ROYAUME DU MAROC
Université Mohammed V - RABAT -
جامعة محمد الخامس – الرباط –
Faculté des sciences
كلية العلوم
CENTRE D'ETUDES DOCTORALES - SCIENCES ET TECHNOLOGIES

N°d'ordre: 3859

# THESE

En vue de l'obtention du ***DOCTORAT***

**Centre de Recherche** : Centre de Recherches Mathématiques et Applications de Rabat
**Structure de Recherche**: Laboratoire de Mathématiques, Informatique et Applications - Sécurité de l'Information
**Discipline** : Mathématiques Appliquées
**Spécialité** : Théorie du Chaos et Cryptographie

**Présentée et soutenue le 21/10/2023 par:**

**Ilias CHERKAOUI**

## Chaos Déterministe et Quelques Aspects Ergodiques de la Théorie des Nombres en Cryptographie

***JURY***

| | | |
|---|---|---|
| **EL Mamoun SOUIDI** | PES, Faculté des Sciences de Rabat, Université Mohammed V, Rabat | Président/ Rapporteur |
| **Hafida BENAZZA** | PES, Faculté des Sciences de Rabat, Université Mohammed V, Rabat | Rapporteur/ Examinateur |
| **Abderrahim EL ABDLLAOUI** | PH, Faculté des Sciences de Rabat, Université Mohammed V, Rabat | Rapporteur/ Examinateur |
| **Mustapha SERHANI** | PES, Faculté des Sciences Juridiques, Economiques et Sociales de Meknès, Université Moulay Ismail, Meknès | Rapporteur/ Examinateur |
| **Fouad ZINOUN** | PES, Faculté des Sciences de Rabat, Université Mohammed V, Rabat | Directeur de thèse |

Année Universitaire: 2023–2024.

✉ Faculté des Sciences, avenue Ibn Battouta, BP. 1014 RP, Rabat –Maroc
☎ 00212(0)5 37 77 18 76 🖷 00212(0)5 37 77 42 61 ; http://fsr.um5.ac.ma/

# Dédicace

J'aimerais adresser mes vifs remerciements à toute personne ayant contribué de près ou de loin à la réalisation de ce travail. Je remercie le Chef et la Secrétaire du Département de Mathématiques, les Entités et Fonctionnaires administratifs, ainsi que et par dessus tout, l'ensemble du Corps Professoral de notre Établissement pour les longues années qu'ils ont investi à me former en l'homme que je suis aujourd'hui.

Et c'est avec une joie incomparable et un coeur ému que je dédie cette thèse entière à mes chers parents pour leur amour inlassable, l'affection inépuisable, leurs précieux conseils, le soutien et encouragement constants étant d'un grand réconfort sans cesse.

Enfin, je remercie tous mes amis qui ont toujours été là pour moi. Leur soutien inconditionnel et leurs encouragements ont été dune grande aide.

À tous ces intervenants, je présente mes remerciements, mon respect et ma gratitude.

# Dedicace

I would like to express my sincere thanks to everyone who has contributed in any way to the completion of this work. I would like to thank the Head and Secretary of the Mathematics Department, the administrative bodies and officials and, above all, the entire teaching staff of our institution for the many years they have invested in shaping me into the man I am today.

And it is with incomparable joy and a moved heart that I dedicate this entire thesis to my dear parents for their untiring love, inexhaustible affection, invaluable advice, constant support and encouragement being a great comfort all the time.

Finally, I'd like to thank all my friends who have always been there for me. Their unconditional support and encouragement have been a great help.

I would like to express my thanks, respect and gratitude to everyone involved.

# Remerciements

Je tiens à remercier mes collègues du Laboratoire de Mathématiques, Informatique et Applications - Sécurité de l'Information (**LabMIA-SI**) du département de mathématiques de l'Université Mohammed V de Rabat, Faculté des Sciences de Rabat, pour leur soutien. Je tiens également à remercier tous mes professeurs pour m'avoir aidé à développer les compétences académiques fondamentales et essentielles.

Je suis très honoré que Monsieur **El Mamoun SOUIDI**, PES à la Faculté des Sciences de Rabat, Université Mohammed V de Rabat, ait accepté de présider le jury de soutenance et d'examiner mon travail de recherche. Ses cours dispensés au sein du Master Codes, Cryptographie et Sécurité de l'Information resteront pour moi une référence dans le domaine. Je les garde précieusement et je lui en suis profondément reconnaissant.

Je tiens à exprimer mes plus vifs remerciements à Monsieur **Fouad ZINOUN**, PES à la Faculté des Sciences de Rabat, Université Mohammed V de Rabat, qui fut pour moi un directeur de thèse attentif et disponible, malgré ses nombreuses charges; sa compétence, sa rigueur scientifique et sa clairvoyance m'avaient énormément appris. Il était et persistera comme l'un des piliers de mes recherches et travaux futurs.

Je tiens notamment à exprimer mon immense sentiment de gratitude à Madame **Hafida BENAZZA**, PES à la Faculté des Sciences de Rabat, Université Mohammed V de Rabat, d'avoir accepté d'être rapporteur de ma thèse. Son expérience en équations

différentielles et systèmes dynamiques est d'un grand apport.

Je tiens également à exprimer ma profonde gratitude à Monsieur **Abderrahim EL ABDLLAOUI**, PH à la Faculté des Sciences de Rabat, Université Mohammed V de Rabat, d'avoir accepté d'être rapporteur de ma thèse. Son enseignement de la théorie qualitative des équations différentielles est d'une grande qualité scientifique. J'en ai énormément appris et je lui suis profondément redevable.

Je suis très sensible à l'honneur et le privilège que me fait Monsieur **Mustapha SERHANI**, PES à la Faculté des Sciences Juridiques, Economiques et Sociales de Meknès, Université Moulay Ismail de Meknès, d'avoir accepté de rédiger un rapport de ma thèse. Outre les systèmes dynamiques, j'estime que son expertise en théorie du contrôle serait très enrichissante pour mes travaux futurs, notamment en ce qui est de la synchronisation des systèmes chaotiques, domaine très prometteur en cryptographie à base de chaos. Qu'ils trouvent ici l'expression de mon grand intérêt et de ma sincère gratitude.

# Acknowledgement

I would like to thank my colleagues at the Laboratory of Mathematics, Informatics and Applications - Information Security (**LabMIA-SI**) in the Department of Mathematics at Mohammed V University in Rabat, Faculty of Sciences of Rabat, for their support. I would also like to thank all my professors for helping me to develop fundamental and essential academic skills.

I am very honoured that Mr. **El Mamoun SOUIDI**, PES at the Rabat Faculty of Sciences, Mohammed V University in Rabat, has agreed to chair the defence jury and examine my research work. His lectures in the Codes, Cryptography and Information Security Master's programme will remain a reference in the field for me. I treasure them and am deeply grateful to him.

I would like to express my warmest thanks to Mr. **Fouad ZINOUN**, PES at the Faculty of Sciences in Rabat, Mohammed V University in Rabat, who was an attentive and available thesis supervisor for me, despite his many duties; his competence, scientific rigour and clear-sightedness taught me a great deal. He was and will remain one of the pillars of my future research and work.

In particular, I would like to express my immense gratitude to Ms. **Hafida BENAZZA**, PES at the Faculty of Sciences, Mohammed V University in Rabat, for agreeing to be the rapporteur for my thesis. Her experience in differential equations and dynamical systems is a great help.

I would also like to express my deep gratitude to Mr. **Abderrahim EL ABDL-LAOUI**, PH at the Faculty of Sciences of Rabat, Mohammed V University in Rabat, for agreeing to be the rapporteur of my thesis. His teaching of the qualitative theory of differential equations is of the highest scientific quality. I have learnt a great deal from him and I am deeply indebted to him.

I am very grateful for the honour and privilege that Mr. **Mustapha SERHANI**, PES at the Faculty of Legal, Economic and Social Sciences of Meknès, Moulay Ismail University of Meknès, has given me by agreeing to write a report on my thesis. In addition to dynamical systems, I believe that his expertise in control theory would be very enriching for my future work, particularly with regard to the synchronisation of chaotic systems, a very promising field in chaos-based cryptography. I would like to express my great interest and sincere gratitude.

# Résumé

La théorie déterministe du chaos intrigue beaucoup, particulièrement en raison de ses interactions mystérieuses avec la théorie ergodique des nombres. Bien que nous utilisions fréquemment des générateurs pseudo-aléatoires certifiés NIST, nous négligeons souvent de comprendre les raisons mathématiques de leur efficacité. Notre thèse se penche en partie sur ce sujet, à ceci près que nous nous concentrons sur les fractions égyptiennes d'irrationnels, au lieu de développements traditionnels tel le développement décimal ou celui en fractions continues, plus couramment utilisés en cryptographie. En adoptant la définition du chaos selon R. Devaney, nous démontrons mathématiquement la chaoticité d'un tel processus, créant ainsi un nouveau cryptosystème venant enrichir l'arsenal de systèmes à base de chaos déjà à notre disposition. Nous explorons ensuite des familles de fonctions injectives et lossy trapdoor (LTF), pratiquement indiscernables sur le plan calculatoire, mais dont l'utilité dans la construction de primitives cryptographiques est avérée. Plus précisément, nous mettons en place une construction efficace d'une variante des fractions égyptiennes pour extraire notre LTF souhaitée. Nous visons ainsi à améliorer l'efficacité du schéma de chiffrement résistant à l'attaque à texte chiffré choisi (IND-CCA), et ce en faisant appel aux notions de tenseurs et de catégories, tout en démontrant l'ergodicité de ce processus. L'aspect pseudo-aléatoire des processus inspirés des fractions égyptiennes contribuera à renforcer la sécurité non seulement dans le schéma IND-CCA, mais aussi dans divers défis liés à l'hypothèse décisionnelle ou calculatoire de Diffie-Hellman (DDH, CDH), ce qui les rend très précieux pour la communication entre plusieurs agents.

**Mots clés:** Théorie du chaos; Théorie ergodique des nombres; Fractions égyptiennes; Cryptographie; LTF.

# Abstract

Deterministic chaos theory is very intriguing, particularly because of its mysterious interactions with ergodic number theory. Although we frequently use NIST-certified pseudorandom generators, we often neglect to understand the mathematical reasons for their effectiveness. Our thesis partly addresses this issue, except that we focus on Egyptian fractions of irrationals, rather than traditional developments such as decimal or continued fractions, which are more commonly used in cryptography. Adopting the definition of chaos according to R. Devaney's definition of chaos, we demonstrate mathematically the chaotic nature of such a process, thus creating a new cryptosystem to add to the arsenal of chaos-based systems already available to us. We then explore families of injective and lossy trapdoor functions (LTFs), which are practically indistinguishable computationally, but which have been shown to be useful in the construction of cryptographic primitives. More precisely, we set up an efficient construction of a variant of Egyptian fractions to extract our desired LTF. Our goal is to improve the efficiency of the chosen ciphertext attack-resistant encryption (IND-CCA) scheme by using the notions of tensors and categories, while demonstrating the ergodicity of this process. The pseudorandom aspect of the processes inspired by Egyptian fractions will contribute to enhancing security not only in the IND-CCA scheme, but also in various challenges related to the Decisional Diffie-Hellman or computational hypothesis (DDH, CDH), making them very valuable for communication between several agents.

**Keywords:** Chaos theory; Ergodic theory of numbers; Egyptian fractions; Cryptography; LTF.

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

> The world and science have their own data, which touch and do not penetrate each other. One shows us what goal we should aim for, the other, the target being given, offers us the means to wait for it.
>
> *Science and Hypothesis*
>
> *Henri Poincaré*

## 1.1 Overview[1]

In the midst of a maelstrom of ontological arguments in the Renaissance, Spinoza[2], a pure advocat of rational and cartesian thought, layed some bedrocks for deterministic philosophy. Although sometime far too complex to entangle, his manoeuvres, skeptical as they seem, used freedom or free-will and determinism interchangeably. Heretofore, Hegel[3] possessed his own doctrine, but still acknowledged determinism, which was even known to affect Leibnitz monadology[4] after his visit to Spinoza's recluse, nay

[1]Antique references are marked as footnotes, while the others are listed in the bibliography.

[2]Baruch Spinoza The Collected Works of Spinoza, Vol. I and II, ed. and trans. by E. Curley. Princeton University Press, 1985.

[3]Georg W. F. Hegel Lectures on the Philosophy of World History: Introduction. Translated by Nisbet, H. B. Cambridge University Press, 1975

[4]Nicholas Rescher Leibniz's Monadology, University of Pittsburgh Press, 1991

altered it once and for all.

Determinism gives the illusion of coincidence when the information at stake is incomplete, like Voltaire[5] once said and here quoted: *chance is a word of void sense; nothing can exist without cause*. Chance may also after a similar event give the appearance of a deterministic process. As a matter of fact, unpredictability does not necessarily involve the intervention of this elusive coincidence. Nonetheless, perfectly deterministic procedures can sometimes have an unexpected outcome, only attainable by carrying out the experiment itself; it is then unpredictable, although no randomness intervenes in the construction.

Throughout history, deterministic perspectives would soon become more and more shaken by later discoveries. To emphasis this stance transitions, a quick dive into astronomy's timeline is to be mentioned:

In ancient Greece and Rome, Ptolemy[6] and Hipparchus[7] focused on finding regular motion of wandering stars (planets) to predict their movement among the fixed stars, coming up with uniform circular motions and epicycles. Until the early seventeenth century, Kepler [1] announced his planetary laws improving the Copernicus model [2], stating that orbits of planets around the sun were elliptical trajectories. Since then, Newton [3] recovered those fixed ellipses in his laws for the orbit of a single planet around the sun; however, he found that adding a second planet like in figure 1.1 to the count would disturb the orbit of the first one, thus revealing the important question of the stability of the solar system. This question will be soon answered at the end of the eighteenth century with Lagrange [4], laying down the differential equations responsible for the elliptic motion variations under the influence of the planetary perturbations. This allowed Laplace [5] to calculate long term variations and periodic terms, convincing him of the solar system's stability and predictability over infinite time, a major keystone of Laplace's determinism.

[5]François-Marie Arouet (known as Voltaire) Le Philosophe ignorant 1766

[6]Alexander Jones The ancient Ptolemy. ln Ptolemy's Science of the Stars in the Middle Ages. Ptolemaeus Arabus et Latinus Studies 1, 13-34, 2020.

[7]Otto E. Neugebauer A History of Ancient Mathematical Astronomy. Pt. 13. Berlin, Heidelberg, New York: Springer Verlag, 1975.

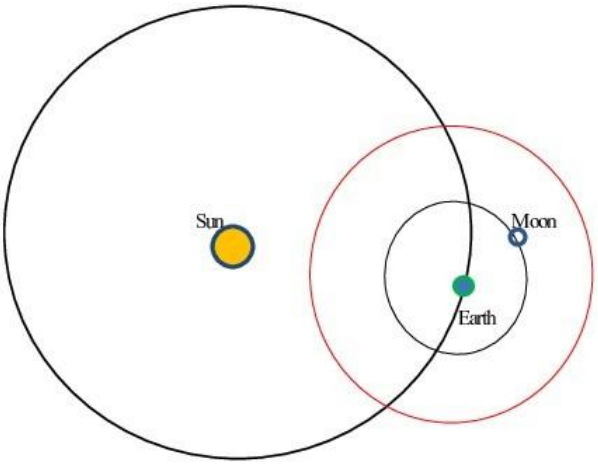



**Figure 1.1:** *The three-body problem in the solar system*

Towards the end of the nineteenth century, Poincaré [7] addressed the stability of the solar system on a simplified model situation, constituting the first contribution in qualitative theory of dynamical systems. He proves that the existence of homoclinic intersections implied sensitive dependence on initial conditions and to predict the gravitational motion of celestial objects, he studies two bodies orbiting one another around their common gravity center, while a smaller body orbits them both (known as the three-body problem). During this computation, he discovered small uncertainties in initial conditions that made his result numbers grow out of range, no matter how hard he tried to shrink those imprecisions, and deducing that the solar system was unpredictable due to the presence of a multitude of variables nearly impossible to specify precisely, giving room for "chaos". Indeed, exactly a century after Poincaré, and thanks the high performance of numerical simulations, the latest works by Laskar [8] has led to different probabilistic scenarios, all confirming Poincaré's prophecy about the unstable nature of the solar system.

A quite simple, yet convincing example of chaoticity can be shown through a game of pool: assume we are in the middle of a conservative game of billiards with no pocket holes, for dynamical reasons, and given a very short distance between two adjacent balls at the start, we can witness an exponential divergence of trajectories with nearby initial conditions (sensitive dependence of initial conditions), shown on figure 1.2. Each

ball collides to another and to the boundary of the billiard table, keeping in mind that the angle of incidence equals the angle of reflection. In 1927, Birkhoff [17] announced that the elliptic shape table billiards were integrable, the collision space for this mapping being foliated by 1-dimensional invariant manifolds. At this point, he conjectured, that strictly convex integrable billiards were elliptic billiards, and as a corollary of the Poincaré-Birkhoff fixed point theorem, it is deduced that on a strictly convex billiard table, there is infinitely many distinct periodic orbits. Thus, it is the shape of the table boundary that renders the behaviour of such Hamiltonian systems, integrable or chaotic.

Astronomy or billiard are not the only illustrations of this dynamical phenomena, as

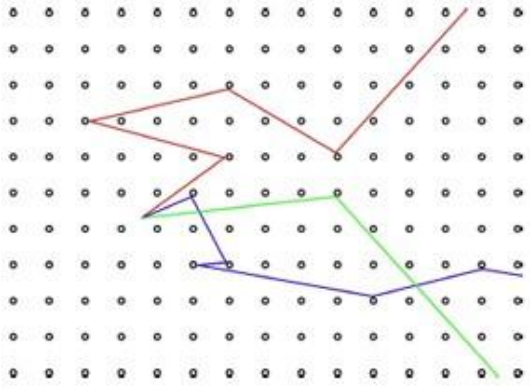


**Figure 1.2:** *Three possible trajectories of a ball colliding to others on a table*

one can find it hidden behind shape deforming: In 1967, shearing the image of a cat with the aim of stretching it, Arnold [9] used a linear transformation via the Fibonacci matrix [10], while adding a modulation to make it a map from a torus to itself, which was found later on to be chaotic; he upped each point with one unit and shifted it two units to the right, as shown on figure 1.3. Now, given a *mod* 1 constraint, the opposite boundaries made the same entity, hence could be connected to get a cylindrical set, and redoing the same process for the other boundary to get a torus. Therefore, regardless of the times this mapping was used, it yielded the same shape inside the square unit. Eventually, since this map was periodic, once it reaches a defined number of iterations, the image is restored back to its original form. This is the beauty in chaos: you do not loose information, it is sitting out there waiting to be retrieved !

In that same year of 1967, Smale [16] suggested a chaotic homeomorphism from the

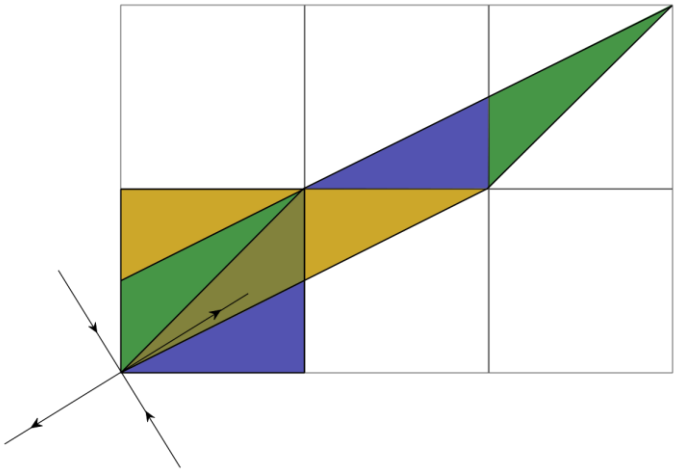


**Figure 1.3:** *Arnold's cat map*

plane to itself. Focusing on the operation inside the unit square, which can be extended to the rest of the plane, the same square is squeezed horizontally, stretched vertically, then bent around in a clockwise direction. Now, if the process is repeated, but in a counter clockwise manner, the transformation would still be symmetrical and invertible, as displayed on figure 1.4. Since clearly one equilibrium does exist, being the origin (saddle point) with one stable and another unstable manifold (eigenspaces matching the $x$ and $y$ axis), then the invariant set of the map is non-empty and lies inside the unit square, where some points remain within for all time.

Mathematically thinking, a chaotic system is usually endowed with sensitivity to initial conditions, which could be referred to as the butterfly effect, meaning that if the start of a path is point A, then starting near point A could lead to a completely different direction, as questioned by Lorenz [14] about the butterfly wings, a question which seemed at the time unnatural for scientists. This gave rise to the new paradigm of strange attractor [15].

Not far from shape transformations and chaos, shortly after the Lorenz attractor discovery, a biochemist named Rössler [131] became fascinated with chaos and began

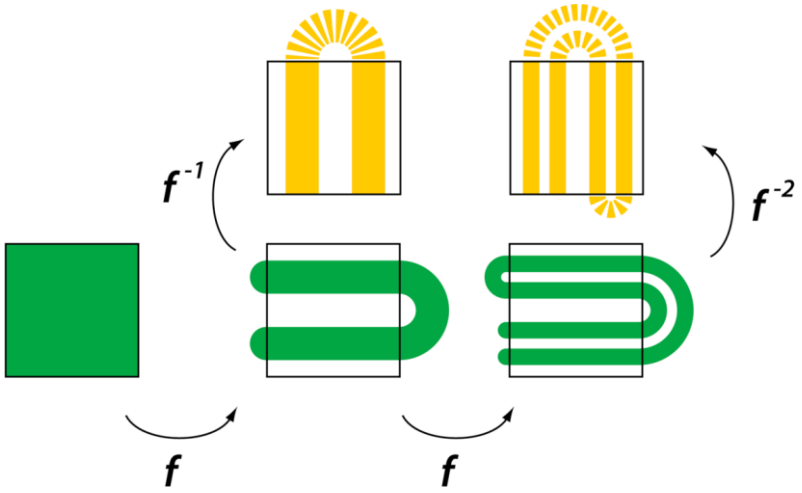



**Figure 1.4:** *Horseshoe map pre-images of the square region*

thinking of a way much simpler to find an equivalent to that attractor in chemistry. Visiting a carnival, Rössler saw the mechanism of the taffy puller, and wondered if he can write an ordinary differential equation doing the same and containing only one nonlinearity. In this attempt, the biochemist noticed that the Lorenz attractor occurred in a bounded region, yet different points are growing apart further and further from each other, an endless expansion that can be mathematically expressed in terms of the Lyapunov exponents [6]. The Rössler attractor is born (figure 1.5).

Interestingly, still within the spirit of chaotic advection, and just like pastry kneading, Baker's map [53] goes from the unit square to itself, by result of flattening and stretching then cutting and stacking (figure 1.6), identically as performed on dough. This map exhibits sensitive dependence to initial conditions, having uncountably many chaotic orbits. When considering the unit square and while the case parameter for the transformation is less than $\frac{1}{2}$ we notice that successive images of the square are all nested inside each other, which insinuates that Baker's map has an attractor, being the limiting set, that pulls in all orbits. Following the map's iterations, any initial condition inside the original square moves around but will be found in the limiting set. This fact is what makes this attractor fractal, with a non-integer Hausdorff dimension. For the same case where the parameter is less than $\frac{1}{2}$, the area of a Baker's map iteration for any given

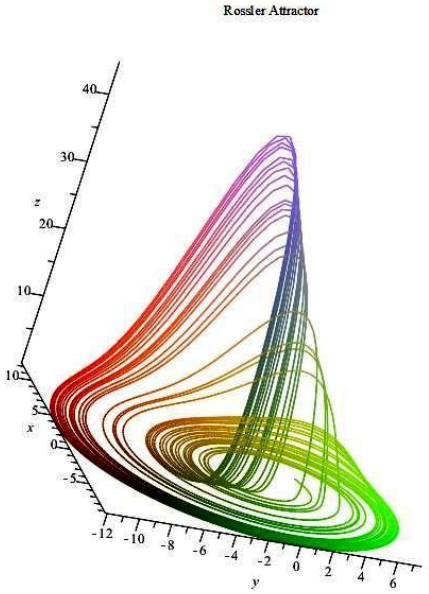



**Figure 1.5:** *Rössler attractor*

region inside the square is less than the area of the higher iteration, meaning we have a shrinking phase space volume, a dissipative system. When the above-mentioned parameter is equal to $\frac{1}{2}$ the map is area-preserving and there is no gap between the stripes, making the square map back to itself. In this case, there is chaos but there is no strange attractor, as for the solar system.

Long after the first spark of chaos, electronic circuits did not escape this scientific breakthrough: in 1983, while collaborating with Matsumoto in Japan, Chua [130] implemented an experiment using a linear resistor, a coil, two capacitors and a nonlinear resistor in the form of operational amplifiers and batteries, allowing later on to get a system of three state variables and five parameters representing the circuit equations. While setting up the circuit, the oscilloscope showed at first wavy signals, but after adjusting the potentiometers and plotting the signals against each other, a double scroll appeared like on figure 1.7, as one of the first Chua attractors indicating chaos.

Chaos is not only encountered in physics, chemistry or economy but also in medical studies: in 1980, medical practitioners had linked the heartbeat of humans to chaotic behavior, which quickly became a study that could help patients with heart problems:

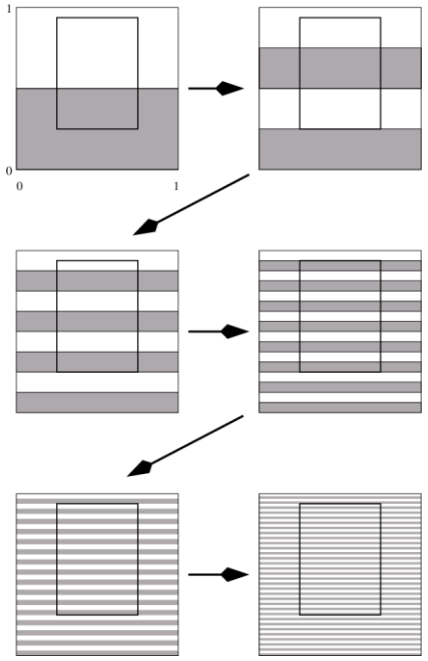



**Figure 1.6:** *Successive iterates of the Baker's map on a set*

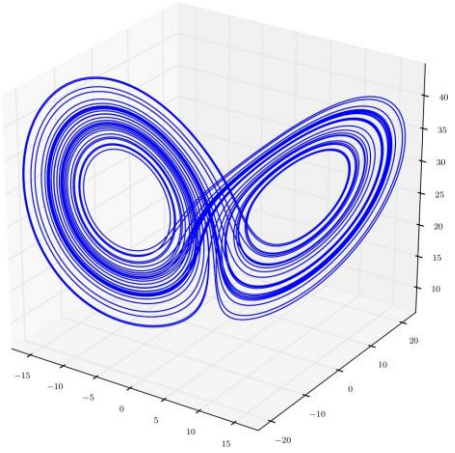



**Figure 1.7:** *Chua's attractor*

Cohen and Kaplan [18] wrote about heart electrical activity becoming chaotic. Lombardi [12] had also published that the fluctuations of palpitations during the normal sinusoidal rhythm of the heartbeat are partially attributed to deterministic chaos, while a decrease in this type of nonlinear variations could be observed in various cardiovascular decays and before ventricular fibrillation.

Finally, and perhaps the most important for our thesis, undoubtable links can be established between chaos and (ergodic) number theory. As mentioned in the abstract, the conjecture of normality (or at least universality) of some fundamental constants like $\pi$ amounts to proving the density in the unit interval of the orbit of $\pi$ via the iterates of the map $10x \bmod 1$, a keystone property in the definition of chaos according to Devaney (see for instance proposition 2.5). In this context, we invite the reader to take profitably this fact to reformulate in dynamical terms some conjectures from number theory. As the main part of this problems is formulated within the decimal or continued fractions expansion, other paths could be explored within Egyptian fractions expansion and its variants, as will be attempted in this thesis.

## 1.2 Thesis Plan and Purpose

In this thesis, it is dealt with some aspects of the interconnection between chaos theory and ergodicity of numbers, as well as their application to cryptography, with the following plan:

- The first chapter was a short overview of how chaos theory became such an interesting field of mathematics.
- The second chapter covers discrete dynamical systems from a topological point of view to reach the definition of chaos according to Devaney [24]. It also defines fundamentals of encryption and chaos-based cryptography.
- The third chapter is about announcing Egyptian fractions and their properties, while reminding the reader of some prerequisites needed for secure multi-party computation and lossy trapdoor functions [74].
- In The fourth chapter, and after having suggested an encryption algorithm on the basis of Egyptian fractions [22], we move on to proving that it is chaos-based, which will be carried out according to definition of Devaney.

- In the fifth chapter, a Lossy Trapdoor Function is then built [23] from a variant of the Engel expansion, using its ergodicity to help improve the security of a Diffie-Hellman mutli-challenge setting.
- The sixth and final chapter is meant to be a conclusion summarising the work done throughout this thesis and proposing new possible routes as perspectives to be followed in future contributions.

## 1.3 Communications and Publications

[19] I. Cherkaoui and F. Zinoun (2018) Encrypting with Egyptian Fractions: From Number Theory to Chaos-Based Cryptography. International Conference on Applied Mathematics, Fez.

[20] I. Cherkaoui and F. Zinoun (2019) On the Egyptian Product-Based Encryption. International Conference on Research in Applied Mathematics and Computer Science, Casablanca.

[21] K. Chetioui, G. Orhanou, H. Bensaid, I. Cherkaoui and Y. Chibi (2019) Formal Verification of Confidentiality in DNSSEC and E-DNSSEC Protocols using pi-calculus and ProVerif. International Workshop on Emerging Networks and Communication, Coimbra, Portugal. (Not included in thesis)

[22] I. Cherkaoui and F. Zinoun (2021) On the use of Egyptian fractions for stream ciphers. Journal of Discrete Mathematical Sciences and Cryptography, Volume 26, Issue 1, p 139-152.

[23] I. Cherkaoui (2021) Diffie-Hellman Multi-Challenge using a New Lossy Trapdoor Function Construction. IAENG International Journal of Applied Mathematics 51(3), pp. 17.

# Chapter 2

# Chaos theory and cryptography fundamentals

> Of all the possible pathways of disorder, nature favors just a few.
>
> *Chaos: Making a New Science*
> *James Gleick*

Generally, throughout this chapter, the focus will mainly revolve around discrete dynamical systems as the papers of which consists the thesis were done in the discrete case. The following section restates certain topological definitions needed in defining chaotic systems. It also introduces the reader to the fundamentals of cryptography.

## 2.1 Topological prerequisites

This part allows us to go over certain topological definitions needed for the definition of chaos. We first recall some types of spaces and functions that will be of interest for what follows:

- The pair $(E, \tau)$ is a **topological space**, where $E$ is a set and $\tau$ is a collection of subsets of $E$, called **open sets** and satisfying:

1) $\varnothing, E \in \tau$ (the empty set and $E$ belong to $\tau$ );
2) The union of members of $\tau$ belong to $\tau$, i.e. the union of open sets are an open set;
3) A finite intersection of members of $\tau$ are in $\tau$, i.e. a finite intersection of open sets are an open set.

- A subset $X \in E$ is **closed** in $(E, \tau)$ if its complement $E \setminus X$ is an open set.
- Let $X$ be a subset of the topological space $(E, \tau)$. The smallest closed set containing $X$ always exists, it is the intersection of all closed sets that contain $X$. This set is referred to as $\overline{X}$ and called **closure**.

A given set can be provided with a variety of topologies, which leads us to the following concept of order:

- A topology $\tau$ over the set $X$ is a **finer** (stronger or larger) topology than $\tau'$, if $\tau' \subset \tau$, the inclusion being understood between sets. The topology $\tau'$ is said to be **coarser** (weaker or smaller) than $\tau$.
  A topology is then finer if it has more open sets. This is a partial order relation.
- Let $(E, \tau)$ be a topological space. A **neighbourhood** of $x \in E$ is every subset of $E$ having an open set that contains $x$.

Given below two main examples of topologies over a set $X$:

- The **discrete topology** over a set $X$ is the topology $\tau = \mathrm{P}(X)$ of all subsets of $X$, which is the largest that contains all subsets as open sets; it is the strongest of all topologies over $X$, where all subsets of $X$ are simultaneously open and closed.
- The **indiscrete topology** is the smallest over $X$; it contains only two open sets: $\tau = \{\varnothing, X\}$.

A convenient method to define topologies is to use **metrics**:

- A **distance** on a set $E$, called a metric, is a map $d : E \times E \longrightarrow \mathrm{R}^{+}$ with the following properties:

  a) **Symmetry:** $\forall\ x, y \in E, d(x,y) = d(y,x)$;

  b) **Indentity of indiscernibles:** $\forall\ x, y \in E, d(x,y) = 0 \Leftrightarrow x = y$;

  c) **Triangle inequality:** $\forall\ x, y, z \in E, d(x,z) \leq d(x,y) + d(y,z)$.

- The pair $(E, d)$ is a **metric space** where $E$ is a set and $d$ a distance on $E$.

