On Rowland's sequence

Fernando Chamizo, Dulcinea Raboso and Serafín Ruiz-Cabello*

Department of Mathematics
Universidad Autónoma de Madrid
28049 Madrid. Spain

fernando.chamizo@uam.es
dulcinea.raboso@uam.es
serafin.ruiz@uam.es

Submitted: Apr 30, 2011; Accepted: May 20, 2011; Published: May 29, 2011
Mathematics Subject Classification: 11A41 (11B37)

Thank you Professor Zeilberger, for the unforgettable Experimental Mathematics Seminars 2009-2010!

Abstract

E. S. Rowland proved that $a_k = a_{k-1} + \gcd(k, a_{k-1})$, $a_1 = 7$ implies that $a_k - a_{k-1}$ is always 1 or prime. Conjecturally this property also holds for any $a_1 > 3$ from a certain k onwards. We state some properties of this sequence for arbitrary values of a_1. Namely, we prove that some specific sequences contain infinitely many primes and we characterize the possible finite subsequences of primes.

1 Introduction

In [4] E. S. Rowland introduced the recursively defined sequence

$$a_k = a_{k-1} + \gcd(k, a_{k-1}), \qquad a_1 = 7. \tag{1}$$

He proved the following suprising result:

Theorem 1.1 (Rowland [4]) *Let $\mathcal{P}$ be the set of prime numbers and $\mathcal{P}_1 = \mathcal{P} \cup \{1\}$. Then $a_k - a_{k-1} \in \mathcal{P}_1$ for every $k > 1$.*

*The first and the third authors are supported by the grant MTM2008-03880 of the Ministerio de Ciencia e Innovación. The second author is also a member of ICMAT.

Unfortunately it is not clear whether the proof applies to all possible values of a_1. Note that $a_1 = 2A$ and $a_1 = 2A + 1$ give the same a_2, so we can restrict ourselves to odd initial conditions. It is easy to check that $a_1 = 1$ and $a_1 = 3$ lead to the sequences $a_k = k$ and $a_k = k + 2$, respectively. Hence, in this paper we only consider the sequences

$$a_k = a_{k-1} + \gcd(k, a_{k-1}) \qquad \text{with } a_1 \text{ odd and greater than 3.} \tag{2}$$

Conjecture 1.2 *For any sequence of the form* (2)*, there exists a positive integer N such that $a_k - a_{k-1} \in \mathcal{P}_1$ for every $k > N$.*

Actually in [4] this conjecture is stated for starting values of the form $a_{k_0} = A$. We consider the former statement more natural (although less general) and, as we shall see, there are some differences between the two situations.

We refer the reader to [1] for some other conjectures about related sequences.

Our approach depends on the introduction of two auxiliary recurrences. They are a version of the 'shortcut' mentioned in [4]. Before giving the actual definitions we motivate them including here a very simple proof of Theorem 1.1 in a stronger form, using the sequences

$$c_n^* = c_{n-1}^* + \mathrm{lpf}(c_{n-1}^*) - 1, \qquad c_1^* = 5 \qquad \text{and} \qquad r_n^* = \frac{c_n^* + 1}{2}, \tag{3}$$

where $\mathrm{lpf}(\cdot)$ denotes the least prime factor. Note that c_n^* is odd for every n.

Proposition 1.3 *Let $\{a_k\}_{k=1}^{\infty}$ be Rowland's sequence* (1)*. Then*

$$a_k - a_{k-1} = \begin{cases} \mathrm{lpf}(c_{n-1}^*) & \text{if } k = r_n^* \text{ for some } n > 1, \\ 1 & \text{otherwise.} \end{cases} \tag{4}$$

