ARITHMETIC DYNAMICS OF RANDOM POLYNOMIALS

PIERRE LE BOUDEC AND NIKI MYRTO MAVRAKI

Abstract. We investigate from a statistical perspective the arithmetic properties of the dynamics of polynomials of fixed degree and defined over the field of rational numbers. To start with, ordering their affine conjugacy classes by height, we show that their average number of rational preperiodic points is equal to zero, thereby proving a strong average version of the uniform boundedness conjecture of Morton and Silverman. Next, inspired by the analogy with the successive minima of a lattice we define the dynamical successive minima of a polynomial. Noting that these quantities are invariant under the action by conjugacy of the affine group we study their average behaviour using the aforementioned ordering by height. In particular, we prove an optimal statistical version of the dynamical Lang conjecture on the canonical height of rational non-preperiodic points.

CONTENTS

1. Introduction 1
2. Preliminaries 6
3. The uniform boundedness conjecture on average 17
4. Dynamical successive minima of random polynomials 19
References 26

1. INTRODUCTION

The arithmetic properties of the dynamics of polynomials defined over the field of rational numbers, despite being intensely investigated, remain largely elusive. Indeed, even in the simplest case of quadratic polynomials, the uniform boundedness conjecture of Morton and Silverman [18] on the number of rational preperiodic points, as well as the dynamical Lang conjecture [22, Conjecture 4.98] on the canonical height of rational non-preperiodic points are wide open. We note that the number of rational preperiodic points of a polynomial defined over \( \mathbb{Q} \) and the set of values of its canonical height at rational points are both invariant under the action by conjugacy of the affine group \( \mathbb{Q} \rtimes \mathbb{Q}^\times \). Given \( d \geq 2 \), it is thus natural to study these conjectures on average over affine conjugacy classes of polynomials of degree \( d \).

Outside a set of dimension \( d - 2 \), we can parametrize these conjugacy classes by a Zariski open subset of \( \mathbb{P}^{d-1}(\mathbb{Q}) \) as follows. We view projective points as primitive vectors of integers so we let \( \mathbb{Z}_{\text{prim}}^d \) be the set of \( (x_1, \ldots, x_d) \in \mathbb{Z}^d \) such that \( \gcd(x_1, \ldots, x_d) = 1 \), and we introduce the set

\[
\mathcal{P}_d = \left\{ \psi_a \in \mathbb{Q}[z] : \begin{array}{c}
a = (a_d, a_{d-2}, \ldots, a_0) \in \mathbb{Z}_{\text{prim}}^d \\
a_d \neq 0, \ a_0 > 0
\end{array} \right\},
\]

where

\[
\psi_a(z) = \frac{a_d}{a_0} z^d + \frac{a_{d-2}}{a_0} z^{d-2} + \cdots + \frac{a_1}{a_0} z + 1. \quad (1.1)
\]

2010 Mathematics Subject Classification. 37P05, 37P15, 37P35.
Key words and phrases. Polynomials, preperiodic points, canonical heights.
We remark that the affine conjugacy class of any degree $d$ polynomial which does not fix the barycenter of its roots in some algebraic closure of $\mathbb{Q}$ contains exactly one element of $\mathcal{P}_d$.

We shall order elements of $\mathcal{P}_d$ using the usual exponential height on projective space. We thus define the height $\mathcal{H}(\psi_a)$ of a polynomial $\psi_a \in \mathcal{P}_d$ as

$$\mathcal{H}(\psi_a) = \max \{|a_d|, |a_{d-2}|, \ldots, |a_1|, a_0\}.$$ 

Finally, given $X \geq 1$ we let

$$\mathcal{P}_d(X) = \{\psi_a \in \mathcal{P}_d : \mathcal{H}(\psi_a) \leq X\}.$$ 

We now proceed to introduce some classical notation. Given a set $S$ and a map $f : S \rightarrow S$, for any $n \geq 1$ we let $f^n$ denote the $n$-th iterate of $f$ and by convention we let $f^0$ be the identity map. Moreover, we let $\text{Prep}_S(f)$ denote the set of preperiodic points of $f$ in $S$, that is

$$\text{Prep}_S(f) = \left\{ s \in S : \exists \ell \geq 0 \exists m \geq 1 \ f^{\ell+m}(s) = f^\ell(s) \right\}.$$ 

The following conjecture is a particular case of the uniform boundedness conjecture of Morton and Silverman [18].

**Conjecture A** (Morton–Silverman). Let $d \geq 2$. There exists $C_d > 0$ such that for any $\psi_a \in \mathcal{P}_d$, we have

$$\#\text{Prep}_Q(\psi_a) \leq C_d.$$ 

We remark that Morton and Silverman actually conjecture that for any $n \geq 1$, the number of preperiodic points of degree $d$ morphisms $\mathbb{P}^n \rightarrow \mathbb{P}^n$ defined over a number field $K$ should be bounded in terms of $d$, $[K : \mathbb{Q}]$ and $n$. In addition, it is worth pointing out that Fakhruddin [5, Remark 2.6] has noticed that the particular case $K = \mathbb{Q}$ of this conjecture implies the general prediction.

We note that Conjecture A is analogous to the celebrated result of Mazur [15] on the number of torsion points on elliptic curves defined over $\mathbb{Q}$. However, its setting critically lacks the structure coming from the group law and therefore remains far out of reach of current technology. Indeed, the best result in the direction of Conjecture A allows the constant $C_d$ to depend on the number of prime numbers where the polynomial $\psi_a \in \mathcal{P}_d$ does not have potentially good reduction (see Benedetto’s work [3, Main Theorem]). In the worst possible case, this only yields

$$\#\text{Prep}_Q(\psi_a) \ll 1 + \log \mathcal{H}(\psi_a),$$

where the implied constant depends at most on $d$. We point out that throughout the article we use the notation $\ll \cdot$ as a convenient replacement for $\epsilon = O(\cdot)$. However, we note that for certain weighted homogeneous one parameter families of polynomials, Ingram [9] has shown that the analog of Conjecture A holds. We also remark that there has been considerable activity in the case of quadratic polynomials. Most notably, Poonen [20] has proved that one can take $C_2 = 8$ in Conjecture A provided that there does not exist a quadratic polynomial with a rational periodic point of exact period larger than 3. Unfortunately, it is only known that if such a point exists then its exact period has to be at least 6 (see [17] and [6]). Finally, Looper [13, 14] has shown that Conjecture A follows from a particular case of Vojta’s conjecture (see also [19]). The interested reader is invited to refer to [2, Section 4] for a recent survey of our current knowledge on the uniform boundedness conjecture.

Averaging Benedetto’s pointwise bound [3, Main Theorem], we deduce that

$$\frac{1}{\#\mathcal{P}_d(X)} \sum_{\psi_a \in \mathcal{P}_d(X)} \#\text{Prep}_Q(\psi_a) \ll (\log \log X) \log \log \log X.$$
Our first result offers a substantial improvement upon this upper bound.

**Theorem 1.1.** Let $d \geq 2$ and $\varepsilon > 0$. Let also $\vartheta_2 = 1/2$ and for $d \geq 3$, let

$$\vartheta_d = \frac{2(d+1)}{5d+1}.$$ 

We have

$$\frac{1}{\# \mathcal{P}_d(X)} \sum_{\psi_a \in \mathcal{P}_d(X)} \# \text{Prep}_Q(\psi_a) \ll \frac{1}{X^{\vartheta_d - \varepsilon}},$$

where the implied constant depends at most on $d$ and $\varepsilon$.

It follows in particular from Theorem 1.1 that when ordered by height, 100% of the affine conjugacy classes of degree $d$ polynomials defined over $\mathbb{Q}$ do not have any rational preperiodic point. Theorem 1.1 can thus be viewed as a strong average version of Conjecture A.

We remark that in view of the upper bound (1.2), it follows from Theorem 1.1 that all higher order moments of the quantity $\# \text{Prep}_Q(\psi_a)$ satisfy the same upper bound as the first moment. In addition, we expect that the contribution from the polynomials $\psi_a \in \mathcal{P}_d$ having a rational periodic point of exact period at least 2 is negligible compared to the contribution from the polynomials having a rational fixed point. As a result, geometry of numbers heuristics lead us to conjecture that there exists $\gamma_d > 0$ such that

$$\frac{1}{\# \mathcal{P}_d(X)} \sum_{\psi_a \in \mathcal{P}_d(X)} \# \text{Prep}_Q(\psi_a) \sim \frac{\gamma_d}{X}. \quad (1.3)$$

As it turns out, we can prove much more precise results than Theorem 1.1 by studying the growth of canonical heights. Recall that the canonical height $\hat{h}_{\psi_a} : \mathbb{Q} \to \mathbb{R}_{\geq 0}$ of a polynomial $\psi_a \in \mathcal{P}_d$ is defined by

$$\hat{h}_{\psi_a}(z) = \lim_{n \to \infty} \frac{1}{d^n} h(\psi_a^n(z)), \quad (1.4)$$

where $h : \mathbb{Q} \to \mathbb{R}_{\geq 0}$ denotes the logarithmic Weil height. We note that for any $z \in \mathbb{Q}$, we have

$$\hat{h}_{\psi_a}(z) = d \cdot \hat{h}_{\psi_a}(z). \quad (1.5)$$

Another crucial property of the canonical height $\hat{h}_{\psi_a}$ is that it vanishes exactly at preperiodic points of $\psi_a$, and more generally it measures how far a given rational number is from being a preperiodic point of $\psi_a$.

In the setting of elliptic curves, Lang [10, page 92] has made a conjecture on the minimum canonical height of non-torsion rational points. Lang’s conjecture is believed to be buried very deep and is only known to hold in special cases (see for instance the works of Silverman [21], and Hindry and Silverman [7]). Nevertheless, we note that the first author [11] has recently established a statistical version of this conjecture for the family of all elliptic curves defined over the field of rational numbers. Silverman [22, Conjecture 4.98] has formulated an intriguing dynamical analog of Lang’s conjecture. Unfortunately, as in the setting of elliptic curves this conjecture remains largely open and only partial or conditional results are known (see for example the works of Ingram [8, 9] and Looper [12, 14]).

Recall that given $\psi_a \in \mathcal{P}_d$ and $u \geq 0$, the set of $z \in \mathbb{Q}$ such that $\hat{h}_{\psi_a}(z) \leq u$ is a set of bounded Weil height and is thus finite. Hence, we can define

$$\lambda_1(\psi_a) = \min \left\{ \hat{h}_{\psi_a}(z) : z \in \mathbb{Q} \setminus \text{Prep}_Q(\psi_a) \right\}. \quad (1.6)$$

In view of [23, page 103], we see that the following conjecture implies Silverman’s prediction [22, Conjecture 4.98] when restricted to our setting.
Conjecture B (Silverman). Let $d \geq 2$. There exists $c_d > 0$ such that for any $\psi_a \in \mathcal{P}_d$, we have

$$\lambda_1(\psi_a) > c_d \log \mathcal{H}(\psi_a).$$

Maybe surprisingly, our next result shows that we can estimate very precisely the quantity $\lambda_1(\psi_a)$ for generic polynomials $\psi_a \in \mathcal{P}_d$.

**Theorem 1.2.** Let $d \geq 2$ and $\epsilon > 0$. We have

$$\lim_{X \to \infty} \frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : \left| \lambda_1(\psi_a) - \frac{\log \mathcal{H}(\psi_a)}{d(d-1)} \right| \leq \epsilon \log \mathcal{H}(\psi_a) \right\} = 1.$$

Theorem 1.2 implies in particular that when ordered by height, 100% of the affine conjugacy classes of degree $d$ polynomials defined over $\mathbb{Q}$ satisfy an optimal version of Conjecture B.

