Polynomials over Finite Fields Free from Large and Small Degree Irreducible Factors

Theodoulos Garefalakis

Department of Mathematics, Royal Hollaway, University of London, Egham, Surrey TW20 0EX, United Kingdom.
E-mail: theo.garefalakis@rhul.ac.uk

and

Daniel Panario

School of Mathematics and Statistics, Carleton University, Ottawa, K1S 5B6, Canada.
E-mail: daniel@math.carleton.ca

We study the number $N_q(n, m_1, m_2)$ of polynomials of degree n over a finite field $\mathbb{F}_q$ with all irreducible factors of degree bigger than m_2 and less than or equal to m_1. Applying the saddle point method, we obtain estimates for $N_q(n, m_1, m_2)$ in the range $m_1 = o(n)$, which have the flavor of de Bruijn [6], Canfield, Erdös and Pomerance [3] and Friedlander [11] for the corresponding problem for integers. Our results have applications in computational number theory and cryptography [12], and include as a particular case the smooth polynomials studied by Odlyzko [17] and others.

1. INTRODUCTION

A well-known area of research in analytic number theory is the study of functions related with the decomposition of numbers into primes. Two of these functions have been largely studied. The Dickman function models numbers without large prime factors [6, 8]. An excellent survey on this topic is due to Hildebrand & Tenenbaum [14]. On the other hand, the Buchstab function covers the study of numbers without small prime factors [2, 5]. Both functions are defined as solutions of some particular difference-differential equations. Furthermore, that is the case for a class of functions related with number theory [15].

The Dickman and Buchstab functions underlay not only the study of numbers without large and small primes but also the one of decomposable

structures without large and small "irreducible" components. This connection appears in [13] for the Dickman function, and in [20, 21] for the Buchstab function.

In this paper we are interested in one particular decomposable structure: polynomials over finite fields. When dealing with polynomials over finite fields, the prime elements are of course the irreducible factors of the polynomial, and the size of the irreducible component is its degree. We study polynomials over finite fields free of large and small degree irreducible factors. Previous results on this direction were obtained by Car [4]. These studies are related to the one of Friedlander [11] for numbers free from small and large primes. It seems also plausible to extend our results to other decomposable combinatorial structures.

Apart from being a natural question form a mathematical standpoint, our interests on this problem also arise from a cryptographical application. As it is well-known, many algorithms for computing discrete logarithms in extensions of finite fields rely on finding smooth polynomials (all the irreducible factors have degree bounded by some value m); see, for instance the excellent surveys by Odlyzko [17, 18]. In this case, the factor base is formed by all irreducible polynomials of degree smaller or equal to m. We are interested in developing a generalized version of the index calculus method for the discrete logarithm problem in $\mathbb{F}_q$, when $q = p^n$, p is a small prime and $n \to \infty$; see [12]. Instead of considering smooth polynomials, let us form our factor base with all irreducible polynomials of degree between given bounds. In order to estimate the asymptotic running time of this version one has to provide estimates for the number of polynomials over $\mathbb{F}_q$ with all their irreducible factors in an interval.

We now describe the structure of the paper. In Section 2, we briefly review the cryptographical application related to this work. Our approach to solve the problem is presented in Section 3. In Section 4, two technical lemmas needed in the proofs of Theorem 6.1 and Corollary 6.1 (the main results of our paper) are presented (Lemmas 4.1 and 4.3). We dissect the study in two cases. The case when the lower bound of the interval is fixed and the upper bound is $o(n)$ is treated in Section 5. Our results (Theorems 5.1 and 5.2) are similar to those of de Bruijn [6] and Canfield, Erdös and Pomerance [3]. Then, in Section 6, we focus on ranges where the lower bound of the interval tends to infinity with n while the upper bound is $o(n)$. Our results are given in the form of an integral representation (Theorem 6.1), which we estimate asymptotically (Corollary 6.1). We also obtain a Canfield, Erdös and Pomerance type result by increasing the range of the estimate; this weakens the result (Theorem 6.2). Due to the nature of the application we have in mind [12], we are interested in the case where the upper bound of the interval is $o(n)$. In Section 7, we comment on other ranges for further work.

We finish the introduction by commenting on the methodology used along this paper. We start with a precise characterization of the polynomials of interest using generating functions (for an introductory survey on this issue, see [9]). Applying some technical lemmas, we give estimates for our generating functions in terms of the exponential integral. Finally, we extract coefficients via Cauchy integrals, which we estimate using saddle point approximations [7, 10].

2. CRYPTOGRAPHICAL APPLICATION

In this section we sketch a cryptographical application that started our interest in the subject of this paper; see [12] for more information on this application.

We are interested on the index calculus method for computing discrete logarithms in $\mathbb{F}_q$, where $q = p^n$, p is a prime, and $n > 1$. The elements in $\mathbb{F}_q$ can be represented as polynomials over $\mathbb{F}_p$ of degree smaller than n.

As it is well-known, the index calculus method depends on finding smooth polynomials, that is, polynomials such that all their irreducible factors have degree smaller than or equal to certain bound m. The algorithm starts by computing a large database formed by the discrete logarithms of all irreducible polynomials of degree smaller than or equal to m. This set of irreducible polynomials is called the *standard factor base*. Until our work, the standard factor base was the only one being used, and it was considered to be the natural choice; see, for instance, Odlyzko [17, 18].

We consider a variant of the index calculus method where the database contains the discrete logarithms of all irreducible polynomials with degree in certain range (m_2, m_1), with $m_2 < m_1$. Thus, the index calculus method depends now on finding polynomials that decompose into irreducible factors with degree in the interval (m_2, m_1). This implies that we have to provide estimates for the number of polynomials over $\mathbb{F}_q$ of degree n that completely decompose into irreducible factors with degree in the interval (m_2, m_1), $m_2 < m_1$. As we proved in [12], the upper bound m_1 of the interval is of the same order as in the original index calculus method. However, the lower bound of the interval, m_2, is a free parameter that can be chosen almost at will. We note that when m_2 is as small as possible we have the original index calculus method, and so, our variant can be seen as a generalization of this method. On the other hand, our method can be extended, under certain technical conditions, to factor bases formed by the union of several intervals (see [12], p. 1259). This greatly increase the possible factor bases to be considered and compared.

In practical terms, our results allow a tradeoff associated with m_2. In fact, smaller values of m_2 imply higher probabilities of success (when finding polynomials that factor into irreducibles with degree between m_2 and

m_1), but the space used in the factor base and the system of congruences to be solved are large. On the other hand, larger values of m_2 mean lower probabilities of success but small size of factor base and size of the system of congruences to be solved. Hence, there is some possible further improvements by tuning the value of m_2.

We should emphasize that our work is only a first step on producing generalized factor bases. Extensive computational experiments should be carried out to draw other conclusions.

3. THE GENERAL TECHNIQUE

We turn now to the problem of counting the number of monic polynomials over a finite field that are free of irreducible factors of small and large degree. We start by fixing the notation. Let q be a prime power, and $\mathbb{F}_q$ be the finite field with q elements. We use boldface letters, e.g., $\mathbf{f}$, $\mathbf{d}$, to denote polynomials over $\mathbb{F}_q$. In particular, $\mathbf{p}$ always denotes an irreducible polynomial, and $\mathbf{1}$ denotes the unit of $\mathbb{F}_q$. We are interested in the number of monic polynomials of degree n over $\mathbb{F}_q$ with all irreducible factors having degree greater than m_2 and less than or equal to m_1, which we denote by $N_q(n, m_1, m_2)$. In general both m_1 and m_2 can (and will) be functions of n.

