UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE MATEMÁTICA
INSTITUTO TÉRCIO PACITTI DE APLICAÇÕES E PESQUISAS COMPUTACIONAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

RODRIGO DOS SANTOS VELOSO MARTINS

RANDOM MAPPINGS AND POLYNOMIALS OVER FINITE FIELDS

Rio de Janeiro
2016

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE MATEMÁTICA
INSTITUTO TÉRCIO PACITTI DE APLICAÇÕES E PESQUISAS COMPUTACIONAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

RODRIGO DOS SANTOS VELOSO MARTINS

RANDOM MAPPINGS AND POLYNOMIALS OVER FINITE FIELDS

Tese de Doutorado submetida ao Corpo Docente do Departamento de Ciência da Computação do Instituto de Matemática, e Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários para obtenção do título de Doutor em Informática.

Orientador: Jayme Luiz Szwarcfiter

Co-orientadores: Luis Menasché Schechter e Daniel Panario

Rio de Janeiro
2016

CBIB Martins, Rodrigo dos Santos Veloso

Random mappings and polynomials over finite fields / Rodrigo dos Santos Veloso Martins. – 2016.
140.: il.

Tese (Doutorado em Informática) – Universidade Federal do Rio de Janeiro, Instituto de Matemática, Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais, Programa de Pós-Graduação em Informática, Rio de Janeiro, 2016.

Orientador: Jayme Luiz Szwarcfiter.
Co-orientadores: Luis Menasché Schechter e Daniel Panario.

1. Mapeamentos aleatórios. 2. Sistemas dinâmicos sobre corpos finitos. 3. Grafo funcional. 4. Polinômios gerais. 5. Heurística de Brent e Pollard. 6. Teste de isomorfismo. – Teses. I. Szwarcfiter, Jayme Luiz (Orient.). II. Panario, Luis Menasché Schechter e Daniel (Co-orient.). III. Universidade Federal do Rio de Janeiro, Instituto de Matemática, Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais, Programa de Pós-Graduação em Informática. IV. Título

CDD

RODRIGO DOS SANTOS VELOSO MARTINS

Random mappings and polynomials over finite fields

Tese de Doutorado submetida ao Corpo Docente do Departamento de Ciência da Computação do Instituto de Matemática, e Institulo Técio Pacitti de Aplicações e Pesquisas Computacionais da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários para a obtenção do título de Doutor em Informática.

Aprovada em: Rio de Janeiro, ____ de _____ de _____ .

Prof. Dr. Jayme Luiz Szwarcfiter (Orientador)

Prof. Dr. Luis Menasché Schechter (Co-orientador)

Prof. Dr. Daniel Panario (Co-orientador)

Prof. Dr. Carlos Gustavo Tamm de Araújo Moreira

Prof. Dra. Sulamita Klein

Prof. Dr. Severino Collier Coutinho

Prof. Dr. Mitre Costa Dourado

Rio de Janeiro

2016

AGRADECIMENTOS

Agradeço primeiramente à minha família, em especial à minha mãe e ao meu irmão, pelo suporte e pela orientação dados ao longo desta caminhada; por caminhada entende-se aqui um processo longo, anterior ao meu ingresso no curso de doutorado ou sequer no curso de graduação. Agradeço à minha mãe por ter não só me incentivado e me munido de todas as ferramentas (acadêmicas) para enfrentar os obstáculos, mas também por ter ido assistir meus jogos de basquete de categoria de base; por ter passado meses acordando cedo no final de semana para nos levar para jogar *baseball*, quando poucos a reprovariam por dizer que tal atividade era besteira; por ter me incentivado à algum investimento na música através de aulas de violão, ainda que de resultado irrisório; e por ter tolerado minhas escolhas profissionais e acadêmicas mesmo quando estas não estavam alinhadas por completo com suas convicções. Agradeço ao meu irmão por ter estabelecido um caminho a ser seguido desde nossa infância, principalmente no âmbito acadêmico e profissional, como na minha escolha por (tentar, num primeiro momento) cursar o Ensino Médio no CEFET-RJ e na escolha pelo magistério. Sem essa orientação, tão natural que se dava de maneira passiva, jamais teria ingressado na universidade com uma visão tão clara do que aquele ambiente tinha para me oferecer, como monitorias e projetos de iniciação científica.

Em uma tentativa de seguir uma ordem cronológica, neste momento eu gostaria de agradecer a toda comunidade do Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, ou CEFET-RJ. O processo de amadurecimento que este colégio me fez passar, e que faz passar centenas de adolescentes todos os anos, jamais será subestimado por mim. Agradeço especialmente aos meus amigos de CEFET, que assinam muitos dos tijolos que construíram a pessoa que sou hoje. Agradeço ao corpo docente desta instituição, que naturalmente desempenha papel preponderante neste processo de amadurecimento, mas gostaria de destacar um professor: Prof. Daniel Sasaki, onde eu via a imagem do professor ideal. Tomada a decisão de seguir a carreira do Magistério, que não poderia ter sido tomada em uma conversa com uma pessoa senão o Prof. Daniel, esta imagem se tornou a referência de minha ambição profissional e acadêmica.

Ciente da flagrante impossibilidade de seguir uma ordem cronológica, volto no tempo agora para agradecer a todas as pessoas que me ajudaram através do esporte, em especial através do basquete e do futebol americano. Aos meus técnicos, dentre os quais eu cito Leandro Rossini, Marcelo Maluco, Mauro, Allan, Leandro, Ivan e Bruno Milano, eu gostaria de dizer me considero evidência de que é possível ensinar muito a um jovem sobre a vida através do esporte e que o trabalho, formal ou informal,

realizado por vocês tem um valor incomensurável na vida de muitas pessoas. Aos meus amigos e companheiros de ACM, Jequiá, Ilha Avalanche e Mamutes fica, além de meus agradecimentos, a maior das saudades que carrego dentro de mim.

Agradeço a todo o corpo docente responsável por minha formação acadêmica matemática, em especial àqueles do Instituto de Matemática da UFRJ. Dentre os quais eu gostaria de destacar: as Prof. Nedir e Maria Darci, nas quais me apoiei nos meus primeiros passos na graduação; a Prof. Luciane Quoos, a quem agradeço por grande parcela do meu conhecimento específico de área, pela capacidade de escrita que hoje possuo e pela orientação ao longo de alguns anos, que me manteve no caminho que me trouxe hoje a um lugar que não trocaria por nenhum outro; agradeço também em especial a todos os professores dos programas de Mestrado e Doutorado que cursei, pelo conhecimento transmitido neste período.

Agradeço ao Professor Luis Menasché por ter acolhido meu esboço de projeto de doutorado e pela sua contribuição no amadurecimento deste. Apesar de meu curso de doutorado não ter seguido no caminho que imaginamos inicialmente, a confiança que o Prof. Luis teve em mim neste estágio inicial foi imprescindível. Agradeço também ao Prof. Jayme Szwarcfiter pelas mesmas razões, e também por ter investido em mim como aluno mesmo quando o caminho a ser seguido não estava mais claro: através do Prof. Jayme tive contato com pessoas e ambientes que me mantiveram não só motivado mas em constante evolução durante o período mais difícil da minha vida acadêmica; hoje vejo que com qualquer orientador senão o Prof. Jayme eu poderia ter abandonado o caminho que segui.

Agradeço ao Prof. Daniel Panario por motivos tantos que sequer tentarei enumerá-los em sua totalidade. Gostaria de agradecer aqui, no entanto, por ter me acolhido como aluno a uma distância de milhares de quilômetros e por ter reformulado minha visão de orientador acadêmico da melhor maneira possível. Agradeço por me ter acolhido em Ottawa e por ter sido, e permanecido, junto com esposa Lucia e seus filhos, um amigo meu e de minha esposa. Agradeço por não só ter recuperado meu sonho de concluir o doutorado mas também por ter me auxiliado a fazê-lo com uma satisfação e sucesso que excedeu minhas perspectivas mais otimistas.

Agradeço finalmente à mulher que hoje é minha esposa. Nossa história juntos se iniciou muito próximo do meu ingresso no doutorado e ela foi sem dúvidas a pessoa que mais me apoiou ao longo deste processo. Agradeço a Deus por ter encontrado uma pessoa com quem eu compartilho um companheirismo tão grande, algo evidenciado pela ajuda que ela me deu nesses anos de doutorando. Incontáveis vezes eu não enxergava nenhuma luz em meio aos meus problemas e ela foi a responsável por me manter de pé, mesmo quando ela mesmo não saberia dizer onde essa luz se encontrava. À minha esposa eu dedico meu o maior dos agradecimentos.

RESUMO

MARTINS, Rodrigo dos Santos Veloso. **Random mappings and polynomials over finite fields**. 2016. 140 f. Tese (Doutorado em Informática) - PPGI, Instituto de Matemática, Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

O comportamento de iterações de funções é uma área de pesquisa em crescimento devido, em parte, a aplicações em criptografia. Frequentemente estamos interessados nas iterações de polinômios sobre corpos finitos, porém o conceito de mapeamentos aleatórios também desempenha papel importante neste campo, definidos como funções escolhidas aleatoriamente de maneira uniforme dentre todas as funções de um conjunto finito nele mesmo. Em algumas aplicações, como na fatoração de inteiros, as propriedades combinatórias de mapeamentos aleatórios fornecem um modelo heurístico para o comportamento de polinômios sobre corpos finitos. O objetivo deste trabalho é estudar as propriedades dinâmicas destas classes, com foco nas relações entre elas. Nossas principais contribuições são a prova de resultados combinatórios sobre polinômios que dão sustentação e estendem a heurística mencionada acima e um algoritmo de pior caso linear que reconhece se dois polinômios têm dinâmicas equivalentes. Em outras palavras, nós fornecemos um teste de isomorfismo linear para o grafo funcional de polinômios sobre corpos finitos. Nós também obtemos estimativas assintóticas sobre a distribuição dos ciclos de mapeamentos cujos graus de entrada estão restritos ao conjunto $\{0, k\}$.

Palavras-chave: Mapeamentos aleatórios, sistemas dinâmicos sobre corpos finitos, grafo funcional, polinômios gerais, heurística de Brent e Pollard, teste de isomorfismo.

ABSTRACT

MARTINS, Rodrigo dos Santos Veloso. **Random mappings and polynomials over finite fields**. 2016. 140 f. Tese (Doutorado em Informática) - PPGI, Instituto de Matemática, Instituto Tércio Pacitti, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

The behavior of iterations of functions is a growing area of research in part due to applications in cryptography. One is frequently interested in the iterations of polynomials over finite fields, but the concept of random mappings is also of interest. These are defined as functions chosen uniformly at random from the set of all functions of a finite set to itself. In some applications, such as the factorization of integers, the combinatorial properties of random mappings provide a heuristic model for the approximation of the behavior of polynomials over finite fields. It is the purpose of this work to study the dynamic properties of these classes, with focus on the connections between them. Our main contributions consist of combinatorial results on polynomials that support and extend the heuristic mentioned above and a worst-case linear-time algorithm that recognizes if two given polynomials have equivalent dynamics. In other words, we provide a linear isomorphism test for the functional graph of polynomials over finite fields. We also obtain asymptotic estimates on the distribution of the cycles of mappings whose indegrees are restricted to the set $\{0, k\}$.

Keywords: random mappings, dynamical systems over finite fields, functional graph, general polynomials, Brent-Pollard heuristic, isomorphism test.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

Let p be an odd prime and f be a polynomial of degree $d \geq 2$ over the integers modulo p, denoted by $\mathbb{F}_p$. We define the functional graph of f as the directed graph $G = (V, E)$ such that $V = [p] = \{0, \ldots, p-1\}$ and $E = \{(x, f(x)), x \in \mathbb{F}_p\}$, where $\mathbb{F}_p$ is represented as $\{0, \ldots, p-1\}$. It is easily seen that the connected components of this digraph are cycles of non-plane trees; see Figure 1.1. The study of iterations of such polynomials and the parameters defined over the corresponding functional graphs is a growing area of research, in part due to some applications in cryptography. Properties of functional graphs of polynomials over finite fields are related to chains of primes (TESKE; WILLIAMS, 2000) and pseudo-random bit generators (BLUM; BLUM; SHUB, 1986), but the most interesting application is perhaps the factorization of integers. One may consider the Lucas-Lehmer test and Pepin's test for the factorization of Mersenne and Fermat numbers, respectively (LEHMER, 1930; LUCAS, 1878). These are based on the iterations of quadratic polynomials, namely the ones of $f(x) = x^2 - 2$ and $f(x) = x^2$.

For the factorization of any integer n, Pollard in (POLLARD, 1975) considers a quadratic polynomial $f(x) = x^2 + a$ and the sequence $(x_k)_{k \geq 0}$ defined by

$$x_{k+1} \equiv f(x_k) \pmod{n},$$

where x_0 is some integer modulo n. A prime factor of n is expected to be found when there is a collision $x_k \equiv x_j \pmod{p}$ for some prime p dividing n. For example, in order to factor $n = 221$, the choice $f(x) = x^2 + 1$ and $x_0 = 7$ in Pollard's algorithm leads to the sequence

$$(x_0, x_1, x_2, x_3, x_4, x_5, x_6) = (7, 50, 70, 39, 196, 184, 44)\,.$$

The computation of $\gcd(x_6 - x_2, 221) = \gcd(-26, 221) = 13$ leads to the factorization

of $n = 221$; see (POLLARD, 1975; BRENT, 1980) for details on the detection of this collision. We stress the connection between the sequence above of integers modulo 221 and the path starting at the node $x_0 = 7$ in the functional graph of $f(x) = x^2 + 1$ modulo 13 (Figure 1.1). The number of steps in the execution of the algorithm depends on the length of this path. This quantity is defined as the *rho length* of the node x_0 in the functional graph of f.

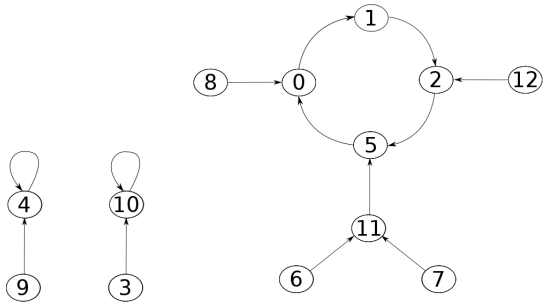

Figure 1.1: Functional graph of $f(x) = x^2 + 1$ over $\mathbb{F}_{13}$.

One might assume uniform distribution on the nodes of a polynomial f and ask oneself what is the average rho length of f. This problem is of prime importance in the analysis of Pollard's factorization method and other cryptographic algorithms. It is estimated in (POLLARD, 1975) that a prime factor p will be detected after about $\sqrt{p}$ steps and, although there is some work in this area, few rigorous results have actually been proved. This prediction is based on previously known statistics on mappings, defined as functions from a finite set to itself. Pollard suggested that a quadratic polynomial like $f(x) = x^2 - 1 \in \mathbb{F}_p[x]$ behaves like a random mapping over $\mathbb{F}_p$, that is, a mapping chosen randomly and uniformly from the class of all the p^p functions over $[p]$ to itself. It is known that, for an element x_0 chosen randomly in $[p]$ and a random mapping φ over $\mathbb{F}_p$, the expected rho length of x_0 under φ is asymptotically $\sqrt{\pi p/2}$. There are a number of papers where estimates on random mappings are obtained, but we single out (ARNEY; BENDER, 1982) because in this work the authors consider uniform distribution on classes of mappings $\varphi : [n] \longrightarrow [n]$ defined by different restrictions on their *indegree distribution*, that is, the sequence

$(n_k)_{k\geq 0}$ given by

$$n_k = \#\{y \in [n] \colon |\varphi^{-1}(y)| = k\},\ k \geq 0.$$

For instance, one might consider mappings $\varphi : [n] \longrightarrow [n]$ such that $n_k = 0$ for $k \geq 3$, since it is known that quadratic polynomials over $\mathbb{F}_p$ have indegrees bounded by 2. The asymptotic average rho length of such a restricted class of mappings is proved in (ARNEY; BENDER, 1982) to be $\sqrt{\pi n/2\lambda}$, where λ is the asymptotic average coalescence of the mappings in hand, defined as follows. The *coalescence* $V(\varphi)$ of a mapping φ is the variance of the distribution of indegrees of its functional graph under uniform distribution on the nodes. For example, a quadratic polynomial $f(x) = x^2$ over $\mathbb{F}_p$, $p > 2$, has indegree distribution given by $n_1 = 1$ and $n_0 = n_2 = (p-1)/2$; see Section 2.4.2. Since the expected preimage size of a random uniform element of $\mathbb{F}_p$ is 1 (see Section 2.3.1), it follows that

$$V(f) = \sum_{x \in \mathbb{F}_p} \frac{1}{p} |f^{-1}(x)|^2 - 1 = \frac{1}{p} + \frac{p-1}{2} \cdot \frac{1}{p} \cdot 4 - 1 = 1 - \frac{1}{p}.$$

We note that the average coalescence of quadratic polynomials is asymptotically equivalent to 1 as p approaches infinity. Thus the results of (ARNEY; BENDER, 1982) support the heuristic by Pollard that quadratic polynomials behave like random unrestricted mappings.

Brent and Pollard observed in (BRENT; POLLARD, 1981) the possible connection between the statistics of a polynomial and its indegree distribution. They conjectured that the expected rho length of a node $x_0 \in [n]$ under a function $\varphi : [n] \longrightarrow [n]$ is given by $\sqrt{\pi n/2V(\varphi)}$, where $V(\varphi)$ is the coalescence of φ. The *factor of non-randomness* of $V(\varphi)$, defined as the ratio of its average rho length and the random mapping estimate $\sqrt{\pi n/2}$, is given by $V(\varphi)^{-1/2}$ according to this heuristic. The Brent-Pollard heuristic was successfully applied in (BRENT; POLLARD, 1981) in the case of polynomials of the form $x^d + c \pmod p$, leading to the factorization of the eighth Fermat number. This heuristic remains an important

element in the design and analysis of some cryptographic algorithms and is known as the *Brent-Pollard heuristic*.

We believe that the Brent-Pollard heuristic plays an important role in the field of cryptographic algorithms. It provides support to adaptations of Pollard's method to other problems such as the discrete logarithm problem (POLLARD, 1978; TESKE, 1998, 2001; VAN OORSCHOT; WIENER, 1999); Pollard himself suggested this application in (POLLARD, 1978). See (GALLANT; LAMBERT; VANSTONE, 2000; WIENER; ZUCCHERATO, 1999) for a few authors that believe that this is the most efficient method against a general instance of the discrete logarithm problem. Several authors have considered the Brent-Pollard heuristic in their work, specially when considering r-adding walks; see for example (BAILEY et al., 2009; BERNSTEIN; LANGE, 2013; BERNSTEIN; LANGE; SCHWABE, 2011; BOS; COSTELLO; MIELE, 2014; BOS et al., 2012; BOS; KLEINJUNG; LENSTRA, 2010; MORAIN, 1998; TESKE, 2001; ZHANG; WANG, 2013).

Our interest in this work lies on combinatorial and number theoretic results on iterations of mappings and polynomials over finite fields. Since the distribution of preimage sizes of a polynomial appears to play an important role in its average rho length, we survey in Chapter 3 the known results on this and give new proofs for cubic and quartic polynomials over finite fields. We improve the error term in one of the cases for quartic polynomials over the finite field with q elements, with $q = p^e$, $p > 3$ and $e > 1$. We consider the class of general polynomials (BIRCH; SWINNERTON-DYER, 1959) and use Cohen's results (COHEN, 1970) to determine their asymptotic indegree distribution. As a consequence, we prove that the coalescence of general polynomials of a fixed degree $d \geq 2$ is asymptotically 1.

In Chapter 4 we focus on the Brent-Pollard heuristic, where a polynomial over a finite field is seen as a random mapping. We use the results of Chapter 3

on the indegree distribution of polynomials to investigate different choices for the class of mappings used in the heuristic. The combination of our work with general polynomials over $\mathbb{F}_p$ and the Brent-Pollard heuristic suggests that, for large values of p, the behaviour of these polynomials is similar to that of random mappings with respect to the average rho length. Our experiments support this heuristic. We use experimental results to show that the erratic behaviour of quadratic polynomials of the form $x^2 - 2 \pmod p$, observed by Pollard in (POLLARD, 1975), is a particular case of a phenomenon observed in Chebyshev polynomials $T_d \in \mathbb{F}_p[x]$ of various degrees $d \geq 2$.

In Chapter 5 we present an algorithm that recognizes in linear time if two polynomials over a finite field have equivalent dynamics. This problem is known as *the graph isomorphism problem (GI)*. Although our algorithm determines correctly the answer to this decision problem in the case of any pair of mappings, the linearity of the analysis holds provided that these mappings have bounded indegrees. If one considers as input of the algorithm two mappings chosen randomly and uniformly from the set of n^n mappings on n nodes, then the average-case bitwise complexity of our algorithm remains subquadratic. Our algorithm is an extension to higher degrees of the algorithm of (KONYAGIN et al., 2016).

In Chapters 3 and 4 we focus on the iterations of polynomials and mappings where the starting point is a node in the corresponding functional graph. In Chapter 6 we investigate the problem treated in (SCHMUTZ, 2011), where the author considers a random mapping $\varphi : [n] \longrightarrow [n]$ and its functional compositions $\varphi^{(\ell)}, \ell \geq 1$, in the space defined by all n^n mappings on n nodes. We investigate the period of this sequence, that is, the least integer $\mathbf{T} \geq 1$ such that $f^{(\ell+\mathbf{T})} = f^{(\ell)}$ for all $\ell \geq n$. The parameter $\mathbf{T} = \mathbf{T}(\varphi)$ is equivalent to the period of the permutation obtained by restricting φ to its cyclic nodes, that is, the least common multiple of the length of all cycles in the functional graph of φ. In (SCHMUTZ, 2011) the author also

considers the parameter $\mathbf{B}$, defined as the product of all cycle lengths of φ, including multiplicities, and compares the asymptotic average value of $\mathbf{T}$ and $\mathbf{B}$ over all mapping on n nodes. We provide in Chapter 6 similar results for the classes of mappings such that the indegree of every node is either 0 or k, for some $k \geq 2$. Our motivation for the treatment of this class of mappings arises from the indegree distribution of quadratic polynomials: the ratio of nodes with indegrees 0, 1 and 2 are asymptotically 1/2, 0 and 1/2.

The work presented in Chapter 6 represents partial results of our initial project on this problem, namely to extend the results of (SCHMUTZ, 2011) to all classes of mappings treated in (ARNEY; BENDER, 1982). The restriction on the indegree distribution of mappings considered in Chapter 6 is a particular case of the one considered in (ARNEY; BENDER, 1982). In Chapter 7 we elaborate on this problem as well as other problems for future research. Obtaining other asymptotic estimates on the parameters $\mathbf{T}$ and $\mathbf{B}$, other than their expected values, could also be of interest; see Section 1 of (SCHMUTZ, 2011). There are interesting open problems in number theory as well: a number of experimental results presented in Chapter 4 could lead to interesting theorems. It is also of our interest to understand the literature on algorithmic aspects of the Galois group of polynomials over finite fields: this might lead to an effective algorithm to recognize general polynomials. We also give our conclusions in Chapter 7.

Part of the results of this thesis, namely the content of Chapters 3 and 4, have been accepted for publication in the International Journal of Number Theory; see (MARTINS; PANARIO, 2016). The research presented in Chapter 5 was submitted for publication in the journal Discrete Applied Mathematics (MARTINS et al., 2015).

2 BACKGROUND

We review in this chapter several mathematical concepts and results needed in this thesis.

2.1 Asymptotic Notation

It is frequently the case that our interest rests in problems with large parameters, such as the factorization of large integers. We are thus interested in the behavior of functions $f(x)$ as x approaches infinity. In this thesis we deal mostly with sequences of real numbers, that is, a real function with domain $\mathbb{N} = \{0, 1, 2, \dots\}$ or $\mathbb{N}^* = \{1, 2, \dots\}$.

Definition 1. *Let $f(n)$ and $g(n)$ be sequences of real numbers. We write:*

(i) $f(n) = O(g(n))$ *as* $n \longrightarrow \infty$ *if there exists* $C > 0$ *such that, for n sufficiently large,*

$$|f(n)| \leq C \cdot |g(n)|;$$

(ii) $f(n) \overset{n\to\infty}{\sim} g(n)$ *if*

$$\lim_{n\to+\infty} \frac{f(n)}{g(n)} = 1;$$

(iii) $f(n) = o(g(n))$ *as* $n \longrightarrow \infty$ *if*

$$\lim_{n\to+\infty} \frac{f(n)}{g(n)} = 0.$$

2.2 Probability Theory

In this section we present the basic concepts of Probability Theory that are necessary for the work presented in this thesis. We refer the reader to (ROHATGI; SALEH, 2011) for a thorough treatment of the results exposed in this section.

2.2.1 Basic Concepts

Let Ω be a set. The definition of a probability space over Ω requires a careful treatment of the class of subsets of Ω, where we study the concept of a σ-field of Ω. This is not necessary when one considers Ω a finite or countable set, that is, a set Ω such that there exists a bijection $\varphi : \Omega \longrightarrow \mathbb{N}$. All the probability spaces that arise in this thesis are defined over finite sets.

Definition 2. *A random experiment is an experiment $\mathcal{E}$ such that*

(i) the set Ω of all possible outcomes of $\mathcal{E}$ are known in advance;

(ii) any particular performance of $\mathcal{E}$ results in an outcome that is not known in advance;

(iii) the experiment $\mathcal{E}$ can be repeated under identical conditions.

If Ω is finite or countable, we define the sample space of $\mathcal{E}$ to be the pair $\left(\Omega, \mathcal{S}\right)$, where $\mathcal{S} = \mathcal{P}(\Omega)$ denotes the class of all subsets of Ω.

One may consider other classes of subsets of Ω instead of $\mathcal{P}(\Omega)$; see Section 2.1 of (CHUNG, 2001).

Definition 3. *Let $\left(\Omega, \mathcal{S}\right)$ be a sample space. A set function $\mathbb{P} : \mathcal{S} \longrightarrow \mathbb{R}$ is a probability function (probability measure) if:*

(i) $\mathbb{P}(A) \geq 0$ *for all* $A \in \mathcal{S}$;

(ii) $\mathbb{P}(\Omega) = 1$;

(iii) if $(A_k)_k$ *is a sequence of subsets of* Ω *such that* $A_k \cap A_j = \emptyset$ *for all* $k \neq j$, *then*

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(A_k).$$

The triple $\left(\Omega, \mathcal{S}, \mathbb{P}\right)$ is a probability space.

Since it is implied that, given a finite or countable set Ω, we consider the class $\mathcal{S} = \mathcal{P}(\Omega)$ of subsets of Ω, we omit $\mathcal{S}$ when we refer to a probability space $\left(\Omega, \mathcal{S}, \mathbb{P}\right)$. Probability spaces possess a number of interesting properties that we do not mention in this text; see (ROHATGI; SALEH, 2011).

Definition 4. *Let Ω be a finite set with n elements and let $\left(\Omega, \mathbb{P}\right)$ be a probability space. If $\mathbb{P}(\{\omega\}) = 1/n$ for all $\omega \in \Omega$, then $\mathbb{P}$ is the uniform (probability) distribution of Ω.*

Our main interest in probability spaces rests on random variables defined over them. We note that our treatment of random variables is also simplified due to the nature of the probability spaces treated in this work. We present below the definition of the Borel σ-field of the real line; we refer the reader to Section 2.1 of (CHUNG, 2001) for a careful treatment of this object.

Definition 5. *The Borel σ-field $\mathcal{B}$ of $\mathbb{R}$ is defined as the class of subsets of $\mathbb{R}$ given by finite or countable unions of intervals of the form $(-\infty, a]$, $a \in \mathbb{R}$, and their complements $\mathbb{R} \backslash (-\infty, a] = (a, +\infty)$.*

Definition 6. *Let $\big(\Omega, \mathbb{P}\big)$ be a finite or countable probability space. A function $X : \Omega \longrightarrow \mathbb{R}$ is a random variable. The function X induces a probability function Q on $(\mathbb{R}, \mathcal{B})$ defined by*

$$Q(B) = \mathbb{P}\big(X^{-1}(B)\big) = \mathbb{P}\big(\{\omega \in \Omega \colon X(\omega) \in B\}\big), \quad B \in \mathcal{B}.$$

Then, Q is the probability distribution of X.

We note that a random variable X defined over Ω may have different probability distributions, depending on the probability function defined over Ω. We omit the reference to the probability space $\big(\Omega, \mathbb{P}\big)$ in the definition of Q in Definition 6 because there is no risk of confusion.

Definition 7. *A random variable X is of discrete type if there exists a countable set $E = \{x_1, x_2, \dots\} \subseteq \mathbb{R}$ such that $\mathbb{P}(X^{-1}(E)) = 1$. The collection of numbers $p_k = \mathbb{P}(X^{-1}(x_k))$, $k \geq 1$, is the probability mass function of X.*

Theorem 1. *If X is a random variable defined on a finite probability space, then X is a discrete random variable.*

Definition 8. *Let X be a random variable defined on a probability space $(\Omega, \mathbb{P})$. The distribution function of X is defined as the function $F : \mathbb{R} \longrightarrow \mathbb{R}$ given by*

$$F(x) = \mathbb{P}\big(\{\omega \in \Omega \colon X(\omega) \leq x\}\big), \quad x \in \mathbb{R}.$$

For simplicity, from this point forward we write $\mathbb{P}(X \leq x)$ for the probability $\mathbb{P}\big(\{w \in \Omega \colon X(\omega) \leq x\}\big)$. Similarly, $\mathbb{P}\big(\{w \in \Omega \colon X(\omega) = x\}\big)$ is denoted by $\mathbb{P}(X = x)$. In order to avoid technical issues regarding convergence of series, from this point forward we restrict ourselves to random variables defined over a finite probability space. In this case we have a probability mass function of the form $\{p_1, \dots, p_s\}$ for some $s \geq 1$, where $p_k = \mathbb{P}(X = x_k)$. The integer s is not of prime interest at this moment, so, for simplicity, we write the probability mass function of a random variable as $\{p_1, p_2, \dots\}$ and leave it implied that this set has finitely many elements.

Definition 9. *Let X be a random variable defined over a finite probability space and let $\{p_1, p_2, \dots\}$ be the probability mass function of X, where $p_k = \mathbb{P}(X = x_k)$. The expectation or expected value of X is defined as*

$$\mathbb{E}[X] = \sum_k x_k p_k.$$

The expected value of a random variable, frequently denoted by μ when there is no risk of confusion, is a real number that describes its central tendency. It can be seen as the average of the values that X may assume, where each such value is weighted by its probability of occurring. We state below an important property of the expectation of a random variable and refer the reader to Chapter 3 of (ROHATGI; SALEH, 2011) for other results.

Theorem 2. *If X is a random variable and $a, b \in \mathbb{R}$, then $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$.*

In order to investigate other parameters associated to a random variable, we need to consider the random variable defined by the composition of a real function and a random variable.

Theorem 3. *Let X be a random variable defined over a finite probability space $(\Omega, \mathbb{P})$ and let $\{p_1, p_2, \dots\}$ be the probability mass function of X, where $p_k = \mathbb{P}(X = x_k)$. If $g : \mathbb{R} \longrightarrow \mathbb{R}$ is a real function, then the composition $g \circ X$ defines a random variable Y over $(\Omega, \mathbb{P})$ such that*

$$\mathbb{E}[Y] = \sum_k g(x_k) p_k.$$

Definition 10. *Let X be a random variable defined over a finite probability space and let $\mu = \mathbb{E}[X]$. The variance of X is defined as $\mathbb{E}[(X - \mu)^2]$ and we write $\sigma^2 = \mathbb{V}[X] = \mathbb{E}[(X - \mu)^2]$. The standard deviation of X is defined as the real number $\sigma = \sqrt{\mathbb{V}[X]}$.*

The variance of a random variable X describes its deviation from its central tendency, that is, how spread are the possible values of X.

Theorem 4. *Let X be a random variable defined over a finite probability space. Then:*

(i) $\mathbb{V}[X] = 0$ *if and only if there exists* $x \in \mathbb{R}$ *such that* $\mathbb{P}(X = x) = 1$;

(ii) $\mathbb{V}[X] = \mathbb{E}[X^2] - \big(\mathbb{E}[X]\big)^2$;

(iii) if $\mu = \mathbb{E}[X]$ *and* $\sigma^2 = \mathbb{V}[X]$, *then the random variable* $Y = (X - \mu)/\sigma$ *satisfies* $\mathbb{E}[Y] = 0$ *and* $\mathbb{V}[Y] = 1$.

The theorem below is known as the Chebyshev-Bienayme inequality and is of prime importance: it provides a bound for the probability that a random variable X assumes a value distant from its mean.

Theorem 5. *Let X be a random variable defined over a finite probability space. If $k \in \mathbb{R}$, then*

$$\mathbb{P}\big(\{|X - \mu| \geq k\}\big) \leq \frac{\sigma^2}{k^2}.$$

2.2.2 Sequences of Random Variables

In this section we discuss the basic concepts regarding sequences of random variables. For instance, consider, for $n \geq 1$, the probability space $(\Omega_n, \mathbb{P}_n)$ defined by the uniform probability distribution on the set of rooted trees on n nodes. Let X_n be the random variable that represents the height of a tree in Ω_n. One might wonder what is the expected value of X_n and, moreover, what is the asymptotic behavior of the sequence $\big(\mathbb{E}[X_n]\big)_n$ as n approaches infinity: does it stay bounded as n approaches infinity? If not, does it grow approximately as $\sqrt{n}$? Or $\log n$? We

investigate similar questions in this thesis and we present in this section the basic definitions and results necessary for our work. We refer the reader to Chapter 6 of (ROHATGI; SALEH, 2011) for more on this topic. See (RUDIN, 1964) for results on sequences of real numbers and sequences of real functions.

Definition 11. *Let $(f_n)_n$ be a sequence of real functions with common domain D. Then $(f_n)_n$ converges pointwise to a function $f : D \longrightarrow \mathbb{R}$ if, for $\epsilon > 0$ and for every $x \in D$, there exists $n_0 = n_0(\epsilon, x) \geq 1$ such that $|f_n(x) - f(x)| < \epsilon$ for all $n \geq n_0$.*

Definition 12. *Let $(X_n)_n$ be a sequence of random variables and let, for $n \geq 1$, $F_n : \mathbb{R} \longrightarrow \mathbb{R}$ be the distribution function of X_n. Let X be a random variable with distribution function F. Then the sequence $(X_n)_n$ converges in distribution to X if the sequence of functions $(F_n)_n$ converges pointwise to F at every point x at which F is continuous. We write $X_n \xrightarrow{L} X$*

Definition 13. *Let $(X_n)_n$ be a sequence of random variables defined on the same probability space $(\Omega, \mathbb{P})$. Then the sequence $(X_n)_n$ converges in probability to a random variable X defined on $(\Omega, \mathbb{P})$ if , for every $\epsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0.$$

We write $X_n \xrightarrow{P} X$.

In Definition 13 it is required that the random variables X_n, $n \geq 1$, are defined in the same probability space. However, we can extend this definition to the case where the random variables X_n are defined in different probability spaces and X is constant.

Definition 14. *Let $(X_n)_n$ be a sequence of random variables defined on possibly distinct probability spaces $(\Omega_n, \mathbb{P}_n)$. Then the sequence $(X_n)_n$ converges in probability to a constant $c \in \mathbb{R}$ if , for every $\epsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}_n(|X_n - c| > \epsilon) = 0.$$

We write $X_n \xrightarrow{P} c$.

2.3 Combinatorics

In this section we present the basic definitions and results concerning mappings. Our focus in this work rests on the dynamic properties of these objects. We stress that polynomials over finite fields and permutations are particular cases of mappings.

2.3.1 Mappings

Definition 15. *A mapping is a function* $\varphi : [m] \longrightarrow [m]$*, where* m *is a positive integer and* $[m] = \{1, \ldots, m\}$.

