# Iterated Galois towers, their associated martingales, and the p-adic Mandelbrot set

Rafe Jones
University of Wisconsin-Madison

**Abstract**

We study the Galois tower generated by iterates of a quadratic polynomial $f$ defined over an arbitrary field. One question of interest is to count the proportion $a_n$ of elements at each level that fix at least one root; in the global field case these correspond to unramified primes in the base field that have a divisor at level $n$ of residue class degree one. We thus define a stochastic process associated to the tower that encodes root-fixing information at each level. We develop a uniqueness result for certain permutation groups, and use this to show that for many $f$ each level of the tower contains a certain central involution. It follows that the associated stochastic process is a martingale, and convergence theorems then allow us to establish a criterion for showing that $a_n$ tends to 0. As an application, we study the dynamics of the family $x^2 + c \in \overline{\mathbb{F}}_p[x]$, and this in turn is used to establish a basic property of the $p$-adic Mandelbrot set.

## 1 Introduction

Let $L$ be a field and $f \in L[x]$. Denote by $f^{\circ n}$ the $n$th iterate of $f$, that is $f^{\circ 1} = f$ and $f^{\circ n} = f \circ f^{\circ n-1}$ for $n \geq 2$. Let $L_n(f)$ be the splitting field over $L$ of $f^{\circ n}$, and let $G_n(f) = \mathrm{Gal}\,(L_n(f)/L)$. The profinite group $G(f) = \varprojlim G_n(f)$ remains rather mysterious in general, having been studied broadly only by Odoni [8]. Even the case of $f$ quadratic remains largely unresolved, although some progress has been made [8, 9, 10, 14]. In this article we use tools from the theory of stochastic processes to study properties of $G(f)$ that have arithmetic applications.

The construction is as follows. Given any field $L$ and a collection $\mathcal{F}$ of separable polynomials $f_1, f_2, \ldots$ in $L[x]$, denote by $L(f_n)$ the splitting field of $f_n$ over $L$, and let $G(f_n) = \mathrm{Gal}\,(L(f_n)/L)$. Suppose that $L(f_n) \supseteq L(f_{n-1})$ for all $n \geq 2$, and let $G(\mathcal{F}) = \varprojlim G(f_n)$. Take $\mathbf{P}$ to be the Haar measure on $G(\mathcal{F})$ with $\mathbf{P}(G(\mathcal{F})) = 1$, and $\psi_n$ to be the natural projection $G(\mathcal{F}) \to G(f_n)$. We define random variables on $G(\mathcal{F})$ by setting $X_n(g)$ to be the number of roots of $f_n$ fixed by $\psi_n(g)$. It follows that

$$\mathbf{P}(X_n > 0) = \frac{1}{\#G(f_n)} \cdot \#\{g \in G(f_n) : g \text{ fixes at least one root of } f_n\}. \tag{1}$$

Recall that a stochastic process is simply an infinite collection of random variables defined on a common probability space. We refer to the random variables in (1) as the *Galois process* of $\mathcal{F}$ and denote it $GP(\mathcal{F})$. In the case $f \in L[x]$ and $\mathcal{F} = \{f, f^{\circ 2}, f^{\circ 3}, \ldots\}$, we write $GP(f)$ instead of $GP(\mathcal{F})$ and $G_n(f)$ instead of $G(f^{\circ n})$. To state our main result on Galois processes, we require the following definition:

**Definition 1.1.** A stochastic process $X_1, X_2, \ldots$ taking values in $\mathbb{Z}$ is a *martingale* if for all $n \geq 2$ and any $t_i \in \mathbb{Z}$,

$$E(X_n \mid X_1 = t_1, X_2 = t_2, \ldots, X_{n-1} = t_{n-1}) = t_{n-1},$$

provided $\mathbf{P}(X_1 = t_1, X_2 = t_2, \ldots, X_{n-1} = t_{n-1}) > 0$.

We also define the *adjusted forward orbit* of a point $l \in L$ under a polynomial $f \in L[x]$ with leading coefficient $a$ to be the set $\{-af(l)\} \cup \{f^{\circ n}(l) : n = 2, 3, \ldots\}$.

**Theorem 1.2.** *Let $L$ be a field of characteristic $\neq 2$, take $f \in L[x]$ of degree two, and suppose that the adjusted forward orbit of the unique finite critical point of $f$ contains no squares. Then $GP(f)$ is a martingale.*

*Remark.* Theorem 1.2 (and also Theorem 1.3) are true as long as, for all $n$, $f^{\circ n}$ is separable and irreducible and $\operatorname{Disc} f^{\circ n}$ is not a square. The hypothesis regarding the critical point ensures this in characteristic $\neq 2$ (see Lemma 4.10) and is easy to check. A version of Theorem 1.2 (allowing the process to begin with $X_k$ for a suitable $k \geq 1$) should remain true if one allows $\operatorname{Disc} f^{\circ n}$ to be a square for finitely many $n$ and $f^{\circ n}$ to be reducible with a number of irreducible factors bounded independently of $n$.

To prove Theorem 1.2 we develop a uniqueness result on sets of fibers for certain morphisms of $G_n(f)$-sets, where $f$ satisfies the conditions of Theorem 1.2. This uniqueness property is used to show that for each $n$ a certain involution must lie in the center of $G_n(f)$. The presence of this involution leads to the proof of Theorem 1.2.

If the hypotheses of Theorem 1.2 are verified, a basic martingale convergence theorem yields

$$\mathbf{P}(\{g \in G(f) : X_1(g), X_2(g), \ldots \text{ is eventually constant}\}) = 1.$$

Let $f$ satisfy the hypotheses of Theorem 1.2, let $L_n(f)$ be the splitting field of the $n$th iterate of $f$, and let $H_n(f) = \operatorname{Gal}(L_n(f)/L_{n-1}(f))$. Then $H_n(f) \cong (\mathbb{Z}/2\mathbb{Z})^m$ for some $0 \leq m \leq 2^{n-1}$, and we call $H_n(f)$ *maximal* if $m = 2^{n-1}$. We show that if $H_n(f)$ is maximal, then for any $u > 0$ and $m < n$, we have

$$\mathbf{P}(X_n = u \mid X_{n-1} = u, \ldots, X_m = u) \leq 1/2,$$

provided $\mathbf{P}(X_{n-1} = u, \ldots, X_m = u) > 0$. This immediately gives:

**Theorem 1.3.** *Let $L$ be a field of characteristic $\neq 2$, take $f \in L[x]$ of degree two, and suppose that the adjusted forward orbit of the unique finite critical point of $f$ contains no squares. Suppose also that $H_n(f)$ is maximal for infinitely many $n$. Then $GP(f)$ converges to $0$, i.e.*

$$\lim_{n \to \infty} \mathbf{P}(X_n = 0) = 1.$$

As an application of Theorem 1.3, we establish a basic property of the $p$-adic Mandelbrot set. This requires the development of some background. Given a field $K$ and an absolute value $|\cdot|$ on $K$, we define the *Mandelbrot set* of $K$ to be

$$M(K) = \{c \in K : 0 \text{ has bounded orbit under iteration of } x^2 + c\}, \tag{2}$$

where we mean bounded with respect to $|\cdot|$.

We consider a subset of $M(K)$ that is motivated by the well-known case $K = \mathbb{C}$. Recall that $\phi \in \mathbb{C}(z)$ is said to be *hyperbolic* if each critical point of $\phi$ tends to an attracting cycle under iteration [6]. We therefore define the *hyperbolic Mandelbrot set* to be

$$\mathcal{H}(K) = \{c \in M(K) : 0 \text{ tends to a formally attracting cycle under iteration of } x^2 + c\}, \tag{3}$$

where by a formally attracting cycle of $f(x) = x^2 + c$ we mean that $|f'| < 1$ at all points of the cycle (see Section 2 for more detailed definitions). When the topology on $K$ induced by $|\cdot|$ gives rise to nontrivial geometry, e.g. $K = \mathbb{C}$ and $K = \mathbb{C}_p$, a formally attracting cycle is also geometrically attracting. We may decompose $\mathcal{H}(K)$ into a disjoint union of open components $\mathcal{H}(K)^{(i)}$ corresponding to $c$ where 0 tends to a formally attracting $i$- cycle . In the complex case these components form some of the most visible features of $M(\mathbb{C})$. For instance, $\mathcal{H}(\mathbb{C})^{(1)}$ is the main cardioid, and $\mathcal{H}(\mathbb{C})^{(2)}$ is the circle tangent to the cardioid on the real axis. Conjecturally, $\mathcal{H}(\mathbb{C})$ is the interior of $M(\mathbb{C})$; this is the simplest case of the celebrated conjecture that hyperbolic rational maps are open and dense in the space of rational maps of given degree [6]. Moreover, both sets are Lebesgue measurable and the measure of $\mathcal{H}(\mathbb{C})$ exceeds 1.503 while the measure of $M(\mathbb{C})$ is less than 1.562 [1].

We consider the size of $\mathcal{H}(K)$ relative to $M(K)$ for $K = \mathbb{C}_p$, the smallest complete, algebraically closed extension of $\mathbb{Q}_p$. The set $M(\mathbb{C}_p)$ proves far less topologically interesting than $M(\mathbb{C})$, as $M(\mathbb{C}_p)$ is just the closed unit disk $\mathcal{O}_p$ in $\mathbb{C}_p$. However, $\mathcal{H}(\mathbb{C}_p)$ is not so simple. Letting $\phi : \mathcal{O}_p \to \overline{\mathbb{F}}_p$ be the reduction homomorphism, we show $\mathcal{H}(\mathbb{C}_p) = \phi^{-1}(\mathcal{H}(\overline{\mathbb{F}}_p))$ for $p \neq 2$. Note that since $\overline{\mathbb{F}}_p$ admits only the trivial absolute value, we have

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

provided $p \neq 2$. Given $\mathcal{C} \subseteq \overline{\mathbb{F}}_p$, we define its density to be:

$$D(\mathcal{C}) = \lim_{s \to 1^+} \frac{\sum_{\alpha \in \mathcal{C}} (\deg \alpha)^{-1} N(\alpha)^{-s}}{\sum_{\alpha \in \overline{\mathbb{F}}_p} (\deg \alpha)^{-1} N(\alpha)^{-s}}, \tag{4}$$

where $\deg \alpha = [\mathbb{F}_p(\alpha) : \mathbb{F}_p]$, and $N(\alpha) = p^{\deg \alpha}$. In a natural sense, $D(\mathcal{H}(\overline{\mathbb{F}}_p))$ measures the density of $\mathcal{H}(\mathbb{C}_p)$. We use Theorem 1.3 to prove:

**Theorem 1.4.** *For $p \neq 2$, $D(\mathcal{H}(\overline{\mathbb{F}}_p)) = 0$.*

In the case $p = 2$, it is trivial to show $\mathcal{H}(\overline{\mathbb{F}}_p) = \overline{\mathbb{F}}_p$, as all points are critical. We remark that there is another notion of density given by $\delta(\mathcal{C}) = \lim_{k\to\infty}(\#\mathcal{C} \cap \mathbb{F}_{p^k}/p^k)$. When $\delta(\mathcal{C})$ exists, then so does $D(\mathcal{C})$ and the two are equal; however, there are sets $\mathcal{C}$ for which $D(\mathcal{C})$ exists and

$\delta(\mathcal{C})$ does not. It is a consequence of Conjecture 6.7 that $\delta(\mathcal{H}(\overline{\mathbb{F}}_p)) = 0$ for $p \neq 2$, and this can be proven unconditionally if $p \equiv 3 \pmod 4$ (see the discussion following Conjecture 6.7).

