ACTA ARITHMETICA
131.3 (2008)

On the (non-)existence of m-cycles for generalized Syracuse sequences

by

JOHN L. SIMONS (Groningen)

1. INTRODUCTION

The $3x+1$ *problem* (also called the *Syracuse problem* or the *Collatz problem*) is defined by a sequence of natural numbers, generated conditionally by $x_{n+1} = \frac{1}{2}x_n$ if x_n is even, and $x_{n+1} = \frac{1}{2}(3x_n + 1)$ if x_n is odd. A famous conjecture says that for all natural numbers x_0 finally the cycle $(1, 2)$ appears. A formal proof is lacking so far in spite of various approaches to the problem [10], [12], [25]. Recently Simons and de Weger [23] proved the non-existence of m-cycles ($m \le 75$) and their approach appears to be generalizable to problems that are similar to the $3x + 1$ problem. This article is structured as follows:

- We first analyze in Section 2 the main line of the approach of Simons and de Weger to the $3x+1$ problem, to see where the special character of the $3x + 1$ problem is utilized and how this can be generalized. We discuss the behaviour of convergents and post-transcendence conditions.
- Secondly, in Sections 3 and 4 we discuss simple generalizations with binary conditions, i.e. the $3x + q$, $px + 1$ and the $px + q$ problem with $\mathrm{GCD}(p, q) = 1$.
- Thirdly, in Sections 5 and 6 we discuss generalizations with ternary (and higher) conditions, i.e. sequences generated by $x_{i+1} = (p_j x_i + q_j)/d$ if $x_i \equiv r_i \pmod{d}$.
- In Section 7 we discuss the Roelants problem [19]. Let p_k be the kth prime. Now a sequence is defined by $x_{n+1} = x_n/p_j$ while $x_n \equiv 0 \pmod{p_j}$ for $j = 1, \ldots, k$, and $x_{n+1} = (p_{k+1}x_n + 1)/2$ else.
- We conclude with general findings and theoretical and computational limitations of the presented approach.

2000 *Mathematics Subject Classification*: 11B83, 11J86.

Key words and phrases: generalized Collatz problem, cycle existence, linear forms in logarithms.

2. THE APPROACH OF SIMONS AND DE WEGER TO THE $3x + 1$ PROBLEM

2.1. Definitions and basic notation. Suppose there exists a cycle for the $3x + 1$ problem. Let there be m local minima $\overline{x}_i$ in the cycle, with indices $t_0, t_1, \ldots, t_{m-1}$. Then there are also m local maxima $\overline{y}_i$, with indices $s_0, s_1, \ldots, s_{m-1}$, any of which lies between two minima. We call such a cycle, that consists of m pairs of increasing sequences with k_i numbers and decreasing sequences with l_i numbers, an *m-cycle*. Further we put

$$K = \sum_{i=0}^{m-1} k_i \quad \text{and} \quad L = \sum_{i=0}^{m-1} l_i.$$

Throughout this section the quantities

$$\Lambda = (K + L)\log 2 - K \log 3 \quad \text{and} \quad \delta = \log_2 3$$

play an important role. To exclude the trivial m-cycle $(1, 2, 1, 2, \ldots)$ we assume that $x_i \geq x_0 > X_0 > 2$, where X_0 is a lower bound from exterior computations [17], [20]. Note that x_i is the ith number and $\overline{x}_i$ is the ith local minimum in a hypothetical cycle.

2.2. Main line of the approach. Simons and de Weger's approach to proving the non-existence of m-cycles for $m \leq 75$ contains the following steps. See [23] for a formal proof.

1. *Derivation of an upper bound for Λ in terms of the local minima $\overline{x}_i$ in a cycle*

(a) Definition of increasing and decreasing subsequences and choice of an appropriate expression for numbers in subsequences. An odd number x_0 can be written as $a_0 2^{k_0} - 1$, which leads to an increasing subsequence of k_0 odd numbers. Then the first even number is $a_0 3^{k_0} - 1$.

(b) Derivation of chain equations between subsequences, resulting in a diophantine matrix equation for the coefficients of such expressions. The chain equation from the ith decreasing subsequence with l_i elements to the $(i + 1)$th increasing subsequence with k_{i+1} elements is $(a_i 3^{k_i} - 1)/2^{l_i} = a_{i+1} 2^{k_{i+1}} - 1$. All chain equations together result in a diophantine system in the coefficients a_i,

$$\begin{pmatrix} -3^{k_0} & 2^{k_1+l_0} & & \\ & -3^{k_1} & 2^{k_2+l_1} & \\ & & \ddots & \ddots \\ 2^{k_0+l_{m-1}} & & & -3^{k_{m-1}} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{m-1} \end{pmatrix} = \begin{pmatrix} 2^{l_0} - 1 \\ 2^{l_1} - 1 \\ \vdots \\ 2^{l_{m-1}} - 1 \end{pmatrix}. \tag{1}$$

An integer solution for the coefficients a_i is a necessary and sufficient condition for the existence of an m-cycle.

(c) Derivation of an inequality for Λ. From the chain equation we find

$$\frac{2^{k_{i+1}+l_i}}{3^{k_i}}\,\frac{a_{i+1}}{a_i} = 1 + \frac{2^{l_i}-1}{3^{k_i}a_i}.$$

Taking the product over all $i = 0, 1, \ldots, m-1$, and using the cyclicity, we get

$$\frac{2^{K+L}}{3^K} = \prod_{i=0}^{m-1}\left(1 + \frac{2^{l_i}-1}{3^{k_i}a_i}\right).$$

Applying $\log(1+x) < x$ to each term in the product, we obtain

$$(2)\qquad 0<\Lambda<\sum_{i=0}^{m-1}\log\left(1+\frac{2^{l_i}-1}{3^{k_i}a_i}\right)<\sum_{i=0}^{m-1}\frac{2^{l_i}-1}{3^{k_i}a_i}<\sum_{i=0}^{m-1}\frac{1}{a_i 2^{k_i}-1}=\sum_{i=0}^{m-1}\frac{1}{\overline{x}_i}.$$

2. *Derivation of an upper bound for Λ in terms of K and m*

(a) Derivation of a lower bound for the maximal $\overline{x}_i$ in a cycle in terms of K and m. From $\overline{x}_i > 2$ it follows that $\overline{x}_i > a_i 2^{k_i-1}$. For the maximal $\overline{x}_i$ in a cycle we find

$$(3)\qquad [\max(\overline{x}_i)]^m > \prod_{i=0}^{m-1}\overline{x}_i > \prod_{i=0}^{m-1} a_i 2^{k_i-1} \geq 2^{K-m}.$$

Consequently, $\max(\overline{x}_i) > 2^{K/m-1}$.

(b) Chaining of magnitudes of local minima in an m-cycle. Based on the chain equation $-3^{k_i}a_i + 2^{k_{i+1}+l_i}a_{i+1} = 2^{l_i}-1$, it can be proved that local minima are approximately of the same size. In particular, $\overline{x}_{i+1} < 2^{\delta-1}\overline{x}_i^{\delta}$ for $i = 0, \ldots, m-1$. This last inequality can be rewritten as

$$(4)\qquad \overline{x}_i^{-1} < 2^{(\delta-1)/\delta}\overline{x}_{i+1}^{-1/\delta} = c\overline{x}_{i+1}^{-1/\delta},$$

from which we derive (taking into account that the worst case appears if $\overline{x}_i < \overline{x}_{i+1}$ for $i = 0, \ldots, m-2$)

$$(5)\qquad \overline{x}_0^{-1} < c^{1+1/\delta+\cdots+1/\delta^{m-2}}\overline{x}_{m-1}^{-1/\delta^{m-1}} < c^{\delta/(\delta-1)}\overline{x}_{m-1}^{-1/\delta^{m-1}} = 2\overline{x}_{m-1}^{-1/\delta^{m-1}}.$$

(c) Approximation of the upper bound for Λ. Inserting (3) and (5) into (2) leads to an upper bound for Λ in terms of K and m:

$$(6)\qquad 0 < \Lambda < \sum_{i=0}^{m-1}\frac{1}{\overline{x}_i} < \frac{m}{\overline{x}_0} < 2^{1+1/\delta^{m-1}}\cdot m\cdot 2^{-K/m\delta^{m-1}}.$$

For large K this upper bound is smaller than $1/K^{2+\varepsilon}$ so according to Roth's lemma [8] there are, for fixed m, a finite number of solutions of (6). For an effective search of these solutions an operational (lower and) upper bound for K is required.

3. *Derivation of a lower bound for K as a function of m.* From (2) a lower bound for K can be derived that applies to specific m-cycles. We have a generalization of a lemma of Crandall [3].

LEMMA 1. *Let p_n/q_n be the nth convergent to δ. Suppose there exists an m-cycle with K odd and L even numbers $x_i \geq x_0 > X_0$. If $q_n + q_{n+1} \leq (\log 2)X_0/m$ then $K \geq K_0(m) = q_{n+1}$.*

Proof. Since $x_i \geq x_0 > X_0$ we find, according to (2),

$$\Lambda < \sum_{i=0}^{m-1} \frac{1}{\overline{x}_i} < \frac{m}{X_0}. \tag{7}$$

Assume $K < q_{n+1}$. From elementary number theory (implicit result of the proof of Theorem 182 of [7]) it follows that

$$\Lambda = (\log 2)|(K+L) - K\delta| \geq (\log 2)|p_n - q_n\delta| > \frac{\log 2}{q_n + q_{n+1}} \geq \frac{m}{X_0}, \tag{8}$$

which contradicts (7). ∎

From distributed computations [17] it is known that for the $3x+1$ problem $X_0 > 2.5 \cdot 2^{60} > 2.88 \cdot 10^{18}$. Lemma 1 implies that if $m \leq 100$ then $q_{31} + q_{32} < 1.2 \cdot 10^{16} < (\log 2)X_0/m$ [23] and consequently $K > 6.2 \cdot 10^{13} \cdot m$. Inserting these values into (7) leads to $\delta K < K + L < (\delta + 8 \cdot 10^{-34})K$, thus $(K+L)/K$ must be a (very) good approximation of δ.

4. *Derivation of an upper bound for K as a function of m*

(a) Application of transcendence theory. From (6) it follows that Λ is exponentially small as a function of K. Transcendence theory (originally a theorem of Baker [1], later refined by Laurent, Mignotte and Nesterenko [13]) states that Λ, being a linear form in logarithms, cannot be too small as a function of K once it is positive. The best result today is due to Rhin [18]:

$$\Lambda > e^{-13.3(0.46057 + \log K)}. \tag{9}$$

(b) Confronting lower and upper bound for Λ. Estimates (6) and (9) provide an upper bound for K. Let $x = K_3(m)$ be the largest solution of

$$e^{-13.3(0.46057 + \log x)} = 2^{1+1/\delta^{m-1}} \cdot m \cdot 2^{-x/m\delta^{m-1}}. \tag{10}$$

Then $K < K_3(m)$.

5. *Upper bound reduction.* Let $x = K_1(m)$ be the largest solution of

$$2^{1+1/\delta^{m-1}} \cdot m \cdot 2^{-x/m\delta^{m-1}} = \frac{\log 2}{2x}. \tag{11}$$

Then from elementary number theory, if $K \geq K_1$ then $(K+L)/K$ must be a convergent of the continued fraction expansion of δ. For convergents we have

$$\frac{\log 2}{(a_{n+1}+2)q_n} < \Lambda < \frac{\log 2}{(a_{n+1})q_n}. \tag{12}$$

From [23] we know that the partial quotients satisfy $a_n \leq 55$ for $n \leq 229$. Note that for $m \sim 500$ we have $K_3(m) \simeq 10^{106} < q_{229} \simeq 10^{114}$. For larger values of (n, m) a similar analysis applies. Let $x = K_2(m)$ be the largest solution of

$$2^{1+1/\delta^{m-1}} \cdot m \cdot 2^{-x/m\delta^{m-1}} = \frac{\log 2}{57x}. \tag{13}$$

Then for $q_n = K \geq K_2(m)$ we have

$$\Lambda < 2^{1+1/\delta^{m-1}} \cdot m \cdot 2^{-K/m\delta^{m-1}} < \frac{\log 2}{57K} \leq \frac{\log 2}{(a_{n+1}+2)K}, \tag{14}$$

which contradicts (12). So $K < K_2(m)$.

Hence two feasible intervals remain for K:

- $K_0 < K < K_1$ in which (non-)convergent solutions K may exist.
- $K_1 \leq K < K_2$ in which (only) convergent solutions K may exist.

If $K_0(m) \geq K_2(m)$ then m-cycles cannot exist. If $K_0(m) < K_2(m)$ then lattice basis reduction techniques may eliminate solutions. Lattice basis reduction ([24, Ch. 3]) is an advanced technique to find/exclude solutions $x_i \in \mathbb{Z}$ of $0 < a_0 + \sum_{i=1}^{n} x_i a_i < \varepsilon$ when $|a_i| < A$.

2.3. Beyond the transcendence conditions for a solution. If $m \geq 76$ the lattice basis reduction algorithm yields K, L values that satisfy the necessary condition for the existence of an m-cycle. Let (K, L, m, X_0) be a quadruple that passes the conditions of transcendence theory. Then in theory all the partitions $(k_0, k_1, \ldots, k_{m-1})$ must be checked for the existence of integer solutions for a_i. A weaker (necessary) check is the condition $a_i \geq 1$. Simons [21] showed (no formal proof) that $\min(a_i) = 1/d$ is maximal for non-integer values of k_i and integer values for l_i defined by $2^{k_{i+1}+1} - 3^{k_i} = d$ and $l_i = 1$ for $i = 0, \ldots, m-2$ and $l_{m-1} = L - (m-1)$, where d is a constant depending on (K, L, m, X_0). He further proves that for (almost) each rounding of k_i to integer values at least one a_i decreases. By calculating d and checking $d > 1$ the non-existence of certain (not all) m-cycles for $m \geq 76$ can be proved.

