COUNTING PERMUTATIONS AND POLYNOMIALS WITH A RESTRICTED FACTORIZATION PATTERN

Arnold Knopfmacher - University of the Witwatersrand, Johannesburg 2050, South Africa

and

Richard Warlimont, Universität Regensburg, 93040 Regensburg, Germany

ABSTRACT

We determine the asymptotic probability that a polynomial of degree n over a finite field with q elements has no more than k irreducible factors of any degree, for each natural number k. In particular we show that for $q = 2$, almost 67% of such polynomials have no more than 2 irreducible factors of any given degree and almost 81% have no more than 3 irreducible factors of any given degree, as $n \to \infty$. Similar results are also shown to hold in the case of the analogous problem for permutations. Here we wish to estimate the asymptotic proportion of permutations of n elements that have no more than k cycles of any given length in their cycle decomposition.

1. Polynomials over a finite field with a restricted factorization pattern

Given a monic polynomial $f(x) \in \mathbb{F}_q[X]$ with $\partial f = n$, let $\alpha_i, 1 \leq i \leq n$, be the number of irreducible factors of degree i (counting multiplicity) that divide f. We call the n-tuple $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ the *factorization pattern* of f, and denote this by (f).

Our aim is to determine the asymptotic probability that a polynomial of degree n has a factorization pattern $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ which satisfies $0 \leq \alpha_i \leq k$, for $1 \leq i \leq n$. That is, the asymptotic probability that a polynomial has no more than k irreducible factors of any degree. We shall use the notation $|(f)| \leq k$ as a shorthand for these conditions on (f). The case $k = 1$, that is, polynomials having only distinct degree factors has previously been considered in [3]. In addition the case of general k has been treated in the more abstract case of certain additive arithmetical semigroups [4]. The present results survey and add to these previous works in the concrete case of polynomials over a finite field.

Many deterministic and probabilistic factorization algorithms for polynomials in $\mathbb{F}_q[X]$ require that a distinct degree factorization of the polynomial be performed as an initial step (see e.g. Shparlinski [6, Chapter 1]. The remainder of the algorithms consist of methods to determine the α_i factors of degree i in f, $1 \leq i \leq n$.

Our results below are therefore of relevance to the problem of finding the computational cost of these algorithms. For example, it follows from Theorem 2 for $q = 2$ that as $n \to \infty$, less than 0.27% of squarefree polynomials of degree n have more than two irreducible factors of any given degree.

Let $k \in \mathbb{N}$. We consider the following three subsets of the set of monic polynomials of degree n in $\mathbb{F}_q[X]$:

$$G_{k,1} = \{f : |(f)| \leq k\}$$

$$G_{k,2} = \{f : |(f)| \leq k \quad \text{and} \quad f \quad \text{is squarefree}\}$$

$$G_{k,3} = \{f : |((f))| \leq k\}$$

where $((f)) = ((\hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_n))$ and $\hat{\alpha}_i$, $1 \leq i \leq n$ is the number of *distinct* irreducible factors of degree i that divide f.

Put

$$\gamma_{k,j}(n) = \#\{f \in G_{k,j} | \partial f = n\}, \quad j = 1, 2, 3.$$

Our estimates for $\gamma_{k,j}(n), j = 1, 2, 3$ are obtained by considering the respective ordinary generating functions for these sequences.

Let $\pi(n) \equiv \pi(n, q)$ denote the number of monic irreducible polynomials of degree n in $\mathbb{F}_q[X]$:

Theorem 1 We have

(a) (1.1)
$$\sum_{n=0}^{\infty} \gamma_{k,1}(n) w^n = \prod_{m=1}^{\infty} \sum_{l=0}^{k} \binom{\pi(m)-1+l}{l} w^{ml}$$

(b) (1.2)
$$\sum_{n=0}^{\infty} \gamma_{k,2}(n) w^n = \prod_{m=1}^{\infty} \sum_{l=0}^{k} \binom{\pi(m)}{l} w^{ml}$$

(c) (1.3)
$$\sum_{n=0}^{\infty} \gamma_{k,3}(n) w^n = \prod_{m=1}^{\infty} \sum_{l=0}^{k} \binom{\pi(m)}{l} \left(\frac{w^m}{1-w^m}\right)^l.$$

