# Introduction

This thesis has four main elements: dynamics, algebraic number theory (in function fields), group theory, and probability. The dynamics provides inspiration for the problem, the algebraic number theory allows an important reformulation, the group theory does the heavy lifting, and the probability theory synthesizes the group theory into a form that solves the problem. Chapter 1 gives the dynamical background, Chapter 2 the number-theoretic translation, and Chapter 3 the group-theoretic results. Chapter 4 deals almost solely with probability, while Chapter 5 applies the probability to give a solution to the problem.

In this introduction, we sketch of the origins of the problem, state the problem, and give a detailed outline of our solution. The origins of the problem and the first stages of the solution lie in the field of algebraic dynamics, which can be broadly defined as the study of function iteration over algebraic/arithmetical sets, such as algebraic number fields and rings, polynomial rings, finite fields, p-adic fields, algebraic curves, etc. This relatively new field evolved naturally from the field of complex dynamics, whose roots stretch back to the work of Julia and Fatou in the early 1900s. Complex dynamics enjoyed an explosion of deep results in the 1980s; one could note in particular Benoit Mandelbrot's popularization of the remarkable set that bears his name, Dennis Sullivan's No Wandering Domains theorem [36], and Curt McMullen's proof of the nonexistence of generally convergent algorithms in degree larger than three [23].

In the 1990s, some attention began to turn to $p$-adic analogues of the blossoming complex theory. Michael Herman and Fields medalist Jean-Christophe Yoccoz [17] had already gone in this direction with their 1983 paper on a non-Archimedean version of Siegel's linearization theorem [34]. Rob Benedetto proved a partial analogue of the No Wandering Domains theorem in 2000 [6] and Juan Rivera-Letelier, a student of Yoccoz, fleshed out much of the theory of $p$-adic Fatou and Julia sets in the early part of the 2000s [29, 30]. An unexpected impetus for exploring $p$-adic dynamics has come from physicists, who began to explore what the world looked like through $p$-adic eyes. The first paper whose title contained the phrase "$p$-adic dynamics" appeared in the Physics literature in 1989 [37].

The work presented in this thesis stems from a question regarding the degree to which $p$-adic dynamics and complex dynamics are analogous. This question deals with the $p$-adic analogue of the complex Mandelbrot set. The complex Mandelbrot set, defined to be

$$M = \{c \in \mathbb{C} : 0 \text{ has a bounded orbit under iteration of } z^2 + c\}$$

not only created a stir among researchers, but its striking images reached an audience far beyond the mathematical community. It has assumed a place as one of the most widely recognized mathematical objects. Thus the $p$-adic analogue of $M$, at first blush, seems to promise a similar treasure trove of complexity. This promise unfortunately proves false, as the analogous set is simply the closed unit disk (for $p \neq 2$).

Peering into the set more closely, however, reveals a particular subset that is more interesting. It is the subset of parameter values $c$ such that $z^2 + c$ is hyperbolic; see below for a definition of hyperbolicity and see (1) for a definition of the set. Its complex analogue has been much studied [1, 14, 19, 22]. This complex analogue is a large subset of $M$, accounting for at least 96% of the area [11], and is conjectured to be the interior of $M$ [24]. The problem this thesis sets out to resolve is to determine the "size" of the hyperbolic subset of the $p$-adic Mandelbrot set. We show that it is in a certain sense a measure zero subset, contrasting sharply with the complex case.

The first hurdle is to say what is meant by size: there is no suitable notion of measure on the $p$-adic analogue of $\mathbb{C}$ because it is not locally compact. The way around this, and the first step in the solution of the above problem, is to use the reduction homomorphism to translate the problem into one of dynamics over $\overline{\mathbb{F}}_p$. In $\overline{\mathbb{F}}_p$ we define density measures closely related to the well-known Dirichlet density and natural density (see e.g. [20]). The proof then follows a path through algebraic number theory, then the Galois theory of function fields, and eventually into the realm of stochastic processes, where it reaches its conclusion. This method of proof appears to be highly unusual, and may be fruitful in answering other density questions regarding dynamically defined sets. We now give a detailed outline of the argument.

In Chapter 1 we begin with some definitions and background. We call a rational function *hyperbolic* if all its critical points are attracted to attracting cycles (see page 8 for definitions of these terms). Hyperbolic maps have many nice properties, and are the subject of the biggest unsolved conjecture in complex dynamics (see page 9 for a statement and [24] for more detail). The map $z^2 + c$ has critical points at infinity and 0, and infinity is an attracting fixed point. Therefore $z^2 + c$ is hyperbolic if and only if 0 is attracted to an attracting cycle. Thus the set

$$\mathcal{H}(\mathbb{C}) = \{c \in M : 0 \text{ is attracted to an attracting cycle of } z^2 + c\} \tag{1}$$

is the *hyperbolic subset* of $M$. We examine the analogous set defined over the field $\mathbb{C}_p$, which is the smallest complete, algebraically closed extension of $\mathbb{Q}_p$; it is therefore the $p$-adic analogue of $\mathbb{C}$. Let $M_p$ be the natural analogue of $M$ over $\mathbb{C}_p$. It's easily seen that $M_p = \{c \in \mathbb{C}_p : |c| \leq 1\}$ provided that $p \neq 2$. (Proposition 1.2). *Throughout this discussion, we take p to be a prime different from* 2.

