Turkish Journal of Analysis and Number Theory, 2017, Vol. 5, No. 4, 139-142
Available online at http://pubs.sciepub.com/tjant/5/4/4

DOI:10.12691/tjant-5-4-4

Some Formulas for the Generalized Kolakoski Sequence *Kol*(*a*, *b*)

Abdallah Hammam*

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

Received April 07, 2017; Revised May 11, 2017; Accepted June 14, 2017

Abstract We present here a new approach to investigate the Kolakoski sequence *Kol*(*a*, *b*). In the first part of this paper, we give some general identities. In the second, we state our main result, which concerns the frequency of the letters in the case where *a* and *b* are odd. Finally, we give an algorithm to compute the term K_n in the particular case of *Kol*(1, 3).

Keywords: *Kolakoski sequence, recursion, recursive formula*

Cite This Article: Abdallah Hammam, "Some Formulas for the Generalized Kolakoski Sequence *Kol*(*a*, *b*)." *Turkish Journal of Analysis and Number Theory*, vol. 5, no. 4 (2017): 139-142. doi: 10.12691/tjant-5-4-4.

1. Introduction

The Kolakoski sequence $Kol(a,b)$ or $(K_n)_{n\geq 1}$ comes from the generalization of the well-known Oldenburger-Kolakoski sequence $Kol(1,2)$ [4,5,8]. It starts with $K_1 = a,$ contains only letters from the alphabet set $\{a,b\}$ and it doesn't change by the run length encoding operator Δ.

The case where *a* and *b* are odd, has been explored by Baake et al [1] who found a connection between the generalized Kolakoski sequence and some deformed model sets. They used Perron-fronebius Theorem and found that the frequency of '3' in $Kol(1,3)$ is " ≈ 0.60 ".

Brlek et al [2] studied smooth words on 2-letter alphabet having the same parity and proved that the frequency of both letters is $\frac{1}{2}$ when *a* and *b* are even. They also presented an expression of the asymptotic density of the letter *b* when *a* = 1 and *b* is odd of the form $\frac{1}{\sqrt{2b-1}-1}$. The case when *a*+*b* is odd is more difficult. Shen [6] has investigated this case but, did focus his work on some "expansion" functions rather than on the frequency of letters. Until now, no expression of the limiting density of the letters is available.

2. Notation

We first introduce some definitions and notation:

Let $\Sigma = \{a,b\}$ where *a* and *b* are positive integers: $a < b,$ be the input alphabet set and let Σ^* be the set of all finite words over Σ. We let Δ denote the run-length encoding operator, such that if W is an element of Σ^*, then $\Delta(W)$ will be the new word containing lengths of blocs of similar digits, in the initial word W. For instance, $\Delta(333355) = 42.$ We now define the primitive operator Δ_a^{-1} in the same way than Hammam [3], by:

If $W = \omega_1\omega_2...\omega_N$ is an element of Σ^*, then $\Delta_a^{-1}(W)$ will be the unique word W' of Σ^* starting from the left by $a,$ and satisfying the encoding condition: $\Delta(W') = W.$ As a consequence, if we let denote W_N for every $N \geq 1,$ the word $K_1K_2...K_N$ with $K_1 = a,$ then $\Delta_a^{-1}(W_N) = W_{S_N}.$ Here, S_N is the partial sum defined by:

$$S_N = \sum_{j=1}^{N} K_j.$$

Therefore, the infinite Oldenburger-Kolakoski word $K = K_1K_2K_3...$ can be seen as a fixed point of the function $\Delta_a^{-1},$ $\left(K = \Delta(K) = \Delta_{K_1}^{-1}(K)\right).$

We also need to define the double sequence $(S(i,N))_{(i,N)\in\mathbb{N}\times\mathbb{N}^*}$ by :

$$\left(\forall N \in \mathbb{N}^*\right) S_{1,N} = S_N = \sum_{j=1}^{j} K_j$$

and

$$\left(\forall i \in \mathbb{N}^*\right) S_{i+1,N} = S_{S_{i,N}}.$$

In the end, we define the density of the letter '*b*' in the word W_N by

$$\rho_N = \frac{\left|\{1 \leq j \leq N : K_j = b\}\right|}{N}.$$

We will try to find $\lim_{N\to\infty} \rho_N$.