Proof: Define $x_1 = 7$, $x_2 = 8$, and $x_k = c_n^* + k + 1$ for $k \in [r_n^*, r_{n+1}^*)$, $n \geq 1$. In this interval; $x_{k-1} = c_n^* + k$ for $k \neq r_n^*$, and $x_{k-1} = c_{n-1}^* + k$ for $k = r_n^* > 3$. Then $x_k - x_{k-1}$ is equal to the right hand side of (4). To deduce $x_k = a_k$, we only need to prove that it is also equal to $\gcd(k, x_{k-1})$. For $k \neq r_n^*$

$$\gcd(k, x_{k-1}) = \gcd(k, c_n^* + k) = \gcd(2k, c_n^*) = \gcd(2(k - r_n^*) + 1, c_n^*)$$

and this is 1, since $2(k - r_n^*) + 1 < 2(r_{n+1}^* - r_n^*) + 1 = \mathrm{lpf}(c_n^*)$. For $k = r_n^*$ the result is the same replacing n by $n - 1$, hence $2(k - r_{n-1}^*) + 1 = 2(r_n^* - r_{n-1}^*) + 1 = \mathrm{lpf}(c_{n-1}^*)$. $\square$

This short proof of Theorem 1.1 suggests that we introduce a general recurrence

$$\begin{cases} r_{n+1} = \min\left\{p + p\lfloor r_n/p \rfloor \ : \ p \mid c_n\right\} \\ c_{n+1} = c_n + \gcd(c_n, r_{n+1}) - 1 \end{cases} \qquad \text{with } r_1 = 1 \text{ and } c_1 = a_1 - 2, \tag{5}$$

where $\lfloor \cdot \rfloor$ denotes the integral part and p denotes a prime number. It is easy to check that $r_n = r_n^*$ and $c_n = c_n^*$ satisfy this recurrence for $n > 1$, where here r_n^* and c_n^* are as in (3). Again c_n is odd for every n. An elementary argument gives an alternative expression to the recurrence for r_n showing that r_{n+1} is the smallest number above r_n being not coprime to c_n (see Lemma 2.1 below and cf. Proposition 3 [4]).

The sequence (2) is determined by (5). Indeed r_n gives the indices k for which $a_k - a_{k-1} \neq 1$. The analogue of Proposition 1.3 is

Proposition 1.4 *The sequence* (2) *satisfies*

$$a_k = c_n + k + 1 \qquad \textit{for} \;\; r_n \leq k < r_{n+1}, \tag{6}$$

where r_n *and* c_n *are defined by* (5). *Moreover,* $a_k - a_{k-1}$ *equals* $\gcd(c_{n-1}, r_n)$ *for* $k = r_n$, *and equals* 1 *otherwise.*

Rowland notes that his proof applies when $a_k = 3k$ for some k (it occurs in (1) when $k = 3$). With our approach it corresponds to $c_n = 2r_n - 1$ for some n which indeed implies $c_l = 2r_l - 1$ for $l > n$. On the other hand, the underlying idea in several number theoretical conjectures (e.g. Schinzel's hypothesis [5] or Hardy-Littlewood k-tuples conjecture [3], [2, IV.2]) is that prime numbers should appear in a sequence if no local divisibility conditions prevent it. Then a natural guess is that c_m is prime for some m. Curiously it seems that the minimal choices of m and n in these claims are always consecutive.

For instance, if $a_1 = 117$ we have

$$\begin{array}{cccccccc} r_1 = 1, & r_2 = 5, & r_3 = 7, & r_4 = 10, & r_5 = 12, & r_6 = 131, & \dots \\ c_1 = 115, & c_2 = 119, & c_3 = 125, & c_4 = 129, & c_5 = 131, & c_6 = 261, & \dots \end{array}$$

Here $c_n = 2r_n - 1$ for the first time when $n = 6$, and the first prime value of c_m occurs for $m = 5$. We have checked every $a_1 < 10^8$ and the experiments suggest

Conjecture 1.5 *Consider the recurrence* (5) *for odd* $a_1 > 3$, *and define*

$$n_0 = \inf\{n \in \mathbb{Z}^+ \; : \; c_n = 2r_n - 1\} \qquad \textit{and} \qquad m_0 = \inf\{n \in \mathbb{Z}^+ \; : \; c_n \textit{ is prime}\},$$

writing conventionally $\inf \emptyset = \infty$ *as usual. Then*

$$\text{(i)} \quad n_0 < \infty, \qquad\qquad \text{(ii)} \quad m_0 < \infty, \qquad\qquad \text{(iii)} \quad n_0 = m_0 + 1 < \infty.$$

In §2 we provide some theoretical support and equivalences. In §3 we include some properties of the set of primes generated by the sequences (2). Any of the three statements in Conjecture 1.5 implies Conjecture 1.2 (Proposition 3.2). In terms of a_k, (iii) implies that the first k for which $a_k - a_{k-1} \neq 1$ and $a_k = 3k$ is necessarily prime (Proposition 2.6).

