Proof of the Collatz Conjecture

Abstract

This paper presents a proof of the Collatz conjecture. By analyzing the dynamics of the original Collatz operations within a stochastic process model, we prove that they lead to contraction due to an inevitable and deterministic lower bound for the ratio a/b of the counter variables. We finally show by strong induction that the original Collatz operations applied to any positive integer n>1 can only produce sequences that contract to 1.

1. The Collatz conjecture

The Collatz conjecture asserts that the Collatz sequence defined by the rule

$$L(n) = \begin{Bmatrix} 3n+1 \quad \text{in case } n \text{ is odd} \\ \frac{n}{2} \quad \text{in case } n \text{ is even} \end{Bmatrix} \tag{1}$$

will eventually reach 1 for any positive integer n.

b: counter number for the amount of the operation (3n+1) in case n is odd

a: counter number for the amount of any operation n/2 in case n is even

k: the total number of operations (steps), k=a+b

1.1 Derivation of an Evolution formula

We derive an evolution formula for the approximation of the resulting number n_k after k=a+b Collatz operations. This formula approximates the numbers n_k that evolve within the sequence by following the defined operations of the Collatz sequence.

We derive the evolution formula as follows: We recursively apply all steps, starting from n_0. The effect of applying b times the multiplication 3n in case n is odd, is:

$$n_0 \cdot 3^b$$

Then we apply all halving steps. But thus we need an additive correction term C, because each (3n+1) operation introduces an additional value 1.

$$n_k = n_0 \cdot \frac{3^b}{2^a} + C \tag{2}$$

The additive correction term C originates as follows:

$$n_k = n_0 \cdot \frac{3^b}{2^a} + \frac{1}{2^a}\sum_{j=0}^{b-1} 3^j \cdot 1 = n_0 \cdot \frac{3^b}{2^a} + \frac{1}{2^a} \cdot \frac{3^b - 1}{2} = n_0 \cdot \frac{3^b}{2^a} + \frac{3^b - 1}{2^{a+1}} \tag{3}$$

1.2 Derivation of the contraction requirement

It is possible to calculate the condition for which the application of an amount of “a” even operations and “b” odd operations defined in the Collatz operations (section 1), leads to contraction, i.e. $n_k<n_0$ after k=a+b total steps.

We start with formula (3), and keep in mind that for contraction we need $n_k<n_0$

$$n_k = n_0 \cdot \frac{3^b}{2^a} + \frac{3^b - 1}{2^{a+1}}$$

We define:

$$T := n_0 \cdot \frac{3^b}{2^a}$$

$$C := \frac{3^b - 1}{2^{a+1}}$$

Then (3) reads as follows:

$$n_k = T + C$$

No we are taking the logarithm on both sides of the equation

$$log_2(n_k) = log_2(T + C)$$

For large n_0 and T»C this means:

$$log_2(n_k) \approx log_2(T) + log_2\left(1 + \frac{C}{T}\right) = log_2(T) + \delta_1 \tag{4}$$

where

$$\delta_1 \coloneqq log_2\left(1 + \frac{C}{T}\right)$$

Now we expand $log_2(T)$

$$log_2(T) = log_2(n_0) + b \cdot log_2(3) - a$$

Combining with (4) leads to

$$log_2(n_k) \approx log_2(T) + log_2\left(1 + \frac{C}{T}\right) = log_2(n_0) + b \cdot log_2(3) - a + \delta_1 \tag{5}$$

If we demand contraction $n_k<n_0$ then:

$$log_2(n_k) < log_2(n_0)$$

Thus

$$log_2(n_0) + b \cdot log_2(3) - a + \delta_1 < log_2(n_0)$$

Subtract $log_2(n_0)$ from both sides and solve for the ratio a/b. Thus we get the condition for the ratio a/b that leads to contraction $n_k<n_0$:

