Journal of Discrete Mathematical Sciences & Cryptography
ISSN 0972-0529 (Print), ISSN 2169-0065 (Online)
Vol. 26 (2023), No. 1, pp. 139–152
DOI : 10.1080/09720529.2021.1923921

On the use of Egyptian fractions for stream ciphers

I. Cherkaoui *
F. Zinoun †
Department of Mathematics and Information Security
Laboratory of Mathematics, Computing and Applications
Faculty of Sciences
Mohammed V University in Rabat
P. O. Box 1014 RP, Rabat
Morocco

Abstract

Within the scope of the mutual interference between modern number theory and chaos-based cryptography, and as it is well-known and already explored for the decimal or the continued fraction expansion of irrational numbers, interesting random-like behaviors seem to be hidden in Egyptian fraction expansions, thus suggesting new chaos-based encryption systems. In fact, at a practical level, and as will be shown through the present study, concatenation of involved denominators and binary expansion generally lead to pseudo-random streams which satisfactorily pass the NIST statistical test suite for randomness. Then, to a certain extent, and for cryptographic purposes, this can be considered as a new tool to complete the conventional class of already-in-use pseudo-random number generators. Some mathematical issues, however, have to be clarified, as for example the chaoticity of the process obtained after converting these denominators to fractional parts of a wandering sequence of real numbers in the unit interval. To the best of our knowledge, and from a cryptographic point of view, previous works on the subject are limited to statistical tests (and subsequent cryptographic analysis) of the resulting one-time pads or stream ciphers, but no rigorous study has been conducted to justify these methods within chaos theory. Considering Egyptian expansions as a working example and using the term "chaotic" in the sense of Devaney, this is what our proposal is aimed at in the present work.

Subject Classification: *11A55, 37A25, 74H65 & 94A60.*

Keywords: *Egyptian fractions, Number theory, Chaos, Cryptography.*

E-mail:* **ilias.cherkaoui15@gmail.com (Corresponding Author)
†E-mail: **zinoun@fsr.ac.ma**

1. Introduction

A wide panoply of cryptosystems have been displayed throughout time and a large entity were based on number theoretic or algebraic concepts, especially notions from modular arithmetics. Chaos is another promising paradigm for cryptography offering a subtle behavior with special characteristics such as ergodicity or extreme sensitivity to initial conditions, thus meeting Shannon's (1948) fundamental criteria of confusion and diffusion. Since the discovery of chaotic synchronization by Pecora and Carroll (1990), and within the scope of dynamical systems, continuous or discrete, and their applications in real systems, man-made or natural, chaos-based cryptography has emerged as a powerful mathematical tool to secure information, from Baptista's (1998) original method to the latest modern-number-theory-based cryptosystems. For a large collection of papers on the subject, see for instance Kocarev and Lian (2011).

In the present work, and being strongly linked to cryptographic methods based on the decimal or the continued fraction expansion of irrational numbers (see for instance Kane [2013] and references therein, Mikram et al. (2012) for encryption with decimal expansion, Masmoudi et al. (2010) for creating a pseudo-random number generator via an Engel expansion, and more recently, in a larger context, Younis et al. (2020) for an interesting use of logic functions to raise the level of complexity and security of the proposed stream cipher), we explore the random-like behavior of a variant of the Egyptian expansion, studied as early as Fibonacci's Liber Abaci (1202) and known today under a general form as Engel expansion. This is the unique non-decreasing sequence of positive integers $(a_i)_{i\in\mathbb{N}^*}$ such that, for a positive real number x, we have $x = \frac{1}{a_1} + \frac{1}{a_1 a_2} + \frac{1}{a_1 a_2 a_3} + \ldots$

