university of groningen

faculty of science
and engineering

mathematics and applied
mathematics

The number of roots of a random p-adic polynomial

Bachelor's Project Mathematics

April 2025

Student: A. Batiray

Daily supervisor: Dr. M.W. Lüdtke

First assessor: Prof. Dr. J.S. Müller

Second assessor: Prof. Dr. O. Lorscheid

Abstract

We determine the probability that a random polynomial of degree n over some complete discrete valuation ring $\mathcal{O}$ with a finite residue field $\mathbb{F}_q$ has exactly r roots in its field of fractions K.

Acknowledgements

I would like to express my sincere gratitude to my assessors, Steffen Müller, and Oliver Lorscheid for supervising this project. My deepest thanks go to my daily supervisor, Martin Lüdtke, who was always available to help me whenever I needed support. His constant assistance played a crucial role in the completion of this work.

I am also incredibly grateful to my family for their unwavering support and for financing my studies, making it possible for me to pursue my education. Their hard work and dedication mean everything to me.

A heartfelt thank you to my friends, Jerzy Olszewski and Kis Levente, for their encouragement and companionship throughout this journey. Finally, I want to express my appreciation to my girlfriend, Havva Malyalızade, for her continuous support during this time.

Contents

1 Introduction

Suppose that $\mathcal{O}$ is a complete discrete valuation ring with a finite residue class field $\mathbb{F}_q$ where q denotes the order of the field. Let K be the field of fractions of $\mathcal{O}$. Fix $n \in \mathbb{Z}_{\geq 0}$. Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ be a random polynomial of degree $\leq n$ having coefficients $a_0, a_1, \ldots, a_n \in \mathcal{O}$. In this paper our aim is to determine the probability that f has exactly r roots in K. We will normalize the additive Haar measure μ on the set of coefficients $\mathcal{O}^{n+1}$ such that $\mu(\mathcal{O}^{n+1}) = 1$, and determine the density $\mu(S_r)$ of the set S_r of degree n polynomials in $\mathcal{O}[x]$ having exactly r roots in K.

The result was first established by Manjul Bhargava, John Cremona, Tom Fisher, and Stevan Gajović for polynomials over the p-adic numbers in their paper [2], and in this paper, we aim to generalize their proof for more general coefficient rings.

Similar to their paper, let us formally define the probabilities, expectations and generating functions required to state our main results. For $0 \leq r \leq n$, let $\rho^*(n,r) := \rho^*(n,r;q)$ denote the density of polynomials of degree n over $\mathcal{O}$ having exactly r roots in K. For $0 \leq d \leq n$, set

$$\rho(n,d) = \sum_{r=0}^{n} \binom{r}{d} \rho^*(n,r). \tag{1}$$

Hence $\rho(n,d)$ is the expected number of sets of size d of K-roots.

Remark. *We will use this notion a lot so we will call sets of size d, d-sets.*

For a fixed n, determining $\rho(n,d)$ for all d is equivalent to determining $\rho^*(n,r)$ for all r, via the binomial inversion formula, [6, Page 192, Trick 3]

$$\rho^*(n,r) = \sum_{d=0}^{n} (-1)^{d-r} \binom{d}{r} \rho(n,d). \tag{2}$$

Equations (1) and (2) are equivalent to the standard observation that a probability distribution is determined by its moments; the formulation in d-sets equivalently in terms of factorial moments is most convenient for our purposes.

Similarly let $\alpha(n,d)$ denote the expected number of d-sets of K-roots of monic polynomials of degree n over $\mathcal{O}$, and let $\beta(n,d)$ denote the expected number of d-sets of K-roots of monic polynomials of degree n over $\mathcal{O}$ that reduce to x^n modulo π where $\pi \in \mathcal{O}$ is a generator of the unique maximal ideal of $\mathcal{O}$. Define the generating functions:

$$\mathcal{A}_d(t) = (1-t) \sum_{n=0}^{\infty} \alpha(n,d) t^n,$$

$$\mathcal{B}_d(t) = (1-t) \sum_{n=0}^{\infty} \beta(n,d) t^n,$$

$$\mathcal{R}_d(t) = (1-t)(1-qt) \sum_{n=0}^{\infty} (q^n + q^{n-1} + \cdots + 1) \rho(n,d) t^n.$$

Then we prove the following theorem.

Theorem 1.1. *Let n, d be any integers such that $0 \leq d \leq n$. Then:*

a) For fixed n and d, the expectations $\alpha(n,d;q), \beta(n,d;q)$ and $\rho(n,d;q)$ are rational functions of q which depend only on n, d, q and satisfy identities:

$$\rho(n,d;q) = \rho(n,d;q^{-1}); \tag{3}$$

$$\alpha(n,d;q) = \beta(n,d;q^{-1}). \tag{4}$$

b) We have the following power series identities in two variables t and u:

$$\sum_{d=0}^{\infty} \mathcal{A}_d(qt)u^d = \left(\sum_{d=0}^{\infty} \mathcal{B}_d(t)u^d\right)^q, \tag{5}$$

$$\sum_{d=0}^{\infty} \mathcal{R}_d(t)u^d = \left(\sum_{d=0}^{\infty} \mathcal{A}_d(qt)u^d\right)\left(\sum_{d=0}^{\infty} \mathcal{B}_d(t)u^d\right) = \left(\sum_{d=0}^{\infty} \mathcal{B}_d(t)u^d\right)^{q+1}, \tag{6}$$

$$\mathcal{B}_d(t) - t\mathcal{B}_d(\frac{t}{q}) = \Phi(\mathcal{A}_d(t) - t\mathcal{A}_d(qt)), \tag{7}$$

where Φ is the operator on power series that multiplies the coefficient of t^n by $q^{-\binom{n}{2}}$.

1.1 Examples

Here we work on an example to show how one can use Theorem 1.1, explicitly calculating $\alpha(n,1)$ for all n. Now we use (5) to get a recursive relation between $\mathcal{A}_d$ and $\mathcal{B}_d$, noting that $\mathcal{A}_0(t) = \mathcal{B}_0(t) = 1$ we expand the sums on both sides.

$$\begin{aligned} 1 + \mathcal{A}_1(qt)u + \cdots &= (1 + \mathcal{B}_1(t)u + \cdots)^q \\ &= 1 + q\mathcal{B}_1(t)u + \cdots \end{aligned}$$

Hence by comparing coefficients wee see that

$$\mathcal{A}_1(qt) = q\mathcal{B}_1(t). \tag{8}$$

Now we use (7),

$$\mathcal{B}_1(t) - t\mathcal{B}_1(\frac{t}{q}) = \Phi(\mathcal{A}_1(t) - t\mathcal{A}_1(qt)). \tag{9}$$

Using equality (8) we set $\mathcal{B}_1(t) = \frac{1}{q}\mathcal{A}_1(qt)$ in (9). Then we have

$$\frac{1}{q}\mathcal{A}_1(qt) - \Phi(\mathcal{A}_1(t)) = \frac{1}{q}t\mathcal{A}_1(t) - \Phi(t\mathcal{A}_1(qt)).$$

Now suppose that a_n is the n'th coefficient of $\mathcal{A}_1(t)$. The equality above with the operator tells us

$$(q^{n-1} - q^{-\binom{n}{2}})a_n = (\frac{1}{q} - q^{n-1-\binom{n}{2}})a_{n-1}.$$

Noting that $a_1 = 1$, we can solve for $a_n, n \geq 1$. Setting $n = 2$, we see that

$$a_2 = \frac{-1}{q+1}.$$

Now setting $n = 3$ we observe

$$(q^2 - q^{-3})a_3 = (\frac{1}{q} - q^{-1})a_2,$$

and hence $a_n = 0$ for $n \geq 3$. Therefore

$$\mathcal{A}_1(t) = t - \frac{1}{q+1}t^2. \tag{10}$$

From how we defined $\mathcal{A}_1(t)$ before, we have

$$\begin{align} A_1(t) &= (1-t)(\alpha(0,1) + \alpha(1,1)t + \alpha(2,1)t^2 + \cdots) \tag{11} \\ &= \alpha(0,1) + (\alpha(1,1) - \alpha(0,1))t + \cdots . \tag{12} \end{align}$$

By comparing coefficients of (10) and (12) we see that

$$\alpha(n,1) = \begin{cases} 0 & n = 0, \\ 1 & n = 1, \\ \frac{q}{q+1} & n \geq 2. \end{cases}$$

2 Valuation Theory

We refer to the entire section in Neukirch's Algebraic Number Theory [7, 3].

2.1 Valuation

Definition 2.1 (Valuation). *A* ***valuation*** *of a field K is a function*

$$|\quad| : K \to \mathbb{R}$$

which satisfies the following properties;

- $|x| \geq 0$, *and* $|x| = 0 \Leftrightarrow x = 0$
- $|xy| = |x||y|$
- $|x + y| \leq |x| + |y|$ *"Triangle Inequality"*

Remark. *There is a trivial valuation on every field K where $|x| = 1$ for all $x \neq 0$, and $|0| = 0$. In our paper, we will always exclude this case.*

A valuation of a field K induces a metric on K. This metric is given by

$$d(x, y) = |x - y|.$$

Proof. Let $x, y, z \in K$

- Note that $\alpha = x - y \in K$. Therefore $d(x, y) = |x - y| = |\alpha| \geq 0$. Moreover $|\alpha| = |x - y| = 0$ if and only if $\alpha = 0$, which implies $x = y$.

