A Cogent Argument that Supports the Conjecture of Keane in Kolakoski Sequence A000002

Abdallah Hammam*
Département de Mathématiques, Faculté des sciences, Université Moulay Ismaïl, 50020 Meknès, Morocco
*Corresponding author: a.hammam@fs.umi.ac.ma

Received February 01, 2019; Revised March 09, 2019; Accepted March 16, 2019

Abstract The aim of our investigation is an attempt to answer two still unsolved questions about Kolakoski sequence \((K_n)_{n \geq 1}\): Is there an explicit expression of the \(n\)th term \(K_n\), and the second one, known as the conjecture of Keane, claims that the asymptotic density of twos, is \(\frac{1}{2}\). In the first section of this paper, we present a new formula for \(K_n\) according to \(K_1, K_2, \ldots, K_p\) where \(p \approx \frac{4}{9}n\). In the second part, we define three sequences satisfying the condition \(U_i V_i = W_i\), and using the fact that \(V_i\) increases at least exponentially while \(W_i\) does not, we conclude that \(U_i\) should converge to zero. Our argument is inductive but so strong to insure the validity of the conjecture in concern with density of twos.

Keywords: Kolakoski sequence, recursive formula, asymptotic density.

2000 Mathematics Subject Classification: Primary 11B83; Secondary 11B85, 11Y55, 40A05.


1. Introduction

The infinite Kolakoski word \(K = K_1 K_2 K_3 \ldots = 12211212211221122121\ldots\) as defined in Sloane’s OEIS [1], is the unique fixed point, starting by \$1\$, \$S$ of the Run Length Encoding operator \(\Delta:\)

\[(\forall n \geq 1) \ K_n \in \{1, 2\} \text{ and } \Delta(K) = K\]

Many questions about this self-generating sequence are still with no answer [2].

The two most known are:

1. Is there an explicit expression of \(K_n\) with respect to \(n\)?
2. Is the asymptotic density of twos, \(\frac{1}{2}\)

Bordellès [3] gave expressions of \(K(K_1+K_2+\ldots+K_p))\) and \(K(K_1+K_2+\ldots+K_p)+1\).

To complete this, we add a new expression of \(K(K_1+K_2+\ldots+K_p)-1\).

Steinsky [4] found a complicated formula which allows to compute \(K_n\) from the former terms \(K_1, K_2, \ldots, K_{n-1}\).

We improve this by using only \(K_1, K_2, \ldots, K_p\) with \(p \approx \frac{4}{9}n\).

Concerning the asymptotic density question, Steinsky [4] presented a curve which disapprove the conjecture, but, in our work, we use the fact that the exponential grows very fast to highly support it.

2. Notation

The successive partial sums

For \(n \geq 1\) and \(i \geq 1\), we define the following very useful partial sums

\[S_{0,n} = n\]

\[S_{1,n} = S_n = \sum_{j=1}^{n} K_j\]

and

\[S_{i+1,n} = S_{S_i,n} \text{.}\]

For example,

\[S_{0,10} = 10, \quad S_{1,10} = 15, \quad S_{2,2} = 7.\]

The density \(\rho_n\) of twos in \(K_1 K_2 \ldots K_n\) and the discrepancy \(\delta(n)\)
For \( n \geq 1 \), by definition, \( n \rho_n \) and \( \delta_n \) are respectively the total number of twos and the difference between the twos and the ones, in the word \( K_1K_2\ldots K_n \).

For instance,

\[
\begin{align*}
123 & \quad \rho = 3 \\
12345 & \quad \delta = 5 \\
2122 & \quad \rho = 3, \delta = 1 \\
12211 & \quad \rho = 1, \delta = 5
\end{align*}
\]

We will also need the classical identities

\[
\delta(n) = (2 \rho_n - 1)n = \sum_{j=1}^{n} (-1)^{j} K_j \\
\delta(S_n) = (2 \rho_{S_{n-1}} - 1)S_n = \sum_{j=1}^{n} (-1)^{j} K_j.
\]

Remark 1. It is well known that

\[
(\forall n \geq 1) \quad \frac{4}{9} \leq \rho_n \leq \frac{5}{9}.
\]

3. A First Expression of \( K_n \)

**Lemma 2.** For each integer \( n \geq 1 \), there exists an integer \( p \leq \frac{9(n+1)}{13} \) such that

\[
(n = S_p) \lor (n = S_{p-1})
\]

**Proof 3.** Let \( n \) be a positive integer. It is clear that

\[
(\forall p \in \mathbb{N}^{+}) \quad S_{p+1} - S_p = K_{p+1} \\
\Rightarrow (\forall p \in \mathbb{N}^{+}) \quad 1 \leq S_{p+1} - S_p \leq 2
\]

