J. Ramanujan Math. Soc. 14 No. 1 (1999) pp. 21–35

The Turán Sieve Method and Some of Its Applications

Yu-Ru Liu
Department of Mathematics
Harvard University
Cambridge, Mass., 02138, USA
E-mail: yrliu@math.harvard.edu

M. Ram Murty*
Department of Mathematics and Statistics
Queen's University
Kingston Ontario, K7L 3N6, Canada
E-mail: murty@mast.queensu.ca

Communicated by V. Kumar Murty

Abstract

We introduce the Turán sieve method and apply it to the probabilistic Galois theory problems in both the rational number field and the function field cases. We estimate the number of polynomials of degree n and height $\leq N$ whose Galois group is a proper subgroup of S_n. For the rational number field case, we get an estimate of $O(N^{n-1/3}(\log N)^2)$ and in the case of the function field over $\mathbb{F}_q$, we get $O(N^{n-1}\log_q N)$.

1. Introduction

In 1934, Paul Turán gave a very simple proof of the celebrated theorem of Hardy and Ramanujan that the normal number of distinct prime factors of a natural number n is $\log\log n$. If $\nu(n)$ denotes the number of distinct prime factors of n, Turán proved that

$$\sum_{n\leq x}(\nu(n)-\log\log n)^2 = O(x\log\log x)$$

* Research partially supported by a Killam Fellowship.

from which the normal order of $\nu(n)$ is easily deduced. The summation on the left hand side looks very much like the variance of the 'random variable' $\nu(n)$. Indeed, this similarity was amplified by Linnik when he developed the powerful large sieve method. However, Turán's original derivation of the Hardy-Ramanujan Theorem has concealed in it, in seed form, an elementary sieve method. Undoubtedly, many of the experts are aware of this and have in fact used the method (dubbed 'the normal order method' in Hooley [5]) to solve cognate problems. Nowhere in the literature is the inherent sieve method in Turán's derivation clearly exposed.

The purpose of this paper is to derive the 'Turán sieve method' and indicate a few of its applications. Most notably, we apply the method to determine the probability that a random polynomial of degree n, either in $\mathbb{Z}[x]$ or $\mathbb{F}_q[t,x]$, has Galois group equal to S_n. This was treated by Gallagher in the $\mathbb{Z}[x]$ case by using the large sieve method. We also treat this case. Though our result is weaker than Gallagher's, our method is far simpler. It is surprising that in the function field case (which is not treated by Gallagher) the elementary Turán sieve gives the correct exponent in the estimate. (See Theorem 3).

There are further applications of the results that are not touched upon in this paper. Most notable is the counting of integral points on algebraic varieties. For example, if V is an irreducible non-linear algebraic variety of dimension n over $\mathbb{Q}$, then one can show using the Turán sieve below that the number of integral points on V whose 'height' (as defined in [10]) is less than N is $O(N^{n-1/3}(\log N)^2)$ for $n \geq 2$. Though better results have been obtained by Cohen in [2] and Serre [10], we feel that our approach is simpler in that it avoids the large sieve. We do not give the details of this here but refer the reader to [10] with the remark that where the large sieve is used, the Turán sieve can be used instead.

We now describe Turán's sieve in an abstract setting.

Let S be a finite set and I be an index set. For each $i \in I$, we use $\Omega(i)$ to denote some specified conditions and define

$$S_i = \{s \in S : s \text{ satisfies } \Omega(i)\}.$$

For each $s \in S$, we define

$$\pi_s(I) = \#\{i \in I : s \text{ satisfies } \Omega(i)\}.$$

Since S is a finite set, we can write

$$|S_i| = \delta_i |S| + e_i,$$

where we think of e_i as an error term measuring the deviation of $|S_i|$ from $\delta_i |S|$. The δ_i's are arbitrary positive real numbers which we think of as approximating the proportion of elements of S satisfying condition $\Omega(i)$.

For different $i, j \in I$, we assume that

$$|S_i \cap S_j| = \delta_i \delta_j |S| + e_{i,j},$$

where $e_{i,j}$ is to be viewed as an error term.

