IRREDUCIBLE POLYNOMIALS OF BOUNDED HEIGHT

LIOR BARY-SOROKER AND GADY KOZMA

ABSTRACT. The goal of this paper is to prove that a random polynomial with i.i.d. random coefficients taking values uniformly in $\{1,\dots,210\}$ is irreducible with probability tending to 1 as the degree tends to infinity. Moreover, we prove that the Galois group of the random polynomial contains the alternating group, again with probability tending to 1.

1. INTRODUCTION

The study of random polynomials has a long history. One direction, that we do no pursue here, is the study of the distribution of the roots of the polynomial. Notable phenomena about the roots are the logarithmic number of real roots, [5, 14, 18] and the fact that they cluster near the unit circle with asymptotically uniform distribution [10, 26, 13]. Additional surprising phenomena have been observed in extensive simulations of the roots of random polynomials with coefficients ± 1, such as the appearance of various approximate Julia sets [3].

When the coefficients of the random polynomial

$$f(X) = X^n + \sum_{i=0}^{n-1} \zeta_i X^i$$

are integral; i.e., $\zeta_0,\dots,\zeta_{n-1} \in \mathbb{Z}$, a natural question is about irreducibility over $\mathbb{Q}$; that is to say, as an element of the ring $\mathbb{Q}[X]$. A more subtle question is about the distribution of the Galois group G_f of the random polynomial f, which is by definition the Galois group of the splitting field of f over $\mathbb{Q}$. This group G_f may be considered as a subgroup of the symmetric group S_n via the action on the roots of f. A basic property of Galois theory that connects irreducibility and Galois groups, is that

f is irreducible if and only if G_f is transitive.

It is believed, and proved in many cases, that with high probability f is irreducible, and in fact 'the most irreducible' in the sense that its Galois group is the full symmetric group.

In the simplest model, *the large box model*, one fixes $n = \deg f$ and the coefficients $\zeta_0,\dots,\zeta_{n-1}$ are i.i.d. random variables taking values uniformly in $\{-L,\dots,L\}$ with $L\to\infty$. We do not know how far back it goes, but it is well known that

$$\lim_{L\to\infty} \mathbb{P}(f \text{ is irreducible}) = 1. \tag{1}$$

The state-of-the-art error term in (1) is given by Kuba [17]: $\mathbb{P}(f \text{ is reducible}) = O(L^{-1})$, $n > 2$. In the more subtle case of Galois groups, the first result goes back at least to Van der Waerden [28] who proved that

$$\lim_{L\to\infty} \mathbb{P}(G_f = S_n) = 1. \tag{2}$$

Van der Waerden's error term in (2) is explicit: $\mathbb{P}(G_f \neq S_n) = O(L^{-\frac{1}{6}})$. It was improved in [8, 12] using sieve methods, and the state-of-the-art result was given recently by Rivin [25] using an elementary method: $\mathbb{P}(G_f \neq S_n) = O(L^{-1+\epsilon})$. This is nearly optimal, since

$$\mathbb{P}(G_f \neq S_n) \geq \mathbb{P}(G_f \leq S_{n-1}) \geq \mathbb{P}(f(0) = 0) = \frac{1}{2L+1}.$$

Another model, which is the focus of investigation of this paper, is *the restricted coefficients model.* In this model, the coefficients of f are i.i.d. random variables taking values uniformly in a *fixed* finite set, and the degree $n = \deg f$ grows to infinity. Two well studied examples are ± 1 coefficients [27, and references within] and $0, 1$ coefficients [21] (with $\zeta_0 = 1$).

In the latter model, Konyagin [15] proves that

$$\lim_{n\to\infty} \mathbb{P}(f \text{ has all irreducible factors of degree} \geq cn/\log n) = 1. \tag{3}$$

This implies that [15, page 334]

$$\mathbb{P}(f \text{ is irreducible}) \geq \frac{c}{\log n}.$$

This is the state-of-the-art result on $\zeta_i \in \{0, 1\}$, although it merely says that the probability does not tend to 0 too rapidly, while the truth is that it tends to 1.

The restricted coefficient model is considered much more difficult than the large box model, mainly because the methods of the large box model are not applicable as they are based on reductions modulo large primes.

Our main result seems to be the first establishment of the analogue of (1) in the restricted coefficients model:

Theorem 1. *Let L be a positive integer divisible by at least 4 distinct primes. Let*

$$f = X^n + \sum_{i=0}^{n-1} \zeta_i X^i$$

be a polynomial, where $\zeta_1, \zeta_2, \ldots$ are i.i.d. random variables taking values uniformly in $\{1, \ldots, L\}$. Then

$$\lim_{n\to\infty} \mathbb{P}(f \text{ is irreducible}) = 1.$$

Note that the smallest L that satisfies the restriction of the theorem is

$$L = 210 = 2 \times 3 \times 5 \times 7.$$

Under the same conditions as in Theorem 1 we can also show that the Galois group of f is either S_n (the whole symmetric group) or A_n (the alternating group, the group of even permutations). We find it worthwhile to note that the irreducibility of f (or, in other words, transitivity of the Galois group) is the part that requires 4 primes — the part that concludes from irreducibility that the Galois group is either S_n or A_n works by reducing modulo one prime, so makes no restrictions on L. Here is the precise formulation:

Theorem 2. *Let f be as in Theorem 1 but for any $L \geq 2$ (i.e. without the restriction that L be divisible by 4 primes). Then*

$$\lim_{n\to\infty} \mathbb{P}(\text{the Galois group of } f \text{ is transitive and different from } A_n \text{ and } S_n) = 0.$$

The proofs of Theorems 1 and 2 are in §§2-3. Section 2 contains well-known facts about the connection between random polynomials and random permutations; and well-known, or at least unsurprising, facts about random permutations. Experts can safely skip to §3 which contains the core of the proof. In §4 we include some heuristics and simulations related to the question that we could not resolve: is the Galois group A_n or S_n?

Acknowledgments. The authors wish to thank Nir Avni for a discussion that turned out to be crucial for the proof of Theorem 1. Noga Alon advised us on random permutations, saving us much time. Igor Rivin shared with us many exciting simulation results. Zeev Rudnick, Ofer Zeitouni and Shoni Gilboa helped with insight and advice.

LBS was partially supported by the Israel Science Foundation. GK was partially supported by the Israel Science Foundation and the Jesselon Foundation.