- Note that $d(x,y) = |x-y| = |-1||y-x| = |y-x| = d(y,x)$. Let us also prove $|1| = |-1| = 1$ in real numbers.
  - $|1| = |1 \cdot 1| = |1||1|$ which implies $|1| = 1$.
  - Now note that $1 = |1| = |(-1)\cdot(-1)| = |(-1)||(-1)|$ which implies $|-1| = 1$ or $|-1| = -1$, since latter cannot be true, $|-1| = 1$.
- $d(x,z) = |x-z| = |(x-y)+(y-z)| \leq |x-y| + |y-z| = d(x,y) + d(y,z)$.

□

Therefore a field K with a valuation is a **Hausdorff topological space** with metric d, detailed information is in sections 2 and 4 of Sutherland's Introduction to Metric and Topological Spaces [8].

Definition 2.2. *The valuation $|\ |$ is called **nonarchimedean** if $|n|$ stays bounded for all $n \in \mathbb{N}$. Otherwise it is called archimedean.*

Remark. *Here* $n = n \cdot 1_K = \underbrace{1+1+\cdots+1}_{n\ times}$.

Proposition 2.3. *The valuation $|\ |$ is nonarchimedean if and only if it satisfies the strong triangle inequality*

$$|x+y| \leq \max\{|x|,|y|\}.$$

Proof. Suppose that the strong triangle inequality holds, then

$$|n| = |1+\cdots+1| \leq 1.$$

Conversely, suppose that $|n| \leq N$ for all $n \in \mathbb{N}$. Let $x, y \in K$ and suppose $|x| \geq |y|$. Then $|x|^v|y|^{n-v} \leq |x|^n$ for $v \geq 0$ and we get

$$|x+y|^n \leq \sum_{v=0}^{n} |\binom{n}{v}||x|^v|y|^{n-v} \leq N(n+1)|x|^n.$$

Therefore

$$|x+y| \leq N^{1/n}(1+n)|x| = N^{1/n}(1+n)^{1/n}\max\{|x|,|y|\}.$$

If we let $n \to \infty$ we get the strong triangle inequality. □

Remark. *The strong triangle inequality implies that*

$$|x| \neq |y| \Rightarrow |x+y| = \max\{|x|,|y|\}.$$

2.2 Exponential Valuation

Definition 2.4 (Exponential Valuation)**.** *An **exponential valuation** of a field K is a function*

$$v : K \to \mathbb{R} \cup \{\infty\}$$

which satisfies the properties:

- $v(x) = \infty \Leftrightarrow x = 0$
- $v(xy) = v(x) + v(y)$
- $v(x + y) \geq \min\{v(x), v(y)\}$

where one uses the convention that for $a \in \mathbb{R}$, $a < \infty, a + \infty = \infty, \infty + \infty = \infty$.

Again we will be excluding the trivial case where $v(x) = 0$ for $x \neq 0$, $v(0) = \infty$ throughout our paper.

Remark. *If $| \ |$ is a valuation of a field K which is nonarchimedean then $| \ |$ induces an exponential valuation v by setting*

$$v(x) = -\log |x|$$

for $x \neq 0$, and $v(0) = \infty$.

In a similar manner if v is an exponential valuation of a field K, v induces a valuation of K by setting

$$|x| = q^{-v(x)}$$

*for some $q \in \mathbb{R}_{>1}$. To distinguish it from v, we call $| \ |$ an associated **multiplicative valuation** or **absolute value**.*

2.3 Valuation Ring

Properties of an exponential valuation v of a field K give rise to a subset of K with some algebraic properties.

Proposition 2.5. *The subset*

$$\mathcal{O} = \{x \in K \mid v(x) \geq 0\} = \{x \in K \mid |x| \leq 1\}$$

is a ring with group of units

$$\mathcal{O}^* = \{x \in K \mid v(x) = 0\} = \{x \in K \mid |x| = 1\}$$

and the unique maximal ideal

$$\mathcal{P} = \{x \in K \mid v(x) > 0\} = \{x \in K \mid |x| < 1\}.$$

$\mathcal{O}$ is an integral domain with field of fractions K and has the property that, for every nonzero $x \in K$, $x \in \mathcal{O}$ or $x^{-1} \in \mathcal{O}$. Such a ring is called a **valuation ring**. Its only maximal ideal is $\mathcal{P} = \{x \in \mathcal{O} \mid x^{-1} \notin \mathcal{O}\}$. The field $\mathcal{O}/\mathcal{P}$ is called **the residue class field** of $\mathcal{O}$.

An exponential valuation v is called **discrete** if it admits a smallest positive value s, which in case one finds

$$v(K^*) = s\mathbb{Z}.$$

It is called **normalized** if $s = 1$. Dividing by s we may always pass to a normalized valuation without changing the invariants $\mathcal{O}, \mathcal{O}^*, \mathcal{P}$. Having done so, an element

$$\pi \in \mathcal{O} \ \textit{ such that } \ v(\pi) = 1$$

is a **prime element**, and every element $x \in K^*$ admits a unique representation

$$x = u\pi^m$$

with $m \in \mathbb{Z}$ and $u \in \mathcal{O}^*$, for if $v(x) = m$, then $v(x\pi^{-m}) = 0$, hence $u = x\pi^{-m} \in \mathcal{O}^*$.

We have a proposition which arises from discrete valuations, but we need to state a definition.

Definition 2.6 (Discrete Valuation Ring)**.** *A principal ideal domain is called a **discrete valuation ring** if it has a unique maximal ideal.*

Proposition 2.7. *If v is a discrete valuation of K, then*

$$\mathcal{O} = \{x \in K \mid v(x) \geq 0\}$$

is a principal ideal domain and hence a discrete valuation ring. Suppose v is normalised, then the nonzero ideals of $\mathcal{O}$ are given by

$$\mathcal{P}^n = \pi^n\mathcal{O} = \{x \in K \mid v(x) \geq n\}, \ \ n \geq 0$$

where π is a prime element, i.e. $v(\pi) = 1$. One has

$$\mathcal{P}^n/\mathcal{P}^{n+1} \cong \mathcal{O}/\mathcal{P}.$$

Proof. Let $I \neq \{0\}$ be an ideal of $\mathcal{O}$ and $x \neq 0$ an element in I with smallest possible value $v(x) = n$. Then $x = u\pi^n$, $u \in \mathcal{O}^*$, so that $\pi\mathcal{O} \subseteq I$. If $y = \epsilon\pi^m \in I$ is arbitrary with $\epsilon \in \mathcal{O}^*$, then $m = v(y) \geq n$, hence $y = (\epsilon\pi^{m-n})\pi^n \in \pi^n\mathcal{O}$. Moreover $a\pi^n \mapsto a \mod \mathcal{P}$ is a surjective map from $\mathcal{P}^n$ to $\mathcal{O}/\mathcal{P}$ with kernel $\mathcal{P}^{n+1}$. □

2.4 Bases

In a discretely valued field K, the chain

$$\mathcal{O} \supseteq \mathcal{P} \supseteq \mathcal{P}^2 \supseteq \mathcal{P}^3 \supseteq \cdots$$

consisting of ideals of the valuation ring $\mathcal{O}$ forms a basis of neighbourhoods of the zero element. If v is a normalized exponential valuation and $|\ \ | = q^{-v}$ $(q > 1)$ an absolute value, then

$$\mathcal{P}^n = \{x \in K \mid |x| < \frac{1}{q^{n-1}}\}.$$

As a basis of neighbourhoods of the element $1 \in K^*$, we obtain in a similar manner the descending chain

$$\mathcal{O}^* = U^{(0)} \supseteq U^{(1)} \supseteq U^{(2)} \supseteq \cdots$$

of subgroups

$$U^{(n)} = 1 + \mathcal{P}^n = \{x \in K^* \mid |1 - x| < \frac{1}{q^{n-1}}\}, \ \ n > 0,$$

of $\mathcal{O}^*$. $U^{(n)}$ is called the n-th **higher unit group** and $U^{(1)}$ the group of **principal units**.

Proposition 2.8.

$$\mathcal{O}^*/U^{(n)} \cong (\mathcal{O}/\mathcal{P}^n)^*$$

and

$$U^{(n)}/U^{(n+1)} \cong \mathcal{O}/\mathcal{P}$$

for $n \geq 1$.

Proof. The proof is given under Proposition 3.10 in Neukirch's book[7, 3]. □

3 Completeness

Completeness or completion of a metric space is usually the next best question to ask after finding a metric on the given set.

