

A RANDOM POLYNOMIAL WITH MULTIPLICATIVE COEFFICIENTS IS ALMOST SURELY IRREDUCIBLE

PÉTER P. VARJÚ AND MAX WENQIANG XU

ABSTRACT. Assume that the Riemann hypothesis holds for Dedekind zeta functions. Under this assumption, we prove that a degree d polynomial with random multiplicative ± 1 coefficients is irreducible in $\mathbb{Z}[x]$ with probability $1 - O(d^{-1/2+\varepsilon})$.

1. INTRODUCTION

The question of how likely it is that a random polynomial in $\mathbb{Z}[x]$ is irreducible has a long history. The first studied model was where the degree of the polynomials is a fixed number and the coefficients are sampled independently and uniformly from growing intervals. This is less relevant to our paper, and we only refer to the recent breakthrough [7] and its references.

Another setting that has gained momentum more recently is where the coefficients are sampled independently from a fixed law and the degree of the polynomials is growing. A sequence of papers [4], [3], [5] established that such random polynomials are irreducible with probability tending to 1 if the common law of the coefficients is uniform enough modulo 4 primes, in particular when the coefficients are uniformly distributed on 35 consecutive elements. See also [2] for results about ± 1 coefficients and special degrees. Using a different method, and assuming the Riemann hypothesis for Dedekind zeta functions, [9] proved that the probability that a random polynomial is irreducible tends to 1 requiring only the necessary condition that the constant coefficient is not 0. This conditionally solved a conjecture of Odlyzko and Poonen [20] and the method also yields better estimates for the probability that the random polynomial is reducible.

In another direction, [12] and [13] proved irreducibility of the characteristic polynomial of random matrices with high probability.

In this paper, we consider other models where the coefficients of the random polynomial are not independent. We define a sequence X_n of ± 1 valued random variables for $n \in \mathbb{Z}_{\geq 1}$ as follows. We let $X_1 = 1$

MWX is supported by a Simons Junior Fellowship from Simons Foundation.

with probability 1. For primes p , we let X_p be independent uniform random variables taking ± 1 values. We consider two models. One is that for $n \in \mathbb{Z}_{\geq 2}$ we let $X_n = X_{p_1} \cdots X_{p_k}$, where $n = p_1 \cdots p_k$ is the prime factorization of n . The other model is that X_n is supported only on square-free integers n defined in the same way, and if n is not square-free, we set $X_n = 0$. For $d \in \mathbb{Z}_{\geq 0}$, and $\{X_n\}_{1 \leq n \leq d}$ being either one of the above two models, we define a random polynomial with multiplicative coefficients as

$$P_d(X) = X_1 x + X_2 x^2 + \dots + X_d x^d.$$

The main result of the paper is the following.

Theorem 1.1. *Suppose that the Riemann hypothesis holds for the Dedekind zeta functions of all number fields. Then for every $\varepsilon > 0$, there is a constant $C = C(\varepsilon)$ such that*

$$\mathbb{P}[P_d(x)/x \text{ is irreducible over } \mathbb{Z}] \geq 1 - Cd^{-1/2+\varepsilon}.$$

We remark that in the model where X_n is supported on square-free n , our bound is close to be sharp, as it is proved in [1] that $x = 1$ is a root with probability at least $Cd^{-1/2-\varepsilon}$.

Polynomials with multiplicative coefficients are of great interest in number theory. The study of their values on the unit circle has a vast literature. See [6] and [16] for recent work in the setting of polynomials with random multiplicative coefficients.

The question of irreducibility was recently studied in the setting of Fekete polynomials in [18] and [19]. For a prime p , the Fekete polynomial F_p is defined as

$$F_p(x) = \sum_{a=1}^{p-1} \left(\frac{a}{p} \right) x^a,$$

where $\left(\frac{a}{p} \right)$ denotes the Legendre symbol. The authors of [18] have made the conjecture that $F_p(x)$ is a product of linear factors corresponding to possible roots at $-1, 0, 1$ and an irreducible polynomial for all p .

Motivated by this, we pose the following problem.

Problem. *Let $f : \mathbb{Z}_{>0} \rightarrow \mathbb{Z}_{>0}$ be a function such that for all d , there is at least one prime with $d < p \leq f(d)$. For $d \in \mathbb{Z}_{>0}$, let p be a random prime in $(d, f(d)]$ sampled uniformly and let*

$$(1.1) \quad F_{d,f}(x) = \sum_{a=1}^d \left(\frac{a}{p} \right) x^a$$

be a random polynomial. What is the asymptotic behaviour of the probability that $F_{d,f}$ is irreducible after removing possible linear factors?

The conjecture in [18] predicts that the probability in question is 1 when $d + 1$ is a prime and $f(d) = d + 1$. If we allow $f(d) > 2^{2\pi(d)}d^4$, then it is an immediate consequence of our main result and known results about the distribution of the Legendre symbol that under the Riemann hypothesis for Dedekind zeta functions, $F_{d,f}$ is irreducible with high probability.

Corollary 1.2. *Let $F_{d,f}(t)$ be defined as in (1.1) and suppose $f(d) > 2^{2\pi(d)}d^4$. Suppose that the Riemann hypothesis holds for Dedekind zeta functions of all number fields. Then for all $\varepsilon > 0$, there is $C = C(\varepsilon)$ such that*

$$\mathbb{P}[F_{d,f}(x)/x \text{ is irreducible over } \mathbb{Z}] \geq 1 - Cd^{-1/2+\varepsilon}.$$

It would be interesting to see what the behaviour is when the range from which the prime is sampled is shorter.

1.1. Outline of the proof. The proof of Theorem 1.1 follows the strategy of [9], which we briefly recall. Fix a polynomial P , and choose a random prime q with a suitably chosen probability distribution. It is a consequence of the prime ideal theorem that if P is irreducible, then it has on average 1 root in \mathbb{F}_q . For different irreducible polynomials these roots rarely coincide, so we can deduce that

$$(1.2) \quad \begin{aligned} \{\text{number of distinct irreducible factors of } P\} \\ \approx \mathbb{E}_q[\text{number of roots of } P \text{ in } \mathbb{F}_q], \end{aligned}$$

where P is a fixed polynomial and the averaging is over a random prime q .