Every positive (ir)rational number has a unique (in)finite Engel expansion, coefficients a_i typically exhibiting exponential growth. For instance, it's easy to see that the Engel expansion of Neper's constant e asks for *all* positive integers, thus describing a dense set in the interval $[0.1, 1[$ if we consider the sequence $(a_i 10^{-\overline{a_i}})_{i\in\mathbb{N}^*}$, $\overline{a_i}$ denoting the number of digits of a_i, as it will be always the case for the rest of the present paper. Indeed, in order to obtain a continuum as a phase space, which is a necessary condition to deal with chaos theory, this is a way among others one can adopt to convey notions from the set of integers to the unit interval. As set in the abstract, for a well-defined function or, abusively, for a random-like (yet deterministic) process, the term "chaotic" will be used

in the sense of Devaney (1989), whose the definition deals with the notions of sensitive dependence on initial conditions, density of the set of periodic points and topological transitivity, in a sense to be clearly specified below.

The paper is globally structured as follows: we first begin with a brief presentation of the Engel expansion algorithm and a recall of some related properties. Subsequently, the corresponding cryptosystem is algorithmically presented and statistically analyzed. As can be seen, and additionally to the fact that it passes the NIST test, our cryptosystem is robust due to the difficulty of extracting the irrational number from a part of its Engel expansion. The underlying recurrence function is then highlighted to be studied from a chaotic point of view. According to Devaney's definition, it will be easily shown that such a function cannot be chaotic while, as expected, it will be the case for the associated random-like sequence described above. Finally, the investigation of another point of view to construct Engel-expansion-based random-like sequences, then efficient pseudo-random number generators, is suggested as a new idea to be developed in future work.

2. The Engel expansion algorithm and some properties

The Engel expansion Erdös et al. (1958) of a positive real number x is the unique non-decreasing sequence of positive integers $a_1, a_2, a_3, a_4, \ldots$ such that

$$x = \frac{1}{a_1} + \frac{1}{a_1 a_2} + \frac{1}{a_1 a_2 a_3} + \ldots$$

The algortihm allowing to retrieve the expansion of x is as follows: let $u_1 := x$ and for $k > 1$, the brackets denoting integer part,

$$a_k := \left[\frac{1}{u_k}\right] + 1 \text{ and } u_{k+1} := u_k a_k - 1.$$

If $u_k = 0$, the algorithm stops.

Example: Here are the steps to get the expansion of 1.175:

$$u_1 = 1.175, a_1 = \left[\frac{1}{1.175}\right] + 1 = 1;$$

$$u_2 = u_1 a_1 - 1 = 1.175 \cdot 1 - 1 = 0.175,$$

$$a_2 = \left[\frac{1}{0.175}\right] + 1 = 6;$$

$$u_3 = u_2 a_2 - 1 = 0.175 \cdot 6 - 1 = 0.05,$$

$$a_3 = \left[\frac{1}{0.05}\right] + 1 = 20;$$

$$u_4 = u_3 a_3 - 1 = 0.05 \cdot 20 - 1 = 0;$$

The algorithm stops. Therefore

$$1.175 = \frac{1}{1} + \frac{1}{1 \cdot 6} + \frac{1}{1 \cdot 6 \cdot 20}$$

and the Engel expansion of $1.175 = [1, 6, 20]$. This is the notation we will use for the Engel expansion in the rest of the paper.

Here are some remarkable properties of the Engel expansion for $x \in]0,1]$:

- Erdös et al. (1958) The sequence of the random variables a_n is a homogenous Markov chain, and its transition probability is given by

$$\begin{cases} P(a_n = k / a_{n-1} = j) = \dfrac{j-1}{k(k-1)} & \text{if} \quad 2 \le j \le k \\ P(a_n = k / a_{n-1} = j) = 0 & \text{if} \quad j > k \end{cases}$$

- Rényi (1962) Let $2 \le k_1 < k_2 < \ldots < k_n < \ldots$ be a strictly increasing sequence of non-negative integers. The sequence (a_n) of the Engel expansion contains for almost every x an infinity of terms of the sequence (k_n) if $\sum_{j=1}^{+\infty} \frac{1}{k_j}$ is divergent, and contains a finite number if $\sum_{j=1}^{+\infty} \frac{1}{k_j}$ is convergent.