The metric spaces that are topological spaces have particular neighbourhoods, denoted **balls**:

- Let $(E, d)$ be a metric space. The **closed ball** centered in the point $P$, of real radius $r$, is the set $\overline{\mathrm{B}}(P, r)$ of points to which the distance to $P$ is less than or equal to $r$:

$$\overline{\mathrm{B}}(P, r) = \{M \in E \,/\, d(M, P) \leq r\}$$

  The **open ball** is the set

$$\mathrm{B}(P, r) = \{M \in E \,/\, d(M, P) < r\}$$

Certain topological spaces, the so-called compact spaces, play a major role in what follows later on. One should first introduce separated spaces and the cover concept, in order to define compactness:

- A **separated space** is a topological space where two distinct points have always two disjoint neighbourhoods (Metric spaces are for instance separated).

- A **cover** of a set $X$ is a set $\mathrm{P}$ of non-empty subsets of $X$ such that the union of these subsets are equal to $X$. The **open cover** is a cover in which the subsets are open.

- A separated space in topology is **compact** whenever it is covered by open sets, it is covered by a finite number of those open sets.

**Proposition 2.1** (Sequential characterisation of compact spaces)**.** *A metric space* $(X, d)$ *is compact if every sequence of X has a convergent subsequence in X (convergence can be defined as usual* $(X, d)$*).*

Let us now introduce the concept of complete spaces, that are metric spaces in which certain sequences converge, as explained below:

- A sequence $(x^n)_{n\in \mathbb{N}}$ of a metric space $(E, d)$ is a **Cauchy sequence** if for every real number $\epsilon > 0$, a natural number $N$ exists such that for all natural numbers $p, q > N$, the distance $d(x^p, x^q)$ is less than $\epsilon$:

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}, \forall\, p, q > N,\ d(x^p, x^q) < \epsilon$$

  .

- A metric space $(E, d)$ is **complete** if every Cauchy sequence in $(E, d)$ has a limit in $(E, d)$.

The continuity notion is generally defined on topological spaces:

- Let $f$ be a map between two topological spaces. It is **continuous** in $x$ if for every neighbourhood $V$ of $f(x)$, there exists a neighbourhood of $x$ which image via $f$ is in $V$.

Continuity can be expressed clearly when considering applications between metric spaces:

- Let $(E, d)$ and $(E', d')$ two metric spaces, $a \in E$ and $f : E \longrightarrow E'$. The map $f$ is continuous in $a$ if :

$$\forall\, \epsilon > 0, \exists\, \eta > 0, \forall\, x \in E, d(x, a) \leq \eta \implies d'(f(x), f(a)) \leq \epsilon$$

**Proposition 2.2.** *Let* $f : (E, d) \longrightarrow (E', d')$ *be a map between two metric spaces.* $f$ *is continuous in* $a \in E$ *if and only if for any sequence* $(x^n)$ *that converges to a, the sequence* $f(x^n)$ *converges to* $f(a)$*.*

As mentioned, the aim of what preceded was to help us recall notions needed to tackle discrete dynamical systems, which are in the midst of our core contribution. So let us get into the discrete case !

## 2.2 Discrete dynamical systems

Let $f : X \to X$ be a map from a topological or metric space $X$ to itself. The sequence of iterates considered is defined through the recurrence relation:

$$\begin{cases} x^0 \in X, \\ x^{n+1} = f(x^n),\ n \in \mathbb{N} \end{cases} \tag{2.1}$$

The behaviour of these iterates relies on the function $f$ and the space where the iterations occur.

A **discrete dynamical system** is a pair $(X, f)$ consisting of

- a non-empty topological space $(X, \tau)$, called *phase space*;
- A function $f : X \to X$, called *transition function*.

The function $f$ can in some cases be reversed, allowing to return back in time, making it a question of reversibility. So a discrete dynamical system $(X, f)$ is **reversible** if $f$ is a bijection.

As will be seen, generally when we have to show the chaoticity of a dynamical system, it is usually a matter of studying a simpler system, which is equivalent in a sense to be specified. This is the notion of **topological equivalence**.

Two dynamical systems $(X, f)$ and $(Y, g)$ are topologically equivalent if there exists a homeomorphism $h : X \to Y$, i.e. a bicontinuous bijection, such that $g \circ h = h \circ f$. In other terms, the following diagram commutes:

Topological equivalence is really an equivalence relation on the set of discrete dynamical systems, preserving topological properties. The two systems have the same dynamics, in the sense that dynamical objects, such as orbits, fixed points, or periodic ones will be preserved up to homeomorphism. So let us address the way a point $x$ evolves during time. It is a matter of **orbit**:

- Given a $x \in X$, the set $\{f^{(n)}(x),\ n \in \mathbb{N}\}$ is called the **orbit** of $x$, and marked as $\gamma_x$.

Chaos theory seeks to discover if the behaviour of discrete dynamical systems can be predicted, i.e. if we can pinpoint the orbit $\gamma_x$ of a given point $x$. In this spirit, the points whose behaviour is the simplest to apprehend are periodic and fixed points:

- A point $p \in X$ is said to be **periodic** of period $k$ or $k$-periodic if $k$ is a positive integer such that $f^{(k)}(p) = p$, and $\forall\ h \in \{1, \ldots, k-1\}, f^{(h)}(p) \neq p$. The orbit $\gamma_p$ is said $k$-periodic or $k$-cycle. Obviously a $k$-periodic point of $f$ is a fixed point of $f^k$.

We refer to the set of $k$-periodic points of $f$ by $Per_k(f)$, and the set of periodic points of any period is denoted by $Per(f)$.
Periodicity can take place after a transition phase more or less long, and the simplest orbits are present in the fixed points, which brings us to the following:

- A point is **ultimately periodic** if there exist two integers $n$ and $p$ such that $f^{(n+p)}(x) = f^{(p)}(x)$. Let $n_0$ be the smallest $n$ satisfying this. The set $\{x, f(x), \ldots, f^{(n_0)}\}$ is called *transient* of $x$, and $n_0$ the *length of the transient*.
- Periodic points of period 1 are called **fixed points** of $f$, and ultimately periodic points of period 1 are also called ultimately fixed points.

It is essential for the course of chaos discussion that we are about to undertake to mention the Sharkovsky order and theorem, dealing with constraints on the presence of periodic points. It is also worth mentioning that this theorem does only require continuity, which is not always the case in the majority of the problems.

**Theorem 2.3** (Sharkovsky [27])**.** *Let* $f : \mathbb{R} \longrightarrow \mathbb{R}$ *be continuous and consider the following order for natural numbers:*

$$1 > 2 > 2^2 > 2^3 > \cdots > 2^n > \cdots > 7 \cdot 2^n > 5 \cdot 2^n > 3 \cdot 2^n > \cdots > 7 \cdot 2 >$$
$$5 \cdot 2 > 3 \cdot 2 > \cdots > 9 > 7 > 5 > 3.$$

*If $f$ has periodic points of period $l$ and $l < k$, then $f$ has also a periodic point of period $k$.*

Following this theorem, it was nearly a decade after, that a special case arose for continuous functions, here below announced as a corollary, and being the first to use the term *chaos* in a mathematical sense:

**Corollary 2.4** (Li & Yorke theorem [28])**.** *Let $f$ be continuous on $[a, b]$, its range contained in $[a, b]$. If $f$ has a 3-periodic point then $f$ has $n$-periodic points for all positive integers $n$.*

In fact, if $f$ has a periodic point of order 3, it would have periodic points of any given period according to Sharkovsky theorem. It implies a certain type of chaos if we take in consideration the multiplicity of periods as a standard.

To bring an element of regularity to chaos, and some first-hand order, we make use of the density concept in topology.

- Let $X$ be a topological space and $A$ a subset of $X$. $A$ is said to be **dense** in $X$ if for every element $x$ of $X$, every neighbourhood of $x$ contains at least a point of $A$.

In a sense, a subset $A$ of a topological space $X$ is dense reflects how for any point $x$ from $X$, a point of $A$ can be found as close to $x$ as possible.

## 2.3 Devaney Chaos

The definition of Devaney for chaos remains one of the most popular among the scientific community. Before to state it, a collection of topological notions have to be first introduced. We begin by the notion regularity in dynamical systems:

- A discrete dynamical system $(X, f)$ is said **regular** if the set of periodic points of $f$ is dense in $X$.

In other terms, in a metric space $(X, d)$, the dynamical system $(X, f)$ is regular if:

$$\forall\ x \in\ X, \forall\ \epsilon > 0, \exists\ p \in\ Per f(f),\ d(x, p) < \epsilon$$

As will be seens through this definition a certain orbit emerges from chaos.

A major topic to study in dynamical systems is characterizing their asymptotic behaviour. Thus, subsets of the phases space that stay stable under the action of the system, are essential in this case study.

- Let $(X, f)$ be a discrete dynamical system. A susbet $A$ of $X$ is said **positively invariant** if $f(A) \subset\ A$, and stricly positively invariant if equal. Obviously, $f^n(A) \subset A, \forall\ n \in\ \mathbb{N}$.

A positively invariant set is then a trap set, once in, it cannot leave.
A pair $(A, f)$ is then considered a sub-dynamical system of $(X, f)$. In other words, the study can be operated on positively invariant parts of $X$, since it is smaller, and maybe easier to manipulate. Some systems are simple to study, since it can be decomposed into positively invariant sub-systems, which is not the case for transitive systems. Decomposable systems are systems that have some specific covers, predictable under the action of $f$, as stated below:

- A discrete dynamical system $(X, f)$ is **decomposable** if there exists a finite open cover of $X$ containing at least two elements such that every open set of the cover be a positively invariant set of $f$.

The independent sub-dynamical systems then act on each element of the cover, and each part of the cover can be studied separately, deducing the complete behaviour of $(\mathbb{X}, f)$ and making the study of the system much simpler.

The opposite notion of decomposability is what we call transitivity, a main component in Devaney chaos.

### 2.3.1 Topological transitivity

- A discrete dynamical system $(X, f)$ is **transitive** if if any two non-empty open sets $A, B \subset X$ can be joined , i.e. $\exists\ k \in \mathrm{N}$ such that

$$f^{(k)}(A) \cap B \neq \emptyset$$

Equivalently, we have

$$\overline{\bigcup_{n \in \mathrm{N}} f^{(n)}(A)} = X$$

In the case where $X$ is compact, we have the following characterisation:

**Proposition 2.5** (Transitivity in compact spaces )**.** *In compact spaces, transitivity is equivalent to having a dense orbit.*

In a sense, since the space is relatively small (compact), transitivity can be seen as a point visiting almost all the space. In addition, in compact spaces, $f$ is imperatively an onto mapping: $f(X) = X$.

Another component of Devaney chaos is sensitivity to initial conditions, a notion strongly related to the concept of stability of orbits introduced here below: Let us start by defining what a stable point means:

- A point $x \in X$ is said to be **stable** if

$$\forall\ \epsilon > 0, \exists\ \delta > 0, \forall\ y \in X, d(x, y) < \delta \Longrightarrow \forall\ n \in \mathrm{N}, d(f^n(x), f^n(y)) < \epsilon$$

In other words, if $y$ is close to $x$, then the orbit of $y$ will be close to the orbit of $x$. In the neighbourhood of a stable point $x$, every point progresses in the same way, and if an error is produced in the initial condition, it would be guaranteed to have just a minimal error between the observed simulation and the theoretical progression. In the normal

course of this discussion, it is natural to define the opposite notion of stability:

$x$ is **unstable** if

$$\exists\ \epsilon_x > 0, \forall\ \delta > 0, \exists\ y \in\ X, \exists\ n \in\ \mathrm{N},\ d(x,y) < \delta \text{ and } d(f^n(x), f^n(y)) \geq \epsilon_x$$

A system with nothing but stable points must be predictable and is said stable, otherwise, it is unstable.

The notion of sensitive dependence on initial conditions is stronger than instability. Hereafter, $\epsilon$ does not depend on the considered point $x$:

### 2.3.2 Sensitivity to initial conditions

- A discrete dynamical system $(X, f)$ is **sensitive to initial conditions** if $\exists\ \epsilon >$ 0such that:

$$\forall\ x \in\ X, \forall\ \delta > 0, \exists\ y \in\ X, \exists\ n \in\ \mathrm{N},\ d(x,y) < \delta \text{ and } d(f^n(x), f^n(y)) \geq \epsilon$$

  we call $\epsilon$ the *sensitivity constant*.

A system is then sensitive if for each $x$, there exists points arbitrarily close to $x$, with their orbits respectively separated by at least $\epsilon$, during the system's progression. This $\epsilon$ is fixed once and for all, it is the same for every point $x$.

*Remark* 2.6. All points neighbouring $x$ are not necessarily distant by $\epsilon$ during the system evolution: it is enough if it exists at least one in each open ball centered in $x$.

The definition which is about to be announced in this part, is the most famous one amongst mathematical chaos definitions available. It is exactly going to be adopted in our work in the following chapters, to judge whether a discrete dynamical system is chaotic or not.

**Definition 2.7** ([24]). *A dynamical system $(X, f)$, or shortly $f$, is* ***chaotic*** *(on $X$) according to Devaney if it is sensitive to initial conditions, regular and transitive.*

*Remark* 2.8. Originally, Devaney stated that in order for the system to be chaotic, sensitivity to initial conditions is necessary, but later it was proved [26] this conditions is redundant:

**Proposition 2.9.** *If a dynamical system is transitive and regular, then it is sensitive to initial conditions, and consequently, is chaotic.*

According to Devaney, the unpredictability of a chaotic system is given through its sensitive dependence to initial conditions, on the other hand, transitivity does not allow to break down the system into two disjoint entities (subsystems), and finally one has an element of order being the regularity: it is indeed order within chaos ! Surprisingly, when $f$ is a continuous function on an interval of R, only transitivity is needed to show chaos:

**Proposition 2.10** ([29] )**.** *Let $I$ be not necessarily a finite interval and $f$ a continuous and transitive map on $I$. The periodic points of $f$ are dense in $I$ and then $f$ is chaotic.*

Practically, it is very rare to attempt proving these properties directly without resorting to a much simpler dynamical system, via topological equivalence:

**Proposition 2.11.** *If two dynamical systems $(X, f)$ and $(Y, g)$ are topologically equivalent and one is chaotic, then the other one is chaotic as well.*

In fact, such a result can be weakened as the transformation $h$ ensuring topological equivalence doesn't need to be invertible, and in this case we speak about *topological semi-equivalence*. Such a fact will be exploited when showing the chaoticity of the logistic function.

A pedagogical example is the well-known logistic map, which can be shown chaotic on $[0, 1]$ after bringing it to the well-known tent map via topological equivalence. Indeed, Baptista method (see end of chapter) took profitably this fact to conceive the first ever chaos-based cryptosystem, a method that cannot be improved in any way without a deep understanding of the logistic function.

### 2.3.3 Chaotic behaviour of the logistic map

Before tackling the logistic function, we first add some notions that will be needed during this analysis, namely attraction and stability in the sense of Lyapunov:

- Let $(X, d)$ be a metric space and $(X, f)$ a dynamical system. A fixed point $x_0$ is **attractive** if it admits a neighbourhood $U$ such that $\forall\ x \in\ U$, the sequence $(f^n(x))_{n\in\ \mathbb{N}}$ converges to $x_0$ when $n \longrightarrow +\infty$.

  The **stable set** $W^s(x_0) :=\ \{x \in\ X, f^n(x) \longrightarrow x_0\}$ is also called the **basin of attraction** of $x_0$.

When the function $f$ is differentiable, the following result, due to Ostrowski [88] and extended to the periodic case, will be of great help for us when studying fixed and periodic points of the logistic map:

**Theorem 2.12.** *If $f$ is a differentiable in a neighbourhood of a fixed (resp. $k$- periodic) point $x$ with $|f'(x)| < 1$ (resp. $|(f^k)'(x)| < 1$), then there exists a neighbourhood $U$ of $x$ with $U \subset\ W^s(x)$, and in the this case, $x$ is an attractive (periodic) point. If $|f'(x)| > 1$ (resp. $|(f^k)'(x)| > 1$), there exists a neighbourhood $V$ of $x$ such that for every $y \in\ V \setminus \{x\}$, there exists $n \in\ \mathbb{N}^*$ such that $f^n(y) \notin\ V$ (resp. $f^{nk}(y) \notin\ V$), and in this case, $x$ is said to be a repulsive (periodic) point.*

When $|f'(x)| \neq 1$ (resp. $|(f^k)'(x)| \neq 1$), the fixed (resp. $k$-periodic) point is said to be *hyperbolic*. It is relatively less sensitive to small changes in the function $f$, oppositely to *non-hyperbolic* points, which need additional informations to conclude about their nature.

#### Analysis of the logistic map

- The **logistic map** on figure 2.1, introduced by Verhulst [89], is defined on the interval $[0, 1]$ by:

$$f(x) = rx(1 - x)$$

where $r$ is a positive real parameter.

The corresponding discrete dynamical system reads as:

$$\begin{cases} x_0 \in [0,1], \\ x_{n+1} = rx_n(1-x_n), \ n \in \mathbb{N} \end{cases} \tag{2.2}$$

This a discrete-time analog of the logistic equation for population growth, where $x_n$ is the dimensionless measure of the population in the $n$-th generation and $r$ is the intrinsic growth rate.

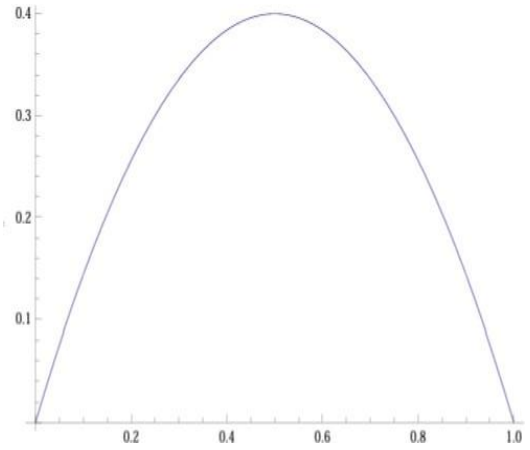



**Figure 2.1:** *The logistic map for* $r = 0.4$

- For $0 \leq r \leq 4$, the map is a parabola with maximum value of $\frac{r}{4}$ at $x = \frac{1}{2}$, as shown on figure 2.2 of the cobweb plot.

In the cobweb diagram, we plot a diagonal line $y = x$ and quadratic curve $y = rx(1 - x)$, and we connect each $x_n$ on diagonal and $x_{n+1}$ on the quadratic curve by a vertical line and each $x_n$ in the quadratic curve to a diagonal line by a horizontal line. In this way, each $y$-value of points on the quadratic curve $y = rx(1 - x)$ indicates the value of the next iteration of $x_{n+1}$.

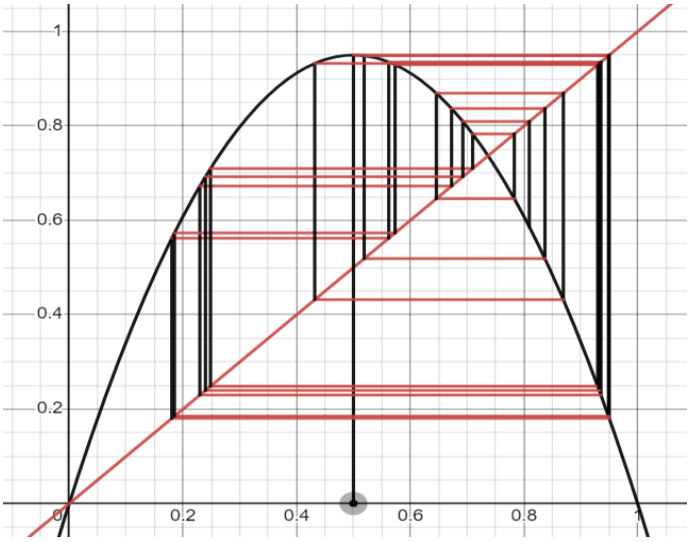



**Figure 2.2:** *The cobweb diagram of the logistic function for* $r = 3.8$ *and* $x_0 = 0.5$

Fixed points of $f$ are solutions of the equation $rx(1-x) = x$, so we have

$$x = 0 \text{ or } x = p_r = 1 - \frac{1}{r}$$

$$f'(0) = r \text{ and } f'(p_r) = 2 - r$$

Thus 0 is a fixed point for all $r$, and $p_r$ is a fixed point only if $r \geq 1$.

1. For $0 < r < 1$, 0 is attractive and $p_r$ repulsive.

2. For $r = 1$, a qualitative change in the dynamics of $f$ occurs, and we speak about *bifurcation*, 0 is non-hyperbolic, $W^s = [0, 1]$ and $W^s(\infty) = ]-\infty, 0[ \cup ]1, +\infty[$. Generally, for $r \geq 1$, $]-\infty, 0[ \cup ]1, +\infty[ \subset W^s(\infty)$.

3. For $1 < r < 3$: 0 is repulsive, $p_r$ is attractive, $W^s(0) = \{0, 1\}$ and $W^s(p_r) = ]0, 1[$.

4. For $r = 3$, a *period-doubling* bifurcation occurs and $p_r$ is (weakly) attractive ($|f'(p_r)| = 1$), 0 is always repulsive and $W^s(p_r) = ]0, 1[$.

5. For $3 < r < 1 + \sqrt{6}$, 0 and $p_r$ are repulsive with $W^s(0) = \{0, 1\}$ and $W^s(p_r)$ being infinite this time. Meanwhile, a 2-periodic orbit occurs and "almost" all points $x \in ]0, 1[$ are *asymptotic* to one of the points $p$ of this orbit ($\lim_{n\to+\infty} f^{2n}(x) = p$).

6. For $r = 1 + \sqrt{6}$, an another period-doubling bifurcation occurs and an attractive 2-periodic orbit is divided into an attractive 4-periodic orbit and a repulsive 2-periodic orbit.

7. For $1 + \sqrt{6} < r < 4$, changes occur quickly. We deduce the apparition of a 3-periodic orbit when $r > 1 + \sqrt{8}$, and then using Sharkovsky theorem there is periodic points of every order.

### Chaos in the logistic map

- For $r = 4$, a classical proof of the chaoticity of $f$ is given below:

First, let us consider the tent map $T$ defined over $[0, 1]$ by :

$$T(x) = \begin{cases} 2x & 0 \leq x \leq \frac{1}{2} \\ 2(1-x) & \frac{1}{2} \leq x \leq 1 \end{cases} \tag{2.3}$$

The cobweb plot of the tent map with the iterations starting from the seed $x_0 = 4 - \pi$ is shown on figure 2.3.

To begin with, we prove that $f$ is semi-equivalent to the tent map $T$, with the continu-

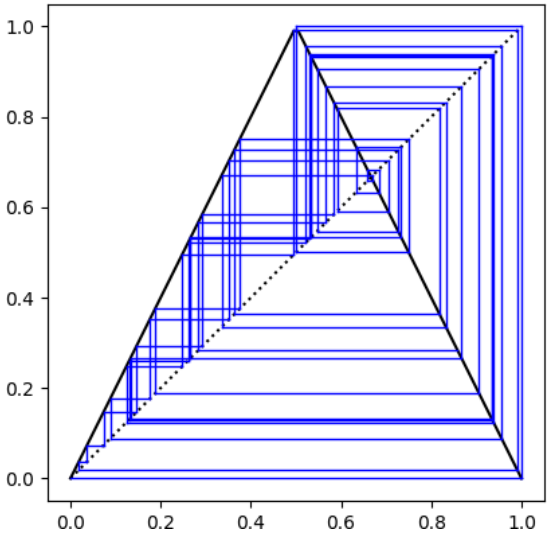



**Figure 2.3:** *Cobweb plot of the tent map*

ous surjection $h$:

$$h : [0,1] \dashrightarrow [0,1]$$
$$x \longrightarrow \sin^2(\frac{\pi}{2}x) = \frac{1}{2}(1 - \cos(\pi x))$$

- For $0 \leq x \leq \frac{1}{2}$:

$$\begin{aligned} h(T(x)) &= \frac{1}{2}(1 - \cos(2\pi x)) \\ &= \sin^2(\pi x) \\ &= 1 - \cos^2(\pi x) \\ &= 4 \frac{1}{2}(1 - \cos(\pi x)) \frac{1}{2}(1 + \cos(\pi x)) \\ &= 4h(x)(1 - h(x)) \\ &= f(h(x)) \end{aligned}$$

- For $\frac{1}{2} \leq x \leq 1$:

$$\begin{aligned} h(T(x)) &= \frac{1}{2}(1 - \cos(2\pi(1 - x))) \\ &= \frac{1}{2}(1 - \cos(2\pi - 2\pi x)) \\ &= \frac{1}{2}(1 - \cos(-2\pi x)) \\ &= \frac{1}{2}(1 - \cos(2\pi x)) \\ &= \ldots \quad \text{same as the case before} \\ &= f(h(x)) \end{aligned}$$

so $h$ and $T$ are semi-conjugated, and the following diagram commutes.

$$\begin{array}{ccc} [0,1] & \xrightarrow{f} & [0,1] \\ \downarrow h & & \downarrow h \\ [0,1] & \xrightarrow{T} & [0,1] \end{array}$$

It is now sufficient to show that the tent map $T$ is chaotic according to Devaney on $I = [0, 1]$.

Since we are working on an interval $I$ and $T$ is a continuous map already, using proposition 2.10 we only have to prove that $T$ is transitive.

In fact, let us prove that $f^k(U) \cap V = \emptyset$ for $k > 0$.
Let: $U = \left]0, \frac{1}{4}\right[$ and $V = \left]\frac{1}{2}, 1\right[$

$$f\left(\left]0, \frac{1}{4}\right[\right) = \left]0, \frac{1}{2}\right[, \quad f^2\left(\left]0, \frac{1}{2}\right[\right) = [0, 1]$$

so

$$f^2\left(\left]0, \frac{1}{2}\right[\right) \cap \left]\frac{1}{2}, 1\right[ \neq \emptyset$$

- For $r > 4$, we show that the set of points which are not in $W^s(\infty)$ is negligible with regard to the Lebesgue measure. In fact, this is a *Cantor set*, that is a compact and totally disconnected set, admitting each of its points as an accumulation point.

In what follows, further details are displayed with figures:

- For $1 \leq r \leq 3$: $x_n$ grows as $n$ increases, reaching a non-zero state, as seen on figure 2.4.

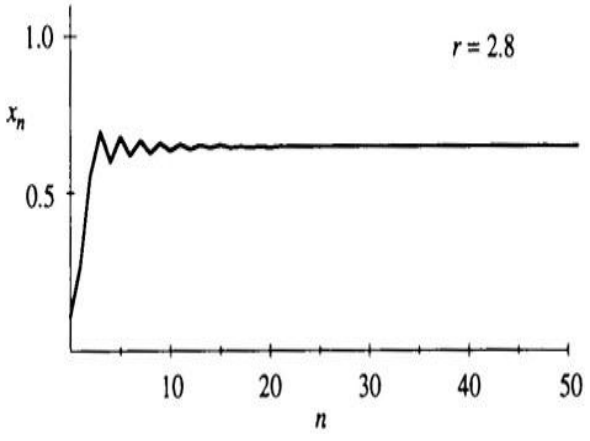



**Figure 2.4:** *Iterations of the logistic map for* $r = 2.8$

- For larger $r$ (e.g. $r = 3.3$) $x_n$ eventually oscilates as on figure 2.5 (2-cycle).

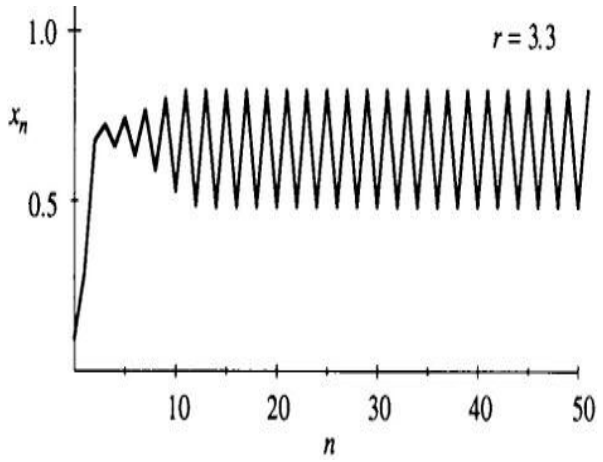



**Figure 2.5:** *Iterations of the logistic map for* $r = 3.3$

- At $r$ still larger (e.g. $r = 3.5$ on figure 2.6), $x_n$ approaches a cycle which repeats every 4 generations (4-cycle).

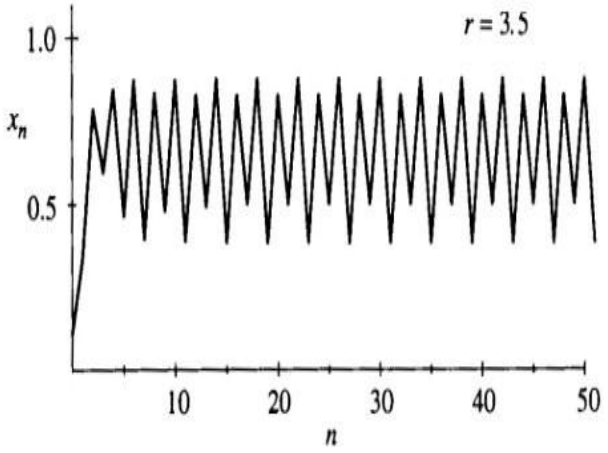



**Figure 2.6:** *Iterations of the logistic map for* $r = 3.5$

Further period doublings for period $8, 16, 32...$ occur as $r$ increases. Numerical experiments show that

$$\begin{aligned} r_1 &= 3 && \text{period 2 is born} \\ r_2 &= 3.449... && \text{period 4 is born} \\ r_3 &= 3.54409... && \text{period 8 is born} \\ r_4 &= 3.5644... && \text{period 16 is born} \\ &... \\ r_\infty &= 3.569946 && \text{period } \infty \text{ is born} \end{aligned}$$

- $r_n$ converges to a limiting value $r_\infty$.
- For large $n$, the distance between successive transitions shrinks by a constant factor, called the **Feigenbaum constant** [97]

$$\delta = \lim_{n\to\infty} \frac{r_n - r_{n-1}}{r_{n+1} - r_n} = 4.669...$$

Now for $r > r_\infty$, what happens is complicated, so for many values of $r$, the sequence $(x_n)$ never settles down to a fixed point or a periodic orbit (the long term behaviour is aperiodic, like shown on figure 2.7). The corresponding cobweb diagram is complicated as well as it can be displayed on figure 2.8, and the orbit diagram on figure 2.9 shows the long-term behaviour for all values of $r$ at once.. [96]

- At $r = 3.4$ the attractor is a 2-cycle.
- As $r$ increases, both branches split, giving a 4-cycle, i.e. a period-doubling bifurcation has occurred.
- A cascade of further period-doublings occurs as $r$ increases, until at $r = r_\infty \approx 3.57$, the map becomes *chaotic* and the attractor changes from a finite to an infinite set of points.

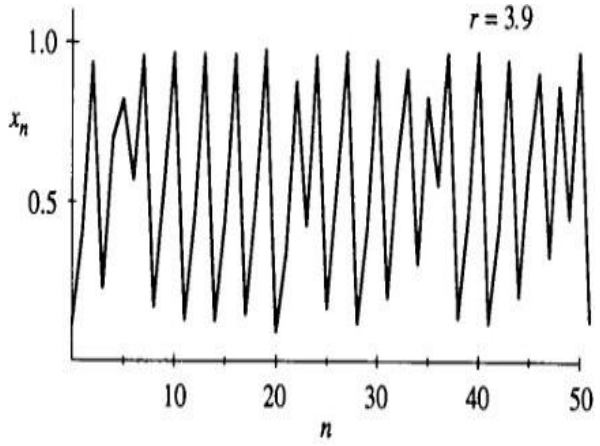



**Figure 2.7:** *Iterations of the logistic map for* $r = 3.9$

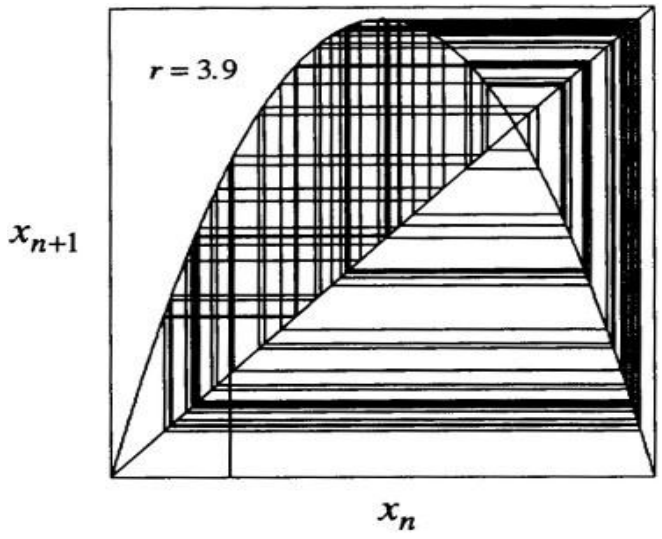



**Figure 2.8:** *Cobweb diagram of the logistic map for* $r = 3.9$

- For $r > r_\infty$, the orbit reveals a mixture of order and chaos, with chaotic clouds of dots.
- The large window near $r \approx 3.83$ contains a stable 3-cycle. A blow-up of part of this window shows that a copy of the orbit diagram reappears in miniature.