For sequences not starting at a_1, the latter primality property and (iii) admit counterexamples. One of the simplest is $a_{59} = 153$ that satisfies Proposition 1.4 allowing in (5) the case $r_1 = 59$ and $c_1 = 93$. The values

$$\begin{array}{lllll} r_1 = 59, & r_2 = 60, & r_3 = 65, & r_4 = 66, & \dots \\ c_1 = 93, & c_2 = 95, & c_3 = 99, & c_4 = 131, & \dots \end{array}$$

show that $a_k = 3k$ for the first time for $k = 66$, which corresponds to $c_4 = 2r_4 - 1$. But neither $r_4 = 66$ nor $c_3 = 99$ are prime.

2 Relation between the conjectures

We start by giving the alternative formula for r_{n+1} and the proof of Proposition 1.4.

Lemma 2.1 *For* $n, m \in \mathbb{Z}^+$

$$\min\left\{p + p\left\lfloor\frac{n}{p}\right\rfloor \;:\; p \mid m\right\} = \min\left\{k > n \;:\; \gcd(k, m) \neq 1\right\}.$$

Proof: Lemma follows from the fact that

$$p + p\left\lfloor\frac{n}{p}\right\rfloor = p\left(1 + \left\lfloor\frac{n}{p}\right\rfloor\right)$$

is the first multiple of p that is greater than n, and that $p \mid \gcd(p + p\lfloor n/p\rfloor, m)$. $\square$

Proof of Proposition 1.4: If $r_n < k < r_{n+1}$ then, by Lemma 2.1, we have that $\gcd(k, c_n) = 1$ and therefore

$$\gcd(k, c_n + k) = 1 = (c_n + k + 1) - (c_n + k) = a_k - a_{k-1}.$$

On the other hand, if $k = r_n$, then $\gcd(r_n, c_{n-1}) \neq 1$ and clearly

$$\gcd(r_n, c_{n-1} + r_n) = \gcd(r_n, c_{n-1}) = (c_n + r_n + 1) - (c_{n-1} + r_n) = a_k - a_{k-1}.$$

This proves (6). Now it is clear that $a_k - a_{k-1}$ is $\gcd(c_{n-1}, r_n)$ for $k = r_n$ and 1 otherwise. $\square$

The following unconditional relation between r_n and c_n plays an important role when relating the conjectures. Compare it with Proposition 1 and Propostion 2 in [4] and the comments given there. Note for instance that after (6) $a_k \geq 3k$ for $k = r_n$.

Proposition 2.2 *Let r_n and c_n be given by (5) with a_1 odd > 3. Then, $r_n \leq (c_n + 1)/2$ for every $n \in \mathbb{Z}^+$. Moreover, the equality for $n > 1$ occurs if and only if $\gcd(c_{n-1}, r_n)$ is a prime p and $p\lfloor r_{n-1}/p\rfloor = (c_{n-1} - p)/2$.*

Proof: We prove the inequality by induction. Clearly is it true for $n = 1$. Assume $r_{n-1} \leq (c_{n-1} + 1)/2$. By definition, $r_n = p + p\lfloor r_{n-1}/p \rfloor$ for some prime $p \mid c_{n-1}$. Apply now the inductive hypothesis

$$r_n = p + p \left\lfloor \frac{r_{n-1}}{p} \right\rfloor \leq p + p \left\lfloor \frac{c_{n-1}+1}{2p} \right\rfloor = p + \frac{c_{n-1} - p}{2} = \frac{c_{n-1} + p}{2}. \tag{7}$$

On the other hand, as $p \mid \gcd(c_{n-1}, r_n)$, then

$$\frac{c_{n-1} + p}{2} \leq \frac{c_{n-1} + (c_{n-1}, r_n)}{2} = \frac{c_n + 1}{2}. \tag{8}$$

Combining (7) and (8) the induction step is finished.