As a result, since $\sum_{i=1}^{+\infty} \frac{1}{p_i}$ is divergent Hardy and Wright (1979), (p_n) being the sequence of prime numbers, we are sure the Engel expansion has an infinity of prime terms for almost every real number of the unit interval.

In the same references, the reader may find more properties of the Engel expansion.

3. Method

3.1 *Encryption scheme*

First of all, Diffie-Hellman protocol will be used to exchange the seed, so as a public key encryption we will use the classic RSA scheme, which the reader is supposed to be acquainted with.

Alice chooses randomly an irrational number in the unit interval. We adopt the following point of view, which is not a common imposed standard Kane (2013), but is rather a suggestion of the irrational we can choose to expand:

For $r \in [3,+\infty[$ and an algebraic number $A \in]1,+\infty[$: $\sqrt[r]{log(A)}$ is transcendental.

Now the encryption operation steps are described as follows with the message denoted m and key as k :

1. When Alice has the intention to send Bob a message she chooses $s \in \mathbb{N}$, calculates $r \equiv s^e [n]$, then sends r to Bob. Bob gets s via his private key d because $s \equiv r^d [n]$.
2. Bob and Alice calculate $\sqrt[e]{log(s)}$.
3. Both sides calculate the Engel expansion of the irrational number $\sqrt[e]{log(s)}$.
4. Both entities choose the a_i coefficients of the expansion.
5. Those a_i coefficients are concatenated then converted to bits, as a part of the keystream.
6. When plaintext bits exceed the keystream, the previous steps are repeated; otherwise, move on to next step.
7. Alice calculates $m_1 := m \oplus k$ and sends it to Bob who gets m because $m := m_1 \oplus k$.

3.2 *Statistical tests*

In order to check how valid is the random aspect of the algorithm, the NIST test suite was performed on our algorithm. Being applied on over a million bits, this NIST suite has fifteen divided specific statistic tests checking the binary sequence, by searching for visual patterns of ones or zeroes Rukhin et al. (2010). Here is a part of the result for the first one thousand of them: To sum up the empirical outcome of statistical tests, there's the frequency of P-values C_1 to C_{10}, then the P-value is obtained through the Chi-square test, and last but not least the proportion of binary