We are actually able to provide a much deeper investigation of the sets of values of canonical heights. Given $\psi_a \in \mathcal{P}_d$, we thus proceed to generalize the definition (1.6) of the quantity $\lambda_1(\psi_a)$ by introducing the dynamical successive minima of the polynomial $\psi_a$. Given $N \geq 1$, inspired by Baker and DeMarco [1, page 9] we say that an element of $\mathbb{Q}^N$ is $\psi_a$-dynamically independent if it does not satisfy any algebraic relation given by a nonzero polynomial $P \in \mathbb{Q}[X_1, \ldots, X_N]$ with the property that

$$\forall (\xi_1, \ldots, \xi_N) \in \mathbb{Q}^N \quad P(\xi_1, \ldots, \xi_N) = 0 \implies P(\psi_a(\xi_1), \ldots, \psi_a(\xi_N)) = 0. \quad (1.7)$$

Motivated by the analogy with the successive minima of a lattice, we introduce the following definition.

**Definition 1.3.** Let $d \geq 2$ and $\psi_a \in \mathcal{P}_d$. For $N \geq 1$, the $N$-th dynamical successive minimum $\lambda_N(\psi_a)$ of the polynomial $\psi_a$ is

$$\lambda_N(\psi_a) = \min \left\{ \max_{i \in \{1, \ldots, N\}} h_{\psi_a}(z_i) : (z_1, \ldots, z_N) \in \mathcal{I}_N(\psi_a) \right\},$$

where $\mathcal{I}_N(\psi_a)$ denotes the set of $\psi_a$-dynamically independent elements of $\mathbb{Q}^N$.

It is worth pointing out that the dynamical successive minima of a polynomial are invariant under the action by conjugacy of the affine group. In addition, it is easy to check that $z \in \mathbb{Q}$ is $\psi_a$-dynamically independent if and only if $z \notin \text{Prep}_{\mathbb{Q}}(\psi_a)$. We thus see that Definition 1.3 in the case $N = 1$ agrees with the definition (1.6).

Given $\psi_a \in \mathcal{P}_d$ and $N \geq 2$, Proposition 4.4 provides us with lower and upper bounds for the quantity $\lambda_N(\psi_a)$ under a mild assumption on $\psi_a$. However, it should not be surprising that the lower bound is only of interest when $\log N$ is large compared to $\log \mathcal{H}(\psi_a)$, and when this is not the case we have to investigate this problem from a statistical point of view. Our methods will prove to be powerful enough to allow us to handle the situation where the integer $N$ grows with the height of the polynomial $\psi_a$, as stated in Proposition 4.5. The full strength of our techniques is demonstrated by Propositions 4.4 and 4.5 but we have decided to defer the statement of these results to Section 4.3 in order to avoid introducing too many technical details here. Instead, we present a direct corollary of Theorems 1.1 and 1.2 and Propositions 4.4 and 4.5 which deals with the case where the integer $N$ is fixed.

**Theorem 1.4.** Let $d \geq 3$ and $\epsilon > 0$. Let also $N \geq 3$. We have

$$\lim_{X \to \infty} \frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_1(\psi_a) > \frac{\log \mathcal{H}(\psi_a)}{d(d-1)} \right\} = 1,$$

and

$$\lim_{X \to \infty} \frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : \frac{\lambda_2(\psi_a)}{\lambda_1(\psi_a)} \leq 1 + \frac{4}{d-1} + \epsilon \right\} = 1.$$
It is especially enlightening to restrict Theorem 1.4 to the case where $d$ is assumed to be large as both statements then become optimal. More precisely, it follows in particular from Theorem 1.4 that for any $\varepsilon > 0$, there exists $D \geq 3$ such that for any $d \geq D$ and $N \geq 3$, we have
\[
\lim_{X \to \infty} \frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : \left| \frac{\lambda_2(\psi_a)}{\lambda_1(\psi_a)} - d \right| \leq \varepsilon d \right\} = 1,
\]
and
\[
\lim_{X \to \infty} \frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : 0 \leq \frac{\lambda_N(\psi_a)}{\lambda_2(\psi_a)} \leq 1 - \varepsilon \right\} = 1.
\]
In other words, if $d$ is somewhat large then when ordered by height, 100% of the affine conjugacy classes of degree $d$ polynomials defined over $\mathbb{Q}$ have their second dynamical successive minimum which is about $d$ times larger than their first, while an arbitrarily large number of the next successive minima essentially have the same size.

We now proceed to give a quick sketch of the proof of our results. To establish Theorems 1.1 and 1.2, as well as the lower bound in Proposition 4.4, and Proposition 4.5, we make key use of the decomposition of the canonical height into local canonical heights, along with the fact that infinity is a superattracting fixed point of any polynomial. Broadly speaking, this allows us to show that most of the time if a polynomial $\psi_a \in \mathcal{P}_d$ has a rational preperiodic point or has an excessively small first dynamical successive minimum, then the integer $a_0$ has to be far from being $k$-free, where $k \geq 2$ is a suitable integer. If there were no restrictions on the choice of $k$, then in Theorem 1.1 one could in theory obtain the essentially optimal saving $X^{1-\varepsilon}$, for any $\varepsilon > 0$. Unfortunately, the decomposition of the canonical height into local canonical heights introduces error terms which reflect the fact that the dynamics of a polynomial $\psi_a \in \mathcal{P}_d$ is hard to analyze whenever the prime factorizations of the integers $a_0$ and $a_d$ are intimately linked. Moreover, these error terms become more and more problematic as the integer $k$ grows, which ultimately prevents us from getting closer to the conjectural estimate (1.3).

Finally, it is worth stressing that along the proof of the upper bound in Proposition 4.4 we crucially appeal to the celebrated work of Medvedev and Scanlon [16]. More precisely, we use the fact that if $\psi_a \in \mathcal{P}_d$ does not belong to a short list of exceptional cases, then the only irreducible algebraic subsets of $\mathbb{A}^N(\overline{\mathbb{Q}})$ defined by polynomials satisfying the property (1.7) are those which are defined by polynomials $P \in \overline{\mathbb{Q}}[X_1, \ldots, X_N]$ of the shape
\[
P(X_1, \ldots, X_N) = X_i - g(X_j),
\]
for some $i, j \in \{1, \ldots, N\}$ and $g \in \overline{\mathbb{Q}}[z]$ commuting with $\psi_a$. For polynomials $\psi_a \in \mathcal{P}_d$ satisfying a mild assumption this allows us to exhibit many $\psi_a$-dynamically independent vectors of rational numbers of controlled Weil height, and we can thus conclude by using a classical upper bound for the difference between the canonical height with respect to $\psi_a$ and the Weil height.

It may be useful to record here that Theorems 1.1 and 1.2 are respectively established in Sections 3 and 4.1, while Propositions 4.4 and 4.5 and Theorem 1.4 are proved in Section 4.3.

We finish this introduction by mentioning that we believe that our techniques are robust enough to prove analogous results for general number fields. However, there is no doubt that generalizing certain of the tools that we use would have made our proofs much longer and much more intricate. For the sake of concision and clarity we have thus decided not to follow this path.
Acknowledgements. The research of the first-named author is integrally funded by the Swiss National Science Foundation through the SNSF Professorship number 170565 awarded to the project Height of rational points on algebraic varieties. The research of the second-named author was also funded by this grant during the period that she spent at the University of Basel. Both the financial support of the SNSF and the perfect working conditions provided by the University of Basel are gratefully acknowledged. It is a pleasure for the authors to thank Laura DeMarco, Dragos Ghioca, Fabien Pazuki and Joe Silverman for their interest and for several insightful comments on an earlier version of this manuscript.

2. Preliminaries

2.1. Inequalities involving the local canonical heights. We start by introducing some classical notation. Given a prime number \( p \), we let \( | \cdot |_p \) denote the usual \( p \)-adic absolute value. We thus have \( |0|_p = 0 \) and if we let \( v_p(z) \) be the \( p \)-adic valuation of \( z \in \mathbb{Q}^\times \) then we have
\[
|z|_p = \frac{1}{p^{v_p(z)}}.
\]

Let \( d \geq 2 \) and \( \psi_n \in \mathcal{P}_d \). Letting \( v \) denote either a prime number \( p \) or \( \infty \), we recall that the local canonical heights \( \hat{\lambda}_{\psi_n,v} : \mathbb{Q} \to \mathbb{R}_{\geq 0} \) of the polynomial \( \psi_n \) are defined by setting
\[
\hat{\lambda}_{\psi_n,v}(z) = \lim_{n \to \infty} \frac{1}{d^n} \log \max \{|\psi_n^i(z)|_v\}.
\]
We will make frequent use of the fact that for any \( z \in \mathbb{Q} \) we have \( \hat{\lambda}_{\psi_n,v}(z) \geq 0 \), and it will also be useful to note that
\[
\hat{\lambda}_{\psi_n,v}(\psi_n(z)) = d \cdot \hat{\lambda}_{\psi_n,v}(z).
\]
Moreover, it follows from the product formula that
\[
\hat{h}_{\psi_n} = \sum_p \hat{\lambda}_{\psi_n,p} + \hat{\lambda}_{\psi_n,\infty}.
\]

In addition, we set
\[
r_v(\psi_n) = \max \left\{ \frac{a_0}{a_d}^{1/(d-1)}, \max_{i \in \{0, \ldots, d-2\}} \frac{|a_i|}{a_d}_v \right\}.
\]

We note here that we simply write \( | \cdot | \) to denote the Archimedean absolute value \( | \cdot |_\infty \).

The three following results are classical but we include their proofs for completeness. They crucially rely on the fact that infinity is a superattracting fixed point of any polynomial. Once combined, they imply in particular that given a polynomial \( \psi_n \in \mathcal{P}_d \), most rational numbers have large canonical height with respect to \( \psi_n \).

Lemma 2.1. Let \( d \geq 2 \) and \( \psi_n \in \mathcal{P}_d \). Let also \( p \) be a prime number. If \( z \in \mathbb{Q}^\times \) satisfies \( |z|_p > r_p(\psi_n) \) then
\[
\hat{\lambda}_{\psi_n,p}(z) - \log |z|_p = \frac{1}{d-1} \log \frac{|a_d|}{|a_0|_p}.
\]

Proof. The assumption \( |z|_p > r_p(\psi_n) \) implies that for any \( i \in \{0, \ldots, d-2\} \) we have
\[
\left| \frac{a_d z^d}{a_0} \right|_p > \left| \frac{a_i}{a_0} z^i \right|_p.
\]

Hence, recalling the definition (1.1) of the polynomial \( \psi_n \) we observe that
\[
|\psi_n(z)|_p = \left| \frac{a_d z^d}{a_0} \right|_p.
\]
Therefore, it follows from the assumption $|z|_p > r_p(\psi_\alpha)$ that $|\psi_\alpha(z)|_p \geq |z|_p$, and an elementary induction thus shows that for any $n \geq 1$, we have

$$|\psi^n_\alpha(z)|_p = \left| \frac{a_d}{a_0} \right|_p (d^{n-1}/(d-1)) |z|_p^{d^n}.$$ 

In addition, the assumption $|z|_p > r_p(\psi_\alpha)$ also yields

$$|\frac{a_d}{a_0}|_p^{1/(d-1)} |z|_p > 1.$$ 

As a result, we conclude that if $n$ is large enough then $|\psi^n_\alpha(z)|_p > 1$. Recalling the definition (2.1) of the local canonical height $\hat{\lambda}_{\psi_\alpha,p}$, we see that this completes the proof. □

The following statement is a straightforward consequence of Lemma 2.1.

**Lemma 2.2.** Let $d \geq 2$ and $\psi_\alpha \in \mathcal{P}_d$. Let also $p$ be a prime number. For any $z \in \mathbb{Q}^\times$, we have

$$\hat{\lambda}_{\psi_\alpha,p}(z) - \log |z|_p \geq -\log r_p(\psi_\alpha).$$

**Proof.** We start by noting that the statement is clear when $|z|_p \leq r_p(\psi_\alpha)$. Moreover, when $|z|_p > r_p(\psi_\alpha)$ it directly follows from Lemma 2.1 and the inequality

$$\frac{1}{d-1} \log \left| \frac{a_d}{a_0} \right|_p \geq -\log r_p(\psi_\alpha),$$

which completes the proof. □

Our next result is the Archimedean analog of Lemma 2.2.

