DISTRIBUTION OF ZETA ZEROES FOR TRIGONAL CURVES OVER A FINITE FIELD

FRANK THORNE AND MAOSHENG XIONG

Abstract. Building on recent work of Zhao [28], we prove two results on random trigonal curves over a finite field $\mathbb{F}_q$. The first, complementing a result of Wood [24], shows that the distribution of the number of $\mathbb{F}_q$-rational points on a random trigonal curve over $\mathbb{F}_q$ converges to a Gaussian distribution as the genus of the curve and the size of $\mathbb{F}_q$ both go to infinity. The second result is that as the genus goes to infinity for fixed $\mathbb{F}_q$, the number of zeroes of the associated zeta function lying in a prescribed arc becomes uniformly distributed, and that the variance again follows a Gaussian distribution.

Zhao's work is still in preparation; we will submit this once his paper is publicly available!

1. Introduction

In an interesting paper, Wood [24] studied the distribution of the number of points on trigonal curves over a finite field $\mathbb{F}_q$. More precisely, let

$$T_g := \{\pi : C \to \mathbb{P}^1 | C \text{ is a smooth, geometrically integral, genus } g \text{ curve with } \pi \text{ degree } 3\}.$$

She proved the following ([24, Theorem 1.1]): Let $\mathbb{F}_q$ have characteristic ≥ 5. Then

$$\lim_{g\to\infty} \frac{\#\{(C,\pi) \in T_g(\mathbb{F}_q) | \#C(\mathbb{F}_q) = m\}}{\#T_g(\mathbb{F}_q)} = \text{Prob}(X_1 + \cdots + X_{q+1} = m)\,,$$

where the X_i's are independent identically distributed random variables and

$$X_i = \begin{cases} 0 & \textit{with probability } \frac{2q^2}{6q^2+6q+6}\,, \\ 1 & \textit{with probability } \frac{3q^2+6}{6q^2+6q+6}\,, \\ 2 & \textit{with probability } \frac{6q}{6q^2+6q+6}\,, \\ 3 & \textit{with probability } \frac{q^2}{6q^2+6q+6}\,. \end{cases} \tag{1}$$

In particular, this shows that as $g \to \infty$, the average number of points on trigonal curves over $\mathbb{F}_q$ is $q+2-\frac{1}{q^2+q+1}$. This stands in contrast to [14, 4, 5, 6], where the average number of points on curves in certain families is shown to be $q+1$, and [8] where this average is $< q+1$.

2000 *Mathematics Subject Classification.* 11G20,11T55,11M38.

Key words and phrases. Zeta functions of curves, Class field theory, Gaussian distribution.

F. Thorne was partially supported by the National Science Foundation under Grant No. DMS-1201330. M. Xiong was supported by the Research Grants Council of Hong Kong under Project Nos. RGC606211 and SBI12SC05.

The key ingredient in Wood's work is a study of the relative probabilities of splitting behaviors in cubic extensions of function fields. These probabilities were computed by Datskovsky and Wright [11], using properties of adelic Shintani zeta functions, but Datskovsky and Wright did not provide upper bounds for the error terms. As Wood pointed out, such an error term is extremely useful for a more detailed analysis.

Recent work on counting functions for cubic fields [2, 13, 21] has established good error terms in the number field setting, along with negative secondary main terms. The results in [2, 21] allow for finitely many splitting conditions to be imposed, and subject to appropriate limitations the error terms are uniform in the splitting conditions. These error estimates were then applied by Martin and Pollack [16] and Cho and Kim [9] to further study statistical properties of cubic number fields.

In the function field setting, the methods of [21] (at least) should work, but only with substantial effort. However, Y. Zhao [28] has demonstrated that the function field version of this problem admits a striking and novel approach using algebraic geometry: trigonal curves may be counted by embedding them into Hirzebruch surfaces and sieving for smoothness. Zhao obtained a negative secondary term in this setting as well, along with error terms uniform in appropriate splitting conditions.[1] (See Theorem 4 for a more precise statement of Zhao's results.)

Using Zhao's result [28], we first complement Wood's work by showing that as $g, q \to \infty$, the distribution of points on trigonal curves converges to a Gaussian.

Theorem 1. *If q and g both tend to infinity, then as C ranges over all elements of $T_g(\mathbb{F}_q)$, the limiting distribution of $\left(\#C(\mathbb{F}_q) - q - 2\right)/\sqrt{q+1}$ is a standard Gaussian with mean zero and variance one.*

That such results might follow from Zhao's work was previously suggested by Wood in [24].

