IRREDUCIBILITY OF RANDOM POLYNOMIALS OF LARGE DEGREE

EMMANUEL BREUILLARD AND PÉTER P. VARJÚ

Abstract. We consider random polynomials with independent identically distributed coefficients with a fixed law. Assuming the Riemann hypothesis for Dedekind zeta functions, we prove that such polynomials are irreducible and their Galois groups contain the alternating group with high probability as the degree goes to infinity. This settles a conjecture of Odlyzko and Poonen conditionally on RH for Dedekind zeta functions.

1. Introduction

Let

$$P(x) = x^d + A_{d-1}x^{d-1} + \ldots + A_1 x + 1 \in \mathbb{Z}[x] \tag{1.1}$$

be a random polynomial with independent coefficients $A_1, \ldots, A_{d-1}$ taking values in 0 and 1 with equal probability. Odlyzko and Poonen [30] conjectured that the probability that P is irreducible in $\mathbb{Z}[x]$ converges to 1 as $d \to \infty$.

The best known lower bound in this problem is due to Konyagin [25] who proved that

$$\mathbb{P}(P \text{ is irreducible}) > \frac{c}{\log d}$$

for an absolute constant $c > 0$.

A strongly related problem was studied by Bary-Soroker and Kozma [3], who proved that

$$\mathbb{P}(x^d + A_{d-1}x^{d-1} + \ldots + A_1 x + A_0 \text{ is irreducible}) \to 1,$$

where $A_0, \ldots, A_{d-1}$ are independent random integers uniformly distributed in $1, \ldots, L$ for a fixed integer L that has at least 4 distinct prime divisors.

In another paper, Bary-Soroker and Kozma [4] studied the problem for bivariate polynomials. See also [31] for a study of the probability that a random polynomial has low degree factors, and [6] for computational experiments on related problems.

2010 *Mathematics Subject Classification.* 11C08 (primary) and 11M41, 60J10 (secondary).

Key words and phrases. random polynomials, irreducibility, Riemann hypothesis, Dedekind zeta function, Markov chains.

EB acknowledges support from ERC Grant no. 617129 'GeTeMo'; PV acknowledges support from the Royal Society.

In this paper we prove the following result.

Theorem 1. *Let P be a random polynomial as in (1.1). Suppose that the Riemann hypothesis holds for the Dedekind zeta function ζ_K for all number fields of the form $K = \mathbb{Q}(a)$, where a is a root of a polynomial with $0, 1$ coefficients.*

Then

$$\mathbb{P}(P \text{ is irreducible in } \mathbb{Z}[x]) \to 1$$

as $d \to \infty$.

See Section 1.2 for more precise results, where we discuss the following finer aspects of the problem

- random polynomials with arbitrary i.i.d. coefficients,
- the rate at which the probability converges to 1,
- relaxation of the assumption of RH,
- Galois groups.

1.1. **Motivation.** Beyond its intrinsic interest, the problem of irreducibility of random polynomials of high degree is motivated by some other problems, which we now briefly discuss.

It is believed to be computationally difficult to determine the prime factorization of integers. On the other hand, polynomial time algorithms are known for computing the factorization of polynomials in $\mathbb{Z}[x]$. Given an integer $N \in \mathbb{Z}_{>0}$, we can write it as $N = P(2)$ for a unique polynomial P with $0, 1$-coefficients. By computing the factorization of P in $\mathbb{Z}[x]$ and evaluating the factors at 2, we can obtain a factorization of N.

The only weakness of this approach is that the polynomial P may be irreducible and thus the factorization of N obtained may be trivial. The problem we study in this paper thus asks for the probability that this procedure returns only a trivial factorization. Therefore, it is desirable to have results, such as those of this paper, proving that this probability converges to 1 very fast.

We will discuss our method in Section 1.3. The method links the problem of irreducibility of random polynomials with mixing times of certain Markov chains, which are mod p analogues of the Bernoulli convolutions we had studied in earlier work (see e.g. [8, 9, 34]). In this paper, we use results available for the Markov chains to study random polynomials, but this can be reversed. In particular, in a forthcoming paper, we will use the results of this paper to obtain new results about the Markov chains.

Our results on irreducibility assume the Riemann hypothesis for Dedekind zeta functions, or at least some information on the zeros. In our last theorem, Theorem 7, we show that conversely irreducibility of random polynomials has (modest) implications about the zeros of Dedekind zeta functions.

1.2. **Results.** Under the full force of the Riemann hypothesis, our best result is the following.

Theorem 2. *Let $P = A_d x^d + \ldots + A_1 x + A_0 \in \mathbb{Z}[x]$ be a random polynomial with independent coefficients. Assume that $A_1, \ldots, A_{d-1}$ are identically distributed with common law μ. Assume further that all coefficients are bounded by $\exp(d^{1/10})$ almost surely. Let $\tau > 0$ be a number such that $\|\mu\|_2^2 := \sum_{x\in\mathbb{Z}} \mu(x)^2 < 1 - \tau$.*

There are absolute (and effective) constants $c, C > 0$ such that if $d \geqslant C/\tau^4$, then with probability at least $1 - \exp(-c\tau d^{1/2}/\log d)$ the following holds for P.

(1) *If RH holds for ζ_K for $K = \mathbb{Q}$ and for all number fields of the form $K = \mathbb{Q}(a)$, where a is a root of P, then $P = \Phi\widetilde{P}$, where $\widetilde{P}$ is irreducible, and Φ is a product of cyclotomic polynomials and x^m for some $m \in \mathbb{Z}_{\geqslant 0}$.*

Moreover with probability at least $1-\exp(-c\tau d^{1/2}/(\log d)^2)$ the following additional property holds for P.

(2) *If RH holds for ζ_K for all number fields of the form $K = \mathbb{Q}(a_1, \ldots, a_m)$, where $a_1, \ldots, a_m$ are any number of roots of P, then $\mathrm{Gal}(\widetilde{P}) \supset \mathrm{Alt}(\deg \widetilde{P})$.*

Here $\mathrm{Alt}(n)$ denotes the alternating group on n elements, $\mathrm{Gal}(P)$ the Galois group of the splitting field of the polynomial P.

There are several remarks in order regarding this theorem. It is natural to allow that the probability laws of A_d and A_0 differ from those of the other coefficients, for example to include the original problem discussed in the beginning of the paper. The exponent $\frac{1}{10}$ has no particular significance and the upper bound $\exp(d^{1/10})$ on the coefficients could be relaxed at the expense of some technical complications in the proof, but we do not pursue this. Nevertheless, the method of proof definitely requires some upper bound in terms of d; it would be interesting to know if this is also necessary for the theorem to hold.

There are certain obstructions to the irreducibility of P that occur with probability higher than the estimate $2\exp(-cd^{1/2}/\log d)$ given in the theorem. In particular, if $\mathbb{P}(A_0 = 0)$ is positive, then $x|P$ with positive probability. Moreover, if ω is a root of unity, then one may think of $P(\omega)$ as the end point of a random walk on $\mathbb{Z}[\omega]$ whose steps are given by $A_j\omega^j$ for $j = 0, \ldots, d$. If we fix $\omega \neq 1$ and μ, then for large values of d, $\mathbb{P}(P(\omega) = 0)$ is proportional to $d^{-r/2}$, where r is the rank of the lattice $\mathbb{Z}[\omega]$.

In summary, the factor Φ may be non-trivial with probability higher than $2\exp(-cd^{1/2}/\log d)$, and its precise behavior can be described by a detailed analysis of random walks on lattices, which we do not pursue here, except for the:

Corollary 3. *Let μ be a probability measure on $\mathbb{Z}$ with finite second moment, which is not supported on a singleton. Let N be a positive integer and U_N be the finite subset of $\mathbb{C}$ consisting of 0 and all of roots unity ω with $[\mathbb{Q}(\omega):\mathbb{Q}] < N$. Let $(A_i)_{i\geqslant 0}$ be a sequence of i.i.d. random variables with common law μ and set $P_d = A_d x^d + \ldots + A_1 x + A_0 \in \mathbb{Z}[x]$. Then, assuming the Riemann hypothesis for Dedekind zeta functions of number fields, as $d \to +\infty$*

$$\mathbb{P}(P_d \text{ is irreducible in } \mathbb{Z}[x]) = 1 - \mathbb{P}(\exists \omega \in U_N, P_d(\omega) = 0) + O_{\mu,N}(d^{-\frac{N}{2}}).$$

In a similar flavor we answer the original problem posed at the beginning of the paper.

Corollary 4. *Let $P(x) = x^d + A_{d-1}x^{d-1} + \ldots + A_1 x + 1 \in \mathbb{Z}[x]$ be a random polynomial with independent coefficients $A_1, \ldots, A_{d-1}$ taking values in 0 and 1 with equal probability. Suppose that the Riemann hypothesis holds for the Dedekind zeta function ζ_K for all number fields of the form $K = \mathbb{Q}(a)$, where a is a root of a polynomial with $0,1$ coefficients.*

Then

$$\mathbb{P}(P \text{ is irreducible in } \mathbb{Z}[x]) = 1 - \sqrt{\frac{2}{\pi d}} + O(d^{-1}),$$

where the implied constant is absolute.

Polynomials of small Mahler measure can also contribute to the error term. In this respect it is also worth pointing out that an exponential bound in the error term in Theorem 2, say of the form $\exp(-cd)$ for some $c > 0$ would easily imply the Lehmer conjecture (arguing, say, as in [9, Lemma 16]).

In the proof of part (2) of Theorem 2, we will show that the Galois group of $\widetilde{P}$ acts k-transitively on its roots with $k > (\log d)^2$. By a well-known fact going back to Bochert and Jordan in the 19-th century, this implies that the Galois group contains the alternating group. In fact, now there are even better results available, which we will discuss in more details in Section 9.1. Using the classification of finite simple groups, it has been proved that all 6-transitive permutation groups contain $\mathrm{Alt}(d)$. However, if we were to rely on this, it would lead only to a very minor improvement in Theorem 2, so we opted for a proof avoiding the classification. Unfortunately, our method cannot distinguish between the symmetric and alternating groups.

Bary-Soroker and Kozma proved that if μ is the uniform distribution on an interval, then with probability tending to 1, the Galois group of P contains $\mathrm{Alt}(d)$ provided P is irreducible. However, our result applies in greater generality and provides a better bound for the probability of exceptions conditionally on the Riemann hypothesis.

Next, we state two results, where the reliance on the Riemann hypothesis is relaxed at the expense of a weakening of the bound.

Theorem 5. *For any numbers $\tau > 0$ and $\alpha > \beta > 3$, there is $c > 0$ such that the following holds. Let $P = A_d x^d + \ldots + A_1 x + A_0 \in \mathbb{Z}[x]$ be a random polynomial with independent coefficients. Assume that $A_1, \ldots, A_{d-1}$ are identically distributed with common law μ. Assume further that all coefficients are bounded by $d^{1/\tau}$ almost surely and $\|\mu\|_2^2 < 1 - \tau$.*

Then with probability at least $1 - 2\exp(-c(\log d)^{\beta-2})$ the following holds for P. Suppose ζ_K has no roots ρ with $|1-\rho| < (\log d)^\alpha/d$ for all $K = \mathbb{Q}(a)$ for any roots a of P. Then $P = \Phi\widetilde{P}$, where $\widetilde{P}$ is irreducible, and Φ is a product of cyclotomic polynomials and x^m for some $m \in \mathbb{Z}_{\geqslant 0}$.

We recall the state of the art in our knowledge about the zeros of Dedekind zeta functions near 1 to motivate the next result. The Dedekind zeta function ζ_K has at most one zero ρ with $|1-\rho| < 4/\log \Delta_K$, where Δ_K is the discriminant of the number field K, see [32, Lemma 3]. If such a zero exists, it must be real, and we call it the exceptional zero of ζ_K. The constant 4 has been improved, see [23] for the latest results. We note that in the setting of Theorems 5 and 6, $\log \Delta_K \leqslant Cd \log d$ for a constant C depending only on τ.

The bounds available for the exceptional zero are much weaker. We know that ζ_K has no zeros ρ with

$$|1-\rho| < \frac{c}{d \cdot d! |\Delta_K|^{1/d}}, \tag{1.2}$$

where d is the degree of K and c is an absolute constant, see [32, proof of Theorem 1']. However, conditionally on Artin's holomorphy conjecture for Artin L-functions, we know by [32, Theorem 4] that ζ_K has no zeros ρ with

$$|1-\rho| < \frac{c}{d \log \Delta_K} + \frac{c}{\Delta_K^{1/d}}.$$

In the next result, we formulate our hypothesis on the zeros of Dedekind functions allowing for an exceptional zero.

Theorem 6. *For any numbers $\tau > 0$, $\alpha > 4$ and $\gamma > 1$ such that $\alpha > 2\gamma + 2$, there is $c > 0$ such that the following holds. Let $P = A_d x^d + \ldots + A_1 x + A_0 \in \mathbb{Z}[x]$ be a random polynomial with independent coefficients. Assume that $A_1, \ldots, A_{d-1}$ are identically distributed with common law μ. Assume further that all coefficients are bounded by $d^{1/\tau}$ almost surely and $\|\mu\|_2^2 < 1 - \tau$.*

Then with probability at least $1 - 2\exp(-c(\log d)^{\alpha-\gamma-2})$ the following holds for P. Suppose ζ_K has at most one root ρ with $|1-\rho| < (\log d)^\alpha/d$ and none with $|1-\rho| < \exp(-c(\log d)^\gamma)$ for all $K = \mathbb{Q}(a)$ for any roots a of P. Then $P = \Phi\widetilde{P}$, where $\widetilde{P}$ is irreducible, and Φ is a product of cyclotomic polynomials and x^m for some $m \in \mathbb{Z}_{\geqslant 0}$.

Most of the interest in our final result is when we know unconditionally that the random polynomial P is irreducible with high probability,

e.g. in the setting of the work of Bary-Soroker and Kozma [3] mentioned above. Then we obtain as a direct consequence of Theorem 7 an unconditional improvement on the bound (1.2) for the exceptional zero of the Dedekind zeta function ζ_K that holds for most number fields K, where K is the sampled by setting $K = \mathbb{Q}(a)$ for a root a of the random irreducible polynomial P.

Theorem 7. *For every $\tau > 0$ and $\alpha > \beta > 3$, there is $c > 0$ such that the following holds. Let $P = A_d x^d + \ldots + A_1 x + A_0 \in \mathbb{Z}[x]$ be a random polynomial with independent coefficients. Assume that $A_1, \ldots, A_{d-1}$ are identically distributed with common law μ. Assume further that all coefficients are bounded by $d^{1/\tau}$ almost surely and $\|\mu\|_2^2 < 1 - \tau$.*

Then with probability at least $1 - 2\exp(-c(\log d)^{\beta-2})$ the following holds for P. There is a root a of P that is not a root of unity, such that $\zeta_{\mathbb{Q}(a)}$ has no zeros ρ with $|1 - \rho| < \exp(-(\log d)^{\alpha+1})$.

1.3. **An outline of the proof.** Our strategy for proving the results stated above aims at finding information about the distribution of the degree sequence in the factorization of the random polynomial P in $\mathbb{F}_p[x]$, and then uses this information to study irreducibility of P in $\mathbb{Z}[x]$ and the Galois group of its splitting field.

Bary-Soroker and Kozma [3] approximated (in a certain sense) the degree sequence in the factorization of a polynomial chosen uniformly at random from degree d monic polynomials in $\mathbb{F}_p[x]$. It is very plausible that such an approximation holds in greater generality not only for the uniform distribution, but we do not know how to prove this. However, we are able to approximate the statistics of the number of low degree factors and this allows us to gain information about the Galois groups using special cases of the Chebotarev density theorem.

The most relevant density theorem for our purposes is the prime ideal theorem, which has the following consequence.

Theorem 8. *Let $P_0 \in \mathbb{Z}[x]$ be a fixed polynomial and let p be a random prime chosen uniformly in a dyadic range $[y, 2y)$. Then*

$$\mathbb{E}[\textit{number of roots of } P_0 \textit{ in } \mathbb{F}_p] \to \{\textit{number of distinct irred. factors of } P_0\} \tag{1.3}$$

as $y \to \infty$.

This observation was suggested as a basis for an algorithm to compute the number of irreducible factors of a polynomial by Weinberger [35].

If the Riemann hypothesis holds for ζ_K for all $K = \mathbb{Q}(a)$, where a is a root of P_0, then the approximation (1.3) is valid once $y > C(\varepsilon)(\log \Delta_{P_0})^{2+\varepsilon}$. A more precise discussion of these ideas including proofs will be given in Sections 2–4. We note that if we wish to approximate the distribution of the full degree sequence of the factorization of P_0 in $\mathbb{F}_p$ using the Chebotarev density theorem, then we need to take

a much larger value for y even if we assume the Riemann hypothesis for all relevant Dedekind zeta functions. Indeed, that would require us to replace the discriminant of P_0 with the discriminant of its splitting field in the above bound, which is potentially much larger, and that would not be sufficient for our purposes.

The next aim of our strategy is to show that

$$\mathbb{E}[\text{number of roots of } P \text{ in } \mathbb{F}_{p_0}] \approx 1, \tag{1.4}$$

where P is a random polynomial in the setting of the above theorems and p_0 is a fixed prime in the range $[y, 2y)$, which is suitably large for the approximation in (1.3) to hold.

If we achieve this goal, then we can randomize the polynomial in (1.3) and the prime in (1.4) and compare the right hand sides to obtain

$$\mathbb{E}[\text{number of distinct irred. factors of } P] \approx 1.$$

Since the number of irreducible factors is always a positive integer, Markov's inequality implies that P has only one irreducible factor with high probability. When we will give the details of the argument, we will choose a slightly different route by estimating the second moments and applying Chebyshev's inequality. Although this is not necessary for Theorems 2, 5 and 6, it does help in that it is enough to make the assumption on the Dedekind zeta functions only for those polynomials for which the conclusion holds. On the other hand, the second moment estimates are necessary for Theorem 7.

To establish (1.4), we fix an element $a \in \mathbb{F}_{p_0}$ and consider an additive random walk on $\mathbb{F}_{p_0}$ whose j'th increment is $A_j a^j$. The endpoint of this walk is $P(a)$. If we can show that the walk mixes rapidly, then we can conclude that

$$\mathbb{P}(P(a) = 0) \approx \frac{1}{p_0}. \tag{1.5}$$

Summing up the probabilities for each $a \in \mathbb{F}_{p_0}$ we arrive at (1.4).

