Statistics about polynomials over finite fields

Mario Alegre, Pedro Juarez, and Hani Pajela

September 29, 2015

Abstract

A recent paper of Jason Fulman uses generating functions to prove enumerative results about square-free polynomials over finite fields. These results had been previously proved in a paper of Church, Ellenberg, and Farb using topology and representation theory. In this paper, we use similar generating function methods to calculate various statistics about how polynomials over finite fields factorize into a product of irreducibles.

Contents

1 Introduction

1.1 Overview

As stated in the abstract, we will utilize generating functions to come up with enumerative results of polynomials in finite fields. In some cases, we will provide explicit numbers whereas, in others, we will consider limits. The two main types of polynomials that this paper is concerned with are m-th-power-free polynomials and then later m-th-power-free polynomials that are nonvanishing at 0.

The next section will define the zeta function as well as provide preliminary counting results that serve as somewhat of a warm up for the general method that will be used throughout the paper. We will consider the number of several types of basic polynomals. The third section will be focused on m-th-power-free polynomials. Beginning with considering the factors of square-free polynomials, we will build a generalization and consider the factors of mth-power-free polynomials. The last section is concerned with m-th-power-free polynomials that are nonvanishing at 0 and quadratic residues. We will explore quadratic residues over the algebraic closure of finite fields, provide statistics about the aforementioned type of polynomials, and count the number of irreducibles with quadratic residue roots.

Before beginning, we would like to touch on the motivation for this endeavor as well as provide a brief explanation of the method.

1.2 Motivation

For a field k, consider the set $\text{Poly}_n(k) = \{f \in k[x] \mid f \textit{ is square free, } \deg(f) = n\}$, the set of all squarefree polynomials in $k[x]$ of degree n. When $k = \mathbb{C}$, we can give this set a natural topology in the following way. Every polynomial in this set has exactly n complex roots by the fundamental theorem of algebra, which must be distinct because of the square free condition. Instead of thinking about them as polynomials, we can think of them as sets of n distinct complex numbers (its roots). Hence we can consider the following subset of $\mathbb{C}^n$: $\text{Conf}_n(\mathbb{C}) = \{(x_1, \ldots, x_n) \in \mathbb{C}^n \mid x_i \neq x_j \ \forall \ i \neq j\}$. This is the set of *ordered* n-tuples of distinct complex numbers, corresponding to one squarefree polynomial of degree n. We give this space the subspace topology inherited from $\mathbb{C}^n$, and then quotient out by the group action of the permutation group S_n on the n coordinates. This turns the space into a set of all *unordered* n-tuples of distinct complex numbers, which is in a bijective correspondence with the set $\text{Poly}_n(\mathbb{C})$ by associating each n-tuple to the polynomial having all of them as roots. Thus we have turned $\text{Poly}_n(\mathbb{C})$ into a topological space via the quotient topology on $\text{Conf}_n(\mathbb{C})$.

Given a representation V of S_n (that is, V a $\mathbb{Q}$-vector space with a group homomorphism $\phi : S_n \to GL(V)$), we can associate a local coefficient system on $\text{Poly}_n(\mathbb{C})$ and ask about its homology. The Grothendieck-Lefschetz trace formula gives us the

following result (Eq. 14 in [1]):

$$\sum_{f \in \mathrm{Poly}_n(\mathbb{F}_q)} \mathrm{Tr}(\phi(\sigma_f)) = q^n \sum_{i=0}^{\infty} (-1)^i \dim H^i(\mathrm{Poly}_n(\mathbb{C}); V) q^{-i}$$

where σ_f is the permutation obtained by the action of the Frobenius map on the set of n roots of f in $\overline{\mathbb{F}_q}$.

Example 1. We can consider the trivial representation, $V = \mathbb{Q}$. The only homomorphism from S_n into $GL(\mathbb{Q})$ is the trivial one, hence the left hand side of the equation simply counts the size of the set $\mathrm{Poly}_n(\mathbb{F}_q)$. That is,

$$|\mathrm{Poly}_n(\mathbb{F}_q)| = q^n \sum_{i=0}^{\infty} (-1)^i \dim H^i(\mathrm{Poly}_n(\mathbb{C}); \mathbb{Q}) q^{-i}$$

Example 2. For the representation $V = \mathbb{Q}^n$ where S_n acts on $\mathbb{Q}^n$ by permuting the coordinates, the trace of any $\phi(\sigma(f))$ is the number of roots which are fixed points of the Frobenius map $z \to z^q$. An element θ is fixed by Frobenius iff $\theta \in \mathbb{F}_q$, hence the formula becomes:

$$\sum_{f \in \mathrm{Poly}_n(\mathbb{F}_q)} (\textit{number of linear factors of } f) = q^n \sum_{i=0}^{\infty} (-1)^i \dim H^i(\mathrm{Poly}_n(\mathbb{C}); \mathbb{Q}^n) q^{-i}$$

In both of these equations, the left hand side is a statistic about a finite set, which is much easier to compute than these homology groups. Section 3 will focus on counting these types of statistics which give information about $\mathrm{Poly}_n(\mathbb{C})$ and related spaces via this formula.

A more refined subset of $\mathrm{Poly}_n(k)$ is this: $B_n(k) = \{f \in \mathrm{Poly}_n(k) \mid f(0) \neq 0\}$. There is an analogous trace formula for the subsets $B_n(\mathbb{C})$ as follows: given a representation V of $(\mathbb{Z}/2\mathbb{Z})^n \rtimes S_n$, we have, for $q \neq 2$:

$$\sum_{f \in B_n(\mathbb{F}_q)} \mathrm{Tr}(\phi(\sigma_f)) = q^n \sum_{i=0}^{\infty} (-1)^i \dim H^i(B_n(\mathbb{C}); V) q^{-i}$$

where σ_f is the permutation obtained by the action of the Frobenius map on the set $\{\sqrt{\theta} \mid f(\theta) = 0\} \subset \overline{\mathbb{F}_q}$. By not including 0 as a root, we ensure that this set has precisely $2n$ elements.

Example 3. We can again consider the trivial representation, $V = \mathbb{Q}$. The only homomorphism from $(\mathbb{Z}/2\mathbb{Z})^n \rtimes S_n$ into $GL(\mathbb{Q})$ is again the trivial one, hence the left hand side of the equation simply counts the size of the set $B_n(\mathbb{F}_q)$. That is,

$$|B_n(\mathbb{F}_q)| = q^n \sum_{i=0}^{\infty} (-1)^i \dim H^i(B_n(\mathbb{C}); \mathbb{Q}) q^{-i}$$

Example 4. We can also consider the representation $V = \mathbb{Q}^{2n}$, with $(\mathbb{Z}/2\mathbb{Z})^n \rtimes S_n$ acting on $\mathbb{Q}^{2n}$ by permuting pairs of coordinates. In this case, the trace of any $\phi(\sigma(f))$ is twice the number of roots whose square roots are also in $\mathbb{F}_q$. That is,

$$\sum_{f \in B_n(\mathbb{F}_q)} 2\cdot(\textit{number of quadratic residue roots of } f) = q^n \sum_{i=0}^{\infty} (-1)^i \dim H^i(B_n(\mathbb{C}); \mathbb{Q}^{2n}) q^{-i}$$

Once again the left hand sides are statistics about finite sets. Section 4 will focus on counting these types of statistics for $B_n(\mathbb{C})$ and related spaces. The left hand side of all these examples are numbers that depend on the value on n. Instead of studying these numbers for particular n, we will study them all at once as a sequence c_n via the method of generating functions.

1.3 Method: generating functions

Given any sequence c_n, we can associate a *generating function* $A(t) = \sum_{i=0}^{\infty} c_n t^n$. This is a formal sum, so convergence in the analytical sense is not an issue. As a result, we will never be substituting a value in for t, but rather keep it as variable to track the different numbers in the sequence c_n. A certain generating function will be used so regularly throughout this paper that we will define a useful shorthand for it.

