#A45 INTEGERS 17 (2017)

STATISTICAL DISTRIBUTION OF ROOTS OF A POLYNOMIAL MODULO PRIMES

Yoshiyuki Kitaoka
Department of Mathematics, Meijo University, Tenpaku, Nagoya, Japan
kitaoka@meijo-u.ac.jp

Received: 5/19/15, Revised: 2/2/17, Accepted: 9/5/17, Published: 10/16/17

Abstract

Let $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ be an irreducible polynomial with integer coefficients. For a prime p for which $f(x)$ is fully splitting modulo p, we consider n roots r_i of $f(x) \equiv 0 \bmod p$ with $0 \le r_1 \le \cdots \le r_n < p$ and propose several conjectures on the distribution of an integer $\lceil \sum_{i \in S} r_i/p \rceil$ for a subset S of $\{1, \ldots, n\}$ when $p \to \infty$.

1. Introduction

Throughout this paper, unless otherwise specified, a polynomial means a monic *irreducible* polynomial of degree > 1 with integer coefficients, and the letter p denotes a prime number. For a polynomial $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ of degree n and a prime number p, we say that $f(x)$ is *fully splitting modulo* p if there are integers $r_1, r_2, \ldots, r_n$ satisfying $f(x) \equiv \prod(x - r_i) \bmod p$. We assume that

$$0 \le r_1, \ldots, r_n < p. \tag{1}$$

Substituting

$$Spl(f, X) := \{p \le X \mid f(x) \text{ is fully splitting modulo } p\}$$

for a positive number X and $Spl(f) := Spl(f, \infty)$, we know that $Spl(f)$ is an infinite set and that the density theorem due to Chebotarev holds; that is,

$$\lim_{X \to \infty} \frac{\#Spl(f, X)}{\#\{p \le X\}} = \frac{1}{[\mathbb{Q}(f) : \mathbb{Q}]},$$

where $\mathbb{Q}$ is the rational number field and $\mathbb{Q}(f)$ is a finite Galois extension field of $\mathbb{Q}$ generated by all roots of $f(x)$ [12].

For $p \in Spl(f)$, the definition of roots r_i with (1) clearly implies that

$$a_{n-1} + r_1 + \cdots + r_n = C_p(f)p \tag{2}$$

for an integer $C_p(f)$. The author has previously studied the statistical distribution of $C_p(f)$ and local roots r_i for $p \in Spl(f)$ [4]–[6], [8], [9]. A basic fact that we need here is as follows.

Proposition 1. *If $p \in Spl(f)$ is sufficiently large, then for any subset S of $\{1, 2, \dots, n\}$ with $\#S = n-1$, we have*

$$\left\lceil \sum\nolimits_{j\in S} r_j/p \right\rceil = C_p(f), \tag{3}$$

where $\lceil x \rceil$ is an integer such that $x \le \lceil x \rceil < x+1$.

A proof of Proposition 1 is given in [5], where it is initially supposed that a sequence of $n!$ points $(r_{\sigma(1)}/p, \dots, r_{\sigma(n-1)}/p)$ for all permutations $\sigma \in S_n$ is uniformly distributed in $[0,1)^{n-1}$ when $p \to \infty$ if a polynomial $f(x)$ is indecomposable. However, this turns out to be false (counterexamples in the case of $n = 6$ are given in [9] and in Section 4 here). Here, a polynomial $f(x)$ is called decomposable if there are polynomials $g(x)$ and $h(x)$ satisfying $f(x) = g(h(x))$ and $1 < \deg h < \deg f$, and indecomposable otherwise. In this paper, we give detailed observations in the case of $n \le 6$. To do so, we introduce an ordering among roots r_i as follows:

$$0 \le r_1 \le \cdots \le r_n < p. \tag{4}$$

This determines roots r_i uniquely. We note that $r_1 = 0$ implies $a_0 \equiv 0 \bmod p$ and (4) is equivalent to $0 < r_1 < \cdots < r_n < p$ for a sufficiently large $p \in Spl(f)$ by the irreducibility of $f(x)$.

In Section 2, we recall observations related to the uniform distribution, and in Section 3, we introduce a new density and give observations in the case of $\deg f \le 5$, where the density is independent of a polynomial if it is irreducible and indecomposable. In Section 4, we give observations in the case of $\deg f = 6$, where the density depends on each polynomial. In Section 5, we give some theoretical results to analyze the data, although it is too far to clarify the whole picture. The data presented in this paper were obtained using pari/gp.[1]

2. Uniform Distribution

Let us recall a fundamental fact about uniform distribution.

Lemma 1. *For a natural number n, the volume of a subset of the unit cube $[0,1)^n$ defined by $\{(x_1, \dots, x_n) \in [0,1)^n \mid x_1 + \cdots + x_n \le x\}$ is given by*

$$U_n(x) := \frac{1}{n!} \sum_{i=0}^{n} (-1)^i \binom{n}{i} \max(x-i, 0)^n,$$

[1] The PARI Group, PARI/GP version `2.7.0`, Bordeaux, 2014, http://pari.math.u-bordeaux.fr/.

and for an integer k with $1 \leq k \leq n$, we have

$$U_n(k) - U_n(k-1) = \frac{1}{n!}\sum_{i=0}^{k}(-1)^i\binom{n+1}{i}(k-i)^n. \tag{5}$$

See [2] for a proof of the first statement, from which identity (5) follows easily. We note that $A(n,k) := n!(U_n(k) - U_n(k-1))$ $(1 \leq k \leq n)$ is called an Eulerian number and satisfies

$$A(1,1) = 1,\ A(n,k) = (n-k+1)A(n-1,k-1) + kA(n-1,k).$$

Necessary values of $A(n,k)$ in this paper are

$n \setminus k$	1	2	3	4	5
2	1	1			
3	1	4	1		
4	1	11	11	1	
5	1	26	66	26	1

and we note that

$$vol\left(\left\{(x_1,\ldots,x_n) \in [0,1)^n \mid \left\lceil \sum x_i \right\rceil = k\right\}\right) = \frac{A(n,k)}{n!}. \tag{6}$$

For a polynomial $f(x)$ of degree n, (2) implies

$$r_1/p + \cdots + r_n/p = C_p(f) - a_{n-1}/p,$$

whose left-hand side is close to an integer $C_p(f)$ when p is large. Thus, the sequence of points $(r_1/p,\ldots,r_n/p)$ is not uniformly distributed in the unit cube $[0,1)^n$ as $p \to \infty$. However, the sequence of $n!$ points $(r_{\sigma(1)}/p,\ldots,r_{\sigma(n-1)}/p)$ for all $\sigma \in S_n$ is expected to be uniformly distributed in $[0,1)^{n-1}$ for the majority of polynomials. This is true without exception in the case of $n = 2$ [1], [13]. If the expectation is true, then the density of the distribution of the value $C_p(f)$ in (2) is given by Lemma 1 as follows:

$$\lim_{X\to\infty}\frac{\#\{p \in Spl(f,X) \mid C_p(f) = k\}}{\#Spl(f,X)} = \frac{A(n-1,k)}{(n-1)!}, \tag{7}$$

since Proposition 1 implies

$$\#\{p \in Spl(f,X) \mid C_p(f) = k\} = \#\{p \in Spl(f,X) \mid \lceil \textstyle\sum_{i\in S} r_i/p \rceil = k\} + O(1) \tag{8}$$

for any subset S of $\{1,\ldots,n\}$ with $\#S = n-1$. Computer experiments support (7) well.

Although we began our study with the distribution of $C_p(f)$, which originated from [3] and [7], it is more interesting in view of (7) and (8) to study the distribution of the value $\lceil(\sum_{i\in S} r_i)/p\rceil$ with the condition (4) on local roots r_i for a given subset S of $\{1,\ldots,n\}$. We provide some observations in the following sections.

3. New Density

We introduce here a new type of distribution. Statements on the density without proof hereinafter are conjectures based on numerical experiments.