2. Generalities

2.1. Properties of random permutations.

Lemma 3. *With probability tending to 1, there is no $l > \log^3 n$ which divides the lengths of two distinct cycles of a random permutation.*

Proof. For any k_1 and k_2, the probability that both are lengths of cycles is bounded above by $1/k_2k_2$ (one may do this easy calculation oneself, or may consult the beginning of the proof of [19, Claim 1]). Summing over k_1 and k_2 in $l\mathbb{Z} \cap [1, n]$ gives that the probability that there are two cycles lengths divisible by l is bounded above by $C(\log n)^2/l^2$. Summing over $l \geq \log^3 n$ gives the lemma. □

The next lemma is well known, and we include its proof for completeness.

Lemma 4. *The lengths of the cycles of a random permutation on n letters can be constructed by the following process. Choose a uniform element x_1 in $\{0, \dots, n-1\}$, and let $n - x_1$ be the first cycle length. Next choose a uniform element x_2 in $\{0, \dots, x_1 - 1\}$ and let $x_1 - x_2$ be the second cycle. Continue until some x_i is 0.*

Proof. Choose some element of $\{1, \dots, n\}$, say 1, and examine its cycle. It is easy to see that the probability that the cycle of 1 has size k is exactly $\frac{1}{n}$, since there are $\binom{n-1}{k-1}$ ways to choose the cycle elements, $(k-1)!$ ways to choose the cycle from the chosen elements, and $(n-k)!$ ways to choose the remaining items. Take the cycle of 1 to be $n - x_1$. The remaining permutation is random on x_1 letters, so continue inductively. □

Lemma 5. *For any $0 \leq a < b \leq 1$, with probability tending to 1, a random permutation has a cycle whose length, l, satisfies the following two requirements:*

- *$l \in [n^a, n^b]$.*
- *l has a prime factor p such that $p > \log^3 n$.*

Proof. Let x_i be as in Lemma 4. Look at the last i such that $x_i > n^b$. Then our conditioning (i.e. that $x_{i+1} \le n^b$) simply gives that x_{i+1} is uniform in $\{1, \dots, n^b\}$, and this is independent of $x_1, \dots, x_i$.

Let l be the cycle added at the $(i+1)^{\text{st}}$ step, i.e. $x_{i+1} - x_{i+2}$. Then l is uniform between 1 and x_{i+1}, itself uniform between 1 and n^b, so $\mathbb{P}(l = \alpha) \le (C \log n)/n^b$ for all $\alpha \in \{1, \dots, n^b\}$. This means that l can be compared to a number uniform between 1 and n^b: if we denote by k such a variable then

$$\mathbb{P}(l \in S) \le C \log n \mathbb{P}(k \in S)$$

for all $S \subset \{1, \dots, n^b\}$.

The probability that k has all its prime divisors smaller than $\log^3 n$ has a well-known estimate. Indeed, the probability that a random number $k \le n^b$ has all its prime factor $\le \log^3 n$ is equal to $n^{-b/3+o(1)}$, see [20, Equation 7.16]. The event that $k \le n^a$ has probability n^{a-b} of course, so the probability of the union is at most $n^{-b/3+o(1)} + n^{a-b}$. Since this tends to 0 as $n \to \infty$ (even after multiplying by the $\log n$ from the comparison of l and k), the lemma is proved. □

2.2. **Polynomials versus Permutations.** In this section we discuss the fact that the cycle structure of a random permutation is similar to the decomposition of a random polynomial to irreducible factors. In a way it goes back to Gauss (who showed that the probability that the random polynomial is irreducible is close to $\frac{1}{n}$, which is the probability that the permutation has only one cycle), and was developed in the literature significantly, say in [1]. Still, we need a few lemmas which we did not find in the literature.

Consider the space Ω of all tuples $(m_1, m_2, \dots)$ of nonnegative integers with finite support; i.e., $m_i \ge 0$ for all i and $m_i = 0$ for all sufficiently large i. We define two sequences of random variables on Ω. First, for $n \ge 1$, let f be a random monic polynomial of degree n in $\mathbb{F}_q[T]$; i.e., the $0, \dots, n-1$ coefficients of f are i.i.d. uniform in $\mathbb{F}_q$ and the n^{th} coefficient is 1; and let $X_n(m_1, m_2, \dots)$ be the probability that f has m_i prime factors of degree i in its prime factorization. In particular, $\sum_i im_i = n$ if $X_n(m_1, m_2, \dots) > 0$. Similarly, let $Y_n(m_1, m_2, \dots)$ be the probability that a random permutation on n letters has m_i cycles of length i in its decomposition to a product of disjoint cycles. Again, $\sum_i im_i = n$ if $Y_n(m_1, m_2, \dots) > 0$. If $c_{i,m}$ denotes the number of possibilities to choose m unordered irreducible polynomials of degree i and if

$$\alpha(i, m) = \frac{c_{i,m}}{q^{im}},$$

then we have the formulas

$$\begin{aligned} Y_n(m_1, m_2, \dots) &= \prod_i \frac{1}{m_i! i^{m_i}}, \qquad \text{and} \\ X_n(m_1, m_2, \dots) &= \prod_i \alpha(i, m_i). \end{aligned} \tag{4}$$

We denote by $\mathbb{P}_{X_n}$ and $\mathbb{P}_{Y_n}$ the probabilities that X_n respectively Y_n induce on Ω.

For $m = 1$ we have the exact formula $i\alpha(i,1) = \sum_{j|i} \mu(i/j)q^j$, with μ the Möbius function, which implies that

$$-\frac{2q^{-i/2}}{i} \leq \alpha(i,1) - \frac{1}{i} \leq 0, \tag{5}$$

see, e.g., [24, Lemma 4]. We can use (5) to get

$$\alpha(i,m) = \frac{1}{m!i^m} \exp\big(O(mq^{-i/2} + m^2 iq^{-i})\big). \tag{6}$$

Indeed, as the number of ways to choose m unordered objects out of x objects with repetition is $\binom{m+x-1}{m}$, one has

$$\alpha(i,m) = \frac{\big(\alpha(i,1) + q^{-i}(m-1)\big)\big(\alpha(i,1) + q^{-i}(m-2)\big)\cdots\alpha(i,1)}{m!}$$

So plugging (5) to this equation we get

$$\begin{aligned}\alpha(i,m) &= \frac{1}{m!i^m}\prod_{j=1}^{m}\big(1 + O(q^{-i/2} + q^{-i}(m-j)i)\big)\\ &= \frac{1}{m!i^m}\exp\big(O(mq^{-i/2} + m^2 iq^{-i})\big),\end{aligned}$$

proving (6).