Definition 1.

$$\sum_{i=0}^{\infty} t^n = \frac{1}{1-t}$$

It can be checked that this definition is well-behaved with elementary addition, multiplication and term by term differentiation. We will use this fact to switch between the infinite power series and its shorthand notation, which is easier to manipulate algebraically. Thus we will use this method of generating functions in this paper to study all the sequences of interest.

2 Counting polynomials using the zeta function

In this section, we will define the zeta function which is an integral part of many of the calculations within the paper. It is important to note that polynomials over finite fields are analogous to integers. Most importantly, monic polynomials are analogs of positive integers and irreducibles are analogs of prime numbers. Using this fact, we can derive an analog of Euler's Product Formula for Riemann's zeta Function.

Using the zeta function and the deduced Euler's Product Formula, we will derive generating functions for as well as calculate the following:

- the number of square-free degree n polynomials over $\mathbb{F}_q$
- the number of mth-degree-free degree n polynomials over $\mathbb{F}_q$

- the number of mth-degree-free degree n polynomials over $\mathbb{F}_q$ that are nonvanishing at 0
- the number of monic irreducibles of degree k over $\mathbb{F}_q$

The characteristics of the polynomials in question will be further elaborated on in their respective subsections.

2.1 The zeta function

Definition 2. Let us define the zeta Function for the affine line $\mathbb{A}$ as

$$Z(\mathbb{A}, t) = \sum_{n=0}^{\infty} (\#\text{ of monic degree } n \text{ polynomials in } \mathbb{F}_q) t^n$$

where $\mathbb{F}_q$ is the finite field of size q. Note that a polynomial satisfying these conditions has the form $x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0$ where each $a_i \in \mathbb{F}_q$. Therefore, the number of degree n polynomials in $\mathbb{F}_q$ is q^n. It then follows that

$$Z(\mathbb{A}, t) = \sum_{n=0}^{\infty} q^n t^n = \frac{1}{1-qt}$$

Remark 1. From here on, will use $Z(t)$ to denote the zeta function.

This zeta function can be easily manipulated and altered to give generating functions of interest to this paper. In fact, this function is featured in all succeeding subsections.

Lemma 1.

$$Z(t) = \prod_{p(x)} (1 + t^{\deg(p(x))} + t^{2\deg(p(x))} + ...)$$

where the product is taken over all monic, irreducible polynomials $p(x)$ in $\mathbb{F}_q$.

Proof. Since every $f \in \mathbb{F}_q[x]$ can be factorized into a product of irreducible polynomials, we can deduce an analog of Euler's Product Formula for the Riemann zeta function:

$$\begin{aligned} Z(t) &= \sum_{n=0}^{\infty} (\#\text{ of monic degree } n \text{ polynomials in } \mathbb{F}_q) t^n \\ &= \prod_{p(x)} (1 + t^{\deg(p(x))} + t^{2\deg(p(x))} + ...) \end{aligned}$$

where the product is taken over all monic, irreducible polynomials $p(x)$ in $\mathbb{F}_q$. □

Remark 2. Unless stated otherwise, we will assume that $p(x)$ is monic and irreducible.

Combining the definition and the lemma above, we see that:

$$Z(t) = \frac{1}{1-qt} = \prod_{p(x)} (1 + t^{\deg(p(x))} + t^{2\deg(p(x))} + ...)$$

2.2 Square-free polynomials

In this subsection, we will consider the number of polynomials over $\mathbb{F}_q$ that are square-free, i.e. every root is distinct. The content of this section is included in Fulman's paper, and the results were well known before Fulman as well, but we have chosen to include it as a warm-up and introduction to the methods that will be used throughout the paper.

Definition 3. Let $\text{Poly}_n(\mathbb{F}_q)$ denote the set of monic square-free degree n polynomials.

Theorem 2. *The number of monic square-free degree n polynomials over $\mathbb{F}_q$ is given by:*

$$|\,\text{Poly}_n(\mathbb{F}_q)| = \begin{cases} q^n - q^{n-1} & \text{if } n \geq 2 \\ q & \text{if } n = 1 \\ 1 & \text{if } n = 0 \end{cases}$$

Proof. To begin, let us define:

$$\begin{aligned} \Phi(t) &= \sum_{n=0}^{\infty} |\,\text{Poly}_n(\mathbb{F}_q)| t^n \\ &= \prod_{p(x)} (1 + t^{\deg(p(x))}) \end{aligned}$$

The second equality holds because there is no multiplicity. The expression is restricting the polynomial of Euler's Product Rule in order to count only the square-free polynomials. Continuing the calculation, we see that

$$\begin{aligned} &= \prod_{p(x)} \frac{1 - t^{2\deg(p(x))}}{1 - t^{\deg(p(x))}} \\ &= \frac{\prod \frac{1}{1-(t^{\deg(p(x))})}}{\prod \frac{1}{1-(t^{2\deg(p(x))})}} \\ &= \frac{\zeta(t)}{\zeta(t^2)} \\ &= \frac{1 - qt^2}{1 - qt} \\ &= (1 - qt^2)(1 + qt + q^2t^2 + q^3t^3 + ...) \\ &= 1 + qt + (q^2 - q)t^2 + (q^3 - q^2)t^3 + ... \end{aligned}$$

It then follows that:

$$|\,\text{Poly}_n(\mathbb{F}_q)| = \begin{cases} q^n - q^{n-1} & \text{if } n \geq 2 \\ q & \text{if } n = 1 \\ 1 & \text{if } n = 0 \end{cases}$$

□

Corollary 3. *The probability that a randomly selected polynomial is square-free is*

$$\frac{\#\ of\ square\text{-}free\ degree\ n\ polys}{\#\ of\ degree\ n\ polys} = \begin{cases} 1 - \frac{1}{q} & if\ n \geq 2 \\ 1 & if\ n = 0\ or\ 1 \end{cases}$$

Remark 3. Observe that statistic in question here is the probability that a randomly selected polynomial is square-free. However, there are infinitely many polynomials. In order for this probability to be well-defined, we define it as: $\frac{\#\text{ of square-free degree } n \text{ polynomials}}{\#\text{ of degree } n \text{ polynomials}}$. Other probabilities mentioned in this section are defined similarly.

Proof. This corollary follows immediately from the theorem above after recalling that the number of monic polynomials over $\mathbb{F}_q$ is q^n. □

The rest of the paper consists of new results, except where stated otherwise.

2.3 mth-power-free polynomials

Now, we will generalize the result found in the previous subsection by considering the number of polynomials over $\mathbb{F}_q$ that are mth-power-free, i.e. every root has multiplicity of at most $m - 1$.

Definition 4. Let $\mathrm{Poly}_n^m(\mathbb{F}_q)$ denote the set of mth-power-free degree n polynomials.