Definition 3.1 (Cauchy Sequence)**.** *Let $(K, |\ \ |)$ be a valued field. A sequence $\{a_n\}_{n \in \mathbb{N}}$ in K is called a* ***Cauchy sequence*** *if for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that*

$$|a_n - a_m| < \epsilon \quad \textit{for all} \quad n, m \geq N.$$

Definition 3.2. *A valued field* $(K, |\ |)$ *is called* ***complete*** *if every Cauchy sequence* $\{a_n\}_{n\in\mathbb{N}}$ *in* K *converges to an element* $a \in K$*, i.e.,*

$$\lim_{n\to\infty} |a_n - a| = 0.$$

Lemma 3.3 (Hensel's Lemma)**.** *Let* K *be a field which is complete with respect to a nonarchimedean valuation* $|\ |$*. Let* $\mathcal{O}$ *be the corresponding valuation ring with maximal ideal* $\mathcal{P}$ *and residue class field* $k = \mathcal{O}/\mathcal{P}$*. A polynomial* $f \in \mathcal{O}[x]$ *is called* ***primitive*** *if* $f(x) \not\equiv 0 \mod \mathcal{P}$*.*

Now suppose that a primitive polynomial $f \in \mathcal{O}[x]$ *admits a modulo* $\mathcal{P}$ *factorization*

$$f(x) \equiv \overline{g}(x)\overline{h}(x) \bmod \mathcal{P}$$

into relatively prime polynomials $\overline{g}, \overline{h} \in k[x]$*, then* $f(x)$ *admits a factorization into polynomials* $g, h \in \mathcal{O}[x]$ *such that* $\deg(g) = \deg(\overline{g})$ *and*

$$g(x) \equiv \overline{g}(x) \bmod \mathcal{P} \quad \text{and} \quad h(x) \equiv \overline{h}(x) \bmod \mathcal{P}.$$

Proof. A detailed proof is given at [7, Lemma 4.6]. □

Remark. *In general for any valued field, one can get a complete valued field by the process of* ***completion****, see [7, Section 4].*

4 Local Fields

Definition 4.1 (Local Fields)**.** *All fields which are complete with respect to a discrete valuation and have a finite residue class field are called (nonarchimedean)* ***local fields.*** *For such a local field, the normalized exponential valuation is denoted by* v_ρ*, and* $|\ |_\rho$ *denotes the absolute value normalized by*

$$|x|_\rho = q^{-v_\rho(x)}$$

where q *is the cardinality of the residue class field.*

Definition 4.2 (Locally Compact)**.** *A topological space* X *is* ***locally compact*** *if every point has a neighborhood which is itself contained in a compact set.*

Proposition 4.3. *A local field* K *is locally compact. Its valuation ring* R *is compact.*

Proof. For a detailed proof, check Proposition 5.1 in Neukirch's book [7, 5]. □

Proposition 4.4. *Let* K *be a local field with discrete valuation ring* $\mathcal{O}$*. Let* $R \subseteq \mathcal{O}$ *be a system of representatives for the residue class field* $\mathbb{F}_q = \mathcal{O}/\mathcal{P}$ *such that* $0 \in R$*, let* $\pi \in \mathcal{O}$ *be a prime element. Then every* $x \neq 0$ *in* K *admits a unique representation as convergent series*

$$x = \pi^m(a_0 + a_1\pi + a_2\pi^2 + \cdots)$$

where $a_i \in R$*,* $a_0 \neq 0$*,* $m \in \mathbb{Z}$*.*

Proof. A detailed proof is given in [7, Proposition 4.4].

□

5 Topological Groups

The following propositions and definitions are taken from [4, Section 9].

Definition 5.1 (Topological Group)**.** *A **topological group** G is a group (say with operation $(x, y) \mapsto xy$) such that G is also equipped with a topology where the operations $(x, y) \mapsto xy$ and $x \mapsto x^{-1}$ are continuous.*

*A **locally compact group** is a topological group whose topology is locally compact and Hausdorff.*

Proposition 5.2. *Let G be a topological group, let e be the identity element of G, and a be an arbitrary element of G.*

- *The functions $x \mapsto ax$, $x \mapsto xa$ and $x \mapsto x^{-1}$ are homeomorphisms of G onto G.*
- *If $\mathscr{U}$ is a base for the family of neighbourhoods of e, then $\{aU : U \in \mathscr{U}\}$ and $\{Ua : U \in \mathscr{U}\}$ are bases for family of the neighbourhoods of a.*
- *If K and L are compact subsets of G, then aK, Ka, KL and K^{-1} are compact subsets of G.*

Proposition 5.3. *Let G be a topological group, let e be the identity element of G, and let U be an open neighbourhood of e.*

- *There is an open neighbourhood V of e such that $VV \subseteq U$.*
- *There is a symmetric open neighbourhood V' of e that is contained in U. (V' is called symmetric if $V' = (V')^{-1}$)*

Proposition 5.4. *Let G be a topological group, and let H be an open subgroup of G. Then H is closed.*

6 Haar Measure

As in Section 5, our reference for this section is [4, 9].

Definition 6.1. *Let G be a locally compact group, and let μ be a nonzero regular Borel measure on G. Then μ is a left **Haar measure** if it is invariant under left translations, in the sense that $\mu(xA) = \mu(A)$ holds for each $x \in G$ and each $A \in \mathscr{B}(G)$. A right Haar measure is defined similarly, in case the group G is abelian then the notion of left and right coincides.*

Remark. *Proposition 5.2 implies that if $x \in G$ and if A is a Borel subset of G, then xA and Ax are Borel subsets of G, thus the Haar measure is well defined on those sets.*

Remark. *In our paper we will be interested in the case where a locally compact group G is abelian, henceforth we will drop the notion of left-right Haar measure.*

Theorem 6.2 (Existence of Haar Measure)**.** *Let G be a locally compact group. Then there is a Haar measure on G.*

Theorem 6.3 (Uniqueness of Haar measure)**.** *Let G be a locally compact group, and let μ and μ' be Haar measures on G. Then there is a positive real number c such that $\mu' = c\mu$.*

Proposition 6.4. *Let G be a locally compact group, and let μ be a Haar measure on G. Then μ is finite if and only if G is compact.*

7 Preliminaries

7.1 Identification on polynomial rings

For a ring R, let $R[x]$ denote the ring of univariate polynomials over R, and for $n \geq 0$, let $R[x]_n$ denote the subset of polynomials of degree at most n, and $R[x]_n^1$ the subset of monic polynomials of degree n.

We identify $R[x]_n^1$ with R^n via

$$x^n + \sum_{i=0}^{n-1} a_i x^i \leftrightarrow (a_0, a_1, \ldots, a_{n-1}).$$

When R is a complete discrete valuation ring with a finite residue field, we will the use measure on $R[x]_n^1$ which is inherited with this identification.

Now suppose that $\mathcal{O}$ is a complete discrete valuation ring with a finite residue field. For $f \in \mathcal{O}[x]$, we denote by $\overline{f}$ its image under reduction modulo π in $\mathbb{F}_q = \mathcal{O}/\mathcal{P} = \mathcal{O}/\pi\mathcal{O}$, where $\mathcal{P} = (\pi)$ is the unique maximal ideal of $\mathcal{O}$. A polynomial with coefficients in $\mathcal{O}$ is primitive if not all its coefficients are divisible by π, or in other words $\overline{f} \neq 0$. For a primitive polynomial $f \in \mathcal{O}[x]$, we define the reduced degree of f to be $\deg(\overline{f})$. Therefore $\deg(f) \geq \deg(\overline{f})$, with equality satisfied if and only if the leading coefficient of f is a unit since $\mathcal{O}$ has a unique maximal ideal which contains all non-units.

7.2 Some counting

Let $\mathcal{O}$ be a complete discrete valuation ring with a finite residue field $\mathbb{F}_q$ where q denotes the order of the field.

A **splitting type of degree** n is a tuple $(d_1^{e_1} d_2^{e_2} \cdots d_t^{e_t})$ where d_j and e_j are positive integers satisfying $\sum d_j e_j = n$. We allow repeats in the list of symbols $d_j^{e_j}$, but the order in which they appear does not matter. Let $S(n)$ denote all the splitting types of degree n. For example $S(2) = \{(1\ \ 1), (1^2), (2)\}$ has three elements, $S(3)$ has five elements, and $S(4)$ has eleven.

We say that a monic polynomial f in $\mathbb{F}_q[x]_n$ of degree n has splitting type $(d_1^{e_1} d_2^{e_2} \cdots d_t^{e_t}) \in S(n)$ if it factors as $f(x) = \prod_{j=1}^{t} f_j(x)^{e_j}$ where f_j are distinct irreducible monic polynomials over $\mathbb{F}_q$ with $\deg(f_j) = d_j$.

We write $\sigma(f)$ for the splitting type of f, and N_σ for the number of monic polynomials in $\mathbb{F}_q[x]_n$ with splitting type σ.

If $\sigma = (d)$, then we write N_d for N_σ, that is the number of degree d irreducible monic polynomials over $\mathbb{F}_q$. Writing μ for the Möbius function [1, Section 2.2], we get a formula for N_d. [5, Section 14.3, Proposition 18]

$$N_d = \frac{1}{d}\sum_{k|d} \mu(k) q^{d/k}.$$

Since there are q^n monic polynomials of degree n in $\mathbb{F}_q[x]$, the probability that a degree n monic polynomial $f \in \mathbb{F}_q[x]$ has splitting type σ, for $\sigma \in S(n)$, is N_σ/q^n. This is a rational function of q, there is a formula given in [2, Section 2.2] for N_σ, where the output for any N_σ is a function of q.

7.2.1 Power series identities involving N_σ

Referring to [2, Section 2.3], we follow the authors' footsteps to get power series identities involving the counts N_σ. They will be necessary for our paper's main goal.

Let x_{de} for $d, e \geq 1$ be indeterminates. For a splitting type $\sigma \in S(n)$ of degree n, let

$$x_\sigma = \prod_{d^e \in \sigma} x_{de}.$$

Polynomials with these indeterminates will be weighted by setting $wt(x_{de}) = de$. Setting $y_0 = 1$, and for $n \geq 1$ we define

$$y_n = \sum_{\sigma \in S(n)} N_\sigma x_\sigma,$$

so that every monomial in y_n has weight n. We set $x_{d0} = 1$ for all $d \geq 1$.