2.4. Remarks. We presented a somewhat modified approach compared to the approach in [23].

1. We use a sharper generalization of Crandall's lemma which results in a larger (stronger) lower bound K_0.

2. We use a weaker formulation for the lower bound for $\prod \overline{x}_i$ which results in a larger (weaker) upper bound K_3.
3. We use a simpler (weaker) upper bound reduction analysis from $K_3(m)$ to $K_2(m)$ which is valid for $m < 500$.
4. From the condition $a_i \geq 1$ we can eliminate (some) hypothetical solutions that pass the transcendence conditions.
5. As will become clear later (see Lemma 13), $px + q$ sequences can generate extra conditions for the cycle length of hypothetical m-cycles. These conditions are ineffective for the $3x + 1$ problem.

3. THE $3x + q$ PROBLEM

3.1. Introduction. The $3x+q$ problem is defined by a sequence, generated conditionally by $x_{n+1} = \frac{1}{2}x_n$ if x_n is even, and $x_{n+1} = \frac{1}{2}(3x_n+q)$ if x_n is odd. If q is composite and $r \mid q$ then each cycle of the $3x+r$ $(px+r)$ problem corresponds to a cycle of the $3x + q$ $(px + q)$ problem. We therefore restrict ourselves to sequences with $q = 1$ or $q \geq 5$ prime and $\mathrm{GCD}(x_i, q) = 1$. Lagarias [11] calls such cycles *primitive* and formulates two conjectures. Let $C_{\mathrm{prim}}(q)$ be the number of primitive cycles. Then $C_{\mathrm{prim}}(q) \geq 1$ and $C_{\mathrm{prim}}(q) < \infty$. Lagarias gives the following numerical results for primitive cycles of the $3x+q$ problem with cycle length $K+L$ starting with minimal x_0.

Smallest element and cycle length for the $3x + q$ problem

q	x_0	$K+L$	q	x_0	$K+L$	q	x_0	$K+L$
1	1	2	5	1	3	7	5	4
11	1	6	13	1	4	17	1	7
19	5	11	23	5	5	29	1	5
31	13	23	37	19	6	41	1	20
43	1	11	47	5	18	53	103	29
59	1	28	61	1	6	67	17	30
71	29	10	73	19	60	79	1	44
83	65	24	89	17	17	97	1	18

We computed the values of K and L for those examples to find that always $2^{K+L} - 3^K \equiv 0 \pmod{q}$, e.g. the $3x + 31$ problem has the 6-cycle $(13 \ldots 17 \ldots 53 \ldots 79 \ldots 67 \ldots 29 \ldots)$, with $K = 12$ and $L = 11$. We have $2^{23} - 3^{12} = 31 \cdot 191 \cdot 1327$. In the rest of this section we will assume that $x_i \geq x_0 > X_0 = 10^6 > mq$ for ease of analysis.

3.2. An upper bound for Λ

1. *Derivation of an upper bound for Λ in terms of the numbers in a cycle*

(a) Definition of increasing and decreasing subsequences and choice of an appropriate expression for numbers in subsequences. An odd number can be written as $x_0 = a_0 2^{k_0} - q$. Consequently, the next number is $a_0 2^{k_0} 3 - q$. After an increasing subsequence of k_0 odd numbers the first appearing even number is $a_0 3^{k_0} - q$, which is the beginning of a decreasing subsequence.

(b) Derivation of chain equations between subsequences, resulting in a diophantine matrix equation for the coefficients of such expressions. The chain equation from the ith decreasing subsequence with l_i elements to the $(i+1)$th increasing subsequence with k_{i+1} elements is $(a_i 3^{k_i} - q)/2^{l_i} = a_{i+1} 2^{k_{i+1}} - q$. All chain equations together result in a diophantine system in the coefficients a_i. A necessary and sufficient condition for the existence of an m-cycle is the existence of a solution (a_i, k_i, l_i) of the diophantine system of equations

$$(15)\qquad \begin{pmatrix} -3^{k_0} & 2^{k_1+l_0} & & \\ & -3^{k_1} & 2^{k_2+l_1} & \\ & & \ddots & \ddots \\ 2^{k_0+l_{m-1}} & & & -3^{k_{m-1}} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{m-1} \end{pmatrix} = \begin{pmatrix} q(2^{l_0}-1) \\ q(2^{l_1}-1) \\ \vdots \\ q(2^{l_{m-1}}-1) \end{pmatrix}.$$

(c) Derivation of an inequality for Λ. Multiplication of all ratios a_{i+1}/a_i leads to an inequality for Λ:

$$(16)\qquad 0 < (K+L)\log 2 - K\log 3 < \sum_{i=0}^{m-1} \frac{q}{a_i 2^{k_i} - q} = \sum_{i=0}^{m-1} \frac{q}{\overline{x}_i}.$$

2. *Derivation of an upper bound for Λ in terms of K and m*

(a) Derivation of a lower bound for the maximal $\overline{x}_i$ in a cycle in terms of K and m. Since $2\overline{x}_i > \overline{x}_i + q = a_i 2^{k_i}$, for the maximal $\overline{x}_i$ in a cycle we find

$$(17)\qquad [\max(\overline{x}_i)]^m > \prod_{i=0}^{m-1} \overline{x}_i > \prod_{i=0}^{m-1} a_i 2^{k_i-1} \geq 2^{K-m}.$$

Consequently, $\max(\overline{x}_i) > 2^{K/m-1}$.

(b) Chaining of magnitudes of minima in a cycle. We have $\overline{y}_i = a_i 3^{k_i} - q$ and thus

$$\overline{x}_{i+1} \leq \frac{\overline{y}_i}{2} < \frac{\overline{y}_i + q}{2} = \frac{a_i 3^{k_i}}{2} = \left(\frac{3}{2}\right)^{k_i} \frac{a_i 2^{k_i}}{2} \leq (a_i 2^{k_i})^{\delta-1} \frac{a_i 2^{k_i}}{2}$$
$$= \frac{(\overline{x}_i + q)^\delta}{2} = \left[1 + \frac{q}{\overline{x}_i}\right]^\delta \frac{\overline{x}_i^\delta}{2} < \left[1 + \frac{q}{X_0}\right]^\delta \frac{\overline{x}_i^\delta}{2} < 2^{\delta-1}\overline{x}_i^\delta.$$

(c) Approximation of the upper bound for Λ. This last inequality can be rewritten as

$$(18)\qquad \overline{x}_i^{-1} < 2^{(\delta-1)/\delta} \overline{x}_{i+1}^{-1/\delta} = c\overline{x}_{i+1}^{-1/\delta},$$

from which we derive (taking into account that the worst case appears if $\overline{x}_i < \overline{x}_{i+1}$ for $i = 0, \ldots, m-2$)

$$(19)\qquad \overline{x}_0^{-1} < c^{1+1/\delta+\cdots+1/\delta^{m-2}}\overline{x}_{m-1}^{-1/\delta^{m-1}} < c^{\delta/(\delta-1)}\overline{x}_{m-1}^{-1/\delta^{m-1}} = 2\overline{x}_{m-1}^{-1/\delta^{m-1}}.$$

Substitution of these results into (16) leads to a lower bound for Λ in terms of K and m:

$$(20)\qquad \Lambda < \sum_{i=0}^{m-1} \frac{q}{\overline{x}_i} < \frac{qm}{\overline{x}_0} < 2^{1+1/\delta^{m-1}} \cdot qm \cdot 2^{-K/m\delta^{m-1}}.$$

Note that this is a generalization of (6).

3.3. A lower bound and upper bound for K. Because of the numerator q in (16), we find, similar to Lemma 1:

LEMMA 2. *$K \geq K_0(m,q) = q_{n+1}$ with n being the maximum index for which*

$$q_n + q_{n+1} \leq \log 2 \cdot \frac{X_0}{qm}.$$

Proof. Assume $K < q_{n+1}$. Then from elementary number theory we have

$$\Lambda = (\log 2)|(K+L) - K\delta| \geq (\log 2)|p_n - q_n\delta| > \frac{\log 2}{q_n + q_{n+1}} \geq \frac{qm}{X_0} > \sum_{i=0}^{m-1} \frac{q}{\overline{x}_i},$$

which contradicts (16). ∎

Applying this, for each m, with the maximal n that satisfies the condition, we find that the lower bound K_0 is a stepwise decreasing function of m and q.

COROLLARY 3. *For $q = 5$ we have:*

- *If $m = 1$ then $K \geq q_{11} = 79\,335$.*
- *If $m = 2$ then $K \geq q_{10} = 31\,867$.*
- *If $3 \leq m \leq 8$ then $K \geq q_9 = 15\,601$.*
- *If $9 \leq m \leq 142$ then $K \geq q_8 = 665$.*
- *If $143 \leq m \leq 386$ then $K \geq q_7 = 306$.*

COROLLARY 4. *For $q = 97$ we have:*

- *If $m \leq 7$ then $K \geq q_8 = 665$.*
- *If $8 \leq m \leq 19$ then $K \geq q_7 = 306$.*
- *If $20 \leq m \leq 76$ then $K \geq q_6 = 53$.*

Since Λ (for the $3x + q$ problem) does not depend on q, inequality (9) applies and we have:

LEMMA 5. *Let $x = K_3(m,q)$ be the largest solution of*

$$(21)\qquad e^{-13.3(0.46057+\log x)} = 2^{1+1/\delta^{m-1}} \cdot qm \cdot 2^{-x/m\delta^{m-1}}.$$

Then $K < K_3(m,q)$.

Let $c_3(m) = K_3(m)/m^2\delta^m$. Then $c_3(m)$ is a decreasing function of m that tends to a constant (independent of X_0 and for $q \le 97$ independent of q) limit value 5.576 as $m \to \infty$.

(22)

K_3-bounds for the $3x + q$ problem

m	$q = 5$		$q = 97$	
	c_3	$K_3 <$	c_3	$K_3 <$
1	64.28	101.9	67.59	107.1
2	40.76	409.6	42.34	425.5
5	22.79	5697.7	23.39	5848.5
10	15.44	154488.1	15.73	157415.8
100	7.05	$7.08 \cdot 10^{24}$	7.07	$7.10 \cdot 10^{24}$
1 000	5.78	$6.04 \cdot 10^{206}$	5.78	$6.04 \cdot 10^{206}$
10 000	5.60	$8.67 \cdot 10^{2008}$	5.60	$8.67 \cdot 10^{2008}$
100 000	5.58	$4.42 \cdot 10^{20012}$	5.59	$4.42 \cdot 10^{20012}$
1 000 000	5.58	$5.47 \cdot 10^{200031}$	5.58	$5.47 \cdot 10^{200031}$

Thus, for fixed q and fixed m, the $3x + q$ problem has a finite number of m-cycles. If $K_0(m, q) \ge K_3(m, q)$ then m-cycles cannot exist. Note that $K_0(m, q)$ heavily depends on X_0 and that $K_3(m, q)$ is an increasing function of m and q. Hence m-cycles may exist for large(r) values of m and q. Let m^+ be the largest value of m for which $K_0(m, q) \ge K_3(m, q)$. We calculated Corollaries 3, 4 and table (22) for $q \le 97$ to find that $m^+ = 6$ for $q = 5, 7$, $m^+ = 3$ for $q = 11, 13$ and $m^+ = 2$ for $q = 17, \ldots, 97$.

3.4. Upper bound reduction. We now use a continued fraction argument as described earlier. Let $x = K_1(m, q)$ be the largest solution of

$$2^{1+1/\delta^{m-1}} \cdot qm \cdot 2^{-x/m\delta^{m-1}} = \frac{\log 2}{2x}. \tag{23}$$

Let $c_1(m) = K_1(m)/m^2\delta^m$. Then $c_1(m)$ is a decreasing function of m that tends to a constant (independent of X_0 and for $q \le 97$ independent of q) limit value 0.419 as $m \to \infty$. If $K > K_1$ then $(K + L)/K$ must be a convergent of δ. Let $x = K_2(m, q)$ be the largest solution of

$$2^{1+1/\delta^{m-1}} \cdot qm \cdot 2^{-x/m\delta^{m-1}} = \frac{\log 2}{57x}. \tag{24}$$

Then $K < K_2(m, q)$. Let $c_2(m) = K_2(m)/m^2\delta^m$. Then $c_2(m)$ is a decreasing function of m that tends to a constant (independent of X_0 and marginally q-dependent) limit value $\simeq 2$ as $m \to \infty$. The reduction factor from K_3 to K_2 is $c_2/c_3 \simeq 0.12$.

K_i-bounds for the $3x+q$ problem

(25)

	$m=1$				$m=10$			
q	$K_0 \geq$	K_1	K_2	K_3	$K_0 \geq$	K_1	K_2	K_3
5	79335	9.02	14.55	101.88	665	13852.66	17094.70	154488.10
97	665	13.93	19.23	107.12	306	16724.51	19934.96	157415.76

We calculated $m^*(q)$, the largest value for which $K_0(m,q) > K_2(m,q)$, to find

Corollary 6. *If the $3x+q$ problem has an m-cycle with $x_i > X_0$ then $m > m^*$ of table* (26).

