ACTA ARITHMETICA
134.3 (2008)

Exotic Collatz cycles

by

John L. Simons (Groningen)

1. Introduction. The generalized Collatz problem (also called the $px+q$ problem with p an odd prime and q odd) is defined by a sequence of natural numbers, generated conditionally by $x \mapsto x/2$ if x is even and by $x \mapsto (px+q)/2$ if x is odd. We restrict ourselves to the case where $p \geq 5$ (prime), q odd and $\mathrm{GCD}(p,q) = 1$. For each factor $c > 1$ of q, either none or all of the numbers x_i in each trajectory (including a hypothetical cycle) must satisfy $\mathrm{GCD}(x_i, q) = c$. Because each cycle of the $px+q$ problem with $\mathrm{GCD}(x_i, q) = c$ corresponds with a cycle of the $px + q/c$ problem, we call cycles with $\mathrm{GCD}(x_i, q) = 1$ *primitive* and any other cycle *non-primitive*. An m-cycle of the $px+q$ problem has m local minima x_i. We call an m-cycle ($m \geq 2$) *trivial* if it is a multiple of an m^*-cycle with $m^* < m$. In a non-trivial m-cycle we have $x_i \neq x_j$ if $i \neq j$ and we assume in what follows that an m-cycle is non-trivial unless explicitly stated otherwise.

Simons [5] has proved that in an m-cycle $(p-2)x_i = a_i 2^{k_i} - q$ and consequently that 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

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

He further derives the following two existence conditions for 1-cycles:

Lemma 1. *A necessary and sufficient condition for the existence of a 1-cycle for the $px+q$ problem is the existence of positive integers k, l and r*

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

Key words and phrases: generalized Collatz problem, cycle existence.

(odd) such that $2^{k+l} - p^k = q \cdot r$ *and the existence of an odd integer* x_0 *such that*

$$x_0 = \frac{p^k - 2^k}{(p-2)r}.$$

COROLLARY 2. *If, and only if,*

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

then there exists a primitive 1*-cycle with minimal element* x_0*.*

Let $C(m,p,q)$ be the number of primitive m-cycles of the $px+q$ problem and let $C(p,q) = \sum_m C(m,p,q)$. Let $B(m,p,q)$ be the number of primitive and non-primitive m-cycles of the $px+q$ problem and let $B(p,q) = \sum_m B(m,p,q)$. Let $S(p)$ be the set of primes q such that

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

has a solution with $k \geq 2$. In this paper we derive the following results:

- For each p, there exist infinitely many $px+q$ problems with $C(1,p,q) \geq 1$.
- For each p and $d > 0$, there exist infinitely many $px+q$ problems with $C(p,q) > d$.
- Let p be fixed. If $S(p)$ is an infinite set (conjectured, no proof) then for each $d > 0$ there exist infinitely many $px+q$ problems (q prime) with $C(p,q) > d$.

Matthews [4] conjectured that (i) if $p = 3$ then all trajectories end in a cycle, (ii) if $p \geq 5$ then almost all trajectories are divergent, (iii) $B(p,q) < \infty$. Lagarias [3] conjectured (i) $C(3,q) \geq 1$, (ii) $C(3,q) < \infty$ and proved that occasionally $C(3,q)$ takes large values. Brox [2] and Simons & de Weger [6] proved $C(m,3,1) < \infty$. Simons [5] proved $C(m,p,q) < \infty$. Belaga & Mignotte [1] numerically showed that there exist $3x+q$ problems with many primitive cycles. Our results formally agree with Matthews' second and third conjecture, however $px+q$ problems with arbitrarily many cycles can be seen as exotic exceptions to the empirical divergent behavior of $px+q$ problems ($p \geq 5$).

