OPTIMAL BOUNDS FOR THE LENGTH OF RATIONAL COLLATZ CYCLES

LORENZ HALBEISEN AND NORBERT HUNGERBÜHLER

ABSTRACT. We consider the arithmetics of Collatz cycles in $\mathbb{Q}[(2)]$. In particular, we prove optimal estimates for the length of a cycle in terms of its minimum. As an application, we derive an improved version of Eliahou's criterion, and we show that the length of (integer) Collatz cycles which do not contain 1, is at least 102 225 496 provided the Collatz conjecture is verified for all initial values $x_0 \leq 212\,366\,032\,807\,211$.

1. INTRODUCTION

For $x \in \mathbb{R}$ let $g_0(x) = \frac{x}{2}$ and $g_1(x) = \frac{3x+1}{2}$. Let $\mathbb{Q}[(2)]$ denote the local ring of fractions of $\mathbb{Z}$ at the prime ideal (2), i.e. the domain of all rational numbers having (written in least terms) an odd denominator. A number $\frac{p}{q} \in \mathbb{Q}[(2)]$ with q odd is considered even or odd according to the parity of the numerator p. Then the Collatz sequence generated by $x_0 \in \mathbb{Q}[(2)]$ is defined by

$$x_n = \begin{cases} g_0(x_{n-1}) & \text{if } x_{n-1} \text{ is even} \\ g_1(x_{n-1}) & \text{if } x_{n-1} \text{ is odd} \end{cases} \tag{1}$$

for $n \in \mathbb{N}$. The structure of the set of Collatz cycles (i.e. periodic Collatz sequences) in $\mathbb{Q}[(2)]$ has been studied by Lagarias in [6]. Of course, if $x_0 \in \mathbb{N}$, then the generated sequence is the usual integer Collatz sequence, and hence, statements about rational Collatz sequences have implications for integer Collatz sequences. We recall that the Collatz conjecture which states that

(C) for all $x_0 \in \mathbb{N}$ there exists an index n with $x_n = 1$,

is still open. Of course (C) is equivalent to the conjunction of (A) and (B):

(A) (1,2) is the only Collatz cycle in $\mathbb{N}$.

(B) Every Collatz sequence in $\mathbb{N}$ is bounded.

As a main reference about the Collatz problem we refer to [5]. For more recent results see [1], [2], [4], [8], [9], [6], [3] and [10]. Partial results are:

(D) The Collatz conjecture holds true for all initial values $x_0 \leq 6.3\ 10^{13}$ (see [7]).

(E) The length of a Collatz cycle in $\mathbb{N}$ which does not contain 1 is at least 17 087 915 (see [3]).

The aim of this article is to present some new techniques which allow a refined analysis of rational (and hence integer) Collatz cycles. In particular, we prove optimal estimates for the length of a positive cycle in terms of its minimum. As an application, we derive an improved version of Eliahou's criterion, and we show that the length of Collatz cycles in $\mathbb{N}$ which do not contain 1, is at least 102 225 496 provided the Collatz conjecture is verified for all initial values $x_0 \leq 212\,366\,032\,807\,211$ (this number is about 3.3 times larger than the actually verified bound mentioned in (D), but Eliahou's original criterion would require to check all initial values up to $2.9\ 10^{14}$ which is about 4.6 times larger).

2. Auxiliary results

Let us start with some notations: Let $S_{l,n}$ denote the set of all 0-1 sequences of length l containing exactly n times a 1, $S_l = \cup_{n=0}^{l} S_{l,n}$ and $S = \cup_{l\in\mathbb{N}} S_l$. To every $s = (s_1, \ldots, s_l) \in S_l$ we associate the affine function $\phi_s : \mathbb{R} \to \mathbb{R}$, $\phi_s = g_{s_l} \circ \ldots \circ g_{s_2} \circ g_{s_1}$. A sequence $(x_0, \ldots, x_l)$ of real numbers x_i is called a pseudo cycle of length l if there exists $s = (s_1, \ldots, s_l) \in S_l$ such that

(i) $\phi_s(x_0) = x_0 \in \mathbb{Q}[(2)]$
(ii) $g_{s_{i+1}}(x_i) = x_{i+1}$ for $i = 0, \ldots, l-1$.