The main idea of the proof is to find the generating function $P_{m_1,m_2}(z)$ of the numbers $N_q(n, m_1, m_2)$, and then estimate the coefficients. We express $P_{m_1,m_2}(z)$ in terms of $r_{m_1}(z) = \sum_{k>m_1} z^k/k$ or $r_{m_1,m_2}(z) = \sum_{k=m_2+1}^{m_1} z^k/k$ depending on the behaviour of m_2. Then we use Lemma 4.1 to express it in terms of the exponential integral. The coefficients of the generating function are given by Cauchy integral, which is estimated via the saddle point method. The application of the saddle point method is the main and most involved part of the proof.

Let $\mathcal{I}$ be the collection of all monic irreducible polynomials over $\mathbb{F}_q$. The set of monic polynomials with irreducible factors between m_2 and m_1 can be symbolically written as

$$\prod_{\mathbf{p}\in\mathcal{I},\ m_2<\deg(\mathbf{p})\le m_1} (\mathbf{1}+\mathbf{p}+\mathbf{p}^2+\ldots)\ -1 = \prod_{\mathbf{p}\in\mathcal{I},\ m_2<\deg(\mathbf{p})\le m_1} (\mathbf{1}-\mathbf{p})^{-1}\ -1.$$

Note that the -1 in the above formula will not affect anything, since $[z^n]\,1 = 0$ for $n \ge 1$. For simplicity we will drop it in the expression of the generating function that follows. Let z be a formal variable. The generating function is obtained immediately via the substitution $\mathbf{p} \to z^{\deg(\mathbf{p})}$

$$P_{m_1,m_2}(z) = \prod_{k=m_2+1}^{m_1} (1-z^k)^{-I_k}. \tag{1}$$

We denote by (m_1, m_2)–polynomials the ones with all irreducible factors of degree bigger than m_2 and less than or equal to m_1. The number of (m_1, m_2)–polynomials is given by the Cauchy coefficient formula

$$\begin{aligned} N_q(n, m_1, m_2) &= [z^n]P_{m_1,m_2}(z) = q^n[z^n]P_{m_1,m_2}\left(\frac{z}{q}\right) \\ &= \frac{q^n}{2\pi i}\int_{\mathcal{C}} P_{m_1,m_2}\left(\frac{z}{q}\right)\frac{dz}{z^{n+1}}, \end{aligned}$$

where the contour $\mathcal{C}$ is chosen to be $z = e^{-\alpha/n+i\theta}$, $-\pi \le \theta \le \pi$, and α is a parameter to be chosen later. The idea for this substitution first appears in the thesis of Gourdon [13]. The change of variable $z = e^{-h/n}$ implies $h = \alpha - in\theta$, and the limits of integration now are $(\alpha + in\pi, \alpha - in\pi)$. Therefore,

$$\begin{aligned} N_q(n, m_1, m_2) &= \frac{q^n}{2\pi i}\int_{\alpha+in\pi}^{\alpha-in\pi} -\frac{1}{n}P_{m_1,m_2}\left(\frac{e^{-h/n}}{q}\right)\frac{dh}{e^{-h}} \\ &= \frac{q^n}{2\pi i}\int_{\alpha-in\pi}^{\alpha+in\pi} P_{m_1,m_2}\left(\frac{e^{-h/n}}{q}\right)\frac{e^h}{n}dh. \end{aligned} \tag{2}$$

The last integral cannot be computed exactly. Instead, we will compute it asymptotically. The function $P_{m_1,m_2}(e^{-h/n}/q)$ will be approximated differently depending on whether m_2 is constant or an increasing function of n. Similarly, α will be chosen differently depending on m_2.

Fixed lower bound. In this case, m_2 is a constant, and one would expect the expression for $N_q(n, m_1, m_2)$ to be similar to that for $N_q(n, m_1)$ of m_1-smooth polynomials. We treat the generating function as follows. The generating function of the set of monic polynomials over $\mathbb{F}_q$ is

$$P(z) = \prod_{k=1}^{\infty}(1 - z^k)^{-I_k} = \frac{1}{1 - qz}.$$

Thus, $P_{m_1,m_2}(z)$ can be expressed as

$$P_{m_1,m_2}(z) = \prod_{k=1}^{m_2}(1 - z^k)^{I_k}P(z)\prod_{k>m_1}(1 - z^k)^{I_k}.$$

The last term in the product can be treated in the same way as in [20], Equation (3.2), and obtain

$$P_{m_1,m_2}(z) = \prod_{k=1}^{m_2} (1-z^k)^{I_k} \frac{1}{1-qz} \exp\left(-r_{m_1}^{[1]}(z) - \frac{r_{m_1}^{[2]}(z)}{2} - \frac{r_{m_1}^{[3]}(z)}{3} - \cdots \right), \tag{3}$$

where

$$r_{m_1}^{[j]}(z) = \sum_{k>m_1} I_k z^{kj}.$$

The well-known estimate $kI_k = q^k + O(q^{k/2})$ implies

$$r_{m_1}^{[1]}\left(\frac{z}{q}\right) = \sum_{k>m_1} \frac{z^k}{k} + O\left(q^{-2m_1/5}\right) \quad \text{for } |z| < q^{1/10},$$

and

$$\sup_{|z|<1} r_{m_1}^{[j]}\left(\frac{z}{q}\right) = O\left(q^{-m_1(j-1)}\right) \quad \text{for } j \geq 2.$$

As it will become clear later, the choice of α in Equation (2) will be such that $\alpha = o(n)$, and $\alpha < 0$. Thus, for n sufficiently large, the condition $|e^{-h/n}| < q^{1/10}$ holds. Then, from the above discussion we conclude

$$\begin{aligned} P_{m_1,m_2}\left(\frac{e^{-h/n}}{q}\right) &= f(h) \cdot \frac{e^{-r_{m_1}^{[1]}(e^{-h/n}/q)+o(1)}}{1-e^{-h/n}} \\ &= f(h) \cdot \frac{e^{-r_{m_1}(e^{-h/n})+o(1)}}{1-e^{-h/n}} \\ &= (1+o(1))f(h) \cdot \frac{e^{-r_{m_1}(e^{-h/n})}}{1-e^{-h/n}}, \end{aligned} \tag{4}$$

where $f(h) = \prod_{k=1}^{m_2}(1-\frac{e^{-kh/n}}{q^k})^{I_k}$, and $r_{m_1}(z) = \sum_{k>m_1} z^k/k$.

Lower bound tending to infinity. In this case, we express the generating function as follows.

$$\begin{aligned} P_{m_1,m_2}(z) &= \prod_{k=m_2+1}^{m_1} (1-z^k)^{-I_k} = \exp\left(-\sum_{k=m_2+1}^{m_1} I_k \log(1-z^k) \right) \\ &= \exp\left(\sum_{j=1}^{\infty} \frac{1}{j} \sum_{k=m_2+1}^{m_1} I_k z^{jk} \right) \end{aligned}$$

$$= \exp\left(r^{[1]}_{m_1,m_2}(z) + \frac{r^{[2]}_{m_1,m_2}(z)}{2} + \cdots\right)$$

where in this case

$$r^{[j]}_{m_1,m_2}(z) = \sum_{k=m_2+1}^{m_1} I_k z^{jk}, \quad j \geq 1.$$

Using now the estimate $kI_k = q^k + O(q^{k/2})$, we obtain for $|z| < q^{1/10}$

$$r^{[1]}_{m_1,m_2}\left(\frac{z}{q}\right) = \sum_{k=m_2+1}^{m_1} \frac{z^k}{k} + O\left(q^{-2m_2/5}\right),$$

and

$$r^{[j]}_{m_1,m_2}\left(\frac{z}{q}\right) = O\left(q^{(-9j/10+1)m_2}\right) = O\left(q^{-4m_2/5}\right), \qquad j \geq 2.$$

Again the choice of α in Equation (2) will be such that $\alpha = o(n)$, so for n large enough we have $|e^{-h/n}| < q^{1/10}$, therefore

$$\begin{aligned} P_{m_1,m_2}\left(\frac{e^{-h/n}}{q}\right) &= e^{r_{m_1,m_2}(e^{-h/n})+o(1)} \\ &= (1+o(1))e^{r_{m_1,m_2}(e^{-h/n})}, \end{aligned} \tag{5}$$

where $r_{m_1,m_2}(z) = \sum_{k=m_2+1}^{m_1} z^k/z$.