Theorem 1 (The Turán sieve). *Let* $\nu = \sum_{i \in I} \delta_i$, *then*

$$\sum_{s \in S} (\pi_s(I) - \nu)^2 = |S| \sum_{i \in I} \delta_i (1 - \delta_i) + \sum_{i,j \in I} e_{i,j} - 2\nu \sum_{i \in I} e_i$$

here we use the convention that $e_{i,i} = e_i$.

A simple consequence follows from the Turán sieve. We immediately get an estimate for the number of elements of S not satisfying any of the conditions $\Omega(i)$ for $i \in I$ (and hence the appellation 'sieve').

Corollary 1.

$$\#\{s \in S : \pi_s(I) = 0\} \leq \frac{|S|}{\nu} + \frac{1}{\nu^2} \sum_{i,j \in I} |e_{i,j}| + \frac{2}{\nu} \sum_{i \in I} |e_i|.$$

Here, let's indicate an elementary application of Corollary 1.

Let $S = \{n \in \mathbb{N} : n \leq x\}$ and I the set of primes $\leq x^{1/4}$. For each $p \in I$, let $\Omega(p)$ be the property "divisibility by p". Then

$$S_p = \#\{n \leq x : n \text{ satisfies } \Omega(p)\} = \left[\frac{x}{p}\right] = \frac{x}{p} + O(1).$$

Thus, $\delta_p = \frac{1}{p}$ and $e_p = O(1)$ in this case.

By Corollary 1,

$$\#\{n \leq x : n \text{ is not divisible by } p \text{ for all } p \leq x^{1/4}\} \leq \frac{x}{\nu} + \frac{x^{1/2}}{\nu^2} + \frac{2}{\nu}x^{1/4}.$$

Now, the left hand side contains at least $\pi(x) - \pi(x^{1/4})$ elements. Also,

$$\nu = \sum_{p \leq x^{1/4}} \frac{1}{p} \gg \log\log x$$

by the following calculation:

Note that

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right)^{-1} \geq \sum_{n \leq x} \frac{1}{n} \geq \log x.$$

Now $\left(1 - \frac{1}{p}\right)^{-1} = 1 + \frac{1}{p} + \frac{1}{p^2} + \cdots$

$$= \left(1 + \frac{1}{p}\right) + \frac{1}{p^2}\left(1 + \frac{1}{p}\right) + \frac{1}{p^4}\left(1 + \frac{1}{p}\right) + \cdots$$

$$= \left(1 + \frac{1}{p}\right)\left(1 + \frac{1}{p^2 - 1}\right).$$

Therefore,

$$\prod_{p \leq x} \left(1 + \frac{1}{p}\right) \gg \log x.$$

Since $e^x \geq 1 + x$, we find

$$\prod_{p \leq x} e^{1/p} = \exp\left(\sum_{p \leq x} \frac{1}{p}\right) \geq \prod_{p \leq x} \left(1 + \frac{1}{p}\right) \gg \log x$$

so that

$$\sum_{p \leq x} \frac{1}{p} \gg \log\log x.$$

Hence,

$$\pi(x) - \pi(x^{1/4}) \ll \frac{x}{\log\log x}.$$

Since $\pi(x^{1/4}) \leq x^{1/4}$, we deduce that

$$\pi(x) = O\left(\frac{x}{\log\log x}\right).$$

This result can also be derived from the sieve of Eratosthenes [9]. The present derivation makes no use of the Möbius function.

