NUMBER OF IRREDUCIBLE POLYNOMIALS IN SEVERAL VARIABLES OVER FINITE FIELDS

ARNAUD BODIN

Abstract. We give a formula and an estimation for the number of irreducible polynomials in two (or more) variables over a finite field.

1. Introduction

Let p be a prime number and $n \geqslant 1$. For $q = p^n$ we denote by $\mathbb{F}_q$ the finite field having q elements. The number of polynomials in $\mathbb{F}_q[x]$ of degree (exactly) d is $N_1(d) = q^{d+1} - q^d$. The number $I_1(d)$ of irreducible polynomials of degree d can be explicitly be computed with the help of the Moebius inversion formula and was already known by Gauss, see [8, p. 93]. Moreover we have an estimation for the proportion of irreducible polynomials among all polynomials of degree d (see [8, Ex. 26-27, p. 142]):

$$\frac{I_1(d)}{N_1(d)} \sim \frac{1}{d}.$$

In particular irreducible polynomials in one variable become more and more rare among the set of polynomials as the degree grows.

Surprisingly the situation is completely different if we look at irreducibility for polynomials in two (or more) variables. We will prove that most of the polynomials of degree d are irreducible and we give an estimate for this proportion as d grows.

Here is the mathematical formulation : let $N_2(d)$ be the number of polynomials in $\mathbb{F}_q[x,y]$ of degree exactly d and $I_2(d)$ the number of irreducible polynomials.

Theorem.

$$1 - \frac{I_2(d)}{N_2(d)} \sim \frac{q+1}{q^d}.$$

In particular it implies that $\frac{I_2(d)}{N_2(d)} \to 1$ as $d \to +\infty$.

Date: November 29, 2007.

For example if $q = 2$, the probability to choose an irreducible polynomial among polynomials of degree d is about $1 - \frac{3}{2^d}$. For $d = 10$ we find:

$$\frac{I_2(10)}{N_2(10)} = \frac{735342418237937154 33}{737509474978192424 96} = 0.997061\ldots$$

that we approach by

$$1 - \frac{3}{2^{10}} = 0.997070\ldots$$

The fact that in several variables almost all polynomials are irreducible is due to L. Carlitz [3]. This work has been expanded to the study of the distribution of irreducible polynomials according not to the degree but to the bi-degree (where the *bi-degree* of $P(x, y)$ is $(\deg_x P, \deg_y P)$) by Carlitz himself [4] and by S. Cohen [5] for more variables. More arithmetical stuff can be found in [6]. More recently such computations have been applied to algorithms of factorization of multivariate polynomials, see [9] and [7].

2. Number of polynomials

We first need to defined what is a normalized polynomial, let $f(x, y) \in \mathbb{F}_q[x, y]$ be a polynomial of degree exactly d :

$$f(x, y) = \alpha_0 x^d + \alpha_1 x^d y + \alpha_2 x^{d-2} y^2 + \cdots + \alpha_d y^d + \text{terms of lower degree.}$$

f is said to be *normalized* if the first non-zero term in the sequence $(\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_d)$ is equal to 1. Of course any polynomial g can be written $g(x, y) = c \cdot f(x, y)$ where f is a normalized polynomial and $c \in \mathbb{F}_q^*$. In particular it implies that the number of normalized polynomials of degree d is the total number of polynomials of degree d divided by $q - 1 = \#\mathbb{F}_q^*$.

The motivation is the following : we will need to factorize polynomials, but unfortunately this factorization is not unique: for example if $g = g_1 \cdot g_2$ is the decomposition of $g \in \mathbb{F}_q[x, y]$ into a product of irreducible factors, then $g = (cg_1) \cdot (c^{-1} g_2)$ is another factorization, for all $c \in \mathbb{F}_q^*$. This phenomenon is problematic when we try to count the number of reducible polynomials. However, now if $f = f_1 \cdot f_2$ is a factorization with f, f_1, f_2 normalized polynomials, then this decomposition is unique (up to permutation).

In the sequel of the text we will count normalized polynomials, normalized irreducible polynomials,... To have the non-normalized results, just multiply by $q - 1$.

Lemma 1. *The number of normalized polynomials of degree exactly d in $\mathbb{F}_q[x,y]$ is*

$$N(d) = \left(\frac{q^{d+1}-1}{q-1}\right) \cdot q^{\frac{d(d+1)}{2}}.$$

For example $N(1) = q(q+1)$, $N(2) = \frac{q^6-q^3}{q-1}$.