So, for any integer \( n \), there are two cases:

\[
(\exists p \in \mathbb{N}) \quad S_p = n.
\]

Or

\[
(\exists p \in \mathbb{N}) \quad S_p < n < S_{p+1} \leq S_p + 2 \\
\Rightarrow n - S_p < 2 \Rightarrow n = S_p + 1.
\]

On the other hand, \((1 + \frac{4}{9}) p \leq S_p \leq (1 + \frac{5}{9}) p\) and

\[
n \geq S_{p-1} \Rightarrow p \leq \frac{9(n+1)}{13}.
\]

**Lemma 4.** For each integer \( p \geq 2 \),

\[
K_{S_p - 1} = \frac{3 - (1)^{p+K_p}}{2}.
\]

**Proof 5.** We just need results in Table 1.

<table>
<thead>
<tr>
<th>( K_{S_p - 1} )</th>
<th>( p ) even and ( K_p = 1 )</th>
<th>( p ) even and ( K_p = 2 )</th>
<th>( p ) even and ( K_p = 3 )</th>
<th>( p ) even and ( K_p = 4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K_{S_p - 1} )</td>
<td>12</td>
<td>22</td>
<td>21</td>
<td>11</td>
</tr>
<tr>
<td>( K_{S_p} )</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Using lemmas above, one can deduce an expression of \( K_n \) in both cases.

**Corollary 6.** For \( n \geq 3 \), let

\[
A = \sum_{j=1}^{\frac{9(n+1)}{13}} j 0^{n-S_j},
\]

and

\[
B = \sum_{j=1}^{\frac{9(n+1)}{13}} j 0^{n-S_{j+1}}.
\]

Then

\[
K_n = \frac{3 + (-1)^{p+K_p}}{2}.
\]

In the next section, we give an improved expression of \( K_n \) according to \( K_1, K_2,\ldots, K_{\frac{9}{4}n} \).

4. A Second Expression of \( K_n \)

**Lemma 7.** For each integer \( n \geq 1 \), there exists an integer \( p \) such that

\[
(n = S_{2p}) \lor (n = S_{2p-1}) \lor (n = S_{2p-2}) \lor (n = S_{2p-3})
\]

**Proof 8.** It is a simple consequence of the fact that

\[
(\forall i \geq 2) \quad S_{2i} - S_{2i-1} = K_{\frac{3 - (1)^{p+K_p}}{2}} \in \{1, 2, 3, 4\}
\]

**Corollary 9.** For \( n \geq 3 \), let

\[
A = \sum_{j=1}^{\frac{8(n+3)}{269}} 0^{n-S_{2j}}, A_1 = \sum_{j=1}^{\frac{8(n+3)}{269}} j 0^{n-S_{2j}},
\]

\[
B = \sum_{j=1}^{\frac{8(n+3)}{269}} 0^{n-S_{2j+1}}, B_1 = \sum_{j=1}^{\frac{8(n+3)}{269}} j 0^{n-S_{2j+1}},
\]

\[
C = \sum_{j=1}^{\frac{8(n+3)}{269}} 0^{n-S_{2j+2}}, C_1 = \sum_{j=1}^{\frac{8(n+3)}{269}} j 0^{n-S_{2j+2}},
\]

and

\[
E = \sum_{j=1}^{\frac{8(n+3)}{269}} 0^{n-S_{2j+3}}, E_1 = \sum_{j=1}^{\frac{8(n+3)}{269}} j 0^{n-S_{2j+3}},
\]