$$\frac{a}{b} > log_2(3) + \frac{\delta_1}{b} = 1.58496 + \frac{\delta_1}{b}$$

where

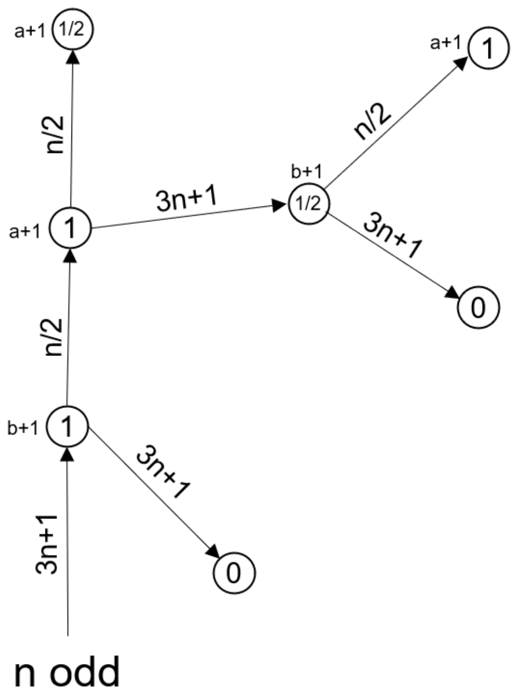
$$\delta_1 = log_2\left(1 + \frac{C}{T}\right) = log_2\left(1 + \frac{\frac{3^b - 1}{2^{a+1}}}{n_0 \cdot \frac{3^b}{2^a}}\right) = log_2\left(1 + \frac{1 - 3^{-b}}{n_0 \cdot 2}\right) \tag{6}$$

The correction term δ_1 shrinks exponentially fast as $n_0 \rightarrow \infty$

2. Recursive stochastic process model

2.1 Probabilistic tree-structure of Collatz operations

We analyze a tree-like structure for the recursive succession of the original Collatz operations.
In the following exemplary tree like structure for an odd integer, the directed branches show the paths of Collatz operations, and the nodes are the probabilities that the operation assigned to the branch is executed. At each node with probability p>0, one of the counter variables “a”, “b” is incremented.



Picture 1: Tree-structure of Collatz operations

In this example, “a” was incremented by 3 in total, and “b” was incremented by 2.
This tree-structure of operations doesn’t take into account the recursive structure of the operations. Thus we need an improved recursive stochastic model.

2.2 Recursive stochastic process model

We appreciate the fact, there exists consecutive operations in case n is odd, i.e. the operation (3n+1) is directly followed by (n/2) with probability 1. We call this halving an "implicite halving". Afterwards there is probability ½, that the next operation is either (3n+1) or probability ½ that the next operation is (n/2), which we call an "explicite halving". Whereas after any (n/2) operation, the probability is ½, to be followed by (n/2) or ½ to be followed by (3n+1).
It is also proven, that the probability for the occurrence of a combined operation (3n+1)/2 and the probability of an explicite halving (n/2), is ½ for each. From this it must follow, that for large k=a+b, the total occurrence of (3n+1) is k/3, the total occurrence of explicite halvings is k/3 and the total occurrence of implicite halvings must be the same as the total occurrence of (3n+1), also k/3. All these statements are based on proven theorems.

In this decisive step we generate a model of the structure of the Collatz process with a Markov-chain approach, for which we need to introduce 5 distinct and recursive main operations O_1 ,..., O_5.

For each of the operations we assign its probability and its impact on the counter variables "a" and "b". We appreciate the fact that there is no deterministic succession or order of these operations. They are only applied recursively by the requirements of the original Collatz process described in section 1.
We also introduce the total occurrence of the operations for large k.

Operation	Description	Probability P(Ox) 1 ≤ x ≤ 5	Total occurrence for large k	Δa	Δb	k
O1	(3n+1) after implicite (n/2)	½	k/3	0	1	1
O2	(3n+1) after explicite (n/2)	½		0	1	2
O3	explicite (n/2) after explicite (n/2)	½	k/3	1	0	3
O4	explicite (n/2) after implicite (n/2)	½		1	0	4
O5	implicite (n/2) after (3n+1)	1	k/3	1	0	5

Table 1: The five main operations O_1 ,..., O_5

These 5 operations O_1 ,..., O_5. cover the whole Collatz process.

We can map these five operations in the order just defined by the Collatz conjecture onto any complete Collatz sequence. As we can see, this model correctly describes the behavior of the counter variables "a" and "b" within the original Collatz counting system. It reconciles that a+b=k.

The main advantage of this approach is, that we are able to calculate the expectation values for the counter variables "a" and "b", without the need for a special probability distribution.

By modeling Δa, and Δb as random variables, we are able to calculate Variances for "a" and "b".