To prove Theorem 1.4, we put $f_c = x^2 + c$ and $f_c^{-\circ n}(0) = \{b \in \overline{\mathbb{F}}_p : f_c^{\circ n}(b) = 0\}$, and introduce sets

$$\mathcal{I}_n = \{c \in \overline{\mathbb{F}}_p : f_c^{-\circ n}(0) \cap \mathbb{F}_p(c) \neq \emptyset\}. \tag{5}$$

We show that $\mathcal{I}_n \supseteq \mathcal{I}_{n+1}$ for all $n \geq 1$ and $\mathcal{H}(\overline{\mathbb{F}}_p) = \bigcap_{n\geq 1} \mathcal{I}_n$. It follows that if $D(\mathcal{I}_n)$ exists for all $n$ and $\lim_{n\to\infty} D(\mathcal{I}_n) = 0$, then $D(\mathcal{H}(\overline{\mathbb{F}}_p)) = 0$. We then use the Tchebotarev density theorem for function fields to show:

**Theorem 1.5.** *Set $L = \mathbb{F}_p(t)$ $(p \neq 2)$, $f = x^2 + t \in L[x]$, and let $L_n(f)$ be the splitting field over $L$ of $f^{\circ n}$, the $n$th iterate of $f$. Put $G_n(f) = \mathrm{Gal}\,(L_n(f)/L)$, and let $\mathcal{I}_n$ be as in* (5)*. Then*

$$D(\mathcal{I}_n) = \frac{1}{\#G_n(f)} \cdot \#\{g \in G_n(f) : g \text{ fixes at least one root of } f^{\circ n}\}.$$

We end with an analysis of the Galois groups of iterates of $f = x^2 + t$ over $L = k(t)$, where $k$ is a field of characteristic different from 2 and $t$ is transcendental over $k$ (cf. [10, 14]). We prove that $G_n(f)$ is maximal when $n$ is squarefree, and Theorem 1.3 applies to show Theorem 1.4.

The layout of the article follows the order in which the original work was done. In Sections 2 and 3 we give background on dynamics and $\mathbb{C}_p$ and prove Theorem 1.5. In Section 4 we introduce Galois processes and prove Theorem 1.2. In Section 5 we prove Theorem 1.3 and discuss the behavior of $GP(f)$ when $G_n(f)$ is maximal for all $n$. In Section 6 we analyze the Galois groups of iterates of $x^2 + t$ and obtain the proof of Theorem 1.4.

## 2 Background on Dynamics and $\mathbb{C}_p$

Let $K$ be a field and $|\cdot|$ an absolute value on $K$. Let $R \in K(x)$. We recall that an $n$-cycle of $R$ is a collection of distinct points $c_1, \ldots, c_n$ such that $R(c_i) = c_{i+1}$ for $1 \leq i \leq n-1$ and $R(c_n) = c_1$. We refer to any point $c$ in a cycle as *periodic* under $R$, and if $c$ is contained in an $n$-cycle we say $c$ has *period $n$*. A cycle $c_1, \ldots, c_n$ is *formally attracting* if $|(R^{\circ n})'(c_i)| < 1$ for any $i$ (equivalently, for all $i$). We use this terminology rather than the more geometrically suggestive "attracting" since we wish to work with fields where there is no nontrivial topology. We say a point $b \in K$ tends to the cycle $c_1, \ldots, c_n$ under iteration of $R$ if given $\epsilon > 0$, there exists $M$ such that $m \geq M$ implies that $\left|R^{\circ nm+i}(b) - c_i\right| < \epsilon$ for $i = 1, 2, \ldots, n$, up to a relabeling of the $c_i$.

Let $c \in K$ and put $f_c = x^2 + c$. Recall from (2) and (3) the definitions of the Mandelbrot set and hyperbolic Mandelbrot set of $K$. We consider the case $K = \mathbb{C}_p$, where $\mathbb{C}_p$ is the smallest complete, algebraically closed extension of $\mathbb{Q}_p$. We use two principal properties of $\mathbb{C}_p$; see [12] for details. First, there is a natural (non-archimedean) absolute value $|\cdot|$ on $\mathbb{C}_p$ that extends the $p$-adic absolute value on $\mathbb{Q}_p$. Second, let $\mathcal{O}_p = \{c \in \mathbb{C}_p : |c| \leq 1\}$ and $m_p = \{c \in \mathbb{C}_p : |c| < 1\}$, and note that $\mathcal{O}_p$ is a subring of $\mathbb{C}_p$ and $m_p$ its unique maximal ideal. Moreover, the quotient $\mathcal{O}_p/m_p$ is isomorphic to $\overline{\mathbb{F}}_p$, the algebraic closure of the finite field with $p$ elements. Denote the natural quotient homomorphism $\mathcal{O}_p \to \overline{\mathbb{F}}_p$ by $\phi$. We call $\phi$ the *reduction homomorphism*.

**Proposition 2.1.** *Let $K$ be a field and $|\cdot|$ a non-archimedean absolute value on $K$. Then $M(K) = \{c \in K : |c| \leq 1\}$. In particular, $M(\mathbb{C}_p) = \mathcal{O}_p$ for all primes $p$.*

*Proof.* Let $f_c = x^2+c$, and suppose $|c| > 1$. A consequence of the strong triangle inequality is that if $|x| \neq |y|$, then $|x+y| = \max\{|x|, |y|\}$. Using this, one easily shows by induction that $|f_c^{\circ n}(0)| = |c|^{2^{n-1}}$, whence $c \notin M(K)$. On the other hand, if $|c| \leq 1$, it follows immediately from the strong triangle inequality that $|f_c^{\circ n}(0)| \leq 1$ for all $n$, showing $c \in M(K)$. ☐

When $p = 2$ and $K = \mathbb{C}_p$, all cycles of $f_c$ contained in $\mathcal{O}_p$ are attracting. Since $f_c$ has good reduction, it follows that cycles of $f_{\phi(c)}$ lift to cycles of $f_c$ (contained in $\mathcal{O}_p$) and also that if $b \equiv a \bmod m_p$ and $a$ is in an attracting cycle, then $b$ tends to this cycle [11, Proposition 4.32]. The orbit of 0 under $f_{\phi(c)}$ is finite and thus is periodic after a certain point, implying that 0 tends to an attracting cycle in $\mathbb{C}_p$. Hence $\mathcal{H}(\mathbb{C}_p) = \mathcal{O}_p$, and clearly also $\mathcal{H}(\overline{\mathbb{F}}_p) = \overline{\mathbb{F}}_p$. *For the remainder of this article we assume that $p \neq 2$.*

We wish to give a characterization of $\mathcal{H}(\mathbb{C}_p)$ via the reduction homomorphism. First we make a few remarks on $\mathcal{H}(K)$ when $K = \overline{\mathbb{F}}_p$. Since $\overline{\mathbb{F}}_p^*$ consists of roots of unity, the only absolute value $K$ admits is the trivial one: $|c| = 1$ for all $c \in K^*$. Under the trivial absolute value, $c \in K$ tends to a formally attracting cycle if and only if $c$ is in fact contained in a formally attracting cycle. For general $K$, we easily derive from the chain rule that $c_1, \ldots, c_n$ is a formally attracting cycle of $R \in K(x)$ if and only if

$$\prod_{i=1}^{n} |R'(c_i)| < 1. \tag{6}$$

In the case $K = \overline{\mathbb{F}}_p$, it follows that a cycle is formally attracting if and only if it contains a critical point. These observations show that

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

We now give the promised characterization of $\mathcal{H}(\mathbb{C}_p)$. This is a consequence of [11, Proposition 4.32], but in our case a more direct argument suffices:

**Proposition 2.2.** *Let $\phi : \mathcal{O}_p \to \overline{\mathbb{F}}_p$ be the reduction homomorphism. Then $\mathcal{H}(\mathbb{C}_p) = \phi^{-1}(\mathcal{H}(\overline{\mathbb{F}}_p))$.*

*Proof.* Suppose first that $c \in \mathcal{H}(\mathbb{C}_p)$, so that 0 tends to the formally attracting cycle $c_1, \ldots, c_n$ under iteration of $f_c = x^2 + c$. Since $p \neq 2$, it follows from (6) that we must have $|c_i| < 1$ for some $i$, whence $\phi(c_i) = 0$. Since the forward orbit of 0 has points arbitrarily close to $c_i$ in $\mathbb{C}_p$, we have that 0 is periodic under iteration of $x^2 + \phi(c)$, whence $\phi(c) \in \mathcal{H}(\overline{\mathbb{F}}_p)$.

Now suppose 0 is periodic of period $n$ under $x^2 + \phi(c)$. Then $f_{\phi(c)}^{\circ n}$ fixes 0, whence $|f_c^{\circ n}(0)| \leq 1$. Note that $f_c^{\circ n}$ is a polynomial in $x^2$ with coefficients in $\mathcal{O}_p$, and since $|f_c^{\circ n}(0)| \leq$

1, it follows by induction that $\left|f_c^{\circ in}(0)\right| \le |f_c^{\circ n}(0)|$ for all $i \ge 1$. Now since $f_c^{\circ n}$ is a polynomial in $x^2$, we have $(f_c^{\circ n}(x) - x)'(0) = -1$, so we can apply Hensel's Lemma to obtain a fixed point $d$ of $f_c^{\circ n}$ with $|d| < 1$. Consider $g(x, y) = (f_c^{\circ n}(x) - f_c^{\circ n}(y))/(x - y)$, which is a polynomial with coefficients in $\mathcal{O}_p$ and without a constant term, since $g$ is divisible by $x + y$. Thus for $|a|, |b| < 1$ we have $|g(a, b)| \le \max\{|a|, |b|\}$. Taking $m \ge 1$, $a = f_c^{\circ(m-1)n}(0)$, and $b = f_c^{\circ(m-1)n}(d) = d$, we have

$$|f_c^{\circ mn}(0) - d| \le |f_c^{\circ(m-1)n}(0) - d| \cdot \max\{|f_c^{\circ(m-1)n}(0)|, |d|\}.$$

Repeating this $m-2$ times, it follows that $|f_c^{\circ mn}(0) - d| < |d| \cdot \prod_{i=1}^{m-1} \max\{|f_c^{\circ in}(0)|, |d|\}$. By the remark at the beginning of this paragraph, the right-hand side of this expression is at most $|d| \cdot (\max\{|f_c^{\circ n}(0)|, |d|\})^{m-1}$. Thus $f_c^{\circ mn}(0)$ tends to $d$ as $m$ grows. Now $d$ is a fixed point of $f_c^{\circ n}$ and thus belongs to a cycle of $f_c$; moreover this cycle is attracting since $f_c^{\circ n}$ being a polynomial in $x^2$ implies $|(f_c^{\circ n})'(d)| \le |d| < 1$. Given $\epsilon > 0$, the continuity of $f_c$ allows us to choose $\delta > 0$ such that $|x - d| < \delta$ implies $\left|f_c^{\circ i}(x) - f_c^{\circ i}(d)\right| < \epsilon$ for all $i$ with $1 \le i \le n$. Thus for $m$ large enough, $\left|f_c^{\circ mn+i}(0) - f_c^{\circ i}(d)\right| < \epsilon$ for all $i$ with $1 \le i \le n$. Therefore the orbit of 0 under $f_c$ converges to the attracting cycle containing $d$, proving that $c \in \mathcal{H}(\mathbb{C}_p)$. □

We discussed on page 3 the decomposition of $\mathcal{H}(K)$ into a disjoint union of open components $\mathcal{H}(K)^{(i)}$ corresponding to $c$ where 0 tends to a formally attracting $i$-cycle. In the case $K = \mathbb{C}_p$, these components are unions of open disks with radius 1. For instance, $f_c$ has a formally attracting fixed point if and only if $f_c(x) - x = x^2 - x + c$ has a root in $m_p$. A quick exercise in Newton polygons shows that this happens if and only if $c \in m_p$, i.e., $\phi(c) = 0$ where $\phi$ is the reduction homomorphism. Thus $\mathcal{H}(\mathbb{C}_p)^{(1)} = \phi^{-1}(0)$. A similar analysis shows that $f_c$ has a formally attracting two-cycle if and only if $\phi(c) = -1$.