2. Generalized Collatz sequences with q odd. From Lemma 1 it follows directly that for each prime p there exist infinitely many q-values for which the $px+q$ problem has a 1-cycle. Simply set $q = 2^{k+l} - p^k$ for each pair (k,l). Then Lemma 1 applies with $r = 1$. Consequently, the $px+q$ problem has a 1-cycle with minimal element $x_0 = (p^k - 2^k)/(p-2)$. For each p we can construct a value for q such that the $px+q$ problem has a primitive 1-cycle.

Lemma 3. *Let $p \geq 5$ be a prime. For each $k \geq 2$ and $l > (\log_2 p - 1)k > k$, let*

$$u = \frac{p^k - 2^k}{p - 2} \quad \textit{and} \quad v = 2^{k+l} - p^k.$$

The $px + q$ problem with $q = v/\mathrm{GCD}(u, v)$ has a primitive 1-cycle.

Proof. Set $x_0 = u/\mathrm{GCD}(u, v)$. Then

$$x_0 = \frac{u}{\mathrm{GCD}(u, v)} = \frac{u \cdot q}{v} = \frac{(p^k - 2^k)q}{(p - 2)(2^{k+l} - p^k)}. \tag{2}$$

Further, we have

$$\mathrm{GCD}(x_0, q) = \mathrm{GCD}\left(\frac{u}{\mathrm{GCD}(u, v)}, \frac{v}{\mathrm{GCD}(u, v)}\right) = 1.$$

We now distinguish two cases:

1. $q = 1$. Then $2^{k+l} - p^k = v = q \cdot v$ and $x_0 = (p^k - 2^k)/(p - 2)v$. Hence Lemma 1 applies.
2. $q \geq 3$. Because $\mathrm{GCD}(x_0, q) = 1$, Corollary 2 applies.

In both cases the $px + q$ problem has a primitive 1-cycle. ∎

Lemma 3 generates for fixed p infinitely many values $q = v/\mathrm{GCD}(u, v)$ for which the $px + q$ problem has a primitive 1-cycle. These q-values are not necessarily different. However, each new pair (k, l) generates a new pair (u, v). Now either the $px + v$ problem has a primitive 1-cycle with k odd and l even elements or the $px + q$ with $q = v/\mathrm{GCD}(u, v)$ problem has such a primitive 1-cycle. Because $C(1, p, v) < \infty$ and $C(1, p, q) < \infty$ the infinite sequence of pairs $\{k, l\}$ must generate an infinite sequence either of $px + v$ problems with a primitive 1-cycle and/or an infinite sequence of $px + q$ problems with a primitive 1-cycle. Hence we have

Corollary 4. *For each p there are infinitely many q-values for which the $px + q$ problem has a primitive 1-cycle.*

Consider as an example the case $p = 5$:

- For $k = 2$ and $l = 3$ we have $u = v = 7$, hence $x_0 = q = 1$. The $5x + 1$ problem has the primitive 1-cycle $(1, 3, 8, 4, 2)$.
- For $k = 2$ and $l = 4$ we have $u = 7$ and $v = 39$, hence $x_0 = 7$ and $q = 39$. The $5x+39$ problem has the primitive 1-cycle $(7, 37, 112, 56, 28, 14)$.
- For $k = 2$ and $l = 5$ we have $u = 7$ and $v = 103$, hence $x_0 = 7$ and $q = 103$. The $5x + 103$ problem has the primitive 1-cycle $(7, 69, 224, 112, 56, 28, 14)$.
- For $k = 3$ and $l = 4$ we have $u = 39$ and $v = 3$, hence $x_0 = 13$ and $q = 1$. The $5x+1$ problem has the primitive 1-cycle $(13, 33, 83, 208, 104, 52, 26)$.

Note that q can be 1, prime or composite. Because the $5x+1$ problem has exactly two 1-cycles [5], for no other pair (k,l) can reduction to the $5x+1$ problem occur.