Theorem 4. *The number of monic square-free degree n polynomials over $\mathbb{F}_q$ is given by:*

$$|\,\mathrm{Poly}_n^m(\mathbb{F}_q)| = \begin{cases} q^n - q^{n-1} & if\ n \geq m \\ q^n & if\ 0 < n < m \\ 1 & if\ n = 0 \end{cases}$$

Proof. First, observe that:

$$\begin{aligned} \Phi(t) &= \sum_{n=0}^{\infty} |\,\mathrm{Poly}_n^m(\mathbb{F}_q)| t^n \\ &= \prod_{p(x)} (1 + t^{\deg(p(x))} + ... + t^{(m-1)\deg(p(x))}) \end{aligned}$$

The second equality holds because the expression is restricting the polynomial of Euler's Product Rule in order to count only polynomials with roots that have multiplicity

of at most $m-1$. Continuing the calculation, we see that

$$\begin{aligned}
&= \prod_{p(x)} \frac{1-t^{m\deg(p(x))}}{1-t^{\deg(p(x))}} \\
&= \frac{\prod \frac{1}{1-t^{\deg(p(x))}}}{\prod \frac{1}{1-t^{m\deg(p(x))}}} \\
&= \frac{\zeta(t)}{\zeta(t^m)} \\
&= \frac{1-qt^m}{1-qt} \\
&= (1-qt^m)(1+qt+q^2t^2+q^3t^3+...) \\
&= 1+qt+q^2t^2+...+q^{m-1}t^{m-1}+(q^m-q^{m-1})t^m+(q^{m+1}-q^m)t^{m+1}+...
\end{aligned}$$

It then follows that:

$$|\operatorname{Poly}_n^m(\mathbb{F}_q)| = \begin{cases} q^n - q^{n-1} & \text{if } n \geq m \\ q^n & \text{if } 0 < n < m \\ 1 & \text{if } n = 0 \end{cases}$$

□

Corollary 5. *The probability of a randomly selected polynomial being mth-power-free is*

$$\frac{\#\ of\ square\text{-}free\ degree\ n\ polys}{\#\ of\ degree\ n\ polys} = \begin{cases} 1-\frac{1}{q^{m-1}} & if\ n \geq m \\ 1 & if\ n < m \end{cases}$$

2.4 mth-power-free polynomials that are nonvanishing at 0

This subsection is concerned with mth-power-free (as above) as well as nonvanishing at 0 polynomials in $\mathbb{F}_q$, i.e. $f(0) \neq 0$.

Definition 5. Let $B_n^m(\mathbb{F}_q)$ denote the set of mth-power-free and nonvanishing at 0 polynomials on $\mathbb{F}_q$

Theorem 6. *The generating function for $|B_n^m(\mathbb{F}_q)|$ is given by*

$$\frac{t-1}{t^m-1}(1+qt+...+q^{m-1}t^{m-1}+(q^m-q^{m-1})t^m+(q^{m+1}-q^m)t^{m+1}+...)$$

Proof. Consider then

$$\begin{aligned}
\Phi(t) &= \sum_{n=0}^{\infty} |B_n^m(\mathbb{F}_q)|t^n \\
&= \prod_{p(x)\neq 0} 1+t^{\deg(p(x))}+t^{2\deg(p(x))}+...+t^{(m-1)\deg(p(x))}
\end{aligned}$$

Note that the product is taken over monic and irreducible polynomials s.t. $p(x) \neq x$. This ensures that only polynomials that are nonvanishing at 0 are accounted for. It then follows that

$$= \frac{1}{1+t+...+t^{m-1}} \prod_{p(x)} \frac{1-t^{m \deg(p(x))}}{1-t^{\deg(p(x))}}$$
$$= \frac{1}{1+t+...+t^{m-1}} \frac{\zeta(t)}{\zeta(t^m)}$$
$$= \frac{t-1}{t^m-1} \frac{1-qt^m}{1-qt}$$
$$= \frac{t-1}{t^m-1}(1+qt+q^2t^2+...+q^{m-1}t^{m-1}+(q^m-q^{m-1})t^m+(q^{m+1}-q^m)t^{m+1}+...)$$

□

2.5 Number of Monic Irreducibles

Recall that an irreducible is a polynomial that cannot be factored any further. As shown in the above calculations (as well as throughout the entirety of the paper), monic irreducibles are integral in computing statistics of polynomials in finite fields. In fact, the enumerative results found in the next section depend on the number of such polynomials. We will now consider the number of degree k monic irreducibles over $\mathbb{F}_q$.

Definition 6. Let $N_k(q)$ denote the number of degree k monic irreducible polynomials over $\mathbb{F}_q$.

This statistic is also well-known, but we have chosen to include it as further practice of the generating function method we are using.
To begin, recall that

$$\zeta_A(t) = \sum q^n t^n = \frac{1}{1-qt} = \prod_{p(x)} \frac{1}{1-t^{\deg(p(x))}}$$

where $p(x)$ is a monic irreducible polynomial in $\mathbb{F}_q$.

Lemma 7.

$$q^k = \sum_{l|k} l N_l(q)$$

Proof. By definition,

$$Z(t) = \frac{1}{1-qt} \tag{2.1}$$

Then, by Lemma 1, $Z(t) = \prod_{p(x)} \frac{1}{1-t^{\deg(p(x))}}$. Note that we can group the product:

$$Z(t) = \frac{1}{1-qt} \prod_{l=1}^{\infty} \left(\frac{1}{1-t^l}\right)^{N_l(q)} \tag{2.2}$$

Let us now consider the two above equations when we apply $\frac{d}{dt}\log$ to them:

$$\begin{aligned}
\frac{d}{dt}\log Z(t) &= \frac{d}{dt}\log\left(\frac{1}{1-qt}\right) \\
&= \frac{d}{dt} - \log(1-qt) \\
&= -\frac{-q}{1-qt} \\
&= \frac{q}{1-qt} \\
\frac{d}{dt}\log Z(t) &= \frac{d}{dt}\log\frac{1}{1-qt}\prod_{l=1}^{\infty}\left(\frac{1}{1-t^l}\right)^{N_l(q)} \\
&= \frac{d}{dt}\sum_{l=1}^{\infty} N_l(q)\log\left(\frac{1}{1-t^l}\right) \\
&= \sum_{l=1}^{\infty} N_l(q)\frac{d}{dt}\log\left(\frac{1}{1-t^l}\right) \\
&= -\sum_{l=1}^{\infty} N_l(q)\frac{d}{dt}\log(1-t^l) \\
&= \sum_{l=1}^{\infty} N_l(q)\frac{lt^{l-1}}{1-t^l}
\end{aligned}$$

Therefore, it then follows that

$$\begin{aligned}
\frac{q}{1-qt} &= \sum_{l=1}^{\infty} N_l(q)\frac{lt^{l-1}}{1-t^l} \\
\frac{qt}{1-qt} &= \sum_{l=1}^{\infty} N_l(q)\frac{lt^{l}}{1-t^l} \\
\sum_{k=1}^{\infty} q^k t^k &= \sum_{l=1}^{\infty} lN_l(q)\sum_{m=1}^{\infty} t^{lm}
\end{aligned}$$

Let us now only consider the right hand side of the previous equation:

$$\text{R.H.S.} = \sum_{l=1}^{\infty}\sum_{m=1}^{\infty} lN_l(q)t^{lm} = \sum_{k=1}^{\infty}\left(\sum_{l|k}^{\infty} lN_l(q)\right)t^k$$

The leftmost equation holds because $k = lm \Leftrightarrow l|m$. Hence

$$\sum_{k=1}^{\infty} q^k t^k = \sum_{k=1}^{\infty}\left(\sum_{l|k}^{\infty} lN_l(q)\right)t^k \implies q^k t^k = \left(\sum_{l|k}^{\infty} lN_l(q)\right)t^k \implies q^k = \sum_{l|k} lN_l(q)$$

□

Definition 7. The Möbius Function $\mu : \mathbb{N} \to \{-1, 0, 1\}$ is defined as follows

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n \text{ is not square-free} \\ (-1)^k & \text{if } n \text{ is a product of } k \text{ primes} \end{cases}$$

As the following theorem uses it, stated below is the Möbius Inversion Formula. Since this is a known formula, no proof will be provided.

