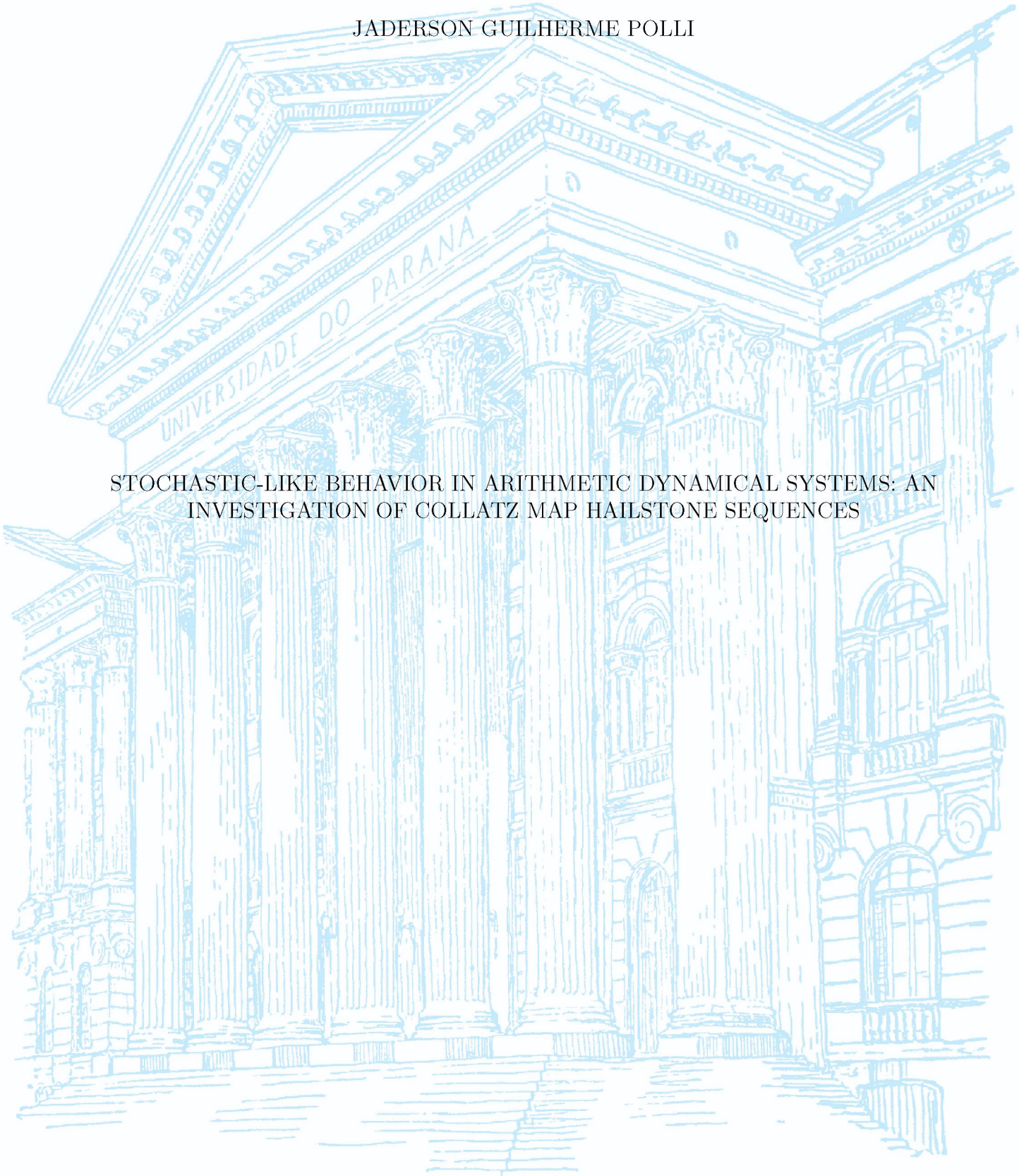
UNIVERSIDADE FEDERAL DO PARANÁ

JADERSON GUILHERME POLLI



STOCHASTIC-LIKE BEHAVIOR IN ARITHMETIC DYNAMICAL SYSTEMS: AN INVESTIGATION OF COLLATZ MAP HAILSTONE SEQUENCES

CURITIBA
2023

JADERSON GUILHERME POLLI

STOCHASTIC-LIKE BEHAVIOR IN ARITHMETIC DYNAMICAL SYSTEMS: AN INVESTIGATION OF COLLATZ MAP HAILSTONE SEQUENCES

Dissertação submetida ao Programa de Pós-graduação em Física, no Setor de Ciências Exatas, na Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Física.

Supervisor: Prof. Dr. Marcos Gomes E. da Luz.

Co-supervisor: Prof Dr. Madras Viswanathan Gandhi

CURITIBA
2023

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS – BIBLIOTECA CIÊNCIA E TECNOLOGIA

Polli, Jaderson Guilherme.
Stochastic-like behavior in Arithmetic Dynamical Systems : an Investigation of Collatz Map Hailstone Sequences. / Jaderson Guilherme Polli. – Curitiba, 2023.
1 recurso on-line : PDF.

Dissertação (Mestrado) – Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-Graduação em Física.
Orientador: Prof. Dr. Marcos Gomes E. da Luz.
Coorientador: Prof. Dr. Madras Viswanathan Gandhi.

1. Física. 2. Dinâmica. 3. Matemática. 4. Análise de series temporais – Processamento de dados. I. Luz, Marcos Gomes E. da. II. Gandhi, Madras Viswanathan. III. Universidade Federal do Paraná. Programa de Pós-Graduação em Física. IV. Título.

Bibliotecário: Nilson Carlos Vieira Júnior CRB-9/1797

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação FÍSICA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **JADERSON GUILHERME POLLI** intitulada: **"Stochastic-like behavior in Arithmetic Dynamical Systems: an Investigation of Collatz Map Hailstone Sequences"** , sob orientação do Prof. Dr. MARCOS GOMES ELEUTÉRIO DA LUZ, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 28 de Fevereiro de 2023.

Assinatura Eletrônica
28/02/2023 17:13:14.0
MARCOS GOMES ELEUTÉRIO DA LUZ
Presidente da Banca Examinadora

Assinatura Eletrônica
01/03/2023 10:58:34.0
CARLOS EDUARDO FIORE DOS SANTOS
Avaliador Interno (UNIVERSIDADE DE SÃO PAULO)

Assinatura Eletrônica
28/02/2023 18:03:00.0
RICARDO LUIZ VIANA
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

AGRADECIMENTOS

Agradeço aos meus pais, Neide Isabel Bassani Polli e Jackson Lucas Polli, por não medirem esforços para que eu chegasse até aqui.

Agradeço a Luna Rhaine Nascimento Oliveira, por todos os cafés na sala, risadas, ideias trocadas — especialmente pela ideia de usar $\log_2$ — pelo carinho, paciência e, sobretudo, agradeço pela compreensão nos momentos mais difíceis da jornada.

Agradeço a Marcos Vinícius Woiski Barcote, pelo companheirismo e amizade.

Agradeço a Eduardo Rosa de Jesus, por manter uma amizade verdadeira independente da distância, sempre disposto a ouvir e compartilhar as agruras e alegrias da vida.

Agradeço ao Professor Dr. Marcos Gomes Eleutério da Luz, meu orientador, por todas as frutíferas reuniões, discussões, e ensinamentos, sem os quais esta dissertação não passaria de um sonho não concretizado.

Agradeço ao Professor Dr. Madras Viswanathan Gandhi, meu co-orientador, pela disponibilidade em contribuir com suas ideias e conhecimentos para pontos chave dos resultados desta dissertação.

Agradeço ao Professor Msc. Antônio João Fidélis, meu orientador de iniciação científica, sem o qual eu não teria chegado até onde cheguei, pela leitura do manuscrito e pelos apontamentos sempre pertinentes.

Agradeço aos membros da banca de pré-defesa, os titulares Professor Dr. Thiago de Lima Prado e Professor Dr. Ernesto Carneiro Pessoa Raposo, e suplente Wilson Marques Junior, pelo aceite em participar do pré-julgamento deste trabalho, bem como por todas as contribuições.

Agradeço aos membros da banca defesa, ao Professor Dr. Ricardo Luiz Viana e ao Professor Dr. Eduardo Fiore dos Santos, pelo aceite em participar da avaliação deste trabalho, bem como por todas as contribuições.

Agradeço à CAPES pelo financiamento deste trabalho.

"There is a story about two friends, who were classmates in high school, talking about their jobs. One of them became a statistician and was working on population trends. He showed a reprint to his former classmate. [...] His classmate was a bit incredulous and was not quite sure whether the statistician was pulling his leg. "How can you know that" was his query. "And what is this symbol here?" "Oh" said the statistician "this is π." "What is that?" "The ratio of the circumference of the circle to its diameter." "Well, now you are pushing your joke too far," said the classmate, "surely the population has nothing to do with the circumference of the circle"." E. P. Wigner, 1950.

RESUMO

Dinâmica Aritmética é uma área de pesquisa emergente, que estuda o comportamento de sistemas em espaços e tempos discretos. A presente dissertação lida com um sistema dinâmico aritmético chamado Mapa de Collatz, uma regra para inteiros positivos n, tais que $n \mapsto n/2$ para n par, e $n \mapsto 3n+1$ para n ímpar. Um renomado e não resolvido problema matemático conjectura que para qualquer inteiro positivo n, finitas iterações do Mapa de Collatz eventualmente atingirão 1. A sequência de inteiros da iteração do Mapa de Collatz a partir de uma condição inicial n_0, até o ponto em que atinge 1, é chamada sequência de granizo. O Mapa de Collatz, apesar de fornecer uma dinâmica muito rica para números naturais, só começou a ser explorado recentemente no contexto de modelos e fenômenos físicos. Este trabalho descreve investigações na tentativa de caracterizar se as sequências de granizo podem ser vistas como um sistema determinístico realizando comportamento do tipo estocástico, buscando iluminar o caminho entre teoria de números e mecânica estatística, através da área de sistemas dinâmicos aritméticos estocásticos. Para fazer isso, análises estatísticas apropriadas em várias sequências foram feitas, utilizando um numeroso conjunto de condições iniciais muito grandes (até a ordem de $n_0 \sim 2^{10000}$). O processo de amostragem de condições iniciais foi conduzido utilizando uma nova representação para inteiros positivos, com conexão direta com 2-ádicas, chamados vetores-$\boldsymbol{m}$. Ao aplicar métodos de análise de séries temporais tais como *Power Spectrum* e *Detrended Fluctuation Analysis*, o comportamento do tipo estocástico é confirmado, reforçando a literatura acerca das sequências de granizo performarem Movimento Browniano Geométrico (MBG). Análises de função de autocorrelação e entropia de von Neumann mostram desvios em relação ao MBG para condições iniciais especiais, indicando fontes de determinismo e previsibilidade dentro da tendência geral estocástica das séries. Estes desvios aparecem na forma de autocorrelações de curto e médio alcance, bem como uma diminuição no valor das entropias para órbitas a partir destas condições iniciais. A entropia de von Neuman também permite a caracterização da estrutura interna da sequência, por meio da análise das componentes dos vetores-$\boldsymbol{m}$, indicando que o processo das sequências de granizo segue o teorema do limite central. Por fim, é possível conceber o Mapa de Collatz como um destruidor de estruturas, criando e reproduzindo padrões aleatórios, e lentamente destruindo toda e qualquer ordem imposta previamente.

Palavras-chave: dinâmica aritmética; sequências de Collatz; comportamento estocástico; análise de séries temporais; complexidade.

ABSTRACT

Arithmetic Dynamics is an emergent research area that studies the behavior of systems in discrete spaces and times. The present dissertation deals with an arithmetic dynamical system called Collatz Map, a rule for positive integers n, stating that $n \mapsto n/2$ for n even, and $n \mapsto 3n+1$ for n odd. A renowned and unsolved mathematical problem conjectures that for any positive integer n, finite iterations of the Collatz Map eventually reach 1. The sequence of integers from iterating the Collatz Map from an initial condition n_0 until reach 1 is often called hailstone sequence. The Collatz Map, besides providing very rich dynamics for natural numbers, only recently has been explored in the context of physical models and phenomena. This work describes investigations trying to characterize whether the hailstone sequences can be regarded as a deterministic system performing stochastic-like behavior, aiming to enlighten the path from number theory to statistical mechanics, in the area of stochastic arithmetic dynamical systems. In order to do that, proper statistical analysis of various sequences were done, utilizing a set of very large initial conditions (up to $n_0 \sim 2^{10000}$). The sampling of initial conditions was conduced by using a new representation for positive integers with direct connection with 2-adics, called $\boldsymbol{m}$-vectors. By employing methods of time series analysis such as Power Spectrum and Detrended Fluctuation Analysis, the stochastic-like behavior is confirmed, reinforcing literature about the hailstone sequences performing Geometric Brownian Motion (GBM). Autocorrelation function and von Neumann entropy analysis shows deviations from GBM for special initial conditions, indicating sources of determinism and predictability inside the general trend. These deviations appears in the form of short- to mid-range autocorrelations and smaller entropy values for the orbits from these specific initial conditions. The von Neumann entropy also allows the characterization of the internal structure of the sequence by the $\boldsymbol{m}$-vectors components analysis, indicating that the hailstone sequence is process following a central limit theorem. Finally, it is possible to conceive the Collatz Map as a structure destroyer, that creates and reproduces random patterns, slowly destructing any previously imposed ordering.

Keywords: arithmetic dynamics; Collatz sequences; stochastic behavior; time series analysis; complexity.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

Chapter 1

Introduction

Philosophy is written in this grand book, the universe, which stands continually open to our gaze. But the book cannot be understood unless one first learns to comprehend the language [...] in which it is written. It is written in the language of mathematics [...]

Galileo Galilei, *The Assayer*, 1623.

The search for simple fundamental rules seems to be placed in the kernel of Physics since William of Occam announced his famous razor: "*Entia non sunt multiplicanda praeter necessitatem*"[1][1]. This lemma has been carried on over natural sciences since the 18th and 19th Centuries, culminating in a quest for a Theory of Everything [2], idealizing to describe the entire universe with one equation. But still, aside from such efforts of simplifying the description of the universe, the interplay between simplicity and complexity has been invariably placed in the search for laws of physics in the various depth levels that the universe can be known [3].

A very fair question, to begin with, is: how does the universe starts with elementary particles at the big bang, and ends up with complex concepts such as life, history, economy, and literature [4]? This question is far from being answered, and the intent of this dissertation is not to answer it but to present a framework on how a simple elementary rule for iterating positive integers can display stochastic-like behavior akin to those found in complex systems.

These findings are related to a fundamental question about the "unreasonable effectiveness of mathematics in natural sciences" [5]. The "unreasonable effectiveness" refers to a large number of mathematical theories developed, on the first hand, without any direct application that, later, has come to play important role in describing physical phenomena — see [6] for further discussions on this topic. The key point here is how the such numerous abstract mathematical structures mimic and explain so well the ontological objects of the concrete world — called by Tegmark [7] as "human baggage" — such as particles, mass, spin, charge, fields, etc.

As a tentative of answering to Wigner, Tegmark [8] pushes to the extreme and introduces the concept of a Mathematical Universe, arguing the *External Reality Hypothesis* (ERH) implies on the *Mathematical Universe Hypothesis* (MUH), meaning the universe is mathematics in a well-defined sense. The idea that the universe is fundamentally mathematical traces back to the Pythagoreans [9], passing by Galileo's quote in this

[1] Hypothesis should not be multiplied unnecessarily.

Chapter's epigraph, has reached the popular culture by the hands of Douglas Adams [10], when 42 was announced as the answer to the *Ultimate Question of Life, the Universe, and Everything* and, by now, is still a bold speculation on the realm of cosmology.

A less philosophical perspective on this topic might be to consider pure mathematical structures and analyze how they might describe or relate to emergent behavior in natural phenomena. As examples, consider the association between arithmetic geometry and gauge field theories [11], differential topology and phase transitions [12] and number theory patterns and dynamical systems, leading to the relatively new area of arithmetic dynamics [13].

Arithmetic dynamical models are those whose evolution takes place in discrete spaces — e.g., in the set of integers or rationals — by the action of a map, so in a discrete-time domain. In a broad context, the area should connect results in the subject of Diophantine equations to discrete versions of Dynamical Systems [13, 14]. However, arithmetic dynamics have already found many applications [15, 16, 17], extrapolating the domain of dynamical systems. This is so because certain arithmetic dynamical systems deals with $p-$adics [18] in ultrametric spaces and fields [19]. By its turn, $p-$adic numbers — see Chapter 3 — can be used in the description of different physical theories such as quantum mechanics [20, 21], quantum logic [22], gravity [23], and string theory [24]. Further, they have been employed in the study of stochastic and complex systems, like self-organized criticality models [25], phase transitions in Ising [26, 27] and Potts [28, 29, 30, 31] models, as well as in Gibbs measures [32, 26, 33], Markov processes [34], and diffusion in a random medium [35] — see [17] for a complete review on applications.

In number theory — also boldly speculated to be the basics of fundamental physics [36] — there is a dichotomy between the intrinsic deterministic features of a numerical series, often given by a specific rule, and the eventual randomness (or more properly pseudo-randomness) the full sequence of these numbers display [37, 38, 39, 40, 41]. This leads to a type of dilemma, in close parallel to the comprehension of some natural processes: despite the constituents of the system evolving by well-defined deterministic laws, sometimes random elements are dominant so it is essential to represent the system by stochastic dynamics and make use of statistical mechanics tools. Furthermore, the presence of seemingly unpredictable dynamics into deterministic systems might provide a certain bridge to chaotic systems, where pure and simple mathematical models are well known to be a "laboratory" for chaos [42]. From these considerations, it is pertinent to question if associated with certain arithmetic dynamical models, made purely of numbers, one could also identify "arithmetic statistical mechanics" processes, connecting number theory, dynamical systems, and more generally physics at al.

The idea of statistical mechanics of pure mathematical or logical systems is not new to science, leading back to the statistic mechanics of cellular automata [43]. The interface between number theory and physics has been growing over the last decades [44, 45, 46, 47, 48, 49, 50, 51]. One of the origins of this interface goes back to Hilbert-Pòlya conjecture [52], although the origin of the conjecture itself is not clear [53]. The Hilbert-Pòlya gives a physical reason why the Riemann hypothesis [54] must be true. It states that the zeros of the Riemann-Zeta function are directly related to the real eigenvalues of an operator $\hat{H}$, which is the Hamiltonian of a physical system. These kind of connections from pure number theoretical problems to physical systems described in the paragraphs above is an indicator to the construction of a more solid path between these areas.

The objective of the present dissertation is way more humble than trying to assign an answer to any ultimate question about life, the universe and everything, but is still related

to connecting a pure mathematical number theoretical problem with natural phenomena. The aforementioned elementary rule for iterating positive integers displaying stochastic-like behavior to be exploited during this dissertation is a very simple rule — for discrete-time t — such that $f(n_t) = n_{t+1}$ for integer n and t. Hence, the main objective of this study is to explore a number-theoretic problem, the Collatz problem, its *hailstone sequences*[2], and its stochastic-like behavior, utilizing the arithmetic dynamics and time series analysis methods, aiming to enlighten the path from arithmetic simple rules and complex/stochastic phenomena. Specifically:

- A brand new arithmetic representation for positive integers is employed as a tool for statistical analysis and for studying the internal dynamics of the numbers;
- Methods of time series analysis, namely Power Spectrum, Detrended Fluctuation Analysis, Autocorrelation, and an entropy measurement, are applied to characterize the stochastic behavior of the system.

By completing the objectives presented above, it is expected an enlightenment of the comprehension of how arithmetic dynamical systems — especially those coming from number theory — can be utilized for modeling stochastic phenomena. Further, there is no intention of extinguishing the discussion on this topic, quite the opposite, the intention is to foment future analysis and discussions. The present dissertation is structured in six chapters, including the current Introduction Chapter, four Chapters of development, and one for Conclusion and Remarks. It also contains seven Appendixes providing supplemental information for the main text.

Chapter 2 contains a concise bibliographic review of the Collatz Map and hailstone sequences, presenting in detail its various formulations, together with the Collatz problem that leads to the infamous Collatz Conjecture. It intents to highlight the most important topics of what has already been done with the Collatz Map in mathematics and physics, followed by a review of topics related to the main study of this dissertation, and finally a very recent application that can be worth analyzing in the future.

Chapter 3 presents the theoretical and mathematical background for the implemented methods of arithmetic representation and time series analysis. It is divided into two sections whose topics are fundamental for the results of Chapters 4 and 5. The first section deals with the construction of an arithmetic representation for natural numbers, starting from algebraic topology concepts, ultrametric spaces, p-adics, and hierarchical structures applied to complex systems, leading to the central representation for the study: the $\boldsymbol{m}$-vectors. The second section reviews methods of time series analysis to be applied in the hailstone sequences and on the new arithmetic representation.

Chapter 4 contains results on the exploration of the Collatz Map's dynamics in the representation of $\boldsymbol{m}$-vectors. It contains results on the $\boldsymbol{M}$-Matrix, which is a matrix containing all information on the dynamics of $\boldsymbol{m}$-vectors, followed by investigations on the dimension of the $\boldsymbol{m}$-vectors and statistical investigation of the components of the vector along the trajectories of the Collatz Map. The Chapter closes with four topics with initial exploration yet to be finished, but presenting promising ideas in a future analysis project. Namely, the toy model of a gas formed by the dynamics of the Collatz Map, algebraic investigations on the new arithmetic representation, the analysis of graphs and networks from the $\boldsymbol{m}$-vectors, and the definition of a chaotic measurement for the Collatz Map.

[2]See page 19 for the explanation on this name.

Chapter 5 displays the main results of this dissertation, about characterizing stochastic-like behavior in the Collatz sequences. It begins with the description of specific methodologies for data sampling and how to apply the methods from Chapter 3 into the hailstone sequences. Results of time series analysis of the hailstone sequences and the $\boldsymbol{m}$-vectors are then presented and their implications are discussed.

Chapter 6 retakes the discussion on the topics from introduction about merging physics and number theory, review the objectives and think over whether they were attained or not, as well as future projections.

This dissertation also contains seven Appendices destined to provide supplemental material for the main text. Appendix A presents definitions of Brownian and Geometric Brownian Motion, important to understand one some features of hailstone sequences. Appendix B presents proofs of lemmas on ultrametric spaces given in Chapter 3. Appendix C provides details on a tool for fastening the Discrete Fourier Transformation called Fast Fourier Transformation (FFT). Appendix D details the process of obtaining one of the main equations of this dissertation connecting the $\boldsymbol{m}$-vectors with the hailstone sequences. Appendix E presents initial conditions in base-10 and its correspondent $\boldsymbol{m}$-vector for a figure in Chapter 4. Appendix F presents transition matrices for the graphs presented in the Future Analysis section of Chapter 4. Finally, Appendix G details the method of sampling the initial conditions for statistical analysis utilizing the $\boldsymbol{m}$-vectors formalism.

Chapter 2

The Collatz Map

Vinte e dois vezes dois, quarenta e três,
Quarenta e três dividido por seis, noventa e um,
Noventa e um vezes três, oitenta e dois,
Oitenta e dois dividido por seis, noventa e um,
...
(Rogério Skylab)

In this chapter, an elementary rule for iterating positive integers will be addressed, the Collatz Map [55]. The Collatz Map is based upon the following rule: Given a natural number n, if it is even, then we divide n by two, if it is odd, we multiply n by three and add one. From this, one construct a map on positive integers $f : \mathbb{Z}^+ \to \mathbb{Z}^+$ as

$$n_{t+1} = f(n_t) = \begin{cases} \dfrac{n_t}{2} & \text{if } n_t \equiv 0 \pmod 2, \\ 3n_t + 1 & \text{if } n_t \equiv 1 \pmod 2. \end{cases} \tag{2.1}$$

The Collatz sequence is generated by iterating the Map from an initial condition $n_0 \in \mathbb{Z}^+$, producing an orbit

$$\mathcal{O}(n_0) \equiv \Big(n_0, f(n_0) = n_1, f^2(n_0) = n_2, \dots f^t(n_0) = n_t, \dots\Big)$$

constituted by integers. We denote $(f^{(t)})$ as the sequence of numbers up to t applications of f. This sequence is frequently called hailstone sequence, because of the multiple ascents and descents that typical trajectories are subjected, just as hailstones are before dropping [56].

The Collatz Map is related to a very simple-to-state conjecture in number theory, that remains unproved: the Collatz Conjecture [57]. The conjecture can be stated as follows.

Collatz Conjecture 2.1. *Starting from any initial condition $n_0 \in \mathbb{Z}^+$, the orbit of the Collatz Map will eventually reach 1. Thereafter, the iterations will cycle through $1, 4, 2, 1, \dots$.*

Therefore, exists a finite S such that $f^S(n_0) = 1$, *i.e.* $\mathcal{O}(n_0) = \{n_0, n_1, \dots, n_S = 1, \dots\}$, where $S(n_0)$ is usually called the *total stopping time* function. Notice that once reaching $n_S = 1$, the orbit becomes periodic $\{\dots, n_S = 1, 4, 2, 1, 4, 2, \dots\}$.

Other formulations of the Collatz Map (2.1) can be found to reach the same conjecture, and some of them are favored in literature. For example, one can perceive that if

$n_t \equiv 1 \pmod 2$, then $n_{t+1} \equiv 1 \times 1 + 1 \equiv 0 \pmod 2$ must be divided by two, hence, one can contract the dynamics by $n_t \mapsto 3n_t + 1 \mapsto (3n_t + 1)/2$ for $n_t \equiv 1 \pmod 2$, considering the map $\tilde{f} : \mathbb{Z}^+ \to \mathbb{Z}^+$

$$n_{t+1} = \tilde{f}(n_t) = \begin{cases} \dfrac{n_t}{2} & \text{if } n_t \equiv 0 \pmod 2 \ , \\ \dfrac{3n_t + 1}{2} & \text{if } n_t \equiv 1 \pmod 2, \end{cases} \tag{2.2}$$

where $\tilde{f}$ is called the Accelerated Collatz Map and the orbit from $n_0 \in \mathbb{Z}^+$ will be denoted as $\tilde{\mathcal{O}}(n_0)$. Another dynamics contracting version of the Collatz Map is defined only on positive odd integers $\tilde{\tilde{f}} : \mathbb{Z}^+_{\text{odd}} \to \mathbb{Z}^+_{\text{odd}}$, by

$$n_{t+1} = \tilde{\tilde{f}} = \frac{3n_t + 1}{2^{k_t}}, \tag{2.3}$$

where $\tilde{\tilde{f}}$ is called the Odd Collatz Map, k_t is the number of factors of 2 contained in $3n_t + 1$, and its orbit will be denoted as $\tilde{\tilde{\mathcal{O}}}(n_0)$. The set of k_t values along the orbit up to time m — denoted by $\gamma_m = (k_0, k_1, \ldots, k_m)$ — is called "m-path".

Working on the map $\tilde{f}$ (2.2) is usually favored in literature. Working with $\tilde{\tilde{f}}$ prohibits substantial analysis because of the variability of k_t [58], but still Sinai [59] has proven some theorems on the m-path, important for stochastic models of the Collatz Map, to be reviewed in Section 2.4.

Further, Kontorovich [60] presents a generalization of the Collatz Map, called (d,g,h)-Maps. The (d,g,h)-Maps are a set of mappings, determined by positive co-primes d and g, with $g > d \geq 2$ and a periodic function $h(n)$ satisfying: (i) $h(n + d) = h(n)$; (ii) $n + h(n) \equiv 0 \pmod d$; (iii) $0 < |h(n)| < g$ for all n not divisible by d. The (d,g,h)-Maps are defined by

$$n_{t+1} = f_{d,g,h}(n_t) = \frac{gn_t + h(n_t)}{d^{k_t}}, \tag{2.4}$$

where k_t is uniquely chosen so the result is not divisible by d. This map is defined in the set of positive integers *not divisible* by d or g: $\Pi = dg\mathbb{Z}^+ + E$ with E the set of integers between 1 and dg that divide neither d nor g. Notice that one finds the Odd Collatz Map $\tilde{\tilde{f}} = f_{2,3,1}$ with $\Pi = 6\mathbb{Z}^+ + \{1,5\}$. The importance of this generalization will come clear in Section 2.4 when dealing with stochastic models and properties of the hailstone sequences.

This Chapter has the objective of reviewing the main literature on the Collatz problem, starting with an overview in Section 2.1, leading to pure mathematical number theory review in Section 2.2, and to application in the theory of dynamical systems in Section 2.3. Then, focusing on the main subject of this dissertation, Section 2.4 reviews stochastic models for the map, and Section 2.5 verse on complexity behavior on hailstone sequences. Finally, Section 2.6 reviews a dictionary between the Collatz process and Quantum Mechanic's operators, aiming for future investigations.

2.1 Overview on the $3x + 1$ problem

The number theoretic problem attached to the Collatz Map is often called the *Collatz problem*. The origin of the Collatz problem is attributed to Lothar Collatz [55, 57]. Collatz circulated the problem at the International Congress of Mathematicians in Cambridge in

n_0	$\mathcal{O}(n_0)$	$\tilde{\mathcal{O}}(n_0)$	$\tilde{\tilde{\mathcal{O}}}(n_0)$
7	$(7, 22, 11, 34, 17, 52, 26, 13,$ $40, 20, 10, 5, 16, 8, 4, 2, 1)$	$(7, 11, 17, 26, 13,$ $20, 10, 5, 8, 4, 2, 1)$	$(7, 11, 17,$ $13, 5, 1)$
15	$(15, 46, 23, 70, 35, 106, 53, 160,$ $80, 40, 20, 10, 5, 16, 8, 4, 2, 1)$	$(15, 23, 35, 53, 40$ $20, 10, 5, 8, 4, 2, 1)$	$(15, 23, 35,$ $53, 5, 1)$
19	$(19, 58, 29, 88, 44, 22, 11, 34, 17,$ $52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1)$	$(19, 29, 44, 22, 11,$ $17, 26, 13, 20, 10,$ $5, 8, 4, 2, 1)$	$(19, 29, 11,$ $17, 13, 5, 1)$
25	$(25, 76, 38, 19, \ldots, 5, 16, 8, 4, 2, 1)$	$(25, 38, 19, \ldots, 5, 8, 4, 2, 1)$	$(25, 19, \ldots, 5, 1)$

Table 2.1: Examples of hailstone sequences for each of the three versions of the Collatz Map: $f, \tilde{f}$ and $\tilde{\tilde{f}}$.

1950. But the origin of the problem is not a consensus. Bryan Thwaites states to have originated the problem, as it is known nowadays in 1952 [61].

This problem also goes under many names, such as *Hasse's algorithm*, because Helmut Hasse discussed the problem with many people, and also published about [62]; *Syracuse problem*, as proposed by Hasse during a visit to Syracuse University [55]; *Kakutani's problem* is attributed to Shizuo Kakutani, who, in a private communication to J. Lagarias in 1981, stated that "for about a month everybody at Yale worked on it, with no result" [63]. Also Stanisław Ulam heard about the problem and circulated around Los Alamos, leading to this problem being called *Ulam's problem* in some circles [61, 64]. In his book on the theme, Lagarias [57] uses the terminology $3x+1$ problem to refer to this multi-named problem. So, to be concise during the text, from now on, the problem of solving the Collatz Conjecture will be called the $3x+1$ *problem*.

To understand why such a problem became so fascinating, one must explore the richness of structures the hailstone sequences display. Playing with some small initial conditions helps to get some acquaintance with the variability of sequences. In Table 2.1 one finds sequences started from small initial conditions for three versions of the Collatz Map: $f, \tilde{f}$ and $\tilde{\tilde{f}}$.

Further, Figure 2.1 presents the plot of comparison between two hailstone sequences with very close initial conditions ($n_0 = 25$ and $n_0 = 27$) for each of the three maps presented. It is important to notice that the map f generates the most biased sequences from all three versions because every ascent is followed directly by a descent, leading to permanent correlations along the sequences. This can be observed clearly in Table 2.1 and clearer in Figures 2.1 (a) and 2.1 (b). By observation of Figures 2.1 (c) and 2.1 (d), is evident that $\tilde{f}$ provides quite rich dynamics, with more variability of ascents and descents than $\tilde{\tilde{f}}$ sequences — compare with Figures 2.1 (e) and 2.1 (f) — and without the predictability of ascents and descents of f.

The current status of the $3x+1$ problem divides into two battlefronts: numerical testing and formal theorems, reinforcing, without necessarily proving the conjecture. In the realm of numerical and computational testing, by 2020, the conjecture has been tested for all initial conditions $n_0 < 2^{68}$ [65]. Using formal number theoretic analysis, recently T. Tao demonstrated that "almost all orbits of the Collatz map attain almost bounded values" [66, 67], while Izadi [68] and Llibre [69] claims to have obtained proofs of the conjecture. The study presented here does not intend to prove or verify any proof of the conjecture,

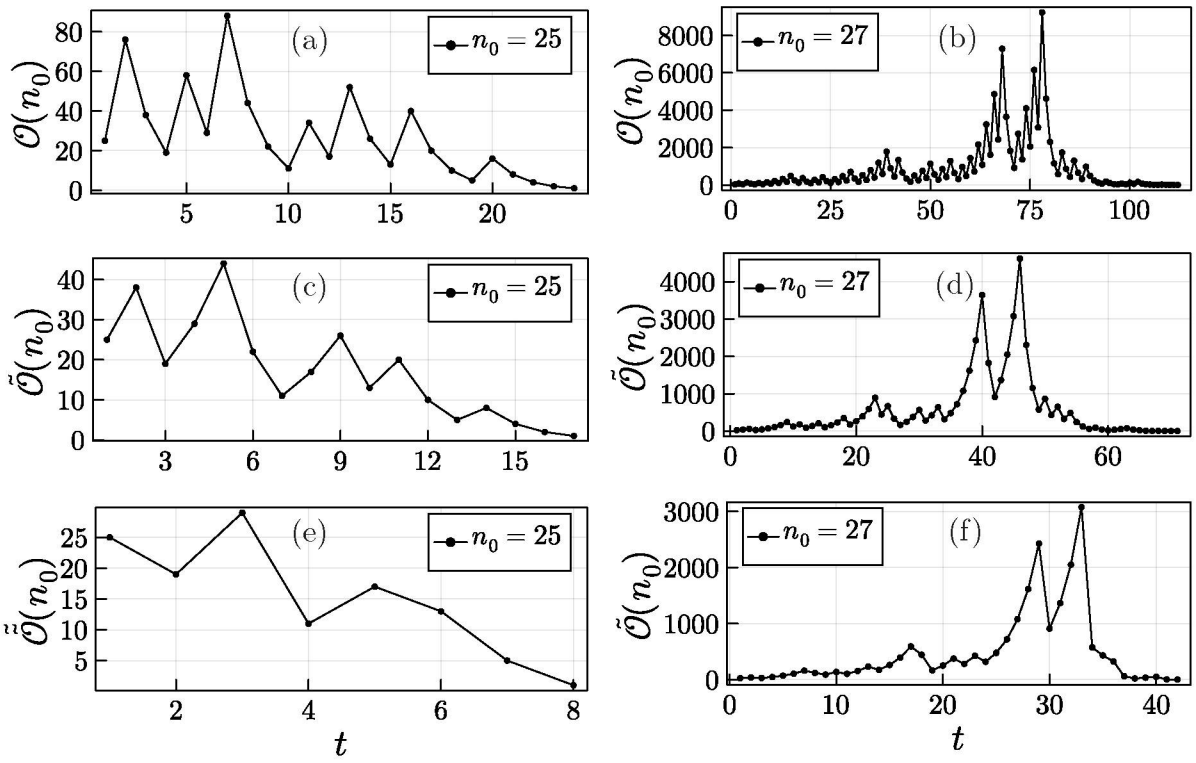


Figure 2.1: Graphic representation of hailstone sequences for two initial conditions evolved by three different versions of the Collatz Map. (a) Orbit $\mathcal{O}(n_0 = 25)$ evolved by the original Collatz Map f (Eq. (2.1)); (b) Orbit $\mathcal{O}(n_0 = 27)$ evolved by the original Collatz Map; (c) Orbit $\tilde{\mathcal{O}}(n_0 = 25)$ evolved by the Accelerated Collatz Map $\tilde{f}$; (d) Orbit $\tilde{\mathcal{O}}(n_0 = 27)$ evolved by the Accelerated Collatz Map $\tilde{f}$; (e) Orbit $\tilde{\tilde{\mathcal{O}}}(n_0 = 25)$ evolved by the Odd Collatz Map $\tilde{\tilde{f}}$; (f) Orbit $\tilde{\tilde{\mathcal{O}}}(n_0 = 27)$ evolved by the Odd Collatz Map $\tilde{\tilde{f}}$.

but instead, to use the diverse dynamics observed in hailstone sequences to reach natural phenomena description.

From the first publications on the $3x + 1$ problem during the 1970's decade up to nowadays, the problem has cut across many fields of mathematics and, more recently, has been applied to model physical systems. In the next sections, a brief literature review on surveys and applications of the Collatz in fields of mathematics and physics will be provided.

2.2 Collatz Map and Number Theory

Number theory is the immediate application of the $3x + 1$ problem, once it is an arithmetic problem, belonging naturally to elementary number theory. It is classified as one of the unsolved problems in number theory by R. K. Guy [70]. The trials of solving the $3x + 1$ problem lead to problems involving exponential Diophantine equations. The main results versing on number theory are surveyed in the papers of Lagarias [55] and Chamberland [58].

To testify the large variability of the transient sequences of $\tilde{f}$, one can obtain various measurements from the orbits. For example, one can look up the behavior of the *total*

Table 2.2: Numerical examples of the total stopping time S, stopping time σ and expansion factor $\mathcal{S}$ for representative initial conditions n_0.

n_0	S	σ	$\mathcal{S}$
3	5	4	2.666
4	2	1	1
5	4	2	1.6
6	6	1	1.336
26	8	1	1
27	70	59	170.962
81	16	2	1.506
82	70	1	56.292
83	70	4	55.614
1024	10	1	1
1025	26	2	1.500
9662	80	1	3.797
9663	118	54	1403.002
9664	18	1	1

stopping time S as a function of the initial condition $S(n_0)$; or, the *stopping time* function can be defined as the least positive $s = \sigma(n_0)$ such that $\tilde{f}^s(n_0) < n_0$; further, one can define the *expansion factor* $\mathcal{S}(n_0)$ as the ratio between the maximum value attained along the orbit $(\sup_{t\geq 0} \tilde{f}^t(n_0))$ and the initial condition n_0.

Table 2.2 displays a few numerical examples of the total stopping time S, stopping time σ and expansion factor $\mathcal{S}$ functions for representative initial conditions. Further, figure 2.2 presents the behavior of the (a) total stopping time $S(n_0)$, (b) stopping time $\sigma(n_0)$ and (c) expansion factor $\mathcal{S}(n_0)$ for all $n_0 \in [3, 10000]$. In Figure 2.2 (a) one observes that there seem to be some classes of numbers with almost the same total stopping time, generating some patterns in the function $S(n_0)$. The same can be observed for the stopping time function and the expansion factor.

Some very interesting analytical results on Number Theory are related to the *stopping time* function. The stopping time $\sigma(n_0)$ simply states the time for an orbit to descend from its initial condition. Riho Terras observed that the stopping time $\sigma(n_0)$ is easier to analyze than the total stopping time $S(n_0)$ function [71]. Thus, one can redefine the original Collatz Conjecture 2.1 in terms of the stopping time as Lagarias [55] presented.

Collatz Conjecture 2.2. *Every integer $n_0 \geq 2$ has a finite stopping time $\sigma(n_0)$.*

This form of stating the conjecture seems, at first glimpse, weaker than the original one because it is versing only about the stopping time, not the *total* stopping time. But if every $2 \leq n_0 \leq \mathcal{K}$ has finite stopping time, then all $n_0 \in [2, \mathcal{K}]$ are eventually mapped into smaller values in the interval $[2, \mathcal{K}]$, then, in the limit $\mathcal{K} \to \infty$, one finds the Collatz Conjecture by an assumption of finite stopping time.

Terras [71, 72] have proven that the set of positive integers with finite stopping time has density one. In the same direction, Garcia [73] has proven that the density of divergent trajectories is zero in $\mathbb{Z}^+$.