then

\[
K_n = \frac{3 + (-1)^{p} \pi p}{2} \left( 2A_1 - (1 + (-1)^{p})B_1 \right).
\]
Proof 10.
It is easy to check that
If \( A = 1 \), then \( n = S_{2,p} \) and \( K_n = \frac{3 + (-1)^{S_p}}{2} \).
If \( A = 0, B = 1 \) then \( n = S_{2,p} - 1 \) and
\[
K_n = \frac{3 + (-1)^{S_p + p}}{2}.
\]
If \( A = B = 0 \) then \( K_n = \frac{3 - (-1)^{S_p}}{2} \).
We can replace the three cases be the next unique formula and using the fact that \( A + 0^A[B + 0^B(C + 0^C E)] = 1 \), we get the desired expression for \( K_n \).
\[
K_n = \frac{3 + (-1)^{S_p} + 0^A B}{2} \left[ \frac{3 + (-1)^{S_p + p}}{2} A + 0^A \frac{3 + (-1)^{S_p}}{2} B + 0^B \frac{3 - (-1)^{S_p}}{2} (C + 0^C E) \right].
\]
This formula has been validated by the code below, written in Maple language.

```maple
N := 100; K1 := 1; K2 := 2; p := 2:
for n from 3 to N do
    A := 0; B := 0; C := 0; E := 0;
    A1 := 0; B1 := 0; C1 := 0; E1 := 0; S1 := 1; S2 := 1:
    for j from 2 to \left( \frac{81(n+3)}{269} \right) do
        S1 := S1 + K1;
        S2 := S2 + K1 * (3 + (-1)^j) / 2:
        A := A + j * \left( 0^{n-S21} \right) / 2; A1 := A + j * \left( 0^{n-S21} \right);
        B := B + j * \left( 0^{n-S21} \right) / 2; B1 := B + j * \left( 0^{n-S21} \right);
        C := C + j * \left( 0^{n-S21} \right) / 2; C1 := C + j * \left( 0^{n-S21} \right);
        E := E + j * \left( 0^{n-S21} \right) / 2; E1 := E + j * \left( 0^{n-S21} \right);
    end do:
    p := A + 0^A * (B1 + 0^B * (C1 + 0^C * E1));
    S_p := add(K_m, m = 1..p):
    p1 := (3 + (-1)^S_p) / 2; p2 := (3 + (-1)^{p+S_p}) / 2:
    p3 := (3 - (-1)^S_p) / 2:
    K_n := A1 * p1 + 0^A * \left( B1 * p2 + 0^B * p3 * (C1 + 0^C * E1) \right):
end do:
print(seq(K_j, j = 1..N))
```

5. A Great Support for the Conjecture of Keane

**Proposition 11.** For an even natural number \( n \geq 2 \) satisfying \( \sum_{j=1}^{n} (-1)^j = 0 \),
\[
\delta(n) = (2\rho_n - 1)n
\]
\[
= \left[ \sum_{j=1}^{n} (-1)^j K_j \frac{1 + (-1)^j}{2} + \sum_{j=1}^{n} (-1)^j \frac{K_j - (-1)^j}{2} \right]
\]
\[
= \frac{1}{2} \left[ \sum_{j=1}^{n} (-1)^j K_j \frac{1 + (-1)^j}{2} - \sum_{j=1}^{n} (-1)^j \frac{K_j - (-1)^j}{2} \right]
\]

**Proof 12.** By definition of the density \( \rho_n \),
\[
(2\rho_n - 1)n = \sum_{j=1}^{n} (-1)^j K_j
\]
\[
= \left[ \sum_{j=1}^{n} (-1)^j K_j \frac{1 + (-1)^j}{2} + \sum_{j=1}^{n} (-1)^j \frac{K_j - (-1)^j}{2} \right]
\]
and
\[
(2\rho_S_n - 1)S_n = \sum_{j=1}^{n} (-1)^j K_j
\]
\[
= \frac{1}{2} \left( \sum_{j=1}^{n} (-1)^j (3 + (-1)^{K_j}) \right) = \frac{1}{2} \left( \sum_{j=1}^{n} (-1)^{j+K_j} \right)
\]
\[
= \frac{1}{2} \left( \sum_{j=1}^{n} (-1)^j K_j \frac{1 + (-1)^j}{2} - \sum_{j=1}^{n} (-1)^j \frac{K_j - (-1)^j}{2} \right)
\]

**Corollary 13.** From the proposition 11 just above, we see that the left hand sides contain a sequence which grows exponentially from \( n \) to \( S_n \approx \frac{3}{2} n \) while, in the right hand side, there is probable change of sign indicating no exponential increasing as illustrated in Figure 1.

Figure 1. An exponential increasing versus a sign changing
6. Concluding Remarks

We presented an optimal expression of the form $K_n = f(K_1, K_2, ..., K_p)$ with $p \approx \frac{4}{9} n$, improving so some former results. About the asymptotic density of twos, We have a strong reason to support Keane's conjecture: Our argument is based on Proposition 12. It uses the fact that if the product $U_i V_j$ of two sequences have a changing sign and if $V_j$ increases exponentially to $+\infty$, then $(U_i)$ should converge to zero. This reasoning has been applied to other Kolakoski sequences: Kol(1,3), Kol(1,5), and Kol(1,7) and we obtained the results predicted by Hammam [5] as illustrated in Figure 2. The blue curve shows clearly that the density of twos, in Kol(1,2) goes to $\frac{1}{2}$.

7. Acknowledgments

The author would like to express his gratitude to the referees for their careful reading of the manuscript and for giving me some of their precious time.

References

[4] B. Steinsky, A recursive formula for the Kolakoski sequence, J. Integer Seq. 9 (2006), Article 06.3.7.