We define the *preimage* of an element $y \in [m]$ under a mapping $\varphi : [m] \longrightarrow [m]$ by

$$\varphi^{-1}(y) = \{x \in [m] \colon \varphi(x) = y\}.$$

It is clear that the size $|\varphi^{-1}(y)|$ of the preimage of y satisfies $0 \leq |\varphi^{-1}(y)| \leq m$ for all $y \in [m]$.

Definition 16. *The preimage size distribution or the indegree distribution of a mapping* $\varphi : [m] \longrightarrow [m]$ *is the vector* $(n_0, n_1, \ldots, n_m)$*, where* n_k *denotes the number of elements with preimage size* k*, for* $k = 0, 1, \ldots, m$.

Theorem 6. *If* $\varphi : [m] \longrightarrow [m]$ *is a mapping and, for* $k \geq 0$*,* n_k *denotes the number of elements with preimage size* k*, then*

$$\sum_{k \geq 0} n_k = \sum_{k \geq 1} k n_k = m.$$

Consider the probability space given by the uniform distribution on $[m]$ and consider the random variable χ that represents the preimage size of an element, that is, $\chi(k) = |\varphi^{-1}(k)|$ for $k = 1, \ldots, m$. It is clear that the expectation of this random variable is 1:

$$\mathbb{E}[\chi] = \sum_{k\in[m]} \frac{1}{m}\chi(k) = \frac{1}{m}\sum_{k\in[m]} |\varphi^{-1}(k)| = 1. \tag{2.1}$$

The variance of χ is called the coalescence of the mapping φ.

Definition 17. *The coalescence of a mapping $\varphi : [m] \longrightarrow [m]$ is*

$$V(\varphi) = \sum_{y\in[m]} \frac{1}{m}|\varphi^{-1}(y)|^2.$$

Let $\varphi : [m] \longrightarrow [m]$ be a mapping and let $x \in [m]$. The sequence $(x_k)_k$ defined by $x_0 = x$ and $x_k = \varphi(x_{k-1})$, $k \geq 1$, is the *orbit* or the *rho path* of x. Since $[m]$ is a finite set, there exist $0 \leq j < k \leq m$ such that $x_k = x_j$. If k, j are the least such integers, we define them to be the *tail length* and the *rho length* of x, respectively. The *cycle length* of x is defined to be $k - j$. The study of these parameters are of crucial importance to several applications. The structure of a mapping is illustrated beautifully by its functional graph, defined next.

Definition 18. *The functional graph associated to the mapping $\varphi : [m] \longrightarrow [m]$ is the directed graph $G = (V, E)$, where $V = [m]$ and $E = \big\{(x, f(x)),\, x \in [m]\big\}$.*

It follows from the discussion above that the connected components of the functional graph of φ consist of a cycle (that may be a loop) whose nodes are roots of trees. It should be noted that these trees, that we shall call *cyclic trees*, are directed from leaves to cyclic nodes.

The *indegree* of a node $y \in [m]$ is the number of edges of the form $(x, y) \in E$.

We note that the preimage size of an element $y \in [m]$ equals the indegree of the corresponding node.

Definition 19. *A random mapping $\varphi : [m] \longrightarrow [m]$ is a mapping chosen uniformly at random among the set $\mathcal{F}_m$ of mappings on m elements.*

It is of interest to consider classes of mappings according to restrictions on their indegree distribution.

Definition 20. *Let $\mathcal{J}$ be a set of non-negative integers that contains the number zero and at least one integer greater than one. A mapping $\varphi : [m] \longrightarrow [m]$ is a $\mathcal{J}$-mapping if $|\varphi^{-1}(y)| \in \mathcal{J}$ for all $y \in [m]$. If $\mathcal{J} = \{0, 1, \ldots, d\}$ for some $d \geq 2$, then φ is a d-mapping.*

Definition 21. *Let $\varphi : [n] \longrightarrow [n]$ be a mapping and let s be its average rho length. The factor of non-randomness of φ is the ratio between s and $\sqrt{\pi n/2}$. If $\mathcal{A}$ is a subset of the class $\mathcal{F}_n$ of mappings on n nodes and $\overline{s}$ represents the average rho length over all mappings of $\mathcal{A}$, then the factor of non-randomness of $\mathcal{A}$ is the ratio between $\overline{s}$ and $\sqrt{\pi n/2}$.*

The quantity $\sqrt{\pi n/2}$ represents the expected rho length of a random uniform mapping on n nodes (ARNEY; BENDER, 1982). Thus the factor of non-randomness of a class $\mathcal{A}$ of mappings can be interpreted as a measure of how similar are the behaviour of random mappings and the mappings of $\mathcal{A}$.

2.3.2 Permutations

Definition 22. *A mapping $\sigma : [m] \longrightarrow [m]$ is a permutation if σ is surjective.*

The set of all permutations on m elements forms a group under the operation of composition; the definition of groups is presented in Section 2.4.1.

Definition 23. *The group of all permutations $\sigma : [m] \longrightarrow [m]$ is the symmetric group on m elements and is denoted by S_m.*

Theorem 7. *Let $\sigma : [m] \longrightarrow [m]$ be a mapping. The following statements are equivalent.*

(i) σ is a permutation;

(ii) σ is injective;

(iii) $|\sigma^{-1}(y)| = 1$ for all $y \in [m]$.

It follows from Theorem 7 that the connected components of the functional graph of a permutation are cycles. In other words, the cyclic trees of permutations contain a single node. Cycles of length 1 of permutations are *fixed points.*

Definition 24. *Let $\sigma : [n] \longrightarrow [n]$ be a permutation. The inverse function of σ is the mapping $\sigma^{-1} : [n] \longrightarrow [n]$ defined by $\sigma^{-1}(x) = y$ if and only if $\sigma(y) = x$, for $x \in [n]$.*

The following result clarifies the asymptotic behaviour of the number of permutation on n elements as n approaches infinity.

Theorem 8 (Stirling's Formula)**.**

$$n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n}\left(1 + O\left(\frac{1}{n}\right)\right).$$

2.4 Algebraic Structures

The focus of this section is to present the basic definitions and properties regarding a field and the set of polynomials in one variable defined by it. Since the latter carries the algebraic structure of a ring and both rings and fields are naturally defined using groups, we present first the concept of a group. We refer the reader to Chapters 1,2 and 3 of LIDL; NIEDERREITER (2008) for the proofs of the results given in this section and more on these algebraic structures; see also Chapter 7 of (IRELAND; ROSEN, 1990). It should be noted that the results presented in this section may not follow the typical order in which they appear in classical textbooks, where results are presented in the logical order that their proofs demand.

2.4.1 Groups and Rings

Definition 25. *A group $(G, *)$ is a set G together with a binary operation $* : G \times G \longrightarrow G$ such that*

(i) $a * (b * c) = (a * b) * c$, *for all* $a, b, c \in G$;

(ii) there is an element $e \in G$, *the identity or the neutral element of* G, *such that* $a * e = e * a = a$, *for all* $a \in G$;

(iii) for each $a \in G$ *there is an element* a', *the inverse of* a, *such that* $a * a' = a' * a = e$.

If $a * b = b * a$ *for all* $a, b \in G$, *then the group* G *is abelian. If* G *has finitely many elements then* G *is a finite group and its number of elements, denoted by* $|G|$, *is its order.*

It is easy to prove that the neutral element of a group G and the inverse of an element $a \in G$ are unique. The binary operation of the groups treated in this work are addition and multiplication. We use the usual notation for these operations, such as $a + a = 2a$ or $a \cdot a = a^2$. The neutral element is usually represented as 0 or 1, if additive or multiplicative notation is used for G, respectively. From this point forward we use multiplicative notation for a generic group G. We often use the notation G for $(G, \cdot)$.

Definition 26. *Let G be a group. If H is a subset of G such that $(H, \cdot)$ is a group, then H is a subgroup of G.*

Definition 27. *Let G be a group. The subgroup of G consisting of all powers of an element $a \in G$ is the subgroup generated by a and is denoted by $\langle a \rangle$. If $\langle a \rangle$ is finite, then its order is the order of a. Otherwise, a is an element of infinite order.*

Theorem 9. *Let G be a finite group. Then the order of every subgroup of G divides the order of G. Moreover, if a is an element of G, then the order of a divides the order of G.*

In some cases the subgroup generated by an element g of a group G is the whole group. In these cases it is possible to write every element $h \in G$ as $h = g^n$ for some $n \geq 0$, where $g^0 = 1$.

Definition 28. *A group $(G, \cdot)$ is cyclic if $G = \langle g \rangle$ for some $g \in G$. The element g is a generator of the group G.*

Example 1. *The set $\mathbb{Z}$ of the integers together with the operation of addition is a cyclic group generated by 1. Its neutral element is 0.*

Definition 29. *Let $\varphi : G \longrightarrow H$ be an application of a group G into a group H. Then, φ is a homomorphism if $\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$ for all $a, b \in G$. If φ is a bijection, then φ is a (group) isomorphism and G and H are isomorphic. An automorphism is an isomorphism of G onto G.*

A group homomorphism $\varphi : G \longrightarrow H$ is an application that preserves the underlying algebraic structure of G and H. If φ is a bijection, then G and H are different presentations of the same group structure.

We introduce now basic definitions and properties regarding *characters* over a finite abelian group; see Section 5.1 of (LIDL; NIEDERREITER, 2008). Let $U = \{z \in \mathbb{C} \colon |z| = 1\}$. We note that U is a multiplicative group, where the inverse of an element $z \in U$ is its complex conjugate $\overline{z}$.

Definition 30. *Let G be a finite abelian group. A character of G is a homomorphism $\chi : G \longrightarrow U$.*

Theorem 10. *Let χ be a character of a finite abelian group G. Then*

(i) $\chi(e) = 1$;

(ii) $\chi(g)$ *is a* $|G|$*-th root of unity for all* $g \in G$*, that is,* $\chi(g)^{|G|} = 1$;

(iii) $\chi(g^{-1}) = \overline{\chi(g)}$ *for all* $g \in G$.

The set of characters of a finite abelian group G together with the operation of composition forms itself a finite abelian group $G^\wedge$ of order $|G|$. The inverse of a character χ in $G^\wedge$ is given by the character $\overline{\chi}$ defined by $\overline{\chi}(g) = \overline{\chi(g)}$, for $g \in G$. The neutral element of $G^\wedge$ is the *trivial character*, defined below.

Definition 31. *Let G be a finite abelian group. The trivial character χ_0 of G is the character given by $\chi_0(g) = 1$, for all $g \in G$.*

Theorem 11. *Let G be a finite abelian group. If $h \neq e$ is an element of G and ψ is a non-trivial character of G, then*

$$\sum_{g \in G} \psi(g) = 0 \quad \textit{and} \quad \sum_{\chi \in G^\wedge} \chi(h) = 0.$$

Definition 32. *Let R be a set and let $+ : R \times R \longrightarrow R$ and $\cdot : R \times R \longrightarrow R$ be binary operations on R. Then, $(R, +, \cdot)$ is a ring if:*

(i) $(R, +)$ *is an abelian group;*

(ii) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ *for all* $a, b, c \in R$*;*

(iii) $a \cdot (b + c) = a \cdot b + a \cdot c$ *and* $(b + c) \cdot a = b \cdot a + c \cdot a$*, for all* $a, b, c \in R$*.*

Example 2. *The set of the integers together with the operations of addition and multiplication is a ring.*

2.4.2 Fields and Polynomials

Definition 33. *If $(R, +, \cdot)$ is a ring and $(R^*, \cdot)$ forms a group, where $R^* = R \backslash \{0\}$, then $(R, +, \cdot)$ is a field. If R has finitely many elements, then $(R, +, \cdot)$ is a finite field.*

If $(R, +, \cdot)$ is a ring (field), we simply write R is a ring (field); the operations of rings and fields in this text are always written additively and multiplicatively and are thus omitted in this context. Fields are usually denoted by $\mathbb{F}$, $\mathbb{K}$ or $\mathbb{L}$. We note that $\mathbb{Q}$ is a field and $\mathbb{Z}$ is not.

Let p be a prime number and consider the application $\varphi : \mathbb{Z} \longrightarrow \mathbb{Z}_p = \{0, 1, \ldots, p-1\}$ defined as follows. For every integer a, write $a = q \cdot p + r$ with $q \in \mathbb{Z}$, $0 \leq r \leq p-1$, and define $\varphi(a) = r$. It is possible to prove that the operations of addition and multiplication in $\{0, 1, \ldots, p-1\}$ defined by $x + y = \varphi(x + y)$ and $x \cdot y = \varphi(x \cdot y)$, $x, y \in \{0, \ldots, p-1\}$, are well defined: if $\varphi(x) = \varphi(x')$ and $\varphi(y) = \varphi(y')$, then $\varphi(x + y) = \varphi(x' + y')$. It can also be proved that $(\mathbb{Z}_p, +, \cdot)$ forms

a field, the *Galois field* of order p denoted by $\mathbb{F}_p$. We often refer to this field as the set of integers modulo p. The proof of this result can be found in Section 2.1 of (LIDL; NIEDERREITER, 2008).

Definition 34. *Let R, S be rings. An application $\varphi : R \longrightarrow S$ is an homomorphism if $\varphi(a)+\varphi(b) = \varphi(a+b)$ and $\varphi(a)\cdot\varphi(b) = \varphi(a\cdot b)$, for all $a, b \in R$. If φ is one-to-one and onto, then φ is an isomorphism and R and S are isomorphic. An isomorphism of a ring onto itself is an automorphism.*

The set of automorphisms of a ring R forms a group under the operation of composition. The neutral element of this group is the identity and the inverse element of an automorphism φ is its inverse φ^{-1}.

Definition 35. *Let $\mathbb{K}$ be a field. If there exists a positive integer n such that $na = 0$ for all $a \in \mathbb{K}$, then the least such integer is the characteristic of $\mathbb{K}$. If no such integer exists, the characteristic of $\mathbb{K}$ is zero.*

Example 3. *The field $\mathbb{Q}$ of the rational numbers has characteristic zero. The field $\mathbb{F}_p$ of the integers modulo p has characteristic p.*

Example 4. *Let $\mathbb{F}_p[i]$ be the set of elements of the form $a + bi$, where $a, b \in \mathbb{F}_p$ and $i^2 = -1$. Then $\mathbb{F}_p[i]$, together with the usual operations of addition and multiplication, is a field with p^2 elements and characteristic p.*

Theorem 12 below states that the cardinality of every finite field is a prime power and provides uniqueness of finite fields in with a given cardinality; see also Theorem 33.

Theorem 12. *Let $\mathbb{K}$ be a finite field with q elements. Then:*

(i) the characteristic of $\mathbb{K}$ is a prime number p;

(ii) there exists an integer $e \geq 1$ *such that* $q = p^e$.

Moreover, if $\mathbb{K}$ *and* $\mathbb{L}$ *are finite fields with* q *elements, then* $\mathbb{K}$ *and* $\mathbb{L}$ *are isomorphic.*

A finite field with q elements is denoted from now on as $\mathbb{F}_q$. There are two groups associated to every finite field $\mathbb{F}_q$: $(\mathbb{F}_q, +)$ and $(\mathbb{F}_q^*, \cdot)$. These are the *additive group* and the *multiplicative group* of $\mathbb{F}_q$. It is not true that all groups are cyclic, but the algebraic structure that a finite field $\mathbb{F}_q$ carries implies that $(\mathbb{F}_q, +)$ and $(\mathbb{F}_q^*, \cdot)$ are both cyclic.

Definition 36. *Let* $\mathbb{F}_q$ *be a finite field. A generator of the group* $(\mathbb{F}_q^*, \cdot)$ *is a primitive element of* $\mathbb{F}_q$.

Characters of finite fields are categorized into additive and multiplicative characters, as they could be defined over either one of the groups $(\mathbb{F}_q, +)$ or $(\mathbb{F}_q^*, \cdot)$. For this thesis it is required only knowledge of basic properties of characters on the multiplicative group $(\mathbb{F}_q^*, \cdot)$ of $\mathbb{F}_q$.

Theorem 13. *Let* $\mathbb{F}_q$ *be a finite field and let* g *be a fixed primitive element of* $\mathbb{F}_q$. *Let* i *be the imaginary complex number. If* ψ *is a multiplicative character of* $\mathbb{F}_q^*$ *then, for some* $j = 0, 1, \ldots, q-2$, ψ *is written as*

$$\psi(g^k) = (e^{2\pi ij/(q-1)})^k, \text{ for } k = 0, 1, \ldots, q-2.$$

Moreover, every such function defines a multiplicative character.

Theorem 14. *The group of multiplicative characters of a finite field* $\mathbb{F}_q$ *is cyclic of order* $q - 1$.

If $\mathbb{F}_q$ is a finite field with q odd and ψ is a generator of the group of multiplicative characters of $\mathbb{F}_q$, then $\eta = \psi^{(q-1)/2}$ is a character of order 2.

We denote the quadratic multiplicative character of a field $\mathbb{F}_q$ of odd characteristic by $\eta(\cdot)$. If g is a primitive element of $\mathbb{F}_q$, then η is defined as in Theorem 13 with $j = (q-1)/2$. We extend η to $\mathbb{F}_q$ by defining $\eta(0) = 0$; this implies that, for $a \in \mathbb{F}_q$,

$$\eta(a) = \begin{cases} 0, & \text{if } a = 0, \\ 1, & \text{if } a = x^2 \text{ for some } x \in \mathbb{F}_q^*, \\ -1, & \text{otherwise.} \end{cases} \tag{2.2}$$

In other words, for $a \in \mathbb{F}_q^*$, $\eta(a) = 1$ if and only if a is the square of an element of $\mathbb{F}_q$. We note that the character defined in Equation (2.2) has order 2. Also, it follows from Theorem 11 that there are as many squares as non-squares in $\mathbb{F}_q^*$.

It follows from Theorem 14 that Equation (2.2) does not define a multiplicative character over a finite field of even characteristic, as the order of every element in a group divides the order of the group; see Theorem 9.

In the case of a finite field $\mathbb{F}_p$ with p an odd prime number, the quadratic character η of $\mathbb{F}_p$ is often referred to as the Legendre symbol of $\mathbb{F}_p$ (integers modulo p) and is denoted by $(\cdot/p)$. For $a \in \mathbb{F}_p$, if $\eta(a) = 1$ then a is a quadratic residue modulo p; otherwise, it is a quadratic non-residue modulo p. Theorem 15 below is known as the Law of Quadratic Reciprocity and represents a valuable tool for handling Legendre symbols.

Theorem 15. *Let p, q be odd prime numbers. Then*

(i) $\left(\dfrac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$.

(ii) $\left(\dfrac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$.

(iii) $\left(\dfrac{p}{q}\right)\left(\dfrac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$.

The classical approach to the proof of Theorem 12 uses polynomials in one variable with coefficients on a field. For this reason it is frequently the case that polynomials are presented in textbooks before the proof of Theorem 12. Our purpose in this section is to simply introduce the basic concepts of fields and polynomials, so there is no harm done in inverting this order in our presentation.

It is known that the set of polynomials $f(x) = a_0 + a_1x + \cdots + a_dx^d$ with coefficients in a field $\mathbb{F}$ forms a ring with the usual operations of sum and product, denoted by $\mathbb{F}[x]$. We refer to Section 1.3 of (LIDL; NIEDERREITER, 2008) for this result.

Definition 37. *Let $f(x) = a_0 + a_1x + \cdots + a_dx^d$ be a polynomial with coefficients in a field $\mathbb{F}$. If $a_d \neq 0$ then d is the degree of f and we write $\deg(f) = d$; moreover, if $a_d = 1$ then f is a monic polynomial.*

If $\mathbb{F}$ is a field, then a rational function over $\mathbb{F}$ is defined as a quotient f/g of polynomials $f, g \in \mathbb{F}[x]$. The set of rational functions over a field $\mathbb{F}$ forms a field under the usual operations of sum and addition. For a precise treatment of this, see the construction of the quotient ring of a ring R by its set of invertible elements in Section II.4 of LANG (2002).

It is possible to prove that the ring of polynomials supports a division with remainder that is analogous to the well known division algorithm in the ring of the integers: if f, g are polynomials over a field $\mathbb{F}$ and $g(x) \neq 0$, then there exist polynomials $q, r \in \mathbb{F}[x]$ such that $0 \leq \deg(r) < \deg g$ and

$$f(x) = q(x)g(x) + r(x). \tag{2.3}$$

Definition 38. *Let f, g be polynomials over a field $\mathbb{F}$ of positive degree. If there exists a polynomial $h \in \mathbb{F}[x]$ such that $f(x) = g(x)h(x)$, then g divides f.*

Definition 39. *Let f be a polynomial over a field $\mathbb{F}$ of positive degree. If $f(x) = g(x)h(x)$ implies that $\deg(g) = 0$ or $\deg(h) = 0$, then f is irreducible. Otherwise, f is a reducible polynomial.*

Theorem 16. *Let $\mathbb{F}$ be a field. If f is a polynomial over $\mathbb{F}$ of positive degree, then f can be written as*

$$f = ap_1^{e_1} \cdots p_s^{e_s},$$

where $a \in \mathbb{F}$, $p_1, \ldots, p_s \in \mathbb{F}[x]$ are monic irreducible polynomials and $e_1, \ldots, e_s$ are positive integers. The expression above is the factorization of f over $\mathbb{F}$ and is unique up to order in which the irreducible factors occur.

Let $\mathbb{F}$ be a field and let $f(x) = a_0 + a_1x + \cdots + x_da^d$ be a polynomial over $\mathbb{F}$. If $w \in \mathbb{F}$, we define the evaluation of f at w as

$$f(w) = a_0 + a_1w + \cdots a_dw^d.$$

It is clear that $f(w)$ is an element of $\mathbb{F}$.

Definition 40. *Let f be a polynomial over a field $\mathbb{F}$. If $w \in \mathbb{F}$ satisfies $f(w) = 0$, then w is a root of f.*

Theorem 17. *An element $w \in \mathbb{F}$ is a root of a polynomial $f \in \mathbb{F}[x]$ if and only if $x - w$ divides f.*

Definition 41. *Let f be a polynomial over a field $\mathbb{F}$ and let $w \in \mathbb{F}$ be a root of f. If n is a positive integer such that $(x - w)^n$ divides f but $(x - w)^{n+1}$ does not, then n is the multiplicity of w. If $n = 1$, then w is a simple root of f; otherwise, w is a multiple root of f.*

Theorem 18. *Let f be a polynomial over a field $\mathbb{F}$ of degree d. If $w_1, \ldots, w_s$ are distinct roots of f in $\mathbb{F}$ with multiplicities $n_1, \ldots, n_s$, then $n_1 + \cdots n_s \leq d$. In particular, f has at most d distinct roots in $\mathbb{F}$.*

Polynomials over a finite field $\mathbb{F}_q$ represent a particular case of mappings on q elements; see Section 2.3.1. We consider the preimage size distribution of polynomials as discussed in Section 2.3.1. It follows from Theorem 18 that, if $f \in \mathbb{F}_q[x]$ is a polynomial of degree d, then $n_k = 0$ for $k > d$.

Definition 42. *Let f be a polynomial over a field $\mathbb{F}_q$. The value set of f is*

$$V_f = \{y \in \mathbb{F}_q \colon y = f(x) \text{ for some } x \in \mathbb{F}_q\}.$$

Definition 43. *A polynomial $f \in \mathbb{F}_q[x]$ is a permutation polynomial if $|V_f| = q$.*

Theorem 19. *If $\sigma \in \mathbb{F}_q[x]$ is a polynomial given by $\sigma(x) = \alpha x + b$ with $\alpha \neq 0$, then σ is a permutation polynomial. Moreover, σ^{-1} is given by $\sigma^{-1}(x) = \alpha^{-1}(x - b)$.*

It is clear that if $f \in \mathbb{F}_q[x]$ is a permutation polynomial, then $f : \mathbb{F}_q \longrightarrow \mathbb{F}_q$ is a bijection. In other words, f acts as a permutation on the elements of $\mathbb{F}_q$. It should be noted that a polynomial $f \in \mathbb{Z}[x]$ gives rise to a polynomial $f_p \in \mathbb{F}_p$, for each prime number p, given by the reduction of its coefficients modulo p. It is possible that, for a given polynomial $f \in \mathbb{Z}[x]$, there exist prime numbers p, p', such that $f_p \in \mathbb{F}_p$ is a permutation but $f_{p'} \in \mathbb{F}_{p'}$ is not.

Theorem 20. *Let f, g be polynomials over a finite field $\mathbb{F}_q$. If $g(x) = af(\alpha x + b) + c$, $a, \alpha, b, c \in \mathbb{F}_q$ with a, α non-zero, then f and g have the same indegree distribution.*

Proof. Let y be an element of $\mathbb{F}_q$. It follows from Theorem 19 that $f(x) = y$ if and only if $g(x') = y'$, where $x' = \alpha^{-1}(x - b)$ and $y' = ay + c$. It follows that $|f^{-1}(y)| = |g^{-1}(y')|$ and, since y is arbitrary,

$$\big|\{y \in \mathbb{F}_q \colon |f^{-1}(y)| = k\}\big| = \big|\{y \in \mathbb{F}_q \colon |g^{-1}(y)| = k\}\big|,$$

for all $k \geq 0$. □

It is known that there are as many squares as non-squares in $\mathbb{F}_q^*$, if q is odd; see the discussion on the quadratic character above. An expression for the indegree distribution of a quadratic polynomial over a field of odd characteristic follows at once from this fact.

Theorem 21. *Let $\mathbb{F}_q$ be a field of odd characteristic. Let $f(x) = x^2 \in \mathbb{F}_q[x]$ and let, for $k \geq 0$, n_k be the number of elements with preimage size k under f. Then,*

$$(n_0, n_1, n_2) = \left(\frac{q-1}{2}, 1, \frac{q-1}{2}\right).$$

Definition 44. *Let a be an element of a finite field $\mathbb{F}_q$ of characteristic $p > 2$. Define, for $k \geq 1$, the Dickson polynomial $g_k(x, a)$ over $\mathbb{F}_q$ as*

$$g_k(x, a) = \sum_{j=0}^{\lfloor k/2 \rfloor} \frac{k}{k-j} \binom{k-j}{j} (-a)^j x^{k-2j}.$$

The Chebyshev polynomial of order k over $\mathbb{F}_q$ is defined as

$$T_k(x) = 2^{-1} g_k(2x, 1).$$

Theorem 22. *The Dickson polynomial $g_k(x, a)$, $a \in \mathbb{F}_q^*$, is a permutation polynomial over $\mathbb{F}_q$ if and only if $\gcd(k, q^2 - 1) = 1$.*

Theorems 23 and 24 below concern character sums with polynomials arguments. These results correspond to Theorems 5.48 and 5.41 of (LIDL; NIEDERREITER, 2008).

Theorem 23. *Let $f(x) = a_2 x^2 + a_1 x + a_0 \in \mathbb{F}_q[x]$ with q odd and $a_2 \neq 0$. Put $D = a_1^2 - 4a_0 a_2$ and let η be the quadratic character of $\mathbb{F}_q$. Then*

$$\sum_{c \in \mathbb{F}_q} \eta(f(c)) = \begin{cases} -\eta(a_2), & \text{if } D \neq 0, \\ (q-1)\eta(a_2), & \text{if } D = 0. \end{cases}$$

Theorem 24. *Let ψ be a multiplicative character of $\mathbb{F}_q$ of order $m > 1$ and let $f \in \mathbb{F}_q[x]$ be a monic polynomial of positive degree d that is not an m-th power of a polynomial. Then for every $a \in \mathbb{F}_q$ we have*

$$\left| \sum_{c \in \mathbb{F}_q} \psi\big(af(c)\big) \right| \leq (d-1)q^{1/2}.$$

2.4.3 Extensions of a Field

Definition 45. *If $\mathbb{F}, \mathbb{K}$ are fields such that $\mathbb{F} \subseteq \mathbb{K}$, then $\mathbb{F}$ is a subfield of $\mathbb{K}$ and $\mathbb{K}$ is an extension (field) of $\mathbb{F}$.*

Definition 46. *A field containing no proper subfields is a prime field.*

The intersection of any family of subfields of a given field $\mathbb{F}$ is a subfield of $\mathbb{F}$. It is clear that such an intersection is a prime field, that is the prime subfield of $\mathbb{F}$.

Theorem 25. *Let $\mathbb{F}$ be a field. The prime subfield of $\mathbb{F}$ is isomorphic to $\mathbb{Q}$ or $\mathbb{F}_p$, according to the characteristic of $\mathbb{F}$ being 0 or a prime number p.*

If $\mathbb{F}$ is a field extension of $\mathbb{K}$, then $\mathbb{F}$ may be seen as a vector space over $\mathbb{K}$. We refer the reader to (ANTON, 2010), for example, for the definition of a vector space, as well as the concept of dimension of a vector space.

Definition 47. *Let $\mathbb{K}$ be a field extension of $\mathbb{F}$. If $\mathbb{K}$ represents a finite-dimensional vector space over $\mathbb{F}$, then $\mathbb{K}$ is a finite extension of $\mathbb{F}$. If we let n be the dimension of $\mathbb{K}$ over $\mathbb{F}$, then n is the degree of $\mathbb{K}$ over $\mathbb{F}$ and write $[\mathbb{K} : \mathbb{F}]$.*

Theorem 26. *If $\mathbb{L}$ is a finite field extension of $\mathbb{K}$ and $\mathbb{K}$ is a finite field extension of $\mathbb{F}$, then*

$$[\mathbb{L} : \mathbb{F}] = [\mathbb{L} : \mathbb{K}][\mathbb{K} : \mathbb{F}].$$

Definition 48. *Let $\mathbb{K}$ be a field extension of $\mathbb{F}$ and let M be a subset of $\mathbb{K}$. We define $\mathbb{F}(M)$ to be the field obtained by the intersection of all subfields of $\mathbb{K}$ containing both $\mathbb{F}$ and M. If M consists of a single element $\alpha \in \mathbb{K}$, then $\mathbb{F}(\alpha)$ is a simple extension of $\mathbb{F}$.*

It is clear that $\mathbb{F}(M)$ is the smallest subfield of $\mathbb{K}$ containing M and $\mathbb{F}$.

Definition 49. *Let $\mathbb{K}$ be a field extension of a field $\mathbb{F}$ and let $\alpha \in \mathbb{K}$. If α satisfies a non-trivial polynomial equation $a_n\alpha^n + \cdots + a_1\alpha + a_0$ with $a_j \in \mathbb{F}$ not all being zero, then α is an algebraic element over $\mathbb{F}$. An extension $\mathbb{L}$ of $\mathbb{F}$ is algebraic if all elements of $\mathbb{L}$ are algebraic over $\mathbb{F}$.*

Definition 50. *Let $\mathbb{K}$ be a field extension of a field $\mathbb{F}$. If $\alpha \in \mathbb{K}$ is not an algebraic element over $\mathbb{F}$, then α is transcendental over $\mathbb{F}$.*

Theorem 27. *Let $\mathbb{K}$ be a field extension of $\mathbb{F}$ and let $\alpha \in \mathbb{K}$. Then α is a transcendental element of $\mathbb{K}$ over $\mathbb{F}$ if and only if the evaluation $f \longmapsto f(\alpha)$, for $f \in \mathbb{F}[x]$, gives an isomorphism of $\mathbb{F}[x]$ onto $\mathbb{F}(\alpha)$.*

Definition 51. *Let $\mathbb{F}(\alpha)$ be a simple algebraic extension of a field $\mathbb{F}$. The minimal polynomial of α over $\mathbb{F}$ is the monic polynomial f of $\mathbb{F}[x]$ of least degree such that $f(\alpha) = 0$. The degree of f is the degree of α over $\mathbb{F}$.*

Theorem 28. *If $\mathbb{K}$ is a finite extension of $\mathbb{F}$, then $\mathbb{K}$ is an algebraic extension of $\mathbb{F}$.*

We are able to describe the structure of a simple extension $\mathbb{F}(\alpha)$ of a field α by considering the following. Let $f \in \mathbb{F}$ and consider the application φ defined on polynomials $g \in \mathbb{F}[x]$ such that $\varphi(g) = r$, with $r \in \mathbb{F}[x]$ as in Equation (2.3). It is

easy to see that the relation $\sim$ defined below is an equivalence relation (see Section 1.1 of (LIDL; NIEDERREITER, 2008)):

$$g \sim h \text{ if and only if } \varphi(g) = \varphi(h).$$

The operations of addition and multiplication on the set of equivalence classes of $\mathbb{F}[x]$ under the relation $\sim$ are well defined and form a ring $R = \mathbb{F}[x]/(f)$, *the ring of polynomials modulo* g.

Theorem 29. *Let $\alpha \in \mathbb{K}$ be an algebraic element of degree n over $\mathbb{F}$ and let f be the minimal polynomial of α over $\mathbb{F}$. Then:*

(i) $\mathbb{F}(\alpha)$ is isomorphic to the ring of polynomials modulo f;

(ii) $[\mathbb{F}(\alpha) : \mathbb{F}] = n$ and $\{1, \alpha, \ldots, \alpha^{n-1}\}$ is a basis of $\mathbb{F}(\alpha)$ over $\mathbb{F}$;

(iii) every $\beta \in \mathbb{F}(\alpha)$ is algebraic over $\mathbb{F}$ and its degree over $\mathbb{F}$ is a divisor of n.

Theorem 30. *Let $\mathbb{F}_q$ be a finite field with $q = p^n$ elements. Then every subfield of $\mathbb{F}_q$ has p^m elements, where m is a positive divisor of n. Conversely, if m is a positive divisor of n, then there exists a unique subfield of $\mathbb{F}_q$ with p^m elements.*

Theorem 31. *Let p be an odd integer, $q = p^e$ and let a be a non-square of $\mathbb{F}_p^*$. Then $\eta(a) = 1$ in $\mathbb{F}_q$ if and only if e is even.*

Proof. Suppose that $\eta(a) = 1$ in $\mathbb{F}_q$ and let $\alpha \in \mathbb{F}_q$ be such that $\alpha^2 = a$. Since a is not a square in $\mathbb{F}_p$, the polynomial $f(x) = x^2 - a \in \mathbb{F}_p[x]$ is irreducible over $\mathbb{F}_p$. Hence f is the minimal polynomial of α over $\mathbb{F}_p$ and, by Theorem 29, we have $[\mathbb{F}_p(\alpha) : \mathbb{F}_p] = 2$. The field $\mathbb{F}_q$ is an extension of $\mathbb{F}_p(\alpha)$, so Theorem 26 implies that e is even.

Suppose that conversely that e is even. It follows from Theorem 30 that $\mathbb{F}_{p^2}$ is a subfield of $\mathbb{F}_q$. Since a is a non-square in $\mathbb{F}_p$, the argument above implies that $\mathbb{F}_{p^2}$

is isomorphic to $\mathbb{F}_p(\alpha)$, where α is a root of $f(x) = x^2 - a$. Therefore, the equation $x^2 - a$ has a solution in $\mathbb{F}_q$ and so $\eta(a) = 1$ in $\mathbb{F}_q$. □

Definition 52. *Let $\mathbb{K}$ be an extension field of $\mathbb{F}$ and let $f \in \mathbb{F}[x]$. Then f splits in $\mathbb{K}$ if f factors completely into linear factors in $\mathbb{K}[x]$, that is,*

$$f(x) = a(x - \alpha_1) \cdots (x - \alpha_d),$$

where a is the leading coefficient of f and $\alpha_1, \ldots, \alpha_d \in \mathbb{K}$ are not necessarily distinct. If f splits in $\mathbb{K}$ and $\mathbb{K} = \mathbb{F}(M)$, with $M = \{\alpha_1, \ldots, \alpha_d\}$, then $\mathbb{K}$ is the splitting field of f over $\mathbb{F}$.