Much less is known about the total stopping time function. A notable phenomenon, observed in Figure 2.2 (a), is that for orbits initiated close to each other, the total stopping time does not vary much. This behavior is related to the fact that nearly initiated

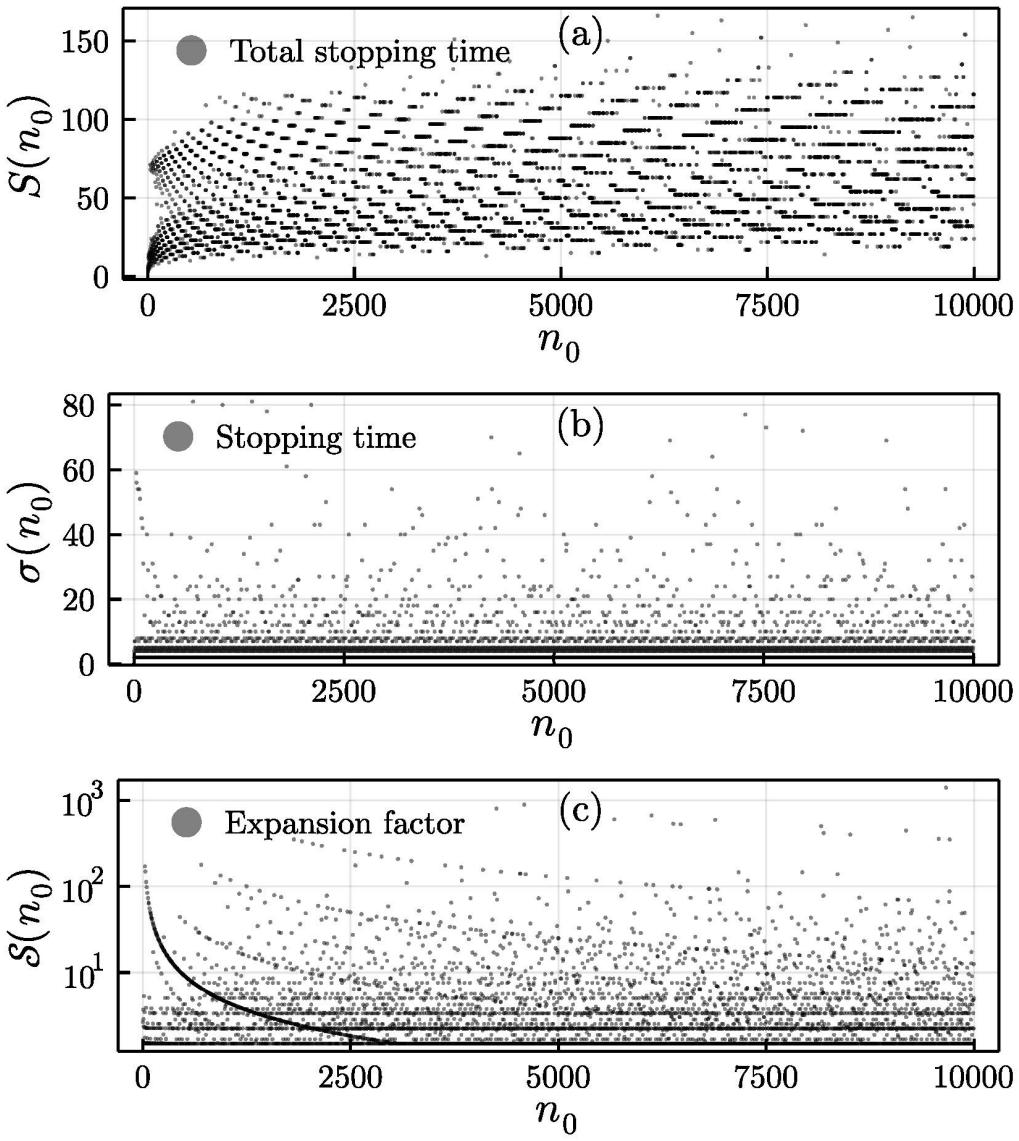


Figure 2.2: Behavior of the (a) total stopping time function $S(n_0)$, (b) Stopping time function $\sigma(n_0)$ and (c) the expansion factor $\mathcal{S}(n_0)$, evidencing the variability of orbit's behavior.

trajectories coalesce within a few steps. For example, the trajectories of $8k+4$ and $8k+5$ coalesce after 3 iterates:

$$8k+4 \mapsto 4k+2 \mapsto 2k+1 \quad \mapsto \quad 3k+2 \tag{2.5}$$

$$8k+5 \mapsto 12k+8 \mapsto 6k+4 \quad \mapsto \quad 3k+2. \tag{2.6}$$

But the strongest result proved so far about the density of the set of positive integers with finite total stopping time only asserts that the counting function of finite total stopping time $\mu_S = \mu\{n_0 : n_0 \leq x \text{ and } S(n_0) < \infty\}$ is related to the integer set counting function x, by a constant $c_4 > 0$ such that $\mu_S > x^{c_4}$, as proved by Crandall [74].

Another formulation of the map can be found by defining the parity of a number

$x(n) = n \pmod 2$, and then rewrite the Accelerated Collatz Map (2.2) as

$$n_{t+1} = \tilde{f}(n_t) = \frac{3^{x_t(n_t)} n_t + x_t(n_t)}{2}, \tag{2.7}$$

then the *parity vector* for the orbit from n_0 is defined as

$$\mathbf{v}_t(n_0) = (x_0(n_0), x_1(n_0), x_2(n_0), \ldots, x_{t-1}(n_0)), \tag{2.8}$$

where

$$x_i(n_0) \equiv \tilde{f}^i(n_0) \pmod 2, \tag{2.9}$$

leading to write the dynamics from n_0 and $x_0(n_0)$ as

$$\begin{aligned}
n_0 &\mapsto f(n_0) = n_1 = \frac{3^{x_0(n_0)} n_0 + x_0(n_0)}{2}, \\
n_1 &\mapsto f(n_1) = n_2 = \frac{3^{x_0(n_0)+x_1(n_0)} n_0}{2^2} + \frac{3^{x_1(n_0)} x_0(n_0)}{2^2} + \frac{x_1(n_0)}{2}, \\
&\vdots \\
n_{t-1} &\mapsto f(n_{t-1}) = n_t = \frac{3^{\sum_{i=0}^{t-1} x_i(n_0)}}{2^t} n_0 + \sum_{i=0}^{t-1} x_i(n_0) \frac{3^{\sum_{j=i+1}^{t-1} x_j}}{2^{t-i}},
\end{aligned} \tag{2.10}$$

then calling

$$\lambda_t(n_0) = \frac{3^{\sum_{i=0}^{t-1} x_i(n_0)}}{2^t} \tag{2.11}$$

$$\rho_t(n_0) = \sum_{i=0}^{t-1} x_i(n_0) \frac{3^{x_{i+1}+\cdots+x_{t-1}}}{2^{t-i}} \tag{2.12}$$

the dynamics can be written as a function of the initial condition as

$$\tilde{f}^t(n_0) = \lambda_t(n_0) n_0 + \rho_t(n_0). \tag{2.13}$$

In this notation, the nonlinear dependence of λ_t and ρ_t on n_0 clarifies why, besides being a linear map, its evolution presents such a complex pattern. The necessary condition for $\tilde{f}^t(n_0) < n_0$ is simply that $\lambda_t(n_0) < 1$, once ρ_t is non-negative.

2.3 Collatz Map and Dynamical Systems

The $3x + 1$ problem can also be studied in a *discrete dynamical system* point of view, once it concerns the iteration on discrete times of a function, the Collatz Map in Equation (2.1). The viewpoint of this problem as a Dynamical System was first given by Wirsching [75], thinking on the Collatz Map as a dynamical system on the state space of $\mathbb{Z}$. From this view, one can think of expanding the domain to larger spaces, such as the real numbers $\mathbb{R}$ and the complex numbers $\mathbb{C}$. Also, other extensions include defining functions on the ring $\mathbb{Z}_2$ of 2-adic integers, or, defining generalized maps on a ring of $p-$adic integers [57]. The ring $\mathbb{Z}_2$ of 2-adic will be addressed in Chapter 3, preceding a new representation for natural numbers in Section 3.1.3. The present Section reviews results of dynamical systems, concerning the dynamics of the Accelerated Collatz Map and variations in larger

spaces.

The extension of the dynamics of the Collatz Map was done toward the integers set $\mathbb{Z}$ by substituting $3x+1$ by $3x+\epsilon$, with $\epsilon = \pm 1$ by Seifert [76].

The extension to the rational set $\mathbb{Q}$ is vastly discussed by Lagarias [77] substituting $3x+1$ by $3x+\kappa$ with $\kappa \equiv \pm 1 \pmod 6$, proving that there are periodic cycles of the map for many κ.

To study the dynamics over integers and rationals, one can look for the set of fixed points $\mathcal{F}_{\tilde{f}^t}$ of the t-th iterate of the Accelerated Collatz Map, that is finding the set of n_j such that

$$\tilde{f}^t(n_j) = n_j, \tag{2.14}$$

leading to the sets of periodic orbits $\mathcal{P}_t^j = \{n_j, \dots, n_{j+t} = n_j\}$.

The periodic orbit

$$\mathcal{P}_2^3 = \{1, 2, \dots\},$$

is the trivial solution, strictly for $n_0 \in \mathbb{Z}^+$, leading to a reformulation of the Collatz Conjecture, by using the concept of a basin of attraction [13], stating that this is the only solution to Equation (2.14) and its variations on the positive integers set.

Collatz Conjecture 2.3. *The Collatz Map and its variations have only one stable periodic orbit in the set of positive integers, whose basin of attraction is the full set of positive integers.*

When allowing the solutions of (2.14) to encompass the complete set of integers $n_j \in \mathbb{Z}$, there are four other cycles: two period-one $\mathcal{P}_1^1 = (0, 0, \dots), \mathcal{P}_1^2 = (-1, -1, \dots)$, one period-three $\mathcal{P}_3^3 = (-5, -7, -10, -5, \dots)$ and one very long [58]

$$(-17, -25, -37, -55, -82, -41, -61, -91, -136, -68, -34, -17, \dots).$$

All these solutions, and more, can be found by solving Equation (2.14) for the set of rational numbers. By rewriting $\tilde{f}^t(n)$ as in Equation (2.13) one finds that Equation (2.14) reads

$$\begin{aligned} (1-\lambda_t) n_j &= \rho_t, \\ \left(1 - \frac{3^{\sum_{i=0}^{t-1} x_i}}{2^t}\right) n_j &= \frac{3^{\sum_{i=0}^{t-1} x_i}}{2^t} \sum_{i=0}^{t-1} x_i \frac{2^i}{3^{\sum_{k=0}^{i} x_k}}. \end{aligned}$$

Then, one can obtain the set of n_j that satisfies the Equation above for each possible parity vector $\mathbf{v}_j = (x_0, x_1, \dots, x_{t-1})$ with length t. There are 2^t possible parity vectors, each of them determines a unique rational solution n_j for Equation (2.14) [55] and there are, consequently, 2^t fixed points of period t.

Although there are 2^t fixed points of period t, when two fixed points are obtained from parity vectors that are only cyclic permutations of each other, these two belong to the same periodic orbit. For example, for $t = 2$, the parity vectors $\mathbf{v}_3 = (0,1)$ and $\mathbf{v}_4 = (1,0)$ are only permutations of each other. The fixed points $n_3 = 2$ and $n_4 = 1$, associated to $\mathbf{v}_3$ and $\mathbf{v}_4$ respectively, are in the same periodic orbit $\mathcal{P}_2^3 = \{2, 1, 2, 1, \dots\}$. The number of periodic orbits of period t is obtained from permutation theory as the number of circular permutations of n beads in a necklace of 2 colors [78] as

$$N_t = \frac{1}{t} \sum_{k=1}^{n} 2^{(k,t)}, \tag{2.15}$$

where $(k,t) = \gcd(k,t)$ is the greatest common divisor between k and t. Equation (2.15) gives rise to a sequence of integers given by

$$\{2, 3, 4, 6, 8, 14, 20, 36, 60, 108, 188, 352, \dots\}$$

with the t-th element corresponding to the number of periodic orbits of period t in the Accelerated Collatz Map extended in the rational set $\mathbb{Q}$.

The extension of the Collatz Map towards the real set $\mathbb{R}$ was studied by Chamberland [79], reviewed by Wirsching [75], by observing that the function defined by

$$\tilde{F}(x) = \frac{x}{2}\cos^2\left(\frac{\pi x}{2}\right) + \frac{3x+1}{2}\sin^2\left(\frac{\pi x}{2}\right) = x + \frac{1}{4} - \frac{2x+1}{4}\cos(\pi x), \quad x \in \mathbb{R}, \tag{2.16}$$

interpolates the Accelerated Collatz Map $\tilde{f}$ for $x \in \mathbb{N}$. With this extension, one can apply methods of one-dimensional discrete dynamical systems to attack the problem, such as the Schwarzian derivative [80] $S(\tilde{F})$, given by

$$S(\tilde{F}(x)) = \frac{\tilde{F}'''(x)}{\tilde{F}'(x)} - \frac{3}{2}\left(\frac{\tilde{F}''(x)}{\tilde{F}'(x)}\right)^2, \tag{2.17}$$

where, for every real $x \geq 0$, $S(\tilde{F}(x)) > 0$.

This enables the application of theorems of periodic points in one-dimensional maps, leading to obtaining three intervals with specific dynamics, and a new formulation of the conjecture. Let $\mu_0, \mu_1, \mu_2, \dots, \mu_n, \dots$ be the fixed points of $\tilde{F}$ on $[0, \infty)$ in increasing order, then Wirsching [75] points out that there are three intervals with specific dynamics: $I_1 := [0, \mu_1]$ invariant under $\tilde{F}$, with $\mu_0 = 0$ an attracting fixed point and μ_1 a repeller; $I_2 := [\mu_1, \mu_2]$ also invariant under $\tilde{F}$, with two attracting cycles

$$A_1 = \{1, 2\}, \quad A_2 = \{1.192531907\dots, 2.13865335\dots\};$$

and a remaining interval $I_3 := [\mu_3, \infty)$, non-invariant under $\tilde{F}$, partitioned to $I_3 = E_{\tilde{F}} \cup R_{\tilde{F}}$, where $E_{\tilde{F}}$ is called the *escape set* for those orbits in this set escape towards I_1 and I_2, and $R_{\tilde{F}}$ the *residual set*, whose orbits stay in I_3. Then, the Collatz Conjecture can be reformulated:

Collatz Conjecture 2.4. $\mathbb{N} \cap R_{\tilde{F}} = \varnothing$.

This indicates that every natural number eventually goes to I_2 and falls into periodic orbits. The conclusions of Chamberland are that (i) the cycles of the Accelerated Collatz Map on $\mathbb{N}$ is an attractive cycle of $\tilde{F}$ on $\mathbb{R}$; (ii) There is a uncountable set $U^0_{\tilde{F}} \in R_{\tilde{F}}$ of unstable bounded orbits and (iii) a set $U^\infty_{\tilde{F}} \in R_{\tilde{F}}$ monotonically increasing divergent trajectory.

2.4 Stochastic models for the Collatz Map

It is an observational fact, perceived in Figure 2.2, that the larger the initial condition n_0, the larger the total stopping time $S(n_0)$ gets, in average. Any tentative of describing the general behavior of the hailstone orbits must, then, be able to describe those starting from as large as possible initial conditions. Hence, one must be able to describe long-term hailstone sequences, starting from any natural number. In this context, stochastic models

are widely used [60, 67] to describe the dynamics and provide rigorous analysis founded on probability theory. This section will deal with describing the hailstone sequences *via* stochastic models. The converse, i.e., to describe stochastic phenomena by hailstone sequences, has barely been explored in literature and, in this dissertation, is reserved for Chapter 5.

Ensembles of orbits started from a small initial condition, as in Figure 2.3(a) do not seem to follow any decay pattern that could benefit from a stochastic approach. On the other hand, the larger the initial condition gets (see Figures 2.3(b) and 2.3(c)) the clearer a central path of decay appears. This indicates that the study of larger orbits might benefit from a statistical approach, based on large numbers law and probabilistic models. These are fair justifications for the stochastic models approach to attack the problem and try to characterize what type of stochastic process is being imitated by the deterministic rules of the Collatz Map.

To rigorously justify a probabilistic approach, first, one needs to look for a result, due to Terras [71], proved in Lagarias [55], called the *t-truncated parity vector*. The theorem states that the t-truncated parity vector $\mathbf{v}_t(n_0) = (x_0, \ldots, x_t)$ of the first t iterates of the Accelerated Collatz Map $\tilde{f}$ is periodic in n_0 with period 2^t. Each of all the 2^t possible $0-1$ vectors with t components occurs exactly once in the initial segment $1 \leq n_0 \leq 2^t$. In order to understand the consequence of this theorem, consider the following intuitive approach. You are set to pick, randomly, numbers from the positive integers set. Once half of the positive integers are even, and half of them are odd, then, the probability of picking, by random, an even or an odd number is the same, $1/2$ for each. The consequence of Theorem t-truncated parity vector is that this random selection is preserved along the Collatz trajectory. That is, by picking any random number n_0 in the positive integers set, the probability, for every $t \geq 0$ on the trajectory $\tilde{f}^t(n_0)$, to be even, is the same as to be odd.

One of the main reviews on stochastic models of the $3x+1$ problems is found in the article of Kontorovich [81], presenting two *random walk* models for the orbit's evolution. These models are the root of obtaining a mathematical characterization of the hailstone sequence process and will be reviewed in the next two subsections.

2.4.1 Multiplicative Random Product (MRP) Model

First, let us consider a model consisting of a random product of independent and identically distributed (iid) variables to simulate the behavior of the Accelerated Collatz Map $\tilde{f}$.

Let Y_t be our dynamical random variable[1], constituted by a random product

$$\mathsf{Y}_t = \mathsf{X}_1 \mathsf{X}_2 \ldots \mathsf{X}_t, \tag{2.18}$$

where X_i are iid variables with distribution

$$\mathsf{X}_i = \begin{cases} \dfrac{3}{2} & \text{with probability } 1/2, \\ \\ \dfrac{1}{2} & \text{with probability } 1/2. \end{cases} \tag{2.19}$$

[1] In the text, mathematical sans serif fonts will be used to denote random variables

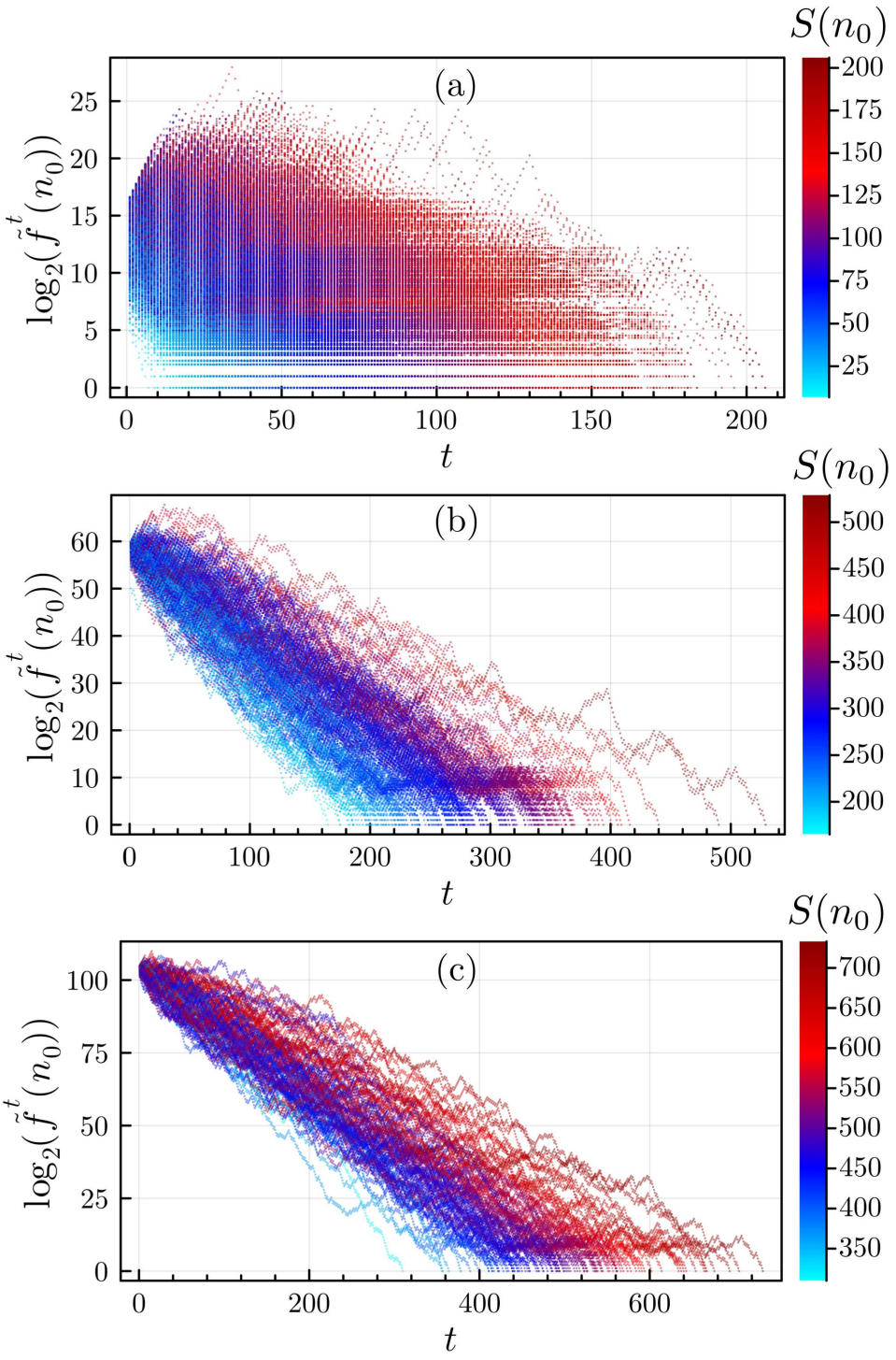


Figure 2.3: (a) Hailstone sequences from 10000 random initial conditions around $10 < n_0 < 10000$ in $\log_2$ scale. (b) Hailstone sequences from 200 random initial conditions around $2^{50} < n_0 < 2^{60}$ in $\log_2$ scale. (c) Hailstone sequences from 100 random initial conditions around $2^{100} < n_0 < 2^{105}$ in $\log_2$ scale. The colors indicate, as labeled in the color bar, the Total Stopping Time $S(n_0)$ of each orbit, cyan (dark red) meaning low (high) S values.

This model is called *Multiplicative Random Product* (MRP) Model, where the dynamical random variable Y_t represents the ratio between the orbit and the initial condition $\dfrac{\tilde{f}^t(n_0)}{n_0}$.

2.4.2 Biased Random Walk (BRW) Model

The second model considers an additive random walker, utilizes the MRP Model and brings much more important results concerning the characterization of the hailstone sequences process. Consider the dynamical deterministic variable y_t given by the natural logarithm of another variable x_t, i.e., $y_t = \ln x_t$ such that

$$y_t = \ln x_t := \log \tilde{f}^t(n), \tag{2.20}$$

meaning that the y_t variable represents the natural logarithm of the t-th step of the hailstone sequence. Then, by simply applying the natural logarithm to $\tilde{f}$ in Equation (2.2), one finds

$$y_{t+1} = \begin{cases} y_t + \ln\frac{3}{2} + e_t & \text{if } x \equiv 1 \pmod 2, \\ y_t + \ln\frac{1}{2} & \text{if } x \equiv 0 \pmod 2, \end{cases} \tag{2.21}$$

with

$$e_t := \log\left(1 + \frac{1}{3x_t}\right),$$

which gets smaller as $|x_t|$ gets larger. Therefore, in the large numbers regime

$$y_{t+1} \approx \begin{cases} y_t + \ln\frac{3}{2} & \text{if } x \equiv 1 \pmod 2, \\ y_t + \ln\frac{1}{2} & \text{if } x \equiv 0 \pmod 2. \end{cases} \tag{2.22}$$

This indicates to an *additive random walk* model that can describe the hailstone sequences in the large numbers regime, where $e_t \ll 1$. Let Z_t be the dynamical random variable in the form

$$\mathsf{Z}_t = \mathsf{Z}_0 + \ln \mathsf{Y}_t, \tag{2.23}$$

where Y_t comes from the MRP model in Equation (2.18), leading to

$$\ln \mathsf{Y}_t = \mathsf{w}_1 + \mathsf{w}_2 \cdots + \mathsf{w}_k, \tag{2.24}$$

with $\mathsf{w}_i = \ln \mathsf{X}_i$, and the steps w_i are distributed as

$$\mathsf{w}_i = \begin{cases} \ln\dfrac{3}{2} & \text{with probability } 1/2, \\ -\ln 2 & \text{with probability } 1/2. \end{cases} \tag{2.25}$$

This is the model of a random walker with steps $\mathsf{w}_i = \ln 3/2$ or $\mathsf{w}_i = -\ln 2$ with equal probability, such that the steps variables can be resumed into a single variable

$$\mathsf{W}_i := -\ln 2 + \mathsf{B}_i \ln 3, \tag{2.26}$$

where B_i is a zero-one random Bernoulli variable. Then, the position of the t-th step of

this random walker will be, starting from $\mathsf{Z}_0 = \ln \mathsf{X}_0$, given by

$$\mathsf{Z}_t = \mathsf{Z}_0 + \sum_{i=1}^{t} \mathsf{W}_i = \ln \mathsf{X}_0 + \sum_{i=1}^{t} (\mathsf{B}_i \ln 3 - \ln 2). \tag{2.27}$$

By t-truncated parity theorem this is a good model for the y_t deterministic variable, i.e., for the natural logarithm of the Collatz Map hailstone sequences: $\left\{\ln \tilde{f}(n_0)\right\}$. This model is called *Biased Random Walker* (BRW) model because it describes a random walker with drift.

2.4.3 Geometric Brownian Motion on Hailstone Sequences

Geometric Brownian Motion (GBM) characterizes processes whose logarithm of the variable of interest follows a Brownian motion [82]. A brief review of mathematical definitions concerning both Brownian and Geometric Brownian Motion is given in Appendix A. GBM is a key ingredient in solidly established models of self-reproducing phenomena, such as population [83], wealth [84, 85] and finances [86], as well as of bacterial cell division (reviewed in [87]).

From the BEW Model, a connection between the Collatz map and the GBM can be established. Consider the model where $\mathsf{Z}_t = \mathsf{Z}_0 + \ln \mathsf{Y}_t$, then Y_t can be written as

$$\mathsf{Y}_t = \exp\left[\mathsf{Z}_t - \mathsf{Z}_0\right] = \exp\left[\sum_{i=1}^{t} \mathsf{W}_i\right], \tag{2.28}$$

and, once the steps variables W_i are iid, Z_t is a Brownian Motion and Y_t is a GBM process with drift given by

$$\mu = \langle \mathsf{W}_i \rangle = -\ln 2 + \frac{1}{2} \ln 3 = \frac{1}{2} \ln\left(\frac{3}{4}\right) \approx -0.14384. \tag{2.29}$$

With this drift, one finds the tendency $\langle \mathsf{Z}(t) \rangle$ of Z_t values as

$$\langle \mathsf{Z}(t) \rangle = \mathsf{Z}_0 + \mu t, \tag{2.30}$$

and the total stopping time, starting from $\mathsf{Z}_0 = \ln \mathsf{X}_0$ will be the time $t = S(\mathsf{X}_0)$ when $\mathsf{Z}_t = 0$

$$S(\mathsf{X}_0) = \frac{1}{\mu} \mathsf{Z}_0 \approx 6.95212 \ln \mathsf{X}_0. \tag{2.31}$$

Once the BRW model describes the hailstone sequences, for a determined initial condition n_0, one finds that $S(n_0) \approx 6.95212 \ln n_0$.

By using the Odd Collatz Map $\tilde{\tilde{f}}$ defined in Equation (2.3), Sinai [59] proved the *Structure Theorem*. The Structure Theorem leads to analytical shreds of evidence the logarithm of the hailstone sequence is Brownian Motion and, consequently, the hailstone sequence itself is GBM. Furthermore, Kontorovich [60] provides a generalization of the Structure Theorem for the (d,g,h)-Maps $f_{d,g,h}$ (Equation (2.4)). This generalization leads to the conclusion that every (d,g,h)-map, in which $3x+1$ is a specific case, performs GBM.

2.5 Complexity and Scale-Free Behavior in the Collatz Map

As this chapter has shown so far, the behavior of the Collatz hailstone sequences is indeed complex, in the common sense of the word "complex", meaning that they are hard to analyze. One of the main features of the scientific concept of complexity is related to the interaction between small and large-scale features of physical systems, such as energy, distances, and time. The Collatz Map is not a physical system, and any direct analogy with one seems naive at first glance. Nevertheless, it is still an open problem to look for the minimum components necessary for complex behavior to arise. Hence, it is very fair to look for features of Complexity in the Collatz Map and, once its rule is very simple, maybe unveil some of the minimum components of complex behavior. In this section, the results of two articles, reporting complexity measures and scale-free behavior in hailstone sequences, are reviewed.

2.5.1 Intermitency and Complexity Indicators

From the aforementioned interaction between small and large-scale features of complex systems, one can mimic this by studying a variety of the Collatz Map, such as described by Casartelli [88]. Considering the Logistic Maps family [42]

$$x_{t+1} = L_\mu(x_t) = \mu x_t(1 - x_t), \tag{2.32}$$

and the Accelerated Collatz Map $\tilde{f}$, one can make different combinations to explore the small-large scale interplay.

For instance, using the Chaotic Map $L_4 : [0,1] \mapsto [0,1]$, let $\lfloor x_t \rfloor$ denote the integer part of x_t and $y_t = x_t - \lfloor x_t \rfloor$, then the map

$$x_{t+1} = C_1(x_t) = \begin{cases} \lfloor (3x_t + 1)/2 \rfloor + L_4(y_t) & \text{if } \lfloor x_t \rfloor \text{ is odd,} \\ \lfloor x_t/2 \rfloor + L_4(y_t) & \text{if } \lfloor x_t \rfloor \text{ is even,} \end{cases} \tag{2.33}$$

is one of the possible combinations, where L_4 plays the small-scale role, while $\tilde{f}$ enhances large-scale features.

Consider other three variations of C_1, namely: C_2, where exact quantities replace the integer parts in C_1,

$$x_{t+1} = C_2(x_t) = \begin{cases} (3x_t + 1)/2 + L_4(y_t) & \text{if } \lfloor x_t \rfloor \text{ is odd,} \\ x_t/2 + L_4(y_t) & \text{if } \lfloor x_t \rfloor \text{ is even,} \end{cases} \tag{2.34}$$

altering the next step evaluation of the integer part; $R_1(t)$ where $L_4(y_t)$ is replaced in C_1 by a noise random variable $r_t \in [0,1]$; and $R_2(t)$ where the same replacement of $L_4(y_t)$ by the random variable r_t is done in the $C_2(t)$ map.

From the analysis of the trajectories of C_1, C_2, R_1 and R_2, Casartelli [88] (a) compares deterministic versus random behavior in small scales, and (b) verifies the existence or not of feedback from large to small-scale. Figure 2.4 shows the orbits of C_1 and C_2, allowing us to compare both. In Figure 2.4(a) one can see that the trajectories present intermittent peaks and some irregularity, while Figure 2.4(b) presents trajectories much more regular than C_1, still with intermittency. Casartelli also tries to find an invariant measure for the maps — see [89] for details on the topic. Although, is important to notice that this

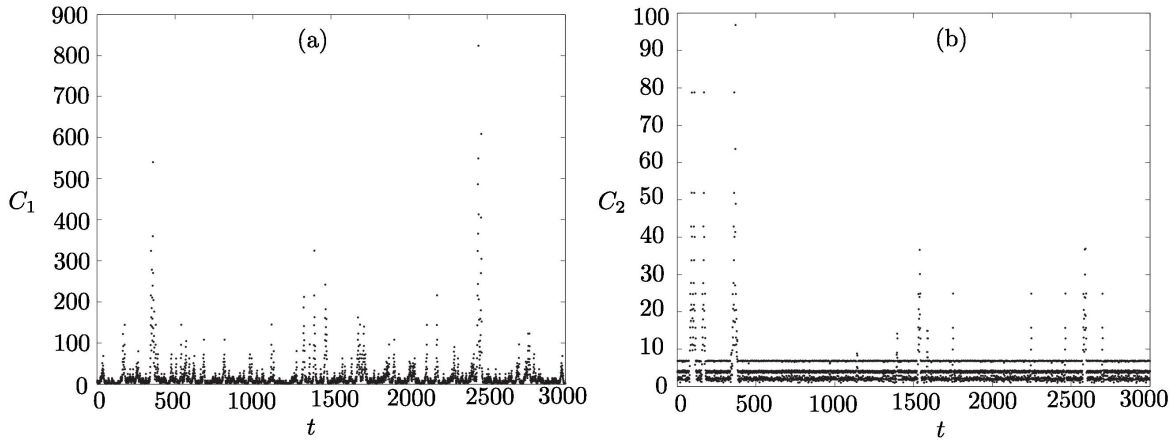


Figure 2.4: Orbits of (a) C_1, coupling the Logistic and Collatz Map as given by Equation (2.33), and (b) C_2, coupling the Logistic and Collatz Maps as in Equation (2.34). Both are adapted from [88].

computation presents some difficulties, especially for C_1, R_1, and R_2 whose trajectories are not even close to being bound. Also quantities such as Lyapunov Exponents (LEs) [90] and dynamical or Kolmogorov-Sinai (KS) entropy [91] are not obvious.

All results presented by [88] are coherent in separating C_1 and R_1 from C_2 and R_2, leading to the conclusion that choosing between random and determinism small scale feedback is irrelevant for these maps. The relevant factor is the coupling between small and large-scale components of the system, once C_1, R_1, and C_2, R_2 differ on this feature. In conclusion, besides these maps aren't physical but mathematical examples, their analysis may prove instructive for more general dynamics. The phenomenology depends not on the deterministic-random choices, but on the way, the large-scale component reacts to small scales.

2.5.2 Scale-Free Behavior in Hailstone Sequences

Following the philosophy of unveiling minimal models that display certain universal ingredients, da Luz [92] looked for typical aspects associated with physical criticalities such as power-laws and Self Organized Criticality (SOC) (see for example [4, 93]) in the hailstone sequences.

The article of da Luz [92] presents the study of six initial conditions, plus another 50 used for statistical analysis. Observing the hailstone sequence for these initial conditions, they indeed resemble one-dimensional (1D) (geometric) random walks (RW) — the GBM — with drift toward an absorbing boundary. This can be observed in Figure 2.5, where Figure 1 of da Luz [92] is reproduced. Thus, da Luz [92] states that if the Collatz Dynamics could be mapped into a biased 1D RW, the Collatz Conjecture would be demonstrated from the famous directed graph structure of the Collatz orbits.

A better insight into the general trend of the hailstone Collatz sequences is obtained by constructing a simple normalized histogram of $\{f^S(n_0)\}$ and generating the distribution $D(f)$. The log-log plot indicates that $D(f)$ follows a power law

$$D(f) = f^{-\alpha}. \tag{2.35}$$

Namely, α is found to be equal to 1, and $D(f) = 1/f$ is a signal commonly associated with

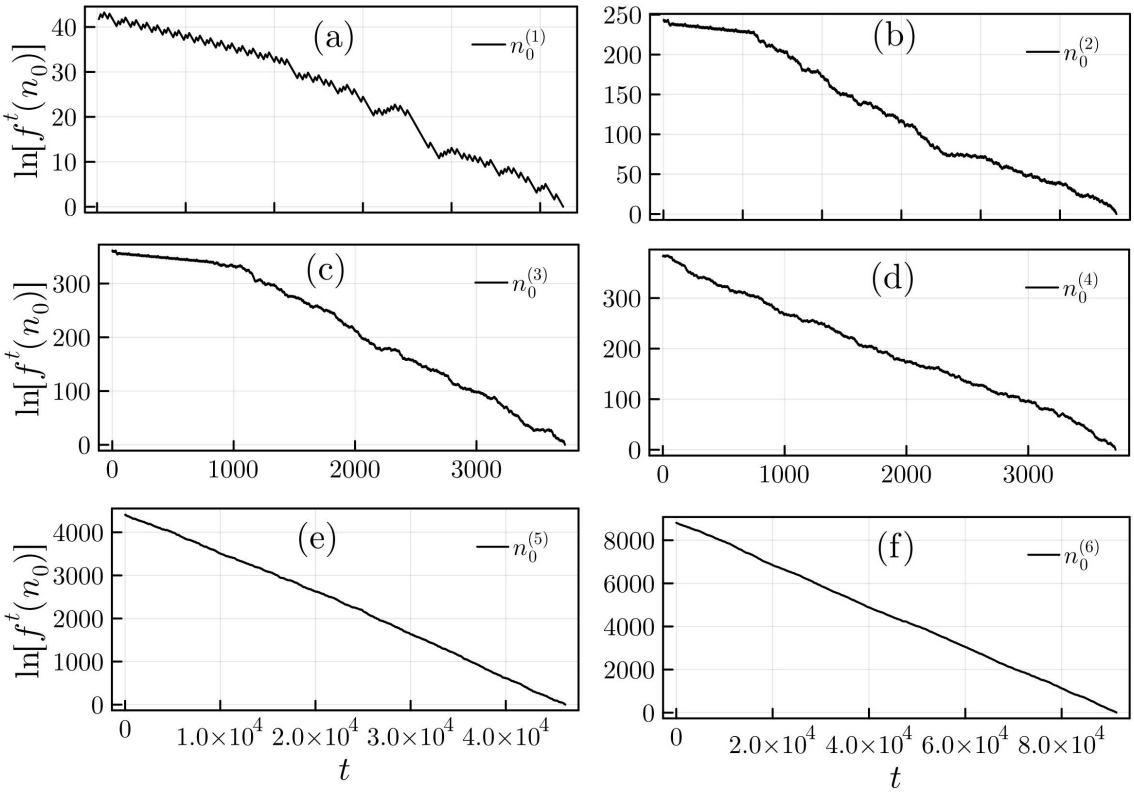


Figure 2.5: Reproduction of plot from the natural logarithm of Collatz hailstone sequences for the six initial conditions considered in [92].

self-organization in different systems. This behavior is plotted by circles in Figure 2.6.

Another variable that is found to follow power law distribution is the length l of avalanches (*i.e.* increasing/decreasing orbit). For instance

$$\mathcal{O}(48) = \{48,24,12,6,3,10,5,16,8,4,2,1\},$$

contains three decreasing avalanches: from 48 to 3 with $l_1 = 45$, from 10 to 5 with $l_2 = 5$, and from 16 to 1 with $l_3 = 15$. So if $n_{j,i}$ and $n_{j,f}$ are, respectively, the first and last terms of an avalanche, then $l_j = n_{j,i} - n_{j,f}$. Then, for larger orbits, it is possible to calculate a distribution $D(l)$, and it follows $D(l) \approx l^{-\alpha_{av}}$, with critical exponent $\alpha_{av} = 1.0$. This behavior is plotted in squares in Figure 2.6.

Finally, [92] looks for the mechanism that underpins the scale-free behavior. The observation of the numerical data leads to the comparison between the hailstone sequences and the motion of a particle with uniform velocity added with some noise. This is a characteristic of GBM. Statistical analysis, supported by the distinction between pure mathematics and physics, concludes that there is no statistically significant evidence against GBM, reinforcing literature results of Section 2.4.3. But the analysis of the avalanche sizes shows that they can be as large as $l = 2^{12}$ for the considered conditions, so the GBM is not the full story.

There are four general remarks and speculations on the implications of the findings. First, the results allow making a numerically testable theoretical prediction joining the GBM drift with the dynamics to estimate the time to reach 1, from an arbitrary initial condition. Second, even though Collatz Dynamics is not an example of SOC, one can establish

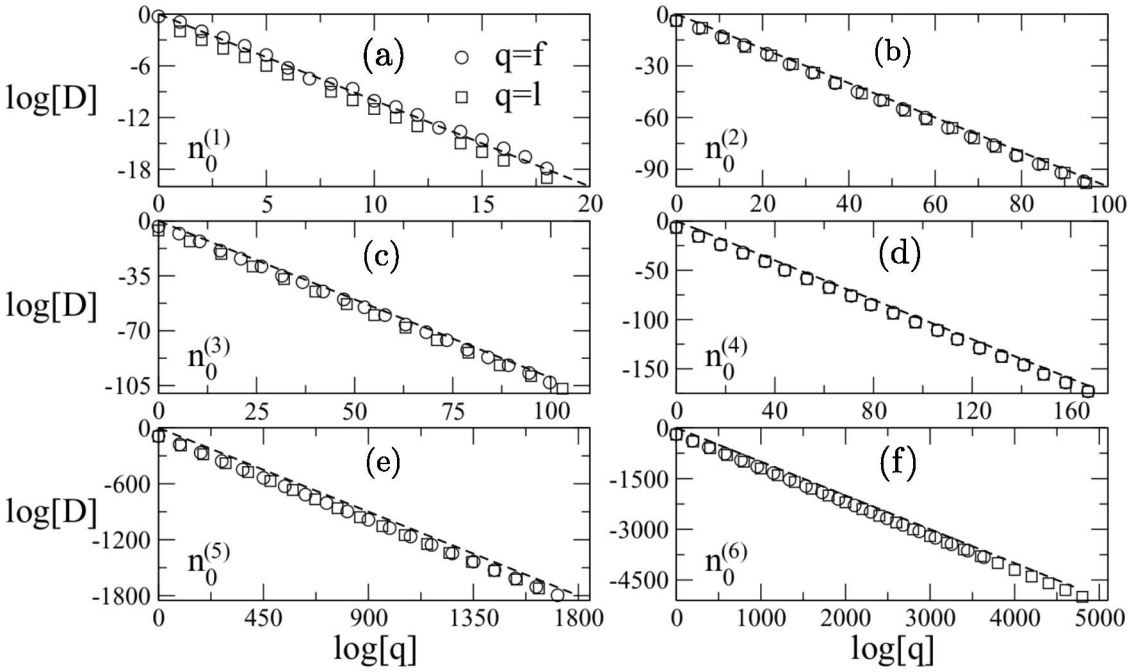


Figure 2.6: Log-log plots of normalized histograms of the hailstone sequences f (circles) and avalanche lengths l (squares). Taken with permission from [92]

some parallels. SOC models alternate between slow input and fast dissipation bursts of energy. For the Collatz map, suppose n_j to be some extensive quantity, proportional to the energy. Then for n large, the increase of energy might correspond to instances where $n \to n' \approx 3n/2$, whereas the fast release corresponds to the avalanches where $n \to n' = n/2^m$. Third, it would be worth asking how and why positive integers yield such remarkable properties under such a simple map. Fourth, an important signature in complex systems is long-range correlation [94, 95] are long-range power-law correlations. Appendix C of [92] demonstrates that highly correlated stretches actually can emerge.

2.6 Quantum Mechanical Oscilator and the Collatz Conjecture

As stated in the last section, it seems naive to try to force an analogy between the Collatz process and physical systems. But still, it is very striking the fact that various characteristics from natural phenomena are present in such a purely mathematical process.

This section encloses the current chapter with the review of a very recent paper from Perelman [96], applying quantum mechanical (QM) oscillator operators formalism into the Collatz Map. The objective of Perelman is not to build a toy model for a QM harmonic oscillator from the Collatz Map, but to make use of the formalism, to throw light into the hailstone sequences.