Fix $n \in \mathbb{N}$ and let $F(x) = x^n + a_1 x^{n-1} + \cdots + a_n \in \mathbb{Z}[x]$. We denote $\mathrm{Gal}(F)$ to be the Galois group of the splitting field of F over $\mathbb{Q}$. Since $F(x)$ is of degree n, $\mathrm{Gal}(F)$ is a subgroup of S_n. The probabilistic Galois theory problem is to estimate the number of monic polynomials of degree n whose Galois group is a proper subgroup of S_n. We define the *height* of F, $H(F)$, to be

$$H(F) = \max\{1, |a_1|, |a_2|, \cdots, |a_n|\}$$

and

$$E_n(N) = \#\left\{F(x) \in \mathbb{Z}[x] : \begin{array}{l} \deg(F) = n,\ F \text{ is monic,} \\ H(F) \le N \text{ and } \mathrm{Gal}(F) \subsetneq S_n \end{array}\right\}.$$

In 1936, van der Waerden [11] gave an upper bound for $E_n(N)$. He proved that

$$E_n(N) \ll N^{n - \frac{c}{\log\log N}}$$

with $c = \frac{1}{6(n-2)}$.

In 1955, Knobloch [7, 8] improved the result of van der Waerden to

$$E_n(N) \ll N^{n-c}$$

with $c = \frac{1}{18n(n!)^3}$.

In 1973, using the technique of the large sieve in several variables, Gallagher [4] sharpened the result to

$$E_n(N) \ll N^{n-1/2} \log N.$$

In this paper, by applying the Turán sieve, we get

Theorem 2. $E_n(N) \ll N^{n-1/3}(\log N)^2$.

Remark 1. Though our result is weaker than Gallagher's, our method of the Turán sieve is far simpler. We also like to point out that what we have called the "Turán sieve method" was already there in rudimentary form in 1934 and this should be compared with the results obtained by van der Waerden and Knobloch cited above.

Remark 2. In the paper of Cohen's [1], he gave a more general version of the probabilistic Galois theory problem. Similarly, we can consider the probabilistic Galois theory problem in the function field over $\mathbb{F}_q$. It is surprising that the Turán sieve will give us the best possible upper bound.

Theorem 3.

$$\begin{aligned} P_n(N) &= \#\{F(x) = x^n + a_1(t)x^{n-1} + a_2(t)x^{n-2} + \cdots + a_n(t) \\ &\qquad \in \mathbb{F}_q[t,x],\ H(F) \le N \text{ and } Gal(F) \subsetneqq S_n\} \\ &\ll N^{n-1}\log_q N. \end{aligned}$$

Here the height of F, $H(F)$, is defined to be

$$H(F) = \max\{q^{\deg(a_i(t))}, 1 \le i \le n\}.$$

Remark. If $q^r \le N < q^{r+1}$, then $P_n(N) = P_n(q^r)$. Hence, without loss of generality, in the case of function field over $\mathbb{F}_q$, we may assume $N = q^r$, a power of q.

2. Proof of the Turán Sieve

In this section, we prove the Turán sieve.

Proof of Theorem 1.

$$\begin{aligned} \sum_{s\in S}(\pi_s(I) - \nu)^2 &= \sum_{s\in S}(\pi_s(I))^2 - 2\nu\sum_{s\in S}\pi_s(I) + |S|\,\nu^2 \\ &= S_1 - S_2 + |S|\,\nu^2 \text{ (say)} \end{aligned}$$

From the definition of $\pi_s(I)$,

$$S_1 = \sum_{s\in S}(\pi_s(I))^2 = \sum_{s\in S}\Big(\sum_{\substack{i\in I\\ s\in S_i}} 1\Big)^2.$$

Interchanging the order of summation and rearranging terms gives:

$$S_1 = \sum_{i,j\in I}\sum_{s\in S_i\cap S_j} 1 = \sum_{\substack{i,j\in I\\ i\neq j}} |S_i\cap S_j| + \sum_{i\in I}|S_i|.$$

By the definition of δ_i, e_i and $e_{i,j}$, one finds

$$\begin{aligned} S_1 &= |S| \sum_{\substack{i,j\in I\\ i\neq j}} \delta_i\delta_j + \sum_{\substack{i,j\in I\\ i\neq j}} e_{i,j} + |S| \sum_{i\in I}\delta_i + \sum_{i\in I} e_i \\ &= |S| \Big(\sum_{i\in I}\delta_i\Big)^2 - |S| \sum_{i\in I}\delta_i^2 + \sum_{i,j\in I} e_{i,j} + |S| \sum_{i\in I}\delta_i. \end{aligned} \tag{1}$$

