{n-1}} x_n - 3^{o_{n-1}} x_0$

$$k_n = 2^{e_{n-1}} x_n - 3^{o_{n-1}} x_0 = 3^{o_{n-1}} \sum_{i=0}^{i=n-1} \frac{2^{e_i}}{3^{o_i}} m_i$$

which describes an identity that applies to all Collatz sequences of finite length.

We refer to this as *the beautiful k* because rather than being an opaque, structure-less k, the formula captures, in a single expression, the intricate twists and turns of any particular Collatz sequence with a simple expression involving only a scaling factor, $3^{o_{n-1}}$ and a series that includes a term, $\frac{2^{e_i}}{3^{o_i}} m_i$ for each term in the underlying sequence, scaled according to its position in the sequence. Only the odd series terms contribute directly to the final sum, but even the even series terms make their presence felt by scaling the subsequent odd terms via the 2^{e_i} exponents.

The sequence k_n

This section we derive an expression k_n in terms of k_{n-1} the identity that relates x_0 and x_{n-1}.

$$k_n = 3^{m_{n-1}+o_{n-2}} (\sum_{i=0}^{i=n-2} \frac{2^{e_i}}{3^{o_i}} m_i + \frac{2^{e_{n-1}}}{3^{o_{n-1}}} m_{n-1})$$

$$k_n = 3^{m_{n-1}} (k_{n-1} + 3^{o_{n-2}} \frac{2^{e_{n-1}}}{3^{o_{n-1}}} m_{n-1})$$

$$k_n = 3^{m_{n-1}} k_{n-1} + 2^{e_{n-1}} m_{n-1}$$

To be well-defined for n, we need an expression for k_0.

Subtituting $n = 1$ we get:

$$k_1 = 3^{m_0} k_0 + 2^{e_0} m_0$$

but we also know:

$$k_1 = 2^{e_0} x_1 - 3^{o_0} x_0$$

So:

$$3^{m_0} k_0 + 2^{e_0} m_0 = 2^{e_0} x_1 - 3^{o_0} x_0$$

Re-arranged:

$k_0 = \frac{2^{e_0}(x_1 - m_0)}{3^{m_0}} - x_0$

Substituting:

$x_1 = \frac{6^{m_0} x_0 + 2 m_0}{2}$

$k_0 = \frac{2^{e_0}(\frac{6^{m_0} x_0 + 2 m_0}{2} - m_0)}{3^{m_0}} - x_0$

$k_0 = \frac{2^{e_0 - 1}(6^{m_0} x_0 + 2 m_0 - 2 m_0)}{3^{m_0}} - x_0$

Simplifying:

$k_0 = 2^{e_0 + m_0 - 1} x_0 - x_0 = (2^{e_0 + m_0 - 1} - 1) x_0$

But $e_0 = 1 - o_0, o_0 = m_0, e_0 + o_0 - 1 = 0$ so:

$k_0 = (2^0 - 1) x_0 = (1 - 1) x_0 = (0) x_0 = 0$

So:

$k_n = 3^{m_{n-1}} k_{n-1} + 2^{e_{n-1}} m_{n-1}, n > 0$

$k_0 = 0, n = 0$

We note that:

- $m_{n-1} = 0 \iff k_n = k_{n-1}$
- $m_{n-1} = 1 \iff k_n = 3k_{n-1} + 2^{e_{n-1}}$
- $k_n >= k_{n-1}$

In [4]:

```
def k(seq):
    for e_ in e(seq):
        if e_["i"] == 0:
            e_["k"] = 0
        else:
            e_["k"]=(3**pred["m"])*pred["k"]+(2**pred["e"])*pred["m
            assert (2**pred["e"])*e_["x"]-(3**pred["o"])*e_["x0"] ==
        yield e_
        pred=e_

results=[e for e in k(collatz(27, 112))]
#results
```

Definition: u, x_u

Recall that:

$x_n = succ^{(n)}(x_0)$

Now define u be the smallest value of n such that $x_n = x_u = succ^{(u)}(x_0) = 1$, if such a value exists and to be undefined otherwise.