Perelman [96] establishes a dictionary between operators associated with the QM harmonic oscillator and the Collatz process, revealing clues as to why the Collatz conjecture most likely is true. The dictionary is based on the binary — also called 2-adic, see Section 3.1.1 on Chapter 3 for details on p-adics — expansion of the natural numbers $n = \sum_{\ell=0}^{N-1} a_{n\ell} 2^\ell$ (for $n < 2^N$) whose coefficients are $a_{n\ell} = \{0, 1\}$. By this expansion, one

constructs a one-to-one correspondence between n and state Ψ_n

$$n = \sum_{\ell=0}^{N-1} a_{n\ell} 2^{\ell} \longleftrightarrow |\Psi_n\rangle \equiv \sum_{\ell=0}^{N-1} a_{n,\ell} \, |\ell\rangle \, . \tag{2.36}$$

The state $|\Psi_n\rangle$ can be represented by a column vector whose entries are the $a_{n,\ell}$ coefficients

$$|\Psi_n\rangle = \begin{pmatrix} a_{n,0} \\ a_{n,1} \\ \vdots \\ a_{n,N-1} \end{pmatrix} . \tag{2.37}$$

By applying the original Collatz Map f for n_t even, when $a_{n_t,0} = 0$, there is a transition between the state $|\Psi_{n_t}\rangle$ towards the state $|\Psi_{n_{t+1}}\rangle = |\Psi_{n_t/2}\rangle$. This is obtained by observing that n_{t+1} is

$$n_{t+1} = \frac{n_t}{2} = 2^{-1} \sum_{\ell=0}^{N-1} a_{n_t\ell} 2^{\ell} = \sum_{\ell=1}^{N-1} a_{n_t\ell} 2^{\ell-1} \tag{2.38}$$

and therefore $|\Psi_{n_{t+1}}\rangle$ will be given by

$$|\Psi_{n_{t+1}}\rangle = |\Psi_{n_t/2}\rangle \equiv \sum_{\ell=1}^{N-1} a_{n_t,\ell} \, |\ell - 1\rangle \, . \tag{2.39}$$

For n_t odd, $a_{n_t,0} = 1$, and there occurs a transition between the state $|\Psi_{n_t}\rangle$ to $|\Psi_{n_{t+1}}\rangle = |\Psi_{3n_t+1}\rangle$. This is obtained by observing that n_{t+1} is

$$n_{t+1} = 3n_t + 1 = 1 + n_t + 2n_t = 1 + \sum_{\ell=0}^{N-1} a_{n_t\ell} 2^{\ell} + \sum_{\ell=0}^{N-1} a_{n_t\ell} 2^{\ell+1}, \tag{2.40}$$

therefore, the transition for $|\Psi_{3n_t+1}\rangle$ will be

$$|\Psi_{3n_t+1}\rangle = |0\rangle + \sum_{l=0}^{N-1} a_{n_t,\ell} \, |\ell\rangle + \sum_{\ell=0}^{N-1} a_{n_t,\ell} \, |\ell + 1\rangle \, . \tag{2.41}$$

Further, Perelman [96] also describes the operators acting on $|\Psi_{n_t}\rangle$ to evolve the system. Given the binary expansion of any even number, one finds a set of non-zero binary coefficients at locations $\ell_1, \ell_2, \ell_3 \ldots$, i.e., $a_{n\ell_1} = a_{n\ell_2} = a_{n\ell_3} = \cdots = 1$, with every other coefficient as zero. By using the well known creation $\mathbf{a}^{\dagger}$ and annihilation $\mathbf{a}$ operators on the energy eigenstates $|\ell\rangle$:

$$\mathbf{a}^{\dagger} \, |\ell\rangle = \sqrt{\ell + 1} \, |\ell + 1\rangle \, , \quad \mathbf{a} \, |\ell\rangle = \sqrt{\ell} \, |\ell - 1\rangle \, ,$$

one finds that the operator $\mathbf{L}_{n_t/2}$ mapping $|\Psi_{n_t}\rangle \to |\Psi_{n_t/2}\rangle$ is

$$\mathbf{L}_{n_t/2} = \mathrm{Diag}\left(\mathbf{1}, \mathbf{1}, \ldots, \frac{\mathbf{a}}{\sqrt{\ell_1}}, \ldots, \frac{\mathbf{a}}{\sqrt{\ell_2}}, \ldots, \frac{\mathbf{a}}{\sqrt{\ell_3}}, \ldots \right) , \tag{2.42}$$

a diagonal $N \times N$ matrix whose entries are the annihilation operator $\mathbf{a}$ and the identity operator $\mathbf{1}$. The identity operators are placed in the vanishing coefficients sites, while the annihilation operators are placed in the matrix sites of the non-vanishing coefficients, ℓ_i,

reducing these bits by unit. Therefore, one may write the operator $\mathbf{L}_{n_t/2} = \mathcal{L}^-$, playing the role of a ladder operator that reduces the bits by unit.

For n odd, the operator $\mathbf{L}_{3n_t+1}$ that maps $|\Psi_{n_t}\rangle$ to $|\Psi_{3n+1}\rangle$ is

$$\begin{aligned}\mathbf{L}_{3n_t+1} = \ &\text{Diag}\,(\mathbf{1},\mathbf{1},\mathbf{1},\ldots,\mathbf{1}) + \text{Diag}\left(\mathbf{a}^\dagger,\mathbf{1},\ldots,\frac{\mathbf{a}^\dagger}{\sqrt{\ell_1+1}},\mathbf{1},\ldots,\frac{\mathbf{a}^\dagger}{\sqrt{\ell_2+1}},\mathbf{1},\ldots,\right)\\ &+\text{Diag}\left(\mathbf{1},\mathbf{1},\ldots,\mathbf{a}^{\ell_1+1},\mathbf{1},\ldots,\mathbf{a}^{\ell_2+1},\mathbf{1},\ldots\right),\end{aligned} \tag{2.43}$$

which is also a diagonal $N \times N$ matrix. One may rewrite $\mathbf{L}_{3n+1}$ operator as

$$\mathbf{L}_{3n_t+1} = \mathbf{I} + \mathcal{L}^+ + \mathcal{P}_{\text{odd}}, \tag{2.44}$$

where $\mathbf{I}$ contains identity operators on the diagonal, $\mathcal{L}^+$ is a ladder operator increasing the bits by unit, while $\mathcal{P}_{\text{odd}}$ is a projection operator mapping the state Ψ_{n_t} (for n_t odd) into the ground state $\mathcal{P}_{\text{odd}}\,|\Psi_{n_t}\rangle = |\Psi_1\rangle = |0\rangle$.

With the above stated, it is clear that the operators are dependent on the eigenstates, hence one can write $\mathbf{L}(\Psi_{n_t})$. Then, the full hailstone sequence from n_0

$$\mathcal{O}(n_0) = \{n_0, n_1, n_2, \ldots, 4, 2, 1\},$$

given by the set of eigenstates

$$\{|\Psi_{n_0}\rangle, |\Psi_{n_1}\rangle, |\Psi_{n_2}\rangle, \ldots, |\Psi_4\rangle, |\Psi_2\rangle, |\Psi_1\rangle = |0\rangle, \},$$

where $\mathbf{L}(\Psi_{n_t})\,|\Psi_{n_t}\rangle = |\Psi_{n_{t+1}}\rangle$ determines the concatenation of operators

$$\mathbf{L}(\Psi_1)\mathbf{L}(\Psi_2)\mathbf{L}(\Psi_4)\ldots\mathbf{L}(\Psi_{n_2})\mathbf{L}(\Psi_{n_1})\mathbf{L}(\Psi_{n_0})\,|\Psi_{n_0}\rangle = \mathcal{P}_{n_0}\,|\Psi_{n_0}\rangle \equiv |0\rangle, \tag{2.45}$$

where $\mathcal{P}$ is either the projection operator for n_0 even, or odd. Equation (2.45) can be rewritten in a null-eigenfunction condition

$$\left(\mathbf{L}(\Psi_1)\mathbf{L}(\Psi_2)\mathbf{L}(\Psi_4)\ldots\mathbf{L}(\Psi_{n_2})\mathbf{L}(\Psi_{n_1})\mathbf{L}(\Psi_{n_0}) - \mathcal{P}_{n_0}\right)|\Psi_{n_0}\rangle = 0. \tag{2.46}$$

With these relations, it is fair to state that if the Collatz conjecture is true, it is equivalent to stating that the equation $\left(\mathbf{L}_*\mathbf{L}_*\ldots\mathbf{L}_* - \mathcal{P}\right)\Psi_n = 0$ has one solution for each positive integer. Then, the Collatz conjecture can be stated in another way, as an operator's identity:

Collatz Conjecture 2.5. *The operator's identity*

$$\mathbf{L}_*\mathbf{L}_*\ldots\mathbf{L}_* - \mathcal{P} = 0$$

for $\mathbf{L}_*$ *either* $\mathbf{L}_{n/2}$ *(Eq. (2.42)) or* $\mathbf{L}_{3n+1}$ *(Eq. (2.43)), has an infinite set of solutions in bijection with the positive integers' set. In other words, every positive integer* n *is associated with a unique projection operator* $\mathcal{P}_n$*.*

Perelman provides a deep discussion on the approaches of operator identities and null eigenfunctions, besides providing very instructive examples in Perelman [96] Appendix.

The results on the quantum mechanical oscilator operator for the Collatz map certainly are exciting and deserve further investigations and might prove useful not only to solve the $3x+1$ problem, but to provide advances in broader mathematics and physics fields. This

section encloses the present chapter, where a vast review of the literature of the Collatz map and its applications was performed. In the next chapter, the pertinent mathematical constructions and methods of analysis will be reviewed in order to provide foundations for the results to be presented about the analysis of the hailstone sequences.

Chapter 3

Mathematical constructions and analysis methods

This chapter has the objective of presenting and developing the mathematical foundations and methods implemented in the present work. It is divided into two blocks that, at a first sight, do not seem to be connected. Both are very important for the development of the results to be presented in Chapters 4 and 5. Section 3.1 encompasses the first block, concerning mathematical foundations justifying a new representation for natural numbers; Section 3.2 enclose the second block, versing on time series analysis methods.

Section 3.1.1 yields a brief review of topological algebra concepts to obtain the definition of an Ultrametric Space. The ultrametric space definition leads to the p-adic numbers. A visual approach to the ultrametric space of p-adics is presented in Section 3.1.2 with hierarchical trees. This visual approach of hierarchical trees leads to applications in complex phenomena. Finally, Section 3.1.3 presents a brand new representation for natural numbers, similar to p-adics for $p = 2$.

The second block begins with definitions, properties and characteristics of time series and stationary stochastic processes. Section 3.2.1 reviews the estimators of Moments and Cumulants of time series. Section 3.2.2 reviews the Pearson Correlation Matrix and the autocorrelation estimation. Section 3.2.3 verse on spectral properties of time series methods, followed by Section 3.2.4 presenting the Detrended Fluctuation Analysis. Finally, section 3.2.5 presents details on an entropy measure between time series.

3.1 Basic Elements for a natural number representation for the Collatz sequences

The signs used for the representation of mathematical objects and concepts are not unique. Consider the example of natural numbers, that can be written as we are used to, i.e., combinations of a closed set of ten Hindu-Arabic signs $(0, 1, 2, 3, 4, 5, 6, 7, 8, 9)$, in contrast with the Roman numerals representation (I, II, III, IV, ..., X, XI,...). As Dijkstra [97] points out, some arithmetic operations such as division get a lot more complicated with Roman numerals than with Hindu-Arabic numerals. From this point of view, one might benefit by finding new representations for natural numbers to perform specific tasks. In this direction, this section reviews the p-adic representation and its ultrametric space in order to introduce a new representation for natural numbers, useful for statistical analysis in the Collatz hailstone sequences.

3.1.1 Ultrametric Spaces and p-adics

The definition of an ultrametric space is related to the notion of a non-archimedean absolute value on a field. In this section, we shall make a digression, adapted from Gouvea [19], from the mathematical concept of a field, towards non-archimedean absolute values, to reach the definition of an ultrametric space.

A *field* $\Bbbk$ is a set provided with two operations $+$ and $\cdot$, such that $\Bbbk$ is a commutative group under $+$ and $\cdot$, with distributive law holding [98]. Adamson's book "Introduction to Field Theory" defines a field $\Bbbk$ as a commutative ring with identity and multiplication inverse [99]. An *absolute value*, also refered as *norm*, on a field $\Bbbk$ is a function $|\cdot| : \Bbbk \to \mathbb{R}^+$ satisfying:

Definition 3.1.1. (Absolute Value)

1. $|x| = 0$ if and only if $x = 0$;
2. $|xy| = |x||y|$ for all $x, y \in \Bbbk$;
3. $|x + y| \leq |x| + |y|$ for all $x, y \in \Bbbk$.

If only these three conditions are satisfied, the absolute value is called *archimedian*. The absolute value on $\Bbbk$ is called *non-Archimedean* if it satisfies an additional condition: the so-called *strong triangle inequality*.

Definition 3.1.2 (Non-Archimedean Absolute value)**.** *An absolute value will be called* non-archimedean *if it obeys the* strong triangle inequality

$$|x + y| \leq \max\{|x|,|y|\}\,. \tag{3.1}$$

The absolute value provides the notion of "size". Hence, it can be used to measure distances between numbers, defining a *metric* to the field $\Bbbk$ induced by the absolute value. With a metric in hands, one can investigate the topology of the field from the definition presented in Gouvea [19]:

Definition 3.1.3 (Metric induced by absolute value)**.** *Let $\Bbbk$ be a field and $|\cdot|$ an absolute value on $\Bbbk$. The distance $d(x,y)$ between two elements $x, y \in \Bbbk$ is*

$$d(x,y) = |x - y|.$$

And the function $d(x,y)$ is called the metric *induced by the absolute value.*

The important metric for our purposes is the metric induced by non-Archimedean absolute values. Non-Archimedean-induced metrics lead to the following lemma about the ultrametric inequality and, subsequently, to the notion of ultrametric spaces:

Lemma 3.1.4 (Ultrametric Inequality)**.** *Let $|\cdot|$ be an absolute value on a field $\Bbbk$, and define a metric by $d(x,y) = |x - y|$. Then, $|\cdot|$ is non-archimedean* if and only if *for any $x, y, z \in \Bbbk$, we have the* ultrametric inequality

$$d(x,y) \leq \max\{d(x,z),\ d(z,y)\}\,.$$

The proof of this Lemma is found in Section B.1 in Appendix B. With this Lemma, we have the definition of an ultrametric space:

Definition 3.1.5 (Ultrametric Space)**.** *An* ultrametric space *is a space whose metric respects the ultrametric inequality on a field* $\Bbbk$ *for every* x, y, $z \in \Bbbk$.

With this definition, we can explore one very particular example of ultrametric space: the compact topological ring of p-adic integers and rationals, providing some applications in complexity and critical phenomena as presented by Boettcher [25] and Sornette [95].

As pointed out in the Introduction, recently a large number of applications of $p-$adics are popping out in physics and many other areas of science [17]. The $p-$adic integers form a compact topological ring $\mathbb{Z}_p$ [18], where p denotes a prime number. Given $a \in \mathbb{Z}$ and p a prime number, then the $p-$adic of a is written as

$$a = a_0 + a_i p + a_2 p^2 + \cdots = \sum_{i \geq 0} a_i p^i, \tag{3.2}$$

with $0 \leq a_i \leq p-1$, and the set $\{a_0, a_1, \ldots, a_n, \ldots\}$ uniquely defining the $p-$adic of a. The set of $p-$adic coincides with the Cartesian space

$$X_p = \{0, 1, \ldots, p-1\}^{\mathbb{N}}, \tag{3.3}$$

of all possible combinations of these coefficients.

In the ring of p-adic integers $\mathbb{Z}_p$, the addition and multiplication are arithmetic and algebraically well defined as presented by Robert [18] and Koblitz [98]. Addition is briefly reviewed in section B.2 of Appendix B.

The concept of *order* of a p-adic, denoted by $v_p = \mathrm{ord}_p$, is important for the definition of its metric. So let $a = \sum_{i \geq 0} a_i p^i, a \neq 0$ be a p-adic integer, denote $v_p(a) = \mathrm{ord}_p(a)$ as the highest power of p which divides a, i.e., the first $a_i > 0$ from the p-adic representation. Then, the p-adic order is a mapping

$$v_p = \mathrm{ord}_p : \mathbb{Z}_p - \{0\} \to \mathbb{N}.$$

For example, let $p = 2$, on the 2-adic ring, for $a = 5$, one has $v_2(5) = 0$, once $2^0 = 1$ is the highest power of two dividing 5; now for $b = 12$, $v_2(12) = 2$, once $2^2 = 4$ is the highest power of two dividing 12.

The absolute value of a p-adic integer $x \in \mathbb{Z}_p$ is denoted by $|x|_p$. It can be defined, for a discrete topology X_p, as

$$|x|_p = \begin{cases} \dfrac{1}{p^{v_p(x)}} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases} \tag{3.4}$$

and then, the p-adic metric is defined, for $x = \{a_0, a_1, \ldots\}, y = \{b_0, b_1, \ldots\} \in X_p$, as

$$d(x,y) = \frac{1}{p^{v_p(x-y)}} \tag{3.5}$$

With this norm defined, the following lemma states the main result of this section about the metric space $(\mathbb{Z}_p, |\ |_p)$.

Lemma 3.1.6. *The metric space* $(\mathbb{Z}_p, |\ |_p)$ *is ultrametric.*

The proof of this lemma is found in Appendix B.

The ring $\mathbb{Z}_p$ is actually a subring of the ring $\mathbb{Q}_p$ of p-adic numbers [100]. This is so because, for given prime p, every rational $x \in \mathbb{Q}$ can be written as $x = p^m y$ for $y \in \mathbb{Z}$ and $m \in \mathbb{Z}$ [19]. Hence, any p-adic representation of an integer can be multiplied by p^m, and then every rational number is also represented as a p-adic. In $\mathbb{Q}_p$, one has multiplication inverse, hitting the final condition to the set be called a field. Any $n \in \mathbb{Q}$ can be written as a p-adic number

$$n = \sum_{i=-M}^{n} a_i p^i, \tag{3.6}$$

with the coeficients $\{a_{-M}, a_{-M+1}, \ldots, a_0, a_1, \ldots a_n\}$ with $0 \leq a_i < p$ are the p-adic number representation in $\mathbb{Q}_p$.

3.1.2 Hierarchical trees structures and Complex Systems

The best way to visualize p-adic space is through hierarchical trees as Rammal pointed out in 1986 [101] and as recent works still do (see for example Figure 1 in [102] and [103], or Figure 2 in [104]). The hierarchical tree visualization captures the ultrametricity of the p-adic integers space by distributing the numbers as leaves in a tree graph and setting the distance between them, as related to the number of generations from the last common edge between them.

This is easily seen in finite ultrametric spaces, such as the p-adic finite additive group $\mathbb{Z}_p/p^{2M}\mathbb{Z}_p$, represented as a rooted tree by Zambrano-Luna [103], with elements of the form

$$a = a_0 + a_1 p + a_2 p^2 + \cdots + a_{2M-1} p^{2M-1}, \tag{3.7}$$

with a_i's coefficients in $0 \leq a_i < p$.

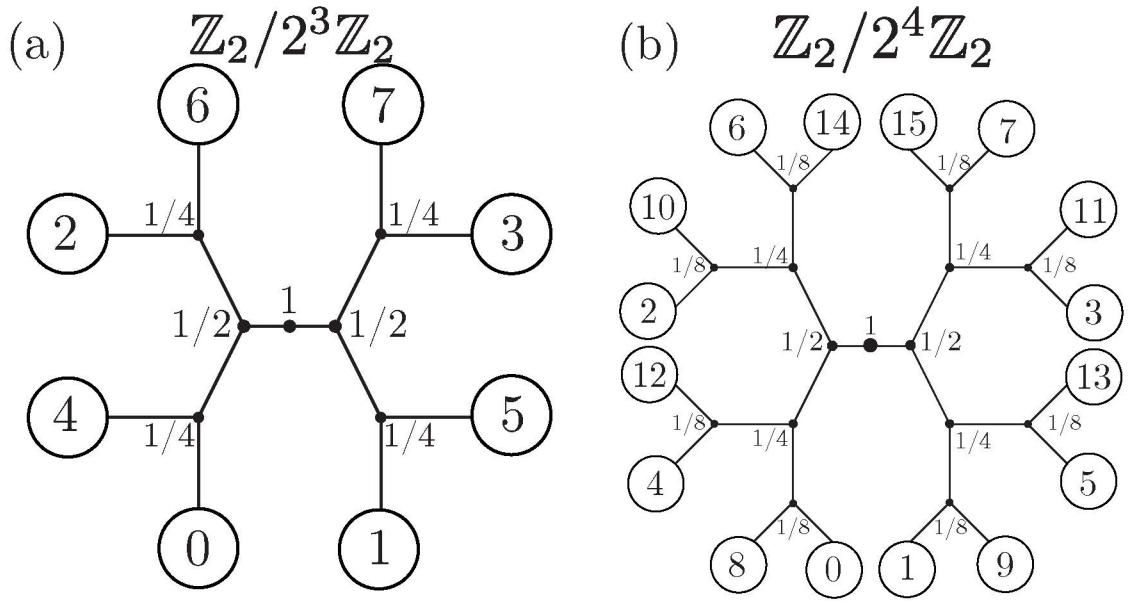


Figure 3.1: Rooted tree representation of (a) $\mathbb{Z}_2/2^3\mathbb{Z}_2$ and (b) $\mathbb{Z}_2/2^4\mathbb{Z}_2$ following the rules of construction from Zambrano-Luna [103]. The labels in the lower level vertices enounce that this is the p-adic distance from leaves whose this is the last common vertex.

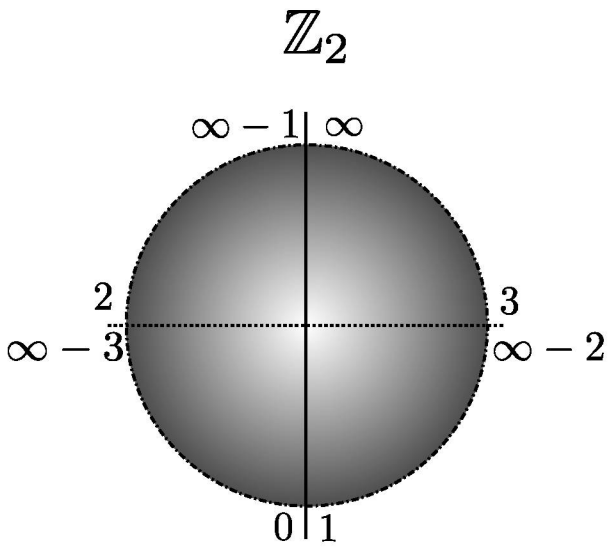


Figure 3.2: Scheme representing the rooted tree of the entire $\mathbb{Z}_2$ set. The line dividing the circle on the left-right and the dashed line dividing the circle on the top-bottom indicates that elements in the left-bottom border are closer to the left-top border than from the right-bottom border and vice-versa.

This space contains p^{2M} elements in a bijection with the top-level vertices of a rooted tree with $2M$ layers. By definition, the root of the tree contains only itself as a vertex on Level 0. Level 1 contains p vertices, corresponding to the possible a_i values. At level ℓ, with $1 \leq \ell \leq 2M$, there are p^ℓ vertices, each corresponding to a truncated $a = a_0 + \dots a_{\ell-1}p^{\ell-1}$ expansion. The connection from vertex in level ℓ corresponding to $a_0 + \cdots + a_{\ell-1}p^{\ell-1}$ with $a_0 + \cdots a_{\ell-2}p^{\ell-2}$ at level $\ell - 1$ occurs only if $a_0 + \cdots + a_{\ell-1}p^{\ell-1} - a_0 + \cdots a_{\ell-2}p^{\ell-2} \equiv 0$ mod $(p^{\ell-1})$, i.e. if the difference between the vertices is divisible by p^{l-1}. Following these rules, Figure 3.1 presents two examples of the 2-adic rooted trees for (a) $\mathbb{Z}_2/2^3\mathbb{Z}_2$ and (b) $\mathbb{Z}_2/2^4\mathbb{Z}_2$, showing the spacial distribution of elements of each set.

Two features of the p-adic numbers are well perceived by observing the graphs in Figure 3.1: (i) The ultrametricity of the space, where the distances between two elements i, j is given by $d = 2^{-v(i-j)}$ with $v(i-j)$ as the level of the last common ancestor between i, j in the tree; (ii) the fractal structure of the p-adic numbers [100], once every vertex gives rise to other p new vertexes in a chain, leading to an infinite graph representing the entire $\mathbb{Z}_p$, homeomorphic to fractal structure. By observing the pattern of numbers distribution, one can infer what will be the representation of $\lim_{M\to\infty} \mathbb{Z}_2/2^{2M}\mathbb{Z}_2 = \mathbb{Z}_2$. This is represented in a scheme in Figure 3.2, where one can perceive the non-trivial distribution of these numbers in the tree, where, for example, 0 and 1, besides being side-by-side, are as far from each other as 0 and ∞. Also, 0 is closer to[1] $\infty - 1$ than it is from 3, and even closer from $\infty - 3$ than from 2.

Hierarchical structures play a very important role in the description of complex systems and critical phenomena, mostly because of the large range of spatial, time, and energy scales that takes part in the complex dynamics [95]. Hence, any description of features of complex phenomena, such as space, time, and energy, would be very enlightened by representations with hierarchical structures. Curiously, ultrametric spaces, such as p-adics,

[1]Here we abuse the notation in order to be clearer in the point of non-triviality of the metric, in this representation, ∞ is given by $\lim_{M\to\infty} 2^{2M} - 1$, hence $\infty - 1$ would be $\lim_{M\to\infty} 2^{2M} - 2$ and so on.

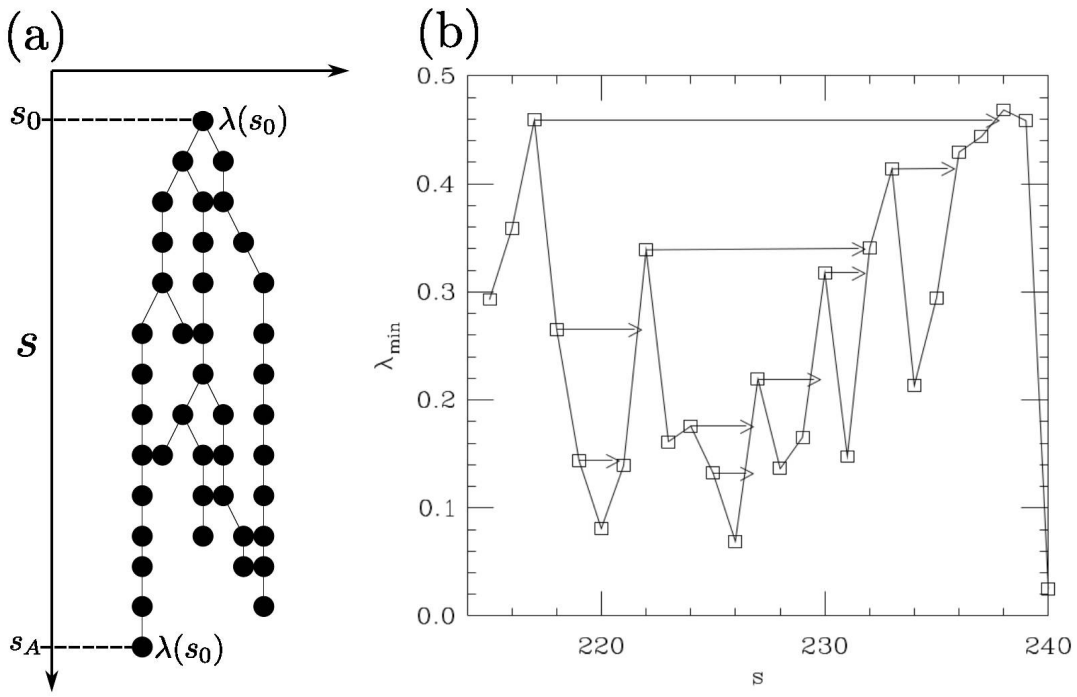


Figure 3.3: (a) Hierarchical tree of avalanches with ultrametric structure. Each time step contains the active sites, one of them is the initial common ancestor $\lambda(s_0)$, that only ends once every sub-avalanche is finished at s_A (Adapted from [25]). (b) Time-series of the evolution of $\lambda_{\min}$, the arrows represent active avalanche sites, that ends when $\lambda_{\min}$ is larger than the λ onset of that avalanche. It shows how a larger avalanche initiates various smaller ones that must end before the bigger one (Taken from [25]).

that at a first sight seems to be purely mathematical and non-realistic, because of their abstraction, fit very well for the description of hierarchical phenomena in complex systems. A few examples such as the p-adic parametrization of the Parisi matrix in the replica method, the method of hierarchical kinetics and the two-dimensional parametrization of the genetic code are reviewed by Kozyrev [105].

Systems presenting Self-Organized-Criticality (SOC), such as the Bak and Sneppen model for evolution [106] and its generalization [25], also can be described by hierarchical trees with ultrametric distances. For example, in the generalized Bak and Sneppen evolutionary model, each species is represented by a single site on a d-dimensional lattice. Each specie contains an array of M traits, represented by M numbers in the unit interval. The dynamics consist of mutating the smallest number λ in the collection of all traits of all species in the system. The mutation is represented by changing λ by a new random number in the unit interval. And as a dynamic effect of this mutation, the neighboring species from the one with the smallest trait also change one, randomly chosen, of the M numbers by another random number. Hence, the minimum trait of the entire system is a function of the generations s, i.e., $\lambda = \lambda(s)$.

An avalanche, in this case, is defined as a sequence of s_A steps in which $\lambda(s_0)$ is a minimum, and $\lambda(s) < \lambda(s_0)$ for every $s_0 \leq s < s_A$. For an avalanche started by $\lambda(s_0)$, at site j, this site is said to be active until the avalanche ends. It happens that, for some $\lambda(s)$ during the avalanche process started in s_0, smaller avalanches are initiated, generating a hierarchical structure of avalanches, that only finishes when the first avalanche ends.

Figure 3.3(a), adapted from Boettcher [25], shows an avalanche hierarchical structure. This hierarchical structure is ultrametric, with the distance between two simultaneous active avalanches given by the distance back in time from the common ancestor. Further, this structure is not only a mere curiosity, but plays a role, once the smaller avalanches behave like barriers against the end of the larger avalanche (see Figure 3.3(b)), and this is very similar to spin glasses dynamics as Sornette [95, p. 430] states.

3.1.3 m-vectors Representation of Natural Numbers

As the previous sections of this chapter have evidenced, new insights into complex systems can be obtained by studying their dynamics in other spaces than usual. As Chapter 2 evidenced, the Collatz hailstone sequences present features very alike to those present in complex systems, such as $1/f$ distributions, GBM, and SOC. The usual space for hailstone sequences is the natural numbers, but their study in different spaces such as p-adic spaces is very common, as T. Tao does in his recent article using 3-adics [67] (see [57] for a vast review on p-adic in $3x+1$ problem). But the p-adic numbers' representation does not extinguish every non-trivial representation of natural numbers, not even enlighten every hidden feature of the Collatz hailstone sequences. Hence, this section shall present an useful representation (first appearing to the author's knowledge in da Luz [92]) for natural numbers, connected with 2-adics, giving rise to a very peculiar space: the $\boldsymbol{m}$-vector's algebraic structure.

The origin of this new representation is not related to p-adics or any trial of studying the hailstone sequences in new spaces, but with the necessity of da Luz [92] to study large orbits and obtain statistically significant results. As already pointed out in Chapter 2, very large orbits are usually generated by very large initial conditions. These are heuristic arguments, and the order of magnitude of the S (the total stopping time) and n_0 can be very flexible. But certainly, the creation of very large and suitable initial conditions is a hard effort. At this point, while investigating large initial condition's orbits, da Luz [92] proposes a different representation for the natural numbers. This new representation consists of writing any natural number $n \in \mathbb{N}^+$ as

$$\begin{aligned} n &= 2^{m_1}\left(2^{m_2}\left(\cdots\left(2^{m_{r-1}}\left(2^{m_r}+1\right)\cdots+1\right)+1\right)+1\right), \\ &= 2^{m_1}+2^{m_1+m_2}+\cdots+2^{m_1+m_2+\cdots+m_{r-1}+m_r}, \end{aligned} \tag{3.8}$$

or

$$n \equiv \boldsymbol{m}(n) = (m_1, m_2, m_3, \ldots, m_{r-2}, m_{r-1}, m_r). \tag{3.9}$$

where $\boldsymbol{m}(n)$ is uniquely determined given n. In other words, the natural numbers can be represented as a unique, from now on called, $\boldsymbol{m}-$vector with r components. The details of the algorithm of construction can be consulted in Appendix A of da Luz [92]. This new representation will be central to the next two chapters of this dissertation.

This representation has a direct connection with the 2-adic. Given a positive integer n with $\boldsymbol{m}$-vector representation given by Equation (3.9), one can write the su$\boldsymbol{m}$-vector. The su$\boldsymbol{m}$-vector, denoted by $\boldsymbol{\Sigma}(\boldsymbol{m})$, consists on a vector whose k-th component, denoted by

$\Sigma_k(\boldsymbol{m})$, is the sum of the first k components of the $\boldsymbol{m}$-vector. Thus, one can write

$$\Sigma_k(\boldsymbol{m}) = \sum_{i=1}^{k} m_i, \tag{3.10}$$

$$\boldsymbol{\Sigma}(\boldsymbol{m}) = (m_1, m_1+m_2, \ldots, , \sum_{i=1}^{r-1} m_i, \sum_{i=1}^{r} m_i). \tag{3.11}$$

The su$\boldsymbol{m}$-vector provides the connection with the 2-adic numbers. Let

$$\boldsymbol{a} = (a_0, a_1, a_2, \ldots, a_{N-2}, a_{N-1}), \tag{3.12}$$

with $a_i = \{0,1\}$, to be the finite 2-adic representation of the same positive integer n with su$\boldsymbol{m}$-vector given by Equation (3.11). Then, one can write $\boldsymbol{a}$ as a function of $\boldsymbol{m}$ using $\boldsymbol{\Sigma}$ as mediator of the process, by

$$\boldsymbol{a}(\boldsymbol{m}): \; a_i(\boldsymbol{m}) = a_i(\boldsymbol{\Sigma}(\boldsymbol{m})) = \begin{cases} 1 & \text{if } i \in \boldsymbol{\Sigma}(\boldsymbol{m}), \\ 0 & \text{if } i \notin \boldsymbol{\Sigma}(\boldsymbol{m}), \end{cases} \tag{3.13}$$

where the only non-vanishing coefficients of the 2-adic are those whose index value is present in the su$\boldsymbol{m}$-vector. In other words, only

$$a_{m_1} = a_{m_1+m_2} = \cdots = a_{\sum_{i=1}^{r-1} m_i} = a_{\sum_{i=1}^{r} m_i} = 1,$$

while every other component is zero.

It is instructive to exemplify this. For example, let $n = 21$, then the $\boldsymbol{m}$-vector, the su$\boldsymbol{m}$-vectors and the 2-adic expansion are

$$\boldsymbol{m}(21) = (0,2,2), \tag{3.14}$$

$$\boldsymbol{\Sigma}(\boldsymbol{m}) = (0,2,4), \tag{3.15}$$

$$\boldsymbol{a}(21) = (1,0,1,0,1), \tag{3.16}$$

and the only non-vanishing components of the 2-adic representation are be a_0, a_2, a_4, as it should from Equation (3.13).

This concludes the first block of this chapter, where the main representation of natural numbers to be exploited in this work was formally presented. Now, the methods of time series analysis to be applied on the hailstone sequences itself and on the time series of $\boldsymbol{m}$-vectors shall be presented.

3.2 Methods of Time Series Analysis

A *time series*, or *signal*[2], is a sequence of observations sequentially ordered on time [107]. Typically, the time series are dependent on the nature of the process generating the sequence of observations. The methods of *time series analysis* are concerned with unfolding the processes behind the signals. The present section reviews the main mathematical tools of time series analysis methods used to obtain the results of Chapter 5.

Time series can be mathematically denoted by functional representations: let $x(t)$ represent the value of a signal on time t, and the set $\boldsymbol{x} = \{x(t)\}$ for $t \in [0,T]$ to be the

[2]Both terms, "time series" and "signals" will be used as synonyms along the text.

complete time series representation up to time T. Signals can be classified depending on the features they display [108, 107].

Concerning the time set where the observations take place, the time series can be of *continuous* or *discrete* time. Continuous time series takes place in continuous time sets, meaning that $\boldsymbol{x}(t)$ is a continuous function of time. Discrete time series takes place in discrete time sets, and the observations are made in specific instants $t \in \{t_1, t_2, \ldots, T\}$. Thence, one can represent a discrete time series with T observations by a vector with T components $\boldsymbol{x} = (x(t_1), x(t_2), \ldots, x(T))$.

Concerning the periodicity of the time series, they can also be classified as *periodic*, for those signals periodically repeated in time, i.e., when $x(t) = x(t + P)$ for any t, or *non-periodic*, when no period is found. The periodicity of the time series will be explored at Section 3.2.3.

Further, time series can be classified according to our knowledge of their future, i.e., the predictability of the time series. If the signal's future and past are completely determined by some known function $F(t)$, then the time series is *deterministic*, and no random noise affects the evolution of $x(t)$. Now, if there are uncertainty elements in the evolution of the signal, they are called *stochastic* or *random* time series.

Besides that hailstone sequences, the main time series to be analyzed in this dissertation, are deterministic time series, their pseudo-random behavior, reviewed in Chapter 2 encourages one to apply stochastic time series methods to analyze them. Hence, the next paragraphs deal with the analysis of stationary stochastic processes, providing a theoretical basis for the methods of forthcoming Sections and the results in the next Chapters.

In stochastic processes, time series evolves according to probabilistic laws. For this reason, one must associate the observation of time series $\boldsymbol{x}$ at one time t as a realization of a random variable z_t with probability distribution $p(\mathsf{z}(t))$. Similarly, observations of the time series $\boldsymbol{x}$ in two times, t_1 and t_2, are realizations of two random variables z_{t_1} and z_{t_2} with *joint probability distribution* $p(\mathsf{z}_{t_1}, \mathsf{z}_{t_2})$. In general, one can describe m observations in a time series, placed in $t_1, t_2, \ldots, t_m$, by m random variables $\mathsf{z}_{t_1}, \mathsf{z}_{t_2}, \ldots, \mathsf{z}_{t_m}$ with joint probabilities distributions $p(\mathsf{z}_{t_1}, \mathsf{z}_{t_2}, \ldots, \mathsf{z}_{t_m})$.

One very important feature of the methods of time series analysis presented in the next sections is that they are developed to analyze *stationary* stochastic signals. According to Box [107], a stochastic process is said to be strictly stationary if none of its properties are affected by a change of time origin. That is, the joint probabilities distributions associated with m observations at times $t_1, \ldots, t_m$, given by $(p(\mathsf{z}_{t_1}), \ldots, p(\mathsf{z}_{t_m}))$ are the same for other m observations at times $t_1 + k, \ldots, t_m + k$, that is $(p(\mathsf{z}_{t_1}), \ldots, p(\mathsf{z}_{t_m})) = (p(\mathsf{z}_{t_1+k}), \ldots, p(\mathsf{z}_{t_m+k}))$, for any k. A process is called mth-order stationary if the joint probabilities associated with m observations are constant.

3.2.1 Moments and Cumulants

In practice, one does not have access to $p(\mathsf{z}_t)$, once the time series is finite. Then, for a discrete signal $\boldsymbol{x}(t) = (x(1), \ldots, x(T))$ with T time-steps, one can estimate the mean

$$\mu = \langle \boldsymbol{x} \rangle = \frac{1}{T} \sum_{t=1}^{T} x(t), \tag{3.17}$$

and the variance

$$\sigma^2 = \langle (\boldsymbol{x} - \langle \boldsymbol{x} \rangle)^2 \rangle = \frac{1}{T} \sum_{t=1}^{T} (x(t) - \langle \boldsymbol{x} \rangle)^2 . \tag{3.18}$$

From now on, the symbol $\langle\cdot\rangle$ will denote the estimated expected value of the quantity inside the brackets.

For the case of time-dependent mean processes, such as Brownian Motion with drift (see Appendix A) one can perform the following transform

$$\boldsymbol{x}' = \frac{\boldsymbol{x} - \langle \boldsymbol{x} \rangle}{\sqrt{\langle (\boldsymbol{x} - \langle \boldsymbol{x} \rangle)^2 \rangle}}, \tag{3.19}$$

to transform any non-stationary process into a first ordered stationary process [108][3].

The mean and variance are specific cases of two more general statistical measurements: the moments and cumulants of a distribution [109]. The estimation of the *moment* of order n of a general process $\boldsymbol{x} = (x_1, \ldots, x_T)$ with T steps is given by the expected value of $\boldsymbol{x}^n$

$$\mu_n = \langle \boldsymbol{x}^n \rangle = \frac{1}{T} \sum_{t=1}^{N} x_i^n. \tag{3.20}$$

The moments can be associated with a moment-generating function of $\boldsymbol{x}$

$$M(t) \equiv M_{\boldsymbol{x}}(t) = \langle \exp(tX) \rangle, \tag{3.21}$$

that can be Taylor expanded around the origin

$$M_{\boldsymbol{x}}(t) = \sum_{n=0}^{\infty} \frac{\mu_n t^n}{n!}. \tag{3.22}$$

From $M_{\boldsymbol{x}}(t)$, the cumulant generating function is obtained as [110]

$$\mathcal{C}(t) \equiv \mathcal{C}_{\boldsymbol{x}}(t) = \log M_{\boldsymbol{x}}(t) = \sum_{n=0}^{\infty} \frac{c_n t^n}{n!}, \tag{3.23}$$

where c_n is the n-th cumulant. This leads to a relationship between the first four moments and cumulants, by extracting coefficients from the expansion, as

$$c_1 = \mu_1 \quad \text{(Mean)}, \tag{3.24}$$
$$c_2 = \mu_2 - \mu_1^2 = \sigma^2 \quad \text{Variance } (\sigma = \text{standard deviation}), \tag{3.25}$$
$$c_3 = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3 \quad \text{(Skewness)}, \tag{3.26}$$
$$c_4 = \mu_4 - 3\mu_2^2 - 4\mu_1\mu_3 + 12\mu_1^2\mu_2 - 6\mu_1^4 \quad \text{(Kurtosis)}. \tag{3.27}$$

One can see that the mean and variance are only the first and second cumulants, respectively. The third and fourth cumulants, skewness and kurtosis, respectively, are important in

[3] This is so once the mean estimation

$$\langle \boldsymbol{x}' \rangle = \langle \boldsymbol{x} - \langle \boldsymbol{x} \rangle \rangle = \langle \boldsymbol{x} \rangle - \langle \langle \boldsymbol{x} \rangle \rangle = 0,$$

and the variance estimation

$$\langle (\boldsymbol{x}' - \langle \boldsymbol{x}' \rangle)^2 \rangle = \langle \boldsymbol{x}'^2 \rangle = \left\langle \left(\frac{\boldsymbol{x} - \langle \boldsymbol{x} \rangle}{\sqrt{\langle (\boldsymbol{x} - \langle \boldsymbol{x} \rangle)^2 \rangle}} \right)^2 \right\rangle = \left\langle \frac{[\boldsymbol{x} - \langle \boldsymbol{x} \rangle]^2}{\langle (\boldsymbol{x} - \langle \boldsymbol{x} \rangle)^2 \rangle} \right\rangle = 1,$$

are both constant.

order to characterize the deviation of the distribution from a normal distribution. In a normal distribution, all cumulants of order higher than 2 are zero.