The number of rational points on a curve over a finite field is determined by the zeta function, and statistical properties of the number of points may be interpreted as properties of the coefficients of the zeta function. Hence a related but more subtle question is to consider statistical properties of zeroes of the zeta function. In the case of hyperelliptic curves, these properties were studied by Faifman and Rudnick [12]. Some other interesting families of curves were studied in [7, 26, 27]. In this paper, combining the method of Faifman and Rudnick [12] and Zhao's work [28], we carry out a similar study.

Now we introduce some notation. For each $(C, \pi) \in T_g(\mathbb{F}_q)$, the Weil zeta function of C has the form

$$Z_C(u) = \frac{P_C(u)}{(1-u)(1-qu)}\,.$$

Here $P_C(u)$ is a polynomial of degree $2g$ with integral coefficients, satisfying the Riemann hypothesis ([22]). Hence we can write

$$P_C(u) = P_C(q^{-s}) = \prod_{j=1}^{2g} \left(1 - q^{1/2-s} e(\theta_{C,j})\right),$$

[1] Zhao's paper is still in preparation!

where $u := q^{-s}$ and $e(\theta_{C,j}) := e^{2\pi i\theta_{C,j}}$ with $\theta_{C,j} \in [-1/2, 1/2)$. We shall study the statistics of the set of angles $\{\theta_{C,j}\}$ as (C, π) varies in the finite set $T_g(\mathbb{F}_q)$, on which we assign the uniform probability. For this purpose, we fix an interval $\mathbf{I} \subset (-\frac{1}{2}, \frac{1}{2})$ of length $|\mathbf{I}|$, and for simplicity as in [12] we assume that $\mathbf{I}$ is symmetric around the origin. Define

$$N_{\mathbf{I}}(C) = \#\left\{j : \theta_{C,j} \in \mathbf{I}\right\}. \tag{2}$$

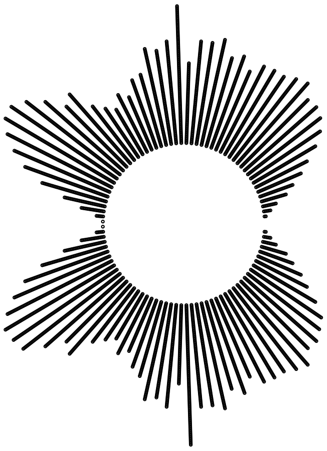
Theorem 2. *Let $\mathbb{F}_q$ have characteristic ≥ 5. Let T_g be defined as above. Let $\mathbf{I} \subset \left(-\frac{1}{2}, \frac{1}{2}\right)$ be a symmetric interval and assume that $g|\mathbf{I}| \to \infty$ as $g \to \infty$. Then for any real numbers a, b, we have*

$$\lim_{g\to\infty} \mathrm{Prob}_{T_g(\mathbb{F}_q)} \left(a < \frac{N_{\mathbf{I}}(C) - 2g|\mathbf{I}|}{\sqrt{\frac{2}{\pi^2}\log(2g|\mathbf{I}|)}} < b \right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}}\,\mathrm{d}x\,.$$

Note that the distribution of zeta zeroes has a slight bias not reflected in this theorem: the mean of $-q^{1/2}\sum_{i=1}^{2g} e(\theta_i)$ converges to $1 - \frac{1}{q^2+q+1}$ rather than zero.

1.1. **Numerical data.** Thanks to work of Rozenhart, Jacobson, and Scheidler [20] and Weir [23] on algorithmically enumerating cubic function fields, we may at least begin to illustrate the distribution of the zeroes of the zeta functions in question. Since the algorithms in question are computationally intensive, we limit ourselves to a single example.

The following picture illustrates the 581, 104 zeroes of the zeta functions associated to the 98, 124 trigonal curves of genus $g \leq 3$ over $\mathbb{P}^1(\mathbb{F}_5)$:



The graph is a histogram of the 581, 104 values $e(\theta_{C,j})$. To produce it, we used a list of cubic function fields kindly provided to us by C. Weir, and then used MAGMA [3] to compute the polynomials $P_C(u)$. Factoring these polynomials over $\mathbb{C}$, we then divided the interval $[0, 2\pi]$ into 100 equal intervals; the length of each radial line represents the number of angles $\theta_{C,j}$ in the corresponding interval. We also computed that the total of all the $e(\theta_i)$ was $-78, 701$, corresponding to an average number of points of $6 + \frac{78701}{98124} = 6.802\ldots$. This may be compared with the larger average value of $6 + \frac{30}{31}$.