Similarly, we can show that

$$S_2 = 2\nu\,|S| \sum_{i\in I}\delta_i + 2\nu \sum_{i\in I} e_i. \tag{2}$$

Combining equation (1) and (2) yields

$$\begin{aligned} S_1 - S_2 + |S|\,\nu^2 &= |S| \left(\sum_{i\in I}\delta_i - \nu\right)^2 + |S| \sum_{i\in I}\delta_i(1-\delta_i) \\ &\quad + \sum_{i,j\in I} e_{i,j} - 2\nu \sum_{i\in I} e_i \\ &= |S| \sum_{i\in I}\delta_i(1-\delta_i) + \sum_{i,j\in I} e_{i,j} - 2\nu \sum_{i\in I} e_i. \end{aligned}$$

□

Proof of Corollary 1. Since $(1-\delta_i) \le 1$, by Theorem 1,

$$\sum_{s\in S} (\pi_s(I) - \nu)^2 \le |S|\,\nu + \sum_{i,j\in I} |e_{i,j}| + 2\nu \sum_{i\in I} |e_i|\,.$$

Notice that

$$\nu^2 \#\{s \in S : \pi_s(I) = 0\} \le \sum_{s\in S} (\pi_s(I) - \nu)^2,$$

Corollary 1 follows. □

3. Probabilistic Galois Theory in the rational number field

We want to get an upper bound for

$$E_n(N) = \#\left\{ F(x) \in \mathbb{Z}[x] : \begin{array}{l} \deg(F) = n,\ F \text{ is monic,} \\ H(F) \le N \text{ and } \mathrm{Gal}(F) \subsetneq S_n \end{array} \right\}.$$

Let $\mathbb{F}_p$ be the finite field of order p and let $\mathbb{F}_p[x]$ be the ring of polynomials in x with coefficients in $\mathbb{F}_p$. It is a Euclidean domain, and hence is a unique factorization domain.

Lemma 1. *Let N_k be the number of monic irreducible polynomials in $\mathbb{F}_p[x]$ with degree k. Then*

$$N_k = \frac{1}{k}\sum_{d|k}\mu(d)p^{\frac{k}{d}},$$

where $\mu(d)$ is the Möbius function.

Proof. Define a set

$$A = \{f(x) \in \mathbb{F}_p[x] : f(x) \text{ is monic}\}.$$

Applying the method used to prove the Euler product for the Riemann zeta-function, we have

$$\begin{aligned}\sum_{f\in A} T^{\deg(f)} &= \prod_{\substack{g\in A \\ g \text{ is irreducible}}} (1-T^{\deg(g)})^{-1} \\ &= \prod_{k=1}^{\infty}(1-T^k)^{-N_k}.\end{aligned}$$

On the other hand, for a fixed degree k, the number of polynomials in A with degree k is p^k. So,

$$\sum_{f\in A} T^{\deg(f)} = \sum_{k=o}^{\infty} p^k T^k = \frac{1}{1-pT}.$$

Hence, we have

$$\frac{1}{1-pT} = \prod_{k=1}^{\infty}(1-T^k)^{-N_k}.$$

Taking the logarithm on both sides and applying the identity

$$-\log(1-x) = \sum_{i=1}^{\infty}\frac{x^i}{i},$$

we get

$$\sum_{i=1}^{\infty}\frac{p^iT^i}{i} = \sum_{i=1}^{\infty}T^i\left(\sum_{jk=i}\frac{N_k}{j}\right).$$

Comparing the coefficient of T^i in the equation, we get

$$p^i = \sum_{k|i} kN_k.$$

Applying the Möbius inversion formula, Lemma 1 follows. □

Given a monic polynomial $F(x) \in \mathbb{Z}[x]$, if the factorization of $F(x) (\mathrm{mod}\ p)$ has r_1 linear factors, r_2 quadratic factors, r_3 cubic factors and etc., we say $F(x)$ has *mod p type* $r = (r_1, r_2, r_3, \cdots)$. For a permutation $\sigma \in S_n$, if σ has r_1 cycles of length 1, r_2 cycles of length 2 and etc., we say σ has *cycle type* $r = (r_1, r_2, \cdots)$. The above two types are closely related: for a fixed type r, if there exists a prime p such that $F(x) (\mathrm{mod}\ p)$ is of type r, then there exists a permutation $\sigma \in \mathrm{Gal}(F) \subseteq S_n$ whose cycle type is r. The converse is also true and this is essentially a consequence of the Chebotarev density theorem. (See [3] and [6])