2.2.1 Expectation values for a and b and the ratio EV(a)/EV(b)

We calculate the expectation values of “a” and “b” just by the counting as described in section 2.1 in case k is large, with the probabilities P(Ox) and its occurences.

$$EV(a) = \left(amount\ of\ explicite\ \frac{n}{2}\right) \cdot P(O3) + \left(amount\ of\ implicite\ \frac{n}{2}\right) \cdot P(O4) + (amount\ of\ 3n+1) \cdot P(O5) = \frac{k}{3} \cdot \frac{1}{2} + \frac{k}{3} \cdot \frac{1}{2} + \frac{k}{3} \cdot 1 = \frac{2}{3}k \tag{7}$$

$$EV(b) = \left(amount\ of\ implicite\ \frac{n}{2}\right) \cdot P(O1) + \left(amount\ of\ explicite\ \frac{n}{2}\right) \cdot P(O2) = \frac{k}{3} \cdot \frac{1}{2} + \frac{k}{3} \cdot \frac{1}{2} = \frac{1}{3}k \tag{8}$$

From (2) and (3) we are able to calculate the ratio of the expectation values EV(a)/EV(b) for large k

$$\frac{EV(a)}{EV(b)} = 2 \tag{9}$$

This must not be misinterpreted as the expectation value for the ratio a/b. But as a first result it shows that $\log_2(3)$<EV(a)/EV(b) can be interpreted as a tendency which could favor contraction.

2.2.2 Variance analysis for Δa, Δb and EV(a)/EV(b)

From our recursive stochastic model, which is shown in table 1, we can directly follow that the probability for an increase in “a” is 3/5, and for “b” it is 2/5.
As we can map this stochastic process model onto any Collatz sequence, especially for large n>>11, we can treat Δa and Δb as strongly dependent and correlated random variables

$$\Delta a = \begin{cases} 0 & no\ increase\ of\ a \\ 1 & increase\ of\ a\ by + 1\ with\ probability\ P(\Delta a) = 3/5 \end{cases}$$

$$\Delta b = \begin{cases} 0 & no\ increase\ of\ b \\ 1 & increase\ of\ b\ by + 1\ with\ probability\ P(\Delta b) = 2/5 \end{cases}$$

Thus we get

$$P(\Delta a) = \frac{3}{5}$$

$$P(\Delta b) = \frac{2}{5}$$

We appreciate the fact that k= Δa+Δb.

No we model both dependent random variables Δa and Δb, by 2 distinct binomial distributions, which are correlated.

$$\Delta a \sim Bi(k;\ P(\Delta a)) = Bi\left(k;\ \frac{3}{5}\right)$$

$$\Delta b \sim Bi(k;\ P(\Delta b)) = Bi\left(k;\frac{2}{5}\right)$$

The approach of a joint distribution for both random variables is avoided in first place for the advance of the following analysis. First we calculate the variances Var(Δa) and Var(Δb).

$$Var(\Delta a) = kP(\Delta a)\big(1 - P(\Delta a)\big) = \frac{6}{25}k \tag{10}$$

$$Var(\Delta a) = kP(\Delta b)\big(1 - P(\Delta b)\big) = \frac{6}{25}k \tag{11}$$

From the way we calculated these variances, one finds that they present an upper variance limit. It is straightforward to conclude that Var(Δa)=Var(a) and Var(Δb)=Var(b).

Now we are able to perform a sensitivity analysis of these variances onto the ratio EV(a)/EV(b):

$$\frac{a}{b} \approx \frac{EV(a) \pm \sqrt{Var(a)}}{EV(b) \pm \sqrt{Var(b)}}$$

We consider the 2 extreme cases of the upper limit for this ratio and the lower limit for it

The upper limit is

$$\left(\frac{EV(a)}{EV(b)}\right)_{max} = \frac{EV(a) + \sqrt{Var(a)}}{EV(b) - \sqrt{Var(b)}} = \frac{\frac{2}{3}k + \sqrt{\frac{6}{25}k}}{\frac{1}{3}k - \sqrt{\frac{6}{25}k}} \tag{12}$$

The lower limit

$$\left(\frac{EV(a)}{EV(b)}\right)_{min} = \frac{EV(a) - \sqrt{Var(a)}}{EV(b) + \sqrt{Var(b)}} = \frac{\frac{2}{3}k - \sqrt{\frac{6}{25}k}}{\frac{1}{3}k + \sqrt{\frac{6}{25}k}} \tag{13}$$

It turns out that these estimations already present rigid upper and lower bounds to the ratio, as we will find out soon. This might be due to the fact, that the variances of the constructed binomial distributions in this case have enhanced limiting properties.