RESULTS FOR THE UNIFORMITY OF P-VALUES AND THE PROPORTION OF PASSING SEQUENCES

generator is </home/ubu/Desktop/tes.txt>

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	P-VALUE	PROPORTION		STATISTICAL TEST
3	1	0	2	1	5	1	2	3	2	0.437274	20/20		Frequency
1	2	1	1	2	0	6	3	4	0	0.066882	20/20		BlockFrequency
2	3	2	1	0	0	3	0	5	4	0.122325	20/20		CumulativeSums
2	1	2	2	0	2	1	6	1	3	0.213309	20/20		CumulativeSums
2	0	3	4	1	2	3	2	1	2	0.739918	20/20		Runs
2	3	1	3	1	3	3	0	3	1	0.739918	20/20		LongestRun
20	0	0	0	0	0	0	0	0	0	0.000000	0/20	*	Rank
0	2	2	5	0	7	0	0	0	4	0.000648	20/20		FFT
7	1	0	3	0	0	3	0	5	1	0.001399	19/20		NonOverlappingTemplate
5	0	0	1	0	0	5	0	5	4	0.002043	19/20		NonOverlappingTemplate
5	0	0	2	0	0	4	0	5	4	0.006196	20/20		NonOverlappingTemplate
3	0	0	2	0	0	4	0	4	7	0.001399	20/20		NonOverlappingTemplate
1	0	0	3	0	0	8	0	6	2	0.000026	20/20		NonOverlappingTemplate
5	0	0	3	0	0	3	0	4	5	0.008879	19/20		NonOverlappingTemplate
2	0	2	2	0	0	6	0	4	4	0.017912	19/20		NonOverlappingTemplate
2	1	0	1	0	0	4	0	7	5	0.000954	19/20		NonOverlappingTemplate
3	0	0	5	0	0	2	0	7	3	0.000954	19/20		NonOverlappingTemplate
1	0	0	2	0	0	6	0	9	2	0.000002	20/20		NonOverlappingTemplate
2	0	0	2	0	0	9	0	6	1	0.000002	19/20		NonOverlappingTemplate
2	0	1	3	0	0	6	0	6	2	0.002971	20/20		NonOverlappingTemplate
2	0	2	4	0	0	4	0	6	2	0.017912	19/20		NonOverlappingTemplate
5	1	0	2	0	0	4	0	4	4	0.025193	20/20		NonOverlappingTemplate
0	0	1	4	0	0	6	0	5	4	0.001399	20/20		NonOverlappingTemplate
4	0	0	3	0	0	5	0	5	3	0.008879	19/20		NonOverlappingTemplate
3	0	3	2	0	0	7	0	1	4	0.004301	20/20		NonOverlappingTemplate
3	0	3	1	0	0	5	0	7	1	0.001399	20/20		NonOverlappingTemplate
5	0	0	2	0	0	3	0	7	3	0.000954	20/20		NonOverlappingTemplate
5	0	0	2	0	0	3	0	4	6	0.002971	17/20	*	NonOverlappingTemplate
2	0	1	2	0	0	5	0	6	4	0.006196	20/20		NonOverlappingTemplate
5	1	0	5	0	0	3	0	3	3	0.025193	20/20		NonOverlappingTemplate
3	0	1	0	0	0	4	0	7	5	0.000439	17/20	*	NonOverlappingTemplate

Figure 1

NIST statistical test

sequences having passed the test is described at the very right column, showing that it passes the majority of tests but fails others as in figure 1. It should be stressed, however, that the overall result depends also on the sample sequence and block size. The test draws the plot of P-values to check if it falls in the range of acceptance ($\hat{p} \pm 3\sqrt{\frac{\hat{p}(1-\hat{p})}{m}}$ with $\hat{p} = 1-\alpha$ and α being the significance level) or verify its uniformity distribution via a histogram (if $P_{value} \geq 0.0001$).

3.3 *Encryption analysis*

For the sake of simplicity, and to save place, we choose to only focus on image encryption. Let's first take a look at what we have accomplished

Figure 2
Clear image

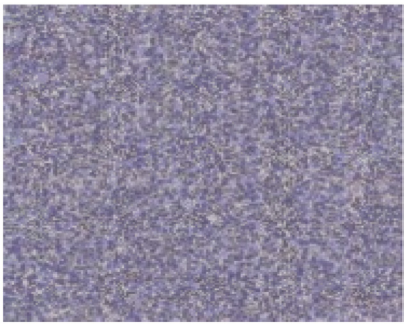

Figure 3
Image cipher

having this input on figure 2. It should be remembered that for the cipher technique, various methods could be considered in order to provide an encrypted image input via pixel adjustments or scrambling according to the key, while dealing with another format before conversion may be of help. Using our encryption scheme described in subsection 3.1 we obtain the image cipher as illustrated on figure 3.

The three RGB components of our encrypted image have uniform histograms, as being displayed on figure 4.