**Lemma 2.3.** Let $d \geq 2$ and $\psi_\alpha \in \mathcal{P}_d$. For any $z \in \mathbb{Q}^\times$, we have

$$\hat{\lambda}_{\psi_\alpha,\infty}(z) - \log |z| \geq -\log r_\infty(\psi_\alpha) - \log d.$$ 

**Proof.** We first note that the statement is clear when $|z| \leq d \cdot r_\infty(\psi_\alpha)$. We now assume that $|z| > d \cdot r_\infty(\psi_\alpha)$ and we follow closely the lines of the proof of Lemma 2.1 to show that

$$\hat{\lambda}_{\psi_\alpha,\infty}(z) - \log |z| \geq \frac{1}{d-1} \log \left| \frac{a_d}{a_0} \right|_p - \frac{1}{d-1} \log 2. \quad (2.5)$$

The assumption $|z| > d \cdot r_\infty(\psi_\alpha)$ implies that for any $i \in \{0,\ldots, d-2\}$ we have

$$\left| \frac{a_d}{a_0}z^i \right| > d^{d-1} \frac{a_d}{a_0} z^i.$$

As a result, recalling the definition (1.1) of the polynomial $\psi_\alpha$ we see that the triangle inequality yields in particular

$$|\psi_\alpha(z)| \geq \left| \frac{a_d}{2a_0} z^d \right|.$$ 

Hence, using the assumption $|z| > d \cdot r_\infty(\psi_\alpha)$ we see that $|\psi_\alpha(z)|_p \geq |z|_p$, and an elementary induction thus shows that for any $n \geq 1$, we have

$$|\psi^n_\alpha(z)| \geq \left| \frac{a_d}{2a_0} \right|_p^{(d^{n-1}/(d-1))} |z|^{d^n}.$$ 

Moreover, the assumption $|z| > d \cdot r_\infty(\psi_\alpha)$ also gives

$$\left| \frac{a_d}{d a_0} \right|_p^{1/(d-1)} |z| > 1.$$
Therefore, we deduce that if \( n \) is large enough then \( |\psi_a^n(z)| > 1 \). Recalling the definition (2.1) of the local canonical height \( \hat{\lambda}_{\psi_a,\infty} \), we see that the lower bound (2.5) follows. We finally remark that the inequality
\[
\frac{1}{d-1} \log \left| \frac{a_d}{a_0} \right| \geq - \log r_{\infty}(\psi_a)
\]
allows us to complete the proof. \( \square \)

2.2. Inequalities involving the global canonical height. For \( k \geq 2 \) and \( n \geq 1 \), we define the \( k \)-free part \( s_k(n) \) of the integer \( n \) by
\[
s_k(n) = \prod_{p \mid n} \frac{1}{[n]_p}.
\]
Throughout the article we say that \( x \in \mathbb{Z} \) and \( y \geq 1 \) are respectively the numerator and the denominator of a given \( z \in \mathbb{Q} \) if and only if we have \( z = x/y \) and \( \gcd(x, y) = 1 \).

Our next result will be the most important tool in the proof of Theorems 1.1 and 1.2 in the case \( d = 2 \).

Lemma 2.4. Let \( \psi_a \in \Phi_2 \). For any \( z \in \mathbb{Q} \), we have
\[
\hat{h}_{\psi_a}(z) \geq \frac{1}{2} \log s_2(a_0).
\]
Proof. Let \( x \in \mathbb{Z} \) denote the numerator of \( z \). We start by writing
\[
\log s_2(a_0) = \sum_{p \mid s_2(a_0)} \log p + \sum_{p \mid x} \log p. \tag{2.6}
\]
We note that given a prime number \( p \) dividing \( s_2(a_0) \), we have \( p \nmid a_2 \) since \( (a_2, a_0) \in \mathbb{Z}_\text{prim}^2 \) so it follows that
\[
r_p(\psi_a) = \frac{1}{p^{a_2}}. \]
Therefore, for each prime number \( p \) dividing \( s_2(a_0) \) and such that \( p \nmid x \) we are in position to apply Lemma 2.1. We deduce in particular that
\[
\sum_{p \mid s_2(a_0)} \log p \leq \sum_{p \mid s_2(a_0)} \hat{\lambda}_{\psi_a, p}(z). \tag{2.7}
\]
Next, given any prime number \( p \) dividing \( s_2(a_0) \) and \( x \), it follows from the definition (1.1) of the polynomial \( \psi_a \) that \( |\psi_a(z)|_p = 1 \). Applying Lemma 2.1 for each prime number \( p \) dividing \( s_2(a_0) \) and \( x \) we thus obtain
\[
\sum_{p \mid s_2(a_0)} \log p = \sum_{p \mid s_2(a_0)} \hat{\lambda}_{\psi_a, p}(\psi_a(z)).
\]
Hence, the equality (2.2) yields
\[
\sum_{p \mid s_2(a_0)} \log p = 2 \sum_{p \mid s_2(a_0)} \hat{\lambda}_{\psi_a, p}(z). \tag{2.8}
\]
Putting together the equalities (2.6) and (2.8) and the upper bound (2.7) we get
\[
\log s_2(a_0) \leq 2 \sum_{p \mid s_2(a_0)} \hat{\lambda}_{\psi_a, p}(z).
\]
We immediately complete the proof by appealing to the equality (2.3). \( \square \)
We now introduce an arithmetic quantity which will play a pivotal role in our work. Given \( \ell, m \in \mathbb{Z} \), we let \( \Delta(\ell(m)) \) denote the largest divisor of \( m \) whose radical divides \( \ell \), that is
\[
\Delta(\ell(m)) = \prod_{p \mid \text{rad} \ell} \frac{1}{|m|_p}.
\] (2.9)
The remainder of this section deals with the case \( d \geq 3 \). Our next result asserts that for a generic polynomial \( \psi_a \in \mathcal{P}_d \), a typical rational number which has small canonical height with respect to \( \psi_a \) also has small Weil height.

**Lemma 2.5.** Let \( d \geq 3 \) and \( \psi_a \in \mathcal{P}_d \). Let also \( z \in \mathbb{Q}^\times \) and let \( x \in \mathbb{Z} \) and \( y \geq 1 \) be respectively the numerator and the denominator of \( z \). We have
\[
\hat{h}_{\psi_a}(z) - \log y \geq -\frac{1}{2} \log \Delta_y(a_d),
\]
and
\[
\hat{h}_{\psi_a}(z) - \log |x| \geq -\frac{1}{2} \log \Delta_y(a_d) - \log r_\infty(\psi_a) - \log d.
\]

**Proof.** We start by noting that since \( a_i \in \mathbb{Z} \) for any \( i \in \{0, \ldots, d-2\} \), the assumption \( d \geq 3 \) implies that for any prime number \( p \), we have
\[
r_p(\psi_a) \leq \frac{1}{|a_d|^{1/2}}.
\] (2.10)
Therefore, applying Lemma 2.2 for each prime number \( p \) dividing \( y \) we obtain
\[
\sum_{p \mid y} \hat{\lambda}_{\psi_a,p}(z) - \log y \geq -\frac{1}{2} \sum_{p \mid y} \log \frac{1}{|a_d|_p}.
\]
Appealing to the equality (2.3) we thus deduce that
\[
\hat{h}_{\psi_a}(z) - \hat{\lambda}_{\psi_a,\infty}(z) - \log y \geq -\frac{1}{2} \log \Delta_y(a_d),
\] (2.11)
which implies the first inequality claimed.

Next, we note that Lemma 2.3 states that
\[
\hat{\lambda}_{\psi_a,\infty}(z) - (\log |x| - \log y) \geq -\log r_\infty(\psi_a) - \log d.
\] (2.12)
Combining the lower bounds (2.11) and (2.12) we obtain the second inequality claimed, which completes the proof. \( \square \)

We now record an immediate consequence of Lemma 2.5 which will be used in the proof of the lower bound in Proposition 4.4. Given \( \psi_a \in \mathcal{P}_d \), it provides us with a lower bound for the difference between the canonical height with respect to \( \psi_a \) and the Weil height.

**Lemma 2.6.** Let \( d \geq 3 \) and \( \psi_a \in \mathcal{P}_d \). For any \( z \in \mathbb{Q} \), we have
\[
\hat{h}_{\psi_a}(z) - h(z) \geq -\frac{1}{2} \log \mathcal{H}(\psi_a) - \log d.
\]

**Proof.** We start by noting that the statement clearly holds in the case \( z = 0 \) so we now assume that \( z \in \mathbb{Q}^\times \). Let \( y \geq 1 \) denote the denominator of \( z \). Lemma 2.5 shows in particular that
\[
\hat{h}_{\psi_a}(z) - h(z) \geq -\frac{1}{2} \log \Delta_y(a_d) - \log r_\infty(\psi_a) - \log d.
\]
But it is clear that \( \Delta_y(a_d) \leq |a_d| \), and moreover the assumption \( d \geq 3 \) implies that
\[
r_\infty(\psi_a) \leq \frac{\mathcal{H}(\psi_a)^{1/2}}{|a_d|^{1/2}}.
\]
We thus have
\[
\frac{1}{2} \log \Delta_p(a_d) + \log r_\infty(\psi_a) \leq \frac{1}{2} \log \mathcal{H}(\psi_a),
\]
which completes the proof. \(\square\)

Given \(\psi_a \in \mathcal{S}_d\) and \(x, y \in \mathbb{Z}\), it is convenient to set
\[
\sigma_a(x, y) = a_dx^d + a_{d-2}x^{d-2}y^2 + \cdots + a_1xy^{d-1}. \tag{2.13}
\]
The following result deals with the case \(d \geq 3\) and will be the key tool in the proof of Theorems 1.1 and 1.2, and Proposition 4.5. It asserts that given a generic polynomial \(\psi_a \in \mathcal{S}_d\) and \(k \in \{2, 3\}\), if the canonical height with respect to \(\psi_a\) of \(z \in \mathbb{Q}^\times\) is small then a large factor of the \(k\)-free part of \(a_0\) has to divide \(\sigma_a(x, y)\). We will make use of the case \(k = 2\) in the proof of Theorem 1.2 and Proposition 4.5. We note that the exponent stated in Theorem 1.1 crucially relies upon the case \(k = 3\) but the case \(k = 2\) is actually sufficient to establish a version of Theorem 1.1 with a weaker exponent.