For $r$ just below the period-3 window, one finds this intermittency route to chaos diagram shown on figure 2.10, where black dots indicate part of the orbit which looks like an attractive 3-cycle.

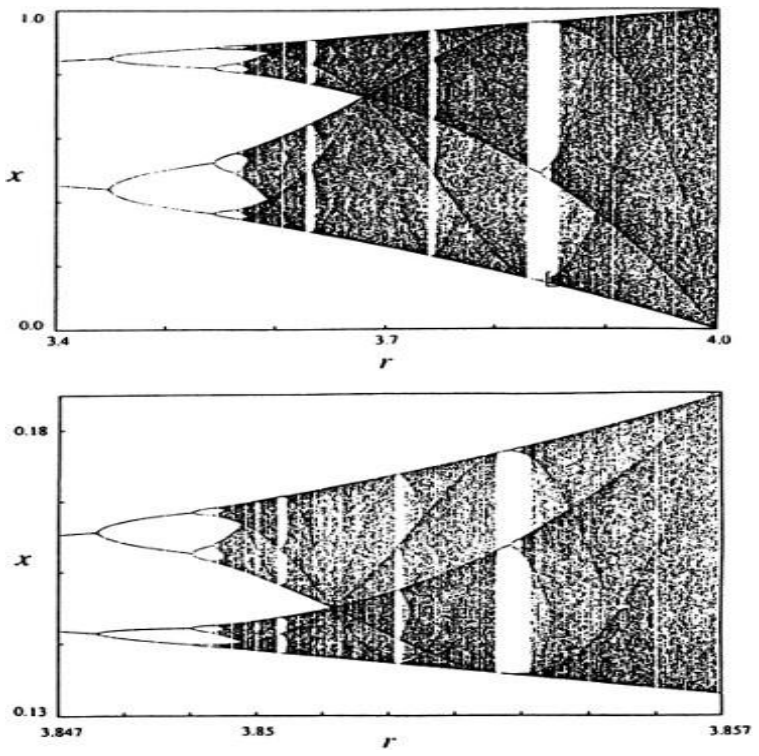



**Figure 2.9:** *The orbit diagram of the logistic map for* $r = 3.85$

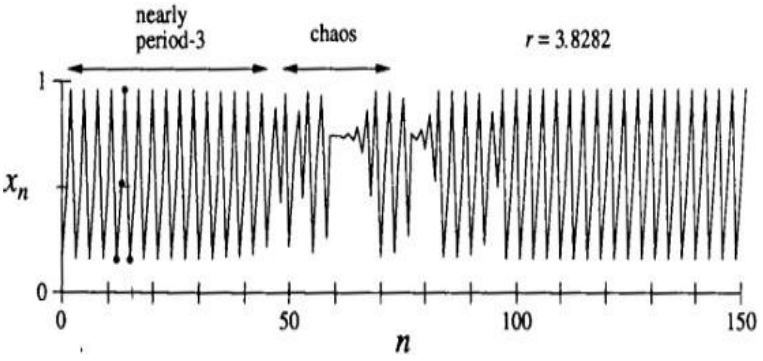



**Figure 2.10:** *Intermittency diagram for the logistic map for* $r = 3.8282$

This analysis of the logistic map is crucial for the understanding of one of the first chaos-based encryption methods, namely the Baptista method developed in a later section. But let us first introduce some fundamentals of cryptography.

## 2.4 Cryptography fundamentals

### 2.4.1 Cryptography prerequisites

In the following part, we are strictly focusing on introducing the reader into the main fundamentals of (chaos-based) cryptography [31], which will be put to use in the next chapters.
A gentle description of certain primitives, block ciphers, stream ciphers along with other notions, is introduced to ease the reader into this discipline.

The objective of cryptography is enabling two individuals or entities (say Alice and Bob) to communicate over insecure channels, to augment and save their privacy against potential adversaries (Oscar or Eve). Initially, Alice detains a *plaintext* (in any data format agreed upon). When she *encrypts* the plaintext via a *key* (generally public), it becomes the *ciphertext* transmitted to Bob over the above mentioned channel. Bob *decrypts* the message using his own secret key. This is the general framework for *asymmetric* or *public-key* cryptography, while *symmetric* cryptography uses the same key to encrypt and decrypt, and which can be exchanged using a public-key protocol. This notions can be formally written using this mathematical expressions:

**Cryptosystem components** A cryptographic system (or shortly a *cryptosystem*) is a quintuplet $(P, C, K, E, D)$ satisfying:

1) A finite set $P$ of possible **plaintexts**.
2) A finite set $C$ of possible **ciphertexts**.
3) The keyspace $K$, a finite set of possible **keys**.
4) For each $k \in K$, there is an encryption rule $e_k \in E$ and a corresponding decryption rule $d_k \in D$. Each $e_k : P \rightarrow C$ and $d_k : C \rightarrow P$ are functions such that $d_k(e_k(x)) = d_k(y) = x$, for every plaintext $x \in P$.

Obviously, the encryption function $e_k$ has to be injective. Schematically, this can be shown on figure 2.11.

In cryptography, each encryption algorithm falls within two categories [32]: first there

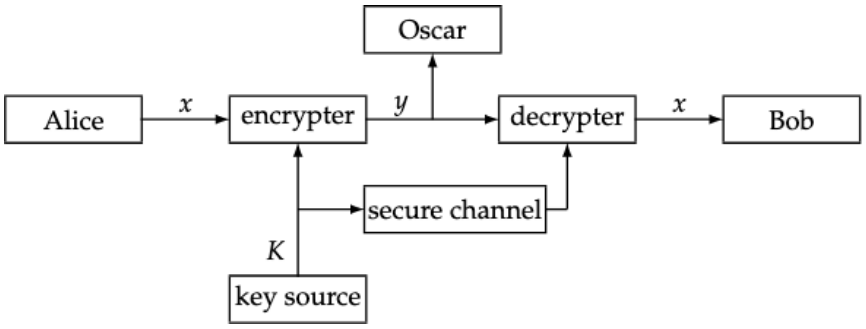



**Figure 2.11:** *The communication channel*

is *block ciphers*, where the text is divided into blocks of fixed size, each (de)encrypted once at a time, while in the second category *stream ciphers* use the key to form a *keystream*, having the same length as the text to be encrypted. Usually, the plaintext and the keystream are expressed in binary format and merged together via the *eXclusive OR* (XOR) function.

More precisely, for block ciphers, successive plaintext elements are encrypted with the same key $k$, starting from the block decomposition of the plaintext $x = x_1x_2..x_n$, we have the ciphertext $y = y_1y_2..y_n = e_k(x_1)e_k(x_2)..e_k(x_n)$, which is called a block cipher cryptosystem.

For stream ciphers, a keystream is generated $k = k_1k_2..k_n$ to encrypt a plaintext string $x_1x_2..x_n$ using $y = y_1y_2..y_n = e_{k_1}(x_1)e_{k_2}(x_2)..e_{k_n}(x_n)$, which is is called a stream cipher cryptosystem.

As for public-key encryption, it consists of a key pair ($k_1$, $k_2$), one public and the other private such that : $e_{k_1}(x) = y$ and $d_{k_2}(y) = x$. The private-key encryption is the case where both keys as identical and kept private, such that : $e_k(x) = y$ and $d_k(y) = x$.

The justification of the public-key cryptography goes back to the founding work of

Diffie and Hellmann [35], from which emerged RSA [34] [33] and ElGamal cryptosystems [113] as more concrete realisations:

### 2.4.2 The Diffie-Hellman problem

As a matter of fact, in the 1970s, Merkle [36] suggested a key distribution system showing it was possible to select a key over open communications channels, in such a fashion that communications security can be maintained, but this proposal is less efficient in transmission cost than what Diffie and Hellman later on suggested. This problem would be useful in understanding RSA [34] which is used for key exchange in the next chapter covering our paper. It is also very useful to go over CDH (Computational Diffie-Hellmann) and DDH (Decisional Diffie-Hellmann), to get a sense of these couple of variations essential to comprehending the contribution in chapter 5.
Diffie and Hellman [35] devised a cryptosystem with two distinct keys, a public one used to encrypt the message and a private one to decrypt it, the *Discrete Logarithm Problem* (DLP) [33] being the core of this idea. Thus, they developed systems in which two parties communicate solely over a public channel, using only publicly known techniques to create a secure connection. This was made possible due to the apparent difficulty of computing logarithms over a finite field with a prime number of elements.

Let us consider defining a general version of the DLP in a subgroup $< \alpha >$ of a group$(G, .)$. Let $(G, .)$ be a multiplicative group, $\alpha \in G$ an element of order $n$ and $\beta \in < \alpha >$. There exists a unique $a \in ¢0, n - 1)$ such that

$$\alpha^a = \beta$$

with $a = \log_\alpha(\beta)$ called the **discrete logarithm** of $\beta$.

The Diffie-Hellman problem is expressed for an abelian group, for instance the multiplicative group of non-zero integers modulo a large prime $p$, but it may become a harder problem when expressed over other groups.

**Diffie-Hellman key exchange protocol** The public parameters are the group $(G, .)$ and $\alpha \in G$ of order $n$. Here is how the key agreement is performed:

1) Alice chooses randomly $a \in ¢0, n - 1)$, and computes :

$$b = \alpha^a$$

then sends $b$ to Bob

2) Bob chooses randomly $\gamma \in ¢0, n - 1)$, and computes

$$\delta = \alpha^\gamma$$

then send $\gamma$ to Alice.

3) Alice computes

$$k = \delta^a$$

and Bob

$$k = b^\gamma$$

ending up with the same keys.

There are some serious limitations of this method, amongst which, the lack of authentication. The Diffie-Hellman scheme by itself is vulnerable to the Man in the Middle attack (MITM), where a malicious attacker is positioned between two communicating parties, hence, it should be used in conjunction with a recognized authentication method, for example digital signatures, to verify the identities of the users over the public communications.

Among variations of Diffie-Hellmann that exist, we list CDH and DDH [37]:

**Computational Diffie-Hellman (CDH)** Let $(G, .)$ be a multiplicative group, $\alpha \in G$ an element of order $n$ and $\beta, \gamma \in < \alpha >$.

$\exists \ \delta \in < \alpha >$ such that

$$\log_\alpha \delta \equiv \log_\alpha(\beta) \log_\alpha(\gamma) \ [n]$$

i.e. given $\alpha^b$ and $\alpha^c$, $\alpha^{bc}$ exists.

The problem here is finding $\delta$, $\alpha^{bc}$.

To solve this problem for either exponent, the discrete logarithm problem is impractical for some groups.

**Decisional Diffie-Hellman (DDH)** Let $(G, .)$ be a multiplicative group, $\alpha \in G$ an element of order $n$ and $\beta, \gamma, \delta \in < \alpha >$.

Do we have the following assumption below ?

$$\log_\alpha \delta \equiv \log_\alpha(\beta) \log_\alpha(\gamma) \ [n]$$

i.e. given $\alpha^b$ and $\alpha^c$ and $\alpha^d$, find out if the assumption below is correct:

$$d \equiv bc \ [n]$$

DDH version is weaker than CDH, since solving CDH allows you to solve the DDH question . DDH remains hard over some groups, such as certain elliptic curves, but not over other groups like multiplicative groups *modp*.

Another pillar of asymmetric encryption, and maybe the most famous one based on the same mathematical problem (DLP) of Diffie-Hellman, is RSA (by Ronald Rivest, Adi Shamir and Leonard Adleman) [34] to be covered next:

### 2.4.3 RSA cryptosystem

RSA is an encryption algorithm having two different keys (public and private), using separate keys for encryption and decryption.

Let $n = pq$, where $p, q \in \mathrm{P}$, the set of prime numbers. Let $\mathrm{P} = \mathrm{C} = \mathrm{Z}_n$, with $\phi(n) =$

$(p-1)(q-1)$ being the *Euler function* of $n$, and let us define the key space

$$\mathrm{K} = \{(n, p, q, a, b) : ab \equiv 1\ [\phi(n)]\}$$

For $\mathrm{K} = (n, p, q, a, b)$, the encryption is

$$y = e_{\mathrm{K}}(x) \equiv x^b\ [n]$$

and decryption

$$d_{\mathrm{K}}(y) = y^a\ [n] = x^{ab}\ [n] = x\ [n]$$

$x, y \in Z_n$.

$(n, b)$ form the public-key and the secret key is $(p, q, a)$.

The drawbacks of RSA is only being able to encrypt data to a maximum amount equal to your key size, minus any padding and header data, so RSA is not often capable of directly encrypting files; one has to use another key size or symmetric encryption.. Actually, this type of algorithms encrypts a random key with RSA, then encrypts the data with a symmetric encryption algorithm, which is not limited in size. In addition, RSA needs much computer power, and is slow in terms of key generation.

In 1985, based on the same Diffie-Hellmann key exchange defined before, and with a slight twist different than RSA, ElGamal suggested a public-key encryption algorithm shown below:

### 2.4.4 ElGamal cryptosystem

This cryptosystem is based on the difficulty of finding discrete logarithm in a cyclic group (group that can be generated by a single element) that is, even if we know $g^a$ and $g^k$, it is very difficult to compute $g^{ak}$.

When Alice wants to communicate a message $m$ to Bob:

1. Bob generates public and private keys:

- Bob chooses a very large number $q$ and a cyclic group $F_q$,
- from the cyclic group $F_q$, he chooses any element $g$ and an element $a$ such that $gcd(a, q) = 1$,
- then he computes $h = g^a$,
- Bob publishes $F_q$, $h = g^a$, $q$, and $g$ as his public key and retains $a$ as private key.

2. Alice encrypts data using Bob's public key :
   - Alice selects an element $k$ from cyclic group $F_q$ such that $gcd(k, q) = 1$,
   - then she computes $p = g^k$ and $s = h^k = g^{ak}$,
   - she multiples $s$ with $m$,
   - then she sends $(p, ms) = (g^k, ms)$.
3. Bob decrypts the message :
   - Bob calculates $r = p^a = g^{ak}$,
   - he divides $ms$ by $r$ to obtain $m$ as $s = r$.

In top of being slow than other encryption algorithms that provide fast processing speed, ElGamal requires larger key sizes to achieve the same level of security as other algorithms. Anyway, for us, algorithms such as RSA or ElGamal will be only used as protocols to exchange the secret key in chaos-based cryptography. As will be seen, the role of such a key can be played by the seed of a chaotic recurrent sequence, like the logistic one.

## 2.5 Chaos-based cryptography

After being well acquainted with modern cryptography, and surely after knowing some of its limitations, it is time to be introduced to chaos-based cryptography [75] perhaps one of the less influential fields in industrial cryptography, but still the one with the utmost importance in terms of meaningful mathematical surplus. The aim

throughout the next two chapters is to concretise it into our proposed algorithms, but first let us take a dive into this world before moving into other details.

### 2.5.1 Chaos-based encryption techniques

Unlike standard encryption mechanisms, many approaches are applied in this field to encrypt data using a chaotic process to initiate the cryptosystem, among these techniques we mention:

a) **Analogue chaos-based cryptosystems** [76] the plaintext is masked with a chaotic signal at a physical level of the communication channel. The optical communication devices is controlled to produce a chaotic waveform that *modulates* the message in a secure way for transmission. Bob receives the signal and *demodulates* it using a chaotic *synchronization techniques*.

b) **Digital chaos-based cryptosystems** [77] the iterative computations of chaotic functions generate digital signals, then substitution and mixing are applied to mask the plaintext, thus using the algorithms initial conditions as control parameters (secret keys).

c) **Digital stream ciphers implemented with chaotic Pseudo-Random Bit Generation (PRBG)** [78]: building a chaotic PRBG requires a $N$-dimensional deterministic discrete-time dynamical system as an iterative map $f : \mathrm{R}^N \rightarrow \mathrm{R}^N$ shown below:

$$x_{k+1} = f(x_k, \Gamma)$$

where $k = 0, 1, ..., n$ is the discrete time, $\Gamma$ the vector of parameters, $x_0$ initial condition and $x_1, ..., x_n$ states of the systems during time. To transforms the states of the system into binary, different numerical algorithm designs are used in chaotic PRBGs, from which we cite the following:

- From each state, bits are extracted along the chaotic orbits.

- The phase space is divided into $r$ sub-spaces, finally the output is a binary $i = 0, 1, ..., r$, that is if the chaotic orbit falls whitin the $i$-th subspace.

- The combination of at least two chaotic systems to generate the pseudo-random numbers as an output.

The keystream is formed of the sequence of generated binaries. The encryption is performed through adding the keystream to the output gotten before via the XOR operation. The initial conditions and the vector of parameters are used as the secret key.

Different perspectives were presented tying cryptography to chaos and how it could be represented or interpreted, among these, Shannon who pointed out this link between the two disciplines [117].

**Shannon's perspective:** In 1948, Shannon stated in his paper [116]: *A more detailed analysis shows that if we assume the constraints imposed by the language are of a rather chaotic and random nature, large crossword puzzles are just possible when the redundancy is 50%*. If one considers a round inside a cryptosystem, it offers us the output *diffusion* (butterfly effect) when applying a slight change in the parameters of the input, and complies with *confusion* (transitivity) since no matter how much the input changes the output has the same distribution. The diffusion property also enables a deviation in a local region to affect the whole phase space. The parameters of the chaotic map may represent the key of the encryption algorithm. However, chaos theory and cryptography differ in a detail: encryption algorithms operate on finite sets, while chaos theory does it with the set of real numbers.

One of the public-key cryptosystems that possess chaotic properties is based on Chebyshev polynomials, which makes it suitable for use in both encryption and digital signature:

### 2.5.2 The Chebyshev polynomials

Chebyshev polynomials [109] (of the first kind)

$$T_n : \mathrm{R} \longrightarrow \mathrm{R}$$

of degree $n$ are defined as follows:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \qquad n \in \mathrm{N}^*,$$

with $T_0 = 1$ and $T_1 = x$.
It is well established that Chebyshev polynomials possess the *semi-group property*

$$T_r(T_s(x)) = T_{rs}(x)$$

for any integers $r$ and $s$.
Hence, these polynomials commute under composition

$$T_s(T_r) = T_r(T_s)$$

The interval $[-1, 1]$ is invariant under the action of the map $T_n : T_n([-1, 1]) = [-1, 1]$. Therefore, the Chebychev polynomial restricted to the interval $[-1, 1]$ is the chaotic map for all $n > 1$ having the unique absolutely continuous invariant measure

$$\mu(x)dx = \frac{dx}{\pi\sqrt{1 - x^2}}$$

with positive Lyapunov exponent $\lambda = \ln n$ (The Lyapunov exponent is a logarithmic measure for the mean expansion rate per iteration, i.e. per unit time, of the distance

between two infinitesimally close trajectories; when positive, it refers to a chaotic system in the sense of Lyapunov). In the case of $n = 2$, the Chebyshev map becomes the logistic map studied before.

The public-key encryption based on Chebyshev polynomials is nothing but a general perspective of the ElGamal cryptosystem:

**The Chebyshev polynomials-based cryptosystem algorithm:** This cryptosystem consists of three phases [75], the first one being the key generation, where Alice follows the steps below as mentioned in the algorithm 1.

**Algorithm 1:** Key Generation

1. Generate a large integer $n$.
2. Choose a random number $x \in [-1, 1]$.
3. Compute $T_n(x)$.
4. Set the public key to $(x, T_n(x))$.
5. Set the private key to $n$.

The encryption mechanism has six components presented in the algorithm 2, where Bob follows those steps to encrypt his message.

**Algorithm 2:** Encryption

1. Get the public-key $(x, T_n(x))$, set by Alice.
2. The message is formatted as a number $M \in [-1, 1]$
3. Generate a large integer $m$.
4. Compute $T_m(x)$.
5. Compute $T_{m.n}(x) = T_m(T_n(x))$. Compute $X = M.T_{m.n}(x)$.

   Send Alice the ciphertext $C = (T_m(x), X)$

The decryption is resumed in two steps, that Alice follows in algorithm 3 to retrieve the plaintext from the ciphertext.

**Algorithm 3:** Decryption

1. Compute $T_{m.n} = T_n(T_m(x))$, using the private key $n$.
2. Compute $M = \frac{X}{T_{m.n}(x)}$, thus getting the message $M$

The Chebyshev polynomials-based cryptosystem is correct because of the semi-group property, so the encryption allows

$$X = M.T_m(T_n(x))$$

and, since those polynomials commute under composition,

$$X = M.T_n(T_m(x))$$

thus,

$$M = \frac{X}{T_{m.n}(x)}.$$

Chebychev polynomials-based cryptosystems as presented here are shown not to be secure, since an adversary can efficiently recover the plaintext from a given ciphertext. An alternative of this is what is called *modified Chebyshev polynomials*, which are both secure and practical, and for which RSA and ELGamal are just particular cases. Fore more details, see [75], chapter 2.

Another independent category chaos-based cryptography is summarised in the following:

### 2.5.3 PRNG-based cryptosystems

The idea consists of generating in a pseudo-random manner a sequence of bits using any chaotic process, then proceeding to a XOR with the plaintext, which is supposed to be already reduced to a binary format. An interesting example of a PRNG-based encryption using a strange attractor is described in the following:

**Lorenz-based PRNG** The well-known Lorenz system [14] is given by the system of continuous differential equations

$$\frac{dx_1}{dt} = \alpha x_1 + \alpha x_2,$$
$$\frac{dx_2}{dt} = -x_1 x_3 + \beta x_1 - x_2,$$
$$\frac{dx_3}{dt} = x_1 x_2 - \rho x_3,$$

where $\alpha$, $\beta$ and $\rho$ are positive control parameters. The Lorenz system is shown to be chaotic for the set of parameters $\alpha = 10$, $\beta = 28$, and $\rho = \frac{8}{3}$.
This system can provide an additional layer during the substitution phase in encryption algorithms [110] or PRNGs, where once considering a scrambled image, the current pixel of that image is multiplexed along the last encrypted pixel (multiplexers add nonlinearity, delay is present as well to improve the statistical test of encryption, since pixels are affected by the previous ones), then it is XORed with the attractor's output. It is recommended to choose 8 bits from the least significant part of each output to ensure the chaocity, after that the output of the Lorenz attractor is mapped to the range from 0 to 255, as shown below:

$$x = mod\ [E(|x_l| \times s),\ 256]$$
$$y = mod\ [E(|y_l| \times s),\ 256]$$
$$z = mod\,[E(|z_l| \times s), 256]$$

with $x_l, y_l, z_l$ the output of the Lorenz attractor, $s$ the scaling factor (here chosen as $10^{12}$ for the selected bits to be highly chaotic), $E(.)$ being the integer part of the number, and *mod* is the remainder .

The third major axis in chaos-based cryptography can be couched in symbolic dynamics, within a generalised version of Baptista's method. In fact, In the last three decades, many different chaotic encryption schemes have been proposed, but in 1998, Baptista [118] proposed the most used symmetric cryptosystem (with regards to chaos) based on ergodic property of the chaotic logistic map, which has attracted much attention from many scholars:

### 2.5.4 (Generalised) Baptista's cryptosystem

Let us consider a chaotic map $F$ on a compact $K$, a bijection $h$ between the partitioned space $K$ and a finite set of symbols $S$, associating to each symbol $s_i$ an element $P_i$ of the partition. Symbolically:

$$F : K \longrightarrow K \quad \text{chaotic,}$$

$$h : S \rightarrow K \quad \text{bijective}$$

$$s_i \longrightarrow h(s_i) = P_i,$$

$$K = \cup_{1 \leq i \leq n} P_i$$

To encrypt a plaintext $s_{i_1} s_{i_2} \ldots s_{i_m}$, which is a finite sequence of unnecessary different symbols, we choose randomly an initial value (the secret key) $k \in K$ and compute the number of iterations $N_1$ of $F$ necessary to reach the subset $h(s_{i_1})$, then starting from $F^{N_1}(k)$ as the new value, we compute the number of iterations $N_2$ to reach $h(s_{i_2})$, and so on. The ciphertext is the string $N_1, N_2 \ldots N_m$. To decrypt, we have just to iterate $F$ according to the integers $N_i$ and to determine the corresponding $h^{-1}(P_i)$.

A classical example as presented originally by Baptista uses the logistic function for $F$, with $r \in [3.57, 4]$, the interval $[0, 1]$ for $K$, and the set of ASCII characters for $S$:

- Let there be $s$ characters to encrypt, the given interval $I = ]0, 1[$ is being divided into $s$ subintervals of length: $\epsilon = \frac{x_{max} - x_{min}}{s}$ such that each interval is associated to a character.

- As for the choice of the interval $[x_{max}, x_{min}]$, it is being dicted by the probability density which implies where the initial value $x_0$ belongs, hence the need for the following plot on figure 2.12.

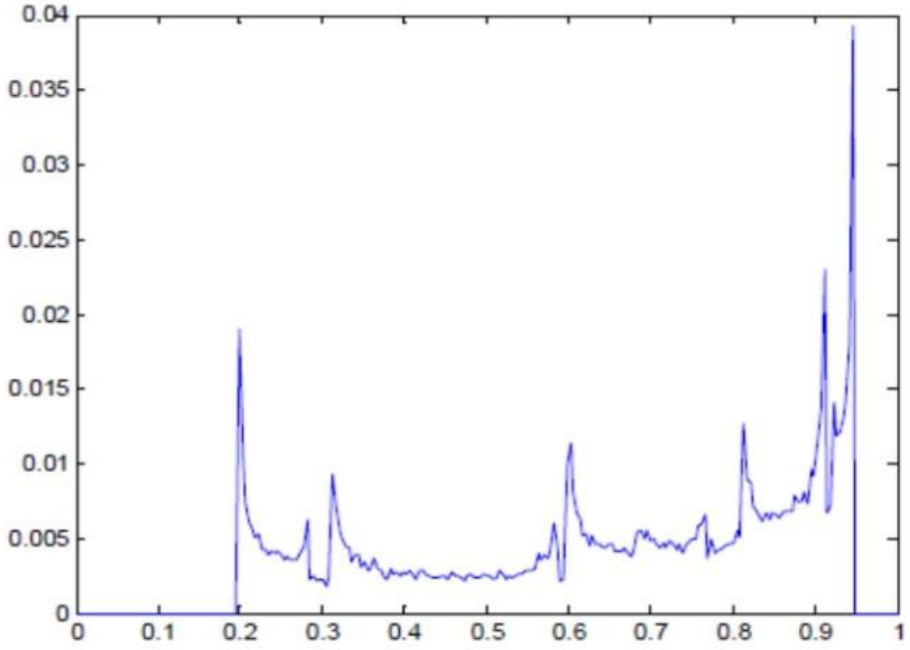



**Figure 2.12:** *Probability density*

- As can be seen, there is high probability between values 0.2 and 0.8. Encryption scheme [118] can be described in the following steps:

  0) Choose a character to encrypt.

  1) Create $s$ intervals corresponding to $s$ characters.

  2) Pick $I_\epsilon$ the interval corresponding to the chosen character.

  3) Choose a secret key $(x_0, r)$.

  4) Apply the logistic map to $x_0$.

  5) Count the number of iterations $N$.

  6) Stop when value of $x_N \in I_\epsilon$.

  7) The integer $N$ is the cipher character.

As an example of text encryption, we choose:

- Initial value: $x_0 = 0.43203125$
- $r = 3.88$
- $s = 256$

- $\epsilon = \frac{0.8 - 0.2}{256}$

Let us consider the following plaintext, which is going to be encrypted:

**"this is the text to be encrypted by the baptista method"**

After transforming the obtained integers into characters, these results are represented in *base64* (group of binary-to-text encoding schemes that represent binary data, more specifically, a sequence of 8-bit bytes, in sequences of 24 bits that can be represented by four 6-bit Base64 digits.) or *unicode* (format processed and stored as binary data using one of several encodings) format, and the final output of the ciphertext is:

**"MTE3MywgNDAzLCA0OTAsIDExMjUsIDMz**

**MC wgNjM1LCAxNDk sIDExMywgNDk5LCA**

**yMSwgMzA3LCA4MDksIDE1NDMsIDE0ODc**

**sIDk1MiwgMTU3NiwgNzIzlCAxNTksIDcxMi**

**wgMzU1LCAyMjM0LCAxMzAyLCAzODUsID**

**EwOSwgMjE1LCAxN"**

The values change when initial conditions are changed as displayed on table 2.1 showing sensitivity to initial conditions:

| | KEYS | | | |
|---|---|---|---|---|
| | $b, x_0$ | $b, x_0$ | $b', x_0$ | $b, x_0'$ |
| | $x_0 = 0.43203125$ $b = 3.88$ | $x_0 = 0.43203125$ $b = 3.88$ | $x_0 = 0.43203125$ $b = 3.7$ | $x_0 = 0.48203125$ $b = 3.88$ |
| M | 719 | M | ÿ | ÿ |
| a | 117 | a | ÿ | Ÿ |
| s | 68 | s | ÿ | \x9f |
| t | 732 | t | ÿ | Ÿ |
| e | 201 | e | \x8c | ñ |
| r | 1428 | r | ª | ¡ |
| | 316 | | & | Ÿ |
| S | 173 | S | 0 | Ÿ |
| D | 1396 | D | x | ÿ |
| B | 950 | B | ù | \x1a |
| D | 594 | D | à | Ÿ |
| | 484 | | ÿ | \x9f |
| - | 5701 | - | ÿ | Ÿ |
| | 252 | | \x18 | Ÿ |
| M | 1851 | M | í | \x83 |
| 1 | 75 | 1 | ÿ | ÿ |

**Table 2.1:** *Sensitivity to initial conditions with* $(b, x_0)$ *being the parameters* $(r, x_0)$ *in the logistic function*

Let us consider the clear image on figure 2.13, which is going to be encrypted. The cryptosystem is based on the same chaotic properties provided by the previous text encryption method, only this time an image is constructed of pixels and each pixel is formed by three components: R (Red), G (Green) and B (Blue), which will be encrypted using the Baptista algorithm. The clear image of Lena is being defined as a matrix of 512 rows and 512 columns.

**Figure 2.13:** *Lena's clear image*

The encrypted image is shown on figure 2.14.

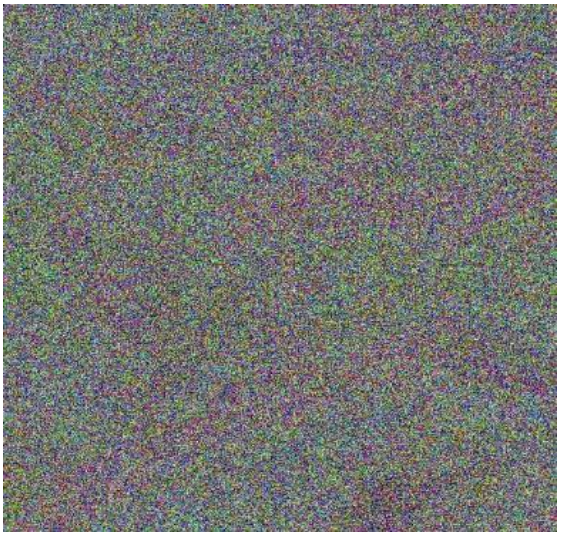


**Figure 2.14:** *Lena's encrypted image*

For transparency, Alvarez et. al. [119] examined encryption schemes of Baptista, and underlined its deficiencies proposing different ways to enhance its security. They found three types of cryptanalysis attacks, *one-time pad attacks* (chosen plain text), *entropy attack* and *key recovery attacks*. Also, a major draw-back of this method is the exponential decay of the repartition of the ciphertext values.

## 2.6 Conclusion

In this chapter, we explored the fascinating world of discrete dynamical systems, chaos theory, and their applications in cryptography. We began by introducing the concept of discrete dynamical systems, which are mathematical models that evolve over discrete time steps. We discussed the key elements of such systems, including the state space, the time evolution function, and the importance of initial conditions.

Next, we delved into chaos theory, which studies the behaviour of dynamical systems that exhibit sensitivity to initial conditions, among others. Chaos theory has revolutionised our understanding of complex systems, revealing their inherent unpredictability and sensitivity to even tiny perturbations. We examined the fundamental characteristics of chaotic systems, such as the butterfly effect, transitivity, and sensitive dependence on initial conditions. Understanding these properties is crucial for the development of chaos-based cryptographic algorithms.

The emerging field of chaos-based cryptography provides an alternative approach based on the unpredictability of chaotic dynamics. One of the key advantages of chaos-based cryptography is its resistance to certain types of attacks, including brute-force attacks and differential attacks. The inherent complexity and sensitivity to initial conditions in chaotic systems make them suitable for generating secure cryptographic keys and ensuring the confidentiality and integrity of data transmission.

Nevertheless, chaos-based cryptography is vulnerable to behaviour analysis attacks and guess of the initial condition [120]. Hence, in these chaotic cryptosystems all the following four types of cryptanalysis attacks are possible: cipher text only attack, chosen cipher text attack, chosen plain text and known plain text.. This motivates us to push forward trying to improve these methods, by suggesting other novel schemes.