3.2.2 Correlation and Autocorrelation Functions

The ability to tell whether a set of observations $\{\boldsymbol{x}, \boldsymbol{y}, \ldots, \boldsymbol{w}\}$ are correlated is fundamental to many areas of science, such as Physics [111], Biology [112], Archaeology [113] and social sciences [114], just citing a few examples — for a review on multivariate statistics see [115] and applications see [116].

One of the most used linear methods to determine correlation is the Pearson Correlation Coeficient [117]. Let $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N)$ be N time series from N processes. Denoting the i-th time series of a process by $\boldsymbol{x}_i = [x_i(1), \ldots, x_i(T)]$, with length T, let $\Delta\boldsymbol{x}_i = x_i(t) - \langle \boldsymbol{x}_i \rangle$ represent the deviations from the estimated mean $\langle \boldsymbol{x}_i \rangle$ of the i-th time series, and $\langle (\Delta \boldsymbol{x}_i)^2 \rangle$ to be the estimated variance of $\boldsymbol{x}_i$. Then, the correlation between $i-$th and $j-$th time-series is given by the Pearson correlation coefficient

$$R_{ij} = \frac{1}{T} \sum_{t=1}^{T} \left(\frac{x_i(t) - \langle \boldsymbol{x}_i \rangle}{\langle \Delta \boldsymbol{x}_i \rangle} \right) \left(\frac{x_j(t) - \langle \boldsymbol{x}_j \rangle}{\langle \Delta \boldsymbol{x}_j \rangle} \right) = \frac{\langle \Delta \boldsymbol{x}_i \Delta \boldsymbol{x}_j \rangle}{\langle \Delta \boldsymbol{x}_i \rangle \langle \Delta \boldsymbol{x}_j \rangle}. \tag{3.28}$$

The Pearson correlation matrix is a $N \times N$ matrix $\boldsymbol{R}$ whose entries R_{ij} are given by Equation (3.28). That is

$$\boldsymbol{R} = \begin{pmatrix} 1 & \cdots & R_{1,N} \\ \vdots & \ddots & \vdots \\ R_{N,1} & \cdots & 1, \end{pmatrix} \tag{3.29}$$

is a hermitian matrix, once the correlation of a time series with itself is maximum, i.e., $R_{ii} = 1$ and $R_{ij} = R_{ji}$.

From the Pearson correlation between different time series, one can think of calculating the Pearson correlation between elements of the same time series. This is the called *autocorrelation* of a time series, enabling one to measure how changes in a signal are related to future changes in the same signal. The estimated autocorrelation at lag τ is obtained from the *autocovariance*. For a signal $\boldsymbol{x} = [x(1), \ldots, x(T)]$ with length T, one can estimate the autocovariance at lag τ as

$$C_\tau = \frac{1}{T} \sum_{t=1}^{N-\tau} \left(x(t) - \langle \boldsymbol{x} \rangle \right) \left(x(t+\tau) - \langle \boldsymbol{x} \rangle \right), \quad \tau = 0, 1, 2, \cdots, \tag{3.30}$$

leading to the autocorrelation

$$R_\tau = \frac{C_\tau}{C_0}. \tag{3.31}$$

This can be estimated by

$$R_\tau = \frac{\sum_{t=1}^{T-\tau} \left(x(t) - \langle \boldsymbol{x} \rangle \right) \left(x(t+\tau) - \langle \boldsymbol{x} \rangle \right)}{\sum_{t=1}^{T} \left(x(t) - \langle \boldsymbol{x} \rangle \right)^2}, \tag{3.32}$$

or, simply the Pearson correlation coefficient, from Equation (3.28), between a time series and itself lagged by τ: that is $R_{t,t+\tau}$. When one lets τ vary, one finds the autocorrelation

function (ACF) estimation, generating a set of autocorrelations with different lags

$$R(\boldsymbol{\tau}) = \{R_{t,t}, R_{t,t+\tau_1}, \ldots, R_{t,t+\tau_k}\}, \tag{3.33}$$

where k can be as long as necessary to define the process.

When estimating the autocorrelation of a time series, one may ask what is the meaning of significant correlations and what is the threshold of no-correlation. Defining a threshold for no-correlation is still an open problem in literature (see [118] and [**felippe-2021**] for discussions). But still, textbook relations can be used to estimate the error of the autocorrelation estimation, $\sigma(R(\boldsymbol{\tau}))$, for a finite time series with length T. It can be done twofold: first, if one considers that the time series values are i.i.d. such that $\sigma(R(\boldsymbol{\tau})) = 1/\sqrt{T}$. But this is only valid for real i.i.d. time series, thence, for trying to investigate the existence or not of correlations, one must consider Bartlett's approximation [107]. The Bartlett's approximation, denoted as $\sigma(R(\boldsymbol{\tau})) = \sigma_{R(\boldsymbol{\tau})}$ is given by

$$\sigma^2_{R(\boldsymbol{\tau})} = \frac{1}{T}\sum_{\nu=-\infty}^{\infty}\left(R_\nu^2 + R_{\nu+\tau}R_{\nu-\tau} - 4R_\tau R_\nu R_{\nu-\tau} + 2R_\nu^2 R_\tau^2\right), \tag{3.34}$$

where one recovers the case $\sigma(R(\boldsymbol{\tau})) = 1/\sqrt{T}$ for i.i.d variables, i.e., for $R_0 = 1, R_{\tau>0} = 0$. For finite time series autocorrelation functions, with length T, one finds the estimation of $\sigma_{R(\boldsymbol{\tau})}$ by setting $R_{\tau<0} = R_{\tau>T} = 0$.

3.2.3 Power Spectrum (PS)

When studying time series, the periodicity of the signal is a very important characteristic of the phenomena behind that particular signal. A very useful way of looking at the periodicity of signals is finding its spectrum, or, the frequency domain signal via *Fourier Analysis* [108]. The Fourier Analysis consists on writing a signal $\boldsymbol{x}$ with length T in terms of harmonic waves as a Fourier Series [119]

$$x(t) = \sum_{m=-\infty}^{\infty} z_m e^{2\pi i m f_0 t}, \tag{3.35}$$

where i is the imaginary unit, f_0 is the fundamental frequency ($f_0 = 1/T$). The coefficients z_m, with m assuming integer values, are called the Fourier coefficients of the signal $\boldsymbol{x}$.

The infinite set of Fourier coefficients $\{z_m\}$ is unique, meaning that two different signals are described by two different sets of coefficients. Obtaining the Fourier coefficients, in practice, means calculating the Fourier Transform [119, 120] of the signal. For discrete time series, this is done by the Discrete Fourier Transform (DFT) for finite time series $\boldsymbol{x}$ with length T, by

$$\hat{\boldsymbol{x}} = \hat{x}(f) = \sum_{t=1}^{T} x(t)e^{-2\pi i f t}, \tag{3.36}$$

for $f = mf_0 = m/T$. Appendix C provides details on a computational technique to fasten the calculation of the DFT called Fast Fourier Transformation (FFT) [121].

Perhaps a more physical approach consists of asking what is the energy associated with each frequency on a signal. Further on this approach, one can ask what is the distribution of energy in the frequency domain, and ask if signals with the same distribution are originated from the same kind of process.

Given a time series $\boldsymbol{x}$, the Power Spectrum $P(f)$ [122] is a measurement of the distribution of energy in the frequency domain. The Power Spectrum $P(f)$ for the signal $\boldsymbol{x}$, with length T, is estimated through

$$P(f) = \frac{1}{T}\left|\hat{x}(f)\right|^2 = \frac{1}{T}\left|\sum_{t=1}^{T} x(t)e^{-2\pi i f t}\right|, \tag{3.37}$$

for $f \in [f_0 = 1/T, 1/2]$.

One of the main power spectrum distributions, often associated with Complexity [4] and self-similarity [123], are power laws $1/f$ signals, characterized by large energy in lower frequencies and power law decay.

When $\beta = 0$, the signal is called white noise, and one finds a constant distribution of energy in the frequency domain $P(f) \propto f^0$. This is a special case when the time series is composed of as many frequencies as possible, all of them with the same energy. Processes following white noise are related to the absence of memory, since the steps are completely uncorrelated, meaning the autocorrelation function is zero for every $\tau > 0$.

In the case of $\beta = 1$, one finds a signal called pink noise, where the energy decays exactly with the inverse of the frequency. This is the classical case of $1/f$ noise, directly associated with self-similar processes in nature [123].

When $\beta = 2$, one has the red noise, often called Brownian, because this is the distribution for the Brownian motion. The Brownian noise is a process composed of uncorrelated increments, i.e., the time series for the increments of the signal is white noise.

3.2.4 Detrended Fluctuation Analysis (DFA)

Another method for time series characterization is obtaining the Hurst Exponent H [124]. This method characterizes the average velocity of variance growing of a process. The Hurst Exponent consists of the generalization for characterizing diffusion, where the relationship between variance and time is a nontrivial power law

$$\langle \Delta \boldsymbol{x}^2 \rangle \propto \Delta t^{2H}, \tag{3.38}$$

for $H \in [0,1]$, one recovers the Brownian motion with $H = 1/2$ for uncorrelated increments, and $H > 1/2$ ($H < 1/2$) characterizing super- (sub-) diffusive processes for correlated increments [111].

But, as stated by Kantz [124], it is not advisable to estimate the Hurst exponent of measured data, once real-world data might carry trends, resulting in either a trivial exponent $H = 1$ or more often, do not exhibit a clear scaling. A method called Detrended Fluctuation Analysis (DFA) has been largely employed to avoid such issues since it was developed by [125] to study the organization of DNA nucleotides. The DFA method was applied to various dynamical phenomena including physiological data [126, 127], financial market [128], and climate data [129] just citing a few examples.

Given a time series $\boldsymbol{x} = (x(1), \ldots, x(T))$, the DFA is calculated by segmenting it into parts of equal length ℓ, for $\ell \in [n_0, T/2]$. That is, from $\boldsymbol{x}$, one finds a set of segments $\boldsymbol{x}_\ell(t) = (x_\ell(t), \ldots, x_\ell(t+\ell))$. The points of each segment with length ℓ are fitted via least square fitting into a function $y_\ell(t)$ and the squared displacement of the segment's points from the fit is taken: $(d_\ell(t))^2 = [x_\ell(t) - y_\ell(t))^2$. The squared displacements are averaged over all the time series, and the square root of this average gives a detrended fluctuation

as a function of the segment length ℓ, given by

$$F(\ell) = \sqrt{\frac{1}{T}\sum_{t=1}^{T}[d_\ell(t)]^2}. \tag{3.39}$$

Finally, the DFA exponent α can be found by fitting $F(\ell) \propto \ell^\alpha$. The α and H exponents must be the same for the same process, but, for real-data time series, α is a much more reliable result. Further, the DFA exponent α and the PS exponent β are related by [130]

$$\beta = 2\alpha - 1 \tag{3.40}$$

and their results complement one another.

3.2.5 Von-Neumann Entropy between Time Series

Entropy is a measure that can be addressed in two main fields: thermodynamics and information theory. The thermodynamic approach is based on the second thermodynamic postulate and its consequences in statistical mechanics [131]. This dissertation will address the information theory approach to entropy, based on the idea that entropy measures the uncertainty of a random variable [132].

The von Neumann Entropy is obtained from the density operator $\boldsymbol{\rho}$ of a mixed state [133] of M pure states $|\psi_j\rangle$ with assigned probabilities p_j of occurence

$$\boldsymbol{\rho} = \sum_{j=1}^{M} p_j \left|\psi_j\right\rangle \left\langle\psi_j\right| . \tag{3.41}$$

The density operator has three properties: (i) Hermitian, (ii) unitary trace, and (iii) is positive semidefinite (i.e., has non-negative eigenvalues).

For a general quantum state $\boldsymbol{\rho}$, John von Neumann defined the entropy by the formula

$$S(\boldsymbol{\rho}) = -\operatorname{tr}(\boldsymbol{\rho}\ln\boldsymbol{\rho}) . \tag{3.42}$$

From decomposition of $\boldsymbol{\rho}$

$$\boldsymbol{\rho} = \sum_{i=1}^{N} \lambda_i \left|i\right\rangle \left\langle i\right| , \tag{3.43}$$

in its eigenstates basis, where λ_i are eigenvalues of $\boldsymbol{\rho}$, the von Neumann Entropy becomes

$$S(\boldsymbol{\rho}) = -\sum_{i=1}^{M} \lambda_i \ln\lambda_i. \tag{3.44}$$

By trying to overcome the threshold for no-correlation cited in Section 3.2.2, [134] proposes a method to calculate the entropy of various time series to testify the entropic brain hypothesis using the von Neumann entropy (see Felippe's [135] Master thesis for more details). Given N time series $\boldsymbol{x}_i = (x_i(1), \ldots, x_i(T))$, Felippe [134] proposes that the Pearson Correlation Matrix (3.29) can be transformed into a matrix with the same properties of the density operator by writing

$$\boldsymbol{\rho}_{\boldsymbol{R}} = \frac{\boldsymbol{R}}{N}, \tag{3.45}$$

where is possible to prove that $\boldsymbol{\rho_R}$ (i) is Hermitian, (ii) has a unitary trace, and (iii) is positive semidefinite (see [134, 135] for details). From $\boldsymbol{\rho_R}$ it is possible to calculate the von Neumann Entropy for the N time series by

$$S(\boldsymbol{\rho_R}) = -\mathrm{tr}\,(\boldsymbol{\rho_R} \ln \boldsymbol{\rho_R}) = -\sum_{j=1}^{N} \lambda_j \ln \lambda_j, \tag{3.46}$$

where λ_j is the $j-$th (from a total of N) eigenvalue of $\boldsymbol{\rho_R}$.

If the time-series are completely non-correlated, $R_{ij} = \delta_{ij}$ and all eigenvalues are $\lambda_i = 1/N$, leading to a maximum entropy value $S_{\max} = \ln N$. Otherwise $S < S_{max}$, indicating some correlation between time series. Results on von Neuman entropy are presented in units relative to the maximum entropy as $\mathbb{S} = S/S_{\max}$.

Chapter 4

Dynamics of m-Vectors

This chapter presents results on characterizing the dynamics of the Collatz Map in the $\boldsymbol{m}$-vectors representation. This is done by extending the $\boldsymbol{m}$-vector representation of a number to a $\boldsymbol{M}$-matrix that contains all the information of the orbit. The $\boldsymbol{M}$-matrix is defined in Section 4.1. The dimension of the $\boldsymbol{m}$-vectors is a piece of important information and is studied in Section 4.2. Further, by studying statistical properties of the $\boldsymbol{M}$-matrix, such as moments and cumulants of its columns, one can characterize the $\boldsymbol{m}$-vectors dynamics, and this is presented in Section 4.3. Future analysis on the $\boldsymbol{m}$-vectors and partial results are presented in Section 4.4

4.1 The M-Matrix

Let $\boldsymbol{m}(t) = (m_1(t), m_2(t), m_3(t), \ldots, m_{r(t)}(t))$ be the $\boldsymbol{m}-$vector representation of the step t of the Accelerated Hailstone Sequence $n_t = \tilde{f}^t(n_0)$. Then one can create a matrix $\boldsymbol{M}(t)$, called $\boldsymbol{M}$-matrix, whose lines are given by the $\boldsymbol{m}$-vectors $\boldsymbol{m}(t)$, and each column i represents the time-series for the component m_i. That is

$$\boldsymbol{M}(t) = \begin{pmatrix} \boldsymbol{m}(1) \\ \boldsymbol{m}(2) \\ \vdots \\ \boldsymbol{m}(t) \end{pmatrix} = \begin{pmatrix} m_1(1) & \cdots & m_{\mathcal{R}}(1) \\ m_1(2) & \cdots & m_{\mathcal{R}}(2) \\ \vdots & \ddots & \vdots \\ m_1(t) & \cdots & m_{\mathcal{R}}(t) \end{pmatrix}, \tag{4.1}$$

where $\mathcal{R}$ is the largest dimension achieved by $\boldsymbol{m}(t)$ during the hailstone sequence. Notice also that if $r(t) < \mathcal{R}$, the matrix components $m_{r'>r(t)}(t)$ are not defined. In this case, we set[1] $m_{r(t)+1}(t) = m_{r(t)+2}(t) = \cdots = m_{\mathcal{R}}(t) = 0$.

The $\boldsymbol{M}(t)$ matrix contains all the information of the hailstone sequence up to time t. Further, if we apply $\boldsymbol{M}(t)$ into a $\mathcal{R}-$dimensional vector $\boldsymbol{u}$ whose components equals unit,

[1] This is only a trick to write a rectangular matrix, these components are not zero, they are undefined, and only m_1 can be zero in this representation.

i.e., $||\boldsymbol{u}|| = \sqrt{\mathcal{R}}$, we find

$$\boldsymbol{M}\boldsymbol{u} = \begin{pmatrix} m_1(1) & \cdots & m_{\mathcal{R}}(1) \\ m_1(2) & \cdots & m_{\mathcal{R}}(2) \\ \vdots & \ddots & \vdots \\ m_1(t) & \cdots & m_{\mathcal{R}}(t) \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{r(1)} m_i(1) \\ \sum_{i=1}^{r(2)} m_i(2) \\ \vdots \\ \sum_{i=1}^{r(t)} m_i(t) \end{pmatrix}, \tag{4.2}$$

and, by numerical comparison with the orbit, one finds that the i-th component of the vector $\boldsymbol{M}\boldsymbol{u}$ goes along with the integer part of the base-2 logarithm of the i-th step of the hailstone sequence. In other words

$$\lfloor \log_2[n_t] \rfloor = \sum_{i=1}^{r(t)} m_i(t). \tag{4.3}$$

This relation can also be obtained for any $n \in \mathbb{Z}^+$ by applying base-2 logarithm to the expansion from Equation (3.9). The details are presented in Appendix D. This allows one to write any hailstone sequence number as

$$\log_2[n_t] = \sum_{i=1}^{r(t)} m_i(t) + \varepsilon(t), \tag{4.4}$$

with $\varepsilon < 1$ given by Equation (D.9).

Equation (4.4) presents an interpretation of the Hailstone Sequence time series at time t as a composition of $r(t) + 1$ time-series: the $r(t)$ integers $m_i(t)$ components plus some noise $\varepsilon(t)$ between 0 and 1. Following this interpretation, one can think of the $\boldsymbol{m}$-vectors as an internal structure of the pseudo-random process. Hence it can be possible to investigate the $m_i(t)$ time series in order to get a better understanding of the hailstone sequence itself. A didactic example of the $\boldsymbol{M}$-matrix applied to the vector $\boldsymbol{u}$ is found in Appendix D.

Figure 4.1 (a-c) presents heat-map representations of the complete $\boldsymbol{M}^{\mathrm{T}}(S)$ from three different initial conditions with random components, respectively:

$$n_0^{(a)} = 2\,842\,073\,768 \equiv \boldsymbol{m}_a(0) = (3,2,2,2,2,2,1,3,2,1,3,1,2,3,2,2)$$

$$n_0^{(b)} = 136\,403\,859\,228\,503\,348 \equiv \boldsymbol{m}_b(0) = (2,2,1,3,2,1,2,2,2,1,1,3,2,1,1,\\ 2,2,3,1,2,3,2,2,1,3,3,3,1,1,1)$$

$$n_0^{(c)} = 62\,511\,978\,391\,312\,874\,001\,289\,897\,106 \equiv \boldsymbol{m}_c(0) = (1,3,3,3,3,1,1,2,3,2,3,3,2,1,3,\\ 3,2,3,2,3,2,3,2,2,3,2,1,1,2,3,\\ 2,3,3,1,2,3,1,1,1,1,1,1,3,3,1).$$

As the initial condition gets larger, the $\boldsymbol{M}$-matrix grows and no pattern seems to appear, pointing that the $\boldsymbol{m}$-vectors time series might be very random.

Figure 4.2 (a) presents the heat-map of $\boldsymbol{M}$-matrix from a very large initial condition whose exact value and correspondent $\boldsymbol{m}$-vector is presented in Appendix E. The noisy characteristic of the $\boldsymbol{M}$-matrix becomes much more apparent, but curiously some interesting patterns can be perceived by zooming into the matrix $M_{t,i}$ components, as Figures 4.2 (b)-(e) shows. Figures 4.2 (b-e) presents four levels of zoom. It is possible to perceive the existence of horizontal stripes where only m_1 changes and the other components do not. These structures appears when the n_t is even, then $m_1(t+1) = m_1(t) - 1$ and

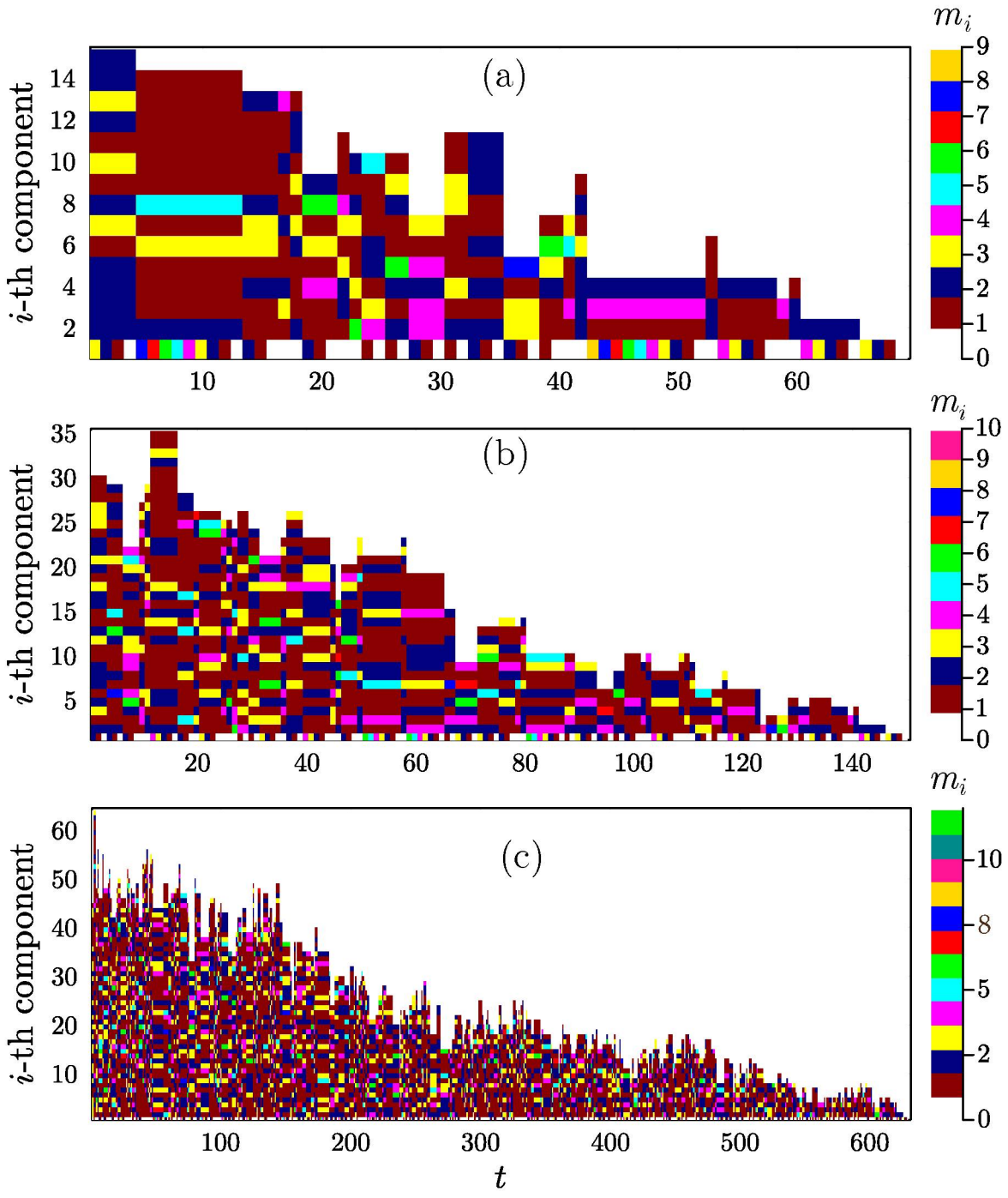


Figure 4.1: Heat-map representations from randomly chosen components of initial $\boldsymbol{m}$-vector. (a) $n_0^{(a)} = 2\,842\,073\,768$, (b) $n_0^{(b)} = 136\,403\,859\,228\,503\,348$, and (c) $n_0^{(c)} = 62\,511\,978\,391\,312\,874\,001\,289\,897\,106$

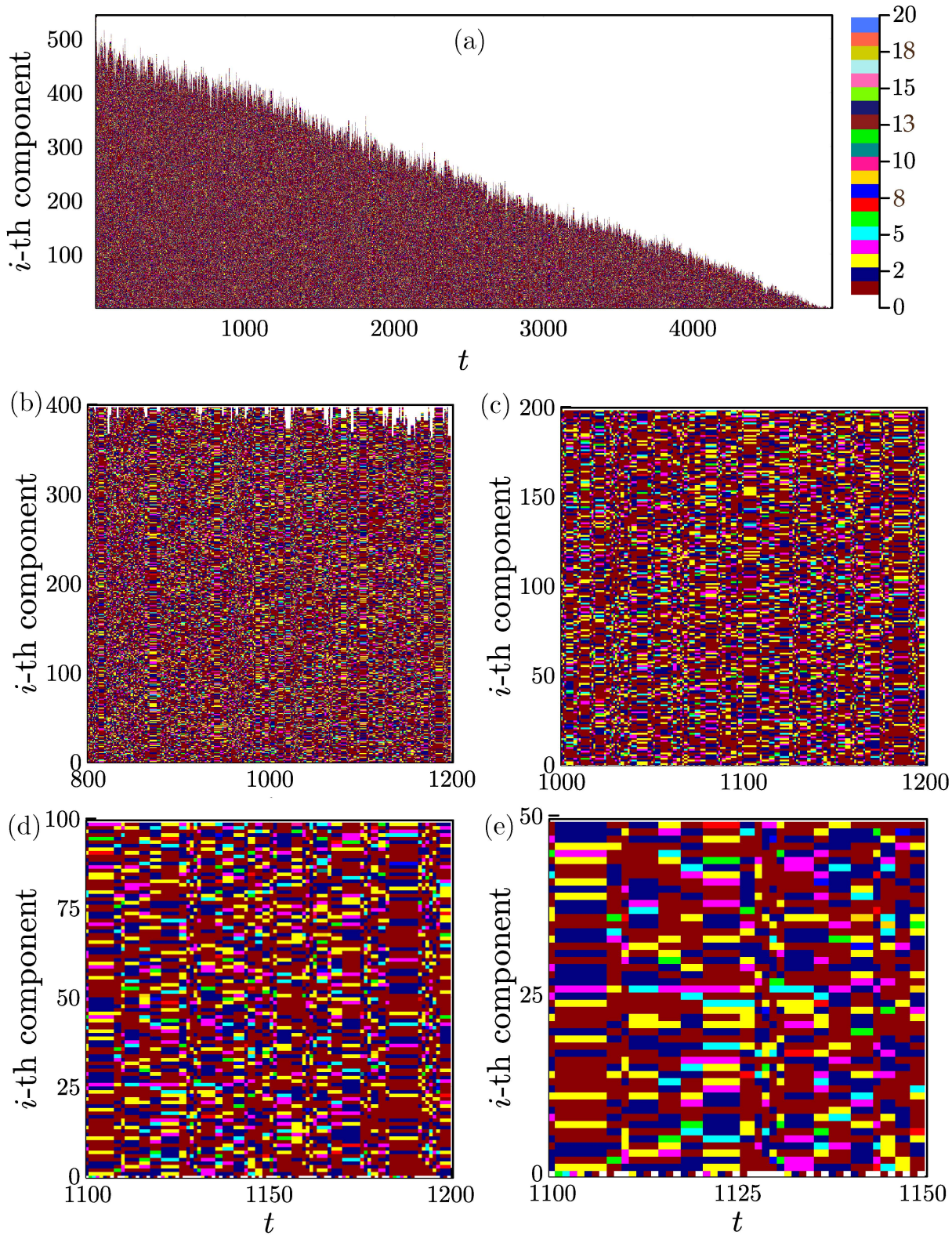


Figure 4.2: (a) Heat-map of the $\boldsymbol{M}$-matrix from initial condition in the order of 2^{968} (see Appendix E for exact value). (b-e) Four levels of zoom into the $\boldsymbol{M}$-matrix: (b) $t \in [800, 1200], i \in [1, 400]$, (c) $t \in [1000, 1200], i \in [1, 200]$, (d) $t \in [1100, 1200], i \in [1, 100]$, (e) $t \in [1100, 1150], i \in [1, 50]$.

$m_{i>1}(t+1) = m_{i>1}(t)$, leading to these horizontal stripes. These stripes become as long as the number of times n_t is divisible by 2. Also, zooming in, besides the presence of the stripes, does not vanish completely the random structure of the $\boldsymbol{M}$-matrix, pointing to the possible existence of scale invariance. The random-structured patterns must be carefully studied, and this will be presented in Chapter 5 by applying von-Neumann Entropy measurements on the $\boldsymbol{m}$-vectors components.

Finally, it is possible to see the dominance of $1, 2$, and 3 — dark red, dark blue, and yellow, respectively — values in the $\boldsymbol{M}$-matrix, leading to questions about what is the distribution of numbers in the $\boldsymbol{M}$-matrix. In the next section, some statistical explorations will characterize formally the distribution of numbers in the $\boldsymbol{m}$-vectors components, as well as how this distribution is affected by initial conditions.

4.2 Dimension of $\boldsymbol{m}$-vectors

By analyzing Figures 4.1 and 4.2 (a) it is clear that the dimension $r(t)$ of the $\boldsymbol{m}$-vector decreases with time. This is a direct consequence of the decrease of the sequence itself, leading to smaller numbers, with smaller $\boldsymbol{m}$-vector dimensions. Figure 4.3 shows the dimension of the $\boldsymbol{m}$-vector $r(n)$ as a function of the base-10 number it represents for $n \in [2^1, 2^{10}]$. The inset of Figure 4.3 details $r(n)$ for $n \in [2^8, 2^{10}]$, revealing a repeated structure from 2^8 up to $2^9 + 2^8$.

A careful analysis shows that the function $r(n)$, in the interval $2^{a-1} \leq n < 2^a$, repeats itself in the interval $2^a \leq n < 2^a + 2^{a-1}$ for every integer $a > 1$. One can see this by comparing $r(n)$ with $r(n + 2^{a-1})$ for $n \in [2^a - 2^{a-1}, 2^a - 1]$: Let

$$n = 2^{m_1} + \cdots + 2^{m_1 + \cdots + m_r},$$

with $r(n) = r_n$, and $m_1 + \cdots + m_r = \sum_i^{r(n)} m_i = a - 1$, then for

$$\begin{aligned} n + 2^{a-1} &= 2^{m_1} + \cdots + 2^{m_1 + \cdots + m_r} + 2^{a-1}, && (4.5) \\ &= 2^{m_1} + \cdots + 2^{a-1} + 2^{a-1}, && (4.6) \\ &= 2^{m_1} + \cdots + 2^a, \Rightarrow r(n + 2^{a-1}) = r_n = r(n). && (4.7) \end{aligned}$$

Further, one can look for the probability distribution function $P(r, \boldsymbol{\Sigma}_r)$ of $\boldsymbol{m}$-vectors with dimension r and sum of components $\boldsymbol{\Sigma}_r(\boldsymbol{m}) = \sum_{i=1}^{r} m_i$. Table 4.1 shows how the base-10 numbers are classified according to r and $\boldsymbol{\Sigma}_r(\boldsymbol{m})$. From Table 4.1 one can conjecture that $D(r, \boldsymbol{\Sigma}_r)$ follows a binomial probability distribution for $p = 1/2$, that is

$$P(r, \boldsymbol{\Sigma}_r) = \frac{1}{2^{\boldsymbol{\Sigma}_r}} \binom{\boldsymbol{\Sigma}_r}{r}. \tag{4.8}$$

Figure 4.4 shows the distribution $D(r, \boldsymbol{\Sigma}_r)$, i.e., the number of $\boldsymbol{m}$-vectors with r components for each $\boldsymbol{\Sigma}_r$, corroborating with the hypothesis that $P(r, \boldsymbol{\Sigma}_r)$ is binomial.

The existence of a maximum on the distributions $D(r, \boldsymbol{\Sigma}_r)$ might lead one to wonder what is the average decaying of r for the Collatz process. By considering one random initial condition from 2^k to 2^{k+1} for each $k = 100, 200, \ldots, 2000$, Figure 4.5 presents how the dimension r decays with t. The decay presents a trend with some noise around. In order to find the general trend for $r(t)$ one can fit the time series $\boldsymbol{r}(t) = (r(0), r(1), \ldots, r(T))$ into a linear function $f(t) = at + b$.

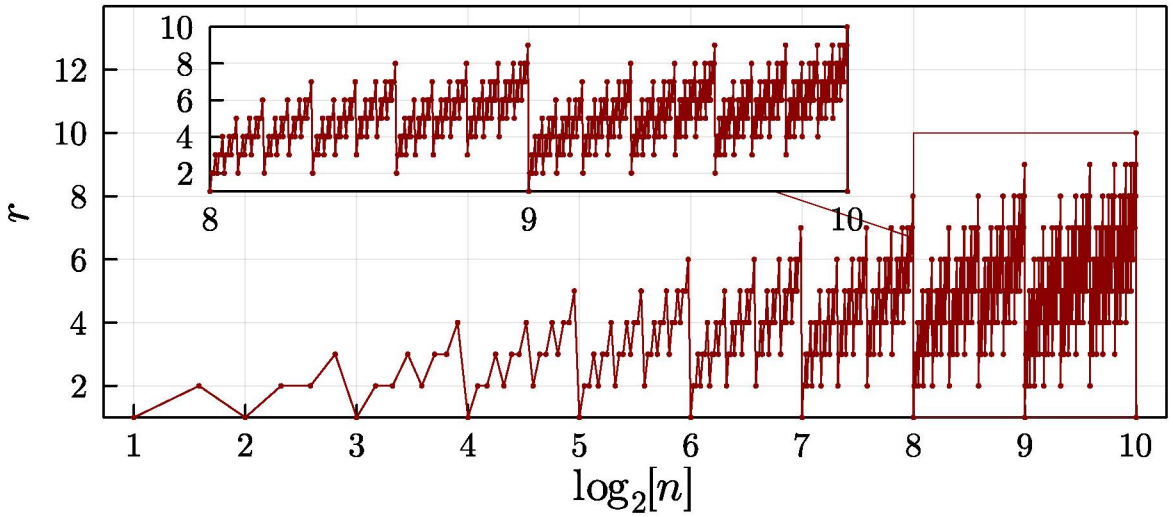


Figure 4.3: Dimension of the $\boldsymbol{m}$-vector, denoted by r, plotted in function of the base-2 logarithm of natural numbers n for $2 \leq n \leq 2^{10}$. The inset shows a zoom into the region of $2^8 \leq n \leq 2^{10}$ to detail the periodic structure between 2^8 and $2^9 + 2^8$.

	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$
$\boldsymbol{\Sigma}_r = 1$	2	3				
$\boldsymbol{\Sigma}_r = 2$	4	5, 6	7			
$\boldsymbol{\Sigma}_r = 3$	8	9, 10, 12	11, 13, 14	15		
$\boldsymbol{\Sigma}_r = 4$	16	17, 18, 20, 24	19, 21, 22, 25, 26, 28	23, 27, 29, 30	31	
$\boldsymbol{\Sigma}_r = 5$	32	33, 34, 36, 40, 48	35, 37, 38, 41, 42, 44, 49, 50, 52, 56	39, 42, 45, 46, 51, 53, 54, 57, 58, 60	47, 55, 59, 61, 62	63

Table 4.1: Classification of base-10 numbers according to r and $\boldsymbol{\Sigma}_r$. Numbers in the same line have the same $\lfloor \log_2 n \rfloor$ value, while numbers in the same column have the same $\boldsymbol{m}$-vector dimension. Numbers in the same cell share both characteristics.

For the present analysis, 100 hailstone time series of $\boldsymbol{r}(t)$ were considered. The time series are started from random initial conditions in the interval $2^{2000} < n_0 < 2^{2001}$. No record of the specific initial condition was made, but these results were exhaustively reproduced in order to be sure that they are as general as possible[2]. The sample of 100 hailstone sequences were analysed, generating the time series $\boldsymbol{r}_i(t) = (r_i(0), r_i(1) \ldots r_i(T))$, for $i = 1, \ldots, 100$. Each $\boldsymbol{r}_i(t)$ time series was fitted via the least square method into $f_i(t) = a_i t + b_i$.

The linear coefficient b_i is compared with each $r_i(0)$, and the relative mean deviation

[2]It is hard to define how general the results are. They are obtained from very large initial conditions ($n_0 \approx 2^{2000}$) to obtain good statistics, but the amount of ignored sample space grows exponentially. The use of random initial conditions aims to draw the most unbiased hailstone sequences possible.

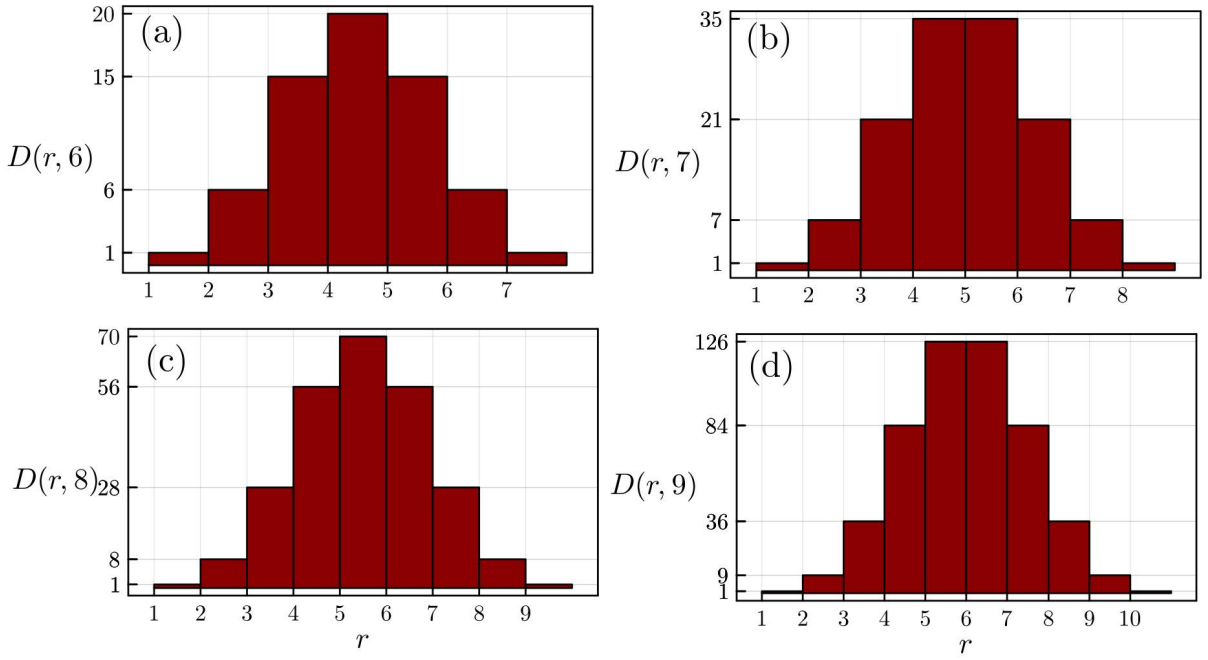


Figure 4.4: Distribution of $\boldsymbol{m}$-vectors dimension r in the ranges of natural numbers (a) $2^6 \leq n < 2^7$, (b) $2^7 \leq n < 2^8$, (c) $2^8 \leq n < 2^9$ and (d) $2^9 \leq n < 2^{10}$, evidencing the binomial distribution for $D(r, \boldsymbol{\Sigma}_r)$.

of b from $r(0)$ is estimated by

$$\langle b - r(0) \rangle = \frac{1}{100} \sum_{i=1}^{100} \frac{|b_i - r_i(0)|}{r_i(0)}, \tag{4.9}$$

obtaining

$$\langle b - r(0) \rangle = 0.02095, \tag{4.10}$$

indicating that $b \approx r(0)$ is a good approximation. Further, the mean of a_i is estimated as

$$\langle a \rangle = -0.103442. \tag{4.11}$$

Finally, a model for the mean decay of $r(t)$ can be formulated as

$$\langle r(t) \rangle \sim r(0) - t/10. \tag{4.12}$$

4.3 Statistical Investigation

This section deals with statistical measures of columns of the $\boldsymbol{M}$-matrix. The analysis of columns is done by estimating the first four cumulants — mean, variance, skewness and kurtosis — of each value in column of the $\boldsymbol{M}$-matrix, i.e., the time series for each m_i component. This study allows us to characterize the distribution of m-values along the $\boldsymbol{m}$-vectors and the $\boldsymbol{M}$-matrix, hence providing a characterization of the internal structure of the hailstone sequences.