Let $f(x)$ be a polynomial of degree n and let p be a prime in $Spl(f)$. We assume the global order (4) on local roots; that is, we number local roots r_i of $f(x)$ modulo p as follows:

$$0 \leq r_1 \leq \cdots \leq r_n < p \quad (f(r_i) \equiv 0 \bmod p).$$

As noted above, we have $0 < r_1 < \cdots < r_n < p$ if p is sufficiently large. Let us consider a more general density than the left-hand side of (7). For a subset S of $\{1, 2, \ldots, n\}$, we define a frequency table $Pr(f, S, X)$ by

$$Pr(f, S, X) := [F_1, \ldots, F_s],$$

where $s := \#S$ and

$$F_k := F_k(f, S, X) = \frac{\#\{p \in Spl(f, X) \mid \lceil \sum_{i \in S} r_i/p \rceil = k\}}{\#Spl(f, X)}. \tag{9}$$

It is clear that the assumption $0 \leq r_i < p$ $(i = 1, \ldots, n)$ implies $F_k = 0$ unless $0 \leq k \leq s$. We see easily that $\lim_{X\to\infty} F_0(f, S, X) = 0$ since primes contributing to the numerator of (9) divide the constant term a_0 of $f(x)$.

Next, note that we may confine ourselves to the case $2 \leq s \leq n - 1$. Suppose that $F_k(f, S, X) \neq 0$ with $s = 1$, say $S = \{i\}$; then, the equation $\lceil r_i/p \rceil = k$ implies $k = 1$ for every sufficiently large p, which implies $\lim_{X\to\infty} F_1(f, S, X) = 1$ and $\lim_{X\to\infty} F_k(f, S, X) = 0\,(k \neq 1)$. When $s = n$, that is, $S = \{1, \ldots, n\}$, we have

$$\lceil \textstyle\sum_{i \in S} r_i/p \rceil = \lceil C_p(f) - a_{n-1}/p \rceil = \begin{cases} C_p(f) & \text{if } a_{n-1} \geq 0, \\ C_p(f) + 1 & \text{if } a_{n-1} < 0, \end{cases}$$

and so this case is reduced to the case of $s = n - 1$ by (8), which has been previously studied [4]–[6], [8].

Assuming that $s = n - 1$ and f is indecomposable, we expect that in the case of $n \leq 5$, a sequence of $n!$ points $(r_{\sigma(1)}/p, \ldots, r_{\sigma(n-1)}/p)$ $(\sigma \in S_n)$ is uniformly distributed as $p \to \infty$, which implies (7). However, this is not the case if $n = 6$, as we will see later.

We use the abbreviation

$$Pr(f, S) := \lim_{X\to\infty} Pr(f, S, X) = \lim_{X\to\infty} [F_1(f, S, X), \ldots, F_s(f, S, X)],$$

assuming that the limit exists, something that the author has no data to refute. The first expectation is as follows.

Conjecture 1. Suppose that $f(x)$ is not equal to $g(h(x))$ for any quadratic polynomial $h(x)$. Then, for every j with $1 \leq j \leq n$, we have

$$Pr(f,S) = [1,1]/2 \text{ for } S = \{j, n+1-j\},$$

where $[1,1]/2$ means $[1/2, 1/2]$ for simplicity; we adopt this notation hereinafter.

We checked the following polynomials. Let BP be a polynomial of degree $n = 4$, 5, or 6 with coefficients equal to 0 or 1, and let α be one of its roots. For a number $\beta = \sum_{i=1}^{n} c_i \alpha^{i-1}$ with $0 \leq c_i \leq 2$, we take a polynomial f of degree n for which β is a root. We skip a reducible polynomial and also a decomposable polynomial, which is in the form $f(x) = g(h(x))$ with $\deg h = 2$. Considering that

$$F_k := F_k(f, S, X) \to 1/2 \quad (k = 1, 2,\ X \to \infty)$$

holds under Conjecture 1, we judge that the expectation is true if

$$|F_1 - F_2| < 0.1$$

for a large number X, since $F_1 + F_2 = 1$. The excluded case is as follows.

Proposition 2. *Suppose that a polynomial $f(x) = x^n + a_{n-1}x^{n-1} + \dots$ is equal to $g(h(x))$ for a quadratic polynomial $h(x)$. Then, for $S = \{j, n+1-j\}$ $(1 \leq j \leq n)$, we have*

$$Pr(f,S) = \begin{cases} [1,0] & \textit{if } a_{n-1} \geq 0, \\ [0,1] & \textit{if } a_{n-1} < 0. \end{cases}$$

Proof. We only have to see that, except for finitely many primes p, the value $\lceil (r_j + r_{n+1-j})/p \rceil$ is equal to 1 or 2 according to whether $a_{n-1} \geq 0$ or $a_{n-1} < 0$, respectively. We note that $\deg g = n/2$ and we may assume that $h(x) = (x+a)^2$ or $h(x) = (x+a)(x+a+1)$ for an integer a according to whether the coefficient of x of $h(x)$ is even or odd, respectively. In the case of $h(x) = (x+a)^2$, $a_{n-1} = an$ is easy, and if $r \in \mathbb{Z}$ $(0 < r < p)$ is a root of $f(x) = g((x+a)^2) \equiv 0 \bmod p$, then $p - r - 2a$ is also one of its roots, and we see that $0 < p - r - 2a < p$ for a sufficiently large p [9]. Hence, by the assumption (4), the sequences $r_1 < \cdots < r_n$ and $p - r_n - 2a < \cdots < p - r_1 - 2a$ are identical. Thus, we have $r_j + r_{n+1-j} = p - 2a = p - 2a_{n-1}/n$, which implies

$$\begin{cases} (r_j + r_{n+1-j})/p \leq 1 & \text{if } a_{n-1} \geq 0, \\ (r_j + r_{n+1-j})/p > 1 & \text{if } a_{n-1} < 0. \end{cases}$$

This completes the proof in the case of $h(x) = (x+a)^2$. In the case of $h(x) = (x+a)(x+a+1)$, noting that $a_{n-1} = (1+2a)n/2$ and both r_i and $p - r_i - 1 - 2a$ $(i = 1, \dots, n)$ are roots, we have $r_j + r_{n+1-j} = p - 1 - 2a = p - 2a_{n-1}/n$ in a similar way as above, which completes the proof. □

For a subset S of $\{1,\dots,n\}$, we put $S^\vee := \{n+1-i \mid i \in S\}$. Then, for $Pr(f,S) = [F_1,\dots,F_s]$, we have

$$Pr(f,S^\vee) = [F_s,\dots,F_1]$$

empirically in many cases, which is equivalent to

$$Pr(f,S^\vee) = Pr(f,S)^\vee, \tag{10}$$

putting $[a_1,\dots,a_s]^\vee := [a_s,\dots,a_1]$.

Proposition 3. *Under the assumption that*

(A) $\sum_{j\in S^\vee} r_j/p$ *is not an integer for every sufficiently large prime* $p \in Spl(f)$,

we have

$$Pr((-1)^n f(-x),S)^\vee = Pr(f(x),S^\vee).$$

Moreover, if $Pr((-1)^n f(-x),S) = Pr(f(x),S)$ *holds, then we have* (10).

Proof. Since we have $f(x) \equiv \prod(x-r_i) \bmod p$ with $0 < r_1 < \cdots < r_n < p$ for a sufficiently large prime p, we get $(-1)^n f(-x) \equiv \prod(x+r_i) \equiv \prod(x-R_i) \bmod p$ with

$$0 < R_1 := p - r_n < \cdots < R_i := p - r_{n+1-i} < \cdots < R_n := p - r_1 < p.$$

Noting an equality $\lceil s-r \rceil = s+1-\lceil r \rceil$ for $s := \#S \in \mathbb{Z}$, $r \notin \mathbb{Z}$, we see that $\lceil \sum_{i\in S} R_i/p \rceil = \lceil s - \sum_{j\in S^\vee} r_j/p \rceil = s+1-\lceil \sum_{j\in S^\vee} r_j/p \rceil$, which implies $F_k((-1)^n f(-x),S,X) = F_{s+1-k}(f(x),S^\vee,X)$. Hence, we have the desired equation $Pr((-1)^n f(-x),S,X) = Pr(f(x),S^\vee,X)^\vee$. □

Remark 1. If f is indecomposable with $n \le 5$, then $Pr(f,S)$ seems to be dependent on only S and $\deg f$, as we see below. Hence, this proposition elucidates (10). Therefore, The assumption (A) is not necessarily true. For example, for $f = x^4+1$, both $r_1 < \cdots < r_4$ and $p - r_4 < \cdots < p - r_1$ are the set of local roots. Hence, we have $r_1 = p - r_4$ and $r_2 = p - r_3$, that is, $\sum_{i\in S} r_i/p = 1$ for $S = \{1,4\},\{2,3\}$. Another example is the polynomial f_3 (cf. Remark 4).