Notice that: if $\mathrm{Gal}(F) \subsetneq S_n$, there exist a conjugacy class C of S_n such that $\mathrm{Gal}(F) \cap C = \emptyset$. Also, each cycle type represent a conjugacy classes of S_n. Combine the above facts: if F is an element of $E_n(N)$, there exists a cycle type r which is never represented by its Galois group. Hence, to obtain an upper bound for $E_n(N)$, it suffices to get an upper bound for

$$E_{r,n}(N) = \#\left\{F(x) \in E_n(N) : \begin{array}{l} F \text{ does } \textit{not} \text{ have } \textit{mod p type } r \\ \text{for all primes p} \geq 3 \end{array}\right\}.$$

Then, summing $E_{r,n}(N)$ over all possible types r will give us an upper bound for $E_n(N)$. In fact, there is a uniform upper bound for $E_{r,n}$ for all types r. Since the number of types r is bounded, to prove Theorem 2, it suffices to prove the following theorem which is also interesting in its own right.

Theorem 4. $E_{r,n}(N) \ll N^{n-1/3}(\log N)^2$.

Before proving Theorem 4, we need the following lemma.

Lemma 2. *Fix a type* $r = (r_1, r_2, \cdots)$, $\sum_k r_k k = n$. *Let* p *be a prime* ≥ 3. *We denote by* $\omega(p)$, *the number of monic polynomials in* $\mathbb{F}_p[x]$ *with*

degree n and of type r. Then

$$\omega(p) \geq \frac{\delta(r)}{3^n} p^n,$$

here $\delta(r) = \prod_k \frac{1}{r_k! k^{r_k}}$.

Proof. From Lemma 1,

$$\begin{aligned} kN_k &= \sum_{d|k} \mu(d) p^{\frac{k}{d}} \\ &\geq p^k - (p^{k-1} + p^{k-2} + \cdots + p + 1) \\ &= p^k - \frac{p^k - 1}{p-1} \geq p^k - 2p^{k-1} \\ &\geq \left(1 - \frac{2}{p}\right) p^k \geq \frac{1}{3} p^k \text{ (since } p \geq 3). \end{aligned}$$

$$\begin{aligned} \text{Hence, } \omega(p) &= \prod_k \binom{N_k + r_k - 1}{r_k} \\ &= \prod_k \frac{1}{r_k!} (N_k + r_k - 1)(N_k + r_k - 2) \cdots N_k \\ &\geq \prod_k \frac{1}{r_k!} N_k^{r_k} \geq \prod_k \frac{1}{r_k!} \left(\frac{p^k}{3k}\right)^{r_k} \\ &\geq \frac{\delta(r)}{3^n} p^n \text{ (since } r_k \leq n). \end{aligned}$$

□

Proof of Theorem 4. The proof is an application of the Turán sieve. Let $S = \{s = (a_1, a_2, \cdots, a_n) \in \mathbb{Z}^n, |a_i| \leq N\}$. For each $s \in S$, we define $F(s) = x^n + a_1 x^{n-1} + \cdots + a_n$. Let $I = \{p : \text{primes and } 3 \leq p \leq z\}$. For each $p \in I$, we define $\Omega(p)$ to be the condition that $F(s) \pmod p$ is of type r, hence,

$$S_p = \{s \in S : F(s) \pmod p \text{ is of type } r\}$$

and $\pi_s(I) = \#\{p \in I : F(s) \pmod p \text{ is of type } r\}$.