Normally, we will use this with m small relative to $q^{i/2}$. For example, if we assume that $m \leq q^{i/4}$, and in general $i \ll q^{i/4}$, so (6) gives

$$\alpha(i,m) = \frac{1}{m!i^m}(1 + O(q^{-i/4})). \tag{7}$$

Two useful equalities, which hold for all $x \leq n$ are:

$$\sum_{\substack{(m_1,\dots,m_n)\\ \sum im_i = x}} \prod_{i=1}^{n} \frac{1}{m_i! i^{m_i}} = 1 \tag{8}$$

$$\sum_{\substack{(m_1,\dots,m_n)\\ \sum im_i \geq x}} \prod_{i=1}^{n} \frac{1}{m_i! i^{m_i}} = n - x + 1 \tag{9}$$

where (8) comes from noting that the terms summed over are exactly the ones which correspond to Y_x, so they are probabilities and sum to 1, and (9) is simply the sum of (8) from x to n. We will also need auxiliary lemmas, that allow us to reduce to the case with $m_i = 0$ for small i:

Lemma 6. *For all $n > 1$, $i \in \{1,\dots,n\}$ and $\lambda \geq 0$,*

$$\mathbb{P}_{X_n}(m_i = \lambda) \leq e^{-c\lambda}, \tag{10}$$

where $c > 0$ is a positive constant. In particular,

$$\mathbb{P}_{X_n}(m_i \geq \lambda) \ll e^{-c\lambda}. \tag{11}$$

Similar estimates hold also for Y_n.

Proof. By (4),

$$\mathbb{P}_{X_n}(m_i = \lambda) = \alpha(i,\lambda) \sum_{\substack{m_i=0 \\ \sum_{j\neq i} m_j = n-\lambda}} \prod_{j\neq i} \alpha(j, m_j).$$

The sum on the right hand side is smaller than the same sum without the restriction $m_i = 0$, which is simply 1 (compare to (8)), so

$$\mathbb{P}_X(m_i = \lambda) \leq \alpha(i,\lambda). \tag{12}$$

Thus it suffices to show that $\alpha(i,\lambda) \leq e^{-c\lambda}$. A similar argument shows that $\mathbb{P}_{Y_n}(m_i = \lambda) \leq 1/\lambda! i^\lambda$ which finishes the Y_n case and we will not return to it.

For $i > 1$, (5) gives

$$\alpha(i,1) \leq \frac{1}{i} \leq \frac{1}{2}.$$

Since $\alpha(i,\lambda) \leq \alpha(i,1)^\lambda$, we get the needed bound

$$\alpha(i,\lambda) \leq \frac{1}{2^\lambda} = e^{-\lambda \log 2}.$$

For $i = 1$ and $\lambda = 2$ there are $\binom{q}{2} + q$ ways to choose two linear polynomials, hence, as $q \geq 2$,

$$\alpha(1,2) = \frac{\binom{q}{2} + q}{q^2} = \frac{1}{2}(1 + 1/q) \leq \frac{3}{4}.$$

This also does the case $i = 1$ and $\lambda > 2$ since

$$\alpha(1, 2\lambda + 1) \leq \alpha(1, 2\lambda) \leq \alpha(1,2)^\lambda \leq e^{-\lambda(\log 4 - \log 3)}.$$

The last remaining case is $i = \lambda = 1$, for which we forgo (12) and estimate $\mathbb{P}(m_1 = 1)$ directly (we just need an estimate uniform in $n \geq 2$ and q). For any linear polynomial p we have

$$\mathbb{P}(p \mid f) = q^{-1}$$

since there are exactly q^{n-1} monic polynomials of degree $n-1$, each one may be multiplied by p to get a monic polynomial of degree n, and these are all different. Similarly, if p_1, p_2 and p_3 are linear polynomials we have

$$\mathbb{P}(p_1p_2 \mid f) = q^{-2} \qquad \text{and} \qquad \mathbb{P}(p_1p_2p_3 \mid f) \leq q^{-3},$$

where the inequality in the second case is simply because we only assumed $n \geq 2$ and if $n = 2$ this probability is 0. Using inclusion-exclusion gives

$$\mathbb{P}(\exists p \text{ linear such that } p \mid f) \leq$$
$$\leq q \cdot q^{-1} - \Big(\tbinom{q}{2} + q\Big) q^{-2} + \Big(\tbinom{q}{3} + q\tbinom{q}{2} + q\Big) q^{-3} = \frac{2}{3} - \frac{3q-5}{6q^2} < \frac{2}{3}.$$

This was the last remaining case so the proof of (10) is done.

Summing over all integers $\geq \lambda$ gives

$$\mathbb{P}_{X_n}(m_i \geq \lambda) = \sum_{\mu=\lambda}^{\infty} \mathbb{P}_{X_n}(m_i = \mu) \leq \sum_{\mu=\lambda}^{\infty} e^{-c\mu} = e^{-c\lambda}\frac{1}{1-e^{-c}},$$

which proves (11). □

In the following lemma we write $d_{TV}(A,B)$ to denote the total variation distance between A and B.

Lemma 7. *Let $X_{n,r}$ be the measure on vectors $(m_r, m_{r+1}, \dots)$ given by restricting X_n, i.e. $X_{n,r}(m)$ is the probability that a random polynomial of degree n has m_r factors of degree r, m_{r+1} factors of degree $r+1$ etc. Let $Y_{n,r}$ be the analogous quantity for Y_n, the measure on cycles of random permutations. Then*

$$d_{TV}(X_{n,r}, Y_{n,r}) \leq C/r.$$

Proof. This is Theorem 5.8 in [1] — to aid the reader in understanding the notation of [1], their Y_j is our m_j for permutations, their C_j is our m_j for polynomials (both are defined on page 349 of [1]) and their notation $\mathscr{L}$ is the standard notation for "the law of a random variable". Let us note that C/r is suboptimal — one may show an exponential decay in r — but we will not need this extra precision. □

3. 4 independent permutations

Lemma 8. *There exists an $\omega : \mathbb{N} \to \mathbb{N}$ with $\lim_{n\to\infty} \omega(n) = \infty$ such that*

$$\lim_{n\to\infty} \mathbb{P}(f \text{ has a divisor of degree } \leq \omega(n)) = 0$$

where f is as in Theorem 1.