# Chapter 3

# Egyptian fractions and Secure Multi-Party Computation

> Lots of people working in cryptography have no deep concern with real application issues. They are trying to discover things clever enough to write papers about.
>
> *Whitfield Diffie*

In this chapter, we provide necessary fundamentals in Egyptian fractions and their properties, on one hand, which will be needed in understanding the first paper. On the other hand, we initiate the reader into trapdoor functions and indistinguishability, inside a context of multiple challenges. Another preliminary, discussing ergodic theory, tensors and categories will be added in order to facilitate a better grasp of the second paper.

## 3.1 Egyptian fractions

In this section, the entire focus will be shifted towards defining Egyptian fractions and the mathematical background necessary to move onto the following chapter, where these fractions would be incorporated into an encryption algorithm.

### 3.1.1 Introduction

The history of *continued fractions* is certainly one of the longest among those of mathematical concepts, since it begins with Euclid's algorithm from the seventh book of his *Elements* (circa 300 B.C.E.), computing the greatest common divisor of two integers. This algorithm can be modified to produce a simple continued fraction for a rational number. For instance, we can use Euclid's algorithm to compute $gcd(21, 51)$ as follows:

$$\begin{array}{ll} 51 = 2 \cdot 21 + 9 & \frac{51}{21} = 2 + \frac{9}{21} \\ 21 = 2 \cdot 9 + 3 & \frac{51}{21} = 2 + \cfrac{9}{2\,9+3} = 2 + \cfrac{1_3}{2 - {}_{+9}} \\ 9 = 3 \cdot 3 & \frac{51}{21} = 2 + \frac{1}{2+\frac{1}{3}} \end{array}$$

Let us define a continued fraction below:

- A generalised **continued fraction** [81] is an expression of the form:

$$x = a_0 + \cfrac{b_0}{a_1 + \cfrac{b1}{a_2 + \cfrac{b_2}{\ddots}}}$$

  Numbers $a_1, ..., b_1...$ are complex or real, and many for simplification may take $b_i = 1$ for any $i$.

- The continued fraction expansion of a number is finite if and only if the number is rational.

- For what follows $x_0 \in \mathrm{R}, x_0 \in\ ]0, 1]$. The regular continued fractions expansion [42, 43] of $x_0$ is given by:

$$\begin{array}{l} \cdot\ u_0 = \left\lfloor \frac{1}{x_0} \right\rfloor \\ \cdot\ x_1 = \frac{1}{x_0} - u_0 \end{array} \tag{3.1}$$

using recurrence it becomes :

$$\begin{aligned} u_n &= \frac{1}{x_n} \\ x_{n+1} &= \frac{1}{x_n} - u_n, \qquad n \in \mathbb{N} \end{aligned} \tag{3.2}$$

For over 2000 B.C., around the time of the construction of the larger Egyptian pyramids, people around the Nile river conceived their system for division and to calculate any rational numbers using only unit fractions (reciprocals), and they have been representing fractions this way $\frac{1}{2}, \frac{1}{3}, \frac{1}{4} \ldots$ , since almost 3000 B.C.

This discovery allowed them to represent any fraction as a sum of different reciprocals, for instance:

$$\frac{3}{4} = \frac{1}{2} + \frac{1}{4}$$

Far later, in 1913, Friedrich Engel, dissected this in detail to find an expansion for rationals called the *Egyptian fraction*, such a representation will be further referred to as an *Egyptian product* or *Engel expansion*.

Though from 1202 C.E. the Engel expansion is not precisely described in the works of Fibonacci [51]; in the Liber Abaci (The Book of Calculations) we find an equation referring to Fibonacci's compound fraction notation, in which a sequence of numerators and denominators sharing the same fraction bar represents an ascending continued fraction:

$$\frac{abcd}{efgh} := \frac{d}{h} + \frac{c}{gh} + \frac{b}{fgh} + \frac{a}{efgh} = \cfrac{d + \cfrac{c + \cfrac{b + \frac{a}{e}}{f}}{g}}{h}.$$

where $a$, $b$, $c$, $d$, $e$, $f$, $g$ and $h$ are integers.

For example, Fibonacci uses the symbol $\frac{111}{345}$ as an abbreviation for:

$$\frac{1 + \frac{1 + \frac{1}{4}}{5}}{3} = \frac{1}{3} + \frac{1}{3 \cdot 4} + \frac{1}{3 \cdot 4 \cdot 5}$$

If this compound fraction has all numerators $0, 1$ , then its is an Engel expansion.

Let us now define an Engel expansion:

**Definition 3.1.** For $x \in \mathbb{R}_{+}^{*}$, the unique non-decreasing sequence of positive integers $\{a_1, a_2, a_3...\}$ such that:

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

is called **Engel expansion** of $x$.

And to illustrate more clearly the shape of this expansion we give a few examples:

**Example 1.** Given below a sequence $a_i$ of the Engel expansion for certain numbers:

$\sqrt{2} := \{1, 3, 5, 5, 16, 18, 78, 102, 120, ...\}$

$e := \{1, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, ...\}$

$\pi := \{1, 1, 1, 8, 8, 17, 19, 300, 1991, 2492, ...\}$

$\ln(2) := \{2, 3, 7, 9, 104, 510, 1413, 2386, ...\}$

- Engel expansions can be written as an ascending variant of a continued fraction[40]:

$$x = \cfrac{1 + \cfrac{1 + \cdots}{\cfrac{a_3}{a_2}}}{a_1}.$$

  with $x$ a positive real number and $(a_1, a_2, a_3, ...)$ positive integers.

- Rational numbers have a finite Engel expansion, while irrational numbers have an infinite Engel expansion.

- The **Pierce expansion** [39], or *alternated Egyptian product*, of a real number $0 < x < 1$ is the unique increasing sequence $\{a_1, a_2, ...\}$ of positive integers $a_i$ such that :

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

- A number $0 < x < 1$ has a finite Pierce expansion if an only if $x$ is rational.

To determine the Engel expansion, we consider an algorithm that also helps compute certain decimal logarithmic values [41]:

- The **Briggs** algorithm [43] is as follows:

$$\begin{cases} u_0 = \left\lfloor \frac{1}{x_0} \right\rfloor + 1 \\ x_1 = u_0 x_0 - 1 \end{cases} \tag{3.3}$$

and using recurrence again:

$$\begin{cases} u_n = \left\lfloor \frac{1}{x_n} \right\rfloor + 1 \\ x_{n+1} = u_n x_n - 1 \end{cases} \tag{3.4}$$

Below are some properties of the Briggs algorithm [42]

a) $\forall\ n \in \mathbb{N}, x_n > 0$

b) $(x_n)$ is a decreasing sequence.

c) $(u_n)$ is an increasing sequence.

d) $\forall\ n \in \mathbb{N}, u_n \geq 2$

- Let $x_0 \in\ ]0, 1]$ and $(u_n)$ be the sequence of the Briggs algorithm, then the serie of general term in the form $\frac{1}{u_0 u_1 \ldots u_n}$ converges, and:

$$x_0 = \frac{1}{u_0} + \frac{1}{u_0 u_1} + \ldots + \frac{1}{u_0 u_1 \ldots u_n} + \ldots$$

- If $x_0 \in\ ]0, 1]$ is written as the expression below:

$$x_0 = \frac{1}{v_0} + \frac{1}{v_0 v_1} + \ldots + \frac{1}{v_0 v_1 \ldots v_n} + \ldots$$

where $(v_n)$ is an increasing sequence of strictly positive integers, then $(u_n) = (v_n)$, sequence given by the Briggs algorithm.

- Let $x_0 \in\ ]0, 1]$ then $x_0$ is rational if and only if the sequence $(u_n)$ of its Engel expansion is constant starting from a certain rank.

*Remark* 3.2. Every rational number also has a unique infinite Engel expansion and using the identity:

$$\frac{1}{n} = \sum_{i=1}^{+\infty} \frac{1}{(n+1)^i}$$

the final digit $n$ in a finite Engel expansion can be replaced by an infinite sequence of $(n + 1)$ without having to change the value.

Engel series can help facilitate many mathematical problems, amongst which we list a major benefit of these expansion, being the approximation of logarithms:

### 3.1.2 Rational approximation of logarithms using Engel series

For simplicity reasons the case study will consist of decimal logarithms [41].

Let us compute $\log a > 1$, $a$ being a positive integer with $2 \leq a \leq 9$.

Let $k_0$ be the smallest positive integer such that

$$a^{k0} > 10$$

Then

$$k_0 \log a > 1$$

and

$$(k_0 - 1) \log a \leq 1$$

so

$$\frac{1}{\log a} + 1 \geq k_0 > \frac{1}{\log a}$$

thus

$$k = \frac{1}{\log a} + 1$$

$k_0$ is then the first term of the Engel expansion of $\log a$.

Let

$$x_1 = k_0 \log a - 1 = \log \frac{a^{k_0}}{10}$$

The same process is done again with $\frac{a^{k_0}}{10}$ instead of $a$. This is done to look for the smallest $k_1$ positive integer, such that

$$\left(\frac{a^{k_0}}{10}\right)^{k_1} > 10$$

$$\Leftarrow\Rightarrow \quad a^{k_0 k_1} > 10^{k_1+1}$$

$k_1$ is the second term of the Engel expansion of $\log a$, so on..

The computation of $\log a$ is basically determining the number of decimal digits, of powers of $a$.

The study of rings of quadratic integers (generalization of the usual integers to quadratic fields, that is an algebraic number field of degree two over the rational numbers) is basic for many questions of algebraic number theory, since they occur in the solutions of many Diophantine equations (typically a polynomial equation in two or more unknowns with integer coefficients, where the only solutions of interest are the integer ones). Therefore, to further expand our knowledge of Engel expansions, we explore another case where they are used on quadratic numbers [44]:

### 3.1.3 Engel expansion of quadratic numbers

The continued fractions expansion of a real number allows the explicit fraction expansion of a quadratic number $x_0$, i.e. solving the equation :

$$ax^2 + bx + c = 0$$

for $a, b, c \in \mathbb{Z}, a \neq 0$.

As for the Engel expansion, it is not the case entirely, but for numbers in the form

$$x_0 = m_0 - \sqrt{m_0^2 - 1}$$

it is possible to obtain such an expansion when $m_0 \in \mathbb{N} \setminus \{0, 1\}$.

Let us consider the sequence $(m_n)$ defined by its first term $m_0$ and the recurrence:

$$m_{n+1} = 2m_n^2 - 1$$

Then

$$m_{n+1}^2 = 4m_n^4 - 4m_n^2 + 1$$

so

$$\sqrt{m_{n+1}^2 - 1} = 2m_n \sqrt{m_n^2 - 1}$$

thereafter

$$\begin{aligned} m_n - \sqrt{m_n^2 - 1} &= m_n - \frac{1}{2m_n} \sqrt{m_{n+1}^2 - 1} \\ &= \frac{2m_n^2 - \sqrt{m_{n+1}^2 - 1}}{2m_n} \\ &= \frac{1}{2m_n} + \frac{1}{2m_n} \left( m_n - \sqrt{m_{n+1}^2 - 1} \right) \end{aligned}$$

Let $x_n = m_n - \sqrt{m_n^2 - 1}$, then

$$x_n = \frac{1}{2m_n} + \frac{1}{2m_n} x_{n+1}$$

or

$$x_{n+1} = 2m_n x_n - 1$$

We already know the Briggs algorithm for $u_n = 2m_n$. Hence :

**Theorem 3.3.** *Let $m_0 \in \mathbb{N} \setminus \{0, 1\}$ and $x_0 = m_0 - \sqrt{m_0^2 - 1}$. The Engel expansion of $x_0$ is given by*

$$x_0 = \sum_{n=0}^{+\infty} \frac{1}{2m_0 \, 2m_1 \ldots 2m_n}$$

*the sequence $(m_n)$ satisfying the recurrence*

$$m_{n+1} = 2m_n^2 - 1$$

Real numbers are deployed in our papers when talking about the Engel expansion.

Hence, for better manipulation skills of such expansions, we go into detail about their relation with continued fractions and some of their properties when dealing with a rational or irrational number. Euler mentioned before [45] how to alter a serie of the form

$$\frac{1}{a} - \frac{1}{ab} + \frac{1}{abc} - \frac{1}{abcd} + \dots$$

into continued fractions. The next part will go precisely over this for the case of Engel series.

### 3.1.4 Engel expansion of real numbers

An essential relation between continued fractions and Engel expansions is shown through the following:

**Proposition 3.4.** *1. With the same sequence conditions as before we have:*

$$\forall n \in \mathbb{N}, \quad \frac{1}{x_n} = 1 - \frac{u_n}{u_{n+1} + \dfrac{1}{x_{n+1}}}$$

2. *Let a rational $x_0 \in \,]0, 1]$ and $(u_n)$ a sequence of its Engel expansion (constant from a rank N). Then*

$$x_0 = \cfrac{1}{u_0 - \cfrac{u_0}{u_1 + 1 - \cfrac{u_1}{u_2 + 1 - \dots \; \dots - \cfrac{u_{N-2}}{u_{N-1} + 1 \dfrac{u_{N-1}}{u_N}}}}}$$

*meaning the Engel algorithm provides an expansion of $x_0$ in limited continued fractions.*

3. *Let $x_0 \in \,]0, 1]$ be irrational and $(u_n)$ a sequence of its Engel expansion. Then the unlimited continued fraction*

$$\cfrac{1}{u_0 - \cfrac{u_0}{u_1 + 1 - \cfrac{u_1}{u_2 + 1 - \cfrac{u_2}{u_3 + 1 - \dots}}}}$$

*converges to* $x_0$.

The Engel algorithm allows the decomposition of $x_0 \in \mathrm{R}$ into a continued fraction that has in common with regular continued fractions the property of being limited when $x_0$ is rational, and unlimited when it is irrational.

In our pursuit of pseudo-randomness and while attempting to generate random-like sequences, the stochastic properties of the Engel series come in handy to explore probable occurrences of certain digits:

### 3.1.5 Stochasticity in Engel series

Let us pick randomly a real number $x_0 \in ]0, 1]$, the unit interval $]0, 1]$ being a probability space endowed with the Lebesgue measure [47, 48].

- The term $u_n$ of the Engel expansion of $x_0$ is then a random variable, with values in $\{2, 3, 4, ...\}$.

- We have $u_n = k, (k = 2, 3, ...)$ if and only if $x_0$ belongs to an interval of the form:

$$\frac{1}{u_0} + \frac{1}{u_0 u_1} + \ldots + \frac{1}{u_0 u_1 \ldots u_{n-1} k} \leq x < \frac{1}{u_0} + \frac{1}{u_0 u_1} + \ldots + \frac{1}{u_0 u_1 \ldots u_{n-1}(k-1)}$$

  with $2 \leq u_0 \leq u_1 \leq \ldots \leq u_{n-1} \leq k$.

- Notice all the intervals are disjoint, therefore

$$\mathrm{P}(u_n = k) = \frac{1}{k(k-1)} \sum_{2 \leq u_0 \leq u_1 \leq \ldots \leq u_{n-1} \leq k} \frac{1}{u_0 u_1 \ldots u_{n-1}}$$

Before understanding the theorem provided next, it is important to know the definition of a *Markov chain* [46], and more precisely, a time-homogeneous Markov chain:

- A discrete-time Markov chain is a sequence of random variables $X_1, X_2, X_3, ...$ with the Markov property, namely that the probability of moving to the next state

depends only on the present state and not on the previous states:

$$Pr(X_{n+1} = x | X_1 = x_1, X_2 = x_2, , X_n = x_n) = Pr(X_{n+1} = x | X_n = x_n)$$

if both conditional probabilities are well defined, that is if

$$Pr(X_1 = x_1, , X_n = x_n) > 0$$

The possible values of $X_i$ form a countable set $S$ called the state space of the chain. Markov chains, are called time-homogeneous when:

$$Pr(X_{n+1} = x | X_n = y) = Pr(X_n = x | X_{n-1} = y) \forall\ n \in \mathrm{N}$$

The probability of the transition is independent of $n$.

**Theorem 3.5.**

- *The sequence of random variables $u_n$ is a homogeneous Markov chain [46], of which the transition probability is given by [125]:*

$$\mathrm{P}\,(u_n = k/u_{n-1} = j) = \frac{j-1}{k(k-1)} si\ 2 \leq j \leq k$$
$$\mathrm{P}\,(u_n = k/u_{n-1} = j) = 0\ si\ j > k$$

*and it is subsequently clear that*

$$P(u_0 = k) = \frac{1}{k(k-1)}$$

- *Let $\epsilon_k$ be the random variable refering to the number of appearences of k inside the Engel expansion. The random variables $\epsilon_k$, $k = 2, 3, ...$ are mutually independent.*

**Lemma 3.6** ([49]). *Let $(A_n)$ be a sequence of mutually independent events and let B be the event: only a finite number of $A_i$ happen simultaneously. Then*

$$P(B) = 1 \Leftarrow\!\Rightarrow \sum_{n=1}^{+\infty} P(A_n)\ \textit{is a convergent serie}$$

**Theorem 3.7** ([48]). *Let* $2 \leq k_1 < k_2 < ... < k_n < ...$ *be a strictly increasing sequence of positive integers. Then, the sequence* $(u_n)$ *of the Engel expansion of* $x_0$ *has for almost every* $x_0$ *an infinity of terms of the sequence* $(k_n)$ *if the serie* $\sum_{j=1}^{+\infty} \frac{1}{k_j}$ *is divergent, and has only a finite number if the serie* $\sum_{j=1}^{+\infty} \frac{1}{k_j}$ *is convergent.*

As a consequence, it is known for a fact that $\sum_{i=1}^{+\infty} \frac{1}{p_i}$ is a divergent serie [50], where $(p_n)$ is the sequence of prime numbers. Thus, choosing a random number $x_0 \in ]0, 1]$, it is quasi-certain its Engel expansion would contain an infinite number of prime terms $u_n$.

Metric properties of such an expansion are vital, therefore, the following part studies the *Hausdorff dimension* [85] of sets related to the growth rate of digits in the Engel expansion:

### 3.1.6 Hausdorff dimension and Engel expansion

Let $(X, d)$ be a metric space. In what follows, we adopt the most prominent construction of Hausdorff measures, which is the *Caratheodory approach* [86]:

- We define the diameter of a set $A \subset X$ as follows:

$$diam(A) := sup \{d(x,y) | x,y \in A\}$$

  if this quantity exists.

- For any $E \subset X$, $\forall\ \delta \in ]0, \infty]$ and $\forall\ \alpha \in [0, \infty[$, we consider the outer measure

$$\mathrm{H}_{\delta}^{\alpha}(E) := \inf \left( \sum_{i=1}^{\infty} (\mathrm{diam}\, E_i)^{\alpha} : E \subset \bigcup_i E_i \quad \text{and} \quad \mathrm{diam}\,(E_i) < \delta \right).$$

- The map $\delta \longmapsto \mathrm{H}_{\delta}^{\alpha}(E)$ is monotone nonincreasing and thus we can define the **Hausdorff $\alpha$-dimensional measure** [85] of $E$ as

$$\mathrm{H}^{\alpha}(E) := \lim_{\delta \to 0} \mathrm{H}_{\delta}^{\alpha}(E)$$

- For $0 \leq s < t < \infty$ and $A \subset X$, we have:
  - $\mathrm{H}^s(A) < \infty \Rightarrow \mathrm{H}^t(A) = 0$
  - $\mathrm{H}^t(A) > 0 \Rightarrow \mathrm{H}^s(A) = \infty$
- The **Hausdorff dimension** $\dim_H(A)$ of a subset $A \subset X$ is then defined as

$$\begin{aligned}\dim_H(A) &:= \sup\{s : \mathrm{H}^s(A) > 0\} = \sup\{s : \mathrm{H}^s(A) = \infty\} \\ &= \inf\ t : \mathrm{H}^t(A) = 0^{\}} = \inf\ t : \mathrm{H}^t(A) < \infty^{\}}\end{aligned}$$

- Let $x = [a_1, a_2, ..]$ denote the Engel expansion of x.
  The coefficients $a_i$ of the Engel expansion typically exhibit exponential growth. More precisely, for almost all numbers in $]0, 1]$:

$$\lim_{n \to +\infty} a_n^{\frac{1}{n}} = e$$

  However, the subset of the interval for which this is not the case is still large enough that

$$dim_H\{x \in\ ]0, 1] : \lim_{n \to +\infty} a_n^{\frac{1}{n}} /= e\} = 1$$

  and for any $\alpha \geq 1$,

$$dim_H\ \ x \in\ ]0, 1] : \lim_{n \to +\infty} a_n^{\frac{1}{n}} = \alpha = 1$$

This Hausdorff dimension may help in the estimation of upper and lower bounds of the Engel series length if needed.

Nowadays there is much interest and need for practical schemes provably achieving security against *chosen-ciphertext attack* (CCA), where an attacker can choose ciphertexts and their decryptions. Attacks presented in the early days of public-key cryptography had highlighted the presence of security threats in this *multi-user setting* [100] (many users, each with a public key, sending each other encrypted data) that were not present in the *single-user setting* (considering a single recipient of encrypted data), arising from

the possibility that a sender might encrypt, under different public keys, plaintexts which although unknown to the attacker, satisfy some known relation to each other.

The scheme we adopt on our paper (chapter 5) is based on a novel *trapdoor* function (TDF: a function easy to compute but difficult to find its inverse) that achieves *indistinguishability* under CCA and better proven-security to cost tradeoffs in the multi-user setting, assuming the DDH problem is hard. This is exactly what secure Multi-Party Computation is revolves around, as stated next:

## 3.2 Secure Multi-Party Computation

Since chapter 5 goes into the construction of a CCA-secure [101] *lossy* trapdoor function (LTF: an 'injective' keyed function having two modes, one invertible and another where some information is lost) [102] in the multi-challenge context (mutli-user), then before going on to more details, one needs to be acquainted with some prerequisites that we define next.

Amongst the major notions that will be encountered later, there is the entropy and Haar measure, but to define them, it is necessary to recall the definitions of Borel $\sigma$-algebra, Hausdorff space, topological group, set translation and left/right invariants:

### 3.2.1 Topological prerequisites

Let $X$ be some set and let $P(X)$ represent its power set. Then a subset $\Sigma \subseteq P(X)$ is called a **$\sigma$-algebra** if it satisfies the following three properties:

1) $X$ is in $\Sigma$ and $X$ is considered to be the universal set (containing all objects, including itself).

2) $\Sigma$ is closed under complementation: If $A$ is in $\Sigma$, then so is its complement $X \setminus A$.

3) $\Sigma$ is closed under countable unions: If $A_1, A_2, A_3, \ldots$ are in $\Sigma$, then so is $A = A_1 \cup A_2 \cup A_3 \cup \ldots$.

Let $M$ denote the phase space of a dynamical system, equipped with $\sigma$-algebra $\mathbb{M}$ and a probability measure $\mu$ defined on $\mathbf{M}$. Consider a finite partition $\xi = \{C_1, C_2, \ldots, C_r\}$

of $M$, generating a stationary random process with values 1, 2, ..., $r$ where

$$w_k(x) = j \quad \text{if } x \in T^{-k}C_j, \quad -\infty < k < \infty$$

The Shannon-MacMillan theorem [108, 114] states there exists

$$h(T,\xi) = \lim_{n\to+\infty} \frac{1}{n} \Sigma_{i_1,\ldots,i_n} \mu(T^{-1}C_{i_1} \cap \ldots T^{-n}C_{i_n}) \ln \mu(T^{-1}C_{i_1} \cap \ldots T^{-n}C_{i_n})$$

The **entropy** $h(T) = \sup_{\xi} h(T,\xi)$ of the dynamical system, with *sup* over every finite partition $\xi$.

*Remark* 3.8. The Kolmogorov entropy [81] mentioned in chapter 5 describes quantitatively PRNGs, with larger entropies being better. PRNGs with equal entropies may present different stochastic properties showing a need for additional metrical invariant.

- Points $x$ and $y$ in a topological space $X$ can be separated by neighbourhoods if there exists a neighbourhood $U$ of $x$ and a neighbourhood $V$ of $y$ such that $U$ and $V$ are disjoint $U \cap V = \emptyset$. $X$ is a **Hausdorff space** if all distinct points in $X$ are pairwise neighbourhood-separable.

- Let $G$ be a group with a Hausdorff topology, with the group operation $g \cdot h$, inverse function $g^{-1}$ and identity $e$. $G$ is a **topological group** if the functions

$$m : G \times G \to G$$

$$m(g,h) = gh$$

and

$$i : G \to G$$

$$i(g) = g^{-1}$$

are continuous functions.

- The **Borel $\sigma$-Algebra** [103] of a topological space $X$ with topology $\mathbb{T}$ is the $\sigma$-algebra $\mathbb{B}(X) := \sigma(\mathbb{T})$. A Borel measure on a space $X$ is a measure on the Borel $\sigma$-algebra of $X$, $\mu : \mathbb{B}(X) \to \mathbb{R} \cup \{\infty\}$ such that if $S \subset X$ is compact, then $\mu(S) < \infty$.

- For any $g \in G$ and $U \subset G$, we define the **set translation** of $U$ as

$$g \cdot U := \{g \cdot u \mid u \in U\}$$

- Let $G$ be a topological group, for any $f \in C(G)$ we define the left and right translation denoted ${}_g f$ and $f_g$ respectively, as

$${}_g f(x) = f(g \cdot x) \text{ and } f_g(x) = f(x \cdot g)$$

- Let $G$ be a compact topological group. A Borel measure $\mu : \mathbb{B}(G) \longrightarrow [0, \infty[$ is **left invariant** (resp. right invariant) if ${}_g\mu(A) = \mu(A)$ (resp. $\mu_g(A) = \mu(A)$) for all $g \in G$ and $A \in \mathbb{B}(G)$, where ${}_g\mu(A) = \mu(g \cdot A)$ (resp. $\mu_g(A) = \mu(A \cdot g)$ ) is the left translation (resp. right translation) of the measure $\mu$.

- Let $G$ be a compact topological group. Then there is a unique left-invariant probability measure $\mu$ on the Borel $\sigma$-Algebra, $\mathbb{B}(G)$ and $\mu$ is also right invariant. This implies that $\mu$ is the unique left and right invariant probability measure.
  $\mu$ is called the **Haar Measure** [104] of $G$.

Since without the proper decryption key, the original data cannot be accessed, data processing can be sometimes outsourced by some companies to a third party, without trusting the third party to properly secure the data. Homomorphic encryption [105] allows this third party to perform calculations on encrypted data, without having to resort to share the secret key.
In the fifth chapter, we discuss the DDH construction over LTF in which this homomorphic encryption concept will be needed to acquire the full intention of our paper, without loosing sense of indistinguishability:

### 3.2.2 Homomorphic encryption and indistinguishability under attack

- An algebraic homomorphism is a map $\phi$ between two groups $(G, \cdot)$ and $(H, *)$ such that:

$$\phi(x \cdot y) = \phi(x) * \phi(y)$$

for $x, y \in G$ and $\phi(x), \phi(y) \in H$.

$G, H$ can be replaced by rings or a similar algebraic structure.

An encryption scheme is **homomorphic** with respect to an operation $\cdot$ on $G$ if

$$\phi^{-1}(\phi(x) * \phi(y)) = \phi^{-1}(\phi(x \cdot y)) = x \cdot y$$

for an operation $*$ on $H$.

- **IND-CPA** [106] is defined between an adversary and a challenger, where the adversary is modeled by a probabilistic polynomial time Turing machine, i.e. it should finish and give back the output within a polynomial number of time steps. In this definition $E(p_k, m)$ represents the encryption of a message $m$ under the key $p_k$:

  1) The challenger generates a key pair $(p_k, s_k)$ with the key size $k$ and publishes $p_k$ to the adversary. The challenger keeps $s_k$.
  2) The adversary submits two distinct chosen plaintexts $m_0, m_1$ to the challenger.
  3) The challenger selects a bit $b \in \{0, 1\}$ uniformly at random, and sends the challenge ciphertext $C = E(p_k, m_b)$ back to the adversary.
  4) The adversary outputs a guess for the value of $b$.

  A cryptosystem is considered **indistinguishable under chosen plaintext attack** if every probabilistic polynomial time adversary possesses only a negligible advantage over random guessing, i.e. if the adversary wins the game with probability $\frac{1}{2} + \epsilon(k)$, where $\epsilon(k)$ is a negligible function in parameter $k$, meaning for every nonzero polynomial function $poly()$ there exists $k_0$ such that $\epsilon(k) \leq \frac{1}{poly(k)}$ for all $k > k_0$.

  The adversary possesses $m_0, m_1$ as knowledge and $p_k$, the encryption of $m_b$ will be a well grounded ciphertext. Due to the probabilistic feature of $E$, thus $m_0, m_1$ are encrypted and the ciphertexts obtained are compared with the challenge ciphertext which does not admit a non-negligible advantage to the adversary.

- In **IND-CCA/IND-CCA2**, additionally to being given the public key, the adversary has a decryption oracle returning the plaintext [101]. In the non-adaptive case, the adversary can ask this oracle until he gets the challenge ciphertext, while in the adaptive situation the adversary proceeds with the decryption oracle despite getting the challenge ciphertext, to prevent it from not passing the challenge ciphertext for decryption.

    1) Challenger generates a key pair $(p_k, s_k)$ relying on the number of bits of the key size $k$ and publishes $p_k$ to the adversary, then keeps $s_k$.
    2) The adversary sends two distinct chosen plaintexts $m_0, m_1$ to the challenger.
    3) Challenger chooses a bit $b \in \{0, 1\}$ uniformly at random, and sends the challenge ciphertext $C = E(p_k, m_b)$ back to the adversary.
    4) The adversary gets the output as a guess for the value of $b$.

In public key cryptography, *trapdoors* are a fundamental tool to build digital signatures or encryption schemes. A trapdoor is a secret piece of information that provides its holder with some special power or advantage, it is fairly easy to stumble upon a door and fall into a trap, however climbing back is hard unless you have the key. Thus a trapdoor as will be seen, is a function computationally easy to calculate and hard to invert, but that becomes invertibility easy with the help of the trapdoor. Now, many if not every such function that we know, is based either on modular arithmetic or lattices in general, therefore, we aim to construct another type in chapter 5 based upon a different process inspired from discrete dynamical systems, namely a lossy trapdoor function. However, leakage of trapdoors has traditionally been considered damaging for practical applications, but our paper emphasises its positive consequences, used to our advantage as it provides a trapdoor to multiple users in order to implement a defensive mechanism, but before going into details in chapter 5, let us first introduce needed concepts to the reader such as one-way functions, trapdoors, and *lossiness*:

### 3.2.3 Trapdoor functions and lossiness

One-way functions and trapdoor functions are often misinterpreted as the same function, one uses the notion of *negligibility*:

- A function $r : \mathrm{N} \to \mathrm{N}$ is **negligible** if:

  $\forall\ p : \mathrm{N} \to \mathrm{N}$ polynomial, $\exists\ \ k_0$ integer such that:

  $$r(k) \leq \frac{1}{p(k)}$$

  for $k \geq k_0$.

- A function $f$ is a **one-way function** if:

  1) Given $x$, it is easy to compute $f(x)$.
  2) Given $y$ in the range of $f$, it is hard to find an $x$ such that $f(x) = y$.

  In other words, the function $f$ is called a one-way function if:

  1) $f$ is polynomial time computable.
  2) Any probabilstic algorithm for inverting $f(x)$ given a random $y = f(x)$ (x at random) has negligible chance of finding a preimage of $y$.

- A function $h$ mapping a message $m$ to a digest is a **one-way hash function** if

  1) $h$ is a one-way function.
  2) Given $m$ and $h(m)$, it is hard to find $m' \neq m$ such that $h(m') = h(m)$.

- A function $f$ is a **trapdoor function** if there exists some secret $k$ such that it is easy to compute $f(x) = y$, but it is a NP-Hard problem to retrieve $x = f^{-1}(y)$ without the secret key $k$.

- A function $f : \{0,1\}^{l(n)} \times \{0,1\}^{n} \to \{0,1\}^{m(n)}$ is a **trapdoor one-way function** (TDOWF) if

  1) $f$ is a one-way function.

2) A public key $y \in \{0,1\}^{l(n)}$, and $f(x, y)$ a function $f_y(x)$ of $x$ mapping $n$ bits to $m(n)$ bits. Then there exists an algorithm for $< y, f_y(x), z >$ giving $x'$ such that $f_y(x') = f_y(x)$, for some trapdoor key $z$ in $\{0,1\}^{k(n)}$.

3) Given $M$ and $f(M)$, it is hard to find $M' \neq M$ such that $f(M') = f(M)$.

- A function $f : \{0,1\}^{l(n)}\{0,1\}^{n} \rightarrow \{0,1\}^{m(n)}$ is a **trapdoor one-way hash function** if $f$ is a trapdoor one-way function and is also a one-way hash function, i.e. if additionally given $M$ and $f(M)$, it is hard to find a message $M' \neq M$ such that $f(M') = f(M)$.

When one switches to lossy mode, he no longer has to deal with polynomial time machines as the proofs become statistical arguments. Thus, a cipher made by an injective key is decrypted, while the one made by a lossy key is statistically independent of the original message, making the both keys indistinguishable from each other.

- A **lossy function** means the size of their image is smaller than the one of their domain.