Lemma 8. *Suppose $g : \mathbb{N} \to \mathbb{C}$. Let $f(n) := \sum_{d|n} g(d)$. Then,*

$$g(n) = \sum_{d|n} f(d)\mu\left(\frac{n}{d}\right) = \sum_{d|n} f\left(\frac{n}{d}\right)\mu(d)$$

Theorem 9. *Let μ be the Möbius Function*

$$N_k(q) = \frac{1}{k}\sum_{l|k} \mu\binom{l}{n} q^l$$

Proof. Let $f(l) = q^l = \sum_{l|k} lN_l(q)$. Hence, $g(l) = lN_l(q)$. It directly follows from the lemmas that

$$kN_k(q) = \sum_{l|k} \mu\left(\frac{k}{l}\right) q^l \implies N_k(q) = \frac{1}{k}\sum_{l|k} \mu\binom{l}{n} q^l$$

□

3 Irreducible factors of polynomials

In the previous section, we counted the number of polynomials that met certain conditions. In this section we will calculate more refined statistics about these polynomials. Specifically, we will calculate the expected number of irreducible factors of degree k of a randomly selected degree n square-free polynomial. We will then calculate a generating function for an arbitrary combination of factors of different degrees in a randomly selected square-free polynomial. Finally, we will seek to calculate the same statistic for a general mth-power-free polynomial. In cases where calculating the expected value is not feasible, we will examine the behavior of our statistic as n goes to infinity. With this goal in mind, we will use the following definitions:

Definition 8. For any $x \in \mathbb{F}_q$, define $\deg x :=$ deg(minimal polynomial of x).

Remark 4. This definition of degree is equivalent to the more useful definition $\deg x = \min\{i | x \in \mathbb{F}_{q^i}\}$. That is, $\deg x$ is the least i such that $x \in \mathbb{F}_{q^i}$. This result is stated without proof.

Definition 9. For any $f \in \mathrm{Poly}_n^m(\mathbb{F}_q)$ define $\mathrm{n_k}(f)$ as the number of distinct degree k irreducible factors of f.

Remark 5. Note that $\mathrm{n_k}(f) = \frac{1}{k} \cdot$ (the number of degree k roots of f).

3.1 Factors of square-free polynomials

The simplest case we will consider is calculating the expected number of degree k factors of a polynomial in $\mathrm{Poly}_n(\mathbb{F}_q)$. We will begin by calculating a generating function for the sequence of $n_k(f)$, and use this generating function to calculate the expected number.

Definition 10. We define the sequence b_n as

$$b_n := \sum_{f \in \mathrm{Poly}_n(\mathbb{F}_q)} \mathrm{n_k}(f)$$

and the generating function of this sequence as

$$B(t) = \sum_{n=0}^{\infty} b_n t^n$$

Now that we have defined our problem in terms of a series, to find b_n we must find $B(t)$.

Proposition 10. *The generating function $B(t)$ of b_n is given by the formula*

$$B(t) = N_k \frac{1 - qt^2}{1 - qt} t^k$$

Proof. Let $\Psi(x,t)$ be defined as

$$\Psi(x,t) = \sum_{n=0}^{\infty} \Big(\sum_{f \in \mathrm{Poly}_n(\mathbb{F}_q)} x^{\mathrm{n_k}(f)} \Big) t^n$$

We can then easily see that

$$\begin{aligned} \frac{\partial \Psi(x,t)}{\partial x}\Big|_{x=1} &= \sum_{n=0}^{\infty} \Big(\sum_{f \in \mathrm{Poly}_n(\mathbb{F}_q)} \mathrm{n_k}(f) \Big) t^n \\ &= \sum_{n=0}^{\infty} b_n t^n \\ &= B(t) \end{aligned}$$

However, if we let $\chi(p(x))$ be defined for any $p(x) \in \mathbb{F}_p[x]$ as

$$\chi(q(x)) = \begin{cases} x & \text{if } \deg p(x) = k \\ 1 & \text{otherwise} \end{cases}$$

Then we can use the Euler Product Rule to obtain

$$\begin{aligned} \Psi(x,t) &= \sum_{n=0}^{\infty} \Big(\sum_{f \in \mathrm{Poly}_n(\mathbb{F}_q)} x^{\mathrm{n_k}(f)} \Big) t^n \\ &= \prod_{p(x)} (1 + \chi(p) t^{\deg p(x)}) \end{aligned}$$

Recall that the polynomial $p(x)$ is monic and irreducible, unless stated otherwise.

$$\begin{aligned}
&= \prod_{\deg p(x)\neq k} (1+t^{\deg p(x)}) \prod_{\deg p(x)=k} (1+xt^{\deg p(x)}) \\
&= \frac{1}{(1+t^k)^{N_k}} \prod_{p(x)} (1+t^{\deg p(x)}) \prod_{\deg p(x)=k} (1+xt^k)
\end{aligned}$$

Note that the value of the first product was already calculated in section 2.2.

$$\begin{aligned}
&= \frac{1}{(1+t^k)^{N_k}} \cdot \frac{1-qt^2}{1-qt} \cdot (1+xt^k)^{N_k} \\
&= \frac{1-qt^2}{1-qt} \Big(\frac{1+xt^k}{1+t^k} \Big)^{N_k}
\end{aligned}$$

Now that we have simplified $\Psi(x,t)$, we take its derivative with respect to x evaluated at $x=1$ to obtain the generating function $B(t)$. Recall that N_k represents the number of monic irreducible polynomials in $\mathbb{F}_q$ of degree i. A formula for N_k was dicussed in section 2.5.

$$\begin{aligned}
B(t) &= \frac{\partial \Psi(x,t)}{\partial x}\Big|_{x=1} \\
&= \frac{1-qt^2}{1-qt} \cdot \frac{\partial}{\partial x} \Big[\frac{(1+xt^k)^{N_k}}{(1+t^k)^{N_k}} \Big]_{x=0} \\
&= \frac{1-qt^2}{1-qt} \cdot \frac{N_k t^k (1+t^k)^{N_k}}{(1+t^k)^{N_k}} \\
&= N_k \frac{1-qt^2}{1-qt} t^k
\end{aligned}$$

□

Using this generating function, we can now find a formula for b_n.

Corollary 11. *For a given k, $b_k = N_k$, $b_{k+1} = qN_k$, and $b_i = N_k(q^{i-k} - q^{i-k-1})$ for $i \geq k+2$.*

Proof. Since the generating function is a formal power series, we can assume that

$$\frac{1}{1+qt} = 1 + qt + q^2t^2 + \dots$$

is true for any t. This gives us that the generating function $B(t)$ is

$$\begin{aligned}
B(t) &= N_k \frac{1-qt^2}{1-qt} t^k \\
&= (N_k t^k - N_k q t^{k+2})(1 + qt + q^2t^2 + \dots) \\
&= N_k t^k + N_k q t^{k+1} + N_k(q^2 - q)t^{k+2} + N_k(q^3 - q^2)t^{k+3} + \dots
\end{aligned}$$

Since b_i is the coefficient of t^i, we can easily see that

$$b_i = \begin{cases} 0 & : i < k \\ N_k & : i = k \\ qN_k & : i = k+1 \\ N_k(q^{i-k} - q^{i-k-1}) & : i \geq k+2 \end{cases}$$

□

Given a formula for b_i, we can now calculate the expected value of $\mathrm{n_k}(f)$ for a randomly selected $f \in \mathrm{Poly}_n(\mathbb{F}_q)$.