Let $\{\boldsymbol{M}\} = \{\boldsymbol{M}^{(1)}, \boldsymbol{M}^{(2)}, \ldots, \boldsymbol{M}^{(N)}\}$ be a set of N $\boldsymbol{M}$-matrices generated from initial $\boldsymbol{m}$-vectors with fixed $r(0)$ random components. Let $M_{t,i}^{(j)}$ denote the t-th row i-th column

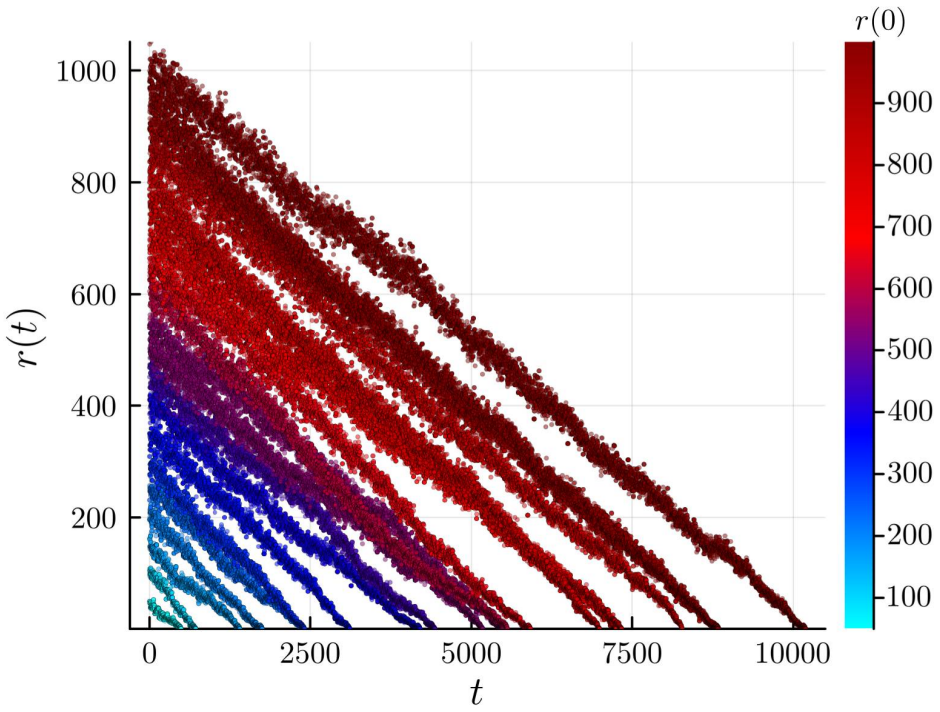


Figure 4.5: Dynamics of the dimension r of $\boldsymbol{m}$-vectors along the hailstone sequences from 20 initial conditions randomly chosen between $2^k < n_0 < 2^{k+1}$ for $k = 100, 200, \ldots, 2000$.

of the j-th $\boldsymbol{M}$-matrix. Once only m_1 can assume zero value, the n-th moment for the 1st column is simply

$$\mu_n^{(j,1)} = \frac{1}{S}\sum_{t=0}^{S}\left(M_{t,1}^{(j)}\right)^n, \tag{4.13}$$

where S is the total stopping time of the hailstone sequence. Then the n-th moment of the column $i > 1$ of the $\boldsymbol{M}$-matrix $\boldsymbol{M}^{(j)}$ can be estimated by

$$\mu_n^{(j,i>1)} = \frac{1}{S_i}\sum_{t=\{t_{M_{t,i}\neq 0}\}}\left(M_{t,i>1}^{(j)}\right)^n, \tag{4.14}$$

where S_i is the number of non-vanishing components $t_{M_{t,i}\neq 0}$ at the column $M_{t,i>1}$ Then the mean of the n-th moment of the j-th column over each $\boldsymbol{M}^{(j)}$ from the set $\{\boldsymbol{M}\}$, is estimated by

$$\langle\mu_n^{(i)}\rangle = \langle\mu_n^{(j,i)}\rangle = \frac{1}{N}\sum_{j=1}^{N}\mu_n^{(j,i)}. \tag{4.15}$$

From these means, one can obtain the mean cumulants of the j-th column by

$$\langle c_1^{(i)}\rangle = \langle\mu_1^{(i)}\rangle, \tag{4.16}$$
$$\langle c_2^{(i)}\rangle = \langle\mu_2^{(i)}\rangle - \langle\mu_1^{(i)}\rangle^2 = \langle\sigma^{(i)}\rangle^2, \tag{4.17}$$
$$\langle c_3^{(i)}\rangle = \langle\mu_3^{(i)}\rangle - 3\langle\mu_1^{(i)}\rangle\langle\mu_2^{(i)}\rangle + 2\langle\mu_1^{(i)}\rangle^3, \tag{4.18}$$
$$\langle c_4^{(i)}\rangle = \langle\mu_4^{(i)}\rangle - 3\langle\mu_2^{(i)}\rangle^2 - 4\langle\mu_1^{(i)}\rangle\langle\mu_3^{(i)}\rangle + 12\langle\mu_1^{(i)}\rangle^2\langle\mu_2^{(i)}\rangle - 6\langle\mu_1^{(i)}\rangle^4. \tag{4.19}$$

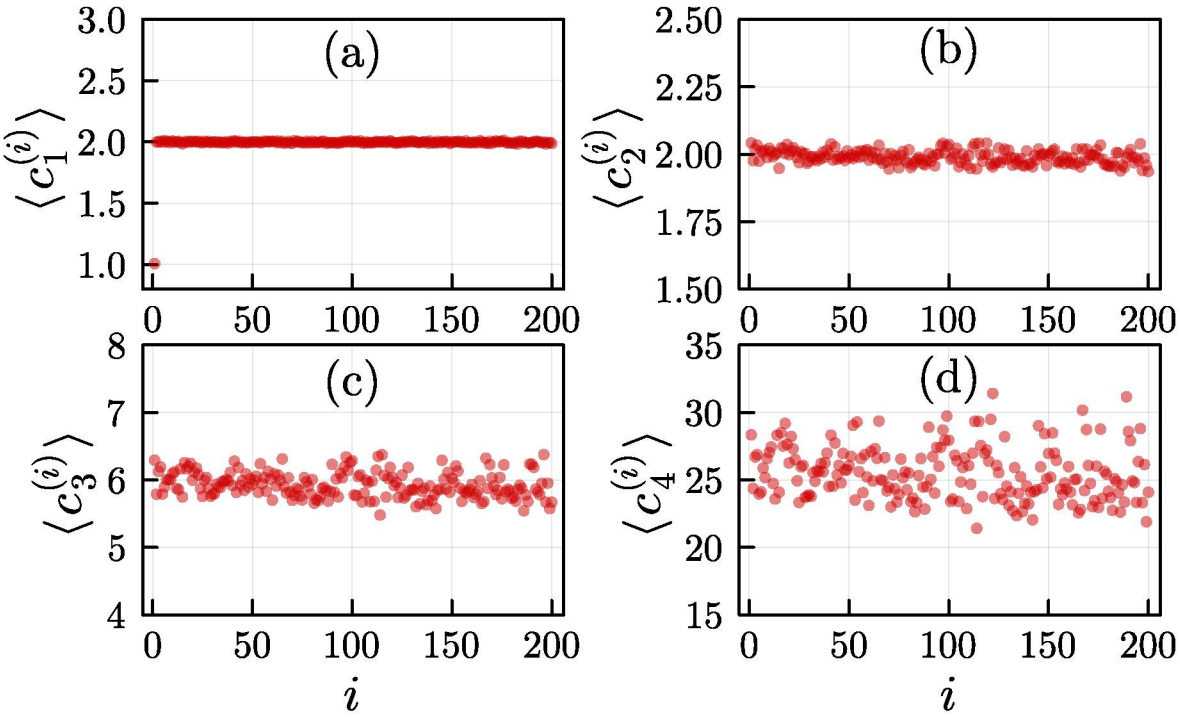


Figure 4.6: First four cumulants for each of 200 first columns of 100 $\boldsymbol{M}$-Matrices, revealing the (a) mean, (b) variance, (c) skewness and (d) kurtosis of the distribution of m-values in the components of $\boldsymbol{m}$-vectors along the hailstone sequences.

Figure 4.6 displays how each cumulant evolves with the column it represents.

By Figure 4.6(a) one can see that the mean $\langle c_1^{(1)} \rangle \approx 1$ and $\langle c_1^{(i>1)} \rangle \approx 2$ for all $i > 1$. The variance — see Figure 4.6(b) — $\langle c_2^{(i)} \rangle \approx 2$ for all i, revealing that almost the same distribution holds for the $\boldsymbol{m}$-vectors components values. The skewness and kurtosis — see Figures 4.6 (c) (d) — are larger than expected for a normal distribution, meaning that there is excess skewness and kurtosis. Positive skewness $\langle c_3^{(i)} \rangle$ means that the distribution is peaked to the left, which is expected given $\boldsymbol{m}$-vectors contain only positive integers with mean $\langle m_{i>1} \rangle = 2$ ($\langle m_1 \rangle = 1$). The excess kurtosis $\langle c_3^{(i)} \rangle$ leads to the conclusion that the distribution is very peaked around the mean, with abrupt decay.

These results raise two questions: (1) what is the distribution function $D(m)$ for the $\boldsymbol{m}$-vectors components values and (2) if the distribution functions for the components are similar to each other, given that the first four moments present only small deviations.

The distribution function can be estimated by counting, for each component, how many of each value appears along the hailstone sequence. From the first cumulant analysis, one finds one distribution for m_1 and another for $m_{i>1}$, but it is expected that they have a similar shape for all i. Figure 4.7 shows the distributions for m_1 (Figure 4.7 (a)) and $m_{i>1}$ (Figure 4.7) over the sample of 100 $\boldsymbol{M}$-Matrices, evidencing that they are indeed similar to each other. In fact, they can be compared with exponential distributions, leading to

$$D(m) = \begin{cases} 2^{-(m+1)}, & \text{if } i = 1, \\ 2^{-m}, & \text{if } i > 1, \end{cases} \tag{4.20}$$

the distribution function for the $\boldsymbol{m}$-vectors components.

From figure 4.6, it is possible to see a great homogeneity in the mean and variance of

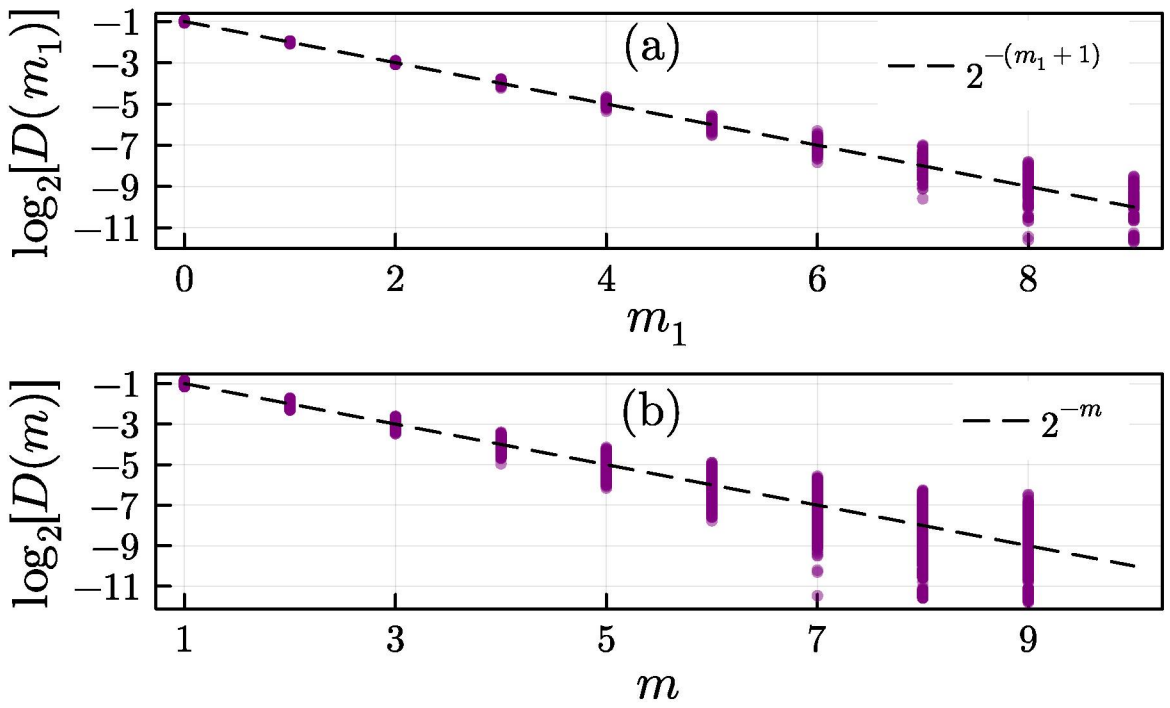


Figure 4.7: Distribution of values of (a) m_1 and (b) $m_{i>1}$ components scattered in the purple dots and comparison with exponential distribution (a) $D(m_1) = 2^{-(m_1+1)}$ and (b) $D(m) = 2^{-m}$.

the components. This is obtained starting from randomly chosen initial components, then, the Collatz process seem to produce certain thermalization in its internal structure, where the values of its components tend to an equilibrium in the long-term (but finite $t < S$ where S is the total stopping time).

4.4 Further points to be analyzed

This section will provide some incomplete but promising results on using the $\boldsymbol{m}$-vectors formalism. Section 4.4.1 presents an investigation on modeling a *Collatz Gas* based on Riemann Gases and ensemble statistical physics considerations. Some explorations on how to evolve properly the map on the $\boldsymbol{m}$-vectors space are provided in section 4.4.2. Finally, the still unsuccessful tentative of directly evolving the $\boldsymbol{m}$-vectors, leads to representing the evolution of $\boldsymbol{m}$-vectors components via graphs and transition matrices in Section 4.4.3.

4.4.1 The Collatz Gas

The interface between number theory and physics has been growing over the last decades as the Introduction has pointed out. One of the origins of this interface traces back to Hilbert-Pòlya conjecture [52], providing a physical reason why the Riemann hypothesis [54] must be true. One of the most curious connections between the Riemann hypothesis and physics is placed within statistical physics, by the construction of Riemann Gases

[136, 137]. The construction starts by defining an ensemble of quantum states of a gas

$$|1\rangle \oplus \{|2\rangle, |3\rangle, |5\rangle, |7\rangle, \ldots, |p_n\rangle, \ldots\} \tag{4.21}$$

where $|1\rangle$ is the vacuum state, and the set of $|p_n\rangle$ are states formed only by prime numbers. This representation is linked to a definition of energy for each state $|p\rangle$ as

$$E_p \equiv \ln p, \tag{4.22}$$

such that the ensemble of energies is given by

$$E_p \in \{0, \ln 2, \ln 3, \ln 5, \ldots, \ln p_n, \ldots\}. \tag{4.23}$$

If we second quantize this system, one can define the existence of creation operators $\hat{\alpha}_p$ and then, non-prime states can be generated from vacuum

$$\hat{\alpha}_{p_i}\hat{\alpha}_{p_j}\cdots\hat{\alpha}_{p_l}|1\rangle = |p_i p_j \cdots p_l\rangle. \tag{4.24}$$

This implies that the created states, and the primes themselves, are bosonic-like, where the creation operators commute $\hat{\alpha}_{p_i}\hat{\alpha}_{p_j} = \hat{\alpha}_{p_j}\hat{\alpha}_{p_i}$. Another possibility would be to create fermionic states, where $\hat{\alpha}_{p_i}\hat{\alpha}_{p_j} = -\hat{\alpha}_{p_j}\hat{\alpha}_{p_i}$.

From this construction, every natural number is associated with an energy state. A partition function Z_R of a gas whose particles can be in states with energy $\ln n$ can be constructed, in the Canonical Ensemble, where the system is in contact with a thermal reservoir with temperature $1/\beta$

$$Z_R = \sum_{i=1}^{\infty} e^{-\beta E_n}, \tag{4.25}$$

but once $E_n = \ln n$, then

$$Z_R = \sum_{i=1}^{\infty} e^{-\beta \ln n} = \sum_{i=1}^{\infty} n^{-\beta} = \zeta(\beta). \tag{4.26}$$

In other words, the partition function Z_R of the system is the Riemann-Zeta-Function! From this partition function, average energy and entropy can be calculated for this gas.

Now, let us consider an ensemble of N non-interacting particles forming another gas, whose Hamiltonian is given by

$$\mathcal{H} \equiv \varepsilon_0 \lfloor \log_2 n \rfloor, \; n = n_0, \tilde{f}(n_0), \tilde{f}^2(n_0), \ldots, 4, 2, 1, \tag{4.27}$$

for $\tilde{f}(n)$ the Accelerated Collatz Map 2.2. Then, given n_0, associated to an energy $E_0 = \varepsilon_0 \lfloor \log_2 n_0 \rfloor$, each particle can assume the following S energy values:

$$\left\{\varepsilon_0 \lfloor \log_2 n_0 \rfloor, \varepsilon_0 \lfloor \log_2 \tilde{f}(n_0) \rfloor, \ldots, 2\varepsilon_0, \varepsilon_0, 0\right\}. \tag{4.28}$$

Hence the partition function for one particle of this system, using base two instead of e, is

$$Z_1 = \sum_{t=0}^{T} 2^{-\beta\varepsilon_0 \lfloor \log_2 n_t \rfloor}. \tag{4.29}$$

By using the relation 4.3, $\lfloor \log_2 n_t \rfloor = \sum_{i=1}^{r(t)} m_i(t)$, the partition function becomes

$$Z_1 = \sum_{t=0}^{S} 2^{-\beta\varepsilon_0 \sum_{i=1}^{r(t)} m_i(t)} = \sum_{t=0}^{S} \prod_{i=1}^{r(t)} 2^{-\beta\varepsilon_0 m_i(t)}, \tag{4.30}$$

or, considering a degenerecy g_m for each m value

$$Z_1 = \sum_{t=0}^{S} \prod_{m=1}^{\max(m(t))} g_m 2^{-\beta\varepsilon_0 m}, \tag{4.31}$$

and once $g_0 = 1$ because, at every time-step, only m_1 can be zero, the partition function can be written as

$$Z_1 = \sum_{t=0}^{T} \prod_{m=1}^{\max(m(t))} g_m 2^{-\beta\varepsilon_0 m}. \tag{4.32}$$

The value of $\max(m(t))$ can be obtained from the previously presented statistical analysis and a still ongoing investigation might lead to a partition function with fruitful energy and entropy measurements for this gas. These results are not concluded yet, hence they will not be presented here.

4.4.2 On how to evolve m-vectors via Collatz Map

The analysis developed here only details more on the discussion presented by [92]. In order evolve the $\boldsymbol{m}$-vectors using the Accelerated Collatz Map $\tilde{f}$, it is useful to rewrite it as

$$n_{t+1} = \tilde{f}(n_t) = \begin{cases} 2^{-1} n_t & \text{if } m_1 > 0; \\ n_t + 2^{-1}(2^0 + n_t) & \text{if } m_1 = 0. \end{cases} \tag{4.33}$$

When $m_1(t) > 0$, the $\boldsymbol{m}$-vector is evolved simply as [3]

$$\boldsymbol{m}(t+1) \;=\; \left(m_1(t) - 1, m_2(t), m_3(t), \ldots, m_{r(t)}(t)\right). \tag{4.34}$$

Now, for $m_1(t) = 0$, in base-10 one finds

$$\begin{aligned} n_t + 2^{-1}(2^0 + n_t) \;=\; & 2^0 + 2^{m_2(t)} + 2^{m_2(t)+m_3(t)} + \cdots + 2^{m_2(t)+m_3(t)+\cdots m_r(t)} + \\ & 2^{-1} + 2^{-1}\left(2^0 + 2^{m_2(t)} + 2^{m_2(t)+m_3(t)} + \cdots + 2^{m_2(t)+m_3(t)+\cdots+m_r(t)}\right), \end{aligned}$$

by reorganizing as

$$\begin{aligned} n_t + 2^{-1}(2^0 + n_t) \;=\; & 2^0 + 2^{m_2(t)} + 2^{m_2(t)+m_3(t)} + \cdots + 2^{m_2(t)+m_3(t)+\cdots+m_r(t)} + \\ & +2^0 + 2^{m_2(t)-1} + 2^{m_2(t)-1+m_3(t)} + \cdots + 2^{m_2(t)-1+m_3(t)+\cdots+m_r(t)}, \end{aligned}$$

[3] Once for $m_1(t) > 0$ one finds

$$n_{t+1} = 2^{-1} n_t = 2^{m_1(t)-1}\left(2^{m_2(t)}\left(\cdots\left(2^{m_{r(t)}(t)} + \cdots + 1\right) + 1\right) + 1\right)$$

.

the $\boldsymbol{m}$-vector becomes, for $m_1(t) = 0$,

$$\boldsymbol{m}(t+1) = (0, m_2(t), m_3(t), \cdots, m_r(t)) \oplus (0, m_2(t) - 1, m_3(t), \cdots, m_r(t)), \tag{4.35}$$

where $\oplus$ denotes the proper form to sum two $\boldsymbol{m}$-vectors with the same dimension. All the discussion so far leads to questions on how to sum two $\boldsymbol{m}$-vectors. Answering this question is still an open problem and no conclusive results were obtained now. But this is a problem that, if solved, given the important description of the internal structure of the hailstone sequences, could lead to new insights into the dynamics of the map, or maybe the conjecture itself.

The problem of defining a proper way for summing the $\boldsymbol{m}$-vectors are attached to the larger effort of defining an algebraic and topological structure for the $\boldsymbol{m}$-vectors. In order to do that, one can lean on the mathematical foundations presented in section 3.1.1 about topological algebra concepts on ultrametric spaces and p-adics, respectively. In addition to defining how to sum two $\boldsymbol{m}$-vectors, one can try to define a field $\mathbb{M}$ for the $\boldsymbol{m}$-vectors, by obtaining the proper multiplication and testing for commutative and distributive laws. Further, one can try to find an absolute value, similar to the p-adic, in order to study the topology of the $\boldsymbol{m}$-vector's field. These are all greater efforts, that go beyond the scope of the present dissertation, but the author hopes that the foundations of Chapter 3, together with these future analysis ideas, might light up the will of performing such tasks.

4.4.3 Graph of m-vectors

Another way to investigate the $\boldsymbol{m}$-vectors' internal structure of the hailstone sequences is via transition matrices and graphs. Given that no deterministic rule has been developed yet to determine the evolution of the components when n_t is odd, and that it is only when n_t is odd that the $\boldsymbol{m}$-vectors components change substantially, a probabilistic view of the problem can be addressed.

First, assuming that the components of the $\boldsymbol{m}$-vectors evolve in a *markovian* process[4] then one can define the probability

$$P_{t+1}(m_i(t+1)|m_i(t)), \tag{4.36}$$

that the i-th component assume the value $m_i(t+1)$ at time $t+1$, given that it has assumed the value $m_i(t)$ at time t.

Second, consider that the transition probability between every pair of possible m_i values does not vary with time. Then, one can write, for any pair $(k^{(i)},\ell^{(i)})$ of possible m_i values

$$P_t(k^{(i)},\ell^{(i)}) = T^{(i)}_{k\ell}, \tag{4.37}$$

where $T^{(i)}_{k\ell}$ is a matrix element of the so-called *transition matrix*[5] $T^{(i)}$, indicating the probability of transition from k to ℓ in the i-th component of the $\boldsymbol{m}$-vector. The transition matrix $T^{(i)}$ can be used as an adjacency matrix to create a directed graph $G^{(i)}$, whose vertexes are all the possible values assumed by the i-th component and an edge is

[4] Characteristic of discrete processes where the state in the next step depends only on the current time state [138]. This assumption for the $\boldsymbol{m}$-vectors components can be enhanced by calculating the autocorrelation function of the components, this is also a possible future analysis.

[5] In this notation, the matrix elements are $T_{k\ell}$ for $k,\ell = 0, 1, 2, \ldots$, to be concise with the possibility of finding zero values along the $\boldsymbol{m}$-vectors components.

constructed from vertex k towards ℓ if the element $T^{(i)}_{k,\ell} \neq 0$ (equivalently an edge is created from ℓ towards k if $T^{(i)}_{\ell,k} \neq 0$).

As a simple example, consider $n_0 = 13$, whose $\boldsymbol{M}$-Matrix is given by

$$\boldsymbol{M}(S) = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 2 & 0 \\ 1 & 2 & 0 \\ 0 & 2 & 0 \\ 3 & 0 & 0 \\ 2 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.38}$$

The first column of $\boldsymbol{M}(S)$ represents the m_1 time series. It is easy to check that there are four possible states $\{0, 1, 2, 3\}$. For state 0, there are two possible transition $0 \mapsto 2$ and $0 \mapsto 3$ that occurs in equal quantities, hence, one can associate equal transition probabilities for each $T^{(1)}_{0,2} = T^{(1)}_{0,3} = 1/2$. For states $k = 1, 2, 3$, there is only one transition $k \mapsto k-1$ and, for each of them, one can associate equal transition probabilities $T_{k,k-1} = 1$. From this, the transition matrix for m_1 can be written as

$$T^{(1)} = \begin{pmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{4.39}$$

The second column of $\boldsymbol{M}(S)$ represents the m_2 time series. In this case, there are only two possible states $\{0, 2\}$. State 0, in this case, represents non-existing components and contains only one possible transition $0 \mapsto 0$, then $T^{(2)}_{00} = 1$. State 2 there are two possible transitions $2 \mapsto 2$ and $2 \mapsto 0$, where the first one occurs in 3 of the 4 transitions and the latter only once, leading to $T^{(2)}_{22} = 3/4$ and $T^{(2)}_{20} = 1/4$. From this, the transition matrix for m_2 can be written as

$$T^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} \end{pmatrix}, \tag{4.40}$$

where the second row and line are null once they represent the transitions for state $\{1\}$, which does not exist for this component.

For the third column of $\boldsymbol{M}(S)$, representing the m_3 time series, there are only $\{0, 1\}$ possible states. For state 0, there is only the transition $0 \mapsto 0$, and for 1 there is only $1 \mapsto 0$ transition, leading to the transition matrix

$$T^{(3)} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}. \tag{4.41}$$

The graphs $G^{(i)}$ for m_1, m_2, and m_3 are very simple in this case and will not be presented here. In order to present a more meaningful example, from larger time series, Appendix F shows the Transition Matrices for the $\boldsymbol{m}$-vectors components starting from $n_0 = 27$, and Figure 4.8 presents each of $G^{(i)}$ graphs of the eight $\boldsymbol{m}$-vectors' components, showing the large variability of connections that can be characterized in future studies. The important statement of this section is the potential analysis that the transition matrices

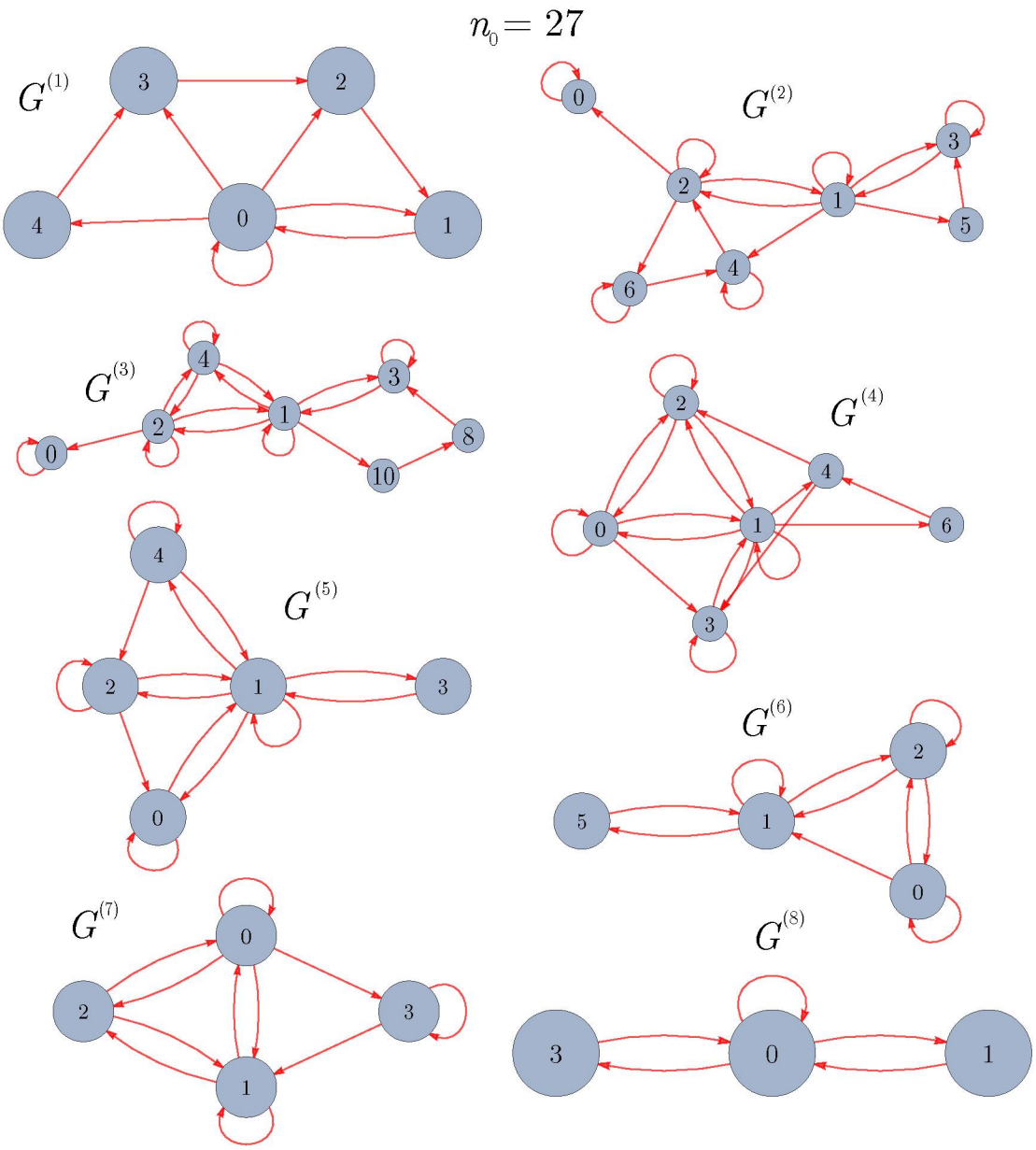


Figure 4.8: $\boldsymbol{m}$-graphs $G^{(j)}$ for $j = 1, 2, \ldots, 8$ representing the transitions of values in the j-th component of $\boldsymbol{m}$-vectors evolved by the Collatz Map from $n_0 = 27$. See Appendix F for the complete transition matrices.

and graphs can generate, enabling a better characterization of the internal structure of the hailstone sequences using network and graph analysis.

4.4.4 Chaos in Hailstone Sequences

Another study that seems very promising is the study of chaotic properties in the transient orbit of the Collatz Map. Chapter 2 has reviewed briefly some studies of the map in Dynamical Systems theory and explorations on defining measures such as Lyapunov exponents for the dynamics. One of the main difficulties of defining a measure such as the Lyapunov exponent for the dynamics are: (i) the space in which the hailstone sequences evolve is discrete, such that the initial separation of orbits can't be infinitesimal to measure

its evolution, and (ii) the orbits are not bounded. The $\boldsymbol{m}$-vector space might prove itself useful in this definition, once it changes the one dimensional space of natural numbers to a space that has connections with non-archimedean spaces with higher dimension. In these higher dimension spaces, if future efforts from section 4.4.2 are fruitful, orbits can be started from $\boldsymbol{m}$-vectors as close to each other as possible, in their space, and maybe some bound can be defined for the orbit. In the paragraphs bellow, two possible studies in measuring the Lyapunov exponent for Collatz Map will be sketched.

Consider a finite subset of the positive integers $\mathbb{B} \in \mathbb{Z}^+$ such that all hailstone sequences with $n_0 \in \mathbb{B}$ are bounded with $\mathcal{S}_{\max}$ the largest expansion factor over all orbits initiated within $\mathbb{B}$. Then consider the set $\mathbb{B}_{\mathcal{S}_{\max}}$ a extension of $\mathbb{B}$ to be the bounding space of all sequneces starting in $\mathbb{B}$.

Now let $\max\left(\mathbb{B}_{\mathcal{S}_{\max}}\right)$ to be the largest element of $\mathbb{B}_{\mathcal{S}_{\max}}$, and consider the transformation

$$n_t \mapsto \frac{n_t}{\max\left(\mathbb{B}_{\mathcal{S}_{\max}}\right)} = x_t \in \left[1/\max\left(\mathbb{B}_{\mathcal{S}_{\max}}\right), 1\right]$$

for $n_t \in \mathbb{B}_{\mathcal{S}_{\max}}$. Then, with this transformation, the hailstone sequence $\mathcal{O}$ transforms as

$$\mathcal{O}(n_0) = \{n_0, n_1, \ldots, n_S = 1\} \mapsto \mathcal{X}(x_0) = \{x_0, x_1, \ldots, x_S = 1/\max\left(\mathbb{B}_{\mathcal{S}_{\max}}\right)\},$$

where the set $\mathcal{X}(x_0)$ is bounded and, if the conjecture is true, x_S can be as small as needed for computing a Finite Size Lyapunov Exponent [139] for the $\mathcal{X}$ orbit, that is isomorphic to the hailstone sequence.

As a second approach, utilizing the $\boldsymbol{m}$-vector space and a – still to be found – proper metric definition $d(\boldsymbol{m}_a, \boldsymbol{m}_b)$, one can calculate two sets of $\boldsymbol{m}$-vectors evolved by the Collatz Map, starting from two $\boldsymbol{m}$-vectors $\boldsymbol{m}(0)$ and $\boldsymbol{m}(0) + \delta(\boldsymbol{m}(0))$ where $\delta(\boldsymbol{m}(0)) \to \min\left(d(\boldsymbol{m}_a, \boldsymbol{m}_b)\right)$ is the minimum distance in the $\boldsymbol{m}$-vector space for the given metric. Then, by applying the methods from Finite Size Lyapunov Exponent, a measurement of the average distancing of the near initiated orbits.

Chapter 5

Stochastic-like characteristics in Collatz hailstone sequences

Up to this point, this dissertation has presented a literature review on the Collatz Map in Chapter 2, a review of mathematical methods in Chapter 3, and an exploratory study on the $\boldsymbol{m}$-vectors evolved by the Collatz Map at Chapter 4. The present Chapter shall present a study on the characterization of the pseudo-random behavior of typical hailstone sequences via time series analysis techniques, reviewed in Section 3.2. Further, a characterization of correlation between the $\boldsymbol{m}$-vectors is presented aiming to obtain entropy measures for the hailstone sequences.

In order to study and characterize the more general hailstone sequences but also capture their variability, a very careful sampling of initial conditions was conducted and must be presented first in Section 5.1. Moreover, a few specific data treatments must be carried over in order to apply correctly the time series analysis methods. These specific data treatments will be explained also in Section 5.1 for the hailstone sequences in base-10 and the $\boldsymbol{m}$-vectors. Then, Section 5.2 contains the main results of this dissertation, which are now submitted for publication. Subsections 5.2.1,5.2.2 and5.2.3 presents the Power Spectrum and DFA, Autocorrelation function, and von Neumann Entropy results, respectively, together with the discussion on the implication of each result.

5.1 Methodology of Data Analysis

It is impossible to analyze every hailstone sequence once the set of natural numbers is infinite. Furthermore, any statistical analysis of these orbits comes from a sample with measure zero in the space of all possible orbits. In order to avoid studying too specific orbits, one needs to find "typical" hailstone sequences to analyze. The concepts of "typical" and "atypical" hailstone sequences are heuristic for our purposes but very useful to understand the sampling made in this study. Atypical orbits consist of too-structured orbits, e.g., with very short lengths or those with a very repetitive succession of ascents and descents. The typical orbits contain fairly heterogeneous patterns of up and down fluctuations. The former should be avoided in the present study, while the latter gives material to statistical analysis, and a sample of initial conditions that leads to them will be used.

The sampling process for the initial conditions n_0 is done by selecting the components of $\boldsymbol{m}-$vectors, generating a $\boldsymbol{m}(0)$. The values are selected in two manners: randomly,

or following a certain structured pattern, namely, all odd, even, oscillating (in a sine- or cosine-like function), elements drawn from lines of the pascal triangle, and a steady increasing or decreasing linear in a saw-tooth pattern.

The code that generates the initial $\boldsymbol{m}$-vector receives the parameter `mVectorSize`$= r_0$, stating the value of the dimension of the initial $\boldsymbol{m}$-vector[1]. For structured types, a parameter of construction called `BlockSize` is given, stating the size b of the building blocks of each structured $\boldsymbol{m}$-vector. For prime, even and odd types, the building block is formed by the first `BlockSize`$= b$ elements of that type. For oscillatory type, `BlockSize`$= b$ indicates the "wavelenght" of the oscillation, while for pascal triangle types, parameter b indicates from which line of the pascal triangle the elements must be drawn (remember that line b contains b elements). Finally, for linear type, `BlockSize`$= b$ it indicates that an arithmetic progression from 1 to b will be used as building block. For instance, if one intends to create a structured initial condition of prime type with $r_0 = 12$ and $b = 3$, then the building block is formed by the first 3 prime numbers $(2, 3, 5)$ and it is repeated $r(0)/b = 4$ times to form $\boldsymbol{m}(0) = (0, 2, 3, 5, 2, 3, 5, 2, 3, 5, 2, 3, 5)$.

In total, 997 initial conditions were generated. All the samplings, hailstone sequences, and statistical analysis were done using Julia Language [140]. The main details of this sampling process are presented in Appendix G.

In this chapter, all the analyses are made using the Accelerated Hailstone Sequence generated by Equation (2.2), then, from now on $n_t = \tilde{f}(n_0)$. The time series in base 10 representation is non-stationary, but the analysis of the semi-logarithmic scale of hailstone sequences — as in Figure 2.5 —, reveals an exponential decay for the orbit towards the unit. Once there is an interpretation for the base-2 logarithmic of the Hailstone Sequence (Equation (4.3)), one must fit the hailstone sequences into a base-2 exponential $n_t \sim n_0\, 2^{\gamma_2 t}$, obtaining a relaxation exponent γ_2 for each initial condition n_0. The exponent γ_2 is obtained simply by connecting only the initial and final points of the sequence, no regression, least squares, spline is performed.

In order to obtain a stationary orbit, one can subtract the logarithm in base-2 of the hailstone sequence from its logarithmic tendency, generating a deviation-from-drift time series $\Delta_{n_0} = (\delta_0, \delta_1, \delta_2, \ldots, \delta_T)$ where $\delta_i = \log_2[n_t] - (\log_2[n_0] - \gamma_2 i)$. Taking the increments of the deviation-from-drift time series $I_i = \delta_{i+1} - \delta_i$, and forming another time series for the increments $\mathfrak{I}_T = (I_0, I_1, \ldots, I_{T-1})$, the analysis of Power Spectrum, Detrended Fluctuation Analysis (DFA), and autocorrelation function were carried over.

Given the increments of the deviation-from-drift $\mathfrak{I}_T$ time series, the FFT is estimated using the FFTW package [141, 142] from Julia Language, and Power Spectrum is obtained by

$$P(f) = \frac{1}{T}|\hat{\mathfrak{I}}_T(f)|^2, \tag{5.1}$$

from Equation 3.37.

Also from the increments time series, the DFA is calculated by dividing the time series into contiguous segments of equal length ℓ, with $\ell \in [4, (T+1)/2]$, leading to the subsets $\mathfrak{I}_l^{(\ell)} = (I_{(l-1)\,\ell}, \ldots, I_{l\,\ell-1})$ with $l = 1, \ldots, (T+1)/\ell$. Then, each $\mathfrak{I}_l^{(\ell)}$ is fitted via least square fitting, obtaining the piece-wise linear functions $y_l(\tau)$, with τ continuously defined in the segment $(l-1)\,\ell \leq \tau \leq l\,\ell - 1$. We then consider the quadratic deviation written as $d_l(t)^2 = (I_t - y_l(\tau = t))^2$, for $t = (l-1)\,\ell, \ldots, l\,\ell - 1$. Finally, we calculate the detrended

[1] Actually, once we create only odd initial conditions $\boldsymbol{m}(0) = (0, m_2, m_3, \ldots, m_{r(0)+1})$, the true dimension of the $\boldsymbol{m}$-vector is $r(0) + 1$, and $r(0)$ is related to the size of the structured part.

fluctuation as

$$F(\ell) = (T+1)^{-1/2} \sqrt{\sum_{l} \sum_{t=(l-1)\,\ell}^{t=l\,\ell-1} d_l(t)^2}. \tag{5.2}$$

This is a detailed version of Equation (3.39).

The exponents β and α for Power Spectrum and DFA, respectively, are evaluated for all orbits in the sample, allowing a characterization of the type of process and correlations governing the dynamics of typical hailstone sequences.

Finally, the autocorrelation function for the steps of the deviation-from-drift time series is estimated using the Equation (3.32).

The $\boldsymbol{m}-$vector representation for the terms of the hailstone sequences allows proper analysis of correlations. Indeed, each exponent $m_i(t)$ can be viewed as a time series and so, the Pearson correlation matrix can be estimated by using Equations (3.28) and (3.29). From it, a recently developed method is employed to ascribe an entropy measure for the full sets of $\{m_i(t)\}$.

Therefore, let $\boldsymbol{m}_i = (m_i(1), m_i(2), \ldots, m_i(s))$ be the time series of the first s steps of the i-th $m-$vector component. Define $\boldsymbol{M}_i = m_i(t) - \langle \boldsymbol{m}_i \rangle$ as the deviations from the estimated mean $\langle \boldsymbol{m}_i \rangle$, with a variance $\sigma_i^2 = \langle \boldsymbol{M}_i^2 \rangle$. The correlation between the i'-th and i''-th terms is given by the Pearson correlation coefficient.

$$\boldsymbol{R}_{i'i''} = \frac{1}{s} \sum_{k=1}^{s} \left(\frac{m_{i'}(k) - \langle \boldsymbol{m}_{i'} \rangle}{\sigma_{i'}} \right) \left(\frac{m_{i''}(k) - \langle \boldsymbol{m}_{i''} \rangle}{\sigma_{i''}} \right) = \frac{\langle \boldsymbol{M}_{i'}\, \boldsymbol{M}_{i''} \rangle}{\sigma_{i'}\, \sigma_{i''}}. \tag{5.3}$$

From the first N components of the $\boldsymbol{m}-$vectors, a $N \times N$ Pearson correlation matrix $\mathbf{R}$ can be calculated with entries $R_{i'i''}$ given by Equation (5.3).