(26)

q	m^*
$q=5$	8
$q=7$	6
$11 \leq q \leq 97$	4

3.5. Elimination of solutions. The classic approach of Simons and de Weger to exclude m-cycles for $m > m^*$ now proceeds with applying a lattice basis reduction algorithm in the interval (K_0, K_2). The $3x+q$ problem ($q > 1$) reveals an extra condition on K and L. As a result of Lemma 13 (see next section) we have: If for the $3x+q$ problem an m-cycle exists, then $2^{K+L} - 3^K \equiv 0 \pmod{q}$. We checked if (K, L) pairs satisfy two conditions

$$(K+L)\log 2 - K\log 3 < \frac{qm}{X_0} \tag{27}$$

and

$$2^{K+L} - 3^K \equiv 0 \pmod{q}. \tag{28}$$

Condition (27) relaxes for increasing m. We found for $m = 10$ that the following hypothetical solutions satisfy these necessary conditions (27) and (28) for the existence of m-cycles:

(29)

q	K	L
5	15601	9126
17	13606	7959
23	12941	7570
41	13606	7959
43	14271	8348
47	9616	5625
53	11611	6792
71	6291	3680
97	7621	4458

As a result we have

LEMMA 7. *The $3x+q$ problem with $5 \leq q \leq 97$ has for $m \leq 10$ and $x_i > X_0$ no m-cycles besides hypothetically those listed in table* (29).

3.6. Numerical results for m-cycles. For the $3x+q$ problem we calculated all m-cycles with minimal element $x_0 \leq X_0$. This computation on a 2.66 Mhz processor took about 50 CPU seconds per q-value.

Number of m-cycles, cycle length, m and minimal element in cycle

(30)

q	n_c	(cycle length, m, $x_{\min}$)
1	1	$(2,1,1)$
5	5	$(3,1,1)$ $(5,1,19)$ $(5,2,23)$ $(27,6,187)$ $(27,7,347)$
7	1	$(4,1,5)$
11	2	$(6,1,1)$ $(14,4,13)$
13	9	$(4,1,1)$ $(24,5,131)$ $(8,1,211)$ $(8,2,259)$ $(8,2,227)$ $(8,2,287)$ $(8,2,251)$ $(8,3,283)$ $(8,3,319)$
17	2	$(7,2,1)$ $(31,9,23)$
19	1	$(11,3,5)$
23	3	$(43,11,41)$ $(5,1,50)$ $(5,2,7)$
29	4	$(5,1,1)$ $(17,5,11)$ $(65,13,3811)$ $(65,16,7055)$
31	1	$(23,6,13)$
37	3	$(6,1,19)$ $(6,2,23)$ $(6,2,29)$
41	1	$(20,4,1)$
43	1	$(11,2,1)$
47	7	$(28,7,25)$ $(18,5,5)$ $(7,1,65)$ $(7,2,89)$ $(7,2,73)$ $(7,2,85)$ $(7,3,101)$
53	1	$(29,9,103)$
59	7	$(28,6,1)$ $(10,1,133)$ $(10,2,181)$ $(10,2,185)$ $(10,3,217)$ $(10,2,149)$ $(10,4,221)$
61	2	$(6,1,1)$ $(66,19,235)$
67	1	$(30,6,17)$
71	7	$(10,3,29)$ $(10,2,31)$ $(27,7,4409)$ $(27,6,2809)$ $(27,7,3985)$ $(27,5,2585)$ $(27,8,4121)$
73	3	$(60,15,19)$ $(12,3,5)$ $(15,4,47)$
79	4	$(44,10,1)$ $(44,12,7)$ $(23,6,233)$ $(23,6,265)$
83	3	$(12,2,109)$ $(12,3,157)$ $(24,6,65)$
89	1	$(17,5,17)$
97	2	$(18,5,1)$ $(9,3,13)$

Thus we have

LEMMA 8. *The $3x+q$ problem with $5 \leq q \leq 97$ (prime) has for $m \leq 10$ no m-cycles besides those listed in table* (29) (*hypothetically*) *and in table* (30).

3.7. Remarks. 1. Because condition (27) sharpens for decreasing m, we can calculate the largest m for which table (29) has no entries. We have $K_2(8, 97) = 5938 < \min(K)$ in table (29), so this happens when $m = 8$. An alternative for Lemma 8 is the exclusion of m-cycles with $m \leq 8$ and referring to table (30) only.

2. Condition (27) relaxes for increasing m. As an example we found for $q = 5$ and $m = 16$ hypothetical pairs ranging from $(15601, 9126)$ to $(379744, 222136)$.

3. For the $3x + 11$ problem we have $2^6 - 3^2 = 55 = 0 \pmod{11}$ and $x_0 = (3^2 - 2^2)/5 = 1$. As a result of Lemma 18 (see next section) we find the cycle $(1, 7, 16, 8, 4, 2)$. This explains the numerical observation in the introduction of this section. Lemma 13 "facilitates" the existence of cycles for large values of q, e.g. the $3x + 17389$ problem has 61 cycles with cycle lengths 46, 92 and 138, while the $3x + 17393$ problem has just one cycle with cycle length 898. The occurrence of many cycles with the same length is explained by m variance, e.g. when for the $3x+13$ problem $K = 5, L = 3$, there (skipping all permutations) can exist one 1-cycle, four 2-cycles ($k_0 = 1$, $l_0 = 1$; $k_0 = 1, l_0 = 2$; $k_0 = 2, l_0 = 1$; $k_0 = 2, l_0 = 2$) and two 3-cycles ($k_0 = 1, k_1 = 1$; $k_0 = 1, k_1 = 2$). See table (30).

4. Corollary 19 (see next section) provides a necessary and sufficient condition for the existence of a 1-cycle with length $k + l$. If (and only if) $x_0 = (3^k - 2^k)q/(2^{k+l} - 3^k)$ is an integer then the $3x+q$ problem has a 1-cycle with minimal element x_0. If q is a Tijdeman prime that can be written as $q = 2^v - 3^w$ with $v > 2$, then the $3x + q$ problem has a primitive 1-cycle. The first such primes are: $5, 7, 23, 29, 31, 37, 47, 61, 101, 127, 229, \ldots$. If q is a non-Tijdeman prime, then the $3x + q$ problem can have a primitive 1-cycle. E.g. the $3x + 59$ problem has the 1-cycle $(133, 229, 373, 589, 913, 1399, 2128, \ldots, 266)$.

5. Matthews [15] and others studied sequences generated conditionally by $x_{n+1} = \frac{1}{2}x_n + q_1$ if x_n is even, and $x_{n+1} = (3x_n + q_2)/2$ if x_n is odd, where q_1 and q_2 are large primes. Via a linear transformation $y_n = x_n - 2q_1$ these sequences are equivalent to $3x + q$ sequences with q prime or composite.

4. THE $px + q$ PROBLEM

4.1. Introduction. This problem is defined by a sequence, generated conditionally by $x_{n+1} = \frac{1}{2}x_n$ if x_n is even, and $x_{n+1} = \frac{1}{2}(px_n + q)$ if x_n is odd. The $px + q$ problem is analyzed by Crandall [3] and Matthews [15]. Without loss of generality we restrict ourselves to the case where $p \geq 5$ (prime), $q = 1$ or $q \geq 3$ (prime) and $p \neq q$. In contrast with the $3x + q$ problem the $px+q$ problem has in general (empirical observation, no formal proof) divergent trajectories. Either none or all of the numbers x_i in any trajectory (including a hypothetical cycle) must satisfy $x_i \equiv 0 \pmod{q}$. Cycles with $\mathrm{GCD}(x_i, q) = 1$ are called *primitive*.

The $px+q$ problem is a non-trivial generalization of the $3x+q$ problem, because there do not exist a and r such that if $x_0 = a2^k - r$ then $x_1 = ap2^{k-1} - r$. For $x_0 = a2^k - b_0$ we have $x_1 = (ap2^k - pb_0 + q)/2 = ap2^{k-1} - b_1$ with $b_1 = (pb_0 - q)/2$. The difference equation for b_k has the general solution

$$b_k = \frac{b_0 - q}{p-2}\left(\frac{p}{2}\right)^k + \frac{q}{p-2}.$$

If (and only if) $q = (p-2)r$ then $b_0 = r$ implies $b_k = r$.

In this section the quantities

$$\Lambda = (K+L)\log 2 - K\log p \quad \text{and} \quad \varrho = \log_2 p$$

play an important role. Note that Λ is defined differently compared to the $3x+q$ problem. We further assume that $x_i \geq x_0 > X_0 = 10^6 > mq$ for ease of analysis.

4.2. An upper bound for Λ

1. *Derivation of an upper bound for Λ in terms of the numbers in a cycle*

(a) Definition of increasing and decreasing subsequences and choice of an appropriate expression for numbers in subsequences. Let a_0 and k_0 satisfy $(p-2)x_0 = a_0 2^{k_0} - q$. Consequently, for the $px+q$ problem an odd number x_0 can be written as $x_0 = (a_0 2^{k_0} - q)/(p-2)$. The next number is $(a_0 2^{k_0-1}p - q)/(p-2)$, which is also an (odd) integer by induction. After an increasing subsequence of k_0 odd numbers the first appearing even number is $(a_0 p^{k_0} - q)/(p-2)$, which is the beginning of a decreasing subsequence.

(b) Derivation of chain equations between subsequences, resulting in a diophantine matrix equation for the coefficients of such expressions. For the $px+q$ problem the chain equation from the ith decreasing subsequence with l_i elements to the $(i+1)$th increasing subsequence with k_{i+1} elements is $(a_i p^{k_i} - q)/2^{l_i} = a_{i+1}2^{k_{i+1}} - q$. All chain equations together result in a diophantine system in the coefficients a_i. A necessary and sufficient condition for the existence of an m-cycle is the existence of a solution (a_i, k_i, l_i) of the diophantine system of equations

$$\begin{pmatrix} -p^{k_0} & 2^{k_1+l_0} & & \\ & -p^{k_1} & 2^{k_2+l_1} & \\ & & \ddots & \ddots \\ 2^{k_0+l_{m-1}} & & & -p^{k_{m-1}} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{m-1} \end{pmatrix} = \begin{pmatrix} q(2^{l_0}-1) \\ q(2^{l_1}-1) \\ \vdots \\ q(2^{l_{m-1}}-1) \end{pmatrix}. \tag{31}$$

(c) Derivation of an inequality for Λ. For the $px+q$ problem this inequality follows from multiplication of all ratios a_{i+1}/a_i and this leads to an inequality for Λ which is

$$(32)\qquad \begin{aligned} 0 &< (K+L)\log 2 - K\log p \\ &< \sum_{i=0}^{m-1} \frac{q}{a_i 2^{k_i} - q} = \frac{q}{p-2}\sum_{i=0}^{m-1}\frac{1}{\overline{x}_i} < \frac{qm}{(p-2)X_0}. \end{aligned}$$

2. *Derivation of an upper bound for Λ in terms of K and m*

(a) Derivation of a lower bound for the maximal local minimum in a cycle in terms of K and m. Based on the expression $(p-2)\overline{x}_i = a_i 2^{k_i} - q$ we find a lower bound for the maximal $\overline{x}_i$:

$$(33)\qquad \begin{aligned} [(p-1)\max(\overline{x}_i)]^m &> \prod_{i=0}^{m-1}[(p-1)\overline{x}_i] > \prod_{i=0}^{m-1}[(p-2)\overline{x}_i + q] \\ &= \prod_{i=0}^{m-1} a_i 2^{k_i} \ge 2^K. \end{aligned}$$

Consequently, $\max(\overline{x}_i) > 2^{K/m}/(p-1)$.

(b) Chaining of magnitudes of minima in a cycle. For the $px+q$ problem we have $(p-2)\overline{y}_i = a_i p^{k_i} - q$ and thus

$$\begin{aligned} \overline{x}_{i+1} &\le \frac{\overline{y}_i}{2} < \frac{\overline{y}_i + \frac{q}{p-2}}{2} = \left(\frac{1}{2}\right)\frac{a_i p^{k_i}}{p-2} = \left(\frac{p}{2}\right)^{k_i}\frac{\overline{x}_i + \frac{q}{p-2}}{2} \\ &= (2^{k_i})^{\varrho-1}\frac{\overline{x}_i + \frac{q}{p-2}}{2} < (a_i 2^{k_i})^{\varrho-1}\frac{(p-2)\overline{x}_i + q}{2(p-2)} = \frac{[(p-2)\overline{x}_i + q]^{\varrho}}{2(p-2)} \\ &= \frac{[(p-2) + q/\overline{x}_i]^{\varrho}}{2(p-2)}\,\overline{x}_i^{\varrho} \le \frac{[(p-2) + q/X_0]^{\varrho}}{2(p-2)}\,\overline{x}_i^{\varrho} < b\overline{x}_i^{\varrho} \end{aligned}$$

where $b = (p-1)^{\varrho}/2(p-2)$.

(c) Approximation of the upper bound for Λ. This last inequality can be rewritten as

$$(34)\qquad \overline{x}_i^{-1} < \frac{p-1}{[2(p-2)]^{1/\varrho}}\,\overline{x}_{i+1}^{-1/\varrho} = c\overline{x}_{i+1}^{-1/\varrho},$$

from which we derive (taking into account that the worst case appears if $\overline{x}_i < \overline{x}_{i+1}$ for $i = 0, \ldots, m-2$)

$$(35)\qquad \overline{x}_0^{-1} < c^{1+1/\varrho+\cdots+1/\varrho^{m-2}}\overline{x}_{m-1}^{-1/\varrho^{m-1}} < c^{\varrho/(\varrho-1)}\overline{x}_{m-1}^{-1/\varrho^{m-1}}.$$

Substitution of these results into (32) leads to an upper bound for Λ in terms of K and m:

$$(36)\qquad \Lambda < \sum_{i=0}^{m-1}\frac{q}{\overline{x}_i} \le \frac{qm}{\overline{x}_0} < \left[\frac{(p-1)^{\varrho}}{2(p-2)}\right]^{\frac{1}{\varrho-1}} \cdot 2^{\frac{\log_2(p-1)}{\varrho^{m-1}}} \cdot \frac{qm}{p-2} \cdot 2^{-K/m\varrho^{m-1}}.$$

Note that for $p = 3$ (hence $\varrho = \delta$) this upper bound is equivalent to (20).