In both cases, in order to estimate the Cauchy integral we need first to estimate the expressions $r_{m_1}(z)$ and $r_{m_1,m_2}(z)$ at least in the cases of interest here. This is the subject of the next two lemmata.

4. TWO TECHNICAL LEMMATA

In this section we prove a pair of technical lemmata that we need for the proof of our main results. It will be crucial to estimate the part of the logarithm series between m_1 and m_2

$$r_{m_1,m_2}(z) = \sum_{k=m_2+1}^{m_1} \frac{z^k}{k}.$$

The following lemma, extension of that in [19], provides an estimate for $r_{m_1,m_2}(z)$ in terms of the exponential integral

$$E(a) = \int_a^\infty \frac{e^{-s}}{s} ds.$$

LEMMA 4.1. *Let $n, m_1, m_2 \in \mathbb{N}$, and assume that $n, m_1, m_2, n/m_1, n/m_2$ tend to infinity. Let $h = -\xi + i\tau$, with $\xi > 0$ and $\xi/n \to 0$.*
If $|\tau| \leq n/m_2$, then

$$r_{m_1,m_2}(e^{-h/n}) = E(m_2 h/n) - E(m_1 h/n) + O\left(\frac{\xi+\tau}{n} e^{\xi m_1/n}\right). \quad (6)$$

For any value of τ,

$$r_{m_1,m_2}(e^{-h/n}) = O\left(E(m_2 h/n) - E(m_1 h/n)\right). \quad (7)$$

Proof. Let $u_1 = n/m_1$ and $u_2 = n/m_2$. By definition of $r_{m_1,m_2}(z)$, we have

$$\begin{aligned} r_{m_1,m_2}(e^{-h/n}) &= \sum_{k=m_2+1}^{m_1} \frac{e^{-kh/n}}{k} = \sum_{k=m_2+1}^{m_1} \int_{h/n}^{\infty} e^{-ky} dy \\ &= \int_{h/n}^{\infty} \left(\sum_{k=m_2+1}^{m_1} e^{-ky} \right) dy \\ &= \int_{h/n}^{\infty} \frac{e^{-m_2 y}}{e^y - 1} dy - \int_{h/n}^{\infty} \frac{e^{-m_1 y}}{e^y - 1} dy \\ &= \int_{h/u_2}^{\infty} e^{-s} \frac{1/m_2}{e^{s/m_2} - 1} ds - \int_{h/u_1}^{\infty} e^{-s} \frac{1/m_1}{e^{s/m_1} - 1} ds. \end{aligned}$$

Consider now the integral

$$\int_{h/u}^{\infty} e^{-s} \frac{1/m}{e^{s/m} - 1} ds = \int_{h/u}^{\infty} \frac{e^{-s}}{s} \frac{s/m}{e^{s/m} - 1} ds,$$

where $u = n/m$, and let $\psi(z) = \frac{z}{e^z - 1}$, which is analytic for $|z| < 2\pi$. Then, the above integral can be written as

$$\begin{aligned} \int_{h/u}^{\infty} \frac{e^{-s}}{s} \psi\left(\frac{s}{m}\right) ds &= \int_{h/u}^{\infty} \frac{e^{-s}}{s} ds + \int_{h/u}^{\infty} \frac{e^{-s}}{s} \left(\psi\left(\frac{s}{m}\right) - 1\right) ds \quad (8) \\ &= E(h/u) + \int_{h/u}^{-h/u} \frac{e^{-s}}{s} \left(\psi\left(\frac{s}{m}\right) - 1\right) ds \\ &\quad + \int_{-h/u}^{\infty} \frac{e^{-s}}{s} \left(\psi\left(\frac{s}{m}\right) - 1\right) ds. \end{aligned}$$

We recall that $\Re(h) = -\xi < 0$. The term $E(h/u)$ is the main term in the approximation of Equation (6). To finish the proof of Equation (6) we need to bound the last two integrals. To that end we use the following sublemma.

LEMMA 4.2. *The function* $\psi(z) = \frac{z}{e^z-1}$ *is analytic for* $|z| < 2\pi$. *We have the following asymptotic estimates.*

1.For $|z| \to 0$, $\psi(z) = 1 + O(z)$.
2.For $|z| \to \infty$, $\psi(z) \to 0$.

Proof. The analyticity of ψ is obvious from its definition. When $|z| \to 0$, expanding e^z we obtain

$$\begin{aligned}\psi(z) - 1 &= \frac{z}{e^z - 1} - 1 = \frac{z}{z + O(z^2)} - 1 \\ &= \frac{1}{1 + O(z)} - 1 = O(z).\end{aligned}$$

When $|z| \to \infty$, clearly $\psi(z) \to 0$. ∎

For the first integral in Equation (8), we observe that for s in the range $(h/u, -h/u)$, we have

$$\begin{aligned}\frac{|s|}{m} &< \frac{1}{m}\frac{|h|}{u} \le \frac{1}{m}\left(m\frac{\xi}{n} + \frac{|\tau|}{u}\right) \\ &\le \frac{\xi}{n} + \frac{|\tau|}{n} \le \frac{\xi}{n} + \frac{1}{m_2},\end{aligned}$$

where in the last step we used the assumption $|\tau| \le u_2$ of the lemma. Since $\xi/n \to 0$, and $m_2 \to \infty$, we get $\frac{|s|}{m} \to 0$. Therefore, using Lemma 4.2 we have

$$\psi\left(\frac{s}{m}\right) - 1 = O\left(\frac{s}{m}\right).$$

This implies

$$\begin{aligned}\left|\int_{h/u}^{-h/u} \frac{e^{-s}}{s}\left(\psi\left(\frac{s}{m}\right) - 1\right) ds\right| &\le \frac{e^{\xi/u}}{|h/u|}\frac{2|h|}{u} O\left(\frac{h}{n}\right) \\ &= O\left(e^{\xi/u}\frac{h}{n}\right) = O\left(e^{\xi/u}\frac{\xi+\tau}{n}\right),\end{aligned}$$

where the last equality holds, since $|h| = O(\xi + \tau)$. For the second integral in Equation (8), we note that in the range $(-h/u, \infty)$ the function ψ is bounded. This follows from Lemma 4.2. Therefore,

$$\left| \int_{-h/u}^{\infty} \frac{e^{-s}}{s} \left(\psi\left(\frac{s}{m}\right) - 1 \right) ds \right| \leq O(1) E(-h/u).$$

Next we need to bound $E(-h/u)$. For that we establish the following bound, as the real part σ of the argument is positive,

$$E(\sigma + i\tau) = \int_{\sigma+i\tau}^{\infty} \frac{e^{-s}}{s} ds = \int_{\sigma+i\tau}^{\sigma} \frac{e^{-s}}{s} ds + \int_{\sigma}^{\infty} \frac{e^{-s}}{s} ds. \tag{9}$$

The second integral is $O(e^{-\sigma}/\sigma)$. The first integral, after the substitution $s = \sigma + iy$, becomes

$$\int_{\sigma+i\tau}^{\sigma} \frac{e^{-s}}{s} ds = ie^{-\sigma} \int_{\tau}^{0} \frac{e^{-iy}}{\sigma + iy} dy.$$

One can check now that the second integral is $O(e^{-\sigma} \log \tau)$.