Proof. The number of monomials of degree lower or equal to d is $\frac{(d+1)(d+2)}{2}$, hence the number of polynomials of degree less or equal than d is

$$N'(d) = q^{\frac{(d+1)(d+2)}{2}}.$$

The number of non-zero homogeneous polynomials of degree d is

$$q^{d+1} - 1.$$

A polynomial of degree exactly d is the sum of a non-zero homogeneous polynomial of degree d with a polynomials of degree $< d$. Hence the number of polynomials of degree exactly d is:

$$\left(q^{d+1}-1\right) \cdot N'(d-1).$$

To get the number of normalized polynomials we divide by $q-1$ and obtain:

$$N(d) = \left(\frac{q^{d+1}-1}{q-1}\right) \cdot N'(d-1) = \left(\frac{q^{d+1}-1}{q-1}\right) \cdot q^{\frac{d(d+1)}{2}}.$$

□

The gap between two consecutive numbers is given by the following lemma.

Lemma 2.

- $\dfrac{N(d)}{N(d+1)} = \dfrac{1}{q^{d+2}} \cdot \left(1 - \dfrac{q-1}{q^{d+2}-1}\right).$
- *In particular* $\dfrac{N(d)}{N(d+1)} \sim \dfrac{1}{q^{d+2}}.$

We will need an upper bound for the product $N(a) \cdot N(b)$.

Lemma 3.

(1) $N(a) \cdot N(b) \leqslant N(a+b)$ *for all* $a \geqslant 1$, $b \geqslant 1$;
(2) $N(a) \cdot N(b) \leqslant q^3 \cdot N(a+b-1)$ *for all* $a \geqslant 1$, $b \geqslant 1$;
(3) $N(a) \cdot N(b) \leqslant q^5 \cdot N(a+b-2)$ *for all* $a \geqslant 3$, $b \geqslant 3$;

Proof. First of all the function defined by $M(d) = q^{d+1} - 1$ verifies $M(a) \cdot M(b) \leqslant qM(a+b)$ for all $a \geqslant 1$, $b \geqslant 1$. Then

$$\begin{aligned}
\frac{N(a) \cdot N(b)}{N(a+b)} &= \frac{M(a) \cdot M(b)}{M(a+b)} \cdot \frac{1}{q-1} \cdot q^{\frac{a(a+1)+b(b+1)-(a+b)(a+b+1)}{2}} \\
&\leqslant \frac{1}{q-1} \cdot q \cdot q^{-ab} = \frac{1}{q-1} \cdot q^{-ab+1} \\
&\leqslant \frac{1}{q-1} \leqslant 1.
\end{aligned}$$

Similar calculus holds for the other bounds. □

3. A formula to compute the number of irreducible polynomials

3.1. **Notations.** We denote by $I(d)$ the number of normalized irreducible polynomials of degree exactly d and by $R(d)$ the number of normalized reducible polynomials of degree exactly d. Of course we have:

$$N(d) = I(d) + R(d).$$

We will decompose the set of polynomials according to the number of irreducible factors. Let $S_k(d)$ be the number of normalized polynomials of degree exactly d having exactly k irreducible (maybe non-distinct) factors. Of course

$$S_1(d) = I(d)$$

and

$$S_2(d) + \cdots + S_d(d) = R(d).$$

3.2. **Torsion product.** Let $(\ell_1, \ldots, \ell_k) \in \mathbb{N}^k$ such that

$$\underbrace{\ell_{i_1} = \cdots = \ell_{i_1+\alpha_1-1}}_{\alpha_1} < \underbrace{\ell_{i_2} = \cdots = \ell_{i_2+\alpha_2-1}}_{\alpha_2} < \ldots < \underbrace{\ell_{i_r} = \cdots = \ell_k}_{\alpha_r}$$

where $i_1 = 1$.

We define the following product:

$$\ell_1 \otimes \ell_2 \otimes \cdots \otimes \ell_k = \binom{\ell_{i_1} + \alpha_1 - 1}{\alpha_1} \times \binom{\ell_{i_2} + \alpha_2 - 1}{\alpha_2} \times \cdots \times \binom{\ell_{i_r} + \alpha_r - 1}{\alpha_r}.$$

In another language this is number of ways to choose k objects from k boxes (combination with repetition), where the i-th box contains ℓ_i objects. Moreover if $\ell_i = \ell_j$ then boxes i and j contain the same objects and if $\ell_i \neq \ell_j$ they contain no common objects.