The study of random walks of this kind goes back at least to Chung, Diaconis and Graham [12], who considered the case $a = 2$. Their work has been extended in several directions by Hildebrand (see [20]), however he mostly focused on the case in which a is a fixed integer independent of p_0. In the setting when a may vary with p_0, the diameter of the underlying graph was considered by Bukh, Harper and Helfgott in an unpublished work, see also [19, footnote 4 on page 372], that is, they considered how large d needs to be taken so that the walk reaches every element of $\mathbb{F}_{p_0}$ with positive probability. Their approach relies on certain estimates of Konyagin [24] pertaining to the Waring problem on finite fields and we will apply the same method. See also [10], where the connection between these random walks and Lehmer's conjecture is explored.

It turns out that the random walk does not mix fast enough for certain choices of the parameter a. Indeed, if $a = 0$, then the walk

does not mix at all. Moreover, if $a = 1$, then the mixing time (i.e. how large d needs to be taken for (1.5) to hold) will be $\approx p_0^2$, as can be seen by the central limit theorem. A similar issue arises if a has low multiplicative order. Therefore, it is useful to exclude certain elements of $\mathbb{F}_{p_0}$ from the count. We say that an element $a \in \mathbb{F}_{p_0}$ is admissible if it is not the root of a cyclotomic polynomial of degree at most $\log p_0$. We can then modify (1.3) by counting admissible primes on the left hand side and non-cyclotomic factors on the right. When we give the details of the argument we will exclude from the admissible elements not only the roots of cyclotomic polynomials but also the roots of polynomials of very small Mahler measure. This allows us to obtain improved bounds.

We are able to show that the mixing time is at most $\log p(\log\log p)^{3+\varepsilon}$ for most of the parameters $a \in \mathbb{F}_p$ in a sufficiently strong sense required by our application.[1] This allows us to set $y = \exp(d/(\log d)^{3+\varepsilon})$ when we apply (1.3). Even if we disregard the effect of the exceptional zero, our current knowledge about the zeros of Dedekind zeta functions would require the larger range $y = \exp(Cd\log d)$. Unfortunately, an argument based on the analysis of the random walks for a fixed parameter $a \in \mathbb{F}_p$ cannot yield a mixing time better than $c\log p$, since the number of points that the random walk can reach grows exponentially with the number of steps. To overcome this barrier, one would need to consider the average distribution of the random walk over the parameters $a \in \mathbb{F}_p$. This however seems to be exceedingly difficult to study.

There is one last issue that we need to consider. The above argument cannot distinguish between irreducible polynomials and proper powers. Indeed, we are able count distinct irreducible factors only. To show that P is not a proper power with high probability, we show that $P(2)$ is not a proper power. To that end, we will use the large sieve together with the classical $a = 2$ case of the above discussed random walks.

Using the above method, we can also obtain information about the Galois group of P. What we discussed so far amounted to showing that the Galois group acts transitively on the complex roots of P. A more general version of this argument can be used to show that the action is k-transitive for large values of k, large enough that it forces the Galois group to contain the alternating group.

Finally, we comment on the proof of Theorem 7. If the Dedekind zeta function has an exceptional zero, then all other zeros are repelled away from 1 by what is known as the Deuring-Heilbronn phenomenon. In the context of Theorem 8, this implies that the left hand side of (1.3) is close to zero for a certain range of primes. This can be contradicted by (1.4).

[1] We can get better results for typical parameters in a weaker sense, which is not suitable for the purposes of this paper. These results will appear in a forthcoming paper.

1.4. **Organization of the paper.** In Sections 2–4 we discuss the prime ideal theorem and use it to obtain estimates for the average number of roots of a polynomial in finite fields related to (1.3). In Sections 5 and 6, we study equidistribution of random walks, we revisit Konyagin's estimates in [24] and the argument suggested by Bukh, Harper and Helfgott. In Section 7, we give an upper bound on the probability that the random polynomial P has a factor of small Mahler measure utilizing some ideas of Konyagin [25]. In Section 8 we use the large sieve to show that P is not the product of a proper power and cyclotomic factors with high probability. In Section 9 we combine the above ingredients to prove the results stated in Section 1.2.

1.5. **Notation.** If K is a number field, we write d_K for its degree and Δ_K for its discriminant. If $P \in \mathbb{Z}[x]$ is a polynomial, we write d_P for its degree, Δ_P for its discriminant and

$$M(P) = a_d \prod_{z_j : |z_j| > 1} |z_j|$$

for its Mahler measure, where a_d is the leading coefficient of P and z_j runs through the complex roots of P taking multiplicities into account. We recall the estimates

$$1 + c\Big(\frac{\log\log d_P}{\log d_P}\Big)^3 \leqslant M(P) \leqslant (a_0^2 + \ldots + a_d^2)^{1/2}, \tag{1.6}$$

where $c > 0$ is an absolute constant $a_0, \ldots, a_d$ are the coefficients of P. See [16] for the inequality on the left hand side and [5, Lemma 1.6.7] for the right hand side.

We write $\sum_p$ for summation over rational primes.

Throughout the paper we use the letters c and C to denote positive numbers whose values may vary at each occurrence. These values are effective and numerical: they could, in principle, be determined by following the arguments. We will use upper case C when the number is best thought to be large, and lower case c when it is best thought to be small. In addition we will use Landau's $O(X)$ notation to denote a quantity that is bounded in absolute value by a constant multiple of X.

1.6. **Acknowledgments.** The authors are grateful to Boris Bukh, Mohammad Bardestani and Peter Sarnak for helpful discussions on various aspects of this work.

2. The prime ideal theorem

Let K be a number field of degree $d = d_K$ with discriminant $\Delta = \Delta_K$ and denote by $\mathcal{O}_K$ its ring of integers. Write ζ_K for the Dedekind zeta function of K. Write $A(n) = A_K(n)$ for the number of prime ideals $\mathfrak{p} \subset \mathcal{O}_K$ with $N_{K/\mathbb{Q}}(\mathfrak{p}) = n$.

The purpose of this section is to compute the average value of $A(p)$ with respect to suitably chosen weights supported on primes. We first consider this question under the assumption that RH holds for ζ_K. In what follows, $\sum_p$ indicates summation over all positive primes in $\mathbb{Z}$.

Proposition 9. *Let $X > 1$ be a number and let*

$$h_X(u) = \begin{cases} 2\exp(-X) & \text{if } u \in (X - \log 2, X], \\ 0 & \text{otherwise.} \end{cases}$$

If RH holds for ζ_K, then

$$\sum_p A(p)\log(p)h_X(\log p) = 1 + O(X^2\log(\Delta)\exp(-X/2)),$$

where the implied constant is absolute.

Proof. We write

$$\psi_K(x) = \sum_{n,m\in\mathbb{Z}_{>0}:n^m\leqslant x} A(n)\log n.$$

There is an absolute constant $C > 0$ such that if *RH* holds for ζ_K, then for all $x > 1$,

$$|\psi_K(x) - x| \leqslant C\sqrt{x}(\log x \log\Delta + d(\log x)^2)$$

See for example [18, Corollary 1.2]. Applying this for $x = \exp(X)$ and $x = \exp(X)/2$, we find that

$$\begin{aligned}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} A(n)\log(n)h_X(\log n^m) =& 2\exp(-X)(\psi_K(\exp(X)) - \psi_K(\exp(X)/2)) \\ =& 1 + O(X^2\log(\Delta)\exp(-X/2)).\end{aligned}$$

Here we used that $d_K \leqslant C(\log\Delta_K)$ by Minkowski's lower bound on the discriminant.

We estimate the contribution of the summands for which n^m is not a prime. First we note that for each of these terms, n^m is a proper power, and there are at most

$$\exp(X/2) + \exp(X/3) + \ldots + \exp(X/\lceil X\rceil) \leqslant C\exp(X/2)$$

such numbers between $\exp(X)$ and $\exp(X)/2$. Each such number can be written in the form n^m in at most X different ways, and $A(n)\log n \leqslant d_K X$. Therefore

$$\sum_{n,m:\ n^m \text{ is not prime}} A(n)\log(n)h_X(\log n^m) \leqslant CX^2\log(\Delta)\exp(-X/2).$$

□

The purpose of the rest of this section is to formulate a variant of this proposition with a milder assumption on the zeros of ζ_K. Readers

only interested in the proof of Theorem 2 may skip to the next section. Everything that follows is well known and classical. We begin by recalling the smooth version of the explicit formula.

Theorem 10. *Let $g \in C^2(\mathbb{R})$ be a function supported in a compact interval contained in $\mathbb{R}_{>0}$. Write*

$$\widehat{g}(s) = \int_{\mathbb{R}} \exp(isu)g(u)du.$$

Then

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} A(n)\log(n)g(m\log(n)) = \widehat{g}(-i) - \sum_{\rho}\widehat{g}(-i\rho),$$

where the summation over ρ is taken over all zeros of ζ_K (including the trivial ones) taking multiplicities into account.

This result is well known but it does not seem to be readily available in this form in standard text books, therefore we give the very short proof for the reader's convenience.

Proof. We note that

$$\frac{\zeta_K'}{\zeta_K}(s) = -\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} A(n)\log(n)n^{-ms}$$

for $\mathrm{Re}(s) > 1$.

Since g is compactly supported and C^2, $\widehat{g}$ is holomorphic and $|\widehat{g}(is)| = O(|\,\mathrm{Im}(s)|^{-2})$ with an implied constant (continuously) depending only on $\mathrm{Re}(s)$. By the Fourier inversion formula, we have

$$\begin{aligned}\int_{\mathrm{Re}(s)=2} n^{-s}\widehat{g}(-is)ds &= \int_{\mathrm{Re}(s)=0} n^{-s}\widehat{g}(-is)ds \\ &= i\int_{-\infty}^{\infty}\exp(-it\log n)\widehat{g}(t)dt = 2\pi i g(\log n).\end{aligned}$$

for each $n \in \mathbb{Z}_{>0}$.

Therefore, we have

$$\frac{-1}{2\pi i}\int_{\mathrm{Re}(s)=2}\frac{\zeta_K'}{\zeta_K}(s)\widehat{g}(-is)ds = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} A(n)\log(n)g(m\log n).$$

Shifting the contour integration to $\mathrm{Re}(s) = -\infty$ we can recover the claimed formula from the residue theorem. We note that $\mathrm{supp}(g) \subset \mathbb{R}_{>0}$, $\widehat{g}(-is)$ decays exponentially as $\mathrm{Re}(-is) \to \infty$ and leave the verification of the rest of the details to the interested reader. □

In the next lemma, we introduce the weight functions that we will use and establish some of their properties. The aim is to find compactly supported weights g such that its Laplace transform $G(s) = \widehat{g}(-is)$ decays fast when $\mathrm{Re}(s) \leqslant 1$ and s is moving away from 1. To achieve

the optimal decay, it is useful to choose g depending on the distance of s from 1 where we wish to make $G(s)$ small. The construction was inspired by Ingham [21].

Lemma 11. *Let $X \in \mathbb{R}_{>0}$ and let $k \in \mathbb{Z}_{>0}$. For $r \in \mathbb{R}_{>0}$, write*

$$I_r(u) = \begin{cases} \frac{1}{r} & \text{if } u \in [-r/2, r/2] \\ 0 & \text{otherwise.} \end{cases}$$

Let

$$g_{X,k}(u) = \exp(-u) \underbrace{I_{X/2k} * \ldots * I_{X/2k}}_{k\text{-fold}}(u - 3X/4), \tag{2.1}$$

$$G_{X,k}(s) = \widehat{g}_{X,k}(-is) = \int_{\mathbb{R}} \exp(su) g(u) du. \tag{2.2}$$

Suppose $k \geqslant 4$ and $X \geqslant 2k$. Then $g_{X,k} \in C^2(\mathbb{R})$ and it is supported in $[X/2, X]$ and we have $g_{X,k}(u) \leqslant \exp(-u)$ for all $u \in \mathbb{R}$. We have $G(1) = 1$ and the following bounds hold for all $s \in \mathbb{C}$ with $\mathrm{Re}(s) \leqslant 1$ and for all $\sigma \in (0,1)$ and $X_1 > X_2$,

$$\begin{aligned} 0 \leqslant 1 - G_{X,k}(\sigma) &\leqslant X(1-\sigma), \\ |G_{X,k}(s)| &\leqslant \Big(\frac{4k}{|1-s|X}\Big)^k, \\ |G_{X,k}(s)| &\leqslant \exp((\mathrm{Re}(s) - 1)X/2), \\ \frac{G_{X_1,k}(\sigma)}{G_{X_2,k}(\sigma)} &\leqslant \exp(-(1-\sigma)(X_1 - X_2)/4). \end{aligned}$$

Proof. The claim $\operatorname{supp} g_{X,k} \subset [X/2, X]$ and $g(u) \leqslant \exp(-u)$ follows immediately from its definition and the assumption $X \geqslant 2k$.

Note

$$\widehat{I}_{X/2k}(s) = \frac{\exp(isX/4k) - \exp(-isX/4k)}{isX/2k}.$$

Then $|\widehat{g}_{X,k}(\sigma)| \leqslant C|\sigma|^{-k}$ for $\sigma \in \mathbb{R}$, where C is a number that depends only on X and k, and it follows that $g_{X,k} \in C^2$ if $k \geqslant 4$.

We also have

$$\widehat{g}_{X,k}(s) = \exp(3i(s+i)X/4)\Big(\frac{\exp(i(s+i)X/4k) - \exp(-i(s+i)X/4k)}{i(s+i)X/2k}\Big)^k.$$

We can write

$$G_{X,k}(s) = \exp(3(s-1)X/4)\Big(\frac{\exp((s-1)X/4k) - \exp(-(s-1)X/4k)}{(s-1)X/2k}\Big)^k. \tag{2.3}$$

Taking the limit $s \to 1$, we get $G_{X,k}(1) = 1$. Using the bound

$$\Big|\frac{\exp(z) - \exp(-z)}{2z}\Big| \leqslant \exp(-\mathrm{Re}(z))$$

with $z = (s-1)X/4k$, which is valid for $\mathrm{Re}(z) \leqslant 0$, we get

$$|G_{X,k}(s)| \leqslant \exp((\mathrm{Re}(s)-1)X/2)$$

if $\mathrm{Re}(s) \leqslant 1$.

Next, we use

$$\Big|\frac{\exp(z)-\exp(-z)}{2z}\Big| \leqslant \frac{\exp(-\mathrm{Re}(z))}{|z|}$$

with $z = (s-1)X/4k$, which is valid for $\mathrm{Re}(z) \leqslant 0$, and we get

$$|G_{X,k}(s)| \leqslant \frac{\exp((\mathrm{Re}(s)-1)X/2)}{(|s-1|X/4k)^k}$$

if $\mathrm{Re}(s) \leqslant 1$. Using $\exp((\mathrm{Re}(s)-1)X/2) \leqslant 1$, we get the claim.

To show

$$\frac{G_{X_1,k}(\sigma)}{G_{X_2,k}(\sigma)} \leqslant \exp(-(1-\sigma)(X_1-X_2)/4),$$

it is enough to prove that $F_1'(Y)/F_1(Y) \leqslant -1$ for $Y \geqslant 0$, where

$$F_1(Y) = \exp(-3Y)\Big(\frac{\exp(Y/k)-\exp(-Y/k)}{2Y/k}\Big)^k.$$

(We use the substitution $Y = (1-\sigma)X/4$ and (2.3)).

This follows at once, if we show that $F_2'(Z)/F_2(Z) \leqslant 1$ for $Z > 0$, where

$$F_2(Z) = \frac{\exp(Z)-\exp(-Z)}{2Z}.$$

To that end, we calculate

$$\frac{F_2(Z)'}{F_2(Z)} = \frac{\exp(Z)+\exp(-Z)}{\exp(Z)-\exp(-Z)} - \frac{1}{Z}$$

and observe that $F_2'(Z)/F_2(Z) \leqslant 1$ is equivalent to

$$2\exp(-Z) \leqslant \frac{\exp(Z)-\exp(-Z)}{Z}.$$

We note that the left hand side is always less than 2 and the right hand side is greater than 2 for $Z > 0$. The latter can be seen, for example by computing the power series expansion of the right hand side. □

We record some well known estimates for the number of roots of ζ_K near $s = 1$. These go back at least to Stark [32].

Lemma 12. *For every $0 \leqslant r \leqslant 1$, we have*

$$|\{\rho : \zeta_K(\rho) = 0, |1-\rho| < r\}| \leqslant \frac{3}{2} + 3r\log|\Delta_K|.$$

There is an absolute constant $C > 0$, such that for every $r > 1$, we have

$$|\{\rho : \zeta_K(\rho) = 0, |1-\rho| < r\}| \leqslant C\log|\Delta_K| + Cd_K r\log r.$$

We count the zeros with multiplicities and include the trivial ones.

Proof. As in the proof of [32, Lemma 3], we have

$$\sum\nolimits'_\rho \frac{1}{\sigma-\rho} < \frac{1}{\sigma-1} + \frac{1}{2}\log|\Delta_K|,$$

where $1 < \sigma \leqslant 2$ is arbitrary and $\sum'_\rho$ indicates summation over an arbitrary subset of non-trivial zeros of ζ_K (taking multiplicities into account) closed under conjugation. If $r < 1/2$, we take $\sigma = 1 + 2r$ and consider the zeros ρ that satisfy $|1-\rho| < r$. For each such ρ, we have

$$\operatorname{Re}\left(\frac{1}{\sigma-\rho}\right) \geqslant \frac{1}{3r},$$

which can be seen easily by finding the image of the disk $\{z : |1-z| < r\}$ under the inversion through $1 + 2r$. This gives us

$$\frac{1}{3r}|\{\rho : |1-\rho| < r\}| < \frac{1}{2r} + \frac{1}{2}\log|\Delta_K|,$$

which yields

$$|\{\rho : |1-\rho| < r\}| < \frac{3}{2} + \frac{3}{2}r\log|\Delta_K|.$$

Taking $\sigma = 1 + r$ for $1/2 \leqslant r \leqslant 1$, the same argument gives

$$|\{\rho : |1-\rho| < r\}| < 2 + r\log|\Delta_K|,$$

which is stronger than our claim since $2r\log|\Delta_K| > 1/2$.

We note that the trivial zeros are among the non-positive integers, and each have multiplicity at most d_K. Moreover, we have

$$|\{\rho : 0 < \operatorname{Re}(\rho) < 1, |\operatorname{Im}(\rho) - t| \leqslant 1\}| \leqslant C\log(\Delta_K) + Cd_K\log(|t|+2)$$

for any $t \in \mathbb{R}$ (see e.g. [27, Lemma 5.4]). These two facts easily imply the second claim. □

Now we formulate a variant of Proposition 9 under a milder assumption on the zeros.