Returning to the lemma now, we conclude that the integral in the range $(-h/u, \infty)$ is $O(e^{-\xi/u} \log(\tau/u))$, which is absorbed by the error term induced by the integral in the range $(h/u, -h/u)$ (recall that $\xi > 0$). Equation (6) now follows by subtraction, considering $u = u_2$ and $u = u_1$, and noting that $u_1 < u_2$. ∎

A similar lemma is needed for the study of *generalized smooth polynomials* (i.e., when m_2 is constant). The next lemma provides an approximation of the remainders of the logarithm series

$$r_m(z) = \sum_{k>m} \frac{z^k}{k}. \tag{10}$$

LEMMA 4.3. *Let $n, m \in \mathbb{N}$, and $h = -\xi + i\tau$, with $\xi > 0$ and $\xi/n \to 0$. If $|\tau| = o(n)$, then*

$$r_m(e^{-h/n}) = E(mh/n) + O\left(\frac{\xi + \tau}{n} e^{\xi m/n} \right). \tag{11}$$

For any value of τ,

$$r_m(e^{-h/n}) = O(E(mh/n)). \tag{12}$$

Proof. The proof is essentially the same as in the previous lemma. One needs only to notice that

$$\sum_{k>m} e^{-ky} = \frac{e^{my}}{e^y - 1}.$$

∎

The choice of the parameters of the above lemmas, although somewhat artificial, are made to fit exactly the saddle point method that is extensively used in the next sections.

5. FIXED LOWER BOUND

From this point on, we distinguish between the cases of m_2 constant, and m_2 tending to infinity. The reason for this distinction is that the expressions for the generating function are sufficiently different, so that the choices that we have to make in the process of estimating the Cauchy integral will be different as well. The method we use for estimating the integral is the saddle point method.

The number of generalized smooth polynomials, i.e., when m_2 is fixed, under certain conditions on m_1, can be expressed in terms of the Dickman function which also governs the behavior of smooth integers [6, 14]. The Dickman function is defined as the unique solution of the following difference-differential equation:

$$\begin{aligned} \rho(u) &= 1, \qquad 0 \le u \le 1 \\ u\rho'(u) &= -\rho(u-1), \qquad u > 1 \\ \rho(u) &\qquad \text{is continuous.} \end{aligned}$$

In our work, it is convenient to consider the integral representation given by the inverse Laplace transform

$$\rho(u) = \frac{1}{2\pi i} \int_{\alpha - i\infty}^{\alpha + i\infty} \frac{-e^{E(s)}}{s} e^{us}\, ds. \tag{13}$$

The following theorem gives an asymptotic estimate for $N_q(n, m_1, m_2)$, and it is similar to de Bruijn's result [6] for integers. The fact that m_2 is constant allows for an elementary argument. However, we take the opportunity to introduce the analytic methods that are fully required in order to prove Theorem 6.1.

THEOREM 5.1. *The number $N_q(n, m_1, m_2)$ of monic polynomials of degree n over $\mathbb{F}_q$ with all irreducible factors with degree greater than m_2 and*

less than or equal to m_1, with m_2 fixed and $\sqrt{n}\log n \ll m_1 \ll n$ satisfies

$$N_q(n, m_1, m_2) \sim q^n f(0)\rho\left(\frac{n}{m_1}\right),$$

where ρ is the Dickman function, and

$$f(h) = \prod_{k=1}^{m_2}\left(1 - \frac{e^{-kh/n}}{q^k}\right)^{I_k}.$$

Proof. From Equation (2) and Equation (4) we have

$$\begin{aligned} N_q(n, m_1, m_2) &= (1 + o(1))\frac{q^n}{2\pi i}\int_{\alpha - in\pi}^{\alpha + in\pi} f(h)\frac{e^{-r_{m_1}(e^{-h/n})}}{n(1 - e^{-h/n})}e^h dh \\ &= (1 + o(1))\frac{q^n}{2\pi i}\int_{\alpha - in\pi}^{\alpha + in\pi} f(h)\frac{e^{-r_{m_1}(e^{-h/n})}}{h}\frac{h/n}{1 - e^{-h/n}}e^h dh. \end{aligned}$$

Let $\phi(z) = \frac{z}{1-e^{-z}}$, that is analytic in $|z| < 2\pi$. We now concentrate on the above integral. One expects that the main contribution to the integral comes from the neighborhood of the real axis. Indeed, let $\delta = o(n)$, $h = \alpha + i\tau$, and consider the integrals

$$J_\delta(n, m_1, m_2) = \int_{\alpha - i\delta}^{\alpha + i\delta} f(h)\frac{e^{-r_{m_1}(e^{-h/n})}}{h}\frac{h/n}{1 - e^{-h/n}}e^h dh,$$

and

$$J_{tail}(n, m_1, m_2) = \int_{\delta < |\tau| \le n\pi} f(h)\frac{e^{-r_{m_1}(e^{-h/n})}}{h}\frac{h/n}{1 - e^{-h/n}}e^h dh.$$

Let us consider first $J_\delta(n, m_1, m_2)$. For the range $(\alpha - i\delta, \alpha + i\delta)$, we have $h/n \to 0$, and $\lim_{z\to 0}\frac{z}{1-e^{-z}} = 1$. As we will see later, $\alpha \gg \delta \sim \sqrt{u_1 \log(u_1)}$, where $u_1 = n/m_1$, and an application of Lemma 4.3 implies

$$J_\delta(n, m_1, m_2) = \exp\left(O\left(\frac{|\alpha|}{n}e^{|\alpha|/u_1}\right)\right)\int_{\alpha - i\delta}^{\alpha + i\delta} f(h)\frac{e^{-E(h/u_1)}}{h}e^h dh.$$

Furthermore, $|h/n| < |\alpha/n| + \pi < 2\pi$, provided that $\alpha/n = o(1)$ (which will be the case for our choice of α, if $\log n/m_1 = o(1)$), and ϕ is analytic

in that range, so it is bounded by a constant. Therefore, for $h = \alpha + i\tau$, using Lemma 4.3 we have

$$J_{tail}(n, m_1, m_2) = O(1) \int_{\delta < |\tau| \leq n\pi} f(h) \frac{e^{-O(E(h/u_1))}}{h} e^h dh.$$

Note now that $f(h)$ is analytic, and bounded by a constant in the range of interest. Specifically, $f(h) = O(1)$ for $\delta < |\tau| \leq n\pi$. Furthermore, the Taylor expansion of $f(h)$ around zero for $|\tau| < \delta$ yields

$$f(h) = f(0) + O\left(\frac{h}{n}\right) = f(0) + o(1).$$

Thus,

$$J_\delta(n, m_1, m_2) = (1 + o(1)) \exp\left(O\left(\frac{|\alpha|}{n} e^{|\alpha|/u_1}\right)\right) f(0) \int_{\alpha - i\delta}^{\alpha + i\delta} \frac{e^{-E(h/u_1)}}{h} e^h dh, \tag{14}$$

and

$$J_{tail}(n, m_1, m_2) = O(1) \int_{\delta < |\tau| \leq n\pi} \frac{e^{-O(E(h/u_1))}}{h} e^h dh. \tag{15}$$

It remains to estimate the above expressions. It is well-known that $e^{-E(s)}/s$ is the Laplace transform $\widehat{\rho}(s)$ of the Dickman function. The integrals in Equations (14) and (15) have been studied by several authors. We briefly sketch here the exposition by Tenenbaum (see [23], pp. 372-376). The key idea is to use the saddle point method to prove that the main contribution comes from $J_\delta(n, m_1, m_2)$, which is then estimated. For the method to go through we have to choose α to be the real solution of the equation

$$(-E(h/u_1) + h - \log h)' = 0.$$

The solution α satisfies $\alpha \sim -u_1 \log(u_1 \log u_1)$. Thus, the condition $\alpha = o(n)$, which was needed throughout the proof, holds if $\log n/m_1 \to 0$. For this choice of α it is then shown that the contribution of the tails is $O(u_1^{-1/2} + \exp(-u_1 \log^{-2} u_1 + 2 \log u_1))$ times the first integral, and is negligible provided $m_1 = o(n)$. The tails along the vertical line up to $+i\infty$ and $-i\infty$ are shown to be negligible. The argument is technical, and we will not go into the details for two reasons: it appears in [23], and a similar argument will be used in the next section for the case $m_2 \to \infty$.

