ACTA ARITHMETICA
114.4 (2004)

The Erdős Theorem and the Halberstam Theorem in function fields

by

Yu-Ru Liu (Waterloo)

1. Introduction. For $n \in \mathbb{N}$, define $\omega(n)$ to be the number of distinct prime divisors of n. The Turán Theorem [9] concerns the second moment of $\omega(n)$ and it implies a result of Hardy and Ramanujan [4] that the normal order of $\omega(n)$ is $\log\log n$. Further development of probabilistic ideas led Erdős and Kac [2] to prove a remarkable refinement of the Hardy–Ramanujan Theorem, namely, the existence of a normal distribution for $\omega(n)$. More precisely, they proved that for $x, \gamma \in \mathbb{R}$,

$$\lim_{x\to\infty} \frac{1}{\#\{n \le x\}} \#\left\{n \le x : \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \le \gamma\right\} = G(\gamma)$$
$$:= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\gamma} e^{-t^2/2}\, dt.$$

Instead of the sequence of all natural numbers, we can consider only the set of primes. Since $\omega(p) = 1$ for each prime p, the normal order of $\omega(p)$ is not $\log\log p$. However, Erdős [1] proved in 1935 that

$$\sum_{p \le x} (\omega(p-1) - \log\log x)^2 \ll \pi(x) \log\log x,$$

where $\pi(x) = \#\{p \text{ prime} : p \le x\}$. This implies that the normal order of $\omega(p-1)$ is $\log\log p$. In 1955, Halberstam [3] improved Erdős's result and proved that

$$\lim_{x\to\infty} \frac{1}{\pi(x)} \#\left\{p \le x : \frac{\omega(p-1) - \log\log p}{\sqrt{\log\log p}} \le \gamma\right\} = G(\gamma).$$

This result can be viewed as a "prime analogue" of the Erdős–Kac Theorem.

Let $\mathbb{F}_q[t]$ be the polynomial ring in one variable over a finite field $\mathbb{F}_q$. Let P be the set of monic irreducible polynomials in $\mathbb{F}_q[t]$. For $m \in \mathbb{F}_q[t]$, let

2000 *Mathematics Subject Classification*: 11N60, 11R09.

Research partially supported by an NSERC discovery grant.

$\deg m$ be the degree of the polynomial m. Also, let $\omega(m)$ denote the number of distinct monic irreducible polynomials dividing m, i.e.,

$$\omega(m) = \sum_{\substack{l \in P \\ l \mid m}} 1.$$

We can formulate analogues of the Erdős Theorem and the Halberstam Theorem in $\mathbb{F}_q[t]$.

THEOREM 1. *Let P be the set of monic irreducible polynomials in $\mathbb{F}_q[t]$. Fix a nonzero polynomial a in $\mathbb{F}_q[t]$. For $n \in \mathbb{N}$, we have*

$$\sum_{\substack{p \in P \\ \deg p \le n}} (\omega(p-a) - \log n)^2 \ll \pi(n) \log n,$$

where $\pi(n) = \#\{p \in P : \deg p \le n\}$.

As a direct consequence of Theorem 1, we have

COROLLARY 1. *Let $\{g_n\}$ be a sequence of real numbers with $g_n \to \infty$ as $n \to \infty$. Then*

$$\#\left\{p \in P : \deg p \le n,\ \left|\frac{\omega(p-a) - \log(\deg p)}{\sqrt{\log(\deg p)}}\right| > g_n\right\} = o(\pi(n)).$$

In particular, given $\varepsilon > 0$, we have

$$\#\{p \in P : \deg p \le n,\ |\omega(p-a) - \log(\deg p)| > \varepsilon \log(\deg p)\} = o(\pi(n)).$$

Thus the normal order of $\omega(p-a)$ is $\log(\deg p)$.

As we see from previous examples, Corollary 1 implies a possibility that the quantity

$$\frac{\omega(p-a) - \log(\deg p)}{\sqrt{\log(\deg p)}}$$

distributes normally. This is indeed the case.

THEOREM 2. *Let P be the set of monic irreducible polynomials in $\mathbb{F}_q[t]$. Fix a nonzero polynomial a in $\mathbb{F}_q[t]$. For $n \in \mathbb{N}$, $\gamma \in \mathbb{R}$, we have*

$$\lim_{n \to \infty} \frac{1}{\pi(n)} \#\left\{p \in P : \deg p \le n,\ \frac{\omega(p-a) - \log(\deg p)}{\sqrt{\log(\deg p)}} \le \gamma\right\} = G(\gamma).$$

2. Proof of Theorem 1. We begin with two facts that are essential for the proof of Theorem 1. Let P be the set of monic irreducible polynomials in $\mathbb{F}_q[t]$. The following facts concern elements of P; their proofs can be found in [8].