Theorem 12. *For a randomly selected $f \in \mathrm{Poly}_n(\mathbb{F}_q)$, the expected number of degree k irreducible factors of f does not depend on n, and is given by the expression $\frac{N_k}{q^{n-1}-q^{k-2}}$ as long as $k \leq n-2$.*

Proof. Since the expected value of a function is given by the sum of the function's values over all elements in the set divided by the total size of the set, we know that for a given $n \in \mathbb{N}$

$$E(\mathrm{n_k}) = \frac{b_n}{|\mathrm{Poly}_n(\mathbb{F}_q)|}$$

Recall that $\mathrm{Poly}_n(\mathbb{F}_q)$ is the set of monic square-free polynomials in $\mathbb{F}_q$. In section 2.2 we showed that the number of monic square-free polynomials is $q^n - q^{n-1}$ for $n \geq 2$. At this point we are forced to split our result into several cases.
If $k \leq n-2$, then, since k must be at least 1, we can guarantee that $n \geq 2$. So our expected value is

$$E(\mathrm{n_k}) = \frac{N_k(q^{n-k} - q^{n-k-1})}{q^n - q^{n-1}} = \frac{N_k}{q^k - q^{k-1}}$$

If $k = n-1$, then we can still guarantee that $n \geq 2$. So our expected value is

$$E(\mathrm{n_k}) = \frac{qN_k}{q^n - q^{n-1}} = \frac{N_k}{q^{n-1} - q^{n-2}}$$

For $k = n$, we will ignore the degenerate case of $n = 1$. Thus, we have

$$E(\mathrm{n_k}) = \frac{N_k}{q^n - q^{n-1}}$$

Note that for a polynomial of degree n, the expected number of factors of degree less than $n-2$ does not depend on n. □

3.2 Generalization

We now move on to consider a more general version of the above problem. Let $\Lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l)$ be a random l-tuple of naturals. Instead of considering the expected number of factors for a single k, we will consider the probability that a

polynomial has λ_i factors of degree i. Once again, we will first calculate a generating function for this statistic. Since calculating an explicit expected value for this statistic is unfeasible, we will instead calculate the limit as n goes to infinity of the expected value.

Definition 11. We define the sequence c_n as

$$c_n := \sum_{f \in \mathrm{Poly}_n(\mathbb{F}_q)} \binom{\mathrm{n}_1(f)}{\lambda_1} \binom{\mathrm{n}_2(f)}{\lambda_2} \cdots \binom{\mathrm{n}_l(f)}{\lambda_l}$$

with attendant generating function

$$C(t) = \sum_{n=0}^{\infty} c_n t^n$$

Note that the sequence we considered in the previous section is a special case of this sequence, where $\lambda_1 = 1$ and $\lambda_i = 0$ for all $i > 1$.

Proposition 13. *The generating function $C(t)$ of c_n can is given by the formula*

$$C(t) = \frac{1 - qt^2}{1 - qt} \prod_{i=1}^{l} \binom{N_i}{\lambda_i} \Big(\frac{t^i}{1 + t^i} \Big)^{\lambda_i}$$

Proof. The general stucture of this proof is identical to the structure of the analogous proof in the previous section, and thus less explanation of individual steps will be provided. Let $\Psi(\mathbf{x}, t)$ be

$$\Psi(\mathbf{x}, t) = \sum_{n=0}^{\infty} \Big(\sum_{f \in \mathrm{Poly}_n(\mathbb{F}_q)} x_1^{\mathrm{n}_1(f)} x_2^{\mathrm{n}_2(f)} \cdots x_l^{\mathrm{n}_l(f)} \Big) t^n$$

It is a fairly straightforward task to show that

$$\frac{\partial^{(\lambda_1 + \lambda_2 + \cdots + \lambda_l)}}{\partial x_1^{\lambda_1} \partial x_2^{\lambda_2} \cdots \partial x_l^{\lambda_l}} \Big[\frac{\Psi(\mathbf{x}, t)}{\lambda_1! \lambda_2! \ldots \lambda_l!} \Big]_{x_1 = 0, x_2 = 0, \ldots, x_l = 0} = C(t)$$

This time, we define $\chi(p)$ as

$$\chi(q(x)) = \begin{cases} x_i & \text{if } \deg q(x) = i \leq l \\ 1 & \text{otherwise} \end{cases}$$

We once again use the Euler Product formula to obtain

$$\begin{aligned}
\Psi(\mathbf{x},t) &= \prod_{p(x)}(1+\chi(p)t^{\deg p(x)}) \\
&= \prod_{\deg p(x)>k}(1+t^{\deg p(x)})\prod_{i=1}^{l}(1+x_it^i)^{N_i} \\
&= \prod_{p(x)}(1+t^{\deg p(x)})\prod_{i=1}^{l}\Big(\frac{1+x_it^i}{1+t^i}\Big)^{N_i} \\
&= \frac{1-qt^2}{1-qt}\prod_{i=1}^{l}\Big(\frac{1+x_it^i}{1+t^i}\Big)^{N_i}
\end{aligned}$$

This in turn gives us

$$\begin{aligned}
C(t) &= \frac{1-qt^2}{1-qt}\prod_{i=1}^{l}\frac{1}{\lambda_i!}\frac{\partial_i^{\lambda}}{\partial x_i^{\lambda_i}}\Big[\frac{(1+x_it^i)^{N_i}}{(1+t^i)^{N_i}}\Big]_{x_1=0,x_2=0,\ldots,x_l=0} \\
&= \frac{1-qt^2}{1-qt}\prod_{i=1}^{l}\binom{N_i}{\lambda_i}\frac{t^{i\lambda_i}(1+t^i)^{N_i-\lambda_i}}{(1+t^i)^{N_i}} \\
&= \frac{1-qt^2}{1-qt}\prod_{i=1}^{l}\binom{N_i}{\lambda_i}\frac{t^{i\lambda_i}}{(1+t^i)^{\lambda_i}}
\end{aligned}$$

□

Since the generating function for c_n is considerably more complicated, and depends on the particular value of each λ_i, we were unable to find a general formula for c_n in terms of $\Lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l)$. However, we can still analyze the behavior of c_n as $n \to \infty$ using the following lemma.

Lemma 14. *Let a_n be a sequence, and $f(t) = \sum_{n=0}^{\infty} a_n t^n$ be the generating function for that sequence. If $f(t)$ can be written as $f(t) = \frac{g(t)}{1-t}$ where $g(t)$ converges at $t = 1$, then $\lim_{a\to\infty} a_n$ exists and is equal to $g(1)$.*

Proof. Given $f(t) = \frac{g(t)}{1-t}$, we can solve for g to get

$$\begin{aligned}
g(t) &= (1-t)f(t) \\
&= (1-t)\sum_{k=0}^{\infty} a_k t^k \\
&= \sum_{k=0}^{\infty} a_k t^k - \sum_{k=0}^{\infty} a_k t^{k+1} \\
&= a_0 + \sum_{k=1}^{\infty}(a_k - a_{k-1})t^k
\end{aligned}$$

Thus, $g(1)$ is

$$\begin{aligned} g(1) &= a_0 + \sum_{k=1}^{\infty}(a_k - a_{k-1}) \\ &= \lim_{n\to\infty} a_0 + \sum_{k=1}^{n}(a_k - a_{k-1}) \\ &= \lim_{n\to\infty} a_0 + (a_1 - a_0) + (a_2 - a_1) + \ldots + (a_n - a_{n-1}) \\ &= \lim_{n\to\infty} a_n \end{aligned}$$

□

Using this lemma, we can find the expected value of c_n as n goes to infinity.