Collecting the previous results we have

$$\begin{aligned} N_q(n, m_1, m_2) &\sim \frac{q^n}{2\pi i} \exp\left(O\left(\frac{|\alpha|}{n} e^{|\alpha|/u_1}\right)\right) f(0) \int_{\alpha - i\delta}^{\alpha + i\delta} \frac{e^{-E(h/u_1)}}{h} e^h dh \\ &\sim q^n \exp\left(O\left(\frac{|\alpha|}{n} e^{|\alpha|/u_1}\right)\right) f(0)\rho(u_1), \end{aligned}$$

where ρ is the Dickman function as defined in Equation (13).

Finally, under the assumption that $\sqrt{n}\log n/m_1 \to 0$, we have

$$O\left(\frac{|\alpha|}{n} e^{|\alpha|/u_1}\right) = O\left(\frac{u_1^2 \log^2 u_1}{n}\right) = O\left(\frac{n \log^2 n}{m_1^2}\right) = o(1).$$

∎

In the spirit of Canfield, Erdös and Pomerance, we can obtain a slightly weaker result that holds for a much larger range of values of m_1. Indeed, replacing the assumption $\sqrt{n}\log n/m_1 \to 0$ by $\log n/m_1 \to 0$, and using the notation $u_1 = n/m_1$, it follows from the proof of Theorem 5.1 that

$$N_q(n, m_1, m_2) \sim q^n \exp\left(\frac{|\alpha|}{n} e^{|\alpha|/u_1}\right) f(0)\rho(u_1).$$

It is known [6] that

$$\rho(u_1) = e^{-(1+o(1))u_1 \log u_1}.$$

Also it can be checked that,

$$e^{\frac{|\alpha|}{n} e^{|\alpha|/u_1}} = e^{O((u_1 \log u_1 \log n)/m_1)} = e^{o(1)u_1 \log u_1}.$$

Therefore, we have the following theorem.

THEOREM 5.2. *The number $N_q(n, m_1, m_2)$ of monic polynomials of degree n over $\mathbb{F}_q$ with all irreducible factors with degree between m_2 and m_1, with m_2 fixed and $\log n \ll m_1 \ll n$, satisfies*

$$N_q(n, m_1, m_2) = q^n e^{-(1+o(1))\frac{n}{m_1} \log \frac{n}{m_1}}.$$

6. LOWER BOUND TENDING TO INFINITY

We turn now to the case when both bounds m_1 and m_2 tend to infinity, with $m_1 = o(n)$. Let $N_q(n, m)$ be the number of m-smooth polynomials. Since $N_q(n, m_1, m_2) \leq N_q(n, m_1)$, the power series

$$P_{m_1,m_2}\left(\frac{z}{q}\right) = \sum_{n=0}^{\infty} N_q(n, m_1, m_2)\frac{z^n}{q^n},$$

is dominated by the power series

$$P_{m_1}\left(\frac{z}{q}\right) = \sum_{n=0}^{\infty} N_q(n, m_1)\frac{z^n}{q^n},$$

and therefore, its radius of convergence is greater or equal to the radius of convergence of P_{m_1}. Since $|z| = e^{-\alpha/n}$ is within the radius of convergence of P_{m_1} for $\alpha \sim -n/m_1 \log(n/m_1 \log n/m_1)$ (see [17, p.74]), it is certainly within the radius of convergence of P_{m_1,m_2} as well. This will turn out to be the choice of α when estimating the integral in Equation (2).

From Equations (2) and (5) we have

$$N_q(n, m_1, m_2) = \frac{(1 + o(1))q^n}{2\pi i}\int_{\alpha-in\pi}^{\alpha+in\pi} e^{r_{m_1,m_2}(e^{-h/n})}\frac{e^h}{n}dh. \qquad (16)$$

We focus on the above integral. Let us denote

$$J(n, m_1, m_2) \;=\; \frac{1}{2\pi i}\int_{\alpha-in\pi}^{\alpha+in\pi} e^{r_{m_1,m_2}(e^{-h/n})}\frac{e^h}{n}dh.$$

As usual in this paper, we write $u_1 = n/m_1$ and $u_2 = n/m_2$. We expect that the main contribution to the integral comes from the neighborhood of the real axis. If the saddle point method is to work, then we hope to approximate the term $r_{m_1,m_2}(e^{-h/n})$ close to the real axis by $E(h/u_2) - E(h/u_1) + O(\frac{-\alpha}{n}e^{-\alpha/u_1})$, according to Lemma 4.1, provided of course that the value of α satisfies the conditions of the lemma. Let

$$f(h) = E(h/u_2) - E(h/u_1) + h.$$

The value of α is determined as the real solution of the equation $f'(h) = 0$, that is,

$$f'(h) = 1 - \frac{e^{-h/u_2}}{h} + \frac{e^{-h/u_1}}{h} = 0.$$

The following lemma will be crucial for the application of the saddle point method.

LEMMA 6.1. *The equation*

$$\frac{e^{-h/u_2}}{h} - \frac{e^{-h/u_1}}{h} = 1 \tag{17}$$

has a negative real solution, $-\xi$*, such that* $\xi \sim u_1 \log(u_1 \log u_1)$*. More precisely,*

$$u_1 \log(u_1 \log u_1) < \xi < u_1 \log(u_1 (\log u_1)^2), \tag{18}$$

where $u_1 \leq c\, u_2$ *for any constant* $c < 1$*, and* $u_1 \to \infty$*.*

Proof. Consider the function $f'(h)$ defined above at the points $h_1 = -u_1 \log(u_1 \log u_1)$, and $h_2 = -u_1 \log(u_1 (\log u_1)^2)$. One can easily check that $f'(h_1) > 0$, and $f'(h_2) < 0$, and since $f'(h)$ is continuous, it follows that it has a zero in (h_2, h_1). ▌

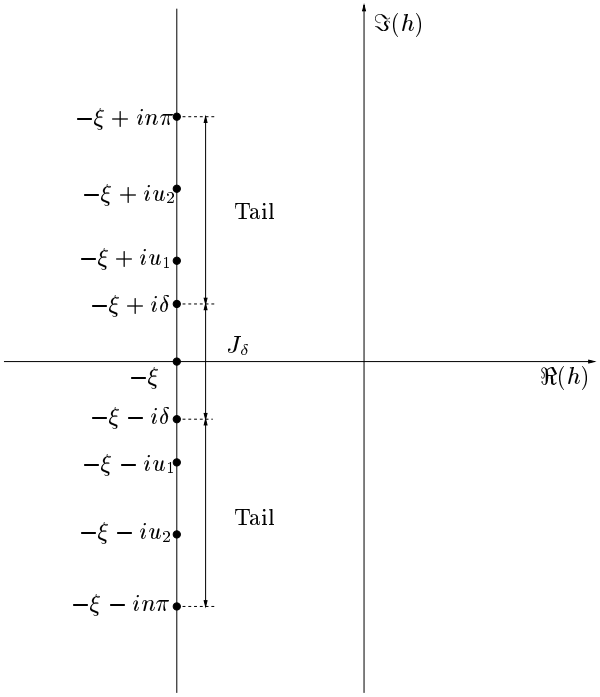

FIG. 1. Decomposition of the integral

To estimate $J(n, m_1, m_2)$ we choose $\alpha = -\xi$. We break $J(n, m_1, m_2)$ as the sum of two integrals along the vertical line with $\Im(h) = -\xi$, as shown in Figure 1

$$\begin{aligned} J(n, m_1, m_2) &= \frac{1}{2n\pi i} \int_{-\xi - i\delta}^{-\xi + i\delta} \exp\left(f(h)\right) dh \\ &\quad + \frac{1}{2n\pi i} \int_{\delta < |\Im(h)| < n\pi} \exp\left(f(h)\right) dh. \end{aligned}$$