**Lemma 2.7.** Let \(d \geq 3\) and \(k \in \{2, 3\}\). Let also \(\psi_a \in \mathcal{S}_d\). Finally, let \(z \in \mathbb{Q}^\times\) and let \(x \in \mathbb{Z}\) and \(y \geq 1\) be respectively the numerator and the denominator of \(z\). We have
\[
d(d-1)\hat{h}_{\psi_a}(z) \geq \frac{d+k-2}{k-1} \log \frac{s_k(a_0)}{\gcd(s_k(a_0), \sigma_a(x, y))} - \log \gcd(s_k(a_0), a_d) - \frac{d-1}{2} \log \Delta_{s_k(a_0)}(a_d).
\]

**Proof.** Let \(w \geq 1\) denote the denominator of \(\psi_a(z)\). We start by writing
\[
\log s_k(a_0) = \sum_{p|s_k(a_0)} \log \frac{1}{|a_0|_p} + \sum_{p|w} \log \frac{1}{|a_0|_p}. \tag{2.14}
\]

For any prime \(p\), the \(p\)-adic valuation of \(s_k(a_0)\) is at most \(k-1\) so we clearly have the inequality
\[
\sum_{p|s_k(a_0)} \log \frac{1}{|a_0|_p} \leq (k-1) \sum_{p|w} \log |\psi_a(z)|_p.
\]

We now remark that the upper bound (2.10) shows that if \(p\) is a prime number dividing \(w\) but such that \(p^2 \nmid a_d\) then we are in position to use Lemma 2.1. Therefore, for each prime number \(p\) dividing \(s_k(a_0)\) and \(w\) we proceed to apply either Lemma 2.1 if \(p^2 \nmid a_d\) or Lemma 2.2 if \(p^2 | a_d\). Using the upper bound (2.10), we deduce in particular that
\[
\sum_{p|s_k(a_0)} \log \frac{1}{|a_0|_p} \leq (k-1) \sum_{p|s_k(a_0)} \hat{h}_{\psi_a,p}(\psi_a(z)) - \frac{k-1}{d-1} \sum_{p|\gcd(s_k(a_0), a_d)} \log \frac{a_d}{a_0} \bigg|_{a_0|_p} \\
+ \frac{k-1}{2} \sum_{p^2|a_d} \log \frac{1}{|a_d|_p}. \tag{2.15}
\]
Moreover, we note that

$$\sum_{p \mid \gcd(s_k(a_0), w)} \log \frac{|a_d|}{|a_0|^p} \geq \sum_{p \mid \gcd(s_k(a_0), w)} \log \frac{1}{|a_0|^p}$$

$$\geq \sum_{p \mid s_k(a_0)} \log \frac{1}{|a_0|^p} - \sum_{p \mid \gcd(s_k(a_0), w)} \log \frac{1}{|a_0|^p}$$

$$\geq \log s_k(a_0) - \sum_{p \mid s_k(a_0)} \log \frac{1}{|a_0|^p} - \sum_{p \mid \gcd(s_k(a_0), w)} \log \frac{1}{|a_0|^p}. \quad (2.16)$$

Putting together the equality (2.14) and the inequalities (2.15) and (2.16), we derive

$$(k - 1) \sum_{p \mid s_k(a_0)} \hat{\lambda}_{\psi_a, p}(\psi_a(z)) \geq \frac{d + k - 2}{d - 1} \log s_k(a_0) - \frac{d + k - 2}{d - 1} \sum_{p \mid s_k(a_0)} \log \frac{1}{|a_0|^p}$$

$$- \frac{k - 1}{d - 1} \sum_{p \mid s_k(a_0)} \log \frac{1}{|a_0|^p} - \frac{k - 1}{2} \sum_{p \mid \gcd(s_k(a_0), w)} \log \frac{1}{|a_0|^p}. \quad (2.17)$$

Next, using the equalities (2.3) and (1.5) we get

$$\sum_{p \mid s_k(a_0)} \hat{\lambda}_{\psi_a, p}(\psi_a(z)) \leq d \cdot \hat{h}_{\psi_a}(z). \quad (2.18)$$

Furthermore, recalling the definition (2.13) of the quantity $\sigma(x, y)$, we see that it follows from the definition (1.1) of the polynomial $\psi_a$ that

$$w = \frac{a_0 y^d}{\gcd(a_0 y^d, \sigma(x, y))}.$$  

We thus deduce that

$$\sum_{p \mid s_k(a_0), p \mid w} \log \frac{1}{|a_0|^p} \leq \log \gcd(s_k(a_0), \sigma(x, y)). \quad (2.19)$$

Putting together the lower bound (2.17) and the upper bounds (2.18) and (2.19), we derive

$$d(d - 1)\hat{h}_{\psi_a}(z) \geq \frac{d + k - 2}{k - 1} \log \frac{s_k(a_0)}{\gcd(s_k(a_0), \sigma(x, y))} - \sum_{p \mid s_k(a_0)} \log \frac{1}{|a_0|^p}$$

$$- \frac{d - 1}{2} \sum_{p \mid \gcd(s_k(a_0), \sigma(x, y))} \log \frac{1}{|a_0|^p}.$$  

In the case $k = 2$, we note that

$$\sum_{p \mid s_k(a_0), p \mid a_d} \log \frac{1}{|a_0|^p} \leq \log \gcd(s_k(a_0), a_d).$$
Since we clearly have
\[ \sum_{p \mid s_k(a_0)} \frac{1}{|a_d|_p} \leq \log \Delta_{s_k(a_0)}(a_d), \]
we see that this completes the proof in the case \( k = 2 \).

In the case \( k = 3 \), we remark that
\[ \sum_{p \mid s_k(a_0)} \frac{1}{|a_0|_p} = \log \gcd(s_k(a_0), a_d) + \sum_{p^2 \mid s_k(a_0)} \frac{1}{|a_d|_p}. \]
Hence, it follows from our assumption \( d \geq 3 \) that
\[ \sum_{p \mid s_k(a_0)} \frac{1}{|a_0|_p} + \frac{d-1}{2} \sum_{p \mid s_k(a_0)} \frac{1}{|a_d|_p} \leq \log \gcd(s_k(a_0), a_d) + \frac{d-1}{2} \log \Delta_{s_k(a_0)}(a_d), \]
which completes the proof in the case \( k = 3 \). \( \square \)

In order to establish Theorem 1.2 and the upper bound in Proposition 4.4 we will also need the following result which, given \( \psi_a \in \mathcal{P}_d \), provides us with an upper bound for the difference between the canonical height with respect to \( \psi_a \) and the Weil height.

**Lemma 2.8.** Let \( d \geq 2 \) and \( \psi_a \in \mathcal{P}_d \). For any \( z \in \mathbb{Q} \), we have
\[ \hat{h}_{\psi_a}(z) - h(z) \leq \frac{1}{d-1} \log \mathcal{H}(\psi_a) + \frac{\log d}{d-1}. \]

**Proof.** For any \( n \geq 0 \) we let \( x_n \in \mathbb{Z} \) and \( y_n \geq 1 \) be respectively the numerator and the denominator of \( \psi_a^n(z) \). Given \( x, y \in \mathbb{Z} \), recall the definition (2.13) of the quantity \( \sigma_n(x, y) \). It follows from the definition (1.1) of the polynomial \( \psi_a \) that for any \( n \geq 0 \), we have
\[ \psi_{a}^{n+1}(z) = \frac{\sigma_n(x_n, y_n) + a_0 y_n^d}{a_0 y_n^d}. \]
We thus have
\[ h(\psi_a^{n+1}(z)) \leq \log \max \left\{ \left| \sigma_n(x, y_n) + a_0 y_n^d \right|, a_0 y_n^d \right\}. \]
But it is clear that
\[ \left| \sigma_n(x, y_n) + a_0 y_n^d \right| \leq d \mathcal{H}(\psi_a) \max \left\{ |x_n|^d, y_n^d \right\}. \]
As a result, we see that
\[ h(\psi_a^{n+1}(z)) \leq \log(d \mathcal{H}(\psi_a)) + d \cdot h(\psi_a^n(z)). \]
Hence, we derive
\[ \frac{1}{dn+1} h(\psi_a^{n+1}(z)) - \frac{1}{dn} h(\psi_a^n(z)) \leq \frac{1}{dn+1} \log(d \mathcal{H}(\psi_a)). \]
Summing this inequality over the integer \( n \), we deduce that for any \( m \geq 0 \), we have
\[ \frac{1}{dm} h(\psi_a^m(z)) - h(z) \leq \frac{1}{dm} \cdot \frac{d^m - 1}{d-1} \log(d \mathcal{H}(\psi_a)). \]
Recalling the definition (1.4) of the canonical height and letting \( m \) tend to \( \infty \), we see that this completes the proof. \( \square \)
2.3. A subfamily of the family of all affine conjugacy classes of polynomials.

It will be useful to remark that a Möbius inversion shows that

\[
\# \mathcal{P}_d(X) = \frac{2^{d-1}}{\zeta(d)} X^d \left( 1 + O \left( \frac{\log X}{X} \right) \right). \tag{2.20}
\]

For \( n \geq 1 \) we let \( \text{rad}(n) \) denote the radical of the integer \( n \), that is

\[
\text{rad}(n) = \prod_{p \mid n} p.
\]

It was proved by de Bruijn [4, Theorem 1] that

\[
\log \left( \sum_{n \leq X} \frac{1}{\text{rad}(n)} \right) \sim \left( \frac{8 \log X \log \log X}{\log \log X} \right)^{1/2}.
\]

The following result is an immediate consequence of this estimate.

**Lemma 2.9.** Let \( \varepsilon > 0 \). We have

\[
\sum_{n \leq X} \frac{1}{\text{rad}(n)} \ll X^{\varepsilon},
\]

where the implied constant depends at most on \( \varepsilon \).

Given \( \ell, m \in \mathbb{Z} \), recall the definition (2.9) of the largest divisor \( \Delta_\ell(m) \) of \( m \) whose radical divides \( \ell \). For \( \delta \in (0, 1) \), we introduce the set

\[
\mathcal{P}_d^{(\delta)}(X) = \left\{ \psi_a \in \mathcal{P}_d(X) : \Delta_{a_0}(a_d) \leq X^\delta \right\}. \tag{2.21}
\]

Our next task is to prove that for any \( \delta \in (0, 1) \), the cardinality of the set \( \mathcal{P}_d^{(\delta)}(X) \) is asymptotically as large as the cardinality of the total set \( \mathcal{P}_d(X) \).

**Lemma 2.10.** Let \( d \geq 2 \). Let also \( \delta \in (0, 1) \) and \( \varepsilon > 0 \). We have

\[
\# \mathcal{P}_d^{(\delta)}(X) = \# \mathcal{P}_d(X) \left( 1 + O \left( \frac{1}{X^{\delta - \varepsilon}} \right) \right),
\]

where the implied constant depends at most on \( d, \delta \) and \( \varepsilon \).

**Proof.** We start by noting that

\[
\# \mathcal{P}_d^{(\delta)}(X) - \# \mathcal{P}_d(X) \ll X^{d-2} \sum_{0 < a_0, |a_d| \leq X} \# \left\{ \ell \mid a_d : \frac{\text{rad}(\ell)}{\ell \cdot a_d} > X^{-\delta} \right\}.
\]

Therefore, we get

\[
\# \mathcal{P}_d^{(\delta)}(X) - \# \mathcal{P}_d(X) \ll X^{d-2} \sum_{X^\delta < \ell \leq X} \frac{1}{\text{rad}(\ell)} \sum_{0 < a_0 \leq X} \frac{1}{\text{rad}(\ell) \cdot a_0} \sum_{0 < |a_d| \leq X} 1.
\]

As a result, Lemma 2.9 gives

\[
\# \mathcal{P}_d^{(\delta)}(X) - \# \mathcal{P}_d(X) \ll X^{d-\delta + \varepsilon}.
\]

Recalling the estimate (2.20), we see that this completes the proof. \( \Box \)
2.4. On the average number of rational numbers with small canonical height.

The following result will be the key tool in the proof of our results and we will invoke it repeatedly. More precisely, we will use it with \( k = 3 \) to obtain the error term stated in Theorem 1.1, and we will apply it with \( k = 2 \) in the proof of Theorem 1.2 and Proposition 4.5.

**Lemma 2.11.** Let \( d \geq 3 \) and \( k \in \{2, 3\} \). Let also \( \delta \in (0, 1) \) and \( \varepsilon > 0 \). Finally, let

\[
\alpha_{d,k} = \frac{(k-1)^2(d-1)}{k(d+k-2)}.
\]

For \( X, T \geq 1 \), we have

\[
\sum_{\psi \in \mathcal{P}^d_{\delta}(X)} \# \left\{ z \in \mathbb{Q}^\times : \hat{h}_{\psi}(z) \leq \log T \right\} \ll X^{d-1 + 1/k + \delta \alpha_{d,k}/2 + \varepsilon} T^{2d \alpha_{d,k}},
\]

and

\[
\sum_{\psi \in \mathcal{P}^d_{\delta}(X)} \# \left\{ z \in \mathbb{Q} : \hat{h}_{\psi}(z) \leq \log T \right\} \ll X^{d-1 + 1/k + \delta \alpha_{d,k}/2 + \varepsilon} T^{2d \alpha_{d,k}},
\]

where the implied constants depend at most on \( d, \delta \) and \( \varepsilon \).