The following table shows the result for max and min ratios of EV(a)/EV(b) from k=3 to 243 exemplary, with some values skipped so the trend can be followed.

k	max EV(a)/EV(b)	min EV(a)/EV(b)
3	18,81	0,62
4	10,31	0,73
5	7,75	0,81
6	6,50	0,88
7	5,75	0,93
8	5,24	0,97
9	4,88	1,01
10	4,60	1,05
11	4,39	1,08
12	4,21	1,11
13	4,06	1,13
14	3,94	1,15
15	3,83	1,17
16	3,74	1,19
17	3,66	1,21
18	3,59	1,23
19	3,53	1,24
20	3,47	1,26
.....	.....	.....
75	2,61	1,56
76	2,61	1,57
77	2,60	1,57
78	2,60	1,57
79	2,59	1,57
80	2,59	1,58
81	2,59	1,58
82	2,58	1,58
83	2,58	1,58
84	2,57	1,59
85	2,57	1,59
.....	.....	.....
215	2,33	1,73
216	2,33	1,73
217	2,33	1,73
218	2,33	1,73
219	2,33	1,73
220	2,33	1,73
221	2,33	1,73
222	2,33	1,73
223	2,33	1,73
224	2,33	1,73
225	2,33	1,73
226	2,33	1,73
227	2,32	1,73
228	2,32	1,73
229	2,32	1,73
230	2,32	1,73
231	2,32	1,74
232	2,32	1,74
233	2,32	1,74
234	2,32	1,74
235	2,32	1,74
236	2,32	1,74
237	2,32	1,74
238	2,32	1,74
239	2,32	1,74
240	2,31	1,74
241	2,31	1,74
242	2,31	1,74
243	2,31	1,74

Table 2: Calculation of max and min EV(a)/EV(b) for 3≤k≤243

The values from the table 2 give a clear indication that the Collatz sequence might produce only ratios for a/b > log2(3).

2.2.3 Derivation of a rigid estimator for EV(a/b) and its variance Var(a/b)

Our main goal in this section is to derive a formula of the form

$$\frac{a}{b} = EV\left(\frac{a}{b}\right) \pm \sqrt{Var\left(\frac{a}{b}\right)} \tag{14}$$

Thus we need to calculate an estimator EV(a/b) and the variance of (a/b).

In order to estimate EV(a/b) we use a second-order Taylor expansion of f(a,b)=a/b:

$$f(a,b) \approx f\big(EV(a),EV(b)\big) + f_a\big(EV(a),EV(b)\big)\big(a - EV(a)\big) + f_b\big(EV(a),EV(b)\big)\big(b - EV(b)\big) + \\ + \frac{1}{2} f_{aa}\big(EV(a),EV(b)\big)\big(a - EV(a)\big)^2 + f_{ab}\big(EV(a),EV(b)\big)\big(a - EV(a)\big)\big(b - EV(b)\big) + \frac{1}{2} f_{bb}\big(EV(a),EV(b)\big)\big(b - EV(b)\big)^2$$

Now we take expectations of this approximation. All the linear terms vanish (since EV(a-EV(a))=0) and we are left with

$$EV\left(\frac{a}{b}\right) \approx \frac{EV(a)}{EV(b)} + \frac{Cov(a,b)}{EV(b)^2} + \frac{EV(a) \cdot Var(b)}{EV(b)^3} \tag{15}$$

In order to continue we just need to derive a reasonable expression for the covariance Cov(a,b). For this we make us of the generalized variance propagation as follows

$$Var(f) = F_x \cdot C_y \cdot F_x^T$$

where F_x is the scaled gradient row vector

$$F_x = \left[\sigma_{x_1} \frac{\partial f}{\partial x_1}, \dots, \sigma_{x_n} \frac{\partial f}{\partial x_n}\right]$$

C_y is the diagonal correlation matrix, with elements r_{xixj}, which need not be symmetric.

Applied to our problem, we take into account that the two random variables Δa, Δb, thus the counter variables a and b are dependent and correlated.