To state the obvious, our graph is not suggestive of the uniform distribution proved for $g \to \infty$ in Theorem 2. Instead, our graph illustrates the 'slow convergence' phenomenon for data concerning cubic fields seen in other works such as [2, 21, 28]. (See also [1] for very striking data on ranks of elliptic curves.)

1.2. **Overview and outline of the paper.** In Section 2 we describe some of the ingredients of our proof. Chief among these is the 'explicit formula', and we also explicitly state the version of Zhao's theorem that we apply here.

In Section 3 we prove Theorem 2, largely following the methods of Faifman and Rudnick [12]. We conclude in Section 4 with the (fairly straightforward) proof of Theorem 1.

Unless otherwise noted, constants implied by the notation $\ll$ and $O(-)$ are allowed to depend on the moment parameter r (to be introduced later), but not on other variables. Occasionally we write $\ll_r$ to emphasize the r-dependence, but this dependence is allowed throughout.

2. Preliminaries

2.1. **Cubic extensions, trigonal curves, zeta functions and the explicit formula.** We first recall some basic facts on zeta functions of trigonal curves over finite fields; interested readers may refer to [19] for more details.

Let $\mathbb{F}_q$ have characteristic > 3. This assumption will let us avoid wild ramification and also inseparable extensions of function fields. For $(C, \pi) \in T_g(\mathbb{F}_q)$, let $\mathbb{K}$ be the function field of C, and let $k := \mathbb{F}_q(x)$ be the rational function field. Then $\mathbb{K}/k$ is a cubic extension. On the other hand, any (geometric) cubic extension $\mathbb{K}/k$ corresponds to a unique element $(C, \pi) \in T_g(\mathbb{F}_q)$. By the Hurwitz formula ([19, Chap. 7]),

$$\deg_k \mathrm{Disc}(\mathbb{K}/\mathrm{k}) = 2g + 4\,.$$

It is known by [11, Theorem I.1] that the number of all such cubic extensions of k is about the size q^{2g+4}, amongst which the number of cyclic extensions is bounded by $O\left(q^{g+2}g^2\right)$ (see [25, Theorem I.3]), since $g \to \infty$, to count elements of $T_g(\mathbb{F}_q)$, it suffices to count

$$E(k, g) := \{\mathbb{K} : \mathbb{K}/k \text{ is a non-cyclic cubic extension and } \deg_k \mathrm{Disc}(\mathbb{K}/k) = 2g + 4\}\,. \tag{3}$$

For $\mathbb{K}/k$ noncyclic, the zeta function $\zeta_{\mathbb{K}}(s)$ of $\mathbb{K}$ is given by

$$\zeta_{\mathbb{K}}(s) = \prod_{\omega \in \mathcal{S}_{\mathbb{K}}} \left(1 - N\omega^{-s}\right)^{-1},$$

where $\mathcal{S}_{\mathbb{K}}$ is the set of all places of $\mathbb{K}$. For each $\mathbb{K}$ we divide $\mathcal{S}_k$ into sets $\mathcal{S}_{111}$, $\mathcal{S}_{12}$, $\mathcal{S}_{11^2}$, $\mathcal{S}_{1^3}$, $\mathcal{S}_3$ depending on the splitting of k in $\mathbb{K}$; e.g. $\mathcal{S}_{111}$ is the set of places totally split in $\mathbb{K}$.[2] For $k = \mathbb{F}_q(x)$, we have $\zeta_k(s) = \prod_{v \in \mathcal{S}_k}(1 - |v|^{-s})^{-1}$, so that

$$\frac{\zeta_{\mathbb{K}}(s)}{\zeta_k(s)} = \prod_{v \in \mathcal{S}_{111}} \left(1 - |v|^{-s}\right)^{-2} \prod_{v \in \mathcal{S}_{12}} \left(1 - |v|^{-2s}\right)^{-1} \prod_{v \in \mathcal{S}_{11^2}} \left(1 - |v|^{-s}\right)^{-1} \prod_{v \in \mathcal{S}_3} \left(1 - |v|^{-3s}\right)^{-1} \left(1 - |v|^{-s}\right),$$

where $|v| := q^{\deg v}$ and $\mathrm{Re}(s) > 1$.

[2] Although we suppress $\mathbb{K}$ from the notation, note that these sets are different for each $\mathbb{K}$.