Proof

(a) The number of monic polynomials of degree n in $\mathbb{F}_q[X]$ that have factorization pattern of type $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is given by

$$\prod_{j=1}^{n} \binom{\pi(j)+\alpha_j-1}{\alpha_j}.$$

Thus
$$\sum_{n=0}^{\infty} \gamma_{k,1}(n) w^n = \sum_{n=0}^{\infty} w^n \sum_{\substack{\alpha_1+2\alpha_2+\cdots=n \\ 0\le\alpha_1\le k, 0\le\alpha_2\le k, \cdots}} \prod_{j=1}^{n} \binom{\pi(j)+\alpha_j-1}{\alpha_j}$$

$$= \left(\sum_{\alpha_1=0}^{k} \binom{\pi(1)+\alpha_1-1}{\alpha_1} w^{\alpha_1}\right)\left(\sum_{\alpha_2=0}^{k} \binom{\pi(2)+\alpha_2-1}{\alpha_2} w^{2\alpha_2}\right)\cdots$$

$$= \prod_{m=1}^{\infty} \sum_{l=0}^{k} \binom{\pi(m)+l-1}{l} w^{ml}.$$

(b) The number of squarefree monic polynomials of degree n in $\mathbb{F}_q[X]$ that have factorization pattern of type $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is given by

$$\prod_{j=1}^{n} \binom{\pi(j)}{\alpha_j}.$$

Hence

$$\sum_{i=0}^{\infty} \gamma_{k,2}(n) w^n = \sum_{n=0}^{\infty} w^n \sum_{\substack{\alpha_1 + 2\alpha_2 \dots = n \\ 0 \le \alpha_1 \le k, 0 \le \alpha_2 \le k, \dots}} \prod_{j=1}^{n} \binom{\pi(j)}{\alpha_j}$$

$$= \left(\sum_{\alpha_1=0}^{k} \binom{\pi(1)}{\alpha_1} w^{\alpha_1} \right) \left(\sum_{\alpha_2=0}^{k} \binom{\pi(2)}{\alpha_2} w^{2\alpha_2} \right) \cdots$$

$$= \prod_{m=1}^{\infty} \sum_{l=0}^{k} \binom{\pi(m)}{l} w^{ml}.$$

(c) This is similar to (b) except that we may replace single occurrence of an irreducible factor by a sequence of such factors. In generating function terms this entails the change of variable $w \to w + w^2 + w^3 + \dots = \frac{w}{1-w}$. Hence

$$\sum_{n=0}^{\infty} \gamma_{k,3}(n) w^n = \sum_{n=0}^{\infty} \gamma_{k,2} \left(\frac{w}{1-w} \right)^n$$

$$= \prod_{m=1}^{\infty} \sum_{l=0}^{k} \binom{\pi(m)}{l} \left(\frac{w^m}{1-w^m} \right)^l.$$

Theorem 1 is now used to derive the asymptotic proportions of polynomials belonging to $G_{k,j}, j = 1, 2, 3 \dots$.

Theorem 2 For each $k \in \mathbb{N}$,

(1.4) $$L_1(q,k) := \lim_{n\to\infty} \frac{\gamma_{k,1}(n)}{q^n} = \prod_{m=1}^{\infty} \left(\sum_{l=0}^{k} \binom{\pi(m) - 1 + l}{l} q^{-ml} \right) \exp\left(-\frac{1}{m}\right)$$

(1.5) $$L_2(q,k) := \lim_{n\to\infty} \frac{\gamma_{k,2}(n)}{q^n} = \prod_{m=1}^{\infty} \left(\sum_{l=0}^{k} \binom{\pi(m)}{l} q^{-ml} \right) \exp\left(-\frac{1}{m}\right)$$

(1.6) $$L_3(q,k) := \lim_{n\to\infty} \frac{\gamma_{k,3}(n)}{q^n} = \prod_{m=1}^{\infty} \left(\sum_{l=0}^{k} \binom{\pi(m)}{l} (q^m - 1)^{-l} \right) \exp\left(-\frac{1}{m}\right).$$

The proof requires the following technical lemma.

Lemma 1

Let $b_m (m \geq 1)$ and $b_{m,\nu} (m \geq 1, \nu \geq 2)$ be complex numbers such that

$$(A) \qquad \sum_{m=1}^{\infty} \left(|b_m - \frac{1}{m}| + \sum_{\nu=2}^{\infty} |b_{m,\nu}| \right) < \infty.$$

Put $b_{m,1} := b_m$. Then the function

$$(B) \qquad f(z) := \prod_{m=1}^{\infty} \left(1 + \sum_{\nu=1}^{\infty} b_{m,\nu} z^{m\nu} \right)$$

is well defined and holomorphic on $|z| < 1$. Let

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \ (|z| < 1).$$

Then

$$(C) \qquad \lim_{n \to \infty} a_n = \prod_{m=1}^{\infty} \left(1 + \sum_{\nu=1}^{\infty} b_{m,\nu} \right) \exp\left(-\frac{1}{m} \right).$$

Proof

Let $0 < r < 1$. Then using (A),

$$\sum_{m=1}^{\infty} \sum_{\nu=1}^{\infty} |b_{m,v}| r^{mv} \leq \sum_{m=1}^{\infty} \left(|b_m - \frac{1}{m}| + \frac{1}{m} r^m + \sum_{\nu=2}^{\infty} |b_{m,\nu}| \right) < \infty.$$

From this it follows that the infinite product in (B) converges uniformly on $|z| \leq r$ and therefore represents a function $f(z)$ which is holomorphic on $|z| < 1$.