4.3. A lower and upper bound for K. With ϱ substituted for δ a lemma similar to Lemma 2 supplies a lower bound $K \geq K_0(m,q)$ for the $px+q$ problem.

LEMMA 9. *$K \geq K_0(m,p,q) = q_{n+1}$ with n being the maximum index for which*

$$q_n + q_{n+1} \leq \log 2 \cdot \frac{p-2}{qm} \cdot X_0.$$

Proof. Assume $K < q_{n+1}$. Then from elementary number theory we have

$$\Lambda = (\log 2)|(K+L) - K\varrho| \geq (\log 2)|p_n - q_n\varrho| > \frac{\log 2}{q_n + q_{n+1}}$$
$$\geq \frac{qm}{(p-2)X_0} > \frac{q}{p-2}\sum_{i=0}^{m-1}\frac{1}{\overline{x}_i},$$

which contradicts (32). ∎

Let p and q be fixed. Applying this, for each m, with the maximal n that satisfies the condition, we find that the lower bound K_0 is a stepwise decreasing function of m, p and q.

COROLLARY 10. *For $p = 5$ and $q = 1$ we have:*

- *If $m = 1$ then $K \geq q_{12} = 1\,838\,395$.*
- *If $2 \leq m \leq 11$ then $K \geq q_{11} = 97\,879$.*
- *If $12 \leq m \leq 21$ then $K \geq q_{10} = 76\,573$.*
- *If $22 \leq m \leq 61$ then $K \geq q_9 = 21\,306$.*
- *If $62 \leq m \leq 97$ then $K \geq q_8 = 12\,655$.*
- *If $98 \leq m \leq 164$ then $K \geq q_7 = 8\,651$.*
- *If $165 \leq m \leq 447$ then $K \geq q_6 = 4\,004$.*

COROLLARY 11. *For $p = 11$ and $q = 19$ we have:*

- *If $m \leq 15$ then $K \geq q_{10} = 16\,503$.*
- *If $16 \leq m \leq 48$ then $K \geq q_9 = 4\,856$.*
- *If $49 \leq m \leq 112$ then $K \geq q_8 = 1\,935$.*
- *If $113 \leq m \leq 169$ then $K \geq q_7 = 986$.*
- *If $170 \leq m \leq 332$ then $K \geq q_6 = 949$.*

Lemma 5 is based on the lower bound of Rhin, which applies to $\Lambda = (K+L)\log 2 - K\log 3$ only. For the $px+q$ problem we have $\Lambda = (K+L) \cdot \log 2 - K\log p$. Then the best result for approximating Λ is from Laurent *et al.* [13]. We apply Corollaire 2 which reads

(37) $$\log|\Lambda| \geq -24.34D^4[\max(\log b' + 0.14, 21/D, 0.5)]^2 \cdot \log A_1 \cdot \log A_2$$

with

$$D = 1, \qquad b' = \frac{K}{D\log A_2} + \frac{K+L}{D\log A_1}, \qquad \log A_1 = \log p, \qquad \log A_2 = 1.$$

Since $X_0 > mq$ from (32) we find the inequality $\varrho K < K+L < \varrho K + 1/\log 2$ and thus $b' = K + (K+L)/\log p < 2.5K$. Substitution into (37) leads to

$$\Lambda > e^{-24.34\cdot\log p\cdot[\max(\log K+1.057,\,21)]^2}. \tag{38}$$

LEMMA 12. *Let $x = K_3(m,p,q)$ be the largest solution of*

$$e^{-24.34\cdot\log p\cdot[\max(\log x+1.057,\,21)]^2} = \left[\frac{(p-1)^\varrho}{2(p-2)}\right]^{\frac{1}{\varrho-1}} \cdot 2^{\frac{\log_2(p-1)}{\varrho^{m-1}}} \cdot \frac{qm}{p-2} \cdot 2^{-x/m\varrho^{m-1}}. \tag{39}$$

Then $K < K_3(m,p,q)$.

Laurent's upper bound is weaker than Rhin's. Let $c_3(m) = K_3(m)/m^3\varrho^m$. Then $c_3(m)$ is a decreasing function of m that tends for fixed p to a constant (independent of X_0 and practically independent of q) limit value as $m \to \infty$. Indeed, we find

(40)

K_3-bounds for the $px+q$ problem

	$p=5,\ q=1$		$p=11,\ q=19$	
m	c_3	$K_3 <$	c_3	$K_3 <$
1	10733.9	24921.9	10735.2	37134.4
2	2683.6	115738.1	2683.9	256934.8
5	429.4	$3.62\cdot10^{6}$	429.4	$2.66\cdot10^{7}$
10	108.1	$4.92\cdot10^{8}$	157.5	$3.87\cdot10^{10}$
100	25.50	$9.80\cdot10^{43}$	49.70	$3.95\cdot10^{61}$
1 000	18.30	$1.29\cdot10^{376}$	39.04	$3.95\cdot10^{549}$
10 000	17.40	$5.35\cdot10^{3671}$	37.69	$4.20\cdot10^{5403}$
100 000	17.29	$1.30\cdot10^{36601}$	37.52	$1.12\cdot10^{53917}$
1 000 000	17.27	$1.01\cdot10^{365868}$	37.49	$2.11\cdot10^{539204}$

Thus, for fixed p, q, m, the $px+q$ problem has a finite number of m-cycles. If $K_0(m,q) \ge K_3(m,p,q)$ then m-cycles cannot exist. Note that $K_0(m,q)$ heavily depends on X_0 and that $K_3(m,p,q)$ is an increasing function of m, p and q.

4.4. Upper bound reduction. From the continued fraction argument, we calculated the values of $K_1(m,p,q)$ and $K_2(m,p,q)$ to reduce the upper bound K_3. In the table below we show the critical values of m for which $K_0 > K_i$.

Minimal values of m for the non-existence of m-cycles of the $px + q$ problem

(41)

$m \to K_0 > K_i$				$m \to K_0 > K_i$				$m \to K_0 > K_i$			
p	q	K_2	K_3	p	q	K_2	K_3	p	q	K_2	K_3
5	1	8	1	7	1	8	2	11	1	7	2
	3	7	1		3	6	2		3	5	2
	7	6	1		5	6	1		5	4	2
	11	6	1		11	4	1		7	4	2
	13	6	1		13	4	1		13	4	1
	17	5	1		17	4	1		17	4	1
	19	5	1		19	4	1		19	4	1

From table (40) it follows that although the K_3 upper bound from Laurent *et al.* is relatively weak, the upper bound K_2 in table (41) is of the same order as for the $3x + q$ problem.

4.5. Elimination of solutions. For m-cycles we have the following

LEMMA 13. *If for the $px + q$ problem a primitive m-cycle exists with K odd and L even elements*, *then* $2^{K+L} - p^K \equiv 0 \pmod{q}$.

Proof. Suppose a primitive m-cycle exists. Then by definition

$$x_{K+L} \equiv \frac{p^K}{2^{K+L}}\, x_0 \pmod{q}.$$

Since $(p, q) = 1$ and $x_0 \not\equiv 0 \pmod{q}$ the lemma is proved. ■

For the $5x + q$ problem we checked for $m = 7$ and $q = 7, \ldots, 19$ in the interval (K_0, K_2) the conditions (32) and (28) to find that the only hypothetical solutions are

(42)

q	K	L
11	29957	39601
17	21306	28165

As a result we have

LEMMA 14. *The $5x+q$ problem with $q = 1$ or $3 \le q \le 19$ has for $m \le 7$ and $x_i > X_0$ no m-cycles besides hypothetically those listed in table* (42).

Directly from table (40) we have

LEMMA 15. *The $7x + q$ problem and the $11x + q$ problem with $q = 1$ or $3 \le q \le 19$ have for $m \le 4$ and $x_i > X_0$ no m-cycles.*

4.6. Numerical results for m-cycles. Finally, we calculated all m-cycles with minimal element $x_0 \le X_0$.

Number of m-cycles, cycle length, m-value and minimal element for each cycle

(43)

p	q	n_c	(cycle length, m, $x_{\min}$)
5	1	3	$(5,1,1)$ $(7,1,13)$ $(7,2,17)$
	3	4	$(3,1,1)$ $(7,2,43)$ $(7,2,53)$ $(7,3,61)$
	7	3	$(35,8,1)$ $(5,2,9)$ $(42,13,57)$
	11	2	$(4,1,1)$ $(14,5,141)$
	13	2	$(6,2,3)$ $(42,10,53)$
	17	1	$(20,5,9)$
	19	1	$(10,2,11)$
7	1	1	$(3,1,1)$
	3	0	
	5	2	$(6,1,3)$ $(31,9,27)$
	11	1	$(34,9,23)$
	13	0	
	17	0	
	19	1	$(18,3,5)$
11	1	0	
	3	0	
	5	1	$(4,1,1)$
	7	3	$(7,1,13)$ $(7,2,15)$ $(7,2,19)$
	13	0	
	17	0	
	19	0	

Thus we have:

LEMMA 16. *The* $5x+q$ *problem with* $q=1$ *or* $3 \le q \le 19$ *has for* $m \le 7$ *no* m*-cycles besides those listed in table* (42) (*hypothetically*) *and in table* (43).

LEMMA 17. *The* $7x+q$ *problem and the* $11x+q$ *problem with* $q=1$ *or* $3 \le q \le 19$ *have for* $m \le 4$ *no* m*-cycles besides those listed in table* (43).

Note that the 1-cycles for the $5x+3$, $5x+11$, $7x+1$ and $11x+5$ problems are trivial (they all start with $x_0 = 1$).

4.7. Existence conditions for m-cycles. Similar to a lemma of Davison [4], the existence of 1-cycles for the $px+q$ problem follows from

LEMMA 18. *A necessary and sufficient condition for the existence of a* (*non-*)*primitive* 1*-cycle for the* $px+q$ *problem is the existence of positive integers* k, l *and* r (*odd*) *such that* $2^{k+l} - p^k = qr$ *and the existence of an odd integer* x_0 *such that* $x_0 = (p^k - 2^k)/(p-2)r$.

Proof. Assume that $2^{k+l} - p^k = qr$ and $x_0 = (p^k - 2^k)/(p-2)r$ is an odd number. Then the general solution of $x_i = (px_{i-1} + q)/2$ is

$$x_i = \frac{\sum_{j=i}^{k-1} p^j 2^{k-1-j} + \sum_{j=0}^{i-1} p^j 2^{k+l-1-j}}{r}. \tag{44}$$

For $i < k$, x_i is an odd number. For $i = k$ we find

$$x_k = \frac{\sum_{j=0}^{k-1} p^j 2^{k+l-1-j}}{r} = 2^l \frac{\sum_{j=0}^{k-1} p^j 2^{k-1-j}}{r} = 2^l x_0. \tag{45}$$

Thus there exists a 1-cycle with k odd and l even elements and with x_0 as minimal element.

Now assume that a 1-cycle exists with k odd and l even elements. Let x_0 be the smallest odd element of the cycle. Since there are k odd elements $x_0, \ldots, x_{k-1}$ followed by an even element x_k, from the chain equation we have

$$a(2^{k+l} - p^k) = q(2^l - 1), \tag{46}$$

from which it follows that $2^{k+l} - p^k = qr$. Further we have

$$a2^k - q = \frac{q(2^l - 1)}{2^{k+l} - p^k} 2^k - q = \frac{2^l - 1}{r} 2^k - q = \frac{2^{k+l} - 2^k - qr}{r} = \frac{p^k - 2^k}{r}. \tag{47}$$

Since $(p-2)x_0 = a2^k - q$ we find that $x_0 = (p^k - 2^k)/(p-2)r$ is an (odd) integer. ∎

Note that if $x_0 = \frac{(p^k - 2^k)q}{(p-2)(2^{k+l} - p^k)}$ then $2^{k+l} - p^k = \frac{\frac{p^k - 2^k}{p-2} q}{x_0} = qr$. Hence we have

Corollary 19. *There exists a primitive* 1*-cycle with minimal element* x_0 *if and only if*

$$x_0 = \frac{(p^k - 2^k)q}{(p-2)(2^{k+l} - p^k)} \quad \textit{and} \quad \mathrm{GCD}(x_0, q) = 1.$$

Corollary 20. *If for the* $px + q$ *problem there exists a solution of the equation* $2^{k+l} - p^k = q$, *then there exists a (non-)primitive* 1*-cycle with* $x_0 = (p^k - 2^k)/(p-2)$.

Consider as an example the $11x + 7$ problem. We have $2^7 - 11^2 = 7$, hence $k = 2$, $l = 5$ and $x_0 = (11^2 - 2^2) \cdot 7/(11-2) \cdot 7 = 13$. We find the 1-cycle (13, 75, 416, 208, 104, 52, 26). For primes $5 \le p \le 97$ we computed the minimal values for q (prime) for which $2^{k+l} - p^k = q$ has a solution with $k \ge 2$. (All solutions of this equation in k and l can be found from an inequality for $\Lambda = (k+l)\log 2 - k \log p$.)

Minimal q with $2^{k+l} - p^k = q$

p	q	p	q	p	q
5	3	31	1087	67	28279
7	79	37	6823	71	126031
11	7	41	367	73	125743
13	1879	43	199	79	1951
17	223	47	30559	83	1303
19	151	53	29959	89	271
23	4217	59	29287	97	32641759
29	7351	61	520567		

4.8. Remarks. 1. The 1-cycle $(1, 2)$ of the $3x + 1$ problem has as a natural generalization the 1-cycle $(1, 2^l, \ldots, 2)$ of the $px + 1$ problem if p is a Mersenne prime $p = 2^{l+1} - 1$. For the $px + q$ problem with $p + q = 2^n$ we call 1-cycles starting with $x_0 = 1$ trivial.