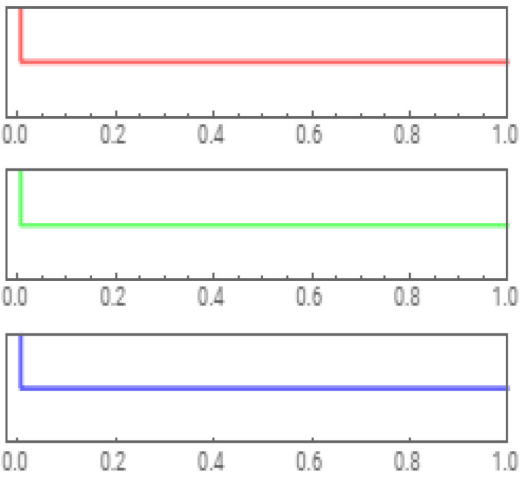

Figure 4
Histogram of each RGB component of the encrypted image

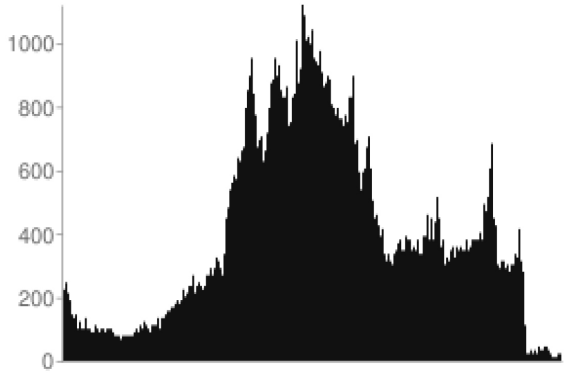

Figure 5

Grayscale histogram of the clear image

Based upon the histogram analysis, we are now sure statistical attacks aren't effective, pixel values being altered using our encryption scheme.

The following histograms on figures 5 and 6 show colors by tone on the abscissas and the number of pixels corresponding to it on the ordonates. The image being processed in grayscale, in the left we have dark colors, gray ones in the middle and light at the right.

Notice that, at least with the naked eye, the original image has gray colors dominance, while the encrypted one has darker colors.

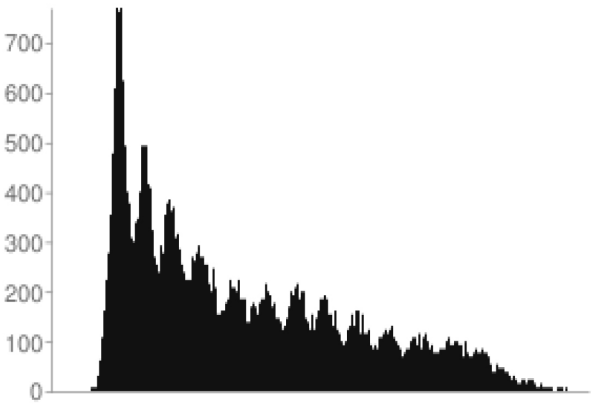

Figure 6

Grayscale histogram of image cipher

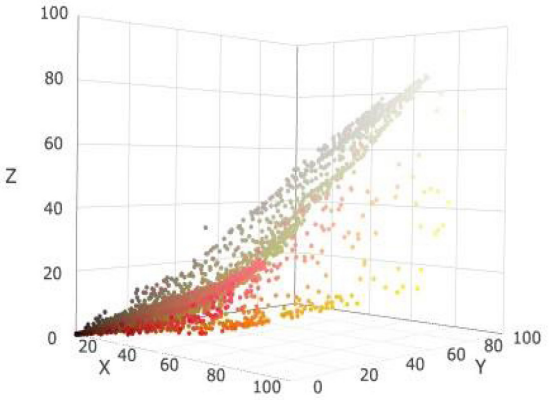

Figure 7

Scatter diagram for the clear image

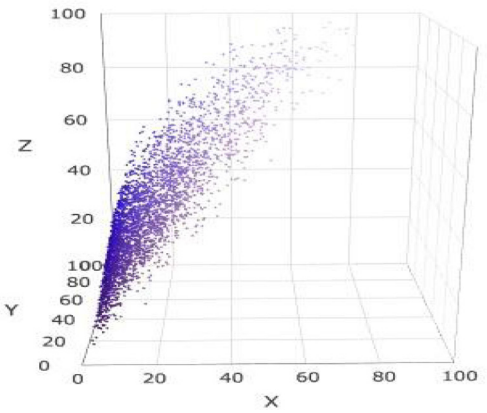

Figure 8

Scatter diagram for the image cipher

The scatter diagrams for RGB distribution of the pixels in both the clear image and image cipher on figure 7 and figure 8 show the decorrelation coefficients between the pixels.