Proposition 7.1. *We have the following identity in $\mathbb{Z}\{x_{de}\}_{d,e\geq 1}[[t]]$:*

$$\sum_{n=0}^{\infty} y_n t^n = \prod_{d=1}^{\infty} (\sum_{e=0}^{\infty} x_{de} t^{de})^{N_d}.$$

Proof. When the right hand side is multiplied out, we see that the coefficient of t^n is a sum of monomials in the x_{de} of weight n. Such product has the form x_σ for some σ in $S(n)$ and the times that each monomial occurs is equal to N_σ which shows the coefficient of t^n is equal to y_n proving the claim. □

Corollary 7.2. *We have the following identity in $\mathbb{Z}[[t]]$:*

$$(1 - qt)^{-1} = \prod_{d=1}^{\infty} (1 - t^d)^{-N_d}.$$

Proof. Setting $x_{de} = 1$ for all d, e, we get $x_\sigma = \prod_{d^e \in \sigma} 1 = 1$, since in Proposition 7.1 we set $x_\sigma = \prod_{d^e \in \sigma} x_{de}$ and $x_{de} = 1$ for all d, e . Therefore $y_n = \sum_{\sigma \in S(n)} N_\sigma = q^n$, since the N_σ actually partitions the set $\mathbb{F}_q[x]^1_n$ which has q^n elements. We have that $\sum_{n=0}^\infty q^n t^n = \sum_{n=0}^\infty (qt)^n = (1-qt)^{-1}$ where we have used the geometric sum formula. We now use Proposition 7.1 again to get $(1-qt)^{-1} = \sum_{n=0}^\infty q^n t^n = \prod_{d=1}^\infty (\sum_{e=0}^\infty t^{de})^{N_d} = \prod_{d=1}^\infty (1-t^d)^{-N_d}$, where we have used the fact that $\sum_{e=0}^\infty t^{de} = (1-t^d)^{-1}$. □

Corollary 7.3. *Let x_e for $e \geq 1$ be indeterminates, and set $x_0 = 1$. Then in $\mathbb{Z}[x_1, x_2, \ldots][[t]]$, we have:*

$$\sum_{n=0}^\infty \sum_{\sigma \in S(n)} N_\sigma (\prod_{1^e \in \sigma} x_e) t^n = (\sum_{n=0}^\infty x_n t^n)^q (1-t)^q (1-qt)^{-1}.$$

Proof. We set $x_{1e} = x_e$ and $x_{de} = 1$ for all $d \geq 2$ in Proposition 7.1. Now we observe that

$$\sum_{n=0}^\infty y_n t^n = \prod_{d=1}^\infty (\sum_{e=0}^\infty x_{de} t^{de})^{N_d} = (\sum_{e=0}^\infty x_e t^e)^{N_1} \prod_{d=2}^\infty (\sum_{e=0}^\infty x_{de} t^{de})^{N_d},$$

where $N_1 = q$ and the indeterminates x_{de} are equal to 1 in the infinite product at the right hand side.

We now make use of the corollary we proved before, note that

$$(1-qt)^{-1} = \prod_{d=1}^\infty (\sum_{e=0}^\infty t^{de})^{N_d} = (\sum_{e=0}^\infty t^e)^{N_1} (\prod_{d=2}^\infty (\sum_{e=0}^\infty t^{de})^{N_d})$$

where

$$(\sum_{e=0}^\infty t^e)^{N_1} = (1-t)^{-q}.$$

Hence

$$\prod_{d=2}^\infty (\sum_{e=0}^\infty t^{de})^{N_d} = (1-qt)^{-1}(1-t)^q$$

Now we have

$$\sum_{n=0}^\infty y_n t^n = \sum_{n=0}^\infty \sum_{\sigma \in S(n)} N_\sigma x_\sigma t^n$$

where

$$x_\sigma = \prod_{1^e \in \sigma} x_e.$$

Therefore

$$\sum_{n=0}^\infty \sum_{\sigma \in S(n)} N_\sigma (\prod_{1^e \in \sigma} x_e) t^n = (\sum_{n=0}^\infty x_n t^n)^q (1-t)^q (1-qt)^{-1}.$$

□

7.3 Resultants, coprime factorizations, and independence

7.3.1 Resultants

We derive some lemmas about resultants of polynomials in $\mathcal{O}[x]$ and their behavior upon reduction modulo π.

Definition 7.4 (Sylvester matrix)**.** *Let $f(x) = a_m x^m + \cdots + a_0$ and $g(x) = b_n x^n + \cdots + b_0$ be two polynomials of degrees m and n, respectively over an arbitrary ring R. The **Sylvester matrix** is an $(m+n)\times(m+n)$ matrix formed by filling the matrix beginning with the upper left corner with coefficients of $f(x)$, then shifting down one row and column to the right filling in the coefficients starting there until they hit the right side. The process is then repeated for the coefficients of $g(x)$. In matrix form it is given as*

$$\mathrm{Syl}(f,g) = \begin{bmatrix} a_m & a_{m-1} & \dots & a_0 & 0 & \dots & 0 \\ 0 & a_m & a_{m-1} & \dots & a_0 & \dots & 0 \\ \vdots & & \ddots & & & \ddots & \vdots \\ 0 & \dots & 0 & a_m & a_{m-1} & \dots & a_0 \\ b_n & b_{n-1} & \dots & b_0 & 0 & \dots & 0 \\ 0 & b_n & b_{n-1} & \dots & b_0 & \dots & 0 \\ \vdots & & \ddots & & & \ddots & \vdots \\ 0 & \dots & 0 & b_n & b_{n-1} & \dots & b_0 \end{bmatrix}.$$

Definition 7.5 (Resultant)**.** *The **resultant** of two polynomials over a commutative ring R is defined as the determinant of their Sylvester matrix. We let $\mathrm{Res}(f,g)$ denote the resultant of two polynomials f, g.*

Lemma 7.6. *Let $f, g \in \mathcal{O}[x]$ have degrees m and n respectively.*

1. *If the leading coefficients of f and g are both units, then $\overline{\mathrm{Res}(f,g)} = \mathrm{Res}(\overline{f}, \overline{g})$.*

2. *If the leading coefficient a_m of f is a unit and $d = \deg(\overline{g}) < n$, then $\overline{\mathrm{Res}(f,g)} = \overline{a_m}^{n-d}\,\mathrm{Res}(\overline{f}, \overline{g})$.*

3. *If the leading coefficients of f and g are both non-units, then $\overline{\mathrm{Res}(f,g)} = 0$.*

Proof. 1. Since the leading coefficients of both f and g are units, they are not contained in the unique maximal ideal, therefore $\overline{a_m} \neq 0, \overline{b_n} \neq 0$. Since in this case $\mathrm{Syl}(\overline{f}, \overline{g})$ is actually the reduction mod π of $\mathrm{Syl}(f,g)$, $\overline{\mathrm{Res}(f,g)} = \mathrm{Res}(\overline{f}, \overline{g})$.

2. Suppose the leading coefficient a_m of f is a unit and $d = \deg(\overline{g}) < n$. This implies that in reduction mod π of $\mathrm{Syl}(f,g)$ all the terms $\overline{b_i}, i > d$ are actually equal to 0. Therefore

$$\overline{\mathrm{Res}(f,g)} = \det(\overline{\mathrm{Syl}(f,g)}) = \overline{a_m}^{n-d}\det(\mathrm{Syl}(\overline{f}, \overline{g})) = \overline{a_m}^{n-d}\,\mathrm{Res}(\overline{f}, \overline{g}).$$

3. This is trivially true since the leading coefficients are non-units, under the reduction map they are equal to 0. Therefore $\overline{\mathrm{Res}(f,g)} = 0$.

□

Corollary 7.7. *Let $f, g \in \mathcal{O}[x]$ have degrees m and n respectively. Then* $\mathrm{Res}(f,g)$ *is a unit if and only if at least one of the leading coefficients of f or g is a unit, and the reductions $\overline{f}, \overline{g}$ are coprime.*

Proof. If $\mathrm{Res}(f,g)$ is a unit in $\mathcal{O}$, then $\mathrm{Res}(f,g) \neq 0$ and also $\overline{\mathrm{Res}(f,g)} \neq 0$. Moreover since R is a principal ideal domain, it is also a unique factorization domain, and a well known property of resultants is that for two polynomials f', g' over a unique factorization domain $\mathrm{Res}(f', g') = 0$ if and only if they share a common root. Therefore by first and second part of Lemma 7.6, the reductions should be coprime while at least one of the polynomials should have a leading coefficient which is a unit.

Conversely if the reductions are coprime then $\mathrm{Res}(\overline{f}, \overline{g}) \neq 0$. Moreover as we assume at least one of the leading coefficients is a unit, we see that $\overline{\mathrm{Res}(f,g)} \neq 0$. Therefore $\mathrm{Res}(f,g)$ is a unit in $\mathcal{O}$.

□

Lemma 7.8. *Let R be an arbitrary ring. For any $d \geq 1$, we identify $R[x]_d^1 \cong R^d$ and $R[x]_d \cong R^{d+1}$ as R-modules.*

1. *The multiplication map $R[x]_m^1 \times R[x]_n^1 \to R[x]_{m+n}^1$ has Jacobian at (f,g) given by* $\mathrm{Res}(f,g)$.