For the curve C whose function field is $\mathbb{K}$, it is known that $Z_C(q^{-s}) = \zeta_{\mathbb{K}}(s)$, that is, the zeta function of C coincides with the zeta function of the function field (see [18, p. 57, Chap. 5]). We have

$$\frac{\zeta_{\mathbb{K}}(s)}{\zeta_k(s)} = \prod_{i=1}^{2g} \left(1 - \sqrt{q} u e\left(\theta_{C,i}\right)\right) , \tag{4}$$

and therefore

Proposition 3 (The explicit formula). *We have, for each integer $n \geq 1$,*
(5)

$$-q^{n/2} \sum_{i=1}^{2g} e(n\theta_{C,i}) = \sum_{\substack{v \in \mathcal{S}_{111} \\ \deg v | n}} 2(\deg v) + \sum_{\substack{v \in \mathcal{S}_{12} \\ \deg v | \frac{n}{2}}} 2(\deg v) + \sum_{\substack{v \in \mathcal{S}_{{}_{11}2} \\ \deg v | n}} (\deg v) + \sum_{\substack{v \in \mathcal{S}_{3} \\ \deg v | \frac{n}{2}}} (\deg v) - \sum_{\substack{v \in \mathcal{S}_{3} \\ \deg v | n}} (\deg v) \, .$$

Proof. Comparing the two expressions for $\frac{\zeta_{\mathbb{K}}(s)}{\zeta_k(s)}$ above, take the logarithmic derivative with respect to s of each and equate the coefficients. □

Understanding the left side of (5) leads to an understanding of the distribution of the $\theta_{C,i}$ on average over C, and therefore (5) explains the relevance of understanding the quantities on the right.

2.2. **The distribution of cubic function fields.** We now see that we must understand the proportions of cubic function fields with splitting behaviors prescribed at finite sets of primes. Wood [24] previously applied work of Datskovsky and Wright [11], and we will apply the following quantitatively stronger result, obtained by Zhao [28]:

Theorem 4 (Zhao [28]). *Define $E(k, g)$ as above. For any finite set of places $\mathcal{S}$, and any set of splitting conditions at the primes in $\mathcal{S}$, define $E(k, g, \mathcal{S})$ to be the subset of $E(k, g)$ consisting of cubic extensions satisfying these splitting conditions.*

Then, there are fixed constants $\delta, A > 0$ for which

$$\frac{|E(k, g, \mathcal{S})|}{|E(k, g)|} = \prod_{v \in S} c_v + O\Big(q^{-\delta g} \prod_{v \in \mathcal{S}} |v|^{A}\Big) , \tag{6}$$

where the implied constant is absolute, and where c_v is given by

$$c_v \left(1 + |v|^{-1} + |v|^{-2}\right) = \begin{cases} 1/6 & \text{for } v \text{ totally split}\,, \\ 1/2 & \text{for } v \text{ partially split}\,, \\ 1/3 & \text{for } v \text{ inert}\,, \\ |v|^{-1} & \text{for } v \text{ partially ramified}\,, \\ |v|^{-2} & \text{for } v \text{ totally ramified}\,. \end{cases}$$

Zhao proved more, most notably a negative secondary term in each of $|E(k, g, \mathcal{S})|$ and $|E(k, g)|$; we have isolated that part of Zhao's results which we need. In our application, $\mathcal{S}$ will contain at most r places of fixed degree $\leq \ell$. Provided that we choose $r < \frac{\delta g}{2A\ell}$, our error term will be $\ll q^{-\delta g/2}$ and asymptotically bounded above by the main term, the implied constant depending on r but not ℓ, g, q, or (except via the r-dependence) $\mathcal{S}$.

It remains to be confirmed (by Zhao) that he obtains such a result. Nevertheless, we are fairly certain that he will, with specific (and fairly good) values of δ and A.

3. Distribution of zeta zeros for $E(k, g)$

3.1. **Preparation.** For a symmetric interval $\mathbf{I} = [-\beta/2, \beta/2]$, $0 < \beta < 1/2$, we let $I_l^{\pm}(x) = \sum_n c^{\pm}(n)e(nx)$ be the two *Beurling-Selberg polynomials* of degree l. These were also applied in [12], and we refer to [17, Chapter 1.2] for more details. These polynomials approximate $\mathbf{1}_{\mathbf{I}}$, the characteristic function of $\mathbf{I}$, with $I_l^- \leq \mathbf{1}_{\mathbf{I}} \leq I_l^+$, and satisfy the following additional properties (with $c(n) := c^{\pm}(n)$):

(i) $c(n) = 0$ if $|n| > l$ and $c(n) = c(-n)$.
(ii) $c(0) = \beta + O\left(l^{-1}\right)$.
(iii) $|c(n)n| \ll 1$ for any $n \in \mathbb{Z}$.
(iv) $\sum_{n \geq 1} nc(n)^2 = \frac{1}{2\pi^2}\log(l\beta) + O(1)$.
(v) $\sum_n c(2n) \ll 1$.
(vi) $|c^+(n) - c^-(n)| \leq \frac{2}{l+1}$ for each n.