If for the $px+q$ problem a 1-cycle exists with $k \geq 2$, then m-cycles with $m \geq 2$ also exist. This follows from the general solution of the system (1):

$$\Delta a_i = q \cdot [p^{k_{i+1}+k_{i+2}+\cdots+k_{i-1}}(2^{l_i}-1) + 2^{k_{i+1}+l_i}p^{k_{i+2}+\cdots+k_{i-1}}(2^{l_{i+1}}-1) + \cdots + 2^{k_{i+1}+l_i+k_{i+2}+l_{i+1}+\cdots+k_{i-1}+l_{i-2}}(2^{l_{i-1}}-1)].$$

If $\Delta = 2^{k+l}-p^k = q$ then each choice for k_i and l_i results in an integral a_i. Given $\sum_{i=0}^{m-1} k_i = k \geq 2$ and $\sum_{i=0}^{m-1} l_i = l \geq 3$ and any choice for k_i, we can choose (cyclic) different values for l_i which result in a new m-cycle.

- The $5x+7$ problem has the non-primitive 1-cycle $(7, 21, 56, 28, 14)$ with $k=2$ and $l=3$. There exists one 2-cycle: ($k_0=k_1=1,\ l_0=1,\ l_1=2$) resulting in $(9, 26, 13, 36, 18)$, which is primitive.
- The $5x+39$ problem has the primitive 1-cycle $(7, 37, 112, 56, 28, 14)$ with $k=2$ and $l=4$. There exist two 2-cycles: ($k_0=k_1=1,\ l_0=1,\ l_1=3$) resulting in $(9, 42, 21, 72, 36, 18)$, which is non-primitive, and ($k_0=k_1=1,\ l_0=l_1=2$) resulting in $(13, 52, 26, 13, 52, 26)$, which is non-primitive and trivial.
- The $5x+103$ problem has the primitive 1-cycle $(7, 69, 224, 112, 56, 28, 14)$ with $k=2$ and $l=5$. There exist two 2-cycles: ($k_0=k_1=1,\ l_0=1,\ l_1=4$) resulting in $(9, 74, 37, 144, 72, 36, 18)$ and ($k_0=k_1=1,\ l_0=2,\ l_1=3$) resulting in $(13, 84, 42, 21, 104, 52, 26)$. Both cycles are primitive.
- The $5x+3$ problem has the non-primitive 1-cycle $(39, 99, 249, 624, 312, 156, 78)$ with $k=3$ and $l=4$. There exist three 2-cycles: ($k_0=1,\ k_1=2,\ l_0=2,\ l_1=2$) resulting in $(51, 129, 324, 162, 81, 204, 102)$, which is non-primitive, and ($k_0=1,\ k_1=2,\ l_0=1,\ l_1=3$) resulting in $(43, 109, 274, 137, 344, 172, 86)$ and ($k_0=1,\ k_1=2,\ l_0=3,\ l_1=1$) resulting in $(53, 134, 67, 169, 424, 212, 106)$, which are both primitive. There exists one 3-cycle: ($k_0=k_1=k_2=1,\ l_0=l_1=1,\ l_2=2$) resulting in $(61, 154, 77, 194, 97, 244, 122)$, which is primitive.

We denote by $B^*(m,p,q)$ (resp. $C^*(m,p,q)$) the number of m-cycles (resp. primitive m-cycles) generated by partitioning from a 1-cycle. Clearly, $B(m,p,q) \geq B^*(m,p,q)$ etc. For special $px+q$ problems, $B(m,p,q)$, $C(2,p,q)$ and consequently $C(p,q)$ can be arbitrarily large.