Apply Lemma 2 and the following observation: fix a monic polynomial $f(x)$ in $\mathbb{F}_p[x]$. The number of monic polynomials $F(x) \in \mathbb{Z}[x]$ with

$H(F) \le N$ and $F(x) \equiv f(x) \pmod p$ is

$$\left(\frac{2N+1}{p} + O(1)\right)^n = \frac{(2N+1)^n}{p^n} + O\left(\left(\frac{N}{p}\right)^{n-1}\right).$$

We have $|S_p| = (2N+1)^n \frac{\omega(p)}{p^n} + O\left(\frac{\omega(p)N^{n-1}}{p^{n-1}}\right)$.

Hence, $\delta_p = \frac{\omega(p)}{p^n}$ and $e_p = O(\frac{\omega(p)N^{n-1}}{p^{n-1}})$.

Also, $e_{p,q} = O\left(\frac{\omega(p)\omega(q)N^{n-1}}{p^{n-1}q^{n-1}}\right)$ if $p \ne q$.

By Lemma 2, $\nu = \sum_{p \le z} \frac{\omega(p)}{p^n} \gg \frac{z}{\log z}$.

Apply Corollary 1,

$$\begin{aligned} E_{r,n}(N) &\le \#\{s \in S, \pi_s(I) = 0\} \\ &\ll \frac{N^n \log z}{z} + \frac{N^n (\log z)^2}{z^2} \sum_{p,q \in I} \frac{pq}{N} \\ &\quad + \frac{N^n \log z}{z} \sum_{p \in I} \frac{p}{N} \text{ (note: } \omega(p) \le p^n) \\ &\ll \frac{N^n (\log z)^2}{z} \left(1 + \frac{z^3}{N}\right). \end{aligned}$$

Choosing $z = N^{1/3}$, we get $E_{r,n}(N) \ll N^{n-1/3}(\log N)^2$. □

4. Probabilistic Galois Theory in the function field over $\mathbb{F}_q$

4.1. Reducible polynomials in $\mathbb{F}_q[t,x]$

Before we start the function field version of the probabilistic Galois theory problem, we estimate first the number of reducible polynomials in $\mathbb{F}_q[t,x]$. More precisely, we find an upper bound for

$$\begin{aligned} R_n(N) = \#\{&F(x) = x^n + a_1(t)x^{n-1} + a_2(t)x^{n-2} + \cdots + a_n(t) \\ &\in \mathbb{F}_q[t,x],\ H(F) \le N \text{ and } F \text{ is reducible}\}. \end{aligned}$$

Fix a monic irreducible polynomial $v(t) \in \mathbb{F}_q[t]$ with degree n_v. We define F_v to be $\mathbb{F}_q[t]/(v(t))$, a finite field with q^{n_v} elements. Let $F_v[x]$ be the ring of polynomials in x with coefficients in F_v. It is a Euclidean domain, and hence is a unique factorization domain.

Give $F(x) = x^n + a_1(t)x^{n-1} + \cdots + a_n(t) \in \mathbb{F}_q[t,x]$. We denote $F(x) (\mathrm{mod}\ v(t)) = x^n + u_1(t)x^{n-1} + \cdots + u_n(t)$, where $a_i(t) \equiv u_i(t) (\mathrm{mod}\ v(t))$ and $\deg(u_i(t)) < \deg(v(t))$. Since $F_v[x]$ is a unique factorization domain, we can factor $F(x) (\mathrm{mod}\ v(t))$ as a product of monic irreducible polynomials.

Lemma 3. *Let N_k be the number of monic irreducible polynomials in $F_v[x]$ with degree k. Then*

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

Proof. The proof is an analogue of Lemma 1, except replacing p by q^{n_v}. □

Theorem 5. $R_n(N) \ll N^{n-1} \log_q N$.