Then we have

$$f(z) = \frac{g(z)}{1-z}$$

where

$$g(z) := \prod_{m=1}^{\infty} \left(1 + \sum_{\nu=1}^{\infty} b_{m,\nu} z^{m\nu}\right) \exp\left(-\frac{1}{m} z^m\right) = \prod_{m=1}^{\infty} \left(1 + (b_m - \frac{1}{m}) z^m + \sum_{\nu=2}^{\infty} c_{m,\nu} z^{mv}\right)$$

where because of (A) we have

$$S := \sum_{m=1}^{\infty} \left(|b_m - \frac{1}{m}| + \sum_{\nu=2}^{\infty} |c_{m,\nu}|\right) < \infty.$$

Therefore

$$g(z) = \sum_{n=0}^{\infty} d_n z^n \qquad (|z| \leq 1)$$

where

$$\sum_{n=0}^{\infty} |d_n| \leq \exp(S) < \infty.$$

Since

$$a_n = \sum_{h=0}^{n} d_h$$

we have

$$\lim_{n\to\infty} a_n = \sum_{h=0}^{\infty} d_h = g(1) =$$

$$\prod_{m=1}^{\infty} \left(1 + \sum_{\nu=1}^{\infty} b_{m,\nu}\right) \exp\left(-\frac{1}{m}\right).$$

Proof of Theorem 2

In Theorem 1 substitute $w = \frac{z}{q}$. Then we get

$$\sum_{n=0}^{\infty} \frac{\gamma_{k,1}(n)}{q^n} z^n = \prod_{m=1}^{\infty} \left(1 + \frac{\pi(m)}{q^m} z^m + \sum_{l=2}^{k} \binom{\pi(m) - 1 + l}{l} q^{-ml} z^{ml}\right)$$

$$\sum_{n=0}^{\infty} \frac{\gamma_{k,2}(n)}{q^n} z^n = \prod_{m=1}^{\infty} \left(1 + \frac{\pi(m)}{q^m} z^m + \sum_{l=2}^{k} \binom{\pi(m)}{l} q^{-ml} z^{ml}\right)$$

$$\sum_{n=0}^{\infty} \frac{\gamma_{k,3}(n)}{q^n} = \prod_{m=1}^{\infty} \left(1 + \frac{\pi(m)}{q^m} z^m + \frac{\pi(m)}{q^m} \frac{z^{2m}}{q^m - z^m} + \sum_{l=2}^{k} \binom{\pi(m)}{l} \left(\frac{z^m}{q^m - z^m}\right)^l\right).$$

The three infinite products are such that we can apply lemma 1: $b_m = \frac{\pi(m)}{q^m}$, using the well known bounds

$$(1.7) \qquad \frac{1}{m} - 2q^{-m/2} < \frac{\pi(m)}{q^m} < \frac{1}{m}.$$

These follow from the formula

$$\pi(n) = \frac{1}{n} \sum_{d/n} \mu(d) q^{n/d} .$$

where $\mu(\cdot)$ denotes the Mobius function.

Here we also must note that

$$\binom{\pi(m) - 1 + l}{l} \quad \text{and} \quad \binom{\pi(m)}{l} = \mathrm{O}(\frac{q^{ml}}{m^l}).$$

The limits (1.4), (1.5) and (1.6) now follow from Lemma 1. ∎

Since exact values of $\pi(m)$ can be determined we can obtain accurate approximations to the limits (1.4), (1.5) and (1.6) for any fixed value of q.
We consider the two extreme cases namely $q = 2$ and $q \to \infty$.

As $q \to \infty$, the bounds (1.7) imply that

$$\frac{\pi(m)}{q^m} \to \frac{1}{m} \quad \text{uniformly for} \quad m = 1, 2, 3, \ldots .$$

Hence as $q \to \infty$, (1.4), (1.5) and (1.6) tend to a common limit

$$L(k) := \prod_{m=1}^{\infty} \left(\sum_{l=0}^{k} \frac{1}{l!} \left(\frac{1}{m}\right)^l \right) \exp\left(-\frac{1}{m}\right). \tag{1.8}$$

We show that this same limit also arises in another context related to permutations (see section 2). Numerical values for $L(k), k = 1, 2, \ldots, 10$ can be found in Table 2 in section 2.