4. An attempt to justify the cryptosystem within chaos theory

Here are some prerequisites which will be used to prove the chaotic behaviour of the process.

4.1 *Definition*

Let f be a function. x is a fixed point of f if $f(x) = x$. It is called a periodic point of f with period $n \geq 2$ if $f^n(x) = x$ and $f^k(x) \neq x$ whenever $0 < k < n$. In other words, a periodic point has period n if it returns to its starting place for the first time after exactly n iterations of f. A periodic point of f with period n is obviously a fixed point of f^n.

The set of all iterates of a point x is called the orbit of x, and if x is a periodic point, then it alongside its iterates are called a periodic orbit or a cycle.

4.2 *Definition*

Let D be a subset of a metric space with metric d. The function $f : D \to D$ is topologically transitive on D if for any open sets U and V that intersect D, there is z in $U \cap D$ and a natural number n such that $f^n(z)$ is in V. Equivalently, f is topologically transitive on D if for any two points x and y in D and any $\varepsilon > 0$, there is z in D such that $d(z, x) < \varepsilon$ and $d(f^n(z), y) < \varepsilon$ for some n.

4.3 *Proposition*

Let D be a subset of a metric spacee and $f : D \to D$. If the periodic points of f are dense in D and there is a point whose orbit under iteration of f is dense in D, then f is topologically transitive on D.

4.4 *Definition*

Let D be a subset of a metric space with metric d. The function $f : D \to D$ exhibits sensitive dependence on initial conditions if there exists a $\delta > 0$ such that for any x in D and any $\varepsilon > 0$, there is a y in D and a natural number n such that $d(x, y) < \varepsilon$ and $d(f^n(x), f^n(y)) > \delta$.

4.5 *Definition (Devaney (1989))*

Let D be a subset of a metric space.

The function $f : D \to D$ is chaotic if

a) the periodic points of f are dense in D (order in chaos),

b) f is topologically transitive (ergodicity), and

c) f exhibits sensitive dependence on initial conditions (butterfly effect).

4.6 *First analysis*

First of all, it will be easily seen that the underlying recurrence function of the Engel expansion, given by

$$f :]0,1] \to]0,1]$$
$$x \mapsto \left(1 + \left\lfloor \frac{1}{x} \right\rfloor\right) x - 1$$

is far from being chaotic. In fact, the sequence (u_n) being monotonic, the set of periodic points of f is empty on one hand, and on the other, f cannot be topologically transitive. Consequently, the property of sensitive dependence on initial conditions will not be of any importance for this study. Besides, it's worth mentioning that the set of fixed points of f (see figure 9) is given by $\mathcal{F} = \left\{1/n \in \mathbb{N}^*\right\}$, and couched in the language of dynamical systems, we have the following formula involving stable sets:

$$W^s(0) =]0,1] \cap (\mathbb{R} \setminus \mathbb{Q})$$
$$\bigcup_{x \in \mathcal{F}} W^s(x) =]0,1] \cap \mathbb{Q}$$

$\mathbb{R}$ and $\mathbb{Q}$ being the field of real and rational numbers, respectively.