Let us remark that:

$$\ell_1 \otimes \cdots \otimes \ell_k \leqslant \ell_1 \times \cdots \times \ell_k.$$

3.3. **Partitions.** Let $\mathcal{P}(k,d)$ be the set of partitions of d into exactly k parts:

$$\mathcal{P}(k,d) = \big\{[d_1,d_2,\ldots,d_k] \mid 1 \leqslant d_1 \leqslant d_2 \cdots \leqslant d_k \text{ and } d_1+d_2+\cdots+d_k = d\big\}.$$

Then the set of partitions of d is:

$$\mathcal{P}(d) = \mathcal{P}(1,d) \cup \mathcal{P}(2,d) \cup \ldots \cup \mathcal{P}(d,d).$$

For example if $d = 5$ we have: $5 = 1+4 = 2+3 = 1+1+3 = 1+2+2 = 1+1+1+2 = 1+1+1+1+1$. Then

$$\mathcal{P}(5) = \big\{[5],[1,4],[2,3],[1,1,3],[1,2,2],[1,1,1,2],[1,1,1,1,1]\big\}.$$

Let $P(d) = \#\mathcal{P}(d)$, the asymptotic behaviour of $P(d)$ is given by a formula of Hardy and Ramanujan:

$$P(d) \sim \frac{1}{4d\sqrt{3}} \exp\left(\pi\sqrt{\frac{2d}{3}}\right).$$

We will need an upper bound, [1, p. 197], for all $d \geqslant 1$:

$$P(d) < \exp\left(\pi\sqrt{\frac{2d}{3}}\right).$$

3.4. **Formula.**

Lemma 4.

$$S_k(d) = \sum_{[d_1,\ldots,d_k]\in\mathcal{P}(k,d)} I(d_1) \otimes I(d_2) \otimes \cdots \otimes I(d_k).$$

Note that if $k \geqslant 2$ then all d_i that appear in this formula verify $d_i < d$.

Proof. In fact a normalized polynomial f of degree d with exactly k irreducible factors can be written $f = f_1 \times \cdots \times f_k$. This decomposition is unique (up to permutation) if we choose the f_i to be irreducible and normalized. If we denote by d_i the degree of f_i we have $d_1 + \cdots + d_k = d$. Then to a factorization we associate a partition $[d_1,\ldots,d_k]$ of d. And the number of polynomials having this partition is exactly $I(d_1) \otimes \cdots \otimes I(d_k)$. □

d	$N(d)$	$I(d)$	$\frac{I(d)}{N(d)}$	$1-\frac{3}{2^d}$
1	6	6	1	-0.5
2	56	35	0.625	0.25
3	960	694	0.72291...	0.625
4	31744	26089	0.82185...	0.8125
5	2064384	1862994	0.90244...	0.90625
6	266338304	253247715	0.95084...	0.95312...
7	68451041280	66799608630	0.97587...	0.97656...
8	35115652612096	34698378752226	0.98811...	0.98828...
9	35993612646875136	35781375988234520	0.99410...	0.99414...
10	73750947497819242496	73534241823793715433	0.99706...	0.99707...

TABLE 1. Number of irreducible polynomials in $\mathbb{F}_2[x,y]$.

3.5. **Algorithm.** Lemma 4 provides an algorithm to compute $I(d)$ recursively.

- Compute $I(1)$ by hand: $I(1) = N(1) = q(q+1)$.
- Assume that you have already computed $I(2), \ldots, I(d-1)$.
- Calculate the sets of partitions $\mathcal{P}(k,d)$, $2 \leqslant k \leqslant d$.
- Apply the recursive formula

$$\begin{aligned} I(d) = N(d) - R(d) &= N(d) - \sum_{k=2}^{d} S_k(d) \\ &= N(d) - \sum_{k=2}^{d} \sum_{[d_1,\ldots,d_k]\in\mathcal{P}(k,d)} I(d_1) \otimes I(d_2) \otimes \cdots \otimes I(d_k). \end{aligned}$$

Contrary to the one variable case it appears in Table 1 that the probability to choose an irreducible polynomials among polynomials of degree d tends to 1 as d tends to infinity. Moreover the speed of this convergence seems to be given by the formula of the introduction.

Some of these numbers appears in Sloane's *Encyclopedia of Integer Sequences* [10], for example the sequence $(I(d))_d = (6, 35, 694, \ldots)$ that gives the number of irreducible polynomials in $\mathbb{F}_2[x,y]$ is referenced as *A115457*. This algorithm is implemented (in any number of variables and in any field) in a Maple sheet available on author's web page [2].

4. Asymptotic value for the number of irreducible polynomials

Lemma 5. *For a partition $[d_1, d_2, \dots, d_k] \in \mathcal{P}(k,d)$ not equal to $[1, d-1]$ we have*

$$I(d_1) \otimes I(d_2) \otimes \cdots \otimes I(d_k) \leqslant q^6 \cdot N(d-2).$$

Proof. First remember from Section 3.2 that $I(d_1) \otimes \cdots \otimes I(d_k) \leqslant I(d_1) \times \cdots \times I(d_k) \leqslant N(d_1) \times \cdots \times N(d_k)$.