2. *The multiplication map $R[x]_m^1 \times R[x]_n \to R[x]_{m+n}$ has Jacobian at (f,g) given by* $\mathrm{Res}(f,g)$.

Proof. A detailed proof is given in [2, Lemma 2.6]. □

Corollary 7.9. *Let $A \subset \mathcal{O}[x]_m^1$, $B \subset \mathcal{O}[x]_n^1$, and $AB \subset \mathcal{O}[x]_{m+n}^1$ be measurable subsets such that multiplication induces a bijection*

$$A \times B \to AB = \{ab \mid a \in A, b \in B\}.$$

If $\mathrm{Res}(g,h) \in \mathcal{O}^*$ *for all $g \in A$ and $h \in B$, then this bijection is measure preserving.*

Proof. We identify $\mathcal{O}[x]_m^1$, $\mathcal{O}[x]_n^1$ and $\mathcal{O}[x]_{m+n}^1$ with $\mathcal{O}^m, \mathcal{O}^n$ and $\mathcal{O}^{m+n}$ respectively. Similarly we also identify $\mathcal{O}^m \times \mathcal{O}^n$ with $\mathcal{O}^{m+n}$. With these identifications we view the map as a map from R^{m+n} to itself. The change of variables formula from measure theory induces a Jacobian factor $\mathrm{Res}(f,g)$. Written explicitly [3, Thrm. 10.1.2],

$$\int_{g \in A} \int_{h \in B} |\mathrm{Res}(g,h)| dh dg = \int_{f \in AB} df.$$

Since $|l| = 1$ for every $l \in \mathcal{O}^*$ and $\mathrm{Res}(g,h)$ is assumed to be a unit, $|\mathrm{Res}(g,h)| = 1$ proving the claim. □

7.3.2 Measure theoretic Hensel's lemma

For $f \in \mathbb{F}_q[x]^1_d$, denote by P_f the set of polynomials in $\mathcal{O}[x]^1_d$ that reduce to f modulo π; and for $n \geq d$, we denote by P^n_f the set of polynomials in $\mathcal{O}[x]_n$ that reduce to f modulo π.

Lemma 7.10. *Suppose that $g, h \in \mathbb{F}_q[x]$ are monic and coprime. Then the multiplication map*

$$P_g \times P_h \to P_{gh}$$

is a measure preserving bijection.

Proof. We first show that P_f is a measurable set for any $f \in \mathbb{F}_q[x]^1_d$. Suppose $g(x) \in P_f$ where $g(x) = a_0 + a_1x + \cdots + a_{d-1}x^{d-1} + x^d \in \mathcal{O}[x]^1_d$, similarly assume $f(x) = b_0 + b_1x + \cdots + b_{d-1}x^{d-1} + x^d \in \mathbb{F}_q[x]^1_d$. Since $\overline{g} = f$, we have that $a_i \equiv b_i$ mod π for all $0 \leq i \leq d$. This implies $a_i = \pi k + b'_i$ for some $k, \in \mathcal{O}$ where b'_i a fixed lift of b_i , therefore $a_i \in b_i + \mathcal{P}$. $x + \mathcal{P}$ is open for all $x \in \mathcal{O}$. Therefore P_f is measurable for all $f \in \mathbb{F}_q[x]^1_d$.

Now let $f \in \mathcal{O}[x]^1_n$ be such that $\overline{f}$ factors in $\mathbb{F}_q[x]$ as $\overline{f} = gh$. Then by Hensel's lemma f factors uniquely in $\mathcal{O}[x]$ as $f = g'h'$, where $g' \in P_g$ and $h' \in P_h$. Therefore the map we have is a bijection of measurable sets, where the Jacobian is a unit in $\mathcal{O}$ (which implies it has a valuation 1). Due to the Corollary 7.9 this map is measure preserving. The resultant is a unit due to Corollary 7.7. □

We use the following variant of previous the lemma to handle polynomials $f \in \mathcal{O}[x]$ whose leading coefficient is not a unit.

Lemma 7.11. *Let $n \geq m \geq 0$ and consider the multiplication map*

$$\mu : \mathcal{O}[x]^1_m \times P^{n-m}_1 \to \{f \in \mathcal{O}[x]_n : \overline{f} \in \mathbb{F}_q[x]^1_m\}. \tag{13}$$

1. *μ is a measure preserving bijection.*

2. *Let r be an integer satisfying $m \leq r \leq n$. In (13) , if we replace the set on the right hand-side with the subset of $f \in \mathcal{O}[x]_n$ also satisfying $\pi^r f(x\pi^{-1}) \equiv x^n (mod\ \pi)$, and replace the second factor of the left-hand side with the subset of $h \in P^{n-m}_1$ satisfying $\pi^{r-m}h(x\pi^{-1}) \equiv x^{n-m} (mod\ \pi)$. Then the restriction of μ to these subsets is still a measure preserving bijection.*

Proof. 1. Let $f \in \mathcal{O}[x]_n$ be such that $\overline{f}$ is monic of degree m. Now note that $\overline{f} = (\overline{1})(x^m + \cdots + \overline{a_0})$. We use Hensel's lemma to see that $f = gh$ where $g \in \mathcal{O}[x]^1_m$ and $h \in P^{n-m}_1$. Therefore (13) is a bijection and measure preserving due to Corollary 7.9 since g is monic and the resultant is a unit. Again the resultant is a unit since the reductions are coprime for g and h, therefore we use 7.7.

2. Let $f \in \mathcal{O}[x]_n$ be such that $\overline{f}$ is monic of degree m, and also that $\pi^r f(x\pi^{-1}) \equiv x^n \mod \pi$; this implies that $\pi^r f(x\pi^{-1}) \in \mathcal{O}[x]$ and $\pi^r f(x\pi^{-1}) - x^n \in \pi\mathcal{O}[x]$. We factor $f = gh$ as before, then

$$\pi^m g(x\pi^{-1}) \cdot \pi^{r-m} h(x\pi^{-1}) = \pi^m f(x\pi^{-1}) \equiv x^n \mod \pi;$$

since $g \in \mathcal{O}[x]^1_m$, we have $\pi^m g(x\pi^{-1}) \in \mathcal{O}[x]$ and $\pi^m g(x\pi^{-1}) \equiv x^m \mod \pi$. Since $\pi^r f(x\pi^{-1}) \in \mathcal{O}[x]$, $\pi^r f(x\pi^{-1}) = \pi^m g(x\pi^{-1}) \cdot \pi^{r-m} h(x\pi^{-1})$ and $\pi^m g(x\pi^{-1}) \in \mathcal{O}[x]$, by Gauss's lemma $\pi^{r-m} h(x\pi^{-1}) \in \mathcal{O}[x]$. Now using the unique factorization in $\mathbb{F}_q[x]$ also $\pi^{r-m} h(x\pi^{-1}) \equiv x^{n-m} \mod \pi$.

Conversely if h satisfies $\pi^{r-m} h(x\pi^{-1}) \equiv x^{n-n} \mod \pi$, then since $\pi^m g(x\pi^{-1}) \equiv x^m \mod \pi$ for all $g \in \mathcal{O}[x]^1_m$, it follows that $f = gh$ satisfies $\pi^r f(x\pi^{-1}) \equiv x^n \mod \pi$. Therefore μ restricts to a bijection between the subsets on each side, and measure preserving as before as well. Similarly the resultant is a unit due to 7.7.

□

7.3.3 Independence Lemmas

We rephrase Lemmas 7.10 and 7.11 to create suitable random variables which are independent.

Corollary 7.12. *Let $f, g \in \mathbb{F}_q[x]$ be coprime monic polynomials. For $f \in P_{gh}$, let ϕ_1 and ϕ_2 denote the projections of P_{gh} onto P_g and P_h, respectively under the bijection $P_{gh} \to P_g \times P_h$. Then the number of K-roots of $f \in P_{gh}$ is $X + Y$, where $X, Y : P_{gh} \to \{0, 1, 2, \ldots\}$ are independent random variables distributed on $f \in P_{gh}$ as the number of K-roots of $\phi_1(f) \in P_g$ and $\phi_2(f) \in P_h$, respectively. The independence is justified due to the measure preserving property of the map in Lemma 7.10.*

Corollary 7.13. *Let $m \leq n$, and let*

$$B_{m,n} := \{f \in \mathcal{O}[x]_n : \overline{f} \in \mathbb{F}_q[x]^1_m\}.$$

For $f \in B_{m,n}$, let ψ_1 and ψ_2 denote the projections of $B_{m,n}$ onto $\mathcal{O}[x]^1_m$ and P_1^{n-m}, respectively under the bijection $B_{m,n} \to \mathcal{O}[x]^1_m \times P_1^{n-m}$. Let $X, Y : B_{m,n} \to \{0, 1, 2, \ldots\}$ be the random variables giving the number of roots of $f \in B_{m,n}$, in $\mathcal{O}$ and in $K \setminus \mathcal{O}$, respectively. Then X and Y are independent random variables distributed on $f \in B_{m,n}$ as the number of K-roots of $\psi_1(f)(x) \in \mathcal{O}[x]^1_m$ and of $\psi_2(f)^{rev}(x) := x^{n-m}\psi_2(f)(\frac{1}{x}) \in P_{x^{n-m}}$, respectively.