Proof. The proof is an application of the Turán sieve.
Let $S = \{s = (a_1(t), a_2(t), \cdots, a_n(t)) \in (\mathbb{F}_q[t])^n,\ q^{\deg(a_i(t))} \leq N\}$. For each $s \in S$, we define $F(s) = x^n + a_1(t)x^{n-1} + \cdots + a_n(t) \in \mathbb{F}_q[t,x]$. Let $I = \{v(t) \in \mathbb{F}_q[t] : v(t) \text{ is monic, irreducible and } 2 \leq \deg(v) = n_v \leq z\}$. For each $v(t) \in I$, we define $\Omega(v)$ to be the condition that $F(s)$ (mod $v(t)$) is irreducible, hence,

$$S_v = \{s \in S : F(s) \ (\mathrm{mod}\ v(t)) \text{ is irreducible}\}$$
and $\pi_s(I) = \#\{v(t) \in I : F(s) \ (\mathrm{mod}\ v(t)) \text{ is irreducible}\}$.

We observe that for a fixed monic polynomial $f(x)$ of degree n in $F_v[x]$, there are $q^{(\log_q N - n_v + 1)n}$ many $s \in S$ such that $F(s) \equiv f(x) \ (\mathrm{mod}\ v(t))$. Hence, we have

$$|S_v| = q^{(\log_q N - n_v + 1)n} N_n \text{ and } |S_v| = \frac{N_n}{q^{n_v n}} |S|.$$

We set $\delta_v = \frac{N_n}{q^{n\nu n}}$ and e_v is 0. Also, $e_{v,w} = 0$ if $v \neq w$. Applying the same method in the proof of Lemma 2,

$$N_n \geq \frac{1}{3n}(q^{N_n})^n.$$

Hence,

$$\nu \geq \sum_{v \in I} \frac{1}{3n} \gg \frac{q^z}{z}.$$

Applying Corollary 1, we get

$$R_n(N) \leq \#\{s \in S, \pi_s(\mathcal{P}) = 0\} \ll \frac{q^{(\log_q N)n} z}{q^z}.$$

Choosing $z = \log_q N$, the theorem follows. □

4.2. Probabilistic Galois theory in $\mathbb{F}_q[t, x]$

We want to get an upper bound for

$$P_n(N) = \#\{F(x) = x^n + a_1(t)x^{n-1} + a_2(t)x^{n-2} + \cdots + a_n(t) \in \mathbb{F}_q[t, x],\ H(F) \leq N \text{ and } \mathrm{Gal}(F) \subsetneq S_n.\}$$

Remark. We notice the fact that not all polynomials in $\mathbb{F}_q[t, x]$ are separable. However, in our case, the number of such polynomials is negligible. Indeed, consider $F(x) \in \mathbb{F}_q[t, x]$ of degree n. By Theorem 5, without loss of generality, we can assume that $F(x)$ is irreducible. Moreover, for an irreducible polynomial $F(x) \in \mathbb{F}_q[t, x]$, $F(x)$ is inseparable if and only if $F(x) = G(x^p)$ for some polynomial G. Hence, if $p \nmid n$, $F(x)$ is always separable. Otherwise, the number of inseparable polynomials is at most $O(N^{\frac{n}{p}})$, which is negligible in our cases. Hence, without loss of generality, we can assume that all $F(x)$ in the correspondence set of $P_n(N)$ are separable, hence $\mathrm{Gal}(F)$ always exist.

For a monic irreducible polynomial $v(t) \in \mathbb{F}_q[t]$, we define *mod* $v(t)$ *type* in the obvious way. As before, to obtain an upper bound for $P_n(N)$, it suffices to get an upper bound for

$$P_{r,n}(N) = \#\{F(x) \in P_n(N) : F \text{ does } \textit{not} \text{ have } \textit{mod } v(t) \textit{ type } r \text{ for all monic irreducible polynomials } v(t) \text{ which is of degree} \geq 2\}.$$

We prove

Theorem 6. $P_{r,n}(N) \ll N^{n-1} \log_q N$.

The analogue of Lemma 2 in the function field is the following:

Lemma 4. *Fix a type* $r = (r_1, r_2, \cdots)$, $\sum_k r_k k = n$. *Let* $\nu(t)$ *be a monic, irreducible polynomial in* $\mathbb{F}_q[t]$ *which is of degree* ≥ 2. *We denote by* $\omega(v)$, *the number of monic polynomials in* $F_v[x]$ *with degree* n *and of type* r. *Then,*

$$\omega(v) \geq \frac{\delta(r)}{3^n}(q^{n_v})^n,$$

here $\delta(r) = \prod_k \frac{1}{r_k! k^{r_k}}$.