If we take a random polynomial P , and show that it has on average 1 root in \mathbb{F}_q , now P and q are both random, then it follows that P is a power of a single irreducible polynomial with high probability. To show this, we fix a prime q and a residue $a \in \mathbb{F}_q$, and show that the value $P(a)$ is equidistributed in \mathbb{F}_q for our random P . In particular, $P(a) = 0 \in \mathbb{F}_q$ will occur with probability approximately $1/q$. Summing this up for a and averaging over q will give the required result.

In the setting of [9], the equidistribution of $P(a)$ in \mathbb{F}_q is related to a Markov chain introduced by Chung, Diaconis and Graham [11]. Due to the dependence of the coefficients, the equidistribution problem cannot be described by a Markov chain in our setup.

Proving equidistribution is the main new contribution of our paper. We do this by conditioning on the values of the coefficients X_p for primes $p < d/2$, and use that the coefficients corresponding to the remaining primes are independent from each other and from the coefficients whose values are influenced by the smaller primes. We build

on [17] (also used in [9]) to prove equidistribution of

$$\sum_{p \geq d/2 \text{ prime}} X_p a^{p-1}.$$

The key difference is that we are summing over the primes, and the argument in [9] requires an arithmetic progression. We bypass this issue by finding many disjoint short arithmetic progressions in the primes using known results from [15], [23], [21] and then apply a version of the argument in [9] for these. This idea has been inspired by [8].

This allows us to prove equidistribution of $P(a)$ for most values of a . For some of the remaining values, equidistribution may fail. It certainly does for $a = -1, 0, 1$ and possibly also for some other low degree roots of unity. For the exceptional residues, we only prove an upper bound of the form $\mathbb{P}[P(a) = 0] \leq Cd^{-1/2+\varepsilon}$ using a classical Littlewood-Offord type bound. This is where the error term in Theorem 1.1 comes from. Here a more precise analysis may yield a better bound giving a more precise estimate for the probability that P is reducible. It may be possible to adapt the arguments in [9] to give stronger estimates for most of the exceptional residues, and the error term may potentially be dominated by the probability that P is divisible by $x - 1$ or $x + 1$. We do not pursue this question.

After this, it remains to show that P is not a proper power of an irreducible polynomial with high probability. We have not been able to adapt the corresponding argument in [9] to our setting. Instead, we extend the equidistribution result to the pair $(P(a), P'(a))$. This allows us to show that P has very few double roots in \mathbb{F}_q on average (much less than 1), and then the main argument can be used to rule out repeated factors of P in $\mathbb{Z}[x]$.

1.2. Organization of the paper. In Section 2, we formulate a statement that makes (1.2) precise. This is a straightforward adaptation of [9], but the result formulated in [9] cannot be applied as a black box. Section 3 contains our equidistribution estimate for random walks with steps $\pm a^i$ where i runs through an index set that contains sufficiently many disjoint arithmetic progressions. We use these estimates to derive bounds for the expected number of roots of P in a finite field in Section 4. We finish the paper by completing the proofs of the main results in Sections 5 and 6.

1.3. Notation. The letters c and C denote positive constants whose values may change from one occurrence to the other.

1.4. Acknowledgement. We would like to thank Jacob Fox, Julian Sahasrabudhe, Peter Sarnak and Joni Teräväinen for helpful discussions on the subject of this paper. Part of the work was done during several visits of MWX at Cambridge University, and the warm hospitality is greatly appreciated.

2. EXPECTED NUMBER OF ROOTS OF A POLYNOMIAL IN A RANDOM FINITE FIELD

Given a number field K , we denote by \mathcal{O}_K its ring of integers and write ζ_K for its Dedekind zeta function. We write $A_K(n)$ for the number of prime ideals $\mathfrak{p} \subset \mathcal{O}_K$ with norm $N_{K/\mathbb{Q}}(\mathfrak{p}) = n$.

For a number $X > 1$, we write

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

Given a polynomial $P \in \mathbb{Z}[x]$ and a rational prime q , we write $B_P(q)$ for the number of distinct roots of P in \mathbb{F}_q . We write \tilde{P} for the product of the irreducible factors of P in $\mathbb{Z}[x]$, and we write Δ_P for the discriminant of P . Given an irreducible polynomial $Q \in \mathbb{Z}[x]$, we write $K_Q = \mathbb{Q}(\alpha)$, where α is a root of Q .

The purpose of this section is to prove the following result.

Proposition 2.1. *Let $d, M \in \mathbb{Z}_{\geq 1}$. Let $P \in \mathbb{Z}[x]$ be a polynomial of degree at most d with coefficients of absolute value at most M . Suppose that for every irreducible factor Q of P , RH holds for ζ_{K_Q} . Let $X \geq 1$. Then*

$$\sum_{q \text{ prime}} B_P(q) \log(q) h_X(\log q) = |\{\text{distinct irreducible factors of } P\}| + O(d^2 X^2 \log(dM) \exp(-X/2)).$$

The implied constant is absolute.

The proof of this follows [9, Proposition 19]. We begin by recalling a quantitative version of the prime ideal theorem under the Riemann hypothesis. This is standard, and this precise formulation can be found in [9, Proposition 9].

Proposition 2.2. *Let K be a number field with discriminant Δ , and suppose RH holds for ζ_K . Let $X > 1$. Then*

$$\sum_{q \text{ prime}} A_K(q) \log(q) h_X(\log q) = 1 + O(X^2 \log |\Delta| \exp(-X/2)),$$

where the implied constant is absolute.