Notice that if $\frac{p}{q} \in \mathbb{Q}$ with $2^r|q$ then $2^r|\tilde{q}$, if $\tilde{q}$ denotes the denominator of $g_i(\frac{p}{q})$ ($i = 0, 1$). Hence, every element of a pseudo cycle is in $\mathbb{Q}[(2)]$. Thus, if $\frac{p}{q}$ and $g_i(\frac{p}{q}) = \frac{\tilde{p}}{\tilde{q}}$ are consecutive elements of a pseudo cycle, then

$$\begin{array}{ll} i = 0 & \text{if } p \text{ is even, since else } \dfrac{\tilde{p}}{\tilde{q}} \notin \mathbb{Q}[(2)] \\ i = 1 & \text{if } p \text{ is odd, since else } \dfrac{\tilde{p}}{\tilde{q}} \notin \mathbb{Q}[(2)]. \end{array}$$

The conclusion of this simple observations is

Lemma 1. *The set of pseudo cycles coincides with the set of Collatz cycles in $\mathbb{Q}[(2)]$. Cycles are either positive or negative.*

Let us now define the function $\varphi\colon S \to \mathbb{N}$ recursively by

$$\begin{aligned} \varphi(\{\}) &= 0 \\ \varphi(s0) &= \varphi(s) \\ \varphi(s1) &= 3\varphi(s) + 2^{l(s)} \end{aligned} \tag{2}$$

where s denotes an arbitrary element of S and $l(s)$ the length of s. φ is easily computed explicitly (see also [6]):

$$\varphi(s) = \sum_{j=1}^{l(s)} s_j 3^{s_{j+1}+\cdots+s_{l(s)}} 2^{j-1}. \tag{3}$$

An easy consequence of the definition (2) is the following decomposition formula which will be convenient afterwards:

$$\varphi(s\bar{s}) = 3^{n(\bar{s})}\varphi(s) + 2^{l(s)}\varphi(\bar{s}). \tag{4}$$

Here $s\bar{s}$ is the concatenation of $s, \bar{s} \in S$, and $n(s)$ denotes here and subsequently the number of 1 in the sequence s. The second lemma shows how φ is used to compute explicitly the function ϕ_s.

Lemma 2 (Lagarias)**.** *For arbitrary $s \in S$ there holds*

$$\phi_s(x) = \frac{3^{n(s)}x + \varphi(s)}{2^{l(s)}}$$

and hence for every $s \in S$ there exists a unique $x_0 \in \mathbb{Q}[(2)]$ which generates a Collatz cycle in $\mathbb{Q}[(2)]$ of length $l(s)$ and which coincides with the pseudo cycle generated by s. x_0 is given by

$$x_0 = \frac{\varphi(s)}{2^{l(s)} - 3^{n(s)}}\,.$$

Proof
The proof is by induction with respect to $l(s)$:

(i) $l(s) = 1$ is easily checked from the definition.

(ii) $l(s) > 1$: 1. case $s = \bar{s}0$:

$$\phi_{\bar{s}0}(x) = \frac{\phi_{\bar{s}}(x)}{2} = \frac{3^{n(\bar{s})}x + \varphi(\bar{s})}{2 \cdot 2^{l(\bar{s})}} = \frac{3^{n(s)}x + \varphi(s)}{2^{l(s)}}$$

2. case $s = \bar{s}1$: analogue. □

For $s \in S_l$ let $\sigma(s)$ denote the orbit of s in S_l generated by the left-shift permutation

$$\lambda_l : (s_1, \ldots, s_l) \mapsto (s_2, \ldots, s_l, s_1),$$

i.e. $\sigma(s) := \{\lambda_l^k(s) \, : \, k = 1, \ldots, l\}$. Furthermore, we define

$$M_{l,n} := \max_{s \in S_{l,n}} \{ \min_{t \in \sigma(s)} \varphi(t)\}.$$

Now, suppose, the Collatz conjecture is verified for all initial values $x_0 \leq m$. If one can then show that

$$\forall n, l < L \, : \quad \frac{M_{l,n}}{2^l - 3^n} \leq m \tag{5}$$

it follows that the length of a Collatz cycle in $\mathbb{N}$ which does not contain 1 is at least L. Before we start with a detailed analysis of the crucial quantity $M_{l,n}$ in the next section, we mention a further technical lemma:

Lemma 3. *Let f_i be a non-increasing sequence of non-negative real numbers, and c_i and d_i be sequences such that for all $n \in \mathbb{N}$*

$$\sum_{i=1}^{n} c_i \leq \sum_{i=1}^{n} d_i \,.$$

Then there holds for all $n \in \mathbb{N}$

$$\sum_{i=1}^{n} c_i f_i \leq \sum_{i=1}^{n} d_i f_i \,.$$

Proof
The stronger assertion

$$\sum_{i=1}^{n} c_i f_i \;\leq\; \sum_{i=1}^{n} d_i f_i - f_n \sum_{i=1}^{n} (d_i - c_i)$$

is easily proved by induction. □

3. The minimum in a Collatz cycle of length l

We start with a monotonicity property of the function φ.