(this lemma does not require L to be divisible by 4 distinct primes)

Proof. This is well known and has many proofs in the literature. By far the best ω was achieved by Konyagin [15] who showed this with $\omega(n) = n/\log n$. The statement in [15] is only for coefficients 0 and 1, but the proof carries through in our case. A simpler argument that gives only $\omega(n) = \sqrt{\log n}$ can be found in [22, Theorem 1.10 and §2.2]. An even simpler argument, with no explicit bound on ω, can be found in [16, §2]. What we tell you three times is true. □

Lemma 9. *Let $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ be 4 independent uniform permutations in S_n. For $i \in \{1, \dots, 4\}$ and $l \leq n$ we define $E_n(i,l)$ to be the event that l can be written as a sum of lengths of cycles of σ_i. Then for all $k < n$,*

$$\mathbb{P}\left(\bigcup_{l=k}^{2k} \bigcap_{i=1}^{4} E_n(i,l)\right) \leq Ck^{-c}$$

where both constants are absolute, in particular independent of n and k.

Further, for an additional parameter $\lambda < 2k$,

$$\mathbb{P}\Big(\bigcup_{l=k}^{2k}\bigcup_{\lambda_1=0}^{\lambda}\cdots\bigcup_{\lambda_4=0}^{\lambda}\bigcap_{i=1}^{4}E_{n-\lambda_i}(i,l-\lambda_i)\Big)\le C(\lambda+1)^4k^{-c}$$

Proof. This lemma is essentially included in [23], but unfortunately not stated as such explicitly. Let us therefore show how to conclude it explicitly. The treatment in [23] first handles an auxiliary model, a random subset $M\subset\mathbb{N}$, where each l is in M independently with probability $\frac{1}{l}$. The subset M is analogous to the lengths of the cycles of a permutation, and the set S of subset sums of M, i.e.

$$S=\Big\{\sum_{a\in A}a : A\subset M\Big\}$$

is analogous to the sizes of the invariant sets of the permutation. The treatment of S revolves around a specific integer-valued random variable which they denote by T. The exact definition of T will not matter for us, but for completeness here it is:

$$T=\inf\Big\{t:\forall s\ge t\ |S\cap[1,s]|<1.01\log s, \sum_{l\in S\cap[1,s/\log s]}l\le s\Big\}$$

(see [23, the first displayed equation of §3 and Lemmas 2.1 and 2.2]). For T they prove the following two properties:

$$\begin{gathered}\mathbb{P}(T>k)\le Ck^{-c}\\ \mathbb{P}\Big(T<\frac{l}{\log l}\text{ and }l\in S\Big)\le Cl^{-1+\ln 2+0.0.2}\end{gathered}\tag{13}$$

See [23, Lemmas 2.1 and 2.2] for the first claim and [23, Lemma 3.1] for the second — the authors of [23] use the notation $m(n)=\lfloor n/\log n\rfloor$, see [23, (2.6)]. From (13) one immediately gets the analogous property for the model S i.e. that if $S_1,\dots,S_4$ are 4 independent copies of S then

$$\mathbb{P}\Big(\bigcup_{l=k}^{2k}\bigcap_{i=1}^{4}\{l\in S_i\}\Big)\le Ck^{-c}$$

because the events $T>k/\log k$ need be counted only once per i (and not once per i and l) while once these events are discarded, the probabilities remaining are $\le l^{-0.28}$ for each i, and taking a fourth power one may simply sum over l from k to 2k.

To pass from S to random permutations the authors of [23] do as follows. For $l<n/\log n$ they use a lemma due to Arratia and Tavaré [23, Lemma 1.8] which allows to couple $S|_{[1,n/\log n]}$ and the invariant sets of a random permutation, also restricted to $[1,n/\log n]$ and succeeds with probability bigger than $1-Cn^{-c\log\log n}$ (recall that a coupling of two random variables is a common probability space realising both variables and we say that a coupling "succeeds" if the two variables are in fact identical for that element of the probability space). This establishes the lemma for $k<n/2\log n$.

For $k > n/2\log n$ we have [23, Lemma 5.2] which claims that

$$\mathbb{P}(T < n/\log n, \text{the coupling succeeds}, E(i,l)) \le C\left(\frac{n}{\log n}\right)^{\ln 2+0.02}\frac{\log^3 n}{n}.$$

In other words, we lose only a polylogarithmic factor compared to the case of small k. Thus the case of large k follows identically, and the main clause of the lemma is proved. The "further" clause is identical. □

Proof of Theorem 1. Let us start the proof with the following reduction: it suffices to show, for every $k < n$, that the probability that f has a divisor of degree between k and $2k$ is smaller than $C/\log^2 k$. Indeed, once this is proved, one may handle divisors of small degree using Lemma 8, and then sum over k running through powers of 2 from $\omega(n)$ (ω from Lemma 8) to n. Let us, therefore, fix one $k < n$ until the end of the proof.

Let $\mathrm{red}_p(f)$ be the polynomial we get by reducing the coefficients of f modulo p. Then $\mathrm{red}_p(f)$ is a random uniform polynomial in $\mathbb{F}_p$, for every $p \mid L$, and the different $\mathrm{red}_p(f)$ are independent. For $r \in \{1,2,3,4\}$ Let X_r be an Ω-valued random variable which takes the value $(m_{1,r}, m_{2,r}, \dots)$ if the reduction of f modulo the r^{th} prime has $m_{i,r}$ irreducible factors of degree i for all i. Let $\mathscr{Q}$ be the event that for some $k \le l < 2k$ we may write $l = \sum i l_{i,r}$ for some $l_{i,r} \le m_{i,r}$ for all $r = 1,2,3,4$. Further, let $\mathscr{B}$ be the event that for some $r \in \{1,2,3,4\}$ and some $i < \log^2 k$ we have $m_{i,r} > \log^2 k$.

Now, by Lemma 6, $\mathbb{P}(\mathscr{B}) \le 4\cdot\log^2 k\cdot Ce^{-c\log^2 k}$ which is negligible. As for $\mathscr{Q}\setminus\mathscr{B}$, it is contained in the event that some $k \le l < 2k$ and some $\lambda_r < \log^6 k$ we may write

$$l - \lambda_r = \sum_{i>\log^2 k} i l_{i,r} \qquad l_{i,r} \le m_{i,r}$$

for all $r \in \{1,2,3,4\}$. Denote this event by $\mathscr{R}$. Then $\mathscr{R}$ is invariant to changes in the first $\log^2 k$ values of m and hence by Lemma 7

$$|\mathbb{P}_{X_n}(\mathscr{R}) - \mathbb{P}_{Y_n}(\mathscr{R})| \le C/\log^2 k.$$

(Formally, Lemma 7 is formulated for a single m and here we have 4, but this is equivalent. The easiest way to see this is probably to use Lemma 7 to construct a coupling between a single polynomial and a single permutation that succeeds with probability $C/\log^2 k$ and then simply couple the 4 polynomials to 4 permutations independently.) Finally, $\mathbb{P}_{Y_n}(\mathscr{R})$ can be estimated by Lemma 9 to get

$$\mathbb{P}_{Y_n}(\mathscr{R}) \le Ck^{-c}\log^{24} k.$$

We conclude that $\mathbb{P}_{X_n}(\mathscr{R}) \le C/\log^2 k$, hence that $\mathbb{P}_{X_n}(\mathscr{Q}\setminus\mathscr{B}) \le C/\log^2 k$, and hence that $\mathbb{P}_{X_n}(\mathscr{Q}) \le C/\log^2 k$. As explained in the first paragraph, this completes the proof of the theorem. □

3.1. **The Galois group.** For a permutation σ and an integer k let $\Psi(\sigma, k)$ be all permutations one may get by changing σ in elements belonging to cycles of σ each of whose length does not exceed k.