If $\gcd(c_{n-1}, r_n)$ is not prime, then we have a strict inequality in (8) and $r_n \neq (c_n+1)/2$. We obtain the same conclusion if $p\lfloor r_{n-1}/p \rfloor \neq (c_{n-1}-p)/2$ using (7). Then, the properties in the statement are necessary conditions for the equality. It is easy to see that the converse is also true. □

Using Lemma 2.1, it is easy to check that (ii) implies (i). Also, trivially (iii) implies (i) and (ii).

Corollary 2.3 *If* (i) *holds and* $\gcd(c_{n_0-1}, r_{n_0}) > r_{n_0-1}$, *then* (iii) *is true.*

We can redefine m_0 with no reference to prime numbers thanks to the following result.

Proposition 2.4 *Given* $n > 1$, $r_n = c_{n-1}$ *if and only if* c_{n-1} *is prime.*

This proposition can be reformulated as the following corollary.

Corollary 2.5 *If* $r_n = c_{n-1}$ *for some* $n > 1$, *then* (i) *and* (ii) *hold true.*

Proof of Proposition 2.4: If c_{n-1} is prime, then Proposition 2.2 implies that $r_{n-1} < c_{n-1}$ and, according to Lemma 2.1, we conclude that r_n has to be c_{n-1}.

For the converse, assume $r_n = c_{n-1}$ and take $m = (c_{n-1} + \operatorname{lpf}(c_{n-1}))/2 = (r_n + \operatorname{lpf}(r_n))/2$. We have that $\gcd(m, c_{n-1}) \neq 1$ and, again by Proposition 2.2, $r_{n-1} < m$. The alternative definition of r_n given by Lemma 2.1 implies that $r_n \leq m$, or equivalently $r_n = \operatorname{lpf}(r_n)$. Hence, $r_n = c_{n-1}$ is prime. □

Proposition 2.6 *Under* (iii), *there exists a prime* p *such that*

$$\inf\{k \,:\, a_k = 3k\} = \frac{p+1}{2} \qquad \text{and} \qquad \inf\{k \,:\, a_k = 3k,\ a_k - a_{k-1} > 1\} = p.$$

Proof: Clearly $a_k = 3k$ is equivalent to $c_n = 2k - 1$. If $a_k - a_{k-1} > 1$, then Proposition 1.4 shows that $k = r_n$ for some n. As r_n is increasing, the minimum is reached in r_{n_0} which is prime by Proposition 2.4.

Without any assumption on $a_k - a_{k-1}$, by Proposition 1.4 for $r_{n_0-1} \le k < r_{n_0}$

$$a_k = 3k \ \Leftrightarrow\ k = \frac{c_{n_0-1}+1}{2} = \frac{p+1}{2}$$

and we know by Proposition 2.2 that actually this value lies on the interval $[r_{n_0-1}, r_{n_0})$.

It only remains to check that $a_k > 3k$ for all $k \le r_{n_0-1}$. Otherwise, if $a_{k-1} \le 3(k-1)$ for some k, then $3 \le 3k - a_{k-1}$. By Proposition 1.4, $a_k - a_{k-1}$ is equal to $\gcd(c_{n-1}, r_n)$ for $k = r_n$ and equals 1 otherwise, therefore it always divides $3k - a_{k-1}$, so $a_k - a_{k-1} \le 3k - a_{k-1}$, and then $a_k \le 3k$. Iterating this process would lead to a contradiction for $k = r_{n_0-1}$. $\square$

Extensive computations show that

$$Q_k = \min_{n<n_0} \frac{c_n+1}{r_n}$$

is by far greater than 2 when a_1 is large. For example when $2^{20} < a_1 < 2^{21}$, the minimum is 340.56. Any improvement of Proposition 2.2 in this direction reduces the equivalence between (i) and (iii) to a finite number of computations. We show an example here using our computer based verification of Conjecture 1.5 for $a_1 < 10^8$.