Lemma 4. *Let $s = (s_1, \ldots, s_l)$ and $t = (t_1, \ldots, t_l)$ be two distinct elements of $S_{l,n}$. If for all $k \in \{1, \ldots, l\}$ there holds $\sum_{i=1}^{k} s_i \leq \sum_{i=1}^{k} t_i$, then $\varphi(s) > \varphi(t)$.*

Proof
Because both sequences s and t contain exactly n times a 1 and because they are distinct, there exists a smallest number k_0 with $s_{k_0} = 0$ and $t_{k_0} = 1$ and a smallest number $k_1 > k_0$ with $s_{k_1} = 1$ and $t_{k_1} = 0$. Now let s' be such that $s'_i = s_i$ for $i \notin \{k_0, k_1\}$, $s'_{k_0} = 1$ and $s'_{k_1} = 0$. Then $\sum_{i=1}^{k} s_i \leq \sum_{i=1}^{k} s'_i \leq \sum_{i=1}^{k} t_i$ (for all $k \in \{1, \ldots, l\}$).

Consider the sequence $\bar{s} = (s_{k_0+1}, \ldots, s_{k_1-1})$. With (2) it is not hard to see that $\varphi(0\bar{s}1) > \varphi(1\bar{s}0)$ and with (4) we get $\varphi(s') < \varphi(s)$. If $s' \neq t$ we can repeat the same procedure with s' as we have done with s and find a sequence s'' such that $\varphi(s') > \varphi(s'')$. After finitely many repetitions we get $\varphi(s) > \varphi(t)$. □

In the next lemma we determin the sequence $\tilde{s}$ for which φ attains the value $M_{l,n}$.

Lemma 5. *Let $n \leq l$ be natural numbers. Let $\tilde{s}_i := \lceil i\frac{n}{l}\rceil - \lceil (i-1)\frac{n}{l}\rceil$ (for $1 \leq i \leq l$), then $\varphi(\tilde{s}) = \min_{t\in\sigma(\tilde{s})}\{\varphi(t)\} = M_{l,n}$.*

To indicate that the constructed $\tilde{s}$ depends on l and n, we will often write $\tilde{s}(l,n)$ for it.

Proof

Let $t \in S_{l,n}$. We can represent $\sum_{i=1}^{k} t_i$ $(1 \leq k \leq l)$ by a staircase (see Figure 1).

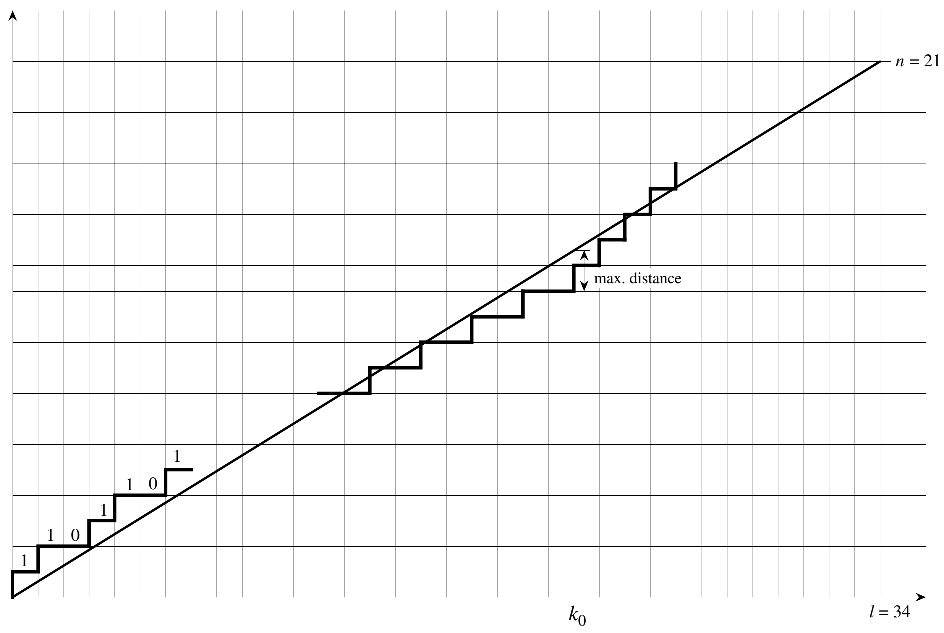

FIGURE 1

Assume there exists a k_0 such that $(k_0\frac{n}{l} - \sum_{i=1}^{k_0} t_i) > 0$ is maximal, then for $t' = \lambda_l^{k_0}(t)$ we have $\sum_{i=1}^{k} t'_i \geq k\frac{n}{l}$ (for all $1 \leq k \leq l$). By construction of $\tilde{s}$ we also have $\sum_{i=1}^{k} t'_i \geq \sum_{i=1}^{k} \tilde{s}_i \geq k\frac{n}{l}$ $(1 \leq k \leq l)$ and by Lemma 4 we get $\varphi(\tilde{s}) > \varphi(t')$ and therefore $\varphi(\tilde{s}) \geq \min_{t'\in\sigma(t)}\{\varphi(t')\}$ for all $t \in S_{l,n}$.