8 Proof of the Theorem

8.1 Proof of Theorem 1(b)

8.1.1 Conditional Expectations

In the introduction we defined the expectations $\alpha(n, d)$ and $\beta(n, d)$. We now make a few definitions for our problem.

Lemma 8.1. *P_f has relative density q^{-n} in $\mathcal{O}[x]_n^1$.*

Proof. We can prove the statement in two ways. Firs suppose that $f(x) = a_0 + a_1x + \ldots + x^n \in \mathbb{F}_q[x]_n^1$. Notice that there are exactly q choices for each a_i with each having probability $\frac{1}{q}$. Therefore the relative density of P_f is $(\frac{1}{q})^n = q^{-n}$.

Another way to prove the statement is to see that if we let $R = \{a_0, \ldots, a_{q-1}\}$ with $a_0 = 0$, be a set of representatives for $\mathbb{F}_q$, then we decompose the local field K as

$$\mathcal{O} = \bigcup_{i=0}^{q-1} (a_i + \mathcal{P}).$$

Since μ is a Haar measure, $\mu(a_i + \mathcal{P}) = \mu(\mathcal{P})$, and normalization $\mu(\mathcal{O}) = 1$ forces $\mu(\mathcal{P}) = q^{-1}$. Using these arguments we see

$$\mu(\mathcal{P}^n) = q^{-n}, \quad \textit{for all } n \in \mathbb{N}_{\geq 0}.$$

□

Definition 8.2.

i) For $f \in \mathbb{F}_q[x]_n^1$, let $\alpha(n, d \mid f)$ denote the expected number of d-sets of K-roots of a polynomial in $P_f \subset \mathcal{O}[x]_n^1$. Since P_f has relative density q^{-n} in $\mathcal{O}[x]_n^1$, we have

$$\alpha(n, d) = q^{-n} \sum_{f \in \mathbb{F}_q[x]_n^1} \alpha(n, d \mid f). \tag{14}$$

Moreover $\beta(n, d) = \alpha(n, d \mid x^n)$.

ii) For $\sigma \in S(n)$, let $\alpha(n, d \mid \sigma)$ be the expected number of d-sets of K-roots of a polynomial in $\mathcal{O}[x]_n^1$ whose splitting type is σ. Therefore

$$\alpha(n, d) = q^{-n} \sum_{\sigma \in S(n)} N_\sigma \alpha(n, d \mid \sigma), \tag{15}$$

and

$$\alpha(n, d \mid \sigma) = N_\sigma^{-1} \sum_{f \in \mathbb{F}[x]_n^1 : \sigma(f) = \sigma} \alpha(n, d \mid f), \tag{16}$$

where $\sigma(f)$ denotes the splitting type of f.

The intuition in Definition 8.2 i) is the fact that the space $\mathcal{O}[x]_n^1$ has a measure which is equal to 1 which is induced by our identification with the space $\mathcal{O}^n$. Each polynomial $f \in \mathbb{F}_q[x]_n^1$ corresponds to a subset of $\mathcal{O}[x]_n^1$ consisting of all the lifts of f which is denoted by P_f. The set of polynomials $\mathcal{O}_n^1[x]$ is covered disjointly by these sets where we have proven their density to be $\mu(P_f) = q^{-n}$. Therefore

$$\alpha(n, d) = \sum_{f \in \mathbb{F}_q[x]_n^1} \mu(P_f)\alpha(n, d \mid f) = q^{-n} \sum_{f \in \mathbb{F}_q[x]_n^1} \alpha(n, d \mid f). \tag{17}$$

Moreover $\beta(n, d)$ is just a specific configuration of α.

Similarly for Definition 8.2 ii), the relative density depending on the splitting type σ is raised by a factor of N_σ for each σ. This also has an additional factor of q^{-n} since P_f has relative density q^{-n} for any $f \in \mathbb{F}_q[x]_n^1$. (One can also remember that the probability of a monic polynomial having splitting type $\sigma \in S(n)$ in $\mathbb{F}_q[x]_n$ is $\frac{N_\sigma}{q^n}$.)

8.1.2 Writing α's in terms of β's

Lemma 8.3. *Let $g, h \in \mathbb{F}_q[x]$ be monic and coprime. Then*

$$\alpha(\deg(gh), d \mid gh) = \sum_{d_1+d_2=d} \alpha(\deg(g), d_1 \mid g) \cdot \alpha(\deg(h), d_2 \mid h), \tag{18}$$

where the sum is over all pairs (d_1, d_2) of non-negative integers summing to d.

Moreover if h has no roots in $\mathbb{F}_q$, then

$$\alpha(\deg(gh), d \mid gh) = \alpha(\deg(g), d \mid g).$$

Proof. The lemma is an implication of Corollary 7.12 with the fact that if X and Y are independent random variables taking values in $\{0, 1, 2, \ldots\}$ then

$$\mathbb{E}\binom{X+Y}{d} = \sum_{d_1+d_2=d} \mathbb{E}\binom{X}{d_1}\mathbb{E}\binom{Y}{d_2}. \tag{19}$$

The implication of Corollary 7.12 tells us the fact that X and Y are the number of K roots of a lift of g and h in P_g and P_h respectively and X and Y are independent. We note that expectation is a linear function, and the fact that each d-set of roots must be formed from some d_1-set from g and d_2-set from h, summing over all possible partitions $d_1 + d_2 = d$. If h has no roots in $\mathbb{F}_q$, then we know that $Y = 0$, so the roots are determined completely by g. □

We recall that $\beta(n, d) = \alpha(n, d \mid x^n)$ is the expected number of d-sets of roots of a monic polynomial of degree n which reduces to x^n modulo π. Using Lemma 8.3, we can express $\alpha(n, d \mid f)$ for monic $f \in \mathbb{F}_q[x]_n$ in terms of $\beta(n', d')$ for suitable n', d'.

Lemma 8.4. *Let $\sigma = (1^{n_1} \cdots 1^{n_k}) \in S(n)$ be a splitting type with exactly $k = m_1(\sigma)$ powers of 1. Then*

$$\alpha(n, d \mid \sigma) = \sum_{d_1+d_2+\cdots+d_k=d} \prod_{i=1}^{k} \beta(n_i, d_i). \tag{20}$$

Proof. Let $f \in \mathbb{F}_q[x]_n^1$ have splitting type σ. To evaluate $\alpha(n, d \mid f)$, we ignore the factors of f of degree greater than 1, since if $f = f_1 f_2$ where $\sigma(f_1) = (1^{n_1} \dots 1^{n_k})$ and f_2 has no linear factors, then $\alpha(n, d \mid f) = \alpha(\deg(f_1), d \mid f_1)$ by Lemma 8.3.

Now let $f = \prod_{i=1}^k l_i^{n_i}$, where l_i are distinct, monic, and of degree 1. Using Lemma 8.3 repeatedly gives

$$\alpha(n, d \mid f) = \sum_{d_1 + \dots + d_k = d} \prod_{i=1}^k \alpha(n_i, d_i \mid l_i^{n_i}).$$

Finally, $\alpha(n_i, d_i \mid l_i^{n_i}) = \alpha(n_i, d_i \mid x^{n_i}) = \beta(n_i, d_i)$, since for fixed $c \in \mathcal{O}$ the map $g(x) \mapsto g(x + c)$ is measure preserving on monic polynomials in $\mathcal{O}[x]$ of given degree. Therefore

$$\alpha(n, d \mid f) = \sum_{d_1 + d_2 + \dots + d_k = d} \prod_{i=1}^k \beta(n_i, d_i) = \alpha(n, d \mid \sigma) \tag{21}$$

where the last equality follows from our assumption of $\sigma = (1^{n_1} \dots 1^{n_k} \dots) \in S(n)$. □

Proof of Theorem 1.1 b).

Proof. Let $\sigma = (1^{n_1} \dots 1^{n_k}) \in S(n)$ be as in the lemma before. We now have

$$\alpha(n, d) = q^{-n} \sum_{\sigma \in S(n)} N_\sigma \alpha(n, d \mid \sigma) = q^{-n} \sum_{\sigma \in S(n)} N_\sigma \sum_{d_1 + \dots + d_k = d} \prod_{i=1}^k \beta(n_i, d_i). \tag{22}$$

Let us now multiply both sides with u^d.

$$\alpha(n, d) u^d = q^{-n} u^d \sum_{\sigma \in S(n)} N_\sigma \sum_{d_1 + \dots + d_k = d} \prod_{i=1}^k \beta(n_i, d_i).$$

Now we sum for d.

$$\sum_{d=0}^n \alpha(n, d) u^d = q^{-n} \sum_{\sigma \in S(n)} N_\sigma \sum_{d=0}^n \sum_{d_1 + \dots + d_k = d} \prod_{i=1}^k \beta(n_i, d_i) u^d,$$

where

$$\sum_{d=0}^n \sum_{d_1 + \dots + d_k = d} \prod_{i=1}^k \beta(n_i, d_i) u^d = \prod_{1^e \in \sigma} (\sum_{d=0}^e \beta(e, d) u^d). \tag{23}$$

Let us explain how the equality above follows. We can reformulate the right hand side as

$$\prod_{1^e \in \sigma} \left(\sum_{d=0}^e \beta(e, d) u^d \right) = \prod_{i=1}^k \left(\sum_{d_i=0}^{n_i} \beta(n_i, d_i) u^{d_i} \right),$$

□

since the parts of σ of the form 1^e are 1^{n_i} for $i = 1, \ldots, k$. Once we multiply out, the coefficient of u^d is in fact $\prod_{d_1+\cdots+d_k=d} \beta(n_i, d_i)$, hence the left hand side is equal to the right hand side in equation (21).