Finally, from $\mathbf{R}$, the von Neumann Entropy is obtained using the methods of Section 3.2.5,

$$S(\boldsymbol{\rho}) = -\mathrm{Tr}\,(\boldsymbol{\rho} \ln[\boldsymbol{\rho}]) = -\sum_{j=1}^{N} \lambda_j \ln \lambda_j, \tag{5.4}$$

where λ_j is the $j-$th (from a total of N) eigenvalue of $\boldsymbol{\rho}$. If the time series are completely non-correlated, $R_{i'i''} = \delta_{i'i''}$ and all eigenvalues are $\lambda_{i'} = 1/N$, leading to the maximum value $S_{\max} = \ln[N]$. Otherwise $S < S_{max}$, indicating some degree of correlation between the time series.

For the present analysis, it is worth mentioning that, by considering the Collatz step n_t as resulting from a sum of $r(t)$ variables at each time step — refer to Equation (4.3) – then if the time-series of the m_i's are non-correlated, this would mean that these variables are associated with non-correlated processes and a kind of central limit theorem would hold for $\ln[n_t]$ as a sum of independent random variables. On the other hand, $m_i(t)$, originated from a dynamical process, must be correlated, and the correlation analysis should identify this.

From the approach above, we must define a sample of N time-series, all with size s. For the numerical ranges used in the present work, we have verified that taking $N = 100$ components already leads to proper averages. In this way, for each orbit, the correlation matrix and entropy are then computed for various sizes s, say, for s ranging from $T/4$ to T. For each s, a value of entropy is obtained. The maximum entropy value between all s sizes is selected as the entropy of that orbit.

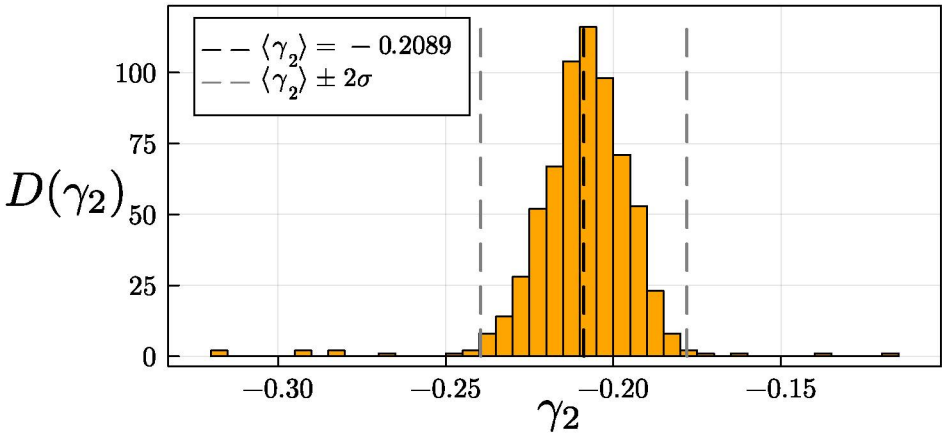


Figure 5.1: Distribution of relaxation exponents γ_2 overall 997 initial conditions with its estimated mean $\langle\gamma_2\rangle$ in black dashed line and $\langle\gamma_2\rangle \pm 2\sigma$ in the gray dashed line.

5.2 Results and Discussion

The first result to be presented is the distribution of γ_2 relaxation exponents for the 997 initial conditions, given that this result is crucial for transforming the proper hailstone sequences into stationary sequences. The estimated mean $\langle\gamma_2\rangle$ was obtained as $\langle\gamma_2\rangle = -0.21$ with estimated standard deviation of $\sigma = 0.02$. The distribution of γ_2 values over all the sample is presented in Figure 5.1, showing that the distribution is well centered with just a few orbits deviating more than 2σ from the mean.

With the estimation of expected $\langle\gamma_2\rangle$, one can estimate the expected total stopping time S from n_0 by

$$\langle S\rangle = -\frac{\log_2[n_0]}{\gamma_2} \approx 4.78858 \log_2 n_0. \tag{5.5}$$

This result is considerably close to that obtained from stochastic models in Equation (2.31)[2].

5.2.1 Power Spectrum and DFA Results

The mean values of the power spectra and DFA exponents for the increments of the deviation-from-drift time series (over our full sample of orbits) are $\langle\beta\rangle = 0.00 \pm 0.16$ and $\langle\alpha\rangle = 0.49 \pm 0.07$. Note that these averages are in good agreement with the relation in Equation (3.40). By separating the time series in types — according to the specific feature of the corresponding $\boldsymbol{m}$-vectors — the resulting $\langle\beta\rangle$ and $\langle\alpha\rangle$, together with their respective standard deviations, are summarized in Table 5.1. In Figures 5.2 (a) and (b) it is displayed typical examples of power spectrum and DFA curves, respectively. Box plots for the distribution of the β and α exponents for each initial condition type are depicted in

[2]To obtain equation (2.31) in $\log_2$ just consider

$$\ln a = \frac{\log_2(a)}{\log_2(e)} \approx \frac{\log_2(a)}{1.4426}.$$

type	$\langle\beta\rangle \pm 2\sigma_\beta$	$\langle\alpha\rangle \pm 2\sigma_\alpha$	$\langle\mathbb{S}\rangle \pm 2\sigma_\mathbb{S}$
Random	-0.01 ± 0.16	0.48 ± 0.06	0.992 ± 0.005
Prime	-0.01 ± 0.16	0.49 ± 0.07	0.98 ± 0.03
Even	-0.02 ± 0.16	0.49 ± 0.07	0.96 ± 0.04
Odd	0.00 ± 0.17	0.48 ± 0.06	0.97 ± 0.06
Oscillatory	0.00 ± 0.10	0.49 ± 0.06	0.98 ± 0.06
Pascal Triangle	0.02 ± 0.22	0.49 ± 0.08	0.97 ± 0.10
Linear	0.00 ± 0.10	0.49 ± 0.06	0.98 ± 0.04

Table 5.1: Summary of the mean β, α and relative von Neuman entropy $\mathbb{S}$ values, with their respective standard deviation for the samples of each type of initial m-vector.

Figures 5.2 (c) and (d), respectively. Remarkably, they present the characteristic behavior of white noise time series.

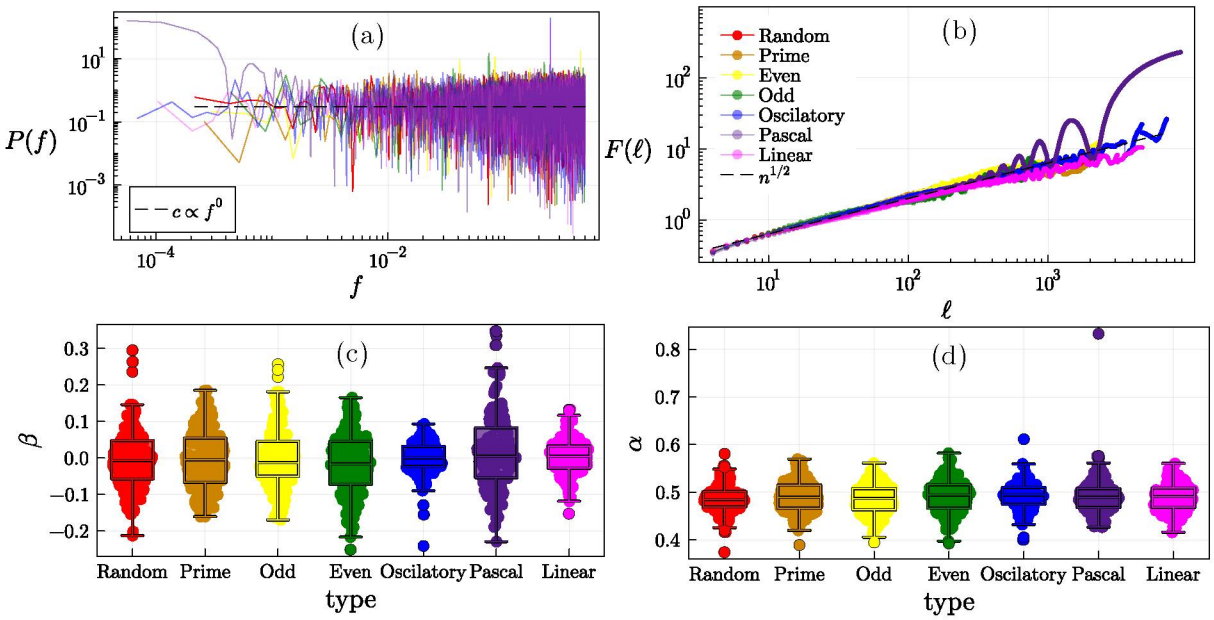


Figure 5.2: For a representative orbit of each $\boldsymbol{m}$-vector type, typical plots of the power spectrum $P(f)$ and of the DFA $F(\ell)$, respectively, in (a) and (b). Note that the behavior in (a), for which $c = 0.3$, indicates the usual trend of white noise. The curves in (b) are also compared with $\ell^{1/2}$. The good fitting once more points to white noise. Box-plot distributions of (c) β and (d) α for the sample orbits of each m-vector type. In (b) the $F(\ell)$ for Pascal type (in purple) is depicts the deviating α dot in Pascal box-plot in (d), obtained for a very specific initial condition $\boldsymbol{m}(n_0) = (0,1,1,\ldots,1,1)$, the Mersenne number $n_0 = 2^{180} - 1$.

The sample of orbits with random initial $\boldsymbol{m}$-vector components can be regarded as the most unbiased sample that can be generated. Consequently, a null hypothesis can be formulated stating that structured types contain the same distribution of α and β that the random type. A t-test was performed using Julia Language's HypothesisTests.jl package to compare Prime, Even, Odd, Oscillatory, Pascal and Linear distributions of α and β with the Random distribution.

For the prime type, the p-value for α (β) is 0.086 (0.78). Although relatively small for α, it is still above the gold standard (of $p > 0.05$). Of course, for β it is considerably high. Thus, the prime has no statistically significant difference from the random type. The same is true for the odd and even types. For the former, the p-values are 0.44 and 0.65 for β and α, respectively. For the latter, they are 0.13 and 0.067, so smaller than for the odd case, but still larger than the gold standard of significance. For Oscilatory, Pascal and Linear types, the p-values for α (β) are 0.09 (0.22), 0.11 ($0.0078 < 0.05$) and 0.65 (0.10), respectively. The only type with statistically significant difference from random type is the β distribution for pascal type. This can be seen in Figure 5.2 (d), where the Pascal type has way more distributed α values, and in Table 5.1 where the standard deviation of Pascal type is more than twice the others. Further, Pascal type seem to contain the largest deviations, hence the $F(\ell)$ with most deviating α value in Figure 5.2 (d) for Pascal type is depicted in Figure 5.2 (b). It refers to a very specific initial condition with 181 components in the m-vector, zero in the first, and one in the rest, i.e. $\boldsymbol{m}(n_0) = (0, 1, 1, \ldots, 1, 1)$ for the Mersenne number $n_0 = 2^{180} - 1$. This type of deviating behavior is, as can be seen from the size of our sample, very rare.

Therefore, the above results strongly support the following

1. The increments of the deviation-from-drift are mostly white noise (characterized by $\beta = 0,\ \alpha = 1/2$), leading to the conclusion that the deviation-from-drift is Brownian and, once the drift is exponential, the original hailstone sequence indeed performs GBM;

2. The Power Spectrum and DFA exponents β and α distributions are very similar for random and structured initial $\boldsymbol{m}$-vectors, indicating, in a first glimpse, that the initial $\boldsymbol{m}$-vector's structure does not play any roll in deviating the series from GBM, except in normally distributed fluctuations.

Random components of $\boldsymbol{m}$-vectors also describe a larger fraction of all the typical orbits than structured, meaning that the results and measurements obtained for them are the results and measurements for *most* typical hailstone sequences. Hence, once the exponents of α and β are significantly equally distributed both in random and structured initial conditions, one can point out with more certainty that GBM is a feature very common in hailstone sequences. These findings agree with the literature review presented in Chapter 2, especially in Sections 2.4.3 and 2.5.2.

5.2.2 Autocorrelation Results

The Power Spectrum and DFA analysis confirms the pseudo-random behavior of the sequences, but still, some determinism must be hidden inside the general trend, once we are dealing with a deterministic system. If the time series increments are pure white noise, then the autocorrelation function must go towards zero after $\tau = 0$. That is, for white noise $R_{\text{WN}}(\tau = 0) = 1$ and $R_{\text{WN}}(\tau > 0) = 0$. In the investigation of the autocorrelation of the increments of the deviation-from-drift time series, it is found that it decays towards zero, with no power law or exponential behavior. Instead, $R(0) = 1$, and for $\tau > 0$ an abrupt decay on auto-correlations is perceived.

To testify whether the auto-correlations are non-significant — Section 3.2.2 — after $\tau = 0$ one must calculate the error of $R(\tau)$ by estimating the standard deviation of R for each τ. The calculation of $\sigma(R(\tau))$ is done twofold: first considering that the time series values are i.i.d. such that $\sigma(R(\tau)) = 1/\sqrt{T}$, where T is the length of the signal; but this is

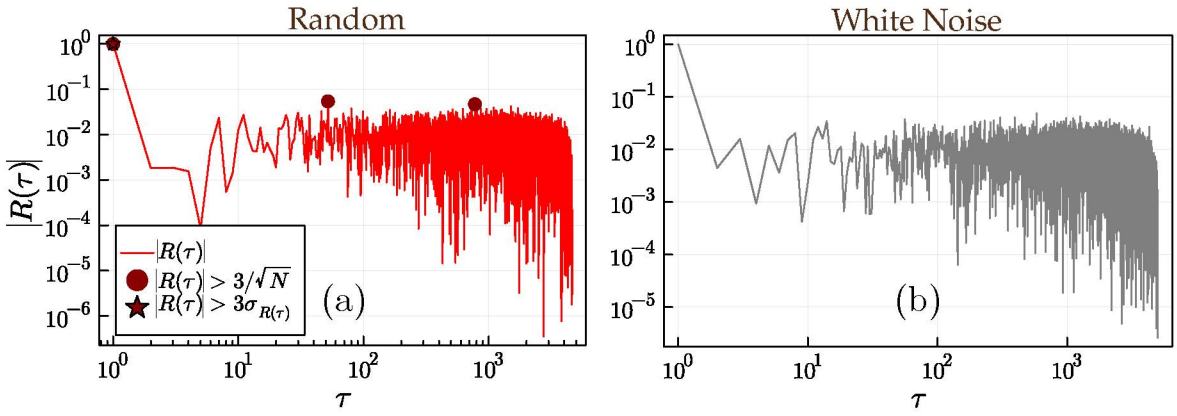


Figure 5.3: Comparison between the autocorrelation function of (a) increments of deviation-from-drift time series for random type initial condition and (b) a pure white noise time series. Dark-red circles indicate significant correlation considering i.i.d. correlations.

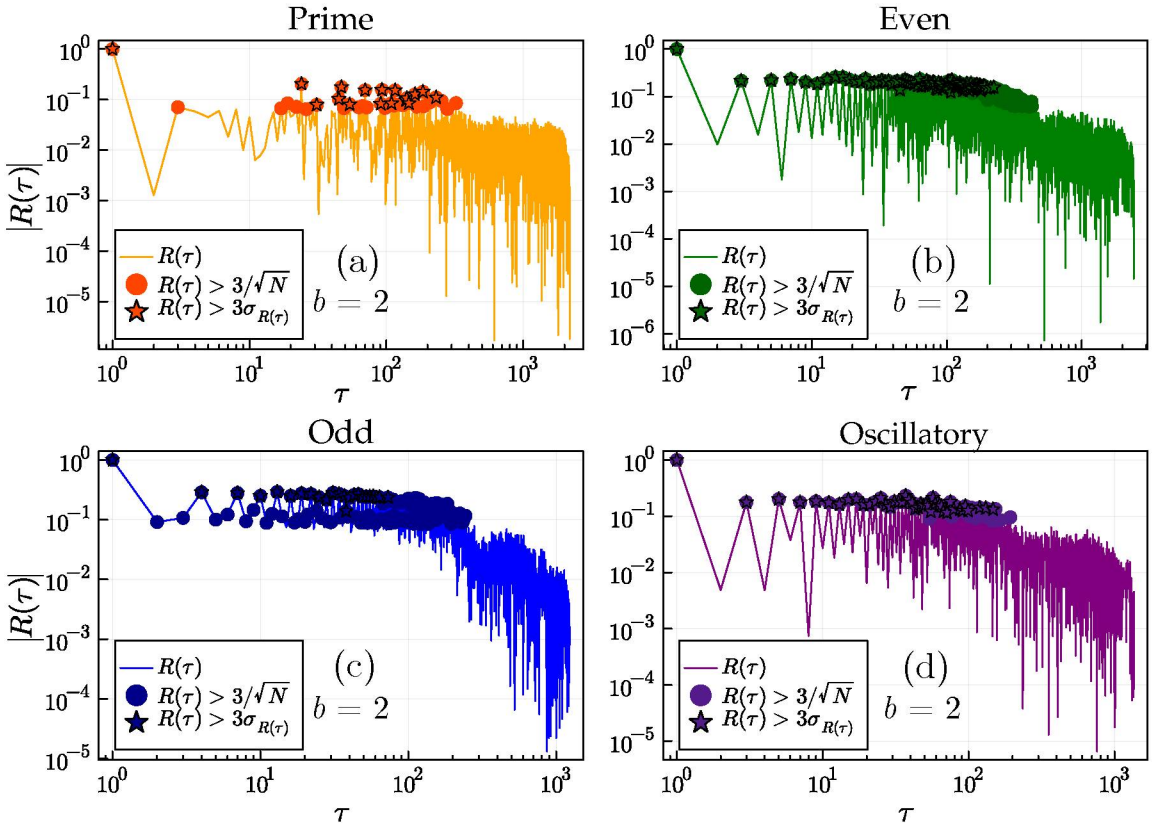


Figure 5.4: Autocorrelation function for structured types with blocksize $b = 2$ for (a) Prime, (b) Even, (c) Odd, and (d) Oscillatory types of initial conditions. Dark circles indicate a significant correlation considering i.i.d. correlations and dark stars indicate significant correlation considering Bartlett's approximation.

not necessarily the case, once we are trying to investigate the existence of correlations, hence, second, we consider Bartlett's approximation from Equation (3.34), denoted as $\sigma(R(\tau)) = \sigma_{R(\tau)}$, for autocorrelation of finite time series.

Figure 5.3(a) shows the autocorrelation function for one representative time series of

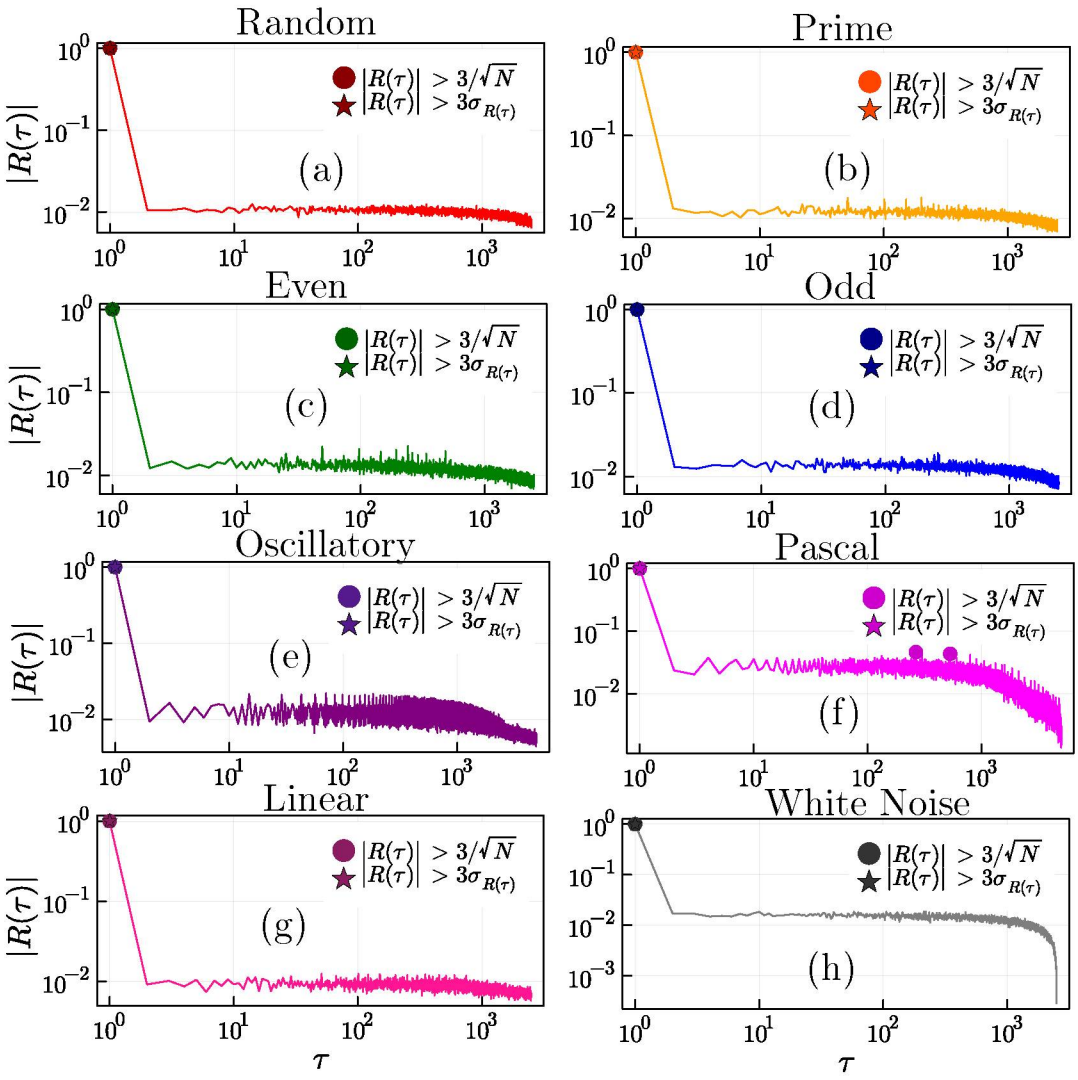


Figure 5.5: Mean autocorrelation functions up to $\tau = 2500$ over the entire sample of each types (a) random, (b) prime, (c) even, (d) odd, (e) oscillatory, (f) pascal triangle and (g) linear, compared with the mean autocorrelation function over 100 white noise time series with $T = 2500$.

random type initial condition in comparison with Figure 5.3(b) for a white noise signal, indicating that the most unbiased case (random type) is indeed very similar to a white noise. This plot was repeated and analyzed — but not shown — for all hailstone sequences of random type and the same behavior was found for all. This means the increments of time series for random type initial condition are uncorrelated. But Figure 5.4 presents representative examples of the autocorrelation function for structured types with the smallest blocksize parameter $b = 2$. In this case, various significant short- to mid-range correlations appears, indicating that for some specific initial conditions there are deviations from the GBM — a feature that confirms the hypothesis raised by [92]. These figures are only representative examples for each type of initial conditions, relevant for the discussion bellow. In order to investigate de general trend, the mean autocorrelation $\langle R(\tau) \rangle$ for each lag τ over the sample of each type is presented in Figure 5.5(a)-(g). Figure 5.5(h) presents the mean autocorrelation over 100 white noise time series, providing comparison between the general trends of each type and stochastic behavior.

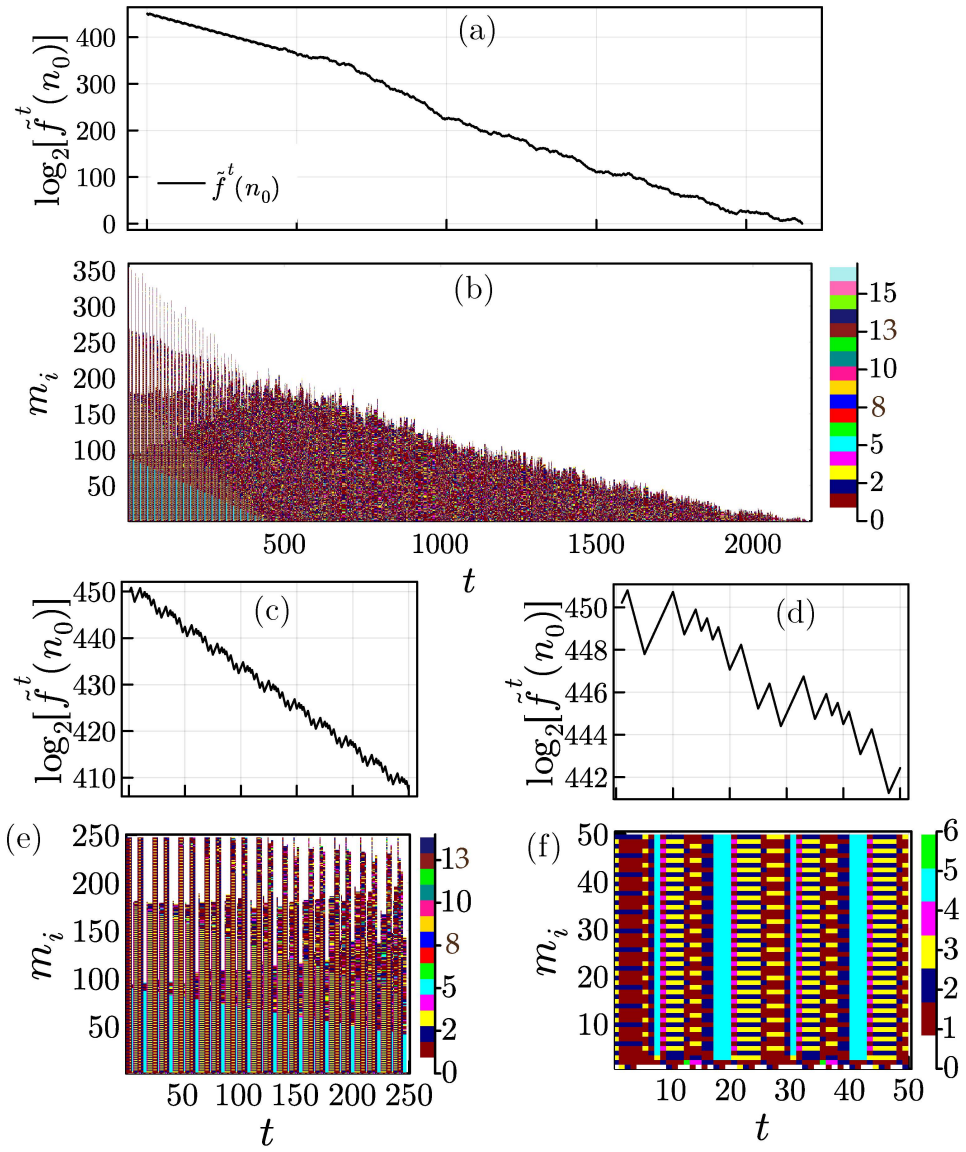


Figure 5.6: (a) Hailstone sequence for a prime type initial condition with $b = 2$ in base-2 log-scale; (b) Heat-map representation of the $\boldsymbol{M}$-Matrix for the hailstone sequence in (a), enabling one to perceive that the initial structure is preserved and directly related to the periodic ascents and descents; Zoom into the first 250 points (c) and from $t = 100$ to $t = 150$ (d) of the hailstone sequence, detailing and evidencing the presence of periodic ups and downs stretches; (e) and (f) zooms into the $\boldsymbol{M}$-Matrix correspondent to the piece of the hailstone sequence in (c) and (d), respectively.

Further, the presence of significant auto-correlations is directly related to the plateaus of periodic ascents and descents in the orbits of specific initial conditions, already pointed out by ref. [92]. This is exemplified in Figures 5.6 (a), 5.6 (c) and 5.6 (d) where the hailstone sequences for a prime type initial condition with $b = 2$ is shown with zoom-ins into the first steps (the same orbit whose auto-correlation is depicted in Figure 5.4(a)). Every observed hailstone sequence with these plateaus of periodic ascents and descents contains significant short- to mid-range correlations. The presence of these correlations and periodic ups and downs strictly in structured initial conditions indicates that the sequences are initial condition dependent, leading to the presence of richer dynamics other

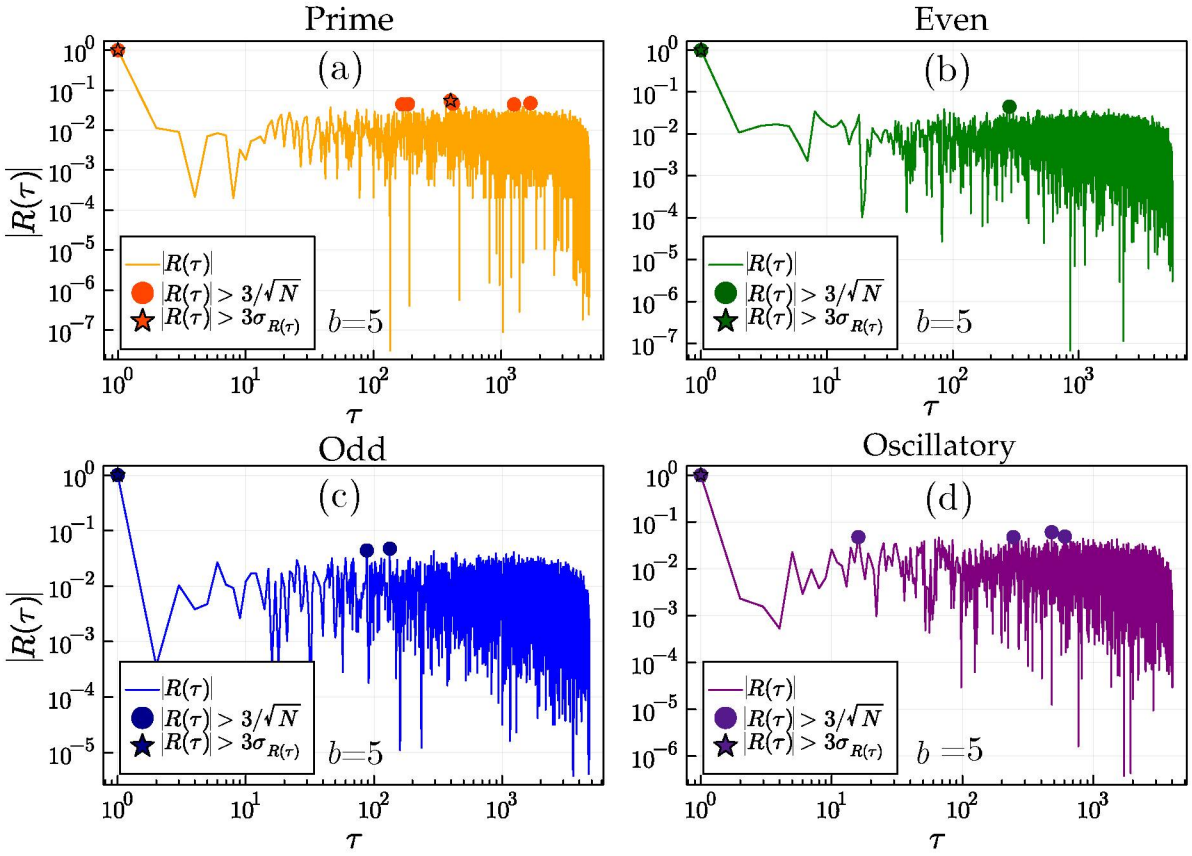


Figure 5.7: Autocorrelation function for structured types with blocksize $b = 5$ for (a) Prime, (b) Even, (c) Odd, and (d) Oscillatory types of initial conditions. Dark circles indicate significant correlation considering i.i.d. correlations and dark stars indicate a significant correlation considering Bartlett's approximation.

than GBM.

These finds could, at a first sight, indicate against the possibility of proving the conjecture by associating the map as a model for a random walker (RW) with an absorbing barrier. But the peaks in auto-correlations presented in Figure 5.4 only appear for some specific structured initial conditions, where the `BlockSize` parameter b (see Appendix G for details) is small, essentially, $b = 2$ and $b = 3$ are the only parameters that create various significant peaks. Figure 5.7 presents the autocorrelation for structured types with larger blocksize parameter $b = 5$, and the short-to-mid range correlations disappear, with the ACF resembling those from random type in Figure 5.3. This indicates that orbits deviating from GBM are very rare, representing only fluctuations around $\beta = 0, \alpha = 1/2$ as Figure 5.2 presents, leading, in fact, to a reinforcement of the idea of proving the conjecture by the association with an RW.

These peaks in auto-correlations and the structured plateaus can be better understood by studying the $\boldsymbol{M}$-Matrix. By observing heat-map representations of the $\boldsymbol{M}$-Matrix in Figures 5.6(b), 5.6 (d) and 5.6 (f), and comparing with the orbit that generates this $\boldsymbol{M}$-Matrix in Figures 5.6 (a), 5.6(c) and 5.6(d), it is possible to see that the structured plateaus in the base-10 representation, only exists while some periodic structure persists in the $\boldsymbol{M}$-matrix. The structure is more evident in Figure 5.8, for an oscillatory type with $b = 2$, and can be slightly identified in Figure 5.9 for a prime type with $b = 5$. It is remarkable by Figures 5.8, and by observing the others structured types with $b = 2$ and

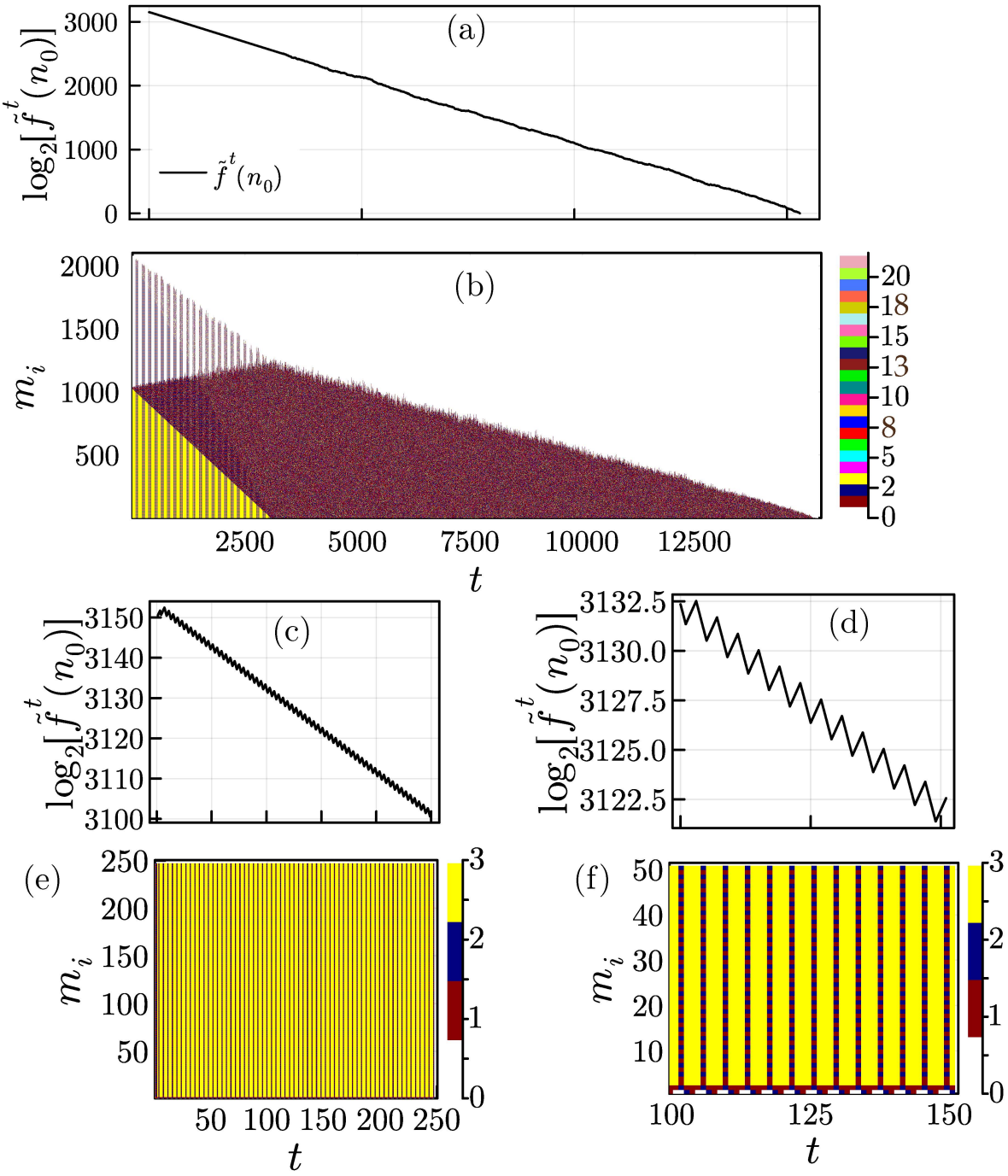


Figure 5.8: (a) Hailstone sequence for for an oscillatory type initial condition with $b = 2$ and initial m-vector size $r = 2100$ (e.g. $n_0 \equiv (0, 1, 2, 1, 2, \ldots, 1, 2, 1, 2)$); (b) $\boldsymbol{M}$-Martix of hailstone sequence in (a), revealing that for the first 2500 steps, some very intricate structure is preserved, being gradually destroyed and replaced by stochastic behavior; Zoom into the first 250 points (c) and from $t = 100$ to $t = 150$ (d) of the hailstone sequence, detailing and evidencing the presence of periodic ups and downs stretches; (e) and (f) zooms into the $\boldsymbol{M}$-Matrix correspondent to the piece of the hailstone sequence in (c) and (d), respectively.

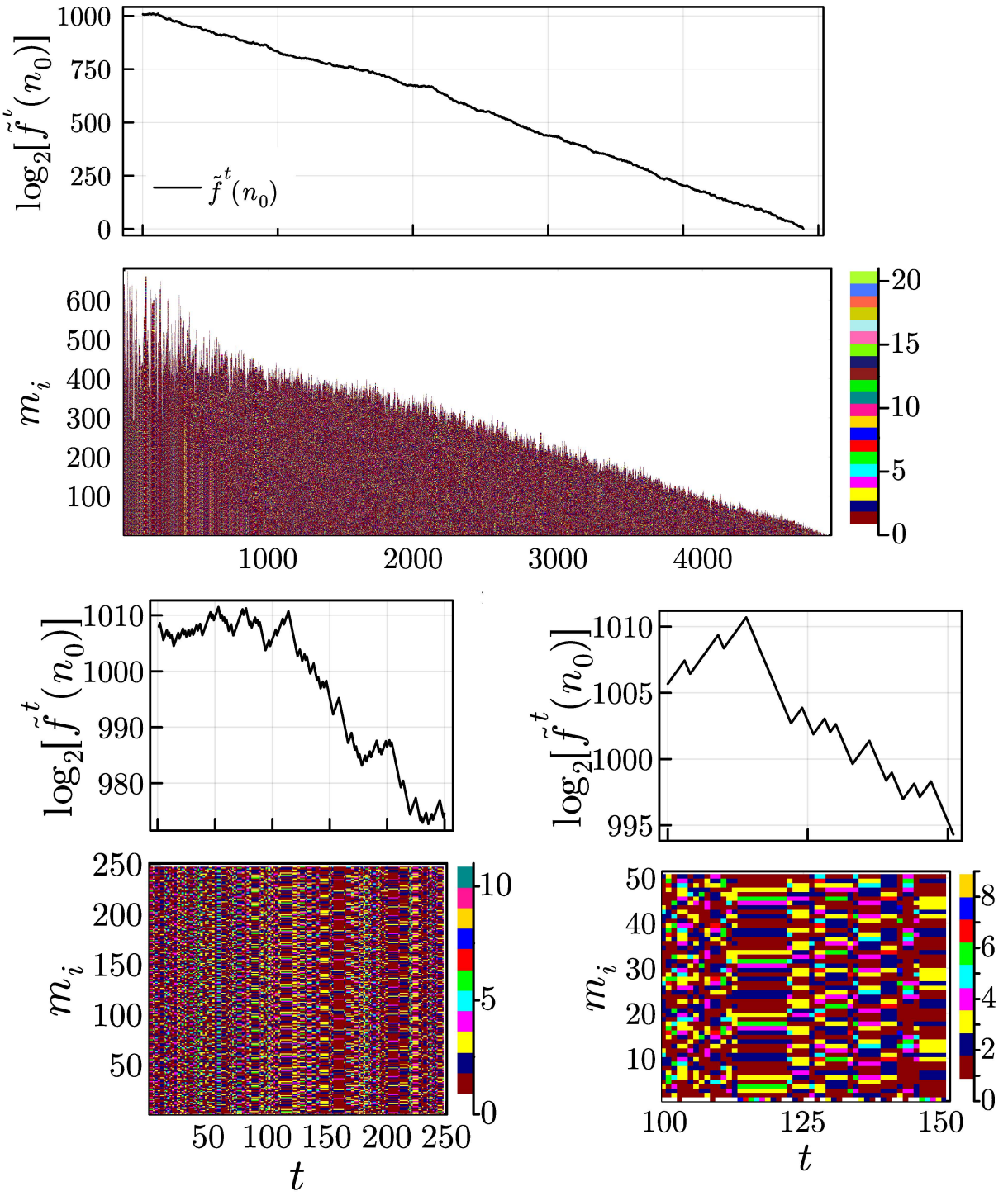


Figure 5.9: (a) Hailstone sequence for a prime type initial condition with $b = 5$ in base-2 log-scale; (b) Heat-map representation of the $\boldsymbol{M}$-Matrix for the hailstone sequence in (a), evidencing some weak structure that does not appears in (a) as a sequence of periodic ups and downs; Zoom into the first 250 points (c) and from $t = 100$ to $t = 150$ (d) of the hailstone sequence, detailing and evidencing the presence of periodic ups and downs stretches; (e) and (f) zooms into the $\boldsymbol{M}$-Matrix correspondent to the piece of the hailstone sequence in (c) and (d), respectively.