Now we must show that $\varphi(\tilde{s}) = \min_{s\in\sigma(\tilde{s})}\{\varphi(s)\}$. Let $\tilde{s}' := \lambda_l^m(\tilde{s})$ for an arbitrary m. Then by construction of $\tilde{s}$ we see that $\sum_{i=1}^{k} \tilde{s}'_i \leq \sum_{i=1}^{k} \tilde{s}_i$ $(1 \leq k \leq l)$ and therefore $\varphi(\tilde{s}) \leq \varphi(\tilde{s}')$ (by Lemma 4). $\square$

From (3) we hence conclude

Corollary 1. *For every l and $n \le l$ there holds*

$$M_{l,n} = \sum_{j=1}^{l} \left(\lceil j\frac{n}{l}\rceil - \lceil (j-1)\frac{n}{l}\rceil\right) 2^{j-1} 3^{n-\lceil j\frac{n}{l}\rceil}.$$

Using this explicit formula we now derive different estimates for $M_{l,n}$. A first rather rough estimate will be useful afterwards.

Proposition 1. *For all l and $n \le n(l) := \lfloor l \log_3 2 \rfloor$ there holds*

$$\frac{M_{l,n}}{2^l - 3^n} \le \frac{n}{l}\,\frac{1}{2-3^{n/l}}\,.$$

Proof
From Corollary 1 we infer that

$$\begin{aligned} M_{l,n} &\le \sum_{j=1}^{l}(\lceil j\frac{n}{l}\rceil - \lceil (j-1)\frac{n}{l}\rceil)2^{j-1}3^{n-j\frac{n}{l}} \\ &= \sum_{i=1}^{l}(\lceil (l-i+1)\frac{n}{l}\rceil - \lceil (l-i)\frac{n}{l}\rceil)2^{l-i}3^{(i-1)\frac{n}{l}} \end{aligned} \tag{6}$$

(where we have changed the order of summation). Observe now, that $f_i := 2^{l-i}3^{(i-1)\frac{n}{l}}$ is a positive decreasing sequence. Furthermore, we have for all $k \le l$

$$\sum_{i=1}^{k}(\lceil (l-i+1)\frac{n}{l}\rceil - \lceil (l-i)\frac{n}{l}\rceil) \le \sum_{i=1}^{k}\frac{n}{l}\,.$$

Applying Lemma 3 we hence conclude

$$\begin{aligned} M_{l,n} &\le \sum_{i=1}^{l} 2^{l-i}3^{(i-1)\frac{n}{l}}\frac{n}{l} \\ &= \frac{n}{l}\,\frac{2^l-3^n}{2-3^{n/l}} \end{aligned}$$

and the claim is proved. □

The estimate (6) above is quite rough and in order to improve the result, we will consider appropriate pieces of the sequence $\tilde{s}(l,n)$.

Proposition 2. *If we denote by $x = \frac{2^l}{3^n}$ and by $m = 2n - l$, then the following holds:*

$$\frac{M_{l,n}}{3^n} \le \alpha_1 x^{1-1/m} + \alpha_2\frac{x^{1-1/m}-1}{x^{1/m}-1} \tag{7}$$

for all $l > 6$ and $n = n(l)$. Here, $\alpha_1 = \frac{2n+3l}{9m}$ and $\alpha_2 = \frac{6n+l}{27m}$.

Proof

For $l = 7, \dots, 159$, the assertion can be checked by direct computation, using Corollary 1. (For $l < 6$, the right hand side of (7) does not make sense.) If $l \geq 160$, then the sequence $\tilde{s}$ (constructed as in Lemma 5) is composed of the sequences 110 (=: Z) and 11010 (=: A). So, let $\tilde{s} = (a_1, \dots, a_m)$, $a_i \in \{A, Z\}$, be of length l. If a denotes the number of A and z the number of Z in $\tilde{s}$, then it follows from

$$\begin{aligned} 5a + 3z &= l \\ 3a + 2z &= n \end{aligned}$$

that $m = a + z = 2n - l$. Then

$$\varphi(\tilde{s}) = 3^{n(\tilde{s})} \sum_{k=1}^{m} (\frac{8}{9})^{\varepsilon_Z(k)} (\frac{32}{27})^{\varepsilon_A(k)} \chi(k) \tag{8}$$

where $\varepsilon_Z(k)$ and $\varepsilon_A(k)$ denote the number of Z and A respectively in the initial sequence $(a_1, \dots, a_{k-1})$ of $\tilde{s}$ and

$$\chi(k) = \begin{cases} \frac{5}{9} & \text{if } a_k = Z \\ \frac{23}{27} & \text{if } a_k = A \end{cases}$$