3. Some Identities for the General Kolakoski Sequence *Kol*(*a*, *b*)

Lemma 1. *For every* $N \geq 1$*:*

$$\left|a^1\right|_N + \left|b^1\right|_N = \left|a^2\right|_N + \left|b^2\right|_N + \frac{1-(-1)^N}{2} \qquad (1)$$

where

$$\left|b^2\right|_N = \left|\left\{1 \leq j \leq N : K_j = b \text{ and } j \text{ even}\right\}\right|$$
$$\left|b^1\right|_N = \left|\left\{1 \leq j \leq N : K_j = b \text{ and } j \text{ odd}\right\}\right|$$
$$\left|a^2\right|_N = \left|\left\{1 \leq j \leq N : K_j = a \text{ and } j \text{ even}\right\}\right|$$
$$\left|a^1\right|_N = \left|\left\{1 \leq j \leq N : K_j = a \text{ and } j \text{ odd}\right\}\right|$$

Proof. This equality simply means that if *N* is even, then in $K_1K_2...K_N$, there are as many odd indices as even ones, and if *N* is odd, the number of odd indices is greater by 1. As a consequence, we have the following useful equality

$$\left|a^2\right|_N + \left|b^2\right|_N = \frac{1}{2}\left(N + \frac{(-1)^N - 1}{2}\right). \qquad (2)$$

Lemma 2. *For every* $N \geq 1$*:*

$$\left[2\rho_{S_N} - 1\right]S_N = \sum_{j=1}^{N} (-1)^j K_j.$$

Proof. For each $N \geq 1$, we have by definition,

$$S_N = \left(a + (b-a)\rho_N\right)N. \qquad (3)$$

When we integrate the word $W_N = K_1K_2K_3...K_N$, the odd indices will become *'aaa...'* and the even will be transformed to *'bbb...'*. So

$$S_{2,N} = \sum_{j=1}^{N} K_j\left(a\frac{1-(-1)^j}{2} + b\frac{1+(-1)^j}{2}\right)$$
$$= \frac{a+b}{2}S_N + \frac{b-a}{2}\sum_{j=1}^{N}(-1)^j K_j$$

but, on the other hand, we have

$$S_{2,N} = \left(a + C\rho_{S_N}\right)S_N$$

with $C = b - a$. Thus, after simplification, we find

$$\left[2\rho_{S_N} - 1\right]S_N = \sum_{j=1}^{N} (-1)^j K_j. \qquad (4)$$

Lemma 3. *If we put C = b − a, then for every* $N \geq 1$:

$$\left[2\rho_{S_{2,N}} - 1\right]S_{2,N} = C\left(\left|b^2\right|_{S_N} - \left|b^1\right|_{S_N}\right) + a\frac{(-1)^N - 1}{2}. \qquad (5)$$

Proof. Written using the cardinals defined above, equation (4) becomes

$$\left[2\rho_{S_N} - 1\right]S_N = b\left(\left|b^2\right|_N - \left|b^1\right|_N\right) + a\left(\left|a^2\right|_N - \left|a^1\right|_N\right)$$
$$= C\left(\left|b^2\right|_N - \left|b^1\right|_N\right)$$
$$+ a\left(\left(\left|b^2\right|_N - \left|b^1\right|_N\right) + \left(\left|a^2\right|_N - \left|a^1\right|_N\right)\right)$$

and by equation (1), we get

$$\left[2\rho_{S_N} - 1\right]S_N = C\left(\left|b^2\right|_N - \left|b^1\right|_N\right) + a\frac{(-1)^N - 1}{2}.$$

Finally, equation (5) is simply obtained by replacing N by S_N and using the fact that $(-1)^N = (-1)^{S_N}$.

4. The Density of *b*'s in *Kol*(*a*, *b*)

When we integrate the finite word $W_N = K_1K_2...K_N$, we get the word $W_{S_N} = \Delta_{K_1}^{-1}\left(W_N\right) = K_1K_2...K_{S_N}$ and the letters change with respect to a certain transformation which depends on the parity of *a* and *b*.