Proposition 13. *Let $\alpha > \beta, \tau \in \mathbb{R}_{>0}$. Let $X = d(\log d)^{-\beta}$ or $X = 2d(\log d)^{-\beta}$ and $k = \lfloor(\log d)^{\alpha-\beta}/10\rfloor$. Let K be a number field of degree at most d and discriminant at most $\exp(\tau^{-1}d\log d)$ in absolute value. Suppose that ζ_K has at most one zero ρ_0 such that $|1-\rho_0| < d^{-1}(\log d)^\alpha$ Then*

$$\sum_p A(p)\log(p)g_{X,k}(\log(p)) = 1 - G_{X,k}(\rho_0) + O(\exp(-c(\log d)^{\alpha-\beta})).$$

When the exceptional zero ρ_0 does not exist the corresponding term should be removed from the formula. The implied constant and c may depend only on α, β and τ.

Proof. In what follows, we will assume that d is sufficiently large depending on α, β and τ. Otherwise, the claim may be made trivial by a sufficient choice of the constants.

The proof is based on the explicit formula in Theorem 10, which gives us

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} A(n)\log(n)g(m\log(n)) = G(1) - \sum_{\rho} G(\rho), \tag{2.4}$$

where $g = g_{X,k}$ and $G = G_{X.k}$.

First, we focus on the left hand side of (2.4) and show that the terms for which n^m is not a prime do not have a significant contribution. We write

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} A(n)\log(n)g(m\log(n)) = \sum_{p}^{\infty}\sum_{l=1}^{\infty} \widetilde{A}(p^l)g(\log(p^l)),$$

where

$$\widetilde{A}(p^l) = \sum_{n,m:n^m=p^l} A(n)\log(n) = \sum_{j:j|l} A(p^j)\log(p^j).$$

We note that $\widetilde{A}(p) = A(p)\log p$ for all primes p and that

$$\widetilde{A}(p^l) \leqslant \sum_{j=1}^{\infty} jA(p^j)\log(p) \leqslant d_K \log(p).$$

(The last inequality is an equality if p is unramified in K.) Therefore, we can write

$$\Big|\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} A(n)\log(n)g(m\log(n)) - \sum_{p} A(p)\log(p)g(\log(p))\Big|$$
$$\leqslant \sum_{p}\sum_{l=2}^{\infty} d_K \log(p)g(\log(p^l)). \tag{2.5}$$

Since the support of $g = g_{X,k}$ is contained in $[\exp(X/2), \exp(X)]$ only those terms contribute in (2.5) for which $p^l \in [\exp(X/2), \exp(X)]$. This means in particular that $\log(p) \leqslant X$ for all such terms and we can write

$$(2.5) \leqslant d_K X \sum_{p} \sum_{\substack{l:l\geqslant 2,\\ p^l\in[\exp(X/2),\exp(X)]}} p^{-l},$$

where we also used $g(\log(x)) \leqslant x^{-1}$. Therefore,

$$(2.5) \leqslant d_K X\Big(\sum_{n=\exp(X/4)}^{\exp(X/2)} n^{-2} + \sum_{n=\exp(X/6)}^{\exp(X/3)} n^{-3} + \sum_{n=2}^{\exp(X/4)} \sum_{l:n^l\in[\exp(X/2),\exp(X)]} n^{-l}\Big).$$

We note that

$$\sum_{l:n^l\in[\exp(X/2),\exp(X)]} n^{-l} \leqslant 2\exp(-X/2)$$

for any n, hence

$$(2.5) \leqslant d_K X(C\exp(-X/4) + C\exp(-X/3) + 2\exp(X/4)\exp(-X/2)),$$

so we can conclude

$$\Big|\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} A(n)\log(n)g(m\log(n)) - \sum_{p} A(p)\log(p)g(\log(p))\Big| \leqslant Cd_K X\exp(-X/4).$$

Now we turn to the right hand side of (2.4) and estimate the contribution of the zeros ρ that satisfy $|1-\rho| > d^{-1}(\log d)^{\alpha}$. We write

$$R_j := \{\rho : \zeta_K(\rho) = 0, 2^j d^{-1}(\log d)^{\alpha} \leqslant |1-\rho| < 2^{j+1}d^{-1}(\log d)^{\alpha}\}$$

for each $j \in \mathbb{Z}_{\geqslant 0}$. We think about this as a multiset with each zero contained in it with its multiplicity.

By Lemma 12, we have

$$|R_j| \leqslant C2^j(\log d)^{\alpha+1}$$

for each j such that $2^{j+1}d^{-1}(\log d)^{\alpha} \leqslant 1$. Here we use that $\log\Delta_K \leqslant \tau^{-1}d\log d$. To consider the case $2^{j+1}d^{-1}(\log d)^{\alpha} > 1$, we note

$$\log(2^{j+1}d^{-1}(\log d)^{\alpha}) \leqslant Cj,$$

and the second part of the same lemma implies that

$$|R_j| \leqslant C(j+1)2^j(\log d)^{\alpha+1} \leqslant C2^{2j}(\log d)^{\alpha+1}.$$

So this last estimate holds for all j.

By Lemma 11 we know that

$$|G(\rho)| \leqslant \Big(\frac{4k}{2^j d^{-1}(\log d)^{\alpha}X}\Big)^k \leqslant \exp(-c(j+1)(\log d)^{\alpha-\beta})$$

for each $\rho \in R_j$. We combine this with the bounds on $|R_j|$ and obtain the following estimates provided d is sufficiently large (depending on α, β and τ)

$$\begin{aligned}\sum_{j=0}^{\infty}\sum_{\rho\in R_j}|G(\rho)| &\leqslant \sum_{j=0}^{\infty} C2^j(\log d)^{\alpha+1}\cdot\exp(-c(j+1)(\log d)^{\alpha-\beta}) \\ &\leqslant \sum_{j=0}^{\infty} C\exp(-c(j+1)(\log d)^{\alpha-\beta}) \\ &\leqslant C\exp(-c(\log d)^{\alpha-\beta}).\end{aligned}$$

We recall that $G(1) = 1$ and using the above estimate, we write

$$|G(1) - \sum_{\rho} G(\rho) - (1-G(\rho_0))| \leqslant C\exp(-c(\log d)^{\alpha-\beta}),$$

where ρ_0 is the unique zero of ζ_K with $|1-\rho_0| < d^{-1}(\log d)^{\alpha}$ if it exists and the term $G(\rho_0)$ should be omitted from the formula if there is no such zero. Combining this with the estimate we gave above for the left hand side, we get the claim of the proposition. □

3. Splitting of prime ideals and roots in finite fields

In this section, we record some facts about the connection between the number of roots a polynomial has in finite fields and the way prime ideals split when we extend $\mathbb{Q}$ by adjoining roots of the polynomial.

We fix two numbers $\kappa \in (0, 1/100)$ and $X \in \mathbb{R}_{>10}$.

Definition 14 (admissible polynomial)**.** *We say that an irreducible polynomial $R \in \mathbb{Z}[x]$ is* (X, κ)-admissible *if $M(R) > \exp(\kappa)$ or $\deg R > 10X$. Otherwise it is called* (X, κ)-exceptional.

By abuse of language and ease of notation in this section we will simply speak of admissible or exceptional polynomials without reference to X and κ, which we assume fixed.

Lehmer's conjecture implies that all exceptional irreducible polynomials are either cyclotomic or equal to x. It follows from a result of Dubickas and Konyagin [17, Theorem 1] that the number of exceptional polynomials of degree d is at most $\exp(\kappa d)$ if d is larger than an absolute constant independent of X.

The reason for excluding polynomials of small Mahler measure is that this will allow us to obtain slightly better results in Sections 5 and 6. We will set the value of κ depending on the common law of the coefficients of the random polynomials so that the probability of a random polynomial having an exceptional and non-cyclotomic factor will be very small. This is proved in Section 7 using the above mentioned estimate for the number of exceptional polynomials. We could opt to make only the low degree cyclotomic polynomials exceptional, but this would not lead to a significant simplification of our arguments.

Definition 15 (admissible residue)**.** *Let p be a prime such that $\log p \in [X/2, X]$. A residue $a \in \mathbb{F}_p$ is said to be* (X, κ)-admissible *if it is not the root of an (X, κ)-exceptional irreducible polynomial.*

Again if X and κ are fixed, as we assume in this section, we will drop the prefix (κ, X) and speak about admissible residues.

Let $P \in \mathbb{Z}[x]$ be a polynomial, F its splitting field and p be a prime. We write $B_P(p)$ for the number of distinct admissible roots of P in $\mathbb{F}_p$. Write $\widetilde{P}$ for the product of the admissible irreducible factors of P. Note that $\widetilde{P}$ is square free. Write Ω for the set of complex roots of $\widetilde{P}$.

Let $m \in \mathbb{Z}_{>0}$ and consider the diagonal action of $G = \mathrm{Gal}(F|\mathbb{Q})$ on Ω^m. We may decompose Ω^m into distinct G-orbits and for each orbit $O \in \Omega^m/G$ pick one representative $\omega := (x_1, \dots, x_m) \in O$ and consider the subfield $K_O = \mathbb{Q}(x_1, \dots, x_m)$. The isomorphism class of K_O is independent of the choice of the representative ω in Ω.

Recall that $A_K(p)$ denotes the number of prime ideals $\mathfrak{p} \subset \mathcal{O}_K$ with norm p.

The purpose of this section is to prove the following.

Proposition 16. *Let $P \in \mathbb{Z}[x]$, let p be a prime such that $p \nmid \Delta_{\widetilde{P}}$ and and $p \nmid \mathrm{Res}(\widetilde{P}, R)$ for any exceptional polynomials R. Let $m \in \mathbb{Z}_{>0}$. Then*

$$B_P(p)^m = \sum_{O \in \Omega^m/G} A_{K_O}(p).$$

Let F be a finite Galois extension of $\mathbb{Q}$ and let $p \in \mathbb{Z}$ be a prime that is unramified in F. Then we write $\mathrm{Frob}_F(p)$ for the (conjugacy class) of the Frobenius element in $\mathrm{Gal}(F|\mathbb{Q})$ at p.

We begin by recalling two standard facts.

Lemma 17 ([13, Theorem 4.8.13])**.** *Let $P \in \mathbb{Z}[x]$ be a polynomial and let F be a finite Galois extension of $\mathbb{Q}$ containing the roots of P. Let p be a prime such that $p \nmid \Delta_P$. Then there is a bijective correspondence between the cycles of $\mathrm{Frob}_F(p)$ acting on the complex roots of P and the irreducible factors of P in $\mathbb{F}_p$, such that the length of a cycle equals the degree of the corresponding irreducible factor.*

Lemma 18 ([29, Chapter 4, Theorem 33])**.** *Let F be a finite Galois extension of $\mathbb{Q}$ with Galois group $G = \mathrm{Gal}(F|\mathbb{Q})$ and let $H \leqslant G$ be a subgroup. Let $p \in \mathbb{Z}$ be a prime, which is unramified in the extension $F|\mathbb{Q}$. Then the number of fixed points of $\mathrm{Frob}_F(p)$ acting on G/H is $A_K(p)$, that is the number of prime ideals in $\mathcal{O}_K$ of norm p, where K is the subfield of F pointwise fixed by H.*

Proof of Proposition 16. Recall that F is the splitting field of P and $\Omega \subset \mathbb{C}$ the set of roots of $\widetilde{P}$. We apply Lemma 17 for $\widetilde{P}$, and see that the number of fixed points of $\mathrm{Frob}_F(p)$ acting on Ω is $B_P(p)$. Here we used that all roots of $\widetilde{P}$ in $\mathbb{F}_p$ are distinct and admissible, because $p \nmid \Delta_{\widetilde{P}}$ and $p \nmid \mathrm{Res}(\widetilde{P}, R)$ for any exceptional R. Therefore, $B_P(p)^m$ is the number of fixed points of $\mathrm{Frob}_F(p)$ acting diagonally on Ω^m.

Consider an orbit O of $\mathrm{Gal}(F|\mathbb{Q})$ in Ω^m and let $K_O = \mathbb{Q}(x_1, \dots, x_m)$ for some representative $\omega := (x_1, \dots, x_m)$ of O. Let H be the stabiliser of ω in G. By the Galois correspondence K_O is the subfield of F fixed by H. Hence the number of fixed points of $\mathrm{Frob}_F(p)$ in O is $A_{K_O}(p)$ by Lemma 18. The claim follows. □

4. Expected number of roots of a polynomial in a random finite field

We combine the results of the previous two sections and deduce the following two results. Below we have kept the notation of Section 3. Recall that the function h_X was defined in Proposition 9, that Ω is the set of roots of $\widetilde{P}$, F the splitting field of P and $G = \mathrm{Gal}(F|\mathbb{Q})$ its Galois group. Given $m \in \mathbb{Z}_{>0}$, $\kappa \in (0, \frac{1}{100})$ and $X > 10$ we will denote by $B_P(p)$ is the set of $(\frac{\kappa}{m}, mX)$-admissible roots of P in $\mathbb{F}_p$ (see Definition 15).

Proposition 19. *Let $d, m \in \mathbb{Z}_{\geqslant 1}$. Let $P \in \mathbb{Z}[x]$ be a polynomial with coefficients at most $\exp(d^{1/10})$ of degree at most d. Suppose that for every G-orbit O on Ω^m, the Dedekind zeta function ζ_{K_O} of the subfield $K_O \leqslant F$ satisfies RH. Let $X \geqslant md^{1/10}$. Then*

$$\sum_p B_P(p)^m \log(p) h_X(\log p) = |\Omega^m/G| + O(\exp(-X/10)).$$

The implied constant is absolute.

Proposition 20. *Let $\alpha > \beta, \tau \in \mathbb{R}_{>0}$. Let $X = d(\log d)^{-\beta}$ or $X = 2d(\log d)^{-\beta}$ and $k = \lfloor (\log d)^{\alpha-\beta}/10 \rfloor$. Let $P \in \mathbb{Z}[x]$ be a polynomial with coefficients at most $d^{1/\tau}$ of degree at most d. Suppose that for every G-orbit $O \subset \Omega$ the Dedekind zeta function ζ_{K_O} of the subfield $K_O \leqslant F$ has at most one root $\rho_{K_O,0}$ such that $|1 - \rho_{K_O,0}| < d^{-1}(\log d)^{\alpha}$. Then*

$$\sum_p B_P(p) \log(p) g_{X,k}(\log p) = \sum_{O \in \Omega/G} (1 - G_{X,k}(\rho_{K_O,0})) + O(\exp(-c(\log d)^{\alpha-\beta})).$$

If the exceptional zero $\rho_{K_O,0}$ does not exist for some O, then the term $G_{X,k}(\rho_{K_O,0})$ should be omitted from the formula. The implied constant and c may depend only on α, β and τ.

We will use the next lemma to estimate the number of primes for which the result of the previous section does not hold.

Lemma 21. *Let $P \in \mathbb{Z}[x]$ be a polynomial of degree at most d with coefficients at most H. Let Q be a polynomial that divides P. Then*

$$|\Delta_Q| \leqslant (Hd)^{2d}.$$

For any irreducible $R \in \mathbb{Z}[x]$ of degree at most d with $M(R) \leqslant 2$, we have

$$|\operatorname{Res}(\widetilde{P}, R)| \leqslant (4Hd)^{2d}.$$

Let K be a number field obtained by adjoining at most m roots of P to $\mathbb{Q}$. Then

$$|\Delta_K| \leqslant (Hd)^{2md^m}.$$

Proof. Recall Mahler's bound on the discriminant of a polynomial $Q \in \mathbb{C}[x]$ of degree n ([28, Theorem 1])

$$|\Delta_Q| \leqslant n^n M(Q)^{2n-2}. \tag{4.1}$$

If Q divides P, $M(Q) \leqslant M(P) \leqslant H(d+1)^{1/2}$ by (1.6) hence Mahler's bound gives:

$$|\Delta_Q| \leqslant d^d (H(d+1)^{1/2})^{2d-2} \leqslant (Hd)^{2d}.$$

Now recall that $|\Delta_{\widetilde{P}R}| = |\Delta_{\widetilde{P}} \Delta_R| \operatorname{Res}(\widetilde{P}, R)^2$. Since R is irreducible and $\widetilde{P}$ is square free, $|\Delta_{\widetilde{P}} \Delta_R| \geqslant 1$ and thus $\operatorname{Res}(\widetilde{P}, R)^2 \leqslant |\Delta_{R\widetilde{P}}|$. Moreover $M(\widetilde{P}R) = M(\widetilde{P})M(R) \leqslant 2M(\widetilde{P})$. So by (4.1) and (1.6) we conclude

$$\operatorname{Res}(\widetilde{P}, R)^2 \leqslant (2d)^{2d} (2H(d+1)^{1/2})^{4d-2} \leqslant (4Hd)^{4d}.$$

Let $\alpha_1, \dots, \alpha_m$ be roots of P and $K = \mathbb{Q}(\alpha_1, \dots, \alpha_m)$. For any two number fields L_1, L_2 we have

$$|\Delta_{L_1 L_2}| \leqslant |\Delta_{L_1}|^{[L_1 L_2 : L_1]} |\Delta_{L_2}|^{[L_1 L_2 : L_2]} \leqslant |\Delta_{L_1}|^{\deg L_2} |\Delta_{L_2}|^{\deg L_1},$$

see e.g. [33]. Using this inductively, we can write

$$|\Delta_K| \leqslant |\Delta_{\mathbb{Q}(\alpha_1)}|^{d^{m-1}} \cdots |\Delta_{\mathbb{Q}(\alpha_m)}|^{d^{m-1}} \leqslant |\Delta_P|^{m d^{m-1}},$$

which proves the claim by the first part. □

Proof of Proposition 19. We apply Proposition 9 for each K_O.

$$\begin{aligned}\sum_{O \in \Omega^m/G} \sum_p A_{K_O}(p) \log(p) h_X(\log p) =& |\Omega^m/G| + O(d^m X^2 \cdot 2md^{m+1/10} \exp(-X/2)) \\ =& |\Omega^m/G| + O(\exp(-X/10)).\end{aligned}$$

Here we used the estimate for Δ_{K_O} from Lemma 21, and the bound $d^m = O(\exp(X/10))$, which follows from our assumption $X \geqslant md^{1/10}$.

We proceed to estimate

$$\Big| \sum_p B_P(p)^m \log(p) h_X(\log(p)) - \sum_{O \in \Omega^m/G} \sum_p A_{K_O}(p) \log(p) h_X(\log(p)) \Big|. \tag{4.2}$$

By Proposition 16, a prime p may contribute to (4.2) only if $p | \Delta_{\widetilde{P}}$ or $p | \operatorname{Res}(\widetilde{P}, R)$ for some $(\frac{\kappa}{m}, mX)$-exceptional irreducible R. As we already noted, [17, Theorem 1] implies that the number of exceptional polynomials is at most $\exp(10\kappa X) \leqslant \exp(X/10)$. Therefore, the number of primes p contributing to (4.2) is at most

$$d^C \exp(X/10).$$

Here we used again the bounds from Lemma 21.