**Proof.** We start by proving the first statement. Recall the definition (2.4) of the quantity \( r_{\infty}(\psi_a) \). We note that for \( \psi \in \mathcal{P}^d_{\delta}(X) \) we have \( |a_i| \leq X \) for any \( i \in \{0, \ldots, d-2\} \), so the assumption \( d \geq 3 \) implies that

\[
r_{\infty}(\psi_a) \leq \frac{X^{1/2}}{|a_d|^{1/2}}.
\]

In addition, we recall that since \( \psi \in \mathcal{P}^d_{\delta}(X) \) we have in particular \( \Delta_{\psi_y}(a_0)(a_d) \leq X^\delta \).

As a result, applying Lemmas 2.5 and 2.7 and letting respectively \( x \in \mathbb{Z} \) and \( y \geq 1 \) denote the numerator and the denominator of \( z \in \mathbb{Q}^\times \), we find that

\[
\# \mathcal{F}^\times_{\psi_a}(T) \leq \# \left\{ z \in \mathbb{Q}^\times : \begin{array}{l}
x \ll \frac{X^{1/2} T \Delta_y(a_d)^{1/2}}{|a_d|^{1/2}}, \\
y \leq T \Delta_y(a_d)^{1/2} \\
\gcd(s_k(a_0), \sigma_a(x,y)) \geq \frac{s_k(a_0)}{A(X,T) \gcd(s_k(a_0), a_d)^{\beta_{d,k}}}
\end{array} \right\},
\]

where we have set

\[
\mathcal{F}^\times_{\psi_a}(T) = \left\{ z \in \mathbb{Q}^\times : \hat{h}_{\psi}(z) \leq \log T \right\},
\]

and

\[
A(X,T) = \left( X^{d/2} T^d \right)^{k \alpha_{d,k} / (k-1)},
\]

and also

\[
\beta_{d,k} = \frac{k-1}{d+k-2}.
\]

We thus have

\[
\sum_{|a_1| \leq X} \# \mathcal{F}^\times_{\psi_a}(T) \ll \sum_{D \geq s_k(a_0)/A(X,T) \gcd(s_k(a_0), a_d)^{\beta_{d,k}}} \sum_{0 < y \leq T \Delta_y(a_d)^{1/2}} \sum_{0 < x \leq X \gcd(x,y) = 1} \sum_{|a_1| \leq X} \sum_{\sigma_a(x,y) \equiv 0 \mod D} 1.
\]

Since \( D \) is cubefree and \( \gcd(x,y) = 1 \), the congruence \( \sigma_a(x,y) \equiv 0 \mod D \) implies that \( \gcd(y^{d-1}, D) \mid \gcd(D, a_d) \).
Using the fact that \( D \leq X \), we thus see that we have in particular
\[
\sum_{|a_1| \leq X} 1 \ll \frac{X \gcd(x, D) \gcd(D, a_d)}{D}.
\]

Therefore, letting \( \tau \) denote the divisor function, the summation over \( x \) yields
\[
\sum_{|a_1| \leq X} \# \mathcal{P}_x^\infty(T) \ll \frac{X^{3/2}T}{|a_d|^{1/2}} \left( \sum_{0 \leq y \leq T} \Delta y(a_d)^{1/2} \right) \sum_{D \mid \mathcal{B}_k(a_0)} \frac{\gcd(D, a_d) \tau(D)}{D}.
\]
(2.24)

It is convenient to introduce the arithmetic function \( \xi \) defined for nonzero \( n \in \mathbb{Z} \) by
\[
\xi(n) = \sum_{\ell \mid n} \frac{\ell}{\text{rad}(\ell)}.
\]
(2.25)

Recalling the definition (2.9) of the largest divisor \( \Delta y(a_d) \) of \( a_d \) whose radical divides \( y \), we deduce
\[
\sum_{0 < y \leq T \Delta y(a_d)^{1/2}} \Delta y(a_d)^{1/2} \ll \sum_{\ell \mid a_d} \ell^{1/2} \sum_{0 < y \leq T \ell^{1/2}} \frac{1}{\text{rad}(\ell)y} \ll T \xi(a_d).
\]

Moreover, we recall that the divisor bound states that for any \( n \geq 1 \) and \( \varepsilon > 0 \), we have
\[
\tau(n) \ll n^\varepsilon.
\]
(2.26)

Therefore, recalling the upper bound (2.24) we see that
\[
\sum_{|a_1| \leq X} \# \mathcal{P}_x^\infty(T) \ll \frac{X^{3/2+\varepsilon}T^2 \xi(a_d)}{|a_d|^{1/2}} \sum_{D \mid \mathcal{B}_k(a_0)} \frac{\gcd(D, a_d) \tau(D)}{D}.
\]

We now split the sum over \( a_0 \) into sums over its \( k \)-full part \( a_0/s_k(a_0) \) and its \( k \)-free part \( s_k(a_0) \). The classical upper bound for the number of \( k \)-full numbers with bounded absolute value asserts that
\[
\# \left\{ b \in \mathbb{Z} : 0 < b \leq Y \forall p \mid b \implies p^k \mid b \right\} \ll Y^{1/k}.
\]
(2.27)

As a result, we deduce that
\[
\sum_{a_0, |a_1| \leq X} \# \mathcal{P}_x^\infty(T) \ll \frac{X^{3/2+1/k+\varepsilon}T^2 \xi(a_d)}{|a_d|^{1/2}} \sum_{D \leq X} \frac{\gcd(D, a_d)}{D} \sum_{s \leq D \mathcal{A}(X, T) \gcd(s, a_d)^{d_{d,k}}} \frac{1}{s^{1/k}}.
\]
But we easily get

$$\sum_{s \leq DA(X,T) \, \gcd(s,a_d)^\beta_{d,k}} \frac{1}{s^{1/k}} \ll \frac{1}{D^{1/k}} \sum_{u \leq A(X,T) \, \gcd(Du,a_d)^\beta_{d,k}} \frac{1}{u^{1/k}}$$

$$\ll \frac{1}{D^{1/k}} \sum_{\ell \mid a_d} \frac{1}{\ell^{1/k}} \sum_{v \leq A(X,T) \, \gcd(Dv,a_d)^\beta_{d,k}/\ell^{1-k_{d,k}}} \frac{1}{v^{1/k}}$$

$$\ll A(X,T)^{1-1/k} \, \gcd(D,a_d)^{(k-1)\beta_{d,k}/k} D^{1/k} \sum_{m \mid a_d} \frac{1}{m^{1-(k-1)\beta_{d,k}/k}}.$$ 

Moreover, we note that it follows from the definition (2.23) of the quantity $\beta_{d,k}$ and our assumptions $d \geq 3$ and $k \in \{2,3\}$ that

$$\beta_{d,k} \leq \frac{1}{k-1}.$$ 

(2.28)

Hence, an application of the divisor bound (2.26) gives

$$\sum_{s \leq DA(X,T) \, \gcd(s,a_d)^\beta_{d,k}} \frac{1}{s^{1/k}} \ll \frac{X^\varepsilon \, A(X,T)^{1-1/k} \, \gcd(D,a_d)^{(k-1)\beta_{d,k}/k}}{D^{1/k}}.$$ 

(2.29)

Recalling the definition (2.22) of the quantity $A(X,T)$, we see that we have obtained

$$\sum_{\psi_{a_d} \in \mathcal{B}_{d}(X)} \# \mathcal{B}_{d,d}(T) \ll \frac{X^{3/2+1/k+6_{d,k}+2+2k_{d,k}} T^{2+do_{d,k}} \xi(a_d)}{a_d^{1/k}} \sum_{D \leq X} \frac{\gcd(D,a_d)^{1+(k-1)\beta_{d,k}/k}}{D^{1+1/k}}.$$ 

But we have

$$\sum_{D \leq X} \frac{\gcd(D,a_d)^{1+(k-1)\beta_{d,k}/k}}{D^{1+1/k}} \ll \sum_{m \mid a_d} \frac{1}{m^{1-(k-1)\beta_{d,k}/k}}.$$ 

Therefore, using the upper bound (2.28) and applying the divisor bound (2.26) once again, we derive

$$\sum_{\psi_{a_d} \in \mathcal{B}_{d}(X)} \# \mathcal{B}_{d,d}(T) \ll X^{d-3/2+1/k+6_{d,k}+2+2k_{d,k}} T^{2+do_{d,k}} \sum_{0 < |a_d| \leq X} \frac{\xi(a_d)}{|a_d|^{1/2}}.$$ 

But recalling the definition (2.25) of the arithmetic function $\xi$, we observe that

$$\sum_{0 < |a_d| \leq X} \frac{\xi(a_d)}{|a_d|^{1/2}} \ll \sum_{T \leq X} \frac{T^{1/2}}{\text{rad}(\ell)} \sum_{0 < |a| \leq X/\ell} \frac{1}{|a|^{1/2}}$$

$$\ll X^{1/2} \sum_{T \leq X} \frac{1}{\text{rad}(\ell)}.$$ 

As a result, we see that an application of Lemma 2.9 completes the proof of the first statement.

Next, we establish the second statement. We start by noting that the assumptions $d \geq 3$ and $k \in \{2,3\}$ imply that

$$2 + do_{d,k} \leq d^2 \alpha_{d,k}.$$
Hence, we see that it suffices to prove that
\[ \sum_{\psi_a \in \mathcal{P}_d(X)} 1 \ll X^{d-1+1/k+\delta a_{d,k}/2+\varepsilon T^d a_{d,k}}. \] (2.30)

The equalities \( \psi_a(0) = 1 \) and (1.5) show that if \( \hat{h}_{\psi_a}(0) \leq \log T \) then \( \hat{h}_{\psi_a}(1) \leq \log T^d. \) In addition, since \( \psi_a \in \mathcal{P}_d(X) \) we have in particular \( \Delta_{\psi_a(a_0)}(a_d) \leq X^d. \) An application of Lemma 2.7 with \( z = 1 \) thus yields
\[
\sum_{\psi_a \in \mathcal{P}_d(X)} 1 \leq \sum_{\psi_a \in \mathcal{P}_d(X)} \# \{ D | s_k(a_0) : D \geq B(X,T) \gcd(s_k(a_0),a_d)^{\beta_{d,k}} \},
\]
where we have set
\[ B(X,T) = \left( X^{\delta/2} T^d \right)^{\kappa_{d,k}/(k-1)}. \] (2.31)

We proceed to carry out the summation over \( a_1 \) first. Using the fact that \( D \leq X, \) we get
\[
\sum_{\psi_a \in \mathcal{P}_d(X)} 1 \leq \sum_{\psi_a \in \mathcal{P}_d(X)} \sum_{0 < a_0, |a_d| \leq X} \sum_{D | s_k(a_0) \geq B(X,T) \gcd(s_k(a_0),a_d)^{\beta_{d,k}}} \sum_{0 \leq a_1 \leq X} \frac{1}{D}.
\]

Once again we handle the summation over \( a_0 \) by first summing over its \( k \)-full part \( a_0/s_k(a_0) \) using the upper bound (2.27). This yields
\[
\sum_{\psi_a \in \mathcal{P}_d(X)} 1 \leq X^{d-2+1/k} \sum_{0 < |a_d| \leq X} \frac{1}{D} \sum_{D \leq X} \sum_{s \leq DB(X,T) \gcd(s,a_d)^{\beta_{d,k}}} \frac{1}{s^{1/k}}.
\]

Recalling the definition (2.31) of the quantity \( B(X,T) \) and appealing to the upper bound (2.29) with \( A(X,T) \) replaced by \( B(X,T) \), we derive
\[
\sum_{\psi_a \in \mathcal{P}_d(X)} 1 \leq X^{d-2+1/k+\delta a_{d,k}/2+\varepsilon T^d a_{d,k}} \sum_{0 < |a_d| \leq X} \sum_{D \leq X} \gcd(D,a_d)^{(k-1)\beta_{d,k}/k} \frac{1}{D^{1+1/k}}.
\]

But using the upper bound (2.28) we easily get
\[
\sum_{0 < |a_d| \leq X} \sum_{D \leq X} \frac{\gcd(D,a_d)^{(k-1)\beta_{d,k}/k}}{D^{1+1/k}} \ll X.
\]

We thus see that the upper bound (2.30) follows, which completes the proof. \( \square \)

3. The uniform boundedness conjecture on average

Our goal in this section is to furnish the proof of Theorem 1.1. The following result is a direct consequence of the work of Benedetto [3] and provides us with a convenient bound for the number of rational preperiodic points of a polynomial \( \psi_a \in \mathcal{P}_d. \)
Lemma 3.1. Let $d \geq 2$ and $\psi_a \in \mathcal{P}_d$. We have

$$\#\text{Prep}_Q(\psi_a) \ll 1 + \log \mathcal{H}(\psi_a),$$

where the implied constant depends at most on $d$.