## 3 Applying the Tchebotarev Density Theorem

In order to prove Theorem 1.4, our overall strategy is to give an upper bound for $D(\mathcal{H}(\overline{\mathbb{F}}_p))$ and show this upper bound is zero. In this section we prove Theorem 1.5, which uses the Tchebotarev Density theorem for function fields to give a practical method for computing the upper bound.

Let $f_c = x^2 + c$, and note that for $c \in \overline{\mathbb{F}}_p$, the forward orbit $\{f_c^{\circ n}(0) : n = 1, 2, \ldots\}$ of 0 is contained in $\mathbb{F}_p(c)$. Clearly 0 is periodic if and only if its backward orbit has points in common with its forward orbit. We thus let $f_c^{-\circ n}(0) = \{b \in \overline{\mathbb{F}}_p : f_c^{\circ n}(b) = 0\}$, and consider the sets

$$\mathcal{I}_n = \{c \in \overline{\mathbb{F}}_p : f_c^{-\circ n}(0) \cap \mathbb{F}_p(c) \ne \emptyset\}.$$

as defined in (5). These sets are useful because they furnish successively better "approximations" of $\mathcal{H}(\overline{\mathbb{F}}_p)$, as we now show:

**Proposition 3.1.** *For each $n \ge 1$, we have $\mathcal{I}_n \supseteq \mathcal{I}_{n+1}$. Moreover, $\mathcal{H}(\overline{\mathbb{F}}_p) = \bigcap_{n \ge 1} \mathcal{I}_n$.*

*Proof.* Let $c \in \mathcal{I}_{n+1}$, and take $b \in \mathbb{F}_p(c)$ such that $f_c^{\circ n+1}(b) = 0$. Then $f_c^{\circ n}(f_c(b)) = 0$ and $f_c(b) \in \mathbb{F}_p(c)$, whence $c \in \mathcal{I}_n$. To show the second statement, let $c \in \mathcal{H}(\overline{\mathbb{F}}_p)$, let $f_c^{\circ m}(0) = 0$, and take $n \geq 1$. Write $n = im - j$ for some $0 < j \leq m$, and note that

$$f_c^{\circ n}(f_c^{\circ j}(0)) = f_c^{\circ im}(0) = 0.$$

Clearly $f_c^{\circ j}(0) \in \mathbb{F}_p(c)$, showing that $f_c^{-\circ n}(0) \cap \mathbb{F}_p(c) \neq \emptyset$. Since $n$ was arbitrary, this shows $c \in \bigcap_{n \geq 1} \mathcal{I}_n$. Now suppose $c \in \bigcap_{n \geq 1} \mathcal{I}_n$, and for each $n$, let $b_n \in f_c^{-\circ n}(0) \cap \mathbb{F}_p(c)$. The finiteness of $\mathbb{F}_p(c)$ implies we must have $b_{n_1} = b_{n_2}$ for some $n_1 < n_2$. Therefore

$$f_c^{\circ n_2 - n_1}(0) = f_c^{\circ n_2 - n_1}(f_c^{\circ n_1}(b_{n_1})) = f_c^{\circ n_2}(b_{n_1}) = f_c^{\circ n_2}(b_{n_2}) = 0.$$

Hence $c \in \mathcal{H}(\overline{\mathbb{F}}_p)$. □

Recall the definition of the density of $\mathcal{C} \subseteq \overline{\mathbb{F}}_p$ given in (4). The following proposition gives a method for showing $D(\mathcal{H}(\overline{\mathbb{F}}_p)) = 0$ using only information about $D(\mathcal{I}_n)$.

**Proposition 3.2.** *Suppose that $D(\mathcal{I}_n)$ exists for all $n$ and $\lim_{n \to \infty} D(\mathcal{I}_n) = 0$. Then $D(\mathcal{H}(\overline{\mathbb{F}}_p))$ exists and equals zero.*

*Proof.* Given $\mathcal{C} \subseteq \overline{\mathbb{F}}_p$, define

$$a_{\mathcal{C}}(s) = \frac{\sum_{\alpha \in \mathcal{C}} (\deg \alpha)^{-1} N(\alpha)^{-s}}{\sum_{\alpha \in \overline{\mathbb{F}}_p} (\deg \alpha)^{-1} N(\alpha)^{-s}}.$$

Since $\mathcal{H}(\overline{\mathbb{F}}_p) \subseteq \mathcal{I}_n$ for all $n$, we have $a_{\mathcal{H}(\overline{\mathbb{F}}_p)}(s) \leq a_{\mathcal{I}_n}(s)$ for $s > 1$. Taking lim sups and using the assumption that $D(\mathcal{I}_n)$ exists gives

$$\limsup_{s \to 1^+} \; a_{\mathcal{H}(\overline{\mathbb{F}}_p)}(s) \leq \limsup_{s \to 1^+} \; a_{\mathcal{I}_n}(s) = \lim_{s \to 1^+} a_{\mathcal{I}_n}(s) = D(\mathcal{I}_n).$$

Since $\lim_{i \to \infty} D(\mathcal{I}_n) = 0$ and $a_{\mathcal{H}(\overline{\mathbb{F}}_p)}(s) \geq 0$ for $s > 1$, it follows that $\limsup_{s \to 1^+} \; a_{\mathcal{H}(\overline{\mathbb{F}}_p)}(s) = 0$. Therefore $\lim_{s \to 1^+} \; a_{\mathcal{H}(\overline{\mathbb{F}}_p)}(s) = 0$, proving that $D(\mathcal{H}(\overline{\mathbb{F}}_p)) = 0$. □

We now wish to use the Tchebotarev Density Theorem for function fields to prove Theorem 1.5, which shows $D(\mathcal{I}_n)$ exists and gives a method for computing it. To do this, we must relate $D(\mathcal{I}_n)$ to the density of a set of primes in $\mathbb{F}_p[t]$.

Let $P$ be the collection of primes in $\mathbb{F}_p[t]$. By the density of a set $S$ of primes of $\mathbb{F}_p[t]$, we mean

$$D(S) = \lim_{s \to 1^+} \frac{\sum_{\mathfrak{p} \in S} \; N\mathfrak{p}^{-s}}{\sum_{\mathfrak{p} \in P} \; N\mathfrak{p}^{-s}}, \tag{7}$$

where $N\mathfrak{p} = p^{\deg \mathfrak{p}}$. Recall from (5) the definition of $\mathcal{I}_n$, and note that $f_c^{-\circ n}(0) \cap \mathbb{F}_p(c) \neq \emptyset$ is equivalent to the factorization of $f_c^{\circ n}(x)$ over $\mathbb{F}_p(c)$ having a linear factor. This in turn is equivalent to $f_t^{\circ n}(x)$ having a linear factor modulo $(\pi_c)$, where $\pi_c \in \mathbb{F}_p[t]$ is the minimal polynomial of $c$. Hence

$$\mathcal{I}_n = \{c \in \overline{\mathbb{F}}_p : f_t^{\circ n} \text{ has a linear factor mod } (\pi_c)\}. \tag{8}$$

Since membership in $\mathcal{I}_n$ depends only on properties of $\pi_c$, it follows that $\mathcal{I}_n$ is invariant under the action of $\mathrm{Gal}\,(\overline{\mathbb{F}}_p/\mathbb{F}_p)$. The following proposition relates the density of Galois-invariant subsets of $\overline{\mathbb{F}}_p$ to the density of related sets of primes in $\mathbb{F}_p[t]$.

**Proposition 3.3.** *Suppose that $\mathcal{C} \subseteq \overline{\mathbb{F}}_p$ is invariant under the action of $\mathrm{Gal}\,(\overline{\mathbb{F}}_p/\mathbb{F}_p)$, and let $B \subseteq P$ be given by $\{(\pi_c) : c \in \mathcal{C}\}$, where $\pi_c$ is the minimal polynomial of $c$. Suppose also that $D(B)$ exists. Then $D(\mathcal{C})$ exists and equals $D(B)$.*

*Proof.* Consider the map $\psi : \overline{\mathbb{F}}_p \to P$ that takes $c$ to $(\pi_c)$. The Galois invariance of $\mathcal{C}$ is equivalent to $\mathcal{C}$ being the full inverse image of $B$ under $\psi$. We thus have

$$\sum_{c \in \mathcal{C}} (\deg c)^{-1} N(c)^{-s} = \sum_{\mathfrak{p} \in B} \sum_{c \in \psi^{-1}(\mathfrak{p})} (\deg c)^{-1} N(c)^{-s}, \tag{9}$$

where we recall $N(c) = p^{\deg c}$. Now for any $c \in \psi^{-1}(\mathfrak{p})$ we have $\deg c = \deg \pi_c = \deg \mathfrak{p}$. Thus $N(c) = N\mathfrak{p}$. Hence the inner sum in the right-hand side of (9) is repeated addition of the same quantity, and the right-hand side becomes $\sum_{\mathfrak{p} \in B} N\mathfrak{p}^{-s}$. Applying the same reasoning to $\sum_{c \in \overline{\mathbb{F}}_p} (\deg c)^{-1} N(c)^{-s}$ gives

$$\frac{\sum_{c \in \mathcal{C}} (\deg c)^{-1} N(c)^{-s}}{\sum_{c \in \overline{\mathbb{F}}_p} (\deg c)^{-1} N(c)^{-s}} = \frac{\sum_{\mathfrak{p} \in B} N\mathfrak{p}^{-s}}{\sum_{\mathfrak{p} \in P} N\mathfrak{p}^{-s}}.$$

Taking limits as $s \to 1^+$ and using the existence of $D(B)$ completes the proof.

□

We now define the following set of primes in $\mathbb{F}_p[t]$:

$$I_n = \{\mathfrak{p} \in P : f_t^{\circ n} \bmod \mathfrak{p} \text{ has at least one linear factor}\}. \tag{10}$$

The following corollary follows immediately from Proposition 3.3 and (8).

**Corollary 3.4.** *For all $n \geq 1$, $D(\mathcal{I}_n) = D(I_n)$.*

We now give a version of a standard result that allows us to apply the Tchebotarev Density theorem to compute $D(I_n)$.