Since for any p we always have $0 \leqslant B_P(p)^m \leqslant d^m$ and

$$0 \leqslant \sum_{O \in \Omega^m/G} A_{K_O}(p) \leqslant \sum_{O \in \Omega^m/G} \deg K_O = \sum_{O \in \Omega^m/G} |O| = |\Omega^m| \leqslant d^m,$$

the contribution of a prime to (4.2) is at most $d^m X \cdot 2 \exp(-X)$. Therefore,

$$(4.2) \leqslant d^{m+C} X \exp(X/10 - X) \leqslant C \exp(-X/10),$$

and the claim follows. □

Proof of Proposition 20. The proof is similar to the previous one. By Lemma 21, we have $|\Delta_{K_O}| \leqslant d^{2(1/\tau+1)d}$ for each orbit O in Ω. Hence Proposition 13 applies to each K_O and we obtain

$$\begin{aligned}&\sum_O \sum_p A_{K_O}(p) \log(p) g_{X,k}(\log p) \\ &\quad = \sum_{O \in \Omega/G} (1 - G_{X,k}(\rho_{K,0})) + O(\exp(-c(\log d)^{\alpha-\beta})).\end{aligned}$$

Since Mahler measure is multiplicative using Dobrowolski's lower bound (1.6) the number $|\Omega/G|$ of irreducible factors of $\widetilde{P}$ is at most $|\Omega/G| \leqslant C(\log d)^4$. Hence $|\Omega/G|$ can be absorbed into $\exp(-c(\log d)^{\alpha-\beta})$. We proceed to estimate

$$\Big|\sum_p B_P(p)\log(p)g_{X,k}(\log(p)) - \sum_{O\in\Omega/G}\sum_p A_{K_O}(p)\log(p)g_{X,k}(\log(p))\Big|. \tag{4.3}$$

We estimate the number of primes p contributing to (4.3) just as we did in the previous proof and find that there are at most

$$2(\tau^{-1}+1)d\log(4d)\exp(X/10)$$

such primes.

Since $g_{X,k}(p) \leqslant p^{-1}$, each such prime contributes to (4.3) at most $dX\exp(-X/2)$. Therefore,

$$(4.3) \leqslant 2(A+1)d\log(4d)\exp(X/10)\cdot dX\exp(-X/2) \leqslant O(\exp(-c(\log d)^{\alpha-\beta})),$$

and the claim follows. □

5. Equidistribution of Random walks

We study equidistribution of certain random walks in this section. The basic example of these is the walk on $\mathbb{F}_p$ started at 0 whose steps are given by $x \mapsto \alpha x \pm 1$, where $\alpha \in \mathbb{F}_p$ is a fixed parameter and the signs $\pm$ are chosen independently at random with equal probabilities at each step. The study of related random walks goes back to [12, 20], but those studies are mostly concerned with the case, when α is a fixed integer independent of p. Much less is known if α is allowed to vary with p.

We will also need to consider direct products of such walks. Before introducing our notation for the general case, we first outline the arguments in the basic setup mentioned above. We write $\nu_\alpha^{(d)}$ for the probability measure on $\mathbb{F}_p$ that is the distribution of the random walk after $d+1$ steps. It is easily seen that $\nu_\alpha^{(d)}$ is the law of the random variable $S_d(\alpha) := \sum_{j=0}^d X_j\alpha^j \in \mathbb{F}_p$, where $X_j \in \{-1,1\}$ are independent unbiased random variables, and we can write its Fourier transform as

$$\widehat{\nu}_\alpha^{(d)}(\xi) := \mathbb{E}\big(\exp(\frac{2i\pi S_d(\alpha)}{p})\big) = \prod_{j=0}^d \cos(2\pi\alpha^j\xi/p).$$

Our first aim is to bound $|\widehat{\nu}_\alpha^{(d)}(\xi)|$ away from 1. Expanding cos in power series at 0, we see that we need to give a lower bound for $\sum([\alpha^j\xi]^\sim)^2$, where $[\cdot]^\sim$ denotes the unique lift of an element in $\mathbb{F}_p$ to $(-p/2, p/2)\cap\mathbb{Z}$. We will use some bounds of Konyagin [24] for this purpose. We note that Bukh, Harper and Helfgott have used similar ideas in unpublished work (see [19, footnote 4 on page 372]) in order to bound the diameter of the graph underlying the random walk.

Once we have bounded away $|\widehat{\nu}_\alpha^{(d)}(\xi)|$ from 1, we can exploit the fact that $\nu_\alpha^{(d)}$ is a convolution product of the form $\nu_\alpha^{(-1,d_1)} * \cdots * \nu_\alpha^{(d_k,d)}$, where $0 \leqslant d_1 \leqslant \ldots d_k \leqslant d$ are some integers and $\nu_\alpha^{(d_1,d_2)}$ is the law of $\sum_{j=d_1+1}^{d_2} X_j \alpha^j \in \mathbb{F}_p$. We can also bound $|\widehat{\nu}_\alpha^{(d_j,d_{j+1})}(\xi)|$ away from 1 in a similar manner. Multiplying these bounds together we can get sufficiently strong bounds for $|\widehat{\nu}_\alpha^{(d)}(\xi)|$ so that we can deduce that the random walk is equidistributed using Parseval's formula.

Since we do not need an equidistribution result for each individual parameter α, we can improve the above argument by giving an initial estimate for the sum of the L^2 norms:

$$\sum_{\alpha\in\mathbb{F}_p} \|\nu_\alpha^{(d)}\|_2^2 = \sum_{\alpha\in\mathbb{F}_p} \frac{1}{2^{2d+2}} |\{(x_0,\ldots,x_d),(x_0',\ldots,x_d') \in \{\pm 1\}^{d+1} : \\ x_0 + \ldots + x_d\alpha^d = x_0' + \ldots + x_d'\alpha^d\}|.$$

Such an initial bound can be given by exploiting the fact that the polynomial equation

$$x_0 + \ldots + x_d\alpha^d = x_0' + \ldots + x_d'\alpha^d$$

may have at most d roots in $\mathbb{F}_p$ unless $x_j = x_j'$ for all j.

The rest of the section is organized as follows. We set out our general framework in the next section and state the equidistribution result we will use later. In Section 5.2, we give a generalized exposition of Konyagin's argument in our setup with some slight quantitative improvements. Then we use it to deduce an estimate for the Fourier coefficients of $\nu_\alpha^{(d)}$. We prove our main equidistribution result (Proposition 23) in Section 5.3. Finally, in Section 5.4 we prove another equidistribution statement that we need exclusively to bound the probability that a random polynomial is a proper power.

In this paper, we focus only on those equidistribution results that we need in our applications. We believe that these random walks are of independent interest and we will study them further in a subsequent paper.

5.1. **The general setting and results.** We use the following notation throughout this section. Let $M \in \mathbb{Z}_{>0}$. Let $p_1,\ldots,p_M$ be distinct primes (say each $\geqslant 5$) and let $m_1,\ldots,m_M \in \mathbb{Z}_{>0}$ be numbers. Let

$$V = \bigoplus_{i=1}^{M} \mathbb{F}_{p_i}^{m_i},$$
$$D = \max_{i=1,\ldots,M} m_i,$$
$$Q = p_1 \cdots p_M,$$

For $\alpha = (\alpha_1,\ldots,\alpha_M) \in V$ we write $\alpha_i := (\alpha_{i,1},\ldots,\alpha_{i,m_i}) \in \mathbb{F}_{p_i}^{m_i}$ and for another $\beta \in V$ we write $\alpha\beta = (\alpha_{i,j}\beta_{i,j})_{i,j}$, so for instance if $n \in \mathbb{Z}$, we have $\alpha^n = (\alpha_{i,j}^n)_{i,j}$.

We have a canonical isomorphism of additive groups between $\oplus_i^M \mathbb{F}_{p_i}$ and $\mathbb{Z}/Q\mathbb{Z}$ given by

$$\Psi : (x_1, \dots, x_M) \mapsto \sum_{i=1}^{M} \psi_i(x_i),$$

where ψ_i denotes the additive homomorphism $\mathbb{F}_{p_i} \hookrightarrow \mathbb{Z}/Q\mathbb{Z}$, $x \mapsto \frac{Q}{p_i}x$. Moreover we have the trace map $\mathrm{tr} : V \to \oplus_i^M \mathbb{F}_{p_i}$ given by

$$\mathrm{tr}(\alpha) = (\sum_{j=1}^{m_1} \alpha_{1,j}, \dots, \sum_{j=1}^{m_M} \alpha_{M,j}).$$

Let $X_0, X_1, \dots, X_d$ be a sequence of independent $\mathbb{Z}$ valued random variables. We assume that $X_1, \dots, X_{d-1}$ are identically distributed and write μ for their common law. We will study the random walk in the additive group $(V, +)$ whose n-th step is $\sum_{j=0}^{n} X_j \alpha^j$ and denote by $\nu_\alpha^{(n)}$ the law of this random element.

Our decision to exempt X_0 and X_d from having the same distribution as the other steps of the walk is motivated by our intention to permit families of random polynomials whose leading and constant terms have a distribution that differs from the rest. Our method would allow us to relax the requirement of identical distribution further by allowing small perturbations of the same law and a small number of exceptional steps. We leave it to the interested reader to formulate such a statement.

Definition 22. *We say that $\alpha \in V$ is* generic *if for each $i \in [1, M]$ the coordinates $(\alpha_{i,j})_{1 \leqslant j \leqslant m_i}$ are non-zero and pairwise distinct.*

For $\kappa \in (0, \frac{1}{100})$, we write $\mathcal{A}_\kappa \subset V$ for the set of generic $\alpha \in V$ such that none of the coordinates $\alpha_{i,j}$ is a root of a polynomial $P \in \mathbb{Z}[x]$ with $\deg P \leqslant 3\log(Q^D)$ and Mahler measure $M(P) \leqslant \exp(\kappa)$.

The aim of this section is to prove the following result, which asserts under suitable conditions that the probability that the random walk is at 0 after d steps is approximately $|V|^{-1}$ on average for parameters $\alpha \in \mathcal{A}_\kappa$.

Proposition 23. *There are absolute constants $c, C > 0$ such that the following holds. Let $d \in \mathbb{Z}_{>0}$, $0 < \tau < 1$. Suppose that*

$$\begin{aligned} d \geqslant & C(\kappa\tau)^{-1} MD \log Q^D (\log\log Q^D)^3, \\ \log(Q^D) \geqslant & \max\left(\frac{1}{\kappa}, \tau^{-1}\right), \\ \|\mu\|_2^2 := & \sum_{x \in \mathbb{Z}} \mu(x)^2 < 1 - \tau. \end{aligned}$$

Suppose further that $\operatorname{supp}\mu \subset (-p_i/2, p_i/2)$ *for each $i = 1, \dots, M$.*

Then

$$\left| \sum_{\alpha \in A} \nu_\alpha^{(d)}(0) - \frac{|A|}{|V|} \right| < \exp\Big(- c \frac{\tau \kappa d}{\log Q^D (\log\log Q^D)^2} \Big)$$

for any $A \subset \mathcal{A}_\kappa$.

Remark. This proposition will be used in Section 6 twice. Once with $M = 1$ and a large prime p and fixed power m. And another time with $M = 2$ and $m_1 = m_2$. For the theorems of the introduction, except part (2) of Theorem 2 about the generic Galois group, it is enough to consider the case $m = 1$.

5.2. **Estimates for the Fourier coefficients of the random walk.** The aim of this section is to revisit an argument of Konyagin from [24] to obtain Proposition 24 below. Then we will use it in Proposition 25 to deduce a bound for the Fourier coefficients of the random walk.

For each $\alpha \in \mathbb{Z}/Q\mathbb{Z}$, we write $\widetilde{\alpha}$ for the unique lift of α in $\mathbb{Z} \cap (-Q/2, Q/2]$.

Proposition 24. *Let notation be as in Section 5.1. Let* $\alpha, \beta \in V$. *Assume that* α *is generic and* $\beta_{i,j} \neq 0$ *for all* i, j. *Write*

$$S_n = \Psi \circ \mathrm{tr}(\beta \alpha^n) \in \mathbb{Z}/Q\mathbb{Z}.$$

Let $L \geqslant 200 \log Q^D \log\log Q^D$ *be an integer and suppose that*

$$\sum_{n=0}^{L} \widetilde{S}_n^2 \leqslant \frac{Q^2}{8 \log(4L)}. \tag{5.1}$$

Then for each $i = 1, \dots, M$ *and* $j = 1, \dots, m_i$, *there is* $P_{i,j} \in \mathbb{Z}[x]$ *of degree at most* $3 \log(Q^D)$ *with Mahler measure at most* $(\log(Q^D))^{30(\log Q^D)/L}$ *such that* $P_{i,j}(\alpha_{i,j}) = 0$.

Write $e_Q(y) = \exp(-2\pi i y/Q)$ for $y \in \mathbb{Z}/Q\mathbb{Z}$. Given $\beta \in V$ the function $\chi_\beta : V \to \mathbb{C}^\times$

$$\chi_\beta : x \mapsto e_Q\Big(\Psi \circ \mathrm{tr}(\beta x)\Big) \tag{5.2}$$

is a complex character of the additive group $(V, +)$ and every character is of this form. Given a measure ν on V, we use the following notation for its Fourier transform:

$$\widehat{\nu}(\beta) = \sum_{x \in V} \chi_\beta(x) \nu(x).$$

We write $\nu_\alpha^{(l_1,l_2)}$ for the law of the random element in V given by $\sum_{n=l_1+1}^{l_2} X_n \alpha^n$. In this notation $\nu_\alpha^{(d)} = \nu_\alpha^{(-1,d)}$.

Proposition 25. *Let notation be as in the beginning of the section. Let* $\alpha, \beta \in V$. *Assume that* α *is generic and* $\beta_{i,j} \neq 0$ *for all* i, j. *Suppose further that* $\mathrm{supp}\, \mu \subset (-p_i/2, p_i/2)$ *for each* $i = 1, \dots, M$.

Let $L \geqslant 200 \log Q^D \log(\log Q^D)$ be an integer and suppose that there are i, j such that $\alpha_{i,j}$ is not a root of an integer polynomial of degree at most $3 \log Q^D$ with Mahler measure at most $(\log Q^D)^{30 \log(Q^D)/L}$.

Then

$$|\widehat{\nu_\alpha^{(l_1,l_2)}}(\beta)| \leqslant \exp\Big(- \frac{1 - \|\mu\|_2^2}{8 \log(4L)} \Big)$$

for all $0 \leqslant l_1 < l_2 < d$ such that $l_2 - l_1 > L$.

First, we focus on the proof of Proposition 24, which closely follows Konyagin [24]. Using a pigeon hole argument, it is easy to find non-zero polynomials $P_1, P_2 \in \mathbb{Z}[x]$ of degree at most $\log Q^D$ with $\pm 1, 0$ coefficients such that $P_1(\alpha_{i,j}) = P_2(\alpha_{i,j}^q) = 0$ for all i, j. Here q is a carefully chosen prime number. The heart of the argument is the idea that when (5.1) holds, it is possible to find P_1 and P_2 in such a way that for each i, j, there is $P_{i,j} \in \mathbb{Z}[x] \backslash \{0\}$ such that $P_{i,j}(\alpha_{i,j}) = 0$ and $P_{i,j}(x) | \operatorname{GCD}(P_1(x), P_2(x^q))$. From this, we will conclude that $\deg(P_{i,j}) \leqslant \deg(P_1)$ and $M(P_{i,j}) \leqslant M(P_2)^{1/q}$.

We begin to implement this strategy. Given a monic irreducible polynomial $P \in \mathbb{Z}[x]$, the next lemma allows us to find a prime q of controlled size such that whenever $P(x) | Q(x^q)$ for another polynomial $Q \in \mathbb{Z}[x]$, we have $M(P) \leqslant M(Q)^{1/q}$.

Lemma 26. *(See also [24, Lemma 2]) Let $\alpha_1, \dots, \alpha_n$ be the roots of an irreducible polynomial $P \in \mathbb{Z}[x]$. Let $s \geqslant 4 \log n$ be a number. If n is larger than some absolute constant, then there is a prime $q \in (s, 2s]$ such that α_i / α_j is not a q-th root of unity for any $i \neq j$.*

Proof. Denote by $\mathcal{P}_s$ the set of primes between s and $2s$. Write $\mathcal{R}$ for the collection of integers r such that there is a root of unity of order r among the numbers α_i / α_j. Suppose by way of contradiction that $\mathcal{P}_s \subset \mathcal{R}$.

We begin with the observation that if $r \in \mathcal{R}$, then there is $1 \leqslant j \leqslant n$ such that α_1 / α_j is a root of unity of order r. Indeed, let i, j be such that α_i / α_j is a root of unity of order r and let σ be an automorphism of $\overline{\mathbb{Q}}$ such that $\sigma(\alpha_i) = \alpha_1$. Then $\sigma(\alpha_i / \alpha_j) = \alpha_1 / \sigma(\alpha_j)$ is a root of unity of order r and this proves the claim.

Suppose $r_1, r_2 \in \mathcal{R}$ are coprime integers. We prove $r_1 r_2 \in \mathcal{R}$. Let i and j be such that α_1 / α_i and α_1 / α_j are roots of unity of order r_1 and r_2 respectively. Then $\alpha_i / \alpha_j = (\alpha_1 / \alpha_i)^{-1} (\alpha_1 / \alpha_j)$ is a root of unity of order $r_1 r_2$, which proves the claim.

Therefore, for each divisor r of $\prod_{p \in \mathcal{P}_s} p$ belongs to $\mathcal{R}$. Since the set of roots $\alpha_1, \dots, \alpha_n$ is invariant under the action of the Galois group, it follows that for any such r all roots of unity of order r appear among the α_i / α_j's. Hence

$$n^2 \geqslant \prod_{p \in \mathcal{P}_s} p = \exp\Big(\sum_{p \in \mathcal{P}_s} \log p \Big).$$

By the prime number theorem, we have

$$n^2 \geqslant \exp(s/2)$$

if n and hence s is sufficiently large. (In fact, we could put here any constant less than 1 in place of 1/2.) This proves the lemma. □

Let $N \geqslant e \geqslant 0$ be integers and $X = (x_0, \ldots, x_N)$ a sequence of integers. Following Konyagin [24], we write $\Lambda_e(X)$ for the set of polynomials $P(x) = a_0 + a_1 x + \ldots + a_e x^e \in \mathbb{Z}[x]$ of degree at most $e \in \mathbb{Z}_{\geqslant 0}$ such that

$$a_0 x_j + \ldots + a_e x_{j+e} = 0$$

holds for all $j = 0, \ldots, N-e$. We denote by $\Lambda(X)$ the set of polynomials P of degree at most N such that $P \in \Lambda_{\deg P}(X)$. We note that $P \in \Lambda_e(x_0, \ldots, x_N)$ if and only if $P \in \Lambda(x_0, \ldots, x_{N-e+\deg P})$.