$b = 3$ (not shown, but exhaustively analyzed) that the initial structure in the $\boldsymbol{M}$-Matrix is preserved as long as there are structured plateaus in the base-10 orbit. In the instant when no structure is left on the $\boldsymbol{M}$-Matrix, the structure in base-10 disappears and GBM takes control. So the dynamics of the $\boldsymbol{m}$-vectors components seem to be as fundamental, or even more, as the base-10 orbit itself. The dynamics of the $\boldsymbol{m}$-vectors components unveil more than it is on the "surface" of base-10 orbits.

5.2.3 von-Neumann Entropy Results

In the last section, it was stated that the presence of short- to mid-range correlations, the plateaus of periodic ascents and descents, and the structures inside the $\boldsymbol{M}$-Matrices are linked. But the plateaus and short- to mid-range correlations only appear when $b = 2$ and $b = 3$, hinting that the structure of larger building blocks has no structure. Still, when one looks into the $\boldsymbol{M}$-matrix of a prime type initial condition with $b = 5$ in Figure 5.9 (a)-(c), one can perceive the existence of some weak structures in the $\boldsymbol{m}$-vectors components time series. This weak structure in the $\boldsymbol{M}$-Matrix does not reveals itself as a structure in the base-10 orbit, nor is perceived by the autocorrelation analysis. Hence, another type of analysis must be done in order to testify if these weak structures represent a significant change in the $\boldsymbol{M}$-Matrix and $\boldsymbol{m}$-vector components dynamics. To perform this analysis, the von Neumann entropy measurement of m-vector components time series is obtained.

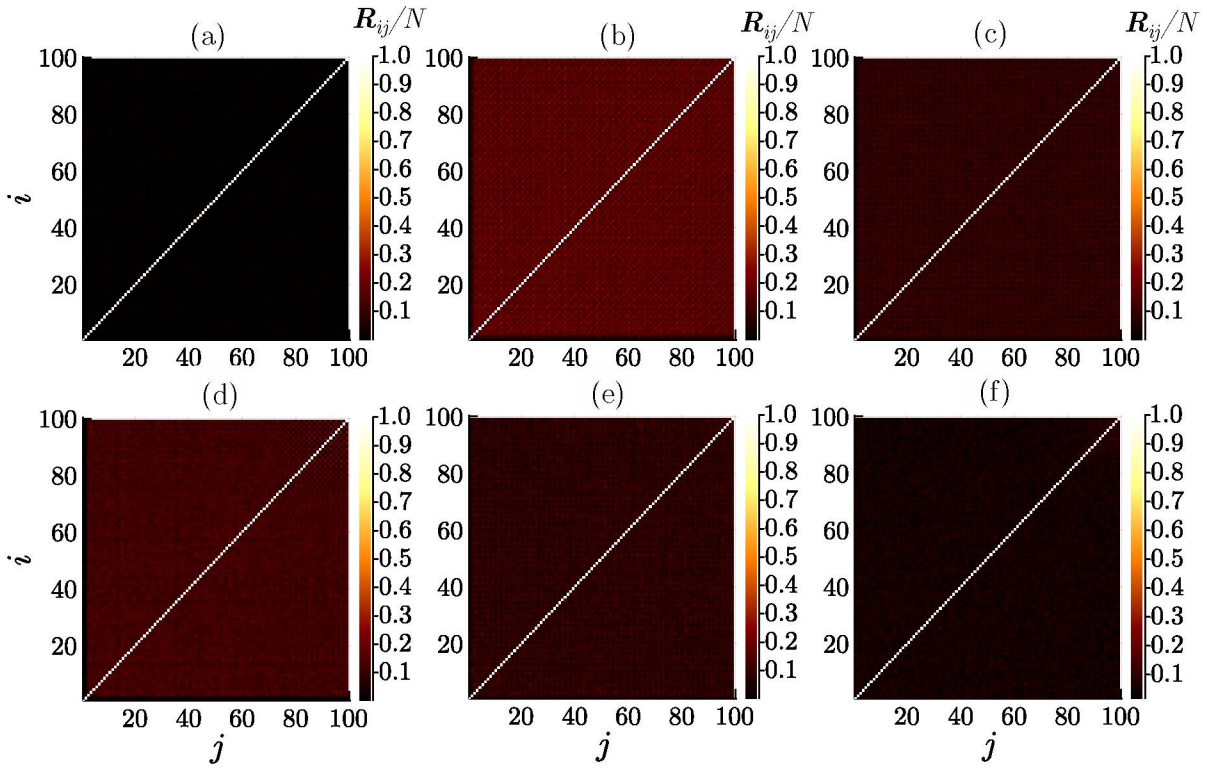


Figure 5.10: Heatmap of the Pearson correlation matrices for examples of each type (a) random (b) prime (c) even, (d) odd, (e) Oscillatory, and (f) Pascal.

In order to obtain the von Neumann entropy, the Pearson Correlation Matrix $\boldsymbol{R}$ must be calculated. In Figure 5.10 examples of the $\boldsymbol{R}$ matrices are presented for each type of initial conditions for $\boldsymbol{m}$-vector initial dimension equal to 2100. For random type, the matrix is very similar to the identity, implying in uncorrelated m-vector's components

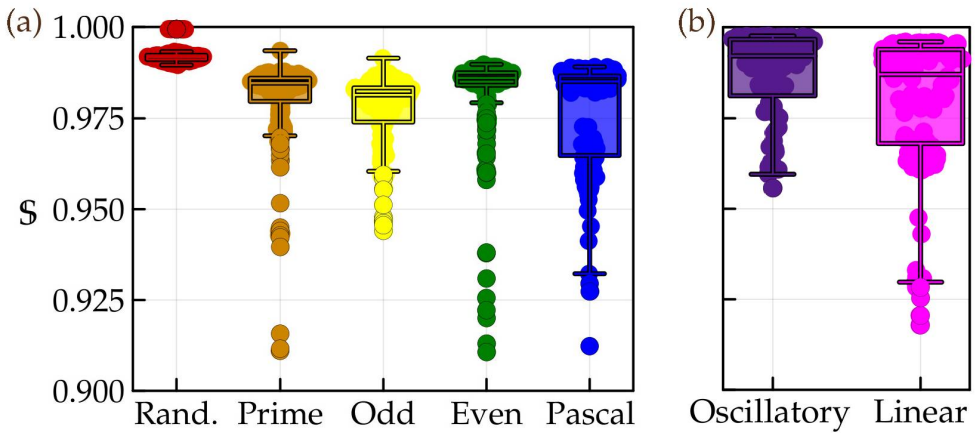


Figure 5.11: Box-plot of the entropy distribution for the (a) Random, prime, odd, even and pascal types, and (b) oscillatory and linear types, evidencing the decrease in the entropy for the structureds in (a) and a different distribution, with presence of lower entropy values in the structureds of (b).

time series, while for structured the non-diagonal terms are larger, implying a correlation between m-vectors time series components. For the structured types, the correlation matrices present some patterns, indicating that the correlations between time series follow a pattern between themselves.

The mean $\boldsymbol{m}$-vectors von Neuman entropy measurements over all orbits in the sample is $\langle \mathbb{S} \rangle = 0.9800$. This result by itself points out in the direction that, independent of any attempt to create organization in the $\boldsymbol{m}$-vectors, its components are still very uncorrelated, leading to an entropy value close to the maximum. For each of the specific types of initial $\boldsymbol{m}$-vectors, the relative entropy values, together with their standard deviations, are summarized in Table 5.1. The mean relative entropy for structured types is lower than for random types.

The distribution of $\mathbb{S}$ values for each type are presented in Figure 5.11 and the structured distributions can be compared, employing a t-test, with the null hypothesis of random type distribution. For all three structures whose samples are large enough, i.e. prime, even, and odd types, one finds $p \ll 0.05$, meaning that the null hypothesis is rejected with statistical significance.

This allows the conclusion that structured initial conditions generate different distributions of entropy, with a lower mean. This lower entropy for structured types is a strong indicator that indeed, the von Neumann entropy method can spot the presence of the weak structures in Fig. 5.9 and these structures represent significant changes in the $\boldsymbol{M}$-Matrix and $\boldsymbol{m}$-vector dynamics. Besides the decrease on the entropy values for some initial conditions, the values continue to be close to $S_{\max}$, meaning that, even for structured types, the $\boldsymbol{m}$-vector's components time series are very uncorrelated.

Chapter 6

Conclusions and Remarks

The area of stochastic arithmetic dynamical systems is still an emergent area heading toward its first steps. By the end of the present dissertation, the author hopes to have helped with a small "knee-step" in the direction of establishing stochastic arithmetic dynamical systems as a research area. In an overview, what has been done and presented in this work, at a first sight, could look like just a bunch of statistical investigations of orbits from a number theory problem with few application. But there is more than that, there is a take-home message attained after various statistical analyses, about how complicated and seemingly unpredictable dynamics can arise from a very simple and deterministic mathematical rule with natural numbers. In this sense, it is possible to affirm that the main objectives of this work, about exploring the hailstone sequences and their stochastic-like behavior using arithmetic dynamics and time series methods was fully completed.

First, the various methods of time series analysis employed to study the base-10 hailstone sequences provide reinforcements to the literature. The power spectrum and DFA analysis are very coherent and complementary with each other, both indicating white noise in the increments of time series, meaning that the deviation-from-drift time series itself is Brownian — and the hailstone sequences are GBM as the literature indicates. The autocorrelation function, besides providing pieces of evidence for deviations from GBM, and indicators of the presence of determinism inside the sequences, confirms that most orbits are indeed uncorrelated, reinforcing power spectrum and DFA findings.

Results presented on Chapter 4 about the dynamics of the map in the $\boldsymbol{m}$-vectors space were very important in order to characterize the internal structure of the natural numbers that compose the hailstone sequences. By defining the $\boldsymbol{M}$-Matrix, it was possible to observe how the $\boldsymbol{m}$-vectors components seem to follow an unpredictable pattern when subjected to the Collatz Map. Equation (4.3) stating that the sum of the $\boldsymbol{m}$-vectors' components equals to the integer part of $\log_2[\tilde{f}^t(n_0)]$ allowed an interpretation of the proper hailstone sequence in base-10 as a sum of various other time series, i.e. the $\boldsymbol{m}$-vectors components time series. Further, the large relative von Neumann entropy measurements presented in Chapter 5 showed that $\boldsymbol{m}$-vectors components are mostly uncorrelated, meaning that these variables are associated with non-correlated processes, and a central limit theorem might hold for $\log_2[\tilde{f}^t(n_0)]$. It is also important to remark that von Neumann's entropy is sensitive to the internal organization of the $\boldsymbol{m}$-vectors in a way that neither autocorrelation nor power spectrum nor DFA is.

Besides the important findings of deviations from GBM by autocorrelation and entropy measurements, randomness is still very common in the set of hailstone sequences.

Consequently, these finds reinforce the possibility of proving the conjecture by associating the map with a random walker with an absorbing barrier. This is so because, even when one tries to force structures and deviates from GBM, the evolution of the map pulls the sequence towards stochastic-like behavior. The calculation of Power Spectrum, DFA, and $R(\tau)$ confirms the randomness taking control in base-10, besides the deviations for some specific initial conditions. But one of the great conclusions here is that the presence of randomness in base-10 hailstone sequences must be related to the presence of randomness in the $\boldsymbol{m}$-vectors time series as Equation (4.3) states. By pure observation of $\boldsymbol{M}$-matrices, it is possible to conceive the Collatz map as a structure destroyer, that creates random patterns in a slow process of destruction of order. The statistical analysis and the measurements made for Collatz hailstone sequences and its $\boldsymbol{M}$-Matrix indicate strongly that this simple mathematical system is one of the most primitive structure underlying complexity and stochastic behavior in Nature.

The findings presented in this dissertation connect two seemingly contradictory views raised in the first paragraphs of the introduction: the Occam's razor and the constant interplay between simplicity and complexity cited by [3]. Occam's razor is contemplated by finding one of the simplest models for stochastic and complex phenomena, hinting that these features, besides appearing in very complicated systems, are rooted in very simple rules. On the other hand, the rise of complex and stochastic behavior surely maintains the interplay between simple and complex, indicating again that these concepts are holding hands.

To conclude, we discuss some remarks on number theory and physics, followed by a brief list of future works.

Although hard to solve in a number theoretic framework, because of the aforementioned dichotomy between the deterministic character of number series and the randomness of these sequences, some number theory problems might be very useful as models for physical systems. One of the most remarkable examples of number theory and physics merging is the Riemann Hypothesis [54]. Therefore, by the present study presented here, we expect to have created another bridge between a number-theoretical hard-to-solve problem, and other areas, specifically, physics and complex systems. The Collatz Map, besides providing very rich dynamics for natural numbers as pointed out in this dissertation, only recently has been explored in the context of physical models and phenomena. In this context, in the present work, several features of the dynamics of natural numbers were explored, as well as a brand new representation, similar to $p-$adics, that unveils natural phenomena emerging from pure mathematics.

From all results presented in this dissertation, the following topics might be interesting of pursuing in future works:

- Describing the Collatz Gas, as subsection 4.4.1 has presented, would provide an enhancement to this bridge between number theory and physics, akin to the Riemman Gas, perhaps leading to insights about proving or disproving the conjecture.

- Related to the Collatz Gas topic, the dictionary between the Collatz Map and the Operators of Quantum Harmonic Oscillator of Perelman [96] seems to be very fruitful and useful in order to give physical meaning to the conjecture.

- Defining the algebraic structures governing the $\boldsymbol{m}$-vector's space, as subsection 4.4.2 has presented, providing some metric and defining how to evolve them by the Collatz Map would fasten the algorithm of generating the $\boldsymbol{M}$-Matrices for very large numbers, and enlighten the apparently random behavior of the $\boldsymbol{M}$-Matrix.

- The characterization of transition graphs of $\boldsymbol{m}$-vectors, as idealized in subsection 4.4.3 would allow another level of understanding of this internal structure, maybe providing connections to another natural phenomena displaying these types of behaviors.
- Defining chaos in discrete arithmetic systems seems like a very non-trivial task. But it has been shown in this dissertation that, besides the stochastic-like behavior, some determinism is still present on the orbits. From these, we ask if the sketches of Lyapunov Exponent drawn in Section 4.4.4 can be a trust measurement of chaos?

These are all interesting questions that urge future works to be answered.

Bibliography

[1] R. B. Lindsay. "The Meaning of Simplicity in Physics". In: *Philosophy of Science* 4.2 (1937), 151–167. DOI: 10.1086/286451 (cit. on p. 15).

[2] S. Hawking. *The theory of everything*. Jaico Publishing House, 2006 (cit. on p. 15).

[3] M. Gell-Mann. "Simplicity and complexity in the description of nature". In: *Engineering and Science* 51.3 (1988), pp. 2–9 (cit. on pp. 15, 85).

[4] P. Bak. *How nature works: the science of self-organized criticality*. Springer Science & Business Media, 2013 (cit. on pp. 15, 33, 51).

[5] E. P. Wigner. "The unreasonable effectiveness of mathematics in the natural sciences. Richard courant lecture in mathematical sciences delivered at New York University, May 11, 1959". In: *Communications on Pure and Applied Mathematics* 13.1 (1960), pp. 1–14. DOI: https://doi.org/10.1002/cpa.3160130102. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.3160130102. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.3160130102 (cit. on p. 15).

[6] M. Dorato. "Why Are (Most) Laws of Nature Mathematical?" In: *Nature's Principles*. Ed. by J. Faye et al. Dordrecht: Springer Netherlands, 2005, pp. 55–75. ISBN: 978-1-4020-3258-5. DOI: 10.1007/1-4020-3258-7_2. URL: https://doi.org/10.1007/1-4020-3258-7_2 (cit. on p. 15).

[7] M. Tegmark. *Our mathematical universe: My quest for the ultimate nature of reality*. Vintage, 2014 (cit. on p. 15).

[8] M. Tegmark. "The mathematical universe". In: *Foundations of physics* 38.2 (2008), pp. 101–150 (cit. on p. 15).

[9] C. S. Calude and K. Svozil. "Spurious, Emergent Laws in Number Worlds". In: *Philosophies* 4.2 (2019), pp. 17–0. ISSN: 2409-9287. DOI: 10.3390/philosophies4020017. URL: https://www.mdpi.com/2409-9287/4/2/17 (cit. on p. 15).

[10] D. Adams. *The Hitchhiker's Guide to the Galaxy: The Complete Trilogy of Five*. Pan Macmillan, 2012 (cit. on p. 16).

[11] M. Kim. "Arithmetic gauge theory: A brief introduction". In: *Modern Physics Letters A* 33.29 (2018), p. 1830012 (cit. on p. 16).

[12] M. Gori et al. "Topological theory of phase transitions". In: *Journal of Physics A: Mathematical and Theoretical* (2022). URL: http://iopscience.iop.org/article/10.1088/1751-8121/ac7f09 (cit. on p. 16).

[13] J. H. Silverman. *The arithmetic of dynamical systems*. Vol. 241. Springer Science & Business Media, 2007 (cit. on pp. 16, 26).

[14] S. Ferenczi, J. Kulaga-Przymus, and M. Lemanczyk. *Ergodic Theory and Dynamical Systems in their Interactions with Arithmetics and Combinatorics*. Springer, 2018 (cit. on p. 16).

[15] A. Vladimir and K. Andrei. *Applied Algebraic Dynamics*. De Gruyter, 2009. isbn: 9783110203011. doi: doi:10.1515/9783110203011. url: https://doi.org/10.1515/9783110203011 (cit. on p. 16).

[16] B. Dragovich et al. "On p-adic mathematical physics". In: *P-Adic Numbers, Ultrametric Analysis, and Applications* 1.1 (2009), pp. 1–17 (cit. on p. 16).

[17] B. Dragovich et al. "p-Adic mathematical physics: the first 30 years". In: *P-Adic numbers, ultrametric analysis and applications* 9.2 (2017), pp. 87–121 (cit. on pp. 16, 41).

[18] A. M. Robert. *A course in p-adic analysis*. Vol. 198. Springer Science & Business Media, 2013 (cit. on pp. 16, 41, 100).

[19] F. Q. Gouvea. *p-adic Numbers: An Introduction*. 3rd. Springer Berlin, Heidelberg, 1997. doi: https://doi.org/10.1007/978-3-642-59058-0 (cit. on pp. 16, 40, 42).

[20] I. Y. Aref'eva, B. G. Dragović, and I. V. Volovich. "On the p-adic summability of the anharmonic oscillator". In: *Physics Letters B* 200.4 (1988), pp. 512–514 (cit. on p. 16).

[21] V. S. Vladimirov, I. V. Volovich, and E. I. Zelenov. *p-adic Analysis and Mathematical Physics*. World Scientific, 1994 (cit. on p. 16).

[22] E. G. Beltrametti and G. Cassinelli. "Quantum mechanics andp-adic numbers". In: *Foundations of Physics* 2.1 (1972), pp. 1–7 (cit. on p. 16).

[23] I. Y. Aref'eva et al. "The wave function of the Universe and p-adic gravity". In: *International Journal of Modern Physics A* 6.24 (1991), pp. 4341–4358 (cit. on p. 16).

[24] I. V. Volovich. "p-Adic string". In: *Classical and Quantum Gravity* 4.4 (1987), p. L83 (cit. on p. 16).

[25] S. Boettcher and M. Paczuski. "Ultrametricity and memory in a solvable model of self-organized criticality". In: *Physical Review E* 54.2 (1996), p. 1082 (cit. on pp. 16, 41, 44, 45).

[26] O. N. Khakimov. "On p-adic Gibbs measures for Ising model with four competing interactions". In: *P-Adic Numbers, Ultrametric Analysis, and Applications* 5.3 (2013), pp. 194–203 (cit. on p. 16).

[27] N. N. Ganikhodjaev, F. M. Mukhamedov, and U. A. Rozikov. "Phase transitions of the Ising model on Z in the p-adic number field". In: *Uzbek. Math. J.* 4 (1998), pp. 23–29 (cit. on p. 16).

[28] M. Farrukh. "A Dynamical System Approach to Phase Transitions for p-Adic Potts Model on the Cayley Tree of Order Two". In: *Reports on Mathematical Physics* 70.3 (2012), pp. 385–406. issn: 0034-4877. doi: https://doi.org/10.1016/S0034-4877(12)60053-6. url: https://www.sciencedirect.com/science/article/pii/S0034487712600536 (cit. on p. 16).

[29] F. Mukhamedov and H. Akin. "Phase transitions for p-adic Potts model on the Cayley tree of order three". In: *Journal of Statistical Mechanics: Theory and Experiment* 2013.07 (2013), P07014 (cit. on p. 16).

[30] M. Farrukh and K. Otabek. "Phase transition and chaos: P-adic Potts model on a Cayley tree". In: *Chaos, Solitons & Fractals* 87 (2016), pp. 190–196. ISSN: 0960-0779. DOI: https://doi.org/10.1016/j.chaos.2016.04.003. URL: https://www.sciencedirect.com/science/article/pii/S096007791630128X (cit. on p. 16).

[31] O. Khakimov and F. Mukhamedov. "Chaotic behavior of the p-adic Potts–Bethe mapping II". In: *Ergodic Theory and Dynamical Systems* (2021), pp. 1–25 (cit. on p. 16).

[32] F. Mukhamedov. "On the existence of generalized Gibbs measures for the one-dimensional p-adic countable state Potts model". In: *Proceedings of the Steklov Institute of Mathematics* 265.1 (2009), pp. 165–176 (cit. on p. 16).

[33] O. N. Khakimov. "p-Adic Gibbs measures for the hard core model with three states on the Cayley tree". In: *Theoretical and Mathematical Physics* 177.1 (2013), pp. 1339–1351 (cit. on p. 16).

[34] S. Ludkowski and A. Khrennikov. "Stochastic processes on non-Archimedean spaces with values in non-Archimedean fields". In: *Markov Process. Related Fields* 9 (Nov. 2001) (cit. on p. 16).

[35] A. V. Antoniouk et al. "A stochastic p-adic model of the capillary flow in porous random medium". In: *Physica A: Statistical Mechanics and its Applications* 505 (2018), pp. 763–777. ISSN: 0378-4371. DOI: https://doi.org/10.1016/j.physa.2018.03.049. URL: https://www.sciencedirect.com/science/article/pii/S0378437118303704 (cit. on p. 16).

[36] I. V. Volovich. "Number theory as the ultimate physical theory". In: *P-Adic Numbers, Ultrametric Analysis, and Applications* 2.1 (2010), pp. 77–87 (cit. on p. 16).

[37] M. Naor and O. Reingold. "Number-Theoretic Constructions of Efficient Pseudo-Random Functions". In: *J. ACM* 51.2 (2004), 231–262. ISSN: 0004-5411. DOI: 10.1145/972639.972643. URL: https://doi.org/10.1145/972639.972643 (cit. on p. 16).

[38] J. Beck. *Inevitable randomness in discrete mathematics*. Vol. 49. American Mathematical Soc., 2009 (cit. on p. 16).

[39] T. Tao. *Structure and randomness: pages from year one of a mathematical blog*. American Mathematical Society Providence, RI, 2008 (cit. on p. 16).

[40] T. Tao. "Structure and Randomness in the Prime Numbers". In: *An Invitation to Mathematics: From Competitions to Research*. Ed. by D. Schleicher and M. Lackmann. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 1–7. ISBN: 978-3-642-19533-4. DOI: 10.1007/978-3-642-19533-4_1. URL: https://doi.org/10.1007/978-3-642-19533-4_1 (cit. on p. 16).

[41] P. Sarnak. "Randomness in Number Theory". In: *Asia Pacific Mathematics Newsletter* 2.3 (2012), pp. 15–19 (cit. on p. 16).

[42] M. M. Robert. "Simple mathematical model with very complicated dynamics". In: *Nature* 261.1 (1976), pp. 459–467 (cit. on pp. 16, 32).

[43] Stephen Wolfram. "Statistical mechanics of cellular automata". In: *Reviews of modern physics* 55.3 (1983), p. 601 (cit. on p. 16).

[44] M. V. Berry. "Riemann's Zeta function: A model for quantum chaos?" In: ed. by H Seligman T. H. and Nishioka. Springer Berlin Heidelberg, 1986, pp. 1–17. isbn: 978-3-540-47230-8 (cit. on p. 16).

[45] K. Irwin. "Toward the unification of physics and number theory". In: *Reports in Advances of Physical Sciences* 3.01 (2019) (cit. on p. 16).

[46] B. Julia. "Statistical theory of numbers". In: *Number theory and physics*. Springer, 1990, pp. 276–293 (cit. on p. 16).

[47] J.-M. Luck, P. Moussa, and M. Waldschmidt. *Number Theory and Physics*. Springer Berlin, Heidelberg, 1990. doi: https://doi.org/10.1007/978-3-642-75405-0. url: https://link.springer.com/book/10.1007/978-3-642-75405-0 (cit. on p. 16).

[48] I. Bakas and M. J. Bowick. "Curiosities of arithmetic gases". In: *Journal of mathematical physics* 32 (7 1991), pp. 1881–1884 (cit. on p. 16).

[49] F. Vericat. "A lattice gas of prime numbers and the Riemann Hypothesis". In: *Physica A: Statistical Mechanics and its Applications* 392 (19 2013), pp. 4516–4522. issn: 0378-4371. doi: https://doi.org/10.1016/j.physa.2013.05.049. url: https://www.sciencedirect.com/science/article/pii/S0378437113004913 (cit. on p. 16).

[50] H. H. Otto. "Phase Transitions Governed by the Fifth Power of the Golden Mean and Beyond". In: *World Journal of Condensed Matter Physics* 10.3 (2020), pp. 135–158 (cit. on p. 16).

[51] D. Schumayer and D. A. W. Hutchinson. "Colloquium: Physics of the Riemann hypothesis". In: *Reviews of Modern Physics* 83 (2 2011), p. 307 (cit. on p. 16).

[52] J. C. Andrade. "Hilbert–Pólya conjecture, zeta functions and bosonic quantum field theories". In: *International Journal of Modern Physics A* 28.17 (2013), p. 1350072 (cit. on pp. 16, 63).

[53] A. M. Odlyzko. *Correspondence about the origins of the Hilbert-Polya conjecture*. 2009 (cit. on p. 16).

[54] H. M. Edwards. *RiemannÆs zeta function*. Academic press, 1974 (cit. on pp. 16, 63, 85).

[55] J. C. Lagarias. "The 3x + 1 Problem and its Generalizations". In: *The American Mathematical Monthly* 92.1 (1985), pp. 3–23. doi: 10.1080/00029890.1985.11971528. url: https://doi.org/10.1080/00029890.1985.11971528 (cit. on pp. 19–23, 26, 28).

[56] C. A. Pickover. *Wonders of Numbers: Adventures in Mathematics, Mind, and Meaning*. Oxford University Press, 2003 (cit. on p. 19).

[57] J. C. Lagarias. *The Ultimate Challenge: The 3x+1 Problem*. American Mathematical Society, 2010 (cit. on pp. 19–21, 25, 45).

[58] M. Chamberland. "A 3x+ 1 survey: Number theory and dynamical systems". In: *The ultimate challenge: the 3x+1 problem*. Vol. 1. 2010, pp. 57–78 (cit. on pp. 20, 22, 26).

[59] Y. Sinai. "Statistical (3x+ 1)–problem". In: *arXiv preprint math/0201102* (2002) (cit. on pp. 20, 31).

[60] A. Kontorovich and Y. Sinai. "Structure Theorem for (d, g, h)-Maps". In: *Bulletin of the Brazilian Mathematical Society* 33 (2002), pp. 213–224 (cit. on pp. 20, 28, 31).

[61] C. Williams et al. "Ulam's Conjecture Continued Again, 1000 Pounds for Proof". In: *PPC Calculator Journal* 9 (1982), pp. 23–24 (cit. on p. 21).

[62] H. Hasse. *Unsolved Problems in Elementary Number Theory*. Lectures at University of Maine (Orono). Mimeographed notes. 1975 (cit. on p. 21).

[63] S. Kakutani. Private Communication to J. Lagarias. 1981 (cit. on p. 21).

[64] C. Böhm and G. Sontacchi. "On the existence of cycles of given length in integer sequences like $x_{n+1} = x_n/2$ if x_n even, and $x_{n+1} = 3x_n + 1$". eng. In: *Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti* 64.3 (Mar. 1978), pp. 260–264. URL: http://eudml.org/doc/290184 (cit. on p. 21).

[65] D. Barina. "Convergence verification of the Collatz problem". In: *The Journal of Supercomputing* 77.3 (2021), pp. 2681–2688 (cit. on p. 21).

[66] T. Tao. "Almost all orbits of the Collatz map attain almost bounded values". In: *arXiv preprint arXiv:1909.03562* (2019) (cit. on p. 21).

[67] T. Tao. "Almost all orbits of the Collatz map attain almost bounded values". In: *Forum of Mathematics, Pi*. Vol. 10. Cambridge University Press. 2022 (cit. on pp. 21, 28, 45).

[68] F. Izadi. "Complete Proof of the Collatz Conjecture". In: *arXiv preprint arXiv:2101.06107* (2021) (cit. on p. 21).

[69] J. Llibre and C. Valls. "A proof of the $3x + 1$ conjecture". In: *arXiv preprint arXiv:2110.12228* (2021) (cit. on p. 21).

[70] R. Guy. *Unsolved problems in number theory*. Vol. 1. Springer Science & Business Media, 2004 (cit. on p. 22).

[71] R. Terras. "A stopping time problem on the positive integers". In: *Acta Arithmetica* 30.3 (1976), pp. 241–252. URL: http://eudml.org/doc/205476 (cit. on pp. 23, 28).

[72] R. Terras. "On the existence of a density". In: *Acta Arithmetica* 35.1 (1979), pp. 101–102 (cit. on p. 23).

[73] M. V. P. Garcia and F. A. Tal. "A note on the generalized 3n+ 1 problem". In: *Acta Arithmetica* 90.3 (1999), pp. 245–250 (cit. on p. 23).

[74] R. E. Crandall. "On the "3x+ 1" problem". In: *Mathematics of computation* 32.144 (1978), pp. 1281–1292 (cit. on p. 24).

[75] G. J. Wirsching. *The dynamical system generated by the 3n+ 1 function*. Springer, 2006 (cit. on pp. 25, 27).

[76] B. G. Seifert. "On the arithmetic of cycles for the Collatz-Hasse ('Syracuse') conjectures". In: *Discrete mathematics* 68.2-3 (1988), pp. 293–298 (cit. on p. 26).

[77] J. C. Lagarias. "The set of rational cycles for the 3x+ 1 problem". In: *Acta Arithmetica* 56.1 (1990), pp. 33–53 (cit. on p. 26).

[78] J. Riordan. *Introduction to combinatorial analysis*. Courier Corporation, 2012 (cit. on p. 26).

[79] M. Chamberland. "A continuous extension of the 3x+ 1 problem to the real line". In: *Dynamics of Continuous, Discrete and Impulsive Systems* 2.4 (1996), pp. 495–509 (cit. on p. 27).

[80] K. T. Alligood et al. "Chaos: an introduction to dynamical systems". In: *SIAM Review* 40.3 (1998), pp. 732–732 (cit. on p. 27).

[81] A. V. Kontorovich and J. C. Lagarias. *Stochastic Models for the 3x+1 and 5x+1 Problems*. 2009. arXiv: 0910.1944 [math.NT] (cit. on p. 28).

[82] S. M. Ross. "10 - Brownian Motion and Stationary Processes". In: *Introduction to Probability Models (Twelfth Edition)*. Ed. by Sheldon M. Ross. Twelfth Edition. Academic Press, 2019, pp. 639–677. isbn: 978-0-12-814346-9. doi: https://doi.org/10.1016/B978-0-12-814346-9.00015-9. url: https://www.sciencedirect.com/science/article/pii/B9780128143469000159 (cit. on pp. 31, 98).

[83] C. A. Braumann. "Population growth in random environments". In: *Bulletin of mathematical biology* 45.4 (1983), pp. 635–641 (cit. on p. 31).

[84] V. Stojkoski et al. "Geometric Brownian motion under stochastic resetting: A stationary yet nonergodic process". In: *Phys. Rev. E* 104 (1 2021), p. 014121. doi: 10.1103/PhysRevE.104.014121. url: https://link.aps.org/doi/10.1103/PhysRevE.104.014121 (cit. on p. 31).

[85] V. Stojkoski and M. Karbevski. "Ergodicity breaking in wealth dynamics: The case of reallocating geometric Brownian motion". In: *Physical Review E* 105.2 (2022), p. 024107 (cit. on p. 31).

[86] F. Black and M. Scholes. "The Pricing of Options and Corporate Liabilities". In: *Journal of Political Economy* 81.3 (1973), pp. 637–654. issn: 00223808, 1537534X. url: http://www.jstor.org/stable/1831029 (visited on 07/20/2022) (cit. on p. 31).

[87] E. Limpert, W. A. Stahel, and M. Abbt. "Log-normal Distributions across the Sciences: Keys and Clues: On the charms of statistics, and how mechanical models resembling gambling machines offer a link to a handy way to characterize log-normal distributions, which can provide deeper insight into variability and probability—normal or log-normal: That is the question". In: *BioScience* 51.5 (May 2001), pp. 341–352. issn: 0006-3568. doi: 10.1641/0006-3568(2001)051[0341:LNDATS]2.0.CO;2. eprint: https://academic.oup.com/bioscience/article-pdf/51/5/341/26891292/51-5-341.pdf. url: https://doi.org/10.1641/0006-3568(2001)051[0341:LNDATS]2.0.CO;2 (cit. on p. 31).

[88] M. Casartelli. "Intermittency from Collatz's itineraries and complexity indicators". In: *Journal of Physics A: Mathematical and General* 35.21 (2002), pp. 4501–4514. doi: 10.1088/0305-4470/35/21/301. url: https://doi.org/10.1088/0305-4470/35/21/301 (cit. on pp. 32, 33).

[89] I. P. Cornfeld and Y. G. Sinai. "Basic Notions of Ergodic Theory and Examples of Dynamical Systems". In: *Dynamical Systems II: Ergodic Theory with Applications to Dynamical Systems and Statistical Mechanics*. Ed. by Y. G. Sinai. Berlin, Heidelberg: Springer Berlin Heidelberg, 1989, pp. 2–27. ISBN: 978-3-662-06788-8. DOI: 10.1007/978-3-662-06788-8_1. URL: https://doi.org/10.1007/978-3-662-06788-8_1 (cit. on p. 32).

[90] Y. B. Pesin. "Characteristic Lyapunov Exponents and Smooth Ergodic Theory". In: *Russian Mathematical Surveys* 32.4 (1977), pp. 55–114. DOI: 10.1070/rm1977v032n04abeh001639. URL: https://doi.org/10.1070/rm1977v032n04abeh001639 (cit. on p. 33).

[91] Y. Grigor'evič Sinai and Y. G. Sinai. *Dynamical systems II: Ergodic theory with applications to dynamical systems and statistical mechanics*. Springer, 1989 (cit. on p. 33).

[92] M. G. E. da Luz et al. "Scale-free behavior in hailstone sequences generated by the Collatz map". In: *Physical Review Research* 3.1 (2021), p. 13073. DOI: 10.1103/physrevresearch.3.013073. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.3.013073 (cit. on pp. 33–35, 45, 65, 77, 78).

[93] P. Bak, C. Tang, and K. Wiesenfeld. "Self-organized criticality: An explanation of the 1/f noise". In: *Phys. Rev. Lett.* 59.4 (1987), pp. 381–384. DOI: 10.1103/PhysRevLett.59.381. URL: https://link.aps.org/doi/10.1103/PhysRevLett.59.381 (cit. on p. 33).

[94] R. Badii. "Complexity: hierarchical structures and scaling in physics". In: 2 (1999) (cit. on p. 35).

[95] D. Sornette. *Critical phenomena in natural sciences: chaos, fractals, selforganization and disorder: concepts and tools*. Springer Science & Business Media, 2006 (cit. on pp. 35, 41, 43, 45).

[96] C. C. Perelman and R. Carbo-Dorca. "The Collatz conjecture and the quantum mechanical harmonic oscillator". In: *Journal of Mathematical Chemistry* 60.1 (2022), pp. 145–160 (cit. on pp. 35–37, 85).

[97] E. W. Dijkstra. "An Interview with Professor Dr. Edsger W. Dijkstra". In: *by Cashman, MW, Datamation* 23.95 (1977), pp. 164–166 (cit. on p. 39).

[98] N. Koblitz. *p-adic Numbers, p-adic Analysis, and Zeta-Functions*. Vol. 58. Springer Science & Business Media, 2012 (cit. on pp. 40, 41).

[99] I. T. Adamson. *Introduction to field theory*. Courier Corporation, 2007 (cit. on p. 40).

[100] S. Albeverio, A. Y. Khrennikov, and V. M. Shelkovich. *Theory of p-adic distributions: linear and nonlinear models*. 370. Cambridge University Press, 2010 (cit. on pp. 42, 43).

[101] R. Rammal, G. Toulouse, and M. A. Virasoro. "Ultrametricity for physicists". In: *Reviews of Modern Physics* 58.3 (1986), p. 765 (cit. on p. 42).

[102] P. Dutta, D. Ghoshal, and A. Lala. "On the Exchange Interactions in Holographic p-adic CFT". In: *Physics Letters B* 773 (May 2017). DOI: 10.1016/j.physletb.2017.08.042 (cit. on p. 42).

[103] B. A. Zambrano-Luna and W. A. Zuniga-Galindo. "p-adic Cellular Neural Networks". In: *Journal of Nonlinear Mathematical Physics* (2022), pp. 1–37 (cit. on p. 42).

[104] B. Dragovich et al. "p-Adic mathematics and theoretical biology". In: *Biosystems* 199 (2021), p. 104288 (cit. on p. 42).

[105] S. V. Kozyrev. "Ultrametricity in the theory of complex systems". In: *Theoretical and Mathematical Physics* 185.2 (2015), pp. 1665–1677 (cit. on p. 44).

[106] P. Bak and K. Sneppen. "Punctuated equilibrium and criticality in a simple model of evolution". In: *Physical review letters* 71.24 (1993), p. 4083 (cit. on p. 44).

[107] G. E. P. Box et al. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015 (cit. on pp. 46, 47, 50).

[108] W. A. Woyczynski and W. A. Woyczyânski. *A first course in statistics for signal analysis*. Springer, 2011 (cit. on pp. 47, 48, 50).

[109] T. A. Severini. *Elements of distribution theory*. Vol. 17. Cambridge University Press, 2005 (cit. on p. 48).

[110] M. Peter. *Cumulants*. Classnotes on author's web page. 2013. URL: http://www.stat.uchicago.edu/~pmcc/courses/stat306/2013/cumulants.pdf (cit. on p. 48).

[111] G. M. Viswanathan et al. *The physics of foraging: an introduction to random searches and biological encounters*. Cambridge University Press, 2011 (cit. on pp. 49, 51).

[112] S. W. McKechnie, P. R. Ehrlich, and R. R. White. "Population genetics of Euphydryas butterflies. I. Genetic variation and the neutrality hypothesis". In: *Genetics* 81.3 (1975), 571A–594 (cit. on p. 49).

[113] C. F. W. Higham, A. Kijngam, and B. F. J. Manly. "An analysis of prehistoric canid remains from Thailand". In: *Journal of Archaeological Science* 7.2 (1980), pp. 149–165 (cit. on p. 49).

[114] J. P. Stevens. *Applied multivariate statistics for the social sciences*. Routledge, 2012 (cit. on p. 49).

[115] B. F. J. Manly and J. A. N. Alberto. *Multivariate statistical methods: a primer*. Chapman and Hall/CRC, 2016 (cit. on p. 49).

[116] R. A. Johnson and et. al. Wichern D. W. "Applied multivariate statistical analysis". In: *New Jersey* 405 (1992) (cit. on p. 49).

[117] J. Benesty et al. "Pearson correlation coefficient". In: *Noise reduction in speech processing*. Springer, 2009, pp. 1–4 (cit. on p. 49).

[118] Y. Liu et al. "Disrupted small-world networks in schizophrenia". In: *Brain* 131.4 (Feb. 2008), pp. 945–961. ISSN: 0006-8950. DOI: 10.1093/brain/awn018. eprint: https://academic.oup.com/brain/article-pdf/131/4/945/16694858/awn018.pdf. URL: https://doi.org/10.1093/brain/awn018 (cit. on p. 50).

[119] E. Butkov. *Mathematical physics*. Addison-Wesley, 1968 (cit. on p. 50).

[120] G. B. Arfken and H. J. Weber. *Mathematical methods for physicists*. 1999 (cit. on p. 50).

[121] J. W. Cooley and J. W. Tukey. "An algorithm for the machine calculation of complex Fourier series". In: *Mathematics of computation* 19.90 (1965), pp. 297–301 (cit. on p. 50).

[122] M. B. Priestley. "Power spectral analysis of non-stationary random processes". In: *Journal of Sound and Vibration* 6.1 (1967), pp. 86–97. ISSN: 0022-460X. DOI: https://doi.org/10.1016/0022-460X(67)90160-5. URL: https://www.sciencedirect.com/science/article/pii/0022460X67901605 (cit. on p. 51).

[123] P. Szendro, G. Vincze, and A. Szasz. "Pink-noise behaviour of biosystems". In: *European Biophysics Journal* 30.3 (2001), pp. 227–231 (cit. on p. 51).

[124] H. Kantz and T. Schreiber. *Nonlinear time series analysis*. Vol. 7. Cambridge university press, 2004 (cit. on p. 51).

[125] C.-K. Peng et al. "Mosaic organization of DNA nucleotides". In: *Phys. Rev. E* 49.2 (1994), pp. 1685–1689. DOI: 10.1103/PhysRevE.49.1685. URL: https://link.aps.org/doi/10.1103/PhysRevE.49.1685 (cit. on p. 51).