2. In the table above, $px + q$ problems with a trivial 1-cycle, e.g. the $5x + 3$, $13x + 3$ and the $13x + 19$ problems are excluded. It remains possible that for a smaller value of q a primitive 1-cycle exists, e.g. the $7x+5$ problem has a 1-cycle with cycle length 6 and minimal element $x_0 = 3$. Note that if a primitive 1-cycle exists, there also exist primitive m-cycles, the argument being the same as in Remark 3 of Section 3.

3. Table (43) shows cycles with "small" cycle lengths and "small" elements only. However, the $5x + 37$ problem has a 61-cycle with cycle length 235 and elements in the interval (109, 1 950 748 496).

4. Lemmas 18 and 13 hold for arbitrary odd primes p and q with $p \neq q$. Hence they explain the numerical results of Lagarias and those listed in tables (30) and (43).

5. Lemma 13 is meaningless for the $px+1$ problem, which in general does not have cycles. A possible explanation is the small value of the numerator ($q = 1$) in the upper bound for Λ and the "relatively" large ratio p/q. Together they lead to a sharp lower bound K_0. The "space" below K_0 for which the number-theoretic approximation does not apply is then too small for the existence of cycles with small cycle length. Matthews [15] argues that $p \geq 5$ is the main reason for divergent behaviour and the non-existence of cycles. This section however gives counterexamples against this argument.

6. Belaga and Mignotte [2] study the $3x + q$ problem and they define for $B\ (= e^{\Lambda}) = 2^{K+L} - 3^K$ a set of numbers $\{A\}$ defined by

$$A = \sum_{i=1}^{K} 3^{K-i} 2^{\sum_{j=1}^{i-1} r_j} \tag{48}$$

with $\sum_{j=1}^{K-1} r_j < K + L$. They prove that a primitive cycle starting with x_0 exists if and only if there is an element A which satisfies $x_0 B = qA$. This lemma follows from the system (15) as a necessary and sufficient condition for the existence of an m-cycle, as our Lemma 18 and Corollary 19, which apply to the $px + q$ problem, follow from the system (31).

5. THE INVERSE COLLATZ PROBLEM

5.1. Introduction. The *inverse Collatz problem* (see Guy [6]) is defined by:

$$x_{i+1} = \begin{cases} \dfrac{3x_i}{2} & \text{if } x_i \equiv 0 \pmod 2, \\ \dfrac{3x_i - 1}{4} & \text{if } x_i \equiv -1 \pmod 4, \\ \dfrac{3x_i + 1}{4} & \text{if } x_i \equiv 1 \pmod 4. \end{cases}$$

Such a sequence is also called a *permutation sequence* because each number has, in contrast with the original Collatz problem, exactly one predecessor. If $x_i > 0$, there are four known cycles: (1), $(2, 3)$, $(4, 6, 9, 7, 5)$ and $(44, 66, 99, 74, 111, 83, 62, 93, 70, 105, 79, 59)$. The behaviour of the sequence $(\ldots, 31, 23, 17, 13, 10, 15, 11, 8, 12, 18, 27, 20, 30, \ldots)$ is unknown. In the rest of this section

$$\Lambda = (K + L)\log 3 - (K + 2L)\log 2$$

plays an important role. We assume that $x_i \geq x_0 > X_0 > 40m + 2$ for ease of analysis.

5.2. An upper bound for Λ

1. *Derivation of an upper bound for Λ in terms of the numbers x_i in a cycle*

(a) Definition of increasing and decreasing subsequences and choice of an appropriate expression for numbers in subsequences. For the inverse Collatz problem an even number can be written as $a2^k$, which leads to $k - 1$ even successors. Then the first odd number is $a3^k$. Odd numbers can be written as $b4^l \pm 1$, which leads to (at least) $l - 2$ odd successors. The next successor is $b3^l \pm 1$ and (depending on the parity of b) can be even or odd. In the last case the decreasing subsequence continues.

(b) Derivation of chain equations between (sub-)subsequences, resulting in a diophantine matrix equation for the coefficients of such expressions. For the inverse Collatz problem the chain equation from the ith increasing subsequence with k_i even elements to the $(i+1)$th decreasing subsequence is $a_i 3^{k_i} = b_{i0} 4^{l_{i0}} \pm 1$. The chain equation from a decreasing sub-subsequence to the next decreasing sub-subsequence is $b_{ij}3^{l_{ij}} \pm 1 = b_{i,j+1}4^{l_{i,j+1}} \mp 1$. The chain equation from a decreasing subsequence to the next increasing subsequence is $b_{ij}3^{l_{ij}} \pm 1 = a_{i+1}3^{k_{i+1}}$. Let $j_k,\ k = 0, \ldots, m-1$, be the number of sub-subsequences in the kth decreasing subsequence. Then all chain equations together result in a diophantine system in the coefficients a_i and b_{ij}

$$A \cdot \begin{pmatrix} a_0 \\ b_{00} \\ \vdots \\ b_{0j_0} \\ a_1 \\ \vdots \\ a_{m-1} \\ \vdots \\ b_{m-1,j_{m-1}} \end{pmatrix} = \begin{pmatrix} \pm 1 \\ \pm 2 \\ \mp 2 \\ \vdots \\ \pm 1 \\ \vdots \\ \pm 1 \\ \vdots \\ \pm 2 \end{pmatrix}$$

where A is the matrix

$$\begin{pmatrix} -3^{k_0} & 4^{l_{00}} & & & & & & & & \\ & -3^{l_{00}} & 4^{l_{01}} & & & & & & & \\ & & \cdots & & & & & & & \\ & & & -3^{l_{0,j_0}} & 2^{k_1} & & & & & \\ & & & & -3^{k_1} & 4^{l_{10}} & & & & \\ & & & & & \cdots & & & & \\ & & & & & & -3^{k_{m-1}} & 4^{l_{m-1,0}} & & \\ & & & & & & & \cdots & & \\ 2^{k_0} & & & & & & & & & -3^{l_{m-1,j_{m-1}}} \end{pmatrix}.$$

An integer solution for the coefficients is a necessary and sufficient condition for the existence of an m-cycle.

(c) Derivation of an inequality for Λ. For the inverse Collatz problem this inequality cannot be derived directly from the chain equations. First, a generalized chain equation linking the coefficients a_i and a_{i+1} must be derived. For the chain equations from a_i to a_{i+1} there are four possible configurations. The decreasing subsequence can start and end with an odd number $\equiv \pm 1 \pmod 4$. We will analyze the two cases that the decreasing subsequence starts with an odd number $\equiv 1 \pmod 4$. Then the corresponding j_i

chain equations can be rewritten as

$$\frac{a_i 3^{k_i}}{b_{i0}4^{l_{i0}}} = 1 + \frac{1}{b_{i0}4^{l_{i0}}},$$
$$\frac{b_{i0} 3^{l_{i0}}}{b_{i1}4^{l_{i1}}} = 1 - \frac{2}{b_{i1}4^{l_{i1}}},$$
$$\frac{b_{i1} 3^{l_{i1}}}{b_{i2}4^{l_{i2}}} = 1 + \frac{2}{b_{i2}4^{l_{i2}}},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\frac{b_{i,j_i-1} 3^{l_{i,j_i-1}}}{a_{i+1}2^{k_{i+1}}} = 1 \pm \frac{1}{a_{i+1}2^{k_{i+1}}}.$$

Multiplication of these equations leads to either

$$\frac{a_i 3^{k_i+l_{i0}+\cdots+l_{i,j_i-1}}}{a_{i+1}2^{k_{i+1}+2(l_{i0}+\cdots+l_{i,j_i-1})}} = \left(1+\frac{1}{b_{i0}4^{l_{i0}}}\right)\cdots\left(1+\frac{2}{b_{i,j_i-1}4^{l_{i,j_i-1}}}\right)\left(1-\frac{1}{a_{i+1}2^{k_{i+1}}}\right) \tag{49}$$

or

$$\frac{a_i 3^{k_i+l_{i0}+\cdots+l_{i,j_i-1}}}{a_{i+1}2^{k_{i+1}+2(l_{i0}+\cdots+l_{i,j_i-1})}} = \left(1+\frac{1}{b_{i0}4^{l_{i0}}}\right)\cdots\left(1-\frac{2}{b_{i,j_i-1}4^{l_{i,j_i-1}}}\right)\left(1+\frac{1}{a_{i+1}2^{k_{i+1}}}\right). \tag{50}$$

Since $b_{i,j}4^{l_{i,j}} \geq b_{i,j+1}4^{l_{i,j+1}} + 2$ we have

$$\left(1-\frac{2}{b_{i,j}4^{l_{i,j}}}\right)\left(1+\frac{2}{b_{i,j+1}4^{l_{i,j+1}}}\right) \geq 1,$$
$$\left(1+\frac{2}{b_{i,j}4^{l_{i,j}}}\right)\left(1-\frac{2}{b_{i,j+1}4^{l_{i,j+1}}}\right) \leq 1.$$

In the first case we find that the right hand side product (RHS) of (49) and (50) satisfies

$$1 - \frac{1}{a_{i+1}2^{k_{i+1}}} < \text{RHS} < 1 + \frac{2}{a_{i+1}2^{k_{i+1}}}. \tag{51}$$

In the second case we find

$$1 - \frac{2}{a_{i+1}2^{k_{i+1}}} < \text{RHS} < 1 + \frac{1}{a_{i+1}2^{k_{i+1}}}. \tag{52}$$

So in both cases we have the inequality

$$1 - \frac{2}{a_{i+1}2^{k_{i+1}}} < \frac{a_i 3^{k_i+l_{i0}+\cdots+l_{i,j_i-1}}}{a_{i+1}2^{k_{i+1}+2(l_{i0}+\cdots+l_{i,j_i-1})}} < 1 + \frac{2}{a_{i+1}2^{k_{i+1}}}. \tag{53}$$

This inequality can be proved in a similar way if the decreasing subsequence starts with an odd number $\equiv -1 \pmod 4$. Multiplication over i leads to the inequality

$$\prod_{i=0}^{m-1}\left[1-\frac{2}{a_i 2^{k_i}}\right] < \frac{3^{K+L}}{2^{K+2L}} < \prod_{i=0}^{m-1}\left[1+\frac{2}{a_i 2^{k_i}}\right]. \tag{54}$$

Taking logs leads to an inequality for $\Lambda = (K+L)\log 3 - (K+2L)\log 2$:

$$-\sum_{i=0}^{m-1}\frac{2}{a_i 2^{k_i}-2} < \Lambda < \sum_{i=0}^{m-1}\frac{2}{a_i 2^{k_i}}. \tag{55}$$

2. *Derivation of an upper bound for Λ in terms of K and m*

(a) Derivation of a lower bound for the maximal local minimum in a cycle in terms of K and m. Based on the expression for the local minima $\overline{x}_i = a_i 2^{k_i}$ we find a lower bound $2^{K/m}$ for the maximal local minimum in a cycle.

(b) Chaining of magnitudes of minima in an cycle. For $i = 0, \ldots, m-1$ we find

$$\begin{aligned}
\overline{x}_{i+1} &= a_{i+1}2^{k_{i+1}} \le b_{i,0}3^{l_{i,0}} + 1 = \left(\frac{3}{4}\right)^{l_{i,0}}\left[b_{i,0}4^{l_{i,0}} + \left(\frac{4}{3}\right)^{l_{i,0}}\right] \\
&= \left(\frac{3}{4}\right)^{l_{i,0}}\left[a_i 3_i^k - 1 + \left(\frac{4}{3}\right)^{l_{i,0}}\right] = \left(\frac{3}{4}\right)^{l_{i,0}} a_i 3^{k_i} + \left[1-\left(\frac{3}{4}\right)^{l_{i,0}}\right] \\
&< \left(\frac{3}{4}\right)^{l_{i,0}} a_i^{\log_2 3} 3^{k_i} + \left[1-\left(\frac{3}{4}\right)^{l_{i,0}}\right] \le \left(\frac{3}{4}\right)[a_i 2^{k_i}]^\delta + \frac{1}{4} = \frac{3}{4}\,\overline{x}_i^\delta + \frac{1}{4} \\
&\le b\overline{x}_i^\delta
\end{aligned}$$

with $b = 3/4 + 1/4X_0^\delta$.