If X were an infinite sequence, then Λ would give rise to a principal ideal in $\mathbb{Z}[x]$. We need a weaker form of this fact that is valid for finite sequences. To this end, we recall the following result.

Lemma 27 (Konyagin [24, Lemma 5])**.** *Let $X = (x_0, \ldots, x_N)$ be a sequence of integers and let $P_1, P_2 \in \Lambda(X)$. If $\deg P_1 + \deg P_2 \leqslant N$, then we have $\gcd(P_1, P_2) \in \Lambda(X)$.*

Corollary 28. *Let $X = (x_0, \ldots, x_N)$ be a sequence of integers and suppose that $\Lambda(X)$ contains a non-zero polynomial of degree at most $N/2$. Then there is a unique (up to multiplication by ± 1) non-zero polynomial $P_0 \in \Lambda(X)$ of minimal degree and with relatively prime coefficients. Furthermore, a polynomial $P \in \mathbb{Z}[x]$ of degree at most $N - \deg P_0$ is contained in $\Lambda(X)$ if and only if $P_0 | P$.*

In the proof of Proposition 24 below, we will use the above results for the sequence $x_n = \widetilde{S}_n$. Under the hypothesis (5.1), we will show that there are many polynomials $P(x) = a_0 + \ldots + a_e x^e$ such that $|a_0 \widetilde{S}_j + \ldots + a_e \widetilde{S}_{e+j}| < Q$ for all j in a certain range. Using the pigeonhole principle, we will find a polynomial that in addition satisfies

$$a_0 S_j + \ldots + a_e S_{e+j} = 0 \tag{5.3}$$

in the same range for j's. These two properties imply that $P \in \Lambda(X)$ for $X = (S_n)$. The next lemma shows that it is enough to satisfy (5.3) for a smaller range of j's.

Lemma 29. *Let $m \in \mathbb{Z}_{>0}$ and let p be a prime. For each $1 \leqslant j \leqslant m$, let $\alpha_j, \beta_j \in \mathbb{F}_p$. Write*

$$T_n = \sum_{j=1}^{m} \beta_j \alpha_j^n$$

for each $i \in \mathbb{Z}_{\geqslant 0}$. Let $P(x) = a_0 + \ldots + a_e x^e \in \mathbb{F}_p[x]$ be a polynomial.

Suppose that the elements α_j are pairwise distinct and $\beta_j \neq 0$ for all j. Suppose further that

$$a_0 T_n + a_1 T_{n+1} + \ldots + a_e T_{n+e} = 0$$

for all $n = 0, \ldots, m-1$.

Then $P(\alpha_j) = 0$ for all $1 \leqslant j \leqslant m$ and

$$a_0 T_n + a_1 T_{n+1} + \ldots + a_e T_{n+e} = 0$$

for all $n \in \mathbb{Z}_{\geqslant 0}$.

Proof. The hypothesis of the lemma implies that

$$\sum_{j=1}^{m} \beta_j \alpha_j^n P(\alpha_j) = 0, \tag{5.4}$$

for each $n = 0, \ldots, m-1$.

We note that the vectors $(\beta_1 \alpha_1^n, \ldots, \beta_m \alpha_m^n)$ for $n = 0, \ldots, m-1$ are linearly independent, as can be seen using Vandermonde determinants. Hence the system of linear equations (5.4) in $P(\alpha_j)$ as variables has only the trivial solution, that is $P(\alpha_j) = 0$ for each j, which proves the first claim. In addition,

$$a_0 T_n + a_1 T_{n+1} + \ldots + a_e T_{n+e} = \sum_{j=1}^{m} \beta_j \alpha_j^n P(\alpha_j) = 0$$

for each $n \in \mathbb{Z}_{\geqslant 0}$ and this establishes the second claim. □

Proof of Proposition 24. Set $E = 3\lfloor \log Q^D \rfloor$. Note that $1 \leqslant E \leqslant L/3$ and $2^E > Q^D \geqslant 3$ and $\lfloor L/6E \rfloor \geqslant 4 \log E$. Our first aim is to show that there is a polynomial $P_1 \neq 0$ of degree at most E such that $P_1 \in \Lambda_E(\{\widetilde{S}_n\}_{n=0}^{L})$.

Let $\xi_0, \ldots, \xi_E$ be a sequence of independent, unbiased ± 1 valued random variables. For any $n = 0, \ldots, L-E$, we have

$$\mathbb{P}\Big(\Big| \sum_{j=0}^{E} \xi_j \widetilde{S}_{j+n} \Big| \geqslant \frac{Q}{2} \Big) \leqslant 2 \exp\Big(- \frac{(Q/2)^2}{2 \sum_0^E \widetilde{S}_{j+n}^2} \Big) \leqslant \frac{1}{2L}$$

by Hoeffding's inequality and our assumption (5.1).

Therefore the set

$$\Omega := \Big\{ x = (x_0, \ldots, x_E) \in \{-1, 1\}^{E+1} : \Big| \sum_{j=0}^{E} x_j \widetilde{S}_{j+n} \Big| < \frac{Q}{2} \quad \text{for all } n = 0, \ldots, L-E \Big\}$$

has cardinality more than $2^{E+1}/2 > Q^D$.

By the pigeonhole principle, it follows that there are $x \neq y \in \Omega$ such that

$$\sum_{j=0}^{E} x_j \widetilde{S}_{j+n} = \sum_{j=0}^{E} y_j \widetilde{S}_{j+n}$$

for all $n = 0, \dots, D-1$. We set $a_j = (x_j - y_j)/2 \in \{-1, 0, 1\}$. It follows from Lemma 29 applied M times to each $T_n^i := \sum_{j=1}^{m_i} \beta_{i,j} \alpha_{i,j}^n$ with the polynomial $P = a_0 + \ldots + a_E x^E \in \mathbb{Z}[x]$ that

$$\sum_{j=0}^{E} a_j S_{j+n} = 0 \in \mathbb{Z}/Q\mathbb{Z}$$

for all $n \in \mathbb{Z}_{\geqslant 0}$. Since $x, y \in \Omega$ and $a_j = (x_j - y_j)/2$, we know that

$$\Big| \sum_{j=0}^{E} a_j \widetilde{S}_{j+n} \Big| < Q/2$$

and hence

$$\sum_{j=0}^{E} a_j \widetilde{S}_{j+n} = 0$$

for any $n = 0, \dots, L - E$. This means that

$$P_1(x) := a_E x^E + \ldots + a_1 x + a_0 \in \Lambda_E(\{\widetilde{S}_n\}_{n=0}^{L}) \subset \Lambda(\{\widetilde{S}_n\}_{n=0}^{[2L/3]}),$$

because $E \leqslant L/3$.

Since $[2L/3] \geqslant 2E$, by Corollary 28 there is a unique (up to multiplication by ± 1) polynomial $P_0 \in \Lambda(\{\widetilde{S}_n\}_{n=0}^{[2L/3]})$ with relatively prime coefficients and of minimal degree. By Lemma 29, $P_0(\alpha_{i,j}) = 0$ for all i and j. Then for each i and j there is an irreducible factor $P_{i,j}$ of P_0 over $\mathbb{Z}$ such that $P_{i,j}(\alpha_{i,j}) = 0$.

Fix i and j. We already know that $\deg P_{i,j} \leqslant E$, and we set out to prove $M(P_{i,j}) \leqslant (D \log Q)^{30D \log Q/L}$. Write $\{\beta_k\}_{k=1}^{\deg P_{i_j}}$ for the roots of $P_{i,j}$. We set $s = \lfloor L/6E \rfloor$. Since $L \geqslant 200 D \log Q \log(D \log Q)$, we have $s \geqslant 4 \log E$. By Lemma 26, there is a prime $q \in (s, 2s]$ such that β_k/β_l is not a q-th root of unity for any $k \neq l$. This means that the numbers $\{\beta_k^q\}_{k=1}^{\deg P_{i,j}}$ are all distinct. Note that $q \leqslant 2s \leqslant L/3E$.

We now employ the same argument as above and find a non-zero polynomial P_2 of the form $P_2(x) = Q_2(x^q)$ for some $Q(x) = b_0 + \ldots + b_E x^E \in \mathbb{Z}[x]$ with $b_j \in \{-1, 0, 1\}$ for $j = 1, \dots, E$ and

$$\sum_{k=0}^{E} b_k \widetilde{S}_{kq+n} = 0$$

for any $n = 0, \dots, L - Eq$. Hence $P_2 \in \Lambda_{Eq}(\{\widetilde{S}_n\}_{n=0}^{L}) \subset \Lambda(\{\widetilde{S}_n\}_{n=0}^{[2L/3]})$, because $Eq \leqslant L/3$.

Since $[2L/3] \geqslant 2Eq$, we have $P_0 | P_2$ by Corollary 28. This means that $\{\beta_k^q\}_{k=1}^{\deg P_{i,j}}$ are all roots of the polynomial Q_2, and since they are all distinct we get:

$$M(P_{i,j})^q \leqslant M(Q_2) \leqslant (E+1)^{1/2}.$$

where the right hand side follows from (1.6). Since $q > s = \lfloor L/6E \rfloor$, $2q \geqslant L/3E$, we get

$$M(P_{i,j}) \leqslant (E+1)^{1/2q} \leqslant (E+1)^{3E/L} \leqslant (D \log Q)^{30D \log Q/L}.$$

□

Remark 30. Let $u_1, \dots, u_m \in \mathbb{R}^n$ be a sequence of vectors with $\|u_j\| \leqslant 1$. A conjecture of Komlós asserts that there is an absolute constant C such that for each such sequence of vectors, there is a sequence of signs $\omega_j = \pm 1$ such that $\|\omega_1 u_1 + \ldots + \omega_m u_m\|_\infty < C$. In this remark, we point out that if this conjecture holds, then assumption (5.1) in Proposition 24, may be relaxed to an upper bound of the form cQ^2, where c is an absolute constant. Unfortunately, the best known result towards Komlós's conjecture in [2] yields no improvement.

We take a sequence $y = (y_0, \dots, y_E) \in \{0,1\}^{E+1}$ and we will apply the conjecture to the vectors $u_j = (y_j \widetilde{S}_j, \dots, y_j \widetilde{S}_{L-E+j})$ for $j = 0, \dots, E$. Under the weakened hypothesis $\sum \widetilde{S}_n^2 \leqslant cQ^2$, Komlós's conjecture implies that for each choice of y, there is $\omega(y) = (\omega_0(y), \dots, \omega_E(y)) \in \{\pm 1\}^{E+1}$ such that

$$\Big| \sum_{j=0}^{E} \omega_j(y) y_j \widetilde{S}_{j+i} \Big| < Q/2$$

for all $i = 0, \dots, L-E$. Now we see that the collection of sequences of the form $(\omega_0(y)y_0, \dots, \omega_E(y)y_E) \in \{-1,0,1\}^E$ may be used in the place of Ω in the proof of Proposition 24 to obtain the same result under the weaker hypothesis.

Now we turn to the proof of Proposition 25.

Given a measure μ on $\mathbb{Z}$ and $x \in V$ we write $\mu.\delta_x$ for the measure on V defined by

$$\sum_{a \in \mathbb{Z}} \mu(a) \delta_{ax}.$$

With this notation, we can write

$$\nu_\alpha^{(l_1,l_2)} = \mu.\delta_{\alpha^{l_1+1}} * \ldots * \mu.\delta_{\alpha^{l_2}}$$

for any $0 \leqslant l_1 < l_2 < d$.

Recall the notation $\widetilde{a}$, which is the unique integer representative of $a \in \mathbb{Z}/Q\mathbb{Z}$ in the interval $(-Q/2, Q/2]$. For typographical convenience we will also use the notation $[a]^\sim$ with the same meaning.

Lemma 31. *Let μ be a measure on $\mathbb{Z}$ and let $x, \beta \in V$. Then*

$$|\widehat{\mu.\delta_x}(\beta)| \leqslant \exp\Big(- \sum_{a_1,a_2 \in \mathbb{Z}} \mu(a_1)\mu(a_2) \Big(\Big[\Psi \circ \mathrm{tr}((a_1-a_2)\beta x) \Big]^\sim \Big)^2 / Q^2 \Big).$$

Proof. By definition, we have

$$|\widehat{\mu.\delta_x}(\beta)|^2 = \sum_{a_1,a_2 \in \mathbb{Z}} \mu(a_1)\mu(a_2) \chi_\beta((a_1-a_2)x)$$

Since $|\widehat{\mu.\delta_x}(\beta)|^2 \in \mathbb{R}$ and $\mathrm{Re}(e_Q(a)) \leqslant \exp(-\widetilde{a}^2/Q^2)$ if $a \in \mathbb{Z}/Q\mathbb{Z}$, the claim now follows from (5.2). □

Proof of Proposition 25. For $\gamma \in V$, write as earlier

$$S_n(\gamma) = \Psi \circ \mathrm{tr}(\gamma\alpha^n) \in \mathbb{Z}/Q\mathbb{Z}.$$

Using

$$|\widehat{\nu_\alpha^{(l_1,l_2)}}(\beta)| = \Big| \prod_{n=l_1+1}^{l_2} \widehat{\mu.\delta_{\alpha^n}}(\beta)\Big|,$$

Lemma 31 implies

$$|\widehat{\nu_\alpha^{(l_1,l_2)}}(\beta)| \leqslant \exp\Big(- \sum_{n=l_1+1}^{l_2} \sum_{a_1,a_2\in\mathbb{Z}} \mu(a_1)\mu(a_2)(\widetilde{S}_n((a_1-a_2)\beta))^2/Q^2\Big).$$

Using $l_2 - l_1 \geqslant L$ and then Proposition 24, we can write

$$\sum_{n=l_1+1}^{l_2} (\widetilde{S}_n((a_1-a_2)\beta))^2 \geqslant \sum_{n=0}^{L} (\widetilde{S}_n((a_1-a_2)\beta\alpha^{l_1+1}))^2 > \frac{Q^2}{2\log(4L)}$$

for all pairs $a_1 \neq a_2$. Since $\mathrm{supp}\,\mu \subset (-p_i/2, p_i/2)$ for all i, we know that $a_1 - a_2$ is non zero in $\mathbb{F}_{p_i}$ for all i whenever $a_1 \neq a_2$ in $\mathbb{Z}$.

We note

$$\sum_{a_1\neq a_2} \mu(a_1)\mu(a_2) = 1 - \|\mu\|_2^2.$$

Therefore

$$|\widehat{\nu_\alpha^{(l_1,l_2)}}(\beta)| \leqslant \exp\Big(- (1-\|\mu\|_2^2)\frac{Q^2}{2\log(4L)} \cdot \frac{1}{Q^2}\Big),$$

as claimed. □

5.3. **Proof of Proposition 23.** We note that $\nu(0) = \sum_{\beta\in V} \widehat{\nu}(\beta)/|V|$ for any measure ν on V. Hence, for any $A \subset \mathcal{A}_\kappa$:

$$\Big| \sum_{\alpha\in A} \nu_\alpha^{(d)}(0) - \frac{|A|}{|V|}\Big| \leqslant \frac{1}{|V|} \sum_{\alpha\in A} \sum_{\beta\in V\setminus\{0\}} |\widehat{\nu}_\alpha^{(d)}(\beta)|. \tag{5.5}$$

We begin by finding a preliminary estimate for

$$\sum_{\alpha\in A} \|\nu_\alpha^{(d_1,d_2)}\|_2^2 \leqslant \sum_{\alpha\in V} \|\nu_\alpha^{(d_1,d_2)}\|_2^2.$$

and then use the Cauchy-Schwarz inequality to convert it into an estimate on the right hand side of (5.5). Let $(A_n)_{n=0,\dots,d}$ and $(A'_n)_{n=0,\dots,d}$ be sequences of independent random variables with the same law as $(X_n)_{n=0,\dots,d}$. We observe that

$$\begin{aligned}\sum_{\alpha\in V} \|\nu_\alpha^{(d_1,d_2)}\|_2^2 &= \sum_{\alpha\in V} \mathbb{P}(A_{d_1+1}\alpha^{d_1+1} + \ldots + A_{d_2}\alpha^{d_2} = A'_{d_1+1}\alpha^{d_1+1} + \ldots + A'_{d_2}\alpha^{d_2}) \\ &= \mathbb{E}(\#\{\alpha \in V : A_{d_1+1}\alpha^{d_1+1} + \ldots + A_{d_2}\alpha^{d_2} = A'_{d_1+1}\alpha^{d_1+1} + \ldots + A'_{d_2}\alpha^{d_2}\}).\end{aligned}$$

If $A_j \neq A'_j$ for at least one $j \in (d_1, d_2]$, then the polynomial

$$(A_{d_1+1} - A'_{d_1+1})x^{d_1+1} + \ldots + (A_{d_2} - A'_{d_2})x^{d_2}$$

has at most $d_2 - d_1$ roots in any given field. This means that for such A_j and A'_j

$$\#\{\alpha \in V : A_{d_1+1}\alpha^{d_1+1}+\ldots+A_{d_2}\alpha^{d_2} = A'_{d_1+1}\alpha^{d_1+1}+\ldots+A'_{d_2}\alpha^{d_2}\} \leqslant (d_2-d_1)^{MD}$$

and

$$\sum_{\alpha\in V} \|\nu_\alpha^{(d_1,d_2)}\|_2^2 \leqslant (d_2-d_1)^{MD}+|V|\sum_{a\in\mathbb{Z}} \mu(a)^{2(d_2-d_1)} \leqslant (d_2-d_1)^{MD}+|V|(1-\tau)^{(d_2-d_1)}.$$