For the partition $[2, d-2]$ it gives

$$I(2) \otimes I(d-2) \leqslant N(2) \cdot N(d-2) \leqslant \frac{q^6 - q^3}{q-1} \cdot N(d-2) \leqslant q^6 \cdot N(d-2).$$

For a partition of type $[a, d-a]$, $a \geqslant 3$ then using Lemma 3-(3) it gives

$$I(a) \otimes I(d-a) \leqslant N(a) \cdot N(d-a) \leqslant q^5 \cdot N(d-2).$$

For a partition of type $[d_1, \dots, d_k]$ with $k \geqslant 3$, we apply twice Lemma 3-(2) and finish using Lemma 3-(1):

$$\begin{aligned} I(d_1) \otimes \cdots \otimes I(d_k) &\leqslant N(d_1) \times N(d_2) \times N(d_3) \times \cdots \times N(d_k) \\ &\leqslant q^3 \cdot N(d_1 + d_2 - 1) \cdot N(d_3) \times \cdots \times N(d_k) \\ &\leqslant q^3 \cdot q^3 \cdot N(d_1 + d_2 + d_3 - 2) \cdot N(d_4) \times \cdots \times N(d_k) \\ &\leqslant q^6 \cdot N(d_1 + d_2 + d_3 - 2 + d_4 + \cdots + d_k) \\ &\leqslant q^6 \cdot N(d-2). \end{aligned}$$

□

Lemma 5 above would enable us to prove that among reducible polynomials those associated to the partition $[1, d-1]$ in number $I(1) \otimes I(d-1)$ are predominant. This is the main idea for the proof of the next Lemma.

Lemma 6. *There exists $d_0 \geqslant 1$ such that for all $d \geqslant d_0$ we have*

$$1 - \frac{1}{d} \leqslant \frac{R(d)}{N(1) \cdot N(d-1)} \leqslant 1 + \frac{1}{d}.$$

Proof. Upper bound.

$R(d) = S_2(d) + \cdots + S_d(d)$ and each $S_k(d)$ is the sum of $I(d_1) \otimes \cdots \otimes I(d_k)$ over all partition $[d_1, \dots, d_k] \in \mathcal{P}(k,d)$. By Lemma 5 and putting apart the partition $[1, d-1]$ we get that $I(d_1) \otimes \cdots \otimes I(d_k) \leqslant q^6 \cdot N(d-2)$.

We recall that $P(d)$ is number of partition of d: $P(d) = \#\mathcal{P}(d) = \#(\mathcal{P}(1,d) \cup \ldots \cup \mathcal{P}(k,d))$, see Section 3.3. Then

$$\begin{aligned} R(d) &\leqslant I(1) \otimes I(d-1) + P(d) \cdot q^6 \cdot N(d-2) \\ &\leqslant N(1) \cdot N(d-1) + \exp\left(\pi\sqrt{\frac{2d}{3}}\right) \cdot q^6 \cdot N(d-2). \end{aligned}$$

Then

$$\begin{aligned} \frac{R(d)}{N(1) \cdot N(d-1)} &\leqslant 1 + \frac{q^6}{N(1)} \cdot \exp\left(\pi\sqrt{\frac{2d}{3}}\right) \cdot \frac{N(d-2)}{N(d-1)} \\ &\leqslant 1 + \frac{q^6}{q(q+1)} \cdot \exp\left(\pi\sqrt{\frac{2d}{3}}\right) \cdot \frac{1}{q^d}. \end{aligned}$$

Then there exists d_0' such that for all $d \geqslant d_0'$

$$(*) \qquad \frac{R(d)}{N(1) \cdot N(d-1)} \leqslant 1 + \frac{1}{d}.$$

Lower bound.

Among reducible polynomials of degree d there are polynomials of type $f_1 \cdot f_2$ where f_1 is an irreducible polynomials of degree 1 and f_2 is irreducible of degree $d-1$. This corresponds to the partition $[1, d-1] \in \mathcal{P}(2,d)$. The number of polynomials corresponding to the partition $[1, d-1]$ is equal to $I(1) \otimes I(d-1) = N(1) \cdot I(d-1)$.