However, the subset

$$\mathcal{H}(\mathbb{C}_p) = \{c \in M_p : 0 \text{ is attracted to an attracting cycle of } z^2 + c\}$$

is not so easily characterized. Letting $\phi : \{|c| \leq 1\} \to \overline{\mathbb{F}}_p$ be the reduction homomorphism (see (1.2)), we establish in Corollary 1.6 that $\mathcal{H}(\mathbb{C}_p) = \phi^{-1}(\mathcal{H}(\overline{\mathbb{F}}_p))$, where

$$\mathcal{H}(\overline{\mathbb{F}}_p) = \{\alpha \in \overline{\mathbb{F}}_p : 0 \text{ is periodic under iteration of } x^2 + \alpha\}.$$

(Note that by periodic we mean that the orbit of 0 is a cycle; some authors refer to this as purely periodic.) We define two notions of density for subsets of $\overline{\mathbb{F}}_p$, called Dirichlet density and natural density, that are closely related to the densities of the same names defined for subsets of primes in $\mathbb{F}_p[x]$. We denote Dirichlet density by $\delta$ and natural density by $D$ (see (1.6) and (1.7) for the definitions). Our main result (Theorem 1.7) is that $\delta(\mathcal{H}(\overline{\mathbb{F}}_p)) = 0$, and its proof is the principal goal of all of our subsequent work. We also establish, using a similar method, a companion result: we show $D(\mathcal{H}(\overline{\mathbb{F}}_p)) = 0$ for $p \equiv 3 \bmod 4$, and we conjecture that $D(\mathcal{H}(\overline{\mathbb{F}}_p)) = 0$ holds for all $p \neq 2$.

In Chapter 2, we take the first steps toward a proof by giving two translations of the problem. Although the definition of $\mathcal{H}(\overline{\mathbb{F}}_p)$ says that the forward orbit of 0 under iteration of $x^2+\alpha$ is a cycle, we focus in Section 2.1 on the *inverse* orbit of 0 under $x^2 + \alpha$. Clearly for any $\alpha \in \overline{\mathbb{F}}_p$ the forward orbit of 0 under iteration of $x^2 + \alpha$ is contained in $\mathbb{F}_p(\alpha)$. If 0 is periodic, however, this forward orbit coincides with one branch of the inverse orbit of 0 in $\overline{\mathbb{F}}_p$, and thus 0 has $n$th preimages in $\mathbb{F}_p(\alpha)$ for each $n \geq 1$. We show (Proposition 2.1) that the converse is also true. Therefore, setting $f_\alpha = x^2 + \alpha$, and denoting by $f_\alpha^{-n}(0)$ the set of $n$th preimages of 0 (in $\overline{\mathbb{F}}_p$) under iteration of $f_\alpha$, we have

$$\mathcal{H}(\overline{\mathbb{F}}_p) = \{\alpha \in \overline{\mathbb{F}}_p : f_\alpha^{-n}(0) \cap \mathbb{F}_p(\alpha) \neq \emptyset \text{ for all } n \geq 1\}.$$

We can therefore define a sequence of sets

$$\mathcal{I}_n = \{\alpha \in \overline{\mathbb{F}}_p : f_\alpha^{-n}(0) \cap \mathbb{F}_p(\alpha) \neq \emptyset\} \tag{2}$$

that serve as progressively better "approximations" of $\mathcal{H}(\overline{\mathbb{F}}_p)$ in the sense that $\mathcal{I}_n \supseteq \mathcal{I}_{n+1}$ and $\mathcal{H}(\overline{\mathbb{F}}_p) = \bigcap_n \mathcal{I}_n$. We show that if $\delta(\mathcal{I}_n)$ exists for all $n$ and $\lim\limits_{n\to\infty} \delta(\mathcal{I}_n) = 0$, then $\delta(\mathcal{H}(\overline{\mathbb{F}}_p)) = 0$.