Lemma 10. *For any $\alpha < 1 - \frac{1+\log\log 2}{\log 2}$ the following holds. Let σ be a random permutation. Then the probability that there exists a transitive subgroup $G \not\geq A_n$ in S_n such that $G \cap \Psi(\sigma, n^\alpha) \neq \emptyset$ goes to 0 as $n \to \infty$.*

We follow Łuczak and Pyber [19] closely (they proved that $\mathbb{P}(\exists G : \sigma \in G) \to 0$ i.e. the same result but without allowing for small perturbations).

Lemma 11. *Let P be the probability that there exists $G \not\geq A_n$ primitive such that $G \cap \Psi(\sigma, n^\alpha) \neq \emptyset$. Then if $\alpha < 0.49$ then $P \to 0$ as $n \to \infty$.*

(the rate of decay may depend on α).

Proof. We follow [19]. The proof of [19] revolves around the notion of the *minimal degree* of a permutation group. Let us define it even though it is classical. For a permutation σ define

$$\deg \sigma = \#\{i \in \{1, \dots, n\} : \sigma(i) \neq i\}$$

i.e. the number of elements moved by σ; and for a group G of permutations we define its minimal degree by

$$\min\deg G = \min_{g \in G \setminus \{1\}} \deg g.$$

Then

Claim 1. *If $G \not\geq A_n$ is primitive then* $\min\deg G \geq (\sqrt{n} - 1)/2$.

Proof. There are two cases to consider. The first is that G is doubly transitive, i.e. for any $a \neq b$ and $c \neq d$ in $\{1, \dots, n\}$ one may find a permutation $\sigma \in G$ such that $\sigma(a) = c$ and $\sigma(b) = d$. This case goes back to [6] who showed that in this case $\min\deg G \geq n/4$, which is bigger than the required $(\sqrt{n} - 1)/2$. The other case is more recent, having been done in [2]: Theorem 0.3 of [2] states that for a primitive non-doubly-transitive permutation group G and for any $a \neq b$ in $\{1, \dots, n\}$ there are at least $(\sqrt{n} - 1)/2$ different values of $c \in \{1, \dots, n\}$ such that a and b are in a different *orbit* of the *stabilizer* of c. Let therefore $g \in G \setminus \{1\}$ and $a \neq b$ be in some non-trivial cycle of g. Then clearly g may not be in the stabilizer of any of the c given from [2, Theorem 0.3], so $\deg g \geq (\sqrt{n} - 1)/2$. □

Returning to the proof of Lemma 11, we apply Lemma 5 to find some cycle of our random permutation σ whose length l is in $[n^\alpha, n^{0.49}]$ and which has a prime divisor p such that $p > \log^3 n$ — Lemma 5 shows that this can be done with probability tending to 1. We apply Lemma 3 to see that p does not divide the length of any other cycle of σ, again with probability tending to 1.

Let now ρ be any permutation in $\Psi(\sigma, n^\alpha)$. Because $l > n^\alpha$, ρ will preserve the cycle of length l from σ. With probability tending to 1, σ has no more than $2\log n$ cycles (see, e.g., [19, Claim 1(i)]). Hence ρ is different from σ in no more than $(2\log n)n^\alpha$ places, and in particular the total length of all cycles of ρ whose length is divisible by p is no more than $l + 2n^\alpha \log n$.

Let therefore M be the product of all primes powers dividing lengths of cycles of ρ, different from powers of p. Then the points not fixed by ρ^M are exactly points which belong to cycles of ρ divisible by p, and by the previous discussion there are no more than $l + 2n^\alpha \log n$ of those (but at least l). In other words, $\deg \rho^M \le Cn^{0.49}$ and $\rho^M \ne 1$. Claim 1 then implies (for n sufficiently large) that ρ cannot belong to any primitive $G \not\geq A_n$, finishing the proof. □

Lemma 12. *Let P be the probability that there exists a transitive imprimitive group $G \le S_n$ such that $G \cap \Psi(\sigma, n^\alpha)$, and let $\alpha < \delta = 1 - \frac{1+\log\log 2}{\log 2}$. Then $P \to 0$ as $n \to \infty$.*

Proof. If G is transitive and imprimitive then there exists a nontrivial block system preserved by G, i.e. one may write $n = rs$ with $r, s > 1$ such that there is a division of $\{1, \dots, n\}$ into disjoint sets $A_1, \dots, A_r$ of common size s such that for every $\rho \in G$ and for any $i \in \{1, \dots, r\}$, $\rho(A_i) = A_j$ for some j. Following [19] we divide the proof to three cases according to the value of r.

Case 1. $2 \le r \le \exp(\log\log n\sqrt{\log n})$. Let G_1 be the event that the random permutation σ has, for each such r, a cycle L_r whose length is $> n^{0.99}$ and is not divisible by r. By [19, Claim 2] $\mathbb{P}(G_1) \to 1$ as $n \to \infty$. Let G_2 be the event that σ has no more than $2\log n$ cycles. By [19, Claim 1(I)] $\mathbb{P}(G_2) \to 1$ as $n \to \infty$. Let $G = G_1 \cap G_2$. The cycle L_r cannot be changed by changing short cycles of σ so it survives in any $\rho \in \Psi(\sigma, n^\alpha)$. Examine the set of A_i which intersect L_r. It cannot be all of the A_i because r does not divide the size of L_r. Since the union of the A_i which intersect L_r is invariant under ρ we found an invariant subset of ρ of size $n(a/r)$ for $1 \le a < r$. If we are also in G_2, then ρ can differ from σ in no more than $2n^\alpha \log n$ points, so the existence of an invariant subset of ρ implies that σ has an invariant subset of size x with $|x - na/r| \le 2n^\alpha \log n$. Denote this event by $B_{a,r}$. For each a and r, $\mathbb{P}(B_{a,r}) \le n^{\alpha-\delta+o(1)}$ by Eberhard, Ford and Green [9]. Summing over a and r (a total of $n^{o(1)}$ possibilities) we see that $\mathbb{P}(\bigcup B_{a,r}) \le n^{\alpha-\delta+o(1)}$. However, under $G \setminus \bigcup B_{a,r}$ no ρ may preserve any partition $A_1, \dots, A_r$ with $2 \le r \le \exp(\log\log n\sqrt{\log r})$ and this case is finished.