Then for $d \geqslant d_0'$:

$$\begin{aligned} R(d) &\geqslant I(1) \otimes I(d-1) \\ &= N(1) \cdot I(d-1) \\ &= N(1)\big(N(d-1) - R(d-1)\big) \\ &\geqslant N(1) \cdot \left(N(d-1) - N(1) \cdot \left(1 + \frac{1}{d-1}\right) \cdot N(d-2)\right) \qquad \text{by } (*) \\ &= N(1) \cdot N(d-1) \cdot \left(1 - N(1) \cdot \left(1 + \frac{1}{d-1}\right) \cdot \frac{N(d-2)}{N(d-1)}\right) \\ &\geqslant N(1) \cdot N(d-1) \cdot \left(1 - N(1) \cdot \left(1 + \frac{1}{d-1}\right) \cdot \frac{1}{q^d}\right) \\ &\geqslant N(1) \cdot N(d-1) \cdot \left(1 - \frac{1}{d}\right) \qquad d \geqslant d_0, \text{ for a } d_0 \geqslant d_0' \end{aligned}$$

□

Proof of the main theorem. We are now ready to prove the theorem of the introduction:

$$\begin{aligned}1-\frac{I(d)}{N(d)} &= \frac{N(d)-I(d)}{N(d)} = \frac{R(d)}{N(d)}\\ &= \frac{R(d)}{N(d-1)}\cdot\frac{N(d-1)}{N(d)}\\ &\sim N(1)\cdot\frac{1}{q^{d+1}} = q(q+1)\cdot\frac{1}{q^{d+1}}\\ &\sim \frac{q+1}{q^d}.\end{aligned}$$

The first equivalence is obtained using Lemma 6 and Lemma 2. □

5. More variables

It is not hard to extend these results to polynomials in $\mathbb{F}_q[x_1,\ldots,x_m]$, with $m \geqslant 2$. In fact only results of section 2 have to be generalized, while the rest of the paper is still valid.

First of all the number $N_m(d)$ of normalized polynomials of degree exactly d in $\mathbb{F}_q[x_1,\ldots,x_m]$ involves some more advanced combinatorics:

$$N_m(d) = \frac{1}{q-1}\cdot\left(q^{\binom{m+d-1}{m-1}}-1\right)\cdot q^{\binom{m+d-1}{m}}.$$

We get that $N_m(1) = \frac{q^{m+1}-q}{q-1}$. Let $I_m(d)$ be the number of normalized irreducible polynomials in $\mathbb{F}_q[x_1,\ldots,x_m]$ of degree exactly d whose asymptotic behaviour of $I_m(d)$ as $d \to +\infty$ is describe by the next result.

Theorem 7.

$$1-\frac{I_m(d)}{N_m(d)} \sim N_m(1)\cdot\frac{N_m(d-1)}{N_m(d)} \sim \frac{q^{m+1}-q}{q-1}\cdot\frac{1}{q^{\binom{m+d-1}{m-1}}}.$$

For example in $\mathbb{F}_2[x,y,z]$ the number $I_3(d)$ of irreducible polynomials verifies:

$$1-\frac{I_3(d)}{N_3(d)} \sim \frac{14}{2^{\frac{(d+1)(d+2)}{2}}}.$$

References

[1] R. Ayoub, *An Introduction to the Analytic Theory of Numbers*, Amer. math. Soc., 1963.

[2] A. Bodin, `http://math.univ-lille1.fr/~bodin`.

[3] L. Carlitz, *The distribution of irreducible polynomials in several indeterminates.* Illinois J. Math. 7 (1963) 371–375.

[4] L. Carlitz, *The distribution of irreducible polynomials in several indeterminates. II.* Canad. J. Math. 17 (1965) 261–266.

[5] S. Cohen, *The distribution of irreducible polynomials in several indeterminates over a finite field.* Proc. Edinburgh Math. Soc. 16 (1968/1969) 1–17.

[6] S. Cohen, *Some arithmetical functions in finite fields.* Glasgow Math. J. 11 (1970) 21–36.

[7] S. Gao and A. Lauder, *Hensel lifting and bivariate polynomial factorisation over finite fields.* Math. Comp. 71 (2002), 1663–1676.

[8] R. Lidl and H. Niederreiter, *Finite fields*, Encyclopedia of mathematics and its application, Cambridge University Press, 1997.

[9] J.-F. Ragot, *Counting polynomials with zeros of given multiplicities in finite fields.* Finite Fields Appl. 5 (1999), 219–231.

[10] N. Sloane, The encyclopedia of integer sequences `http://www.research.att.com/~njas/sequences/`.

E-mail address: `Arnaud.Bodin@math.univ-lille1.fr`

Laboratoire Paul Painlevé, Mathématiques, Université Lille 1, 59655 Villeneuve d'Ascq Cedex, France