For any fixed value of q, $L_1(q,k)$ and $L_3(q,k) \to 1$ as $k \to \infty$, while $L_2(q,k) \to 1 - \frac{1}{q}$, the proportion of squarefree polynomials of degree n in $\mathbb{F}_q[X]$.

k	$L_1(2,k)$	$L_2(2,k)$	$L_3(2,k)$
1	0.39673411	0.39673411	0.66559658
2	0.66784706	0.49869547	0.99727609
3	0.80812272	0.49997319	0.99994395
4	0.88958533	0.49999955	0.99999906
5	0.93724695	0.49999999	0.99999999
6	0.96478145	0.5	1.
7	0.98045331	0.5	1.
8	0.98925397	0.5	1.
9	0.99413967	0.5	1.
10	0.99682593	0.5	1.

A comparison of the proportions in a given row suggests that it is best to apply factorization methods for polynomials in $\mathbb{F}_2[X]$ to squarefree polynomials. For example we see from the table that about 19.2% of polynomials of degree n, n large, have more than 3 factors of the same degree, but only 0.0054% of squarefree polynomials have this property. Since the squarefree part of a polynomial is easily determined, it seems preferable to do this as a first step prior to applying a factorization algorithm.

As $q \to \infty$, the proportion of squarefree polynomials tends to 1 and the benefit of applying an algorithm to a squarefree polynomial rather than any polynomials of degree n, falls away.

2. Permutations with a restricted cycle type

It is well known that every permutation of n letters has a unique factorization into disjoint cycles (apart from order). We say that $\pi \in S_n$ has cycle type $\lambda = (1^{\lambda_1} 2^{\lambda_2} \dots n^{\lambda_n})$ if π has $\lambda_i \equiv \lambda_i(\pi)$ cycles of length $i, 1 \leq i \leq n$, in its cycle factorization.

In this section we investigate the sizes of the subsets of permutations of n letters which belong to the sets

$$S_n^k := \{\pi \in S_n, \lambda_i(\pi) \leq k \qquad (1 \leq i \leq n)\} .$$

Put

$$\sigma_k(n) = \#\{\pi \in S_n^k\}.$$

The special case $k = 1$, corresponds to permutations of n letters with distinct cycle lengths. This problem is discussed by Wilf [7,p.106]. In addition a more precise asymptotic estimate for the coefficients of this particular generating function is derived in Greene and Knuth [2].

Let

$$\text{(2.1)} \qquad \exp_k(x) = \sum_{m=0}^{k} \frac{x^m}{m!}.$$

Then we have

Lemma 2 The exponential generating function for $\{\sigma_k(n)\}$ is

$$\text{(2.2)} \qquad \sum_{n=0}^{\infty} \sigma_k(n) x^n / n! = \prod_{m=1}^{\infty} \exp_k\left(\frac{x^m}{m}\right).$$

Note that if we let $k \to \infty$ then $\exp_k(x) \to e^x$ and $\prod_{m=1}^{\infty} \exp_k\left(\frac{x^m}{m}\right) \to \frac{1}{1-x}$. It follows that $\sigma_\infty(n) = n!$, as expected.

Proof

The number of permutations in S_n having $\lambda = (1^{\lambda_1}2^{\lambda_2}\dots n^{\lambda_n})$ as their cycle type is

$$\frac{n!}{\prod_{j\geq 1}(\lambda_j! j^{\lambda_j})}.$$

Thus

$$\sum_{n=0}^{\infty} \sigma_k(n) x^n/n! = \sum_{n=0}^{\infty} x^n \sum_{\substack{\lambda_1+2\lambda_2+\dots=n \\ 0\leq\lambda_1\leq k, 0\leq\lambda_2\leq k,\dots}} \frac{1}{\prod_j(\lambda_j! j^{\lambda_j})}$$

$$= \left(\sum_{\lambda_1=0}^{k} \frac{x^{\lambda_1}}{1^{\lambda_1}\lambda_1!}\right)\left(\sum_{\lambda_2=0}^{k} \frac{(x^2)^{\lambda_2}}{2^{\lambda_2}\lambda_2!}\right)\cdots.$$

$$= \exp_k(x)\exp_k\left(\frac{x^2}{2}\right)\exp_k\left(\frac{x^3}{3}\right)\cdots.$$