Lemma 2.3. *Let $P \in \mathbb{Z}[x]$, and let q be a rational prime with $q \nmid \Delta_{\tilde{P}}$. Then*

$$(2.1) \quad B_P(q) = \sum_{Q|P \text{ irreducible}} A_{K_Q}(q).$$

This lemma is standard. It is closely related to the $m = 1$ case of [9, Proposition 16]. The difference compared to that result is that there the roots of certain exceptional polynomials are not counted in $B_P(p)$ and they are also not counted on the right hand side of (2.1). While this is not formally permitted in [9], the proof works verbatim if we take the empty set for the exceptional polynomials.

Proof of Proposition 2.1. We first estimate Δ_P . It is represented by a determinant of size $2d - 1$ with entries bounded by dM (divided by the leading coefficient). Therefore,

$$|\Delta_P| \leq (2d - 1)^{2d-1} (dM)^{2d-1} \leq (2d)^{4d} M^{2d}.$$

This is also an upper bound for $\Delta_{\tilde{P}}$ and the discriminants for all the number fields that we may obtain by adjoining a root of P to \mathbb{Q} .

If $q \nmid \Delta_{\tilde{P}}$ is a prime, then

$$B_P(q) = \sum_{Q|P \text{ irreducible}} A_{K_Q}(q)$$

by Lemma 2.3. If $q|\Delta_{\tilde{P}}$, then we just use the trivial bounds $0 \leq B_P(q) \leq d$, and $0 \leq \sum_{Q|P \text{ irreducible}} A_{K_Q}(q) \leq d$ to deduce that

$$\left| B_P(q) - \sum_{Q|P \text{ irreducible}} A_{K_Q}(q) \right| \leq d,$$

and we have at most

$$\frac{\log |\Delta_{\tilde{P}}|}{X - \log 2} \leq \frac{10d(\log dM)}{X}$$

number of primes q for which this holds.

Therefore,

$$\sum_{q \text{ prime}} \left| B_P(q) - \sum_{Q|P \text{ irreducible}} A_{K_Q}(q) \right| \log(q) h_X(\log q) \leq \frac{Cd^2(\log dM)}{\exp(X)}$$

for an absolute constant C .

Using Proposition 2.2 to estimate the sums of $A_{K_Q}(q)$, we get

$$\begin{aligned} \left| \sum_{q \text{ is prime}} B_P(q) \log(q) h_X(\log q) - |\{\text{distinct irreducible factors of } P\}| \right| \\ \leq CdX^2 \log((2d)^{4d} M^{2d}) \exp(-X/2) + Cd^2 (\log dM) \exp(-X) \\ \leq Cd^2 X^2 \log(dM) \exp(-X/2). \end{aligned}$$

□

3. EQUIDISTRIBUTION ESTIMATE

We denote by $M(P)$ the Mahler measure of a polynomial $P \in \mathbb{Z}[x]$. Let $l \in \mathbb{Z}_{\geq 1}$ and let q be a prime. We say that a polynomial P is (l, q) -exceptional if $\deg(P) \leq l$ and $M(P) \leq q^{1/(l+1)^2}$. If q is a prime, then an element of \mathbb{F}_q is l -exceptional if it is the root of an (l, q) -exceptional polynomial.

The purpose of this section is to prove the following result and a weaker estimate that is valid for all non-zero residues, which we formulate at the end of the section.

Proposition 3.1. *Let $l, d \in \mathbb{Z}_{\geq 3}$ and let q be a prime that is suitably large in terms of l . Let $K \in \mathbb{Z}_{\geq 1}$. Let $a \in \mathbb{F}_q$ be an element such that a^k is not l -exceptional for any $k = 1, \dots, K$. Let $I \subset [0, \dots, d]$ be a set that contains $q^{5/(l+1)}$ pairwise disjoint arithmetic progressions of length $3l^3$ with common difference at most K . Let X_i be independent uniform ± 1 valued random variables, and let*

$$Y = \sum_{i \in I} X_i(a^i, ia^{i-1})$$

be a random element of \mathbb{F}_q^2 . Then

$$\left| \mathbb{P}[Y = x] - \frac{1}{q^2} \right| < q^{-10}$$

for all $x \in \mathbb{F}_q^2$.

The strategy of the proof is to estimate the Fourier coefficients of Y . For an element $x \in \mathbb{F}_q^2$, we write $|x|$ for the smallest absolute value of an integer in the residue class of x . We will show that if a^k is not an exceptional residue and $(\xi_1, \xi_2) \in \mathbb{F}_q^2 \setminus (0, 0)$, then $|\xi_1 a^j + \xi_2 j a^{j-1}|$ cannot be small for all values of j in a suitably long arithmetic progression of step size k . This is the content of the next lemma. Once we have this, we may use the assumption that I contains many disjoint arithmetic progressions to find many indices j for which $|\xi_1 a^j + \xi_2 j a^{j-1}|$ is not

small. This will allow us to estimate the Fourier transform using a product formula that follows from the independence of X_j .

Lemma 3.2. *Let q be a prime, let $a \in \mathbb{F}_q$, and let $(\xi_1, \xi_2) \in \mathbb{F}_q^2 \setminus (0, 0)$. Let $l \in \mathbb{Z}_{\geq 3}$, let $r > l^2$ be a prime, and let $k \in \mathbb{Z}_{>0}$. Suppose that we have*

$$|\xi_1 a^{jk+j_0} + \xi_2(jk + j_0)a^{jk+j_0-1}| < \frac{q^{1-2/(l+1)}}{l+1}$$

for all $j = 0, \dots, l(r+1)$. Then a^k is an l -exceptional residue.

The proof of this lemma follows an argument of Konyagin [17]. Writing b_j for the integer with the smallest absolute value in the residue class of $\xi_1 a^{jk+j_0} + \xi_2(jk + j_0)a^{jk+j_0-1}$, we will show that under the assumption of the lemma, b_j satisfies two linear recurrence relations. We will use this to show that b_j also satisfies the linear recurrence corresponding to the greatest common divisor of the polynomials associated to the original recurrences, and hence a is a root of this polynomial. One of the polynomials will be used to control the degree, while the other will be used to control the Mahler measure of the greatest common divisor.