Before giving a sufficient condition to (A), let us recall a relation between the decomposition of a polynomial $f(x)$ modulo p and that of p to the product of prime ideals over $F := \mathbb{Q}(\alpha)$, where α is a root of $f(x)$. Denote the ring of integers of F by O_F and prime ideals lying above p by $\mathfrak{p}_i$. Suppose that $p \in Spl(f)$ is sufficiently large and $r_1,\dots,r_n$ are roots of $f(x)$ modulo p; then, we have the decomposition of p: $pO_F = \mathfrak{p}_1 \cdots \mathfrak{p}_n$ and we may suppose that, by renumbering

$$\mathfrak{p}_i = (\alpha - r_i)O_F + pO_F \text{ and } O_F/pO_F \cong O_F/\mathfrak{p}_1 \oplus \cdots \oplus O_F/\mathfrak{p}_n, \tag{11}$$

in particular $\alpha \equiv r_i \bmod \mathfrak{p}_i$. The isomorphism in (11) is given by

$$\beta \bmod pO_F \mapsto (\beta \bmod \mathfrak{p}_1,\dots,\beta \bmod \mathfrak{p}_n)$$

and

$$O_F/\mathfrak{p}_i \cong \mathbb{Z}/p\mathbb{Z}.$$

Moreover, p splits fully over F if and only if p splits fully over the field $\mathbb{Q}(f)$ generated by all roots of $f(x)$.

Proposition 4. *If the condition* (A) *for S with $\#S = s$ does not hold, then a sum of some s roots of $f(x)$ is zero, that is, $f(x) = (x^s + 0 \cdot x^{s-1} + \dots)(x^{n-s} + a_{n-1}x^{n-s-1} + \dots)$.*

Proof. The assumption means that there are infinitely many primes p such that $\sum_{j \in S^\vee} r_j \equiv 0 \bmod p$. Let α be a root of $f(x)$ and put $F = \mathbb{Q}(\alpha)$, and let $K := \mathbb{Q}(f)$ be a field generated by all roots of $f(x)$. For a sufficiently large prime $p \in Spl(f)$ and roots $r_1, \dots, r_n$ of $f(x)$ modulo p with (4), let $\mathfrak{p}_i$ be a prime ideal of F defined by (11) and let $\mathfrak{p}_i = \mathfrak{P}_{i,1} \dots \mathfrak{P}_{i,g}$ ($g = [K : F]$) be the decomposition of $\mathfrak{p}_i$ to the product of prime ideals over K. The congruence $\alpha \equiv r_i \bmod \mathfrak{p}_i$ implies $\alpha \equiv r_i \bmod \mathfrak{P}_{i,1}$. Taking an automorphism σ_i of K over $\mathbb{Q}$ such that $\mathfrak{P}_{i,1}^{\sigma_i} = \mathfrak{P}_{1,1}$, we have $\alpha^{\sigma_i} \equiv r_i \bmod \mathfrak{P}_{1,1}$. Hence, $\sum_{i \in S^\vee} \alpha^{\sigma_i} \equiv \sum_{i \in S^\vee} r_i \equiv 0 \bmod \mathfrak{P}_{1,1}$ for infinitely many prime numbers $p \in Spl(f)$. Although automorphisms σ_i depend on p, we can choose an appropriate infinite subset of $Spl(f)$ so that automorphisms σ_i are independent of p. Hence, we have $\sum_{i \in S^\vee} \alpha^{\sigma_i} \equiv \sum_{i \in S^\vee} r_i \equiv 0 \bmod \mathfrak{P}_{1,1}$ for infinitely many primes p, which implies $\sum_{i \in S^\vee} \alpha^{\sigma_i} = 0$. Since α^{σ_i} are distinct roots of $f(x)$ by $\alpha^{\sigma_i} \equiv r_i \bmod \mathfrak{P}_{1,1}$, we complete the proof. □

Let us give some observations in the cases of $n = 3, 4, 5$. The case of $n = 6$ is discussed in the following section.

In the case of $n = 3$, Conjecture 1 and (8) give

$$Pr(f, S) = [1, 1]/2 \text{ if } \#S = 2.$$

In the case of $n = 4$, supposing that f is irreducible and indecomposable, We conjecture

$$\begin{aligned} Pr(f, \{1, 2\}) &= Pr(f, \{3, 4\})^\vee = [5, 1]/6, \\ Pr(f, \{1, 3\}) &= Pr(f, \{2, 4\})^\vee = [5, 1]/6, \\ Pr(f, \{1, 4\}) &= Pr(f, \{2, 3\}) = [1, 1]/2, \\ Pr(f, S) &= [1, 4, 1]/6 \text{ if } s := \#S = 3. \end{aligned}$$

We note

$$\binom{n}{s}^{-1} \sum_{\#S = s} Pr(f, S) = \begin{cases} [1, 1]/2! & \text{if } s = 2, \\ [1, 4, 1]/3! & \text{if } s = 3, \end{cases}$$

where $\binom{n}{s}$ is the number of subsets S with $\#S = s$. This suggests that a sequence of points $[r_i, r_j]/p$ ($i \neq j$) (resp. $[r_i, r_j, r_k]/p$ ($i \neq j, j \neq k, i \neq k$)) is uniformly distributed in $[0,1)^2$ (resp. $[0,1)^3$) (cf. (6)). Thus, the identity (10) holds. We checked the following polynomials. Let BP be an irreducible polynomial of degree 4 with coefficients equal to 0 or 1, and let α be one of its roots. For a number $\beta = \sum_{i=1}^{4} c_i\alpha^{i-1}$ with $0 \le c_i \le 2$, we take a polynomial f for which β is a root, but skip a reducible polynomial and a decomposable one. We observe the behavior of values $6[F_1, F_2] - [5, 1]$ for $S = \{1, 2\}$ for increasing X, for example. If the above conjecture is true, then it converges to $[0, 0]$ when $X \to \infty$. Defining an integer X_j by $\#Spl(f, X_j) = 1000j$, we observe values $|6F_1 - 5| + |6F_2 - 1|$ at $X = X_j$. If they are less than 0.01 for successive integers $X = X_j, \ldots, X_{j+100}$ for some j, we conclude that the above is true.

In the case of $n = 5$ we adopt the following d-adic approximation method to find a candidate for the limit. First, we take the polynomial $f = x^5 - 10x^3 + 5x^2 + 10x + 1$, which defines a unique subfield of degree 5 in a cyclotomic field $\mathbb{Q}(\exp(2\pi i/25))$, and define an integer X_j by $\#Spl(f, X_j) = 1000j$ as before. Suppose that a sequence of vectors c_m converges to a rational vector $\boldsymbol{a} = [a_1, \ldots, a_s]/b$ ($a_i, b \in \mathbb{Z}$) and let D be a finite set of integers including b. Then, for a large integer m, the error $\sum_i |dc_m[i] - r(dc_m[i])|$ is minimal at $d = b$, where $r(x)$ denotes the nearest integer to x. Noting this, to guess the limit from a sequence $\{c_m\}$ given by computer experiments, we begin by guessing a set D including the denominator b of $\boldsymbol{a}$ by some means or other. In this case, we take for D $\{d \mid 0 < d \le 500$ and a prime divisor of d is $2, 3$ or $5\}$. Second, we look for an integer $d = d_0 \in D$ that gives the minimum of errors $\sum_i |dc_m[i] - r(dc_m[i])|$ ($d \in D$). Then, d_0 is a candidate for the denominator. We checked that there is an integer j such that for successive integers $X = X_j, \ldots, X_{j+10^5}$, both the integer $d = d_0$ determined above and rounded integers of elements of $d \cdot Pr(f, S, X_i)$ are stable. In this case, the minimum error is less than 0.01 for $X = 10^{10}$ and the conjecture holds. The identity (10) holds and we have