**Proposition 3.5.** *Let $R = \mathbb{F}_p[t], L = \mathbb{F}_p(t)$, and $f \in L[x]$. Let $L_n(f)$ be the splitting field of $f^{\circ n}$ over $L$, and let $G_n(f) = \mathrm{Gal}\,(L_n(f)/L)$. There exists a finite set $S$ of primes in $R$ (including the ramified primes) such that if $\mathfrak{p}$ is not in $S$ and $(\mathfrak{p}, L_n(f)/L) \subset G_n(f)$ is the Artin conjugacy class of $\mathfrak{p}$, then the following holds. If $f_1 f_2 \cdots f_r$ is the factorization into irreducibles of $f^{\circ n} \bmod \mathfrak{p}$, then any element of $(\mathfrak{p}, L_n(f)/L)$ acts on the roots of $f^{\circ n}$ as a product $\sigma_1 \cdots \sigma_r$ of disjoint cycles, with $\sigma_i$ having length $\deg f_i$.*

*Proof.* Let $\beta$ be a root of $f^{\circ n}$, set $L_\beta = L(\beta)$, and let $R_\beta$ be the integral closure of $R$ in $L_\beta$. A standard result in algebraic number theory [7, Theorem 4.12, p. 177] states that for all primes $\mathfrak{p}$ not contained in a finite set $S'$, we have

$$\mathfrak{p}R_\beta = \mathfrak{P}_1\mathfrak{P}_2\cdots\mathfrak{P}_r, \tag{11}$$

where $\mathfrak{P}_i$ is a prime in $R_\beta$ with residue class degree $d(\mathfrak{P}_i/\mathfrak{p}) = \deg f_i$. Another standard result (see [7, Lemma 7.13, p. 391] for a proof easily adapted to the function field case) states that if $\mathfrak{p}$ is in addition unramified, then (11) implies that any element of $(\mathfrak{p}, L_n(f)/L)$ acts on the roots of $f^{\circ n}$ as a product $\sigma_1\cdots\sigma_r$ of disjoint cycles, with $\sigma_i$ having length $d(\mathfrak{P}_i/\mathfrak{p})$. $\square$

Using the notation of Proposition 3.5, let $f = f_t = x^2 + t$ and let $U_n$ be the set of primes in $R$ that are unramified in $L_n(f_t)$. Note that conjugacy preserves the lengths of the cycles in the disjoint cycle decomposition, so if one element of $(\mathfrak{p}, L_n(f_t)/L)$ fixes a root of $f_t^{\circ n}$, then all do. Put

$$J_n = \{\mathfrak{p} \in U_n : \text{each } \sigma \in (\mathfrak{p}, L_n(f_t)/L) \text{ fixes at least one root of } f_t^{\circ n}\}. \tag{12}$$

It follows immediately from Proposition 3.5 and (10) that $D(J_n) = D(I_n)$. To compute $D(J_n)$, we use the Tchebotarev Density theorem, which we now state.

**Theorem 3.6** ((Tchebotarev))**.** *Let $L/K$ be a Galois extension of function fields, and denote by $U_K$ the set of primes of $K$ unramified in $L$. For $\mathfrak{p} \in U_K$, let $(\mathfrak{p}, L/K)$ be the Artin conjugacy class of $\mathfrak{p}$. Fix a conjugacy class $C$ of $G = \mathrm{Gal}\,(L/K)$. Then for all $k \geq 1$,*

$$D(\{\mathfrak{p} \in U_K : (\mathfrak{p}, L/K) = C\}) = \frac{\#C}{\#G}.$$

For a proof, see [13, Chapter 9].

*Proof of Theorem 1.5.* Recall the definition of $J_n$ from (12). Using Proposition 3.2, Proposition 3.3, and $D(I_n) = D(J_n)$, it suffices to find $D(J_n)$. Let $\mathfrak{C}$ denote the collection of conjugacy classes of $G_n(f)$ each of whose elements fixes at least one root of $f_t^{\circ n}$. Using Theorem 3.6, we have

$$D(J_n) = \sum_{C\in\mathfrak{C}} \frac{\#C}{\#G_n(f)} = \frac{1}{\#G_n(f)} \sum_{C\in\mathfrak{C}} \#C = \frac{1}{\#G_n(f)} \# \left\{ \bigcup_{C\in\mathfrak{C}} C \right\}.$$

Since $g$ fixes a root of $f_t^{\circ n}$ if and only if every element of its conjugacy class does the same, $\bigcup_{C\in\mathfrak{C}} C = \{g \in G_n(f) : g \text{ fixes at least one root of } f_t^{\circ n}\}$. $\square$

*Example* 3.7. Consider the case $n = 2$. Label the roots of $f_t^{\circ 2} = (x^2+t)^2 + t$ as follows:

$$\sqrt{-t+\sqrt{-t}} \longleftrightarrow 1 \qquad -\sqrt{-t+\sqrt{-t}} \longleftrightarrow 2$$
$$\sqrt{-t-\sqrt{-t}} \longleftrightarrow 3 \qquad -\sqrt{-t-\sqrt{-t}} \longleftrightarrow 4$$

We show (Corollary 6.6) that under this labeling $G_2(f)$ is a subgroup of $S_4$ of order 8 that contains $\{e, (1\ \ 2), (3\ \ 4), (1\ \ 2)(3\ \ 4)\}$ as well as four elements that interchange the sets $\{1,2\}$ and $\{3,4\}$ and therefore have no fixed points. Hence $D(\mathcal{I}_2) = 3/8$. In fact, we know more. Let $k$ be large, choose $c \in \mathbb{F}_{p^k}$ at random, and let $i_c = \#\{f_c^{-\circ 2}(0) \cap \mathbb{F}_p(c)\}$. Then $i_c = 2$ with probability $1/4$ and $i_c = 4$ with probability $1/8$.

*Remark.* There is a second version of Theorem 3.6 that gives the stronger conclusion $\#\{\mathfrak{p} \in U_K : \deg \mathfrak{p} = k \text{ and } (\mathfrak{p}, L/K) = C\} = \frac{p^k}{k}(\#C/\#G + O(p^{-k/2}))$ [13, Theorem 9.13B]. With this conclusion, one can replace $D(\mathcal{I}_n)$ by $\delta(\mathcal{I}_n)$ (see p. 3) in Theorem 1.5, and thus also in Theorem 1.4. However, this stronger version of Theorem 3.6 requires the hypothesis that $L/K$ be *geometric*, i.e. that if $k$ is the field of constants in $K$, then $\overline{k} \cap L = k$. Determining the geometricity of the fields generated by roots of $f_t^{\circ n}$ appears to be a difficult problem. Indeed, the most natural approach may be to first prove Conjecture 6.7 (see the discussion following the conjecture).

# 4 Galois processes and Galois martingales

Let $L$ be a field of characteristic $\neq 2$, and for $f \in L[x]$, let $L(f)$ denote the splitting of $f$ over $L$. By a *tower of polynomials over* $L$ we mean a sequence $f_1, f_2, \ldots$ such that $L(f_n) \supseteq L(f_{n-1})$ for $n \geq 2$. We call the tower separable if $f_n$ is separable over $L$ for all $n \geq 1$.

Let $\mathcal{F} = f_1, f_2, \ldots$ be a separable tower of polynomials, and put $L_\infty = \bigcup_{n=1}^{\infty} L(f_n)$. Denote $\operatorname{Gal} L(f_n)/L$ by $G(f_n)$, and let

$$G(\mathcal{F}) = \operatorname{Gal} L_\infty/L \cong \lim_{\leftarrow} G(f_n).$$

Let $\mathbf{P}$ be the Haar measure on the compact group $G(\mathcal{F})$, normalized so that $\mathbf{P}(G(\mathcal{F})) = 1$. Letting $\mathcal{B}$ be the Borel sigma algebra, the triple $(G(\mathcal{F}), \mathbf{P}, \mathcal{B})$ is then a probability space. Denote by $\psi_n$ the natural projection $G(\mathcal{F}) \to G(f_n)$, and define random variables $X_n$ on $G(\mathcal{F})$ as follows:

$$X_n(g) = \text{number of roots of } f_n \text{ fixed by } \psi_n(g).$$

The data $(G(\mathcal{F}), \mathbf{P}, \mathcal{B}, \{X_n\}_{n\geq 1})$ by definition give a stochastic process, which we call the *Galois process* of the tower $\mathcal{F}$, and denote $GP(\mathcal{F})$. Intuitively, this process resembles a random walk through successively higher levels of the group $G(\mathcal{F})$. Positions at each level are assigned a value based on the number of roots of $f_n$ left fixed. Note that it follows from basic properties of Haar measure that

$$P(X_1 = t_1, \ldots, X_n = t_n) = \frac{1}{\#G(f_n)} \# \left\{ g \in G(f_n) : g \text{ fixes } t_i \text{ roots of } f_i \text{ for } i = 1, 2, \ldots, n \right\}. \tag{13}$$

*Example* 4.1. Let $L = \mathbb{Q}$, $f_1 = x^2 + x + 1$, and $f_2 = x^3 - 2$. Clearly $L(f_2) \supseteq L(f_1)$. Define a separable tower $\mathcal{F}$ by setting $f_n = f_2$ for all $n \geq 3$. Then $G(\mathcal{F}) \cong G(f_2)$, which is the full symmetric group on the roots of $f_2$. Since the Haar measure $\mathbf{P}$ is invariant under multiplication by an element of $G(\mathcal{F})$, it follows that $\mathbf{P}(\psi_n^{-1}(g)) = 1/\#G(f_n)$ for all $g \in G(f_n)$. Thus

$$\mathbf{P}(X_2 = i) = \begin{cases} 1/6 & \text{if } i = 3 \\ 1/2 & \text{if } i = 1 \\ 1/3 & \text{if } i = 0 \end{cases}$$

One easily sees that the kernel of the quotient map $G(f_2) \to G(f_1)$ is isomorphic to the alternating group $A_3$. Thus if $g \in A_3$, $X_1(g) = 2$, while if $g \in G(f_2) \setminus A_3$, then $X_1(g) = 0$. Because $G(f_2) \setminus A_3$ is composed entirely of transpositions, we have $\mathbf{P}(X_2 = 1 \mid X_1 = 0) = 1$. Therefore $GP(\mathcal{F})$ is not a martingale (see Definition 1.1).

Consider the case where $f \in L(x)$ has the property that all iterates $f^{\circ n}$ are separable over $L$ and $\mathcal{F} = f, f^{\circ 2}, f^{\circ 3}, \ldots$. This is the case of greatest interest to us. In this situation we write $GP(f)$ instead of $GP(\mathcal{F})$ and $G_n(f)$ instead of $G(f^{\circ n})$. We now develop preparatory material for proving Theorem 1.2. Recall that a $G$-set is a set $T$ on which $G$ acts, and a map $\phi : T \to T'$ is a morphism of $G$-sets if $\phi(\sigma(t)) = \sigma(\phi(t))$ for all $\sigma \in G$ and $t \in T$. We define a notion we use throughout:

**Definition 4.2.** Let $G$ be a group and $T$ a $G$-set. By a *fiber system* [1] on $T$, we mean the collection of fibers of a surjective morphism $\phi : T \to T'$ of $G$-sets.

Note that a fiber system on $T$ gives a partition of $T$, and the sets belonging to this partition are permuted by $G$; indeed, these properties characterize fiber systems. By way of illustration, we offer the following:

**Proposition-Definition 4.3.** Let $L$ be a field and $f \in L[x]$ a polynomial with all iterates separable. Let $R_n$ denote the roots of $f^{\circ n}$ and $R_{n-1}$ the roots of $f^{\circ n-1}$. Then $f : R_n \to R_{n-1}$ is a surjective morphism of $G_n(f)$-sets. It defines a fiber system on $R_n$ that we call the *fundamental fiber system*.

*Proof.* We need only check that $f(\sigma(\beta)) = \sigma(f(\beta))$ for any $\sigma \in G_n(f)$ and $\beta \in R_n$. This clearly holds since $\sigma \in \mathrm{Gal}\,(L(f^{\circ n})/L)$. □

For instance, let $L = k(t)$, $f(x) = x^2 + t$, and $n = 2$, and use the labelings of Example 3.7. Then the fundamental fiber system on $R_2$ is $\{\{1, 2\}, \{3, 4\}\}$. Note that the fundamental fiber system consists of sets each containing $\deg f$ elements.

The proof of Theorem 1.2 makes crucial use of a uniqueness result on fiber systems for $G$-sets when $G$ is a certain kind of permutation group. Specifically, let $f$ be quadratic with separable and irreducible iterates, and suppose $G_n(f)$ contains at least one odd permutation. We wish to show that the fundamental fiber system is the only fiber system on $R_n$ (considered as a $G_n(f)$-set) that consists of two-element sets. The next few definitions and lemmas build up to this result (Corollary 4.9).