We now have that

$$\sum_{d=0}^{n} \alpha(n,d)u^d = q^{-n} \sum_{\sigma \in S(n)} N_\sigma \prod_{1^e \in \sigma} \left(\sum_{d=0}^{e} \beta(e,d)u^d \right). \tag{24}$$

Now we multiply both sides with (qt^n) and sum over n. For a moment let us focus only on the right hand side of (22). On the right hand side we then get

$$\sum_{n=0}^{\infty} \sum_{\sigma \in S(n)} N_\sigma \prod_{1^e \in \sigma} \left(\sum_{d=0}^{e} \beta(e,d)u^d \right) t^n. \tag{25}$$

If we set

$$\prod_{1^e \in \sigma} \left(\sum_{d=0}^{e} \beta(e,d)u^d \right) t^n = \prod_{1^e \in \sigma} x_e t^n,$$

we can use Corollary 7.3, after that we get

$$\sum_{n=0}^{\infty} \sum_{\sigma \in S(n)} N_\sigma \prod_{1^e \in \sigma} \left(\sum_{d=0}^{e} \beta(e,d)u^d \right) t^n = \frac{\left(\sum_{d=0}^{\infty} \left(\sum_{n=0}^{\infty} \beta(n,d)t^n \right) u^d \right)^q (1-t)^q}{(1-qt)}.$$

On the left hand side (after multiplying with $(qt)^n$ and summing over n) we have

$$\sum_{n=0}^{\infty} \sum_{d=0}^{n} \alpha(n,d)u^d (qt)^n = \sum_{d=0}^{\infty} \left(\sum_{n=0}^{\infty} \alpha(n,d)(qt)^n \right) u^d.$$

Therefore we have the equality

$$\sum_{d=0}^{\infty} \left(\sum_{n=0}^{\infty} \alpha(n,d)(qt)^n \right) u^d = \left(\sum_{d=0}^{\infty} \left(\sum_{n=0}^{\infty} \beta(n,d)t^n \right) u^d \right)^q (1-t)^q (1-qt)^{-1}.$$

Now multiplying both sides with $1 - qt$ we have,

$$\sum_{d=0}^{\infty} \mathcal{A}_d(qt)u^d = \sum_{d=0}^{\infty} (1-qt) \left(\sum_{n=0}^{\infty} \alpha(n,d)(qt)^n \right) u^d = \tag{26}$$

$$\left(\sum_{d=0}^{\infty} \left(\sum_{n=0}^{\infty} \beta(n,d)t^n \right) (1-t)u^d \right)^q = \left(\sum_{d=0}^{\infty} \mathcal{B}_d(t)u^d \right)^q, \tag{27}$$

proving (5).

8.1.3 Writing ρ's in terms of α's and β's

In the introduction we defined $\rho(n,d)$ to be the expected number of d-sets of K-roots of polynomials $f \in \mathcal{O}[x]$ of degree n. Our aim in this section is to prove equation 6 in Theorem 1.1. Before we start, we need to work out some facts.

We start by noting that any $f \in \mathcal{O}[x]$ such that $f \neq 0$ can be decomposed as $f = \pi^e f_{\text{prim}}$ where $e \geq 0$ is the greatest possible integer and $f_{\text{prim}} \in \mathcal{O}[x]$ is a primitive polynomial. In the case $e = 0$, the polynomial f itself is already primitive and $f_{\text{prim}} = f$. Moreover the set of roots of f is equal to f_{prim}, since $f(\epsilon) = 0$ if and only if $f_{\text{prim}}(\epsilon) = 0$ for any root $\epsilon \in K$ of f. Since $\rho(n,d)$ only depends on n, d and q, the restriction will not change when we restrict this expectation to primitive polynomials.

Let $f \in \mathcal{O}[x]$ be a primitive polynomial of degree n. Let $m = \deg(\overline{f})$ be the reduced degree of f. For fixed $0 \leq m \leq n$, we get the density of primitive polynomials $f \in \mathcal{O}[x]_n$ with reduced degree m with the following statements. We first note that there are exactly $q^{n+1} - 1$ choices of a polynomial of degree $\leq n$ in $\mathbb{F}_q[x] \setminus \{0\}$. Similarly if $f \in \mathbb{F}_q[x]_m$ with degree m, then we have $q - 1$ choices for the m'th coefficient, and q choices for all other m many coefficients. Now we derive a density $\frac{q-1}{q^{n+1}-1} q^m$. Conditioning on the value of m, we have

$$\rho(n,d) = \frac{q-1}{q^{n+1}-1} \sum_{m=0}^{n} q^m \rho(n,d,m), \tag{28}$$

where $\rho(n,d,m)$ denotes the expected number of d-sets of K-roots of f as $f \in \mathcal{O}[x]_n$ runs over polynomials of degree n with reduced degree m. This expectation will not change when we restrict to f whose reduction mod π is monic, since the behavior of $\rho(n,d,m)$ depends on the distribution of polynomials with given reduced degree m (and their roots). If we restrict ourselves to only ones that reduce to a monic polynomial, then technically we are changing a factor of $q - 1$ on all of them, which lets the expectation stay the same.

Lemma 8.5. *We have*

$$\rho(n,d,m) = \sum_{d_1+d_2=d} \alpha(m,d_1)\beta(n-m,d_2).$$

Proof. Corollary 7.13 tells us that the number of roots from reduction and the number of lifted roots are independent. The number of $\mathcal{O}$-roots is determined by the polynomial's reduction mod π. Similarly the number of roots appearing in K but not $\mathcal{O}$ is determined by the additional $n - m$ higher-degree terms. So we can decompose $\rho(n,d,m)$ with contributions from roots that are already present in the reduction and the roots that come from the lifting. $\alpha(m,d_1)$ is the expected number of d_1-sets of roots coming from the reduction of the polynomial mod π. In a similar manner $\beta(n-m,d_2)$ are the expected number of d_2-sets of roots that arise from lifting the roots. Since they are independent we can use (19) to get $\rho(n,d,n) = \sum_{d_1+d_2=d} \alpha(m,d_1)\beta(n-m,d_2)$. □

Proof of equation (6).

$$\begin{aligned}
\sum_{d=0}^{\infty}\mathcal{R}_d(t)u^d &= \sum_{d=0}^{\infty}(1-t)(1-qt)\sum_{n=0}^{\infty}(q^n+q^{n-1}+\cdots+1)\rho(n,d)t^nu^d \\
&= (1-t)(1-qt)\sum_{d=0}^{\infty}u^d\sum_{n=0}^{\infty}(q^n+\cdots+1)t^n\frac{q-1}{q^{n+1}-1}\sum_{m=0}^{n}q^m\rho(n,d,m) \\
&= (1-t)(1-qt)\sum_{d=0}^{\infty}u^d\sum_{n=0}^{\infty}t^n\sum_{m=0}^{n}q^m\sum_{d_1+d_2=d}\alpha(m,d_1)\beta(n-m,d_2) \\
&= \sum_{d=0}^{\infty}u^d\sum_{n=0}^{\infty}t^n\sum_{m=0}^{n}q^m\sum_{d_1+d_2=d}(1-qt)\alpha(m,d_1)(1-t)\beta(n-m,d_2) \\
&= \sum_{d=0}^{\infty}u^d\sum_{n=0}^{\infty}t^n\sum_{m=0}^{n}q^m\sum_{d_1,d_2}(1-qt)\alpha(m,d_1)(1-t)\beta(n-m,d_2) \\
&= \sum_{d=0}^{\infty}\sum_{n=0}^{\infty}t^n\sum_{m=0}^{n}q^m\sum_{d_1,d_2}(1-qt)u^{d_1}\alpha(m,d_1)(1-t)u^{d_2}\beta(n-m,d_2) \\
&= \sum_{d=0}^{\infty}\sum_{n=0}^{\infty}\sum_{m=0}^{n}q^m\sum_{d_1,d_2}(1-qt)u^{d_1}\alpha(m,d_1)(1-t)u^{d_2}\beta(n-m,d_2)t^n \\
&= \left(\sum_{d=0}^{\infty}\mathcal{A}_d(qt)u^d\right)\left(\sum_{d=0}^{\infty}\mathcal{B}_d(t)u^d\right).
\end{aligned}$$

□

8.1.4 Writing β's in terms of α's

In this section we aim to prove equation (7) of Theorem 1.1.

Fixing d we set $\alpha_n := \alpha(n,d)$ and $\beta_n := \beta(n,d)$. We express β_n in terms of α_s for $s \le n$ with the following lemma.