The following simple lemma will be used to construct polynomials such that b_j satisfies the corresponding linear recurrences mod q , and also at the end of the proof to conclude that a is a root of the greatest common divisor.

Lemma 3.3. *Let $a \in \mathbb{F}_q$ and let $\alpha_0, \dots, \alpha_l \in \mathbb{F}_q$. Consider the equations*

$$(3.1) \quad \sum_{j=0}^l \alpha_j (\xi_1 a^{j+j_0} + \xi_2(j+j_0)a^{j+j_0-1}) = 0,$$

where $\xi_1, \xi_2 \in \mathbb{F}_q$ for $j_0 \in \mathbb{Z}_{\geq 0}$ with the conventions $0 \cdot 0^{-1} = 0$ and $0^0 = 1$.

Then the following hold.

- (1) If equation (3.1) holds for $j_0 = 0$ and $j_0 = 1$ and $(\xi_1, \xi_2) \neq (0, 0)$, then a is a root of the polynomial $\alpha_0 + \alpha_1 x + \dots + \alpha_l x^l$.
- (2) If a is a double root of the polynomial $\alpha_0 + \alpha_1 x + \dots + \alpha_l x^l$, then equation (3.1) holds for all $j_0 \in \mathbb{Z}_{\geq 0}$ and all $\xi_1, \xi_2 \in \mathbb{F}_q$.

Proof. We begin with the first claim. If $\xi_2 \neq 0$, we subtract a times equation (3.1) for $j_0 = 0$ from the same equation for $j_0 = 1$. We get

$$\sum_{j=0}^l \alpha_j (\xi_1(a^{j+1} - a^{j+1}) + \xi_2((j+1)a^j - ja^j)) = 0,$$

which reduces to

$$\sum_{j=0}^l \alpha_j \xi_2 a^j = 0,$$

and proves the claim upon dividing the equation by ξ_2 .

If $\xi_2 = 0$ and then necessarily $\xi_1 \neq 0$, we get the claim if we divide (3.1) for $j_0 = 0$ by ξ_1 . This proves the first claim.

Next, we turn to the second claim. Using that a is a root of the polynomial $x^{j_0}(\alpha_0 + \dots + \alpha_l x^l)$ and of its derivative, we get the equations

$$\begin{aligned} \sum_{j=j_0}^l \alpha_j a^{j+j_0} &= 0, \\ \sum_{j=j_0}^l (j+j_0) \alpha_j a^{j+j_0-1} &= 0. \end{aligned}$$

Taking a linear combination of these equations with coefficients ξ_1, ξ_2 , we get (3.1). This proves the second claim. \square

Let $X = (x_0, \dots, x_N)$ be a sequence of integers. We write $\Lambda(X)$ for the set of polynomials $P(x) = \alpha_0 + \dots + \alpha_d x^d$ of degree d for some $d \leq N$ such that

$$\sum_{j=0}^d \alpha_j x_{j+j_0} = 0$$

for all $j_0 = 0, \dots, N-d$.

The next lemma, which we quote from [17, Lemma 5], will be used to show that the sequence b_j satisfies the linear recurrence relation corresponding to the greatest common divisor of the two polynomials that we will construct.

Lemma 3.4. *Let $X = (x_0, \dots, x_N)$ be an integer sequence. Suppose $P_1, P_2 \in \Lambda(X)$ and $\deg(P_1) + \deg(P_2) \leq N$ then we have $\gcd(P_1, P_2) \in \Lambda(X)$.*

One of the polynomials that we will construct will be a polynomial in x^r for a suitable number r . The next lemma gives an estimate for the Mahler measures of low degree divisors of such a polynomial. This result is standard, but we include the proof for the sake of completeness.

Lemma 3.5. *Let $P_1, P_2 \in \mathbb{Z}[x]$ be non-zero polynomials and let $r > \deg(P_1)^2$ be a prime and suppose $P_1(x) | P_2(x^r)$. Then $M(P_1)^r \leq M(P_2)$*

Proof. Let ζ be a primitive r 'th root of unity. For each root α of P_1 , $\alpha\zeta^j$ is a root of $P_2(x^r)$ for all $j = 0, \dots, r-1$ with the same multiplicity

as α is a root of P_1 . Moreover, these are distinct numbers because $\deg(\alpha/\alpha') < \deg(P_1)^2$ for any two roots α, α' of P_1 and $\deg(\zeta^j) \geq r - 1$ for $j = 1, \dots, r - 1$.

Therefore, denoting the places of \mathbb{Q} by $M(\mathbb{Q})$, we have

$$\begin{aligned} M(P_2) &= \prod_{v \in M(\mathbb{Q})} \prod_{\alpha \in \overline{\mathbb{Q}_v}: P_2(\alpha)=0} \max(1, |\alpha|_v) \\ &\geq \prod_{v \in M(\mathbb{Q})} \prod_{\alpha \in \overline{\mathbb{Q}_v}: P_1(\alpha)=0} \prod_{j=0}^{r-1} \max(1, |\alpha \zeta^j|_v) \\ &= M(P_1)^r. \end{aligned}$$

□

Proof of Lemma 3.2. If $a = 0$, the conclusion holds, so we assume that $a \neq 0$. We observe that

$$\begin{aligned} \xi_1 a^{jk+j_0} + \xi_2(jk + j_0)a^{jk+j_0-1} &= \xi_1 a^{j_0}(a^k)^j + \xi_2(jk + j_0)a^{j_0-1}(a^k)^j \\ &= (\xi_1 a^{j_0} + \xi_2 j_0 a^{j_0-1})(a^k)^j + \xi_2 k a^{k+j_0-1} j (a^k)^{j-1}. \end{aligned}$$

If we replace a by a^k , ξ_1 by $\xi_1 a^{j_0} + \xi_2 j_0 a^{j_0-1}$ and ξ_2 by $\xi_2 k a^{k+j_0-1}$, this reduces the lemma to the case $j_0 = 0$ and $k = 1$, so we will only consider that case. It is easy to check that the new values of ξ_1 and ξ_2 are not both 0 if the original values were not.