We set $d_0 = \lceil -\log(|V|)/\log(1-\tau)\rceil$, and obtain

$$\frac{1}{|V|}\sum_{\alpha\in V}\sum_{\beta\in V} |\widehat{\nu}_\alpha^{(d,d+d_0)}(\beta)|^2 = \sum_{\alpha\in V} \|\nu_\alpha^{(d,d+d_0)}\|_2^2 \leqslant 2d_0^{MD}$$

for all d. We note that $\nu_\alpha^{(2d_0)} = \nu_\alpha^{(-1,0)} * \nu_\alpha^{(0,d_0)} * \nu_\alpha^{(d_0,2d_0)}$. Therefore for each $\alpha \in V\backslash\{0\}$, we have, since $|\widehat{\nu}_\alpha^{(-1,0)}(\beta)| \leqslant 1$,

$$\begin{aligned}\frac{1}{|V|}\sum_{\beta\in V} |\widehat{\nu}_\alpha^{(2d_0)}(\beta)| \leqslant& \frac{1}{|V|}\sum_{\beta\in V} |\widehat{\nu}_\alpha^{(0,d_0)}(\beta)\cdot\widehat{\nu}_\alpha^{(d_0,2d_0)}(\beta)| \\ \leqslant& \Big[\frac{1}{|V|}\sum_{\beta\in V} |\widehat{\nu}_\alpha^{(0,d_0)}(\beta)|^2\Big]^{1/2}\cdot\Big[\frac{1}{|V|}\sum_{\beta\in V} |\widehat{\nu}_\alpha^{(d_0,2d_0)}(\beta)|^2\Big]^{1/2}\end{aligned}$$

This gives us by another application of Cauchy-Schwarz

$$\begin{aligned}&\frac{1}{|V|}\sum_{\alpha\in A}\sum_{\beta\in V} |\widehat{\nu}_\alpha^{(2d_0)}(\beta)| \\ &\quad\leqslant\Big[\frac{1}{|V|}\sum_{\alpha\in A}\sum_{\beta\in V} |\widehat{\nu}_\alpha^{(0,d_0)}(\beta)|^2\Big]^{1/2}\cdot\Big[\frac{1}{|V|}\sum_{\alpha\in A}\sum_{\beta\in V} |\widehat{\nu}_\alpha^{(d_0,2d_0)}(\beta)|^2\Big]^{1/2} \\ &\quad\leqslant 2d_0^{MD}. \qquad (5.6)\end{aligned}$$

Now we set $d_1 = \lceil \frac{200}{\kappa}\log(Q^D)\log\log(Q^D)\rceil$. If $\alpha \in A \subset \mathcal{A}_\kappa$, then $\alpha_{i,j}$ is not a root of a polynomial of degree at most $3\log Q^D$ with Mahler measure at most $(\log Q^D)^{30\log(Q^D)/d_1}$, and we also have $d_1 \geqslant 200\log Q^D\log(\log Q^D)$. Therefore, we can apply Proposition 25 with $L = d_1$ and get

$$|\widehat{\nu}_\alpha^{(d,d+d_1)}(\beta)| \leqslant \exp\Big(-\frac{\tau}{8\log(4d_1)}\Big) \qquad (5.7)$$

for all d, $\alpha \in \mathcal{A}_\kappa$ and $\beta \in V\backslash\{0\}$. (If β has some 0 coordinates, then V splits as a direct sum $V = V_0 \oplus V_1$, with $\beta \in V_1$ having no non-zero coordinate in V_1, and we need to apply the proposition to V_1 and the projected random walk on V_1 modulo V_0.)

Now suppose that $d > 2d_0 + Kd_1$ for some $K \in \mathbb{Z}_{\geqslant 0}$ and write

$$|\widehat{\nu}_\alpha^{(d)}(\beta)| \leqslant |\widehat{\nu}_\alpha^{(2d_0)}(\beta)|\cdot|\widehat{\nu}_\alpha^{(2d_0,2d_0+d_1)}(\beta)|\cdots|\widehat{\nu}_\alpha^{(2d_0+(K-1)d_1,2d_0+Kd_1)}(\beta)|.$$

We combine (5.6) with (5.7) and obtain

$$\frac{1}{|V|}\sum_{\alpha\in\mathcal{A}_\kappa}\sum_{\beta\in V}|\widehat{\nu}_\alpha^{(d)}(\beta)| \leqslant 2d_0^{MD}\exp\Big(-K\frac{\tau}{8\log(4d_1)}\Big).$$

By the assumption on d in the proposition, we can take $K > d/2d_1$ and a simple calculation yields that

$$2d_0^{MD} < \exp\Big(K\frac{\tau}{16\log(4d_1)}\Big)$$

and hence we obtain the claim of the proposition. In the interest of these calculations, it is useful to note that the lower bounds on $\log Q^D$ in terms of κ and τ that we assumed in the proposition implies that

$$\max(\log d_0, \log d_1) \leqslant C\log\log Q^D.$$

5.4. **The case $\alpha = 2$.** In this section, we consider the special case $V = \mathbb{F}_{p_1}\oplus\ldots\oplus\mathbb{F}_{p_M}$ and $\alpha_{i,1} = 2$ for all i. We write $\nu_2^{(d_1,d_2)}$ for the measure $\nu_\alpha^{(d_1,d_2)}$ with the above choice of V and α. We will use this case later to estimate the probability that $P(2)$ is a proper power for a random polynomial, which, in turn, yields an estimate for the probability that P is a proper power of a polynomial.

Our main result is the following.

Proposition 32. *Let $\tau > 0$ and assume $\|\mu\|_2^2 \leqslant 1-\tau$. Suppose further that $\operatorname{supp}\mu \subset (-p_i/2, p_i/2)$ for each $i = 1,\ldots,M$.*

There is an absolute constant $C > 0$ such that for all $x \in V$ and $d \geqslant \frac{1}{\tau}(C\log(Q))^2$, we have

$$|\nu_2^{(d)}(x) - Q^{-1}| \leqslant Q^{-10}.$$

The study of this case goes back to Chung, Diaconis and Graham [12], who obtained very precise estimates for the mixing time, which are much better than the bound $\log(Q)^2$ implied by the above result. However, our application requires strong bounds for the distance between $\nu_2^{(d)}$ and the uniform distribution, which was not considered in [12]. Nevertheless, our proof draws on the ideas of [12] heavily.

We begin with a lemma on the Fourier coefficients of $\nu_2^{(d)}$. Its proof relies on Lemma 31 and on the elementary fact that a sequence of the form $\widetilde{\beta}, \widetilde{\beta\cdot 2}, \ldots [\beta\cdot 2^{\lfloor\log_2 Q\rfloor}]^\sim$ cannot stay below $Q/4$.

Lemma 33. *Let $q = \lfloor\log_2(Q)\rfloor$. Then for any $l \in [0, d-q)$ and $\beta \in V\backslash\{0\}$, we have*

$$|\widehat{\nu_2}^{(l,l+q)}(\beta)| \leqslant \exp\Big(-\frac{1-\|\mu\|_2^2}{16}\Big).$$

Proof. By Lemma 31, we have

$$|\widehat{\nu_2}^{(l,l+q)}(\beta)| = \prod_{j=1}^{q} |\widehat{\nu_2}^{(l+j-1,l+j)}(\beta)|$$

$$\leqslant \exp\Big(- \sum_{n=l+1}^{l+q} \sum_{a_1,a_2\in\mathbb{Z}} \mu(a_1)\mu(a_2) R(a_1,a_2,n)^2/Q^2\Big),$$

where

$$R(a_1,a_2,n) = \Big[\Psi((a_1-a_2)2^n\beta)\Big]^{\sim}.$$

We note that $R(a_1,a_2,n+1) \equiv 2R(a_1,a_2,n) \mod Q$. Therefore, if $|R(a_1,a_2,n)| < Q/4$, then $|R(a_1,a_2,n+1)| = 2|R(a_1,a_2,n)|$. Now, it is easy to see that for any $a_1 \neq a_2$, there is $n \in [l+1, l+q]$ such that $|R(a_1,a_2,n)| \geqslant Q/4$, and the claim follows. □

Proof of Proposition 32. We note that

$$\nu_2^{(d)}(x) - Q^{-1} = \frac{1}{|Q|} \sum_{\beta\in V\setminus\{0\}} \widehat{\nu}_2^{(d)}(\beta)\chi_\beta(x),$$

hence it is enough to prove that for all $\beta \in V\setminus\{0\}$,

$$|\widehat{\nu}_2^{(d)}(\beta)| < Q^{-10}.$$

To that end, we choose an integer $L \leqslant d/q$ and write

$$\nu_2^{(d)} = \nu_2^{(-1,0)} * \nu_2^{(0,q)} * \ldots * \nu_2^{((L-1)q,Lq)} * \nu_2^{(Lq,d)}$$

and note that Lemma 33 implies

$$|\widehat{\nu}_2^{(d)}(\beta)| < \exp(-\tau L/16).$$

This yields the desired estimate, if we set $L = \lceil 160\log(Q)/\tau\rceil$, which is permitted if the constant C is taken sufficiently large. □

6. Expected number of roots of a random polynomial

In this section, we use the results of the previous section to calculate the expected number of roots of a typical polynomial in $\mathbb{F}_p$ for a random prime. In the proofs of our main result, we will compare these with the formulae in Section 4.

Let $m, d \geqslant 1$, $\kappa \in (0, \frac{1}{100})$ and $X > 10$. For a random polynomial $P \in \mathbb{Z}[x]$ of degree at most d, we will now estimate the number $B_P(p)$ of admissible roots of P in $\mathbb{F}_p$ on average over the prime p. Here and below a residue modulo p will be called admissible if it is $(\frac{\kappa}{m}, mX)$-admissible in the notation of Definition 15. For the irreducibility results it will be sufficient to set $m = 1$, but for information on the Galois groups, we will need to consider larger values of m. Nevertheless m will not exceed a fixed power of $\log d$.

We suppose that the coefficients of P, except for the leading coefficient and the constant term, are identically distributed and write μ

for their common law. The notation $\mathbb{E}_P$ is used to denote expectation with respect to the law of the random polynomial P. Our purpose in this section is the prove the following result.

Proposition 34. *There are absolute constants $c_0, C_0 > 0$ such that the following holds. Let $\tau, \kappa > 0$, $d, m \in \mathbb{Z}_{>0}$ and let μ be a probability measure on $\mathbb{Z}$ supported on $[-\exp(d^{1/10}), \exp(d^{1/10})]$. Assume $\|\mu\|_2^2 < 1 - \tau$. Let X be a number such that*

$$100m^2 \max\{\kappa^{-1}, \tau^{-1}, d^{1/10}\} < X < \frac{\kappa\tau}{C_0} \frac{d}{(m \log(md))^3} \tag{6.1}$$

and let $g : \mathbb{R} \to \mathbb{R}_{\geqslant 0}$ be a function such that $\operatorname{supp} g \subset [X/2, X]$, $g(x) \leqslant 2\exp(-x)$ for all x.

Then with probability at least

$$1 - \exp(-\frac{X}{3}) - \exp(\frac{X}{40} - \frac{c_0\tau\kappa d}{mX(\log(md))^2})$$

the following holds for P:

$$|\sum_p B_P(p)^m \log(p) g(\log p) - B_m w| < \frac{1}{2},$$

where B_m stands for the m-th Bell number and $w = \sum_p \log(p) g(\log p)$.

Recall that the Bell number B_m is the number of equivalence classes on a set with m elements. We begin by recording the following consequence of Proposition 23.

Lemma 35. *There are absolute constants $c_0, C_0 > 0$ such that the following holds. Let $\tau, \kappa > 0$ and $m, d \in \mathbb{Z}_{>0}$. Suppose that the probability measure μ on $\mathbb{Z}$ is supported on $[-\exp(d^{1/10}), \exp(d^{1/10})]$ and that $\|\mu\|_2^2 < 1 - \tau$. Let X be such that*

$$10 \max\{\kappa^{-1}, \tau^{-1}, d^{1/10}\} < X < \frac{\kappa\tau}{C_0} \frac{d}{(m \log(md))^3},$$

and let $p, p_1 \neq p_2 \in [\exp(X/2), \exp(X)]$ be primes. Then

$$|\mathbb{E}_P[B_P(p)^m] - B_m| \leqslant B_m \cdot \mathrm{Err}(X, d, m),$$
$$|\mathbb{E}_P[B_P(p_1)^m B_P(p_2)^m] - B_m^2| \leqslant B_m^2 \cdot \mathrm{Err}(X, d, m),$$

where $\mathrm{Err}(X, d, m) = 40m^2 \exp(-\frac{X}{5}) + \exp(-\frac{c_0\tau\kappa d}{mX(\log(md))^2})$.

Proof. We write A_p for the set of $(\frac{\kappa}{m}, mX)$-admissible elements of $\mathbb{F}_p$. In the notation of Section 5 we take $M = 1$ and $V = \mathbb{F}_p^m$. Then

$$\mathbb{E}[B_P(p)^m] = \sum_{\alpha \in (A_p)^m} \nu_\alpha^{(d)}(0).$$

We decompose $(A_p)^m$ as a disjoint union of subsets $(A_p)^m(\varepsilon)$ for which Proposition 23 applies. To this end, we write $\mathcal{E}_m$ for the set of equivalence relations on the set $\{1, \dots, m\}$. For each $\varepsilon \in \mathcal{E}_m$, we let

$V(\varepsilon)$ be the subgroup of V formed by the equations $\alpha_i = \alpha_j$ whenever $(i,j) \in \varepsilon$, and write $(A_p)^m(\varepsilon)$ for the subset of $(A_p)^m \cap V(\varepsilon)$ made of those m-tuples α such that $\alpha_i = \alpha_j$ if and only if $(i,j) \in \varepsilon$.

Given $\varepsilon \in \mathcal{E}_m$ we may apply Proposition 23 to the group $V(\varepsilon) \simeq \mathbb{F}_p^{m_\varepsilon}$, where m_ε is the number of equivalence classes in ε and obtain

$$\left| \sum_{\alpha \in (A_p)^m(\varepsilon)} \nu_\alpha^{(d)}(0) - \frac{|(A_p)^m(\varepsilon)|}{p^{m_\varepsilon}} \right| < \exp(-c_0 \frac{\tau\kappa}{mX} \frac{d}{(\log(md))^2}). \tag{6.2}$$

As we have already noted, the number of polynomials of degree at most $10mX$ and Mahler measure at most $\exp(\kappa/m)$ is at most $\exp(X/10)$ by [17, Theorem 1]. Therefore, $|\mathbb{F}_p \backslash A_p| \leqslant 10mX \exp(X/10)$ and

$$0 \leqslant 1 - \frac{|(A_p)^m(\varepsilon)|}{p^{m_\varepsilon}} \leqslant 10m^2 X \exp(X/10)/p \leqslant 20m^2 \exp(-X/5).$$

Now summing up (6.2) for $\varepsilon \in \mathcal{E}_m$, we arrive at the first claim.

The proof of the second claim is entirely similar using Proposition 23 for the random walk on $V = \mathbb{F}_{p_1}^m \oplus \mathbb{F}_{p_2}^m$. We leave the details to the reader. □

Lemma 36. *Let $X > 10$ and let $g : \mathbb{R} \to \mathbb{R}_{\geqslant 0}$ be a function such that $\operatorname{supp} g \subset [X/2, X]$, $g(x) \leqslant 2\exp(-x)$ for all x. Then*

$$w_2 := \sum_p (\log(p) g(\log p))^2 \leqslant 8X^2 \exp(-X/2)$$

$$w := \sum_p \log(p) g(\log p) \leqslant 4X^2.$$

Proof. A simple calculation yields

$$\sum_p (\log(p) g(\log p))^2 \leqslant \sum_{\exp(X/2) \leqslant n \leqslant \exp(X)} 4\frac{(\log n)^2}{n^2} \leqslant 8X^2 \exp(-X/2),$$

$$\sum_p \log(p) g(\log p) \leqslant \sum_{\exp(X/2) \leqslant n \leqslant \exp(X)} 2\frac{\log n}{n} \leqslant 4X^2.$$

□

Proof of Proposition 34. Let $Z = \sum_p B_P(p)^m \log(p) g(\log p) - B_m w$. We bound $\mathbb{E}(Z^2)$, then apply Chebyshev's inequality

$$\mathbb{P}_P(|Z| \geqslant \frac{1}{2}) \leqslant 4\mathbb{E}_P(Z^2)$$

to prove the claim. Setting $h(x) = \log(x) g(\log x)$ we compute:

$$Z^2 = \sum_{p_1, p_2} (B_P(p_1) - B_m)(B_P(p_2) - B_m) h(p_1) h(p_2)$$

so

$$\begin{aligned}\mathbb{E}_P(Z^2) &= \sum_{p_1,p_2} \mathbb{E}_P((B_P(p_1)-B_m)(B_P(p_2)-B_m))h(p_1)h(p_2)\\ &= \sum_{p_1,p_2} \mathbb{E}_P(B_P(p_1)B_P(p_2))h(p_1)h(p_2)\\ &\quad - 2B_m \sum_{p_1} h_{p_1} \sum_{p_2} \mathbb{E}_P(B_P(p_2))h(p_2) + B_m^2\Big(\sum_p h(p)\Big)^2\\ &= \sum_{p_1,p_2} \mathbb{E}_P(B_P(p_1)B_P(p_2)-B_m^2)h(p_1)h(p_2)\\ &\quad - 2B_m \sum_{p_1} h(p_1) \sum_{p_2} \mathbb{E}_P(B_P(p_2)-B_m)h(p_2)\\ &= \sum_{p_1\neq p_2} \mathbb{E}_P(B_P(p_1)B_P(p_2)-B_m^2)h(p_1)h(p_2)\\ &\quad - 2B_m w \sum_p \mathbb{E}_P(B_P(p)-B_m)h(p) + \sum_p \mathbb{E}_P(B_P(p)^2-B_m^2)h(p)^2.\end{aligned}$$

We use Lemma 35 to bound the first two terms and the crude bounds $B_m \leqslant 2^{m^2} \leqslant \exp(X/100)$ and $B_P(p) \leqslant d^m \leqslant \exp(X/100)$ for the third:

$$\mathbb{E}_P(Z^2) \leqslant \exp(X/50)(3\,\mathrm{Err}(X,d,m)w^2 + 2w_2).$$

Finally we use Lemma 36 and obtain:

$$4\mathbb{E}_P(Z^2) \leqslant \exp(-X/3) + \exp\Big(\frac{X}{40} - \frac{c_0\tau\kappa d}{mX(\log(md))^2}\Big).$$

□

7. Polynomials of small Mahler measure

In this section, we estimate the probability that the random polynomial P is divisible by a non-cyclotomic polynomial of small Mahler measure. The following result and the ideas in its proof are inspired by Konyagin's paper [25].

Proposition 37. *Let $P = A_d x^d + \ldots + A_1 x + A_0 \in \mathbb{Z}[x]$ be a random polynomial with independent coefficients, and write μ_j for the law of A_j. Let $\tau > 0$ be a number. We assume*

$$\operatorname{supp}\mu_j \subset [-\exp(d^{1/10}), \exp(d^{1/10})]$$

for all j and $\|\mu_j\|_2^2 \leqslant 1-\tau$ for all $j \neq 0, d$.