The formula (8) is easily proved by induction using the following formulas which are special cases of (4):

$$\begin{aligned} \varphi(A*) &= 3^{n(A*)} (\frac{\varphi(*)}{3^{n(*)}} \tfrac{32}{27} + \tfrac{23}{27}) \\ \varphi(Z*) &= 3^{n(Z*)} (\frac{\varphi(*)}{3^{n(*)}} \tfrac{8}{9} + \tfrac{5}{9}) \end{aligned}$$

Here, $*$ denotes an arbitrary 0-1 sequence. From

$$\frac{2\varepsilon_Z(k) + 3\varepsilon_A(k)}{3\varepsilon_Z(k) + 5\varepsilon_A(k)} \geq \frac{n}{l} \tag{9}$$

we obtain

$$\varepsilon_Z(k) \geq \varepsilon_A(k) \, \frac{5n - 3l}{2l - 3n}$$

and on the other hand, from $\varepsilon_A(k) + \varepsilon_Z(k) = k - 1$ and (9) it follows that

$$\varepsilon_A(k) \leq (k-1) \, \frac{2l - 3n}{2n - l} \, .$$

Hence we can estimate as follows:

$$\begin{aligned} \frac{\varphi(\tilde{s})}{3^n} &= \sum_{i=1}^{m} (\frac{8}{9})^{\varepsilon_Z(m-i+1)} (\frac{32}{27})^{\varepsilon_A(m-i+1)} \chi(m-i+1) \\ &\leq \sum_{i=1}^{m} \left((\frac{8}{9})^{\frac{5n-3l}{2l-3n}} \frac{32}{27} \right)^{(m-i)\,\frac{2l-3n}{2n-l}} \chi(m-i+1) \\ &= \sum_{i=1}^{m} x^{1-i/m} \chi(m-i+1) \, . \end{aligned} \tag{10}$$

Notice that the sequence $x^{1-1/m}$ is positive and decreasing, since $x = \frac{2^l}{3^n} \in]1,3[$. On the other hand, we have for arbitrary $k \in \{1, \dots, m\}$

$$\sum_{i=1}^{k} \chi(m-i+1) = \frac{1}{27}(15E_Z(k) + 23E_A(k))$$

where $E_Z(k)$ and $E_A(k)$ denote the number of Z and A respectively in $(a_{m-k+1}, \dots, a_m)$. From

$$\frac{2E_Z(k) + 3E_A(k) + 1}{3E_Z(k) + 5E_A(k)} \geq \frac{n}{l}$$

and $E_Z(k) + E_A(k) = k$ we derive that

$$15E_Z(k) + 23E_A(k) \leq k\,\frac{6n+l}{m} + \frac{8l}{m}.$$

Thus

$$\sum_{i=1}^{k} \chi(m-i+1) \leq \sum_{i=1}^{k} d_i$$

if we set $d_1 := \alpha_1$ and $d_i := \alpha_2$ for $i \geq 2$. Application of Lemma 3 to (10) and summation of the resulting geometric series finishes the proof. □

Remark 1: Using similar arguments, one can estimate the growth of $M_{l,n}$. In fact, for all $l, n = n(l)$ there holds:

$$\frac{1}{20} \leq \frac{M_{l,n}}{n\,3^n} \leq \frac{7}{10}$$

and for $l > 10^3$ one can even replace $\frac{7}{10}$ by $\frac{1}{2}$.

Since we want to improve a criterion of Eliahou in the next section, it is now necessary to state one more technical lemma:

Lemma 6. *For all $x \in]1,3[$ and all $l > 160$, there holds for $\alpha = 0.9$, $n = n(l)$ and α_1, α_2, and m as in Proposition 2*

$$f(x) := \left(\alpha_1 x^{1-1/m} + \alpha_2 \frac{x^{1-1/m}-1}{x^{1/m}-1}\right) \frac{3(x^{1/n}-1)}{x-1} \leq \alpha$$

Remark 2: We will see in the proof that the value of α could be improved to $\lim_{z\to 1} f(z) = 0.842\dots$ for l large enough.