Since $g|\mathbf{I}| = g\beta \to \infty$ as $g \to \infty$, we can choose integers $l = l(g)$ in such a way that

$$\frac{g}{l} \to \infty,\ l\beta \to \infty,\ \text{ and } \frac{g}{l} \ll (\log l\beta)^{1/4} \quad \text{as } g \to \infty.$$

For any $\mathbb{K} \in E(k, g)$, let $\{\theta_{\mathbb{K},i} : 1 \leq i \leq 2g\}$ be the angles of $\zeta_{\mathbb{K}}(s)$, as in (4). From the monotonicity of $I_l^{\pm}$, we have $N_l^-(\mathbb{K}) \leq N_{\mathbf{I}}(\mathbb{K}) \leq N_l^+(\mathbb{K})$, where

$$N_l^{\pm}(\mathbb{K}) = \sum_{i=1}^{2g} I_l^{\pm}\left(\theta_{\mathbb{K},i}\right), \quad N_{\mathbf{I}}(\mathbb{K}) = \left\{i : \theta_{\mathbb{K},i} \in \mathbf{I}\right\}.$$

Define $N_l(\mathbb{K}) := N_l^{\pm}(\mathbb{K})$. Then

$$N_l(\mathbb{K}) \ = \ \sum_{i=1}^{2g} I_l\left(\theta_{\mathbb{K},i}\right) = \sum_{n\in\mathbb{Z}} c(n) \sum_{i=1}^{2g} e\left(n\theta_{\mathbb{K},i}\right).$$

From the explicit formula (5) we obtain
(7)

$$N_l(\mathbb{K}) = 2gc(0) - 2\sum_{1\leq n\leq l} c(n)q^{-n/2}\left\{\sum_{\substack{v\in\mathcal{S}_{111}\\ \deg v|n}} 2\deg v + \sum_{\substack{v\in\mathcal{S}_{12}\\ \deg v|\frac{n}{2}}} 2\deg v + \sum_{\substack{v\in\mathcal{S}_{11^2}\\ \deg v|n}} \deg v - \sum_{\substack{v\in\mathcal{S}_{3}\\ \deg v|n}} \deg v + \sum_{\substack{v\in\mathcal{S}_{3}\\ \deg v|\frac{n}{3}}} 3\deg v\right\}.$$

The contribution of the terms with $\deg v \leq \frac{n}{3}$ to the expression inside the braces above is

$$\ll \sum_{1\leq n\leq l} |c(n)|q^{-n/2}q^{n/3} \ll 1.$$

We therefore obtain

$$N_l(\mathbb{K}) \;=\; 2g\beta + T_l(\mathbb{K}) + \triangle_1(\mathbb{K}) + \triangle_2(\mathbb{K}) + O\left(\frac{g}{l}\right) ,$$

where

$$T_l(\mathbb{K}) := -2\sum_{v\in\mathcal{S}_k} c(\deg v)|v|^{-1/2}(\deg v) f_{\mathbb{K}}(v) ,$$

with

$$f_{\mathbb{K}}(v) := \begin{cases} 2 & : \quad \text{if } v \in S_{111} , \\ -1 & : \quad \text{if } v \in S_3 , \\ 0 & : \quad \text{otherwise} ; \end{cases} \tag{8}$$

the term $\triangle_1(\mathbb{K})$, corresponding to pairs (v, n) with $v \in \mathcal{S}_{11^2}$ and $\deg v = n$, is given by

$$\triangle_1(\mathbb{K}) := -2\sum_{v\in\mathcal{S}_{11^2}} c(\deg v)|v|^{-1/2}(\deg v) ;$$

and $\triangle_2(\mathbb{K})$, corresponding to (v, n) with $\deg v = \frac{n}{2}$, is given by

$$\triangle_2(\mathbb{K}) := -2\sum_{v\in\mathcal{S}_k} c(2\deg v)|v|^{-1}(\deg v) g_{\mathbb{K}}(v) ,$$

where $g_{\mathbb{K}}(v) : \mathcal{S} \to \{2, 1, -1, 0\}$ is defined similarly to $f_{\mathbb{K}}(v)$.