  Let us assume having an input message $x$ of $n$ bits and $r$ exists such that

$$|ImF| < 2^r$$

  in lossy mode, with the image consisting of the residual leakage (leaks bits from the input), then if given less than $n - r$ bits of $x$, $f$ cannot be inverted.

Intuitively, the family of LTFs has two modes, or branches, injective mode, which has a trapdoor, and lossy mode [74] which is guaranteed to have a small image size. This implies that with high probability the preimage of an element in the image will be a large set. Formally we have:

- Let $\mathrm{X} = \{X_\lambda\}_{\lambda \in \mathrm{N}}$ and $\mathrm{Y} = \{Y_\lambda\}_{\lambda \in \mathrm{N}}$ denote two ensembles of random variables indexed by $\lambda$.

  Given an algorithm $\mathrm{A}$, we define its **advantage** in distinguishing between $X$ and $Y$ as

$$|P_r(\mathrm{A}(X_\lambda) = 1) - P_r(\mathrm{A}(Y_\lambda) = 1)|$$

with the probability being over random values $X_\lambda, Y_\lambda$ and randomness of $\mathbb{A}$.

- $X$ and $Y$ are **computationally indistinguishable**, denoted $X \overset{s}{\approx} Y$, if the advantage of any probabilistic polynomial-time (PPT) algorithm $\mathbb{A}$ is $negl(\lambda)$ ($negl(\lambda)$ standing for negligible in $\lambda$).
- A tuple $(T, f, f^{-1})$ of PPT algorithms is called a family of $(n, k)$-**lossy trapdoor functions** if:
  1) $T(1^\lambda, 1)$ outputs $t, r$ with $t$ a function index, and $r$ its trapdoor ($1^\lambda$ a security parameter). The function $f(t, .)$ is injective deterministic on $\{0,1\}^n$, and
  $$f^{-1}(r, f(t, x)) = x$$
  for all $x$.
  2) $T(1^\lambda, 0)$ outputs $(t, \perp)$ where $t$ is a function index ($\perp$ is a distinguished symbol indicating decryption failure) and $f(t, .)$ is a function on $\{0, 1\}^n$, where the image of $f(t, .)$ has size at most $2^{n-k}$.
  3) The first outputs of $T(1^\lambda, 1)$ and $T(1^\lambda, 0)$ are computationally indistinguishable.

When studying chaos, ergodic theory comes in handy in dynamical systems, especially knowing ergodicity can be a strong existence proof of topological transitivity. This Theory combines techniques and examples from probability theory, statistical mechanics, vector fields on manifolds, group actions of homogeneous spaces.. Ergodic theory has surprising applications in fields as diverse as number theory, which is the part that interests us in our work. The transformation built in chapter 5 is shown to be ergodic relatively to the Lebesgue measure, hence, it is only fair to dedicate this fragment to present an elementary introduction for ergodic theory:

### 3.2.4 Some notions in ergodic theory

Ergodic theory studies the asymptotic average behaviour of systems evolving in time.

- The collection of all states of the system form a space $X$, and the evolution is represented by either

  1) The *tranformation* $T : X \to X$, with $Tx$ representing the state of the system at time $t = 1$, where at time $t = 0$ the system is initially in state $x$.
  2) When an evolution is continuous (or possesses a spacial structure), the evolution is then described by inspecting a group of transformations $G$ (e.g: $Z^2$, R, $R^2$) acting on $X$, meaning every $g \in G$ is identified with a transformation $T_g : X \to X$ and $T_{gg'} = T_g \circ T_{g'}$

**Example 2.** The orbit of a point $x \in X$ under a transformation $T : X \to X$ is $\{T^n(x) \mid n \in \mathrm{N}\}$. The structure of the orbit can explain a lot concerning the original point $x$.

- Let us express T for the quotient group $\mathrm{R/Z} = \{x + \mathrm{Z} \mid x \in \mathrm{R}\}$ identifiable with a circle, being a topological space it can result as a quotient space of [0, 1] by identifying 0 to 1. A bijection between T and [0, 1[ results by forwarding the coset $x + Z$ to $Frac(x)$ the fractional part of $x$.

  Let us define $T : \mathrm{T} \to \mathrm{T}$ by $T(x) \equiv 10x\,[1]$.

  Thus, $x \in T$ is rational if and only if the orbit of $x$ under $T$ is finite.

  First, consider $x = \frac{p}{q}$ is rational: the orbit of $x$ is some subset of $0, \frac{1}{q}, \ldots, \frac{q-1}{q}$.

  Contrariwise, when the orbit is finite then there are integers $m, n$ where $1 \leq n < m$ for which $T^m(x) = T^n(x)$.

  Therefore, $10^m x = 10^n x + k$ for some $k \in \mathrm{N}$, then $x$ is rational.

- A **measure-preserving system** $(X, \mathrm{B}, \mu, T)$ is a finite measure space $(X, \mathrm{B}, \mu)$ equipped with a measureable $T : X \to X$ that is measure-preserving (equivalenty, $T$-invariant)

$$\mu\left(T^{-1}A\right) = \mu(A)$$

  for all $A \in \mathrm{B}$.

- A system is called **ergodic** if the following property is verified:

  For any $A \in \mathbb{B}$ such that $A = T^{-1}A$, $\mu(A) \in \{0, 1\}$.

  Then, if we denote by $L^p$ the vector space of functions which their $p$ exponent is Lebesgue integrable, then ergodicity is equivalent to:

  $f \in L^2, \quad f \circ T = f \Rightarrow f$ is constant almost everywhere.

**Example 3.** Consider $T_\mu : [0, 1[ \to [0, 1[$, $T_\mu x = x + \mu \, mod \, 1$. $T_\mu$ preserves Lebesgue and Haar measures.

- If $\mu = p/q$ is rational, the system is not ergodic, knowing if $A \subseteq ]0, 1/q[$ then

$$\cup_{i=1}^{q}(A + i/q) \text{ is } T_\mu\text{-invariant.}$$

  If $\mu$ is irrational, then $T_\mu$ is ergodic since if

$$f(x) = \sum_n a_n e^{2\pi i n x}$$

  is $T_\mu$ invariant, then

$$f(x) = f(x + \mu) = \sum_n a_n e^{2\pi i \mu} e^{2\pi i n x}$$

  and

$$a_n \left( e^{2\pi i \mu} - 1 \right) = 0$$

  for all $n \neq 0$.

  Then $a_n = 0$ for all $n \neq 0$ because $\mu$ is irrational.

**Theorem 3.9** (Birkhoff Pointwise Ergodic Theorem [84])**.** *Let* $(X, \mathbb{B}, \mu, T)$ *be a measure-preserving system. For any integrable* $f : X \to \mathbb{C}$*, the time average*

$$f^*(x) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x)$$

*exists for almost every* $x \in X$.

*The time average* $f^*$ *is T-invariant,* $f^* \in L^1$ *and* $\int f d\mu = \int f^* d\mu$.

*If T is ergodic with respect to* $\mu$, *then the time average is constant and equal to the space average*

$$f^*(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) = \frac{1}{X} \int f d\mu(x) = \mu$$

*for almost every* $x \in X$.

**Proposition 3.10.** *Let* $U : L^1(X) \longrightarrow L^1(X)$ *be positive* $(f \geq 0 \Rightarrow Uf \geq 0)$ *with* $\|U\| \leq 1$ *and let* $f \in L^1$ *be real valued.*

*If* $f_0 = 0$, $f_n = \sum_{i=0}^{n-1} U^i f$ *for* $n \geq 1$ *and* $F_N(x) = \max\{f_n(x) : 0 \leq n \leq N\}$ *(pointwise maximum), then*

$$\int_{\{F_N > 0\}} f d\mu \geq 0$$

*for* all *N.*

**Corollary 3.11.** *Let* $(X, \mathbb{B}, \mu, T)$ *be a measure preserving system and* $g \in L^1$ *real-valued.*

*If* $A \in \mathbb{B}$ *is T-invariant, then*

$$\int_{B_\alpha \cap A} g d\mu \geq \alpha \mu(B_\alpha \cap A)$$

*where*

$$B_\alpha = \left\{ x : \sup_{n\geq 1} \frac{1}{n} \sum_{i=0}^{n-1} g(T^i x) > \alpha \right\}, \quad \alpha \in \mathbb{R}.$$

## 3.3 Conclusion

Cryptography has evolved to include more sophisticated methods ensuring the highest level of security, including chaos theory and ergodic theory of numbers, as a fascinating field of study exploring unpredictability and complexity of nonlinear systems. Hence, after noticing some chaotic properties in Engel expansions or what might be called Egyptian product, it was noticed they provide a unique and efficient way of representing real numbers as an infinite series of fractions, allowing the use of these expansions for high-end encryption and decryption speed, with a robust resistance to

attacks, generating unique keys for each transmission and showing promising results. It also, rendered the ability to improve the security of trapdoor functions easier, which depend on the complexity of the function and the secrecy of the secret key. Those results are crucial for secure multi-party computation in a world heading towards a more and more complex distributed environment computation wise, as it becomes more interconnected to multiple collaborating parties working on sensitive data that is supposed to be private and cannot be revealed.

# Chapter 4

# On the Use of Egyptian Fractions for Stream Ciphers[1]

> Chaos is a science of process rather than state, of becoming rather than being.
>
> *Chaos: Making a New Science*
>
> *James Gleick*

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,

[1]The text, as well as for the next chapter, is included in thesis exactly as it is published in the corresponding journal.

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.

## 4.1 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 [108] fundamental criteria of confusion and diffusion. Since the discovery of chaotic synchronization by [115], 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 original methods [118] to the latest modern-number-theory-based cryptosystems. For a large collection of papers on the subject, see for instance [75].

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 [121] and references therein, [122] for encryption with decimal expansion, [123] for creating a pseudo-random number generator via an Engel expansion, and more recently, in a larger context, [124] 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 [51] 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 $x = \frac{1}{a_1} + \frac{1}{a_1 a_2} + \frac{1}{a_1 a_2 a_3} + \dots$

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 the *whole* 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 [24], 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.

## 4.2 The Engel expansion algorithm and some properties

The Engel expansion [125] 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 algorithm allowing to retrieve the expansion of $x$ is as follows:

let $u_1 := x$ and for $k > 1$, the brackets denoting integer part,

$a_k := \left\lfloor \frac{1}{u_k} \right\rfloor + 1$ 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:

$u_1 = 1.175$, $a_1 = \left\lfloor \frac{1}{1.175} \right\rfloor + 1 = 1$;

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

$a_2 = \left\lfloor \frac{1}{0.175} \right\rfloor + 1 = 6$;

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

$a_3 = \left\lfloor \frac{1}{0.05} \right\rfloor + 1 = 20$;

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

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 $x \in ]0, 1]$:

- [125] The sequence of the random variables $a_n$ is a homogenous Markov chain, and its transition probability is given by

$$P(a_n = k / a_{n-1} = j) = \frac{j-1}{k(k-1)} \quad if \quad 2 \leq j \leq k$$
$$P(a_n = k / a_{n-1} = j) = 0 \quad if \quad j > k$$

- [48] Let $2 \leq k_1 < k_2 < ... < k_n < .....$ 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 [50], $(p_n)$ 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.

## 4.3 Method

### 4.3.1 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 [121], but is rather a suggestion of the irrational we can choose to expand:

For $r \in [3, +\infty[$ and an algebraic number $A \in ]1, +\infty[$: $\sqrt[r]{\overline{log(A)}}$ 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 $\sqrt[e]{\overline{log(s)}}$.
3. Both sides calculate the Engel expansion of the irrational number $\sqrt[e]{\overline{log(s)}}$.
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 $m := m_1 \oplus k$ .

### 4.3.2 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 [52]. 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 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 4.1. 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 $\alpha$ being the significance level) or verify its uniformity distribution via a histogram (if $P_{value} \geq 0.0001$).

### 4.3.3 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 having this input on figure 4.2. 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

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 4.1:** *NIST statistical test*

**Figure 4.2:** *Clear image*

image cipher as illustrated on figure 4.3.

The three RGB components of our encrypted image have uniform histograms, as being displayed on figure 4.4.

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

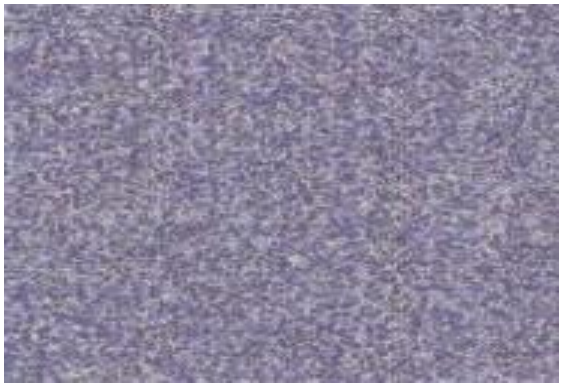


**Figure 4.3:** *Image cipher*

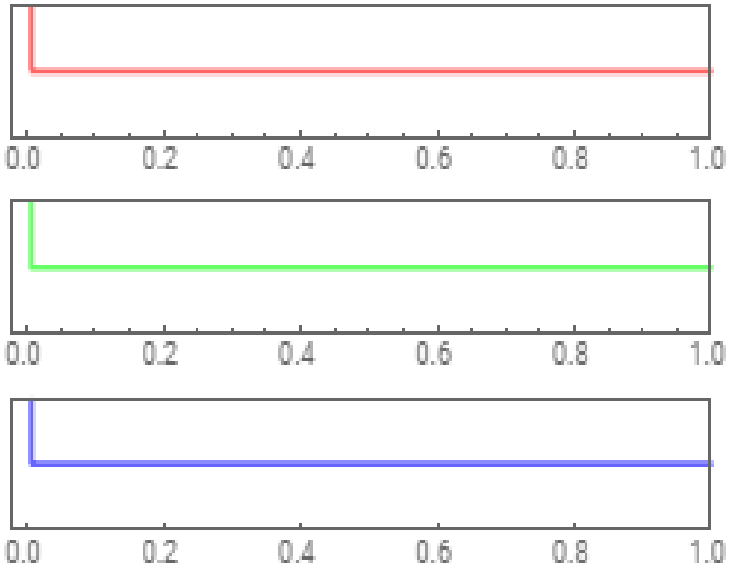



**Figure 4.4:** *Histogram of each RGB component of the encrypted image*

The following histograms on figures 4.5 and 4.6 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.

The scatter diagrams for RGB distribution of the pixels in both the clear image and image cipher on figure 4.7 and figure 4.8 show the decorrelation coefficients between the pixels.

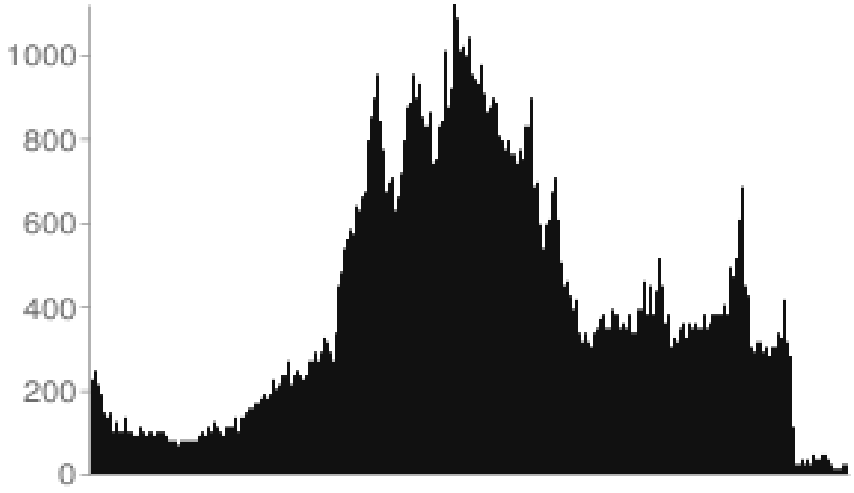



**Figure 4.5:** *Grayscale histogram of the clear image*

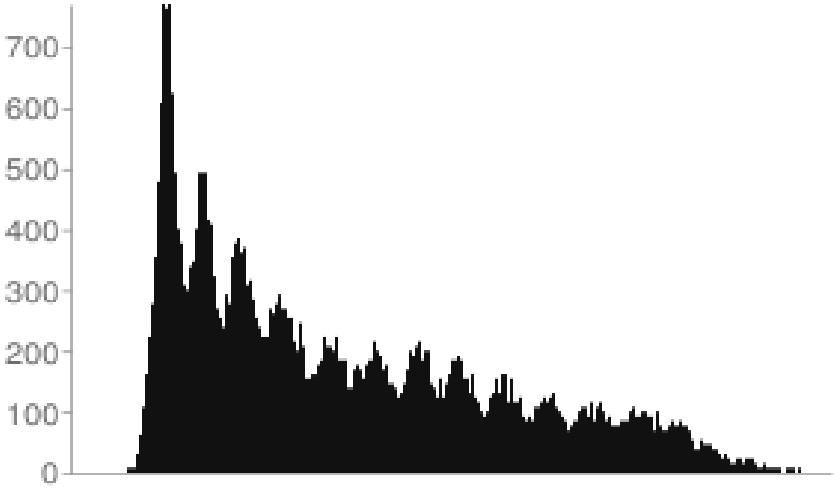



**Figure 4.6:** *Grayscale histogram of the image cipher*

## 4.4 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.

**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 $f^n$.

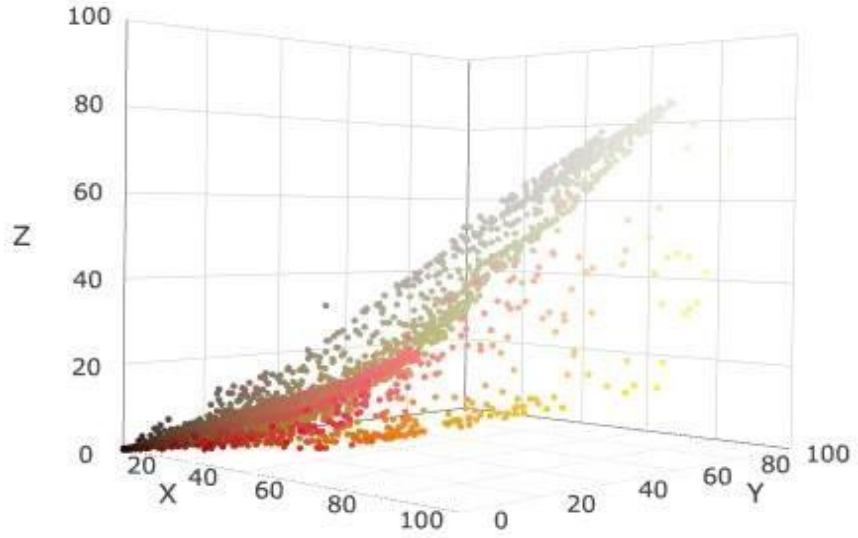



**Figure 4.7:** *Scatter diagram for the clear image*

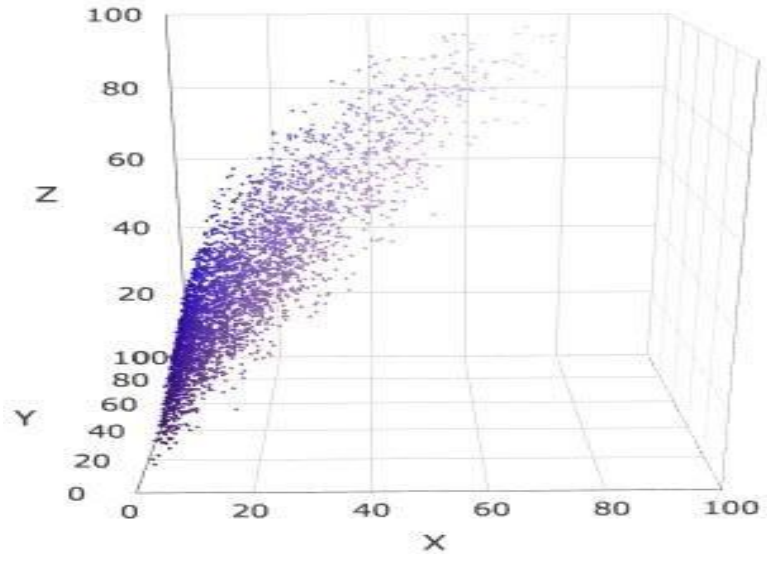



**Figure 4.8:** *Scatter diagram for the image cipher*

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.

**Definition.** Let $D$ be a subset of a metric space with metric $d$. The function $f : D \longrightarrow D$ 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 $\epsilon > 0$, there is $z$ in $D$ such

that $d(z,x) < \epsilon$ and $d(f^n(z),y) < \epsilon$ for some $n$.

**Proposition 1.** Let $D$ be a subset of a metric space and $f: D \longrightarrow D$. 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 $D$.

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

**Definition** ([24])**.** Let $D$ be a subset of a metric space.
The function $f: D \longrightarrow D$ 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).

### 4.4.1 First analysis

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

$$\begin{aligned} f: ]0,1] &\longrightarrow ]0,1] \\ x &\rightarrowtail (1 + \frac{1}{x})x - 1 \end{aligned}$$

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 4.9) is given by $\mathbb{F} = \frac{1}{n}/n \in \mathbb{N}^*$ , and couched in the language of dynamical systems, we have the following formula involving stable sets:

$$W^s(0) = ]0,1] \cap (\mathrm{R}\backslash\mathrm{Q})$$
$$\bigsqcup_{x \in \mathrm{F}} W^s(x) = ]0,1] \cap \mathrm{Q}$$

R and Q being the field of real and rational numbers, respectively.

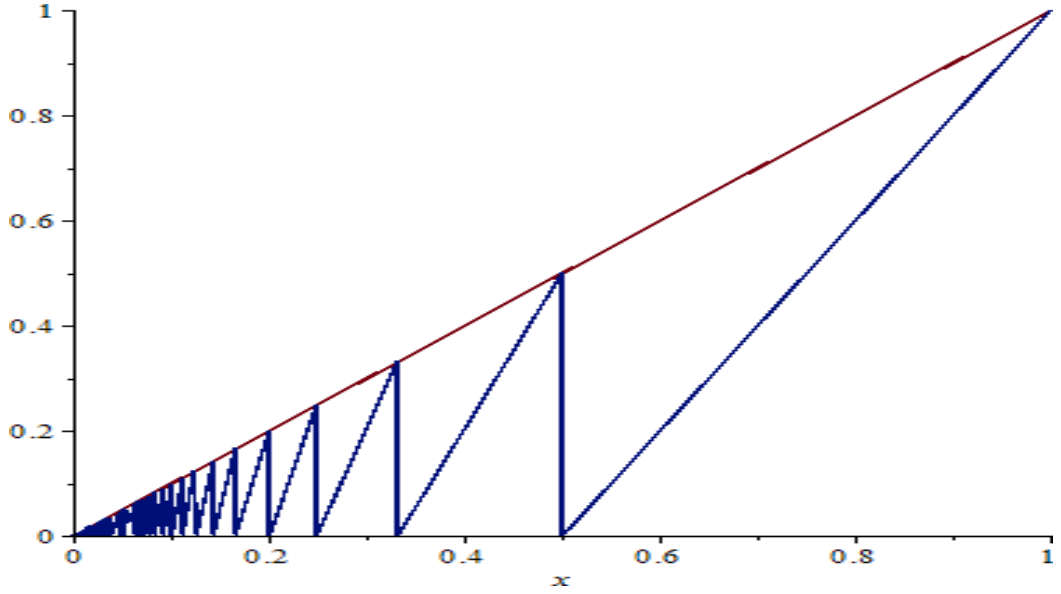



**Figure 4.9:** *Graph of f*

### 4.4.2 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 $[a_1, a_2, ..., a_n]$, 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 the proposition, 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 $\epsilon > 0$. 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 $N \in \mathrm{N}$ such that $\frac{1}{a_1 a_2 ... a_N} < \frac{\epsilon}{2}$.

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 $|x - y| < \epsilon$ 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.

### 4.4.3 Engel expansion distribution

The histogram of what distributions of $0.a_i$ throughout the intervals resembles is shown on figure (4.10), upon which we can notice an almost invariant distribution of the iterates on the interval $[0.1, 1[$.

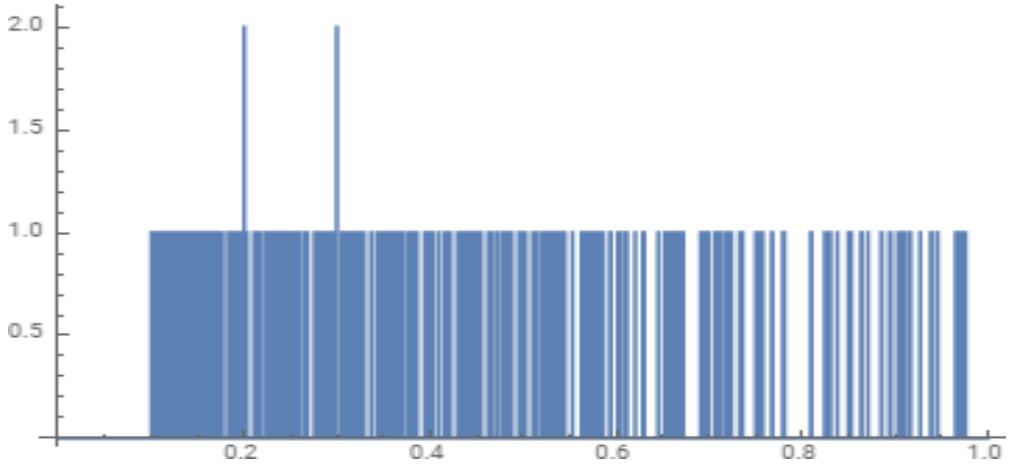



**Figure 4.10:** *Distribution of iterates*

Now, showing the dependence on initial conditions, let's take on figure 4.11 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.

## 4.5 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.

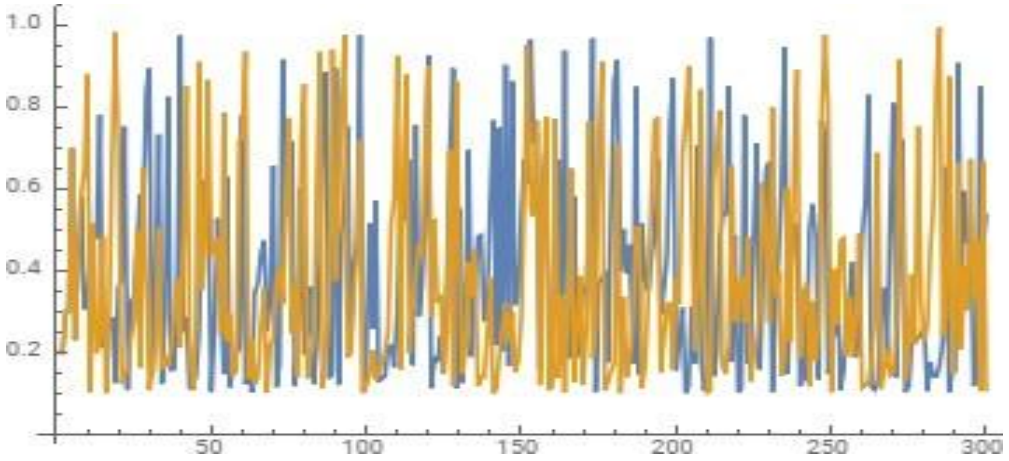



**Figure 4.11:** *Sensitive dependence of the process on initial conditions*

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.

# Chapter 5

# Diffie-Hellman Multi-Challenge using a New Lossy Trapdoor Function Construction based on chaos

> Every secret creates a potential failure point.
>
> *Bruce Schneier*

Trapdoor functions contributed since their announcement in the evolvement cryptography as we know it, especially the lossy mode, by helping reduce the residual leakage for an optimal rate, but to make it more resilient cryptographically: generic constructions were made based on graph isomorphism, or other NP-hard problems defended by the zero-knowledge proof, such were used in Indistinguishability under Chosen-Plaintext Attack (IND-CPA), Computationnal Diffie-Hellman (CDH), or Decisional Diffie-Hellman (DDH).

Once schemes like Indistinguishability under Chosen-Ciphertext Attack (IND-CCA) were adopted it became clear it cannot simulate a decryption using Lossy Trapdoor

Functions (LTF); the problem with existing trapdoor functions in general is partial information leakage, lack of randomness and multiple messages insecurity.

In the light of the following issues came the idea to present through this paper a simple but important fix, in the note of randomness a new Variate of the Engel expansion (VEE) is chosen, providing a pseudo-random bit sequence as an output, the reason being to recover the seed of the algorithm for an attacker, it is considered a hard number theory problem, and surely after the new construction in this paper, another NP-complete problem emerging from tensors the scheme is more secure. As for the strenghtening evidence of how it can be trusted, it seems more robust to supply a proof of its ergodicity as being done in this article, instead of semantic security analysis, to prove the efficiency of the new construction resolving the issues surrounding multi-challenge using a lossy trapdoor function.

## 5.1 Introduction

First and foremost, it is essential to tackle the definition of the Engel expansion [125].

Let $x$ be a positive real number such that:

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

The unique non-decreasing sequence of positive integers $a_1, a_2, a_3, a_4, \dots$ is called the Engel expansion.

The Engel expansion of $x$ can be obtained through executing the fllowing algorithm:

- let $u_1 = x$,
- $a_k = \frac{1}{u_k} + 1$ and $u_{k+1} = u_k a_k - 1$
- If $u_k = 0$ the algorithm stops.

Let $\overline{a_i}$ be the number of digits of $a_i$ and $T_E$ the Variate Engel Expansion (VEE) defined by the following expression:

$$x = \frac{1}{a_1 a_1 10^{-\overline{a_1}}} + \frac{1}{a_1 a_2 a_2 10^{-\overline{a_2}}} + \frac{1}{a_1 a_2 a_3 a_3 10^{-\overline{a_3}}} + \ldots$$

with the same initial conditions and starting domain required for the Engel expansion.

### 5.1.1 Brief introduction to chaos-based cryptogrtaphy

Chaos-based cryptography is interesting due to the broadband power spectrum of chaotic signals, high rates of information transmission, and efficiency at sufficiently low signal-to-noise ratio, chaos is a behaviour of a nonlinear system, looking random, with no stochastic reason [75]. To encrypt using this method keys are generated with chaotic maps or in this case the ergodic nature of the chaotic trajectory, emerging from a seed intializing the system at first.

### 5.1.2 Random number generators

The main one used primarly is the pseudo-random number generator (PRNG) which is periodic and deterministic and the other is the true random number generator (TRNG). When dealing with cryptography, a PRNG is called cryptographically strong if an intruder intercepts information generated by the PRNG, but still doesn't have the possibility to reconstruct the remaining data of the output.

### 5.1.3 Ergodic theory

Ergodic theory is the study of the asymptotic average behavior of systems evolving in time. The collection of all states of the system form a space X, and the evolution is represented by a transformation $T : X \to X$, where $Tx$ is the state of the system at time $t = 1$, when the system (at time $t = 0$) was initially in state $x$.

### 5.1.4 One way function

#### 5.1.4.1 Negligible function

A function $r : \mathrm{N} \rightarrow \mathrm{N}$ is negligible if $\forall \ p : \mathrm{N} \rightarrow \mathrm{N}$ polynomial, $\exists \ k_0$ integer such that: $r(k) \leq \frac{1}{p(k)}$ for $k \geq k_0$ .

#### 5.1.4.2 One way function definition

A function $f$ is called a *one-way function* if:

1) $f$ is polynomial time computable.

2) Any probabilstic algorithm for inverting $f(x)$ given a random $y = f(x)$ (x at random) has negligible chance of finding a preimage of $y$.

### 5.1.5 Trapdoor function

A *trapdoor* function is given an input $m$ is easy to compute the result, but the reversible process is a NP-Hard problem except if we know a special piece of information being the secret.

### 5.1.6 Lossiness

When you switch to lossy mode, you no longer have to deal with polynomial time machines as the proofs become statistical arguments. Thus, a cipher made by an injective key is decrypted, while the one made by a lossy key is statistically independent of the original message, making the both keys indistinguishable from each other.
A lossy function means the size of their image is smaller than the one of their domain. Let us assume having an input message $x$ of $n$ bits and $r$ exists such that $|ImF| < 2^r$ in lossy mode, with the image consisting of the residual leakage (leaks bits from the input), if given less than $n - r$ bits of $x$, then $f$ cannot be inverted.

### 5.1.7 Plan

This paper is divided into several sections as follows:

- Related works: where the aim is to mention previous works in similar domains which gather the essence of our contribution.
- LTF construction: where the construction of the trapdoor function starts its building blocks towards the one-way function as a first step.
- Preliminary: which is a section englobing multiple prerequesites necessary to understand how the issue at hand is approached.
- Effectiveness proof: is a section in which we use the prerequisites we already mentioned before, to prove some properties such as ergodicity.
- Using the VEE as LTF for the DDH: This is the final attained objective where the Diffie-Hellman assumption and trapdoor function are gathered via the first one-way function to solve the problem of Multi-Challenges in IND-CCA.
- Conclusion: It is the final section summing up the focus of our contribution in this paper and possible perspectives.