Proposition 15. *The expected value of c_n as n goes to infinity is given by*

$$\lim_{n\to\infty} \frac{c_n}{|\mathrm{Poly}_n(\mathbb{F}_q)|} = \prod_{i=1}^{l} \binom{N_i}{\lambda_i} \Big(\frac{1}{1+q^i}\Big)^{\lambda_i}$$

Proof. First, will will split our limit into the quotient of two easier-to-calculate limits

$$\lim_{n\to\infty} \frac{c_n}{|\mathrm{Poly}_n(\mathbb{F}_q)|} = \frac{\lim_{n\to\infty} \frac{c_n}{q^n}}{\lim_{n\to\infty} \frac{|\mathrm{Poly}_n(\mathbb{F}_q)|}{q^n}}$$

So, we have

$$\begin{aligned} \lim_{n\to\infty} \frac{|\mathrm{Poly}_n(\mathbb{F}_q)|}{q^n} &= \lim_{n\to\infty} \frac{q^n - q^{n-1}}{q^n} \\ &= \lim_{n\to\infty} 1 - \frac{1}{p} \\ &= 1 - \frac{1}{p} \end{aligned}$$

Next, we note that the generating function $f(u)$ for $\frac{c_n}{q^n}$ is given by $C(\frac{u}{q})$

$$\begin{aligned} f(u) &= C(\frac{u}{q}) \\ &= \frac{1-\frac{u^2}{q}}{1-u} \prod_{i=1}^{l} \binom{N_i}{\lambda_i} \Big(\frac{(\frac{u}{q})^i}{1+(\frac{u}{q})^i}\Big)^{\lambda_i} \\ &= \frac{1-\frac{u^2}{q}}{1-u} \prod_{i=1}^{l} \binom{N_i}{\lambda_i} \Big(\frac{u^i}{q^i+u^i}\Big)^{\lambda_i} \end{aligned}$$

This in turn gives us that

$$g(u) = (1-u)f(u) = (1-\frac{u^2}{q}) \prod_{i=1}^{l} \binom{N_i}{\lambda_i} \Big(\frac{u^i}{q^i+u^i}\Big)^{\lambda_i}$$

and

$$g(1) = (1 - \frac{1}{q}) \prod_{i=1}^{l} \binom{N_i}{\lambda_i} \Big(\frac{1}{1+q^i}\Big)^{\lambda_i}$$

So, we can conbclude that

$$\lim_{n\to\infty} \frac{c_n}{|\mathrm{Poly}_n(\mathbb{F}_q)|} = \prod_{i=1}^{l} \binom{N_i}{\lambda_i} \Big(\frac{1}{1+q^i}\Big)^{\lambda_i}$$

□

3.3 Factors of mth-power-free polynomials

So far we have only considered these statistics in relation to square-free polynomials. We will now move on to general mth-power-free polynomials. As this is a generalization of the previous problem, many of the concepts and even the general structure of proofs remains similar to the previous two subsections.

Remark 6. Recall that $\mathrm{Poly}_n^m(\mathbb{F}_q)$ is the set of mth-power-free degree n polynomials in $\mathbb{F}_q$, and $\mathrm{n_k}(f)$ is the number of *distinct* degree k irreducible factors of f. Now that we are no longer considering exclusively square-free polynomials, we may have cases where a certain factor is repeated multiple times. In this case, $\mathrm{n_k}$ counts only the number of unique degree k factors of f. For example, $\mathrm{n_1}(f) = 2$ for $f(x) = (x-1)^2(x+1)^2$.

Definition 12. Recall that $\mathrm{Poly}_n^m(\mathbb{F}_q)$ denotes the set of monic mth-power-free polynomials of degree n over $\mathbb{F}_q$. We define the sequence d_n as

$$d_n := \sum_{f\in\mathrm{Poly}_n^m(\mathbb{F}_q)} \binom{\mathrm{n_1}(f)}{\lambda_1} \binom{\mathrm{n_2}(f)}{\lambda_2} \cdots \binom{\mathrm{n_l}(f)}{\lambda_l}$$

with attendant generating function

$$D(t) = \sum_{n=0}^{\infty} d_n t^n$$

Note that the previous two sections are a special case of this section, where $m = 2$.

Proposition 16.

$$D(t) = \frac{1-qt^m}{1-qt} \prod_{i=1}^{l} \binom{N_i}{\lambda_i} \Big(\frac{t^{im}-t^i}{t^{im}-1}\Big)^{\lambda_i}$$

Proof. Let $\Psi(\mathbf{x}, t)$ be

$$\Psi(\mathbf{x}, t) = \sum_{n=0}^{\infty} \Big(\sum_{f\in\mathrm{Poly}_n(\mathbb{F}_q)} x_1^{\mathrm{n_1}(f)} x_2^{\mathrm{n_2}(f)} \cdots x_l^{\mathrm{n_l}(f)} \Big) t^n$$

Once again, it can be seen that

$$\frac{\partial^{(\lambda_1+\lambda_2+\cdots+\lambda_l)}}{\partial x_1^{\lambda_1}\partial x_2^{\lambda_2}\cdots\partial x_l^{\lambda_l}}\Big[\frac{\Psi(\mathbf{x},t)}{\lambda_1!\lambda_2!\ldots\lambda_l!}\Big]_{x_1=0,x_2=0,\ldots,x_l=0} = D(t)$$

Let $\chi(p)$ be

$$\chi(q(x)) = \begin{cases} x_i & \text{if } \deg q(x) = i \leq k \\ 1 & \text{otherwise} \end{cases}$$

It is at this point that the proof stops being identical to the one in section 3.2. Since our polynomials are no longer square-free, but rather mth-power-free, we must consider more terms applying the Euler Product Formula. Thus, we get

$$\begin{aligned}\Psi(\mathbf{x},t) &= \prod_{p(x)} 1+\chi(p)(t^{\deg p}+t^{2\deg p}+\ldots+t^{(m-1)\deg p}) \\ &= \prod_{p(x)}(1+t^{\deg p}+t^{2\deg p}+\ldots+t^{(m-1)\deg p})\prod_{i=1}^{l}\Big(\frac{1+\chi(p)(t^i+t^{2i}+\ldots+t^{(m-1)i})}{1+t^i+t^{2i}+\ldots+t^{(m-1)i}}\Big)^{N_i} \\ &= \frac{1-pt^m}{1-pt}\prod_{i=1}^{l}\Big(\frac{1+x_i(t^i+t^{2i}+\ldots+t^{(m-1)i})}{1+t^i+t^{2i}+\ldots+t^{(m-1)i}}\Big)^{N_i}\end{aligned}$$

This in turn gives us

$$\begin{aligned}C(t) &= \frac{1-qt^m}{1-qt}\prod_{i=1}^{l}\frac{1}{\lambda_i!}\frac{\partial_i^{\lambda}}{\partial x_i^{\lambda_i}}\Big[\frac{(1+x_i(t^{\deg p}+t^{2\deg p}+\ldots+t^{(m-1)\deg p}))^{N_i}}{(1+t^{\deg p}+t^{2\deg p}+\ldots+t^{(m-1)\deg p})^{N_i}}\Big]_{x_1=0,x_2=0,\ldots,x_l=0} \\ &= \frac{1-qt^m}{1-qt}\prod_{i=1}^{l}\binom{N_i}{\lambda_i}\Big(\frac{t^{im}-t^i}{t^{im}-1}\Big)^{\lambda_i}\end{aligned}$$

□

Once again, although we cannot compute a general expression for d_n, we can calculate its limit.

Proposition 17.

$$\lim_{n\to\infty}\frac{c_n}{|\mathrm{Poly}_n(\mathbb{F}_q)|} = \prod_{i=1}^{l}\binom{N_i}{\lambda_i}\Big(\frac{1-q^m}{1-q^{im}}\Big)^{\lambda_i}$$

Proof. Once again, we split our limit into the quotient of two limits. This gives us

$$\begin{aligned}\lim_{n\to\infty}\frac{|\mathrm{Poly}_n^m(\mathbb{F}_q)|}{q^n} &= \lim_{n\to\infty}\frac{q^n-q^{n-(m-1)}}{q^n} \\ &= \lim_{n\to\infty}1-\frac{1}{p^{m-1}} \\ &= 1-\frac{1}{p^{m-1}}\end{aligned}$$

and

$$\begin{aligned}\lim_{n\to\infty}\frac{d_n}{q^n} &= (1-u)D(\frac{u}{q})|_{u=1}\\ &= (1-\frac{u^m}{q^{m-1}})\prod_{i=1}^{l}\binom{N_i}{\lambda_i}\Big(\frac{(\frac{u}{q})^{im}-(\frac{u}{q})^i}{(\frac{u}{q})^{im}-1}\Big)^{\lambda_i}\Big|_{u=1}\\ &= (1-\frac{1}{q^{m-1}})\prod_{i=1}^{l}\binom{N_i}{\lambda_i}\Big(\frac{1-q^m}{1-q^{im}}\Big)^{\lambda_i}\end{aligned}$$