Proof

Using $l\log_3 2 - 1 \leq n \leq l\log_3 2$ we first observe that

$$\begin{aligned}
\alpha_1 &\leq \frac{1}{9}\,\frac{2\log 2 + 3\log 3}{2\log 2 - \log 3 - \frac{2\log 3}{l}} \to \frac{1}{9}\,\frac{2\log 2 + 3\log 3}{2\log 2 - \log 3} = 1.808\dots \\
\alpha_2 &\leq \frac{1}{27}\,\frac{6\log 2 + \log 3}{2\log 2 - \log 3 - \frac{2\log 3}{l}} \to \frac{1}{27}\,\frac{6\log 2 + \log 3}{2\log 2 - \log 3} = 0.676\dots
\end{aligned}$$

It can be checked elementary that (for $l > 160$) f is concave for $x > 1$. Hence, $f(x) \leq \lim_{z\to 1} f(z) + \lim_{z\to 1} f'(z)(x-1)$. Calculating these values, on gets

$$\lim_{z\to 1} f(z) = \frac{3}{n}(\alpha_1 + \alpha_2(m-1)) \to 3\alpha_2(2-\log_2 3) = 0.842\ldots \quad \text{as } l \to \infty$$

$$\lim_{z\to 1} f'(z) = \frac{3}{2n^2}(\alpha_1 + \alpha_2(m-1)) + \frac{3}{2mn}(\alpha_1(m-2) - 2\alpha_2(m-1)) \to 0 \quad \text{as } l \to \infty$$

and the claim follows easily from these facts. □

We close this section with a further monotonicity lemma. We show that $M_{l,n}$ is increasing in n for every fixed l:

Lemma 7. *If $n' < n$ then $M_{l,n'} < M_{l,n}$.*

Proof
We may assume that $n = n' + 1$. Let $s = \tilde{s}(l,n)$ and $s' = \tilde{s}(l,n')$ denote the 0-1 sequences as constructed in Lemma 5 with the pairs l,n and l,n' respectively. From the construction it follows that $t' = 1s' \in S_{l+1,n}$ dominates $t = 0s \in S_{l+1,n}$ in the sense that $\sum_{i=1}^{k} t_i \leq \sum_{i=1}^{k} t'_i$ for all $k \in \{1,\ldots,l+1\}$. Hence, Lemma 4 implies that $\varphi(t) > \varphi(t')$. Now (4) implies $\varphi(s') < \varphi(s) - \frac{1}{2}3^{n'}$ and this proves the claim. □

4. Comparison of different criteria

4.1. **Crandall's criterion.** The classical (implicit) estimate for the length of Collatz cycles is due to Crandall [1]:

Lemma 8 (Crandall)**.** *If $m > 1$ is the minimum of a Collatz cycle generated by a sequence $s \in S_{l,n}$ with $n < n(l)$, then*

$$m < \frac{n(3+\frac{1}{m})^{n-1}}{2^l - 3^n}.$$

Together with well-known facts about continued fractions this leads to Crandall's Theorem

Theorem 1 (Crandall)**.** *Let $\frac{p_k}{q_k}$ be convergents to $\log_2 3$ then for $k > 4$*

$$n > \min(q_k, \frac{2m}{q_k + q_{k+1}})$$

with m and n as in Lemma 8.

Using the fact, that the Collatz conjecture was verified for all initial values $x_0 \leq 2^{40}$, Lagarias [5] obtained as lower bound for the length of non-trivial Collatz cycles the value $l \geq 301\,994$.

4.2. **Eliahou's criterion.** In [3] Eliahou improved Crandall's estimate and obtained the following criterion.

Theorem 2 (Eliahou). *If $k(m)$ denotes the smallest integer k such that*

$$\frac{k}{n(k)} \leq \log_2(3 + \frac{1}{m})$$

then there holds for every Collatz cycle C in $\mathbb{N}$

$$|C| \geq k(\min C).$$

Using the theory of continued fractions Eliahou showed that $k(2^{40}) = 17\,087\,915$ (this can also be checked directly from the definition).

4.3. **Eliahou's criterion improved.**

Theorem 3. *For every positive Collatz cycle C in $\mathbb{Q}[(2)]$ there holds*

$$|C| \geq k(\frac{\min C}{\alpha})$$

provided $|C| > 160$.

Proof
Let C be a positive Collatz cycle generated by some $s \in S_{l,n}$, $l > 160$. Then by definition we have

$$\min C \leq \frac{M_{l,n}}{2^l - 3^n} \leq \frac{M_{l,n(l)}}{2^l - 3^{n(l)}} \tag{11}$$

where we used the monotonicity Lemma 7 in the last step. Using the estimate in Proposition 2 and Lemma 6, we have

$$\frac{M_{l,n(l)}}{2^l - 3^{n(l)}} \leq f(x)(3x^{1/n(l)} - 3)^{-1} \leq \alpha(2^{l/n(l)} - 3)^{-1}$$

where we used the notation of Lemma 6. From the last line and (11) we finally conclude

$$\frac{|C|}{n(|C|)} \leq \log_2(3 + \frac{\alpha}{\min C})$$

and the theorem is proved. □

Theorem 3 tells in particular, that Eliahou's estimates on the length of cycles (e.g. the one above in Section 4.3) are true with the weaker assumption that the Collatz conjecture only needs to be checked up to some value which is 10% smaller than the originally demanded value.