Case 2. $\exp(\log\log n\sqrt{\log n}) \le r \le n\exp(-\log\log n\sqrt{\log n})$. Let G be the event that the random permutation has a cycle L whose length is divided by some prime $p > n\exp(-\log\log n\sqrt{\log n})$. By [19, Claim 4] $\mathbb{P}(G) \to 1$ as $n \to \infty$. The cycle L will appear also in any $\rho \in \Psi(\sigma, n^\alpha)$, and of course prevents ρ from preserving any partition $A_1, \dots, A_r$ with r as above. Hence this case is also finished.

3. $n\exp(-\log\log n\sqrt{\log n}) \le r < n$. Let G_1 be the same event from case 1, i.e. the event that the random permutation σ has, for each $s \le \exp(\log\log n\sqrt{\log n})$, a cycle L_s whose length is $> n^{0.99}$ and is not divisible by s. By [19, Claim 2], $\mathbb{P}(G_1) \to 1$ as $n \to \infty$. Let G_2 be the event that any two cycles M_1 and M_2 of σ satisfy $\mathrm{lcd}(M_1, M_2) \le n^{0.9}$ (here and until the end of the lemma we do not distinguish between cycles and their lengths in the notation). By [19, Claim 1 (ii)], $\mathbb{P}(G_2) \to 1$ as $n \to \infty$ too. Fix now some r as above and let $s = n/r$. Under G_1, there exists a cycle L_s as above. Because $L_s > n^{0.99}$, it will be preserved in any $\rho \in \Psi(\sigma, n^\alpha)$. Assume ρ preserves a partition $A_1, \dots, A_r$ and examine the collection of blocks A_i which intersect L_s. Their union cannot be covered completely by L_s (because s does not divide L_s) hence it must contain at least one additional cycle, denote it by M. But then $\mathrm{lcd}(M, L_s)$ is (at least) the number of such blocks A_i and in particular $\mathrm{lcd}(M, L_s) > L_s \exp(-\log\log n\sqrt{\log n})$. This means, for n sufficiently large, that $M > n^{0.9}$

and hence M appears also in σ. But the appearance of both M and L_s in σ contradicts the event G_2. Hence we get that the event that for some r as above, some $\rho \in \Psi(\sigma, n^\alpha)$ preserves some partition $A_1, \dots, A_r$, is contained in $(G_1 \cup G_2)^c$. We get that the probability of this event also goes to zero with n. The case is finished, and so is the lemma. □

Proof of Lemma 10. Lemmas 11 and 12 do all the work. □

Proof of Theorem 2. Recall that $\Psi(\sigma, k)$ denotes the set of permutations which differ from σ only in elements which belong to cycles shorter than k. By Lemma 10, with probability tending to 1, there is no transitive subgroup $G \not\geq A_n$ such that $G \cap \Psi(\sigma, n^\alpha) \neq \emptyset$. (here σ is a random permutation and α is some arbitrary number in $\left(0, 1 - \frac{1+\log\log 2}{\log 2}\right)$ whose exact value will play no role). Let us reformulate this in the notations of § 2.2: for a tuple $m = (m_1, \dots, m_n)$ let $E(m)$ be the event that there exists a transitive subgroup $G \not\geq A_n$ and an element $g \in G$ such that g has exactly m_i cycles of length i for all $i \geq n^\alpha$. Then the promised reformulation is:

$$\lim_{n\to\infty} \mathbb{P}_{Y_n}(E) = 0.$$

Further, E is clearly invariant to changing cycles of σ shorter than n^α. Hence we may apply Lemma 7 to it. We get that

$$\lim_{n\to\infty} \mathbb{P}_{X_n}(E) = 0. \tag{14}$$

Now let $f = X^n + \sum_{i=0}^{n-1} \zeta_i X^i$, $\zeta_i \in \{1, \dots, L\}$ be a random polynomial as in the theorem and let $p \mid L$ be a prime number. In particular, $\bar{f} := \mathrm{red}_p(f)$ is a random monic polynomial of degree n in $\mathbb{F}_p[X]$. Let N be the splitting field of f over $\mathbb{Q}$ in $\mathbb{C}$, $R \subseteq N$ the set of roots of f, and $G = \mathrm{Gal}(N/\mathbb{Q}) \leq \mathrm{Sym}(R)$. By Theorem 1, with probability tending to 1 we have that f is irreducible, which is equivalent to say that G is transitive.

Let O be the ring of integers of N. By Chevalley's Theorem [11, Proposition 2.3.1] there exists a prime ideal $\mathfrak{P}$ of O that lies over p, i.e. $\mathfrak{P} \cap \mathbb{Q} = p\mathbb{Z}$. Choose one such $\mathfrak{P}$ arbitrarily. The map $O \to O/\mathfrak{P}$ takes $\mathbb{Z}$ to $\mathbb{F}_p$ so if we write, in $O[X]$, $f = \prod_{\rho\in R}(X - \rho)$, then we get for $\bar{f}$, our reduction of f to $\mathbb{F}_p$, that $\bar{f} = \prod_{\rho\in R}(X - \bar{\rho})$, where $\bar{\rho}$ is the image of ρ under the map $O \to O/\mathfrak{P}$.

We may write $\bar{f} = \phi\psi$, with relatively prime $\phi, \psi \in \mathbb{F}_p[X]$, such that ϕ is squarefree and ψ is squarefull (i.e., the multiplicity of each irreducible factor of ψ is at least 2). The probability that $\bar{f}$ has a square of degree k dividing it, is $q^{-k/2}$; hence with probability tending to 1, $\deg\psi \leq n^\alpha$ (with α as above). We decompose R as $R = R_\phi \cup R_\psi$, with $R_\phi = \{\rho \in R : \phi(\bar{\rho}) = 0\}$ and $R_\psi = R \setminus R_\psi$. So the map $\rho \in R_\phi \mapsto \bar{\rho}$ surjects onto the roots of ϕ. Since ϕ and ψ are relatively prime, we cannot have $\rho \in R_\phi$ such that $\psi(\bar{\rho}) = 0$. Thus, since ϕ is squarefree, the map $\rho \mapsto \bar{\rho}$ is a bijection from R_ϕ onto the roots of ϕ.