LEMMA 5. *For each prime $p \geq 5$, m and $d > 0$ there exist infinitely many $q = 2^{k+l}-p^k$ for which $B(m,p,q) > d$.*

Proof. For fixed p,k,l,m, a (cyclic) new pair of partitions

$$[(k_0, k_1, \ldots, k_{m-1}), (l_0, l_1, \ldots, l_{m-1})]$$

generates a new m-cycle. Note that q follows from p, k, l. For fixed p and m, $B^*(m, p, q)$ is an increasing function of k and l. So for any p, m and $d > 0$ we can choose k^*, l^* such that if $k > k^*$ and $l > l^*$ then $B^*(m, p, q) > d$. Consequently, $B(m, p, q) \geq B^*(m, p, q) > d$. ∎

LEMMA 6. *For each prime $p \geq 5$ and $d > 0$ there exist infinitely many $q = 2^{k+l} - p^k$ for which $C(p, q) \geq C(2, p, q) > d$.*

Proof. For each $d > 0$ we can choose k, l (hence q) such that $\mathrm{GCD}(k, l) = 1$ and $B^*(2, p, q) > 2d$. Let $x_0(j)$ be a local minimal element of the jth 2-cycle. From the matrix system (1) it follows for $k_0 = k - 1$, $k_1 = 1$, $l_0 = j$, $l_1 = l - j$ that

$$a_0(j) = (p-2)2^j + (2^{l+1} - p). \tag{3}$$

Using $(p - 2)x_0 = a_0 2^{k_0} - q$ we find

$$x_0(j) = \frac{p^k - 2^k}{p - 2} + (2^j - 1) \cdot 2^{k-1}. \tag{4}$$

We observe that $x_0(0)$ is the global minimum of the 1-cycle, hence $x_0(j+1) - x_0(j) = 2^{j+k}$ for all j, from which it follows that if the jth 2-cycle is non-primitive, then the $(j-1)$th and the $(j+1)$th 2-cycle is primitive. Such a primitive 2-cycle is non-trivial because $\mathrm{GCD}(k, l) = 1$. As a consequence we find $C(p, q) \geq C(2, p, q) \geq C^*(2, p, q) \geq \frac{1}{2}B^*(2, p, q) > d$. ∎

For $m \geq 3$ we cannot prove that $C(m, p, q)$ is arbitrarily large. For fixed p, m, d and each $d^* > 0$ we can choose q such that $B(m, p, q) \geq B^*(m, p, q) > d^*$. Now suppose $C(m, p, q) \leq d \leq d^*$. Then for this pair (p, q) the number of non-primitive m-cycles is $B(m, p, q) - C(m, p, q) > d^* - d \geq 0$. Each non-primitive m-cycle reduces to a different primitive m-cycle for a $px + r$ problem with $r \mid q$. Now $C(m, p, r) < \infty$ and consequently $B^*(m, p, q) \leq B(m, p, q) = \sum_r C(m, p, r) < \infty$. This upper bound for $B^*(m, p, q)$ could conflict with the lower bound d^* if the number of different factors of q is small enough. In such cases the assumption $C(m, p, q) \leq d \leq d^*$ cannot be true, so $C(m, p, q)$ is arbitrarily large. However if the number of different factors of q is sufficiently large, the contradiction of the lower and upper bound of $C(m, p, q)$ cannot be reached.

3. Generalized Collatz sequences with q an odd prime. Recall that $S(p)$ is the set of primes q such that $2^{k+l} - p^k = q$ has a solution with $k \geq 2$, thereby excluding the trivial 1-cycle $(1, 2^{l-1}, \ldots, 2)$ for $p+q = 2^l$. Such $px+q$ problems form an interesting subset because of the empirical exceptionality of non-primitive cycles, i.e. $B(m, p, q) \simeq C(m, p, q)$. Also, for these $px+q$ problems, $C(2, p, q)$ and consequently $C(p, q)$ can be arbitrarily large.