Theorem 32. *If $\mathbb{F}$ is a field and $f \in \mathbb{F}[x]$, then there exists a splitting field of f over $\mathbb{F}$. Moreover, if $\mathbb{K}$ and $\mathbb{L}$ are splitting fields of f over $\mathbb{F}$, then there exists an isomorphism $\varphi : \mathbb{K} \longrightarrow \mathbb{L}$ that maps roots of f roots into each other and $\varphi(a) = a$ for all $a \in \mathbb{F}$.*

Splitting fields contribute to the classification and the description of the structure of finite fields; see Theorem 12.

Theorem 33. *If $\mathbb{F}_q$ is a finite field with $q = p^e$ elements, then $\mathbb{F}_q$ is isomorphic to the splitting field of $x^q - x$ over $\mathbb{F}_p$. In particular, every $a \in \mathbb{F}_q$ satisfies $a^q = a$.*

Definition 53. *If every polynomial $f \in \mathbb{F}[x]$ of degree $d \geq 1$ has a root in $\mathbb{F}$, then $\mathbb{F}$ is algebraically closed.*

It follows that, if $\mathbb{F}$ is an algebraically closed field, then every polynomial f over $\mathbb{F}$ splits in $\mathbb{F}$. It can be proved that, for every field $\mathbb{F}$, there exists an extension $\mathbb{K}$ of $\mathbb{F}$ such that $\mathbb{K}$ is algebraically closed; see Theorem 2.5 of (LANG, 2002).

Theorem 34 (Section V.2 of (LANG, 2002))**.** *Let $\mathbb{F}$ be a field. Then there exists an extension $\overline{\mathbb{F}}$ of $\mathbb{F}$ that is algebraically closed and algebraic over $\mathbb{F}$. Moreover, if $\mathbb{K}$ and $\mathbb{L}$ are algebraic extensions of $\mathbb{F}$ that are algebraically closed, then there exists an isomorphism $\varphi : \mathbb{K} \longrightarrow \mathbb{L}$ such that the restriction of φ to $\mathbb{F}$ is the identity.*

Theorem 34 states that for every field $\mathbb{F}$ there exists an unique (up to isomorphism) algebraically closed field $\mathbb{K} \supseteq \mathbb{F}$ such that every element of $\mathbb{K}$ is the root of a polynomial over $\mathbb{F}$. This extension is the algebraic closure of $\mathbb{F}$.

Definition 54. *Let $\mathbb{F}$ be a field. If $\mathbb{K}$ is an algebraic extension of $\mathbb{F}$ that is algebraically closed, then $\mathbb{K}$ is the algebraic closure of $\mathbb{F}$.*

Next we present the concept of the Galois group of an extension of fields. We refer the reader to Section VI.1 of (LANG, 2002) for more on this topic.

Definition 55. *Let $\mathbb{K}$ be a field extension of $\mathbb{F}$. If σ is an automorphism of $\mathbb{K}$ such that the restriction of σ to $\mathbb{F}$ gives the identity of $\mathbb{F}$, then σ is an automorphism of $\mathbb{K}$ over $\mathbb{F}$ or an automorphism of the extension $\mathbb{K}\big|_{\mathbb{F}}$.*

It is easy to see that the set of automorphisms of a field extension $\mathbb{K}\big|_{\mathbb{F}}$ forms a group under the operation of composition.

Definition 56. *Let f be a polynomial over a field $\mathbb{F}$ with no multiple roots and let $\mathbb{K}$ be the splitting field of f over $\mathbb{F}$. The group of automorphisms of $\mathbb{K}$ over $\mathbb{F}$ is denoted by $Gal\left(\mathbb{K}\big|_{\mathbb{F}}\right)$ and is the Galois group of $\mathbb{K}\big|_{\mathbb{F}}$ or the Galois group of f over $\mathbb{F}$.*

Let $f(x) = a_d x^d + \cdots + a_1 x + a_0$ be a polynomial over a field $\mathbb{F}$ with no multiple roots and let $\mathbb{K}$ be the splitting field of f over $\mathbb{F}$. Then $\mathbb{K} = \mathbb{F}(M)$ with $M = \{\alpha_1, \ldots, \alpha_d\}$ being the set of roots of f over $\mathbb{K}$. Let σ be an automorphism of $\mathrm{Gal}(\mathbb{K}\big|_{\mathbb{F}})$. Since $\sigma(a) = a$ for all $a \in \mathbb{F}$, the automorphism σ is determined by the image of $\sigma(\alpha_i)$, for $i = 1, \ldots, d$. Moreover, the fact that σ fixes the elements of $\mathbb{F}$ implies that

$$0 = \sigma(0) = \sigma(a_d \alpha_i^d + \cdots a_1 \alpha_i + a_0) = a_d \sigma(\alpha_i)^d + \cdots a_1 \sigma(\alpha_i) + a_0,$$

so σ maps a root of f to another root of f. For this reason we identify the elements of $\mathrm{Gal}(\mathbb{K}\big|_{\mathbb{F}})$ with the permutations on d elements. The group $\mathrm{Gal}(\mathbb{K}\big|_{\mathbb{F}})$ is thus seen as a subgroup of the symmetric group S_d on d elements; see Definition 23.

Definition 57. *Let $f \in \mathbb{F}[x]$ be a polynomial of degree $d \geq 2$ and suppose that f is written as $f(x) = a(x-\alpha_1)\cdots(x-\alpha_d)$ over the splitting field of f over $\mathbb{F}$. The discriminant of $D(f)$ of f is defined as*

$$D(f) = a_0^{2d-2} \prod_{1 \leq i < j \leq d} (\alpha_i - \alpha_j)^2.$$

It is clear from Definition 57 that $D(f) = 0$ if and only if f has a multiple root. It is possible to prove that $D(f) \in \mathbb{F}$ for every polynomial $f \in \mathbb{F}[x]$. There are known expressions for the discriminant of polynomials of small degree in terms of their coefficients. We present the formula for the discriminant of quadratic and cubic polynomials in Theorem 35; we give the discriminant of quartic polynomials only in the case of interest in the remainder of this thesis.

Theorem 35. *Let $\mathbb{F}$ be a field and f be a polynomial over $\mathbb{F}$.*

(i) If $f(x) = ax^2 + bx + c$, then $D(f) = b^2 - 4ac$.

(ii) If $f(x) = ax^3 + bx^2 + cx + d$, then $D(f) = b^2c^2 - 4b^3d - 4ac^3 - 27a^2d^2 + 18abcd$.

(iii) If $f(x) = x^4 + ax^2 + bx + c$, then $D(f) = 256c^3 - 128a^2c^2 + (16a^4 + 144ab^2)c - (4a^3b^2 + 27b^4)$.

The theorem below gives further illustration of the applications of the concept of discriminant of a polynomial. It relates the factorization of a polynomial over $\mathbb{F}_q$ into irreducible factors with its discriminant.

Theorem 36 (Pellet-Stickelberger). *Let f be a monic polynomial over $\mathbb{F}_q$ with q odd, $d = \deg f$ and discriminant $D \neq 0$. If ℓ is the number of irreducible factors in the factorization of f over $\mathbb{F}_q$, then*

$$\eta(D) = (-1)^{d-\ell},$$

where η is the quadratic character of $\mathbb{F}_q$.

Theorem 36 was used in (SWAN, 1962) to prove that there are no irreducible polynomials of the form $x^n + x^k + 1$ over $\mathbb{F}_2$ with degree multiple of 8. Polynomials of this form have important applications in cryptography, such as the efficient representation of a finite field in a computer; see Section 1 of (VON ZUR GATHEN, 2003), for example. We use Theorem 36 in Chapter 3 to investigate the indegree distribution of polynomials over finite fields.

3 DISTRIBUTION OF INDEGREES OF POLYNOMIALS OVER FINITE FIELDS

In Chapter 1 we introduced the motivation for this thesis, namely the cryptographic problems related to iterations of mappings and polynomials over finite fields. Understanding Pollard's method for the factorization of integers remains an important problem of this field 40 years past its publication. The rho length of a node represents just one of the parameters of interest defined over the functional graph of polynomials over a finite field. The Brent-Pollard heuristic (BRENT; POLLARD, 1981) and known asymptotic results on the distribution of these parameters (ARNEY; BENDER, 1982; FLAJOLET; ODLYZKO, 1990; KNUTH, 2011a) lead us to consider the indegree distribution of polynomials over finite fields as a determinant aspect of their statistics.

Previous authors have considered the indegree distribution of polynomials over finite fields in their work, many of them motivated by the problem of determining the value set of polynomials (see Definition 42). If f is a polynomial over $\mathbb{F}_p$, then it is easy to prove that $|f^{-1}(y)| \leq \deg(f)$ for all $y \in \mathbb{F}_p$; see Theorem 18. Hence it is natural to consider d-mappings, defined in Section 2.3.1 as mappings with indegrees bounded by d, as a model for the Brent-Pollard heuristic. However, in the functional graph of a quadratic polynomial modulo an odd prime p, only one node has indegree 1, while the remaining $p-1$ nodes are split in half between nodes with indegree 0 and 2; see Theorem 21. Therefore one might consider more elaborate restrictions on the indegrees of the mappings in the Brent-Pollard heuristic, such as $\mathcal{J}$-mappings; see Definition 20. In this chapter we exhibit known results on the indegree distribution of polynomials over finite fields and present our contributions to the field; they are motivated by our interest in explaining the randomness of

polynomials when seen as mappings.

In Sections 3.1 and 3.2 we survey the known results on the indegree distribution of cubic and quartic polynomials over finite fields, giving new proofs on known results. We also improve one of the cases for quartic polynomials over $\mathbb{F}_q$, with $q = p^e$, $p > 3$ and $e > 1$. Our main contributions are presented in Section 3.3. We determine the asymptotic indegree distribution of the class of general polynomials (BIRCH; SWINNERTON-DYER, 1959) using the results of (COHEN, 1970). This allows us to prove that the asymptotic coalescence of these polynomials is 1. This is an important result in the context of the Brent-Pollard heuristic; the implications are treated in Chapter 4. We prove under a plausible assumption that the indegree distribution and the coalescence of a typical polynomial over $\mathbb{F}_p$ of a fixed degree is dominated by the behavior of the corresponding general polynomials. These results also provide support for our experimental results concerning the Brent-Pollard heuristic, presented in Chapter 4.

In this chapter we investigate the indegree distribution $(n_0, \ldots, n_d)$ of a polynomial $f(x) = a_d x^d + \cdots + a_1 x + a_0$, $a_d \neq 0$, over $\mathbb{F}_q$, $q = p^e$ with p odd. Theorem 20 implies that we can assume without loss of generality that $a_d = 1$ and $a_{d-1} = a_0 = 0$. In the case of a quadratic polynomial, the distribution of (n_0, n_1, n_2) is easily derived from known properties of the quadratic character on finite fields; see Theorem 21. We study next the indegree distribution of cubic and quartic polynomials, before proceeding to a general polynomials of degree $d \geq 2$.

The results of this chapter have been accepted for publication in the International Jounal of Number Theory. The submitted version is available in (MARTINS; PANARIO, 2016).

3.1 Cubic Polynomials

The distribution of the indegrees of a cubic polynomial $f(x) = x^3+bx$ over $\mathbb{F}_q$, $q = p^e$ with $p > 3$, is well known, although it is not easy to determine who was the first to state and prove this. It is easily derived from the size of the value set of the polynomial; see Section 8.3 of (MULLEN; PANARIO, 2013). It was established in (STERNECK, 1908) in the case of $q = p$ a prime number. The result for the general case is stated without proof in (UCHIYAMA, 1954) and it is proved apparently for the first time in (TURNWALD, 1995). We give a proof using Pellet-Stickelberger parity theorem (LIDL; NIEDERREITER, 2008; MULLEN; PANARIO, 2013). For the remainder of this section, unless stated otherwise, $\mathbb{F}_q$ denotes a finite field of characteristic $p > 3$.

It follows from Theorem 36 that the number N_r of solutions of $x^3+bx-r = 0$ is partially determined by its discriminant $D_r = -4b^3 - 27r^2$:

$$N_r = \begin{cases} 0 \text{ or } 3, & \text{if } \eta(D_r) = 1, \\ 1, & \text{if } \eta(D_r) = -1. \end{cases} \tag{3.1}$$

We observe that if $D_r = 0$, we cannot conclude anything using Equation (3.1). If we write

$$\Delta_r = -3^{-1} \cdot D_r = 3^{-1}4b^3 + (3r)^2, \tag{3.2}$$

by the multiplicativity of the quadratic character, we conclude that determining the value that D_r assumes is essentially equivalent to studying Δ_r.

Lemma 1. *We have $\eta(-3^{-1}) = 1$ in $\mathbb{F}_q$ if and only if $q \equiv 1 \,(mod\ 3)$.*

Proof. Let p be the characteristic of $\mathbb{F}_q$ and write $q = p^e$. We note that, if $q \equiv 1 \pmod 3$, then either $p \equiv 1 \pmod 3$ or $p \equiv 2 \pmod 3$ with e even and Theorem 31 implies that -3^{-1} is a square in $\mathbb{F}_q$. Conversely, if $q \equiv 2 \pmod 3$, then $p \equiv 2$

(mod 3) and e is an odd integer. Since $-3^{-1} = (-3)^{-1}$, we have

$$\left(\frac{-3^{-1}}{p}\right) = \left(\frac{-3}{p}\right)^{-1} = \left(\frac{-3}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{3}{p}\right)$$

It follows from Theorem 15 that

$$\left(\frac{-3^{-1}}{p}\right) = (-1)^{\frac{p-1}{2}}(-1)^{\frac{3-1}{2}\frac{p-1}{2}}\left(\frac{p}{3}\right) = \left(\frac{p}{3}\right).$$

Therefore -3^{-1} is a non-square in $\mathbb{F}_p$ and, by Theorem 31, $\eta(-3^{-1}) = -1$. □

The cases $b \neq 0$ and $b = 0$ should be treated separately, as $\eta(\Delta_r)$ has a trivial behaviour in the latter. We consider first the case $b = 0$. It is well known that, if $p \equiv 2 \pmod 3$, then f is a permutation polynomial; see Section 8.1 of (MULLEN; PANARIO, 2013). We give a proof of this fact.

Proposition 1. *For $q = p^e$ with $q \equiv 2 \pmod 3$, the polynomial $f(x) = x^3$ acts as a permutation on $\mathbb{F}_q$.*

Proof. From (3.2) and Lemma 1 we have $\eta(D_r) = -1$ for every non-zero $r \in \mathbb{F}_q$. Hence, by (3.1), every non-zero node in the functional graph of f has indegree 1. The result follows from the fact that the equation $x^3 = 0$ has exactly one solution in $\mathbb{F}_q$. □

Proposition 2. *If $f(x) = x^3 \in \mathbb{F}_q[x]$ with $q \equiv 1 \pmod 3$, then $(n_0, \ldots, n_3)$ is given by*

$$V = \left(\frac{2}{3}(q-1), 1, 0, \frac{1}{3}(q-1)\right).$$

Proof. The equation $x^3 - r = 0$ has discriminant $D_r = -27r^2$. By Lemma 1 we have that $\eta(D_r) = 1$ for every non-zero $r \in \mathbb{F}_q$. By (3.1) this means that $|f^{-1}(r)| = 0$ or 3 for every non-zero r. The equation $x^3 = 0$ has exactly one solution over $\mathbb{F}_q$, hence $n_1 = 1$ and $n_2 = 0$. The result follows from Theorem 6. □

We turn now our attention to the case $b \neq 0$. This implies that $\alpha = 3^{-1}4b^3 \neq 0$. Since (3.1) does not give us any information on the elements $r \in \mathbb{F}_q$ such that $D_r = 0$, we treat them separately with the following lemma.

Lemma 2. *Let $f(x) = x^3 + bx \in \mathbb{F}_q[x]$, $b \neq 0$, and let $\alpha = 3^{-1}4b^3$. Then*

$$n_2 = \begin{cases} 0, & \text{if } \eta(-3b) = -1, \\ 2, & \text{if } \eta(-3b) = 1. \end{cases}$$

Moreover, if $\eta(-3b) = 1$, the elements r with indegree 2 are given by $r = \pm 3^{-1}a$, where a is such that $a^2 = \alpha$.

Proof. The discriminant of the equation $x^3 + bx = r$ is zero if and only if it has multiple roots. This means that either there are $\beta, \gamma \in \mathbb{F}_q$ such that $x^3 + bx - r = (x - \beta)^2(x - \gamma)$ or there is $\beta \in \mathbb{F}_q$ such that $x^3 + bx - r = (x - \beta)^3$. The latter cannot hold if $b \neq 0$ because the coefficient of x^2 in $(x - \beta)^3$, for $\beta \in \mathbb{F}_q^*$, is non-zero. Hence, we have $D_r = \Delta_r = 0$ if and only if r has indegree 2 in the functional graph of f. Moreover, $\Delta_r = 0$ if and only if $(3r)^2 = -\alpha$ and this has a solution for r over $\mathbb{F}_q$ if and only if $\eta(-\alpha) = \eta(-3b) = 1$. If this is the case and a is such that $a^2 = -\alpha$, we have $\Delta_r = 0$ if and only if $3r = \pm a$, as desired. □

Once we determine n_1, the values of n_0 and n_3 are easily computed using Theorem 6. The following lemma shows that n_1 is given by the values assumed by the Legendre symbol on the images of a quadratic polynomial.

Lemma 3. *Let $f(x) = x^3 + bx \in \mathbb{F}_q[x]$, $b \neq 0$, and $\alpha = 3^{-1}4b^3$. Let us also consider*

$$L_{-1} = \#\left\{x \in \mathbb{F}_q \colon \eta(x^2 + \alpha) = -1\right\},$$

$$L_1 = \#\left\{x \in \mathbb{F}_q \colon \eta(x^2 + \alpha) = 1\right\}.$$

Then the number of nodes with indegree 1 in the functional graph of f is given by

$$n_1 = \begin{cases} L_{-1}, & \text{if } q \equiv 1 \pmod 3, \\ L_1, & \text{otherwise.} \end{cases}$$

Proof. Theorem 19 implies that $3r$ runs through all the elements in $\mathbb{F}_q$ when r does the same. Since $\alpha = 3^{-1}4b^3$ is a constant for a given polynomial f, we have that

$$L_{-1} = \#\left\{k \in \mathbb{F}_q \colon \eta(\Delta_r) = -1\right\},$$

$$L_1 = \#\left\{k \in \mathbb{F}_q \colon \eta(\Delta_r) = 1\right\}.$$

The result follows by Lemma 1 that gives, for every $r \in \mathbb{F}_q$ such that $D_r \neq 0$,

$$\eta(D_r) = \begin{cases} \eta(\Delta_r), & \text{if } q \equiv 1 \pmod 3, \\ -\eta(\Delta_r), & \text{otherwise.} \end{cases}$$

□

Next we consider the sum

$$\sum_{x \in \mathbb{F}_q} \eta(x^2 + \alpha). \tag{3.3}$$

Its terms are equal to 1, -1 or 0. Since the number of terms equal to 0 is easily described through $\eta(-3b)$, the values of L_1 and L_{-1} are determined by the value of the sum (3.3).

Proposition 3. *Let $f(x) = x^3 + bx$, $b \neq 0$, be a polynomial over $\mathbb{F}_q$. Then the number n_1 of nodes with indegree equal to 1 in the functional graph of f is given by*

$$n_1 = \frac{1}{2}\big(q - 1 + \eta(-3) - \eta(-3b)\big).$$

Proof. It follows from Theorem 23 that

$$\sum_{x \in \mathbb{F}_q} \eta(x^2 + \alpha) = -\eta(1) = -1,$$

and, as a consequence, $L_{-1} = L_1 + 1$. We have $\eta(x^2 + \alpha) = 0$ for some $x \in \mathbb{F}_q$ if and only if $\eta(-\alpha) = \eta(-3b) = 1$ and, if so, there are two such elements. Hence, $L_{-1} + L_1 = q - (1 + \eta(-3b))$ and

$$(L_{-1}, L_1) = \left(\frac{1}{2}\big(q - \eta(-3b)\big), \frac{1}{2}\big(q - 2 - \eta(-3b)\big)\right).$$

The value of n_1 follows from Lemma 3. □

The distribution of $(n_0, \ldots, n_3)$ for the case $b \neq 0$ follows from Lemma 2, Proposition 3 and Theorem 6. The next theorem summarizes the description of the distribution of indegrees of the functional graph of a cubic polynomial over a finite field of characteristic $p > 3$.

Theorem 37. *If $f(x) = x^3 + bx \in \mathbb{F}_q$, $q = p^e$ with $p > 3$, then $(n_0, \ldots, n_3)$ is given by*

(i) $\left(\frac{1}{3}\big(q - \eta(-3)\big), \frac{1}{2}\big(q - 1 + \eta(-3) - \eta(-3b)\big), 1 + \eta(-3b), \frac{1}{6}\big(q - 3 - 3\eta(-3b) - \eta(-3)\big)\right)$, *if $b \neq 0$;*

(ii) $\left(\frac{2}{3}(q-1), 1, 0, \frac{1}{3}(q-1)\right)$, *if $b = 0$ and $q \equiv 1 \pmod 3$;*

(iii) $(0, p, 0, 0)$, *if $b = 0$ and $q \equiv 2 \pmod 3$.*

Corollary 1. *If $f(x) = x^3 + bx \in \mathbb{F}_q$, $q = p^e$ with $p > 3$, then its coalescence satisfies*

(i) $V = 1 - \big(1 + \eta(-3b) + \eta(-3)\big)/q$, *if $b \neq 0$;*

(ii) $V = 2 - 2/q$, *if $b = 0$ and $q \equiv 1 \pmod 3$;*

(iii) $V = 0$, *if $b = 0$ and $q \equiv 2 \pmod 3$.*

Remark. Theorem 8.3.4 and Remark 8.3.6 of (MULLEN; PANARIO, 2013) state that the size of the value set V_f of a cubic polynomial $f \in \mathbb{F}_q[x]$ that is not a permutation satisfies

$$\lceil q/3 \rceil \leq \#V_f \leq q - \lceil (q-1)/3 \rceil .$$

If $q \equiv 1 \pmod 3$, then the lower bound is achieved by the polynomials in case (ii). The polynomials in case (i) achieve equality in the upper bound whether $q \equiv 1$ or $2 \pmod 3$.

If $b \neq 0$, then

$$(n_0, \ldots, n_3) = \left(\frac{1}{3}(q + c_0), \frac{1}{2}(q + c_1), c_2, \frac{1}{6}(q + c_3) \right),$$

where $c_0, \ldots, c_3$ are bounded as q approaches infinity. The asymptotic distribution in this case satisfies $(n_0/q, \ldots, n_3/q) \sim (1/3, 1/2, 0, 1/6)$, as q approaches infinity. If $b = 0$ and $q \equiv 1 \pmod 3$, then the asymptotic distribution satisfies $(n_0/q, \ldots, n_3/q) \sim (2/3, 0, 0, 1/3)$, as q approaches infinity. The asymptotic distribution of preimage sizes of polynomials over finite fields is considered in the next section, where we study heuristics for the approximation of polynomials over finite fields by random mappings.

The characteristic 3 case is approached by the same arguments. To determine the distribution of $(n_0, \ldots, n_3)$ for $f(x) = x^3 + bx$, we note that the discriminant of $x^3 - bx - r$ is, in this case, given by $D(r) = -4b^3$. This implies $\eta(D(r)) = \eta(-b)$. If $\eta(-b) = 1$, then it follows from (3.1) and Theorem 6 that $(n_0, \ldots, n_3) = (2q/3, 0, 0, q/3)$. If $\eta(-b) = -1$, then, by (3.1), we have $n_1 = q$, that is, f is a permutation polynomial. See Proposition 4.6 of (TURNWALD, 1995) for a proof of the size of the value set of a cubic polynomial in this case.

3.2 Quartic Polynomials

Let $f(x) = x^4 + ax^2 + bx$ be a polynomial over $\mathbb{F}_q$, $q = p^e$ with $p > 3$. As we will see, the following division into cases comes up naturally when studying the distribution of $(n_0, \ldots, n_4)$:

(i) $a = b = 0$ and $q \equiv 1 \pmod 4$;

(ii) $a = b = 0$ and $q \equiv 3 \pmod 4$;

(iii) $a \neq 0$ and $b = 0$;

(iv) $b \neq 0$.

Cases (i) and (ii) correspond to the polynomial $f(x) = x^4$. Precise results for the size of the value set of such polynomials are well know (see Section 8.3 of (MULLEN; PANARIO, 2013)), but we also give a proof of this in this section.

The problem of estimating $(n_0, \ldots, n_4)$ for $q = p$ was studied in (MCCANN; WILLIAMS, 1967), where the authors give asymptotic estimates as q approaches infinity. The cases (i)-(iii) are estimated with an error term of $O(1)$, but the authors say it is possible to give precise values in these cases. The distribution in case (iv) was estimated with an error term of $O(p^{1/2})$. Theorems 2.2 and 2.3 of (SUN, 2006) give an explicit expression for the value set of a polynomial $x^4+ax^2+bx \in \mathbb{F}_p[x]$ with $b \neq 0$, but it does not appear to be easy to derive an explicit enumeration result from this. See (SUN, 2006) also for connections with elliptic curves and several special cases where the value set of such a polynomial is determined explicitly.

The case $q = p^e$ with $e > 1$ was solved in (COHEN, 1970) by Cohen with an error term of $O(q^{1/2})$ in all cases, as q approaches infinity. We improve the results

of (COHEN, 1970) and (MCCANN; WILLIAMS, 1967) in case (iii) by giving exact results. We also give a new proof for case (iv). This proof is short and elementary: we use results that are proved here using elementary techniques and a result of WILLIAMS (1967) on the value set of these polynomials. Lemma 4 below is used in this case as well as in cases (i)-(iii).

Lemma 4. *If $f(x) = x^4 + ax^2 + bx \in \mathbb{F}_q[x]$ then $n_3 \leq 3$. Moreover, if $b = 0$, then*

$$n_3 = \begin{cases} 0, & \text{if } a = 0, \\ \frac{1}{2}\big(1 + \eta(-a)\big), & \text{otherwise.} \end{cases}$$

Proof. Let f be a quartic polynomial as above and assume that $f(x) - r = 0$ has exactly three distinct roots in $\mathbb{F}_q$. It follows that f factors completely into linear factors over $\mathbb{F}_q$ with exactly one squared linear factor. As a consequence, a necessary condition for r to be a node of indegree 3 is that its discriminant $D(r)$, given by

$$D(r) = -256r^3 - 128a^2r^2 - (16a^4 + 144ab^2)r - (4a^3b^2 + 27b^4), \tag{3.4}$$

be zero. This is a cubic polynomial in r, hence $n_3 \leq 3$.

Assume now that $b = 0$ and write $f(x) = g(x^2)$, with $g(w) = w^2 + aw$. If $g(w) - r$ has 0 or 1 distinct solutions in $\mathbb{F}_q$, then $|f^{-1}(r)| \leq 2$. If $g(w) - r = 0$ has distinct solutions ω_1, ω_2 in $\mathbb{F}_q$, then the roots of $f(x) - r$ are given by $x^2 = \omega_1$ and $x^2 = \omega_2$. If ω_1 and ω_2 are non-zero, then $|f^{-1}(r)|$ is even. Therefore, the preimage size of an element r is 3 if and only if $\omega_1 = 0$ and $\eta(\omega_2) = 1$. The condition $\omega_1 = 0$ implies $r = 0$ and $\omega_2 = -a$, hence we have $|f^{-1}(r)| = 3$ if and only if $r = 0$ and $\eta(-a) = 1$. □

We derive estimates with different arguments in the cases $b = 0$ and $b \neq 0$. In the first case we have $f(x) = g(x^2)$ for some quadratic polynomial $g(w) \in$

$\mathbb{F}_q[w]$. This allows us to not only give more elementary proofs but to obtain explicit enumeration results.

Case $\mathbf{b = 0}$. If $f(x) = g(x^2)$ for some quadratic polynomial $g(w) \in \mathbb{F}_q[w]$, we prove in Lemma 5 below that $n_1 \leq 1$; we give a precise result on this. Moreover, we determine precisely the cardinality of its value set. Since $n_0 = q - \#V_f$, the distribution of $(n_0, \ldots, n_4)$ follows using Theorem 6.

Lemma 5. *If $f(x) = x^4 + ax^2 \in \mathbb{F}_q[x]$ then*

$$n_1 = \begin{cases} 1, & \text{if } a = 0, \\ \frac{1}{2}\big(1 - \eta(-a)\big), & \text{if } a \neq 0. \end{cases}$$

Proof. If $a = 0$, the equation $x^4 = r$ has exactly one root in $\mathbb{F}_q$ if and only if $r = 0$. Let $a \neq 0$ and write $f(x) = g(x^2)$ with $g(w) = w^2 + aw = (w + a/2)^2 - a^2/4$. The polynomial g has exactly one root in $\mathbb{F}_q$ if and only if $r = -a^2/4$. The roots of f in $\mathbb{F}_q$ are in this case given by $z^2 = -a/2$, hence $|f^{-1}(-a^2/4)| = 0$ or 2.

If g has distinct roots ω_1, ω_2 in $\mathbb{F}_q$, then $f(x) = 0$ if and only if $x^2 = \omega_1$ or $x^2 = \omega_2$. This gives exactly one root for f in $\mathbb{F}_q$ if and only if $\omega_1 = 0$ and $\eta(\omega_2) = -1$. We have $g(0) = 0$ if and only if $r = 0$ and $\omega_2 = -a$. It follows that there is an element r with preimage size 1 in the case $a \neq 0$ if and only if $\eta(-a) = -1$ and, if so, $n_1 = 1$. □

It is enough to determine the value of $n_0 = q - \#V_f$ to finish the case $b = 0$. The following two lemmas determine the size of the value set of such polynomials. The first one is a standard result and can be found in Section 8.3 of (MULLEN; PANARIO, 2013), but we give a proof for completeness using our framework.

Lemma 6. *The value set of a polynomial $f(x) = x^4 \in \mathbb{F}_q[x]$ has cardinality $1 + (q - 1)/\gcd(q - 1, 4)$.*

Proof. We have $\#V_f = 1 + \#V_f^*$, where $V_f^* = \{w^2 \colon\ w \in \mathbb{F}_q^* \text{ is a square}\}$. The set of all squares of the multiplicative group $\mathbb{F}_q^*$ has cardinality $(q-1)/2$; see Section 2.4.2. The cardinality of V_f^* is given by the difference between $(q-1)/2$ and the number of collisions $w_1^2 = w_2^2$, for w_1, w_2 distinct squares of $\mathbb{F}_q$. This happens if and only if $w_2 = -w_1$, that is, if $\eta(-1) = 1$. Theorems 15 and 31 imply that this is true if and only if $q \equiv 1 \pmod 4$, so

$$\#V_f^* = \begin{cases} (q-1)/4, & \text{if } q \equiv 1 \pmod 4, \\ (q-1)/2, & \text{if } q \equiv 3 \pmod 4, \end{cases}$$

that is, $\#V_f^* = (q-1)/\gcd(q-1, 4)$. □

The value set of quartic polynomials $f(x) = x^4 + ax^2 \in \mathbb{F}_q[x]$ with $a \neq 0$ has already been determined in the case $q = p$ (STERNECK, 1908); see also (MCCANN; WILLIAMS, 1967). As far as we know, we are the first to give a precise result in the case $q = p^e$ with $p > 3$ and $e > 1$.

Proposition 4. *If $f(x) = x^4 + ax^2 \in \mathbb{F}_q[x]$ with $a \neq 0$, then*

$$\#V_f = \frac{1}{8}\big(3q + 4 - 2\eta(-a) + \eta(-1) + 2\eta(-2a)\big).$$

Proof. We have

$$\#V_f = \#\{w^2 + aw \colon\ w \in \mathbb{F}_q \text{ is a square}\}$$

and, by completing squares, we can write

$$\#V_f = \#\left\{\left(w + \frac{a}{2}\right)^2 - \frac{a^2}{4} \colon\ w \in \mathbb{F}_q \text{ is a square}\right\}.$$

It is easy to see that, if g, h are polynomials over $\mathbb{F}_q$ such that $h(x) = g(x) + \alpha$, for some $\alpha \in \mathbb{F}_q$, then $\#V_g = \#V_h$. Hence,

$$\#V_f = \#\left\{\left(w + \frac{a}{2}\right)^2 \colon\ w \in \mathbb{F}_q \text{ is a square}\right\}.$$

Since the number of squares in $\mathbb{F}_q$ is $(q+1)/2$, if we determine the number of distinct squares w_1, w_2 of $\mathbb{F}_q$ such that $(w_1 + a/2)^2 = (w_2 + a/2)^2$, we are able to estimate $\#V_f$. We note that this happens if and only if

$$w_1 + \frac{a}{2} = -\left(w_2 + \frac{a}{2}\right), \text{ that is, } w_1 + a = -w_2. \tag{3.5}$$

We treat the case where $w_2 = 0$ separately and determine the number of squares w of $\mathbb{F}_q$ such that $\eta(w + a) = \eta(-1)$.

Let $S = \sum_{x \in \mathbb{F}_q} \eta(x^2 + a)$. We use Theorem 23 to determine how many terms are equal to 0, 1 and -1 in S. This allows us to determine how many squares w of $\mathbb{F}_q$ are such that $\eta(w + a) = \eta(-1)$. We treat the term $\eta(a)$ separately because, with the exception of this term, there is a 2-to-1 correspondence between the terms of S and the squares of $\mathbb{F}_q$. Write $S = \eta(a) + S'$ and let N_1 and N_{-1} be the number of terms equal to 1 and -1 in S'. Since the sum S' contains $q - 1 - \big(1 + \eta(-a)\big)$ non-zero terms, it follows from Theorem 23 that

$$\begin{cases} N_1 + N_{-1} = q - 1 - \big(1 + \eta(-a)\big), \\ N_1 - N_{-1} = -\big(1 + \eta(a)\big). \end{cases}$$

Thus $N_1 = \big(q - 3 - \eta(-a) - \eta(a)\big)/2$ and $N_{-1} = \big(q - 1 - \eta(-a) + \eta(a)\big)/2$. Therefore, the number of terms in S' equal to $\eta(-1)$ is $\big(q - 1 - (1 + \eta(-1)) - \eta(-a) - \eta(-1)\eta(a)\big)/2$. This implies that the number of non-zero squares w such that $\eta(w+a) = \eta(-1)$ is $\big(q-2-2\eta(-a)-\eta(-1)\big)/4$. Since $\eta(a) = \eta(-1)$ if and only if $\eta(-a) = 1$, it follows that the number N of squares w such that $\eta(w+a) = \eta(-1)$ is given by

$$N = \frac{1}{4}\big(q - 2 - 2\eta(-a) - \eta(-1)\big) + \frac{1}{2}\big(1 + \eta(-a)\big) = \frac{1}{4}\big(q - \eta(-1)\big). \tag{3.6}$$

We note that Equation (3.5) holds for $w_2 = 0$ if and only if $\eta(-a) = 1$. Hence the total number of pairs w_1, w_2 such that (3.5) holds, with w_1, w_2 not necessarily distinct, is $N + \big(1 + \eta(-a)\big)/2$. We note that (3.5) holds for $w_1 = w_2$ if and only if

$w_1 = -a/2$, and this holds if and only if $\eta(-2a) = 1$. It follows from (3.6) that the number of pairs of distinct squares w_1, w_2 such that (3.5) holds is

$$N + \frac{1}{2}\big(1 + \eta(-a)\big) - \frac{1}{2}\big(1 + \eta(-2a)\big) = \frac{1}{4}\big(q + 2\eta(-a) - \eta(-1) - 2\eta(-2a)\big).$$

Therefore,

$$\begin{aligned} \#V_f &= \frac{q+1}{2} - \frac{1}{2} \cdot \frac{1}{4}\big(q + 2\eta(-a) - \eta(-1) - 2\eta(-2a)\big) \\ &= \frac{1}{8}\big(3q + 4 - 2\eta(-a) + \eta(-1) + 2\eta(-2a)\big), \end{aligned}$$

as desired. □

Case b $\neq$ 0. The case of polynomials $f(x) = x^4 + ax^2 + bx$ with $b \neq 0$ is of a different nature because we cannot write $f(x) = g(x^2)$ for any $g \in \mathbb{F}_q[x]$. In addition to the equations of Theorem 6, we estimate the sum $n_1 + 4n_2 + 9n_3 + 16n_4$ up to an error term of $O(q^{1/2})$. The argument for this estimate is analogous to the one of (MCCANN; WILLIAMS, 1967).