Let b_j be the smallest integer in absolute value in the residue class of $\xi_1 a^j + \xi_2 j a^{j-1}$ for $j \in \mathbb{Z}_{\geq 0}$. By assumption, we have

$$|b_j| < \frac{q^{1-2/(l+1)}}{l+1}$$

for all $j = 0, \dots, l(r+1)$.

Let $\alpha_0, \dots, \alpha_l, \beta_0, \dots, \beta_l \in \mathbb{Z}$ with $|\alpha_j|, |\beta_j| < q^{2/(l+1)}$ for all j be such that a is a double root of both polynomials $P_1(x) = \alpha_0 + \dots + \alpha_l x^l$ and $P_2(x) = \beta_0 + \beta_1 x^r + \dots + \beta_l x^{rl}$, and both P_1 and P_2 are non-zero. Such coefficients exist by the pigeon-hole principle. This is a finite field version of Siegel's lemma.

We assume, as we may, that α_l and β_l are both non-zero, for otherwise we could multiply P_1 or P_2 by suitable powers of x .

Now by Lemma 3.3, it follows that

$$(3.2) \quad \sum_{j=0}^l \alpha_j b_{j+j_0} \equiv 0 \pmod{q}$$

for all $j_0 = 0, \dots, lr$ and

$$(3.3) \quad \sum_{j=0}^l \beta_j b_{jr+j_0} \equiv 0 \pmod{q}$$

for all $j_0 = 0, \dots, l$. For the second claim, we use the lemma for a coefficient sequence of length $lr + 1$ with $r - 1$ zeroes inserted between consecutive β_j .

By the triangle inequality and the upper bounds on the α_j , β_j and b_j , we have that the left hand sides of (3.2) and (3.3) are less than

$$(l+1) \frac{q^{1-2/(l+1)}}{l+1} q^{2/(l+1)} = q$$

in absolute value, hence they are 0. Therefore, $P_1, P_2 \in \Lambda(b_0, \dots, b_{l(r+1)})$.

By Lemma 3.4, we have $\gcd(P_1, P_2) \in \Lambda(b_0, \dots, b_{l(r+1)})$. By Lemma 3.3, a is then a root of $\gcd(P_1, P_2)$. We clearly have $\deg(\gcd(P_1, P_2)) \leq \deg(P_1) \leq l$ and

$$M(\gcd(P_1, P_2)) \leq M(P_2)^{1/r} \leq ((l+1)q^{2/(l+1)})^{1/r} \leq q^{1/(l+1)^2}$$

by Lemma 3.5 and $l \geq 3$, provided q is sufficiently large in terms of l . Therefore, $P := \gcd(P_1, P_2)$ is an (l, q) -exceptional polynomial and a is an l -exceptional residue. \square

Proof of Proposition 3.1. We consider the Fourier transform of the distribution of Y , which we compute as

$$\begin{aligned} \widehat{Y}(\xi_1, \xi_2) &= \mathbb{E}[\exp(2\pi i(\xi_1, \xi_2) \cdot Y/q)] \\ &= \prod_{j \in I} \mathbb{E}[\exp(2\pi i X_j(\xi_1 a^j + \xi_2 j a^{j-1})/q)] \\ &= \prod_{j \in I} \cos(2\pi(\xi_1 a^j + \xi_2 j a^{j-1})/q). \end{aligned}$$

We note the elementary inequality

$$\cos(2\pi b/q) \leq \exp(-(\pi^2/2)(|2b|/q)^2).$$

Suppose $(\xi_1, \xi_2) \neq (0, 0)$. We apply Lemma 3.2 with a prime $r \in [l^2, 2l^2]$ and conclude that every arithmetic progression in $\mathbb{Z}_{\geq 0}$ of length

$$l(r+1) + 1 \leq l(2l^2 + 1) \leq 3l^3$$

with common difference at most K contains an element j such that

$$|2\xi_1 a^j + 2\xi_2 j a^{j-1}| \geq \frac{q^{1-2/(l+1)}}{l+1},$$

and hence

$$\cos(2\pi(\xi_1 a^j + \xi_2 j a^{j-1})/q) \leq \exp(-(\pi^2/2)q^{-4/(l+1)}/(l+1)^2).$$

Since there are more than

$$q^{5/(l+1)} \geq 10 \log q \cdot (2/\pi^2) q^{4/(l+1)} (l+1)^2$$

disjoint arithmetic progressions in I , we have

$$|\widehat{Y}(\xi_1, \xi_2)| \leq q^{-10}$$

provided $(\xi_1, \xi_2) \neq (0, 0)$.

Using the Fourier inversion formula

$$\mathbb{P}[Y = x] = \frac{1}{q^2} \sum_{\xi_1, \xi_2 \in \mathbb{F}_q} \exp(-2\pi i(\xi_1, \xi_2) \cdot x) \widehat{Y}(\xi_1, \xi_2)$$

and $\widehat{Y}(0, 0) = 1$, we conclude

$$\left| \mathbb{P}[Y = x] - \frac{1}{q^2} \right| \leq q^{-10}$$

as required. \square

We finish this section by recording a classical Littlewood-Offord type bound that shows that the random walk spreads out substantially even when a is an exceptional residue. In the case $a = 1$, the result is essentially sharp.

Proposition 3.6. *Let q be an odd prime, and let $a \in \mathbb{F}_q^\times$. Let I be a set of positive integers with $|I| \leq q$. Let X_i be independent uniform ± 1 valued random variables for $i \in I$, and let*

$$Y = \sum_{i \in I} X_i a^i$$

be a random element of \mathbb{F}_q . Then

$$\mathbb{P}[Y = x] \leq 129|I|^{-1/2}.$$

for all $x \in \mathbb{F}_q$.