$$\begin{aligned}
&Pr(f, \{1,2\}) = Pr(f, \{4,5\})^\vee = [137, 7]/144,\\
&Pr(f, \{1,3\}) = Pr(f, \{3,5\})^\vee = [11, 1]/12,\\
&Pr(f, \{1,4\}) = Pr(f, \{2,5\})^\vee = [17, 7]/24,\\
&Pr(f, \{1,5\}) = [1, 1]/2,\\
&Pr(f, \{2,3\}) = Pr(f, \{3,4\})^\vee = [29, 19]/48,\\
&Pr(f, \{2,4\}) = [1, 1]/2,\\
&Pr(f, \{1,2,3\}) = Pr(f, \{3,4,5\})^\vee = [71, 67, 6]/144,\\
&Pr(f, \{1,2,4\}) = Pr(f, \{2,4,5\})^\vee = [11, 12, 1]/24,\\
&Pr(f, \{1,2,5\}) = Pr(f, \{1,4,5\})^\vee = [7, 39, 2]/48,\\
&Pr(f, \{1,3,5\}) = [1, 22, 1]/24,\\
&Pr(f, \{2,3,4\}) = [1, 7, 1]/9,
\end{aligned}$$

$$\begin{aligned}
&Pr(f,\{2,3,5\}) = Pr(f,\{1,3,4\})^{\vee} = [1,17,6]/24,\\
&Pr(f,S) = [1,11,11,1]/24 \text{ if } \#S = 4.
\end{aligned}$$

We note that

$$\binom{n}{s}^{-1} \sum_{\#S=s} Pr(f,S) = \begin{cases} [1,1]/2! & \text{if } s=2,\\ [1,4,1]/3! & \text{if } s=3,\\ [1,11,11,1]/4! & \text{if } s=4. \end{cases}$$

To check other polynomials, we consider that an element a of $Pr(f,S,X)$ converges to a candidate A/B $(A,B \in \mathbb{Z})$ if we have $A = r(a \cdot B)$. By this method, we checked the above for any irreducible polynomial f of degree 5 that has a root $\sum_i c_i \alpha^{i-1}$ $(0 \leq c_i \leq 2)$, where α is a root of an irreducible polynomial with coefficients equal to 0 or 1.

4. The Case of Degree 6

In the case of $n \leq 5$, the classification of being decomposable or not is enough to consider densities. However, in the case of $n = 6$, it is not enough and indecomposable polynomials have been divided into at least four types so far First, we give some examples.

Example 1. For the indecomposable polynomial $f = f_1 = x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$, we expect

$$\begin{aligned}
&Pr(f,\{1,2\}) = Pr(f,\{5,6\})^{\vee} = [39,1]/40,\\
&Pr(f,\{1,3\}) = Pr(f,\{4,6\})^{\vee} = [14,1]/15,\\
&Pr(f,\{1,4\}) = Pr(f,\{3,6\})^{\vee} = [17,3]/20,\\
&Pr(f,\{1,5\}) = Pr(f,\{2,6\})^{\vee} = [23,7]/30,\\
&Pr(f,\{1,6\}) = [1,1]/2,\\
&Pr(f,\{2,3\}) = Pr(f,\{4,5\})^{\vee} = [19,5]/24,\\
&Pr(f,\{2,4\}) = Pr(f,\{3,5\})^{\vee} = [3,1]/4,\\
&Pr(f,\{2,5\}) = [1,1]/2,\\
&Pr(f,\{3,4\}) = [1,1]/2,
\end{aligned}$$

$$\begin{aligned}
&Pr(f,\{1,2,3\}) = Pr(f,\{4,5,6\})^{\vee} = [251,106,3]/360,\\
&Pr(f,\{1,2,4\}) = Pr(f,\{3,5,6\})^{\vee} = [67,52,1]/120,\\
&Pr(f,\{1,2,5\}) = Pr(f,\{2,5,6\})^{\vee} = [37,82,1]/120,\\
&Pr(f,\{1,2,6\}) = Pr(f,\{1,5,6\})^{\vee} = [16,73,1]/90,\\
&Pr(f,\{1,3,4\}) = Pr(f,\{3,4,6\})^{\vee} = [37,82,1]/120,
\end{aligned}$$

$$\begin{aligned}
Pr(f,\{1,3,5\}) &= Pr(f,\{2,4,6\})^{\vee} = [27,92,1]/120,\\
Pr(f,\{1,3,6\}) &= Pr(f,\{1,4,6\})^{\vee} = [13,104,3]/120,\\
Pr(f,\{1,4,5\}) &= Pr(f,\{2,3,6\})^{\vee} = [11,46,3]/60,\\
Pr(f,\{2,3,4\}) &= Pr(f,\{3,4,5\})^{\vee} = [17,50,5]/72,\\
Pr(f,\{2,3,5\}) &= Pr(f,\{2,4,5\})^{\vee} = [5,16,3]/24,
\end{aligned}$$

$$\begin{aligned}
&Pr(f,\{1,2,3,4\}) = Pr(f,\{3,4,5,6\})^{\vee} = [25,68,26,1]/120,\\
&Pr(f,\{1,2,3,5\}) = Pr(f,\{2,4,5,6\})^{\vee} = [20,73,26,1]/120,\\
&Pr(f,\{1,2,3,6\}) = Pr(f,\{1,4,5,6\})^{\vee} = [7,178,53,2]/240,\\
&Pr(f,\{1,2,4,5\}) = Pr(f,\{2,3,5,6\})^{\vee} = [10,83,26,1]/120,\\
&Pr(f,\{1,2,4,6\}) = Pr(f,\{1,3,5,6\})^{\vee} = [1,89,29,1]/120,\\
&Pr(f,\{1,2,5,6\}) = [1,59,59,1]/120,\\
&Pr(f,\{1,3,4,5\}) = Pr(f,\{2,3,4,6\})^{\vee} = [5,83,31,1]/120,\\
&Pr(f,\{1,3,4,6\}) = [1,59,59,1]/120,\\
&Pr(f,\{2,3,4,5\}) = [1,23,23,1]/48,\\
&Pr(f,S) = [1,26,66,26,1]/120 \text{ if } \#S = 5.
\end{aligned}$$

Remark 2. In this case, the identity (10) holds. To look for conjectural values, we adopt the 10-adic approximation besides the d-adic one in the previous section. That is, we observe the minimum of errors $\sum_i |c_m[i] - r(d \cdot c_m[i])/d|$ and $\sum_i |d \cdot c_m[i] - r(d \cdot c_m[i])|$ $(d \in D)$ for a sequence of vectors c_m. In this case, we take for D $\{d \mid 1 \le d \le 500$ and a prime divisor of d is $2, 3$ or $5\}$. Let p_j be the smallest prime number in $Spl(f)$ larger than $10^9 j$. To the extent of $p_j < 10^{11}$ and $j > 30$, the values of $Pr(f, S, p_j)$ support the above conjecture by this double-checking method.

We have

$$\binom{n}{s}^{-1} \sum_{\#S=s} Pr(f,S) = \begin{cases} [1,1]/2! & \text{if } s = 2,\\ [1,4,1]/3! & \text{if } s = 3,\\ [1,11,11,1]/4! & \text{if } s = 4,\\ [1,26,66,26,1]/5! & \text{if } s = 5, \end{cases}$$

and

$$Pr(f,\{1,2,5\}) = Pr(f,\{1,3,4\}), \tag{12}$$

$$Pr(f,\{2,5,6\}) = Pr(f,\{3,4,6\}), \tag{13}$$

$$Pr(f,\{2,3,4\}) + Pr(f,\{1,5,6\}) = Pr(f,\{3,4,5\}) + Pr(f,\{1,2,6\}), \tag{14}$$

$$\begin{aligned}
Pr(f,\{2,3,5\}) + Pr(f,\{1,4,6\}) &= Pr(f,\{1,3,5\}) + Pr(f,\{2,4,6\}) =\\
Pr(f,\{2,3,6\}) + Pr(f,\{1,4,5\}) &= Pr(f,\{2,4,5\}) + Pr(f,\{1,3,6\}),
\end{aligned} \tag{15}$$