Then the probability that there is a non-cyclotomic polynomial Q with $\log M(Q) < \tau/10$ dividing P is at most $2\exp(-c\tau d^{4/5})$, where $c > 0$ is an absolute constant.

The exponent 4/5 is not optimal and there is a trade-off between it and the bound imposed on the coefficients of P. Since any improvement of this bound would have no effect on our theorems, we leave it to the

interested reader to find the optimal bound that can be derived from the proof.

We give two simple Lemmata that estimate the probability that a fixed single polynomial Q divides a random polynomial. Both of them are implicitly contained in [25]. The first one is useful when $\deg Q$ is large.

Lemma 38. *Let $P = A_d x^d + \ldots + A_1 x + A_0 \in \mathbb{Z}[x]$ be a random polynomial with independent coefficients, and write μ_j for the law of A_j. Let $Q \in \mathbb{Z}[x]$ be a polynomial of degree $n \leqslant d$.*

Then

$$\mathbb{P}_P(Q|P) \leqslant \|\mu_0\|_\infty \cdots \|\mu_{n-1}\|_\infty.$$

Proof. Write R for the remainder of $A_d x^d + \ldots + A_n x^n$ modulo Q in $\mathbb{Q}[x]$. If $Q|P$, then $R = -A_{n-1}x^{n-1} - \ldots - A_0$. Therefore, the probability of $Q|P$ conditioned on the value of $A_d x^d + \ldots + A_n x^n$ is bounded by the maximal probability of $A_0, \ldots, A_{n-1}$ taking any given value, which is precisely the claimed bound. □

Lemma 39. *Let $P = A_d x^d + \ldots + A_1 x + A_0 \in \mathbb{Z}[x]$ be a random polynomial with independent coefficients, and write μ_j for the law of a_j. Let $H \in \mathbb{Z}_{>0}$ and $\tau > 0$ be numbers. We assume $\operatorname{supp} \mu_j \subset [-H, H]$ for all j and $\|\mu_j\|_2^2 \leqslant 1 - \tau$ for all $j \neq 0, d$. Let $Q \in \mathbb{Z}[x]$ be a non-cyclotomic irreducible polynomial.*

Then for d larger than some absolute constant:

$$\mathbb{P}(Q|P) \leqslant \exp(-c\tau d(\log H + \log d)^{-1}(\log d)^{-3}),$$

where $c > 0$ is some absolute constant.

Proof. Let

$$s = \frac{\log(2H(d+1)^{1/2})}{c(\log\log d)^3/(\log d)^3},$$

where c is a sufficiently small constant so that

$$\log(M(Q)) > c(\log\log d)^3/(\log d)^3.$$

The existence of such a constant follows by Dobrowolski's bound [16].

By Lemma 26, there is a prime $q \in (s, 2s]$ such that the ratio of any two roots of Q is not a root of unity of order q. Let $P_1, P_2 \in \mathbb{Z}[x]$ be two polynomials with coefficients of absolute value at most H that differ only in some of the coefficients of monomials of the form x^{qj} for $j \in \mathbb{Z}_{\geqslant 0}$. If $Q|P_1 - P_2$, then each number $z\omega$ is a root for $P_1 - P_2$, where z is a root of Q and ω is a q-th root of unity. And by our choice of q all $z\omega$ are distinct as z ranges over the roots of Q and ω over the q-th roots of unity. This implies that

$$M(P_1 - P_2) \geqslant M(Q)^q > 2H(d+1)^{1/2},$$

which is impossible by (1.6). This means that for any given choice of integers $b_j \in [-H, H]$ for those $j \leqslant d$ that are not a multiple of

q, in each class of $\mathbb{Z}[x]$ modulo Q there is at most one polynomial $P = a_0 + \ldots + a_d x^d$ with $a_j = b_j$ for all such j.

Hence conditioning on the value of a_j for all indices j that are not multiples of q, the probability of $Q|P$ is bounded by the probability that the rest of the coefficients take any particular given value. Therefore

$$\mathbb{P}(Q|P) \leqslant \|\mu\|_\infty^{\lfloor d/q \rfloor - 1}.$$

□

Proof of Proposition 37. We fix a small number $\varepsilon > 0$. Let $j \geqslant 0$ be an integer and write $\mathcal{Q}_j$ for the set of non-cyclotomic irreducible polynomials Q with $\deg Q = j$ and $\log M(Q) < \tau/10$. By the estimate of Dubickas and Konyagin [17, Theorem 1], we have $|\mathcal{Q}_j| \leqslant \exp(\tau j/10)$ if j is sufficiently large.

Using Lemma 38, we then have

$$\mathbb{P}(\exists Q \in \mathcal{Q}_j : Q|P) \leqslant \exp(\tau j/10) \cdot \exp(-\tau j/2) \leqslant \exp(-\tau j/10)$$

for each j. By Lemma 39 applied with $H = \exp(d^{1/10})$

$$\mathbb{P}(\exists Q \in \bigcup_{j<d^{4/5}} \mathcal{Q}_j : Q|P) \leqslant \exp(\tau d^{4/5}/10) \cdot \exp(-\tau d^{4/5}) \leqslant \exp(-\tau d^{4/5}/10).$$

provided d is sufficiently large depending on an absolute constant.

Summing up the above bounds we get

$$\mathbb{P}(\exists Q \in \bigcup \mathcal{Q}_j : Q|P) \leqslant \exp(-\tau d^{4/5}/10) + \sum_{j \geqslant d^{4/5}} \exp(-\tau j/10),$$

which proves the claim. □

8. Proper powers

In this section, we estimate the probability that a random polynomial P is of the form ΦQ^m with $m > 1$, where Φ is the product of cyclotomic factors.

Proposition 40. *Let $P = A_d x^d + \ldots + A_1 x + A_0 \in \mathbb{Z}[x]$ be a random polynomial with independent coefficients. Assume that $A_1, \ldots, A_{d-1}$ are identically distributed with common law μ. Assume further that all coefficients are bounded by $\exp(d^{1/10})$ almost surely. Let $\tau > 0$ be a number such that $\|\mu\|_2^2 < 1 - \tau$.*

Then there are absolute constants $c, C > 0$ such that the probability that $P = \Phi Q^k$, where Φ is a product of cyclotomic polynomials, $Q \in \mathbb{Z}[x]$ and $k \geqslant 2$, is less than $2\exp(-c(\tau d)^{1/2})$, provided d is larger than C/τ^4.

In the next two lemmas we keep the assumptions of Proposition 40. The first is a reformulation of Proposition 32.

Lemma 41. *There is an absolute constant $c_0 > 0$ such that the following holds. Let $q < \exp(c_0(\tau d)^{1/2})$ be a product of distinct primes larger than $2\exp(d^{1/10})$. Then for every $a \in \mathbb{Z}$, we have*

$$|\mathbb{P}_P[P(2) \equiv a \mod q] - q^{-1}| < q^{-10}.$$

Lemma 42. *Fix $R \in \mathbb{Z}[x]$, and fix an integer $2 \leqslant k \leqslant d^{1/5}$. Then*

$$\mathbb{P}_P[P = RQ^k \text{ for some } Q \in \mathbb{Z}[x]] \leqslant \exp(-c(\tau d)^{1/2}),$$

where $c > 0$ is an absolute constant.

In the proof that follows, we will use the upper bound on m in only one place, where we apply the prime number theorem in arithmetic progressions. It would be sufficient to impose a significantly milder upper bound on m, but we will see that $P = RQ^k$ may hold with $k > d^{1/5}$ only if Q is cyclotomic.

Proof of Lemma 42. If $R(2) = 0$ and R divides P, then $P(2) = 0$. Picking a prime q in the interval $(\frac{1}{2}\exp(c_0(\tau d)^{1/2})/2, \exp(c_0(\tau d)^{1/2}))$, Lemma 41 implies that $\mathbb{P}_P[P(2) = 0] < 2/q$. So we can safely assume in the rest of the proof that $R(2) \neq 0$. We note also that $R(2) \leqslant P(2) \leqslant \exp(d^{1/10})2^{d+1}$.

We denote by $\mathcal{P}$ the collection of primes

$$p \in [\frac{1}{2}\exp(c_0(\tau d)^{1/2}/2)/2, \exp(c_0(\tau d)^{1/2}/2)]$$

such that $p \nmid R(2)$ and $k|p-1$. It follows from the prime number theorem in arithmetic progressions [14, Chp. 20, (10)] that there are more than

$$|\mathcal{P}| \geqslant \exp(c_0(\tau d)^{1/2}/4)$$

such primes if d is sufficiently large (i.e. τd larger than an effective constant: we are counting primes between $x/2$ and x that are congruent to 1 modulo k with k allowed to take any value $\ll (\log x)^{2/5}$ say).

For each $p \in \mathcal{P}$ we denote by X_p the random variable that is equal to 1 if

$$P(2) \equiv R(2)a^k \mod p$$

for some $a \in \mathbb{Z}/p\mathbb{Z}$ and that is equal to 0 otherwise. If $P = RQ^k$ for some $Q \in \mathbb{Z}[x]$, then clearly $X_p = 1$ for all $p \in \mathcal{P}$.

It follows from Lemma 41 applied first to $q = p_1$ and then to $q = p_1p_2$ that

$$\mathbb{E}_P[X_{p_1}] = \mathbb{E}_P[X_{p_1}^2] = \frac{(p_1-1)/k+1}{p_1} + O(\exp(-9c_0(\tau d)^{1/2}/2))$$

$$\mathbb{E}_P[X_{p_1}X_{p_2}] = \frac{(p_1-1)/k+1}{p_1}\cdot\frac{(p_2-1)/k+1}{p_2} + O(\exp(-9c_0(\tau d)^{1/2}/2))$$

for any $p_1 \neq p_2 \in \mathcal{P}$. Therefore, writing $Y = \sum_{p\in\mathcal{P}} X_p$, since $k \geqslant 2$,

$$\mathbb{E}_P Y \leqslant \frac{2}{3}|\mathcal{P}|$$

and the variance $\mathbb{V}ar(Y) = \mathbb{E}_P Y^2 - (\mathbb{E}_P Y)^2$ is bounded by

$$\begin{aligned}\mathbb{V}ar(Y) = &\sum_{p\in\mathcal{P}} \Big(\frac{(p-1)/k+1}{p} - \Big(\frac{(p-1)/k+1}{p}\Big)^2\Big)\\ &+ O(|\mathcal{P}|^2 \exp(-9c_0(\tau d)^{1/2}/2))\\ \leqslant &\frac{2}{3}|\mathcal{P}| + 1 < |\mathcal{P}|\end{aligned}$$

provided d is sufficiently large. We conclude from Chebyshev's inequality that

$$\mathbb{P}_P(Y = |\mathcal{P}|) \leqslant \mathbb{P}_P(Y - \mathbb{E}_P Y \geqslant \frac{1}{3}|\mathcal{P}|) \leqslant \mathbb{V}ar(Y)\big(\frac{3}{|\mathcal{P}|}\big)^2 < \frac{9}{|\mathcal{P}|},$$

which proves the lemma. □

Proof of Proposition 40. Boyd and Montgomery [7] gave an asymptotic formula for the number of polynomials Φ in $\mathbb{Z}[x]$ of degree n that are the product of their cyclotomic factors. In particular, they proved that there are at most $\exp(C_0 n^{1/2})$ such polynomials, where C_0 is an absolute constant ($C_0 = 4$ works for large enough n).

For a fixed Φ we may apply Lemma 38 and conclude that the probability that Φ divides P is at most $\exp(-\tau \deg(\Phi)/2)$. Therefore, the probability that $P = \Phi Q^k$ for some Φ with

$$\deg \Phi \geqslant 4\frac{C_0}{\tau} d^{1/2}$$

is at most $\exp(-C_0 d^{1/2})$.

We consider now the probability that $P = \Phi Q^k$ with Φ having smaller degree. We can assume that Q is not a product of cyclotomic factors, otherwise it can be absorbed into Φ and it is covered by the previous case. We note that if $P = \Phi Q^k$, then by (1.6)

$$M(Q) = M(P)^{1/k} \leqslant \exp(d^{1/10}/k)(d+1)^{1/2k}.$$

Since Q is not a product of cyclotomic factors, this implies that $k \leqslant d^{1/5}$ (say) by Dobrowolski's bound (1.6).

Again by [7] the number of polynomials in the role of Φ that are not covered by the previous case is at most

$$\exp\big(2C_0(C_0/\tau)^{1/2} d^{1/4}\big).$$

Now we can use Lemma 42 to estimate the probability of $P = \Phi Q^k$ for individual choices of Φ and k and conclude the proof. □

9. Proof of the main results

We first give a simple lemma that allows us to decide when a permutation group is m-transitive. Recall that the Bell number B_m is the number of equivalence relations on a set with m elements.

Lemma 43. *Let G be a permutation group acting on a set Ω and let $m \in \mathbb{Z}_{>0}$. Suppose $|\Omega| \geqslant m$. The number $|\Omega^m/G|$ of orbits of G acting diagonally on Ω^m satisfies*

$$|\Omega^m/G| \geqslant B_m$$

with equality if and only if the action of G on Ω is m-transitive.

Proof. If G is m-transitive, then its orbits on Ω^m are in one-to-one correspondence with equivalence relations on the set of coordinates. Given an equivalence relation on the m coordinates, the corresponding orbit is the set of tuples in Ω^m whose coordinates are equal if and only if they are related by the equivalence relation. Since $|\Omega| \geqslant m$, all equivalence relations can occur. Hence $|\Omega^m/G| = B_m$.

Now in the general case $G \leqslant \mathrm{Sym}(\Omega)$, so each orbit of G is contained in an orbit of $\mathrm{Sym}(\Omega)$. Thus $|\Omega^m/G| \geqslant |\Omega^m/\mathrm{Sym}(\Omega)| = B_m$.

If G is not m-transitive, then the orbit of the full symmetric group $\mathrm{Sym}(\Omega)$ consisting of tuples with distinct coordinates splits into multiple orbits of G, hence $|\Omega^m/G| > B_m$. □

9.1. **Proof of Theorem 2.** We set $\kappa = \tau/100$, $m = 1$ and let $X > 10$. Recall that we denote by $\widetilde{P}$ the product of the (κ, X)-admissible irreducible factors of P and that Ω is the set of complex roots of $\widetilde{P}$ (see Definition 14). We aim to show that the Galois group G of the splitting field of $\widetilde{P}$ acts transitively on Ω with high probability.

Recall that h_X is the function $h_X(u) = 2e^{-X}1_{(X-\log 2, X]}(u)$. It follows from the prime number theorem that

$$w := \sum_p \log(p) h_X(\log p) \to 1$$

as $X \to \infty$. We apply Proposition 34 for $g = h_X$ with $m = 1$. It applies if X is large enough and we conclude that

$$|\sum_p B_P(p) \log(p) h_X(\log p) - 1| \leqslant \frac{2}{3} \tag{9.1}$$

holds for any $X \in [100d^{1/10}, \frac{\tau^2}{100C_0}(\log d)^{-3}d]$ with probability at least

$$1 - \exp(-X/3) - \exp(X/40 - \frac{c_0\tau^2}{100}d(\log d)^{-2}/X).$$

provided $d > 100/\tau$ say. Taking $X = \tau(c_0 d/5)^{1/2}/\log d$ this bound becomes $\geqslant 1 - 2\exp(-X/40)$ provided $\tau^4 d$ is large enough. We now assume that (9.1) holds for P, and ζ_K satisfies RH for all $K = \mathbb{Q}(a)$ for any root a of P. By Proposition 19 applied with $g = h_X$, we then have

$$\sum_p B_P(p) \log(p) h_X(\log p) = |\Omega/G| + O(\exp(-X/10)).$$

If d is sufficiently large, we can conclude that

$$|1 - |\Omega/G|| < 1$$

under the above assumptions on P. We therefore conclude that $|\Omega/G| = 1$, and hence G acts transitively on Ω, i.e. $\widetilde{P}$ is irreducible.

By Proposition 37, with probability at least $1 - 2\exp(-c\tau d^{4/5})$, any exceptional factor of P is cyclotomic. If that holds in addition to the hypothesis we have already made, then $P = \Phi\widetilde{P}^k$, where $\Phi \in \mathbb{Z}[x]$ is a product of a power of x and cyclotomic polynomials, and $k \in \mathbb{Z}_{>0}$.

By Proposition 40, we know that $k = 1$ with probability at least $1 - 2\exp(-c(\tau d)^{1/2})$. This establishes part (1) of Theorem 2.

The proof of part (2) is similar, but we need to also consider moments of $B_P(p)$ of order $m > 1$ in order to show that $|\Omega^m/G| = B_m$ and hence conclude, by Lemma 43, that G acts m-transitively on Ω. An old fact, going back to Bochert and Jordan [22] in the 19-th century, asserts that every degree d permutation group that is at least $(30\log d)^2$-transitive must contain the alternating group $\mathrm{Alt}(d)$. A simple proof of a slightly better bound can be found in [1] (see also [15, Theorem 5.5.B] where Wielandt's stronger bound $6\log d$ is proved). Using the classification of finite simple groups it is now known that there is a bound independent of d and indeed every 6-transitive group contains $\mathrm{Alt}(d)$ (see [11, Corollary 5.4]). But we choose not to rely on the classification, since, at the expense of loosing a $\log(d)$ factor in the probability of exceptions, we can avoid it. In fact if instead we use Wielandt's bound (whose proof is more involved) we can get the slightly better bound $\exp(-c\tau d^{1/2}/(\log d)^{3/2})$ in (2) of Theorem 2.

So let $m \geqslant 1$, $\kappa = \tau/100$ and $X > 10$ and consider $\widetilde{P}$ the product of the irreducible $(\frac{\kappa}{m}, mX)$-admissible factors of P and as earlier $B_P(p)$ the number of $(\frac{\kappa}{m}, mX)$-admissible roots of P in $\mathbb{F}_p$.