(c) Approximation of the upper bound for Λ. This last inequality can be rewritten as

$$\overline{x}_i^{-1} < b^{1/\delta}\overline{x}_{i+1}^{-1/\delta} = c\overline{x}_{i+1}^{-1/\delta}, \tag{56}$$

from which we derive (taking into account that the worst case appears if $\overline{x}_i < \overline{x}_{i+1}$ for $i = 0, \ldots, m-2$)

$$\begin{aligned}
\overline{x}_0^{-1} &< c^{1+1/\delta+\cdots+1/\delta^{m-2}}\overline{x}_{m-1}^{-1/\delta^{m-1}} < c^{\delta/(\delta-1)}\overline{x}_{m-1}^{-1/\delta^{m-1}} \\
&= b^{1/(\delta-1)}\overline{x}_{m-1}^{-1/\delta^{m-1}}.
\end{aligned} \tag{57}$$

Substitution of these results into (55) leads to an upper bound for Λ in terms of K and m:

$$\Lambda < \sum_{i=0}^{m-1}\frac{2}{\overline{x}_i} < \frac{2m}{\overline{x}_0} < \frac{2mb^{1/(\delta-1)}}{2^{K/m\delta^{m-1}}}. \tag{58}$$

In a similar way we find a lower bound for Λ and together this leads to

$$(59)\qquad -\frac{2mb^{1/(\delta-1)}}{2^{K/m\delta^{m-1}}-2} < \Lambda < \frac{2mb^{1/(\delta-1)}}{2^{K/m\delta^{m-1}}}.$$

5.3. A lower bound and an upper bound for K. Suppose there exists an m-cycle with K odd numbers and L even numbers with $x_0 > X_0$. Let p_n/q_n be the nth convergent to δ. Similar to Lemma 1 we have

Lemma 21. *If $q_n + q_{n+1} \le (\log 2)(X_0 - 2)/2m$ then $K + L \ge q_{n+1}$.*

Proof. Since $x_i \ge x_0 > X_0$ we find, according to (55),

$$(60)\qquad |\Lambda| < \sum_{i=0}^{m-1} \frac{2}{x_i - 2} < \frac{2m}{X_0 - 2}.$$

Assume $K + L < q_{n+1}$. Then

$$\begin{aligned} |\Lambda| &= (\log 2)|(K + 2L) - (K + L)\delta| \\ &\ge (\log 2)|p_n - q_n\delta| > \frac{\log 2}{q_n + q_{n+1}} \ge \frac{2m}{X_0 - 2}, \end{aligned}$$

which contradicts (60). ∎

Since $X_0 > 40m + 2$ we have $|\Lambda| < 1/20$ from (60). We find

$$(\log 3 - \log 2)K \ge (2\log 2 - \log 3)L - 1/2 > 0.237L \quad \text{if} \quad L \ge 1.$$

Thus $2K > L$ and $3K > K + L \ge q_{n+1}$. Applying this, for each m, with the maximal n that satisfies the condition, we find that the lower bound $K_0(m)$ is a stepwise decreasing function of m.

Corollary 22.

- *If $m \le 2$ then $3K > q_{13} = 190\,537$.*
- *If $m = 3$ then $3K > q_{12} = 111\,202$.*
- *If $4 \le m \le 6$ then $3K > q_{11} = 79\,335$.*
- *If $7 \le m \le 14$ then $3K > q_{10} = 31\,867$.*
- *If $15 \le m \le 42$ then $3K > q_9 = 15\,601$.*
- *If $43 \le m \le 712$ then $3K > q_8 = 665$.*

Rhin's lemma applies with absolute values, thus $|\Lambda| > e^{-13.3(0.46057+\log K)}$. Inequality (59) gives an upper bound

$$|\Lambda| < \frac{2mb^{1/(\delta-1)}}{2^{K/m\delta^{m-1}}-2}.$$

Let $x = K_3(m)$ be the largest solution of

$$(61)\qquad e^{-13.3(0.46057+\log x)} = \frac{2mb^{1/(\delta-1)}}{2^{x/m\delta^{m-1}}-2}.$$

Then $K < K_3(m)$. Let $c_3(m) = K_3(m)/m^2\delta^m$. Then $c_3(m)$ is a decreasing function of m that tends to a constant (independent of X_0) limit value 5.576 as $m \to \infty$.

K_3-bounds for the inverse Collatz problem

(62)

m	c_3	$K_3 <$
1	61.13	96.9
2	39.40	395.9
3	30.49	1092.4
4	25.53	2577.5
5	22.33	5584.9
6	20.09	11466.4
10	15.23	152399.8
100	7.03	$7.06 \cdot 10^{24}$
1 000	5.78	$6.03 \cdot 10^{206}$
10 000	5.60	$8.67 \cdot 10^{2008}$
100 000	5.58	$4.42 \cdot 10^{20012}$
1 000 000	5.58	$5.47 \cdot 10^{200031}$

Combining the lower bound from Corollary 22 and the upper bound from table (62) shows that a hypothetical m-cycle with $x_i > X_0$ has $m \geq 7$.

5.4. Upper bound reduction. We calculated K_1 and K_2 from the continued fraction argument. Up to $K_3(10) = 155399.8$ the champion partial quotient is $a = 23$. Let $x = K_2(m)$ be the largest solution of

$$\frac{2mb^{1/(\delta-1)}}{2^{x/m\delta^{m-1}} - 2} = \frac{\log 2}{25x}. \tag{63}$$

Then $K < K_2(m)$. Let $c_2(m) = K_2(m)/m^2\delta^m$. Then $c_2(m)$ is a decreasing function of m that tends to a constant (independent of X_0) limit value as $m \to \infty$. The reduction factor from K_3 to K_2 is $c_2/c_3 \simeq 0.1$.

K_2-bounds for the inverse Collatz problem

(64)

m	c_2	$K_2 <$
1	5.40	8.6
2	3.68	37.0
3	2.89	103.6
4	2.43	245.3
5	2.12	531.3
6	1.91	1099.4
7	1.74	2146.0
8	1.61	4114.7
9	1.51	7723.3
10	1.42	14255.3

Combining the lower bound from Corollary 22 and the upper bound from table (64), we find that if the inverse Collatz problem has an m-cycle with $x_i > X_0$ then $m > 9$.

5.5. Numerical results for m-cycles. We checked that for $x_0 \leq X_0$ all divergent trajectories have $m \geq 10$ and that no other cycles existed than those listed in the introduction. Thus we have

LEMMA 23. *The inverse Collatz problem has for $m \leq 9$ no m-cycles other than* $(1), (2, 3), (4, 6, 9, 7, 5)$ *and* $(44, 66, 99, 74, 111, 83, 62, 93, 70, 105, 79, 59)$.

5.6. Remarks. 1. Our results are sharper than the results of [6].

2. Because for the inverse Collatz problem each x_i has only one predecessor, limit cycles with $m \leq 9$ cannot exist when a divergent trajectory has ≥ 10 pairs of increasing and decreasing subsequences.

3. The asymptotic argument for divergence, namely that the average multiplication factor of the inverse Collatz sequence is $3^4/2^6 \simeq 1.26$, looses its value because divergent trajectories (also) depart from ∞.

4. Our theoretical approach cannot exclude the existence of m-cycles for $m \geq 10$ with minimal element $x_0 \leq X_0$ because the lower and upper bounds of (55) cannot be sharpened. So the behaviour of the sequence $(\ldots, 31, 23, 17, 13, 10, 15, 11, 8, 12, 18, 27, 20, 30, \ldots)$ remains an open question.

6. THE GENERALIZED COLLATZ PROBLEM

6.1. Introduction. Such sequences of integers x_n are introduced among others by Matthews [15] and defined by the conditional recurrence relation

$$x_{n+1} = \frac{p_i x_n + (q - p_i) r_i}{q} \quad \text{if } x_n \equiv i \pmod{q} \tag{65}$$

where p_i, r_i $(i = 0, \ldots, q-1)$ and q satisfy $(p_i, q) = 1$, $r_i \equiv i \pmod{q}$. Note that $(q - p_i) r_i \in \mathbb{Z}$ but $r_i \in \mathbb{Q}$. Depending on the sign of $p_i - q$, this recurrence relation generates an increasing or decreasing subsequence, so a 1-cycle may consist of an arbitrary number of increasing subsequences followed by an arbitrary number of decreasing subsequences. We will derive an upper bound in terms of the numbers in a cycle for Λ (which is a linear form in logarithms that depends on the actual m-cycle) for a 1-cycle and sketch the general approach for an m-cycle, where appropriate. For ease of analysis, we assume $x_n \geq x_0 > X_0 > \max(r_i, mq)$.

6.2. An upper bound for Λ

1. *Derivation of an upper bound for Λ in terms of the numbers in a cycle*

(a) Let the 1-cycle start with the minimal element $x_0 \equiv i \pmod{q}$. Then a subsequence starts with $x_0 = a_i q^{k_i} + r_i$. Because $r_i \in \mathbb{Q}$ also $a \in \mathbb{Q}$ to ensure $x_0 \in \mathbb{Z}$. The k_ith successor is $a_i p_i^{k_i} + r_i$, which can be rewritten as $a_j q^{k_j} + r_j$.

(b) The corresponding chain equation can be written as

$$\frac{a_i p_i^{k_i}}{a_j q^{k_j}} = 1 + \frac{r_j - r_i}{a_j q^{k_j}}. \tag{66}$$

Multiplication of all chain equations leads to

$$\frac{p_i^{k_i} p_j^{k_j} \cdots p_l^{k_l}}{q^{k_i+k_j+\cdots+k_l}} = \left(1 + \frac{r_j - r_i}{a_j q^{k_j}}\right) \cdots \left(1 + \frac{r_i - r_s}{a_s q^{k_s}}\right). \tag{67}$$

(c) Let d^+ be the maximum of the positive values $(r_j - r_i) \ldots (r_i - r_s)$ and d^- be the minimum of the negative values $(r_j - r_i) \ldots (r_i - r_s)$. Then

$$\prod \left[1 + \frac{d^-}{a_i q^{k_i}}\right] < \frac{p_i^{k_i} p_j^{k_j} \cdots p_l^{k_l}}{q^{k_i+k_j+\cdots+k_l}} < \prod \left[1 + \frac{d^+}{a_i q^{k_i}}\right] \tag{68}$$

where the left hand product runs over all terms with $r_i - r_j$ negative and the right hand product runs over all terms with $r_i - r_j$ positive. Taking logs leads to an inequality for Λ.

2. *Derivation of an upper bound for Λ in terms of K and m*

(a) Derivation of a lower bound for the maximum element in a cycle. For a possible m-cycle we have

$$[\max(a_i q^{k_i})]^m > \prod_{i=0}^{m-1} a_i q^{k_i} > q^K. \tag{69}$$

(b) Chaining of magnitudes of numbers in a cycle. Let the (minimal) start element of an increasing subsequence be $\overline{x}_i$ and the maximal element be $\overline{y}_i$. If $a_i = 1$ then $\overline{x}_i = \overline{y}_i^{\beta_i}$ with $\beta_i = (\log q^{k_i} + r_i)/(\log p_i^{k_i} + r_i)$. If $a_i \geq 2$ then, since $(\log x)/(\log x - a_i)$ is a decreasing function of x, we have

$$\frac{\log a_i q^{k_i} + r_i}{\log a_i p_i^{k_i} + r_i} > \frac{\log a_i q^{k_i} + a_i r_i}{\log a_i p_i^{k_i} + a_i r_i} = \frac{\log a_i + \log(q^{k_i} + r_i)}{\log a_i + \log(p_i^{k_i} + r_i)} > \beta_i, \tag{70}$$

hence $\overline{x}_i \geq \overline{y}_i^{\beta_i}$ for increasing subsequences. If the increasing part of a 1-cycle contains v increasing subsequences and $\beta = \max(\beta_i)$ then $\overline{x}_i \geq (\overline{y}_{i+v-1})^{\beta^v}$. For a decreasing subsequence from $\overline{y}_i$ to $\overline{x}_{i+1}$ we have

$$\overline{x}_{i+1} \leq \frac{p_j \overline{y}_i + (q - p_j) r_j}{q} \leq \frac{p_j \overline{y}_i}{q}.$$

If the decreasing part of a 1-cycle contains w decreasing subsequences and $r = \max(p_j/q)$ we have $\overline{x}_{i+1} \leq r^w \overline{y}_i$. From a local minimum $\overline{x}_i$ to the next local minimum $\overline{x}_{i+1}$ we have $\overline{x}_i \geq [\overline{x}_{i+1}/r^w]^{\beta^v}$.

(c) Approximation of the upper bound for Λ. Similar to the approach for $px + q$ problem these expressions can be substituted into (68). So it is possible (though not straightforward due to the variety in rest values mod q) to derive an upper bound for Λ in terms of K and m.

6.3. A lower bound and upper bound for K. The (generalized) lemma of Crandall is no longer applicable, since it depends on the condition that $\Lambda = A\log u + B\log v$. A (trivial) lower bound for K can be established from exhaustive search of small solutions for K. Since Λ is in general a linear form in ≥ 3 logarithms, also the sharp lower bounds for two logarithms from Rhin and Laurent *et al.* are no longer applicable. Baker's original result and later refinements for more than two logarithms remain applicable to find an upper bound for K. Upper bound reduction based on convergents fails, since it requires a linear form in two logarithms, but lattice basis reduction algorithms (e.g. the L^3-algorithm [24, p. 41]) remain applicable. So a lower bound and an upper bound for K can be derived and as a consequence for fixed m the number of hypothetical cycles is bounded.

6.4. Theoretical and numerical results for cycles

6.4.1. *Introduction.* Instead of analyzing the general case (which is complex because of the variety in p_i, r_i and q) we discuss two examples of Matthews [15]. The first example is a classical generalized Collatz problem. The second example is a special case where (if we only consider cycles) degeneration to a binary conditioned "$px + q$" problem occurs.

6.4.2. *A Collatz problem with four conditions.* As a first example we discuss a sequence generated by

$$x_{i+1} = \begin{cases} \dfrac{x_i}{4} & \text{if } x_i \equiv 0 \pmod 4, \\ \dfrac{3x_i - 3}{4} & \text{if } x_i \equiv 1 \pmod 4, \\ \dfrac{5x_i - 2}{4} & \text{if } x_i \equiv 2 \pmod 4, \\ \dfrac{17x_i - 3}{4} & \text{if } x_i \equiv 3 \pmod 4. \end{cases}$$

Matthews conjectures that there exist 16 cycles with starting values in the range $(-750, 5127)$. In our definition we have $r_0 = 0$, $r_1 = -3$, $r_2 = 2$ and $r_3 = \frac{3}{13}$. Indeed, if $x_0 = 7 = \frac{22}{13}\cdot 4^1 + \frac{3}{13}$ we find $x_1 = 29 = \frac{22}{13}\cdot 17^1 + \frac{3}{13}$ etc. Our approach leads to $\Lambda = A\log 2 + B\log 3 + C\log 5 + D\log 17$ and can in principle be used to exclude m-cycles up to any fixed upper bound for m. Because of the absence of simple upper bound reduction methods, we leave this example for future research.