Fact 1 ([8, p. 14]). *For $d \in \mathbb{N}$, we have*

$$\#\{p \in P : \deg p = d\} = \frac{q^d}{d} + O(q^{d/2}).$$

The next fact concerns arithmetic progressions of irreducible polynomials in function fields. It is a theorem of Kornblum [5].

Fact 2 ([8, p. 40]). *Let a, m be polynomials in $\mathbb{F}_q[t]$ that are relatively prime. For any $\varepsilon > 0$ and $d \in \mathbb{N}$, we have*

$$\#\{p \in P : \deg p = d,\ p \equiv a \pmod{m}\} = \frac{1}{\phi(m)} \cdot \frac{q^d}{d} + O(q^{d(1+\varepsilon)/2}),$$

where $\phi(m)$ is the cardinality of $(\mathbb{F}_q[t]/m\mathbb{F}_q[t])^*$.

Before proving Theorem 1, we consider its analogous version for monic irreducible polynomials of a fixed degree.

Lemma 1. *Let a be a fixed nonzero polynomial and p a monic irreducible polynomial in $\mathbb{F}_q[t]$. For $d \in \mathbb{N}$, we have*

$$\sum_{\deg p = d} (\omega(p-a) - \log d)^2 \ll \frac{q^d}{d} \log d.$$

Proof. Let δ be a constant with $0 < \delta < 1$ which will be chosen later. Let l be a monic irreducible polynomial. Notice that

$$\omega(p-a) = \sum_{\substack{l \mid (p-a) \\ \deg l \le \delta d}} 1 + \sum_{\substack{l \mid (p-a) \\ \delta d < \deg l \le d}} 1 = \omega_\delta(p-a) + O(1/\delta),$$

where

$$\omega_\delta(p-a) = \sum_{\substack{l \mid (p-a) \\ \deg l \le \delta d}} 1.$$

By Facts 1 and 2, we have

$$\begin{aligned}
\sum_{\deg p = d} \omega(p-a) &= \sum_{\deg p = d} (\omega_\delta(p-a) + O(1/\delta)) \\
&= \sum_{\deg l \le \delta d} \sum_{\substack{\deg p = d \\ p \equiv a \,(\mathrm{mod}\, l)}} 1 + O(q^d/d) \\
&= \sum_{\deg l \le \delta d} \left(\frac{1}{q^{\deg l} - 1} \cdot \frac{q^d}{d} + O(q^{d(1+\varepsilon)/2}) \right) + O(q^d/d).
\end{aligned}$$

By choosing $\delta < 1/2$, Fact 1 implies that

$$\begin{aligned}\sum_{\deg p=d} \omega(p-a) &= \frac{q^d}{d} \sum_{\deg l \le \delta d} \frac{1}{q^{\deg l}} + O(q^d/d) \\ &= \frac{q^d}{d} \sum_{k \le \delta d} \frac{1}{q^k} \left(\frac{q^k}{k} + O(q^{k/2}) \right) + O(q^d/d) \\ &= \frac{q^d}{d} \log d + O(q^d/d).\end{aligned}$$

Now, consider $\sum_{\deg p=d} \omega^2(p-a)$. Write

$$\begin{aligned}\sum_{\deg p=d} \omega^2(p-a) &= \sum_{\deg p=d} (\omega_\delta(p-a) + O(1/\delta))^2 \\ &= \sum_{\deg p=d} \omega_\delta^2(p-a) + O(q^d \log d/d).\end{aligned}$$

We have

$$\begin{aligned}\sum_{\deg p=d} \omega_\delta^2(p-a) &= \sum_{\substack{\deg l_1, \deg l_2 \le \delta d \\ l_1 \ne l_2}} \sum_{\substack{\deg p=d \\ p \equiv a \,(\mathrm{mod}\, l_1 l_2)}} 1 + \sum_{\deg l \le \delta d} \sum_{\substack{\deg p=d \\ p \equiv a \,(\mathrm{mod}\, l)}} 1 \\ &= \sum_{\deg l_1, \deg l_2 \le \delta d} \left(\frac{1}{\phi(l_1 l_2)} \cdot \frac{q^d}{d} + O(q^{d(1+\varepsilon)/2}) \right) \\ &\quad + O(q^d \log d/d).\end{aligned}$$

By choosing $0 < \delta < 1/4$, we have

$$\begin{aligned}\sum_{\deg p=d} \omega^2(p-a) &= \frac{q^d}{d} \sum_{\deg l_1, \deg l_2 \le \delta d} \frac{1}{q^{\deg l_1} \cdot q^{\deg l_2}} + O(q^d \log d/d) \\ &= \frac{q^d}{d} (\log d)^2 + O(q^d \log d/d).\end{aligned}$$

Combining all the above results and choosing $\delta = 1/5$, we obtain

$$\begin{aligned}&\sum_{\deg p=d} (\omega(p-a) - \log d)^2 \\ &\quad = \sum_{\deg p=d} \omega^2(p-a) - 2\log d \sum_{\deg p=d} \omega(p-a) + (\log d)^2 \sum_{\deg p=d} 1 \\ &\quad \ll \frac{q^d \log d}{d}.\end{aligned}$$

Thus Lemma 1 follows.