Proof of Theorem 6. The proof is again an application of the Turán sieve.

Let $S = \{s = (a_1(t), a_2(t), \cdots, a_n(t)) \in (\mathbb{F}_q[t])^n,\ q^{\deg(a_i(t))} \leq N\}$. For each $s \in S$, we define $F(s) = x^n + a_1(t)x^{n-1} + \cdots + a_n(t) \in \mathbb{F}_q[t, x]$. Let $I = \{v(t) \in \mathbb{F}_q[t] : v(t)$ is monic, irreducible and $2 \leq \deg(v) = n_v \leq z\}$. For each $v(t) \in I$, we define $\Omega(v)$ to be the condition that $F(s) \pmod{v(t)}$ is of type r. By Lemma 4 and Corollary 1,

$$|S_v| = \frac{\omega(v)}{q^{n_v n}} |S|.$$

The rest of the proof is just the same as the proof of Theorem 5, except replacing N_n by $\omega(v)$. □

5. Concluding Remarks

Given $F(x) = x^n + a_1 x^{n-1} + \cdots + a_n \in \mathbb{Z}[x]$ with $H(F) \leq N$, if we assume $a_n = 0$, Gal(F) is a proper subgroup of S_n. Since the order of such elements is N^{n-1}, we have

$$E_n(N) \gg N^{n-1}.$$

By the same argument, we can show that $P_n(N) \gg N^{n-1}$. In 1936, van der Waerden [11] conjectured that $E_n(N) \ll N^{n-1}$.

By developing the large sieve in function fields over finite fields, it is possible to prove the following:

Theorem 7. *The number of the set*

$$\left\{F(x) = x^n + a_1(t)x^{n-1} + \cdots + a_n(t) \in \mathbb{F}_q[t,x] : \begin{array}{l}\deg(a_i(t)) \le N \\ \text{and } Gal(F) \subsetneq S_n\end{array}\right\}$$

is $\ll N^{n-1}$.

The details are involved and we relegate this to a further paper.

References

[1] S. D. Cohen, The distribution of the Galois groups of integral polynomials, *Illinois Journal of Mathematics* **23** (1979), 135–152.

[2] S.D. Cohen, The distribution of Galois groups and Hilbert's irreducibility theorem, *Proc. London Math. Soc.* **43** (1981), 227–250.

[3] G. Frobenius, *Über Beziehungen zwischen den Primidealen eines algebraischen Körpers und den Substitutionen seiner Gruppe*, S. B. Akad. Wiss. Wien 134 (1925), 69–80.

[4] P. X. Gallagher, The large sieve and probabilistic Galois theory, *Analytic Number Theory* (Proc. Sympos. Pure Math., Vol. 24), Amer. Math. Soc., Providence, R.I. (1973), 91–101.

[5] C. Hooley, *Applications of Sieve Methods to the Theory of Numbers*, Cambridge University press (1976).

[6] N. Jacobson, *Basic algebra I*, Freeman (1996), 302–304.

[7] H. W. Knobloch, *Zum Hilbertschen Irreduzibilitätssatz*, Abh. Math. Sem. Univ. Hamberg. 19 (1955), 176–190.

[8] H. W. Knobloch, *Die Seltenheit der reduziblen Polynome*, Jber. Deutsch. Math. verein. 59 (1956), 12–19.

[9] M. R. Murty and N. Saradha, On the sieve of Eratosthenes, Canadian Journal of Mathematics **39** (1987), 1107–1122.

[10] J.-P. Serre, *Lectures on Mordell-Weil Theorem*, Aspects of Mathematics, Friedr. Viewig and Sohn (1989).

[11] B. L. van der Waerden, Die Seltenheit der reduziblen Gleichungen und der Gleichungen mit Affekt, *Monatsh. Math* **43** (1936), 133–147.