4.7 *Second analysis*

For reasons that will become clear afterwards, our phase space will be from now on the interval $I = [0.1,1[$. For a rational with Engel expansion

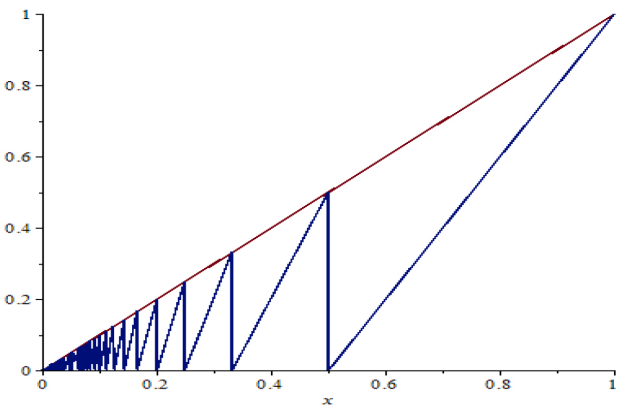

Figure 9

Graph of f

$[a_1, a_2, ..., a_n]$, we define its orbit as being the cycle obtained by repeating the sequence $(x, 0.a_1, 0.a_2, ..., 0.a_n)$ and for x irrational, we associate the sequence $(x, 0.a_1, 0.a_2, ...)$.

It's then easy to see that, by construction, any orbit is included in I and the set of periodic points is dense in I. According to proposition 4.3, we also have the topological transivity since the orbit of $e-2$ is dense in I.

As for sensitive dependence on initial conditions, let x be irrational and $\varepsilon > 0$. We have the Engel expansion $x = [a_1, a_2, ..., a_n, ...]$. Since $(a_n)_{n\geq 0}$ is a non-decreasing sequence, there's $N \in \mathbb{N}$ such that $\frac{1}{a_1 a_2 ... a_N} < \frac{\varepsilon}{2}$.

Let $y = [a_1, a_2, ..., a_{N-1}, b_N, a_{N+1}, a_{N+2}, ...]$, with $b_N := a_N + 10^{\overline{a_N}}$ if the first digit of a_N belongs to $2,9$, and $b_N := a_N + 20^{\overline{a_N}}$ if that digit is 1.

We have $|x - y| < \varepsilon$ and $|0.a_N - 0.b_N| > \delta = 0.1$.

The same reasoning holds for rationals, although their Engel expansion is finite. Thus, we obtain the sensitive dependence on initial conditions, and conclude that this process is chaotic.

4.8 *Engel expansion distribution*

The histogram of what distributions of $0.a_i$ throughout the intervals resembles is shown on figure (10), upon which we can notice an almost invariant distribution of the iterates on the interval $[0.1, 1[$.

Now, showing the dependence on initial conditions, let's take on figure 11 the plot of the orbits starting from two initial conditions x_0 and x_1 with a slight difference and see how their curves part ways.