Now, the map $G \to \mathrm{Gal}((N/\mathfrak{P})/\mathbb{F}_p)$ is onto (see, e.g., [11, Lemma 6.1.1(a)]) and hence there exists an element $\tau \in G$ which maps to the Frobenius element $x \mapsto x^p$ i.e. satisfying

$$\tau\rho \equiv \rho^p \mod \mathfrak{P},$$

for all $\rho \in R$. Thus τ acts on R_ϕ the same as the Frobenius map acts on the roots of ϕ. The cycle lengths of the latter is the same as the degrees of the irreducible factors of ϕ (this is classical, and follows from the fact that Galois groups over a finite field are generated by the Frobenius element, and the roots of each irreducible factor is an orbit of the Galois group.) The rest of the cycles of τ are of total size $\leq n^\alpha$. All in all, we get that the cycle lengths of τ distribute the same as of the degrees of the irreducible factors of $\bar{f}$ up to cycles of length $\leq n^\alpha$. By (14), with probability tending to 1, the element τ can not lie in a transitive group other than A_n or S_n. But it lies in G, so $G = A_n$ or $G = S_n$. □

4. Heuristics and simulations about the A_n vs. S_n problem

Conjecture. *Let f be as in Theorem 1. Then*

$$\lim_{n\to\infty} \mathbb{P}(\textit{the Galois group of } f \textit{ is } S_n) = 1.$$

Let us first explain why this conjecture does not follow from our methods. Indeed, considering reductions modulo p one gets 4 elements of the Galois group (the lifts of the Frobenius elements), whose conjugation classes are independent, and close to uniform in the sense explained above, i.e. with some deviations in the very smallest cycles. Now, 4 independent uniform permutations have probability exactly $\frac{1}{16}$ to all be in A_n. The comparison techniques described in §2 can be used to conclude that, for a random polynomial, all 4 lifts of the Frobenius element belong to A_n with probability close to $\frac{15}{16}$. Hence we get that the lower bound for the limit of the probabilities in the conjecture is at least bounded away from 0, but not quite 1.

To differentiate S_n from A_n, one may use the discriminant. Recall the definition of the discriminant $\Delta(f) = \prod_{i<j}(\alpha_i - \alpha_j)^2$, where $\alpha_1, \dots, \alpha_n$ are the complex roots of f. As a symmetric expression in the roots, $\Delta(f)$ is an integer. The basic Galois theoretic property of $\Delta(f)$, for separable f, is that $G_f \leq A_n$ if and only if $\Delta(f)$ is a perfect square. Therefore, in order to show in Theorem 2 that $\mathbb{P}(G_f = S_n) \to 1$, we have to show that $\mathbb{P}(\Delta(f) \neq \square) \to 1$.

Hence it makes sense to study $\Delta(f)$ for a random f. Simulations done by Igor Rivin show that $\log|\Delta(f)|$ has an asymptotically normal law, with average and variance both linear in n. It would be interesting to prove that rigorously, maybe even for the case that f has gaussian coefficients (in the gaussian case, extremely fine estimates have been shown for the distribution of the zeroes of f, we covered some references in the introduction). This gives the following crude heuristic: the discriminant is a random very large (exponential in n) integer, so the probability that it is a square should be very small. We performed simulations of the probability that the discriminant is a square, and it seems to decay exponentially in the degree, though there are also arithmetic effects: for example, it seems the discriminant of $\sum_{i=0}^n \pm x^i$ can never be a square when $n \equiv 2$ or $4 \bmod 8$.

Another fact discovered during simulations is that the sign is approximately evenly distributed (though it seems the inhomogeneity does not decay as $n \to \infty$ and depends on arithmetic properties of n). As Ofer Zeitouni remarked to us, the sign of the discriminant is simply $(-1)^{(\text{number of non-real roots})/2}$ since the contribution of $(\lambda_i - \lambda_j)^2$ is positive if they are both real, and if λ_j is non-real then the contribution of $(\lambda_i - \lambda_j)^2(\lambda_i - \overline{\lambda_j})^2$ is also positive, and similarly in the case that they are both non-real. The only negative contributions

degree	jump	degree	jump	degree	jump	degree	jump	degree	jump
9	4	29	2	49	4	69	1	89	2
13	3	33	8	53	2	73		93	
17	2	37		57		77	2	97	3
21	10	41	1	61	5	81			
25	12	45	11	65	2	85	14		

TABLE 1. The arithmetic progressions of allowable values of k. Holes indicate values of the degree where only one value of k is allowed.

come from $(\lambda_i - \overline{\lambda_i})^2$. A lot is known about the number of real zeroes, but to the best of our knowledge their parity has not been studied.

Other interesting phenomena discovered in simulations relate to the powers of 2 that may divide $\Delta(\sum \pm x^i)$. Denote by k the maximal number such that 2^k divides the discriminant. Then there are many connections between k and the degree, n. Let us present simulations for $n \leq 100$:

- If n is even then $k = 0$. Indeed, this is because modulo 2 our polynomial is always $P = \sum_{i=0}^n x^i$ and then $P = P'(1 + x) + x^n$ and they cannot have a joint zero, so the discriminant is odd.
- If n is odd then always $k \geq n - 1$. Shoni Gilboa gave a beautiful proof of this fact, which we will only sketch: write $\Delta(f) = \det(MM^*)$ where M is the Vandermonde matrix α_i^j and α_i are still the roots of P. The entries of MM^* can be related to the coefficients of P using the *Newton identities* and it is a simple inductive check that all entries turn out to be odd integers. Subtracting the first row from all the others one can pull out 2^{n-1} and still get an integeral matrix.
- The tail of the distribution of k is much fatter than we expected. Values of 120 and more are easily observed in simulations (say with 10^5 runs). This is especially true if $n \equiv 3 \bmod 4$.
- If $n \equiv 7 \bmod 8$ then k cannot take the values n and $n + 2$.
- If $n \equiv 3 \bmod 8$ then k cannot take the values n and $n + 4$. Sometimes the value $n+10$ is also prohibited: in our simulations the values of n for which this happened were 27, 43, 51, 67, 75, 91 (notice the absence of 99, so this is not related to modulo 24).
- If $n \equiv 1 \bmod 4$ then k is much more restricted. Occasionally ($n = 37, 57, 73, 81, 93$), it may only take the value $n - 1$. Sometimes it is restricted to two values ($n = 9, 21, 25, 33, 45, 85$). And very typically it is restricted to an arithmetic progression (starting from $n - 1$), see table 1 for the jumps. The jump always divides $n - 1$, so the arithmetic progression can be thought of as starting from 0. Notice two values of n, 41 and 69, for which the jump is 1, i.e. all values are allowed. But these are the only exceptions that came up in our simulations.