Nearly all the contribution to the integral will come from the neighborhood of the point $-\xi$, with the contribution of the tail being negligible. To show that, we first estimate

$$\frac{1}{2n\pi i} \int_{-\xi - i\delta}^{-\xi + i\delta} e^{r_{m_1, m_2}(e^{-h/n})} \frac{e^h}{n} dh,$$

for a suitable δ. One can easily check now that the conditions of Equation (6) in Lemma 4.1 are satisfied — in fact the conditions were chosen to fit the proof. Substituting the estimate $f(h)$ for $r_{m_1, m_2}(e^{-h/n}) + h$, we obtain

$$\frac{1 + o(1)}{2n\pi i} \int_{-\xi - i\delta}^{-\xi + i\delta} \exp\left(f(h)\right) dh.$$

We need to estimate

$$J_\delta(n, m_1, m_2) = \frac{1}{2n\pi i} \int_{-\xi - i\delta}^{-\xi + i\delta} \exp\left(f(h)\right) dh. \tag{19}$$

We need the first three derivatives at $h = -\xi$. By the definition of ξ, we have

$$f'(-\xi) = 0.$$

For the second derivative we can write

$$\begin{aligned} f''(-\xi) &= \frac{1}{-\xi} \left(\frac{e^{\xi/u_2}}{u_2} - \frac{e^{\xi/u_1}}{u_1} + 1 \right) \\ &\sim \frac{1}{u_1 \log(u_1 \log u_1)} \left(\frac{(u_1 \log u_1)^{u_1/u_2}}{u_2} - \log u_1 + 1 \right) \\ &\sim \frac{-\log u_1}{u_1 \log(u_1 \log u_1)} \sim \frac{-1}{u_1}, \end{aligned} \tag{20}$$

where we used the facts that $\xi \sim u_1 \log(u_1 \log u_1)$, that $-\xi$ satisfies Equation (17), and that $u_2 \geq c u_1$ for some $c < 1$. Finally, the third derivative

can be computed and shown to be

$$f^{'''}(-\xi) = O\left(\frac{e^{\xi/u_1}}{u_1^2\xi}\right) = O\left(\frac{1}{u_1^2}\right).$$

It follows that

$$f(-\xi+it) = f(-\xi) - \frac{f^{''}(-\xi)}{2}t^2 + O\left(f^{'''}(-\xi)t^3\right).$$

Taking $\delta = u_1^{1/2}\log u_1$ and using Lemma 6.1, we obtain $t^3 f^{'''}(-\xi) = O\left(u_1^{-1/2}(\log u_1)^3\right)$. Under the change of variable $h = -\xi + it$, we have

$$\begin{aligned} J_\delta(n, m_1, m_2) &= \frac{1}{2n\pi}\int_{-\delta}^{\delta} \exp\left(f(-\xi) - t^2 f^{''}(-\xi)/2 + t^3 O\left(f^{'''}(-\xi)\right)\right) dt \\ &= \frac{\exp\left(f(-\xi)\right)}{2n\pi}\int_{-\delta}^{\delta}\left(1 + t^3 O\left(f^{'''}(-\xi)\right)\right)\exp\left(-t^2 f^{''}(-\xi)/2\right) dt. \end{aligned}$$

The term containing t^3 is $o(1)$. Therefore, we obtain

$$\begin{aligned} J_\delta(n, m_1, m_2) &= (1+o(1))\frac{\exp\left(f(-\xi)\right)}{2n\pi}\int_{-\delta}^{\delta}\exp\left(-t^2 f^{''}(-\xi)/2\right) dt \\ &= (1+o(1))\frac{\exp\left(f(-\xi)\right)}{n\sqrt{2\pi|f^{''}(-\xi)|}}. \end{aligned} \tag{21}$$

For the tails now, we will use the following technical lemma.

LEMMA 6.2. *For $h = -\xi + it$, $t \in \mathbb{R}$, and $g(h) = O(E(h/u_2) - E(h/u_1)) + h$, we have*

$$\exp(g(h)) = \exp\left(-\xi + o\left(u_1\right) + it\right), \qquad |t| \geq u_2;$$

$$\exp(\Re(f(h))) \leq \exp\left(f(-\xi) - \frac{Kt^2}{u_1}\right), \qquad |t| \leq u_1;$$

$$\exp(\Re(f(h))) \leq \exp\left(f(-\xi) - \frac{u_1}{(\log u_1)^2 + \pi^2} + O(1)\right), \qquad u_1 < |t| < u_2. \tag{22}$$

Proof. We note that $E(s) = o(e^{-\sigma}/|\tau|)$, for $s = \sigma + i\tau$ (see [23], p. 373). Therefore, for $|t| \geq u_2$

$$\begin{aligned}\exp(g(h)) &= \exp\left(O\left(e^{\xi/u_2}u_2/|t|\right) - O\left(e^{\xi/u_1}u_1/|t|\right) - \xi + it\right) \\ &= \exp\left(-\xi + o\left(u_1\right) + it\right),\end{aligned}$$

where we need to assume $u_1 \log u_1 = o(u_2)$, which is satisfied if $m_2 = o(m_1/\log n)$.

The second equation is more involved. We start by checking that (see [1], 5.1.37)

$$\begin{aligned}f(-\xi) - \Re(f(h)) &= \int_0^1 \frac{e^{\xi y/u_1}(1-\cos(ty/u_1))}{y}dy \\ &\quad - \int_0^1 \frac{e^{\xi y/u_2}(1-\cos(ty/u_2))}{y}dy \\ &\geq \int_0^1 \frac{e^{\xi y/u_1}(\cos(ty/u_2)-\cos(ty/u_1))}{y}dy. \qquad (23)\end{aligned}$$

For $|t| \leq u_1$ and an appropriate constant $K^{'}$, we have

$$\cos(ty/u_2) - \cos(ty/u_1) \geq \frac{K^{'}t^2y^2}{u_1^2}. \qquad (24)$$

Using Equations (23) and (24), we have for a constant K

$$\begin{aligned}f(-\xi) - \Re(f(h)) &\geq \int_0^1 e^{\xi y/u_1}\frac{K^{'}t^2y}{u_1^2}dy = \frac{K^{'}t^2}{u_1^2}\int_0^1 ye^{\xi y/u_1}dy \\ &= \frac{K^{'}t^2}{u_1^2}\frac{e^{\xi/u_1}(\xi-u_1)u_1+u_1^2}{\xi^2} \sim \frac{Kt^2}{u_1}.\end{aligned}$$

The bound in Equation (22) can be shown in a similar way. We work in the range $u_1 < |t| < u_2$. We observe that

$$\begin{aligned}f(-\xi) - \Re(f(h)) &= \int_0^1 \frac{e^{\xi y/u_1}(1-\cos(ty/u_1))}{y}dy \\ &\quad - \int_0^1 \frac{e^{\xi y/u_2}(1-\cos(ty/u_2))}{y}dy \\ &\geq \frac{u_1}{(\xi/u_1)^2+\pi^2} + O(1). \qquad (25)\end{aligned}$$