Lemma 8.6. *We have*

$$\beta_n = q^{-\binom{n}{2}}\alpha_n + (q-1)\sum_{0\le s<r<n} q^{-\binom{r+1}{2}}q^s\alpha_s. \tag{29}$$

Proof. We first recall that β_n is the expected number of d-sets of K-roots of $f \in P_{x^n}$, but we now note that since $\mathcal{O}$ is integrally closed, any K-root is in fact an $\mathcal{O}$-root. Now let us show all such roots must lie in $\pi\mathcal{O}$. Suppose that $\epsilon \in \mathcal{O}$ is a root of $f \in P_{x^n}$. Since $f(\epsilon) = 0$, $\overline{f}(\epsilon) = \epsilon^n \equiv 0 \bmod \pi$. This is possible if $\epsilon = \pi k$ for some $k \in \mathcal{O}$. Furthermore since a root must be in $\pi\mathcal{O}$, if $\epsilon = \pi\epsilon'$ is a root of f where $\epsilon' \in \mathcal{O}$, then ϵ' is necessarily a $\mathcal{O}$-root of $f(\pi x)$. Now for each $f \in P_{x^n}$, we associate a pair of integers (r,s) with $0 \le s \le r \le n$ as follows. We consider $f(\pi x)$, and let r be the largest integer such that $\pi^r \mid f(\pi x)$, so that $1 \le r \le n$. Indeed since for $f = a_0 + a_1x + \cdots x^n \in P_{x^n}$ we know that

$a_i \equiv 0 \bmod \pi$ for all i. Therefore $f(\pi x) = a_0'\pi + a_1'\pi^2 x + \cdots \pi^n x^n$ for $a_i' \in \mathcal{O}$ for all i. Hence $1 \leq r \leq n$. We let s to be the reduced degree of $g(x) = \pi^{-r} f(\pi x)$. Then either $0 \leq s < r < n$, or $s = r = n$. Let us also show this fact. Suppose that $f(\pi x) = a_0'\pi + a_1'\pi^2 x + \cdots + a_{n-1}'\pi^n x^{n-1} + \pi^n x^n$ as before. Suppose $1 \leq r \leq n$ is found and consider:

$$\begin{aligned}\pi^{-r} f(\pi x) =& a_0'\pi^{1-r} + a_1'\pi^{2-r} x + \cdots + a_{n-1}'\pi^{n-r} x^{n-1} + \pi^{n-r} x^n \\ =& \pi^{n-r}(a_{n-1} x^{n-1} + x^n) + \sum_{i=0}^{n-2} a_i' \pi^{i+1-r} x^i.\end{aligned}$$

If $r = n$ then it is clear that $s = n = r$. On the other hand if $r < n$, then s depends on the fact when $i + 1 - r = 0$ which in that case $s = i_0$ for the highest i_0 satisfying the equality, and since $i + 1 = r < n$, we have that $0 \leq s < r < n$.

The relative density of the subset of $f \in P_{x^n}$ such that $\pi^r \mid f(\pi x)$ is $\pi^{-\binom{r}{2}}$. First note that we do not need to consider the coefficients of x^n and x^{n-1} in $f(\pi x)$ since they are divisible by π^r for all r. Now note that from before $a_i = a_i'\pi$ are the coefficients of the indeterminates in f. So $\pi^r \mid f(\pi x)$ if and only if $a_i'\pi^{i+1} = a_i \pi^i = \pi^r k'$ for some $k' \in \mathcal{O}$. So for $0 \leq i \leq r - 2$, we require the coefficient of x^i in f to be divisible by π^{r-i} and not only π. Now since we know that the probability that the coefficient of x^i is divisible by π^{r-i} is $\frac{1}{q^{r-i}}$ we have the density

$$1/q^{r-2} \times 1/q^{r-3} \times \cdots \times 1/q = q^{((r-2)+(r-3)+\cdots+1)} = q^{\frac{-r(r-1)}{2}} = q^{-\binom{r}{2}}.$$

Given $r < n$, the condition that $\pi^{-r} f(\pi x)$ has reduced degree at least s gives out $r - s - 1$ more divisibility conditions, since there are $r - s - 1$ coefficients that we put the condition on (-1 comes from excluding x^s). Each contribution gives out a factor of q^{-1} and since coefficient of x^s must not be divisible by π we have a factor of $(1 - 1/q)$ as well. Therefore the density of f such that the reduced degree is exactly s is

$$q^{-(r-s-1)}(1 - 1/q) = q^{s-r}(q - 1).$$

Therefore the relative density of $f \in P_{x^n}$ with parameters (r, s) is given by

$$q^{-\binom{r}{2}} q^{s-r}(q - 1) = q^{-\binom{r+1}{2}} q^s (q - 1)$$

for $0 \leq s < r < n$. If $r = n$, then $s = r$ as we have shown earlier, therefore the density with parameters (n, n) is $q^{-\binom{n}{2}}$.

If $s = r = n$, then $g = \pi^{-n} f(\pi x)$ is distributed as an arbitrary element of $\mathcal{O}[x]_n^1$, while if $s < r < n$ then g is subject to the conditions that $\overline{g}$ has degree s, and moreover $\pi^r f(x\pi^{-1}) = f(x) \equiv x^n \bmod \pi$. In both cases, given r and s the conditional expected number of d-sets of $\mathcal{O}$-roots of $f \in P_{x^n}$ is α_s independent of r. In the case $s < r < n$, we use Lemma 7.11(b) and consider the restriction of the random variable X in Corollary 7.13 to the appropriate subset. Using all this information we have

$$\beta_n = q^{-\binom{n}{2}}\alpha_n + \sum_{0\le s<r<n} q^{-\binom{r+1}{2}} q^s (q-1)\alpha_s.$$

□

Proof of (7). Taking equation (29) for n and $n-1$ and subtracting we have

$$\begin{aligned}\beta_n - \beta_{n-1} = q^{-\binom{n}{2}}\alpha_n - q^{-\binom{n-1}{2}}\alpha_{n-1} \\ + (q-1)\left[\sum_{0\le s<r<n} q^{-\binom{r+1}{2}} q^s\alpha_s - \sum_{0\le s<r<n-1} q^{-\binom{r+1}{2}} q^s\alpha_s\right].\end{aligned}$$

We observe that

$$\sum_{0\le s<r<n} q^{-\binom{r+1}{2}} q^s\alpha_s = q^{-\binom{n}{2}}\sum_{s=0}^{n-2} q^s\alpha_s + \sum_{0\le s<r<n-1} q^{-\binom{r+1}{2}} q^s\alpha_s,$$

therefore

$$q^{\binom{n}{2}}(\beta_n - \beta_{n-1}) = (\alpha_n - q^{n-1}\alpha_{n-1}) + (q-1)\sum_{s=0}^{n-2} q^s\alpha_s. \tag{30}$$

Now we take Equation (30) for n and $n-1$ and subtract which yields

$$\begin{aligned}q^{\binom{n}{2}}(\beta_n - \beta_{n-1}) - q^{\binom{n-1}{2}}(\beta_{n-1} - \beta_{n-2}) = (\alpha_n - q^{n-1}\alpha_{n-1}) \\ - (\alpha_{n-1} - q^{n-2}\alpha_{n-2}) \\ + (q-1)\sum_{s=0}^{n-2} q^s\alpha_s - (q-1)\sum_{s=0}^{n-3} q^s\alpha_s.\end{aligned}$$

We first simplify the right hand side

$$\begin{aligned}\alpha_n - \alpha_{n-1} + q^{n-2}\alpha_{n-2} - q^{n-1}\alpha_{n-1} + q^{n-1}\alpha_{n-2} - q^{n-2}\alpha_{n-2} = \\ = (\alpha_n - \alpha_{n-1}) - q^{n-1}(\alpha_{n-1} - \alpha_{n-2}),\end{aligned}$$

and now noting $\binom{n-1}{2} = \binom{n}{2} + (1-n)$ we also simplify left hand side

$$\begin{aligned}q^{\binom{n}{2}}(\beta_n - \beta_{n-1}) - q^{\binom{n}{2}} q^{1-n}(\beta_{n-1} - \beta_{n-2}) = \\ = q^{\binom{n}{2}}[(\beta_n - \beta_{n-1}) - q^{1-n}(\beta_{n-1} - \beta_{n-2})].\end{aligned}$$

Therefore we have

$$q^{\binom{n}{2}}[(\beta_n - \beta_{n-1}) - q^{1-n}(\beta_{n-1} - \beta_{n-2})] = (\alpha_n - \alpha_{n-1}) - q^{n-1}(\alpha_{n-1} - \alpha_{n-2}),$$

where this equality implies equality of the coefficients of t^n on both sides of (7).

□

References

[1] Tom M Apostol. *Introduction to analytic number theory*. Springer Science & Business Media, 2013.

[2] Manjul Bhargava, John Cremona, Tom Fisher, and Stevan Gajović. The density of polynomials of degree n over $\mathbb{Z}_p$ having exactly r roots in $\mathbb{Q}_p$. *Proceedings of the London Mathematical Society*, 124(5):713–736, 2022.

[3] Nicolas Bourbaki. *Variétés différentielles et analytiques: fascicule de résultats*, volume 8. Springer Science & Business Media, 2007.

[4] D.L. Cohn. *Measure Theory: Second Edition*. Birkhäuser Advanced Texts Basler Lehrbücher. Springer New York, 2013.

[5] D.S. Dummit and R.M. Foote. *Abstract Algebra*. Wiley, 2003.

[6] Ronald L Graham, Donald E Knuth, and Oren Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley Professional, 1994.

[7] J. Neukirch and N. Schappacher. *Algebraic Number Theory*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2013.

[8] W.A. Sutherland. *Introduction to Metric and Topological Spaces*. An open university set book. Clarendon Press, 1975.