$$Pr(f,\{1,2,5,6\}) = Pr(f,\{1,3,4,6\}). \tag{16}$$

$\mathbb{Q}(f) = \mathbb{Q}(\exp(2\pi i/7))$ is obvious. The sequence $[1, 59, 59, 1]$ for $S = \{1, 2, 5, 6\}$ and $\{1, 3, 4, 6\}$ is given by $T_1(4, i)$ $(i = 1, 2, 3, 4)$, where $T_1(n, k)$ $(1 \leq k \leq n)$ is defined by

$$T_1(1,1) = 1, T_1(n,k) = (4n-4k+1)T_1(n-1,k-1) + (4k-3)T_1(n-1,k),$$

and the sequence $[1, 23, 23, 1]$ for $S = \{2, 3, 4, 5\}$ is given by $T_2(4, i)\,(i = 1, 2, 3, 4)$ where $T_2(n, k)$ $(1 \leq k \leq n)$ is defined by

$$T_2(1,1) = 1, T_2(n,k) = (2n-2k+1)T_2(n-1,k-1) + (2k-1)T_2(n-1,k).$$

Example 2. For the indecomposable polynomial $f = f_2 = x^6 - 2x^5 + 11x^4 + 6x^3 + 16x^2 + 122x + 127$, we expect

$$\begin{aligned}
&Pr(f,\{1,2\}) = Pr(f,\{5,6\})^\vee = [139,5]/144,\\
&Pr(f,\{1,3\}) = Pr(f,\{4,6\})^\vee = [127,17]/144,\\
&Pr(f,\{1,4\}) = Pr(f,\{3,6\})^\vee = [7,2]/9,\\
&Pr(f,\{1,5\}) = Pr(f,\{2,6\})^\vee = [25,11]/36,\\
&Pr(f,\{1,6\}) = [1,1]/2,\\
&Pr(f,\{2,3\}) = Pr(f,\{4,5\})^\vee = [3,1]/4,\\
&Pr(f,\{2,4\}) = Pr(f,\{3,5\})^\vee = [107,37]/144,\\
&Pr(f,\{2,5\}) = [1,1]/2,\\
&Pr(f,\{3,4\}) = [1,1]/2,
\end{aligned}$$

$$\begin{aligned}
&Pr(f,\{1,2,3\}) = [49,23,0]/72, && Pr(f,\{4,5,6\}) = [0,1,3]/4, \quad (*)\\
&Pr(f,\{1,2,4\}) = [37,35,0]/72, && Pr(f,\{3,5,6\}) = [0,23,49]/72, \quad (*)\\
&Pr(f,\{1,2,5\}) = [5,13,0]/18, && Pr(f,\{2,5,6\}) = [0,89,55]/144, \quad (*)\\
&Pr(f,\{1,2,6\}) = [28,115,1]/144, && Pr(f,\{1,5,6\}) = [0,13,5]/18, \quad (*)\\
&Pr(f,\{1,3,4\}) = [5,13,0]/18, && Pr(f,\{3,4,6\}) = [0,89,55]/144, \quad (*)\\
&Pr(f,\{1,3,5\}) = [1,3,0]/4, && Pr(f,\{2,4,6\}) = Pr(f,\{1,3,5\})^\vee,\\
&Pr(f,\{1,3,6\}) = [16,121,7]/144, && Pr(f,\{1,4,6\}) = [1,115,28]/144, \quad (*)\\
&Pr(f,\{1,4,5\}) = [3,12,1]/16, && Pr(f,\{2,3,6\}) = [0,3,1]/4, \quad (*)\\
&Pr(f,\{2,3,4\}) = [35,101,8]/144, && Pr(f,\{3,4,5\}) = [0,13,5]/18, \quad (*)\\
&Pr(f,\{2,3,5\}) = [29,95,20]/144, && Pr(f,\{2,4,5\}) = [8,101,35]/144, \quad (*)
\end{aligned}$$

$$\begin{aligned}
&Pr(f,\{1,2,3,4\}) = Pr(f,\{3,4,5,6\})^\vee = [31,77,36,0]/144,\\
&Pr(f,\{1,2,3,5\}) = Pr(f,\{2,4,5,6\})^\vee = [19,89,36,0]/144,\\
&Pr(f,\{1,2,3,6\}) = Pr(f,\{1,4,5,6\})^\vee = [0,3,1,0]/4,
\end{aligned}$$

$$\begin{aligned}
&Pr(f,\{1,2,4,5\}) = Pr(f,\{2,3,5,6\})^\vee = [1,26,9,0]/36,\\
&Pr(f,\{1,2,4,6\}) = Pr(f,\{1,3,5,6\})^\vee = [0,107,37,0]/144,\\
&Pr(f,\{1,2,5,6\}) = [0,1,1,0]/2,\\
&Pr(f,\{1,3,4,5\}) = Pr(f,\{2,3,4,6\})^\vee = [0,25,11,0]/36,\\
&Pr(f,\{1,3,4,6\}) = [0,1,1,0]/2,\\
&Pr(f,\{2,3,4,5\}) = [0,1,1,0]/2,\\
&Pr(f,S) = [0,1,2,1,0]/4 \text{ if } \#S = 5.
\end{aligned}$$

Remark 3. Conjectural values are determined by the double-checking method above. In this case, the identity does not hold for lines with tag $(*)$ for $\#S = 3$. In this case, we have

$$\binom{n}{s}^{-1} \sum_{\#S=s} Pr(f,S) = \begin{cases} [1,1]/2 & \text{if } s = 2,\\ [3,13,4]/20 & \text{if } s = 3,\\ [1,19,19,1]/40 & \text{if } s = 4,\\ [0,1,2,1,0]/4 & \text{if } s = 5, \end{cases}$$

and (12), (13), and (16) hold. Putting $t(n,m) = 2A(n+1,m+1) - \binom{n}{m}$, we see that $[1,19,19,1] = [t(3,0),t(3,1),t(3,2),t(3,3)]$. The polynomial $x^2/4 + 2x^3/4 + x^4/4$ corresponding to $[0,1,2,1,0]/4$ for $\#S = 5$ above is equal to $(x/2 + x^2/2)^2$. That is, the generating polynomial of $Pr(f,S)$ is identical to the square of the generating polynomial of densities of the two-dimensional uniform distribution (cf. Section 4). This shows that a sequence of points $(r_{\sigma(1)},\ldots,r_{\sigma(5)})/p$ for $\sigma \in S_6$ is not uniformly distributed in $[0,1)^5$.

Let α be a root of f. Then, we have $\mathbb{Q}(\alpha) = \mathbb{Q}(f) = \mathbb{Q}(\exp(2\pi i/7))$, and over a quadratic subfield $M_2 = \mathbb{Q}(\sqrt{-7})$ of $\mathbb{Q}(\alpha)$, f has a divisor $g_3(x) := x^3 - x^2 + (\sqrt{-7}+5)x + 3\sqrt{-7} + 8$, for which the coefficient of x^2 is the rational number -1. This is an example of the first case of Proposition 5 below.

The densities for the polynomial $f_2(-x)$ are the same as those for the next polynomial $f_3(x)$; that is,

$$Pr(f_2(-x),S) = Pr(f_3(x),S). \tag{17}$$

Example 3. For the indecomposable polynomial $f = f_3 = x^6 - 2x^3 + 9x^2 + 6x + 2$, we expect that

$$Pr(f_3,S) = Pr(f_2,S^\vee)^\vee. \tag{18}$$

Remark 4. Conjectural values are determined by the double-checking method. Let us make a remark from a theoretical viewpoint. Because we can check that the polynomial $f_2(x)$ satisfies the assumption (A) by using Proposition 4, we have $Pr(f_2(-x),S)^\vee = Pr(f_2(x),S^\vee)$, and hence, (17) and (18) are equivalent. Polynomials $f_3(x)$ and $f_3(-x)$ have the same densities, and assumption (A) on S is not satisfied for either polynomial if $\#S = 3$ and $S \neq \{1,3,5\}, \{2,4,6\}$. Let α be a root of f. Then, $\mathbb{Q}(\alpha)$ is a splitting field of the polynomial $x^3-3x-14$, which is the composite field of $\mathbb{Q}(\sqrt{-1})$ and a field defined by $x^3-3x-14=0$. Over a quadratic subfield $M_2 = \mathbb{Q}(\sqrt{-1})$ of $\mathbb{Q}(\alpha)$, f has a divisor $g_3(x) := x^3+0\cdot x^2-3\sqrt{-1}x-\sqrt{-1}-1$, whose second leading coefficient is a rational number 0. This is also an example of the first case of Proposition 5.