4.4. **Optimal criterion.** The following criterion (which is according to Lemma 5 optimal, since for every given cycle length there exists a cycle, namely the cycle generated by $\tilde{s}$ in Lemma 5, with equality in (12) below) has the disadvantage that it is relatively costly to check it. However we will include an example below which shows how to handle this problem.

Theorem 4. *If $L(m)$ denotes the smallest integer L such that for $n = n(L)$*

$$\frac{\sum_{j=1}^{L}(\lceil j\frac{n}{L}\rceil - \lceil (j-1)\frac{n}{L}\rceil 2^{j-1}3^{n-\lceil j\frac{n}{L}\rceil}}{2^L - 3^n} \geq m$$

then there holds for every positive Collatz cycle C in $\mathbb{Q}[(2)]$

$$|C| \geq L(\min C). \tag{12}$$

As an example, suppose that the Collatz conjecture is verified for all initial values $x_0 \leq m = 212\,366\,032\,807\,211$. This is about three times larger than the momentary world record of [7] which is $6.3\ 10^{13}$. We want to prove that than the length of a nontrivial Collatz cycle is at least $L = 102\,225\,496$.

First step: Using Theorem 2 or 3 and either the tables in [3] which list the function k or by direct computation, we conclude, that $L \geq 17\,087\,915$.

Second step: A simple *Mathematica* procedure can check that

$$\frac{n(l)}{l}\,\frac{1}{2 - 3^{n(l)/l}} < m$$

or alternatively that

$$\frac{1}{2}\,n(l) < m\,(\frac{2^l}{3^{n(l)}} - 1)$$

for all values of $l \in \{p_{16}, \ldots, p_{18}-1\} \setminus (A_1 \cup A_2)$ where $A_1 = \{k \cdot p_{16} : k = 1, \ldots, 5\}$ and $A_2 = \{kp_{16} + p_{14} : k = 3, 4, 5\}$ (where $p_{14} = 301\,994$, $p_{16} = 17\,087\,915$ and $p_{18} = 102\,225\,496$). The reason, why p_{17} is not dangerous is simply that $2^{p_{17}} - 3^{q_{17}}$ is negative. Hence $M_{l,n(l)}/(2^l - 3^{n(l)}) < m$ (according to Lemma 1 or Remark 1) for the mentioned set of values of l.

Third step: Using Corollary 1 we can check by direct computation, that

$$\frac{M_{l,n}}{2^l - 3^n} < m$$

for $l = p_{16}$ and $n = n(l)$.

Observe that $q_{16}\log_3 2 = q + \varepsilon$ holds for some number $q \in \mathbb{N}$ and $\varepsilon \in]0, 1/100[$ and this implies that $kn(q_{16}) = n(kq_{16})$ for $k = 1, 2, \ldots, 100$. Hence the sequence $\tilde{s}(l, n(l))$ constructed in Lemma 5 for $l = kq_{16}$ is just the concatenation of k copies of the sequence $\tilde{s}(l, n(l))$ constructed for $l = q_{16}$. Then from (4) it follows that

$$\frac{M_{kq_{16},n(kq_{16})}}{2^{kq_{16}} - 3^{n(kq_{16})}} = \frac{M_{q_{16},n(q_{16})}}{2^{q_{16}} - 3^{n(q_{16})}} < m$$

for $k = 1, 2, \ldots, 100$. This rules out in particular the whole set A_1 without further computation. The case $l \in A_2$ can either be handled with a similar argument or by direct verification.

Fourth step: From the first three steps we infer that

$$\frac{M_{l,n(l)}}{2^l - 3^{n(l)}} < m$$

for all $l \in \{1, \ldots, p_{18} - 1\}$. By the monotonicity Lemma 7 we finally conclude that in fact

$$\frac{M_{l,n}}{2^l - 3^n} < m$$

for all $l \in \{1, \ldots, p_{18} - 1\}$ and all $n \leq n(l)$. This proves the claim.

Notice that we would have used the value $m = 2.9\ 10^{14}$ in order to prove the same result by Eliahou's criterion.