We are not sure how all this reflects on the probability that the discriminant be a square, but we thought it is interesting enough to mention.

Finally, we studied the probability that the determinant of a random matrix with ± 1 entries is a square. Since determinant and discriminant share both the 'd' in the beginning and the 'minant' at the end, this seems quite relevant (also, of course, the discriminant has at least two formulas as the determinant of a matrix whose entries are functions of the coefficients: one as the determinant of MM^* that we mentioned above, and another that comes from the fact that it is the resultant of P and P', giving a $(2n+1)\times(2n+1)$ matrix whose entries are multiples of the coefficients of P). In this case we did manage to prove the following:

Theorem 13. *Let M be an $n\times n$ matrix with i.i.d. entries taking the value 0 with probability $\frac{1}{2}$ and the values 1 and -1 with probability $\frac{1}{4}$ each. Then*

$$\lim_{n\to\infty}\mathbb{P}(\exists k\in\mathbb{Z} \textit{ s.t. } \det M = k^2) = 0.$$

The proof of Theorem 13 is too long to include here, and we will publish it in a complementary paper [4].

References

[1] Richard Arratia, Andrew D. Barbour and Simon Tavaré, *On random polynomials over finite fields.* Math. Proc. Cambridge Philos. Soc. 114:2 (1993), 347–368. Available at: cambridge.org/13FBE1

[2] László Babai, *On the order of uniprimitive permutation groups.* Ann. of Math. (2) 113:3 (1981), 553–568. Available at: jstor.org/2006997

[3] John Baez, The Beauty of Roots. Available at: ucr.edu/baez/roots

[4] Lior Bary-Soroker and Gady Kozma, Supplement to the article "irreducible polynomials with bounded height". In preparations.

[5] André Bloch and George Pólya, On the roots of certain algebraic equations. Proc. Lond. Math. Soc., II. Ser. 33 (1931), 102–114. Available at: wiley.com/s2-33.1.102

[6] Alfred Bochert, *Ueber die Classe der transitiven Substitutionengruppen* [German: On the class of transitive substitution groups]. Math. Ann. 40:2 (1892), 176–193. springer.com/BF01443562

[7] John D. Bovey, *The probability that some power of a permutation has small degree.* Bull. London Math. Soc. 12:1 (1980), 47–51. Available at: cambridgejournals.org/12.1.47

[8] Rainer Dietmann, On the distribution of Galois groups. Mathematika, 58:1 (2012), 35–44. Available at: cambridge.org/AC9AE5

[9] Sean Eberhard, Kevin Ford and Ben Green, Invariable generation of the symmetric group. Duke Math. J. 166:8 (2017), 1573–1590. Available at: projecteuclid.org/1486695668

[10] Paul Erdős, and Pál Turán, On the distribution of roots of polynomials. Ann. Math. 51 (1950), 105–119. Available at: jstor.org/1969500

[11] Michael D. Fried and Moshe Jarden, Field Arithemtic, 3rd edition. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [German: Results in Mathematics and Related Areas. 3rd Series], 11. Springer-Verlag, Berlin, 2008.

[12] Patrick X. Gallagher, On the distribution of primes in short intervals. Mathematika 23:1 (1976), 4–9. Available at: cambridge.org/11858C

[13] Ildar Ibragimov and Ofer Zeitouni, On roots of random polynomials. Trans. Amer. Math. Soc. 349:6 (1997), 2427–2441. Available at: ams.org/01766-2

[14] M. Kac, On the average number of real roots of a random algebraic equation, Bull. Amer. Math. Soc. 49 (1943) 314–320. Available at: ams.org/07912-8

[15] Sergei V. Konyagin, On the number of irreducible polynomials with 0,1 coefficients. Acta Arith. 88:4 (1999), 333–350. Available at: pldml.icm.edu.pl/aa8844

[16] Gady Kozma and Ofer Zeitouni, *On common roots of random Bernoulli polynomials*. Int. Math. Res. Not. IMRN 2013, no. 18, 43344347. Available at: oxfordjournals.org/rns.164.full

[17] Gerald Kuba, On the distribution of reducible polynomials. Mathematica Slovaca, 59:3 (2009), 349–356. Available at: degruyter.com/0131-6

[18] John E. Littlewood and Albert C. Offord, On the number of real roots of a random algebraic equation. II. Proc. Cambridge Philos. Soc. 35, (1939), 133-148. Available at: cmabridge.org/45392A

[19] Tomasz Łuczak and László Pyber, *On random generation of the symmetric group*. Combin. Probab. Comput. 2:4 (1993), 505–512. Available at: cambridge.org/C6FE584

[20] Hugh L. Montgomery and Robert C. Vaughan, Multiplicative number theory. I. Classical theory. Cambridge Studies in Advanced Mathematics, 97. Cambridge University Press, Cambridge, 2007.

[21] Andrew M. Odlyzko and Bjorn Poonen, Zeros of polynomials with 0, 1 coefficients, Enseign. Math. 39:3-4 (1993), 317–348. Available at: e-priodica.ch/1993:39#566

[22] Sean O'Rourke and Philip Matchett Wood, *Low-degree factors of random polynomials*. Preprint, available at: arXiv:1608.01938

[23] Robin Pemantle, Yuval Peres and Igor Rivin, Four random permutations conjugated by an adversary generate S_n with high probability. Random Structures Algorithms 49:3 (2016), 409–428. Available at: wiley.com/rsa.20632, arXiv:1412.3781

[24] Paul Pollack, Irreducible polynomials with several prescribed coefficients. Finite Fields and Their Applications, 22 (2013), 70–78. Available at: sciencedirect.com/000336

[25] Igor Rivin, Galois groups of generic polynomials. Preprint (2015). Available at: arXiv:1511.06446

[26] Larry A. Shepp and Robert J. Vanderbei, The complex zeros of random polynomials. Trans. Amer. Math. Soc. 347 (1995), 4365–4384. Available at: ams.org/1308023-8

[27] Some guy on the street, Irreducible polynomials with constrained coefficients, MathOverflow. Available at: mathoverflow.net/7969

[28] Bartel L. van der Waerden, Die Seltenheit der reduziblen Gleichungen und der Gleichungen mit Affekt. Monatsh. Math. Phys., 43:1 (1936), 133–147. Available at: springer.com/BF01707594

Raymond and Beverly Sackler School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel

E-mail address: barylior@post.tau.ac.il

Department of Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel

E-mail address: gady.kozma@weizmann.ac.il