4.1. The Density of *b*'s in *Kol*(2*m*+1, 2*m*+2*n*+1)

In this case, the integration operation is defined by the following matrix

$$\begin{bmatrix} \left|b^2\right|_{S_N} \\ \left|b^1\right|_{S_N} \\ \left|a^2\right|_{S_N} \\ \left|a^1\right|_{S_N} \end{bmatrix} = \begin{bmatrix} m+n+1 & 0 & m+1 & 0 \\ m+n & 0 & m & 0 \\ 0 & m+n & 0 & m \\ 0 & m+n+1 & 0 & m+1 \end{bmatrix}\begin{bmatrix} \left|b^2\right|_N \\ \left|b^1\right|_N \\ \left|a^2\right|_N \\ \left|a^1\right|_N \end{bmatrix}.$$

This matrix comes from the fact that S_N and N have the same parity.

Proposition 4. *If* $\lim_{N\to\infty} \rho_N$ *exists* $(= L)$*, then L is a root of the following third order equation*

$$(2x-1)(Cx+a)^2 = \frac{C}{2} \qquad (6)$$

where C = b − a.

Proof. Using lemma 3, we have

$$\left(2\rho_{S_{2,N}} - 1\right)S_{2,N} = C\left(\left|b^2\right|_{S_N} - \left|b^1\right|_{S_N}\right) + a\frac{(-1)^{S_N} - 1}{2}.$$

Using the matrix integration and the fact that S_N and N have the same parity, we find that

$$\left(2\rho_{S_{2,N}} - 1\right)S_{2,N} = C\left(\left|b^2\right|_N + \left|a^2\right|_N\right) + a\frac{(-1)^N - 1}{2}$$

and finally, by equation (2),

$$\left(2\rho_{S_{2,N}}-1\right)S_{2,N}=\frac{C}{2}N+\frac{a+b}{2}\frac{(-1)^N-1}{2}.$$

And if we use equation (3), we obtain

$$\begin{aligned}&\left(2\rho_{S_{2,N}}-1\right)\left[C.\rho_{S_N}+a\right]\left[C.\rho_N+a\right]\\&=\frac{C}{2}+\frac{a+b}{2}\frac{(-1)^N-1}{2N}.\end{aligned}\qquad(7)$$

As a consequence, if we assume that $\lim_{N\to\infty}\rho_N exists(=L)$, then $\lim_{N\to\infty}\rho_{S_{2,N}}=\lim_{N\to\infty}\rho_{S_N}=L$ and when $N\to\infty$, equation (7) gives

$$(2L-1)[C.L+a]^2=\frac{b-a}{2}.$$

If we put $y = Cx + a$, equation (6) will take the form

$$\left(y-\frac{a+b}{2}\right)y^2=\frac{(b-a)^2}{4}.$$

And a graphical study shows that this equation has always a real solution $y>\frac{a+b}{2}$ corresponding to a density $L>\frac{1}{2}$ given by

$$L=\left(\frac{D}{6}+\frac{2*A^2}{3D}+\frac{A}{3}-a\right).\frac{1}{C}$$

with $A=\frac{a+b}{2}, B=\frac{C^2}{4}$ and

$$D=\left(108*B+8*A^3+12*\sqrt{81*B^2+12*B*A^3}\right)^{1/3}.$$

For instance, with a = 1 and b = 3, we find the same value of Baake et al [1] ($L\approx 0.60278$).

4.2. The Density of *b*'s in *Kol*(2*m*, 2*m* + 2*n*)

In this case, the integration operation is defined by the following matrix

$$\begin{bmatrix}\left|b^2\right|_{S_N}\\\left|b^1\right|_{S_N}\\\left|a^2\right|_{S_N}\\\left|a^1\right|_{S_N}\end{bmatrix}=\begin{bmatrix}m+n&0&m&0\\m+n&0&m&0\\0&m+n&0&m\\0&m+n&0&m\end{bmatrix}\begin{bmatrix}\left|b^2\right|_N\\\left|b^1\right|_N\\\left|a^2\right|_N\\\left|a^1\right|_N\end{bmatrix}.$$

Using lemma 3 with S_N instead of N, we obtain

$$\left[2\rho_{S_{2,N}}-1\right]S_{2,N}=0$$

since S_N is always even.

This proves that if $\lim_{N\to\infty}\rho_N$ exists, it should be $\frac{1}{2}$.