Indeed, we refer to the proof in [23] p. 374 that the first integral in Equation (25) is greater than or equal to $u_1/((\xi/u_1)^2+\pi^2)$. Next we give a

lower bound for the second integral

$$\int_0^1 \frac{e^{\xi y/u_2}(1-\cos(ty/u_2))}{y}dy \;=\; \frac{t}{u_2}\int_0^1 e^{\xi y/u_2}\frac{1-\cos(ty/u_2)}{ty/u_2}dy. \quad (26)$$

For the range $u_1 < |t| < u_2$, and since $0 < y < 1$, we have $|ty/u_2| < 1$. For $x < 1$, from the Taylor expansion of $cos(x)$ we have

$$cos(x) = 1 - O(x^2),$$

which implies that

$$\frac{1-cos(x)}{x} = O(x).$$

Applying this to the integral in Equation (26), we obtain

$$\begin{aligned}\int_0^1 \frac{e^{\xi y/u_2}(1-\cos(ty/u_2))}{y}dy \;&=\; O(1)\frac{t}{u_2}\int_0^1 \frac{ty}{u_2}e^{\xi y/u_2}dy \\ &= O(1)\frac{t^2}{u_2^2}\int_0^1 ye^{\xi y/u_2}dy \le O(1)\frac{t^2}{u_2^2}\int_0^1 e^{\xi y/u_2}dy \\ &= O(1)\frac{t^2}{u_2^2}\frac{u_2}{\xi}(e^{\xi/u_2}-1) = \frac{t^2}{u_2^2}\frac{u_2}{\xi}O\left(\frac{\xi}{u_2}\right) \le O(1)\end{aligned}$$

where we used the assumption $u_1 \log u_1 = o(u_2)$, which implies that $\xi/u_2 \to 0$, and therefore $e^{\xi/u_2} - 1 = O(\xi/u_2)$. Also in this range $|t/u_2| < 1$ which proves Equation (25) and concludes the proof. ▌

We break up the tails in three parts: $|t| \ge u_2$, , $u_1 < |t| < u_2$ and $\delta < |t| \le u_1$. For the first range, the contribution is negligible due to the above lemma, and the following easy facts:

$$u_1 = O\left(E\left(\frac{-\xi}{u_2}\right) - E\left(\frac{-\xi}{u_1}\right)\right),$$

and

$$f(-\xi) = -\xi + E\left(\frac{-\xi}{u_2}\right) - E\left(\frac{-\xi}{u_1}\right).$$

We concentrate now in the range $\delta < |t| \le u_1$. By the above lemma, the tail in this range is upper bounded by

$$\int_{\delta<|t|\le u_1} \exp\left(f(-\xi) - \frac{K}{u_1}t^2\right)dt \;\le\; \exp\left(f(-\xi)\right)\int_\delta^\infty e^{-\frac{K}{u_1}t^2}dt$$

$$\begin{aligned} &\leq \frac{\exp\left(f(-\xi)\right)\sqrt{u_1}}{2\sqrt{K}} \int_{\delta^2 K/u_1}^{\infty} e^{-y}\frac{dy}{\sqrt{y}} \\ &\leq \frac{\exp\left(f(-\xi)\right)u_1}{2\delta K} e^{-K\delta^2/u_1} \\ &= \frac{\exp\left(f(-\xi)\right)\sqrt{u_1}}{2K\log(u_1)} e^{-K(\log u_1)^2} \end{aligned}$$

which is clearly acceptable.

For the intermediate range $u_1 < |t| < u_2$, the tails are bounded by

$$\exp\left(e^{\xi/u_1}\frac{u_2}{n}\right)\exp\left(f(-\xi)\right)e^{-\frac{u_1}{(\log u_1)^2+\pi^2}}(u_2-u_1),$$

which again is acceptable provided that the above expression is $o(\sqrt{u_1})$. This holds by the hypothesis of Theorem 6.1 since $m_2 \gg \log^3 n$ and $m_2 \gg \sqrt{nm_1}e^{-n/(m_1(\log n)^2)}$. Putting all pieces together, we have proven that

$$N_q(n,m_1,m_2) = \frac{q^n}{2n\pi i} e^{O\left(e^{\xi/u_1}(\xi+\delta)/n\right)} \int_{-\xi-i\infty}^{-\xi+i\infty} e^{E(h/u_2)-E(h/u_1)}e^h dh.$$

Under the condition $\sqrt{n}\log n \ll m_1$, we have

$$e^{O\left(e^{\xi/u_1}(\xi+\delta)/n\right)} = 1 + o(1).$$

We have proven the following theorem.

THEOREM 6.1. *The number $N_q(n,m_1,m_2)$ of monic polynomials over $\mathbb{F}_q$ with all irreducible factors with degree between m_1 and m_2, $m_1, m_2 \to \infty$ satisfying $\sqrt{n}\log n \ll m_1 \ll n$, and $\max\{\log^3 n, \sqrt{nm_1}e^{-n/(m_1(\log n)^2)}\} \ll m_2 \ll m_1/\log n$, is asymptotically*

$$N_q(n,m_1,m_2) \sim \frac{q^n}{2n\pi i}\int_{-\xi-i\infty}^{-\xi+i\infty} e^{E(m_2h/n)-E(m_1h/n)}e^h dh,$$

where ξ is given in Equation (18) of Lemma (6.1).

The proof of Theorem 6.1 gives more than the integral form stated. The following corollary gives the asymptotic estimate obtained in the proof of Theorem 6.1. The estimate is stated in terms of the integral

$$Ei(x) = \int_{-\infty}^{x} \frac{e^t}{t}\,dt \qquad (x>0).$$

The reason for this is that $Ei(x)$ is a real valued function, and the final result is more natural expressed in that way.

COROLLARY 6.1. *The number $N_q(n, m_1, m_2)$ of monic polynomials over $\mathbb{F}_q$ with all irreducible factors with degree between m_1 and m_2, $m_1, m_2 \to \infty$, $\sqrt{n}\log n \ll m_1 \ll n$, and $\max\{\log^3 n, \sqrt{nm_1}e^{-n/(m_1(\log n)^2)}\} \ll m_2 \ll m_1/\log n$, is asymptotically*

$$N_q(n, m_1, m_2) \sim \frac{q^n\sqrt{m_1}}{\sqrt{2n\pi}} \exp\left(Ei(\xi m_1/n) - Ei(\xi m_2/n) - \xi\right),$$

where ξ is given in Equation (18) of Lemma (6.1).

Proof. From the proof of Theorem 6.1 we have

$$N_q(n, m_1, m_2) \sim q^n J_\delta(n, m_1, m_2).$$

This combined with Equation (21) implies that

$$N_q(n, m_1, m_2) \sim \frac{q^n \exp\left(f(-\xi)\right)}{n\sqrt{2\pi|f''(-\xi)|}}.$$

As it was pointed out in Equation (20), $f''(-\xi) \sim -1/u_1$, where again $u_1 = n/m_1$. Thus, it only remains to estimate $f(-\xi)$

$$\begin{aligned} f(-\xi) &= E(-\xi/u_2) - E(-\xi/u_1) - \xi \\ &= -Ei(\xi/u_2) + Ei(\xi/u_1) - \xi, \end{aligned}$$

where the second equality holds since $E(x + i0) = -Ei(-x) - i\pi$ (see [1], 5.1.7). ∎

Again, as for the generalized smooth polynomials, one can extend the range of the estimate considerably, by weakening the result.