Proof. Let \tilde{X}_i for $i \in I$ be a sequence of independent random variables taking the value 0 with probability $3/4$ and each of ± 1 with probability $1/8$.

By [10, Lemma 12] applied with $\mu = 1/4$, we have

$$\mathbb{P}\left[\sum_{i \in I} \tilde{X}_i a^i = x\right] \leq 64(|I|/4)^{-1/2} + q^{-1} \leq 129|I|^{-1/2}$$

for all $x \in \mathbb{F}_q$.

By [22, Corollary 7.12] applied with $\mu' = 1$ and $\mu = 1/4$, we have

$$\mathbb{P}\left[\sum_{i \in I} X_i a^i = x\right] \leq \mathbb{P}\left[\sum_{i \in I} \tilde{X}_i a^i = 0\right] \leq 129|I|^{-1/2}$$

for all $x \in \mathbb{F}_q$. \square

4. EXPECTED NUMBER OF ROOTS OF A RANDOM POLYNOMIAL IN A FINITE FIELD

Recall that X_j for $j \in \mathbb{Z}_{\geq 0}$ is a random multiplicative sequence. To simplify notation, we introduce the random polynomial

$$R(x) := P_{d+1}(x)/x = X_1 + X_2 x + \dots + X_{d+1} x^d.$$

Recall also that $B_Q(q)$ is the number of distinct roots of a polynomial Q in the finite field \mathbb{F}_q .

The purpose of this section is to deduce the following estimates for the expected number of roots and double roots of R .

Proposition 4.1. *Fix $l \in \mathbb{Z}_{>0}$ and $\varepsilon > 0$. Let q be an odd prime and $d \in \mathbb{Z}_{>0}$ that are both sufficiently large in terms of l and such that $l \geq \varepsilon^{-1} \log q / \log d$ and $q > d$. Then*

$$\begin{aligned} |\mathbb{E}[B_R(q)] - 1| &< d^{-1/2+\varepsilon}, \\ \mathbb{E}[|\{a \in \mathbb{F}_q : a \text{ is a double root of } R\}|] &< d^{-1/2+\varepsilon}. \end{aligned}$$

We comment on the role of the parameter l in the above statement and in the proof. We will need to apply this proposition such that q is a sufficiently large power of d so that the error term in the prime ideal theorem (or in Proposition 2.1 to be precise) is smaller than our claimed estimate in the main theorem. In fact, $q > d^5$ will be sufficient for this purpose. We can then set l depending on ε and $\log q / \log d$. For the proposition to hold, we need that d and q is large enough depending on l . This is because l ultimately controls the length of the arithmetic progressions we need in the set I , so we need d to be large enough to satisfy the condition of the Green Tao theorem.

The proof follows easily from Propositions 3.1 and 3.6 and the Theorem of Green and Tao [15] that the primes contain arbitrarily long arithmetic progressions. We will use the following version that provides control for the step size of the progressions, which we recall from [23, Theorem 5] and [21, Theorem 1.3].

Theorem 4.2. *For every $L \in \mathbb{Z}_{>0}$, there is a constant $C = C(L)$ such that the following holds for all $d \in \mathbb{Z}_{>0}$ that is sufficiently large in terms of L . Let A be a subset of the primes in $[1, d]$ with $|A| > d/10 \log d$.*

Then A contains an arithmetic progression of length L with common difference less than $C(\log d)^C$.

We also need the following simple lemma to count the number of exceptional residues for which Proposition 3.6 will be applied.

Lemma 4.3. *For every l , there is a constant $C = C(l)$ such that the number of (l, q) -exceptional polynomials is at most $Cq^{1/(l+1)}$.*

Proof. The coefficients of a polynomial $Q \in \mathbb{Z}[x]$ of degree at most l are bounded by $CM(Q)$ for some constant C depending only on l . Therefore, an (l, q) -exceptional polynomial has coefficients bounded by $Cq^{1/(l+1)^2}$ due to the definition of being (l, q) -exceptional. The claim follows by raising this bound to the power $l + 1$. \square

Proof of Proposition 4.1. Let $K = C_1(\log d)^{C_1}$, where C_1 is the constant C in Theorem 4.2 applied with $L = 3l^3$. We first consider the probability that some $a \in \mathbb{F}_q$ is a root or a double root of R under the condition that a^k is not l -exceptional for any $1 \leq k \leq K$. To this end, we will apply Proposition 3.1 with the set of primes in $[d/2, d]$ in the role of I .

We first show that I contains the required number of arithmetic progressions. We apply Theorem 4.2 repeatedly to find arithmetic progressions of length $3l^3$ with common difference at most K . We start this with all primes in $[d/2, d]$ in the role of A at first, then we remove from A all arithmetic progressions that we find to apply Theorem 4.2 again for this reduced set.

This process can be run more than $d/(10 \log(d)l^3)$ times before the number of elements in A falls below the threshold in the theorem, so we find at least this many arithmetic progressions. Since $l \geq 5 \log q / \log d$,

$$q^{5/(l+1)} \leq \frac{d}{10 \log(d)l^3},$$

and we have enough arithmetic progressions to apply Proposition 3.1.

Next we write

$$(R(a), R'(a)) = Y + Z,$$

where

$$Y = \sum_{i \in I} X_i(a^{i-1}, (i-1)a^{i-2})$$

and

$$Z = \sum_{i \in [1, d+1] \setminus I} X_i(a^{i-1}, (i-1)a^{i-2}).$$

We observe that Y and Z are independent. So, conditioning on the value of Z , we can write

$$\left| \mathbb{P}[R(a) = x_1, R'(a) = x_2] - \frac{1}{q^2} \right| \leq \max_{z \in \mathbb{F}_q^2} \left| \mathbb{P}[Y = (x_1, x_2) - z] - \frac{1}{q^2} \right| < q^{-10}$$

for all $x_1, x_2 \in \mathbb{F}_q$, where we applied Proposition 3.1. We conclude

$$(4.1) \quad \left| \mathbb{P}[R(a) = 0] - \frac{1}{q} \right| \leq q^{-9}$$

$$(4.2) \quad \mathbb{P}[a \text{ is a double root of } R] \leq \frac{2}{q^2}.$$

Now we prove a bound that is valid for all $a \in \mathbb{F}_q^\times$. We write $R(a) = Y_1 + Z_1$, where Y_1 and Z_1 are the first components of the vectors Y and Z defined above. Conditioning on the value of Z_1 and then using Proposition 3.6 we can write

$$\mathbb{P}[R(a) = 0] \leq \max_{z \in \mathbb{F}_q} \mathbb{P}[Y_1 = -z] \leq 129|I|^{-1/2} \leq 1000(d/\log d)^{-1/2}.$$

Note that $R(0) = 1$ almost surely, so the probability that 0 is a root is 0, so the above bound is valid even for $a = 0$.