Proposition 2.7 *Assume* (i) *and* $\left(2 + \frac{1}{2500}\right) r_n < c_n + 1$ *for* $n < n_0$. *Then* (iii) *holds.*

Proof: By Proposition 2.2 we have $\gcd(c_{n_0-1}, r_{n_0}) = p$. Also, for some j and l,

$$\begin{aligned} r_{n_0-1} &= pj + l, & r_{n_0} &= p(j+1), \\ c_{n_0-1} &= p(2j+1), & c_{n_0} &= 2p(j+1) - 1. \end{aligned}$$

For the sake of brevity write $K = 3500$. If $j > K$, then

$$\frac{c_{n_0-1}+1}{r_{n_0-1}} \le \frac{p(2j+1)+1}{pj} = 2 + \frac{1}{j} + \frac{1}{3j} < 2 + \frac{1}{2500},$$

which does not match our assumption. So we can suppose $1 \le j \le K$, since $j = 0$ clearly implies (iii). We distinguish several cases.

If $p \le 4K - 3$ then $c_{n_0-1} < 10^8 - 2$ and it corresponds to some $a_1 < 10^8$ for which (iii) was checked with a computer.

The remaining case has $p > 4K - 3$. If $l < p - 2K$, there always exists $r_{n_0-1} < m < r_{n_0}$ being a multiple of $2j+1$, hence $\gcd(m, c_{n_0-1}) \ne 1$, and this contradicts Lemma 2.1. Then we have $l \ge p - 2K$ and

$$\frac{c_{n_0-1}+1}{r_{n_0-1}} \le \frac{p(2j+1)+1}{p(j+1) - 2K} = 2 + \frac{4K - p + 1}{p(j+1) - 2K}.$$

Comparing with the assumed inequality we should have

$$p(j+1) - 2K < 2500\big(4K - p + 1\big),$$

which is impossible for $j > 0$ and $p > 4K - 3$. $\square$

Proposition 2.8 *Given N there exists a_1 such that $m_0 > N$.*

Proof: Let a_1 be such that $m_0 < \infty$. Take $a_1' = a_1 + M$ with $M = c_{m_0}!$. We claim that the sequences (5) corresponding to a_1' are

$$r_j' = r_j \qquad \text{and} \qquad c_j' = c_j + M \qquad \text{for} \quad j \le m_0.$$

Clearly c_j is a nontrivial factor of c_j' then $m_0' > m_0$ and iterating this process N times we can obtain $a_1^{(N)}$ whose m_0 exceeds at least in N the m_0 corresponding to a_1.

To prove the claim note that $\gcd(k, c_j + M) = \gcd(k, c_j)$ for any $k \le r_{m_0}$ because in fact k divides M and appeal to Lemma 2.1. $\square$

3 Primes

Proposition 3.1 *Under* (i) *we have*

$$c_n = c_{n-1} + \operatorname{lpf}(c_{n-1}) - 1 \quad \textit{and} \quad r_n = (c_n + 1)/2,$$

for $n > n_0$.

Proof: We are going to prove that given $n \ge n_0$, if $r_n = (c_n + 1)/2$, then

$$c_{n+1} = c_n + \operatorname{lpf}(c_n) - 1 \quad \text{and} \quad r_{n+1} = (c_{n+1} + 1)/2. \tag{9}$$

As (i) assures $r_{n_0} = (c_{n_0} + 1)/2$, the result follows by an inductive argument.

By Lemma 2.1,

$$r_{n+1} = \min\{l \ge 1 \ : \ \gcd(r_n + l, c_n) \ne 1\},$$

where $\gcd(r_n + l, c_n) = \gcd((c_n + 1 + 2l)/2, c_n) = \gcd(1 + 2l, c_n)$, as c_n is odd. Hence, $1 + 2l = \operatorname{lpf}(c_n)$. Then $r_{n+1} = r_n + (\operatorname{lpf}(c_n) - 1)/2$ and

$$c_{n+1} = c_n + \gcd\left(c_n, \frac{c_n + 1}{2} + \frac{\operatorname{lpf}(c_n) - 1}{2}\right) - 1 = c_n + \operatorname{lpf}(c_n) - 1,$$

and we obtain (9). $\square$

Proposition 3.2 *Under* (i), (ii) *or* (iii) *Conjecture 1.2 is true. Moreover, $\{a_k - a_{k-1}\}_{k=1}^{\infty}$ contains infinitely many distinct primes.*