Proposition 5. *Let $f(x) = x^4 + ax^2 + bx$ be a polynomial over $\mathbb{F}_q$ with $b \neq 0$. Then*

$$n_1 + 4n_2 + 9n_3 + 16n_4 = 2q + O(q^{1/2}).$$

Proof. Define $N_r = |f^{-1}(r)|$, for $r \in \mathbb{F}_q$. We note that, if $\mathcal{N}_j$ denotes the set of elements of $\mathbb{F}_q$ with preimage size j, then

$$\sum_{j=1}^{4} j^2 n_j = \sum_{j=1}^{4} \sum_{x \in \mathcal{N}_j} j^2 = \sum_{j=1}^{4} \sum_{x \in \mathcal{N}_j} |f^{-1}(x)|^2 = \sum_{x \in \mathbb{F}_q} |f^{-1}(x)|^2 = N_f,$$

where N_f is the number of solutions (x, y) in $\mathbb{F}_q$ to the equation $f(x) = f(y)$. If N'_f is the number of such solutions with $x \neq y$, then

$$N_f = q + N'_f. \tag{3.7}$$

We note that, if $x \neq y$, then $f(x) = f(y)$ if and only if

$$(x+y)(x^2+y^2+a) = -b. \tag{3.8}$$

Since this is a cubic equation, the number of solutions (x, y) of (3.8) with $x = y$ is at most 3. Hence, the total number N''_f of solutions of (3.8) is such that

$$N'_f \leq N''_f \leq N'_f + 3. \tag{3.9}$$

We estimate N''_f by noting first that, since $b \neq 0$, $x + y$ and $x^2 + y^2 + a$ are both non-zero. Therefore we can write N''_f as

$$N''_f = \sum_{t \in \mathbb{F}_q^*} N_{f,t}, \tag{3.10}$$

where $N_{f,t}$ is the number of solutions of the equations

$$\begin{cases} x^2 + y^2 + a = t, \\ x + y = -bt^{-1}. \end{cases}$$

This is equivalent to

$$x^2 + (-bt^{-1} - x)^2 + a = t,$$

that is,

$$x^2 + bt^{-1}x + 2^{-1}(b^2t^{-2} + a - t) = 0.$$

We complete the square and note that this is equivalent to

$$(x + 2^{-1}bt^{-1})^2 = 2^{-2}b^2t^{-2} - 2^{-1}(b^2t^{-2} + a - t).$$

It follows that the number $N_{f,t}$ of solutions to this equation is given by

$$N_{f,t} = 1 + \eta\big(2^{-2}b^2t^{-2} - 2^{-1}(b^2t^{-2} + a - t)\big) = 1 + \eta(b^2t^{-2} - 2(b^2t^{-2} + a - t)),$$

that is,

$$N_{f,t} = 1 + \eta(2t - 2a - b^2t^{-2}) = 1 + \eta(2t^3 - 2at^2 - b^2). \tag{3.11}$$

We used above the fact that the quadratic character η is multiplicative and $\eta(4) = \eta(t^2) = 1$. It follows from (3.10), (3.11) and Theorem 24 that

$$N_f'' = \sum_{t\in\mathbb{F}_q^*} \big(1 + \eta(2t^3 - 2at^2 - b^2)\big) = q - 1 + \sum_{t\in\mathbb{F}_q^*} \eta(2t^3 - 2at^2 - b^2) = q + O(q^{1/2}).$$

The result follows by (3.9) and (3.7). □

A proof of the indegree distribution of polynomials of the form $x^4+ax^2+bx \in \mathbb{F}_q[x]$, $b \neq 0$, follows easily now. It is proved in (WILLIAMS, 1967) that the value set of these polynomials satisfies $\#V_f = 5q/8 + O(q^{1/2})$, hence $n_0 = 3q/8 + O(q^{1/2})$. Since $n_3 = O(1)$ by Lemma 4, all that remains to do is to estimate n_1, n_2 and n_4. These follow at once from the equations obtained in Theorem 6 and Proposition 5.

Theorem 38. *Let $f(x) = x^4 + ax^2 + bx$ be a polynomial over $\mathbb{F}_q$, $q = p^e$ with $p > 3$. Then $(n_0, \ldots, n_4)$ is given by*

(i) $\left(\frac{3}{4}(q-1), 1, 0, 0, \frac{1}{4}(q-1)\right)$, *if $a = b = 0$ and $q \equiv 1 \pmod 4$;*

(ii) $\left(\frac{1}{2}(q-1), 1, \frac{1}{2}(q-1), 0, 0\right)$, *if $a = b = 0$ and $q \equiv 3 \pmod 4$;*

(iii) $\Big(\frac{1}{8}\big(5q - 4 + 2\eta(-a) - t\big), \frac{1}{2}\big(1 - \eta(-a)\big), \frac{1}{4}\big(q + t\big), \frac{1}{2}\big(1 + \eta(-a)\big),$
$\frac{1}{8}\big(q - 4 - 2\eta(-a) - t\big)\Big)$, *if $a \neq 0, b = 0$, where $t = \eta(-1) + 2\eta(-2a)$;*

(iv) $\left(\frac{3q}{8} + O(q^{1/2}), \frac{q}{3} + O(q^{1/2}), \frac{q}{4} + O(q^{1/2}), O(1), \frac{q}{24} + O(q^{1/2})\right)$, *if $b \neq 0$, as q approaches infinity.*

Remark. The polynomials in case (i) are minimum value set polynomials; see Remark 8.3.6 of (MULLEN; PANARIO, 2013).

It should be clear that the asymptotic distribution of indegrees in the case of quartic general polynomials satisfies $(n_0/q, \ldots, n_4/q) \sim (3/8, 1/3, 1/4, 0, 1/24)$, as q approaches infinity. If $a \neq 0$, then the asymptotic distribution of polynomials of the form $f(x) = x^4 + ax^2$ satisfies $(n_0/q, \ldots, n_4/q) \sim (5/8, 0, 1/4, 0, 1/8)$, as q approaches infinity.

We are able to derive precise results for the coalescence of quartic polynomials in the first three cases. This is done in the corollary below. The coalescence of the quartic polynomials in case (iv) are a particular case of the results of the next section; see Theorems 42 and 45.

Corollary 2. *If $f(x) = x^4 + ax^2 + bx$ is a polynomial over $\mathbb{F}_q$, $q = p^e$ with $p > 3$, then its coalescence satisfies*

(i) $V = 3 - 3/q$, *if* $a = b = 0$ *and* $q \equiv 1 \pmod 4$;

(ii) $V = 1 - 1/q$, *if* $a = b = 0$ *and* $q \equiv 3 \pmod 4$;

(iii) $V = 2 - (3 + t)/q$, *if* $a \neq 0, b = 0$, *where* $t = \eta(-1) + 2\eta(-2a)$.

3.3 General Polynomials

In this section we consider general polynomials. This class was introduced in (BIRCH; SWINNERTON-DYER, 1959). The authors were interested in the value set of polynomials over finite fields and proved the result that we state in Theorem 39. We recall that $\overline{\mathbb{K}}$ denotes the algebraic closure of a finite field $\mathbb{K}$ and S_d is the symmetric group on d elements.

Definition 58. *Let f be a polynomial over $\mathbb{F}_q$ of degree $d \geq 2$. Let t be a transcendental element over $\mathbb{F}_q$ and $\overline{G}$ be the Galois group of $f(x) - t$ over $\overline{\mathbb{F}}_q(t)$. We say*

that f is a general polynomial if $\overline{G} = S_d$.

Theorem 39 ((BIRCH; SWINNERTON-DYER, 1959)). *For general polynomials $f \in \mathbb{F}_q[x]$ we have*

$$\#V_f = \left(1 - \frac{1}{2!} - \cdots - \frac{(-1)^d}{d!}\right) q + O(\sqrt{q}).$$

We note that an estimate on the size of the value set of polynomials $f \in \mathbb{F}_q[x]$ represents a partial result on the indegree distribution of such functions. Cohen (COHEN, 1970) approaches this problem, for a polynomial $f(x)$ of any degree $d \geq 2$, by considering $f(x) - t$ as a polynomial in x over the field $\mathbb{F}_q(t)$, as Birch and Swinnerton-Dyer do in (BIRCH; SWINNERTON-DYER, 1959). He relates the factorization of a polynomial f into irreducible factors with the decomposition of permutations of S_d into cycles.

In order to state Cohen's theorem, we establish the following notation. Let f be a polynomial of degree $d \geq 2$ over $\mathbb{F}_q$. If f has exactly a_j irreducible factors of degree j in its factorization over $\mathbb{F}_q$, for $j = 1, \ldots, d$, then f has *cycle pattern* $\Lambda = 1^{a_1} \ldots d^{a_d}$; a permutation $\sigma \in S_d$ that decomposes into a_j cycles of length j, for $j = 1, \ldots, d$, has same cycle pattern $\Lambda = 1^{a_1} \ldots d^{a_d}$. Let t be a transcendental element over $\mathbb{F}_q$ and G be the Galois group of $f(x) - t$ over $\mathbb{F}_q(t)$, with splitting field K. For any $\sigma \in G$, we define K_σ to be the subfield of K fixed under σ. We define k to be the largest algebraic extension of $\mathbb{F}_q$ in K and $G^* = \{\sigma \in G \colon K_\sigma \cap k = \mathbb{F}_q\}$. For any H contained in S_d and any cycle pattern Λ, we denote by H_Λ the set of permutations of H with cycle pattern Λ.

Cohen considers the factorization of $f(x) - rg(x)$ for $r \in \mathbb{F}_q$ and fixed $f, g \in \mathbb{F}_q[x]$. We state his result in the particular case $g(x) = 1$ of interest in this paper.

Theorem 40. *Let f be a polynomial over $\mathbb{F}_q$ of degree $d \geq 2$. Then the number $\pi(f,q)$ of elements $r \in \mathbb{F}_q$ such that $f(x) - r$ has cycle pattern Λ satisfies*

$$\pi(f,q) = \frac{|G^*_\Lambda|}{|G^*|}q + O(q^{1/2}),$$

where the implied constant depends only on d.

As far as we know, there is no general method to determine the Galois group of interest. For the case where $G^* = S_d$, the asymptotic estimate for $\pi(f,q)$ follows from the ratio of permutations of S_d with a given cycle pattern. This is guaranteed to hold if the Galois group of $f(x) - t$ over $\overline{\mathbb{F}}_q(t)$ is S_d, that is, if f is a general polynomial. We focus on this case for the remainder of this section.

Theorem 41. *If f is a general polynomial over $\mathbb{F}_q$ of degree $d \geq 2$, then the number of nodes with indegree k in its functional graph satisfies $n_k = P_{d,k} \cdot q + O(q^{1/2})$, where*

$$P_{d,k} = \frac{1}{k!}\sum_{\ell=0}^{d-k}\frac{(-1)^\ell}{\ell!}.$$

Moreover, the implied constant depends only on d.

Proof. It follows from Theorem 40 that the number of nodes with indegree k in the functional graph of a general polynomial $f \in \mathbb{F}_q[x]$ of degree d, for $k = 0, \ldots, d$, is given by $n_k = P_{d,k} \cdot q + O(q^{1/2})$, for some $0 \leq P_{d,k} < 1$. If $r \in \mathbb{F}_q$ is such that the discriminant $D(r)$ of $f(x) - r$ is non-zero, then $f(x) - r$ has no multiple roots; see Definition 57. It follows from Theorem 17 that the preimage size of $r \in \mathbb{F}_q$ is given by the number of linear factors in the factorization of $f(x) - r$. Since $D(r)$ is a polynomial in r of degree at most $d - 1$, it follows that we have $D(r) = 0$ for at most $d - 1$ elements $r \in \mathbb{F}_q$. Thus $P_{d,k}$ is given by the ratio of permutations σ in S_d that have exactly k fixed points, that is, k cycles of length 1.

Explicit expressions for $P_{d,k}$ for every $d \geq 1$ and $0 \leq k \leq d$, are well known. We use exponential generating functions for this; see (FLAJOLET; SEDGEWICK, 2009) for a nice treatment of the subject. If $(T_d)_{d\geq 1}$ is the counting sequence of permutations on d elements, that is, $T_d = d!$, then its exponential generating function EGF is defined as $T(z) = \sum_d T_d \cdot z^d/d!$. Therefore,

$$T(z) = \sum_{d=1}^{\infty} z^d = \frac{1}{1-z}.$$

Using the symbolic method one interprets the class of permutations as sets of cycles, where these combinatorial constructions correspond to the exponential and logarithmic functions, respectively; see Page 120 of (FLAJOLET; SEDGEWICK, 2009). We are thus able to write

$$T(z) = \exp\left(\log \frac{1}{1-z}\right) = \frac{1}{1-z}. \tag{3.12}$$

This alternative expression for $T(z)$ allows us to introduce a new variable to keep track of the number of fixed points. The bivariate EGF of the double sequence $T_{d,k}$ is defined as

$$T(z,u) = \sum_{d=1}^{\infty}\sum_{k=0}^{\infty} T_{d,k} u^k \frac{z^d}{d!} = \sum_{d=1}^{\infty}\sum_{k=0}^{\infty} P_{d,k} u^k z^d. \tag{3.13}$$

Fixed points (cycles of length 1) correspond to the term z in the expansion of $\log(1-z)^{-1}$. We subtract the term z from this expansion and add it back coupled with the new variable u; see Page 175 of (FLAJOLET; SEDGEWICK, 2009). We conclude from Equation (3.12) that

$$T(z,u) = \exp\left(\log \frac{1}{1-z} - z + uz\right) = \frac{1}{1-z} e^{(u-1)z}. \tag{3.14}$$

It is clear from Equation (3.13) that $P_{d,k}$ is given by the coefficient of $u^k z^d$ in the expansion of Equation (3.14) as a power series. It follows from $e^{uz} = \sum_k (uz)^k/k!$ that the coefficient of u^k in $T(z,u)$ is given by

$$[u^k]T(z,u) = \frac{z^k}{k!}\frac{e^{-z}}{1-z}. \tag{3.15}$$

The coefficient of z^d in Equation (3.15) is given by the following expansion:

$$\frac{z^k e^{-z}}{1-z} = z^k \sum_{j\geq 0} z^j \sum_{\ell \geq 0} \frac{(-1)^\ell}{\ell!} z^\ell = z^k \sum_{j,\ell \geq 0} \frac{(-1)^\ell}{\ell!} z^{j+\ell},$$

that is,

$$\frac{z^k e^{-z}}{1-z} = z^k \sum_{d\geq 0} \left(\sum_{\ell=0}^{d} \frac{(-1)^\ell}{\ell!} \right) z^d = \sum_{d \geq k} \left(\sum_{\ell=0}^{d-k} \frac{(-1)^\ell}{\ell!} \right) z^d. \tag{3.16}$$

The result follows from Equations (3.13), (3.15) and (3.16). □

Using generating functions we are also able to prove that the coalescence of general polynomials is asymptotically 1; see Definition 17. We use this result in connection to the Brent-Pollard heuristic in Chapter 4.

Theorem 42. *The coalescence $V(f)$ of general polynomials $f \in \mathbb{F}_q[x]$ of fixed degree $d \geq 2$ satisfies*

$$V(f) = 1 + O(q^{-1/2}),$$

as q approaches infinity, where the implied constant depends only on d.

Proof. For a general polynomial f over $\mathbb{F}_q$, consider the random variable Z defined on the elements of $\mathbb{F}_q$ as $Z(r) = |f^{-1}(r)|$, where all the elements of $\mathbb{F}_q$ are equally likely. If $\mathcal{N}_k = \{r \in \mathbb{F}_q \colon\ |f^{-1}(r)| = k\}$, it follows from Equation (2.1) that

$$\begin{aligned} V(f) = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2 &= \sum_{r \in \mathbb{F}_q} \frac{1}{q} Z(r)^2 - 1 \\ &= \sum_{r \in \mathbb{F}_q} \frac{1}{q} |f^{-1}(r)|^2 - 1 = \sum_{k=0}^{d} \sum_{r \in \mathcal{N}_k} \frac{1}{q} k^2 - 1. \end{aligned}$$

It follows from Theorem 41 that

$$V(f) + 1 = \sum_{k=0}^{d} \left(P_{d,k} \cdot q + O(q^{1/2}) \right) \frac{1}{q} k^2 = \sum_{k=0}^{d} P_{d,k} k^2 + O(q^{-1/2}) \sum_{k=0}^{d} k^2,$$

that is,

$$V(f) + 1 = \sum_{k=0}^{d} P_{d,k} k^2 + O(q^{-1/2}). \tag{3.17}$$

If $T(z, u)$ is the EGF defined in Equation (3.13), then

$$(u\partial_u)^2 T(z, u) = \sum_{d=0}^{\infty} \sum_{k=1}^{d} k^2 P_{d,k} u^k z^d. \tag{3.18}$$

Thus we are able to write the sum in the right-hand side of Equation (3.17) using the coefficient of z^d in the evaluation of Equation (3.18) at $u = 1$:

$$V(f) + 1 = [z^d](u\partial_u)^2 T(z, u)\Big|_{u=1} + O(q^{-1/2}). \tag{3.19}$$

We note that the connection between the second derivative of a bivariate generating function and the variance of the random variable that it describes is highlighted in Section 3.2 of (FLAJOLET; SEDGEWICK, 2009). Using Equation (3.14) we obtain

$$[z^d](u\partial_u)^2 T(z, u)\Big|_{u=1} = [z^d] \left(\frac{z}{1-z} + \frac{z^2}{1-z} \right) = 2, \tag{3.20}$$

for $d \geq 2$. The result follows from Equations (3.19) and (3.20). □

Theorem 41 implies that general polynomials present a very specific asymptotic distribution of indegrees. The experiments presented in Chapter 4 suggest that most polynomials over $\mathbb{F}_q$ of a given degree $d \geq 2$ have similar indegree distribution. This is proved next as a conditional result, where we assume that the probability that a random uniform polynomial $f \in \mathbb{F}_q$ of degree $d \geq 2$ is general is asymptotically 1. It is proved in (WILLIAMS, 1967) that this assumption holds for the polynomials whose value set satisfies the estimate in Theorem 39. Although this is not proved for general polynomials, we note that Birch and Swinnerton-Dyer mention in (BIRCH; SWINNERTON-DYER, 1959) that *"We therefore obtain simple sufficient condition for $\overline{G}$ to be the full symmetric group. Though not necessary, these are in fact satisfied by almost all polynomials."*

Theorem 43. *Let $d \geq 2$ and consider the probability space defined by the uniform measure on the set $\Omega_{q,d}$ of polynomials of degree $d \geq 2$ over $\mathbb{F}_q$. Assume that the number N of general polynomials of the form $f(x) = a_d x^d + \cdots + a_1 x + a_0 \in \mathbb{F}_q[x]$ satisfies $N = q^{d+1}\big(1 + O(1/q)\big)$. Let $X_{q,k}(f)$, $f \in \Omega_{q,d}$, be the random variable representing the number of elements with preimage size k under f and let*

$$P_{d,k} = \frac{1}{k!} \sum_{\ell=0}^{d-k} \frac{(-1)^\ell}{\ell!}.$$

Then, the sequence $(X_{q,k}/P_{d,k} \cdot q)_q$ converges in probability to 1:

$$\frac{X_{q,k}}{P_{d,k} \cdot q} \xrightarrow{P} 1.$$

Moreover, the convergence in probability is uniform with respect to k in the following sense: for every $\epsilon > 0$, there exists a constant $M = M(d, \epsilon)$ such that, for sufficiently large q,

$$\mathbb{P}\left[\left|\frac{X_{q,k}}{P_{q,k} \cdot q} - 1\right| \geq \epsilon\right] \leq \frac{M}{q} = o(1),$$

for $k = 0, 1, \ldots, d$.

Proof. We prove that the expectation and the standard deviation of $X_{q,k}$, denoted by $\mu_{q,k}$ and $\sigma_{q,k}$ respectively, satisfy $\sigma_{q,k} = o(\mu_{q,k})$ as q approaches infinity. All big-Oh estimates in this proof are taken as q approaches infinity. Let $\Omega^g_{q,d}$ be the set of general polynomials in $\Omega_{q,d}$. Using Theorem 41 and the fact that $X_{q,k}(f) \leq q$ for every $f \in \Omega_{q,d}$, it follows that

$$\begin{aligned}\mu_{q,k} &= \sum_{f \in \Omega_{q,d}} \frac{1}{|\Omega_{q,d}|} X_{q,k}(f) = \sum_{f \in \Omega^g_{q,d}} \frac{1}{|\Omega_{q,d}|} X_{q,k}(f) + \sum_{f \in \Omega_{q,d} \setminus \Omega^g_{q,d}} \frac{1}{|\Omega_{q,d}|} X_{q,k}(f) \\ &= \sum_{f \in \Omega^g_{q,d}} \frac{1}{q^{d+1}} \big(P_{d,k} \cdot q + O(q^{1/2})\big) + \sum_{f \in \Omega_{q,d} \setminus \Omega^g_{q,d}} \frac{1}{q^{d+1}} O(q).\end{aligned}$$

It follows from our assumption that

$$\begin{aligned}\mu_{q,k} &= q^{d+1}\big(1 + O(1/q)\big) \frac{1}{q^{d+1}} \big(P_{d,k} \cdot q + O(q^{1/2})\big) + O(q^d) \frac{1}{q^{d+1}} O(q) \\ &= P_{d,k} \cdot q + O(q^{1/2}).\end{aligned} \tag{3.21}$$

Furthermore, the variance of $X_{q,k}$ satisfies

$$\begin{aligned}\sigma_{q,k}^2 &= \mathbb{E}[(X_{q,k}-\mu_{q,k})^2] \\ &= \sum_{f\in\Omega_{q,d}^g} \frac{1}{|\Omega_{q,d}|}(X_{q,k}(f)-\mu_{q,k})^2 + \sum_{f\in\Omega\setminus\Omega_{q,d}^g} \frac{1}{|\Omega_{q,d}|}(X_{q,k}(f)-\mu_{q,k})^2.\end{aligned}$$

Using Equation (3.21) and the fact that $|X_{q,k}(f)-\mu_{q,k}| \le 2q$ for every $f \in \Omega_{q,d}$, it follows that

$$\begin{aligned}\sigma_{q,k}^2 &= \sum_{f\in\Omega_{q,d}^g} \frac{1}{q^{d+1}}\big(O(q^{1/2})\big)^2 + \sum_{f\in\Omega\setminus\Omega_{q,d}^g} \frac{1}{q^{d+1}}O(q^2) \\ &= q^{d+1}\big(1+O\left(1/q\right)\big)\frac{1}{q^{d+1}}O(q) + O(q^d)\frac{1}{q^{d+1}}O\big(q^2\big),\end{aligned}$$

hence $\sigma_{q,k}^2 = O(q)$. We note that

$$\begin{aligned}\left|\frac{X_{q,k}}{P_{d,k}\cdot q}-1\right| &= \left|\frac{X_{q,k}}{\mu_{q,k}}\frac{1}{1+O(q^{-1/2})}-1\right| \\ &= \left|\frac{X_{q,k}}{\mu_{q,k}}-1+O(q^{-1/2})\right|\frac{1}{(1+O(q^{-1/2}))},\end{aligned}$$

and, by the triangle inequality,

$$\left|\frac{X_{q,k}}{P_{d,k}\cdot q}-1\right| \le \left(\left|\frac{X_{q,k}}{\mu_{q,k}}-1\right|+O(q^{-1/2})\right)\frac{1}{(1+O(q^{-1/2}))}.$$

Let $\epsilon > 0$. The equation above implies that

$$\mathbb{P}\left[\left|\frac{X_{q,k}}{P_{d,k}\cdot q}-1\right| \ge \epsilon\right] \le \mathbb{P}\left[\left|\frac{X_{q,k}}{\mu_{q,k}}-1\right| \ge \epsilon(1+O(q^{-1/2}))\right].$$

Since the standard deviation of $X_{q,k}/\mu_{q,k}$ is $\sigma_{q,k}/\mu_{q,k}$, it follows by the Chebyshev-Bienayme Inequality (Theorem 5) that

$$\mathbb{P}\left[\left|\frac{X_{q,k}}{P_{d,k}\cdot q}-1\right| \ge \epsilon\right] \le \left(\frac{1}{\epsilon(1+O(q^{-1/2}))}\frac{\sigma_{q,k}}{\mu_{q,k}}\right)^2. \tag{3.22}$$

The big-Oh estimate in Theorem 41 does not depend on k, thus the big-Oh estimates in this proof do not depend on k. In particular, there exists a constant $M_1 = M_1(d)$ such that $\sigma_{q,k}^2 \le M_1 q$ for sufficiently large q. We note that Equation (3.21) implies

that, if $P_d = \min_k P_{d,k}$, then, for sufficiently large q, $\mu_{q,k} \geq P_d \cdot q/2$. It follows from Equation (3.22) that, for sufficiently large q,

$$\mathbb{P}\left[\left|\frac{X_{q,k}}{P_{d,k}\cdot q} - 1\right| \geq \epsilon\right] \leq \left(\frac{1}{\epsilon/2}\right)^2 \frac{M_1 q}{(P_d \cdot q/2)^2} = \frac{M}{q},$$

where $M = 16M_1/\epsilon^2 {P_d}^2$. □

Next we prove, under the same assumption, that the coalescence of a typical polynomial $f \in \mathbb{F}_q[x]$ of degree $d \geq 2$ coincides with the one of general polynomials. This result is also relevant in the context of the Brent-Pollard heuristic and will be further explored in Chapter 4.

Theorem 44. *Let $Y_q(f)$ be the random variable representing the coalescence of a polynomial $f \in \Omega_{q,d}$, Y_q defined on the probability space described in Theorem 43. Assume that the number N of general polynomials of the form $f(x) = a_d x^d + \cdots + a_1 x + a_0 \in \mathbb{F}_q[x]$ satisfies $N = q^{d+1}\big(1 + O(1/q)\big)$. Then $Y_q \xrightarrow{P} 1$.*

Proof. We note that the coalescence of polynomials of a fixed degree $d \geq 2$ over $\mathbb{F}_q$ is bounded as q approaches infinity. Indeed,

$$V(f) = \sum_{x\in\mathbb{F}_q} \frac{1}{q}|f^{-1}(x)| = \sum_{k=0}^{d}\sum_{x\in\mathcal{N}_k} \frac{1}{q}|f^{-1}(x)| = \sum_{k=0}^{d} \frac{|\mathcal{N}_k|}{q}k \leq \sum_{k=0}^{d} k = \frac{d(d+1)}{2}.$$

The result follows from arguments similar to those in the proof of Theorem 43. □

We note that, as far as we known, there is no characterization for general polynomials, although in (WILLIAMS, 1967) the author obtains a characterization for the polynomials of degree $d = 2$, 3 and 4 whose value set satisfies the estimate in Theorem 39. A combination of the results of (BIRCH; SWINNERTON-DYER, 1959) and (COHEN, 1970) provides a result in this direction for degrees $d = 2$, 3 and 4.

Theorem 45. *Let p be a prime number and $q = p^e$, $e \geq 1$.*

(i) If $p > 2$ and f is a quadratic polynomial over $\mathbb{F}_q$, then f is general.

(ii) If $p > 3$ and $f(x) = x^3 + bx$ is a cubic polynomial over $\mathbb{F}_q$, then f is general if and only if $b \neq 0$.

(iii) If $p > 3$ and $f(x) = x^4 + ax^2 + bx$ is a quartic polynomial over $\mathbb{F}_q$, then f is general if and only if $b \neq 0$.

We refer the reader to (BIRCH; SWINNERTON-DYER, 1959) and Section 6 of (COHEN, 1970) for the proof of items (ii) and (iii) of Theorem 45. Item (i) follows at once from Lemma 3 of (BIRCH; SWINNERTON-DYER, 1959).

3.4 Conclusion

In this chapter we give new proofs of known results on the indegree distribution of cubic and quartic polynomials. We give an exact result in the case of polynomials of the form $x^4 + ax^2 \in \mathbb{F}_q[x]$, $a \neq 0$, with $q = p^e$, $p > 3$ and $e > 1$, improving a previous result that carried an error term of the order of $q^{1/2}$. Our main contributions in this chapter are given in Section 3.3, where we prove results on the asymptotic indegree distribution of general polynomials. It is important to stress the connections between polynomials and random mappings. We prove in Theorem 41 that general polynomials present a very specific asymptotic indegree distribution. Moreover, we prove under a plausible assumption that the indegree distribution of polynomials of a fixed degree converges in probability to the limit one of general polynomials; see Theorem 43. Random mappings present a similar concentration property. It is proved in (ARNEY; BENDER, 1982) that the random variables $X_{n,k}$

defined on $\mathcal{J}$-mappings as the ratio of nodes with preimage size k have asymptotic normal distribution with variance proportional to its mean. These random variables are concentrated around the mean, so random $\mathcal{J}$-mappings also present a strong sense of a typical indegree distribution.

We prove that the coalescence of general polynomials of a fixed degree $d \geq 2$ is asymptotically 1 and that this is expected from a typical polynomial of degree d; this result is conditional as well, see Theorem 42. It is remarkable that the expected coalescence of random uniform mappings coincides asymptotically with the coalescence of general polynomials. According to the Brent-Pollard heuristic, this equality implies that general polynomial and random uniform mappings present the same expected rho length. We investigate this conjecture in Chapter 4, where we present our numerical results on the average rho length of polynomials over finite fields.

4 BRENT POLLARD HEURISTIC

Our main interest in this thesis lies in the connections between the dynamics of mappings and polynomials over finite fields. The heuristic proposed by Pollard in (POLLARD, 1975), built upon in (BRENT; POLLARD, 1981), raises questions about the randomness of polynomials as mappings in a given class. More precisely, does a typical quadratic polynomial modulo p represent a typical random mapping on p nodes? One might wonder the same in the case where a quadratic polynomial is seen as a 2-mapping or a $\{0,2\}$-mapping, as discussed in Chapters 1 and 3; see Section 2.3.1 for the definition of these classes of mappings. Relevant results on this area would represent substantial advances on the field of cryptographic algorithms; see the discussion in Chapter 1. However, the problems mentioned above are still open in many ways. The most notable result is due to Bach (BACH, 1991), where an estimate is obtained for the probability of a collision on less than k steps for families of quadratic polynomials.

Brent and Pollard in (BRENT; POLLARD, 1981) conjectured that the expected rho length of a node $x_0 \in [n]$ under a function $\varphi : [n] \longrightarrow [n]$ is given by $\sqrt{\pi n/2V(\varphi)}$, where $V(\varphi)$ is the coalescence of φ; see Definition 17. The argument for the heuristic is that one may regard the function φ as a random element of a set M of similar functions, where M is contained in the class $\mathcal{M}$ of mappings with the same indegree distribution as φ. If the set M consists of an "adequate sample" of the class $\mathcal{M}$, then the expected behaviour of φ should be similar to that of a random element of $\mathcal{M}$. It is proved in (ARNEY; BENDER, 1982) that the factor of non-randomness of a random uniform element of the class of mappings with a given indegree distribution is asymptotically $V^{-1/2}$, where V is the corresponding coalescence; see Section 2.3.1 for the definition of factor of non-randomness. In practice,

when estimating the average rho length of a given polynomial f modulo p, one identifies it as an element of both the classes M and $\mathcal{M}$ of polynomials and mappings with the same indegree distribution as f, respectively. This leads to the prediction of an average rho length of $\sqrt{\pi p/2V(f)}$ for the polynomial f. This heuristic was successfully applied in the case where M is the set of polynomials of the form $x^d + c \pmod p$, leading to the factorization of the eighth Fermat number; see Section 3 of (BRENT; POLLARD, 1981) for experimental results on the heuristic in this case. See Chapter 1 for further applications of the Brent-Pollard heuristic.

We investigate the heuristic of approximating statistics of polynomials over finite fields by the ones of classes of mappings, focusing on the average rho length. We use the results of Chapter 3 on the indegree distribution of polynomials over finite fields to determine classes M of polynomials with the same distribution and we investigate choices for the class $\mathcal{M}$ such that the Brent-Pollard heuristic provides a good prediction. Arney and Bender give asymptotic results in (ARNEY; BENDER, 1982) on classes of $\mathcal{J}$-mappings, where one restricts the indegrees of mappings to a fixed subset of the integers; see Definition 20. For example, quadratic polynomials modulo p are best approximated by $\{0, 2\}$-mappings; see the discussion in Chapters 1. This was already observed in (DUCHON et al., 2004). We stress that the case $\mathcal{J} = \mathbb{N}$ corresponds to the unrestricted case and $\mathcal{J} = \{0, 1, \ldots, d\}$ corresponds to d-mappings. Many authors have considered such classes before and derived estimates for many parameters (ARNEY; BENDER, 1982; DRMOTA; SORIA, 1997; GITTENBERGER, 1997; MACFIE; PANARIO, 2012), including the average rho length: if $\mathbb{E}_n[V]$ denotes the average coalescence over $\mathcal{J}$-mappings of size n and $\lambda = \lim_n \mathbb{E}_n[V]$, then the average rho length of a $\mathcal{J}$-mapping is asymptotically $\sqrt{\pi n/2\lambda}$, as n goes to infinity. This means that the *factor of non-randomness* of this class (Definition 21) is $\lambda^{-1/2}$. This estimate supports the heuristic suggested by Brent and Pollard.

Although many questions remain unanswered, previous authors have succeeded in the application of the Brent-Pollard heuristic in the case of polynomials of the form $x^m + c \pmod p$ and $x^2 + a \pmod p$ (BRENT; POLLARD, 1981; POLLARD, 1975). We stress that in the latter there is no mention to the coalescence of quadratic polynomials. The combination of the results of Section 3.3 and the Brent-Pollard heuristic suggests that general polynomials present a random behaviour for large values of p with respect to the average rho length. Our experiments support this heuristic. We use experimental results to show that the erratic behaviour of quadratic polynomials of the form $x^2 - 2 \pmod p$ is a particular case of a phenomenon observed in Chebyshev polynomials $T_d \in \mathbb{F}_p[x]$ of degree $d \geq 2$. We use the theoretical results on the distribution of indegrees of polynomials over finite fields to identify the class M of polynomials with the same indegree distribution of $T_d \in \mathbb{F}_p[x]$, for $d = 3, 4$. We observe that the cubic Chebyshev polynomial is general but T_4 belongs to a different class. We show numerically that the Brent-Pollard heuristic, as described above, provides a good prediction for this class. The particular nature of the quadratic Chebyshev polynomial was observed in (VASIGA; SHALLIT, 2004) and this was extended for higher degrees in (GASSERT, 2014).

We note that the Brent-Pollard heuristic clearly does not apply to a permutation polynomial $f \in \mathbb{F}_p[x]$: since $|f^{-1}(a)| = 1$ for every $a \in \mathbb{F}_p$, its coalescence $V(f)$ is zero. This can be attributed to the fact that the estimates of (ARNEY; BENDER, 1982) for $\mathcal{J}$-mappings require that the set $\mathcal{J}$ of allowed preimage sizes contains an integer greater than 1.

The results of this chapter have been accepted for publication in the International Jounal of Number Theory. The submitted version is available in (MARTINS; PANARIO, 2016).