The number of elements $a \in \mathbb{F}_q$ for which the bounds (4.1) and (4.2) do not apply is at most $K^2 l C_2 q^{1/(l+1)}$, where C_2 is the constant C in Lemma 4.3. We see that the expected number of exceptional roots of R is less than

$$1000C_1^2 C_2 l (\log d)^{2C_1+1/2} q^{1/(l+1)} d^{-1/2} < d^{-1/2+\varepsilon}/2$$

because $l \geq \varepsilon^{-1} \log q / \log d$.

Summing up (4.2) for $a \in \mathbb{F}_p$ that are not exceptional and combining with the above bound, we get

$$\mathbb{E}[|\{a \in \mathbb{F}_q : a \text{ is a double root of } R\}|] \leq \frac{2}{q} + d^{-1/2+\varepsilon}/2 < d^{-1/2+\varepsilon}.$$

Summing up (4.1) and combining with the bound for exceptional roots, we get

$$\left| \mathbb{E}[|\{a \in \mathbb{F}_q : R(a) = 0\}|] - \frac{q}{q} \right| \leq \frac{K^2 l C_2 q^{1/(l+1)}}{q} + q^{-8} + d^{-1/2+\varepsilon}/2 < d^{-1/2+\varepsilon}.$$

□

5. PROOF OF THEOREM 1.1

Fix some $\varepsilon > 0$, and an integer $l > 5\varepsilon^{-1}$. Let d be sufficiently large in terms of l , and let $X = 5 \log d$. Recall $R(x) = P_{d+1}(x)/x$. Proposition 4.1 gives

$$(5.1) \quad |\mathbb{E}[B_R(q)] - 1| < d^{-1/2+\varepsilon}$$

for all primes q with $\log q \in [X - \log 2, X]$.

Applying Proposition 2.2 for $K = \mathbb{Q}$, we get

$$(5.2) \quad \sum_{q \text{ prime}} \log(q) h_X(\log q) = 1 + O(X^2 \exp(-X/2))$$

Summing up (5.1), we get

$$\begin{aligned} \sum_{q \text{ prime}} B_R(q) \log(q) h_X(\log q) &= 1 + O(X^2 \exp(-X/2)) + O(d^{-1/2+\varepsilon}) \\ &= 1 + O(d^{-1/2+\varepsilon}). \end{aligned}$$

Now we apply Proposition 2.1 with $M = 1$ and get

$$\begin{aligned} \mathbb{E}[|\{\text{distinct irreducible factors of } R\}|] &= 1 + O(d^{-1/2+\varepsilon}) + O(d^2 X^2 \log(d) \exp(-X/2)) \\ &= 1 + O(d^{-1/2+\varepsilon}). \end{aligned}$$

Since R has always at least 1 irreducible factor, Markov's inequality implies that the probability that it has more than 1 is less than $Cd^{-1/2+\varepsilon}$.

It remains to prove that R is not a proper power of a single irreducible polynomial with high probability. Write $\tilde{B}_R(q)$ for the number of elements of \mathbb{F}_q that are roots of R with multiplicity at least 2. When R is a proper power, $B_R(q) = \tilde{B}_R(q)$.

Using that R always has at least one irreducible factor, it follows from Proposition 2.1 that

$$\sum_{q \text{ prime}} B_R(q) \log(q) h_X(\log q) \geq 1/2$$

for all R . Therefore,

$$\sum_{q \text{ prime}} \mathbb{E}[\tilde{B}_R(q)] \log(q) h_X(\log q) \geq \mathbb{P}[R \text{ is a proper power}]/2.$$

Using Proposition 4.1 and (5.2), we get

$$2d^{-1/2+\varepsilon} > \mathbb{P}[R \text{ is a proper power}]/2,$$

as required.

6. PROOF OF COROLLARY 1.2

We deduce the corollary from the main theorem and the following result about the distribution of Legendre symbols due to Granville and Soundararajan.

Theorem 6.1 ([14, Proposition 9.1]). *Suppose the Riemann hypothesis holds for all Dirichlet L -functions. Let $x, d \in \mathbb{Z}_{>0}$. Let $\omega_r = \pm 1$ for each prime $r \leq d$, and let $\mathcal{P}(x, \{\omega_r\})$ denote the set of primes $p \leq x$ such that $\left(\frac{r}{p}\right) = \omega_r$ for each $r \leq d$. Then for $d \leq x^{1/2}$, we have*

$$\sum_{p \in \mathcal{P}(x, \{\omega_r\})} \log p = \frac{x}{2^{\pi(d)}} + O(x^{1/2}(d + \log x)^2).$$

Proof of Corollary 1.2. In the statement of the corollary, the prime p is selected from $[d+1, f(d)]$ uniformly. We modify the distribution of p in such a way that p is selected from $[1, f(d)]$ with probability

$$(6.1) \quad \frac{\log p}{\sum_{p \in [1, f(d)]} \log p} \leq C \frac{\log p}{f(d)}.$$

This does not affect the validity of the conclusion for the following reason. The probability that $p < f(d)^{1/10}$ is smaller than $Cd^{-9/10}$ in both models, which is much smaller than the claimed bound for the probability that $F_{d,f}$ is reducible, so we may ignore this event. Conditioning on $p > f(d)^{1/10}$ the ratio between the probabilities that p takes a given value in the two models is bounded between two absolute constants, so the probability of the event of reducibility will also only change by a constant factor. For the rest of the proof, we consider p to be random with distribution (6.1).