By Proposition 34 applied to $g = h_X$ we get that

$$|\sum_p B_P(p)^m \log(p) h_X(\log p) - B_m w| < \frac{1}{2} \tag{9.2}$$

with probability at least

$$1 - \exp(-X/3) - \exp(X/40 - c_0\tau^2 d/(100mX(\log(md))^2)$$

provided X is in the interval allowed by (6.1). We now set $m = \lceil 30(\log d)^2 \rceil$ and $X = \sqrt{c_0\tau^2/5}\cdot d/(m(\log(md))^2)$. Note then that when $\tau^4 d$ is large enough X is in the allowed interval and that (9.2) holds with probability at least $1 - 2\exp(-X/40)$. Assume now that (9.2) holds for P and that ζ_K satisfies RH for all $K = \mathbb{Q}(a_1, \ldots, a_m)$ for any choice of m roots of P. By Proposition 19 we then get

$$\sum_p B_P(p)^m \log(p) h_X(\log p) = |\Omega^m/G| + O(\exp(-X/10)).$$

If d is large enough $|w - 1| = O(X^2\exp(-X/2))$ by Proposition 9 (assuming RH for $\zeta_{\mathbb{Q}}$). Since $B_m \leqslant 2^{m^2} \leqslant \exp(X/100)$ this implies

that

$$|B_m - |\Omega^m/G|| \leqslant \frac{1}{2} + B_m|1-w| + O(\exp(-X/10)) < 1$$

as soon as d is large enough, and hence that $|\Omega^m/G| = B_m$. So by Lemma 43 G acts m-transitively on Ω and by the 19-th century transitivity bound recalled earlier, since $\deg \widetilde{P} \leqslant d$, G contains the alternating group $\mathrm{Alt}(\deg \widetilde{P})$. Finally as in part (1), except for a small set of exceptions $P = \Phi\widetilde{P}$, and this completes the proof of the theorem.

9.2. **Proof of Corollaries 3 and 4.** The following lemma is implicitly contained in [25, pp. 345]

Lemma 44. *Let ω_n be the n-th cyclotomic polynomial of degree $\varphi(n)$. Then for all n, d,*

$$\mathbb{P}[\omega_n|P_d] \leqslant \big(C(\mu)\frac{n}{d}\big)^{\varphi(n)/2},$$

where $C(\mu) > 0$ depends only on μ.

Proof. Write $Q_d = \sum_{j=0}^{n-1} B_j x^j$, where $B_j := \sum_{i \equiv j \mod n, 0 \leqslant i \leqslant d} A_i$. Note that if $\omega_n|P_d$, then $\omega_n|Q_d$, and hence Lemma 38 implies that

$$\mathbb{P}[\omega_n|P_d] \leqslant \prod_0^{\varphi(n)-1} \|\mu_j\|_\infty,$$

where μ_j is the law of B_j, which is the sum of roughly $\lfloor d/n \rfloor$ i.i.d. variables with common law μ. Since μ has a finite second moment, there is a constant $C(\mu) > 0$ such that we have $\|\mu_j\|_\infty \leqslant (C(\mu)n/d)^{1/2}$, as follows say from the local limit theorem. The claim follows. □

We apply this lemma for different ranges of n. If $N \leqslant \varphi(n) \leqslant 100N$, then n is bounded in terms of N and

$$\mathbb{P}[\omega_n|P] \leqslant (C(\mu)n/d)^{N/2} = O_{N,\mu}(d^{-N/2})$$

If $100N < \varphi(n) \leqslant d^{1/2}$, then $n \leqslant Cd^{1/2}\log\log d$ for some absolute constant $c > 0$ and

$$\mathbb{P}[\omega_n|P] \leqslant (C(\mu)c\log\log d/d^{1/2})^{50N}$$

If $d^{1/2} \leqslant \varphi(n) \leqslant d$, then

$$\mathbb{P}[\omega_n|P] \leqslant \|\mu\|_\infty^{d^{1/2}/2}$$

by Lemma 38. Summing over all such n's we get:

$$\mathbb{P}[\omega_n|P \text{ for some } n \text{ with } \varphi(n) \geqslant N] = O_{\mu,N}(d^{-N/2}).$$

In order to apply Theorem 2, we need to truncate the coefficients. But

$$\mathbb{P}[\max_{0\leqslant i\leqslant d}|A_i| > e^{d^{1/10}}] \leqslant (d+1)\mathbb{P}[|A_0| > e^{d^{1/10}}] \leqslant (d+1)e^{-2d^{1/10}}\mathbb{E}[|A_0|^2]$$

by Chebyshev's inequality. The proof of Corollary 3 now follows by combining the above inequalities with Theorem 2.

To get Corollary 4 take $N = 2$ and observe that the law of P in the statement is designed to make sure that $x \nmid P$ always. Note also that Lemma 44 still holds even though A_0 and A_d are not distributed like the other A_i's, so the above estimates continue to hold. Since P has non-negative coefficients and at least two positive ones $P(1) > 0$. So it is only left to estimate $\mathbb{P}[\omega_2|P] = \mathbb{P}[P(-1) = 0]$. Looking at $P(-1)$ yields a random walk on $\mathbb{Z}$ and it is therefore a simple matter to verify that $\mathbb{P}[P(-1) = 0] = \sqrt{\frac{2}{\pi d}} + O(d^{-1})$, as desired.

9.3. **Proof of Theorem 5.** The proof is identical to that of part (1) of Theorem 2, except that we take $X = d(\log d)^{-\beta}$ and apply Proposition 20 instead of Proposition 19. We note that the exceptional zeros are not present by the assumptions of the theorem, so the right hand side of the displayed formula in Proposition 20 becomes

$$|\Omega/G| + O(\exp(-c(\log d)^{\alpha-\beta})).$$

9.4. **Proof of Theorem 6.** Set $\beta := \alpha - \gamma$. As in the proof of Theorem 2, we set $\kappa = \tau/10$ for the admissibility parameter (see Def. 14). By Proposition 37 with probability at least $1 - 2\exp(-c_\tau d^{4/5})$ every irreducible factor of P has Mahler measure at least $\exp(\kappa)$. We may thus assume that P has this property, and let $\widetilde{P}$ be the product of the non-cyclotomic irreducible factors of P. As before Ω is the set of roots of $\widetilde{P}$ and G the Galois group of the splitting field of P.

We use Proposition 34 with $X = X_1 = 2d(\log d)^{-\beta}$ and $X = X_2 = d(\log d)^{-\beta}$ for the functions $g = g_{X_1,k}$ and $g = g_{X_2,k}$, respectively, where $k = \lfloor (\log d)^{\alpha-\beta}/10 \rfloor$. We can conclude that both

$$\sum_p B_P^i(p) \log(p) g_{X_i,k}(\log p) = w_i + O(\exp(-c(\log d)^{\beta-2})) \tag{9.3}$$

hold for each $i = 1, 2$ with probability at least $1 - 2\exp(-c(\log d)^{\beta-2})$, where

$$w_i = \sum_p \log(p) g_{X_i,k}(\log p), \tag{9.4}$$

and $B_P^i(p)$ is the set of (X_i, κ)-admissible roots of P in $\mathbb{F}_p$.

We note that $|1 - w_i| < C\exp(-c(\log d)^{\alpha-\beta})$ as can be seen for example from Proposition 13 applied for $K = \mathbb{Q}$, since the Riemann zeta function $\zeta_{\mathbb{Q}}$ has no zeros in a sufficiently small neighborhood of 1. (Significantly better bounds can be obtained by the proof of Proposition 13, but this is not needed.)

Now we assume that P satisfies (9.3) for both $i = 1, 2$. We apply Proposition 20 and obtain for $i = 1, 2$

$$\sum_p B_P^i(p)\log(p)g_{X_i,k}(\log p) = \sum_{O\in\Omega/G}(1-G_{X_i,k}(\rho_{K_O,0}))+O(\exp(-c(\log d)^{\alpha-\beta})) \tag{9.5}$$

Now we combine the above estimates and $|w_1-w_2| < C\exp(-c(\log d)^{\alpha-\beta})$ to get

$$\sum_{O\in\Omega/G}(G_{X_2,k}(\rho_{K_O,0})-G_{X_1,k}(\rho_{K_O,0})) \leqslant C(\exp(-c(\log d)^{\alpha-\beta})+\exp(-c(\log d)^{\beta-2})).$$

We note that

$$\begin{aligned}(G_{X_2,k}(\rho_{K_O,0}) - G_{X_1,k}(\rho_{K_O,0})) =& \Big(1 - \frac{G_{X_1,k}(\rho_{K,0})}{G_{X_2,k}(\rho_{K_O,0})}\Big)G_{X_2,k}(\rho_{K_O,0})\\ \geqslant& (1 - \exp(-(1 - \rho_{K_O,0})(X_1 - X_2)/4))G_{X_2,k}(\rho_{K_O,0})\\ \geqslant& c\exp(-c_0(\log d)^{\gamma})d(\log d)^{-\beta}G_{X_2,k}(\rho_{K_O,0}).\end{aligned}$$

Here we used the bound on $G_{X_1,k}/G_{X_2,k}$ from Lemma 11 and then the assumption on the exceptional zeros from the theorem, and the constant c_0 is the constant c in that bound. Therefore, we can conclude that

$$\sum_{O\in\Omega/G} G_{X_2,k}(\rho_{K_O,0}) \leqslant C(\exp(-c(\log d)^{\alpha-\beta}) + \exp(-c(\log d)^{\beta-2}))$$

if we choose c_0 sufficiently small, since $\alpha - \beta, \beta - 2 \geqslant \gamma$ and $\gamma > 1$.

We combine the last estimate with (9.3), (9.4) and (9.5) and we can write

$$|\Omega/G| = 1 + O(\exp(-c(\log d)^{\alpha-\beta}) + \exp(-c(\log d)^{\beta-2})),$$

hence $|\Omega/G| = 1$ as it is an integer. Therefore $\widetilde{P}$ is irreducible. Now we can finish the proof by applying Proposition 40.

9.5. **Proof of Theorem 7.** We pick a number $\alpha' \in (\beta, \alpha)$. We use Proposition 34 with $g = g_{X,k}$, where $X = d(\log d)^{-\beta}$ and $k = \lfloor(\log d)^{\alpha'-\beta}/10\rfloor$. We get that

$$\sum_p B_P(p)\log(p)g_{X,k}(\log p) = w + O(\exp(-c(\log d)^{\beta-2})),$$

holds with probability at least $1 - C\exp(-c(\log d)^{\beta-2})$, where

$$w = \sum_p \log(p)g_{X,k}(\log p).$$

Moreover, as before, using Proposition 13 for the field of rational numbers we see that $|w - 1| = O(\exp(-c(\log d)^{\alpha-\beta}))$. Hence

$$\sum_p B_P(p)\log(p)g_{X,k}(\log p) = 1 + O(\exp(-c(\log d)^{\min\{\beta-2,\alpha-\beta\}})) \tag{9.6}$$

with probability at least $1 - C\exp(-c(\log d)^{\beta-2})$.

According to the Deuring-Heilbronn phenomenon if the Dedekind zeta function ζ_K of a number field K has a real zero very close to 1, then it cannot have other zeros nearby 1. More precisely (see [26, Theorem 5.1]) there is a positive, absolute, effectively computable constant $c_0 > 0$ such that for every number field K if ζ_K has a real zero $\rho_{K,0}$, then every other zero ρ satisfies:

$$|1-\rho| \geqslant \frac{c_0}{\log(2^d\Delta_K)} \log\Big(\frac{c_0}{|1-\rho_{K,0}|\log(2^d\Delta_K)}\Big).$$

So assume, by contradiction, that ζ_K has a zero $\rho_{K,0}$ with $|1-\rho_{K,0}| < \exp(-(\log d)^{\alpha+1})$ for each $K = \mathbb{Q}(a)$ for each non-zero complex root a of P, which is not a root of unity (this is void and hence always holds if P is a product of cyclotomic polynomials or factors of the type x^m). Note then that $|1-\rho_{K,0}| < 1/(4\log|\Delta_K|)$ (because by Lemma 21 $|\Delta_K| \leqslant d^{(1+\tau^{-1})2d}$) and hence by [32, Lemma 3] $\rho_{K,0}$ must be real and is the unique Siegel zero of ζ_K. Thus every other zero ρ satisfies

$$|1-\rho| \geqslant \frac{c}{d\log d}\log\Big(\frac{c\exp((\log d)^{\alpha+1})}{d\log d}\Big) \geqslant \frac{(\log d)^{\alpha'}}{d}$$

provided d is sufficiently large.

So we can apply Proposition 20 and, using Lemma 11, write

$$\begin{aligned}\sum_p B_P(p)\log(p) g_{X,k}(\log p) &= \sum_{O\in\Omega/G}(1-G_{X,k}(\rho_{K_O,0})) + O(\exp(-c(\log d)^{\alpha'-\beta}))\\ &\leqslant |\Omega/G| X\exp(-(\log d)^{\alpha+1}) + C\exp(-c(\log d)^{\alpha'-\beta})\\ &\leqslant d^2\exp(-(\log d)^{\alpha+1}) + C\exp(-c(\log d)^{\alpha'-\beta}).\end{aligned}$$

But this is incompatible with (9.6).

References

[1] L. Babai and Á. Seress. On the degree of transitivity of permutation groups: a short proof. *J. Combin. Theory Ser. A* 45 (2):310–315, 1987. ↑42

[2] W. Banaszczyk. Balancing vectors and Gaussian measures of n-dimensional convex bodies. *Random Structures Algorithms* 12 (4):351–360, 1998. ↑29

[3] L. Bary-Soroker and G. Kozma. Irreducible polynomials of bounded height, 2017. arXiv:1710.05165v1. ↑1, 6

[4] ———. Is a bivariate polynomial with ±1 coefficients irreducible? Very likely!. *Int. J. Number Theory* 13 (4):933–936, 2017. ↑1

[5] E. Bombieri and W. Gubler. *Heights in Diophantine geometry*. New Mathematical Monographs, vol. 4. Cambridge University Press, Cambridge, 2006. ↑9

[6] C. Borst, E. Boyd, C. Brekken, S. Solberg, M. M. Wood, and P. M. Wood. Irreducibility of random polynomials, 2017. Experimental Mathematics, to appear. ↑1

[7] D. W. Boyd and H. L. Montgomery. Cyclotomic partitions. In Number theory (Banff, AB, 1988), pages 7–25. 1990. ↑40

[8] E. Breuillard and P. P. Varjú. On the dimension of Bernoulli convolutions, 2016. arXiv:1610.09154v2. ↑2

[9] ———. Entropy of bernoulli convolutions and uniform exponential growth for linear groups, 2018. arXiv:1510.04043v3, to appear in J. d'Analyse Math. ↑2, 4
[10] ———. On the Lehmer conjecture and counting in finite fields, 2018. Preprint. ↑7
[11] P. J. Cameron. Finite permutation groups and finite simple groups. *Bull. London Math. Soc.* 13 (1):1–22, 1981. ↑42
[12] F. R. K. Chung, P. Diaconis, and R. L. Graham. Random walks arising in random number generation. *Ann. Probab.* 15 (3):1148–1165, 1987. ↑7, 21, 32
[13] H. Cohen. *A course in computational algebraic number theory*. Graduate Texts in Mathematics, vol. 138. Springer-Verlag, Berlin, 1993. ↑18
[14] H. Davenport. *Multiplicative number theory*. Graduate Texts in Mathematics, vol. 74. Springer-Verlag, New York-Berlin, Second, 1980. Revised by Hugh L. Montgomery. ↑39
[15] J. D. Dixon and B. Mortimer. *Permutation groups*. Graduate Texts in Mathematics, vol. 163. Springer-Verlag, New York, 1996. ↑42
[16] E. Dobrowolski. On a question of Lehmer and the number of irreducible factors of a polynomial. *Acta Arith.* 34 (4):391–401, 1979. ↑9, 37
[17] A. Dubickas and S. V. Konyagin. On the number of polynomials of bounded measure. *Acta Arith.* 86 (4):325–342, 1998. ↑17, 20, 35, 38
[18] L. Grenié and G. Molteni. Explicit versions of the prime ideal theorem for Dedekind zeta functions under GRH. *Math. Comp.* 85 (298):889–906, 2016. ↑10
[19] H. A. Helfgott. Growth in groups: ideas and perspectives. *Bull. Amer. Math. Soc. (N.S.)* 52 (3):357–413, 2015. ↑7, 21
[20] M. V. Hildebrand. *Rates of convergence of some random processes on finite groups*. ProQuest LLC, Ann Arbor, MI, 1990. Thesis (Ph.D.)–Harvard University. ↑7, 21
[21] A. E. Ingham. A Note on Fourier Transforms. *J. London Math. Soc.* 9 (1):29–32, 1934. ↑12
[22] C. Jordan. Nouvelles recherches sur la limite de transitivité des groupes qui ne contiennent pas le groupe alterné. *Journal de Math. Pures et Appliquées*:35–60, 1895. ↑42
[23] H. Kadiri. Explicit zero-free regions for Dedekind zeta functions. *Int. J. Number Theory* 8 (1):125–147, 2012. ↑5
[24] S. V. Konyagin. Estimates for Gaussian sums and Waring's problem modulo a prime. *Trudy Mat. Inst. Steklov.* 198:111–124, 1992. ↑7, 9, 21, 24, 25, 26
[25] ———. On the number of irreducible polynomials with $0, 1$ coefficients. *Acta Arith.* 88 (4):333–350, 1999. ↑1, 9, 36, 37, 43
[26] J. C. Lagarias, H. L. Montgomery, and A. M. Odlyzko. A bound for the least prime ideal in the Chebotarev density theorem. *Invent. Math.* 54 (3):271–296, 1979. ↑46
[27] J. C. Lagarias and A. M. Odlyzko. Effective versions of the Chebotarev density theorem. In Algebraic number fields: L-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975), pages 409–464. 1977. ↑14
[28] K. Mahler. An inequality for the discriminant of a polynomial. *Michigan Math. J.* 11:257–262, 1964. ↑19
[29] D. A. Marcus. *Number fields*. Springer-Verlag, New York-Heidelberg, 1977. Universitext. ↑18
[30] A. M. Odlyzko and B. Poonen. Zeros of polynomials with $0, 1$ coefficients. *Enseign. Math. (2)* 39 (3-4):317–348, 1993. ↑1

[31] S. O'Rourke and P. M. Wood. Low-degree factors of random polynomials, 2016. arXiv:1608.01938v1. ↑1
[32] H. M. Stark. Some effective cases of the Brauer-Siegel theorem. *Invent. Math.* 23:135–152, 1974. ↑5, 13, 14, 46
[33] H. Tôyama. A note on the different of the composed field. *Kōdai Math. Sem. Rep.* 7:43–44, 1955. ↑20
[34] P. P. Varjú. Absolute continuity of Bernoulli convolutions for algebraic parameters, 2017. arXiv:1602.00261v3. ↑2
[35] P. J. Weinberger. Finding the number of factors of a polynomial. *J. Algorithms* 5 (2):180–186, 1984. ↑6

Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK
E-mail address: breuillard@maths.cam.ac.uk

Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK
E-mail address: pv270@dpmms.cam.ac.uk