In Section 2.2, we prove that $\delta(\mathcal{I}_n)$ exists and give a method of computing it using the Galois groups of certain algebraic extensions of $\mathbb{F}_p(x)$. Our main tool in this endeavor is the Tchebotarev Density theorem for function fields. In order to use it, we find a set of primes in $\mathbb{F}_p[x]$ that is expressible in terms of the Artin symbol (see page 25 for a definition), and whose Dirichlet density (in the sense of (2.3)) is equal to $\delta(\mathcal{I}_n)$. Note that the set $f_\alpha^{-n}(0)$ consists of the roots of $f_\alpha^n$, and thus $f_\alpha^{-n}(0) \cap \mathbb{F}_p(\alpha) \neq \emptyset$ if and only if the factorization of $f_\alpha^n$ over $\mathbb{F}_p(\alpha)$ contains at least one linear term. There is a thus a close relationship between $\mathcal{I}_n$ and the following set of primes in $\mathbb{F}_p[x]$:

$$I_n = \{\mathfrak{p} \subseteq \mathbb{F}_p[x] : f_x^n \text{ has a linear factor mod } \mathfrak{p}\},$$

where $f_x = y^2 + x \in \mathbb{F}_p(x)[y]$. Indeed, we show that $\delta(\mathcal{I}_n) = \delta(I_n)$, where this second Dirichlet density is the usual one for sets of primes in a function field (2.3). We then use some standard arguments in algebraic number theory to show that $I_n$ differs by only finitely many primes from a set of primes defined in terms of the Artin symbol (2.14). This allows us to apply the Tchebotarev Density theorem. Let $K_n$ be the splitting field of $f_x^n$ over $K = \mathbb{F}_p(x)$, and $G_n = \mathrm{Gal}\,(K_n/K)$. In Theorem 2.18 we show that $\delta(\mathcal{I}_n)$ exists for all $n$ and equals

$$\frac{1}{\#G_n} \# \{g \in G_n : g \text{ fixes at least one root of } f_x^n\}. \tag{3}$$

The same statement hold for $D(\mathcal{I}_n)$, provided that the extensions $K_n/K$ are geometric for all $n$ (see Definition 2.14), a statement we conjecture to hold for $p \neq 2$ but which we can only show when $p \equiv 3 \bmod 4$ (Corollary 3.40). To illustrate Theorem 2.18, we give here a few values of $\delta(\mathcal{I}_n)$. We can describe the $n = 1$ case completely: the roots of $f_x$ are $\{\sqrt{-x}, -\sqrt{-x}\}$, which we label $\{a_1, a_2\}$. Clearly we have $G_1 = \{e, (a_1 \;\; a_2)\}$, whence $\delta(\mathcal{I}_1) = 1/2$. In Example 2.19, page 29, we work out the case $n = 2$, showing that $\delta(\mathcal{I}_2) = 3/8$. Only with significantly more work (Theorem 3.2, (5.28), and Corollary 5.11) can we show $\delta(\mathcal{I}_3) = 39/128$. Moreover we note in Chapter 3 (see the discussion on page 56) that for $n > 7$, $\delta(\mathcal{I}_n)$ may depend on the prime $p$, and cannot in general be easily computed.

In Chapter 3 we undertake an analysis of the groups $G_n$. We do this through the study of $H_n = \mathrm{Gal}\,(K_n/K_{n-1})$. In (3.2), we show that $K_n$ is obtained from $K_{n-1}$ by adjoining the square roots of $2^{n-1}$ elements. Thus $|H_n| \leq 2^{2^{n-1}}$, and we call $H_n$ *maximal* if this inequality is an equality. One principal result of the chapter is that for all $p \neq 2$, $H_n$ is maximal for $n$ squarefree, and if $p \equiv 3 \bmod 4$ then $H_n$ is maximal for all $n$ (Theorem 3.2). The other result that is of central importance in later chapters is Corollary 3.23 at the end of Section 3.2 (see below for explanation).

In Section 3.1, we prove some basic properties about $f_x^n$, and we introduce the polynomials $p_n \in \mathbb{F}_p[x]$, defined by $p_1 = x$ and $p_n = p_{n-1}^2 + x$ for $n \geq 2$, which play an important role. We show that the discriminant of $f_x^n$ is a product of powers of $p_i$ for $i \leq n$ (Proposition 3.9) and use this to prove that $G_n$ is not alternating, i.e. composed of even permutations (Corollary 3.10). In Section 3.2 we examine the center of $G_n$. We note that $G_n$ is a 2-group, and we establish a series of propositions about 2-groups acting on certain sets. This culminates in Theorem 3.22, which states that if $G_n$ is not alternating then a certain permutation of the roots of $f_x^n$ must lie in $G_n$. Indeed, as shown in Corollary 3.23, this permutation must lie in $H_n$, proving that $H_n$ is nontrivial for all $n$.