Let $T$ be a set and $\mathfrak{S}$ a partition of $T$. Denote by $\mathrm{Perm}(T, \mathfrak{S})$ the set of all permutations of $T$ that act as permutations on $\mathfrak{S}$. Thus if a group $G$ acts on $T$ and $\mathfrak{S}$ is a fiber system for $G$ then $G \subseteq \mathrm{Perm}(T, \mathfrak{S})$.

[1] Some authors use the terminology "block system."

**Definition 4.4.** Let $T$ be a set and $\mathfrak{S}$ a partition of $T$. A *permutation associated to* $\mathfrak{S}$ is a permutation $\sigma \in \mathrm{Perm}(T, \mathfrak{S})$ whose orbits are precisely the subsets belonging to $\mathfrak{S}$. In the case where $\mathfrak{S}$ is composed entirely of two-element sets, we denote by $\sigma_{\mathfrak{S}}$ the unique permutation associated to $\mathfrak{S}$.

**Proposition 4.5.** *Let $T$ be a set, $\mathfrak{S}$ a partition of $T$, and $\sigma$ a permutation associated to $\mathfrak{S}$. Then $\tau\sigma\tau^{-1}$ is a permutation associated to $\mathfrak{S}$ for each $\tau \in \mathrm{Perm}(T, \mathfrak{S})$. In particular, if $\mathfrak{S}$ is composed of two-element subsets, then*

$$\sigma_{\mathfrak{S}} \in Z(\mathrm{Perm}(T, \mathfrak{S})),$$

*the center of* $\mathrm{Perm}(T, \mathfrak{S})$.

*Proof.* Let $\tau \in \mathrm{Perm}(T, \mathfrak{S})$, and $S \in \mathfrak{S}$ with $\#S = k$. Then $\tau\sigma\tau^{-1}$ must map $S$ to itself. Note that $\sigma$ acts on $S$ as a $k$-cycle since by definition $S$ is an orbit of $\sigma$. Since conjugation preserves cycle decomposition type and both $\sigma$ and $\tau\sigma\tau^{-1}$ permute $S$, it follows that $\tau\sigma\tau^{-1}$ acts on $S$ as a $k$-cycle. Therefore $S$ is an orbit of $\tau\sigma\tau^{-1}$. □

**Lemma 4.6.** *Let $G$ be a group acting on a set $T$. Suppose that $\mathfrak{S} = \{S_1, S_2, \ldots, S_m\}$ is a fiber system for $G$ with $\#S_i = 2j$ for $1 \leq i \leq m$. Then $G$ is isomorphic to a subgroup of the wreath product $\mathrm{Sym}(2j) \wr \mathrm{Sym}(\mathfrak{S}) = \mathrm{Sym}(2j)^{\mathfrak{S}} \rtimes \mathrm{Sym}(\mathfrak{S})$ and if $g \mapsto ((\delta_1, \ldots, \delta_m), \sigma)$ then the signature of the action of $g$ on $T$ is $\prod_{i=1}^{m} \mathrm{sgn}\, \delta_i$.*

*Proof.* Since $\mathfrak{S}$ is a fiber system for $G$, we have $G \subseteq \mathrm{Perm}(T, \mathfrak{S})$. It therefore suffices to prove the Lemma for $\mathrm{Perm}(T, \mathfrak{S})$. Each $\tau \in \mathrm{Perm}(T, \mathfrak{S})$ induces a permutation $\tau'$ on $\mathfrak{S}$. Fix an ordering of the elements in each $S_i$, and suppose that $\tau(S_i) = S_k$. Say $S_i = \{s_1, \ldots, s_{2j}\}$ and $S_k = \{t_1, \ldots, t_{2j}\}$. Then $t_n \mapsto \tau(s_n)$ is a permutation of $S_k$, which we denote by $\delta_i$. The map $\mathrm{Perm}(T, \mathfrak{S}) \to \mathrm{Sym}(2j) \wr \mathrm{Sym}(\mathfrak{S}) : \tau \mapsto ((\delta_1, \ldots, \delta_m), \tau')$ is readily seen to be an isomorphism.

Suppose $\tau \in \mathrm{Perm}(T, \mathfrak{S})$ satisfies $\delta_i = \mathrm{id}$ for $i = 1, \ldots, m$, and let $C = (S_{i_1} \ \cdots \ S_{i_l})$ be a cycle of $\tau'$ and $\Sigma = S_{i_1} \cup \cdots \cup S_{i_l}$. Consider the action of $\tau$ on $\Sigma$. Since $C$ is an $l$-cycle, the orbit of any $s \in \Sigma$ must have length at least $l$, but clearly $\tau^l(s) = s$ for all $s$. Thus $\tau$ acts on $\Sigma$ as a product of $2j$ $l$-cycles, whence this action is even. The same holds for all cycles of $\tau'$, implying that the signature of $\tau$ is 1.

If $\tau \in \mathrm{Perm}(T, \mathfrak{S})$ satisfies $\tau' = \mathrm{id}$, then $\tau$ has the same cycle structure as the product $\delta_1 \cdots \delta_m \in \mathrm{Sym}(2jm)$. Thus the signature of $\tau$ is $\prod_{i=1}^{m} \mathrm{sgn}\, \delta_i$. Since any $((\delta_1, \ldots, \delta_m), \sigma) \in \mathrm{Sym}(2j) \wr \mathrm{Sym}(\mathfrak{S})$ admits the decomposition $((\delta_1, \ldots, \delta_m), \mathrm{id}) \cdot ((\mathrm{id}, \ldots, \mathrm{id}), \sigma)$, the lemma is proved. □

**Theorem 4.7.** *Let $T$ be a set with $2m$ elements, and let*

$$\mathfrak{S} = \{S_1, \ldots, S_m\} \qquad \text{and} \qquad \mathfrak{U} = \{U_1, \ldots, U_m\}$$

*be partitions of $T$ with $\#S_i = \#U_i = 2$ for all $i$. Let $\sigma_{\mathfrak{S}}$ and $\sigma_{\mathfrak{U}}$ be the permutations associated to $\mathfrak{S}$ and $\mathfrak{U}$, respectively, and suppose that $\sigma_{\mathfrak{U}} \in \mathrm{Perm}(T, \mathfrak{S})$. If $\sigma_{\mathfrak{S}} \neq \sigma_{\mathfrak{U}}$, then any subgroup of $\mathrm{Perm}(T, \mathfrak{S}) \cap \mathrm{Perm}(T, \mathfrak{U})$ that acts transitively on $T$ is alternating, i.e., composed entirely of even permutations.*

*Proof.* Let $G \leq \mathrm{Perm}(T, \mathfrak{S}) \cap \mathrm{Perm}(T, \mathfrak{U})$ act transitively on $T$. From Proposition 4.5, we have that $G$ centralizes $H = \langle \sigma_{\mathfrak{U}}, \sigma_{\mathfrak{S}} \rangle$ in $\mathrm{Sym}(T)$ and also $\sigma_{\mathfrak{S}}$ commutes with $\sigma_{\mathfrak{U}}$ (the latter since $\sigma_{\mathfrak{U}} \in \mathrm{Perm}(T, \mathfrak{S})$). Hence $H$ has order 4, because $\sigma_{\mathfrak{S}} \neq \sigma_{\mathfrak{U}}$. Note that by definition $\sigma_{\mathfrak{U}}$ and $\sigma_{\mathfrak{S}}$ have no fixed points in $T$, so that if an orbit of the action of $H$ on $T$ has fewer than four elements then we must have $\sigma_{\mathfrak{U}}(t) = \sigma_{\mathfrak{S}}(t)$ for some $t \in T$. But $G$ centralizes $H$ and acts transitively on $T$, implying $\sigma_{\mathfrak{U}}(t) = \sigma_{\mathfrak{S}}(t)$ for all $t \in T$, which contradicts $\sigma_{\mathfrak{S}} \neq \sigma_{\mathfrak{U}}$.

Since $G$ centralizes $H$, it follows immediately that the set $V = \{V_i\}$ of orbits of the action of $H$ on $T$ is a fiber system for $G$. From Lemma 4.6 we have that $G$ injects into $\mathrm{Sym}(4) \wr \mathrm{Sym}(V)$. Let $g \in G$ and fix $\{t_i\}$ such that $V_i = \{t_i, \sigma_{\mathfrak{U}}(t_i), \sigma_{\mathfrak{S}}(t_i), \sigma_{\mathfrak{U}}\sigma_{\mathfrak{S}}(i)\}$ for each $i$. Suppose $g(V_i) = V_j$, and let $\delta_i$ be as in the proof of Lemma 4.6. Again since $G$ centralizes $H$, one easily verifies that $\delta_i$ is the identity if $g(t_i) = t_j$ and is the product of two transpositions otherwise. Hence the signature of $\delta_i$ is 1, and it follows from Lemma 4.6 that $\mathrm{sgn}\, g = 1$. □

If the group $G$ has nontrivial center, we have another source of nontrivial fiber systems:

**Proposition-Definition 4.8.** Let $G$ be a group acting on a set $T$, suppose that $\sigma \in Z(G)$, and let $\mathfrak{S}$ be the partition of $T$ given by the orbits of $\sigma$. Then $\mathfrak{S}$ is a fiber system for $T$ and $\sigma$ is a permutation associated to $\mathfrak{S}$. We call $\mathfrak{S}$ a *central fiber system.*

*Proof.* Let $S \in \mathfrak{S}$ and $\tau \in G$. Write $S = \{\sigma^n(s) : n \geq 1\}$, and note that $\tau(S) = \{\tau\sigma^n(s) : n \geq 1\} = \{\sigma^n(\tau(s)) : n \geq 1\}$, which is again an element of $\mathfrak{S}$. Thus $\mathfrak{S}$ is a $G$-set, and indeed is the fiber system associated to the natural morphism $T \to \mathfrak{S}$ of $G$-sets. Clearly $\sigma$ is a permutation associated to $\mathfrak{S}$. □

A salient feature of central fiber systems is that at least one associated permutation must lie in the group $G$. This is not necessarily the case for the fundamental fiber system defined in Proposition-Definition 4.3. This feature of central fiber systems is precisely what we need to establish our uniqueness result on fiber systems consisting of two-element sets. In the process, we show that the central involution $\sigma_{\mathfrak{C}}$ associated to the fundamental fiber system must lie in $G_n(f)$. This provides a vital step in the proof of Theorem 1.2.

**Corollary 4.9.** *Let $L$ be a field, $f \in L[x]$ a quadratic polynomial with $f^{\circ n}$ separable and irreducible over $L$, and suppose that* $\mathrm{Disc}\, f^{\circ n}$ *is not a square in $L$. Then there is a unique fiber system of two-element sets on $R_n$, the set of roots of $f^{\circ n}$ (considered as a $G_n(f)$-set). In particular, if $\mathfrak{C}$ is the fundamental fiber system defined in Proposition-Definition 4.3, then the permutation $\sigma_{\mathfrak{C}}$ associated to $\mathfrak{C}$ is contained in $G_n(f)$.*

*Proof.* The splitting field $L(f^{\circ n})$ of $f^{\circ n}$ is obtained from $L(f^{\circ n-1})$ by adjoining roots of $f(x) - \alpha$ for each root $\alpha$ of $f^{\circ n-1}$. Since $\deg f = 2$, it follows that $\mathrm{Gal}\,(L(f^{\circ n})/L(f^{\circ n-1}))$ is an elementary abelian 2-group. Hence $G_n(f)$ is a 2-group, and therefore has nontrivial center. Since $Z(G_n(f))$ is again a 2-group, there must be $\delta \in Z(G_n(f))$ of order two. Suppose that $\delta$ fixes $r \in R_n$. Since $\delta$ belongs to the center of $G_n(f)$, this gives $\delta\sigma(r) = \sigma(r)$ for all $\sigma \in G_n(f)$. The irreducibility of $f^{\circ n}$ implies that $G_n(f)$ acts transitively on $R_n$, whence $\delta$ is the identity, a contradiction. Therefore $\delta$ has no fixed points, implying that the associated central fiber system $\mathfrak{D}$ (see Proposition-Definition 4.8) consists of two-element sets.