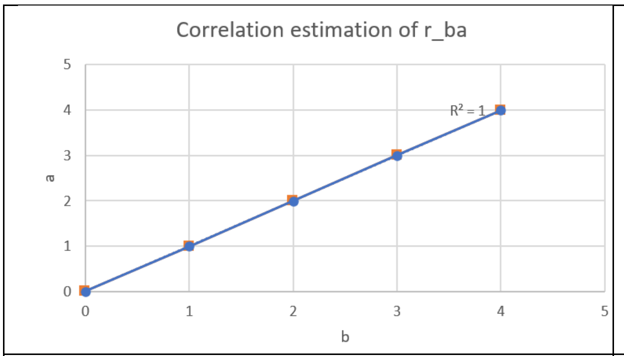
The generalized variance propagation turns out for f(a,b)=a/b, to be

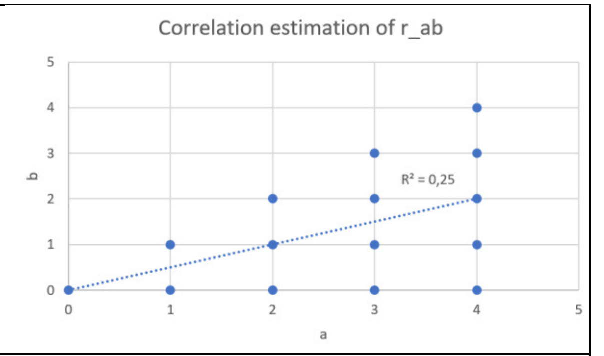
$$Var(f(a,b)) = \left(\frac{\partial f}{\partial a}\right)^2 Var(a) + \left(\frac{\partial f}{\partial b}\right)^2 Var(b) + \frac{\partial f}{\partial b}\frac{\partial f}{\partial a}\sqrt{Var(a)}\sqrt{Var(b)}(r_{ab} + r_{ba}) \tag{16}$$

where in the last expression after the "+" sign we find the Covariance as

$$Cov(a,b) = \sqrt{Var(a)}\sqrt{Var(b)}(r_{ab} + r_{ba}) \tag{17}$$

We need an estimation of r_{ab} and r_{ba} and find it by analyzing the behavior of a and b with respect to the Collatz operations:





We find r_ba=1	and r_ab=0.5

In order to calculate (16) and (17) symbolically we need following additional expressions

$$f(a,b) = \frac{a}{b}$$

$$\frac{\partial f}{\partial a} = \frac{1}{b} = \frac{1}{EV(b)} = \frac{3}{k}$$

$$\frac{\partial f}{\partial b} = -\frac{a}{b^2} = -\frac{EV(a)}{EV(b)^2} = -\frac{6}{k}$$

$$Var(a) = \frac{6}{25}k$$

$$Var(b) = \frac{6}{25}k$$

Our final result is

$$\frac{a}{b} = EV\left(\frac{a}{b}\right) \pm \sqrt{Var\left(\frac{a}{b}\right)} = \left(2 + \frac{231}{25k}\right) \pm \frac{\sqrt{108}}{5\sqrt{k}}$$

The following table shows the result for max and min expectation values for EV(a/b) from k=3 to 243 exemplary, with some values skipped so the trend can be followed.

k	EV(a/b)	Sigma(a/b)	EV(a/b)+sigma(R)	EV(a/b)-sigma(R)
3	5,08	1,20	6,28	3,88
4	4,31	1,04	5,35	3,27
5	3,85	0,93	4,78	2,92
6	3,54	0,85	4,39	2,69
7	3,32	0,79	4,11	2,53
8	3,16	0,73	3,89	2,42
9	3,03	0,69	3,72	2,33
10	2,92	0,66	3,58	2,27
11	2,84	0,63	3,47	2,21
12	2,77	0,60	3,37	2,17
13	2,71	0,58	3,29	2,13
14	2,66	0,56	3,22	2,10
15	2,62	0,54	3,15	2,08
16	2,58	0,52	3,10	2,06
17	2,54	0,50	3,05	2,04
18	2,51	0,49	3,00	2,02
19	2,49	0,48	2,96	2,01
20	2,46	0,46	2,93	2,00
...	...	...	...	...
75	2,12	0,24	2,36	1,88
76	2,12	0,24	2,36	1,88
77	2,12	0,24	2,36	1,88
78	2,12	0,24	2,35	1,88
79	2,12	0,23	2,35	1,88
80	2,12	0,23	2,35	1,88
81	2,11	0,23	2,35	1,88
82	2,11	0,23	2,34	1,88
83	2,11	0,23	2,34	1,88
84	2,11	0,23	2,34	1,88
85	2,11	0,23	2,33	1,88
...	...	...	...	...
215	2,04	0,14	2,18	1,90
216	2,04	0,14	2,18	1,90
217	2,04	0,14	2,18	1,90
218	2,04	0,14	2,18	1,90
219	2,04	0,14	2,18	1,90
220	2,04	0,14	2,18	1,90
221	2,04	0,14	2,18	1,90
222	2,04	0,14	2,18	1,90
223	2,04	0,14	2,18	1,90
224	2,04	0,14	2,18	1,90
225	2,04	0,14	2,18	1,90
226	2,04	0,14	2,18	1,90
227	2,04	0,14	2,18	1,90
228	2,04	0,14	2,18	1,90
229	2,04	0,14	2,18	1,90
230	2,04	0,14	2,18	1,90
231	2,04	0,14	2,18	1,90
232	2,04	0,14	2,18	1,90
233	2,04	0,14	2,18	1,90
234	2,04	0,14	2,18	1,90
235	2,04	0,14	2,17	1,90
236	2,04	0,14	2,17	1,90
237	2,04	0,14	2,17	1,90
238	2,04	0,13	2,17	1,90
239	2,04	0,13	2,17	1,90
240	2,04	0,13	2,17	1,90
241	2,04	0,13	2,17	1,90
242	2,04	0,13	2,17	1,90
243	2,04	0,13	2,17	1,90