Proof: We can always suppose (i) is true because it is in principle less general. By Propositions 1.4 and 3.1, for $k = r_n$ with $n > n_0$, we have

$$\begin{aligned} a_k - a_{k-1} = \gcd(c_{n-1}, r_n) &= \gcd\left(c_n + 1 - \operatorname{lpf}(c_{n-1}), \frac{c_n+1}{2}\right) \\ &= \gcd(\operatorname{lpf}(c_{n-1}), c_n + 1) = \operatorname{lpf}(c_{n-1}). \end{aligned}$$

It remains to be proved that the set $\{a_k - a_{k-1}\}_{k=1}^{\infty}$ contains infinitely many primes. Let P be the product of the primes being smaller than N, with N such that $P > c_{n_0}$. Let n be the only integer satisfying $c_n < P \le c_{n+1}$. If we put $c_n = pq$ with $p = \operatorname{lpf}(c_n)$ and use Proposition 3.1, then $pq < P \le pq + p - 1$, and so $0 < P - pq < p$. Now, as $P - pq$ cannot be a multiple of p, we deduce that p has to be greater than N.

Therefore, given N we have found n such that $a_{r_n+1} - a_{r_n} = \operatorname{lpf}(c_n) = p > N$. Letting N tend to infinity, we obtain an unbounded sequence of primes and the result follows. □

Not all possible sequences of primes do actually appear. For instance it is obvious that (4) and (3) prevent from getting the same prime twice as consecutive values of $a_k - a_{k-1} \ne 1$. It motivates the following definition.

Definition: We say that a finite sequence of k odd primes $C_k = \{p_1, p_2, \dots, p_k\}$ is a *Rowland's chain* if there exists $c_1^* > 1$ such that $p_n = \operatorname{lpf}(c_n^*)$ for $1 \le n \le k$, where $c_n^* = c_{n-1}^* + \operatorname{lpf}(c_{n-1}^*) - 1$. We associate to C_k the shifted partial sums

$$S(n) = \sum_{j<n} (p_j - 1) \qquad \text{and} \qquad S(1) = 0.$$

The following is a characterization of Rowland's chains.

Proposition 3.3 *A finite sequence of odd primes $C_k = \{p_1, p_2, \dots, p_k\}$ is a Rowland's chain if and only if the following three conditions are satisfied:*

a) $S(m) \equiv S(n) \pmod{p_n}$ *when* $p_n = p_m$.

b) $S(m) \not\equiv S(n) \pmod{p_n}$ *when* $p_n < p_m$.

c) *For any prime q the set $\{S(j) \pmod{q} \ : \ p_j > q\}$ does not contain all residue classes modulo q.*

Of course in the third condition the set is empty except for q less than the maximum of C_k and it is also trivially satisfied if $q > k$, so this characterization allows one to verify whether C_k is a Rowland's chain in a finite number of steps. For instance $\{3, 19, 5, 3\}$ is a Rowland's chain because $S(1) = 0$, $S(4) = 24$ imply a). The rest of the values, $S(2) = 2$, $S(3) = 20$ imply that neither $S(1)$ nor $S(4)$ are congruent to $S(2)$ or $S(3)$ (mod 3), and $S(2) \not\equiv S(3)$ (mod 5), which is b). Finally c) does not need a verification because (excluding the trivial case $q = 2$) if the set is nonempty $q \ge 5$ and we only have

4 residue classes. On the other hand $\{17, 5, p\}$ is not a Rowland's chain for any $p > 3$ because it violates c) for $q = 3$.