6.4.3. *A Collatz problem with three (two) conditions.* As a second example we analyze the existence of m-cycles for a "tantalizing" problem of

Matthews

$$x_{i+1} = \begin{cases} 2x_i & \text{if } x_i \equiv 0 \pmod 3, \\ \dfrac{7x_i+2}{3} & \text{if } x_i \equiv 1 \pmod 3, \\ \dfrac{x_i-2}{3} & \text{if } x_i \equiv 2 \pmod 3, \end{cases}$$

which has for $x_0 > 0$ many divergent trajectories and for $x_i < 0$ two cycles (-1) and $(-2,-4)$. In our definition we have $p_0 = 6$, $r_0 = 0$, $p_1 = 7$, $r_1 = 1/2$, $p_2 = 3$, $r_2 = 1$. So formally this problem does not meet our definition (65), because $(p_0, q) = q \neq 1$. Note, however, that the equation $x_{i+1} = 2x_i$ if $x_i \equiv 0 \pmod 3$ has as result that a hypothetical m-cycle only contains numbers which are $\neq 0 \pmod 3$. By considering such cycles the requirements of (65) are satisfied.

1. *A lower bound for K.* Because of the restriction to $x_i \neq 0 \pmod 3$, we can find, similar to Lemma 2, a lower bound $K > K_0(m) = q_{n+1}$ with for q_n the maximum value such that $q_n + q_{n+1} \leq (\log 3)2X_0/m$. Again we found that the lower bound $K_0(m)$ is a stepwise decreasing function of m.

Corollary 24.

- *If* $m = 1$ *then* $K > q_{10} = 439\,341$.
- *If* $m = 2$ *then* $K > q_9 = 330\,356$.
- *If* $3 \leq m \leq 9$ *then* $K > q_8 = 108\,985$.
- *If* $10 \leq m \leq 309$ *then* $K > q_7 = 3\,401$.

2. *An upper bound for Λ.* Recall that $x_0 = \overline{x}_0$. A number $x_0 \equiv 1 \pmod 3$ can be written as $2x_0 = a_0 \cdot 2 \cdot 3^{k_0} - 1$ with $a_0 \equiv 2 \pmod 3$. Then $2x_{k_0} = a_0 \cdot 2 \cdot 7^{k_0} - 1$ with $x_{k_0} \equiv 2 \pmod 3$. Thus $(a_0 \cdot 2 \cdot 7^{k_0} - 1)/2 = b_0 \cdot 3^{l_0} - 1$ with $b_0 \equiv 2 \pmod 3$ and $x_{k_0+l_0} = b_0 - 1 = (a_0 \cdot 2 \cdot 3^{k_1} - 1)/2$. A similar argument applies to $\overline{x}_i$ (the ith local minimum). If b_i is eliminated we find the chain equation

$$-7^{k_i} \cdot 2a_i + 3^{l_i}(3^{k_{i+1}} \cdot 2a_{i+1} + 1) = 1, \tag{71}$$

from which it follows that the coefficients a_i must satisfy the diophantine system

$$\begin{pmatrix} 7^{k_0} & -3^{k_1+l_0} & & \\ & 7^{k_1} & -3^{k_2+l_1} & \\ & & \ddots & \ddots \\ -3^{k_0+l_{m-1}} & & & 7^{k_{m-1}} \end{pmatrix} \begin{pmatrix} 2a_0 \\ 2a_1 \\ \vdots \\ 2a_{m-1} \end{pmatrix} = \begin{pmatrix} 3^{l_0}-1 \\ 3^{l_1}-1 \\ \vdots \\ 3^{l_{m-1}}-1 \end{pmatrix}.$$

The ith chain equation can be rewritten as

$$\frac{7^{k_i} \cdot a_i}{3^{k_{i+1}+l_i} \cdot a_{i+1}} = 1 + \frac{3^{l_i}-1}{3^{k_{i+1}+l_i} \cdot 2a_{i+1}} < 1 + \frac{1}{3^{k_{i+1}} \cdot 2a_{i+1}}. \tag{72}$$

Multiplication (with indices taken mod 3) and taking logs leads to an inequality for $\Lambda = K\log 7 - (K+L)\log 3$,

$$0 < \Lambda < \sum_{i=0}^{m-1} \log\left[1 + \frac{1}{3^{k_i}\cdot 2a_i}\right] < \sum_{i=0}^{m-1} \frac{1}{3^{k_i}\cdot 2a_i} = \sum_{i=0}^{m-1} \frac{1}{2\overline{x}_i + 1}. \tag{73}$$

For the maximal $\overline{x}_i$ we have

$$(2\max(\overline{x}_i)+1)^m > \prod_{i=0}^{m-1}(2\overline{x}_i+1) = \prod_{i=0}^{m-1} a_i\cdot 2\cdot 3^{k_i} > 2^m 3^K. \tag{74}$$

Hence $\max(\overline{x}_i) > 3^{K/m} - 1/2 > 3^{K/m}/3$. Further we find (with $\varrho = \log_3 7$)

$$\begin{aligned}\overline{x}_{i+1} \le \frac{\overline{y}_i}{3} = \frac{a_i\cdot 2\cdot 7^{k_i} - 1}{6} &< (3^{k_i})^{\varrho-1}\,\frac{\overline{x}_i + 1/2}{3} < (a_i\cdot 2\cdot 3^{k_i})^{\varrho-1}\,\frac{\overline{x}_i+1/2}{3}\\ = \frac{(\overline{x}_i+1/2)^{\varrho}}{3} &= \frac{(1+1/2\overline{x}_i)^{\varrho}\overline{x}_i^{\varrho}}{3} < \frac{(1+1/2X_0)^{\varrho}\overline{x}_i^{\varrho}}{3} = b\overline{x}_i^{\varrho}.\end{aligned}$$

This last inequality can be rewritten as

$$\overline{x}_i^{-1} < b^{1/\varrho}\overline{x}_{i+1}^{-1/\varrho} = c\overline{x}_{i+1}^{-1/\varrho}, \tag{75}$$

from which we derive (taking into account that the worst case appears if $\overline{x}_i < \overline{x}_{i+1}$ for $i = 0,\dots,m-2$)

$$\overline{x}_0^{-1} < c^{1+1/\varrho+\dots+1/\varrho^{m-2}}\overline{x}_{m-1}^{-1/\varrho^{m-1}} < c^{\varrho/(\varrho-1)}\overline{x}_{m-1}^{-1/\varrho^{m-1}}. \tag{76}$$

Substitution of these results into (73) leads to a lower bound for Λ in terms of K and m:

$$\Lambda < \sum_{i=0}^{m-1}\frac{1}{2\overline{x}_i+1} < \frac{m}{2\overline{x}_0} < \frac{mb^{1/(\varrho-1)}}{2}\,3^{-(K-m)/m\varrho^{m-1}}. \tag{77}$$

3. *A lower bound for Λ.* Since $\Lambda = K\log 7 - (K+L)\log 3$, the best result for approximating Λ is from Laurent *et al.* [13]. We apply Corollaire 2 with $D = 1$, $b' = (K+L)/\log A_2 + K/\log A_1$, $\log A_1 = \log 3$ and $\log A_2 = 7$. Since $\Lambda > 0$ we have $K + L < \varrho K$ and thus $b' < 1.8K$. Substitution into (37) leads to

$$\Lambda > e^{-52.09\cdot[\max(\log K + 1.728,\, 21)]^2}. \tag{78}$$

LEMMA 25. *Let $x = K_3(m)$ be the largest solution of*

$$e^{-52.09\cdot[\max(\log x+1.728,\, 21)]^2} = \frac{mb^{1/(\varrho-1)}}{2}\,3^{-(x-m)/m\varrho^{m-1}}. \tag{79}$$

Then $K < K_3(m)$.

Let $c_3(m) = K_3(m)/m^3\delta^m$. Then $c_3(m)$ is a decreasing function of m that tends to a constant (independent of X_0) limit value 0.294 as $m\to\infty$. We calculated K_1 and K_2 from the continued fraction argument. Up to

$K_3(10) = 3.59 \cdot 10^7$ the champion partial quotient is $a = 32$. Let $x = K_2(m)$ be the largest solution of

$$\frac{mb^{1/(\varrho-1)}}{2} 3^{-(x-m)/m\varrho^{m-1}} = \frac{\log 2}{34x}. \tag{80}$$

Then $K < K_2(m)$. Let $c_2(m) = K_2(m)/m^2\delta^m$. Then $c_2(m)$ is a decreasing function of m that tends to a constant (independent of X_0) limit value as $m \to \infty$. The reduction factor from K_3 to K_2 is $c_2/c_3 \simeq 1/100m$.

K_2- and K_3-bounds for the generalized Collatz problem

(81)

m	$K_2 <$	$K_3 <$
1	2.4	20907.8
5	422.5	$1.03 \cdot 10^6$
9	10501.3	$9.15 \cdot 10^6$
10	21985.7	$3.59 \cdot 10^7$

Combining the lower bound from Corollary 24 and the upper bound from table (81), we find that if the generalized Collatz problem has an m-cycle with $x_i > X_0$ then $m > 9$. We checked that for $0 < x_0 \le X_0$ always a divergent trajectory occurs. Thus we have

LEMMA 26. *If the generalized Collatz problem has for $x_i > 0$ an m-cycle then $m \ge 10$.*

6.5. Remarks. 1. The divergent character of the second Matthews problem vanishes if the recurrence relation $x_{i+1} = 2x_i$ if $x_i \equiv 0 \pmod 3$ is replaced by $x_{i+1} = 2x_i/3$ if $x_i \equiv 0 \pmod 3$. Then it remains true that the above analyzed m-cycles do not exist, but other m-cycles with elements $x_i \equiv 0 \pmod 3$ may exist.

2. Recently Farkas [5] analyzed new variants of the Collatz problem. Consider a sequence of natural numbers conditionally generated by

$$x_{n+1} = \begin{cases} \dfrac{x_n}{2} & \text{if } x_n \equiv 0, 2 \pmod 4, \\ \dfrac{x_n+1}{2} & \text{if } x_n \equiv 1 \pmod 4, \\ \dfrac{3x_n+1}{2} & \text{if } x_n \equiv 3 \pmod 4. \end{cases} \tag{82}$$

In [22] isomorphy between sequences is defined (isomorphic sequences share cycle behaviour) and it is shown that (some) Farkas sequences are isomorphic to the $3x+1$ sequence.

7. THE ROELANTS PROBLEM

7.1. Introduction. H. Roelants [19] and others [14], [16] have analyzed the following *problem* R_{p_k}. Let k be fixed and let p_k be the kth prime. Then

$$x_{n+1} = \begin{cases} x_n/p_j & \text{while } x_n \equiv 0 \pmod{p_j} \text{ for } 1 \le j \le k, \\ (p_{k+1}x_n + 1)/2 & \text{else.} \end{cases} \tag{83}$$

If $x_0 = 1$ then $x_1 = (p_{k+1} + 1)/2$, which only contains factors p_j with $j \le k$, hence $x_2 = 1$ and $(1, (p_{k+1} + 1)/2)$ is a trivial cycle. For most R_{p_k} problems there exist non-trivial cycles. Roelants found that e.g. $p_k = 61$ has at least three non-trivial cycles, for $x_0 = 97, 199, 26833$, and $p_k = 281$ has at least eleven non-trivial cycles, for $x_0 = 521 \ldots \simeq 3 \cdot 10^7$. Murad [16], who independently studied these sequences, calculated for primes < 1000 all sequences with $x_0 < 10^8$. He found (numerical observation) that for 60 primes, e.g. $(3, 5, 7, 29, 41, 89, \ldots, 997)$ the R_{p_k} problem has no non-trivial cycles. The R_2 problem is equivalent to the $3x+1$ problem and for $p_k \ge 3$ the R_{p_k} problem is a non-trivial generalization of the $3x + 1$ problem. The $px + q$ problem suggests that an appropriate expression for numbers in an increasing subsequence could be $x_0 = (a2^s - 1)/(p_{k+1} - 2)$. Then $x_1 = (a2^{s-1}p_{k+1} - 1)/(p_{k+1} - 2)$ etc. In general, however, for $2 < j < k$ there exists an x_i with $i \ge 1$ for which $x_i \equiv 0 \pmod{p_j}$. As will be made clear in the following sections, our approach for finding cycles is applicable for specific m-cycles only, and for one specific cycle in general several Λ expressions apply. We sketch the applicability of our approach for 1-cycles of the R_3 problem, with $x_i > 0$.

7.2. An upper bound for Λ

1. *Definition of increasing and decreasing subsequences and choice of an appropriate expression for numbers in subsequences.* Any number can be written as one of the following expressions:

$$\begin{aligned} x_0 &= a2^r3^s, \\ x_0 &= a2^{2r}3^s - 1, \\ x_0 &= a2^{2r+1}3^s - 1, \\ x_0 &= a2^r3^s + 1. \end{aligned}$$

For the R_3 problem numbers of the first type lead to a decreasing subsequence and numbers of the last three types lead to an increasing subsequence.

Hence 1-cycles can be of complex composition, and may lead to a matrix system of variable dimension (> 1). For ease of analysis, we restrict ourselves to simple 1-cycles which contain in the increasing subsequence numbers of one type only.

2. *Derivation of chain equations.* Note that for the last type x_1 is a multiple of 3 and that $x_2 > x_0$, which gives rise to a chain equation. Adjustment of the definition of the recurrence relation for numbers $\neq 0 \pmod 6$ overcomes this ambiguity. As a consequence, a simple 1-cycle has only one unknown coefficient and there are no chain equations.