THEOREM 6.2. *The number $N_q(n, m_1, m_2)$ of monic polynomials over $\mathbb{F}_q$ with all irreducible factors between m_1 and m_2, with $m_1, m_2 \to \infty$, $m_1 e^{-n/m_1} \ll m_2 \leq cm_1$ for any constant $c < 1$, and $2(\log n)^2 \leq m_1 \ll n$ satisfies*

$$N_q(n, m_1, m_2) = q^n e^{-(1+o(1))\frac{n}{m_1}\log\frac{n}{m_1}}.$$

Proof. Let $u_1 = n/m_1$ and $u_2 = n/m_2$. The number $N_q(n, m_1, m_2)$ was estimated in terms of four integrals that correspond to the ranges $|t| \leq \delta$, $\delta < |t| \leq u_1$, $u_1 < |t| < u_2$, and $|t| \geq u_2$. From the proof of Theorem 6.1, it is clear that the main integral corresponding to the range

$|t| \leq \delta$ is $e^{-(1+o(1))u_1 \log u_1}$ with the only assumption that $\log n \ll m_1 \ll n$. Moreover, the tail integrals that correspond to the ranges $|t| \geq u_2$ and $\delta < |t| \leq u_1$ are $e^{-(1+o(1))u_1 \log u_1}$ under no further assumption. The rest of the conditions come from the range $u_1 < |t| < u_2$. In that range the tail is

$$\begin{aligned} & \exp\left(e^{\xi/u_1}\frac{u_2}{n}\right)\exp\left(f(-\xi)\right)e^{-\frac{u_1}{(\log u_1)^2+\pi^2}}(u_2-u_1) \\ = \; & \exp\left(f(-\xi)\right)e^{-\frac{u_1}{(\log u_1)^2+\pi^2}+\frac{u_1 u_2 \log u_1}{n}+\log(u_2-u_1)}. \end{aligned}$$

We know that $f(-\xi) = -(1+o(1))u_1 \log u_1$, and all the other terms in the exponent are $o(1)u_1 \log u_1$, except maybe $\log u_2$. In order to ensure this, we need to impose the condition $m_2 \gg m_1 e^{-n/m_1}$. The theorem now follows. ∎

7. CONCLUSION AND FURTHER WORK

We studied the problem of estimating the number of polynomials of degree n over $\mathbb{F}_q$ that have irreducible factors of degree greater than m_2 and less than or equal to m_1. Our results hold for certain ranges of values for m_1 and m_2. For example, if $m_2 = 0$, then we have proven

$$N_q(n, m_1, m_2) \sim q^n \rho\left(\frac{n}{m_1}\right), \tag{27}$$

only if $\sqrt{n}\log n << m_1$. Can the range of m_1 for which the Dickman function appears, be extended? If not, one might be able to prove that for smaller values of m_1, $N_q(n, m_1, m_2)$ is asymptotically different than the expression in Equation (27), by proving a stronger version of Lemma 4.1. Using this, one could also extend the range for m_1 in Theorem 6.1.

Also the case $m_1 = n/c$, $c \geq 1$, was not considered at all. When $c = 1$, that is $m_1 = n$, it has been proven [20] that he number of polynomials with irreducible factors with degrees in that range are related with the Buchstab function. When $c > 1$, using the same arguments as in Section 6, one finds that Equation (16) holds. Furthermore, the approximation of Lemma 4.1 will work equally well. However, the saddle point method fails because of technical reasons (basically due to the fact that the corresponding value of ξ is constant). It would be interesting to solve the case $m_1 = n/c$, $c > 1$, and thus complete the range.

Another interesting problem is the study of similar estimates for other decomposable structures. These studies would provide estimates, for instance, for the number of permutations that decompose into cycles of length within a certain interval, or the unmber of 2-regular graphs that decompose into connected components of size between two bounds, and so on.

In addition to the previous mathematical questions, we are interested in the application of the results of this paper to cryptography. A first step in this direction has appeared in [12].

ACKNOWLEDGMENTS

This work was done while the authors were with the Department of Computer Science of the University of Toronto. Daniel Panario was supported by NSERC of Canada grant number 238757. We want to thank Prof. Manstavičius for useful comments about his related work.

REFERENCES

1. M. Abramowitz and I. Stegun. *Handbook of mathematical functions.* Dover, New York, 1970.
2. A.A. Buchstab. Asymptotic estimates of a general number theoretic function. *Mat. Sbornik*, 44:1239–1246, 1937.
3. E. R. Canfield, P. Erdös, and C. Pomerance. On a problem of Oppenheim concerning 'Factorisatio Numerorum'. *J. Number Theory*, 17:1–28, 1983.
4. M. Car. Théorèmes de densité dans $\mathbb{F}_q[x]$. *Acta Arith.*, 48:145–165, 1987.
5. N.G. de Bruijn. On the number of uncancelled elements in the sieve of Eratosthenes. *Indag. Math.*, 12:247–256, 1950.
6. N.G. de Bruijn. On the number of positive integers $\leq x$ and free of prime factors $> y$. *Indag. Math.*, 13:2–12, 1951.
7. N.G. de Bruijn. *Asymptotic methods in analysis.* Dover, New York, 1981.
8. K. Dickman. On the frequency of numbers containing prime factors of a certain relative magnitude. *Ark. Mat. Astr.Fys.*, 22:1–14, 1930.
9. P. Flajolet, X. Gourdon, and D. Panario. The complete analysis of a polynomial factorization algorithm over finite fields. *J. of Algorithms*, 40:37–81, 2001.
10. P. Flajolet and R. Sedgewick. Analytic combinatorics. In preparation: http://algo.inria.fr/flajolet/Publications/books.html
11. J. Friedlander. Integers free of large and small primes. *Proc. London Math. Soc.*, 3:565–576, 1976.
12. T. Garefalakis and D. Panario. The index calculus method using non-smooth polynomials. *Mathematics of Computation*, 70:1253–1264, 2001.
13. X. Gourdon. *Combinatoire, algorithmique et géométrie des polynômes.* PhD thesis, École Polytechnique, 1996.
14. A. Hildebrand and G. Tenenbaum. Integers without large prime factors. *J. de Théorie des Nombres*, 5:411–484, 1993.
15. A. Hildebrand and G. Tenenbaum. On a class of differential-difference equations arising in number theory. *J. DÁnalyse Mathématique*, 61:145–179, 1993.
16. E. Manstavičius. Semigroup elements free of large prime factors. In *New Trends in Probability and Statistics*, pages 135–153. VSP/TRV, 1992.
17. A. M. Odlyzko. Discrete logarithms in finite fields and their cryptographic significance. *Advances in Cryptology: Proceedings of EUROCRYPT 84, T. Beth, N. Cot, and I. Ingemarsson (eds.), Springer-Verlag, Lecture Notes in Computer Science*, 209:224–314, 1985.

18. A. M. Odlyzko. Discrete logarithms and smooth polynomials. In G.L. Mullen and P. J.-S. Shiue, editors, *Finite fields: theory, applications and algorithms*, pages 269–278. Contemporary Mathematics 168, Amer. Math. Soc., 1994.

19. D. Panario, X. Gourdon, and P. Flajolet. An analytic approach to smooth polynomials over finite fields. In *ANTS: 3rd International Algorithmic Number Theory Symposium*, volume 1423 of *LNCS*, pages 226–236. Springer-Verlag, 1998.

20. D. Panario and L.B. Richmond. Analysis of Ben-Or's polynomial irreducibility test. *Random Structures and Algorithms*, 13:439–456, 1998.

21. D. Panario and L.B. Richmond. Smallest components in decomposable structures: exp-log class. *Algorithmica*, 29:205–226, 2001.

22. K. Soundararajan. Asymptotic formulas for the counting function of smooth polynomials. Preprint.

23. G. Tenenbaum. *Introduction to analytic and probabilistic number theory*. Cambridge University Press, 1996.