Proof: Note that, according to the definition of Rowland's chain, $c_n^* = c_1^* + S(n)$ and C_k is a Rowland's chain if and only if there exists c_1^* satisfying for $1 \le n \le k$

$$c_1^* + S(n) \equiv 0 \pmod{p_n} \quad \text{and} \quad c_1^* + S(n) \not\equiv 0 \pmod{q} \quad \text{for every } q < p_n. \tag{10}$$

If $p_n = p_m$ then $c_1^* + S(n) \equiv c_1^* + S(m) \equiv 0 \pmod{p_n}$ implies a). On the other hand, the Chinese remainder theorem assures that under these conditions there exists a solution to the system formed by the first set of equations of (10).

Let q be any prime less that the maximum of C_k. Then the equations in (10) involving q are

$$c_1^* + S(m) \not\equiv 0 \pmod{q} \quad \text{for } m \in \{j \ : \ p_j > q\},$$

and if $q \in C_k$, say $q = p_n$, we have to add also

$$c_1^* + S(n) \equiv 0 \pmod{q}.$$

In the first case there exists a solution $\pmod{q}$ if and only if $S(m)$ does not cover all residue classes. This is c). In the second case we also need $S(m) \not\equiv S(n) \pmod{p_n}$ and this is b).

Finally note that once we have checked that the repeated equations are coherent, the Chinese remainder theorem can be used to find an arithmetic progression of possibilities for c_1^*. □

We know, thanks to the second part of Proposition 3.2, that the sequence of primes cannot be periodic. But the situation is even more restrictive; it cannot repeat blocks.

Corollary 3.4 *If $p_1, \dots, p_k$ are distinct primes, then $C_{2k} = \{p_1, p_2, \dots, p_k, p_1, p_2, \dots, p_k\}$ is not a Rowland's chain.*

Proof: Note that $\lambda = S(n+k) - S(n)$ is constant for $1 \le n \le k$. Then Proposition 3.3 a) implies that this value is divisible by every p_n, and hence λ is a multiple of $p_1 p_2 \dots p_k$. But this is impossible, since this product is greater than λ. □

In the most of the cases Proposition 3.3 imposes severe restrictions to construct Rowland's chains with few given distinct primes and large k. But on the other hand it is possible to find rather long chains for special choices of the primes. For instance, using the first five odd primes there are no chains of length greater than 10 but we have

$$C_{27} = \{3, 5, 3, 23, 3, 5, 3, 653, 3, 5, 3, 23, 3, 5, 3, 3603833, 3, 5, 3, 23, 3, 5, 3, 653, 3, 5, 3\}$$

of length 27 that only involves the primes $3, 5, 23, 653$ and 3603833. In fact it is maximal for this set of primes (there is another valid maximal chain of the same length). It corresponds to $c_1^* = 1550303031682203$.

Remark: The proof of Corollary 3.4 gives $\mathrm{lcm}(p_1, p_2, \ldots, p_k) \mid \sum_{j=1}^{k}(p_j - 1)$, even admitting repeated primes. We expect that also in this case consecutive identical blocks cannot appear because Proposition 3.3 would impose too strong conditions. On the other hand, we think that it is possible to concatenate arbitrarily long identical blocks inserting one prime between them (see the chain C_{27}).

Acknowledgments: We are indebted to the referee for the careful reading, the help to improve the exposition and the intriguing questions. We thank Angélica Benito for the reading of the manuscript and her suggestions. The first author spent his 2009-10 sabbatical at Rutgers and would like to express his deepest gratitude to Professor Zeilberger for his kindness and the stimulating talks.

References

[1] B. Cloitre. 10 conjectures in additive number theory. ArXiv:1101.4274, 2011.

[2] T. Gowers, J. Barrow-Green, and I. Leader, editors. *The Princeton companion to mathematics*. Princeton University Press, Princeton, NJ, 2008.

[3] A. Odlyzko, M. Rubinstein, and M. Wolf. Jumping champions. *Experiment. Math.*, 8(2):107–118, 1999.

[4] E. S. Rowland. A natural prime-generating recurrence. *J. Integer Seq.*, 11(2):Article 08.2.8, 13, 2008.

[5] A. Schinzel and W. Sierpiński. Sur certaines hypothèses concernant les nombres premiers. *Acta Arith. 4 (1958), 185–208; erratum*, 5:259, 1958.