Example 4. For an indecomposable polynomial $f = f_4 = x^6-9x^5-3x^4+139x^3+93x^2-627x+1289$, we expect

$$\begin{aligned}
Pr(f,\{1,2\}) &= Pr(f,\{5,6\})^{\vee} = [277,11]/288,\\
Pr(f,\{1,3\}) &= Pr(f,\{4,6\})^{\vee} = [661,59]/720,\\
Pr(f,\{1,4\}) &= Pr(f,\{3,6\})^{\vee} = [38,7]/45,\\
Pr(f,\{1,5\}) &= Pr(f,\{2,6\})^{\vee} = [559,161]/720,\\
Pr(f,\{1,6\}) &= [1,1]/2,\\
Pr(f,\{2,3\}) &= Pr(f,\{4,5\})^{\vee} = [33,7]/40,\\
Pr(f,\{2,4\}) &= Pr(f,\{3,5\})^{\vee} = [47,13]/60,\\
Pr(f,\{2,5\}) &= [1,1]/2,\\
Pr(f,\{3,4\}) &= [1,1]/2,
\end{aligned}$$

$$\begin{aligned}
Pr(f,\{1,2,3\}) &= Pr(f,\{4,5,6\})^{\vee} = [475,164,9]/648,\\
Pr(f,\{1,2,4\}) &= Pr(f,\{3,5,6\})^{\vee} = [649,416,15]/1080,\\
Pr(f,\{1,2,5\}) &= Pr(f,\{2,5,6\})^{\vee} = [314,751,15]/1080,\\
Pr(f,\{1,2,6\}) &= Pr(f,\{1,5,6\})^{\vee} = [208,857,15]/1080,\\
Pr(f,\{1,3,4\}) &= Pr(f,\{3,4,6\})^{\vee} = [314,751,15]/1080,\\
Pr(f,\{1,3,5\}) &= Pr(f,\{2,4,6\})^{\vee} = [15,56,1]/72,\\
Pr(f,\{1,3,6\}) &= Pr(f,\{1,4,6\})^{\vee} = [433,2726,81]/3240,\\
Pr(f,\{1,4,5\}) &= Pr(f,\{2,3,6\})^{\vee} = [539,2520,181]/3240,\\
Pr(f,\{2,3,4\}) &= Pr(f,\{3,4,5\})^{\vee} = [722,2375,143]/3240,\\
Pr(f,\{2,3,5\}) &= Pr(f,\{2,4,5\})^{\vee} = [639,2314,287]/3240.
\end{aligned}$$

$$\begin{aligned}
Pr(f,\{1,2,3,4\}) &= Pr(f,\{3,4,5,6\})^{\vee} = [53,175,56,4]/288,\\
Pr(f,\{1,2,3,5\}) &= Pr(f,\{2,4,5,6\})^{\vee} = [101,469,140,10]/720,\\
Pr(f,\{1,2,3,6\}) &= Pr(f,\{1,4,5,6\})^{\vee} = [17,268,70,5]/360,\\
Pr(f,\{1,2,4,5\}) &= Pr(f,\{2,3,5,6\})^{\vee} = [24,261,70,5]/360,\\
Pr(f,\{1,2,4,6\}) &= Pr(f,\{1,3,5,6\})^{\vee} = [5,277,73,5]/360,
\end{aligned}$$

$$\begin{aligned}
&Pr(f,\{1,2,5,6\}) = [1,35,35,1]/72,\\
&Pr(f,\{1,3,4,5\}) = Pr(f,\{2,3,4,6\})^{\vee} = [27,515,168,10]/720,\\
&Pr(f,\{1,3,4,6\}) = [1,35,35,1]/72,\\
&Pr(f,\{2,3,4,5\}) = [7,137,137,7]/288,\\
&Pr(f,S) = [1,14,42,14,1]/72 \text{ if } \#S = 5.
\end{aligned}$$

Remark 5. The conjectural values above were determined by the double-checking method for $p < 10^{13}$ and $D = \{d \mid d \le 4000 \text{ and a prime divisor of } d \text{ is } 2, 3 \text{ or } 5\}$. The identities (10) and (12)–(16) hold.

In this case, we expect

$$\binom{n}{s}^{-1} \sum_{\#S=s} Pr(f,S) = \begin{cases} [1,1]/2! & \text{if } s=2,\\ [1,4,1]/3! & \text{if } s=3,\\ [1,11,11,1]/4! & \text{if } s=4,\\ [1,14,42,14,1]/72 & \text{if } s=5. \end{cases}$$

Substituting $T(m,n) := (mn)! \prod_{k=0}^{n-1}(k!/(k+m)!)$ $(m,n \ge 1)$ [11], which is called a multidimensional Catalan number, we see $T(m, 6-m) = 1, 14, 42, 14, 1$ according to $m = 1, 2, 3, 4, 5$, respectively.

Let α be a root of f. Then, over a cubic subfield M_3 defined by $\beta^3-\beta^2-2\beta+1 = 0$ of $\mathbb{Q}(\alpha)$, f has a divisor $g_2(x) := x^2+(6\beta-5)x+9\beta^2-15\beta+8$, whose discriminant is the rational number -7. This is an example of the second case of Proposition 5. We have $\mathbb{Q}(\alpha) = \mathbb{Q}(\exp(2\pi i/7))$ and $Pr(f_4(-x), S) = Pr(f_4(x), S)$.

Remark 6. With respect to the polynomials on pp. 86–87 in [9], the densities defined here for the polynomials in cases (1)–(5) are equal to the ones given in Examples 2, 3, 1, 1, and 4, respectively.

We can consider a more general density. For a real function $t = t(x_1, \dots, x_n)$, we define $Pr(f,t,X) := [\dots, F_0, F_1, \dots]$ by

$$F_k := \frac{\#\{p \in Spl(f,X) \mid \lceil t(r_1/p, \dots, r_n/p) \rceil = k\}}{\#Spl(f,X)}$$

and put $Pr(f,t) := \lim_{X\to\infty} Pr(f,t,X)$.

For example, for $f = x^3 - 3x + 1$, we expect

$$Pr(f, 4x_i)_{[1..4]} = \begin{cases} [9,5,2,0]/16 & (i=1),\\ [3,5,5,3]/16 & (i=2),\\ [0,2,5,9]/16 & (i=3), \end{cases}$$

where $v_{[n..m]}$ means a subsequence $[v_n, \dots, v_m]$ for $v = [\dots, v_0, v_1, \dots]$. The identity $Pr(f, 4x_1)[4] = Pr(f, 4x_3)[1] = 0$ is not difficult to prove.

5. Arithmetic Aspects

We recall that in this paper, a polynomial is supposed to be an irreducible monic one with integer coefficients, and hereinafter, we neglect the global order (4). To analyze the case of $\deg f = 6$, we prepare the following.

Proposition 5. *Let $f(x) = x^{2m} + a_{2m-1}x^{2m-1} + \dots$ be a polynomial of even degree $2m$ and let α be a root of $f(x)$ and put $F = \mathbb{Q}(\alpha)$. Let p be a sufficiently large prime number in $Spl(f)$, and let $r_1, \dots, r_{2m} \in \mathbb{Z}$ be roots of $f(x)$ modulo p, that is,*

$$f(x) \equiv \prod_{i=1}^{2m}(x - r_i) \bmod p. \tag{19}$$

(1) *Suppose that F contains a quadratic subfield M_2 and that the coefficient of x^{m-1} of the monic minimal polynomial $g_m(x)$ of α over M_2 is a rational integer a. Then, for the decomposition $pO_{M_2} = \mathfrak{p}_1\mathfrak{p}_2$ to the product of prime ideals $\mathfrak{p}_i$ of M_2, we can renumber the roots r_i so that*

$$g_m(x) \equiv \prod_{i=1}^{m}(x - r_i) \bmod \mathfrak{p}_1, \quad g_m(x) \equiv \prod_{i=m+1}^{2m}(x - r_i) \bmod \mathfrak{p}_2, \tag{20}$$

and we have the linear relation

$$r_1 + \dots + r_m \equiv r_{m+1} + \dots + r_{2m} \equiv -a \bmod p. \tag{21}$$

Moreover, we have $f(x) = x^{2m} + 2ax^{2m-1} + \dots$.