Table 3: Calculation of EV(a/b) for $3 \leq k \leq 243$

3. The Proof of the Collatz conjecture

3.1 Conclusive Step: Forced Contraction

The contraction requirement $n_k<n_0$ resulted in the fulfillment of a needed inequality for the ratio a/b in (6) as follows:

$$\frac{a}{b} > log_2(3) + \frac{\delta_1}{b}$$

where

$$\delta_1 := log_2\left(1 + \frac{C}{T}\right) = log_2\left(1 + \frac{1 - 3^{-b}}{n_0 \cdot 2}\right) \quad (6)$$

So the correction term δ_1 shrinks to zero exponentially fast as b, $n_0 \to \infty$

In section 2.2.2 we derived a rigor lower bound inequality (13), which must fulfill (6) in order to get contraction, which means $n_k<n_0$

$$\left(\frac{EV(a)}{EV(b)}\right)_{min} = \frac{EV(a) - \sqrt{Var(a)}}{EV(b) + \sqrt{Var(b)}} = \frac{\frac{2}{3}k - \sqrt{\frac{6}{25}k}}{\frac{1}{3}k + \sqrt{\frac{6}{25}k}} > log_2(3) + \frac{\delta_1}{b} \quad (14)$$

For k>=84 the lower bound on EV(a)/EV(b) already exceeds log2(3).
Thus, from the already calculated sequences it can be confirmed, that the lower bound derived in formula (13) forces contraction by exceeding the needed value in (6), thus fulfilling inequality (14). This rigorously and deterministically proves that from already known Collatz sequences for large enough starting integer value, the inequality (14) predicts contraction correctly for any larger starting integer.

3.2 Inductive Step: Full proof of the Collatz Conjecture

We will now prove that for any integer n>1, the Collatz sequence starting at n eventually reaches 1 by strong induction.

We define the Collatz Reachability Statement:

P(n): The Collatz sequence starting at n eventually reaches 1.

Base case:

We verify manually by applying the Collatz operations (see section 1)

P(n=4): 4→2→1

Inductive hypothesis:

If we have shown that P(j) holds for all 2≤j≤n. That is, for all j less than or equal to n, the Collatz sequence starting at j reaches 1. A large enough n, from already calculated sequences can easily be found to confirm the point from which the fulfillment of inequality (14) starts to predict contraction correctly.

Inductive step: Prove P(n+1)

From the conclusive step in section 3.1 it is shown that the Collatz operations applied to any positive integer n+1 will finally lead to a positive number n<n+1 due to the fact that contraction is forced (see 3.1). Then by the inductive hypothesis, since all smaller numbers down to 2 are shown to reach 1, so must n+1.

We have finally and fully proven that the Collatz operations defined in Section 1 applied to any positive integer n>1 can only produce sequences that contract to 1.

References

[1] Jeffrey C. Lagarias & Alan Weiss (1992)
"The $3x + 1$ problem: two stochastic models." Annals of Applied Probability, 2(1), 229–261.

[2] Jeffrey C. Lagarias (1985)
"The $3x + 1$ problem and its generalizations." American Mathematical Monthly, 92(1), 3–23.