Which gives us

$$\begin{aligned}\lim_{n\to\infty}\frac{c_n}{|\mathrm{Poly}_n(\mathbb{F}_q)|} &= \frac{\lim_{n\to\infty}\frac{d_n}{q^n}}{\lim_{n\to\infty}\frac{|\mathrm{Poly}_n^m(\mathbb{F}_q)|}{q^n}}\\ &= \prod_{i=1}^{l}\binom{N_i}{\lambda_i}\Big(\frac{1-q^m}{1-q^{im}}\Big)^{\lambda_i}\end{aligned}$$

□

4 Statistics about polynomials non-vanishing at 0 and quadratic residues.

In this section we consider the set $B_n^m(\mathbb{F}_q) = \{f \in \mathrm{Poly}_n^m(\mathbb{F}_q) \mid f(0) \neq 0\}$. This is equivalent to saying that the irreducible x does not divide the polynomial. For $m = 2$, we will simply use $B_n(\mathbb{F}_q)$. Because 0 is no longer a root, we can ask more refined questions about its statistics, specifically about quadratic residues. We will use the same methods to compute some statistics about this set.

Throughout this section we will assume that q is a prime number not equal to 2.

4.1 Quadratic residues on $\overline{\mathbb{F}_q}$

Recall that we for any $\theta \in \overline{\mathbb{F}_q}$, we defined $\deg(\theta)$ to be the degree of its minimal polynomial over $\mathbb{F}_q$. This definition is also equivalent to $\deg(\theta)$ being the minimum i such that $x \in \mathbb{F}_{q^i}$.

Definition 13. A nonzero $\theta \in \overline{\mathbb{F}_q}$ is a *quadratic residue* (QR) if $\deg(\theta) = \deg(\sqrt{\theta})$. Otherwise it is a *quadratic nonresidue* (NQR)

By $\sqrt{\theta}$ we mean either of the solutions to $x^2 - \theta$, which have the same degree. (Note that $\mathbb{F}_q(\sqrt{\theta})$ and $\mathbb{F}_q(-\sqrt{\theta})$ are contained in each other, hence are the same field). From the second definition of degree, we always have $\deg(\theta) \leq \deg(\sqrt{\theta}) \leq 2\deg(\theta)$. This is because if $\sqrt{\theta} \in \mathbb{F}_{q^k}$, then $\theta \in \mathbb{F}_{q^k}$. On the other hand, if $\theta \in \mathbb{F}_{q^k}$, then the polynomial $x^2 - \theta$ generates at most a quadratic extension of $\mathbb{F}_{q^k}$.

For any $\theta \in \mathbb{F}_q$, it is usually said that θ is a quadratic residue if there is some $c \in \mathbb{F}_q$ such that $c^2 = \theta$, and a quadratic nonresidue otherwise. Our definition extends to all of $\overline{\mathbb{F}_q}$, and agrees with the usual one for $\theta \in \mathbb{F}_q$. For example, for $3 \in \mathbb{F}_5$, $\deg(3) = 1$, but $x^2 - 3$ is irreducible over $\mathbb{F}_5$. Hence $\deg(\sqrt{3}) = 2$, so 3 is a NQR over $\mathbb{F}_5$.

Lemma 18. *Let $f(x)$ be an irreducible polynomial over $\mathbb{F}_q$, $f(x) \neq x$. Then either all of the roots of f are QR, or they all are NQR.*

Proof. It suffices to show that one root being QR implies that all of them are QR. Let $\deg(f) = n$. Then, we know that $\mathbb{F}_{q^n}$ is the splitting field of f, with cyclic Galois group of order n generated by the Frobenius map $x \to x^q$. The n roots of f must be distinct since f is irreducible, and all have degree n. Since the Frobenius map sends roots to roots and the Galois group must be transitive on the roots, the set of distinct roots is $\{\theta, \theta^q, \ldots, \theta^{q^{n-1}}\}$ for some $\theta \in \mathbb{F}_{p^n}$. For ease of notation, let $x_k = \theta^{q^k}$.

Now, assume that f has a root which is QR. Without loss of generality, we may pick $\theta = x_0$ to be the root. Then $\exists\, \gamma \in \mathbb{F}_{q^n}$ such that $\gamma^2 = x_0$. Hence, $\gamma^{q^k} \in \mathbb{F}_{q^n}$, and $\left(\gamma^{q^k}\right)^2 = x_k$ for any k. Hence we have that $\deg(\sqrt{x_k}) \leq n = \deg(x_k)$. Combining this with the previous observation that $\deg(\sqrt{\alpha}) \geq \deg(\alpha)$ we get that $\deg(x_i) = \deg(\sqrt{x_i})$ for all roots x_i of f, hence all are QR. □

We say that an irreducible polynomial $p(x)$ is QR if all of its roots are QR, and NQR otherwise. This leads us to the following three definitions:

Definition 14.

1. For any $f \in B_n^m(\mathbb{F}_q)$ and $k \in \mathbb{N}$, define $X_k(f) = \#$ of *distinct* irreducible factors g_i of f such that g_i is QR of degree k.
2. $\alpha_k :=$ total number of QR irreducibles in $\mathbb{F}_q[x]$ of degree k
3. $\beta_k :=$ total number of NQR irreducibles in $\mathbb{F}_q[x]$ of degree k

Note that by Theorem 13 we always have $\alpha_k + \beta_k = N_k$, except when $k = 1$ because we include x as an irreducible, but we don't classify 0 as QR or NQR. The computation of α_k (and β_k, by our knowledge of N_k) will be considered later.

4.2 Statistics about m-th power free polynomials non vanishing at 0

With these tools, one very general statistic about this set we can ask is the following. Given a $\Lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l)$ as before, define

$$e_n = \sum_{f \in B_n^m(\mathbb{F}_q)} \prod_{k=1}^{l} \binom{X_k(f)}{\lambda_k}$$

We will derive the generating function for such a sequence in its most general form. Let the generating function be

$$E(t) = \sum_{n=0}^{\infty} e_n t^n$$

Theorem 19.

$$E(t) = \frac{1-qt^m}{1-qt}\frac{t-1}{t^m-1}\prod_{i=1}^{l}\binom{\alpha_i}{\lambda_i}\left(\frac{t^{im}-t^i}{t^{im}-1}\right)^{\lambda_i}$$

Proof. The method will be very similar to the computations before. Let $\Psi(\mathbf{u}, t)$ be

$$\Psi(\mathbf{u}, t) = \sum_{n=0}^{\infty}\left(\sum_{f\in B_n^m(\mathbb{F}_q)} u_1^{X_1(f)}\dots u_l^{X_l(f)}\right)t^n$$

Once again, it can be seen that

$$\frac{\partial^{(\lambda_1+\dots+\lambda_l)}}{\partial u_1^{\lambda_1}\cdots\partial u_l^{\lambda_l}}\left[\frac{\Psi(\mathbf{u},t)}{\lambda_1!\dots\lambda_l!}\right]_{u_1=0,\dots,u_l=0} = E(t)$$

Again, let $\chi(p)$ be:

$$\chi(p) = \begin{cases} u_i & \text{if } \deg(p) = i \le l, \text{ and } p \text{ is QR;} \\ 1 & \text{otherwise;} \end{cases}$$

Turning Ψ into a product formula is very similar as before, but notice that no polynomial in $B_n^m(\mathbb{F}_q)$ has the irreducible x as a factor, so we divide the term it contributes out. Again, $p(x)$ ranges over every irreducible in $\mathbb{F}_q[x]$.