Let $\mathfrak{S}$ be another fiber system for $G_n(f)$ consisting of two element sets. Note that $\delta \in G_n(f) \subseteq \mathrm{Perm}(T, \mathfrak{S})$. Finally, since $\mathrm{Disc}\, f^{\circ n}$ is not a square in $L$, basic Galois theory tells us $G_n(f)$ cannot be alternating. We then apply Theorem 4.7 to get $\mathfrak{S} = \mathfrak{D}$. In particular, $\mathfrak{C} = \mathfrak{D}$, implying that $\sigma_{\mathfrak{C}} = \delta \in G_n(f)$. □

Recall from the introduction that the adjusted forward orbit of a point $l$ under $f \in L[x]$ with leading coefficient $a$ is $\{-af(l)\} \cup \{f^{\circ n}(l) : n = 2, 3, \ldots\}$.

**Lemma 4.10.** *Let $L$ be a field of characteristic $\neq 2$, $f \in L[x]$ quadratic, and $\gamma \in L$ the unique finite critical point of $f$. Suppose that the adjusted forward orbit of $\gamma$ contains no squares in $L$. Then for all $n$, $f^{\circ n}$ is separable and irreducible and* $\mathrm{Disc}\, f^{\circ n}$ *is not a square in $L$.*

*Proof.* We first show that $\mathrm{Disc}\, f^{\circ n}$ is not a square; this implies $f^{\circ n}$ is separable. Let $f(x) = ax^2 + bx + c$. For $n = 1$ we note that $-4af(\gamma) = \mathrm{Disc}\, f$, so that $-af(\gamma)$ not a square implies $\mathrm{Disc}\, f$ not a square. For $n \geq 2$, it follows from [8, Lemma 3.1, part iv] that $\mathrm{Disc}\, f^{\circ n} = 2^{2^n}(\mathrm{Disc}\, f^{\circ n-1})^2 \mathrm{Res}(f', f^{\circ n})$, where $\mathrm{Res}(f', f^{\circ n})$ denotes the resultant of $f'$ and $f^{\circ n}$. From the definition of resultant (see [8, p. 393]), $\mathrm{Res}(f', f^{\circ n}) = f^{\circ n}(\gamma)$. Thus $\mathrm{Disc}\, f^{\circ n}$ is not a square.

To show that $f^{\circ n}$ is irreducible, first note that the case $n = 1$ is covered by the previous paragraph. For $n \geq 2$ we use Capelli's Lemma [8, p. 387], which implies that $f^{\circ n}$ is irreducible if and only if for any root $\alpha$ of $f^{\circ n-1}$, we have $f(x) - \alpha$ irreducible over $L(\alpha)$. This is equivalent to $b^2 - 4ac + 4a\alpha$ not being a square in $L(\alpha)$, which must hold if $N_{L(\alpha)/L}(b^2 - 4ac + 4a\alpha)$ is not a square in $L$. But

$$\begin{aligned} N_{L(\alpha)/L}(b^2 - 4ac + 4a\alpha) &= (-4a)^{2^{n-1}} \prod_{\alpha \text{ root of } f^{\circ n-1}} \left( -\frac{b^2}{4a} + c \right) - \alpha \\ &= (4a)^{2^{n-1}} f^{\circ n-1}(-b^2/4a + c) = (4a)^{2^{n-1}} f^{\circ n-1}(f(\gamma)). \end{aligned}$$

Thus $f^{\circ n}(\gamma)$ not a square in $L$ implies $f^{\circ n}$ is irreducible. □

*Proof of Theorem 1.2.* We must show that

$$E(X_n \mid X_1 = t_1, \ldots, X_{n-1} = t_{n-1}) = t_{n-1}, \tag{14}$$

where $t_1, \ldots, t_{n-1}$ are integers with $\mathbf{P}(X_1 = t_1, \ldots, X_{n-1} = t_{n-1}) > 0$. By definition, the left-hand side of (14) is

$$\sum_k k \cdot \frac{\mathbf{P}(X_1 = t_1, \ldots, X_{n-1} = t_{n-1}, X_n = k)}{\mathbf{P}(X_1 = t_1, \ldots, X_{n-1} = t_{n-1})}. \tag{15}$$

Put

$$
\begin{aligned}
S &= \{g \in G_n(f) : g \text{ fixes } t_i \text{ roots of } f^{\circ i} \text{ for } 1 \le i \le n-1\} \\
S_k &= \{g \in S : g \text{ fixes } k \text{ roots of } f^{\circ n}\}
\end{aligned}
$$

From the basic property of $\mathrm{GP}(f)$ given in (13), the expression in (15) is equal to

$$\sum_k k \cdot \frac{\#S_k}{\#S}$$

This in turn may be rewritten as

$$\frac{1}{\#S} \sum_{g \in S} (\text{number of roots of } f^{\circ n} \text{ fixed by } g)\,. \tag{16}$$

Note that if $h \in H_n(f) \stackrel{\text{def}}{=} \mathrm{Gal}\,(L(f^{\circ n})/L(f^{\circ n-1}))$, then $h$ fixes the roots of $f^{\circ i}$ for $1 \le i \le n-1$. Thus $S$ is invariant under multiplication by $H_n(f)$, whence $S$ is a union of cosets of $H_n(f)$. Recall from Proposition-Definition 4.3 the fundamental fiber system $\mathfrak{C}$ for $G_n(f)$ defined by the morphism $f : R_n \to R_{n-1}$ of $G_n(f)$-sets. Let $\sigma_{\mathfrak{C}}$ be the permutation associated to $\mathfrak{C}$. From Corollary 4.9 and Lemma 4.10 we have $\sigma_{\mathfrak{C}} \in G_n(f)$, and thus $\sigma_{\mathfrak{C}} \in H_n(f)$.

Now take $g_0 H_n(f) \subseteq S$. Note that the group $\{e, \sigma_{\mathfrak{C}}\}$ acts by right multiplication on the set $g_0 H_n(f)$, dividing it into two-element orbits. We analyze the number of roots of $f^{\circ n}$ fixed by the elements of such an orbit. Let $g \in g_0 H_n(f)$, let $\alpha$ be a root of $f^{\circ n-1}$, and note that if $g(\alpha) \neq \alpha$ then neither $g$ nor $g\sigma_{\mathfrak{C}}$ have any fixed points in $f^{-1}(\alpha)$. On the other hand, if $g(\alpha) = \alpha$ then $g(f^{-1}(\alpha)) = f^{-1}(\alpha)$. Since by definition $\sigma_{\mathfrak{C}}$ exchanges the elements of $f^{-1}(\alpha)$, we have that $g$ fixes the elements of $f^{-1}(\alpha)$ if and only if $g\sigma_{\mathfrak{C}}$ exchanges them. It follows that

$$\#\{\text{roots of } f^{\circ n} \text{ fixed by } g\} + \#\{\text{roots of } f^{\circ n} \text{ fixed by } g\sigma_{\mathfrak{C}}\} = 2 \cdot \#\{\text{roots of } f^{\circ n-1} \text{ fixed by } g\}.$$

By the definition of $S$, all $g \in g_0 H_n(f)$ fix $t_{n-1}$ elements of $f^{\circ n-1}$. Therefore we have

$$\sum_{g \in g_0 H_n(f)} \#\{\text{roots of } f^{\circ n} \text{ fixed by } g\} = t_{n-1} \cdot \#H_n(f).$$

Since $S$ is a union of cosets of $H_n(f)$, the expression in (16) equals $t_{n-1}$. □

Martingales are important chiefly because they often converge in the following sense:

**Definition 4.11.** Let $X_1, X_2, \ldots$ be a stochastic process defined on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The process *converges* if

$$\mathbf{P}\left(\omega \in \Omega : \lim_{n \to \infty} X_n(\omega) \text{ exists}\right) = 1.$$

We give one simple martingale convergence theorem (see e.g. [2, Section 12.3] for a proof).

**Theorem 4.12.** *Let $M = (X_1, X_2, \ldots)$ be a martingale whose random variables take non-negative real values. Then $M$ converges.*

Since the random variables in $GP(f)$ take nonnegative integer values, we immediately have the following:

**Corollary 4.13.** *Let $L$ be a field and $f \in L[x]$ a quadratic polynomial satisfying the hypotheses of Theorem 1.2. Then*

$$\mathbf{P}(\{g \in G(f) : X_1(g), X_2(g), \ldots \text{ is eventually constant}\}) = 1.$$

## 5 Quadratic Galois Processes Under Maximality Assumptions

In this section, let $L$ be a field and $f \in L[x]$ a quadratic polynomial with all iterates separable over $L$. Throughout the section, all quantities are assumed chosen so that conditional probabilities are defined.

Let $L(f^{\circ n})$ be the splitting field of the $n$th iterate of $f$, and let $H_n(f) = \mathrm{Gal}\,(L(f^{\circ n})/L(f^{\circ n-1}))$.

**Proposition-Definition 5.1.** For each $n \geq 1$, $H_n(f) \cong (\mathbb{Z}/2\mathbb{Z})^m$ for some $0 \leq m \leq 2^{n-1}$. We call $H_n(f)$ *maximal* if $m = 2^{n-1}$.

*Proof.* Let $R_{n-1}$ denote the set of roots of $f^{\circ n-1}$. Over $L(f^{\circ n-1})$, $f^{\circ n}$ factors as $\prod_{\alpha \in R_{n-1}} f(x) - \alpha$. Since $\#R_{n-1} = 2^{n-1}$, the extension $L(f^{\circ n})/L(f^{\circ n-1})$ is the compositum of at most $2^{n-1}$ quadratic extensions. □

The next result gives, for $n$ with $H_n(f)$ maximal, an explicit expression of the probability distribution of $X_n$ given past behavior. However, the Lemma does not hold for all possible past behaviors: we must assume that the value of $X_{n-1}$ is known.

**Lemma 5.2.** *Let $L$ be a field, $f \in L[x]$ a quadratic polynomial with all iterates separable over $L$, $H_n(f) = \mathrm{Gal}\,(L(f^{\circ n})/L(f^{\circ n-1}))$, and $(\Omega, \mathcal{F}, \mathbf{P}, (X_n)_{n \geq 0}) = \mathrm{GP}(f)$. Suppose that $H_n(f)$ is maximal, and let $m_1 < m_2 < \cdots < m_k$, be positive integers with $m_k = n - 1$. Then for any positive integers $t_1, \ldots, t_k$ we have*

$$\mathbf{P}(X_n = u \mid X_{m_1} = t_1, \ldots, X_{m_k} = t_k) = \begin{cases} \binom{t_k}{w} \frac{1}{2^{t_k}} & \text{if } u = 2w \text{ for some } 0 \leq w \leq t_k \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

*Proof.* To prove the Lemma, we must compute

$$\frac{\mathbf{P}(X_{m_1} = t_1, \ldots, X_{m_k} = t_k, X_n = u)}{\mathbf{P}(X_{m_1} = t_1, \ldots, X_{m_k} = t_k)}. \tag{18}$$

Note that by our standing assumption that all quantities are chosen so that conditional probabilities are defined, $t_1, \ldots, t_k$ are such that the denominator of (18) is nonzero. Put

$$\begin{aligned} T &= \{g \in G_n(f) : g \text{ fixes } t_i \text{ roots of } f^{\circ m_i} \text{ for } 1 \leq i \leq k\} \\ T_u &= \{g \in T : g \text{ fixes } u \text{ roots of } f^{\circ n}\} \end{aligned}$$

From (13), we see that we must compute $\#T_u/\#T$. Since the denominator of (18) is nonzero, $\#T \neq 0$ as well. As in the proof of Theorem 1.2, note that $T$ is invariant under multiplication by $H_n(f)$, whence it is a union of cosets of $H_n(f)$. Let $R_n$ be the set of roots of $f^{\circ n}$, and recall from Proposition-Definition 4.3 that the sets $f^{-1}(\alpha)$, where $\alpha$ is a root of $f^{\circ n-1}$, form a partition of $R_n$.