■

We now apply Lemma 1 to (2.2) to deduce

Theorem 3 For each $k \in \mathbb{N}$,

(2.3) $$L(k) := \lim_{n\to\infty} \sigma_k(n)/n! = \prod_{m=1}^{\infty} \exp_k\left(\frac{1}{m}\right)\exp\left(-\frac{1}{m}\right).$$

■

As shown for example by Wilf [7]

$$L(1) = \prod_{m=1}^{\infty}\left(1+\frac{1}{m}\right)e^{-1/m} = e^{-\gamma} \approx 0.5614\dots$$

where γ is Euler's constant. For $k \geq 2$, clearly

$$L(1) \leq L(k) \leq L(\infty) = 1.$$

For large k we have the following accurate upper and lower bounds from (2.3),

(2.4) $$L(k) \leq 1 - e^{-1}\sum_{m=k+1}^{\infty}\frac{1}{m!}$$

and

$$L(k) \geq 1 - \sum_{m=1}^{\infty} \exp\Big(-\frac{1}{m}\Big) \sum_{l=k+1}^{\infty} \frac{1}{l!}\Big(\frac{1}{m}\Big)^l$$

$$= 1 - \sum_{l=k+1}^{\infty} \frac{1}{l!} \sum_{m=1}^{\infty} \exp\Big(-\frac{1}{m}\Big)/m^l$$

$$> 1 - \sum_{l=k+1}^{\infty} \frac{1}{l!} \sum_{m=1}^{\infty} \frac{1}{m^l} = 1 - \sum_{l=k+1}^{\infty} \zeta(l)/l!$$

$$> 1 - \zeta(2) \sum_{l=k+1}^{\infty} \frac{1}{l!}. \tag{2.5}$$

In particular we deduce that as $k \to \infty$,

$$1 - L(k) \sim \frac{1}{(k+1)!} . \tag{2.6}$$

The table below shows the rapid convergence of $L(k)$ to 1 even for small k. In particular we see that only 10.3% of permutations $\pi \in S_n$ have more than 2 cycles of any given length and less than 2.2% have more than 3 cycles of any length. Approximations obtained from the exact formula (2.3) for $L(k)$ are compared with the values from the upper and lower bounds (2.4) and (2.5).

k	Lower bound	$L(k)$	Upper bound
1	-0.18152625	0.56145948	0.73575888
2	0.64094078	0.89687126	0.9196986
3	0.91509646	0.97864866	0.98101184
4	0.98363538	0.99613197	0.99634015
5	0.99734317	0.99938983	0.99940582
6	0.9996278	0.99991568	0.99991676
7	0.99995417	0.99998969	0.99998975
8	0.99999497	0.99999887	0.99999887
9	0.9999995	0.99999989	0.99999989
10	0.99999996	0.99999999	0.99999999

Table 2

As noted at the end of section 1, the proportion of polynomials of degree n in $G_{k,1}$, $G_{k,2}$, and $G_{k,3}$ as $q \to \infty$ have the same limiting value in each case, namely $L(k)$. As pointed out by the referee one might already have predicted this coincidence of the limits from the results of Cohen [1]. For example Cohen notes for that for fixed n and large q, the proportion of polynomials of degree n with factorization pattern $(\lambda_1, \lambda_2, \ldots \lambda_n)$ is asymptotically equal to the proportion of permutations in S_n with cycle type $(1^{\lambda_1}, 2^{\lambda_2} \ldots, n^{\lambda_n})$. Greene and Knuth [2] exploited this fact for the case $k = 1$, in their determination of $L_1(q, 1)$, for large q.

REFERENCES

1. Cohen S D (1970), The distribution of polynomials over finite fields, *Acta Arith.* **17** (1970), 255-271.

2. Greene D H and Knuth D E (1990), *Mathematics for the Analysis of Algorithms*, 3rd edn. Birkhauser, Boston.

3. Knopfmacher A and Warlimont R (in press). Distinct degree factorizations for polynomials over a finite field. *Trans. Amer. Math. Soc.*.

4. Knopfmacher A and Warlimont R, Distinct degree factorizations in additive arithmetical semigroups, manuscript.

5. Lidl R and H Niederreiter, *Finite Fields*, Addison-Wesley, Reading, MA 1983.

6. Shparlinski I E (1992), *Computational and Algorithmic Problems in Finite Fields*, Kluwer, Dordrecht.

7. Wilf H S, *generatingfunctionology*, 2nd edn., Academic Press, New York, 1994.

email: arnoldk@gauss.cam.wits.ac.za

(Received 2/3/95; revised 9/6/95)