$$\begin{aligned}\Psi(\mathbf{u},t) &= \frac{\prod_{p(x)} 1+\chi(p)\left(t^{\deg(p)}+\dots+t^{(m-1)\deg(p)}\right)}{1+t+\dots+t^{m-1}} \\ &= \frac{\prod_{p(x)}(1+t^{\deg p}+\dots+t^{(m-1)\deg p})}{1+t+\dots+t^{m-1}}\prod_{i=1}^{l}\left(\frac{1+u_i(t^i+\dots+t^{(m-1)i})}{1+t^i+\dots+t^{(m-1)i}}\right)^{\alpha_i} \\ &= \frac{1-qt^m}{1-qt}\frac{t-1}{t^m-1}\prod_{i=1}^{l}\left(\frac{1+u_i(t^i+\dots+t^{(m-1)i})}{1+t^i+\dots+t^{(m-1)i}}\right)^{\alpha_i}\end{aligned}$$

This in turn gives us

$$\begin{aligned}E(t) &= \frac{1-qt^m}{1-qt}\frac{t-1}{t^m-1}\prod_{i=1}^{l}\frac{1}{\lambda_i!}\frac{\partial_i^{\lambda}}{\partial u_i^{\lambda_i}}\left[\frac{1+u_i(t^i+\dots+t^{(m-1)i})}{1+t^i+\dots+t^{(m-1)i}}\right]_{u_1=0,\dots,u_l=0}^{\alpha_i} \\ &= \frac{1-qt^m}{1-qt}\frac{t-1}{t^m-1}\prod_{i=1}^{l}\binom{\alpha_i}{\lambda_i}\left(\frac{t^{im}-t^i}{t^{im}-1}\right)^{\lambda_i}\end{aligned}$$

□

Example 4 from the introduction corresponds to this sequence e_n with $\lambda_1 = 1$ and $m = 2$. From Theorem 18, we have the generating function for the desired sequence:

$$\frac{1-qt^2}{1-qt}\frac{t\alpha_1}{(1+t)^2} = \sum_{n=0}^{\infty}\left(\sum_{f\in B_n(\mathbb{F}_q)} X_1(f)\right)t^n$$

Which we can expand out for any n.

These results rely on us knowing the value of α_k for arbitrary k. This is the subject of the next section.

4.3 Counting the number of irreducibles with quadratic residue roots.

In this section we will prove the following result:

Theorem 20.

$$\alpha_k = \begin{cases} (q-1)/2 & \text{if } k = 1; \\ N_k/2 & \text{if } k \text{ is odd, } k \neq 1; \\ \frac{1}{2}\left(N_k - \sum_{i=1}^{j}\frac{N_{k/2^i}}{2^i}\right) & \text{if } k = 2^j m, \text{ where } 2 \nmid m; \end{cases}$$

Proof. We will look at the splitting field of such polynomials. For any irreducible of degree k, there are exactly k distinct roots of degree k by Lemma 17. Conversely, given any element θ of degree k, its minimal polynomial has a unique splitting field, namely $\mathbb{F}_{q^k}$, and hence is an irreducible polynomial of degree k. Therefore there is an k to 1 correspondence between elements of degree k and irreducibles of degree k. Our goal here is to separate out the elements which are QR to obtain the irreducibles that contribute to the count of α_k.

Let $S_k = \{\theta \in \overline{\mathbb{F}_q} \mid \deg(\theta) = k\}$. This is a set of size $|S_k| = kN_k$ by the above correspondence. For any $\theta \in S_k$ we must also have $-\theta \in S_k$, since they have the same degree. This is because $\mathbb{F}_q(\theta)$ and $\mathbb{F}_q(-\theta)$ are contained in each other, hence are the same field. Since $q \neq 2$, $\pm\theta$ are *distinct* elements.

Let $S_k^2 = \{\theta^2 \mid \theta \in S\}$. This is a set of size $\frac{kN_k}{2}$, because each pair of additive inverses in S_k square to the same number. We are ultimately interested in $S_k^2 \cap S_k = \{\theta \mid \deg(\theta) = \deg(\sqrt{\theta}) = k\}$. This is precisely the number of elements θ of degree k which are QR. Using the correspondence of elements to irreducibles, we have $|S_k^2 \cap S_k| = k\alpha_k$. The size of this set depends on the parity of k:

- $k = 1$.

  As stated previously, for $\theta \in \mathbb{F}_q$ (elements of degree 1), the notion of QR is the usual one. It is well known that there are exactly $\frac{q-1}{2}$ quadratic residues in $\mathbb{F}_q$, hence $\alpha_1 = \frac{q-1}{2}$.

- k is odd, $k \neq 1$.

  Let $\theta \in S_k^2$, and suppose that $\theta \notin S_k$. Then it must lie in some proper subextension $E \subset \mathbb{F}_{q^k}$. The polynomial $x^2 - \theta$ is irreducible over E, and thus generates a particular quadratic extension of E. However, this extension must be $\mathbb{F}_{q^k}$ from our definition of θ, and we would have $[\mathbb{F}_{q^k} : E] = 2$, which is impossible if k is odd. Hence $S_k^2 \subset S_k$, and we must have $k\alpha_k = \dfrac{kN_k}{2}$.

- k is even.

  A recursive formula for α_k will be given in this case. For even k there will always be a unique subextension $\mathbb{F}_{q^{k/2}} \subset \mathbb{F}_{q^k}$ of degree 2. Now let $\theta \in S_k^2$, and assume that $\deg(\theta) = k/2$. Then the polynomial $x^2 - \theta$ is irreducible over $\mathbb{F}_{q^{k/2}}$, because its two roots lie in the quadratic extension $\mathbb{F}_{q^k}$. Hence θ is a NQR of degree $k/2$. Conversely, given a θ which is NQR of degree $k/2$, the polynomial $x^2 - \theta$ generates the same quadratic extension $\mathbb{F}_{q^n}$, with two distinct roots, $\pm\sqrt{\theta}$. Both of these elements must be degree k and square to θ, so $\theta \in S_k^2$. There are precisely $\frac{k\beta_{k/2}}{2}$ such elements in S_k^2, so we remove them from the count and are left with the number of QR elements of degree k. Hence, $k\alpha_k = \frac{kN_k}{2} - \frac{k\beta_{k/2}}{2}$

  Using the relation $\alpha_k + \beta_k = N_k$ for any $k \neq 1$, we get to the following recursive formula:

  $$\alpha_k = \frac{1}{2}\left(N_k - N_{k/2} + \alpha_{k/2}\right)$$

  Plugging the formula back into itself will remove a power of 2 every time. For $k = 2^j m$ where $2 \nmid m$, we plug it back into intself $(j-1)$ times, at which point we are left with α_m, which we already know since m is odd. This recursion simplifies to the formula above.

□

Acknowledgment

We would like to extend our gratitude to Dr. Peter May for organizing and running the UChicago Math Summer REU, which was a fantastic experience and opportunity. We would also like to thank Dr. Benson Farb for not only allowing and encouraging us to work with his group but also inspiring us with his enthusiasm. Finally, a huge thank you to our mentor Weiyan Chen, for teaching us and guiding us through the REU.

References

[1] Church, T., Ellenberg, J., and Farb, B., *Representation stability in cohomology and asymptotics for families of varieties over finite fields*, Contemporary Mathematics620(2014), 1-54

[2] Fulman, J. *A generating function approach to counting theorems for square-free polynomials and maximal tori*. arXiv:1410.3540.

Department of Mathematics,
University of Chicago,
5734 S. University Ave.
Chicago, IL 60637, U.S.A.

E-mail: Mario Alegre: alegre@uchicago.edu, Pedro Juarez: pdrjuarez@uchicago.edu, Hani Pajela: hanipajela@uchicago.edu