In Section 3.3 we use abelian Kummer Theory to show that $H_n$ is maximal if and only if $p_n$ is a square in $K_{n-1}$ (Theorem 3.27). This relies heavily on Corollary 3.23. In Section 3.4 we use the work of the previous sections to show that $H_n$ is maximal if and only if a certain element $\Phi_n \in \mathbb{F}_p[x]$ is a not a square in $K = \mathbb{F}_p(x)$ (Theorem 3.38). Specifically, $\Phi_n$ is the primitive part of $p_n$:

$$\Phi_n = \prod_{d|n} (p_d)^{\mu(n/d)}.$$

One can easily show that the degree of $\Phi_n$ is odd when $n$ is squarefree (Corollary 3.30), establishing the maximality of $H_n$ for squarefree $n$. When $n$ is not squarefree one cannot rule out the possibility that $\Phi_n$ is a square in $K$. However, the facts that $\Phi_n$ is separable over $\mathbb{Q}$ for all $n$ (see the proof of Theorem 3.2, page 54) and the degree of $\Phi_n$ grows like $2^n$ suggest this is unlikely. We thus conjecture that $H_n$ is maximal for all $n$. In section 3.5 we adapt an argument of Stoll [35] to show that if $p \equiv 3 \bmod 4$ then $H_n$ is maximal for all $n$.

The results of Chapter 3 provide some insight into the structure of $H_n$ and therefore $G_n$, but there is no obvious way to use them to compute the limit as $n \to \infty$ of the expression in (3). In Chapters 4 and 5 we build a stochastic process where the main results of Chapter 3 have a natural interpretation. The tools of the theory of stochastic processes then allow us to prove this process is eventually 0 with probability 1, and this is enough to establish Theorem 1.7.

Chapter 4 focuses on proving the process we seek exists and fleshing out definitions and standard results from the theory of stochastic processes. In Section 4.1 we recall that a discrete-time stochastic process (or simply *process* for short) is a sequence $\{X_n\}_{n\geq 0}$ of random variables defined on a common probability space. We consider only processes whose random variables take positive-integer values. Such a process can be thought of as a game of chance, with $X_n$ denoting a gambler's score at turn $n$. We prove that there exists a process where the probability of the gambler having score $t$ at turn $n$ is determined by the structure of $G_n$. Specifically,

$$\mathbf{P}(X_n = t) = \frac{1}{\#G_n} \# \{g \in G_n : g \text{ fixes } t \text{ roots of } f_x^n\} \tag{4}$$

We actually prove an even stronger property (see (4.3)). From (4) and the remark immediately before (3) it follows that

$$\delta(\mathcal{I}_n) = \mathbf{P}(X_n > 0). \tag{5}$$

For more on why probability theory is a relatively natural tool in this context, see the introduction to Chapter 4 on page 59.

The main work of section 4.1 is to show that a process satisfying (4) exists. We call this process the Galois process of the iterates of $f$, or GP($f$) for short. (Because there is no possibility of ambiguity, we drop the $x$ from previous notation and simply write $f = y^2 + x$.) In Section 4.2 we give some probabilistic background, including definitions of martingales and Markov chains and one version of the basic martingale convergence theorem (Theorem 4.9). In Section 4.3 we present the basic theory of branching processes, which not only is useful in Chapter 5 but also illustrates some of the definitions of Section 4.2.

In Chapter 5 all of the threads come together. In Section 5.1 we use Corollary 3.23, which guarantees the existence of a certain type of element in each $H_n$, to establish that GP($f$) is a martingale. Thus GP($f$) is the first known example of a class of processes we call *Galois martingales* (see the remark on page 80), giving some justification for the first two words of this thesis' title. The martingale convergence theorem then shows that with probability 1 the sequence $\{X_n(\omega)\}_{n\geq 0}$ is eventually constant (where $\omega$ is any element in the underlying probability space).

In Section 5.2 we compute, under the assumption that $H_n$ is maximal, the conditional distribution of $X_n$ given $X_{n-1} = t$ for any value of $t$. Thus, using the metaphor of the gambler, when $n$ is such that $H_n$ is maximal we have explicit information about the probability that the gambler's score goes from $t$ to $t'$ at turn $n$. Using Theorem 3.2, which says that $H_n$ is maximal when $n$ is squarefree, we show that for any $t \geq 1$ and $m \geq 1$,

$$\mathbf{P}\left(X_n = t \text{ for all } n \geq m\right) = 0.$$

This is Theorem 5.8. It quickly follows that GP($f$) converges to 0 with probability 1. We deduce from this in Section 5.3 that $\lim\limits_{n\to\infty} \mathbf{P}(X_n > 0) = 0$, and then from (5) and the remark following (2) we conclude that $\delta(\mathcal{H}(\overline{\mathbb{F}}_p)) = 0$, which proves Theorem 1.7. Finally, we show that under the assumption that $H_n$ is maximal for all $n$, GP($f$) is a particularly simple branching process. We also give explicit values for $\delta(\mathcal{I}_n)$ under this assumption (see (5.28) and Corollary 5.11).