Consider a coset $g_0 H_n(f) \subset T$. By the maximality of $H_n(f)$, there exists $h_\alpha \in H_n(f)$ that exchanges the elements of $f^{-1}(\alpha)$ and fixes the elements of $f^{-1}(\alpha')$ for all $\alpha' \neq \alpha$. Let $Q$ be the set of roots of $f^{\circ n-1}$ fixed by $g_0$, and put $M = f^{-1}(Q) \subseteq R_n$. Since $g_0 \in T$, $g_0$ fixes $t_k$ roots of $f^{\circ n-1}$, whence $\#M = 2t_k$.

Now let $J$ be the subgroup of $H_n(f)$ that fixes each element of $M$. The maximality of $H_n(f)$ shows

$$\#J = 2^{(2^{n-1}-t_k)}. \tag{19}$$

Take $h \in H_n(f)$. Since $h$ fixes all roots of $f^{\circ n-1}$, we have $g_0 h(f^{-1}(\alpha)) = f^{-1}(g_0(\alpha))$ for all $\alpha$. Thus $g_0 h$ cannot fix any element of $R_n - M$. On the other hand, elements of $J$ fix all members of $M$, so it follows that every element of a set of the form $g_0 hJ$ has the same number of fixed points in $R_n$.

Since $M = \bigcup_{\alpha \in Q} f^{-1}(\alpha)$, we can write any $h \in H_n(f)$ as

$$j \prod_{\alpha \in Q} (h_\alpha)^{e_\alpha},$$

where $j \in J$ and $e_\alpha = 0$ or $1$ for each $\alpha$. Thus any coset $g_0 hJ$ may be written uniquely as $g_0 J \prod_{\alpha \in Q} (h_\alpha)^{e_\alpha}$. Moreover, all elements of this coset have exactly

$$2t_k - \sum_{\alpha \in Q} 2e_\alpha \tag{20}$$

fixed points in $R_n$ (recall $\#Q = t_k$). The number of ways (20) can equal $u$ is precisely $\binom{t_k}{w}$ if $u = 2w$ for some $0 \leq w \leq t_k$ and zero otherwise. Note that from (19) and the maximality of $H_n(f)$ we have $\#J/\#H_n(f) = 2^{-t_k}$. Hence the proportion of elements of $g_0 H_n(f)$ contained in $T_u$ is $\binom{t_k}{w} 2^{-t_k}$. The Lemma now follows from the fact that $T$ is a union of cosets of $H_n(f)$. □

Note that Lemma 5.2 remains valid if the $t_i$ are allowed to be 0. Indeed it is easy to see directly that $X_m = 0$ implies $X_n = 0$ for all $n > m$, and in the case $t_k = 0$, the Lemma gives $P(X_n = 0) = 1$.

We give two consequences of Lemma 5.2. The first requires the Markov property, which a stochastic process $X_1, X_2, \dots$ satisfies if

$$\mathbf{P}(X_n = u \mid X_{m_1} = t_1, \dots, X_{m_k} = t_k) = \mathbf{P}(X_n = u \mid X_{m_k} = t_k) \tag{21}$$

for $n$, any $m_1 < \cdots < m_k < n$ and any $u, t_i$. Such a stochastic process is called a *Markov chain*. Lemma 5.2 shows that, for $n$ with $H_n(f)$ maximal, $GP(f)$ obeys a restricted version of the Markov property at stage $n$ (since $m_k = n-1$ is required). However, if $H_n(f)$ is maximal for all $n$, it is a straightforward exercise to show $GP(f)$ is a Markov chain.

The second consequence of Lemma 5.2 is that when $H_n$ is maximal and for any $m < n$ and $1 \le w \le 2^{m-1}$, we have

$$\mathbf{P}(X_n = 2w \mid X_m = 2w, \ldots, X_{n-1} = 2w) = \binom{2w}{w} \frac{1}{4^w}. \tag{22}$$

We now give an upper bound for the right-hand side of (22).

**Lemma 5.3.** *Suppose $H_n(f)$ is maximal. Then for any $m < n$ and $u > 0$ we have*

$$\mathbf{P}(X_n = u \mid X_m = u, \ldots, X_{n-1} = u) \le \frac{1}{2}.$$

*Proof.* First note that if $u$ is not of the form $2w$ for some $1 \le w \le 2^{m-1}$ then $P(X_n = u) = 0$ from Lemma 5.2 and we are done. Thus we assume $u$ is of this form. From (22) we need only show that $c_w \stackrel{\text{def}}{=} \binom{2w}{w}\frac{1}{4^w} \le \frac{1}{2}$ for all $w \ge 1$. Note that

$$\frac{c_{w+1}}{c_w} = \frac{1}{4}\frac{(2w+2)(2w+1)}{(w+1)^2} = \frac{4w^2+6w+2}{4w^2+8w+4}.$$

The right-hand side of this equation is less than 1 for $w \ge 1$. Since $c_1 = 1/2$, the Lemma follows. □

*Proof of Theorem 1.3.* By Theorem 1.2, $GP(f)$ is a martingale, and thus is eventually constant with probability 1 (see Corollary 4.13). Therefore it remains only to show that for any $m \ge 0$ and $u > 0$,

$$\mathbf{P}\left(\bigcap_{i=m}^{\infty} X_i = u\right) = 0.$$

Clearly

$$\mathbf{P}\left(\bigcap_{i=m}^{\infty} X_i = u\right) \le \lim_{j\to\infty} \mathbf{P}\left(\bigcap_{i=m}^{j} X_i = u\right),$$

and note that the sequence on the right-hand side is decreasing. Let $C_i = \{X_i = u\}$, and suppose $\mathbf{P}(C_{j-1} \cap \cdots \cap C_m) \ne 0$ (otherwise we're done). We have

$$\mathbf{P}\left(\bigcap_{i=m}^{j} C_i\right) = \mathbf{P}(C_m)\mathbf{P}(C_{m+1} \mid C_m)\cdots\mathbf{P}(C_j \mid C_m \cap \cdots \cap C_{j-1}). \tag{23}$$

By Lemma 5.3, if $H_n(f)$ is maximal then

$$\mathbf{P}(C_n \mid C_m \cap \cdots \cap C_{n-1}) \le 1/2.$$

Let $S = \{n \in \mathbb{N} : H_n(f) \text{ maximal}\}$. Then (23) yields

$$\mathbf{P}\left(\bigcap_{i=m}^{j} X_i = u\right) \leq \left(\frac{1}{2}\right)^{\#(S \cap \{m,\ldots,j\})}$$

The infinitude of $S$ now gives $\lim_{j\to\infty} \mathbf{P}\left(\bigcap_{i=m}^{j} X_i = u\right) = 0$. □

Note that it follows from Theorem 1.3 that $\lim_{m\to\infty} \mathbf{P}(\{X_m > 0\}) = 0$.

We close this section with an examination of $GP(f)$ under the assumption that $H_n(f)$ is maximal for all $n$. This situation arises rather frequently, for example when $L = \mathbb{Q}$ and $f = x^2 + a$ for many values of $a$ [14]. It appears likely that $H_n(f)$ is maximal for all $n$ also in the case that concerns us, namely $L = k(t)$, $\operatorname{char} k \neq 2$, and $f = x^2 + t$; see Conjecture 6.7. The following definition is adapted from [4]. Statements about conditional probabilities apply only when the conditional probabilities are well-defined, and the sum of zero random variables is taken to be zero with probability 1.

**Definition 5.4.** A Markov chain $X_1, X_2, \ldots$ is *time-homogeneous* if $\mathbf{P}(X_n = u \mid X_{n-1} = t)$ depends only on $u$ and $t$. By a *branching process* we mean a time-homogenous Markov chain $X_1, X_2, \ldots$ taking nonnegative integer values such that the random variable $(X_n \mid X_{n-1} = t)$ has the same distribution as the sum of $t$ independent copies of $X_1$.

**Proposition 5.5.** *Suppose that $H_n(f)$ is maximal for all $n$. Then $GP(f)$ is a branching process with $\mathbf{P}(X_1 = 0) = 1/2$ and $\mathbf{P}(X_1 = 2) = 1/2$.*

*Proof.* Clear from Lemma 5.2 and the discussion of the Markov property immediately following. □

It is interesting to note that R.W.K. Odoni observes in [8, p. 398] that branching processes share many properties with iterated wreath products. This observation is a forerunner of Proposition 5.5, since it follows from [10, Lemma 1.1] that $H_n(f)$ maximal for all $n$ implies $G_n(f)$ is the $n$-fold iterated wreath product of $\mathbb{Z}/2\mathbb{Z}$.

Branching processes are very well-understood; see [3, Sec. 7.1] for a readable introduction and [4] for a detailed account. Here we merely state some results of interest in our case.

**Proposition 5.6.** *Let $X_1, X_2, \ldots$ be the branching process of Proposition 5.5. Let $a_n = \mathbf{P}(X_n = 0)$ and $b_n = \mathbf{P}(X_n > 0) = 1 - a_n$. Then*

1. *$a_n$ is given by the evaluation at $z = 0$ of the $n$th iterate of $\frac{1}{2} + \frac{1}{2}z^2$.*
2. *As $n \to \infty$, we have*

$$b_n = \frac{2}{n}\left\{1 - \frac{\log n}{n} - \frac{\alpha}{n} + O((\log n)^2/n^2)\right\}$$

*for some constant $\alpha$. In particular, $b_n \downarrow 0$.*

*Proof.* For 1, see [3, Theorem 7.2]. A proof of a much more general theorem than 2 can be found in [4, p. 21]. For a simpler, direct proof of 2 see [9, p. 5]. □

# 6 The Galois groups of iterates of $x^2 + t$

In this section we use the same notation as in Section 5, only with the following specializations: let $k$ be a field with char $k \neq 2$, $t$ be transcendental over $k$, $L = k(t)$, and $f(x) = x^2 + t$. We also put $A = k[t]$. As $f$ is fixed throughout this section, we write $G_n$ and $H_n$ in place of $G_n(f)$ and $H_n(f)$, respectively. Our goal is an in-depth examination of $H_n$, along the lines of that found for the characteristic 0 case in [10, 14]. At the end of the section we apply our results to give a proof of Theorem 1.4.

Let $\{p_n : n = 1, 2, 3\}$ be the adjusted forward orbit of the critical point 0, i.e. $p_1 = -t$ and $p_n = f^{\circ n}(0)$ for $n \geq 2$. Note that $p_n$ is a square in $L_n$ for all $n$; this is clear for $n = 1$, and follows for $n \geq 2$ because $p_n$ is the product of the roots of $f^{\circ n}$, which occur in an even number of $\pm$ pairs. We define a related sequence $\Phi_n$:

$$\Phi_n = \prod_{d|n} (p_d)^{\mu(n/d)} \in L. \tag{24}$$

We shall show that $p_n$ and $\Phi_n$ have much to do with the maximality of $H_n$. First we establish some divisibility properties of these sequences.