To prove the corollary, it is enough to show that for any fixed $\{\omega_r\} \in \{-1, 1\}^{r \leq d \text{ prime}}$ the probability that $\left(\frac{r}{p}\right) = \omega_r$ for all r is bounded by an absolute constant times

$$\mathbb{P}[X_r = \omega_r \text{ for all } r] = 2^{-\pi(d)}.$$

Therefore, it is enough to show that

$$\sum_{p \in \mathcal{P}(f(d), \{\omega_r\})} \frac{\log p}{f(d)} \leq \frac{C}{2^{\pi(d)}}$$

for some constant $C > 0$. This clearly follows from Theorem 6.1. \square

REFERENCES

- [1] R. Angelo, M. Aymone, J. Campos-Vargas, O. Klurman, and M. W. Xu, *Crossing probabilities and random multiplicative functions* (2025). In progress. [↑2](#)

- [2] L. Bary-Soroker, D. Hokken, G. Kozma, and B. Poonen, *Irreducibility of Littlewood polynomials of special degrees*, 2024. arXiv:2308.04878v2. [↑1](#)
- [3] L. Bary-Soroker, D. Koukoulopoulos, and G. Kozma, *Irreducibility of random polynomials: general measures*, Invent. Math. **233** (2023), no. 3, 1041–1120. MR4623542 [↑1](#)
- [4] L. Bary-Soroker and G. Kozma, *Irreducible polynomials of bounded height*, Duke Math. J. **169** (2020), no. 4, 579–598. MR4072635 [↑1](#)
- [5] P.-A. Bazin, *Irreducibility of random polynomials of $\mathbb{Z}[X]$* , Int. Math. Res. Not. IMRN **11** (2025), Paper No. rnaf136, 21. MR4914934 [↑1](#)
- [6] J. Benatar, A. Nishry, and B. Rodgers, *Moments of polynomials with random multiplicative coefficients*, Mathematika **68** (2022), no. 1, 191–216. MR4405975 [↑2](#)
- [7] M. Bhargava, *Galois groups of random integer polynomials and van der Waerden’s conjecture*, Ann. of Math. (2) **201** (2025), no. 2, 339–377. MR4878222 [↑1](#)
- [8] J. Bourgain, *On the distribution of the residues of small multiplicative subgroups of \mathbb{F}_p* , Israel J. Math. **172** (2009), 61–74. MR2534239 [↑4](#)
- [9] E. Breuillard and P. P. Varjú, *Irreducibility of random polynomials of large degree*, Acta Math. **223** (2019), no. 2, 195–249. MR4047924 [↑1, 3, 4, 5, 6](#)
- [10] M. Campos, M. Jenssen, M. Michelen, and J. Sahasrabudhe, *Singularity of random symmetric matrices revisited*, Proc. Amer. Math. Soc. **150** (2022), no. 7, 3147–3159. MR4428895 [↑12](#)
- [11] F. R. Chung, P. Diaconis, and R. L Graham, *Random walks arising in random number generation*, The Annals of Probability **15** (1987), no. 3, 1148–1165. [↑3](#)
- [12] S. Eberhard, *The characteristic polynomial of a random matrix*, Combinatorica **42** (2022), no. 4, 491–527. MR4492162 [↑1](#)
- [13] A. Ferber, V. Jain, A. Sah, and M. Sawhney, *Random symmetric matrices: rank distribution and irreducibility of the characteristic polynomial*, Math. Proc. Cambridge Philos. Soc. **174** (2023), no. 2, 233–246. MR4545205 [↑1](#)
- [14] A. Granville and K. Soundararajan, *The distribution of values of $L(1, \chi_d)$* , Geom. Funct. Anal. **13** (2003), no. 5, 992–1028. MR2024414 [↑17](#)
- [15] B. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. (2) **167** (2008), no. 2, 481–547. MR2415379 [↑4, 13](#)
- [16] S. Hardy, *Bounds for exponential sums with random multiplicative coefficient*, Int. Math. Res. Not. IMRN **22** (2024), 14138–14156. MR4830075 [↑2](#)
- [17] S. V. Konyagin, *Estimates for Gaussian sums and Waring’s problem modulo a prime*, Trudy Mat. Inst. Steklov. **198** (1992), 111–124. MR1289921 [↑4, 8, 9](#)
- [18] J. Mináč, T. T. Nguyen, and N. D. Tân, *Fekete polynomials, quadratic residues, and arithmetic*, J. Number Theory **242** (2023), 532–575. MR4490459 [↑2, 3](#)
- [19] ———, *On the arithmetic of generalized Fekete polynomials*, Exp. Math. **33** (2024), no. 4, 723–754. MR4830683 [↑2](#)
- [20] A. M Odlyzko and B. Poonen, *Zeros of polynomials with 0, 1 coefficients*, L’Enseignement Mathématique **39** (1993), no. 3-4, 317–348. [↑1](#)
- [21] X. Shao, *Narrow arithmetic progressions in the primes*, Int. Math. Res. Not. IMRN **2** (2017), 391–428. MR3658139 [↑4, 13](#)
- [22] T. Tao and V. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2006. MR2289012 [↑13](#)

- [23] T. Tao and T. Ziegler, *Narrow progressions in the primes*, Analytic number theory, 2015, pp. 357–379. MR3467408 ↑4, 13

CENTRE FOR MATHEMATICAL SCIENCES, WILBERFORCE ROAD, CAMBRIDGE CB3 0WA, UK

Email address: pv270@dpmms.cam.ac.uk

COURANT INSTITUTE OF MATHEMATICAL SCIENCES, 251 MERCER STREET, NEW YORK 10012, USA

Email address: maxxu1729@gmail.com