## 5.2 Related work

One-way trapdoor functions are one of the most fundamental cryptographic primitives, especially lossy trapdoor functions LTFs attracting a lot of attention since the contribution of the pioneers [74], unleashed a wave of similar works on LTFs. [126] thought of a new technique to shrink the public key of matrix construction of [74]. [127] and [128] showed LTFs imply correlated-product TDFs and adaptive TDFs.
After being introduced by [74], Lossy trapdoor functions have become more popular in the recent years, due to the multiple varieties they can offer and how it can benefit other concepts like extending it to the identity-based setting, and trying other constructions more efficient hence the design suggested in this paper.
A previous paper [107] did investiga a novel computational problem "the Composite Residuosity Class Problem, and its applications to public-key cryptography" in which he suggested a new trapdoor mechanism, he also came up in the same work with a

trapdoor permutation and two homomorphic probabilistic encryptions.

Another paper can be mentionned here as well [129] showing techniques used for generic constructions of fully-secure IBE ( Identity-Based Encryption) and selectively-secure HIBE (Hierarchical IBE).

Another technical novelty was back when the paper [126] proposed a compact encoding technique for generating compressed representations for some sequences of group elements using public parameters, which also focuses on shkrinking the discrete-log lossy trapdoor functions key size.

## 5.3 LTF construction

Random proceses cause electronic noise, varrying a signal from its digital position in time, this "jitter" would later serve us in generating random numbers.

The idea is to assemble with a XOR operation multiple outputs coming from inverter ring oscillators. Instead of Brownian noise, so $\phi_i$ is the $i$-th term of the VEE (Variate Engel Expansion).

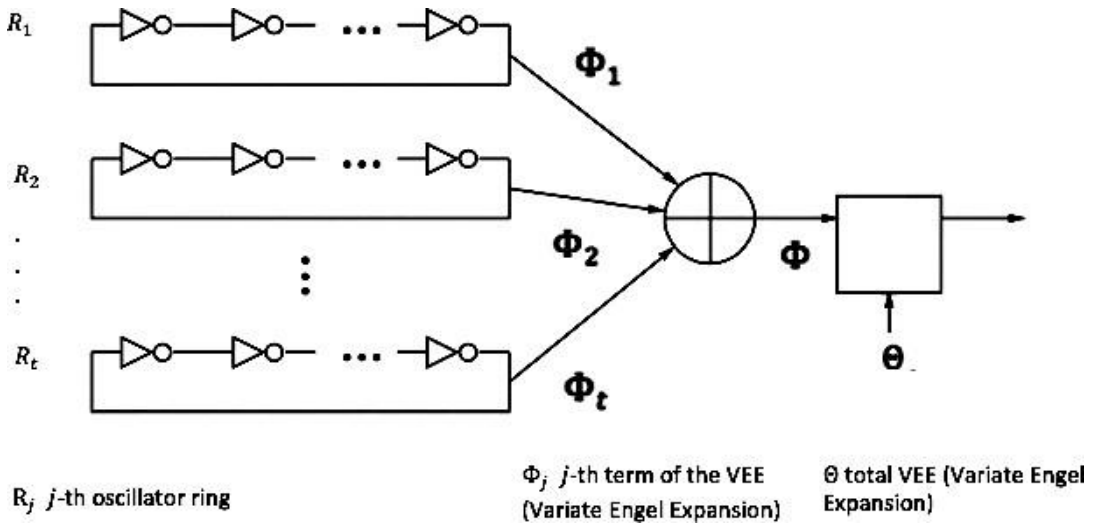



**Figure 5.1:** *Circuit diagram ring design of the LTF*

The events $\phi_1$, $\phi_2$, ...$\phi_r$ fill the oscillation period of the signal $\phi$ which is subdivided into $r$ equal time intervals "urns" less than the jitter boundaries of one ring output, providing a random bit when sampling it as the time lapse is shortened between events as in

figure 5.1.

Now in each of those rings the output value is made following the process we choose to detail afterwards. Once we choose the seed $x$ and number $N$ of iterations, we may follow both strategies below:

1) Divide $[0, 1[$ using the Variate Engel Expansion (VEE) to get an ideal true random number generator allowing the harvest of random integers.

2) Divide $[0, 1[$ using a partition $P_n$ of $2^n$ intervals, to get a non-ideal true random number generator with $2^n$ random integers.

Given $X$ an irrational number on which we count on applying the expansion, the procedure is like the following described algorithm:

During the initialisation, the counter $C$ and key $K$ are set to null giving a clear assumption that the generator is not seeded at first as shown in algorithm 4.

Then the internal entity of the PRNG generates at random a number of blocks.

**Algorithm 4:** Initialisation

**Input:** Initialise

Allocate a real number value to $x$ and

an integer one to $N$ /* $N$ being the $N$-th term of the VEE of $x$ */

**Output:** Generator state

$(K, C) \leftarrow (0, 0)$

$(X, N) \leftarrow (x, n)$ /* $x \in \mathrm{R}$ *and* $n \in \mathrm{N}$ */

**GenerateRandom** function check first if $C$ is not null, as the generator is not seeded in algorithm 5, then the loop starts with $\epsilon$ in $r$ and appends blocks that are computed into r building the output value by the VEE.

Now we know if we take a rational $x_0 \in ]0, 1]$, and $u_n$ is a serie of its Engel expansion starting from rank $N$, then Engel's algorithm offers an expansion of $x_0$ as a limited continued fraction [45].

**Algorithm 5:** GenerateRandom

**Input:** GenerateRandom
/* The VEE($j$) of the iteration $j$ is loaded and given the couple ($x, n$) */
G: Generator state has the VEE
k: block number
**Output:** Pseudorandom string
**while** $C = 0$ **do**
  $r \leftarrow \epsilon$ /* empty string */
  Append block
  **for** *i=1,...,k* **do**
    $r \leftarrow VEE(i)$
    $C \leftarrow C + 1$
  return $r$

$$x_0 = \cfrac{1}{u_0 - \cfrac{u}{u_{1+} - \cfrac{1 \quad 0 \quad u_1}{u_2 + 1 - \ldots \ldots} \quad \cfrac{}{u_{N-2}}}}$$
$$u_{N-1} + 1 - \frac{u_{N-1}}{u_N}$$

While a mapping of continued fractions is given by:

$$x_{n+1} = T(x_n) = \frac{1}{x_n} - \left[\frac{1}{x_n}\right]$$

Gauss found this probability distribution:

$$p(x) = \frac{1}{(1+x)\ln 2}$$

The amplification sensitivity measured by the Kolmogorov entropy [81] is as follows:

$$h = \int_0^1 \ln\left|\frac{dT}{dx}\right| p(x)dx$$

Which results in the following for the mapping T:

$$h = \int_0^1 \frac{-2\ln(x)}{(1+x)\ln 2} dx = \frac{\pi^2}{6\ln 2}$$

Knowing $h$ is non-zero the mapping is considered chaotic so we can deduce Engel Continued Fractions (ECF) is therefore at least sensitive dependent on initial conditions.

The algorithm we chose following the function $F$ is similar to the $r$-adic Rényi transformation in shifting [80]: $S(x) = rx(mod1)$ which is already chaotic for $0 \leq x \leq 1$ and $r > 1$.

Due to the relation between regular continued fractions and ECF we can deduce that the VEE of $x$ has the approximation [79] :

$$\tilde{x} : x \dashrightarrow \frac{p}{q}$$

with $p, q \in \mathbb{Q}$.

For $n \geq 1$ and $1 \leq k \leq a_{n+1} - 1$ the mediants are defined by:

$$\tilde{f} : \frac{p}{q} \dashrightarrow \frac{kp_n(x) + p_{n-1}(x)}{kq_n(x) + q_{n-1}(x)} 10^{\overline{a_n}}$$

which is the finite-precision approximating function, with $\frac{p_n(x)}{q_n(x)}$ the convergents of $x$ in regular continued fractions (RCF).

Let $|f(x) - \tilde{f}(x)| = \epsilon(x)$, then if for all x: $\epsilon(x) << 1$ we can conclude that $\tilde{f}$ shadows $f$ [75] for the pseudo-chaotic approximation $\tilde{f} o \tilde{x}$.

Hence, the use of the ergodic property for this map, to create a nonlinear PRNG which is the main concern and core of our intended LTF, aiming towards a low-complexity implementation and strong statistical test results.

## 5.4 Preliminary

### 5.4.1 Prerequisites

**Definition.** Let $(\Omega_1, F_1, P_1)$ and $(\Omega_2, F_2, P_2)$ be probability spaces and $T$ a transformation:

1. $T$ is measurable if $\forall \; E \in \; F_2 \Rightarrow \; T^{-1}E \in \; F_1$

2. A measurable transformation $T$ is non-singular if $\forall\ E \in F_2 : P_2(E) = 0 \Rightarrow P_1(T^{-1}E) = 0$

3. A measurable non-singular transformation $T$ is ergodic if $T^{-1}E = E$ for $E \in F \Rightarrow P(E) = 0$ or $P(E) = 1$

**Theorem 5.1** (Theorem:[83]). *Let $E$ a Lebesgue measurable subset of $[0,1]$ with $P(E) > 0$ and Lebesgue measure is $\lambda(B_n) = \prod_{j=1}^{n} \frac{1}{k_j}$ $\forall\ n \in \mathbb{N}^*$, $J$ a collection of subintervals of $[0,1]$:*

1. *Every open subinterval of $[0,1]$ is almost a denumerable union of disjoint elements of $J$ ($P$ almost surely)*

2. *$\forall\ B \in J, P(EB) \geq c\,P(E)$ with constant $c > 0 \Rightarrow P(E) = 1$*

**Theorem 5.2.** *Let us define $B_n$ with the expression below:*

$$B_n = B_n(k_1, k_2, \ldots) = \{x \in (0,1] / a_1(x) = k, a_2(x) = k_2, \ldots a_n(x) = k_n\}\ \forall\ k_1, \ldots k_n \in \mathbb{N}^*$$

*for $a_i, i = 1, \ldots, n$ being the coefficients of the Variate Engel Expansion sequence.*
*The set of $B_n$ is bounded and its bounds are:*

$$M_n = sup\, B_n(k_1, k_2, \ldots, k_n) = \frac{10^{k_1}}{k_1} + \frac{10^{k_2}}{(k_1+1)k_2} + \ldots + \frac{10^{k_n}}{(k_1+1)(k_2+1)\ldots(k_n+1)k_n}$$

*and*

$$m_n = inf\, B_n(k_1, k_2, \ldots, k_n) = \frac{10^{k_1}}{k_1} + \frac{10^{k_2}}{(k_1+1)k_2} + \ldots + \frac{10^{k_n}}{(k_1+1)(k_2+1)\ldots(k_n^{-1}+1)k_n} + \frac{10^{k_{n-1}}}{(k_1+1)(k_2+1)\ldots k_{n-1}}$$

*Proof.* If $a_n(x) = k_n$ then $r_{n-1}(x) = \frac{1}{k_n} - \frac{1}{k_{n+1}} r\ (x), n \in \mathbb{N}^*$

where

$$r_n(x) = \frac{1}{a_{n+1}(x)} + \frac{1}{a_{n+1}(x) + 1\ a_{n+2}(x)} + \ldots$$

with $a_{n+1} \in \mathrm{N}^* \quad \forall \ m \geq 1$ and $x \in B_{k_1 k_2 .. k_n}$

Therefore if $x \in B_{k_1 k_2 .. k_n}$ then:

$$x = \frac{10^{k_1}}{k_1} + \frac{10^{k_2}}{(k_1+1)k_2} + .. + 10^{k_n} \frac{1}{(k_1+1)(k_2+1)..(k_{n-1}+1)k_n}$$

$$+ 10^{k_{n+1}} \frac{1}{(k_1+1)(k_2+1)..(k_n+1)a_{n+1}(x)} + 10^{k_{n+2}} \frac{1}{(k_1+1)..(k_n+1)(a_{n+1}(x)+1)a_{n+2}(x)}$$

$$+ ..$$

$$= \frac{10^{k_1}}{k_1} + \frac{10^{k_2}}{(k_1+1)k_2} + .. + \frac{10^{k_n}}{(k_1+1)(k_2+1)..(k_{n-1}+1)k_n} + \frac{10^{k_{n+2}}}{(k_1+1)(k_2+1)..(k_n+1)}.$$

$$\frac{1}{a_{n+1}(x)} + \frac{1}{(a_{n+1}(x)+1)a_{n+2}(x)} + ..$$

$$= \frac{10^{k_1}}{k_1} + \frac{10^{k_2}}{(k_1+1)k_2} + .. + \frac{10^{k_n}}{(k_1+1)(k_2+1)..(k_{n-1}+1)k_n}$$

$$+ \frac{10^{k_n}}{(k_1+1)(k_2+1)..(k_n+1)} r_n(x)$$

Now we are facing two situations:

1) First scenario $n = 2k+1, k = 0, 1, 2, ..$

   If $r_n(x) = 0$, then:

$$m_n = inf\ B_n(k_1, k_2, ..., k_n)$$

$$= \frac{10^{k_1}}{k_1} + \frac{10^{k_2}}{(k_1+1)k_2} + .. \frac{10^{k_n}}{(k_1+1)(k_2+1)..(k_{n-1}+1)k_n}$$

and if $r_n(x) = 1$, then

$$M_n = sup\ B_n(k_1, k_2, ..., k_n)$$

$$= \frac{10^{k_1}}{k_1} + \frac{10^{k_2}}{(k_1+1)k_2} + .. \frac{10^{k_n}}{(k_1+1)(k_2+1)..(k_{n-1}+1)k_n} + \frac{10^{k_n}}{(k_1+1)..(k_n+1)}$$

2) Second scenario $n = 2k, k = 0, 1, 2, ..$

If $r_n(x) = 0$, then:

$$m_n = inf\ B_n(k_1, k_2, ..., k_n)$$

$$= \frac{10^{k_1}}{k_1} + \frac{10^{k_2}}{(k_1+1)k_2} + .. \frac{10^{k_n}}{(k_1+1)..(k_{n-1}+1)k_n} + \frac{10^{k_{n-1}}}{(k_1+1)..(k_{n-1}+1)}$$

while if $r_n(x) = 1$, then:

$$M_n = sup\ B_n(k_1, k_2, ..., k_n)$$

$$= \frac{10^{k_1}}{k_1} + \frac{10^{k_2}}{(k_1+1)k_2} + .. + \frac{10^{k_n}}{(k_1+1)..(k_{n-1}+1)k_n}$$

□

## 5.5 Effectiveness Proof

Ergodicity is the chaotic property equivalent to the cryptographic confusion of Shannon in information theory, where the output has the same distribution for all inputs, making the keystream sequence unpredictable, and kept secret with absence of redundancy.

In order to ensure the chaotic behaviour of the PRNG, ergodicity is a must, knowing with this property at hand trajectories have an invariant distribution unattached to the

initial state, and visiting all intervals of all sizes. Thus, what follows in the paper is the ergodicity establishment of the function $T_E$ defined in the beginning.

**Theorem 5.3.** *The built up transformation based on the Variate Engel expansion $T_E$ is ergodic relatively to the Lebesgue measure $\lambda$.*

*Proof.* Let us define a function $\psi_n = \psi_n(k_1, k_2, ..., k_n), \psi_n [0,1] \to B_n$,

$$\psi_n(v) = \Sigma_{j=1}^{n} \frac{10^{k_j}\lambda(B_{j-1})}{k_j} + 10^{\overline{k_n}}.v.\lambda(B_n) = \Sigma_{j=1}^{n} \frac{10^{kj}}{k_1 k_2 .. k_{j-1}} \left(1 + \frac{10^{k_{j-1}}}{k_j}\right) + 10^{\overline{k_n}}.v.\prod_{j=1}^{n} \frac{1}{k_j}$$

if $x \in B_n$ then:

$$\begin{aligned} x &= \sum_{j=1}^{\infty} \frac{10^{kj}}{a_1 a_2 .. a_{j-1} a_j} \\ &= \sum_{j=1}^{n} \frac{10^{\overline{kj}}\lambda(B_{j-1})}{k_j} + \lambda(B_n). \sum_{j=n+1}^{\infty} \frac{10^{\overline{kj}}}{a_{n+1} a_2 .. a_{j-1} a_j} \\ &= \psi_n(T_E^n(x)) \end{aligned}$$

then $\psi_n = T_E^n : B_n \to I$

and $M_n = \psi_n(1), m_n = \psi_n(0), \forall \; n = 2, 4..$
with $\psi_n(0) = \frac{10^{k_1}}{k_1} + \frac{10^{\overline{k_2}}}{k_1 k_2} + .. + \frac{10^{\overline{k_n}}}{k_1 .. k_n}$
and $\psi_n(1) = \Sigma_{j=1}^{n} \frac{10^{kj}\lambda(B_{j-1})}{k_j} + 10^{\overline{k_n}}\lambda(B_n)$
(If $n$ is odd we invert)

So for any interval $]a, b[ \subseteq I$ we have:

$$\begin{aligned} \lambda(T_E^n]a, b\,[ \cap B_n) &= \lambda(\psi_n[a, b\,] \cap B_n) \\ &= |\psi_n(b) - \psi_n(a)| \\ &= (b - a)\lambda(B_n) \\ &= \lambda]a, b\,[.\lambda(B_n) \end{aligned}$$

thus

$$\lambda(T_E^{-n}E \cap B_n) = \lambda(E)\lambda(B_n) \qquad (*)$$

No matter the set inside the boolean ring $R$ of all finite disjoint unions of intervals $]a, b] \subset I$ the equation is still valid for all borel set $E$ in $I$.Let now $E$ be a Borel set in $I$ such that:

$T_E^{-1}E = E$ then: $T^{-n}E = E, \forall\ n \geq 1$

and (*) $\Rightarrow$ $\lambda(E \cap B_n) =$

$\lambda(E)\lambda(B_n)$

or $\lambda(E \cap B_n) = K\lambda(B_n)$ with $K = \lambda(E) > 0$

If $C$ is the collection of all cylinders $B_n$, $n > 1$, and $a_{j+1} > a_j$, $a_j > 1\ \forall\ j \geq 1$, then any open subinterval of (0, 1 ] would be denumerable at most as a disjoint union of elements of $C$, therefore:

$\lambda(E \cap B) = K\lambda(B)$, $\forall\ B = B_n$ a set of a countable disjoint union. Hence:

From the first property of theorem 5.1 we have $\lambda(E) = 1$ and by the third assertion of definition 5.4.1.

$\Rightarrow$ $T_E^{-1}E = E \Rightarrow$ $P(E) = 1 \Rightarrow$ $T_E$ is ergodic. □

### 5.5.1 Product's sequence correlation

Using the wavelet scalogram as in the figure 5.2 it is shown there is no consistent correlation in the product resulting from the expansion of the specific sequence being chosen, where the horizontal axis represents the time, the vertical axis represents the scale, with which normally correlation is found by measuring energy, where the wavelet transform is defined like the following:

$$Wf(u,s) := \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{+\infty} f(t)\, \psi^*_{u,s}(t)dt$$

where

$$\psi_{u,s} := \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right)$$

$u \in$ R, $s > 0$.

Hence the scalogram of $f, S$ is:

$$S(s) := \|Wf(u,s)\| = \left( \int_{-\infty}^{+\infty} |Wf(u,s)|^2 du \right)^{\frac{1}{2}}$$

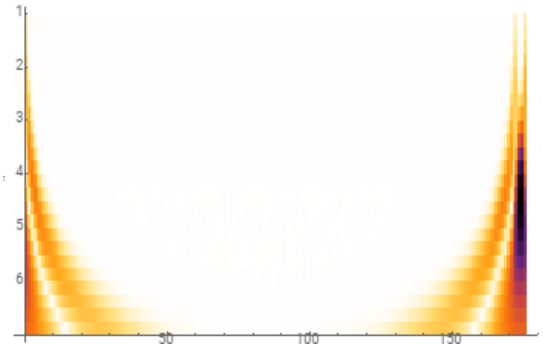



**Figure 5.2:** *Wavelet scalogram for an egyptian product*

Using the innerscalogram which is the normalized scalogram [98], the observer can deduce the scale index obtained by dividing the minimal value of the last one by its maximum, will lead obviously here close to 1 for this highly non periodic expansion [99].

As for the wavelot plot of our Egyptian product, it can be noticed on figure 5.3 that when cross-correlating the wavelet transform with this signal there is no spots at the first rows that may show matches, so it may happen at high number rows randomly due to the specification of the VEE algorithm, thus proving the point.

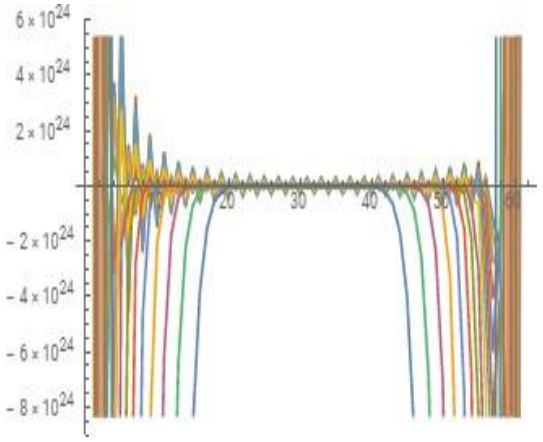



**Figure 5.3:** *Wavelet plot for an egyptian product*

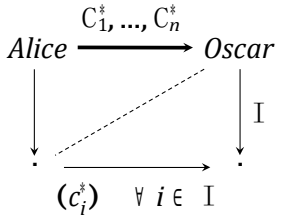
## 5.6 Using the VEE as LTF for the DDH

In this section, the Variant Engel Expansion is being put use as a Lossy Trapdoor Function for the Decisional Diffie-Hellman problem in order to establish a well put together encryption.

### 5.6.1 Multi-Challenge solution in IND-CCA

When dealing with one sender or user in IND-CCA, the trapdoor is used in one challenge making easy to perform hashes or encryption via a one-way function; the issue at hands occurs when dealing with multiple users forcing the encryption modules to handle Multi-Challenges.

Selective Openings do target this topic :




As you can see, there is no indistinguishability with this adaptive corruption of multiple senders, giving away open $(c_i^*)$ $\forall i \in I$ while sending the ciphertexts $C_1^*, \ldots, C_n^*$, which makes the attacker well aware of important informationnal parts on the public key $p_k$ and ciphers $(c_i^*)$, since the randomness uses openings in committment.

Now let's consider in what follows this keyed function:

$$x \longrightarrow f_{k_i}(x)$$

with $k_1$: the key corresponding to the invertible mode

and $k_2$: the key corresponding to the lossy mode

where $k_1 \approx k_2$ and the VEE will be used as the trapdoor function $f$.

In the case of invertibility, an invertible key is being called upon while when needing lossiness the construction of the function guarantees that the image set is much smaller than the preimage set ($f_{k_2}(x)$ $X$).

So getting back to the issue, knowing the attacker gets the LTF key and image $p_k$, $c^*$ from the sender, then although switching LTF to lossy mode would deny the eavesdropper from reaching information on the messages, if the sender is operating under IND-CCA the decryption oracle is unable to function in lossy mode, due to its limitation to work either under a cipher using fully invertible mode or lossy, and cannot alternate between the two.

To prevent this from happening while keeping the encryption functional, tags like $t^*$ are being introduced [107] that switch the function $f_{k,t}$ to lossy mode only for one special tag.

Let f be an $n$-degree polynomial function such that :

$$f(t) = \sum f_i t^i$$

with $f_i$ being the output of the VEE.

and the only tags non-null are $t_1^*, ..., t_n^*$.

then $k = (p_k, C_0 = E_{p_k}(f_0), ..., C_n = E_{p_k}(f_n))$
and $f_{k,t}(x) = (\prod_i c^{t^i})^* = E_{p_k}(f(t)X)$. Now due to the number of challenges to encode $n$ lossy tags the space complexity is linear, and the Selective Openings chosen-ciphertext attack (SO-CCA) model would secure the public key echange (PKE) but will make the public key $p_k$ larger, so the sender will have to consider each $t_i^*$ sampled by the trapdoor function is corresponding to a ciphertext challenge, because there are many superpolynomial lossy tags.

### 5.6.2 DDH construction over LTF

The adopted approach her is where matrices are used instead of single bits, hence the encryption will be performed over a matrix $M$ as the message, the $E$ denotes the DDH encryption scheme and $t$ is the lossy tag used alongside the trapdoor function (TDF).

$$t \to E(M) = \begin{bmatrix} E(M_{1,1}) & \dots & E(M_{1,n}) \\ \vdots & \ddots & \vdots \\ E(M_{n,1}) & \dots & E(M_{n,n}) \end{bmatrix}$$

then the function becomes $f_{k,t}(x) = E(M) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

So

$$f_{k,t}(x) = \begin{bmatrix} \prod_i E(M_{1,i})^{x_i} \\ \vdots \\ \prod_i E(M_{n,i})^{x_i} \end{bmatrix} = E(MX)$$

Notice $f_{k,t}$ is lossy $\Leftarrow \Rightarrow$ $M$ is non-invertible

$$\Leftarrow \Rightarrow \quad det(M) = 0$$

This $det(M)$ can be used to encode complex computations when being cubic, but this is not the aim in this section.

In Diffie-Hellman, the model relies on a number $g \in G$ considered as a group generator such that if $M$ is a bit sequence of the message:

$$M \in Z_p^{n \times n} \Rightarrow [M] = g^M$$

Giving

$$[M] \in G^n$$

Allowing any integer matrix $M$ to be encoded as $[M]$, therefore any input of the TDF as bits $x \in \{0, 1\}$ can be encoded as $[Mx]$.

The slight twist here is to substitute the exponentiation $g^M$ by $E(M)$ which is additively homomorphic and the pairing becomes the multiplication introduced earlier (Paillier) when dealing with matrices.

In order, to establish the earlier method for a multi-dimensionnal purpose, tensors are

considered, and instead of matrix multiplication tensor product is adopted.

Let there be the field $\mathrm{G} = GF2$ and vector spaces

$$F \cong \mathrm{G}^{d1}, M \cong \mathrm{G}^{d2}, X \cong \mathrm{G}^{d3}$$

Then E is a 3-tensors space :

$$\mathrm{E} = F \otimes M \otimes X$$

and let $F^\star = \mathrm{Hom}(F, \mathrm{G})$

Given the action of the group $G = GL(F) \times GL(M) \times GL(X)$ on E if $(e_1, \ldots, e_{d_1})$ is the considered basis for a space $U$, the dual basis is $e^\star, \ldots, e^\star$ for $U^\star$ where $e^\star e_j = \delta_{ij}$.

The (E; $n$)-tensor is as follows:

$$E = \sum_{i=1}^{n} f_i \otimes m_i \otimes x_i$$

for vectors $f_1, \ldots, f_n \in F$ being the decomposition of $f$ the Variant Engel Expansion (VEE) one-way function $m_1, \ldots, m_n \in M$ those of the message sequence and $x_1, \ldots, x_n \in X$ an input sequence of bits.

The difficulty the attacker would encounter arrises from the bilinear inversion problem which is NP-complete:

Given a bilinear mapping $E$ and $z \in X$, find $x$ and $y$ such that $E(x, y) = z$.

Let $d_1 = |F| = d_2 = |M|, d_3 = |X|$ and E the space of bilinear mappings

$$E : F^\star \times M^\star \longrightarrow X$$

$$E(x, y) = E \cdot (x \otimes y)$$

So

$$(E(x,y))_k = \sum_{i,j} E_{i,j,k} x_i y_j \qquad \text{for } k \in \{1,\ldots,m\}$$

The public key is $E$ and the private key is the decomposition $f_1,\ldots,f_{d_3} \in F, v_1,\ldots,v_{d_3} \in M, w_1,\ldots,w_{d_3} \in X$.

The one-way function is the bilinear map

$$T : \mathrm{F}^{2d1} \longrightarrow \mathrm{F}^{d3}, (x,y) \longrightarrow T(x,y) = z$$

taking $m$ and $n$ two factors $(a_i)$ of the VEE expansion such that $n < m$, leaving the problem NP-Hard, thus the robustness of the encryption scheme.

## 5.7 Conclusion

Ergodic theory is a gathering of number theory, probability theory, group actions of homogeneous spaces and other fields. An additional concept may arise and can be relied on, the one of asymptotic average weak independence much stronger which is *mixing*, deriving from the *Birkhoff's ergodic theorem*.

This notion presented the opportunity to entangle LTFs from another point of view, the seed of the system became the parameter of the key generation algorithm, while the VEE designed in this paper outputs the one-way trapdoor function, and the chaotic behaviour made the lossy mode accessible through the ergodic property.

Over the line of this work, this new construction combines both aspects of number theory problem and a chaos theory: the use of a lossy trapdoor function was aimed towards fixing revolving issues surrounding multi-party communications, while using the DDH for exchange, thus making a sound proof to what may become a IND-CCA resistant scheme, especially if studied in the future along distributed systems or adopted in certain communication protocols.

# Chapter 6

# Conclusion and some perspectives

> The tool which serves as intermediary between theory and practice, between thought and observation, is mathematics; it is mathematics which builds the linking bridges and gives the ever more reliable forms.
>
> *Radio broadcast (8 Sep 1930)*
>
> *David Hilbert*

Some chaos-based digital cryptosystems were suggested and extensively studied. The main idea was to take profitably the statistical ergodic properties of Egyptian fractions to design NIST-Certified PRNGs for stream cipher applications. Interestingly, and perhaps the most important, Devaney's chaos has been detected inside such a process, leading to a mathematical justification of the designed cryptosystems. In this attempt to unveil the hidden secrets of the random-like behaviour in Egyptian fractions expansions, we were inspired by previous explorations of the decimal and continued fractions expansions of irrational numbers, to find ourselves confronted to a mutual interference between modern number theory and chaos-based cryptography.

We are still waiting for good news from the Berkley Laboratory about the normality of $\pi$, although such a type of problems reveals to be as difficult as the Riemann hypothesis or the Syracuse conjecture. It could also be, who knows, Gödel-undecidable !

Realistically, for future work, we won't try to prove the normality of $\pi$ or the irrationality of the Feigenbaum constant, but to push forward the present study of Egyptian fractions in order to design stronger cryptosystems and to bring to light new chaotic functions. For example, another approach would consist in considering a concatenation of the involved denominators as fractional part of a number in the unit interval, then shifting alongside the sequence, a process which seems to be chaotic from some primary diagnoses.

Another major result was deduced from the chaotical aspect of the process made for the prior PRNG, which is based on a variant of the Engel expansion with a slight play on terms, the initial condition of the system still being the parameter of the key generation. This allowed us to create a lossy trapdoor function in the context of multi-party computation, proved to be ergodic in our case, thus conceiving a DDH exchange with an IND-CCA robust scheme.

Many challenges remain therefore to be addressed to pave the way towards high performance chaos and number-theory-based methods, tackling large scale applications in cryptography, and essentially, *post-quantum cryptography*, hopefully linking other mathematical disciplines together, as Hilbert once dreamed ...

# Appendices

# Appendix A

# Statistical tests

This part, is fundamental to grasping the knowledge of an essential randomness generation test suite, put together by the NIST scientists, that we incorporated into our paper as exhibited in chapter 4.

## A.1 Overview: NIST test suite

The statistical test results are always displayed on a table of $p$ rows and $q$ columns, in a way such that:

- $p$ represents the number of rows, which denotes the number of statistical tests applied to the sequence generated.
- $q$ represents the number of columns $q = 13$ with:
  - The columns from $C1$ to $C10$ constitutes the frequency of 10 $P$-values.
  - The column 11 is the $P$-value from the chi-square test 11.
  - The column 12 embodies the proportion of binary sequences that passed
  - Finally, the column 13 is the corresponds statistical test.

The appropriate values of the parameters for each particular test recommended by NIST is shown on A.1, where the recommended size bitstreams for each particular test. Here below the technical and a more detailed description for the tests adopted.

| Test # | Test name | $n$ |
|---|---|---|
| 1. | Frequency | $n \geq 100$ |
| 2. | Frequency within a Block | $n \geq 100$ |
| 3. | Runs | $n \geq 100$ |
| 4. | Longest run of ones | $n \geq 128$ |
| 5. | Rank | $n > 38\,912$ |
| 6. | Spectral | $n \geq 1000$ |
| 7. | Non-overlapping T. M. | $n \geq 8m - 8$ |
| 8. | Overlapping T.M. | $n \geq 10^6$ |
| 9. | Maurer's Universal | $n > 387\,840$ |
| 10. | Linear complexity | $n \geq 10^6$ |
| 11. | Serial | |
| 12. | Approximate Entropy | |
| 13. | Cumulative sums | $n \geq 100$ |
| 14. | Random Excursions | $n \geq 10^6$ |
| 15. | Random Excursions Variant | $n \geq 10^6$ |

**Figure A.1:** *Recommended size of bits n for each NIST test*

### A.1.1 Frequency (Monobit) Test

The test determines if the number of ones and zeros in a sequence is almost the same being a random sequence, by inspecting how close is the fraction of ones to $\frac{1}{2}$.

The null hypothesis is the key element in this test [54]:

In a sequence of independent identically distributed Bernoulli random variables ($X$ or $\epsilon$, with $X = 2\epsilon - 1$, then $S_n = X_1 + ... + X_n = 2(\epsilon_1 + ... + \epsilon_n) - n$), the probability of ones is $\frac{1}{2}$.