Lemma 7. *For fixed p, consider the set of $px+q$ problems with $q \in S(p)$. If $S(p)$ is an infinite set, then for each $d > 0$ there are infinitely many $px+q$ problems with $C(2,p,q) > d$.*

Proof. Let $q \in S(p)$ satisfy $B(2,p,q) \geq B^*(2,p,q) \simeq (k-1)(l-1)/2 > 2d$. Consider the set of 2-cycles generated by partitioning from the 1-cycle with k odd and l even elements, starting with $x_0 = (p^k - 2^k)/(p-2)$. Similarly to the proof of Lemma 6 we have $C^*(2,p,q) \geq \frac{1}{2}B^*(2,p,q)$. Consequently, $C(2,p,q) \geq C^*(2,p,q) \geq \frac{1}{2}B^*(2,p,q) > d$. Since $S(p)$ is an infinite set, for infinitely many $k^* > k$ and $l^* > l$ another q^* (prime) results for which the same reasoning applies. ∎

Note that $C(2,p,1) < \infty$ and only exceptionally (empirical fact, no proof) $C(2,p,1) > 0$. For a generalization of the proof of Lemma 7 to $C(m,p,q) > d$, it is required that for every $k > 1$ there exist infinitely many primes $q \in S(p)$. Since $C(p,q) \geq C(2,p,q)$ we have without this requirement:

Corollary 8. *For fixed p, consider the set of $px+q$ problems with $q \in S(p)$. If $S(p)$ is an infinite set, then for each $d > 0$ there are infinitely many $px+q$ problems with $C(p,q) > d$.*

Trivial m-cycles can exist by definition, but they need not be generated by partitioning if q is prime.

Lemma 9. *If $2^{k+l} - p^k = q$ with p, q prime and $k \geq 2$, for infinitely many p all the m-cycles of the $px+q$ problem, generated by partitions from a 1-cycle, are non-trivial.*

Proof. For the existence of a trivial m-cycle we must have $\mathrm{GCD}(k,l) = c > 1$. Then $k = c \cdot v$ and $l = c \cdot w$ and $q = [2^{v+w}]^c - [p^v]^c = [2^{v+w} - p^v] \cdot [2^{(v+w)(c-1)} + \cdots + p^{v(c-1)}]$. Now q is prime and the second factor is $> p+2$, so $2^{v+w} - p^v = 1$ hence $v = 1$ and $p = 2^{w+1} - 1$. Consequently, if p is not a Mersenne prime, then all generated m-cycles are non-trivial. If $p = 2^{w+1} - 1$ and $q = \sum_{j=0}^{c-1} \binom{c-1}{j} 2^{(v+w)(c-1-j)} p^{vj}$ are both prime, then a non-trivial non-primitive cycle with length $v+w$ exists. ∎

As a consequence of Corollary 8 and Lemma 9 we have

Corollary 10. *Consider the $px+q$ problem with $p \geq 5$ prime and $q \in S(p)$. If $S(p)$ is an infinite set, then for each p and $d > 0$ there are infinitely many $px+q$ problems with $C(p,q) > d$.*

Simons [5] gives for $p < 100$ the smallest $q \in S(p)$. They are shown in Table 1.

Exotic candidates are $px+q$ problems with large values for k and l. For example, the $23x + 4217$ problem with $k = 3$ and $l = 11$ has 26 primitive cycles, and the $97x + 32641759$ problem with $k = 3$ and $l = 22$ has 92 primitive cycles.

Table 1. $px + q$ problem (q minimal) with primitive cycles

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		

We computed $C(p, q)$ for $p = 5, 7$ and $q \in S(p)$ ($q < 10^6$). These are presented in Table 2.

Table 2. $C(5, q)$ and $C(7, q)$ as a function of k and l

q	k	l	$C(5,q) \geq$	q	k	l	$C(7,q) \geq$
3	3	4	3	79	2	5	3
7	2	3	3	463	2	7	5
103	2	5	3	1999	2	9	5
131	3	5	9	5791	4	9	62
487	2	7	4	30367	4	11	95
971	5	7	66	32719	2	13	7
1423	4	7	37	130729	3	14	40
8167	2	11	6	131023	2	15	8
13259	5	9	173	521887	4	15	206
32143	4	11	95				
130447	4	13	140				
259019	5	13	489				

Notice that $B^*(p, q)$ is an increasing function of k and l. These tables confirm that $C(p, q) \simeq B(p, q) \geq B^*(p, q)$ is an increasing function of k and l.