Now, Theorem 1 follows directly from Lemma 1:

Proof of Theorem 1. By Lemma 1, we have

$$\begin{aligned}
&\sum_{\deg p \le n} (\omega(p-a) - \log n)^2 \\
&= \sum_{d \le n} \sum_{\deg p = d} (\omega(p-a) - \log d + \log d - \log n)^2 \\
&\ll \sum_{d \le n} \sum_{\deg p = d} (\omega(p-a) - \log d)^2 + \sum_{d \le n} \sum_{\deg p = d} (\log d - \log n)^2 \\
&\ll \sum_{d \le n} \frac{q^d}{d} \log d + \sum_{1 \le d \le n/2} \sum_{\deg p = d} (\log n)^2 + \sum_{n/2 < d \le n} \sum_{\deg p = d} (\log d - \log n)^2.
\end{aligned}$$

The third term of the last inequality is

$$\sum_{n/2 < d \le n} \sum_{\deg p = d} (\log d - \log n)^2 \ll (\log 2)^2 \sum_{n/2 < d \le n} \sum_{\deg p = d} 1 \ll \pi(n).$$

The second term can be estimated by

$$\sum_{1 \le d \le n/2} \sum_{\deg p = d} (\log n)^2 = (\log n)^2 \pi(n/2) \ll \pi(n).$$

The first term is the main term. It is bounded by

$$\sum_{d \le n} \frac{q^d}{d} \log d \ll \log n \sum_{d \le n} \#\{p \in P : \deg p = d\} \ll \pi(n) \log n.$$

Combining all the above estimates, we obtain

$$\sum_{\deg p \le n} (\omega(p-a) - \log n)^2 \ll \pi(n) \log n.$$

Hence, Theorem 1 follows. We have thus obtained an analogue of the Erdős Theorem in $\mathbb{F}_q[t]$.

3. Proof of Theorem 2. In this section, we shall prove that the quantity

$$\frac{\omega(p-a) - \log(\deg p)}{\sqrt{\log(\deg p)}}$$

distributes normally. This follows from Theorem 1 of [6]. Instead of stating that theorem in its general form, we state below its consequence in $\mathbb{F}_q[t]$. Let P be the set of monic irreducible polynomials in $\mathbb{F}_q[t]$. For $m \in \mathbb{F}_q[t]$, define $N(m) := q^{\deg m}$. Take $X = \{q^z : z \in \mathbb{Z}\}$. Let S be an infinite subset of $\mathbb{F}_q[t]$. For $x \in X$, define

$$S(x) = \{m \in S : N(m) \le x\}.$$

We assume that S satisfies the following condition:

(C) $$|S(x^{1/2})| = o(|S(x)|) \quad \text{for all } x \in X.$$

Let f be a map from S to M. For each $l \in P$, we write

$$\frac{1}{|S(x)|}\,\#\{m \in S(x) : l \mid f(m)\} = \lambda_l(x) + e_l(x),$$

where $\lambda_l = \lambda_l(x)$ can be thought of as the main term (and is usually chosen to be independent of x) and $e_l = e_l(x)$ is an error term. For any sequence of distinct elements $l_1, \dots, l_u \in P$, we write

$$\frac{1}{|S(x)|}\,\#\{m \in S(x) : l_i \mid f(m) \text{ for all } i = 1, \dots, u\} = \lambda_{l_1} \cdots \lambda_{l_u} + e_{l_1 \dots l_u}(x).$$

We will use $e_{l_1 \dots l_u}$ to abbreviate $e_{l_1 \dots l_u}(x)$ below.

Suppose that for all $x \in X$, there exists a constant β with $0 < \beta \le 1$ and $y = y(x) < x^{\beta}$ such that the following conditions hold:

(1) $\#\{l \in P : N(l) > x^{\beta},\ l \mid f(m)\} = O(1)$ for each $m \in S(x)$.

(2) $$\sum_{y < N(l) \le x^{\beta}} \lambda_l = o((\log\log x)^{1/2}).$$

(3) $$\sum_{y < N(l) \le x^{\beta}} |e_l| = o((\log\log x)^{1/2}).$$

(4) $$\sum_{N(l) \le y} \lambda_l = \log\log x + o((\log\log x)^{1/2}).$$

(5) $$\sum_{N(l) \le y} \lambda_l^2 = o((\log\log x)^{1/2}).$$

(6) For $r \in \mathbb{N}$ and $u = 1, \dots, r$, we have

$$\sum\nolimits'' |e_{l_1 \dots l_u}| = o((\log\log x)^{-r/2}),$$

where $\sum''$ extends over all u-tuples $(l_1, \dots, l_u)$ with $N(l_i) \le y$ and l_i are all distinct.