4.1 Known Theoretical Results and Overview of the Experiments

It is important to stress that the reductions considered in the proofs of Chapter 3 do not apply in general when studying the structure of functional graphs of polynomials. While the problem of characterizing the equivalence of the complex dynamics of two polynomials $F, G \in \mathbb{C}[x]$ can be treated with the concept of moduli space (see Section 2.1 of (DEMARCO; PILGRIM, 2011) or Section 4.4 of (SILVERMAN, 2007)), in the finite field scenario the situation appears to be trickier. For instance, determining the number of non-isomorphic polynomials $f \in \mathbb{F}_p[x]$ of a fixed degree is an open problem; see (KONYAGIN et al., 2016). Two polynomials $f, g \in \mathbb{F}_p[x]$ of degree $d \geq 2$ have isomorphic functional graphs if there is an affine transformation $\sigma(x) = \alpha x + \beta$ such that $g = \sigma^{-1} \circ f \circ \sigma$. However, one is not able to make the same reductions as in the complex case because $\mathbb{F}_p$ is not algebraically closed. Using conjugation by affine polynomials we are able to prove that a polynomial $f(x) = a_d x^d + \cdots + a_1 x + a_0 \in \mathbb{F}_p[x]$, $d \geq 3$, admits an isomorphic counterpart with $a_{d-1} = 0$, but the reduction to monic polynomials does not appear to be as simple. We restrict our experiments to monic polynomials nonetheless.

The focus of our experiments is to verify the randomness of general polynomials implied by the Brent-Pollard heuristic and Theorem 42. Since the Brent-Pollard heuristic predicts the average rho length of a class of polynomials based exclusively on its coalescence, we consider in our experiments the class M of polynomials with the same indegree distribution as general polynomials. We run experiments for other classes M of polynomials as well. We compare the numerical data with known results for several random mapping models $\mathcal{M}$, in order to determine if M presents the typical behaviour of $\mathcal{M}$. Such classes of mappings include unrestricted mappings and mappings with a fixed indegree distribution in the following sense: given a finite sequence $(n_0, \ldots, n_d)$ such that $0 \leq n_k \leq 1$ for $k = 0, \ldots, d$, we consider mappings

$\varphi : [n] \longrightarrow [n]$ such that

$$\frac{\#\{y \in [n] \colon |\varphi^{-1}(y)| = k\}}{n} = n_k, \text{ for } k = 0, \ldots d.$$

We note that Theorem 6 must be satisfied. This is motivated by Theorem 41, where we prove that general polynomials present a very particular asymptotic indegree distribution. Mappings with indegree distribution fixed to the limit one of general polynomials of degree d are called here *d-general mappings*. We also consider $\mathcal{J}$-mappings, where we study how efficient is the heuristic of approximating the behaviour of a polynomial f by random mappings that are restricted only in the sense of its *allowed indegrees*; see the discussion in Section 2.3.1.

Let $\mathcal{M}_n$ be a class of mappings on n nodes with indegrees restricted according to one of the models discussed above. Let $\mathcal{M} = \bigcup_n \mathcal{M}_n$ and let λ be the constant representing the asymptotic average coalescence of the mappings in $\mathcal{M}$. It is proved in (ARNEY; BENDER, 1982) that the average rho length of a random node $x \in [n]$ of a mapping $\varphi \in \mathcal{M}_n$, over all mappings $\varphi \in \mathcal{M}_n$ and all $x \in [n]$, is asymptotically equivalent to $\sqrt{\pi n/2\lambda}$. This means that, on average, the non-randomness factor of a mapping in $\mathcal{M}$ is $\lambda^{-1/2}$. We exhibit in Table 4.1 the asymptotic value of these quantities for some classes of mappings.

We compute the numerical average rho length $\overline{s}$ of different classes of polynomials $M \subseteq \mathbb{F}_p[x]$ and compute the ratio $\overline{s}/\sqrt{p}$. Due to the results of (ARNEY; BENDER, 1982), we expect this quantity to be bounded. We compare the value of $\overline{s}/\sqrt{p}$ and the constants involved in the random mapping estimates, namely $\sqrt{\pi/2\lambda}$ for random mappings of a class with asymptotic average coalescence λ; see Table 4.1. We also compute the factor of non-randomness of the numerical results, that is, the ratio between $\overline{s}/\sqrt{p}$ and the constant $\sqrt{\pi/2} \approx 1.2533$ involved in the estimate of random unrestricted mappings.

Class of mappings	λ	$\sqrt{\pi/2\lambda}$	Factor of non-randomness
$\mathcal{J} = \mathbb{N}$	1	1.2533	1
$\mathcal{J} = \{0, 2\}$	1	1.2533	1
$\mathcal{J} = \{0, 1, 3\}$	1.0196	1.2412	0.9903
$\mathcal{J} = \{0, 1, 2, 4\}$	0.8366	1.3703	1.0933
$\mathcal{J} = \{0, 1, 2, 3, 5\}$	0.8941	1.3254	1.0575
$\mathcal{J} = \{0, 1, 2\}$	0.5858	1.6375	1.3066
$\mathcal{J} = \{0, 1, 2, 3\}$	0.8428	1.3652	1.0893
$\mathcal{J} = \{0, 1, \ldots, 4\}$	0.9507	1.2854	1.0256
$\mathcal{J} = \{0, 1, \ldots, 5\}$	0.9875	1.2612	1.0063
$\mathcal{J} = \{0, 2, 4\}$	1.3923	1.0622	0.8475
d-general mappings, $d \geq 2$	1	1.2533	1

Table 4.1: The coalescence and rho-length estimate of many classes of mappings.

It is natural to pay special attention to the polynomials x^d and T_d, where T_d is the Chebyshev polynomial of degree d, due to connections to the group of endomorphisms of the multiplicative group $(\mathbb{F}_p^*, \cdot)$; see Sections 1.6, 6.1 and 6.2 of (SILVERMAN, 2007). This is mentioned in the quadratic case in (POLLARD, 1975). Experimental results on the behavior of polynomials of the form $x^d + c \in \mathbb{F}_p[x]$ were obtained in (BRENT; POLLARD, 1981). They show indeed that these polynomials deviate significantly from the average behavior over $\mathbb{F}_p$ but their numerical result agree with the particular indegree distribution of these polynomials; see Theorems 37 and 38. We investigate if the Brent-Pollard heuristic provides a good prediction for the Chebyshev polynomials: not only we run experiments focused on these polynomials, but we avoid them when selecting polynomials with similar indegree distribution.

We note that our focus in this work is to analyze the average rho length of polynomials over finite fields, as opposed to improving an algorithm that involves this parameter, such as Pollard's rho algorithm for the factorization of integers. We compute the rho length $s(x)$ of *all* nodes $x \in \mathbb{F}_p$ for each polynomial $f \in \mathbb{F}_p[x]$ considered; this allows us to compute the exact average rho length of each

polynomial. This computation is highly costly, hence the range of primes considered in this work is small when compared to other experiments in computational number theory. However, this is enough for our purposes, namely to illustrate the efficiency of the heuristics discussed above.

4.2 Numerical Results

In this section we analyze our numerical results, presented in Table 4.2. In the description of a class of polynomials $a_d x^d + \cdots + a_1 x + a_0 \in \mathbb{F}_p[x]$, we use the notation a_k^* to indicate that the coefficient a_k was chosen randomly among the non-zero elements of $\mathbb{F}_p$. For the classes that correspond to general polynomials we considered 80 random primes p in the interval $[10^3, 10^5]$ and computed the average rho length over all nodes of $2\lfloor \log p \rfloor$ polynomials over $\mathbb{F}_p$. For the classes of unrestricted polynomials of degrees 3 and 4, we considered the same range and number of primes but $(\log p)^2$ polynomials over $\mathbb{F}_p$. For $f_3, f_4 \in \mathbb{Z}[x]$ given in Table 4.2, we considered their reductions modulo the first 100 primes greater than 10^5.

The first line of Table 4.2 shows our experimental results on the average rho length of quadratic polynomials over $\mathbb{F}_p$. The numerical results suggest that, on average, the class of quadratic polynomials behave indeed like random mappings with respect to its rho length. This was already observed experimentally by Pollard in (POLLARD, 1975). We note that the asymptotic indegree distribution of quadratic polynomials over $\mathbb{F}_p$ is equivalent to that of $\{0, 2\}$-mappings. The asymptotic coalescence of this class is 1, hence the behavior of quadratic polynomials can be approximated by both classes of unrestricted mappings and $\{0, 2\}$-mappings. However, approximating quadratic polynomials by $\{0, 1, 2\}$-mappings does not provide a good heuristic; see Table 4.1.

Class of polynomials	Average value of $\overline{s}/\sqrt{p}$	Ave. error of heuristic
$x^2 + a^*$	1.2454	0.9937
$x^3 + b^*x + c$	1.2595	1.0049
$x^3 + bx + c$	1.2543	1.0009
$f_3(x) = x^3 + x + 1$	1.2565	1.0025
$x^4 + ax^2 + b^*x + c$	1.2495	0.9969
$x^4 + ax^2 + bx + c$	1.2554	1.0016
$x^4 + a^*x^2$	0.8796	0.9925
$f_4(x) = x^4 + x^2$	0.8912	1.0056

Table 4.2: Experimental average rho length of polynomials over $\mathbb{F}_p[x]$.

Our experiments suggest that cubic polynomials $x^3 + b^*x + c$ behave like random mappings. We note that this class corresponds to cubic general polynomials; see Theorem 45. Unrestricted mappings and 3-general mappings have asymptotic average coalescence 1, so both classes represent good heuristic models for these polynomials. One must be careful using intermediate classes such as $\{0, 1, 2, 3\}$- and $\{0, 1, 3\}$-mappings, see Table 4.1. It should be noted that not all cubic polynomials are general. Polynomials of the form $x^3 + c$ that are not permutation polynomials have coalescence approximately 2 and, once this is taken into account, the heuristic provides a good prediction; see (BRENT; POLLARD, 1981). Nevertheless, one can have confidence predicting that a random cubic polynomial $x^3 + bx + c$ behaves like a random mapping. We attribute this fact to the results proved in Theorems 43 and 44. The polynomial f_3 is a fine example of what one expects from the reductions of a given polynomial modulo different primes: results that, on average over many prime numbers, show the typical behaviour of its class but with significant variance around this prediction.

Table 4.2 also shows that, on average, the behavior of the polynomials $x^4 + ax^2 + b^*x + c$ is similar to that of random unrestricted mappings or random 4-general mappings with respect to their average rho length. This corresponds to the class of quartic general polynomials. As suggested by Theorems 43 and 44, this is also

observed in the numerical results for quartic polynomials x^4+ax^2+bx+c. The classes of $\{0,1,2,4\}$-mappings and $\{0,1,2,3,4\}$-mappings are not as adequate for such a heuristic. We also analyze the class $x^4+ax^2 \in \mathbb{F}_p[x]$, $a \neq 0$, and the polynomial $f_4 \in \mathbb{Z}[x]$ whose reductions modulo p lie in this class. This is motivated by the fact that $T_4 \in \mathbb{F}_p[x]$ belongs to this class. The results are shown in the last two lines of Table 4.2: the prediction that arises from Corollary 2 and the Brent-Pollard heuristic is very accurate in this case as well. We note that these polynomials do not represent typical elements of the classes of $\{0,2,4\}$- or $\{0,1,2,3,4\}$-mappings; see Table 4.1.

Table 4.3 gives a clear indication that Chebyshev polynomials T_d do not behave as the fellow polynomials in their class. We computed the average rho length of T_3 and T_4 for the first 1000 primes greater than 10^5; for T_4, T_5 and T_6 we considered only the first 100 such primes. The cases where $T_d \in \mathbb{F}_p[x]$ is a permutation polynomial were avoided according to the characterization given by Theorem 7.16 of (LIDL; NIEDERREITER, 2008). The coalescence of these polynomials was obtained numerically for degrees greater than four, as no classification according to indegrees is known in these cases; the classification of general polynomials of degree $d \geq 5$ is an open problem as well. Our results confirm that these polynomials present an erratic behavior when compared to their class. This comparison is immediate for $T_3(x) = x^3 - 3x$ and $T_4(x) = x^4 - 4x^2 + 2$: see Theorems 37 and 38, Table 4.2 and the discussion above. For $d = 5, 6, 7$, we compared the numerical results of $T_d \in \mathbb{F}_p[x]$ with polynomials with the same indegree distribution, obtained by the formula $a \cdot T_d(x+b)+c$, with $a, b, c \in \mathbb{F}_p$ not all zero. We chose $2\lfloor \log p \rfloor$ such triples (a, b, c) for this comparison.

Class of polynomials	Ave. coalescence	$\overline{s}/\sqrt{p}$	Ave. error of heuristic
$T_3(x) = x^3 - 3x$	Corollary 1	14.6304	11.6733
$T_4(x) = x^4 - 4x^2 - 2$	Corollary 2	8.8558	9.9927
$T_5(x)$	1.9999	9.2637	10.4530
$aT_5(x+b)+c$	1.9995	0.8877	1.0016
$T_6(x)$	2.9999	11.7737	16.2710
$aT_6(x+b)+c$	2.9995	0.7320	1.0116
$T_7(x)$	2.9998	9.8664	13.6352
$aT_7(x+b)+c$	2.9993	0.7365	1.0178

Table 4.3: Experimental results on the average rho length of Chebyshev polynomials $T_d(x) \in \mathbb{F}_p[x]$.

4.3 Conclusion

The Brent-Pollard heuristic provides valuable support to the design and analysis of some cryptographic algorithms, but few rigorous results have been proved so far. We provide further heuristic arguments that support this heuristic, both for quadratic polynomials and for higher degrees. Our experiments show that the prediction that quadratic polynomials over finite fields behave like random mappings can be extended to general polynomials of higher degree, based on the fact that the asymptotic coalescence of these polynomials is 1. This provides a new perspective on the study of the behavior of quadratic polynomials, since all of them are general.

We proved in Chapter 3 that, under a plausible assumption, the indegree distribution of polynomials of a fixed degree is dominated by general polynomials. This explains why the numerical factor of non-randomness of a random polynomial of degree d modulo a large prime p is 1. Characterizations of polynomials of a fixed degree $d \geq 2$ according to indegree distribution provide a better method for choosing polynomials over a finite field, as different classes appear to have different average rho lengths, depending on its coalescence. Such characterizations are still an open problem for $d \geq 5$. If one is interested in statistics of a given polynomial,

the knowledge of its coalescence provides a better prediction.

It is important to stress that the Brent-Pollard heuristic is very accurate only *on average over all polynomials in an appropriately chosen class.* One might encounter polynomials whose behavior deviate significantly from this prediction, as expected from random mappings: the average rho length of some mappings differ significantly from the mean. For example, the polynomial $x^3 + 285x \in \mathbb{F}_p[x]$, $p = 36671$, has coalescence approximately 1, but its factor of non-randomness is approximately 2.5. We expect to encounter several examples as this one for each prime number p.

The classes of polynomials $x^d + c$ and T_d have anomalous behavior, as one could expect from previously known results on their dynamics; see Sections 6.1 and 6.2 of (SILVERMAN, 2007), for example. However, it should be emphasized that the dynamics of these classes represent distinct situations in the following sense. The Chebyshev polynomial $T_d \in \mathbb{F}_p[x]$ may be general or not; see Theorem 45. Nevertheless, it can be seen from Tables 4.2 and 4.3 that Chebyshev polynomials do not behave as the Brent-Pollard heuristic suggests. Polynomials of the form $x^d + c \in \mathbb{F}_p[x]$ are predictably not typical, but their average rho length agrees with the Brent-Pollard prediction; see the experiments in (BRENT; POLLARD, 1981).

It is interesting to note that estimating the asymptotic average rho length of Chebyshev polynomials remains an open problem, but some work has been done in this direction. In (CHOU; SHPARLINSKI, 2004; VASIGA; SHALLIT, 2004) the authors study the average tail length of Chebyshev polynomials over $\mathbb{F}_p$. Our experiments support the interpretation of (VASIGA; SHALLIT, 2004) that this represents a bounded average tail length for Chebyshev polynomials $T_d \in \mathbb{F}_p[x]$ over different primes p. In contrast, it is proved in (ARNEY; BENDER, 1982) that the expected tail length of a class of mappings with asymptotic average coalescence λ

is $\sqrt{\pi n/8\lambda}$. The results of (CHOU; SHPARLINSKI, 2004; VASIGA; SHALLIT, 2004), coupled with our numerical results, suggest that the average cycle length of a random node of $T_d \in \mathbb{F}_p[x]$ is much larger than that of random mappings.

Our experiments suggest that the behavior of polynomials of a given class is dictated by its average coalescence, as conjectured in (BRENT; POLLARD, 1981). For instance, general polynomials of degree $d = 2, 3$ and 4 have different asymptotic distribution of indegrees, but the same asymptotic coalescence 1. These classes present an experimental average rho length of approximately $\sqrt{\pi p/2}$. The results of (ARNEY; BENDER, 1982) support this heuristic. It is proved that, for a class $\mathcal{M}$ of mappings as the ones treated in this work, the asymptotic distribution of the number of cyclic nodes and local parameters such as the tail length, cycle length and rho length of a random node depend only on the asymptotic average coalescence of $\mathcal{M}$. The distribution of the number of components of a random mapping and the component size and tree size of a random node do not depend asymptotically on the restrictions on the indegrees. It is remarkable that so much structure of classes of mappings is preserved by their coalescence. This appears to hold because, under appropriate conditions, the moments of a random variable determine its distribution function uniquely; see Section 3.3 of (ROHATGI; SALEH, 2011). It is clear that, for any class $\mathcal{F}$ of mappings, the first moment of its asymptotic average indegree distribution $\overline{\mathcal{N}}(\mathcal{F})$ is 1. The second moment of $\overline{\mathcal{N}}(\mathcal{F})$ defines the dominant term in the asymptotic distribution of various parameters (ARNEY; BENDER, 1982). It seems that higher order moments of $\overline{\mathcal{N}}(\mathcal{F})$ determine the terms of smaller order in the asymptotic distribution of such parameters. This would imply that, for example, approximating general polynomials of degree $d \geq 2$ by d-general mappings instead of unrestricted mappings provides a better prediction for primes of intermediate size.

5 ISOMORPHISM OF FUNCTIONAL GRAPHS

In this chapter we focus on the problem of determining if two mappings have equivalent dynamics. This problem is known as the *graph isomorphism problem (GI)*, whose most popular version concern undirected graphs. Given two undirected graphs G, H with sets of vertices $V(G)$ and $V(H)$, GI consists on the decision problem of determining if there is a bijection $\varphi : V(G) \longmapsto V(H)$ such that (u, v) is an edge of G if and only if $(\varphi(u), \varphi(v))$ is an edge of H. Among the several applications of this problem we highlight the ones related to organic chemistry, where graphs represent molecular links (CONE; VENKATARAGHAVAN; MCLAFFERTY, 1977).

We focus on the restriction of the graph isomorphism problem to functional graph of mappings. A linear isomorphism test for the functional graphs of quadratic polynomials over a finite field appears in (KONYAGIN et al., 2016). We extend this algorithm to *d-mappings*, defined as mappings $f : [n] \longrightarrow [n]$ such that, for some $d \geq 2$, $|f^{-1}(y)| \leq d$ for all $y \in [n]$. In other words, we give a linear isomorphism test for functional graphs with bounded indegrees; our analysis is done under the bitwise model. We note that a polynomial of degree d over a finite field is a particular case of a d-mapping. If one considers as input of the algorithm two mappings chosen randomly and uniformly from the set of n^n mappings on n nodes, then the average-case bitwise complexity of our algorithm remains subquadratic.

The authors in (KONYAGIN et al., 2016) also provide algorithms for isomorphism of unrestricted mappings. Their time complexities are given under different models by $O(k_* n)$ and $O(c_* n)$, where k_* represents the size of the largest component observed in the functional graphs and c_* represents the maximum multiplicity that a component size occurs in the functional graphs. See Section 3.4

of (KONYAGIN et al., 2016) for details. Even though c_* is expected to be very small, both parameters may assume the value n. Therefore the worst-case time complexities of both algorithms are quadratic on n.

It should be noted that there is a linear time reduction from the directed graph isomorphism problem to the undirected graph isomorphism problem, discussed above. Graphs with bounded degree represent a case where there is a known polynomial-time algorithm for GI (LUKS, 1982) whose complexity in its original publication is $O(n^{ck \log k})$, where $c > 1$ is a constant and k is the degree of the graph. This algorithm is unlikely to be useful in practice (MCKAY; PIPERNO, 2014).

The graph isomorphism problem is known to be in the class NP, that is, it is known to be verifiable in polynomial time. However, it is not known if GI is in P (solvable in polynomial time) or NP-complete. It is believed that, if P $\neq$ NP, then GI may be one of the problems in the intermediate complexity classes (FORTIN, 1996). See (BABAI, 2015) for an unpublished preprint where the author claims to have obtained an algorithm that solves GI in quasipolynomial time. We refer to (READ; CORNEIL, 1977) for a survey of the important results known up to 1977.

Many researchers have made progress in restricted versions of GI. In the cases of connected planar graphs and rooted trees, there are algorithms whose worst case time complexity is linear in the number of vertices. There are known algorithms that solve this problem in polynomial time in several restrictions of GI; see (FORTIN, 1996) and the references therein.

This chapter is organized as follows. In Section 5.1 we present preliminary definitions and give an overview of the isomorphism test for quadratic polynomials presented in (KONYAGIN et al., 2016). Our algorithm is presented in Section 5.2. Section 5.3 contains the analysis of our algorithm and is split into two subsections. In

the first one we give a preliminary analysis, from which the worst-case and average-case of our algorithm follows; these are given in the second subsection of Section 5.3. Section 5.4 contains our conclusions and possibilities of future work.

The results of this chapter have been submitted for publication in August of 2015. (MARTINS et al., 2015).

5.1 Preliminaries

Let $f : [n] \longrightarrow [n]$ be a mapping. It is known that the connected components of the functional graph of f consist of a cycle (that may be a loop) whose nodes are roots of trees that are directed from leaves to cyclic nodes; these are called *cyclic trees*. Mappings with restrictions on preimage sizes present similar structure. See for example Section II.5 of (FLAJOLET; SEDGEWICK, 2009).

Our approach to the aforementioned restriction of GI consists of creating a binary encoding for the graphs at hand. It is thus necessary to keep track of binary operations in the analysis of our algorithm. For this reason, we analyze our algorithm under the *bitwise model* instead of the usual uniform cost model; see Chapter 1 of (AHO; HOPCROFT; ULLMAN, 1974). In the latter, it is assumed that the execution of basic instructions requires one unit of time, while in the bitwise model the length of the binary representation of the arguments of such an operation are taken into account. For instance, the cost of adding two integers m and n is assumed to be 1 in the uniform cost model, whereas in the bitwise model the time complexity of this operation is $\log m + \log n$. The notation $O_B(\cdot)$ is frequently used to indicate an analysis under the bitwise model. We use the notation $O(\cdot)$ for simplicity since in this paper we focus only on the bitwise model.

Let p be a prime number and $e \geq 1$ and consider the finite field $\mathbb{F}_q$ on $q = p^e$ elements. Our algorithm is a generalization of the one in (KONYAGIN et al., 2016) for quadratic polynomials over finite fields that we now summarize. The authors in (KONYAGIN et al., 2016) assume that the input of such an algorithm contains the adjacency list of each mapping. Hence the input has size $O(n \log n)$. In short, the authors take advantage of the fact that, if f is a quadratic polynomial over $\mathbb{F}_q$ and q is odd, then, with the exception of a single node with indegree 1, every node in the graph has indegree 0 or 2; see Theorem 21. They use Floyd's cycle detection technique to run the orbit of unassigned nodes and compute the cyclic vertices of the graphs; see Section 3.1 of (KNUTH, 2011b). Next a depth-first search is performed starting at each cyclic node in order to compute the connected components of $\mathcal{G}_f$ and $\mathcal{G}_h$. They then add a dummy child to the node of indegree 1 and observe that every cyclic node of the graph is the root of a binary unordered tree, with the exception of the cyclic node itself: it has just one child. This fact is used to encode each tree efficiently by assigning 0 or 1 to a node if it has 0 or 2 children, respectively. The connected components of the functional graphs are encoded by concatenating the encoding of each tree in an appropriate order. When this is done for $\mathcal{G}_f$ and $\mathcal{G}_h$, they use prefix trees or tries (see Section 6.3 of (KNUTH, 2011a)) to compare the encoding of their connected components and determine if the functional graphs are isomorphic. This process can be implemented with linear complexity in time and space.

We focus on generalizing (KONYAGIN et al., 2016) to d-mappings. The main obstacle that one may identify for extending this algorithm to d-mappings $f, h : [n] \longrightarrow [n]$, $d \geq 2$, is that the distribution of indegrees in this case is not as simple as the one for quadratic polynomials over a finite field of odd characteristic. It would take a major change in the graph to have d-ary cyclic trees. See Chapter 3 for results on the indegree distribution of cubic, quartic and general polynomials. Moreover, after comparing the encoding of the modified connected components using

prefix trees, the same procedure would not be enough to determine if the original functional graphs are isomorphic.

We are able to keep linear bitwise complexity in time and space by considering, first and foremost, an additional phase to the ones described in (KONYAGIN et al., 2016). In this additional phase, called Phase 0, we compute the indegree of each node in $\mathcal{G}_f$ and $\mathcal{G}_h$ by running through their adjacency list. We store them in vectors Deg_f and Deg_h of length n. We add dummy children to the nodes of $\mathcal{G}_f$ and $\mathcal{G}_h$ so that every node has indegree either 0 or d, giving rise to functional graphs $\mathcal{G}'_f$ and $\mathcal{G}'_h$. We encode the connected components of $\mathcal{G}'_f$ and $\mathcal{G}'_h$ in a similar fashion, modifying this phase of the algorithm due to the insertion of dummy children. We compare the encoding of the modified components of $\mathcal{G}'_f$ and $\mathcal{G}'_h$ using prefix trees as well. Finally, to be able to decide whether $\mathcal{G}_f$ and $\mathcal{G}_h$ are isomorphic, the nodes in the prefix tree keep track of how many dummy children each interior node in the corresponding connected components has. This is also done using prefix trees.

5.2 The algorithm

Our algorithm is described in phases, numbered from 0 to 3, and this section is divided into subsections accordingly. We denote by $d = d(f, h)$ the maximum indegree observed among the nodes of unrestricted mappings f, h.

5.2.1 Phase 0

In Phase 0 of our algorithm we run through the adjacency lists of $\mathcal{G}_f$ and $\mathcal{G}_h$ to compute the vectors Deg_f and Deg_g: the k-th coordinate of these vectors are defined as the indegree of the node labeled k in the corresponding functional graphs.

5.2.2 Phase 1

In Phase 1 we compute the connected components of $\mathcal{G}_f$ and $\mathcal{G}_h$, identifying the cyclic nodes and the cyclic trees of each functional graph. We do this in an alternative way. We select a node in $\mathcal{G}_f$ and compute its orbit using the adjacency list of the mapping f; this is assumed to be contained in the input of the algorithm. We can keep track of the previous nodes in the current orbit with a binary vector of length n, where the k-th coordinate is 1 if and only if the node k is a previous node in the orbit. In order to identify the cyclic nodes of this component, we run through the cyclic part of the orbit one more time. Next we choose a node that has not been visited previously and repeat the process. If we keep track of the nodes that have been visited in previous orbits, it is required that we visit each cyclic node of $\mathcal{G}_f$ a constant number of times.

5.2.3 Phase 2

Next, in Phase 2, we consider modified functional graphs $\mathcal{G}'_f$ and $\mathcal{G}'_h$ obtained by introducing dummy children to the nodes with indegree $1, \ldots, d-1$ of $\mathcal{G}_f$ and $\mathcal{G}_h$, so that the indegree of every node in the graphs is 0 or d. The cyclic nodes should be handled separately: for every cyclic node with indegree $m > 1$, we add $d + 1 - m$ dummy children. We note that the cyclic trees of $\mathcal{G}'_f$ and $\mathcal{G}'_h$ are d-ary. The connected components of the modified graphs can be encoded recursively as follows. Let $s \circ t$ denote the concatenation of binary strings s and t and let $L(\cdot)$ denote the label of an object. For each vertex v, we assign the bit 0 to v if v is a leaf of $\mathcal{G}'_f$; otherwise, we assign to v the binary string $1 \circ L\big(T_1(v)\big) \circ \cdots \circ L\big(T_d(v)\big)$, where $T_1(v), \ldots, T_d(v)$ are the subtrees of v in $\mathcal{G}'_f$ and:

(i) $L\big(T_1(v)\big) \geq \cdots \geq L\big(T_d(v)\big)$ in lexicographic order;

(ii) if $L\big(T_{i_1}(v)\big) = \cdots = L\big(T_{i_s}(v)\big)$, $i_1 \leq \cdots \leq i_s$, and $u_{i_1}, \ldots, u_{i_s}$ represent their respective roots, then $Deg[u_{i_1}] \geq \cdots \geq Deg[u_{i_s}]$.

We use the encoding of the cyclic trees of $\mathcal{G}'_f$ to create an encoding for the connected components. The encoding of the cyclic trees of $\mathcal{G}'_f$ are concatenated according to their cyclic ordering in the graph, with the first node being the one that determines the greatest value for the encoding of the corresponding component; this is done using the ideas of Algorithm 2 of (KONYAGIN et al., 2016).

We note that if $\mathcal{G}_f$ and $\mathcal{G}_h$ are isomorphic, so are $\mathcal{G}'_f$ and $\mathcal{G}'_h$, but the converse is not true. The structure of the connected components of $\mathcal{G}_f$ and $\mathcal{G}_h$ should be encoded unambiguously in the encoding of the connected components of $\mathcal{G}'_f$ and $\mathcal{G}'_h$. The condition (i) is enough to verify whether $\mathcal{G}'_f$ and $\mathcal{G}'_h$ are isomorphic, but in order to solve the original problem we keep track of the $\mathcal{G}_f$-indegree of each interior node in the encoding of a connected component C' of $\mathcal{G}'_f$. We do this by attaching to $L(C')$ the vector $\mathcal{V}^{C'}$ defined as follows:

(i) The length of $\mathcal{V}^{C'}$ is the number of 1's in the encoding of C'.

(ii) If the k-th bit 1 in the encoding of C' represents the node ℓ of $\mathcal{G}_f$, then $\mathcal{V}^{C'}[k] = Deg_f[\ell]$.

We keep track of the insertion of dummy children with this vector. We note that this procedure imposes an ordering on the subtrees of an interior node, so we must ensure that the encoding $\big(L(C'), \mathcal{V}^{C'}\big)$ of a component C' presents structural consistency: this is done by means of item (ii) of the labeling description.

We illustrate the necessity of ordering isomorphic subtrees of a node in $\mathcal{G}'_f$ with respect to their $\mathcal{G}_f$-indegree in Figures 5.1 and 5.2. In these figures we have

isomorphic trees where dummy nodes are inserted in order to create ternary trees. The binary encoding of the modified trees coincide, as they are both 1100010000. However, the sequence of indegrees of internal nodes are distinct: the bits that are equal to 1 correspond to nodes with indegrees $3, 1, 2$ and $3, 2, 1$, respectively. This would lead the algorithm to respond that these graphs are non-isomorphic. In Figure

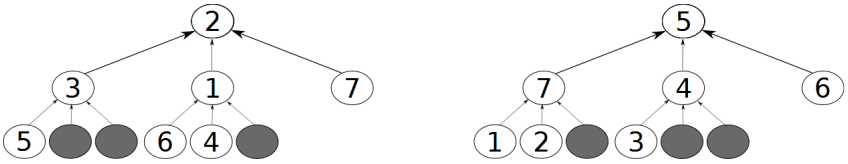

Figure 5.1: Isomorphic trees with siblings ordered according to their labels.

5.2, the isomorphic modified subtrees of the interior nodes of both trees are ordered according to their $\mathcal{G}_f$-indegree. The encoding of the modified trees remains the same and the sequence of indegrees of internal nodes are now both $3, 2, 1$.

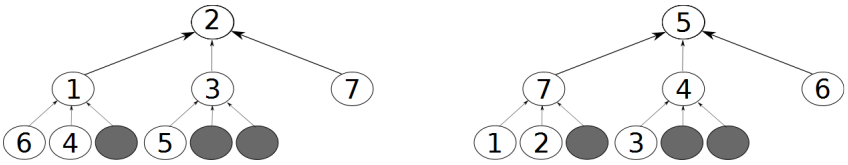

Figure 5.2: Isomorphic trees with siblings ordered according to their labels and indegrees.

The following lemma proves that the pair $\left(L(C'), \mathcal{V}^{C'}\right)$ encodes perfectly the structure of a connected component C of $\mathcal{G}_f$.

Lemma 7. *Let $\mathcal{G}_f$ be the functional graph of a mapping $f : [n] \longrightarrow [n]$. Let $\mathcal{C}$ be the set of isomorphism classes of connected components of $\mathcal{G}_f$ and, for each $C \in \mathcal{C}$, let $\Psi(C) = \left(L(C'), \mathcal{V}^{C'}\right)$, where $L(C')$ and $\mathcal{V}^{C'}$ are as detailed above. Then $\Psi(\cdot)$ is a bijection between the sets $\mathcal{C}$ and $\mathcal{C}' = \left\{\left(L(C'), \mathcal{V}^{C'}\right), C \in \mathcal{C}\right\}$.*

Proof. It is enough to prove that this correspondence is injective. Let C_1, C_2 be connected components of $\mathcal{G}_f$ such that C_1' and C_2' are isomorphic and $\mathcal{V}^{C_1'} = \mathcal{V}^{C_2'}$.

Let A be the set of vertices of C_1. We identify the set of vertices of C_1' as the union of A and A', where the latter represents the dummy nodes of C_1'.

Since C_1' and C_2' are isomorphic, there is a bijection ϕ between the nodes of C_1' and C_2' such that (x, y) is an arc of C_1' if and only if $\big(\phi(x), \phi(y)\big)$ is an arc of C_2', for x, y nodes of C_1'. If x is an interior node of C_1', then $\phi(x)$ is an interior node of C_2' and, since $\mathcal{V}^{C_1'} = \mathcal{V}^{C_2'}$, $\phi(x)$ has the same number of dummy children as x. Therefore ϕ can be chosen so that $\phi(\alpha)$ is a dummy node of C_2' if and only if α is a dummy node of C_1': it is enough to permute the leaves in the original graphs with the dummy nodes of C_1', C_2'.

Let $(a, f(a))$ be an arc of C_1. It follows that $\phi(a)$ is not a dummy node of C_2', so the arc $\big(\phi(a), f(\phi(a))\big)$ of C_2' is also present in C_2. Therefore the restriction $\phi\big|_A$ of ϕ to the set A is a 1-to-1 application onto the set nodes of C_2 such that $(a, f(a))$ is an arc of C_1 if and only if $(\phi(a), f(\phi(a)))$ is an arc of C_2. □

5.2.4 Phase 3

As in (KONYAGIN et al., 2016), we use prefix trees or tries to compare binary strings. We create a prefix tree $\mathcal{T}_f$ with the encoding of the connected components C of $\mathcal{G}_f$. This can be done using the pairs $\big(L(C'), L(\mathcal{V}^{C'})\big)$; see Lemma 7. We insert the encodings $L(C')$ in the prefix tree $\mathcal{T}_f$ and attach to the terminal node v of such an insertion a new prefix tree, denoted by $\mathcal{T}_f(v)$, where the labelings $L(\mathcal{V}^{C'})$ are inserted and thus compared. If no previous insertion of a connected component C_0' with $\mathcal{V}^{C_0'} = \mathcal{V}^{C'}$ has terminated at the corresponding node, we create a counter for it, initialized with the value 1. If such an insertion has taken place before, we just increment the corresponding counter. It follows that isomorphic components C' of $\mathcal{G}_f'$ are grouped in nodes of $\mathcal{T}_f$. Each such class splits into subclasses of isomorphic

components C of $\mathcal{G}_f$, represented by nodes of the corresponding tries $\mathcal{T}_f(v)$.