4.3. The Density of *b*'s in *Kol*(2*m*+1, 2*m*+2*n*)

In this case, the integration operation is defined by the following matrix

$$\begin{bmatrix}\left|b^2\right|_{S_N}\\\left|b^1\right|_{S_N}\\\left|a^2\right|_{S_N}\\\left|a^1\right|_{S_N}\end{bmatrix}$$

$$=\begin{bmatrix}m+n&0&m&m+1&0&0\\m+n&0&m+1&m&0&0\\0&m+n&0&0&m+1&m\\0&m+n&0&0&m&m+1\end{bmatrix}\begin{bmatrix}\left|b^2\right|_N\\\left|b^1\right|_N\\\left|a^{21}\right|_N\\\left|a^{22}\right|_N\\\left|a^{12}\right|_N\\\left|a^{11}\right|_N\end{bmatrix}.$$

The structure of the matrix changes at each iteration. This shows the complexity of this case and explains the difficulty to find the density.

5. Expression the Term K_n for *a* = 1 and *b* = 3

Proposition 5. *Let* $N\ge 1$ *and* p *be the unique integer satisfying* $S(2,p)\le N<S(2,p+1)$ *then*

$$K_N=\begin{cases}2+(-1)^p & \text{if } N-S(2,p)=0\\2-(-1)^p & \text{if } N-S(2,p)=1\\1 & \text{if } 4\le N-S(2,p)\le 6\\3 & \text{otherwise}\end{cases}.$$

Table 1. How the 1's of W_n are transformed by Δ_1^{-1}

W_N	1^1	1^2
W_{S_N}	1^1	3^2
$W_{S_{2,N}}$	1^1	333^2
$W_{S_{3,N}}$	1	333111333

Table 2. How the 3's of W_n are transformed by Δ_1^{-1}

W_N	3^1	3^2
W_{S_N}	111^1	333^2
$W_{S_{2,N}}$	131^1	333111333^2
$W_{S_{3,N}}$	13331	333111333131333111333

Proof. For every $p > 0$, we have that $S_{p+1} - S_p = K_p \leq 3$ and in the worst case,

$$S_{2,p+1} = S_{S_{p+1}} = S_{3+S_p}$$

which gives

$$\begin{aligned} & S_{2,p+1} - S_{2,p} \\ & = S_{3+S_p} - S_{2+S_p} + S_{2+S_p} - S_{1+S_p} + S_{1+S_p} - S_{S_p} \\ & \leq 3+3+3=9. \end{aligned}$$

Using a disjunction of cases argument and Table 1 and Table 2, we derive the expression of term K_N.

As an example, for $N = 10$, we have $S_{2,2} = 10$ and $S_{2,3} = 13$, so $p = 2$ and $K_N = 2 - (-1)^p = 1$.

6. Concluding Remarks

The approach adopted here is, in a certain way, similar to the method used bay Sing [7] in the case where a and b are even. The difference is that we treat the terms of the sequence as numbers instead of string of characters. We also revealed the dependence of the sequence on the parity of the partial sum S_N.

We obtained our main result of proposition 4, by a simple recursive relation, and without any additionnal condition on *a* and *b*.

We understand now why the case where *a* and *b* have not the same parity is complex.

Our next investigation will focus on this case.

Acknowledgements

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

References

[1] M. Baake and B. Sing, Kolakoski (3,1) is a (deformed) model set, *Canad. Math. Bull.*47(2):168-90 (2004).

[2] S. Brlek, D. Jamet and G. Paquin, Smooth words on 2-letter alphabet having same parity, *Theoritical Computer Science, Elsevier*, 2008, 393 (1-3), pp 166-181.

[3] A. Hammam, Some new Formulas for the Kolakoski Sequence A000002. *Turkish Journal of Analysis and Number Theory* . 2016; 4(3):54-59.

[4] W. Kolakoski, Problem 5304: Self Generating Runs, *Amer. Math. Monthly* 72 (1965), 674.

[5] Rufus Oldenburger, Exponent trajectories in symbolic dynamics, *Trans Amer. Math. Soc.* 46 (1939), 453-466.

[6] B. Shen, A uniformness of the Kolakoski sequence, graph, connectivity and correlations (2017) *arXiv :1703.00180*.

[7] B. Sing, Kolakoski(2m,2n) are limit-periodic Model Sets, *arXiv* 0207037.

[8] N. J. A. Sloane, The On-line Encyclopedia of Integer Sequences, published electronically at http://oeis.org.