[126] C.-K. Peng et al. "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series". In: *Chaos: an interdisciplinary journal of nonlinear science* 5.1 (1995), pp. 82–87 (cit. on p. 51).

[127] A. L. Goldberger et al. "Fractal dynamics in physiology: alterations with disease and aging". In: *Proceedings of the national academy of sciences* 99.suppl_1 (2002), pp. 2466–2472 (cit. on p. 51).

[128] T. Lux and M. Marchesi. "Scaling and criticality in a stochastic multi-agent model of a financial market". In: *Nature* 397.6719 (1999), pp. 498–500 (cit. on p. 51).

[129] Y. Ashkenazy et al. "Nonlinearity and multifractality of climate change in the past 420,000 years". In: *Geophysical research letters* 30.22 (2003) (cit. on p. 51).

[130] C. Heneghan and G. McDarby. "Establishing the relation between detrended fluctuation analysis and power spectral density analysis for stochastic processes". In: *Physical review E* 62.5 (2000), p. 6103 (cit. on p. 52).

[131] S. R. A. Salinas. *Introdução à Física Estatística*. second. EDUsp, 1999 (cit. on pp. 52, 97).

[132] M. T. C. A. J. Thomas and A. T. Joy. *Elements of information theory*. Wiley-Interscience, 2006 (cit. on p. 52).

[133] J. J. Sakurai and J. Napolitano. *Modern quantum mechanics. 2-nd edition*. 2014, p. 39 (cit. on p. 52).

[134] Helcio Felippe et al. "Threshold-free estimation of entropy from a Pearson matrix". In: *Europhysics Letters* (2023) (cit. on pp. 52, 53).

[135] F. Helcio. "Pearson matrices as density operators: A test of the entropic brain hypothesis using the von Neumann entropy". MA thesis. Natal: University of Rio Grande do Norte, 2021 (cit. on pp. 52, 53).

[136] W. Christoph, P. Steffen, and H. Martin. "Factorising numbers with a Bose–Einstein condensate". In: *Physica A: Statistical Mechanics and its Applications* 341 (2004), pp. 586–606. ISSN: 0378-4371. DOI: https://doi.org/10.1016/j.physa.2004.05.047. URL: https://www.sciencedirect.com/science/article/pii/S0378437104007162 (cit. on p. 64).

[137] J. G. Dueñas and N. F. Svaiter. "Thermodynamics of the bosonic randomized Riemann gas". In: *Journal of Physics A: Mathematical and Theoretical* 48 (31 July 2015), p. 315201. DOI: 10.1088/1751-8113/48/31/315201. URL: https://doi.org/10.1088/1751-8113/48/31/315201 (cit. on p. 64).

[138] T. Tomé. *Dinâmica Estocástica e Irreversibilidade Vol. 35*. Edusp, 2001 (cit. on pp. 66, 98).

[139] Massimo Cencini and Angelo Vulpiani. "Finite size Lyapunov exponent: review on applications". In: *Journal of Physics A: Mathematical and Theoretical* 46.25 (2013), p. 254019 (cit. on p. 69).

[140] J. Bezanson et al. "Julia: A fresh approach to numerical computing". In: *SIAM Review* 59.1 (2017), pp. 65–98. DOI: 10.1137/141000671. URL: https://epubs.siam.org/doi/10.1137/141000671 (cit. on p. 71).

[141] M. Frigo and S. G. Johnson. "The Design and Implementation of FFTW3". In: *Proceedings of the IEEE* 93.2 (2005). Special issue on "Program Generation, Optimization, and Platform Adaptation", pp. 216–231. DOI: 10.1109/JPROC.2004.840301 (cit. on p. 71).

[142] M. D. Maas. *Using the FFTW Library in julia*. 2022. URL: https://www.matecdev.com/posts/julia-fft.html (cit. on p. 71).

[143] T. Editors of Encyclopaedia Britannica. *Robert Brown*. 2021. URL: https://www.britannica.com/biography/Robert-Brown-Scottish-botanist (cit. on p. 97).

[144] R. Brown. "XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies". In: *The Philosophical Magazine* 4.21 (1828), pp. 161–173. DOI: 10.1080/14786442808674769. eprint: https://doi.org/10.1080/14786442808674769. URL: https://doi.org/10.1080/14786442808674769 (cit. on p. 97).

[145] E. Nelson. *Dynamical Theories of Brownian Motion*. 2nd. Priceton University Press, 2001 (cit. on p. 97).

Appendix A

Brownian and Geometric Brownian Motion (GBM)

Brownian Motion receives its name from Robert Brown (1773–1858), Scottish Botanic and Physicist [143]. Brown was not the discoverer of the movement, once in the 19th Century, anyone with a microscope could detect the random movement of small particles in water. Nonetheless, Brown's description of the movement discarded other theories, such as those that assigned random movement to life, while studying the movement of flowers' pollen and inorganic particles [144]. His theory about movement conjectured that matter is composed of small particles, displaying fast and irregular movement originating in the particles themselves [145].

From the 19th Century, until 1905, the Kinetic Movement Theory has shown itself the most plausible explanation for the particles' movements. However, until that time, it was very hard to define velocity for particles in Brownian Motion and compare it with the theoretical and experimental energy equipartition theorem. Einstein's theory of Brownian Motion was successful in establishing this relationship.

Einstein's argumentation about the Brownian Motion involves considering a probability density $\rho(x,t)$ that a particle is in the position x at time t. Hence, Einstein finds the diffusion equation

$$\frac{\partial \rho}{\partial t} = D\frac{\partial^2 \rho}{\partial x^2} \tag{A.1}$$

where D is the diffusion coefficient of the movement, related to the probability of finding a particle around its initial position in a posterior instant of time. Considering that the particle is at $x = 0$ when $t = 0$, the solution for $\rho(x,t)$ is found as

$$\rho(x,t) = \frac{1}{(4\pi D t)^{3/2}} \exp\left(-\frac{|x|^2}{4Dt}\right) \tag{A.2}$$

with $|x|$ the distance between the particle and the origin.

Einstein's work is very important because it provides a quantitative measure for diffusion, D, allowing great advances in the atomic structure of matter theory [131].

A.1 Statistical Measures of Brownian Motion

The Brownian Motion presents specific characteristics, that can be found in variables following markovian processes [138], such as symmetric *random walks* with equal probability of unitary step in each direction $P_{i,i+1} = P_{i,i-1} = 1/2$. Let $X(t)$ a variable following the symmetric *random walk* describer above, with steps of size Δx given at every Δt time, so

$$X(t) = \Delta x \left(\sum_{i=1}^{t/\Delta t} X_i \right), \tag{A.3}$$

where $X_i = \pm 1$.

In this process, once $\langle X_i \rangle = 0, \text{var}(X_i) = \langle X_i^2 \rangle = 1$ then

$$\langle X(t) \rangle = 0, \tag{A.4}$$

$$\langle (\Delta X(t))^2 \rangle = (\Delta x)^2 \left[\frac{t}{\Delta t} \right], \tag{A.5}$$

and if one lets Δx and Δt go to 0 such that $\Delta x = \sigma\sqrt{\Delta t}$, one finds the stochastic process called Brownian Motion where the variance $\langle (\Delta X(t))^2 = \sigma^2 t$. Ross [82] defines the Brownian Motion as follows.

Definition A.1. *A stochastic process* $\{X(t), t \geq 0\}$ *is said to be a Brownian Motion if*

1. $X(0) = 0$;
2. $\{X(t), t \geq 0\}$ *has stationary and independent increments;*
3. *for every* $t > 0$, $X(t)$ *follows the normal distribution with mean* 0 *and variance* $\sigma^2 t$.

From its definition, one concludes that the density function of $X(t)$ is

$$f_t(x) = \frac{1}{\sqrt{2\pi t}} \exp\left(-x^2/2t\right), \tag{A.6}$$

which has the same behavior from Equation (A.2) obtained by Einstein for the particle's density under diffusion with $D = 1$, called standard Brownian Motion.

According to reference [82], one says that $\{X(t), t \geq 0\}$, with drift coefficient μ and variance $\sigma^2 t$, is a *Brownian motion with Drift* if

$$X(t) = \sigma B(t) + \mu t, \tag{A.7}$$

for $B(t)$ a variable under standard Brownian Motion.

A.2 Geometric Brownian Motion (GBM)

If $\{Y(t), t \geq 0\}$ is a continuous variable in a Brownian Motion process with a drift coefficient μ and variance σ^2, then a process $\{X(t), t \geq 0\}$ defined by

$$X(t) = \exp\left(Y(t)\right), \tag{A.8}$$

is called *Geometric Brownian Motion* (GBM) with drift [82].

For a discrete variable X_n written as

$$X_n = r_n r_{n-1} \cdots r_1 r_0 X_0, \tag{A.9}$$

with r_i iid, then, taking $\ln(X_n)$ one finds

$$\ln(X_n) = \sum_{i=1}^{n} \ln(r_i) + \ln(X_0). \tag{A.10}$$

Once the steps $\ln(r_i)$ are iid, then $Y_n = \ln(X_n)$, when compared to Equation (A.3) will come closer to a Brownian Motion with Drift for an appropriate normalization — equivalent to $\Delta t \to 0$ and $\Delta x = \sigma\sqrt{\Delta t}$ —, hence

$$X_n = \exp[\ln(X_n)] = \exp[Y_n], \tag{A.11}$$

will become a Geometric Brownian Motion.

Appendix B

Proofs on Ultrametric Spaces

B.1 Proof of Ultrametric Inequality Lemma 3.1.4

Proof. We want to prove the if-and-only-if relation between non-Archimedean absolute value and the ultrametric inequality, so we must go both ways.

To go one way, from non-Archimedean absolute value to ultrametric inequality, applying the non-archimedean property to

$$d(x,y) = |x-y| = |(x-z)+(z-y)| \tag{B.1}$$

one finds

$$\begin{aligned} d(x,y) &= |(x-z)+(z-y)| \leq \max\{|x-z|,|z-y|\} \\ d(x,y) &\leq \max\{d(x,z),d(z,y)\}\,. \end{aligned} \tag{B.2}$$

And the metric respects the ultrametric inequality.

The converse, from ultrametric inequality to non-Archimedean absolute value, is found by taking $y=-y_1, z=0$ in Equation (B.2), leading to

$$\begin{aligned} d(x,-y_1) = |x+y_1| &\leq \max\{d(x,0),d(0,-y_1)\} = \max\{|x-0|,|0,-(-y_1)|\} \\ |x+y_1| &\leq \max\{|x|,|y_1|\}\,, \end{aligned} \tag{B.3}$$

and the space has non-archimedean absolute value. □

B.2 Proof that the space $(\mathbb{Z}_p, |\ |_p)$ is ultrametric (Lemma 3.1.6)

First of all, let us define the + operation for this set as in [18]. The sum of $a=\sum_{i\geq 0} a_i p^i$ and $b=\sum_{i\geq 0} b_i p^i$ is defined component-wise

$$(a+b)_i = \begin{cases} a_i+b_i & \text{if } a_i+b_i < p, \\ a_i+b_i-p & \text{if } a_i+b_i \geq p, \end{cases} \tag{B.4}$$

and the second case adds a *carry* to the next component.

The proof consists on showing that the absolute value $|\ |_p$ on Equation (3.4), rewritten here for $x \neq 0$

$$|x|_p = \frac{1}{p^{v_p(x)}}, \tag{B.5}$$

induces the metric $d(x,y)$ from Equation (3.5)

$$d(x,y) = \frac{1}{p^{v_p(x-y)}}, \tag{B.6}$$

and follows the ultrametric equality from Lemma 3.1.4

$$d(x,y) \leq \max\left\{d(x,z),\ d(z,y)\right\}. \tag{B.7}$$

This is done by demonstrating that the absolute value $|\ |_p$ is non-Archimedean on $\mathbb{Z}_p$, i.e., it obeys the strong triangle inequality at Equation (3.1)

$$|x+y|_p \leq \max\left(|x|_p, |y|_p\right), \tag{B.8}$$

for all $x, y \in \mathbb{Z}_p$.

Proof. To prove that any $x, y \in \mathbb{Z}_p$ obeys the strong inequality for the metric $|\ |_p$, lets write

$$x = a_{v_p(x)} p^{v_p(x)} + \sum_{i > v_p(x)} a_i p^i, \tag{B.9}$$

$$y = a_{v_p(y)} p^{v_p(y)} + \sum_{j > v_p(y)} a_j p^j. \tag{B.10}$$

Then, the sum $x + y$ is

$$x + y = a_{v_p(x)} p^{v_p(x)} + a_{v_p(y)} p^{v_p(y)} + \sum_{i > v_p(x)} a_i p^i + \sum_{j > v_p(y)} a_j p^j, \tag{B.11}$$

$$= a_{v_p(x+y)} p^{v_p(x+y)} + \sum_{k > v_p(x+y)} a_k p^k \tag{B.12}$$

where $v_p(x+y)$, i.e., the order of $x + y$ must be

$$v_p(x+y) \geq \min\left(v_p(x), v_p(y)\right). \tag{B.13}$$

The equal sign comes from when $v_p(x) \neq v_p(y)$, or $v_p(x) = v_p(y)$ with $a_{v_p(x)} + a_{v_p(y)} \neq p$, then $v_p(x+y)$ must be *equal* to the minimum between both. The $>$ sign comes from when $v_p(x) = v_p(y)$ and $a_{v_p(x)} + a_{v_p(y)} = p$ (case when the $v_p(x)$ component vanishes), then $v_p(x+y)$ must be *larger* than $v_p(x)$ and $v_p(y)$.

Now, by using Inequality (B.13), the absolute value of the sum $|x+y|$ becomes

$$|x+y|_p = \frac{1}{p^{v_p(x+y)}} = p^{-v_p(x+y)}, \tag{B.14}$$

then

$$-v_p(x+y) = \log_p(|x+y|_p). \tag{B.15}$$

And once $-v_p(x+y) \leq \max(-v_p(x), -v_p(y))$, one finds

$$\log_p(|x+y|_p) \leq \max(-v_p(x), -v_p(y)), \tag{B.16}$$
$$p^{-v_p(x+y)} \leq p^{\max(-v_p(x), -v_p(y))}, \tag{B.17}$$

and once $v_p(x)$ and $v_p(y)$ are both non-negative integers, then

$$p^{-v_p(x+y)} \leq \max(p^{-v_p(x)}, p^{v_p(y)}), \tag{B.18}$$

and one finds

$$|x+y|_p \leq \max(|x|_p, |y|_p), \tag{B.19}$$

the strong triangle inequality. Hence $|\ |_p$ is non-Archimedean on $\mathbb{Z}_p$, and the space $(\mathbb{Z}_p, |\ |_p)$ is ultrametric □

Appendix C

Fast Fourier Transformation (FFT) Method

The present Appendix presents details of a tool for performing the Fast Fourier Transformation method for discrete time series. Given a discrete time series $\boldsymbol{x} = \{x_0, x_2, \ldots, x_{T-1}\}$, its *Discrete Fourier Transform* (DFT) is given by the set of $\hat{x}_f$ such that

$$\hat{x}_f = \frac{1}{T}\sum_{t=0}^{T-1} x_t e^{-2\pi i t f/T}.$$

That is, the sequence

$$\hat{\boldsymbol{x}} = \{\hat{x}_0, \ldots, \hat{x}_{T-1}\}, \tag{C.1}$$

is the DFT of the signal $\boldsymbol{x}$.

Notice that for the computation of each DFT term, one has to perform T multiplications, and there are T terms to be computed. Consequently, the Computational Complexity (CC) of the entire $\hat{\boldsymbol{x}}$ calculation is $CC_{\hat{\boldsymbol{x}}} = T^2$. In 1965, Cooley and Turkey popularized an algorithm called *Fast Fourier Transform (FFT)*, that cleverly groups terms in the sum. This algorithm was first discovered by Carl Friedrich Gauss at the beginning of the 19th century.

For the specific case where $T = 2^n$, let $\omega_T = e^{-2\pi i/T}$ be a complex Tth root of unity. For $M = \frac{T}{2}$, first defining some important relations

$$(\omega_{2M})^{2tf} = \left(e^{-2\pi i/2M}\right)^{2tf} = \left(e^{-2\pi i/M}\right)^{tf} = (\omega_M)^{tf}, \tag{C.2}$$

$$(\omega_M)^{M+f} = \underbrace{\left(e^{-2\pi i/M}\right)^{M}}_{1}\left(e^{-2\pi i/M}\right)^{f} = (\omega_M)^{f}, \tag{C.3}$$

$$(\omega_{2M})^{M+f} = \underbrace{\left(e^{-2\pi i/2M}\right)^{M}}_{-1}\left(e^{-2\pi i/2M}\right)^{f} = \underbrace{-e^{-\pi i f/M}}_{\omega_{2M}^{f}} = (\omega_{2M})^{f} \tag{C.4}$$

With these relations above, Equation (C.1) can be written as

$$\begin{aligned}
\hat{x}_f &= \frac{1}{2M}\sum_{t=0}^{2M-1} x_t e^{-\pi i t f/M} \\
&= \frac{1}{2}\left(\frac{1}{M}\sum_{t=0}^{M-1} x_{2t} e^{-2\pi i t f/M} + \frac{1}{M}\sum_{t=0}^{M-1} x_{2t+1} e^{-2\pi i t f/M - \pi i f/M}\right), \\
&= \frac{1}{2}\left(\underbrace{\frac{1}{M}\sum_{t=0}^{M-1} x_{2t}\omega_M^{tf}}_{\hat{x}_f^{\text{even}}} + \underbrace{\frac{1}{M}\sum_{t=0}^{M-1} x_{2t+1}\omega_M^{tf}}_{\hat{x}_f^{\text{odd}}} \cdot \omega_{2M}^{f}\right), \qquad \text{(C.5)}
\end{aligned}$$

or

$$\hat{x}_f = \frac{1}{2}\left(\hat{x}_f^{\text{even}} + \hat{x}_f^{\text{odd}} \cdot \omega_{2M}^{f}\right), \qquad \text{(C.6)}$$

And with this relation, the following can be obtained

$$\begin{aligned}
\hat{x}_{f+M} &= \frac{1}{2}\left(\frac{1}{M}\sum_{t=0}^{M+1} x_{2t}\underbrace{\omega_M^{f+M}}_{\omega_M^f} + \sum_{t=0}^{M+1} x_{2t+1} e^{\pi i (f+M)(2t+1)/M}\right), \qquad \text{(C.7)} \\
&= \frac{1}{2}\left(\frac{1}{M}\sum_{t=0}^{M+1} x_{2t}\omega_M^{f} + \sum_{t=0}^{M+1} x_{2t+1} e^{-2\pi i (f+M)t/M}\underbrace{\omega_{2M}^{f+M}}_{-\omega_{2M}^f}\right), \qquad \text{(C.8)} \\
&= \frac{1}{2}\left(\frac{1}{M}\sum_{t=0}^{M+1} x_{2t}\omega_M^{f} + \sum_{t=0}^{M+1} x_{2t+1}\omega_M^{tf}\underbrace{e^{-2\pi i t}}_{1}(-\omega_{2M}^f)\right), \qquad \text{(C.9)} \\
\hat{x}_{f+M} &= \frac{1}{2}\left(\hat{x}_f^{\text{even}} - \hat{x}_f^{\text{odd}} \cdot \omega_{2M}^{f}\right). \qquad \text{(C.10)}
\end{aligned}$$

Hence, the $\hat{x}_f$ terms can be calculated only for $f = 0, 1, 2, \ldots, M-1 = T/2-1$, and all the terms $f = M, M+1, \ldots, T$ can be calculated by relation (C.10). This procces can be done recursivelly for $M = T/2^n$, reducing the Computational Complexity from $CC_{\hat{\boldsymbol{x}}} = T^2 = 2^{2n}$ as referred before, to $CC_{FFT(\hat{\boldsymbol{x}})} = \frac{T}{2}\log_2 T = n\,2^{n-1}$.

Appendix D

On details and examples for Equation (4.3)

In order to show that Equation (4.3) hold true, first, apply base-2 logarithm to Equation (3.9), one finds

$$\begin{aligned}\log_2 n &= \log_2\left[2^{m_1}(2^{m_2}(2^{m_3}(\cdots(2^{m_{r-1}}(2^{m_r}+1)+1)\ldots)+1)+1)\right] \\ &= m_1+\log_2\left[2^{m_2}(\cdots(2^{m_{r-1}}(2^{m_r}+1)+1)\cdots+1)+1\right] \end{aligned} \tag{D.1}$$

the argument of logarithm in Equation (D.1) can be written as $2^{m_2}\xi_{3,t}+1$ for

$$\xi_{3,t} = 2^{m_3}(2^{m_4}(\cdots 2^{m_{r-1}}(2^{m_r}+1)\cdots+1)+1). \tag{D.2}$$

This leads to

$$\begin{aligned}\log_2 n &= m_1+\log_2\left[2^{m_2}\xi_{3,t}+2^0\right] \\ &= m_1+\log_2\left[2^{m_2}(\xi_{3,t}+2^{-m_2})\right] = m_1+m_2+\log_2\left[\xi_{3,t}+2^{-m_2}\right] \end{aligned} \tag{D.3}$$

by substituting $\xi_{3,t}$ from Equation (D.2)

$$\log_2 n = m_1+m_2+\log_2\left[2^{m_3}(2^{m_4}(\cdots 2^{m_{r-1}}(2^{m_r}+1)\cdots+1)+1)+2^{-m_2}\right], \tag{D.4}$$

one can proceed in the same manner by setting $\xi_{4,t} = 2^{m_4}(\cdots 2^{m_{r-1}}(2^{m_r}+1)\cdots+1)$ and writing the logarithm argument as $2^{m_3}\xi_{4,t}+1$ leads to

$$\begin{aligned}\log_2 n &= m_1+m_2+\log_2\left[2^{m_3}\xi_{4,t}+2^0+2^{-m_2}\right] \\ &= m_1+m_2+m_3+\log_2\left[\xi_{4,t}+2^{-m_3}+2^{-m_2-m_3}\right]. \end{aligned} \tag{D.5}$$

This process can be carried over until reach

$$\begin{aligned}\log_2 n &= m_1+m_2+\ldots+m_{r-1}+\log_2\left[(2^{m_r}+1)+2^{-m_{r-1}}+\cdots+2^{-m_2-\cdots-m_{r-1}}\right] \\ &= m_1+m_2+\ldots+m_{r-1}+m_r+\log_2\left[1+2^{-m_r}+\cdots+2^{-m_2-\cdots+m_r}\right]. \end{aligned} \tag{D.6}$$

It is possible to see that the sum of powers of two inside the logarithm assume the largest value when $m_2 = m_3 = \ldots = m_{r-1} = m_r = 1$, i.e., the exponents are the smallest. In this

case, one finds

$$2^{-m_r} + 2^{-m_r - m_{r-1}} + \cdots + 2^{-m_2 - \ldots - m_r} = 2^{-1} + 2^{-2} \ldots 2^{-(r-1)} = \sum_{n=1}^{r-1} 2^{-n}, \tag{D.7}$$

a well-known geometric series, that converges monotonically increasing to 1 when $r \to \infty$, leading to the conclusion that

$$2^{-1} + 2^{-2} \ldots 2^{-(r-1)} = \sum_{n=1}^{r-1} 2^{-n} < 1 \tag{D.8}$$

and

$$\varepsilon = \log_2 \left[1 + 2^{-m_r} + 2^{-m_r - m_{r-1}} \ldots 2^{-m_2 - \cdots - m_r} \right] < 1. \tag{D.9}$$

This allows one to write any natural number as

$$\log_2 n = \sum_{i=1}^{r} m_i + \varepsilon \tag{D.10}$$

with $\varepsilon < 1$, or, to the conclusion that, in fact

$$\lfloor \log_2 n \rfloor = \sum_{i=1}^{r} m_i, \tag{D.11}$$

and this is exactly Equation (4.3).

As an example, let's consider the orbit from $n_0 = 29$, $\boldsymbol{m}(0) = (0, 2, 1, 1)$. The orbit is simply

$$\mathcal{O}(29) = (29, 44, 22, 11, 17, 26, 13, 20, 10, 5, 8, 4, 2, 1)\,.$$

In the $\boldsymbol{m}-$vectors representation, $\boldsymbol{m}(0) = (0, 2, 1, 1)$, $\boldsymbol{m}(1) = (2, 1, 2)$, $\boldsymbol{m}(2) = (1, 1, 2)$, $\boldsymbol{m}(3) = (0, 1, 2)$, $\boldsymbol{m}(4) = (0, 4)$, $\boldsymbol{m}(5) = (1, 2, 1)$, $\boldsymbol{m}(6) = (0, 2, 1)$, $\boldsymbol{m}(7) = (2, 2)$, $\boldsymbol{m}(8) = (1, 2)$, $\boldsymbol{m}(9) = (0, 2)$, $\boldsymbol{m}(10) = (3)$, $\boldsymbol{m}(11) = (2)$, $\boldsymbol{m}(12) = (1)$, $\boldsymbol{m}(13) = (0)$. Then the $\boldsymbol{M}(S)$ (recall that S is the time step when $f^{(S)}(n_0) = 1$) matrix is represented by

$$\boldsymbol{M}(S) = \begin{pmatrix} 0 & 2 & 1 & 1 \\ 2 & 1 & 2 & 0 \\ 1 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 4 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 2 & 1 & 0 \\ 2 & 2 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{D.12}$$

Now, by applying the $\boldsymbol{M}(T)$ matrix into a $4-$dimensional unitary vector $\boldsymbol{u} = (1, 1, 1, 1)$,

one gets

$$\boldsymbol{M}(T)\boldsymbol{u} = \begin{pmatrix} 0 & 2 & 1 & 1 \\ 2 & 1 & 2 & 0 \\ 1 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 4 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 2 & 1 & 0 \\ 2 & 2 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 4 \\ 3 \\ 4 \\ 4 \\ 3 \\ 4 \\ 3 \\ 2 \\ 3 \\ 2 \\ 1 \\ 0 \end{pmatrix} \tag{D.13}$$

and then, when comparing $\boldsymbol{M}(T)\boldsymbol{u}$ with the orbit $\mathcal{O}(19)$, it is readily checked that the components of $\boldsymbol{M}(T)\boldsymbol{u}$ are exactly $\lfloor \log_2 \mathcal{O}(19) \rfloor$.

Appendix E

Initial condition of M-Matrix of Figure 4.2

The base-10 value of the initial condition is of the order of 2^{968}, once $\sum_{i=1}^{r(0)} m_i(0) = 968$. The exact value is

$$
\begin{aligned}
n_0 = \; & 2\,894\,406\,445\,530\,459\,127\,801\,439\,587\,166\,010\,579\,412\,407\,598\,002\,616\,655\,783\,985 \\
& 410\,974\,349\,944\,946\,862\,195\,562\,886\,153\,303\,107\,407\,672\,503\,643\,211\,159\,480\,461 \\
& 019\,488\,145\,877\,004\,681\,620\,273\,117\,409\,567\,837\,533\,855\,974\,893\,812\,026\,396\,228 \\
& 244\,923\,764\,634\,974\,437\,339\,099\,166\,982\,117\,371\,483\,584\,490\,457\,851\,959\,726\,133 \\
& 386\,848\,412\,321\,139\,544\,078\,287\,519\,197\,379\,890\,553\,765\,278\,925\,825.
\end{aligned}
$$

The initial $\boldsymbol{m}$-vector is

$$
\begin{aligned}
\boldsymbol{m}(0) = \; & (0, 10, 4, 3, 8, 6, 3, 9, 8, 3, 6, 7, 10, 7, 6, 7, 2, 10, 10, 5, 1, 7, 5, 1, 10, 4, 1, 5, 5, 6, 3, \\
& 2, 9, 7, 7, 5, 4, 1, 5, 1, 10, 5, 6, 2, 3, 1, 10, 6, 1, 10, 9, 8, 1, 8, 10, 3, 9, 1, 2, 3, 4, 6, 7, \\
& 5, 6, 4, 6, 3, 5, 10, 2, 1, 4, 7, 4, 4, 2, 7, 7, 1, 4, 5, 8, 4, 8, 3, 2, 7, 8, 6, 7, 10, 8, 2, 6, 5, \\
& 10, 3, 5, 5, 9, 3, 2, 2, 4, 2, 9, 9, 4, 7, 3, 4, 7, 2, 3, 9, 7, 9, 7, 10, 6, 1, 9, 9, 7, 5, 5, 7, 6, \\
& 6, 5, 9, 9, 4, 10, 6, 2, 3, 1, 7, 9, 6, 5, 9, 2, 2, 7, 6, 6, 2, 3, 4, 9, 9, 3, 9, 6, 4, 1, 3, 1, 8, \\
& 3, 1, 7, 1, 7, 9, 6, 7, 8, 5, 5, 5, 7, 4, 2, 8, 3, 2, 3)
\end{aligned}
$$

Appendix F

Transition Matrices of m-vectors

The Transition Matrices $T^{(i)}$ for each of the eighth first graphs displayed in Figure 4.8 are

$$T^{(1)} = \begin{pmatrix} \frac{24}{41} & \frac{10}{41} & \frac{3}{41} & \frac{3}{41} & \frac{1}{41} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \tag{F.1}$$

$$T^{(2)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{13}{18} & \frac{1}{12} & \frac{5}{36} & \frac{1}{36} & \frac{1}{36} & 0 \\ \frac{1}{16} & \frac{3}{16} & \frac{11}{16} & 0 & 0 & 0 & \frac{1}{16} \\ 0 & \frac{6}{7} & 0 & \frac{1}{7} & 0 & 0 & 0 \\ 0 & 0 & \frac{2}{3} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{3}{4} \end{pmatrix} \tag{F.2}$$

$$T^{(3)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{13}{18} & \frac{1}{6} & \frac{1}{36} & \frac{1}{18} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} \\ \frac{1}{15} & \frac{2}{5} & \frac{7}{15} & 0 & \frac{1}{15} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{2}{5} & 0 & \frac{3}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \tag{F.3}$$

$$T^{(4)} = \begin{pmatrix} \frac{11}{17} & \frac{1}{8} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ \frac{4}{17} & \frac{21}{32} & \frac{1}{4} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{17} & \frac{1}{8} & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{17} & \frac{1}{32} & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{32} & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{32} & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{F.4}$$

$$T^{(5)} = \begin{pmatrix} \frac{5}{7} & \frac{2}{7} & 0 & 0 & 0 \\ \frac{4}{25} & \frac{12}{25} & \frac{4}{25} & \frac{3}{25} & \frac{2}{25} \\ \frac{4}{9} & \frac{1}{9} & \frac{4}{9} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & \frac{1}{5} & \frac{1}{5} & 0 & \frac{3}{5} \end{pmatrix} \tag{F.5}$$

$$T^{(6)} = \begin{pmatrix} \frac{49}{53} & \frac{2}{53} & \frac{2}{53} & 0 & 0 & 0 \\ 0 & \frac{5}{9} & \frac{1}{3} & 0 & 0 & \frac{1}{9} \\ \frac{4}{7} & \frac{1}{7} & \frac{2}{7} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{F.6}$$

$$T^{(7)} = \begin{pmatrix} \frac{27}{29} & \frac{1}{29} & \frac{1}{58} & \frac{1}{58} \\ \frac{2}{7} & \frac{3}{7} & \frac{2}{7} & 0 \\ \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \tag{F.7}$$

$$T^{(8)} = \begin{pmatrix} \frac{33}{34} & 1 & 0 & 1 \\ \frac{1}{68} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{68} & 0 & 0 & 0 \end{pmatrix} \tag{F.8}$$

Appendix G

Method for sampling the set of initial conditions

The process of generating suitable initial conditions was very careful and systematic and will be described below. The source codes for this and all other analysis can be found in the following GitHub repository. The code that creates the initial conditions receives these four arguments:

- `|type|` for the type of components, it can be assigned as `|Random|`, for random initial components, and `prime`, `even`, `odd`, `Pascal Triangle`, `Oscilatory` or `Linear` for structured initial components;
- `mVectorSize` $= r_0$ for the length of the $\boldsymbol{m}-$vector;
- `MaxRand` used when `type = "Random"`, defining the maximum random value of the $m-$vector components;
- `BlockSize` $= b$ used for structured initial components, describing the size of the building block of the $\boldsymbol{m}-$vector, e.g, for `BlockSize` $= b$ the m_i components will be formed by blocks with b (denoted generically as $[\varepsilon_1, \ldots \varepsilon_b]$) elements of that `type`, repeating until $i =$ `mVectorSize`. This imposes that `mVectorSize` must be divisible by b.

This method for sampling allows one to estimate how large are the initial conditions being created. Using Equation (4.3) it is possible to estimate $\log_2 n_0$. For the structured initial conditions, given `mVectorSize` $= r_0$ and `BlockSize` $= b$ one finds

$$\lfloor \log_2 n_0 \rfloor = \frac{r_0}{b} \sum_{i=1}^{b} \varepsilon_i, \tag{G.1}$$

where the sum goes along the ε_i elements of that `type`. For `type = Random`, it is possible to estimate the sum of components by their average value:

$$\lfloor \log_2 n_0 \rfloor \approx r_0 \langle m_i \rangle. \tag{G.2}$$

Once the decay of the orbit towards the asymptotic state is exponential $n_t = n_0 2^{-\gamma_2 t}$,

the average total stopping time $\langle T \rangle$, starting from n_0 is approximately

$$\langle T \rangle = \frac{1}{\gamma_2} \lfloor \log_2 n_0 \rfloor \approx 5 \lfloor \log_2 n_0 \rfloor \tag{G.3}$$

where $\gamma_2 \approx 1/5$ was used from results of transforming the sequences into stationary described in section 5.1. This enables one to estimate the orbit length from the estimation of $\log_2 n_0$.

For the `type = Random`, all the components of the $\boldsymbol{m}-$vectors are randomly chosen values with a maximum value given by `MaxRand` argument. In total, 168 initial conditions of `Random` type were generated, all of them with `MaxRand` $= 10$, then $\langle m_i \rangle = 5.5$. From all 168, there are 152 with `mVectorSize` $= 180$ and 16 with `mVectorSize` $= 2100$.

For `type = prime`, the components of the initial $m-$vector will be prime numbers. The `BlockSize` $= b$ argument states that the first b primes will be used to generate the building blocks. In total $b!$ different initial conditions are generated from different permutations of the first b prime numbers as building blocks. For example, `type = prime` and $b = 3$ creates the following 3! building blocks: `[2, 3, 5]`, `[2, 5, 3]`, `[3, 2, 5]`, `[3, 5, 2]`, `[5, 2, 3]`, `[5, 3, 2]`. Then, each building block is concatenated with itself until fill the entire `mVectorSize` $= r_0$ components to form one initial condition. As another example, consider the building block `[2, 3, 5]` with $r_0 = 12$, then $\vec{m}(n_0) = (0, 2, 3, 5, 2, 3, 5, 2, 3, 5, 2, 3, 5)$. Notice that $m_1 = 0$, and this is a rule applied for all initial conditions, in order to create odd n_0 and avoid more bias.

For `type = even` (`type = odd`) the $m-$vectors components will be even (odd) numbers. The `BlockSize` $= b$ argument states that the first b even (odd) numbers will be used to form the building blocks. In total, $b!$ different initial conditions are generated from different permutations of the first b even (odd) numbers as building blocks. For example, `|type= even|` and $b = 3$ creates the following 3! building blocks: `[2, 4, 6]`, `[2, 6, 4]`, `[4, 2, 6]`, `[4, 6, 2]`, `[6, 2, 4]`, `[6, 4, 2]`. Each building block is also concatenated with itself until fill the entire `mVectorSize` components to form one initial condition.

For each of the three types described above, it was used $b = 2, 3, 4, 5$, generating $2! + 3! + 4! + 5! = 152$ initial conditions with `mVectorSize` $= 180$. Another 4, one for each $b = 2, 3, 4, 5$ value, were created with `mVectorSize` $= 2100$ in order to include very large orbits in our study. In total, 156 initial conditions were generated for each of `Prime`, `Even` and `Odd` types.

For `type = Pascal Triangle`, the building blocks of size `BlockSize` $= b$ are formed by elements of the bth row of the Pascal Triangle. In this case, the rows contains repeated elements, and then the permutations are taken discarding repetition. For example, $b = 3$ takes elements from 3rd row of the pascal triangle `[1, 2, 1]` with 3 different permutations, generating the following building blocks: `[1, 2, 1]`, `[1, 1, 2]`, `[2, 1, 1]`. Then each building block of the row is then concatenated to itself until fill the entire `mVectorSize` components to form one initial condition. For this type, $b = 2, 3, 4, 5, 6$ were used, creating 130 initial conditions with `mVectorSize` $= 180$. Another 4, one for each $b = 2, 3, 4, 5$ were created with `mVectorSize` $= 2100$.

The `type = Oscillatory` is formed by sine-like building blocks `[1, 2,..., A, A-1, ..., 2]`, and cosine-like building blocks `[A, A-1, ..., 2, 1, 2, ..., A-1]` where `A` is the amplitude of the block. The amplitude is directly related to the `blocksize` b by $A = (b + 2)/2$, and the `mVectorSize` $= r_0$ is set according to the amplitude in order to obtain control on the size of the initial condition. For this type, both sine-

and cosine-like, b needs to be even and no permutation is allowed, then it was chosen $b = 2, 4, \ldots, 100$, generating 50 initial conditions for each sine- and cosine-like oscillatory types, creating 100 initial conditions with $r_0 = (\lfloor \log_2 n_0 \rfloor b)/\sum \varepsilon_i$, given $\sum \varepsilon_i = b^2/4 + b$ and fixed $\lfloor \log_2 n_0 \rfloor = 3000$. Another 8 initial conditions were generated for $b = 2, 4, 6, 8$ in sine-like building blocks, 4 with $r_0 = 180$ and 4 with $r_0 = 2100$.

Finally, the `type` = `Linear` creates a building blocks that are arithmetic progressions from 1 to `BlockSize` $= b$, and another building block with its reverse sequence, from b to 1. For $b = 30$, the building block is `[1, 2, 3, ..., 29, 30]`, and the reversed is `[30, 29, ..., 2, 1]`. This arithmetic progression is concatenated with itself until fill the entire `mVectorSize`$= r_0$ components, that is also set according to the `blocksize` parameter to control the initial condition size. For this type $b = 3, 4, \ldots 60$ was used, creating 57 initial conditions for ascendant linear and 57 for descendent linear with $r_0 = (\lfloor \log_2 n_0 \rfloor b)/\sum \varepsilon_i$, given $\sum \varepsilon_i = b(b+1)/2$ and fixed $\lfloor \log_2 n_0 \rfloor = 2000$. Another 3 initial conditions were generated for $b = 30, 60, 180$ with $r_0 = 180$ fixed.

All information given in the last paragraphs is summarized in Table G.1, together with information about the magnitude of the initial condition from $\lfloor \log_2 n_0 \rfloor$ from Eqs. (G.1) and (G.2), and estimations of total stopping time from Eq. (G.3).

Table G.1: Summary of the sample of orbits and initial conditions. For each type, it is presented: the size of the initial $m-$vector (`mVectorSize`), the size of the building block for structured $m-$vectors (`BlockSize`), the average of the components for random $m-$vectors or the sum of the building blocks for structured ($\langle m_i \rangle$ or $\sum \varepsilon_i$), the magnitude of the initial condition ($\lfloor \log_2 n_0 \rfloor$), the average total stopping time ($\langle S \rangle$) and the number of initial conditions of that type generated

`type`	$b =$ `BlockSize`	`mVectorSize`	$\lfloor \log_2 n_0 \rfloor$	$\langle T \rangle$	n$^{\text{o}}$ of i.c.
`Random`	-	180	990	4950	152
	-	2100	11550	57750	16
					Subtotal = 168
`Prime`	2	180	450	2250	2
	2	2100	5250	26250	1
	3	180	600	3000	6
	3	2100	7000	35000	1
	4	180	765	3825	24
	4	2100	8925	44635	1
	5	180	1008	5040	120
	5	2100	11760	55800	1
					Subtotal = 156
`Even`	2	180	540	2700	2
	2	2100	6300	31500	1
	3	180	720	3600	6
	3	2100	8400	42000	1
	4	180	900	4500	24
	4	2100	10500	52500	1
	5	180	1080	5400	120
	5	2100	12600	63000	1
					Subtotal = 156
`Odd`	2	180	360	1800	2
	2	2100	4200	21000	1
	3	180	540	2700	6
	3	2100	6300	31500	1
	4	180	720	3600	24
	4	2100	8400	42000	1
	5	180	900	4500	120
	5	2100	10500	52000	1
					Subtotal = 156
`Oscilatory` (Sine-like)	$2, 4, 6, \ldots, 100$	$\dfrac{\lfloor \log_2 n_0 \rfloor b}{\sum \varepsilon_i}$	3000	15000	50
	$2, 4, 6, 8$	180	$45b + 180$	$5\lfloor \log_2 n_0 \rfloor$	4
	$2, 4, 6, 8$	2100	$45b + 180$	$5\lfloor \log_2 n_0 \rfloor$	4
`Oscilatory` (Cosine-like)	$2, 4, 6, \ldots, 100$	$\dfrac{\lfloor \log_2 n_0 \rfloor b}{\sum \varepsilon_i}$	3000	15000	50
					Subtotal = 108

`Pascal`	2	180	180	900	1
	2	2100	2100	10500	1
	3	180	240	1200	3
	3	2100	2800	14000	1
	4	180	360	1800	6
	4	2100	4200	21000	1
	5	180	576	2880	30
	5	2100	6720	33600	1
	6	180	960	4800	90
	6	2100	11200	56000	1
					Subtotal = 134
`Linear` (ascendant)	3 to 60	$\frac{\lfloor \log_2 n_0 \rfloor b}{\sum \varepsilon_i}$	2000	10000	57
`Linear` (decreasing)	3 to 60	$\frac{\lfloor \log_2 n_0 \rfloor b}{\sum \varepsilon_i}$	2000	10000	57
`Linear`	30, 60, 180	180	$90(b+1)$	$5\lfloor \log_2 n_0 \rfloor$	3
					Subtotal = 119
					Total = 997