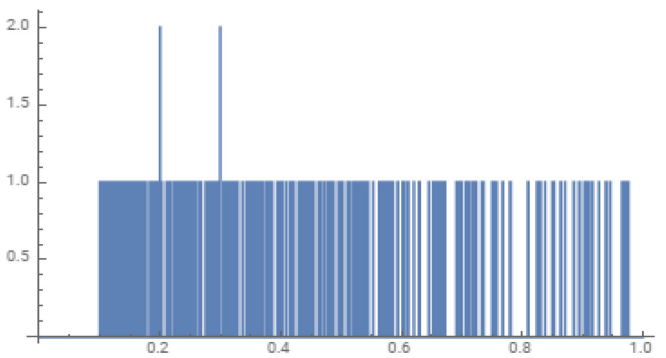

Figure 10

Distribution of iterates

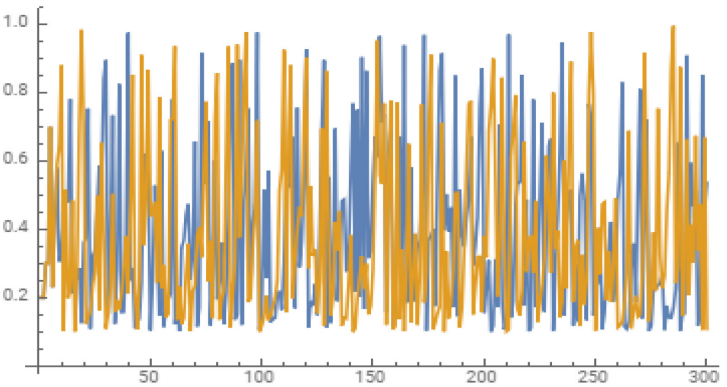

Figure 11

Sensitive dependence of the process on initial conditions

5. Conclusion

In the present work, an Egyptian-expansion-based cryptosystem was proposed and analyzed from a statistical point of view and a chaotic one. From the first experiments, we deduce that secure encryption schemes can emerge from an advanced study of Egyptian fractions.

For future work, we suggest another approach consisting on considering the concatenation of the coefficients as fractional parts and shifting alongside the sequence. From a first diagnosis, it seems that such a process leads to *normal* numbers for almost every initial condition, as for instance it is still conjectured for some irrationals.

References

[1] Shannon, C. E. (1948) 'A Mathematical Theory of Communication', *The Bell System Technical Journal*, Vol. 27, pp. 379-423, 623-656.

[2] Pecora, L. M. and Carroll, T. L. (1990) 'Synchronization in chaotic systems', *Phys. Rev. Lett.*, 64, 821.

[3] Baptista, M. S. (1998) 'Cryptography with chaos', *Physics Letters A*, 240(12), 50 - 54.

[4] Kocarev, L. and Lian, S. (2011) 'Chaos-Based Cryptography: Theory, Algorithms and Applications', *Studies in Computational Intelligence*, Springer, Volume 354.

[5] Kane, A. M. (2013) 'On the use of continued fractions for stream ciphers', *IACR Cryptology ePrint Archive*, In Proceedings of Security and Management 2009, Las Vegas.

[6] Mikram, J., Zinoun, F. and Hamri, M. (2012) 'An Encryption Algorithm Based on the Decimal Expansion of Irrationals', *Applied Mathematical Sciences*, Vol. 6, no. 70, 3475 - 3494.

[7] Masmoudi, A., Puech, W., Bouhlel, M. S. (2010) 'An Efficient PRBG Based on Chaotic Map and Engel Continued Fractions', *J. Software Engineering & Applications*, 3, 1141-1147.

[8] Hussain A. Younis, Israa M. Hayder, Israa Shakir Seger and Hameed Abdul-Kareem Younis, (2020) 'Design and implementation of a system that preserves the confidentiality of stream cipher in non-linear flow coding', *Journal of Discrete Mathematical Sciences and Cryptography*, 23:7, 1409-1419. DOI: 10.1080/09720529.2020.1714890.

[9] Fibonacci, L. (trad. Laurence E.Sigler)(2002) 'Fibonacci's Liber Abaci: A Translation Into Modern English of Leonardo Pisano's Book of Calculation', *Springer-Verlag*.

[10] Devaney, R. L. (1989) 'An introduction to chaotic dynamical systems', *Addison-Wesley*.

[11] Erdös, P., Rényi, A., and Szüsz, P. (1958) 'On Engel's and Sylvester's series', *Ann. Univ. Sci. Budapest.*, Eötvös Sect. Math. 1, 7-32.

[12] Rènyi, A. (1962) 'A new approach to the theory of Engel's series.', *Ann. Univ. Sci. Budapest*, Sectio Math. 5 (1962), 25-32.

[13] Hardy, G. H. and Wright, E. M. (1979) 'An introduction to the theory of numbers', *Fifth ed*, The Clarendon Press Oxford University Press.

[14] Rukhin, A., Soto, J., Nechvatal, J., Smid, M., Barker, E., Leigh, S., Levenson, M., Vangel, M., Banks, D., Heckert, A. and Dray, J. (2010) 'A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications', *National Institute of Standards and Technology*, computer security, Special Publication 800-22, Revision 1a.

Received October, 2020

Revised February, 2021