**Lemma 6.1.** *Let $q \in A$ be irreducible, let $v_q$ be the valuation corresponding to $q$, and suppose $v_q(p_n) = e \geq 1$. Then for all $m \geq 1$, we have $v_q(p_{mn}) = e$.*

*Proof.* Induction on $m$. The case $m = 1$ is trivial. Suppose inductively that $v_q(p_{(m-1)n}) = e$. Note that $p_{mn} = f^{\circ(m-1)n}(p_n)$, and also $f^{\circ(m-1)n}$ is a polynomial in $x^2$. Thus we can write

$$f^{\circ(m-1)n}(x) = x^2 g(x) + f^{\circ(m-1)n}(0) = x^2 g(x) + p_{(m-1)n},$$

for some $g \in L[x]$. Hence putting $x = p_n$ we have

$$p_{mn} = p_n^2 g(p_n) + p_{(m-1)n}.$$

Now $v_q\left[(p_n)^2(g(p_n))\right] \geq 2e$, and by our inductive hypothesis $v_q(p_{(m-1)n}) = e$. Since $e \geq 1$, the first summand vanishes to higher order at $q$ than the second, so we conclude $v_q(p_{mn}) = e$. ☐

**Proposition 6.2.** *For each $n$, $\Phi_n$ is a polynomial, and the $\Phi_n$ are pairwise relatively prime.*

*Proof.* Let $q \in A$ be irreducible, and let $m = \min\{n \geq 1 : q \mid p_n\}$. By Lemma 6.1, we have $v_q(p_n) = e$ if $m \mid n$ and $v_q(p_n) = 0$ otherwise. Thus

$$v_q(\Phi_n) = \sum_{d|n} v_q(p_d) \cdot \mu(n/d) = e \cdot \sum_{dm|n} \mu(n/dm),$$

and this last expression is $e$ if $n = m$ and 0 otherwise. Hence $\Phi_n$ is a polynomial and moreover $v_q(\Phi_n) > 0$ for only one $n$. Thus the $\Phi_n$ are pairwise relatively prime. ☐

**Proposition 6.3.** *For each $n$,* Disc $f^{\circ n} = a^2 p_n$ *for some $a \in A$.*

*Proof.* See the proof of Lemma 4.10, first paragraph. □

Our main result in this section has to do with the maximality of $H_n$ (See Proposition-Definition 5.1). We first give two preparatory results.

**Lemma 6.4.** *Let $n \geq 1$. Then $H_n$ is maximal if and only if $p_n$ is not a square in $L_{n-1}$.*

*Proof.* Identical to the argument in [14, Lemma 1.6]. □

**Theorem 6.5.** *Let $n \geq 1$. Then $H_n$ is maximal if and only if $\Phi_n$ is not a square in $L$.*

*Proof.* The case $n = 1$ is clear, so we take $n \geq 2$. Suppose that $\Phi_n$ is a square in $L$, and note that it follows from (24) that

$$p_n = \prod_{d|n} \Phi_d. \tag{25}$$

Now $p_m$ is a square in $L_m$ for all $m \geq 1$, and a quick induction allows one to deduce that $\Phi_m$ is also a square in $L_m$ for all $m \geq 1$. Thus from (25) we have that $p_n$ is a square in $L_{n-1}$, whence by Lemma 6.4 $H_n$ is not maximal.

Now suppose $\Phi_n$ is not a square in $L$. We claim the squarefree part of $\Phi_n$ has positive degree. To see this, note that $p_n$ is monic and of even degree for $n \geq 2$ while $p_1$ has odd degree and leading coefficient $-1$. Thus from (24) we have that $\Phi_n$ is monic and of even degree if $\mu(n) = 0$ and has leading coefficient $-1$ and odd degree otherwise. If $\mu(n) = 0$ then $\Phi_n$ is a monic non-square in $L$ and thus its squarefree part has positive degree. If $\mu(n) \neq 0$ then $\Phi_n$ has odd degree and thus its squarefree part has positive degree as well.

Now let $q \in A$ be an irreducible dividing the squarefree part of $\Phi_n$. Since $\Phi_n$ is relatively prime to $\Phi_m$ for $m < n$, $q$ cannot divide $\operatorname{Disc} f^{\circ n-1}$ by Proposition 6.3 and (25). Now a prime $\mathfrak{p} \subset A$ not dividing $(\operatorname{Disc} f^{\circ n-1})$ cannot be ramified in $L(\alpha)$, where $\alpha$ is a root of $f^{\circ n-1}$. From [7, Corollary 2, p. 157] it follows that $\mathfrak{p}$ is unramified in $L_{n-1}$. Therefore $(q)$ is unramified in $L_{n-1}$. Thus the squarefree part of $p_n$ has an irreducible factor unramified in $L_{n-1}$, whence $p_n$ cannot be a square in $L_{n-1}$. By Lemma 6.4 $H_n$ is therefore maximal. □

*Remark.* Recall $L = k(t)$, and let $F$ be the prime subfield of $k$. Since $\Phi_n \in F[t]$, the roots of $\Phi_n$ in $\overline{k}$ must lie in $\overline{F}$, whence all factors of $\Phi_n$ in $k[t]$ must have coefficients in $\overline{F}$. Thus if $\Phi_n$ is a square in $k[t]$, then in fact it is a square in $\overline{F}[t]$, and since $F$ is perfect it follows that $\Phi_n$ must be a square in $F[t]$. To show the last assertion, note that if $\Phi_n$ is not a square in $F[t]$, then the squarefree part of $\Phi_n$ has positive degree (same argument as in the proof of Theorem 6.5), and thus is divisible by an irreducible polynomial in $F[t]$. Since $F$ is perfect, this irreducible polynomial is separable, and thus cannot become a square in $\overline{F}[t]$, showing that $\Phi_n$ is not a square in $\overline{F}[t]$. We have now shown that $\Phi_n$ is a square in $L$ if and only if it is a square in $F[t]$, so that only the characteristic of $L$ is relevant in this matter. In particular, if $\Phi_n$ is not a square in $L$, then $H_n$ remains maximal if we replace $L$ by $\overline{k}(t)$. Therefore if $\Phi_n$ is not a square for all $n \leq m$, then $[L_n : k(t)] = [\overline{k}L_n : \overline{k}(t)]$, whence $L_n/L$ is geometric.

**Corollary 6.6.** *If $n$ is squarefree, then $H_n$ is maximal.*

*Proof.* From (24), $n$ squarefree implies $\deg \Phi_n$ odd. The Corollary now follows from Theorem 6.5. □

*Proof of Theorem 1.4.* Let $k = \mathbb{F}_p$ with $p \neq 2$, $L = k(t)$, and $f(x) = x^2+t$. By Proposition 3.2 it is enough to show $\lim_{n\to\infty} D(\mathcal{I}_n) = 0$. Let $X_1, X_2, \ldots$ be the Galois process of $f$, and note that by Theorem 1.5 and (13) we have $P(X_n > 0) = D(\mathcal{I}_n)$. From Lemma 6.1 and the fact that $v_t(p_1) = 1$, we have $v_t(p_n) = 1$ for all $n$. Hence the adjusted forward critical orbit of $f$ contains no squares. Finally, by Corollary 6.6 we have $H_n(f)$ maximal for infinitely many $n$. Theorem 1.3 then applies to show $\lim_{n\to\infty} \mathbf{P}(X_n = 0) = 1$, which implies $\lim_{n\to\infty} \mathbf{P}(X_n > 0) = 0$. □

We close with a conjecture.

**Conjecture 6.7.** Let char $k \neq 2$, $L = k(t)$, and $f(x) = x^2 + t$. Then $H_n$ is maximal for all $n \geq 1$.

Thanks to Propositions 5.5 and 5.6, Conjecture 6.7 would give a simpler proof of Theorem 1.4. It would also give near-complete information about $GP(f)$ and provide very precise estimates for $D(\mathcal{I}_n)$ for large $n$ (see part 2 of Proposition 5.6). Moreover, if Conjecture 6.7 is true, then it follows from the remark just before Corollary 6.6 that $L_n/L$ is geometric for all $n$. Thus the strong form of the Tchebotarev density theorem for function fields [13, Theorem 9.13B] applies to show that $\delta(\mathcal{H}(\overline{\mathbb{F}}_p)) = 0$ for $p \neq 2$ (see the discussion on p. 10)

One approach to proving Conjecture 6.7 is to use Theorem 6.5 and show $\Phi_n$ is not a square in $L$ for all $n \geq 1$. In the characteristic zero case one can show $\Phi_n$ is separable for all $n$ by reducing mod 2. In the case char $k \equiv 3 \bmod 4$ one can show $\Phi_n$ is not a square for all $n$ by adapting the argument in [14, Sections 2, 3] (see [5, Section 3.5] for details). The remaining cases are still open, though calculations for several small primes $p \equiv 1 \bmod 4$ have shown $\Phi_n$ is not a square for $n \leq 2000$.

## acknowledgements

This paper contains some of the results of the author's Ph.D. thesis, written at Brown University with Joseph Silverman. The author would like to thank Prof. Silverman for his generous support and guidance. Further thanks go to Michael Rosen and Rob Benedetto, for many useful discussions on function fields, $p$-adic dynamics, and other matters. Thanks also to Bas Edixhoven for suggesting a way of making the proofs of Lemma 4.6 and Theorem 4.7 shorter and more conceptual.

## References

[1] Yuval Fisher and Jay Hill. Bounding the area of the mandelbrot set. Availalbe at http://citeseer.ist.psu.edu/35134.html.

[2] Geoffrey R. Grimmett and David R. Stirzaker. *Probability and random processes*. Oxford University Press, New York, third edition, 2001.

[3] John Haigh. *Probability models*. Springer Undergraduate Mathematics Series. Springer-Verlag London Ltd., London, 2002.

[4] Theodore E. Harris. *The theory of branching processes*. Die Grundlehren der Mathematischen Wissenschaften, Bd. 119. Springer-Verlag, Berlin, 1963.

[5] Rafe Jones. *Galois martingales and the hyperbolic subset of the $p$-adic Mandelbrot set*. PhD thesis, Brown University, 2005.

[6] Curtis T. McMullen. Frontiers in complex dynamics. *Bulletin of the AMS*, 31(2):155–172, 1994.

[7] Władysław Narkiewicz. *Elementary and analytic theory of algebraic numbers*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, third edition, 2004.

[8] R. W. K. Odoni. The Galois theory of iterates and composites of polynomials. *Proc. London Math. Soc. (3)*, 51(3):385–414, 1985.

[9] R. W. K. Odoni. On the prime divisors of the sequence $w_{n+1} = 1 + w_1 \cdots w_n$. *J. London Math. Soc. (2)*, 32(1):1–11, 1985.

[10] R. W. K. Odoni. Realising wreath products of cyclic groups as Galois groups. *Mathematika*, 35(1):101–113, 1988.

[11] Juan Rivera-Letelier. *Dynamique des fonctions rationnelles sur des corps locaux*. PhD thesis, Universite Paris 6, 2001.

[12] Alain M. Robert. *A course in $p$-adic analysis*, volume 198 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000.

[13] Michael Rosen. *Number theory in function fields*, volume 210 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002.

[14] Michael Stoll. Galois groups over **Q** of some iterated polynomials. *Arch. Math. (Basel)*, 59(3):239–244, 1992.