(2) *Suppose that F contains a subfield M_m of degree m and that the discriminant of the monic minimal quadratic polynomial $g_2(x)$ of α over M_m is a rational integer D. Then, we can renumber the roots r_i so that we have*

$$g_2(x) \equiv (x - r_i)(x - r_{i+m}) \bmod \mathfrak{p}_i \quad (i = 1, \dots, m) \tag{22}$$

for the decomposition $pO_{M_m} = \mathfrak{p}_1 \dots \mathfrak{p}_m$ to the product of prime ideals, and we have the quadratic relation

$$(r_i - r_{i+m})^2 \equiv D \bmod p \quad (i = 1, \dots, m). \tag{23}$$

Moreover, F contains a quadratic field $\mathbb{Q}(\sqrt{D})$.

Proof. We number the roots r_i of $f(x) \equiv 0 \bmod p$ and prime ideals $\mathfrak{P}_i$ of F lying above p so that $\alpha \equiv r_i \bmod \mathfrak{P}_i$. Let us prove case (1) above. First, we note that the degree of $g_m(x) \in O_{M_2}[x]$ is m. We may assume that $pO_{M_2} = \mathfrak{p}_1\mathfrak{p}_2$ and $\mathfrak{p}_1 O_F = \mathfrak{P}_1 \dots \mathfrak{P}_m$ and $\mathfrak{p}_2 O_F = \mathfrak{P}_{m+1} \dots \mathfrak{P}_{2m}$, which imply $\mathfrak{P}_i \cap M_2 = \mathfrak{p}_1$ $(i = 1, \dots, m)$ and

$\mathfrak{P}_i \cap M_2 = \mathfrak{p}_2 \, (i = m+1, \ldots, 2m)$. The assumptions $g_m(\alpha) = 0$ and $\alpha \equiv r_i \bmod \mathfrak{P}_i$ imply $g_m(r_i) \equiv 0 \bmod \mathfrak{P}_i$; hence,

$$g_m(r_i) \in \mathfrak{P}_i \cap M_2 = \begin{cases} \mathfrak{p}_1 & (i = 1, \ldots, m), \\ \mathfrak{p}_2 & (i = m+1, \ldots, 2m), \end{cases}$$

which concludes (20). Therefore, the definition of a implies $a + \sum_{i=1}^{m} r_i \in \mathfrak{p}_1 \cap \mathbb{Z} = p\mathbb{Z}$ and $a + \sum_{i=m+1}^{2m} r_i \in \mathfrak{p}_2 \cap \mathbb{Z} = p\mathbb{Z}$ since $a, r_i \in \mathbb{Z}$; hence, we get (21). Equations $a_{2m-1} + \sum_{i=1}^{2m} r_i \equiv 0 \bmod p$ and (21) imply $a_{2m-1} \equiv 2a \bmod p$; hence, $a_{2m-1} = 2a$, since p is sufficiently large. Thus, we have $f(x) = x^{2m} + 2ax^{2m-1} + \ldots$.

Next, let us prove case (2) above. Let $g_2(x) = x^2 + Ax + B \, (A, B \in O_{M_m})$. The assumption $g_2(\alpha) = 0$ implies $g_2(r_i) \equiv 0 \bmod \mathfrak{P}_i$, that is, $g_2(r_i) \in \mathfrak{P}_i$ $(i = 1, \ldots, 2m)$. By renumbering, we may assume that

$$pO_{M_m} = \mathfrak{p}_1 \ldots \mathfrak{p}_m, \quad \mathfrak{p}_i O_F = \mathfrak{P}_i \mathfrak{P}_{i+m} \quad (i = 1, \ldots, m).$$

Then we have

$$g_2(r_i) \in \mathfrak{P}_i \cap M_m = \mathfrak{p}_i, \quad g_2(r_{i+m}) \in \mathfrak{P}_{i+m} \cap M_m = \mathfrak{p}_i \quad (i = 1, \ldots, m),$$

that is, (22). Therefore, we have

$$D \equiv (r_i + r_{i+m})^2 - 4r_i r_{i+m} \equiv (r_i - r_{i+m})^2 \bmod \mathfrak{p}_i \quad (i = 1, \ldots, m).$$

Since D and r_i are rational integers and $\mathfrak{p}_i \cap \mathbb{Z} = p\mathbb{Z}$, we have (23). Since the difference $\sqrt{D}$ of α and its conjugate over M_m is in F and D is a rational integer, F contains a quadratic field $\mathbb{Q}(\sqrt{D})$. □

A sufficient condition for the assumption in (1) is as follows.

Proposition 6. *If $f(x) = g(h(x))$ holds for a polynomial $g(x)$ of degree 2 and a polynomial $h(x)$ of degree $m \, (> 1)$, then the assumption in* (1) *of Proposition 5 is satisfied.*

Proof. Let α be a root of $f(x)$. Substituting $\beta := h(\alpha)$ and $M_2 := \mathbb{Q}(\beta)$, we have $g(\beta) = f(\alpha) = 0$; hence, M_2 is a quadratic field and $g(x)$ is $(x - \beta)(x - \overline{\beta})$ for a conjugate $\overline{\beta} \in M_2$ of β over $\mathbb{Q}$. Then, $g_m(x) := h(x) - \beta$ satisfies $f(x) = g_m(x)(h(x) - \overline{\beta})$, $g_m(\alpha) = 0$, and the second leading coefficient of $g_m(x)$, which is equal to that of $h(x)$, is rational. If $g_m(x)$ is reducible over M_2, there is a decomposition $g_m(x) = k_1(x)k_2(x)$ with $k_i(x) \in M_2[x]$ and $\deg k_i > 1$. Thus, $f(x) = (h(x) - \beta)(h(x) - \overline{\beta}) = g_m(x)\overline{g_m(x)}$ is divisible by a polynomial $k_i(x)\overline{k_i(x)} \in \mathbb{Q}[x]$, which contradicts the irreducibility of $f(x)$. □

Let us make a few comments on the relationship between (20) and the densities considered above.

Lemma 2. *Keep the case* (1) *in Proposition 5 and assume that* $0 \leq r_i < p$ $(1 \leq i \leq 2m)$. *Substituting* $C_p(g, \mathfrak{p}_1) := (a + \sum_{i=1}^{m} r_i)/p \in \mathbb{Z}$, $C_p(g, \mathfrak{p}_2) := (a + \sum_{i=m+1}^{2m} r_i)/p \in \mathbb{Z}$, *we have* $C_p(f) = C_p(g, \mathfrak{p}_1) + C_p(g, \mathfrak{p}_2)$ *and*

$$C_p(g, \mathfrak{p}_1) = \lceil (r_1 + \cdots + r_{m-1})/p \rceil, \; C_p(g, \mathfrak{p}_2) = \lceil (r_{m+1} + \cdots + r_{2m-1})/p \rceil \tag{24}$$

except for finitely many primes p.