Proof. The set of prime numbers where the polynomial $\psi_a$ does not have potentially good reduction (see [3, Definition 2.1]) is contained in the set of prime divisors of $a_0a_d$. Therefore, letting $\omega(n)$ denote the number of prime numbers dividing an integer $n \geq 1$ and appealing to Benedetto’s result [3, Main Theorem], we get

$$\#\text{Prep}_Q(\psi_a) \ll 1 + \omega(a_0|a_d|) \log (1 + \omega(a_0|a_d|)).$$

Invoking the classical upper bound

$$\omega(n) \ll \frac{\log n}{\log \log n},$$

we deduce

$$\#\text{Prep}_Q(\psi_a) \ll 1 + \log(a_0|a_d|).$$

Since $a_0, |a_d| \leq \mathcal{H}(\psi_a)$, we see that this completes the proof. \hfill \Box

We are now ready to reveal the proof of Theorem 1.1.

Proof of Theorem 1.1. We start by handling the case $d = 2$. It follows from Lemma 2.4 that if a polynomial $\psi_a \in \mathcal{P}_2$ has a rational preperiodic point then the integer $a_0$ must be squareful. As a result, combining Lemma 3.1 with the upper bound (2.27), we get

$$\sum_{\psi_a \in \mathcal{P}_2(X)} \#\text{Prep}_Q(\psi_a) \ll X^{3/2} \log X.$$

Recalling the estimate (2.20) we thus derive

$$\frac{1}{\#\mathcal{P}_2(X)} \sum_{\psi_a \in \mathcal{P}_2(X)} \#\text{Prep}_Q(\psi_a) \ll \frac{\log X}{X^{1/2}},$$

which completes the proof of Theorem 1.1 in the case $d = 2$.

We now deal with the case $d \geq 3$. Given $\delta \in (0, 1)$, recall the definition (2.21) of the set $\mathcal{P}_d^{(\delta)}(X)$. Combining Lemmas 2.10 and 3.1 we deduce that for any $\delta \in (0, 1)$, we have

$$\sum_{\psi_a \in \mathcal{P}_d(X)} \#\text{Prep}_Q(\psi_a) = \sum_{\psi_a \in \mathcal{P}_d^{(\delta)}(X)} \#\text{Prep}_Q(\psi_a) + O \left( X^{d-\delta+\epsilon} \right).$$

Applying Lemma 2.11 with $T = 1$, we see that for $k \in \{2, 3\}$, we have

$$\sum_{\psi_a \in \mathcal{P}_d(X)} \#\text{Prep}_Q(\psi_a) \ll X^{d+\epsilon} \left( \frac{1}{X^{1-1/k-\delta d,k/2}} + \frac{1}{X^{\delta}} \right).$$

An easy calculation shows that the optimal choice of $\delta \in (0, 1)$ is given by

$$\delta = \frac{2(k-1)(d+k-2)}{k^2(d+1)-2k+d-1}.$$

Choosing $k = 3$ and using the estimate (2.20), we see that this completes the proof of Theorem 1.1 in the case $d \geq 3$. \hfill \Box
4. DYNAMICAL SUCCESSIVE MINIMA OF RANDOM POLYNOMIALS

4.1. A statistical version of the dynamical Lang conjecture. Our aim in this section is to establish Theorem 1.2. The following result provides us with a sharp upper bound for the quantity $\lambda_1(\psi_a)$ under a mild assumption on the polynomial $\psi_a \in \mathcal{P}_d$.

**Lemma 4.1.** Let $d \geq 2$ and $\psi_a \in \mathcal{P}_d$. If $0 \notin \text{Prep}_Q(\psi_a)$ then

$$\lambda_1(\psi_a) \leq \frac{1}{d(d-1)} \log \mathcal{H}(\psi_a) + \frac{\log d}{d(d-1)}.$$

**Proof.** An application of Lemma 2.8 gives

$$\hat{h}_{\psi_a}(1) \leq \frac{1}{d-1} \log \mathcal{H}(\psi_a) + \frac{\log d}{d(d-1)}.$$

Therefore, the equalities $\psi_a(0) = 1$ and (1.5) yield

$$\hat{h}_{\psi_a}(0) \leq \frac{1}{d(d-1)} \log \mathcal{H}(\psi_a) + \frac{\log d}{d(d-1)}.$$

which completes the proof since by assumption $0 \notin \text{Prep}_Q(\psi_a)$. $\square$

We now have all the tools required to prove Theorem 1.2.

**Proof of Theorem 1.2.** Appealing to Lemma 4.1 we get

$$\# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_1(\psi_a) > \left( \frac{1}{d(d-1)} + \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} \leq \sum_{\psi_a \in \mathcal{P}_d(X)} \# \text{Prep}_Q(\psi_a).$$

It follows in particular that

$$\# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_1(\psi_a) > \left( \frac{1}{d(d-1)} + \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} \ll \sum_{\psi_a \in \mathcal{P}_d(X)} \# \text{Prep}_Q(\psi_a).$$

Therefore, an application of Theorem 1.1 yields

$$\lim_{X \to \infty} \frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_1(\psi_a) > \left( \frac{1}{d(d-1)} + \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} = 0.$$

We thus conclude that in order to complete the proof of Theorem 1.2, it suffices to prove that for any $\varepsilon \in (0, 1/d^2)$, we have

$$\lim_{X \to \infty} \frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_1(\psi_a) < \left( \frac{1}{d(d-1)} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} = 0. \quad (4.1)$$

We start by handling the case $d = 2$. Invoking Lemma 2.4 we deduce that

$$\# \left\{ \psi_a \in \mathcal{P}_2(X) : \lambda_1(\psi_a) < \left( \frac{1}{2} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} \leq \sum_{\psi_a \in \mathcal{P}_2(X)} \frac{1}{X^{1/2}}.$$

Summing first over the squareful part $a_0/s_2(a_0)$ of the integer $a_0$ using the upper bound (2.27), we get

$$\# \left\{ \psi_a \in \mathcal{P}_2(X) : \lambda_1(\psi_a) < \left( \frac{1}{2} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} \ll X^{3/2} \sum_{s \leq X^{1/2}} \frac{1}{s^{1/2}}.$$

Appealing to the estimate (2.20) we thus derive

$$\frac{1}{\# \mathcal{P}_2(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_2(X) : \lambda_1(\psi_a) < \left( \frac{1}{2} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} \ll \frac{1}{X^\varepsilon}.$$

The equality (4.1) follows, which completes the proof of Theorem 1.2 in the case $d = 2$. 

ARITHMETIC DYNAMICS OF RANDOM POLYNOMIALS 19
We now deal with the case \( d \geq 3 \). An application of Lemma 2.10 shows that for any \( \delta \in (0, 1) \), we have
\[
\# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_1(\psi_a) < \left( \frac{1}{d(d-1)} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} = \# \mathcal{E}_{d,\varepsilon}^{(\delta)}(X) + O \left( X^{d-\delta+\varepsilon} \right),
\]
(4.2)
where we have set
\[
\mathcal{E}_{d,\varepsilon}^{(\delta)}(X) = \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_1(\psi_a) < \left( \frac{1}{d(d-1)} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\}.
\]
Moreover, it is clear that for any \( \psi_a \in \mathcal{E}_{d,\varepsilon}^{(\delta)}(X) \), we have
\[
\# \left\{ z \in \mathbb{Q} : \hat{h}_{\psi_a}(z) < \left( \frac{1}{d(d-1)} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} \geq 1.
\]
It follows that
\[
\# \mathcal{E}_{d,\varepsilon}^{(\delta)}(X) \leq \sum_{\psi_a \in \mathcal{P}_d(X)} \# \left\{ z \in \mathbb{Q} : \hat{h}_{\psi_a}(z) < \left( \frac{1}{d(d-1)} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\},
\]
which implies in particular
\[
\# \mathcal{E}_{d,\varepsilon}^{(\delta)}(X) \leq \sum_{\psi_a \in \mathcal{P}_d(X)} \# \left\{ z \in \mathbb{Q} : \hat{h}_{\psi_a}(z) < \log X^{1/(d(d-1)-\varepsilon)} \right\}.
\]
Appealing to Lemma 2.11 with \( k = 2 \), we thus get
\[
\# \mathcal{E}_{d,\varepsilon}^{(\delta)}(X) \ll X^{d-\varepsilon(-1+d(d-1)/2)+\delta(d-1)/4d}.
\]
Since \( \varepsilon \in (0, 1/d^2) \) we can choose \( \delta = \varepsilon d^2 \). Therefore, recalling the estimate (4.2), we obtain
\[
\# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_1(\psi_a) < \left( \frac{1}{d(d-1)} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} \ll X^{d-\varepsilon(-1+d(d-1)/4)}.
\]
As a result, using the estimate (2.20) we eventually derive
\[
\frac{1}{\# \mathcal{P}_d(X)} \# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_1(\psi_a) < \left( \frac{1}{d(d-1)} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} \ll \frac{1}{X^{\varepsilon(-1+d(d-1)/4)}}.
\]
The equality (4.1) thus follows, which completes the proof of Theorem 1.2 in the case \( d \geq 3 \).

4.2. Dynamically independent vectors of rational numbers. In this section we say that two polynomials \( f, g \in \mathbb{Q}[z] \) are conjugate if there exists \((\alpha, \beta) \in \mathbb{Q}^* \times \mathbb{Q} \) such that
\[
g(z) = \frac{f(\alpha z + \beta)}{\alpha} - \frac{\beta}{\alpha}.
\]
In addition, given \( \ell \geq 1 \) we follow Medvedev and Scanlon [16, Definition 2.23] and we define the \( \ell \)-th Chebyshev polynomial \( C_{\ell} \in \mathbb{Z}[z] \) as the unique polynomial satisfying the functional equation
\[
C_{\ell} \left( z + \frac{1}{z} \right) = z^\ell + \frac{1}{z^\ell}.
\]
(4.3)
Furthermore, we call a polynomial of the form \(-C_\ell\) for some \( \ell \geq 1 \) a negative Chebyshev polynomial. Our first task is to establish the following result.