We encode the vector $\mathcal{V}^{C'}$ associated to a component C' of $\mathcal{G}'_f$ as follows. For every component C' of $\mathcal{G}'_f$ we consider, for $\ell = 1, \ldots, d-1$, the binary vector $V_\ell^{C'}$ with $|V_\ell^{C'}| = |\mathcal{V}^{C'}|$ and $V_\ell^{C'}[k] = 1$ if and only if the k-th bit 1 of $\mathcal{V}^{C'}$ represents a node with indegree ℓ. The labeling $L(\mathcal{V}^{C'})$ of $\mathcal{V}^{C'}$ is defined as

$$L(\mathcal{V}^{C'}) = V_1^{C'} \circ V_2^{C'} \circ \cdots \circ V_{d-1}^{C'}.$$

We check if the graphs $\mathcal{G}_f$ and $\mathcal{G}_h$ are isomorphic operating with counters. When we insert in the prefix tree $\mathcal{T}_f$ a pair $(L(C'), \mathcal{V}^{C'})$, we increment (or create) the counter in the terminal node of the corresponding trie $\mathcal{T}_f(v)$. Next, we repeat the process for the pairs $\left(L(C''), \mathcal{V}^{C''}\right)$ of $\mathcal{G}_h$ with the following modification: we decrement the counter of the terminating nodes of the insertions. We note that if the insertion of a pair $\left(L(C''), \mathcal{V}^{C''}\right)$ of $\mathcal{G}_h$ leads to a node where a counter has not been initialized previously, then $\mathcal{G}_f$ and $\mathcal{G}_h$ are not isomorphic. The functional graphs $\mathcal{G}_f, \mathcal{G}_h$ are isomorphic if and only if all counters are equal to zero at the end of the process.

We assume that $\mathcal{G}_f$ and $\mathcal{G}_h$ have the same number of connected components; otherwise, $\mathcal{G}_f$ and $\mathcal{G}_h$ are not isomorphic. We note that there is a counter with a positive value at the end of the process if and only if:

(i) there is a counter with a negative value, or

(ii) if the insertion of a connected component C'' of $\mathcal{G}'_h$ leads to a new external node.

We are able to detect the occurrence of (i) or (ii) during the process of insertion of the pairs $\left(L(C''), \mathcal{V}^{C''}\right)$ in the system of prefix trees. Therefore it is not necessary to

run through the external nodes of the "system" of tries to check if all counters are zero.

5.3 Analysis of the algorithm

The input of our algorithm is considered to be the adjacency list of the functional graphs $\mathcal{G}_f, \mathcal{G}_h$ of mappings $f, h : [n] \longrightarrow [n]$. Therefore the input size is proportional to $n \log n$. We recall that d is the parameter that represents the maximum indegree observed among the nodes of $\mathcal{G}_f$ and $\mathcal{G}_h$. This is an important parameter on the analysis of performance of the algorithm; it determines the number of dummy nodes inserted in $\mathcal{G}_f$ and $\mathcal{G}_h$ for the creation of $\mathcal{G}'_f$ and $\mathcal{G}'_h$. In Section 5.3.1 we consider d as a parameter in the analysis and we conclude in Section 5.3.2 providing the worst-case bitwise complexity in the case where d is bounded. The average-case analysis of our algorithm is done under the assumption that f, h are mappings chosen uniformly at random from the set of all mappings on n nodes. This is done in Section 5.3.2 as well.

5.3.1 Preliminaries

For the remainder of this section, f, h represent mappings on n elements and d represents the maximum indegree among the nodes of $\mathcal{G}_f$ and $\mathcal{G}_h$.

Proposition 6. *The worst-case bitwise time and space complexities for Phase 0 are* $O(n \log n)$ *and* $O(n \log d)$*, respectively.*

Proof. The time complexity for identifying the image of an element of $[n]$ is $O(\log n)$, giving a total time complexity of $O(n \log n)$ for Phase 0. Since the indegree of every

node of $\mathcal{G}_f$ or $\mathcal{G}_h$ is at most d, it follows that the space complexity of this phase is $O(n \log d)$. □

Proposition 7. *The worst-case bitwise time and space complexities for Phase 1 are* $O(n \log n)$.

Proof. As mentioned in Section 5.2.2, each node requires a constant number of visits in the implementation of Phase 1. The time complexity of each visit is $O(\log n)$, so the time complexity of Phase 1 is $O(n \log n)$.

Since the number of connected components in a mapping on n nodes is no greater than n, the space complexity for identifying the connected components of $\mathcal{G}_f$ is $O(n \log n)$. The same arguments show that the space complexity for identifying cycles of $\mathcal{G}_f$ is $O(n \log n)$ as well. The result follows from the fact that the space complexity for the cyclic trees of $\mathcal{G}_f$ is proportional to the product of their number of nodes and the space required for the labels that each node carries. □

Proposition 8. *The worst-case bitwise time complexity for Phase 2 is* $O(d^2 n \log n)$.

Proof. The bulk of the cost for the implementation of Phase 2 is given by the ordering of the labels of the subtrees $T_1(v), \ldots, T_d(v)$ of an internal node v of $\mathcal{G}'_f$. It is necessary to order $L\big(T_1(v)\big), \ldots, L\big(T_d(v)\big)$ in lexicographic order and then order the maximal runs $L\big(T_{i_1}(v)\big) = \cdots = L\big(T_{i_s}(v)\big)$ as detailed in Section 5.2.3.

We analyze first the cost of the lexicographic ordering throughout the creation of the labeling of a d-ary cyclic tree T. We assume that T is a complete and full d-ary tree, for this is the worst case for the creation of the label $L(T)$. Let $h \geq 0$ be

the height of T. Its number of nodes is given by

$$|T| = \frac{d^{h+1} - 1}{d - 1}. \tag{5.1}$$

Let C_k be the number of bit comparisons required for the ordering of the children of the nodes of $\mathcal{G}_f$ in the k-th level of T and let L_k be the length of each label in this level. Then, for $k = 0, \ldots, h$,

$$L_k = 1 + d + \cdots + d^{h-k} = \frac{d^{h-k+1} - 1}{d - 1}. \tag{5.2}$$

There are d^k nodes in the k-th level of T, grouped in d-uples of siblings. In order to sort lexicographically d bit strings of length L_k, we must carry no more than dL_k bit comparisons. It follows from Equation (5.2) that

$$C_k \leq \frac{d^k}{d} dL_k \leq d^k \frac{d^{h-k+1}}{d} \frac{d}{d-1} \leq 2d^h.$$

As a consequence, the total number C_T of bit comparisons required for the lexicographic ordering of the nodes of T throughout the creation of $L(T)$ is

$$C_T = \sum_{k=1}^{h} C_k \leq \sum_{k=1}^{h} 2d^h = 2hd^h.$$

Equation (5.1) implies that $h = \log_d \big(|T|(d-1) + 1\big) - 1$, so

$$C_T \leq 2\big[\log_d(d|T| + d|T|)\big] \frac{1}{d}\big(d|T| + d|T|\big) = O(|T| \log |T|). \tag{5.3}$$

Next we analyze the number C_T^* of bit operations required for ordering sibling nodes of T according to their $\mathcal{G}_f$-indegree, as described in Section 5.2.3. We use counting sort for this process; see Section 8.2 of (CORMEN et al., 2009). The time complexity for ordering m integers in the range $[1, d]$ is $O(m \log m + d \log m + m \log d)$. If T has s interior nodes with $m_1, \ldots, m_s$ children each, then C_T^* satisfies

$$C_T^* = O\left(d \sum_{i=1}^{s} \log m_i + \sum_{i=1}^{s} (m_i + m_i) \log |T|\right). \tag{5.4}$$

It is a fact that (see Section 2 of (STEELE, 2004), for example)

$$\left(\prod_{i=1}^{s} m_i\right)^{1/s} \leq \frac{1}{s}\sum_{i=1}^{s} m_i. \tag{5.5}$$

It follows from Equations (5.4) and (5.5) that

$$C_T^* = O\left(d \cdot s\left(\log\frac{1}{s} + \log|T|\right) + 2|T|\log|T|\right) = O(d|T|\log|T|). \tag{5.6}$$

Let C be a connected component of $\mathcal{G}_f$ and consider the associated component C' of $\mathcal{G}'_f$. Equations (5.3) and (5.6) imply that the total time complexity for the creation of the label $L(C')$ is

$$O\left(\sum_{T\subseteq C'}(C_T + C_T^*)\right) = O\left(\sum_{T\subseteq C'}(|T|\log n + d|T|\log n)\right) = O(d|C'|\log n).$$

Therefore, the total time complexity of Algorithms 2.a and 2.b for creating the pairs $\left(L(C'), \mathcal{V}^{C'}\right)$, $C' \subseteq \mathcal{G}'_f$, is

$$\sum_{C'\in\mathcal{G}'_f} d|C'|\log n = O(d|\mathcal{G}'_f|\log n) = O(d^2 n\log n).$$

□

Proposition 9. *The worst-case bitwise space complexity for Phase 2 is $O(dn)$.*

Proof. The length of the binary encoding of the functional graphs $\mathcal{G}'_f$ and $\mathcal{G}'_h$ is given by the number of nodes in each functional graph, so the space complexity for the binary encoding of each graph is $O(dn)$. Also, the vector $\mathcal{V}^{C'}$ has length $O(|C'|)$, where each coordinate requires $O(\log d)$ space. Therefore the space complexity for Phase 2 is $O(dn)$. □

Proposition 10. *The worst-case bitwise time complexity for Phase 3 is $O(dn + n\log n)$.*

Proof. Let C be a connected component of $\mathcal{G}_f$. We analyze first the cost of creating the labeling $L(\mathcal{V}^{C'})$. For $j = 1, \ldots, d-1$, we determine the vector $V_j^{C'}$ as described in Section 5.2.4 and concatenate it to right of the string $L(\mathcal{V}^{C'})$, initialized as an empty string. The time complexity of the creation of $V_1^{C'}$ is $O(|C|)$. If $|C|_k$ denotes the number nodes of C with indegree k, then Theorem 6 gives

$$\sum_{k\geq 0} |C|_k = \sum_{k\geq 1} k|C|_k = n. \tag{5.7}$$

If the nodes with indegree $k < j$ were marked previously once we start the computations for determining $V_j^{C'}$, $j = 2, \ldots, d-1$, then Equation (5.7) implies that the time complexity for the creation of $L(\mathcal{V}^{C'})$ is

$$O\left(|C| + \sum_{j=2}^{d-1}\left[\sum_{k=1}^{j-1}|C|_k + \sum_{k=j}^{d-1}|C|_k \log d\right]\right) = O(d|C|). \tag{5.8}$$

The worst-case complexity for both the search of a binary string s in a trie $\mathcal{T}$ and its inclusion in $\mathcal{T}$ is $O(|s|)$; see Section 6.3 of (KNUTH, 2011a). The labeling $L(\mathcal{V}^{C'})$ consists of the concatenation of $d-1$ binary vectors of length $|\mathcal{V}^{C'}|$, so the time complexity for the insertion of $|C|$ and $L(\mathcal{V}^{C'})$ in the corresponding prefix trees are $O(|C|)$ and $O(d|C|)$, respectively.

Let $C_1, \ldots, C_r$ be the connected components of $\mathcal{G}_f$. Once a search for $L(\mathcal{V}^{C'_i})$ is run in the corresponding trie, we create a counter with the value 1 in a newly created node or increment the counter in an existing node; we denote the cost of the appropriate operation by c_i. It follows from Equation (5.8) that the time complexity of Phase 3 is

$$O\left(\sum_{i=1}^{r}(d|C_i| + c_i)\right) = O\left(d|\mathcal{G}_f| + \sum_{i=1}^{r} c_i\right). \tag{5.9}$$

The worst-case bitwise complexity complexity for $\sum_i c_i$ is given by the case where $\mathcal{G}'_f$ has n connected components with the same encoding, for it is cheaper to set a

counter with value 1 than to add 1 to an existing counter. Therefore,

$$\sum_{i=1}^{r} c_i = O\left(\sum_{i=1}^{r} \log i\right) = O\left(r \log r - r\right) = O(n \log n). \tag{5.10}$$

It follows from Equations (5.9) and (5.10) that the worst-case time complexity for the insertion of $\left(L(C'), L(\mathcal{V}^{C'})\right)$ in $\mathcal{T}_f$ is $O(dn + n \log n)$. The insertion of the connected components of $\mathcal{G}_h$ in $\mathcal{T}_f$ have the same time complexity. □

Proposition 11. *The worst-case bitwise space complexity for Phase 3 is* $O(dn)$.

Proof. The functional graph of f has at most $|\mathcal{G}_f|$ connected components and the addition of dummy nodes does not create new components. Therefore $\mathcal{G}'_f$ has at most $|\mathcal{G}_f|$ connected components and the trie $\mathcal{T}_f$ for the encoding of the connect components of $\mathcal{G}'_f$ has at most $|\mathcal{G}_f|$ nodes. The worst case space complexity for $\mathcal{T}_f$ is that of a trie on $|\mathcal{G}_f|$ nodes. Indeed, every time the encoding of connected components of $\mathcal{G}'_f$ overlap we increment the counter of the corresponding node in $\mathcal{T}_f$. This demands at most one extra bit in the counter, so the cost of increasing a counter is less than the cost of creating a distinct branch in the trie. It follows that (KNUTH, 2011a) the worst case space complexity for $\mathcal{T}_f$ is $O(n)$.

For each component C'_i of $\mathcal{G}'_f$, the label $L(\mathcal{V}^{C'_i})$ consists of the concatenation of $d-1$ vectors of length $|\mathcal{V}^{C'_i}|$. Therefore the space complexity for the labels $L(\mathcal{V}^{C'_i})$, $i = 1, \ldots, r$, is

$$O\left(\sum_{i=1}^{r} d|C_i|\right) = O(dn).$$

By arguments similar to the ones presented above, the worst case for the prefix trees $\mathcal{T}_f(v)$ is that where there is no coincidence of label $L(C'_i)$ and $L(\mathcal{V}^{C'_i})$. The prefix trees $\mathcal{T}_f(v)$, in this case, require altogether $O(r) = O(|\mathcal{G}_f|) = O(n)$ of space. □

5.3.2 Worst-case and average-case analysis

If $d \geq 2$ is a fixed integer and f, h are d-mappings on n elements, then the maximum indegree observed in $\mathcal{G}_f, \mathcal{G}_h$ remains bounded as n approaches infinity. The worst-case analysis of our algorithm in this case follows at once from Propositions 6 to 11.

Theorem 46. *The algorithm described in Section 5.2 provides an isomorphism test for functional graphs of d-mappings on n nodes whose worst-case time and space bitwise complexities are $O(n \log n)$. In particular, this algorithm has linear bitwise complexity in time and space.*

Let $\mathcal{F}_n$ be the set of all mappings on n nodes. We study the average case bitwise complexity of our algorithm under the assumption that we are given two random uniform mappings $f, h \in \mathcal{F}_n$. It should be noted that a naive algorithm for testing functional graphs for isomorphism has linear average-case time complexity. It is possible to compute in linear time parameters of a functional graph such as the degree sequence or the sequence of cycles sizes. It is very unlikely that these parameters coincide for two random uniform mappings, so in most cases this should be enough. It would be interesting to obtain precise combinatorial results on these predictions. Instead, in the following, we prove that our algorithm has subquadratic average time and space bitwise complexities.

The average-case analysis of our algorithm follows from results on the *balls-in-bins-model*, where n balls are thrown into m bins. In the case of a mapping $f \in \mathcal{F}_n$, we have $m = n$ and the allocation of the i-th ball in the j-th bin represents the choice j for the image of i: $f(i) = j$. The number of balls in the most filled bin at the end of the process is the *capacity* of the system; it represents the maximum indegree in the functional graph of the corresponding mapping. Proposition VIII.10

of (FLAJOLET; SEDGEWICK, 2009) gives an asymptotic estimate for the expected value of the capacity of such an object.

Theorem 47 (Proposition VIII.10 of (FLAJOLET; SEDGEWICK, 2009))**.** *Let f, h be random mappings chosen uniformly and independently from $\mathcal{F}_n$. Then the expectation of the maximum indegree $d = d(f, h)$ observed in the functional graphs $\mathcal{G}_f, \mathcal{G}_h$ satisfies*

$$\mathbb{E}[d] \leq \log n,$$

for sufficiently large n. Moreover, the probability $\mathbb{P}_n[d > 2\log n]$ that d exceeds $2\log n$ tends to 0 as n approaches infinity.

The combination of Theorem 47 and Propositions 6 to 11 allows us to bound the average-case bitwise time and space complexities of our algorithm. We stress that the asymptotic result on $\mathbb{P}_n[d > 2\log n]$ implies that it is very unlikely that two random uniform mappings f, h represent an input that deviates significantly from our average-case analysis.

Theorem 48. *The algorithm described in Section 5.2 provides an isomorphism test for functional graphs of mappings on n nodes whose average-case time and space bitwise complexities are $O(n\log^3 n)$ and $O(n\log n)$. In particular, this algorithm has subquadratic bitwise complexity in time and space.*

5.4 Conclusions and further work

We present in this chapter an isomorphism test with linear time and space bitwise complexities for the functional graphs $\mathcal{G}_f, \mathcal{G}_h$ of d-mappings $f, h : [n] \longrightarrow [n]$. In particular, it provides a linear isomorphism test for the functional graph of polynomials over finite fields. If f, h are random uniform mappings on n nodes, then the ratio between the bitwise average-case complexity of our algorithm and its input

size is $O(\log n)$. This is attributed to the fact that the expectation of the maximum indegree d observed in $\mathcal{G}_f, \mathcal{G}_h$ is $O(\log n)$.

It would be interesting to conduct a more precise analysis of our algorithm in order to estimate the constant implied by the big-Oh notation. This could provide valuable insight on how practical our algorithm is. An implementation of our algorithm would also be interesting in that regard.

The algorithm presented in this section may be improved using a parenthesis-based codification for trees; this could lead to a linear isomorphism test that requires no restriction on the indegrees of the mappings.

6 PERIODS OF ITERATIONS OF MAPPINGS

Let Ω_n be the set of all mappings $f : [n] \longrightarrow [n]$, where $[n] = \{1, \ldots, n\}$, and consider the probability space defined by the uniform distribution on the elements of Ω_n. Let $f \in \Omega_n$. Since Ω_n is finite, the sequence of functional compositions $f^{(m)} = f \circ f^{(m-1)}$, $m \geq 1$, must cycle back to itself. Let $\mathbf{T}(f)$ be the period of this sequence, that is, the least integer $T \geq 1$ such that $f^{(m+T)} = f^{(m)}$ for all $m \geq n$. The parameter $\mathbf{T}(f)$ equals the order of the permutation obtained by restricting the mapping f to its cyclic vertices. It is proved in (HARRIS, 1973) that the distribution of $\log \mathbf{T}$, when centered around $h_n = (\log n)^2/8$ and normalized by $b_n = (\log n)^{3/2}/\sqrt{24}$, converges in distribution to the Gaussian distribution:

$$\lim_{n\to\infty} \mathbb{P}_n \left[\frac{\log \mathbf{T} - h_n}{b_n} \leq x \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt. \tag{6.1}$$

An asymptotic estimate for the expected value of $\mathbf{T}$ over all mappings on n nodes is obtained in (SCHMUTZ, 2011):

$$\mathbb{E}_n[\mathbf{T}] = \exp\left(k_0 \sqrt[3]{\frac{n}{\log^2 n}} \big(1 + o(1)\big) \right), \tag{6.2}$$

where $k_0 \approx 3.36$. We note that Equation (6.1) does not imply convergence of moments. Also, for X a random variable and g a continuous real function, in general $g\big(\mathbb{E}[X]\big)$ does not equal $\mathbb{E}\big[g(X)\big]$; see Chapters 3 and 6 of (ROHATGI; SALEH, 2011).

The parameter $\mathbf{T}$ can be proven to be the least common multiple of the cycle lengths of the components of the functional graph of f. If $\mathbf{B}(f)$ is the product of all cycle lengths of f including multiplicities and $\gcd_f$ is the grand common divisor of the length of all cycles of f, then

$$\mathbf{T}(f) = \frac{\mathbf{B}(f)}{\gcd_f}.$$

One might consider $\mathbf{B}$ as an approximation for $\mathbf{T}$. For instance, the probability that $\log \mathbf{T}$ and $\log \mathbf{B}$ differ by at least $\log^{3/2} n$ satisfies

$$\mathbb{P}_n\big[\log \mathbf{B} - \log \mathbf{T} \geq \log^{3/2} n\big] \leq \frac{c(\log\log n)^2}{\log^{1/2} n} \to 0,$$

as n approaches infinity; see Proposition 1.2 of (SCHMUTZ, 2011). However, it is proved in (SCHMUTZ, 2011) that the expectation of $\mathbf{B}$ deviates significantly from the expectation of $\mathbf{T}$:

$$\mathbb{E}_n[\mathbf{B}] = \exp\left(\frac{3}{2}\sqrt[3]{n}\big(1 + o(1)\big)\right).$$

In this chapter we derive similar results for the classes of $\{0, k\}$-mappings, defined as mappings with indegrees restricted to the set $\{0, k\}$ for some $k \geq 2$; see Definition 20.

The research on estimates on random mappings is motivated in part due to the use of these objects as a model for the statistics of polynomials. This is the Brent-Pollard heuristic. It was introduced by Pollard in the analysis of his factorization method: he conjectured that quadratic polynomials behave like random mappings with respect to their average rho length (POLLARD, 1975). However, $\{0, 2\}$-mappings could provide a better model for quadratic polynomials due to the similarities between the indegree distribution of these classes; see the discussion in Chapter 1. It is thus of interest to estimate the asymptotic behavior of parameters of mappings with certain restrictions on their indegrees, as discussed in Chapter 1 and Section 2.3.1. The class of $\{0, k\}$-mappings provides a good heuristic model for quadratic polynomials ($k = 2$) and also for polynomials of the form $x^k + a \in \mathbb{F}_p[x]$ with $p \equiv 1 \pmod{k}$; see (BRENT; POLLARD, 1981) and Theorems 37 and 38. We denote this class of polynomials in this chapter by $\{0, k\}$-*polynomials*.

The discussion in Chapter 4 suggests that unrestricted mappings and $\{0, 2\}$-mappings represent equally accurate models for the expected rho length of quadratic

polynomials. It is curious that the knowledge of the indegree distribution of these polynomials, as discussed in Chapter 3, does not represent an improvement on the heuristic. An asymptotic estimate on a different parameter, such as $\mathbf{B}$ and $\mathbf{T}$, represents thus an interesting problem: it could provide a significant deviation on the asymptotic behavior of these classes of mappings or reinforce the similarities between them. We prove in Sections 6.2 and 6.3 that the latter is the case for $\mathbf{B}$ and $\mathbf{T}$.

6.1 Preliminary Results

Let $\Omega_n^{\{0,k\}}$ be the set of $\{0,k\}$-mappings. Let $\mathcal{Z} = \mathcal{Z}(f)$ be the set of cyclic nodes of a mapping $f \in \Omega_n^{\{0,k\}}$ and denote $\mathbf{Z} = |\mathcal{Z}|$. We can write the expected value of $\mathbf{B}$ over $\Omega_n^{\{0,k\}}$ as

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{B}] = \sum_{m=1}^{n} \mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m]\mathbb{E}_n^{\{0,k\}}[\mathbf{B}|\mathbf{Z} = m]. \tag{6.3}$$

To avoid confusion, we index probabilities and expected values by the set of allowed indegrees of the class of mappings in question: $\mathbb{N}$ in the case of (SCHMUTZ, 2011) or $\{0,k\}$ in our case. It is a well known fact that the restriction of a random unrestricted mapping f to its set $\mathcal{Z}$ of cyclic nodes is a uniform random permutation of $\mathcal{Z}$, but this also holds for $\{0,k\}$-mappings; see Lemma 1 of (ARNEY; BENDER, 1982). If we let μ_m be the expected value of the product of the cycle lengths of a uniform random permutation of the symmetric group S_m on m elements, then Equation (6.3) implies

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{B}] = \sum_{m=1}^{n} \mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m]\mu_m. \tag{6.4}$$

A similar expression for $\mathbb{E}_n^{\{0,k\}}[\mathbf{T}]$ is derived by the same arguments. The author in (SCHMUTZ, 2011) combines an exact result for $\mathbb{P}_n^{\mathbb{N}}[\mathbf{Z} = m]$ with the following lemmas to estimate the expected values of $\mathbf{B}$ and $\mathbf{T}$ asymptotically in the case of

unrestricted mappings.

Lemma 8. *Let μ_m be the expected value of the product of the cycle lengths of a random uniform permutation of S_m. Then*

$$\mu_m \overset{m\to\infty}{\sim} \frac{\exp(2\sqrt{m})}{2\sqrt{\pi e}m^{3/4}}.$$

In particular, for any $\epsilon > 0$, there is an N_ϵ such that, for all $m > N_\epsilon$,

$$\exp((2-\epsilon)\sqrt{m}) < \mu_m < \exp((2+\epsilon)\sqrt{m}).$$

Lemma 9. *Let M_m be the expected order of a random permutation of S_m. Let $\beta_0 = \sqrt{8I}$, where*

$$I = \int_0^\infty \log\log\left(\frac{e}{1-e^{-t}}\right) dt.$$

Then

$$\log M_m = \beta_0\sqrt{\frac{m}{\log m}} + O\left(\frac{\sqrt{m}\log\log m}{\log m}\right).$$

In this chapter, we obtain an asymptotic estimate for $\mathbb{E}_n^{\{0,k\}}[\mathbf{B}]$ following the same strategy as in (SCHMUTZ, 2011), that we describe next. Equation (6.4) and Lemma 8 imply that, for every $1 \le m_0 \le n$,

$$\mathbb{E}_n^{\mathbb{N}}[\mathbf{B}] \ge \mathbb{P}_n^{\mathbb{N}}[\mathbf{Z} = m_0]\exp((2-\epsilon)\sqrt{m_0}). \tag{6.5}$$

It is chosen an integer m_0 that provides asymptotically the best lower bound of this form. If m_* is the integer that maximizes $\mathbb{P}_n^{\mathbb{N}}[\mathbf{Z} = m]\mu_m$ for $1 \le m \le n$, then by the same arguments we obtain the upper bound

$$\mathbb{E}_n^{\mathbb{N}}[\mathbf{B}] \le n \cdot \mathbb{P}_n^{\mathbb{N}}[\mathbf{Z} = m_*]\exp((2+\epsilon)\sqrt{m_*}). \tag{6.6}$$

We prove in Section 6.2 that the lower and upper bound in Equations (6.5) and (6.6) provide the following estimate:

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{B}] = \exp\left(\frac{3}{2}\left(\frac{n}{\lambda}\right)^{1/3}(1+o(1))\right),$$

where $\lambda = k-1$. We adopt the same strategy for the expected value of $\mathbf{T}$ and obtain

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{T}] = \exp\left(k_0 \left(\frac{n}{\lambda}\right)^{1/3} \frac{1}{\log^{2/3} n}(1+o(1))\right),$$

where k_0 is a constant made explicit in Section 6.3. The following theorem gives an exact result on the distribution of the number of cyclic nodes in $\{0,k\}$-mappings (RUBIN; SITGREAVES, 1953). We note that a $\{0,k\}$-mapping of size n satisfies $n = kh$ for some $h \geq 1$. Also, the coalescence of a $\{0,k\}$-mapping $\varphi : [n] \longrightarrow [n]$ is given by $\lambda = k-1$; see Definition 17. This result can be found in Table I of (ARNEY; BENDER, 1982) but can also be derived easily from the fact that $|\varphi^{-1}(y)| = 0$ or k for all $y \in [n]$.

Theorem 49. *Let $\lambda = k-1$. A random uniform $\{0,k\}$-mapping on $n = kh$ nodes has exactly $m \in [1,h]$ cyclic nodes with probability*

$$\mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m] = \lambda k^{m-1} \binom{h-1}{m-1} \binom{n-1}{m}^{-1}.$$

It is possible to extend the quantity above to real numbers using the Gamma function $\Gamma(\cdot)$, since $n! = \Gamma(n+1)$ for any integer $n \geq 1$; see Chapter 6 of (ABRAMOWITZ; STEGUN, 1965). It follows from Theorem 49 that, if $\lambda = k-1$,

$$\begin{aligned} \mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m] &= \lambda k^{m-1} \frac{(h-1)!}{(m-1)!(h-m)!} \frac{m!(n-m-1)!}{(n-1)!} \\ &= \lambda m k^{m-1} \frac{\Gamma(h)}{\Gamma(h-m+1)} \frac{\Gamma(n-m)}{\Gamma(n)}. \end{aligned} \tag{6.7}$$

The following results concerning the Gamma function can be found in Chapter 6 of (ABRAMOWITZ; STEGUN, 1965). They are used in the calculations of Section 6.2 and 6.3.

Lemma 10. *The Gamma function satisfies*

$$\log\Gamma(y) = y\log y - y - \frac{1}{2}\log y + \frac{1}{2}\log(2\pi) + o(1).$$

Moreover, if $\Psi(z)$ *is the derivative of* $\log\Gamma(z)$*, then, as* y *approaches infinity,*

$$\Psi(y) = \log y + O\left(\frac{1}{y}\right) \quad \textit{and} \quad \Psi'(x) = \sum_{k=0}^{\infty} \frac{1}{(x+k)^2}.$$

Lemma 11. *Let* $\lambda = k - 1$*. The distribution of the number of cyclic nodes on a* $\{0, k\}$*-mapping on* $n = kh$ *nodes satisfies*

$$\mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = x] = \frac{\lambda x}{n} \exp\left(-\frac{\lambda x^2}{2n} + O\left(\frac{x^3}{n^2}\right) + o(1)\right),$$

if $x = o(n)$ *as* n *approaches infinity.*

Proof. If $S_1 = \log\Gamma(h) - \log\Gamma(h - x + 1)$ and $S_2 = \log\Gamma(n - x) - \log\Gamma(n)$, then

$$\begin{aligned} \log \mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = x] &= \log\lambda + \log x + (x-1)\log k + S_1 + S_2 \\ &= \log\left(\frac{\lambda x}{k}\right) + x\log k + S_1 + S_2. \end{aligned} \tag{6.8}$$

Next we estimate S_1 and S_2 separately. It follows from Lemma 10 that

$$\begin{aligned} S_1 &= h\log h - h - \frac{1}{2}\log h + \frac{1}{2}\log(2\pi) \\ &\quad - (h-x+1)\log(h-x+1) + (h-x+1) + \frac{1}{2}\log(h-x+1) - \frac{1}{2}\log(2\pi) + o(1) \\ &= h\log h - \frac{1}{2}\log h - (h-x+1)\left[\log h + \log\left(1 - \frac{x-1}{h}\right)\right] - x + 1 \\ &\quad + \frac{1}{2}\left[\log h + \log\left(1 - \frac{x-1}{h}\right)\right] + o(1) \\ &= h\log h - \frac{1}{2}\log h - (h-x+1)\left[\log h - \frac{x-1}{h} - \frac{(x-1)^2}{2h^2} + O\left(\frac{x^3}{h^3}\right)\right] - x + 1 \\ &\quad + \frac{1}{2}\left[\log h - \frac{x-1}{h} - \frac{(x-1)^2}{2h^2} + O\left(\frac{x^3}{h^3}\right)\right] + o(1). \end{aligned}$$

Since x/h and x^2/h^2 approach zero as n approaches infinity, we obtain

$$\begin{aligned} S_1 &= h\log h - \frac{1}{2}\log h - h\log h + x - 1 + \frac{(x-1)^2}{2h} + x\log h - \frac{x(x-1)}{h} - \log h \\ &\quad - x + 1 + \frac{1}{2}\log h + O\left(\frac{x^3}{h^2}\right) + o(1), \end{aligned}$$

that is,

$$S_1 = x \log h - \log h - \frac{x^2}{2h} + O\left(\frac{x^3}{h^2}\right) + o(1). \tag{6.9}$$

Using the same arguments we conclude that

$$\begin{aligned}
S_2 &= (n-x)\log(n-x) - (n-x) - \frac{1}{2}\log(n-x) - n\log n + n + \frac{1}{2}\log n + o(1) \\
&= (n-x)\left[\log n + \log\left(1 - \frac{x}{n}\right)\right] + x - n \\
&\quad - \frac{1}{2}\left[\log n + \log\left(1 - \frac{x}{n}\right)\right] - n\log n + n + \frac{1}{2}\log n + o(1) \\
&= (n-x)\left[\log n - \frac{x}{n} - \frac{x^2}{2n^2} + O\left(\frac{x^3}{n^3}\right)\right] + x - n \\
&\quad - \frac{1}{2}\left[\log n - \frac{x}{n} - \frac{x^2}{2n^2} + O\left(\frac{x^3}{n^3}\right)\right] - n\log n + n + \frac{1}{2}\log n + o(1) \\
&= n\log n - x - \frac{x^2}{2n} - x\log n + \frac{x^2}{n} + x - n \\
&\quad - \frac{1}{2}\log n - n\log n + n + \frac{1}{2}\log n + O\left(\frac{x^3}{n^2}\right) + o(1),
\end{aligned}$$

that is,

$$S_2 = \frac{x^2}{2n} - x\log n + O\left(\frac{x^3}{n^2}\right) + o(1). \tag{6.10}$$

It follows from Equations (6.8), (6.9) and (6.10) that

$$\begin{aligned}
\log \mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = x] &= \log\left(\frac{\lambda x}{k}\right) + x\log k + x\log h - \log h - \frac{x^2}{2h} \\
&\quad + \frac{x^2}{2n} - x\log n + O\left(\frac{x^3}{n^2}\right) + o(1) \\
&= \log\left(\frac{\lambda x}{n}\right) - \frac{\lambda x^2}{2h} + O\left(\frac{x^3}{n^2}\right) + o(1),
\end{aligned}$$

as desired. □

6.2 Expected Value of B

It is clear from Equation (6.4) that, if m_* is the integer that maximizes $\mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m]\mu_m$ for $1 \leq m \leq n$ and m_0 is an integer in $[1, n]$, then

$$\mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m_0]\mu_{m_0} \leq \mathbb{E}_n^{\{0,k\}}[\mathbf{B}] \leq n\mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m_*]\mu_{m_*}.$$

Let $\epsilon > 0$. It follows from Lemma 8 that, for sufficiently large n,

$$\mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m_0]e^{(2-\epsilon)\sqrt{m_0}} \leq \mathbb{E}_n^{\{0,k\}}[\mathbf{B}] \leq n\mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m_*]e^{(2+\epsilon)\sqrt{m_*}}, \tag{6.11}$$

provided that m_* approaches infinity when so does n; we defer the proof of this claim to the proof of Theorem 50. In light of Equation (6.7), in order to obtain upper and lower bounds by Equation (6.11) we consider the function

$$U_{n,\epsilon}(x) = \lambda x k^{x-1} \frac{\Gamma(h)}{\Gamma(h-x+1)} \frac{\Gamma(n-x)}{\Gamma(n)} \exp((2+\epsilon)\sqrt{x}), \tag{6.12}$$

where ϵ denotes a real number that may be positive or negative. If x_* is the point that maximizes $U_{n,\epsilon}(x)$ for $x \in (0, n)$ and m_0 is an integer in $[1, n]$, then Equation (6.11) implies, for sufficiently large n, that

$$U_{n,-\epsilon}(m_0) \leq \mathbb{E}_n^{\{0,k\}}[\mathbf{B}] \leq n \cdot U_{n,\epsilon}(x_*). \tag{6.13}$$

In order to simplify the calculations that follow, we consider $H_{n,\epsilon}(x) = \log U_{n,\epsilon}(x)$ and note that x_* is a local maximum of $U_{n,\epsilon}(x)$ if and only if it is a local maximum of $H_{n,\epsilon}(x)$. It is known that the inverse of the Gamma function has simple zeroes in the non-positive integers, so the function $H_{n,\epsilon}(x)$ is not well defined for $x \in \{h+1, h+2, \dots\}$. We consider the range $[1, h]$ and note that this is not an issue because $\mathbb{P}_n[\mathbf{Z} = m] = 0$ for $m > h$. Indeed, every cyclic node in a $\{0, k\}$-mapping $\varphi : [n] \longrightarrow [n]$ has indegree k, hence n and m must satisfy $k \cdot m \leq n$.

We recall that a real function is differentiable in the range $[1, h]$ if and only if it is differentiable in $(1-\delta, h+\delta)$ for some $\delta > 0$.