5. Final remarks

All estimates on the length of Collatz cycles given so far are valid for rational Collatz cycles although they have been stated originally for integer Collatz cycles. But since we have seen that the minimum of rational cycles grows at least linearly in terms of their length (compare Remark 1), such an approach *cannot* be successful to prove (A). The only chance to achieve further progress in that question would hence involve number theoretical arguments, e.g. of the following kind:

If ρ_l denotes the right-shift permutation acting on S_l, i.e.

$$\rho_l\colon (s_1, \ldots, s_l) \mapsto (s_l, s_1, \ldots, s_2),$$

then we obtain from (4) the following formulas which describe the effect of ρ_l on φ:

$$\begin{aligned} \varphi(\rho_l(s0)) &= \varphi(0s) = 2\varphi(s0) \\ \varphi(\rho_l(s1)) &= \varphi(1s) = \tfrac{1}{3}(2\varphi(s1) + 3^{n(s1)} - 2^{l(s1)}) \end{aligned} \tag{13}$$

Then we have the following lemma.

Lemma 9. *Let $s \in S_l$, $s \neq (0, \ldots, 0)$, $\bar{s}1 \in \sigma(s)$ and $2^{l(s)} - 3^{n(s)} > 0$. Then*

$$\gcd(\varphi(\bar{s}1), \varphi(\rho_l(\bar{s}1))) = \gcd(\varphi(\bar{s}1), 2^{l(s)} - 3^{n(s)}) \tag{14}$$

and in particular

$$\gcd(\varphi(t) \,:\, t \in \sigma(s)) = \gcd(\varphi(s), 2^{l(s)} - 3^{n(s)}). \tag{15}$$

Proof
Let $x = \varphi(\bar{s}1)$, $l = l(s)$, $n = n(s)$. Then we have from (13) $\varphi(\rho_l(\bar{s}1)) = \frac{1}{3}(2x + 3^n - 2^l)$ and hence

$$\begin{aligned} \gcd(x, \tfrac{1}{3}(2x + 3^n - 2^l)) &= \gcd(x, 2x + 3^n - 2^l) = \\ &= \gcd(x, 3^n - 2^l) \end{aligned}$$

Note that in the first line above we used that $\frac{1}{3}(2x + 3^n - 2^l) \in \mathbb{N}$ implies that x is not a multiple of 3. Observing then that

$$\gcd(\varphi(s'0), \varphi(\rho_l(s'0))) = \gcd(\varphi(s'0), 2\varphi(s'0)) = \varphi(s'0)$$

we get (15). □

Using Lemma 9 we obtain that $\neg$(A) is equivalent to

(A′) There exists a non-periodic $s \in S_l$, $l > 3$, such that

$$\gcd\{\varphi(t) \,:\, t \in \sigma(s)\} = 2^{l(s)} - 3^{n(s)}.$$

Acknowledgement: We would like to thank Jeffrey Lagarias and Douglas Rogers for their useful hints and comments about this subject as well as Gary Leavens and Mike Vermeulen for communicating their result [7]. We are grateful to the referee for several bibliographical hints.

References

[1] R. E. Crandall: On the $3x + 1$ problem. Math. Comp. **32** (1978), 1281–1292
[2] J. M. Dolan, A. F. Gilman, S. Manickam: A generalization of Everett's result on the Collatz $3x + 1$ problem. Adv. in Appl. Math. **8** (1987), 405–409
[3] S. Eliahou: The $3x + 1$ problem: New lower bounds on nontrivial cycle lengths. Discrete Math. **118** (1993), 45–56
[4] I. Krasikov: How many numbers satisfy the $3x + 1$ conjecture? Internat. J. Math. Sci. **12(4)** (1989), 791–796
[5] J. C. Lagarias: The $3x + 1$-problem and its generalizations. American Math. Monthly **92** (1985), 3–23
[6] J. C. Lagarias: The set of rational cycles for the $3x + 1$ problem. Acta Arith. **56** (1990), 33–53
[7] G. Leavens, M. Vermeulen: private communication
[8] J. W. Sander: On the $(3N + 1)$-conjecture. Acta Arith. **55** (1990), 241–248
[9] B. G. Seifert: On the arithmetic of cycles for the Collatz-Hasse (Syracuse) conjectures. Discrete Math. **68** (1988), 293–298
[10] G. Wirsching: An improved estimate concerning $3n + 1$ predecessor sets. Acta Arith. **63** (1993), 205–210

Lorenz Halbeisen, Mathematik Departement, ETH Zürich, CH-8092 Zürich (Switzerland)

E-mail address: `halbeis@@math.ethz.ch`

Norbert Hungerbühler, University of Minnesota, IMA, 514 Vincent Hall, 206 Church Street S.E., Minneapolis, MN 55455 (U.S.A.)

E-mail address: `hungerbu@@ima.umn.edu`