Using De Moivre-Laplace theorem, the distribution of the binomial sum is normalized by $\sqrt{n}$, approximated by a standard normal distribution.

With help of the Central Limit Theorem:

$$\lim_{n\to\infty} P\left(\frac{S_n}{\sqrt{n}} \leq x\right) = \Phi(x) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$$

So for $z$ positive

$$P\left(\frac{S_n}{\sqrt{n}} \leq x\right) = 2\Phi(x) - 1$$

Relying on

$$s = \frac{|S_n|}{\sqrt{n}}$$

the observed value $|s(ob) = \frac{|X_1 + \cdot + X_n|}{\sqrt{n}}|$ is evaluated and the *P*-value is computed being :

$$2(1 - \Phi(|s(ob)|)) = err \quad \frac{|s(ob)|}{\sqrt{n}}$$

with *err* the error function

$$err(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} e^{-t^2} dt$$

.

### A.1.2 Frequency Test within a Block

The test detects localized deviations from the 50% frequency of ones, which is the perfect case, and this by fragmenting the test sequence into a number of nonoverlapping subsequences and using the chi-square test for a homogeneous match of empirical frequencies to $\frac{1}{2}$.

*P*-values that are little represent a sign of large deviations from the equal proportion of 1's and 0's in at least one of the substrings [55]. The string of ones and zeros or $-1$ and 1 is organised into disjoint substrings, for each one of them, we calculate the proportion of ones.

The chi-square test compares the proportions to the perfect case scenario, with degrees of freedom exactly being the number of substrings. *N* being the number of substrings and $l$ the length of each one such that $n = Nl$ and thus the probability of 1's estimated by the observed relative frequency of ones, $\pi_i$, with $i = 1, ..., N$.

$$\chi^2(ob) = 4l \sum_{i=1}^{N} \left(\pi_i - \frac{1}{2}\right)^2$$

this sum under the randomness hypothesis has $\chi^2$- distribution with *N* degress of freedom. The *P*-value then becomes

$$\frac{\int_{\chi^2(ob)}^{+\infty} e^{-\frac{t}{2}} t^{\frac{N}{2}-1} \, dt}{\Gamma(N/2)2^{N/2}} = \frac{\int_{\chi^2(ob)/2}^{+\infty} e^{-t} t^{\frac{N}{2}-1} \, dt}{\Gamma(N/2)} = \gamma\left(\frac{N}{2}, \frac{\chi^2(ob)}{2}\right)$$

with $\Gamma$ and $\gamma$ denoting the incomplete gamma function, with

$$\gamma(s,x) = \int_0^x t^{s-1} e^{-t} dt$$

being the lower incomplete gamma function, and the upper one is

$$\Gamma(s,x) = \int_0^{+\infty} t^{s-1} e^{-t} dt$$

### A.1.3 Runs test

It is a nonparametric test considering substrings of consecutive ones and consecutive zeros as runs, to see if there is a fast oscillation [56].

Consider $V_n$ the distribution of the total number of runs, and the fixed proportion $\pi = \frac{\epsilon_i}{n}$, such that $|\pi - \frac{1}{2}| \leq \frac{2}{\sqrt{n}}$ we have

$$\lim_{n \to +\infty} P \left( \frac{V_n - 2n\pi(1-\pi)}{2\sqrt{n\pi(1-\pi)}} \leq x \right) = \Phi(x)$$

For $k = 1, ..., n-1$, $V_n = \sum_{k=1}^{n-1} r(k) + 1$, with

$$\begin{aligned} r(k) &= 0 \; ; \epsilon_k = \epsilon_{k+1} \\ r(k) &= 1 \; ; \epsilon_k \neq \epsilon_{k+1} \end{aligned} \tag{A.1}$$

and the reported $P$-value would be

$$err \left( \frac{V_n(ob) - 2n\pi(1-\pi)}{2\sqrt{2n\pi(1-\pi)}} \right)^!$$

When $V_n(ob)$ is bigger it means the oscillation in the string of $\epsilon$'s faster.

### A.1.4 Test for the Longest Run of Ones in a Block

Let $n$ be the string length, $n = MN$ divided into $N$ substrings each of length $M$, and let $v_0, v_1, ..., v_k$ be the frequencies, with $v_0 + v_1 + ... + v_k = N$. When $r$ ones and $M - r$

zeros are in the $m$-bit block, with [57] $U = min \quad m - r + 1, \frac{r}{m+1}$ then,

$$P\ v \quad m\ r \quad \frac{1}{\quad} \overset{U}{\underset{1}{)}} \overset{i}{} \quad M - r + 1, \ , \ M - i(m+1),$$

$$\overset{M}{(\ \leq \ |\ ) = \ \ \sum_{i=0}(-\ \ \bullet \ \ \ \ \bullet}$$

$$i \qquad M - r$$

and

$$P(v \leq m) = \sum_{r=0}^{M} \overset{M}{r} P\overset{(}{v} \leq m \mid r)\frac{1}{2^M}$$

So the $\mathrm{X}^2$-statistic conjoints the empirical frenquencies $v_i$ for $i = 0, ..., k$ in [58]:

$$\mathrm{X}^2 = \sum_{i=0}^{k} \frac{(v_i - N\pi_i)^2 N\pi_i}{}$$

which is a distribution having $k$ freedom degrees, then the $P - value$ is

$$\frac{\int_{\mathrm{X}^2(ob)}^{+\infty} e^{-\frac{t}{2}} t^{\frac{k}{2}-1}\, dt}{\Gamma(k/2)2^{k/2}} = \gamma \quad \frac{k}{2}, \frac{\mathrm{X}^2(ob)}{2}$$

The statistic $\mathrm{X}^2$ shows if the sequence has crowded ones, and when it is random $v_n$'s are little.

### A.1.5 Binary Matrix Rank Test

Matrices are built from successive zeroes and ones of the sequence to verify if there is linear dependence between the rows or columns, the details and what is left of this part is well detailed on [59, 60, 61]. When $\mathrm{X}^2$ is bigger it means the deviation of the rank distribution from the one of a random sequence is also bigger while for random sequences it should be smaller.

The rank $R$ of the $M \times Q$ matrix full of random binary has values $r = 0, 1, 2, ..., m$, with $m = min(M, Q)$ and

$$\underset{r}{p} = \quad 2^{r(Q+M-R)-MQ} \prod_{i=0}^{r-1} \frac{(1 - 2^{i-Q})(1 - 2^{i-M})}{1 - 2^{i-r}}$$

for the test suite the probability is fixed $M = Q = 32$, $n = M^2N$, with $N$ the size of the sequence and $M$ the parameter. The choice of $M$ and $N$ is such that $n - NM^2$ is little, then

$$p_M \approx \prod_{i=1}^{+\infty} 1 - \frac{1}{2^i} \approx 0.2888...$$

$$p_{M-1} \approx 2p_M \approx 0.5776...$$

$$p_{M-2} \approx \frac{4p_M}{9} \approx 0.1284...$$

the rest of the probabilities are less than or equal to 0.005 meaning little once $M \geq 10$.

### A.1.6 Discrete Fourier Transform (Spectral) Test

The Fourier test inspired from the discrete tranform finds periodicity in the sequence [62].

Consider $x^k$ the $k$-th bit, for $k = 1, ..., n$, bits are coded $-1, 1$ supposably, while considering the function

$$f_i = \sum_{k=1}^{n} x_k \exp(2\pi i(k-1)\frac{j}{n})$$

with $\exp(\frac{2\pi ikj}{n}) = \cos(\frac{2\pi kj}{n}) + i\sin(\frac{2\pi kj}{n})$, for $j = 0, ..., n-1$, and $i = \sqrt{-1}$. due to the real to complex value transform symmetry it is only logical to take in consideration values just from 0 to $\frac{n}{2} - 1$. The modulus of the complex number $f_j$ is denoted by $mod_j$, $N_1$ the number of peaks less than $h = \sqrt{\log(\frac{1}{0.05})n}$, considering only the first $\frac{n}{2}$ peaks, with $N = \frac{(N_1 - N_0)}{\sqrt{n\frac{0.95\times 0.05}{4}}}$, then the $P$- value would be

$$2(1 - \Phi(|d|)) = err \quad \frac{|d|}{\sqrt{2}}$$

where $\Phi$ represents the cumulative probability function of the standard normal distribution.

### A.1.7 Non-Overlapping Template Matching Test

This session consists in detecting sequences with aperiodic pattern [63].
Consider $B = (\epsilon^0_1, ..., \epsilon^0_m)$ a sequence of 1's and 0's with length $m$, and the set of periods

of $B$

$$B = \{j, 1 \leq j \leq m-1, \epsilon^0_{j+k} = \epsilon^0_k, k = 1, ..., m-j\}$$

Let $W = W(m, M)$ be the number of occurrences of a pattern B, the statistic $W$ is

$$W = \sum_{i=1}^{n-m+1} I(\epsilon_{i+k-1} = \epsilon^0_k k = 1, ..., m)$$

For $I(\epsilon_{i+k-1} = \epsilon^0, k = 1, ..., m)$ are $m$-dependent random variables, thus Central Limit Theorem is applicable, and the mean and variance of the approximating normal distribution woul be

$$\mu = \frac{n-m+1}{2^m}$$

and

$$\sigma^2 = n\left(\frac{1}{2^m} - \frac{2m-1}{2^{2m}}\right)$$

Let $W_j = W_j(m, M)$, $W_j$ has normal distribution then

$$\chi^2(ob) = \sum_{j=1}^{N} \frac{(W_j - \mu)^2}{\sigma^2}$$

has almost a $\chi^2$- distribution of $N$ freedom degrees, with the $P$-value being

$$1 - P\left(\frac{N}{2}, \frac{\chi^2(ob)}{2}\right)$$

### A.1.8 Overlapping Template Matching Test

This session detects $m$-runs of 1's, so let $n = MN$ dividing the string into $N$ subparts of length $M$, $W_j = W_j(m, n)$ is the number of runs of 1's of length $m$ in the $j$-th block, with $W_j$ having compound Poisson distribution as an asymptotic distribution [64]

$$E\exp\{tW_j\} \rightarrow \exp\{\frac{\lambda(e^t - 1)}{2 - e^t}\}$$

once $(M - m + 1)2^{-m} \to \lambda > 0$ with $t \in \mathbb{R}$

For $u \geq 1$ with $\eta = \frac{\lambda}{2}$, $\Phi$ confluent hypergeometric function :

$$P(U = u) = \frac{e^{-\eta}}{2^n} \sum_{l=1}^{u} \binom{u-1}{l-1} \frac{\eta^l}{l!} = \frac{\eta e^{-2\eta}}{2^u} \Phi(u + 1, 2, \eta)$$

The distribution function has a complement

$$L(u) = P(U > u) = e^{-\eta} \sum_{l=1}^{u} \frac{n^l}{l!} \Delta(l, u)$$

where

$$\Delta(l, u) = \sum_{k=l}^{u} \frac{1}{2^k} \binom{k-1}{l-1}$$

and with $v_0, v_1 ..., v_k$ the cell frequencies such thqt $v_0 + v_1 + ... + v_k = N$ the value becomes

$$X^2 = \sum_{i=0}^{k} \frac{(v_i - N\pi_i)^2}{N\pi_i}$$

and the $P$-value formula stays the same as the recent test.

### A.1.9 Maurer's Universal Statistical Test

This section looks for statistical defects coming from an ergodic stationary source which is of finite memory [65, 66]. Let us take a sequence of bits partitioned into $L$-bits blocks, $Q$ initial blocks and $K$ blocks for testing, for maximisation let $K = \left\lfloor \frac{n}{L} \right\rfloor - Q$ The test calculates the logarithm of distances and averages it

$$f_n = \frac{1}{N} \sum_{i=Q+1}^{Q+K} \log_2(i)$$

then

$$Ef_n = 2^{-L} \sum_{i=1}^{+\infty} (1 - 2^{-L})^{i-1} \log_2 i$$

For a geometric random variable $G$ with $1 - 2^{-L}$ as a parameter, the variance is

$$Var(f_n) = c(L, K)\frac{Var(\log_2 G)}{K}$$

where

$$c(L, K) = 0.7 - \frac{0.8}{L} + \left(1.6 + \frac{12.8}{L}\right) K^{-\frac{4}{L}}$$

amd finally the $P$-value is

$$err = \left(\frac{f_n - E(L)}{\sqrt{var(f_n)}}\right)$$

### A.1.10 Linear Complexity Test

This section employs linear complexity to check if there is indeed a random aspect [67], wich emanate as a principle from keystream generators, for instance, Linear Feedback Shift Registers (LFSR).

Linear complexity $L(s_n)$ of a periodic sequence $(s_n)$ over a field F is the smallest positive integer $L$ such that there are constants $c_0, ..., c_{L-1} \in$ F with

$$s_{n+L} = c_{L-1}s_{n+L-1} + ... + c_0 s_n, \; n \geq 0$$

For a positive integer $N$ the $N$-th linear complexity $L(s_n, N)$ of a sequence $(s_n)$ over F is the smallest positive integer $L$ such that there are constants $c_0, ..., c_{L-1} \in$ F satisfying

$$s_{n+L} = c_{L-1}s_{n+L-1} + ... + c_0 s_n$$

with $0 \leq n \leq N - L - 1$.

### A.1.11 Serial Test

The test checks the uniformity of distributions of patterns. Let $v_{i_1,...,i_m}$ be the freauency of string of bits $(\epsilon_1, ..., \epsilon_n, \epsilon_l, ..., \epsilon_{m-l})$, thus the function

$$\psi_m^2 = \frac{2^m}{n}\sum_{i_1...i_m}\left(v_{i_1...i_m} - \frac{n}{2^m}\right)^2 = \frac{2^m}{n}\sum_{i_1...i_m} v_{i_1...i_m}^2 - n$$

so the generalised serial statistics [68, 69, 70] are

$$\nabla \psi_m^2 = \psi_m^2 - \psi_{m-1}^2$$

and

$$\nabla^2 \psi_m^2 = \psi_m^2 - 2\psi_{m-1}^2 + \psi_{m-2}^2$$

In this case $\psi_0^2 = \psi_{-1}^2 = 0$, so $\nabla \psi_m^2$ has a $\chi^2$-distribution having $2^{m-1}$ freedom degrees, while $\nabla^2 \psi_m^2$ has $2^{m-2}$ freedom degrees.

For $m \leq \lfloor \log_2(n) - 2 \rfloor$, the $2m$ $P$-values are

$$P_{value1} = \gamma(2^{m-2}, \frac{\nabla \psi_m^2}{2})$$

and

$$P_{value2} = \gamma(2^{m-3}, \frac{\nabla^2 \psi_m^2}{2})$$

with $\nabla \psi_m^2$ converges to the $\chi^2$-distribution.

### A.1.12 Approximate Entropy Test

Let us denote by $\epsilon_1, ..., \epsilon_n$, a sequence of independent and identically distributed random variables [71], $\epsilon_j \in \{1, ..., s\} \forall j \in \{1, ..., n\}$.

For $Y_i(m) = (\epsilon_i, ..., \epsilon_{i+m-1})$, $1 \leq i \leq n - m + 1$, let

$$C_i^m = \frac{1}{n - m + 1} \#\{j : 1 \leq j \leq n - m + 1, Y_j(m) = Y_i(m)\} = \pi_i$$

and

$$\phi^{(m)} = \frac{1}{n + 1 - m} \sum_{i+1}^{n+1-m} \log C_i^m$$

knowing $C_i^m$ is the relative frequency of patterns $Y_i(m)$ in the string and $-\phi^{(m)}$ is the entropy of the empirical distribution of all $2^m$ possible patterns

$$\phi^{(m)} = \sum_{l=1}^{2^m} \pi_i \log \pi_i$$

with $\pi_i$ the relative frequency of pattern $l = (i_1, ..., i_m)$ in the string, then approximate entropy of order $m$, $m > 1$ is

$$ApEn(m) = \phi^{(m)} - \phi^{(m+1)}$$

So to sum up small values of approximate entropy reflect strong regularity, while the opposite means irregularity.

### A.1.13 Cumulative Sums (Cusum) Test

This section provides either large values meaning there are many ones or zeros in the beginning of the sequence, while the opposite means those zeros and ones are separated evenly.

Let $S_k' = X_n + ... + X_{n-k+1}$, then

$$\lim_{n\to\infty} P\left(\frac{\max_{1\le k\le n}|S_k|}{\sqrt{n}} \le z\right) = \sqrt{\frac{1}{2n}}\int_{k=-\infty}^{\infty}(-1)^k \exp\left(-\frac{(u-2kz)^2}{2}\right) du$$

$$= \frac{4}{\pi}\sum_{j=0}^{\infty}\frac{(-1)^j}{2j+1}\exp\left(-\frac{(2j+1)^2\pi^2}{8z^2}\right) = H(z), z > 0.$$

with $z = max_{1\le k\le n}\frac{|S_k'(ob)|}{\sqrt{n}}$, $P$-value is $1 - H\left(max_{1\le k\le n}\frac{|S_k'(ob)|}{\sqrt{n}}\right) = 1 - G\left(max_{1\le k\le n}\frac{|S_k'(ob)|}{\sqrt{n}}\right)$ with $G(z)$ be defined below

$$G(z) = \frac{1}{\sqrt{2\pi}}\int_{-z}^{z}\sum_{k=-\infty}^{\infty}(-1)^k \exp\left(-\frac{(u-2kz)^2}{2}\right) du$$

$$= \sum_{k=-\infty}^{\infty}(-1)^k[\Phi((2k+1)z) - \Phi((2k-1)z)]$$

$$= \Phi(z) - \Phi(-z) + 2\sum_{k=1}^{\infty}(-1)^k[\Phi((2k+1)z) - \Phi((2k-1)z)]$$

$$= \Phi(z) - \Phi(-z) - 2\sum_{k=1}^{\infty}[2\Phi((4k-1)z) - \Phi((4k+1)z) - \Phi((4k-3)z)]$$

$$\approx \Phi(z) - \Phi(-z) - 2[2\Phi(3z) - \Phi(5z) - \Phi(z)]$$

$$\approx 1 - \sqrt{\frac{4}{2\pi z}}\exp\left(-\frac{z^2}{2}\right), z \to \infty$$

with $\phi(x)$ being the standard normal distribution, and $H(z)$ series converging rapidly and used for small values of $z$ as it is the same as $G(z)$ for all $z$.

and thus [58]:

$$P\left(\max_{1\le k\le n}|S_k|\ge z\right)=1-\sum_{k=-\infty}^{\infty}P((4k-1)z<S_n<(4k+1)z)+\sum_{k=-\infty}^{\infty}P((4k+1)z<S_n<(4k+3)z)$$

### A.1.14 Random Excursions Test

This part takes an extensive look at successive sums of the binary bits as a random walk, to check for deviations from the distribution of the number of visits of any integer value [58, 72].

Let $S_k = X_1 + ... + X_k$ the random walk being a sequence of excursions

$$(i, ..., l): \; S_{i-1} = S_{l+1} = 0, \;\; S_k \neq 0 \quad for \; 1 \le k \le l$$

Consider $J$ the number of these excursions, the limiting distribution for $J$ is

$$\lim_{n\to+\infty} P\left(\frac{J}{\sqrt{n}}<z\right)=\sqrt{\frac{2}{\pi}}\int_0^z e^{-\frac{u^2}{2}}du,\; z>0$$

then if $J$ is large the test proceeds, and if the $P$-value is little

$$P(J<J(ob))\approx\sqrt{\frac{2}{\pi}}\int_0^J e^{-\frac{u^2}{2}}du=P\left(\frac{1}{2},\frac{J^2(ob)}{2n}\right)$$

### A.1.15 Random Excursions Variant Test

Let $\xi_J(x)$ be the number of visits to $x$ for $J$ excursions. $S_k$ gets a new value at every zero, $\xi_(x)$ is a sum of independent identically distributed variables [72] with the same distribution as $\xi(x) = \xi_J(x)$, so the limiting distribution of $\xi_(x)$ is normal

$$\lim_{J\to\infty} P\left(\frac{\sqrt{\xi_J(x)-J}}{J(4|x|-2)}<z\right)=\Phi(z),$$

The randomness hypothesis is not taken into consideration if the following value is small of the *P*-value

$$\operatorname{erfc}\left(\frac{|\xi_I(x)(obs) - I|}{\sqrt{2J(4|x| - 2)}}\right)$$

# Appendix B

# Baptista Method encryption

In this part, the attention is solely on the programming aspect of the Baptista method in chapter 2.

## B.1 Code for text encryption

Below is presented the source code script of the algorithm programmed using Python.

```
Nmax , x0 , b, xmin , xmax , eps ;
message_a_code , messagecode , message ;
for char in message ;
        message_a_code . append (ord (char))
xn= x0 ;
message_crypt =[ ];
for i in message_a_code ;
    N=0;
    while (xn >= xmin + eps * i) and (xn < xmin + eps * (i +1)):
            xn= b*xn*(1-xn);
            N=N+1;
    while (xn <= xmin + eps * i) or (xn > xmin + eps * (i +1)):
            xn= b*xn*(1-xn);
            N=N+1;
    message_crypt . append (N);
for char in message_crypt :
     messagecode . append (chr (char));
xn = x0;
message_a_decode , messagedecode ;
for i in range (1, len (message_crypt )+1):
```

```
      for k in range (1 ,
            message_crypt[i-1]+1):
            xn= b*xn*(1-xn);
      for l in range(0,256):
            if (xn>=xmin+eps*l) or(
               xn<xmin+eps*(l+1)):
                break ;
      message_a_decode.append(l);
msgdecrypt= message_a_decode;
for   char in   msgdecrypt:
       messagedecode.append(chr(char))
```

## B.2 Code for image encryption

The Python source code script to encrypt the image using the Baptista method is shown below.

```
mu, x0, s, interval_min, interval_max;
r= (interval_max-interval_min) / s;

def f(xn):
       return mu * xn * (1-xn)
img,pix;
width, height= img.width, img.height;

rgb, enc_rgb;
for i in range(width):
      for j in range ( height ):
            rgb.append(pix[i,j])

x=x0;
for triplet in rgb:
      pixel;
      for val in triplet:
            x_min = interval_min + r * val;
            x_max= interval_min +
                        r * (val+1);
            N=0;
            while x < x_min or x > x_max:
                      x = f(x);
                      N+=1;
```

```
            pixel . append ( N % s)
        enc_rgb.append(tuple(pixel));

encImage  =  Image.new(
                img.mode(widthmheight));
encImage.putdata(enc_rgb)
```

# Appendix C

# Encryption technical details

## C.1 Egyptian product PRNG

This section, helps understand certain components of chapter 4 and 5, contributing to integral approach of the papers.

### C.1.0.1 Text Stream cipher algorithm module

Let $m$ be the message with $k$ the key generated from the previously mentionned PRNG based on egyptian fractions, the plain text is $p_1, p_2, ..., p_n$ , and $C_1, C_2, ..., C_n$ the cipher.
Let $E$ be the function

$$E(k, x) = (x \times k_1) \oplus k_2 \oplus k_3 \oplus k_4$$

where $k$ is the key of 128-bits $k = k_1k_2k_3k_4$ for 32-bits $k_i$ , $x$ is a 32-bit string, $\oplus$ being the bit-wise exlusive-or, $+$ and $\times$ are $mod 2^{32}$ addition and multiplication.
For the encryption of every 32 bits from the plaintext, here are the steps to follow:

1) 128-bit key sequence is generated by our PRNG and then $k_i = k_1k_2k_3k_4$ for 32-bit $k_i$.

2) The value of $A_i$ is

$$A_i = E(k_i, C_{i-1} \oplus P_{i-1})$$

where $x$'s place is taken by , $C_{i-1} \oplus p_{i-1}$, with , $C_{i-1}$ and $p_{i-1}$ equal to 32 bits of the rpevious cleartext and encrypted text, respectively.

3) The value of $B_i$ is

$$B_i = E(k_i, A_i \oplus p_{i-2})$$

the value of $x$ is substitued by $A_i \oplus p_{i-2}$, with $p_{i-2}$ equal to 32 bits of the plaintext and $A_i$ already computed before.

4) This step the value of $D_i$ becomes

$$D_i = E(k_i, B_i \oplus C_{i-2})$$

the value of $x$ equals $B_i \oplus C_{i-2}$, with $C_{i-2}$ equaling 32 bits of the ciphertext before and $B_i$ already computed.

5) A ciphertext of 32 bits is given by

$$C_i = p_i \oplus D_i$$

with $p_i$ equal to 32 bits of the cleartext and $D_i$ was already computed.

These steps from 2 to 5 can be gathered in one formula

$$C_i = p_i \oplus E(k_i, E(k_i, E(k_i, C_{i-1} \oplus p_{i-1} \oplus p_{i-2}) \oplus C_{i-2})$$

and from step 2 to 4 the value of $C_0, C_{-1}, p_0, p_{-1}$ can be considered equal to $k_1, k_2, k_3, k_4$. The decryption and encryption are both identical except for the last step where

$$p_i = C_i \oplus D_i$$

during the decryption.

### C.1.0.2 Image Encryption module

This module in charge of the image encryption using our PRNG, relies on confusion processing which is heavily based on the *XOR* operation of a pixel matrix of the image with the number of pseudorandom sequence $x$, and also a scrambling processing

method which is the use of the pseudorandom sequencegiven by our egyptian fractions predefiend algorithm.

The confusion method is stated as below:

1) Let the initial condition $x_0$ be chosen randomly into the our encryption Engel expansion predefined algorithm, and compute the pseudorandom sequence number $x$.

2) Compute the elements of the pseudorandom sequence in $(x_i \times 256) mod\ 256$, then do a binary conversion to get a $M \times N$ sequence.

3) The pixel color component sequence of the image is encrypted and obtain the color component sequence vector $G$ of the image.

4) $g_i$ being the first element in $G$, the $XOR$ operation follows $x \oplus g_i$, and for the rest of the elements of $G$ is follows

$$i(k) = x^{(k)} \oplus \left\{ \left\lfloor x^{(k)} + g_k \right\rfloor mod N \right\} \oplus i(k+1)$$

with $k$ being the $k$ image pixel.

5) The pixel sequence is now reversed and the $M \times N$ elements are adjusted to the first position, and $M \times (N-1)$ to the second one, then the prior formula performs the obfuscation.

As for the scrambling process it uses the steps below:

1) The pseudorandom sequence number $x$ of the previous process is now considered the final set $x$ of the pseudorandom sequence.

2) $i$ is homogenised and get an empty vector $Y$ of size $M \times N$, the corresponding $x$ is extended to the integer domain space of $(0, M \times N)$ when $x$ is homogenised, then $y$ is the output vector.

3) The pixel is scrambled for $i$ after the confusion method while the encrypted image $i$ and $y$ vector are both used in the process. The value of the $i$-th pixel and the one of the $y_i$-th pixel in $i$ are swapped.

4) The output is put through positive order confusion and reverse order confusion following to the steps 4 and 5 in the confusion method, thus getting the terminal encrypted image *i*.

The decryption is starts by scrambling then confusing, as for confusing inverse processing here is the method

$$g_k = \{x(k) \oplus i(k) \oplus i(k-1) - x(k)\} \, mod \, N$$

# Appendix D

# Technical analysis: Egyptian product encryption

## D.1 Image encryption analysis

Images in JPEG format are color images with 24-bit color resolution. 24-bit image is the combination of R, G and B colors (8-bits each) ranged from 0 to 255, where 0 being black and 255 the red, green and blue respectively in each color domain. 8-bits R, G and B color channel value are concatenated to get a 24-bit RGB color value in JPEG image format. R, G and B color channels are 8-bit in resolution converted back to 24-bit color channel by making remaining color channels as black, and each color channel has its own histogram and the color image can be viewed into different color channels each on a separate segment.

### D.1.1 Distribution histogram

The three RGB components of our encrypted image had uniform histograms. The taller the hump in the graph, the more pixels reside at that particular tonal range. Based upon the histogram analysis, we are now sure statistical attacks aren't effective, pixel values being altered using our encryption scheme.

### D.1.2 Grayscale histogram

The values of the three channels of each pixel of the colored image are used to determine what the grayscale of the pixel is. The histogram of an image in image processing refers to a histogram of the pixel intensity values. The grayscale histogram is a graph showing the number of pixels in an image at each different intensity value found in that image. The image histograms showed colors by tone on the abscissas and the number of pixels corresponding to it on the ordinates. 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.

### D.1.3 Scatter plot

In chapter 5, this scatter plot in root cause analysis employs dots as values for numeric variables: their positions on the axis indicates values for an data points, in order to depict the relation among variables. In the present case, this tool is a three-dimensional graphical representation for a set of data, determining the correlation between variables, the stronger the correlation, the tighter the points will be closer to the line. However statistically speaking correlation does not imply causation, as the pattern may happen to be coincidental. For the third variable that indicates categorical values, the encoding is through point color. When the points are scattered over the chart, the degree of correlation between the variables is less and vice-versa. The value of a variable determines the relative position of the symbol along the $X$-axis and the value of a second variable determines the relative position of the symbol along the $Y$-axis, and so on for the $Z$-axis.

The first plot related to the RGB composure for the clear image shows strong positive correlation, since the value of $Z$ increases as the values of $X$ and $Y$ increase, while on the side of the cipher image the plot tends to display a weak positive correlation from up far, since most will guess the value of $Z$ increases slightly as the values of $X$ and $Y$

increase, but as you zoom in closer and closer you see that there exist no correlation, as no connection is demonstrated among the variables.

Finally wou could say for sure, the scatter diagrams for RGB distribution of the pixels for both the clear image and image cipher showed the decorrelation coefficients between the pixels.

# Appendix E

# The wavelet scalogram

To prove a certain aspect of chaos, there is plenty of approaches and methods in the midst of this variety of chaos definitions, there is a method that was added as proof in chapter four when testing our constructed function, therefore we define the wavelet scalogram [90] and some of its properties.

The continuouos wavelet transform (CWT) [91] establishes a time-frequency decomposition of $f$ in the time-frequency plane, that is why one of the key factors that revamped Wavelet theory as a substitute technique to the windowed Fourier transform (WFT) is the possibility to study time and frequency concomitantly with the parameters of time $u$ and scale $s$ in the CWT, with a resolution relying on the scale of interest.

- The **windowed Fourier transform** is defined by

$$Sf(u,\xi) = < f, g_{u,\xi} > = \int_{-\infty}^{+\infty} f(t)g(t-u)e^{-i\xi t}dt$$

  It uses an atom which is the product of a sinusoidal wave with a finite energy symmetric window $g$. The windowed Fourier transform family of atoms is obtained by time translations and frequency modulations of the original window:

$$g_{u,\xi} = e^{i\xi t}g(t-u)$$

  This atom has a frequency center $x$ and is symmetric with respect to $u$.

- A function $\psi \in L^2(\mathrm{R})$ centered in the neighborhood of $t = 0$, with zero average, i.e. $\int_{\mathrm{R}} \psi = 0$ and normalized $||\psi|| = 1$ is said to be a **wavelet** function. A family of time-frequency atoms $\psi_{u,s}$ can be obtained via Scaling $\psi$ by a positive quantity $s$, and translating it by $u \in \mathrm{R}$, which is described in the formula below

$$\psi_{u,s}(t) := \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right), \quad u \in \mathrm{R}, s > 0.$$

For $f \in L^2(\mathrm{R}$, the continuous wavelet transform (CWT) of $f$ at time $u$ and scale $s$ is

$$Wf(u,s) := < f, \psi_{u,s} > = \int_{-\infty}^{+\infty} f(t)\psi^*_{u,s}(\mathrm{t})\, dt,$$

- The **scalogram** of $f$ is the function represents the energy of $Wf$ at a scale $s$:

$$(s) := \left| Wf(s,u) \right|_{\mathrm{S}} := \left( \int_{-\infty}^{+\infty} ||Wf(s,u)\,^2 du \right)^{\frac{1}{2}}$$

For all scale $s$ there is $S(s) \geq 0$ , if $S(s) > 0$ it is said that the signal $f$ has details at scale $s$.

Thus, the scalogram [92] allows the detection of the most representative scales or frequencies of a signal, that is, the scales that contribute the most to the total energy of the signal.

The scalogram detects the representative scales of a signal that contribute the most to the total energy of the signal.

In a time interval $[t_0, t_1]$, the windowed scalogram is given by

$$\mathrm{S}_{[t_0,t_1]}(s) := \left| \left| Wf(s,u)_{[t_0,t_1]} \right| \right|_{\mathrm{S}} = \left( \int_{t_0}^{t_1} |Wf(s,u)|^2 du \right)^{\frac{1}{2}}$$

- A function $f : X \to V$ on a topological space with values in a vector space $V$ or a pointed set with the basepoint (called 0) has compact support or is **compactly supported** if the closure of its support, the set of points where it is non-zero, is a compact subset, i.e. the subset $\overline{f^{-1}(V \setminus \{0\})}$is a compact subset of $X$.

*Remark* E.1. A signal $f$ is ready to be numerically studied if first $f$ is defined over a finite time interval $I = [a, b]$. Boundary emerge from the support of $\psi_{u,s}$ when it overlaps $t = a$ or $t = b$, which we can overcome by methods with large amplitude coefficients at the boundary or computationally complex. When the wavelet function $\psi$ is compactly supported and the interval is not small, you can study wavelet coefficients which were not affected by boundary effects and the inner scalogram is a solution.