**Lemma 4.2.** Let \( d \geq 2 \) and \( \psi_a \in \mathcal{P}_d \). If \( 0 \notin \text{Prep}_Q(\psi_a) \) then no iterate of \( \psi_a \) is conjugate to a monomial, a Chebyshev polynomial or a negative Chebyshev polynomial.
Proof. We start by recalling that for any \( m \geq 1 \) the polynomial \( \psi^m_a \) has degree \( d^m \). We reason by contradiction and we thus assume that for some \( n \geq 1 \) and for some polynomial \( \varphi(z) \in \{ z^{d^n}, C_{d^n}(z), -C_{d^n}(z) \} \), there exists \( (\alpha, \beta) \in \mathbb{Q}^2 \) such that

\[
\psi^n_a(z) = \frac{\varphi(\alpha z + \beta)}{\alpha} - \frac{\beta}{\alpha}.
\]

But it is immediate to check by induction that the coefficient of degree \( d^n - 1 \) of the polynomial \( \psi^n_a(z) \) is equal to 0. Since \( \varphi \) also has this property, we deduce that we must have \( \beta = 0 \) and therefore

\[
\psi^n_a(z) = \frac{\varphi(\alpha z)}{\alpha}.
\]

If either \( d \) is odd or \( \varphi(z) = z^{d^n} \) then \( \varphi(0) = 0 \) and thus also \( \psi^n_a(0) = 0 \), which contradicts our assumption that \( 0 \not\in \text{Prep}_Q(\psi_a) \). Next, if \( d \) is even then the polynomial \( C_{d^n} \) is even and is therefore conjugate to \( -C_{d^n} \). As a result, we can now assume that \( d \) is even and \( \varphi(z) = C_{d^n}(z) \). Observing that it follows from the definition (4.3) of the polynomial \( C_{d^n} \) that \( C_{d^n}(0) = 2(-1)^{d^n/2} \), we see that the equality (4.4) shows that we must have

\[
\psi^n_a(z) = (-1)^{d^n/2} \psi^n_a(0) \left( \frac{2}{\psi^n_a(0)} z \right).
\]

In addition, the definition (4.3) of the polynomial \( C_{d^n} \) also gives \( C_{d^n}(2) = 2 \), which yields

\[
\psi^{2n}_a(0) = (-1)^{d^n/2} \psi^n_a(0).
\]

But the equality (4.5) shows that the polynomial \( \psi^n_a(z) \) is even, so composing the equality (4.6) by \( \psi^n_a \) we eventually obtain \( \psi^{3n}_a(0) = \psi^{2n}_a(0) \). Once again this contradicts our assumption that \( 0 \not\in \text{Prep}_Q(\psi_a) \), which completes the proof.

Given a polynomial \( \psi_a \in \mathcal{P}_d \) satisfying a mild assumption, the following result allows us to exhibit many \( \psi_a \)-dynamically independent vectors of rational numbers by appealing to the work of Medvedev and Scanlon [16]. This will play a key role in the proof of the upper bound in Proposition 4.4.

**Lemma 4.3.** Let \( d \geq 2 \) and \( \psi_a \in \mathcal{P}_d \). Let also \( N \geq 1 \) and \( q_1, \ldots, q_N \) be distinct prime numbers coprime to \( a_0a_d \). If \( 0 \not\in \text{Prep}_Q(\psi_a) \) then the vector

\[
\left( \frac{1}{q_1}, \ldots, \frac{1}{q_N} \right)
\]

is \( \psi_a \)-dynamically independent.

**Proof.** We reason by contradiction and we thus assume that there exists a nonzero polynomial \( P \in \mathbb{Q}[X_1, \ldots, X_N] \) satisfying the property (1.7) and

\[
P \left( \frac{1}{q_1}, \ldots, \frac{1}{q_N} \right) = 0.
\]

We introduce the algebraic set

\[
W = \{ (\xi_1, \ldots, \xi_N) \in A^N(\mathbb{Q}) : P(\xi_1, \ldots, \xi_N) = 0 \},
\]

and we let \( V_0 \) be an irreducible component of \( W \) such that

\[
\left( \frac{1}{q_1}, \ldots, \frac{1}{q_N} \right) \in V_0.
\]

We also let

\[
\Psi_a = (\psi_a, \ldots, \psi_a) \in \mathbb{Q}[z]^N.
\]
The fact that the polynomial $P$ satisfies the property (1.7) is equivalent to saying that
\[
\Psi_a(W) \subseteq W.
\]
Therefore, since the set $\Psi_a(V_0)$ is irreducible there exists an irreducible component $V_1$ of $W$ such that
\[
\Psi_a(V_0) \subseteq V_1.
\]
Iterating this process we recursively define a sequence $(V_j)_{j \geq 0}$ of irreducible components of $W$. Next, we select $n \geq 0$ and $m > n$ such that $V_m = V_n$. By construction, the algebraic set $V_n$ is irreducible and satisfies
\[
\Psi_{a-n}(V_n) \subseteq V_n.
\]
In addition, by assumption we have $0 \notin \text{Prep}_Q(\psi_a)$ so it follows from Lemma 4.2 that the polynomial $\psi^{m-n}_a$ is not conjugate to a monomial, a Chebyshev polynomial or a negative Chebyshev polynomial. As a result, we are in position to appeal to the work of Medvedev and Scanlon [16, Theorem]. We deduce that there exist $R \geq 1$ and $(i_1, j_1), \ldots, (i_R, j_R) \in \{1, \ldots, N\}^2$, and also $g_1, \ldots, g_R \in \mathbb{Q}[z]$ commuting with $\psi^{m-n}_a$ such that
\[
V_n = \{ (\xi_1, \ldots, \xi_N) \in \mathbb{A}^N(\overline{\mathbb{Q}}) : \forall r \in \{1, \ldots, R\} \xi_{i_r} = g_r(\xi_{j_r}) \}. \tag{4.8}
\]
Let $r \in \{1, \ldots, R\}$. Recalling that $V_0$ satisfies the assumption (4.7), we see that by construction we have
\[
\Psi_a(\frac{1}{q_1}, \ldots, \frac{1}{q_N}) \subseteq V_n.
\]
Hence, since the polynomials $g_r$ and $\psi^{m-n}_a$ commute, for any $L \geq 0$ we have
\[
\psi_a^{L(m-n)+n}\left(\frac{1}{q_{i_r}}\right) = g_r(\psi_a^{L(m-n)+n}\left(\frac{1}{q_{j_r}}\right)). \tag{4.9}
\]
But by assumption $q_{i_r}$ is coprime to $a_0a_q$ so it is straightforward to check that for any $s \geq 0$, we have
\[
\left|\psi_a^s\left(\frac{1}{q_{i_r}}\right)\right|_{q_{i_r}} = q_{i_r}^{d_s}.
\tag{4.10}
\]
It follows that for any $L \geq 0$, we have
\[
\left|g_r\left(\psi_a^{L(m-n)+n}\left(\frac{1}{q_{j_r}}\right)\right)\right|_{q_{i_r}} = q_{i_r}^{d(L(m-n)+n)}.
\]
Since $m > n$ this implies in particular that
\[
\lim_{L \to \infty} \left|\psi_a^{L(m-n)+n}\left(\frac{1}{q_{i_r}}\right)\right|_{q_{i_r}} = \infty.
\]
Using again the assumption that $q_{i_r}$ is coprime to $a_0$ we deduce that $q_{i_r} = q_{j_r}$, and therefore $i_r = j_r$ since the prime numbers $q_1, \ldots, q_N$ are distinct. The equality (4.9) thus shows that
\[
\left\{ \psi_a^{L(m-n)+n}\left(\frac{1}{q_{i_r}}\right) : L \geq 0 \right\} \subseteq \{ z \in \mathbb{Q} : z - g_r(z) = 0 \}.
\]
But the fact that $m > n$ and the equality (4.10) imply that the left-hand side is an infinite set, so we finally obtain $g_r(z) = z$.

We have thus proved that for any $r \in \{1, \ldots, R\}$, we have $i_r = j_r$ and $g_r(z) = z$. Recalling the equality (4.8) we see that it eventually follows that $V_n = \mathbb{A}^N(\overline{\mathbb{Q}})$ and therefore $W = \mathbb{A}^N(\overline{\mathbb{Q}})$. This contradicts the fact that the polynomial $P$ is nonzero, which completes the proof. \qed
4.3. Dynamical successive minima of higher order. Our purpose in this section is to establish Propositions 4.4 and 4.5 and Theorem 1.4. Given \( N \geq 1 \) and \( \psi_a \in \mathcal{P}_d \), recall that the \( N \)-th dynamical successive minimum \( \lambda_N(\psi_a) \) of the polynomial \( \psi_a \) was introduced in Definition 1.3. Our first task is to complement Lemma 4.1 by providing upper and lower bounds for the quantity \( \lambda_N(\psi_a) \) for \( N \geq 2 \) under a mild assumption on the polynomial \( \psi_a \in \mathcal{P}_d \).

**Proposition 4.4.** Let \( d \geq 3 \) and \( \psi_a \in \mathcal{P}_d \). Let also \( N \geq 2 \) and \( \varepsilon > 0 \). If \( 0 \notin \text{Prep}_d(\psi_a) \) then

\[
\frac{1}{2} \log \frac{N}{\mathcal{H}(\psi_a)} + O(1) < \lambda_N(\psi_a) \leq (1 + \varepsilon) \log N + \left( \frac{1}{d-1} + \varepsilon \right) \log \mathcal{H}(\psi_a) + O(1),
\]

where the implied constants depend at most on \( d \) and \( \varepsilon \).

**Proof.** We start by proving the lower bound. Let \((z_1, \ldots, z_N) \in \mathbb{Q}^N\) be a \( \psi_a \)-dynamically independent vector. It is clear that for any \( i, j \in \{1, \ldots, N\} \) such that \( i \neq j \) we have \( z_i \neq z_j \). As a result, since

\[
\# \left\{ z \in \mathbb{Q} : h(z) \leq \frac{1}{2} \log N - \log 2 \right\} < N,
\]

it follows that there exists \( i_0 \in \{1, \ldots, N\} \) such that

\[
h(z_{i_0}) > \frac{1}{2} \log N - \log 2.
\]

Therefore, Lemma 2.6 yields

\[
\hat{h}_{\psi_a}(z_{i_0}) > \frac{1}{2} \log N - \frac{1}{2} \log \mathcal{H}(\psi_a) - \log d - \log 2,
\]

which completes the proof of the claimed lower bound.

We now turn to the proof of the upper bound. We start by noting that Lemma 2.8 shows that for any prime number \( p \), we have

\[
\hat{h}_{\psi_a}(\frac{1}{p}) \leq \frac{1}{d-1} \log \mathcal{H}(\psi_a) + \log p + \frac{\log d}{d-1}.
\] (4.11)

By assumption we have \( 0 \notin \text{Prep}_d(\psi_a) \) so we are in position to appeal to Lemma 4.3. Recall that \( \omega(n) \) denotes the number of prime numbers dividing an integer \( n \geq 1 \). In addition, for any \( m \geq 1 \) we let \( p_m \) be the \( m \)-th prime number. Combining Lemma 4.3 and the inequality (4.11) we deduce that

\[
\lambda_N(\psi_a) \leq \frac{1}{d-1} \log \mathcal{H}(\psi_a) + \log p_{N+\omega(a_0|a_d|)} + \frac{\log d}{d-1}.
\] (4.12)

Furthermore, using Chebyshev’s result stating that for any \( m \geq 1 \) we have \( p_{m+1} \leq 2p_m \), we derive

\[
p_{N+\omega(a_0|a_d|)} \leq 2^{\omega(a_0|a_d|)} p_N.
\]

Hence, the divisor bound (2.26) and Chebyshev’s upper bound \( p_m \ll m \log m \) yield in particular

\[
p_{N+\omega(a_0|a_d|)} \ll \mathcal{H}(\psi_a)^{\varepsilon} N^{1+\varepsilon}.
\]

Recalling the upper bound (4.12), we see that the claimed upper bound follows, which completes the proof. \( \square \)

It may be worth pointing out that the same proof shows that the upper bound in Proposition 4.4 also holds if \( d = 2 \).

Given \( d \geq 3 \) and \( \psi_a \in \mathcal{P}_d \) satisfying \( 0 \notin \text{Prep}_d(\psi_a) \), Proposition 4.4 implies in particular that whenever \( N > \mathcal{H}(\psi_a)^{1+\kappa} \) for some \( \kappa > 0 \), we have

\[
\log N \ll \lambda_N(\psi_a) \ll \log N,
\]
where the implied constants depend at most on $d$ and $\kappa$.

We now proceed to study the much harder situation where $\log N$ is small compared to $\log \mathcal{H}(\psi_a)$. For this purpose, for any $t \geq 0$ we define the integer
\[ N_t(\psi_a) = 1 + \left\lfloor \frac{\log \mathcal{H}(\psi_a)}{t} \right\rfloor. \tag{4.13} \]

The following result is the culmination of our statistical investigation of the size of dynamical successive minima of polynomials.