An expression for x_0 in terms of $\{i \in [0, n-1] : m_i\}$

In the final analysis, an expression for k is only useful if it leads to progress towards the Collatz conjecture:

Now let's forget k, and consider the identity:

$2^{e_{n-1}} x_n - 3^{o_{n-1}} x_0 = 3^{o_{n-1}} \sum_{i=0}^{i=n-1} \frac{2^{e_i}}{3^{o_i}} m_i$

Let's assume that the conjecture is true. Given the identity holds and substituting $x_n = x_u = 1$ then:

$\text{Collatz} \implies \forall x_0 \in \mathbb{N} : \exists u \in \mathbb{N}, m_i \in [0,1] : 2^{e_{u-1}} - 3^{o_{u-1}} x_0 = 3^{o_{u-1}} \sum_{i=0}^{i=u-1} \frac{2^{e_i}}{3^{o_i}} m_i$

Rearranging and factoring out common terms we get:

$2^{e_{u-1}} = 3^{o_{n-1}} (x_0 + \sum_{i=0}^{i=u-1} \frac{2^{e_i} m_i}{3^{o_i}})$

Solving for x_0 we get:

$\text{Collatz} \implies \forall x_0 \in \mathbb{N} : \exists u \in \mathbb{N}, m_i \in [0,1] : x_0 = \frac{2^{e_{u-1}}}{3^{o_{u-1}}} - \sum_{i=0}^{i=u-1} \frac{2^{e_i} m_i}{3^{o_i}}$

Congruency of paths

Two Collatz sequences of the same length are congruent if, and only if, their identities are equal, that is:

$$x_0 \ldots x_n \cong y_0 \ldots y_n \iff 2^{e_{n-1}} x_n - 3^{o_{n-1}} x_0 = 2^{e_{n-1}} y_n - 3^{o_{n-1}} y_0 = 3^{o_{n-1}} \sum_{i=0}^{i=n-1} \frac{2^{e_i}}{3^{o_i}} m_i$$

By definition, two paths are in the same congruency class if, and only if, they are congruent by the defintion above.

Path class identifiers

Let $p_{n,k}$ be defined as:

$$p_{n,k} : 2^{\mathbb{N}} \to \mathbb{N} : 2^n + \sum_{i=0}^{i=n-1} 2^i m_i$$

$p_{n,k}$ can be interpreted as identifier of the class of paths of length $n + 1$ which are congruent to k.

The isolated 2^n term serves to encode the length of the path and thus helps to distinguish the class path identifiers for paths like $8 \to 4 \to 2 \; (1000) \text{ and } 4 \to 2 \; (100)$.

---

* *whatever the importance of this contribution, the title is technically not wrong - these adjectives also describe paths that fail to reach their final objective*

Appendix A: an example: $x_0 = 3 \to x_7 = 1$

Consider $x_0 = 3$:

$\{x_k\} = \{3, 10, 5, 16, 8, 4, 2, 1\}$

$n = 7$

$\{m_k\} = \{1, 0, 1, 0, 0, 0, 0, 1\}$

$\{o_k\} = \{1, 1, 2, 2, 2, 2, 2, 3\}$

$\{e_k\} = \{0, 1, 1, 2, 3, 4, 5, 5\}$

The 8th term of $\{x_k\}$ is x_7=1.

Substituting into the identity we have:

$2^{e_6} x_7 - 3^{o_6} x_0 = 2^{e_0} 3^{o_6 - o_0} + 2^{e_2} 3^{o_6 - o_2}$

$2^{e_6} x_7 - 3^{o_6} x_0 = 2^{e_0} 3^{o_6 - o_0} + 2^{e_2} 3^{o_6 - o_2}$

$2^5 x_7 - 3^2 x_0 = 2^0 3^{2-1} + 2^1 3^{2-2}$

$2^5 x_7 - 3^2 x_0 = 2^0 3^1 + 2^1 3^0$

$32x_7 - 9x_0 = 3 + 2$

$32 \times 1 - 9 \times 3 = 3 + 2$

$5 = 5$

Revision History

version	date	notes
1.0.5	2023-10-14	fixed alternative recurrence relations - thanks A! Added notes about path congruency and path class identifiers, notes about identity if Conjecture is true. Added expressions for the sequence k_n. Defined u, x_u. Corrected deriviations of a_n, c_n, d_n. Provided implementations of a_n, c_n, d_n, e_n and k_n
1.0.4	2023-10-14	reworded abstract to reflect the fact the identity is already well known
1.0.3	2023-10-14	cleanup some typos and added acknowledgments and dedications
1.0.2	2023-10-14	a beautiful k formulation
1.0.1	2023-10-12	slight cleanups of formulae.
1.0	12023-10-10	initial version - no revision history