**Definition E.2.** The **inner scalogram** of $f$ at scale $s$ is

$$\mathrm{S}^{inner}(s) := \mathrm{S}_{J(s)}(s) = ||Wf(s,u)_{J(s)}|| = \left(\int_{c(s}^{d(s)} |Wf(s,u)|^2 du\right)^{\frac{1}{2}}$$

with $J(s) = [c(s), d(s)] \subseteq I$ the maximal subinterval in $I$ for which the support of $\psi_{u,s}$ is included in $I$ for all $u \in J(s)$. The length of $I$ shoud be big in order for $J(s)$ to be nonempty and not very small,meaning $b - a >> sl$, with $l$ the length of the support of $\psi$. The values of the inner scalogram are incomparable at distinct scales since $J(s)$ is scale $s$ dependent, from which the necessity the normalisation of the inner scalogram

$$\overline{\mathrm{S}}^{inner}(s) = \frac{\mathrm{S}^{inner}(s)}{(d(s) - c(s))^{\frac{1}{2}}}$$

**Theorem E.3.** *Let $f : \mathrm{R} \to C$ be a $T$-periodic function in $L^2([0, T])$, and let $\psi$ be a compactly supported wavelet. Then $Wf(u, 2T) = 0$ for all* R.

**Corollary E.4.** *Let $f : I = [a, b] \to \mathrm{C}$ a $T$-periodic function in $L^2([a, a + T])$. If $\psi$ is a compactly supported wavelet, then the normalized inner scalogram of $f$ at scale $2T$ is zero.*

*Remark* E.5. The scalogram of a $T$-periodic signal vanishes at all $2kT$ scales $k \in \mathrm{N}$, you can work with scales greater than a fundamental scale $s_0$, so a signal with details at an arbitrarily large scale is non-periodic. The most representative scale of a signal $f$ will be the scale $s_{max}$ for which the scalogram reaches its maximum value. If $S(s)$ never becomes too small compared to $S(s_{max})$ for $s > s_{max}$, then the signal is numerically non-periodic in $[s_0, s_1]$.

The **scale index** of $f$ in the scale interval $[s_0, s_1]$ is

$$i_{scale} := \frac{\mathbb{S}(s_{min})}{\mathbb{S}(s_{max})}$$

with $s_{max}$ is the smallest scale such that $S(s) \leq S(s_{max})$, $\forall \ s \in [s_0, s_1]$, and $s_{min}$ the smallest scale such that $\mathbb{S}(s_{min}) \leq \mathbb{S}(s) \ \forall \ s \in [s_{max}, s_1]$.

A bounded signal is considered chaotic [87] if

a) it has sensitive dependence on the initial conditions

b) is non-periodic or does not converge to a periodic orbit.

# Appendix F

# The categorical, tensorial and algebraic approach

In this part, the attention is shined upon the mathematicals tools, needed when using Lossy Trapdoor Function for the Decisional Diffie-Hellman problem in order to establish a well put together encryption, where we consider a tensor space and tensor products replacing the ordinary used matrices.

**General linear group** If $V$ is a vector space over the field $F$, the **general linear group** of $V$, written $GL(V)$ or $Aut(V)$, is the group of all automorphisms of $V$, i.e. the set of all bijective linear transformations $V \to V$, together with functional composition as group operation. If $V$ has finite dimension $n$, then $GL(V)$ and $GL(n, F)$ are isomorphic. The isomorphism is not canonical; it depends on a choice of basis in $V$. Given a basis $(e_1, ..., e_n)$ of $V$ and an automorphism $T$ in $GL(V)$, we have then for every basis vector ei that

$$T(e_i) = \sum_{j=1}^{n} a_{ij} e_j$$

for some constants $a_{ij}$ in $F$; the matrix corresponding to $T$ is then just the matrix with entries given by the $a_{ij}$.

**Galois field** A *Galois field* [93] is a field with a finite field order, i.e. number of elements, also called a finite field. The order of a finite field is always a prime or a power

of a prime. For each prime power, there exists exactly one (with the usual caveat that exactly one means exactly one up to an isomorphism finite field $GF(p^n)$, often written $\mathrm{F}(p^n)$.

$GF(p)$ is called the prime field of order $p$, and is the field of residue classes modulo $p$, where the $p$ elements are denoted 0, 1, ..., $p-1$. $a = b$ in $GF(p)$ means the same as $a = b(modp)$.

**Tensor product of vectors [94]** The vectors $x, y$ are of length $M, N$, respectively, their *tensor product* $x \otimes y$ is the $M \times N$-matrix given by $(x \otimes y)_{ij} = x_i y_j$, i.e. $x \otimes y = xy^T$.

*Remark* F.1. Consider the standard bases for $R^M$ and $R^N$ to be

$$\boldsymbol{\varepsilon}_M = \{e_i\}_{i=0}^{M-1} \text{ and } \boldsymbol{\varepsilon}_N = \{e_i\}_{i=0}^{N-1}$$

Then the standard basis for the set of $M \times N$-matrices $L_{M,N}(\mathrm{R})$ is

$$\boldsymbol{\varepsilon}_{M,N} = \{e_i \otimes e_j\}_{(i,j)=(0,0)}^{M-1,N-1}$$

- Let $S : \mathrm{R}^M \longrightarrow \mathrm{R}^M$ and $T : \mathrm{R}^N \longrightarrow \mathrm{R}^N$ be matrices.

  The linear mapping $S \otimes T : L_{M,N}(\mathrm{R}) \longrightarrow L_{M,N}(\mathrm{R})$ is the tensor product of matrices $S$ and $T$ when given by the linear extension

$$(S \otimes T)(e_i \otimes e_j) = (Se_i) \otimes (Te_j)$$

- Consider a linear space $E$ over a field $K$. The vector space tensor product $\otimes_{\lambda=1}^{k} E$ is said to be a **tensor space** of degree $k$. A tensor space of type $(r, s)$ is a vector space tensor product between $r$ copies of vector fields and $s$ copies of the dual vector fields, called *one-forms*.

**Example 4.**

$$T^{(3,1)} = TM \otimes TM \otimes TM \otimes T^* M$$

is the vector bundle of $(3,1)$ tensors on a manifold $M$. Tensors of type $(r,s)$ form a vector space.

**Categories**

$$Hom(A, G) = h : A \longrightarrow G | h \text{ homomorphism}$$

$Hom(A, G)$ is a group under function addition. The dual homomorphism to $f : A \to B$ is the homomorphism

$$f^* : Hom(A, G) \leftarrow Hom(B, G)$$

defined by

$$f^*(\psi) = \psi \circ f : A \to B \to G$$

- A **category** [95] $\mathbb{C}$ consists of objects $A, B, C, ...$ and morphisms $f : A \to B$ (where $A$ and $B$ are objects).

  If f: $A \to B$ and $g : B \to C$ are morphisms, we have a notion of composition, i.e. there is a morphism $gf = g \circ f : A \to C$, such that:

  - If $f : A \to B$, $g : B \to C$, $h : C \to D$, then $(hg)f = h(gf)$. (*Associativity*)
  - For each object $A$ there is a morphism $1_A : A \to A$ such that for each morphism $f : A \to B$, we have $f1_A = 1_B f = f$. (*Identity*)

- A proper class is a class that is not a set (class is a collection of sets). A category is *locally small* if each of its hom-sets is a small set, i.e. is a set instead of a proper class.

- Consider two objects $x$ and $y$ in a (locally small) category, there is a set $hom(x, y)$ (hom-set), whose elements are morphisms from $x$ to $y$. Given a morphism $f$ in this hom-set, we write $f : x \to y$ to indicate that it goes from $x$ to $y$. A morphism is in general between objects in any $n$-category.

*Remark* F.2. For each objects $(A, B)$ , the collection of morphisms $f : A \to B$ should be a set not a proper class.

**Example 5.** Here are a few object examples:

- *Sets*: The objects are sets and the morphisms are functions.
- *Groups*: The objects are groups and the morphisms are group homomorphisms.
- *Rings* : The objects are rings and the morphisms are ring homomorphisms.

**Functors**

- A **functor** is a function between categories mapping objects to objects and morphisms to morphisms.
- A functor $F$ is **covariant** if it preserves the directions of arrows, meaning every arrow $f : A \longrightarrow B$ is mapped to an arrow $F(f) : F(A) \longrightarrow F(B)$.
- A functor $F$ is called **contravariant** if it reverses the directions of arrows, which means, every arrow $f : A \longrightarrow B$ is mapped to an arrow $F(f) : F(B) \longrightarrow F(A)$.

The assignment of the dual homomorphism

$$A \longrightarrow Hom(A, G) \text{ and } f \longrightarrow f^*$$

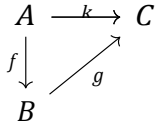
is a contravarient functor from the category of abelian groups and homomorphisms to itself, because if $i : A \longrightarrow A$ is the identity map on $A$, then

$$i^*(\psi) = \psi \circ i = \psi$$

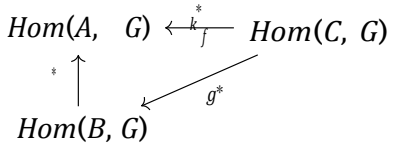
is the identity map on $Hom(A, G)$.
And if we have $f : A \longrightarrow B$, $g : B \longrightarrow C$, $\psi : C \longrightarrow G$, then

$$(f^* \circ g^*)(\psi) = f^*(g^*(\psi)) = f^*(\psi \circ g) = \psi \circ g \circ f = (g \circ f)^*(\psi)$$

This means when this diagram commutes



the following diagram commutes also



# Bibliography

[1] J. Kepler 1621 Epitome Astronomiae Copernicanae.

[2] F. Ragep, Jamil 2001 Tusi and Copernicus: The Earth's Motion in Context. Science in Context, Cambridge University Press, 14 (12): 145163.

[3] D. Lynden-Bell, R. Lynden-Bell 1997 On the Shapes of Newton's Revolving Orbits. Notes and Records of the Royal Society of London. 51 (2): 195198.

[4] R. A. Broucke 1969 Periodic Orbits in the Elliptic Restricted Three-Body Problem. Pasadena, California, Jet Propulsion Laboratory, California Institute of Technology.

[5] J. Laskar 1995 The Stability of the Solar System from Laplace to the Present. In General History of Astronomy, R. Taton , C. Wilson eds, vol. 2B, pp. 240-248.

[6] Y. B. Pesin 1977 Characteristic Lyapunov Exponents and Smooth Ergodic Theory. Russian Math. Surveys. 32 (4): 55114.

[7] H. Poincaré 1892 Les Méthodes Nouvelles de la Mécanique Céleste. Paris: Gauthier-Villars.

[8] J. Laskar 2012 Is the Solar System Stable ? arXiv:1209.5996

[9] V. Arnold, A. Avez 1967 Problèmes Ergodiques de la Mécanique Classique (in French). Paris: Gauthier-Villars.

[10] R. Honsberger 1985 The Matrix Q. Mathematical Gems III. Washington, DC: Math. Assoc. Amer., pp. 106-107.

[11] H. L. D. S. Cavalcante, M. Oria, D. Sornette, E. Ott, D. J. Gauthier 2013 Predictability and suppression of extreme events in complex systems. arXiv:1301.0244.

[12] F. Lombardi 2000 Chaos Theory, Heart Rate Variability, and Arrhythmic Mortality. Circulation, 101 :8,2000. <http://circ.ahajournals.org/cgi/contentifullI10111/8>.

[13] D. H. Bailey, R. E. Crandall 2001 On the Random Character of Fundamental Constant Expansions. Experiment. Math. 10 (2) 175 - 190.

[14] E. N. Lorenz 1972 Predictability: Does the Flap of a Butterflys Wings in Brazil Set off a Tornado in Texas. American Association for the Advancement of Science.

[15] D. Ruelle and F. Takens 1971 On the nature of turbulence, Comm. math. Phys. 20, 167-192; 23, 343-344.

[16] S. Smale 1967 Differentiable dynamical systems. Bulletin of the American Mathematical Society. 73 (6): 747-817.

[17] G. Birkhoff 1927 Dynamical Systems. American Mathematical Society Colloquium Publication, 9.

[18] D. T. Kaplan, R. J. Cohen 1990 Is fibrillation chaos?. Originally published 1 Oct https://doi.org/10.1161/01.RES.67.4.886 Circulation Research. 67:886892.

[19] I. Cherkaoui and F. Zinoun (2018) Encrypting with Egyptian Fractions: From Number Theory to Chaos-Based Cryptography. International Conference on Applied Mathematics, Fez.

[20] I. Cherkaoui and F. Zinoun (2019) On the Egyptian Product-Based Encryption. International Conference on Research in Applied Mathematics and Computer Science, Casablanca.

[21] K. Chetioui, G. Orhanou, H. Bensaid, I. Cherkaoui and Y. Chibi (2019) Formal Verification of Confidentiality in DNSSEC and E-DNSSEC Protocols using pi-calculus and ProVerif. International Workshop on Emerging Networks and Communication, Coimbra, Portugal. (Not included in thesis).

[22] I. Cherkaoui and F. Zinoun (2021) On the use of Egyptian fractions for stream ciphers. Journal of Discrete Mathematical Sciences and Cryptography, Volume 26, Issue 1, p 139-152.

[23] I. Cherkaoui (2021) Diffie-Hellman Multi-Challenge using a New Lossy Trapdoor Function Construction. IAENG International Journal of Applied Mathematics. 51(3), pp. 17

[24] R. L. Devaney 1989 An Introduction to Chaotic Dynamical Systems. Addison-Wesley.

[25] M. W. Hirsch, S. Smale 1974 Differential equations, dynamical systems and linear algebra, Academic Press.

[26] J. Banks, J. Brooks, G. Cairns, P. Stacey 1992 On devaneys definition of chaos. Amer. Math. Monthly, 99 :332334.

[27] A. N. Sharkovskii 1964 Co-existence of cycles of a continuous mapping of the line into itself. Ukrainian Math. J. 16: 6171.

[28] T. Li, J. A. Yorke 1975 Period Three Implies Chaos. The American Mathematical Monthly, Vol. 82, No. 10, pp. 985-992 (8 pages). Published By: Taylor & Francis, Ltd.

[29] M. Vellekoop and R. Berglund 1994 On Intervals, Transitivity = Chaos. The American Mathematical Monthly Vol. 101, No. 4, pp. 353-355 (3 pages) Published By: Taylor and Francis, Ltd.

[30] J. C. Sprott 2003 Chaos and Time-Series Analysis. Oxford University Press.

[31] D. Buell 2021 Fundamentals of Cryptography Introducing Mathematical and Algorithmic Foundations. UTICS.

[32] O. Goldreich 2001 The Foundations of Cryptography, Volume 1. Cambridge University Press.

[33] F. I. Blake, T. Garefalakis 2004 On the complexity of the discrete logarithm and Diffie-Hellman problems. Journal of Complexity. Festschrift for Harald Niederreiter, Special Issue on Coding and Cryptography. 20(2): 148170.

[34] R. Rivest, A. Shamir, L. Adleman 1978 A Method for Obtaining Digital Signatures and Public-Key Cryptosystems. Communications of the ACM. 21 (2): 120126.

[35] W. Diffie, M. E. Hellman 1976 New Directions in Cryptography. IEEE Transactions on Information Theory. 22 (6): 644654.

[36] R. C. Merkle 1978 Secure communications over insecure channels. Communications of the ACM , vol. 21, pp. 294-299.

[37] A. Joux, K. Nguyen 2003 Separating decision Diffie-Hellman from computational Diffie-Hellman in cryptographic groups. Journal of Cryptology, 16 (4): 239247.

[38] F. Engel 1913 Entwicklung der Zahlen nach Stammbruechen. Verhandlungen der 52. Versammlung deutscher Philologen und Schulmaenner in Marburg, pp. 190191.

[39] T. A. Pierce 1929 On an algorithm and its use in approximating roots of algebraic equations. Am. Math. Monthly, 36 (10): 523-525. JSTOR 2299963.

[40] C. Kraaikamp, J. Wu, 2004 On a new continued fraction expansion with non-decreasing partial quotients. Monatshefte für Mathematik 143 (4): 285298, doi:10.1007/s00605-004-0246-3.

[41] C. Naux 1966 Histoire des logarithmes de Neper à Euler, A. Blanchard.

[42] H.Lebesgue 1987 Leçons sur les constructions géométriques. Jacques Gabay.

[43] G. Valiron, 1948 Cours d'analyse mathématiques. Masson.

[44] G. Stratemeyer 1931 Entwicklung positiver Zahlen nach Stammbrüchen (Dissertation). Mitteil, des Mathem. Seminars d. Universität Gieben. Bd II, Heft 20.

[45] L. Euler 1796 Introduction à l'analyse infinitésimale. Chez Barrois, l'an IV-V, volumes 1 et 2, 788 pages.

[46] A. A. Markov 1906 Rasprostranenie zakona bol'shih chisel na velichiny, zavisyaschie drug ot druga. Izvestiya Fiziko-matematicheskogo obschestva pri Kazanskom universitete, 2-ya seriya, tom 15, pp. 135156.

[47] J. Galambos, J 1976 Representations of real numbers by infinite series. Springer-Verlag.

[48] A. Rényi 1962 A new approach to the theory of Engel's series. Ann. Univ. Sci. Budapest, Sectio Math. 5, p. 25-32.

[49] J. Neveu 1964 Bases mathématiques du calcul et des probabilités. Masson.

[50] G.H. Hardy, E.M. Wright 1979 An introduction to the theory of numbers. Oxford Science Publications.

[51] L.Fibonacci 2002 Liber Abaci. translated by Sigler, Laurence E., Springer-Verlag, 2002, ISBN 0-387-95419-8 .

[52] A. Rukhin, J. Soto, J. Nechvatal, M. Smid, E. Barker, S. Leigh, M. Levenson, M. Vangel, D. Banks, A. Heckert, J. Dray 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.

[53] R. Sturman 2012 The Role of Discontinuities in Mixing. Advances in Applied Mechanics

[54] K. L. Chung 1979 Elementary Probability Theory with Stochastic Processes. New York: Springer-Verlag (especially pp. 210-217).

[55] N. Maclaren 1993 Cryptographic Pseudo-random Numbers in Simulation. Cambridge Security Workshop on Fast Software Encryption. Cambridge, U.K.: R. Anderson, pp. 185-190.

[56] J. D. Gibbons 1985 Nonparametric Statistical Inference, 2nd ed. New York: Marcel Dekker, (especially pp. 50-58).

[57] F. N. David, D. E. Barton 1962 Combinatorial Chance. New York: Hafner Publishing Co., p. 230

[58] P. Revesz 1990 Random Walk in Random and Non-Random Environments. Singapore: World Scientific.

[59] G. Marsaglia 1996 DIEHARD: a battery of tests of randomness. http://www.stat.fsu.edu/pub/diehard/.

[60] I. N. Kovalenko 1972 Distribution of the linear rank of a random matrix. Theory of Probability and its Applications. 17, pp. 342-346.

[61] G. Marsaglia, L. H. Tsay 1985 Matrices and the structure of random number sequences. Linear Algebra and its Applications. Vol. 67, pp. 147-156

[62] R. N. Bracewell 1986 The Fourier Transform and Its Applications. New York: McGraw-Hill.

[63] A. D. Barbour, L. Holst, and S. Janson 1992 Poisson Approximation. Oxford: Clarendon Press (especially Section 8.4 and Section 10.4).

[64] O. Chrysaphinou, S. Papastavridis 1988 A Limit Theorem on the Number of Overlapping Appearances of a Pattern in a Sequence of Independent Trials. Probability Theory and Related Fields, Vol. 79 (1988), pp. 129-143.

[65] U. M. Maurer 1992 A Universal Statistical Test for Random Bit Generators. Journal of Cryptology. Vol. 5, No. 2, 1992, pp. 89-105

[66] J. S Coron, D. Naccache 1998 An Accurate Evaluation of Maurer's Universal Test. Proceedings of SAC '98 (Lecture Notes in Computer Science). Berlin: Springer-Verlag, 1998.

[67] H. Gustafson, E. Dawson, L. Nielsen, W. Caelli 1994 A computer package for measuring the strength of encryption algorithms. Computers and Security. 13, pp. 687-697.

[68] I. J. Good 1953 The serial test for sampling numbers and other tests for randomness. Proc. Cambridge Philos. Soc.. 47, pp. 276-284.

[69] M. Kimberley 1987 Comparison of two statistical tests for keystream sequences. Electronics Letters. 23, pp. 365-366.

[70] D. E. Knuth 1998 The Art of Computer Programming. Vol. 2, 3rd ed. Reading: Addison-Wesley, Inc., pp. 61-80.

[71] A. Rukhin 2000 Approximate entropy for testing randomness. Journal of Applied Probability. Vol. 37.

[72] M. Baron, A. L. Rukhin 1999 Distribution of the Number of Visits For a Random Walk. Communications in Statistics: Stochastic Models. Vol. 15, pp. 593-597

[73] T. J. Stieltjes 1884 Quelques recherches sur la théorie des quadratures dites mécaniques, Ann. Sci. Ecole Norm. Sup. (3) 1, 409426.

[74] C. Peikert and B. Waters 2008 Lossy trapdoor functions and their applications. In STOC.

[75] L. Kocarev, S. Lian 2011 Chaos-Based Cryptography, Theory, Algorithms and Applications. Studies in Computational Intelligence, springer, Volume 354.

[76] G. Alvarez, S. Li 2006 Some basic cryptographic requirements for chaos-based cryptosystems. International journal of bifurcation and chaos, World Scientific.

[77] D. Arroyo, G. Alvarez, V. Fernandez 2009 A basic framework for the cryptanalysis of digital chaos-based cryptography. 6th International Multi-Conference on Systems, Signals and Devices, Publisher: IEEE.

[78] F. Dridi, S. El Assad, W. E. H. Youssef, M. Machhout, R. Lozi 2021 The Design and FPGA-Based Implementation of a Stream Cipher Based on a Secure Chaotic Generator. Applied Sciences, Vol. 11, Pages 625.

[79] H. Hu, Y. Yu, Y. Zhao 2017 A note on approximation efficiency and partial quotients of Engel continued fractions. International Journal of Number Theory. Vol. 13, No. 9 (2017) 24332443, World Scientific Publishing Company.

[80] A. Lasota, M.C Mackey 1994 Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics, Second Edition, Applied Mathematical Sciences, vol. 97.

[81] J. D. Barrow 2000 Chaos in Numberland: The secret life of continued fractions, plus.maths.org.

[82] L. Euler 1987 Introduction à l'analyse infinitésimale. ACL-Editions, 1987.

[83] C. Ganatsiou 1997 On the stochastic behaviour of the digits in the modified Engel-type alternating series representations for real numbers. IBSG Proceedings 5, IProceedings of the Workshop on Global Analysis, Differential Geometry, Lie Algebras, Aristotle University of Thessaloniki, July 1997, Balkan Society of Geometry, Geometry Balkan Press, Bucharest - Romania, 33-39.

[84] G.D. Birkhoff 1931 Proof of the ergodic theorem. Proc. Nat. Acad. Sci. USA, 17, pp. 656660.

[85] Y. Liu, J. Wu 2001 Hausdorff dimensions in Engel expansions. Mathematics, Acta Arithmetica.

[86] G. Edgar 1991 Measure, Topology and Fractal Geometry. Springer-Verlag New York Inc.

[87] S. H. Strogatz 1994 Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry and Engineering. Addison-Wesley Publishing.

[88] A. M. Ostrowski 1973 Solutions of Equations in Euclidean and Banach Spaces. Academic Press, New York.

[89] P. Verhulst 1838 Notice sur la loi que la population poursuit dans son accroissement. Correspondance mathématique et physique, no 10, 1838, p. 113-121.

[90] R. Benitez,V. J. Bolos, M. E. Ramirez 2010 A wavelet-based tool for studying non-periodicity. Comput. Math. Appl. 60, no. 3, 634641.

[91] C. Chandre, S. Wiggins, T. Uzer 2003 Time-frequency analysis of chaotic systems. Phys. D181, no. 3-4, 171196.

[92] B. Donald, P. Walden, A. T. Walden 2000 Wavelet Methods for Time Series Analysis. Cambridge University Press.

[93] A. I. Skopin, 1994 Galois field. Encyclopedia of Mathematics, EMS Press.

[94] N. Bourbaki 1989 Elements of mathematics, Algebra I. Springer-Verlag.

[95] S. Awodey 2006 Category theory. Oxford logic guides, vol. 49, Oxford University Press.

[96] S. Chen, S. Feng, W. Fu, Y. Zhang 2014 Logistic Map: Stability and Entrance to Chaos To cite this article. J. Phys. Conf. Ser, 012009.

[97] M. J. Feigenbaum 1980 The Metric Universal Properties of Period Doubling Bifurcations and the Spectrum for a Route to Turbulence. Ann. New York. Acad. Sci. 357, 330-336.

[98] A. Akhshani, A. Akhavan, A. Mobaraki, S.C. Lim, Z. Hassan 1994 Pseudo random number generator based on quantum chaotic map. Commun Nonlinear Sci Numer Simulat. 19, 101-111.

[99] V. J. Bolo, R. Benitez 2013 The wavelet scalogram in the study of time series XXIII Congreso de Ecuaciones Diferenciales y Aplicaciones XIII Congreso de Matematica Aplicada. pp. 1-8.

[100] M. Bellare, A. Boldyreva, S. Micali 2000 Public-key encryption in a multi-user setting: Security proofs and improvements. In Eurocrypt.

[101] R. Canetti, S. Halevi, J. Katz 2004 Chosen-ciphertext security from identity-based encryption. In Eurocrypt.

[102]X. Boyen, Q. Li 2017 All-but-many lossy trapdoor functions from lattices and applications. In Crypto.

[103] S. M. Srivastava 1991 A Course on Borel Sets. Springer Verlag.

[104]A. Haar 1933 Der Massbegriff in der Theorie der kontinuierlichen Gruppen. Annals of Mathematics, 2, vol. 34, no. 1, pp. 147169.

[105]K. Munjal, R. Bhatia 2022 A systematic review of homomorphic encryption and its contributions in healthcare industry. Complex and Intelligent Systems.

[106]D. Chakraborty, F. Rodríguez-Henríquez 2008 Cryptographic Engineering. Çetin Kaya Koç (ed.).

[107]P. Paillier 1999 Public-key cryptosystems based on composite-degree residuosity classes Advances in Cryptology-EUROCRYPT'99, Lecture Notes in Computer Science. vol. 1592, ed. J. Stern. Springer-Verlag, Berlin, 223238.

[108]C. E. Shannon 1948 A Mathematical Theory of Communication, Bell System Technical Journal 27, 379-423.

[109] T. J. Rivlin 1990 Chebyshev Polynomials. Wiley, New York.

[110]A. G. Radwan, S. K. Abd-El-Hafiz, S. H. Abd-El-Haleem 2012 Image encryption in the fractional-order domain. International conference on engineering and technology (ICET). p. 16.

[111]S. H. Abd ElHaleem, A. G. Radwan, S. K. Abd-El-Hafiz 2013 Design of pseudo random keystream generator using fractals. IEEE international conference on electrical circuits & systems (ICECS). p. 87780.

[112]G. Boeing 2016 Visual Analysis of Nonlinear Dynamical Systems: Chaos, Fractals, Self-Similarity and the Limits of Prediction. Systems, 4 (4), 37, doi:10.3390/systems4040037.

[113]T. E. Gamal 1985 A public-key cryptosystem and a signature scheme based on discrete logarithms. IEEE Trans. Inf. Theory IT-31(4), 469472.

[114]B. McMillan 1953 The Basic Theorems of Information Theory, Ann. Math. Stat. 24, 196-219.

[115]L. M. Pecora, T. L. Carroll 1990 Synchronization in chaotic systems, Phys. Rev. Lett., 64, 821.

[116]C. E. Shannon 1948 A Mathematical Theory of Communication. Bell System Technical Journal. First published: July 1948

[117]C. E. Shannon 1949 Communication theory of secrecy systems. Bell Sys. Tech. J. 28, 656715.

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

[119]G. Alvarez, F. Montoya, M. Romera and G. Pastor 2003 Cryptanalysis of an ergodic chaotic cipher. Physics Letters A, 311(23), 172-179.

[120]S. Li, X. Mou and Y. Cai 2001 Improving security of a chaotic encryption approach. Physics Letters A, 290(3-4), 127-133.

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

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

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

[124]H. A. Younis, I. M. Hayder, I. S. Seger, H. A-K. Younis 2020 Design and implementation of a system that preserves the confidentiality of stream cipher in nonlinear flow coding, Journal of Discrete Mathematical Sciences and Cryptography, 23:7, 1409-1419.

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

[126]X. Boyen, B. Waters, 2010 Shrinking the keys of discrete-log-type lossy trapdoor functions. Applied Cryptography and Network Security, 8th International Conference, ACNS 2010, 2010, pp. 3552.

[127]A. Rosen, G. Segev 2009 Chosen-ciphertext security via correlated products. Theory of Cryptography, 6th Theory of Cryptography Conference, TCC 2009, in: LNCS, vol. 5444, Springer, 2009, pp. 419436.

[128]E. Kiltz, P. Mohassel, A. O'Neill 2010 Adaptive trapdoor functions and chosen-ciphertext security. Advances in Cryptology EUROCRYPT 2010, LNCS, vol. 6110, Springer, 2010, pp. 673692.

[129]N. Döttling, S. Garg 2017 Identity-based encryption from the Diffie-Hellman assumption. Advances in Cryptology - CRYPTO, pages 537569, 2017.

[130]T. Matsumoto 1984 A Chaotic Attractor from Chua's Circuit. IEEE Transactions on Circuits and Systems. IEEE. CAS-31 (12): 10551058.

[131] O. E. Rössler, "An equation for continuous chaos", Physics Letters, 1976

***Résumé :***

La théorie déterministe du chaos intrigue beaucoup, particulièrement en raison de ses interactions mystérieuses avec la théorie ergodique des nombres. Bien que nous utilisions fréquemment des générateurs pseudo-aléatoires certifiés NIST, nous négligeons souvent de comprendre les raisons mathématiques de leur efficacité. Notre thèse se penche en partie sur ce sujet, à ceci près que nous nous concentrons sur les fractions égyptiennes d'irrationnels, au lieu de développements traditionnels tel le développement décimal ou celui en fractions continues, plus couramment utilisés en cryptographie. En adoptant la définition du chaos selon R. Devaney, nous démontrons mathématiquement la chaoticité d'un tel processus, créant ainsi un nouveau cryptosystème venant enrichir l'arsenal de systèmes à base de chaos déjà à notre disposition. Nous explorons ensuite des familles de fonctions injectives et lossy trapdoor (LTF), pratiquement indiscernables sur le plan calculatoire, mais dont l'utilité dans la construction de primitives cryptographiques est avérée. Plus précisément, nous mettons en place une construction efficace d'une variante des fractions égyptiennes pour extraire notre LTF souhaitée. Nous visons ainsi à améliorer l'efficacité du schéma de chiffrement résistant à l'attaque à texte chiffré choisi (IND-CCA), et ce en faisant appel aux notions de tenseurs et de catégories, tout en démontrant l'ergodicité de ce processus. L'aspect pseudo-aléatoire des processus inspirés des fractions égyptiennes contribuera à renforcer la sécurité non seulement dans le schéma IND-CCA, mais aussi dans divers défis liés à l'hypothèse décisionnelle ou calculatoire de Diffie-Hellman (DDH, CDH), ce qui les rend très précieux pour la communication entre plusieurs agents.

***Mots-clés :*** Théorie du chaos; Théorie ergodique des nombres; Fractions égyptiennes; Cryptographie; LTF.

***Abstract:***

Deterministic chaos theory is very intriguing, particularly because of its mysterious interactions with ergodic number theory. Although we frequently use NIST-certified pseudorandom generators, we often neglect to understand the mathematical reasons for their effectiveness. Our thesis partly addresses this issue, except that we focus on Egyptian fractions of irrationals, rather than traditional developments such as decimal or continued fractions, which are more commonly used in cryptography. Adopting the definition of chaos according to R. Devaney's definition of chaos, we demonstrate mathematically the chaotic nature of such a process, thus creating a new cryptosystem to add to the arsenal of chaos-based systems already available to us. We then explore families of injective and lossy trapdoor functions (LTFs), which are practically indistinguishable computationally, but which have been shown to be useful in the construction of cryptographic primitives. More precisely, we set up an efficient construction of a variant of Egyptian fractions to extract our desired LTF. Our goal is to improve the efficiency of the chosen ciphertext attack-resistant encryption (IND-CCA) scheme by using the notions of tensors and categories, while demonstrating the ergodicity of this process. The pseudorandom aspect of the processes inspired by Egyptian fractions will contribute to enhancing security not only in the IND-CCA scheme, but also in various challenges related to the Decisional Diffie-Hellman or computational hypothesis (DDH, CDH), making them very valuable for communication between several agents.

***Keywords :*** Chaos theory; Ergodic number theory; Egyptian fractions; Cryptography; LTF.

Année Universitaire: 2022– 2023