3. *Derivation of an inequality for Λ*

(a) Let $x_0 = a2^{2r}3^s - 1$ with $r > 0$ and $s > 0$. Then $x_r = a3^s5^r - 1$, which is even and not a multiple of 3. A simple 1-cycle exists if

$$a3^s5^r - 1 = 2^p(a2^{2r}3^s - 1). \tag{84}$$

If a solution exists then

$$0 < 1 - \frac{5^r}{2^{p+2r}} \leq \frac{2^p - 1}{2^{p+2r}3^s} < \frac{1}{2^{2r}3^s},$$

from which a necessary condition for Λ follows:

$$0 < (p + 2r)\log 2 - r\log 5 < \frac{1}{a2^2r3^s - 1} = \frac{1}{x_0}. \tag{85}$$

(b) Let $x_0 = a2^{2r+1}3^s - 1$ with $r > 0$ and $s > 0$. Then $x_r = 2a3^s5^r - 1$, $x_{r+1} = a3^s5^{r+1} - 2$ and $x_{r+2} = (a3^s5^{r+2} - 9)/2$, which may be even and is a multiple of 3. A simple 1-cycle exists if

$$\frac{a3^s5^{r+2} - 9}{2} = 2^p3^q(a2^{2r+1}3^s - 1). \tag{86}$$

If a solution exists (with $s = 1$, hence $q = 1$) then

$$0 < 1 - \frac{5^{r+2}}{2^{p+2r+2}3} \leq \frac{2^{p+1} - 3}{a2^{p+2r+2}3} < \frac{1}{a2^{2r+1}3},$$

from which a necessary condition for Λ follows:

$$0 < (p + 2r + 2)\log 2 + \log 3 - (r + 2)\log 5 < \frac{1}{a2^{2r+1}3 - 1} = \frac{1}{x_0}. \tag{87}$$

If a solution exists (with $s \geq 2$, hence $q = 2$) then

$$0 < 1 - \frac{5^{r+2}}{2^{p+2r+2}3^2} \leq \frac{2^{p+1} - 1}{a2^{p+2r+2}3^2} < \frac{1}{a2^{2r+1}3^2},$$

from which a necessary condition for Λ follows:

$$0 < (p + 2r + 2)\log 2 + 2\log 3 - (r + 2)\log 5 < \frac{1}{a2^{2r+1}3^2 - 1} = \frac{1}{x_0}. \tag{88}$$

(c) Let $x_0 = a2^r3^s + 1$ with $s > r > 0$. Then $x_{2r} = a3^{s-r}5^r + 1$. A simple 1-cycle exists if

$$a3^{s-r}5^r + 1 = 2^p(a2^r3^s + 1). \tag{89}$$

If a solution exists then

$$0 < \frac{5^r}{2^{p+r}3^r} - 1 \leq \frac{2^p - 1}{a2^{p+r}3^s} < \frac{1}{a2^r3^s},$$

from which a necessary condition for Λ follows:

$$0 < r\log 5 - (p+r)\log 2 - s\log 3 < \frac{1}{a2^r3^s} < \frac{b}{x_0} \tag{90}$$

with $b = X_0/(X_0-1)$.

(d) Let $x_0 = a2^r3^s + 1$ with $s = r > 0$. Then $x_{2r} = a5^r + 1$. A simple 1-cycle exists if

$$a5^r + 1 = 2^p3^q(a2^r3^s + 1). \tag{91}$$

If a solution exists then

$$0 < \frac{5^r}{2^{p+r}3^{q+s}} - 1 \le \frac{2^p - 1}{a2^{p+r}3^{q+s}} < \frac{1}{a2^r3^s},$$

from which a necessary condition for Λ follows:

$$0 < r\log 5 - (p+r)\log 2 - (q+s)\log 3 < \frac{1}{a2^r3^s} < \frac{b}{x_0} \tag{92}$$

with $b = X_0/(X_0-1)$.

(e) Let $x_0 = a2^r3^s + 1$ with $r > s > 0$. Then $x_{2s} = a2^{r-s}5^r + 1$. A simple 1-cycle exists if

$$a2^{r-s}5^r + 1 = 3^q(a2^r3^s + 1). \tag{93}$$

If a solution exists then

$$0 < \frac{5^r}{2^s3^{q+s} - 1} \le \frac{2^p - 1}{a2^s3^{q+s}} < \frac{1}{a2^r3^s},$$

from which a necessary condition for Λ follows:

$$0 < r\log 5 - s\log 2 - (q+s)\log 3 < \frac{1}{a2^r3^s} < \frac{b}{x_0} \tag{94}$$

with $b = X_0/(X_0-1)$.

In all cases we find a general necessary condition for Λ:

$$0 < |A_1\log 2 + A_2\log 3 - A_3\log 5| < \frac{1}{a2^r3^s} < \frac{b}{x_0} \tag{95}$$

where $A_i > 0$ depends on p, q, r, s as defined above.

7.3. A lower and an upper bound for K. The generalized theorem of Crandall cannot be used for finding an effective lower bound, so only brute force calculations remain. Transcendence theory supplies a (weak) upper bound (probably $K_3 \sim m^3 \max(\varrho_i)^m$). There is no simple upper bound reduction method, but lattice basis reduction algorithms remain applicable. As for the first example of the generalized Collatz problem we leave this problem for future research.

7.4. Remarks. Any specific 1-cycle for the R_2 problem can be analyzed in a similar way, but any such 1-cycle requires a new matrix of chain equations. For the general R_{p_k} problem there seems to be no appropriate expression for numbers in a hypothetical cycle, which supports analysis of all m-cycles for fixed m. This clearly demonstrates one of the limits of our approach. The similarity in the equations for Λ for the R_2 problem suggests that for an m-cycle of the general R_{p_k} problem for $\Lambda = |A_1 \log 2 + \cdots + A_{k-1} \log p_{k-1} - A_k \log p_k|$ an inequality of the type $0 < \Lambda < \sum b/\overline{x}_i$ holds. An expression for the maximal local minimum in a cycle can be found in a similar way as for the $px + q$ problem. So we conjecture that our approach can in principle be applied to find m-cycles for the Roelants problem.

8. CONCLUSIONS

8.1. General. For all these generalized Syracuse sequences a general approach applies.

1. Brute force computation gives lower bounds for K and x_i and may reveal some cycles below those bounds. These lower bounds are necessary for the number-theoretic analysis to exclude other cycles.
2. Based on a non-trivial "appropriate" expression for numbers in a possible m-cycle above those bounds, an upper bound for $\Lambda = \sum A_i \log p_i$ in terms of K and m is derived. Such expression depends on the type of generalization. Combination of this upper bound with a lower bound from transcendence theory leads to an upper bound K_3 for K. This guarantees a finite number of m-cycles for any fixed m.
3. If Λ is a linear form in two logarithms then elementary number theory gives a lower bound K_0. In this case there exist an upper bound reduction to K_2 based on convergents and an efficient search technique in an interval (K_1, K_2) with $K_0 < K_1 < K_2 < K_3$.
4. If $\Lambda = A \log x + B \log y + \cdots + C \log z$ then derivation of an upper bound for Λ for given m-cycles remains feasible. The efficient search techniques from elementary number theory collapse. Lattice reduction basis techniques remain applicable.
5. For the $3x+q$ problem a heuristic argument excludes the existence of divergent trajectories and we have an efficient algorithm for finding all m-cycles for any fixed m. Here K_1 seems positively related to q and defines the "space" for small cycles. This explains why for large q in general non-trivial cycles exist.
6. For the $px + q$ problem in general divergent trajectories exist. However, also for the $px + q$ problem this approach supplies an efficient algorithm for finding all m-cycles for any fixed m. Again K_1 seems positively related to q and defines the "space" for small cycles. This explains (partly) why the $px+1$ problem has in general no non-trivial cycles, and the $px+q$ problem with $p < q$ exceptionally has non-trivial cycles.

7. For the $3x+q$ problem and the $px+q$ problem additional conditions for the existence of m-cycles follow from elementary number theory (residue class ring arithmetic).
8. For other generalizations derivation of the structure of 1-cycles (and m-cycles) remains feasible. The derivation of the diophantine system for the coefficients a_i and the subsequent exclusion of m-cycles may become numerically (very) complex.

8.2. Limitations of the approach

8.2.1. *Dependence on the appropriate expression for x_i.* Crucial in our approach is the derivation of an expression for Λ as a linear form in logarithms. This expression follows from the expression for $\overline{x}_i$ and the chain equations for a specific m-cycle. As a result our approach fails for generalizations such as the "real" $3x+1$ problem (introduced by Konstadinidis [9]) defined by $x_i = x_i/2$ if $\lfloor x_i \rfloor$ is even, and $x_{i+1} = (3x_i+1)/2$ if $\lfloor x_i \rfloor$ is odd. Trajectories of this problem are bounded (heuristic argument, no formal proof) and must end in a limit cycle, but an "appropriate" expression for $\overline{x}_i$ cannot be found. Konstadinidis proves without using transcendence theory that cycles of the "real" $3x+1$ problem are integer cycles.

8.2.2. *The dependence on brute force computation.* Brute force computation plays an important role as an incubator for our approach. The upper bound K_3 theoretically depends (practically does not depend) on the value X_0 (e.g. once $X_0 > mq$ for the $px+q$ problem). The lower bound K_0 heavily depends on X_0 via Crandall's generalized lemma. For an efficient search we need $K_0 \approx K_1$. The post-transcendence analysis of $a_i \geq 1$ explicitly depends on X_0. The bounds K_3 and K_1 increase with m. To analyze the existence of m-cycles up to an upper bound $m \leq M$ thus requires a certain value of X_0 that increases with M. So the application range of the approach directly depends on X_0.

Acknowledgements. The author thanks Benne de Weger for generously supplying his transcendence theory *Mathematica* notebook for calculating the various lower and upper bound values for the cycle length and the anonymous referee for carefully reading and checking the manuscript.

References

[1] A. Baker, *Transcendental Number Theory*, Cambridge Univ. Press, Cambridge, 1975.

[2] E. G. Belaga and M. Mignotte, *Cyclic structure of dynamical systems associated with the* $3x+d$ *extensions of Collatz problem*, http://www-irma.u-strasbg.fr/irma/publications/2000/00017.shtml.

[3] R. E. Crandall, *On the* $3x+1$ *problem*, Math. Comp. 32 (1978), 1281–1292.

[4] J. L. Davison, *Some comments on an iteration problem*, in: Proceedings of the 6th Manitoba Conference on Numerical Mathematics and Computing (Winnipeg, Man., 1976), Utilitas Math., Winnipeg, Man., 1977, 155–159.
[5] H. Farkas, *Variants of the* $3N + 1$ *Conjecture and Multiplicative Semigroups*, Contemp. Math. 387, Amer. Math. Soc., Providence, RI, 2005.
[6] R. K. Guy, *Unsolved Problems in Number Theory*, 2nd ed., Springer, Berlin, 1994.
[7] G. H. Hardy and E. M. Wright, *The Theory of Numbers*, 4th ed., Oxford Univ. Press, Oxford, 1975.
[8] M. Hindry and J. H. Silverman, *Diophantine Geometry*, *An Introduction*, Grad. Texts in Math. 201, Springer, Berlin, 2000.
[9] P. B. Konstadinidis, *The real* $3x + 1$ *problem*, Acta Arith. 122 (2006), 35–44.
[10] J. C. Lagarias, *The* $3x + 1$ *problem and its generalizations*, Amer. Math. Monthly 92 (1985), 3–23.
[11] —, *The set of rational cycles for the* $3x + 1$ *problem*, Acta Arith. 56 (1990), 33–53.
[12] —, *The* $3x+1$ *problem*: *An annotated bibliography*, http://arxiv.org/abs/math.NT/0309224.
[13] M. Laurent, M. Mignotte et Yu. Nesterenko, *Formes linéaires en deux logarithmes et déterminants d'interpolation*, J. Number Theory 55 (1995), 285–321.
[14] J. Lesieutre, *On a generalization of the Collatz conjecture*, http://web.mit.edu/course/other/rsi/compendium/edit2004/Final/lesieutre-john-mit-both.pdf.
[15] K. R. Matthews, *The generalized* $3x + 1$ *mapping*, http://www.numbertheory.org/pdfs/survey.pdf.
[16] A. Murad, *A generalized Collatz sequence*, http://www.primepuzzles.net/puzzles/.
[17] T. Oliveira e Silva, *Computational verification of the* $3x + 1$ *conjecture*, http://www.ieeta.pt/˜tos/3x+1.html.
[18] G. Rhin, *Approximants de Padé et mesures d'irrationalité*, in: Progr. Math. 71, Birkhäuser, 1987, 155–164.
[19] H. Roelants, personal communication.
[20] E. Roosendaal, *Status of distributed computation on the* $3x + 1$ *problem*, http://personal.computrain.nl/eric/wondrous/index.html.
[21] J. L. Simons, *Post-transcendence conditions for the existence of* m*-cycles of the* $3x + 1$ *problem*, submitted, 2006.
[22] —, *On isomorphism between Farkas sequences and Collatz sequences*, http://www.rug.nl/staff/j.l.simons/Ontheisomorphy.pdf.
[23] J. L. Simons and B. M. M. de Weger, *Theoretical and computational bounds for* m*-cycles of the* $3n + 1$*-problem*, Acta Arith. 117 (2005), 51–70; see for latest results http://www.win.tue.nl/˜bdeweger/research.html.
[24] B. M. M. de Weger, *Algorithms for Diophantine Equations*, CWI Tract 65, Centre for Math. and Comput. Sci., Amsterdam, 1990.
[25] G. J. Wirsching, *The Dynamical System Generated by the* $3n + 1$ *Function*, Lecture Notes in Math. 1681, Springer, Berlin, 1998.

University of Groningen
P.O. Box 800, 9700 AV Groningen, The Netherlands
E-mail: j.l.simons@rug.nl

Received on 5.12.2006
and in revised form on 7.8.2007 (5339)