Proposition 12. *For each $n \geq 1$, there exists a unique point x_* that maximizes the function $H_{n,\epsilon}(x)$ for $x \in [1, h]$, where $h = n/k$. Moreover, $H_{n,\epsilon}(x_*)$ and $H_{n,\epsilon}(\lfloor x_* \rfloor)$ are both given by*

$$\left(1 + \frac{\epsilon}{2}\right)^{4/3} \frac{3}{2} \left(\frac{n}{\lambda}\right)^{1/3} (1 + o(1)),$$

where $\lambda = k - 1$.

Proof. The function $\log \Gamma(x)$ is infinitely differentiable for $x > 0$ (ABRAMOWITZ; STEGUN, 1965). Hence, for each $n \geq 1$, $H_{n,\epsilon}(x)$ is infinitely differentiable for $x \in [1, h]$. It follows from Equation (6.12) that, for $x \in [1, h]$,

$$\begin{aligned} H'_{n,\epsilon}(x) &= \frac{1}{x} + \log k + \frac{d}{dx} \log \left(\frac{\Gamma(n - x)}{\Gamma(h - x + 1)} \right) + \left(1 + \frac{\epsilon}{2}\right) x^{-1/2} \\ &= \log k + \frac{1}{x} + \left(1 + \frac{\epsilon}{2}\right) x^{-1/2} - \Psi(h - x + 1)(-1) + \Psi(n - x)(-1) \\ &= \log k + \frac{1}{x} + \left(1 + \frac{\epsilon}{2}\right) x^{-1/2} + \Psi(h - x + 1) - \Psi(n - x). \end{aligned} \tag{6.14}$$

Using Lemma 10 we obtain

$$\begin{aligned} H'_{n,\epsilon}(1) &= \log k + 2 + \frac{\epsilon}{2} + \Psi(h) - \Psi(n - 1) \\ &= \log k + 2 + \frac{\epsilon}{2} + \log h + O\left(\frac{1}{h}\right) - \log(n - 1) + O\left(\frac{1}{n - 1}\right) \\ &= \log k + 2 + \frac{\epsilon}{2} + \log h - \log n - \log\left(1 - \frac{1}{n}\right) + O\left(\frac{1}{n}\right) \\ &= \log k + 2 + \frac{\epsilon}{2} + \log h - \log k - \log h + O\left(\frac{1}{n}\right) = 2 + \frac{\epsilon}{2} + O\left(\frac{1}{n}\right). \end{aligned}$$

Therefore $H'_{n,\epsilon}(1) > 0$ for sufficiently large values of n. On the other hand,

$$\begin{aligned} H'_{n,\epsilon}(h) &= \log k + \frac{1}{h} + \left(1 + \frac{\epsilon}{2}\right) h^{-1/2} + \Psi(1) - \Psi(n - h) \\ &= \log k + \frac{1}{h} + \left(1 + \frac{\epsilon}{2}\right) h^{-1/2} + \Psi(1) - \log(n - h) + O\left(\frac{1}{n - h}\right). \end{aligned}$$

Since

$$\frac{1}{n - h} = \frac{1}{k - 1} \cdot \frac{1}{h} = O\left(\frac{1}{h}\right),$$

it follows that

$$H'_{n,\epsilon}(h) = \log k + \Psi(1) - \log(k-1) - \log h + O\left(\frac{1}{h^{1/2}}\right) = -\log h + O(1).$$

Hence $n = kh$ implies that $H'_{n,\epsilon}(h) < 0$ for sufficiently large n. This proves the existence of a point x_* that is a local maximum of $H_{n,\epsilon}(x)$. Also,

$$H''_{n,\epsilon}(x) = -x^{-2} - \left(\frac{1}{2} + \frac{\epsilon}{4}\right) x^{-3/2} - \Psi'(h-x+1) + \Psi'(n-x).$$

Since $h - x + 1 < n - x$, it follows from Lemma 10 that $\Psi'(n-x) < \Psi'(h-x+1)$ and thus $H''_{n,\epsilon}(x) < 0$ for $x \in [1, h]$. This proves that x_* is unique.

We obtain next a heuristic estimate for x_* as n approaches infinity. Using Lemma 10 and the same arguments as in the proof of Lemma 11, one proves that, for $x = o(n)$,

$$\Psi(n-x) = \log n - \frac{x}{n} + O\left(\frac{x^2}{n^2}\right) + O\left(\frac{1}{n}\right), \tag{6.15}$$

and

$$\Psi(h-x+1) = \log h - \frac{x}{h} + O\left(\frac{x^2}{h^2}\right) + O\left(\frac{1}{n}\right). \tag{6.16}$$

It follows from Equations (6.14), (6.15) and (6.16) that

$$\begin{aligned} H'_{n,\epsilon}(x) &= \frac{1}{x} + \left(1 + \frac{\epsilon}{2}\right) x^{-1/2} - \frac{x}{h} + \frac{x}{n} + O\left(\frac{x^2}{n^2}\right) + O\left(\frac{1}{n}\right) \\ &= \frac{1}{x} + \left(1 + \frac{\epsilon}{2}\right) x^{-1/2} - \frac{(k-1)x}{n} + O\left(\frac{x^2}{n^2}\right) + O\left(\frac{1}{n}\right). \end{aligned} \tag{6.17}$$

We recall that $\lambda = k - 1$ and consider the equation

$$\frac{1}{x} + \left(1 + \frac{\epsilon}{2}\right) x^{-1/2} - \frac{\lambda x}{n} = 0,$$

that is,

$$\left(1 + \frac{\epsilon}{2}\right) x^{-1/2} \left(1 + O\left(\frac{1}{\sqrt{x}}\right)\right) = \frac{\lambda x}{n}.$$

The equation above suggests that

$$x_* = \left(\left(1 + \frac{\epsilon}{2}\right) \frac{n}{\lambda}\right)^{2/3} \big(1 + o(1)\big). \tag{6.18}$$

In order to confirm that Equation (6.18) holds, we prove that

$$H'_{n,\epsilon}\left(\left[\left(1+\frac{\epsilon}{2}\right)\frac{n}{\lambda}\right]^{2/3}(1+\delta_n)\right)<0 \tag{6.19}$$

and

$$H'_{n,\epsilon}\left(\left[\left(1+\frac{\epsilon}{2}\right)\frac{n}{\lambda}\right]^{2/3}(1-\delta_n)\right)>0, \tag{6.20}$$

for some small $\delta_n = o(1)$ to be determined. We observe that Equation (6.17) implies

$$\begin{aligned}
&H'_{n,\epsilon}\left(\left[\left(1+\frac{\epsilon}{2}\right)\frac{n}{\lambda}\right]^{2/3}(1+\delta_n)\right)\\
&=\left(\left(1+\frac{\epsilon}{2}\right)\frac{n}{\lambda}\right)^{-2/3}(1+\delta_n)^{-1}+\left(1+\frac{\epsilon}{2}\right)\left(\left(1+\frac{\epsilon}{2}\right)\frac{n}{\lambda}\right)^{-1/3}(1+\delta_n)^{-1/2}\\
&\quad-\frac{(k-1)}{n}\left(\left(1+\frac{\epsilon}{2}\right)\frac{n}{\lambda}\right)^{2/3}(1+\delta_n)+O\big(n^{-2/3}\big)+O\big(n^{-1}\big)\\
&=\left(1+\frac{\epsilon}{2}\right)^{2/3}\left(\frac{n}{\lambda}\right)^{-1/3}(1+\delta_n)^{-1/2}-\left(1+\frac{\epsilon}{2}\right)^{2/3}\left(\frac{n}{\lambda}\right)^{-1/3}(1+\delta_n)\\
&\quad+O\big(n^{-2/3}\big)\\
&=\left(1+\frac{\epsilon}{2}\right)^{2/3}\left(\frac{n}{\lambda}\right)^{-1/3}\left[(1+\delta_n)^{-1/2}-(1+\delta_n)+O\big(n^{-1/3}\big)\right]\\
&=\left(1+\frac{\epsilon}{2}\right)^{2/3}\left(\frac{n}{\lambda}\right)^{-1/3}\left[\left(1-\frac{1}{2}\delta_n+O(\delta_n^2)\right)-(1+\delta_n)+O\big(n^{-1/3}\big)\right]\\
&=\left(1+\frac{\epsilon}{2}\right)^{2/3}\left(\frac{n}{\lambda}\right)^{-1/3}\left(-\frac{3}{2}\delta_n+O(\delta_n^2)+O\big(n^{-1/3}\big)\right).
\end{aligned}$$

It is of our interest to write

$$H'_{n,\epsilon}\left(\left[\left(1+\frac{\epsilon}{2}\right)\frac{n}{\lambda}\right]^{2/3}(1+\delta_n)\right)=\left(1+\frac{\epsilon}{2}\right)^{2/3}\left(\frac{n}{\lambda}\right)^{-1/3}\left(-\frac{3}{2}\delta_n+o(\delta_n)\right),$$

as this would allow us to determine if the left-hand side of the equation above is positive or negative, depending on the value of δ_n. We set $\delta_n = n^{-1/4}$ and conclude that

$$H'_{n,\epsilon}\left(\left[\left(1+\frac{\epsilon}{2}\right)\frac{n}{\lambda}\right]^{2/3}(1+n^{-1/4})\right)<0, \tag{6.21}$$

for sufficiently large n. One proves similarly that, for sufficiently large n, we have

$$H'_{n,\epsilon}\left(\left[\left(1+\frac{\epsilon}{2}\right)\frac{n}{\lambda}\right]^{2/3}(1-\delta_n)\right)=\left(1+\frac{\epsilon}{2}\right)^{2/3}\left(\frac{n}{\lambda}\right)^{-1/3}\left(\frac{3}{2}\delta_n+o(\delta_n)\right).$$

Hence, for sufficiently large n,

$$H'_{n,\epsilon}\left(\left[\left(1+\frac{\epsilon}{2}\right)\frac{n}{\lambda}\right]^{2/3}(1-n^{-1/4})\right) > 0. \tag{6.22}$$

Equations (6.21) and (6.22) imply that Equation (6.18) holds indeed.

We estimate now the value of $H_{n,\epsilon}(x_*)$. We recall that $H_{n,\epsilon}(x_*) = \log U_{n,\epsilon}(x_*)$, where $U_{n,\epsilon}(x_*)$ is defined in Equation (6.12). If $S_1^* = \log\Gamma(h) - \log\Gamma(h - x_* + 1)$ and $S_2^* = \log\Gamma(n - x_*) - \log\Gamma(n)$, then

$$\begin{aligned} H_{n,\epsilon}(x_*) &= \log\lambda + \log x_* + (x_* - 1)\log k + S_1^* + S_2^* + (2+\epsilon)\sqrt{x_*} \\ &= x_*\log k + S_1^* + S_2^* + (2+\epsilon)\sqrt{x_*} + O(\log n). \end{aligned} \tag{6.23}$$

The same arguments in the calculations leading up to Equations (6.9) and (6.10), combined with the fact that $x^3/n^2 = O(1)$, allow us to prove that

$$S_1^* = -\frac{x_*^2}{2h} + x_*\log h + O(1), \tag{6.24}$$

and

$$S_2^* = \frac{x_*^2}{2n} - x_*\log n + O(1). \tag{6.25}$$

It follows from Equations (6.23), (6.24) and (6.25) that

$$\begin{aligned} H_{n,\epsilon}(x_*) &= x_*\log k - \frac{x_*^2}{2h} + x_*\log h + \frac{x_*^2}{2n} - x_*\log n + (2+\epsilon)\sqrt{x_*} + O(\log n) \\ &= -(k-1)\frac{x_*^2}{2n} + (2+\epsilon)\sqrt{x_*} + O(\log n). \end{aligned}$$

Hence, by Equation (6.18),

$$\begin{aligned} H_{n,\epsilon}(x_*) &= \left(\frac{n}{\lambda}\right)^{1/3}\left[-\frac{1}{2}\left(\frac{2+\epsilon}{2}\right)^{4/3} + (2+\epsilon)\left(\frac{2+\epsilon}{2}\right)^{1/3}\right]\big(1+o(1)\big) \\ &= \left(1+\frac{\epsilon}{2}\right)^{4/3}\left(\frac{n}{\lambda}\right)^{1/3}\left[-\frac{1}{2}+2\right]\big(1+o(1)\big) \\ &= \left(1+\frac{\epsilon}{2}\right)^{4/3}\frac{3}{2}\left(\frac{n}{\lambda}\right)^{1/3}\big(1+o(1)\big), \end{aligned}$$

as desired. The estimate of $H_{n,\epsilon}(\lfloor x_* \rfloor)$ follows easily from the fact that

$$\lfloor x_* \rfloor = x_* - \{x_*\} = x_* + O(1) = x_*\big(1 + o(1)\big).$$

□

Theorem 50. *The expected value of* $\mathbf{B}$ *over all* $\{0,k\}$*-mappings on* n *nodes satisfies*

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{B}] = \exp\left(\frac{3}{2}\left(\frac{n}{\lambda}\right)^{1/3}(1+o(1))\right),$$

where $\lambda = k - 1$.

Proof. We recall that the bounds in Equation (6.11) hold provided that the integer $m_* = m_*(n)$ that maximizes $\mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m]\mu_m$ for $1 \le m \le n$ tends to infinity when so does n. We prove this claim next. Indeed, if there exists $C > 0$ and a subsequence $(m_*(n_j))_j$ such that $m_*(n_j) \le C$ for all $j \ge 1$, then $\mathbb{P}_{n_j}^{\{0,k\}}[\mathbf{Z} = m]\mu_m$ is bounded for $j \ge 1$. However, it follows from Lemmas 8 and 11 that, for $m = \lfloor n^{1/2} \rfloor$,

$$\mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m]\mu_m \sim \frac{\lambda n^{1/2}}{n \cdot 2\sqrt{\pi e} n^{3/8}} \exp\left(-\frac{\lambda n}{2n} + 2n^{1/4}\right) = \left(\frac{\lambda^2 e^{-k}}{4\pi}\right)^{1/2} \frac{e^{2n^{1/4}}}{n^{7/8}}.$$

Thus, for $m = \lfloor n^{1/2} \rfloor$, $\mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m]\mu_m$ approaches infinity when so does n. This is a contradiction, so we have indeed that $m_*(n) \to \infty$ as $n \to \infty$.

Let $h = n/k$. We recall that $\mathbb{P}_n[\mathbf{Z} = m] = 0$ for $m > h$. It follows from Equation (6.13) that

$$\max_{1 \le m \le n} \mathbb{P}_n[\mathbf{Z} = m]\mu_m = \max_{1 \le m \le h} \mathbb{P}_n[\mathbf{Z} = m]\mu_m \le n \cdot \max_{1 \le x \le h} U_{n,\epsilon}(x) = n \cdot U_{n,\epsilon}(x_*).$$

Since $n = \exp(\log n)$, using Proposition 12 we conclude that

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{B}] \le \exp\left(\left(1 + \frac{\epsilon}{2}\right)^{4/3} \frac{3}{2}\left(\frac{n}{\lambda}\right)^{1/3}(1+o(1))\right). \tag{6.26}$$

We prove now that by letting $\epsilon \to 0$ in Equation (6.26) we conclude that

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{B}] \le \exp\left(\frac{3}{2}\left(\frac{n}{\lambda}\right)^{1/3}(1+o(1))\right). \tag{6.27}$$

Let $a_n = \log \mathbb{E}_n^{\{0,k\}}[\mathbf{B}]$, $n \geq 1$. We prove that there exists a sequence w_n such that $w_n = o(n^{1/3})$ and, for sufficiently large n,

$$a_n - \frac{3}{2}\left(\frac{n}{\lambda}\right)^{1/3} \leq w_n. \tag{6.28}$$

We prove this by way of contradiction. Suppose that Equation (6.28) does not hold, that is, for every sequence $(w_n)_n$ such that $w_n = o(n^{1/3})$ there exists a subsequence $(a_{n_\ell})_\ell$ such that

$$a_{n_\ell} - \frac{3}{2}\left(\frac{n_\ell}{\lambda}\right)^{1/3} > w_{n_\ell},$$

for all $\ell \geq 1$. This implies that, for all $\ell \geq 1$,

$$a_{n_\ell} - \frac{3}{2}\left(\frac{n_\ell}{\lambda}\right)^{1/3} > 0$$

and that the ratio between the left-hand side of the inequality above and $n_\ell^{1/3}$ does not approach 0 as $\ell \to \infty$. It follows that there exists $\delta > 0$ and a subsequence $(n_{\ell_j})_j$ such that, for all $j \geq 1$,

$$\frac{1}{n_{\ell_j}^{1/3}}\left(a_{n_{\ell_j}} - \frac{3}{2}\left(\frac{n_{\ell_j}}{\lambda}\right)^{1/3}\right) > \delta,$$

that is,

$$a_{n_{\ell_j}} > \frac{3}{2}\left(\frac{n_{\ell_j}}{\lambda}\right)^{1/3} + \delta n_{\ell_j}^{1/3}.$$

It follows from Equation (6.26) that, for some sequence $e_n = o(n^{1/3})$,

$$\frac{3}{2}\left(\frac{n_{\ell_j}}{\lambda}\right)^{1/3} + \delta n_{\ell_j}^{1/3} < \left(1+\frac{\epsilon}{2}\right)^{4/3}\frac{3}{2}\left(\frac{n_{\ell_j}}{\lambda}\right)^{1/3} + e_{n_{\ell_j}},$$

for sufficiently large j. This means that

$$\left(\frac{n_{\ell_j}}{\lambda}\right)^{1/3}\left[\frac{3}{2} + \delta\lambda^{1/3} - \frac{3}{2}\left(1+\frac{\epsilon}{2}\right)^{4/3}\right] < e_{n_{\ell_j}}.$$

The function $x \longmapsto x^{4/3}$ is differentiable in a neighborhood of $x = 1$, hence, by the Mean Value Theorem, there exists $\xi \in (0, \epsilon/2)$ such that

$$\left(1+\frac{\epsilon}{2}\right)^{4/3} = 1 + \frac{4}{3}\xi.$$

It follows that

$$\begin{aligned} e_{n_{\ell_j}} &> \left(\frac{n_{\ell_j}}{\lambda}\right)^{1/3}\left[\frac{3}{2}+\delta\lambda^{1/3}-\frac{3}{2}\left(1+\frac{4}{3}\xi\right)\right] \\ &= \left(\frac{n_{\ell_j}}{\lambda}\right)^{1/3}\left(\delta\lambda^{1/3}-2\xi\right) > \left(\frac{n_{\ell_j}}{\lambda}\right)^{1/3}(\delta-\epsilon)\,. \end{aligned}$$

Since Equation (6.26) holds for every $\epsilon > 0$ and the subsequence $(a_{n_{\ell_j}})_j$ does not depend on ϵ, we have, for $\epsilon = \delta/2$, that $n_{\ell_j}^{1/3} < 2\lambda^{1/3}\delta^{-1}e_{n_{\ell_j}}$ for sufficiently large j. This is a contradiction since $e_n = o(n^{1/3})$, so Equation (6.28) holds indeed. This proves Equation (6.27).

If $\epsilon > 0$ and $m_0 = \lfloor x_* \rfloor$, then Equation (6.13) and Proposition 12 imply

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{B}] \geq U_{n,-\epsilon}(m_0) = \exp\left(\left(1-\frac{\epsilon}{2}\right)^{4/3}\frac{3}{2}\left(\frac{n}{\lambda}\right)^{1/3}(1+o(1))\right). \tag{6.29}$$

One proves similarly that

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{B}] \geq \exp\left(\frac{3}{2}\left(\frac{n}{\lambda}\right)^{1/3}(1+o(1))\right) \tag{6.30}$$

holds for sufficiently large n. Hence,

$$\frac{3}{2}\left(\frac{n}{\lambda}\right)^{1/3}(1+o(1)) \leq \log \mathbb{E}_n^{\{0,k\}}[\mathbf{B}] \leq \frac{3}{2}\left(\frac{n}{\lambda}\right)^{1/3}(1+o(1)),$$

so

$$o\left(n^{1/3}\right) \leq \log \mathbb{E}_n^{\{0,k\}}[\mathbf{B}] - \frac{3}{2}\left(\frac{n}{\lambda}\right)^{1/3} \leq o(n^{1/3}),$$

as desired. □

Corollary 3. *Let $\mathbb{E}_n^{\mathcal{J}}[\mathbf{B}]$ denote the expected value of the product of the length of the cycles of a random uniform $\mathcal{J}$-mapping. Then its estimates for unrestricted mappings and $\{0,2\}$-mappings are asymptotically equivalent in logarithm:*

$$\log \mathbb{E}_n^{\{0,2\}}[\mathbf{B}] \sim \log \mathbb{E}_n^{\mathbb{N}}[\mathbf{B}] \text{ as } n \to \infty.$$

The class of $\{0,k\}$-mappings represents an interesting heuristic model for $\{0,k\}$-polynomials, that is, polynomials of the form $x^k + a \in \mathbb{F}_p[x]$ with $p \equiv 1$

(mod k). Based on this heuristic Corollary 3 one predicts that quadratic polynomials present the same expected behavior as unrestricted mappings.

We note that the logarithm of the asymptotic expected value of $\mathbf{B}$ over the classes of $\{0,k\}$-mappings presents the same order of growth for different values of k. However, the implied constant grows increasingly smaller as k grows larger:

$$\frac{\log \mathbb{E}_n^{\{0,k\}}[\mathbf{B}]}{n^{1/3}} \sim M_k \text{ as } n \to \infty,$$

where $M_k = (k-1)^{-1/3} \cdot 3/2$. It is an interesting problem to analyze if this behavior is observed in $\{0,k\}$-polynomials as well; see (BRENT; POLLARD, 1981) for numerical evidence of a similar behavior with respect to the average rho length of $\{0,k\}$-polynomials.

6.3 Expected Value of T

Let $\mathbb{E}_n^{\{0,k\}}[\mathbf{T}]$ be the average value of $\mathbf{T}$ over $\Omega_n^{\{0,k\}}$. We can write, as in Equations (6.3) and (6.4),

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{T}] = \sum_{m=1}^{n} \mathbb{P}_n^{\{0,k\}}[\mathbf{Z}=m]\mathbb{E}_n^{\{0,k\}}[\mathbf{T}|\mathbf{Z}=m] = \sum_{m=1}^{n} \mathbb{P}_n^{\{0,k\}}[\mathbf{Z}=m]M_m, \qquad (6.31)$$

where M_m is the expected order of a uniform random permutation of S_m. We use Theorem 49 to obtain an expression for $\mathbb{P}_n^{\{0,k\}}[\mathbf{Z}=m]$ and Lemma 9 for an asymptotic estimate for the values of M_m. Let $\widetilde{m}$ be the integer that maximizes $\mathbb{P}_n[\mathbf{Z}=m]M_m$ for $1 \leq m \leq n$. Once again, we estimate the expected value of $\mathbf{T}$ by noting that, for all $m_0 \in [1,n]$,

$$\mathbb{P}_n^{\{0,k\}}[\mathbf{Z}=m_0]M_{m_0} \leq \mathbb{E}_n^{\{0,k\}}[\mathbf{T}] \leq n\mathbb{P}_n^{\{0,k\}}[\mathbf{Z}=\widetilde{m}]M_{\widetilde{m}}.$$

Consider, for $n \geq 1$ and for any real number $\epsilon \in (-1, 1)$, the functions $\phi_{n,\epsilon} : (1, n) \longrightarrow \mathbb{R}$ given by

$$\phi_{n,\epsilon}(x) = \lambda x k^{x-1} \frac{\Gamma(h)}{\Gamma(h-x+1)} \frac{\Gamma(n-x)}{\Gamma(n)} \exp\left(\beta_\epsilon \sqrt{\frac{x}{\log x}}\right), \tag{6.32}$$

where $\beta_\epsilon = \beta_0 + \epsilon$ and β_0 is the constant defined in Lemma 9. We introduce such functions $\phi_{n,\epsilon}$ because the big-Oh term in the statement of Lemma 9 does not allow us to bound M_m by above or below using the main term; we increment or decrement the constant β_0 multiplying the main term by a small real number to bound the quantity M_m as we see fit. We extend the factorials in the expression of Theorem 49 using Gamma functions, as in Section 6.2.

Proposition 13. *For each $n \geq 1$ and $\epsilon \in (-1, 1)$, there exists a unique point $\widetilde{x}$ that maximizes the function $x \longmapsto \phi_{n,\epsilon}(x)$ for $x \in (1, n)$. Moreover, if $k_\epsilon = \beta_\epsilon^{4/3} 3^{5/3} 2^{-3}$ and $\lambda = k - 1$, then*

$$\phi_{n,\epsilon}(\widetilde{x}) = \exp\left(k_\epsilon \left(\frac{n}{\lambda}\right)^{1/3} \frac{1}{\log^{2/3} n}(1 + o(1))\right).$$

Proof. We consider the function $\Phi_{n,\epsilon}(x) = \log \phi_{n,\epsilon}$ and prove that it assumes a unique maximum in the interval $(1, n)$. We stress that the calculations in this proof are very similar to the ones of Proposition 12, since

$$\begin{aligned} \Phi_{n,\epsilon}(x) &= \log(\lambda x) + (x-1)\log k + \log \frac{\Gamma(h)}{\Gamma(h-x+1)} + \log \frac{\Gamma(n-x)}{\Gamma(n)} \\ &\quad + \beta_\epsilon \sqrt{\frac{x}{\log x}} \\ &= H_{n,\epsilon}(x) - (2+\epsilon)\sqrt{x} + \beta_\epsilon \sqrt{\frac{x}{\log x}}. \end{aligned} \tag{6.33}$$

We have

$$\begin{aligned} \Phi'_{n,\epsilon}(x) &= \frac{1}{x} + \log k + \frac{d}{dx} \log\left(\frac{\Gamma(n-x)}{\Gamma(h-x+1)}\right) + \frac{\beta_\epsilon}{2} \sqrt{\frac{\log x}{x}} \frac{\log x - 1}{\log^2 x} \\ &= \log k + \frac{1}{x} + \Psi(h-x+1) - \Psi(n-x) + \frac{\beta_\epsilon}{2\sqrt{x \log x}} \left(1 - \frac{1}{\log x}\right). \end{aligned} \tag{6.34}$$

We obtain in the calculations that follow a point $\widetilde{x}$ that is a local maximum of $\Phi''_{n,\epsilon}(x)$. In order to prove that this point coincides with the one mentioned in the statement of Proposition 13, we note that

$$\begin{aligned}
\Phi''_{n,\epsilon}(x) &= -\frac{1}{x^2} - \Psi'(h-x+1) + \Psi'(n-x) \\
&\quad + \frac{\beta_\epsilon}{2}\left[-\frac{1}{2}\frac{\log x + 1}{(x\log x)^{3/2}}\left(1 - \frac{1}{\log x}\right) + (x\log x)^{-1/2}\frac{1}{x\log^2 x}\right] \\
&= -\frac{1}{x^2} - \Psi'(h-x+1) + \Psi'(n-x) \\
&\quad - \frac{\beta_\epsilon}{2}\left[\frac{1}{2}(x^3\log x)^{-1/2}\left(1 + \frac{1}{\log x}\right)\left(1 - \frac{1}{\log x}\right) - (x^3\log x)^{-1/2}\frac{1}{\log^2 x}\right] \\
&= -\frac{1}{x^2} - \Psi'(h-x+1) + \Psi'(n-x) \\
&\quad - \frac{\beta_\epsilon}{4}(x^3\log x)^{-1/2}\left[1 - \frac{1}{\log^2 x} - \frac{2}{\log^2 x}\right],
\end{aligned}$$

that is,

$$\Phi''_{n,\epsilon}(x) = -\frac{1}{x^2} - \Psi'(h-x+1) + \Psi'(n-x) - \frac{\beta_\epsilon}{4}(x^3\log x)^{-1/2}\left[1 - \frac{3}{\log^2 x}\right]. \quad (6.35)$$

Lemma 10 implies that, if $0 < x_1 < x_2$, then $\Psi'(x_1) > \Psi'(x_2)$. It follows that $-\Psi'(h-x+1) + \Psi'(n-x) < 0$, since $h - x + 1 < n - x$ for $h = n/k$, $k \geq 2$. The terms $1/x^2$ and $(x^3\log x)^{-1/2}$ are positive for $x > 1$, hence $1 - 3\log^{-2} x > 0$ implies $\Phi''_{n,\epsilon}(x) < 0$ for $x > 1$. It follows that $\Phi''_{n,\epsilon}(x) < 0$ for all $x > \exp(\sqrt{3})$.

We note that Equation (6.34) implies that $\Phi'_{n,\epsilon}(x) = 0$ if and only if

$$\log k + \frac{1}{x} + \Psi(h-x+1) - \Psi(n-x) + \frac{\beta_\epsilon}{2\sqrt{x\log x}}\left(1 - \frac{1}{\log x}\right) = 0,$$

where, if $x = o(n)$, Equations (6.15) and (6.16) imply that

$$\Psi(h-x+1) - \Psi(n-x) \sim -\log k - \frac{x}{h} + \frac{x}{n}.$$

We proceed heuristically in order to obtain an intuition for the asymptotic behavior of the point $\widetilde{x} \in (1, n)$ that maximizes $\Phi_{n,\epsilon}(x)$. Assuming that the estimate above

holds as an equality, the equation $\Phi'_{n,\epsilon}(x) = 0$ is equivalent to

$$\frac{1}{x} - \frac{(k-1)x}{n} + \frac{\beta_\epsilon}{2\sqrt{x \log x}} \left(1 - \frac{1}{\log x}\right) = 0,$$

and multiplying this equation by x we obtain

$$\frac{\beta_\epsilon}{2} \left(\frac{x}{\log x}\right)^{1/2} \left(1 - \frac{1}{\log x} + \frac{2}{\beta_\epsilon} \left(\frac{\log x}{x}\right)^{1/2}\right) = \frac{\lambda x^2}{n}.$$

This is equivalent to

$$(x^3 \log x)^{1/2} = \frac{\beta_\epsilon}{2} \frac{n}{\lambda} \left(1 - \frac{1}{\log x} + \frac{2}{\beta_\epsilon} \left(\frac{\log x}{x}\right)^{1/2}\right).$$

If the function $\Phi_{n,\epsilon}(x)$ assumes indeed a unique maximum $\widetilde{x}$ for $x \in (1, n)$ and $\widetilde{x}$ approaches infinity when so does n, we expect to have

$$(\widetilde{x}^3 \log \widetilde{x})^{1/2} = \frac{\beta_\epsilon}{2} \frac{n}{\lambda} (1 + o(1)),$$

that is,

$$\widetilde{x}^3 \log \widetilde{x} = \frac{\beta_\epsilon^2}{4} \left(\frac{n}{\lambda}\right)^2 (1 + o(1)). \tag{6.36}$$

We use the bootstrapping method to obtain an approximation for the solution of Equation (6.36); see Section 4.1.2 of (GREENE; KNUTH, 2007). If not for the term $\log \widetilde{x}$ in Equation (6.36), the solution would present asymptotic behavior $\widetilde{x} = cn^{2/3}(1+o(1))$ as $n \to \infty$, for some real number $c > 0$. Thus Equation (6.36) suggests that $\widetilde{x} \sim c_1 n^{2/3} \log^{c_2} n$, for some constants c_1, c_2. This implies $\log \widetilde{x} \sim \frac{2}{3} \log n$ as n approaches infinity and

$$\widetilde{x}^3 \frac{2}{3} \log n = \frac{\beta_\epsilon^2}{4} \left(\frac{n}{\lambda}\right)^2 (1 + o(1)),$$

that is,

$$\widetilde{x}^3 = \frac{3}{8} \beta_\epsilon^2 \left(\frac{n}{\lambda}\right)^2 \frac{1}{\log n} (1 + o(1)).$$

Hence,

$$\widetilde{x} = \beta_\epsilon^{2/3} \sqrt[3]{\frac{3}{8}} \left(\frac{n}{\lambda}\right)^{2/3} \frac{1}{\log^{1/3} n} (1 + o(1)). \tag{6.37}$$

We prove now what was obtained heuristically in Equation (6.37). We define

$$\widetilde{t} = \beta_\epsilon^{2/3} \sqrt[3]{\frac{3}{8}} \left(\frac{n}{\lambda}\right)^{2/3} \frac{1}{\log^{1/3} n} \tag{6.38}$$

and prove that, for some $\delta_n = o(1)$ to be determined,

$$\phi'_{n,\epsilon}\big(\widetilde{t}(1+\delta_n)\big) < 0 < \phi'_{n,\epsilon}\big(\widetilde{t}(1-\delta_n)\big). \tag{6.39}$$

Equation (6.39) implies $\widetilde{x} = \widetilde{t}\big(1+O(\delta_n)\big)$, as desired. We note that

$$\begin{aligned}
&\frac{\beta_\epsilon}{2}\left(\frac{1}{\widetilde{t}(1+\delta_n)\log\big(\widetilde{t}(1+\delta_n)\big)}\right)^{1/2} \\
&= \frac{\beta_\epsilon}{2}\widetilde{t}^{\,-1/2}(1+\delta_n)^{-1/2}\left[\log\left(\beta_\epsilon^{2/3}\sqrt[3]{\frac{3}{8}}\left(\frac{n}{\lambda}\right)^{2/3}\frac{1}{\log^{1/3} n}\right) + \log(1+\delta_n)\right]^{-1/2} \\
&= \frac{\beta_\epsilon}{2}\widetilde{t}^{\,-1/2}(1+\delta_n)^{-1/2}\left[\frac{2}{3}\log\left(\frac{n}{\lambda}\right) + \log\log^{-1/3} n + O(1) + O(\delta_n)\right]^{-1/2} \\
&= \frac{\beta_\epsilon}{2}\widetilde{t}^{\,-1/2}(1+\delta_n)^{-1/2}\left(\frac{2}{3}\log n + O\big(\log\log n\big)\right)^{-1/2} \\
&= \frac{\beta_\epsilon}{2}\widetilde{t}^{\,-1/2}\left(\frac{3/2}{(1+\delta_n)\log n}\right)^{1/2}\left(1+O\left(\frac{\log\log n}{\log n}\right)\right)^{-1/2} \\
&= \frac{\beta_\epsilon}{2}\left(\beta_\epsilon^{2/3}\sqrt[3]{\frac{3}{8\lambda^2}}\frac{n^{2/3}}{\log^{1/3} n}\right)^{-1/2}\left(\frac{3/2}{(1+\delta_n)\log n}\right)^{1/2}\left(1+O\left(\frac{\log\log n}{\log n}\right)\right) \\
&= \sqrt[3]{\frac{3\lambda\beta_\epsilon^2}{8}}\frac{1}{n^{1/3}\log^{1/3} n}(1+\delta_n)^{-1/2}\left(1+O\left(\frac{\log\log n}{\log n}\right)\right).
\end{aligned} \tag{6.40}$$

Equations (6.15) and (6.16) imply that

$$\log k + \Psi\big(h-\widetilde{t}(1+\delta_n)+1\big) - \Psi\big(n-\widetilde{t}(1+\delta_n)\big) = -\frac{\lambda\widetilde{t}(1+\delta_n)}{n} + O\left(\frac{\widetilde{t}^2}{n^2}\right).$$

Hence,

$$\begin{aligned}
&\Phi'_{n,\epsilon}(\widetilde{t}(1+\delta_n)) \\
&= \frac{1}{\widetilde{t}(1+\delta_n)} - \frac{\lambda \widetilde{t}(1+\delta_n)}{n} \\
&\quad + \frac{\sqrt[3]{\frac{3\lambda\beta_\epsilon^2}{8}}}{n^{1/3}\log^{1/3} n}(1+\delta_n)^{-1/2}\left(1+O\left(\frac{\log\log n}{\log n}\right)\right) + O\left(\frac{1}{n^{2/3}\log^{2/3} n}\right) \\
&= -\frac{\lambda}{n}\widetilde{t}(1+\delta_n) + \frac{\sqrt[3]{\frac{3\lambda\beta_\epsilon^2}{8}}}{n^{1/3}\log^{1/3} n}(1+\delta_n)^{-1/2}\left(1+O\left(\frac{\log\log n}{\log n}\right)\right) \\
&= \frac{\sqrt[3]{\frac{3\lambda\beta_\epsilon^2}{8}}}{n^{1/3}\log^{1/3} n}\left[-(1+\delta_n) + (1+\delta_n)^{-1/2}\left(1+O\left(\frac{\log\log n}{\log n}\right)\right)\right],
\end{aligned}$$

that is,

$$\begin{aligned}
&\Phi'_{n,\epsilon}(\widetilde{t}(1+\delta_n)) \\
&= \frac{\sqrt[3]{\frac{3\lambda\beta_\epsilon^2}{8}}}{n^{1/3}\log^{1/3} n}\left[-1-\delta_n + \left(1-\frac{1}{2}\delta_n + O(\delta_n^2)\right)\left(1+O\left(\frac{\log\log n}{\log n}\right)\right)\right] \qquad (6.41) \\
&= \frac{\sqrt[3]{\frac{3\lambda\beta_\epsilon^2}{8}}}{n^{1/3}\log^{1/3} n}\left[-\frac{3}{2}\delta_n + O(\delta_n^2) + O\left(\frac{\log\log n}{\log n}\right)\right].
\end{aligned}$$

We choose $\delta_n = (\log\log n)^2/\log n$ as in the proof of Proposition 12, in order to have

$$\Phi'_{n,\epsilon}(\widetilde{t}(1+\delta_n)) = \sqrt[3]{\frac{3\lambda\beta_\epsilon^2}{8}}\frac{1}{n^{1/3}\log^{1/3} n}\left(-\frac{3}{2}\delta_n + o(\delta_n)\right). \qquad (6.42)$$

This proves that $\Phi'_{n,\epsilon}(\widetilde{t}(1+\delta_n)) < 0$ for sufficiently large n. One proves similarly that $\Phi'_{n,\epsilon}(\widetilde{t}(1-\delta_n)) > 0$, so Equation (6.37) holds indeed.