**Proposition 4.5.** Let $d \geq 3$ and $\varepsilon > 0$. Let also $t \geq 0$. We have
\[ \lim_{X \to \infty} \frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : \left( \frac{1 + 2t}{d + 3} - \varepsilon \right) \log \mathcal{H}(\psi_a) < \lambda_{N_t(\psi_a)}(\psi_a) \right\} = 1. \]

**Proof.** Our goal is to prove that for any $\varepsilon \in (0, 1/2d)$, we have
\[ \lim_{X \to \infty} \frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_{N_t(\psi_a)}(\psi_a) \leq \left( \frac{1 + 2t}{d + 3} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} = 0. \tag{4.14} \]

First, we note that Lemma 2.10 shows that for any $\delta \in (0, 1)$, we have
\[ \# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_{N_t(\psi_a)}(\psi_a) \leq \left( \frac{1 + 2t}{d + 3} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} = \# \mathcal{J}^{(d)}_{d, \varepsilon, t}(X) \]
\[ + O \left( X^{d-\delta+\varepsilon} \right), \]

where we have set
\[ \mathcal{J}^{(d)}_{d, \varepsilon, t}(X) = \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_{N_t(\psi_a)}(\psi_a) \leq \left( \frac{1 + 2t}{d + 3} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\}. \]

Next, we observe that by Definition 1.3, for any $\psi_a \in \mathcal{J}^{(d)}_{d, \varepsilon, t}(X)$ we clearly have
\[ \# \left\{ z \in \mathbb{Q}^x : \hat{h}_{\psi_a}(z) \leq \left( \frac{1 + 2t}{d + 3} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} \geq N_t(\psi_a) - 1. \]

Recalling the definition (4.13) of the quantity $N_t(\psi_a)$, we thus see that
\[ \# \mathcal{J}^{(d)}_{d, \varepsilon, t}(X) \ll \sum_{\psi_a \in \mathcal{P}_d(X)} \frac{1}{\mathcal{H}(\psi_a)^t} \cdot \# \left\{ z \in \mathbb{Q}^x : \hat{h}_{\psi_a}(z) \leq \left( \frac{1 + 2t}{d + 3} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\}. \]

We now proceed to prove that
\[ \# \mathcal{J}^{(d)}_{d, \varepsilon, t}(X) \ll X^{d-\varepsilon(d-1)/4 + \delta(d-1)/4d}. \tag{4.16} \]

We handle separately the cases $t = 0$ and $t > 0$. First, we see that in the case $t = 0$ we have
\[ \# \mathcal{J}^{(d)}_{d, \varepsilon, 0}(X) \ll \sum_{\psi_a \in \mathcal{P}_d^{(d)}(X)} \# \left\{ z \in \mathbb{Q}^x : \hat{h}_{\psi_a}(z) \leq \log X^{1/(d+3)\varepsilon} \right\}. \]

As a result, an application of Lemma 2.11 with $k = 2$ gives
\[ \# \mathcal{J}^{(d)}_{d, \varepsilon, 0}(X) \ll X^{d-\varepsilon(d+1)/2 + \delta(d-1)/4d}, \]

and the upper bound (4.16) in the case $t = 0$ follows. We now deal with the case $t > 0$. An application of partial summation gives
\[ \# \mathcal{J}^{(d)}_{d, \varepsilon, t}(X) \ll \int_1^\infty \sum_{\psi_a \in \mathcal{P}_d^{(d)}(X) : \mathcal{H}(\psi_a)^t < u} \# \left\{ z \in \mathbb{Q}^x : \hat{h}_{\psi_a}(z) \leq \left( \frac{1 + 2t}{d + 3} - \varepsilon \right) \log \mathcal{H}(\psi_a) \right\} \frac{du}{u^2}. \]
We deduce that
\[\# \mathcal{G}_{d, t}^{(d)}(X) \ll \int_{\mathbb{Q}_d} \sum_{\psi_a \in \mathcal{G}_{d, t}^{(d)}(X)} \# \left\{ z \in \mathbb{Q}_d^\infty : \hat{h}_\psi(z) \leq \log \left( X^{1/(d+3) - \epsilon/2u^2/(d+3) - \epsilon/2t} \right) \right\} \frac{du}{u^2}.\]

Therefore, an application of Lemma 2.11 with \( k = 2 \) shows that for \( t > 0 \), we have
\[\# \mathcal{G}_{d, t}^{(d)}(X) \ll X^{d-\epsilon(d-1)/4 + \delta(d-1)/4d} \int_{\mathbb{Q}_d} \frac{du}{u^{1+\epsilon(d-1)/4t}},\]
and the upper bound (4.16) in the case \( t > 0 \) follows.

In addition, since \( \epsilon \in (0, 1/2d) \) we can choose \( \delta = \epsilon d/2 \). As a result, recalling the estimate (4.15), we see that
\[\# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_{N_d}(\psi_a) \leq \left( \frac{1}{d+3} - \epsilon \right) \log \mathcal{H}(\psi_a) \right\} \ll X^{d-\epsilon(d-1)/8}.\]

Appealing to the estimate (2.20), we eventually get
\[\frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_{N_d}(\psi_a) \leq \left( \frac{1}{d+3} - \epsilon \right) \log \mathcal{H}(\psi_a) \right\} \leq \frac{1}{X^{\epsilon(d-1)/8}}.\]
The equality (4.14) thus follows, which completes the proof. \( \square \)

We now proceed to show that Theorem 1.4 is a direct consequence of Theorems 1.1 and 1.2 and Propositions 4.4 and 4.5.

**Proof of Theorem 1.4.** To start with, we note that the upper bound in Proposition 4.4 shows that
\[\# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_2(\psi_a) \leq \left( \frac{1}{d-1} + \epsilon \right) \log \mathcal{H}(\psi_a) \right\} \leq \sum_{\psi_a \in \mathcal{P}_d(X)} 1 + O(1).\]

It follows in particular that
\[\# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_2(\psi_a) \leq \left( \frac{1}{d-1} + \epsilon \right) \log \mathcal{H}(\psi_a) \right\} \leq \sum_{\psi_a \in \mathcal{P}_d(X)} \# \text{Prep}_Q(\psi_a).\]

Appealing to Theorem 1.1 we thus obtain
\[\lim_{X \to \infty} \frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : \lambda_2(\psi_a) > \left( \frac{1}{d-1} + \epsilon \right) \log \mathcal{H}(\psi_a) \right\} = 0.\]
Therefore, an application of Theorem 1.2 yields
\[\lim_{X \to \infty} \frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : \frac{\lambda_2(\psi_a)}{\lambda_1(\psi_a)} \leq d + \epsilon \right\} = 1. \quad (4.17)\]
In addition, choosing \( t = 0 \) in Proposition 4.5 and appealing again to Theorem 1.2, we derive
\[\lim_{X \to \infty} \frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : \frac{d(d-1)}{d+3} - \epsilon < \frac{\lambda_2(\psi_a)}{\lambda_1(\psi_a)} \right\} = 1.\]
As a result, noting that
\[\frac{d(d-1)}{d+3} = d \left( 1 - \frac{4}{d+3} \right),\]
we see that we have obtained the first part of Theorem 1.4.

Next, we remark that proceeding exactly as in the proof of the equality (4.17) we get
\[\lim_{X \to \infty} \frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : \frac{\lambda_N(\psi_a)}{\lambda_1(\psi_a)} \leq d + \epsilon \right\} = 1.\]
Appealing to the first part of Theorem 1.4 and noting that
\[
\left(1 - \frac{4}{d+3}\right)^{-1} = 1 + \frac{4}{d-1},
\]
we thus derive
\[
\lim_{X \to \infty} \frac{1}{\# \mathcal{P}_d(X)} \cdot \# \left\{ \psi_a \in \mathcal{P}_d(X) : \frac{\lambda_N(\psi_a)}{\lambda_2(\psi_a)} \leq 1 + \frac{4}{d-1} + \varepsilon \right\} = 1.
\]
Since we clearly have \(\lambda_N(\psi_a) \geq \lambda_2(\psi_a)\) for any \(\psi_a \in \mathcal{P}_d\), we see that this completes the proof of Theorem 1.4. □

References
1. M. Baker and L. De Marco, *Special curves and postcritically finite polynomials*, Forum Math. Pi 1 (2013), e3, 35.
2. R. Benedetto, P. Ingram, R. Jones, M. Manes, J. H. Silverman, and T. J. Tucker, *Current trends and open problems in arithmetic dynamics*, Bull. Amer. Math. Soc. (N.S.) 56 (2019), no. 4, 611–685.
3. R. L. Benedetto, *Preperiodic points of polynomials over global fields*, J. Reine Angew. Math. 608 (2007), 123–153.
4. N. G. de Bruijn, *On the number of integers \(\leq x\) whose prime factors divide \(n\)*, Illinois J. Math. 6 (1962), 137–141.
5. N. Fakhruddin, *Questions on self maps of algebraic varieties*, J. Ramanujan Math. Soc. 18 (2003), no. 2, 109–122.
6. E. V. Flynn, B. Poonen, and E. F. Schaefer, *Cycles of quadratic polynomials and rational points on a genus-2 curve*, Duke Math. J. 90 (1997), no. 3, 435–463.
7. M. Hindry and J. H. Silverman, *The canonical height and integral points on elliptic curves*, Invent. Math. 93 (1988), no. 2, 419–450.
8. P. Ingram, *Lower bounds on the canonical height associated to the morphism \(\phi(z) = z^d + c\)*, Monatsh. Math. 157 (2009), no. 1, 69–89.
9. , *Canonical heights and preperiodic points for certain weighted homogeneous families of polynomials*, Int. Math. Res. Not. IMRN (2019), no. 15, 4859–4879.
10. S. Lang, *Elliptic curves: Diophantine analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 231, Springer-Verlag, Berlin-New York, 1978.
11. P. Le Boudec, *A statistical view on the conjecture of Lang about the canonical height on elliptic curves*, Trans. Amer. Math. Soc. 372 (2019), no. 12, 8347–8361.
12. N. Looper, *A lower bound on the canonical height for polynomials*, Math. Ann. 373 (2019), no. 3–4, 1057–1074.
13. , *Dynamical uniform boundedness and the abc-conjecture*, Invent. Math. 225 (2021), no. 1, 1–44.
14. , *The uniform boundedness and dynamical Lang conjectures for polynomials*, arXiv:2105.05240v2 (2021).
15. B. Mazur, *Modular curves and the Eisenstein ideal*, Inst. Hautes Études Sci. Publ. Math. (1977), no. 47, 33–186 (1978), With an appendix by Mazur and M. Rapoport.
16. A. Medvedev and T. Scanlon, *Invariant varieties for polynomial dynamical systems*, Ann. of Math. (2) 179 (2014), no. 1, 81–177.
17. P. Morton, *Arithmetic properties of periodic points of quadratic maps*, Acta Arith. 62 (1992), no. 4, 343–372.
18. P. Morton and J. H. Silverman, *Rational periodic points of rational functions*, Internat. Math. Res. Notices (1994), no. 2, 97–110.
19. C. Panraksa, *Rational periodic points of \(x^d + c\) and Fermat-Catalan equations*, arXiv:2105.03715v6 (2021).
20. B. Poonen, *The classification of rational preperiodic points of quadratic polynomials over \(\mathbb{Q}\): a refined conjecture*, Math. Z. 228 (1998), no. 1, 11–29.
21. J. H. Silverman, *Lower bound for the canonical height on elliptic curves*, Duke Math. J. 48 (1981), no. 3, 633–648.
22. , *The arithmetic of dynamical systems*, Graduate Texts in Mathematics, vol. 241, Springer, New York, 2007.
23. ———, *Moduli spaces and arithmetic dynamics*, CRM Monograph Series, vol. 30, American Mathematical Society, Providence, RI, 2012.

**Departement Mathematik und Informatik, Fachbereich Mathematik, Spiegelgasse 1, 4051 Basel, Switzerland**

*Email address: pierre.leboudec@unibas.ch*

**Harvard University, Department of Mathematics, Science Center Room 325, 1 Oxford Street, Cambridge, MA 02138, USA**

*Email address: mavraki@math.harvard.edu*