It was proved in [6] that there is a generalization of the Erdős–Kac Theorem in $\mathbb{F}_q[t]$.

THEOREM 3 (Theorem 1 in [6]). *Let P and X be as before. Let S be a subset of $\mathbb{F}_q[t]$ satisfying condition* (C). *Let $f : S \to \mathbb{F}_q[t]$. Suppose there exists a constant β with $0 < \beta \le 1$ and $y = y(x) < x^{\beta}$ such that conditions* (1) *to* (6) *hold. Then for $\gamma \in \mathbb{R}$, we have*

$$\lim_{x \to \infty} \frac{1}{|S(x)|}\,\#\left\{m \in S(x) : \frac{\omega(f(m)) - \log\log N(m)}{\sqrt{\log\log N(m)}} \le \gamma\right\} = G(\gamma).$$

Now, we are ready to prove Theorem 2. Let $S = P$ and $f : p \mapsto p - a$. By Fact 1, condition (C) follows. Choose $y = x^{1/\log\log x}$ and let β be any constant such that $0 < \beta < 1/2$. Since for $N(p) \le x = q^n$ with x large (say $> N(a)$), we have

$$\#\{l \in P : N(l) > x^\beta,\ l \mid (p-a)\} \le 1/\beta,$$

condition (1) is satisfied. For a monic irreducible polynomial l, Fact 2 implies that

$$\#\{p \in P : \deg p \le n,\ p \equiv a \pmod{l}\} = \frac{1}{\phi(l)}\,\pi(n) + O(\pi(n)^{1/2+\varepsilon}).$$

Take $\lambda_l = 1/\phi(l)$. Lemmas 1 and 2 in [7] state that

$$\sum_{N(l)\le x} \frac{1}{N(l)} = \log\log x + O(1), \qquad \sum_{N(l)\le x} \frac{1}{N(l)^2} \ll 1.$$

Thus conditions (2), (4), and (5) follow. Also, we have

$$\sum_{y<N(l)\le x^\beta} |e_l| \ll \pi(n)^{-1/2+\varepsilon} \cdot \pi(n)^\beta \ll 1,$$

since $\beta < 1/2$. Thus, condition (3) follows. For distinct primes $l_1, \dots, l_u$ with $N(l_i) \le y$, by Fact 2, we have

$$|e_{l_1 \dots l_u}| \ll \pi(n)^{-1/2+\varepsilon}.$$

Since $y = o(x^\varepsilon)$, condition (6) is satisfied. Combining all the above results, Theorem 3 implies that

$$\lim_{n\to\infty} \frac{1}{\pi(n)}\,\#\left\{p \in P : \deg p \le n,\ \frac{\omega(p-a) - \log(\deg p)}{\sqrt{\log(\deg p)}} \le \gamma\right\} = G(\gamma).$$

We have thus obtained an analogue of the Halberstam Theorem in $\mathbb{F}_q[t]$.

Acknowledgements. I would like to thank Prof. B. Mazur and Prof. R. Murty for their comments about this work.

References

[1] P. Erdős, *On the normal order of prime factors of* $(p-1)$ *and some related problems concerning Euler's* ϕ*-function*, Quart. J. Math. Oxford 6 (1935), 205–213.

[2] P. Erdős and M. Kac, *The Gaussian law of errors in the theory of additive number theoretic functions*, Amer. J. Math. 62 (1940), 738–742.

[3] H. Halberstam, *On the distribution of additive number theoretic functions*, *I*, *II*, *III*, J. London Math. Soc. 30 (1955), 43–53; 31 (1956), 1–14, 15–27.

[4] G. H. Hardy and S. Ramanujan, *The normal number of prime factors of a number n*, Quart. J. Pure Appl. Math. 48 (1917), 76–97.

[5] H. Kornblum, *Über die Primfunktionen in einer arithmetischen Progression*, Math. Z. 5 (1919), 100–111.

[6] Y.-R. Liu, *A generalized Erdős–Kac Theorem and its prime analogues*, submitted.

[7] —, *A generalization of the Turán Theorem and its applications*, Canad. Math. Bull., to appear.

[8] M. Rosen, *Number Theory in Function Fields*, Springer, 2002.
[9] P. Turán, *On a theorem of Hardy and Ramanujan*, J. London Math. Soc. 9 (1934), 274–276.

Department of Pure Mathematics
University of Waterloo
Waterloo, ON, Canada N2L 3G1
E-mail: yrliu@math.uwaterloo.ca

Received on 26.2.2003 (4474)