4. Remarks. 1. The infinity of $S(p)$ is a conjecture (as the Mersenne conjecture). We found that $3, 7, 103, 131, 487, 971, 1423, 8167, 13259, 32143, 130447, 259019, 1706527, 4191179, 16699091, 16774091$, 18280739, 33163807, $\ldots \in S(5)$. This $S(p)$ conjecture is milder than the Mersenne conjecture because $k \geq 2$ instead of $k = 0$ is required for the primality of q.

2. If $q \in S(p)$ and $C(1,p,q)=1$, then $B^*(p,q)$ can easily be calculated. Let $v(a,b)$ be the number of (cyclic) different and $w(a,b)$ be the total number of b-partitions of a. For fixed k and l (note that $l > k$) there are: one 1-cycle, $v(k,2)\cdot w(l,2)$ 2-cycles etc. We checked that in many cases (not always) this lower bound equals the empirically found number of primitive cycles for $p=5,7$. Exceptions are the $5x+7$ problem with no primitive 1-cycle and the $3x+463$ problem with an extra 3-cycle.

3. For the $px+q$ problem, the minimal elements in found cycles of equal length show a regular pattern. Let $x_0(m,j)$ be the minimal element of the jth primitive m-cycle. From the matrix system (1) it follows for $m=2, k_0=k-1, l_0=j$ that

$$x_0(2,j)=\frac{p^k-2^k}{p-2}+(2^j-1)\cdot 2^{k-1}. \tag{5}$$

A similar expression can be found for $x_1(2,j)$. For small j we have $x_0<x_1$ and consequently successive minima of 2-cycles for $j=1,\dots,l-1$ differ by successive powers of 2. Because the minimal element is $\min(x_0,x_1)$ this regular pattern can be disturbed for larger j.

4. For the $23x+4217$ problem all cycles have length 14 with $k=3$ and $l=11$. We found next to $x_0(1,1)=579$ the sequence of minima: $x_0(2,j)=583, 591, 607, 639, 703, 831, 1087$ for 2-cycles. The next elements of this sequence: $1599, 2623, 4671$, appear to be local minima in a 2-cycle, while the global minima are $929, 729, 629$. They form the beginning of a similar sequence for the global minima of 3-cycles. So next to $x_0(2,9)=929$ we found $x_0(3,j)=961, 1025, 1152, 1409$ for 3-cycles etc.

5. For the $7x+521887$ problem, $k=4$ and $l=15$. From partitioning we found 204 cycles ($m\le 4$) with length 19 and with $x_0(m,j)$ in a regular pattern. We found two extra cycles ($m=5,6$) with length 38 and with a minimal element smaller than $x_0(1,1)$.

6. In this paper we have proved that $C(p,q)$ can be arbitrarily large. In view of the regular behavior of the minimal x_i over the cycles we did not apply transcendence theory (as in [5]) to exclude cycles with large cycle length. This leaves theoretically open that $C(5,q)$ and $C(7,q)$ are greater than the numbers indicated in Table 2.

References

[1] 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.

[2] T. Brox, *Collatz cycles with few descents*, Acta Arith. 92 (2000), 181–188.

[3] J. C. Lagarias, *The set of rational cycles for the* $3x+1$ *problem*, ibid. 56 (1990), 33–53.

[4] K. R. Matthews, *The generalized* $3x+1$ *mapping*, http://www.numbertheory.org/pdfs/survey.pdf.
[5] J. L. Simons, *On the (non-)existence of* m*-cycles for generalized Syracuse sequences*, Acta Arith. 131 (2008), 217–254.
[6] J. L. Simons and B. M. M. de Weger, *Theoretical and computational bounds for* m*-cycles of the* $3n+1$*-problem*, ibid. 117 (2005), 51–70; see for latest results http://www.win.tue.nl/~bdeweger/research.html.

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

Received on 1.3.2007
and in revised form on 15.5.2008 (5399)