We estimate now the value of $\Phi_{n,\epsilon}(\widetilde{x})$. Using Equation (6.33) we obtain

$$\begin{aligned}
\Phi_{n,\epsilon}(\widetilde{x}) &= \log\lambda + \log\widetilde{x} + (\widetilde{x}-1)\log k + \log\Gamma(h) - \log\Gamma(h-\widetilde{x}+1)) \\
&\quad + \log\Gamma(n-\widetilde{x}) - \log\Gamma(n) + \beta_\epsilon\sqrt{\frac{\widetilde{x}}{\log\widetilde{x}}} \\
&= \widetilde{x}\log k + \log\Gamma(h) - \log\Gamma(h-\widetilde{x}+1) \\
&\quad + \log\Gamma(n-\widetilde{x}) - \log\Gamma(n) + \beta_\epsilon\sqrt{\frac{\widetilde{x}}{\log\widetilde{x}}} + O(\log n).
\end{aligned} \qquad (6.43)$$

If we let $\widetilde{S}_1 = \log\Gamma(h) - \log\Gamma(h-\widetilde{x}+1)$ and $\widetilde{S}_2 = \log\Gamma(n-\widetilde{x}) - \log\Gamma(n)$, then Equations (6.9) and (6.10) imply

$$\begin{aligned}
\Phi_{n,\epsilon}(\widetilde{x}) &= \widetilde{x}\log k - \frac{\widetilde{x}^2}{2h} + \widetilde{x}\log h + \frac{\widetilde{x}^2}{2n} - \widetilde{x}\log n + \beta_\epsilon\sqrt{\frac{\widetilde{x}}{\log\widetilde{x}}} + O(\log n)\\
&= -(k-1)\frac{\widetilde{x}^2}{2n} + \beta_\epsilon\sqrt{\frac{\widetilde{x}}{\log\widetilde{x}}} + O(\log n)\\
&= -\frac{1}{2}\left(\beta_\epsilon^{2/3}\sqrt[3]{\frac{3}{8}}\right)^2\left(\frac{n}{\lambda}\right)^{1/3}\frac{1}{\log^{2/3}n}\big(1+o(1)\big)\\
&\quad + \beta_\epsilon\left(\beta_\epsilon^{2/3}\sqrt[3]{\frac{3}{8}}\right)^{1/2}\left(\frac{n}{\lambda}\right)^{1/3}\frac{1}{\log^{1/6}n}\left(\frac{2}{3}\log n + O(\log\log n)\right)^{-1/2} \qquad (6.44)\\
&= -\frac{1}{2}\left(\beta_\epsilon^{2/3}\sqrt[3]{\frac{3}{8}}\right)^2\left(\frac{n}{\lambda}\right)^{1/3}\frac{1}{\log^{2/3}n}\big(1+o(1)\big)\\
&\quad + \beta_\epsilon\left(\beta_\epsilon^{2/3}\sqrt[3]{\frac{3}{8}}\right)^{1/2}\left(\frac{n}{\lambda}\right)^{1/3}\frac{1}{\log^{2/3}n}\left(\frac{3}{2}\right)^{1/2}\big(1+o(1)\big)\\
&= k_\epsilon\left(\frac{n}{\lambda}\right)^{1/3}\frac{1}{\log^{2/3}n}\big(1+o(1)\big),
\end{aligned}$$

where

$$k_\epsilon = -\frac{1}{2}\left(\beta_\epsilon^{2/3}\sqrt[3]{\frac{3}{8}}\right)^2 + \beta_\epsilon\left(\beta_\epsilon^{2/3}\sqrt[3]{\frac{3}{8}}\right)^{1/2}\left(\frac{3}{2}\right)^{1/2} = \left(-\frac{1}{8}+\frac{1}{2}\right)\beta_\epsilon^{4/3}3^{2/3},$$

as desired. □

Theorem 51. *Let*

$$I = \int_0^\infty \log\log\left(\frac{e}{1-e^{-t}}\right)dt.$$

The expected value of $\mathbf{T}$ *over all* $\{0,k\}$*-mappings on* n *nodes satisfies*

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{T}] = \exp\left(k_0\left(\frac{n}{\lambda}\right)^{1/3}\frac{1}{\log^{2/3}n}(1+o(1))\right),$$

where $\lambda = k-1$ *and* $k_0 = (3I)^{2/3}3/2$.

Proof. It follows from Equation (6.31) that, if $1 \leq m_0 \leq n$ and $\widetilde{m}$ is the integer that maximizes $\mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m]M_m$ for $1 \leq m \leq n$, then

$$\mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m_0]M_{m_0} \leq \mathbb{E}_n^{\{0,k\}}[\mathbf{T}] \leq n\mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = \widetilde{m}]M_{\widetilde{m}}. \tag{6.45}$$

Let $m_0 = \lfloor \widetilde{t} \rfloor$ and $\epsilon \in (-1, 0)$. Then

$$m_0 = \widetilde{t} - \{\widetilde{t}\} = \widetilde{t} + O(1) = \widetilde{t}\big(1 + o(1)\big).$$

Hence, by Equation (6.45) and the calculations in the proof of Proposition 13,

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{T}] \geq \exp\left(k_\epsilon \left(\frac{n}{\lambda}\right)^{1/3} \frac{1}{\log^{2/3} n}(1 + o(1))\right).$$

The equation above holds for every $\epsilon \in (-1, 0)$, so an argument similar to the one in the proof of Theorem 50 shows that

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{T}] \geq \exp\left(k_0 \left(\frac{n}{\lambda}\right)^{1/3} \frac{1}{\log^{2/3} n}(1 + o(1))\right), \tag{6.46}$$

where using the fact that $\beta_0 = \sqrt{8I}$ (see Lemma 9) we obtain

$$k_0 = \beta_0^{4/3} 3^{2/3} \frac{3}{8} = (\sqrt{8I})^{4/3} 3^{2/3} \frac{3}{8} = 4I^{2/3} 3^{2/3} \frac{3}{8} = \frac{3}{2}(3I)^{2/3}.$$

An argument similar to the one in the proof of Theorem 50 proves that the integer $\widetilde{m}$ that maximizes $\mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m]M_m$ for $1 \leq m \leq n$ satisfies $\widetilde{m} = o(n^{2/3})$, $m \to \infty$. Therefore, for any fixed $\epsilon \in (0, 1)$,

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{T}] \leq n \max_{1 \leq m \leq n} \mathbb{P}_n^{\{0,k\}}[\mathbf{Z} = m]M_m \leq n\Phi(\widetilde{x}),$$

for n sufficiently large, with $\widetilde{x}$ as in Proposition 13. Therefore, by Proposition 13,

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{T}] \leq n \exp\left(k_\epsilon \left(\frac{n}{\lambda}\right)^{1/3} \frac{1}{\log^{2/3} n}(1 + o(1))\right) = \exp\left(k_\epsilon \left(\frac{n}{\lambda}\right)^{1/3} \frac{1}{\log^{2/3} n}(1 + o(1))\right).$$

By letting $\epsilon \to 0$ we conclude that

$$\mathbb{E}_n^{\{0,k\}}[\mathbf{T}] \leq \exp\left(k_0 \left(\frac{n}{\lambda}\right)^{1/3} \frac{1}{\log^{2/3} n}(1 + o(1))\right). \tag{6.47}$$

The theorem follows from Equations (6.46) and (6.47). □

Corollary 4. *Let $\mathbb{E}_n^{\mathcal{J}}[\mathbf{T}]$ denote the expected value of the least common multiple of the length of the cycles of a random uniform $\mathcal{J}$-mapping. Then its estimates for unrestricted mappings and $\{0,2\}$-mappings are asymptotically equivalent in logarithm:*

$$\log \mathbb{E}_n^{\{0,2\}}[\mathbf{T}] \sim \log \mathbb{E}_n^{\mathbb{N}}[\mathbf{T}] \text{ as } n \to \infty.$$

As mentioned in Section 6.2, the class of $\{0,k\}$-mappings may be used as a model for $\{0,k\}$-polynomials. Corollary 4 suggests that the period of the sequence of functional compositions $f^{(k)} = f^{(k-1)}$, $k \geq 1$, of a quadratic polynomial $f \in \mathbb{F}_p[x]$ has expected behavior similar to that of unrestricted mappings for large values of p.

The analysis of the estimate in Theorem 51 for different values of k is similar to that in Section 6.2. The logarithm of the asymptotic expected value of $\mathbf{T}$ over the classes of $\{0,k\}$-mappings over $\mathbb{F}_p$ present the same order of growth, but the implied constant grows increasingly smaller as k grows larger:

$$\frac{\log \mathbb{E}_n^{\{0,k\}}[\mathbf{T}]}{n^{1/3} \log^{-2/3} n} \sim M_k \text{ as } n \to \infty,$$

where $M_k = k_0 \cdot \lambda^{-1/3}$. Experimental results and estimates on the behavior of the parameter $\mathbf{T}$ over $\{0,k\}$-polynomials represent an interesting problem as well.

6.4 Conclusion

In this chapter we consider the functional compositions $f^{(k)} = f \circ f^{(k-1)}$ of a $\{0,k\}$-mapping f. We give asymptotic estimates on the expected value of the period $\mathbf{T}$ of this sequence; this parameter is given by the least common multiple of the length of the cycles of f. We also provide estimates for the expected value of a related parameter $\mathbf{B}$, given by the product of the length of the cycles of f. We follow the arguments of (SCHMUTZ, 2011). The ratio between the logarithm of our

results and the respective estimates for unrestricted mappings satisfy

$$\frac{\log \mathbb{E}^{\{0,k\}}[\mathbf{T}]}{\log \mathbb{E}^{\mathbb{N}}[\mathbf{T}]} \sim \lambda^{-1/3} \quad \text{and} \quad \frac{\log \mathbb{E}^{\{0,k\}}[\mathbf{B}]}{\log \mathbb{E}^{\mathbb{N}}[\mathbf{B}]} \sim \lambda^{-1/3},$$

where $\lambda = k - 1$ is the coalescence of $\{0, k\}$-mappings. It is interesting to note that these ratios are asymptotically equivalent. Moreover, the estimates for $\{0, k\}$-mappings are identical to the ones for unrestricted mappings except for the substitution $n \leftarrow n/\lambda$. The same occurs with the estimates of (ARNEY; BENDER, 1982) on other parameters of $\mathcal{J}$-mappings such as the total number of cyclic nodes and the rho length of a node.

We consider the class of $\{0, k\}$-mappings in this chapter motivated in large part by the heuristic model for iterations of quadratic polynomials over $\mathbb{F}_p$; see Chapters 1 and 4. The indegree distribution of these polynomials is very accurately approximated by $\{0, 2\}$-mappings. The asymptotic coalescence of these mappings is $\lambda = 1$, thus the estimates for $\log \mathbb{E}^{\{0,2\}}[\mathbf{T}]$ and $\log \mathbb{E}^{\{0,2\}}[\mathbf{B}]$ are asymptotically equivalent to their unrestricted counterparts. It would be interesting to evaluate numerically if quadratic polynomials and unrestricted mappings present similar statistics with respect to the parameters $\mathbf{T}$ and $\mathbf{B}$, as is the case with the average rho length; see Chapter 4.

We stress how the ratios between the results for $\{0, k\}$-mappings and unrestricted mappings behave as k varies. The coalescence of a $\{0, k\}$-mapping is $k - 1$, thus these ratios approach zero as k grows larger. In summary, for $k = 2$ the behavior of $\{0, k\}$-mappings is asymptotically equivalent to that of unrestricted mappings (with respect to the parameters $\mathbf{T}$ and $\mathbf{B}$) and, as k grows larger, the estimates for $\{0, k\}$-mappings grow increasingly smaller in comparison. Experimental results on the average value of $\mathbf{T}$ and $\mathbf{B}$ over $\{0, k\}$-polynomials would represent an interesting point in this discussion: they would allow us to compare the behavior of these classes of polynomials with the estimates obtained in this chapter for $\{0, k\}$-mappings.

7 CONCLUSION AND FUTURE WORK

In this thesis we present our original results on the field of discrete dynamical systems. In Chapters 3 and 4 we give our contributions to the understanding of Pollard's classical rho method (POLLARD, 1975), whose running time for the factorization of a composite integer n is heavily impacted by the dynamics of quadratic polynomials modulo prime numbers. It remains an open problem to fully understand the heuristic used in the original analysis of this algorithm: it is conjectured that the average rho length of a quadratic polynomial $x^2 + a \pmod{p}$ is similar to that of a random uniform mapping $\varphi : [p] \longrightarrow [p]$. This heuristic was generalized to other classes of polynomials in (BRENT; POLLARD, 1981), where it is suggested that the numerical factor of non-randomness of a given class of polynomial is determined by its coalescence.

We consider the class of general polynomials in this thesis and obtain in Chapter 3 an asymptotic estimate on the indegree distribution of these polynomials. As a consequence, we prove that the asymptotic coalescence of these polynomials is 1. This suggests that this class of polynomials behaves as random mappings with respect to its average rho length. Our experiments in Chapter 4 confirm this prediction. Moreover, we prove under a plausible assumption that the indegree distribution of a random uniform polynomial $f \in \mathbb{F}_p[x]$ is dominated by the distribution of general polynomials; we prove a similar result for the coalescence of random polynomials. This explains why the average rho length of the class of all polynomials $f \in \mathbb{F}_p[x]$ of a given degree is very similar to the asymptotic result on random mappings. We believe that our results provide a new perspective to Pollard's original heuristic that quadratic polynomials behave as random mappings, since all quadratic polynomials are general.

Our contributions in this thesis include an isomorphism test for the functional graph of mappings; it is an extension of the algorithm of (KONYAGIN et al., 2016). In other words, we present an algorithm that recognizes if two mappings have equivalent dynamics and are thus equally adequate for the applications mentioned in this thesis, such as Pollard's method. Our algorithm is executed in linear time and space if the mappings in hand have bounded indegrees; we stress that polynomials over finite fields represent a particular case of this class of mappings. The average-case analysis of our algorithm remains subquadratic for unrestricted mappings. Although this analysis is not as interesting as the worst-case one, as mentioned in Section 5.3, it is worthwhile noting that it involves the expected value of the maximum indegree observed in the nodes of a random mapping; see Theorem 47.

In Chapter 6 we provide an asymptotic estimate for the expected value of the parameter $\mathbf{T}$ over a random uniform $\{0,k\}$-mapping f, defined as the period of the functional composition $f^{(k)} = f \circ f^{(k-1)}$, $k \geq 1$. We calculate an analogous asymptotic estimate for the related parameter $\mathbf{B}$, the product of the length of the cycles of f. It is remarkable that the logarithm of the expected value of $\mathbf{T}$ and $\mathbf{B}$ over $\{0,k\}$-mappings coincide with analogous quantities for unrestricted mappings (SCHMUTZ, 2011), except for the substitution $n \leftarrow n/\lambda$, where λ is the coalescence of $\{0,k\}$-mappings. The same holds for other parameters defined on a random mapping f, such as its total number of cyclic nodes and the expected length of the periodic and non-periodic parts of a random node of f (ARNEY; BENDER, 1982). It would be interesting to investigate $\{0,k\}$-mappings as a heuristic model for $\{0,k\}$-polynomials (see Chapter 6), but unfortunately this will be left as an important aspect of our future research.

There are a number open problems that are an integral part of our projects for the future. As mentioned in Chapter 1, the results presented in Chapter 6 of this

thesis represent partial results of our initial project in the area, namely to extend the results of (SCHMUTZ, 2011) to $\mathcal{J}$-mappings. It is proved in (ARNEY; BENDER, 1982) that the probability that a random uniform $\mathcal{J}$-mapping $f : [n] \longrightarrow [n]$ has m cyclic nodes satisfies

$$\mathbb{P}_n^{\mathcal{J}}[\mathbf{Z} = m] \sim \frac{\lambda m}{n} \exp\left(-\frac{\lambda m^2}{2n}\right),$$

as n, m approach infinity and $m = o(n^{2/3})$, where λ is the asymptotic average coalescence of $\mathcal{J}$-mappings. The restriction on the range of m represents an obstacle in the use of the arguments of Chapter 6 for $\mathcal{J}$-mappings: it is necessary to prove that the maximum of the corresponding functions in $[1, n]$, denoted in Sections 6.2 and 6.3 respectively by x_* and $\widetilde{x}$, are both $o(n^{2/3})$. The literature on tail estimates could lead to advances on this problem; see (HWANG, 1996, 1998), for example. Moreover, we expect that the point x_* in the calculations for the estimate of $\mathbf{B}$ is of the order of $n^{2/3}$ in the case of $\mathcal{J}$-mappings as well. The estimate on the distribution of $\mathbf{Z}$ of (ARNEY; BENDER, 1982) would have to be extended; the methodology of (DRMOTA; SORIA, 1997) could provide some insight on how this could be done.

We also consider an interesting open problem to obtain a more precise asymptotic estimate on the expected values of $\mathbf{T}$ and $\mathbf{B}$. From the results of (SCHMUTZ, 2011) one obtains the limit value of the $\log \mathbb{E}_n[\mathbf{T}]$ and $\log \mathbb{E}_n[\mathbf{B}]$, but not for the expected values themselves. We note that in the method used in (SCHMUTZ, 2011) we write the expected value of $\mathbf{T}$ (and $\mathbf{B}$) as a sum and bound it from above using the maximum of the summands; this leads to a term of smaller order in Theorems 1.3 and 1.4 of (SCHMUTZ, 2011) given by $\log n$. Therefore, through this method one cannot refine the error term in Theorems 1.3 and 1.4 of (SCHMUTZ, 2011) down to $o(1)$. Thus one should somehow consider the sum itself in one's calculations, or the part of it providing the bulk of the quantity of interest. Research on applications of Laplace's method (SEDGEWICK; FLAJOLET, 2013) could be fruitful for the refinement of the estimates of (SCHMUTZ, 2011).

It is also of our interest to provide asymptotic results on the distribution of other random variables defined over the space of random uniform $\mathcal{J}$-mappings. According to the authors in (KONYAGIN et al., 2016), there are no known results on the distribution of the following parameter. Let $f : [n] \longrightarrow [n]$ be a random uniform (unrestricted) mapping and let $C_1, \ldots, C_t$ be the size of the components of its functional graph $\mathcal{G}_f$, counting multiplicities. Define $\mathbf{M}$ to be the multiplicity of the most popular component size of $\mathcal{G}_f$, that is, the value $1 \leq \mathbf{M} \leq n$ that has the greatest number of occurrences in the sequence $C_1, \ldots, C_t$. The authors in (KONYAGIN et al., 2016) present experimental results on this parameter and observe consistency in its average value over all mappings $f : [n] \longrightarrow [n]$ for large values of n. We have thus far been able to determine the cumulative generating function (CGF) of this parameter:

$$\Xi(z) = \sum_{h \geq 0} \left[F(z) - \prod_{i \geq 1} T_h \left[\exp\left(K_i \frac{z^i}{i!} \right) \right] \right],$$

where K_i represents the number of connected mappings on i nodes and $T_h[\cdot]$ is the operator defined by the truncation of a power series in its h-th term. It is known that an asymptotic estimate for the expected value of $\mathbf{M}$ could be derived from an appropriate analysis of $\Xi(z)$ as a power series on the complex numbers. See Page 159 of (FLAJOLET; SEDGEWICK, 2009) for the definition of cumulative generating functions and Page 217 for the CGF of a extremal parameter, such as $\mathbf{M}$. The work of previous authors on the analysis of the CGF of different parameters could help us advance on this problem; see for example (GATHEN; PANARIO; RICHMOND, 2012) and the references therein.

Our work with general polynomials in Chapters 3 and 4 leads us to believe that it is of great interest to develop an algorithm that decides efficiently if a given polynomial $f \in \mathbb{F}_p[x]$ is general. There is some work done in the area and our projects include a through investigation of the literature on the field; see (SUTHERLAND, 2015), for an example.

There are also various number theoretic open problems that would represent relevant contributions to the applications mentioned in this thesis. Chapter 4 contains a number of them, such as proving the apparent random behavior of general polynomials; if one is able to prove that our assumption on the number of general polynomials of a given degree over $\mathbb{F}_q$ indeed holds, it could extend such a result to the class of all polynomials $f \in \mathbb{F}_p[x]$ of a given degree. We believe that our contributions in Chapters 3 and 4 could motivate authors to use results from Galois theory to prove the aforementioned randomness conjecture.

We highlight in Chapter 4 the particular nature of Chebyshev polynomials, observed previously by a number of authors. As mentioned in Section 4.3, some work has been done in the statistics of Chebyshev polynomials $T_d \in \mathbb{F}_p[x]$ (CHOU; SHPARLINSKI, 2004; VASIGA; SHALLIT, 2004). The authors of both of these papers obtain asymptotic estimates on the tail length of a random node of the functional graph of Chebyshev polynomials. They mention the expected cycle length of a random node as an interesting open problem, though a difficult one. Obtaining analogous results in extensions of $\mathbb{F}_p$ remains an interesting open problem. One could also consider using the same techniques for other rational functions.

BIBLIOGRAPHY

ABRAMOWITZ, M.; STEGUN, I. A. **Handbook of mathematical functions**. New York: Dover, 1965.

AHO, A. V.; HOPCROFT, J. E.; ULLMAN, J. D. **The design and analysis of computer algorithms**. Cambirdge: Addison-Wesley, 1974.

ANTON, H. **Elementary linear algebra**. New York: John Wiley & Sons, 2010.

ARNEY, J.; BENDER, E. Random mappings with constraints on coalescence and number of origins. **Pacific Journal of Mathematics**, Berkeley, v.103, n.2, p.269–294, 1982.

BABAI, L. **Graph isomorphism in quasipolynomial time**. arXiv:1512.03547.

BACH, E. Toward a theory of Pollard's rho method. **Information and Computation**, Cambridge, v.90, n.2, p.139–155, 1991.

BAILEY, D. V. et al. **Breaking ECC2K-130**. IACR Cryptology ePrint Archive.

BERNSTEIN, D. J.; LANGE, T. Two grumpy giants and a baby. In: ANTS X: PROCEEDINGS OF THE TENTH ALGORITHMIC NUMBER THEORY SYMPOSIUM, Berkeley. **Proceedings...** Mathematical Sciences Publishers, 2013. p.87––111. (Open Book Series, v.1).

BERNSTEIN, D. J.; LANGE, T.; SCHWABE, P. On the correct use of the negation map in the Pollard rho method. In: PUBLIC KEY CRYPTOGRAPHY - PKC 2011, New York. **Proceedings...** Springer, 2011. p.128–146. (Lecture Notes in Computer Science, v.6571).

BIRCH, B. J.; SWINNERTON-DYER, H. P. F. Note on a problem of Chowla. **Acta arithmetica**, Warszawa, v.5, n.4, p.417–423, 1959.

BLUM, L.; BLUM, M.; SHUB, M. A simple unpredictable pseudo-random number generator. **SIAM Journal on computing**, Philadelphia, v.15, n.2, p.364–383, 1986.

BOS, J. W.; COSTELLO, C.; MIELE, A. Elliptic and hyperelliptic curves: a practical security analysis. In: PUBLIC-KEY CRYPTOGRAPHY - PKC 2014, New York. **Proceedings...** Springer, 2014. p.203–220. (Lecture Notes in Computer Science, v.8383).

BOS, J. W. et al. Solving a 112-bit prime elliptic curve discrete logarithm problem on game consoles using sloppy reduction. **International Journal of Applied Cryptography**, Genève, v.2, n.3, p.212–228, 2012.

BOS, J. W.; KLEINJUNG, T.; LENSTRA, A. K. On the use of the negation map in the Pollard rho method. In: ALGORITHMIC NUMBER THEORY, New York. **Proceedings...** Springer, 2010. p.66–82. (Lecture Notes in Computer Science, v.6197).

BRENT, R. P. An improved Monte Carlo factorization algorithm. **BIT Numerical Mathematics**, New York, v.20, n.2, p.176–184, 1980.

BRENT, R. P.; POLLARD, J. M. Factorization of the eighth Fermat number. **Mathematics of Computation**, Providence, v.36, n.154, p.627–630, 1981.

CHOU, W.-S.; SHPARLINSKI, I. E. On the cycle structure of repeated exponentiation modulo a prime. **International Jounal of Number Theory**, Singapore, v.107, n.2, p.345–356, 2004.

CHUNG, K. L. **A course in probability theory**. New York: Academic press, 2001.

COHEN, S. The distribution of polynomials over finite fields. **Acta Arithmetica**, Warszawa, v.17, n.3, p.255–271, 1970.

CONE, M. M.; VENKATARAGHAVAN, R.; MCLAFFERTY, F. W. Computer-aided interpretation of mass spectra. 20. Molecular structure comparison program for the identification of maximal common substructures. **Journal of the American Chemical Society**, Washington, v.99, n.23, p.7668–7671, 1977.

CORMEN, T. H. et al. **Introduction to algorithms**. 3rd.ed. Cambridge: MIT press, 2009.

DEMARCO, L.; PILGRIM, K. Critical heights on the moduli space of polynomials. **Advances in Mathematics**, Amsterdam, v.226, n.1, p.350–372, 2011.

DRMOTA, M.; SORIA, M. Images and preimages in random mappings. **SIAM Journal on Discrete Mathematics**, Philadelphia, v.10, n.2, p.246–269, 1997.

DUCHON, P. et al. Boltzmann samplers for the random generation of combinatorial structures. **Combinatorics, Probability and Computing**, Cambirdge, v.13, n.4-5, p.577–625, 2004.

FLAJOLET, P.; ODLYZKO, A. M. Random mapping statistics. In: ADVANCES IN CRYPTOLOGY - EUROCRYPT '89, New York. **Proceedings...** Springer, 1990. p.329–354. (Lecture Notes in Computer Science, v.434).

FLAJOLET, P.; SEDGEWICK, R. **Analytic combinatorics**. Cambirdge: Cambridge University press, 2009.

FORTIN, S. **The graph isomorphism problem**. Edmonton: University of Alberta, 1996. 24p.

GALLANT, R.; LAMBERT, R.; VANSTONE, S. Improving the parallelized Pollard lambda search on anomalous binary curves. **Mathematics of Computation**, Providence, v.69, n.232, p.1699–1705, 2000.

GASSERT, T. A. Chebyshev action on finite fields. **Discrete Mathematics**, Cambridge, v.315, n.6, p.83–94, 2014.

GATHEN, J. von zur; PANARIO, D.; RICHMOND, B. Interval partitions and polynomial factorization. **Algorithmica**, New York, v.63, n.1-2, p.363–397, 2012.

GITTENBERGER, B. On the number of predecessors in constrained random mappings. **Statistics & probability letters**, Cambridge, v.36, n.1, p.29–34, 1997.

GREENE, D. H.; KNUTH, D. E. **Mathematics for the analysis of algorithms**. New York: Springer, 2007.

HARRIS, B. The asymptotic distribution of the order of elements in symmetric semigroups. **Journal of Combinatorial Theory, Series A**, Cambridge, v.15, n.1, p.66–74, 1973.

HWANG, H.-K. Large deviations for combinatorial distributions I. Central limit theorems. **The Annals of Applied Probability**, Beachwood, v.6, n.1, p.297–319, 1996.

HWANG, H.-K. Large deviations of combinatorial distributions II. Local Limit Theorems. **Annals of Applied Probability**, New York, v.8, n.1, p.163–181, 1998.

IRELAND, K.; ROSEN, M. I. **A classical introduction to modern number theory**. New York: Springer, 1990. v.84.

KNUTH, D. E. **The art of computer programming**: sorting and searching. Cambirdge: Addison-Wesley, 2011. v.3.

KNUTH, D. E. **The art of computer programming**: seminumerical algorithms. Cambirdge: Addison-Wesley, 2011. v.2.

KONYAGIN, S. V. et al. Functional graphs of polynomials over finite fields. **Journal of Combinatorial Theory Series B**, Cambridge, v.116, p.87–122, 2016.

LANG, S. **Algebra**. New York: Springer, 2002. (Graduate Texts in Mathematics).

LEHMER, D. H. An extended theory of Lucas' functions. **Annals of Mathematics**, New York, v.31, n.3, p.419–448, 1930.

LIDL, R.; NIEDERREITER, H. **Finite fields**. Cambridge: Cambridge University Press, 2008. (Encyclopedia of Mathematics and its Applications, v.20).

LUCAS, E. Théorie des fonctions numériques simplement périodiques. **American Journal of Mathematics**, New York, v.1, n.4, p.289–321, 1878.

LUKS, E. M. Isomorphism of graphs of bounded valence can be tested in polynomial time. **Journal of Computer and System Sciences**, Cambridge, v.25, n.1, p.42–65, 1982.

MACFIE, A.; PANARIO, D. Random mappings with restricted preimages. In: PROGRESS IN CRYPTOLOGY - LATINCRYPT 2012, New York. **Proceedings...** Springer, 2012. p.254–270. (Lecture Notes in Computer Science, v.7533).

MARTINS, R. S. V. et al. **An isomorphism test for functional graphs**. Preprint submitted for publication in Discrete Applied Mathematics.

MARTINS, R. S. V.; PANARIO, D. **On the heuristic of approximating polynomials over finite fields by random mappings**. Accepted for publication in the International Journal of Number Theory, DOI: 10.1142/S1793042116501219, 2016.

MCCANN, K.; WILLIAMS, K. The distribution of the residues of a quartic polynomial. **Glasgow Mathematical Journal**, Cambridge, v.8, n.2, p.67–88, 1967.

MCKAY, B.; PIPERNO, A. Practical graph isomorphism, II. **Journal of Symbolic Computation**, Cambridge, v.60, p.94–112, 2014.

MORAIN, F. Primality proving using elliptic curves: an update. In: ALGORITHMIC NUMBER THEORY, New York. **Proceedings...** Springer, 1998. p.111–127. (Lecture Notes in Computer Science, v.1423).

MULLEN, G. L.; PANARIO, D. **Handbook of finite fields**. Boca Raton: CRC Press, 2013.

POLLARD, J. M. A Monte Carlo method for factorization. **BIT Numerical Mathematics**, New York, v.15, n.3, p.331–334, 1975.

POLLARD, J. M. Monte Carlo methods for index computation (mod p). **Mathematics of computation**, Providence, v.32, n.143, p.918–924, 1978.

READ, R. C.; CORNEIL, D. G. The graph isomorphism disease. **Journal of Graph Theory**, New York, v.1, n.4, p.339–363, 1977.

ROHATGI, V. K.; SALEH, A. M. E. **An introduction to probability and statistics**. New York: John Wiley & Sons, 2011. v.910.

RUBIN, H.; SITGREAVES, R. **Probability distributions related to random transformations of a finite set**. Stanford: Applied Mathematics and Statistics Laboratory, 1953. 51p.

RUDIN, W. **Principles of mathematical analysis**. New York: McGraw-Hill, 1964.

SCHMUTZ, E. Period lengths for iterated functions. **Combinatorics, Probability and Computing**, Cambridge, v.20, n.2, p.289–298, 2011.

SEDGEWICK, R.; FLAJOLET, P. **An introduction to the analysis of algorithms**. Cambirdge: Addison-Wesley, 2013.

SILVERMAN, J. H. **The arithmetic of dynamical systems**. New York: Springer, 2007.

STEELE, J. M. **The Cauchy-Schwarz master class**: an introduction to the art of mathematical inequalities. Cambridge: Cambridge University Press, 2004.

STERNECK, R. D. von. Über die Anzahl inkongruenter Werte, die eine ganze Funktion dritten Grades annimmt. **Sitzungsber. Akad. Wiss. Wien (2A)**, Vienna, v.114, p.711–717, 1908.

SUN, Z.-H. On the number of incongruent residues of $x^4 + ax^2 + bx$ modulo p. **Journal of Number Theory**, Cambridge, v.119, n.2, p.210–241, 2006.

SUTHERLAND, N. Computing Galois groups of polynomials (especially over function fields of prime characteristic). **Journal of Symbolic Computation**, Cambridge, v.71, n.C, p.73–97, 2015.

SWAN, R. G. Factorization of polynomials over finite fields. **Pacific Journal of Mathematics**, Berkeley, v.12, n.3, p.1099–1106, 1962.

TESKE, E. A space efficient algorithm for group structure computation. **Mathematics of Computation**, Providence, v.67, n.224, p.1637–1663, 1998.

TESKE, E. On random walks for Pollard's rho method. **Mathematics of computation**, Providence, v.70, n.234, p.809–825, 2001.

TESKE, E.; WILLIAMS, H. C. **A note on Shanks's chains of primes**. New York: Springer, 2000.

TURNWALD, G. A new criterion for permutation polynomials. **Finite Fields and Their Applications**, Cambridge, v.1, n.1, p.64–82, 1995.

UCHIYAMA, S. Sur le nombre des valeurs distinctes d'un polynôme à coefficients dans un corps fini. **Proceedings of the Japan Academy**, Tokyo, v.30, n.10, p.930–933, 1954.

VAN OORSCHOT, P. C.; WIENER, M. J. Parallel collision search with cryptanalytic applications. **Journal of cryptology**, New York, v.12, n.1, p.1–28, 1999.

VASIGA, T.; SHALLIT, J. On the iteration of certain quadratic maps over GF (p). **Discrete Mathematics**, Cambridge, v.277, n.1, p.219–240, 2004.

VON ZUR GATHEN, J. Irreducible trinomials over finite fields. **Mathematics of Computation**, Providence, v.72, n.244, p.1987–2000, 2003.

WIENER, M. J.; ZUCCHERATO, R. J. Faster attacks on elliptic curve cryptosystems. In: SELECTED AREAS IN CRYPTOGRAPHY, New York. **Proceedings...** Springer, 1999. p.190–200. (Lecture Notes in Computer Science, v.1556).

WILLIAMS, K. S. On general polynomials. **Canadian Mathematical Bulletin**, Ottawa, v.10, n.4, 1967.

ZHANG, F.; WANG, P. Speeding up elliptic curve discrete logarithm computations with point halving. **Designs, codes and cryptography**, New York, v.67, n.2, p.197–208, 2013.