Proof. The definition (2) of $C_p(f)$ implies $C_p(f)p = 2a + \sum_{i=1}^{2m} r_i = (a + \sum_{i=1}^{m} r_i) + (a + \sum_{i=m+1}^{2m} r_i)$, i.e., $C_p(f) = C_p(g, \mathfrak{p}_1) + C_p(g, \mathfrak{p}_2)$. Substituting $k = \lceil (r_1 + \cdots + r_{m-1})/p \rceil$, we have $(r_1 + \cdots + r_{m-1})/p \leq k < (r_1 + \cdots + r_{m-1})/p + 1$. Hence, $C_p(g, \mathfrak{p}_1) - (r_m + a)/p \leq k < C_p(g, \mathfrak{p}_1) - (r_m + a)/p + 1$, and so

$$-(r_m + a)/p \leq k - C_p(g, \mathfrak{p}_1) < -(r_m + a)/p + 1.$$

If $k - C_p(g, \mathfrak{p}_1) \leq -1$ holds, then we have $-(r_m + a)/p \leq -1$; hence, $1 \leq p - r_m \leq a$. If this inequality holds for infinitely many primes, there is an integer r between 1 and a such that $p - r_m = r$ for infinitely many primes, which implies $f(-r) \equiv f(r_m) \equiv 0 \bmod p$, hence the contradiction $f(-r) = 0$. Thus, $k - C_p(g, \mathfrak{p}_1) \geq 0$ holds. Next, suppose that $k - C_p(g, \mathfrak{p}_1) \geq 1$ holds for infinitely many primes. Then, we have $(r_m + a)/p < 0$, and hence, $0 \leq r_m < -a$ for infinitely many primes, which is also a contradiction similar to the above. Hence, we have $k - C_p(g, \mathfrak{p}_1) = 0$. Another equality is similarly proved. □

Keeping and continuing the above, (8) implies

$$\lim_{X \to \infty} F_k(f, S, X) = \lim_{X \to \infty} \frac{\#\{p \in Spl(f, X) \mid C_p(g, \mathfrak{p}_1) + C_p(g, \mathfrak{p}_2) = k\}}{\#Spl(f, X)}$$

for any subset S with $\#S = n - 1$. We note that there are $2(m!)^2$ ways of choosing points $(r_1/p, \ldots, r_{m-1}/p, r_{m+1}/p, \ldots, r_{2m-1}/p) \in [0, 1)^{2(m-1)}$ by two ways for $\mathfrak{p}_1, \mathfrak{p}_2$ and $m!$ ways of choosing $r_1, \ldots, r_{m-1}$ (resp. $r_{m+1}, \ldots, r_{2m-1}$) from $r_1, \ldots, r_m$ (resp. $r_{m+1}, \ldots, r_{2m}$). *If, therefore, a sequence of all* $2(m!)^2$ *points* $(r_1/p, \ldots, r_{m-1}/p, r_{m+1}/p, \ldots, r_{2m-1}/p) \in [0, 1)^{2(m-1)}$ *for every prime* $p \in Spl(f)$ *distributes uniformly when* $p \to \infty$, then by (24) we have

$$\begin{aligned}
&\lim_{X \to \infty} F_k(f, S, X) \\
&= vol(\{(x_1, \ldots, x_{2(m-1)}) \in [0, 1)^{2(m-1)} \mid \lceil \sum_{l=1}^{m-1} x_l \rceil + \lceil \sum_{l=m}^{2(m-1)} x_l \rceil = k\}) \\
&= \sum_{i+j=k} vol\{(x_1, \ldots, x_{m-1}) \in [0, 1)^{m-1} \mid \lceil \sum_{l=1}^{m-1} x_l \rceil = i\} \times \\
&\qquad\qquad vol\{(x_m, \ldots, x_{2(m-1)}) \in [0, 1)^{m-1} \mid \lceil \sum_{l=m}^{2(m-1)} x_l \rceil = j\}.
\end{aligned}$$

The volume $vol\{(x_1,\dots,x_{m-1})\in[0,1)^{m-1}\mid\lceil\sum_{l=1}^{m-1}x_l\rceil=i\}$ is given by Eulerian numbers as above. In the case of $m=3$ for now, by

$$vol\{(x_1,x_2)\in[0,1)^2\mid\lceil x_1+x_2\rceil=i\}=\begin{cases}0 & \text{if} \quad i\le 0,\\ 1/2 & \text{if} \quad i=1,2,\\ 0 & \text{if} \quad i\ge 2,\end{cases}$$

we have

$$\lim_{X\to\infty}F_k(f,S,X)=\begin{cases}1/4 & \text{if } k=2,4,\\ 1/2 & \text{if } k=3,\\ 0 & \text{otherwise.}\end{cases}$$

This elucidates $Pr(f,S)=[0,1,2,1,0]/4$ at $\#S=5$ in the cases of Examples 2 and 3.

In the case of $\deg f=4$, the assumption in (1) of Proposition 5 and that of being decomposable are equivalent as follows.

Proposition 7. *Let $M_2=\mathbb{Q}(\sqrt{D})$ $(D\in\mathbb{Q})$ be a quadratic field and $f(x)\in\mathbb{Q}[x]$ be a polynomial of degree $\ell+2$. Suppose that $f(x)=g(x)h(x)$ with $g(x)=x^2+ax+b_1+b_2\sqrt{d}, h(x)\in M_2[x]$ with $a,b_1,b_2\in\mathbb{Q}$. Then, $\ell=2$ and $f(x)$ is equal to $(x^2+ax+b_1)^2-b_2^2\,d$, in particular, decomposable, which implies that in* (1) *of Proposition 5, the polynomial f is decomposable if* $\deg f=4$.

Proof. We note that the irreducibility of a polynomial f implies $b_2\neq 0$. Write $h(x)=x^\ell+h_1(x)+\sqrt{d}\,h_2(x)$ $(h_1,h_2\in\mathbb{Q}[x])$; then, we have

$$\begin{aligned}f(x)&=(x^2+ax+b_1+b_2\sqrt{d})(x^\ell+h_1(x)+\sqrt{d}\,h_2(x))\\&=(x^2+ax+b_1)(x^\ell+h_1(x))+b_2dh_2(x)\\&+\sqrt{d}[b_2(x^\ell+h_1(x))+(x^2+ax+b_1)h_2(x)]\in\mathbb{Q}[x].\end{aligned}$$

Thus, we have $b_2(x^\ell+h_1(x))+(x^2+ax+b_1)h_2(x)=0$, and hence, $x^\ell+h_1(x)=-b_2^{-1}(x^2+ax+b_1)h_2(x)$, which implies $h(x)=-b_2^{-1}(x^2+ax+b_1-b_2\sqrt{d})h_2(x)$. Thus, we have $f(x)=-b_2^{-1}((x^2+ax+b_1)^2-b_2^2\,d)h_2(x)$. Since $f(x)$ is irreducible and monic, we have $f(x)=(x^2+ax+b_1)^2-b_2^2\,d$. □

References

[1] W. Duke, J.B. Friedlander, and H. Iwaniec, Equidistribution of roots of a quadratic congruence to prime moduli, *Ann. of Math.* **141** (1995), 423-441.

[2] W. Feller, An Introduction to Probability Theory and Its Applications, Vol. 2, J. Wiley, New York, 1966.

[3] T. Hadano, Y. Kitaoka, T. Kubota, and M. Nozaki, Densities of sets of primes related to decimal expansion of rational numbers, in *Number Theory: Tradition and Modernization*, 67-80, Springer Science + Business Media, Inc, 2006.

[4] Y. Kitaoka, A statistical relation of roots of a polynomial in different local fields, *Math. of Comp.* **78** (2009), 523-536.

[5] Y. Kitaoka, A statistical relation of roots of a polynomial in different local fields II, in Number Theory : Dreaming in Dreams (Series on Number Theory and Its Application, Vol. 6), 106-126, World Scientific, 2010.

[6] Y. Kitaoka, A statistical relation of roots of a polynomial in different local fields III, *Osaka J. Math.* **49** (2012), 393-420.

[7] Y. Kitaoka, Introduction to Algebra (in Japanese), Kin-en-Shobo, 2012.

[8] Y. Kitaoka, A statistical relation of roots of a polynomial in different local fields IV, *Uniform Distribution Theory* **8** (2013), no.1, 17-30.

[9] Y. Kitaoka, Statistical distribution of roots of a polynomial modulo prime powers, in Number Theory: Plowing and Starring through High Wave Forms (Series on Number Theory and Its Application, Vol. 11), 75-94, World Scientific, 2013.

[10] Y. Kitaoka, Statistical distribution of roots modulo primes of an irreducible and decomposable polynomial of degree 4, *Uniform Distribution Theory* **10** (2015), no.2, 1-10.

[11] The On-Line Encyclopedia of Integer Sequences, `https://oeis.org/`.

[12] J.P.Serre, Quelques applications du théorème de densité de Chebotarev, *I.H.E.S.* **54** (1981), 323-401.

[13] Á. Tóth, Roots of quadratic congruences, *Internat. Math. Res. Notices* **14** (2000), 719-739.