We denote by $\langle \bullet \rangle$ the mean value of any quantity defined on $E(k, g)$, that is, let $\chi : E(k, g) \to \mathbb{C}$ be a map, then

$$\langle \chi \rangle := \frac{1}{\#E(k, g)} \sum_{\mathbb{K}\in E(k,d)} \chi(\mathbb{K}) . \tag{9}$$

3.2. **Moment computations.** For each positive integer r we now compute the rth moments of the various quantities above. In what follows, all implied constants may depend on r.

Proposition 5. *We have, for each r,*

$$\langle (\Delta_1)^r \rangle \ll_r 1 . \tag{10}$$

Proof. We have

$$\begin{aligned} \langle (\triangle_1)^r \rangle \;&=\; (-2)^r \sum_{v_1,\ldots,v_r\in\mathcal{S}_k} \prod_{i=1}^{r} c(\deg v_i)|v_i|^{-1/2}(\deg v_i)\frac{1}{\#E(k, g)} \sum_{\substack{\mathbb{K}\in E(k,g) \\ v_1,\ldots,v_r\in\mathcal{S}_{11^2}}} 1 \\ &\ll\; \sum_{\substack{1\le t\le r \\ \lambda_1+\cdots+\lambda_t=r \\ \lambda_i\ge 1}} \sum_{\substack{v_1,\ldots,v_t\in\mathcal{S}_k \\ \text{distinct}}} \prod_{i=1}^{t} \left\{|c(\deg v_i)|\,|v_i|^{-1/2}(\deg v_i)\right\}^{\lambda_i} \times \frac{\sum_{\substack{\mathbb{K}\in E(k,g) \\ v_1,\ldots,v_t\in\mathcal{S}_{11^2}}} 1}{\#E(k, g)} . \end{aligned}$$

The average over $E(k,g)$ is $\ll \prod_{i=1}^{t} |v_i|^{-1}$ by Zhao's theorem, so that

$$\langle (\Delta_1)^r \rangle \ll \left(\sum_{v \in \mathcal{S}_k} |c(\deg v)|\, |v|^{-3/2} (\deg v) \right)^r \ll 1\,. \tag{11}$$

□

Proposition 6. *We have, for each r,*

$$\langle (T_l)^r \rangle = \frac{r!\delta(r/2)}{\pi^r (r/2)!} \log^{r/2}(l\beta) + O_r\left(\log^{-1+r/2}(l\beta) \right), \tag{12}$$

where $\delta(c) = 1$ if $c \in \mathbb{Z}$ and $\delta(c) = 0$ if $c \notin \mathbb{Z}$.

Proof. For each r, we have

$$\begin{aligned} \langle (T_l)^r \rangle &= (-2)^r \sum_{\substack{1 \le t \le r \\ \lambda_1 + \cdots + \lambda_t = r \\ \lambda_i \ge 1}} \frac{r!}{t! \prod_i \lambda_i!} \sum_{\substack{v_1, \ldots, v_t \in \mathcal{S}_k \\ \text{distinct}}} \prod_{i=1}^{t} \left\{ c(\deg v_i) |v_i|^{-\frac{1}{2}} (\deg v_i) \right\}^{\lambda_i} \\ &\quad \times \frac{1}{\# E(k,g)} \sum_{\substack{\mathbb{K} \in E(k,g) \\ v_1, \ldots, v_t \in \mathcal{S}_k}} f_{\mathbb{K}}(v_1)^{\lambda_1} \cdots f_{\mathbb{K}}(v_t)^{\lambda_t}\,. \end{aligned}$$

We will find that the main contribution comes from the terms where each λ_i is equal to 2. In the first place, suppose that some λ_i is equal to 1. Then, the corresponding main terms from Zhao's theorem will cancel, leaving only an error term

$$\ll q^{-\delta g} \Big(\sum_{\substack{v \in \mathcal{S}_k \\ \deg v \le l}} |v|^{-1/2} \Big)^r \ll q^{-\delta g + lr/2} \ll 1.$$

We next study the terms where $\lambda_i = 2$ for each i (hence $r = 2s$ must be even). These terms are given by

$$\begin{aligned} \langle (T_l)^{2s} \rangle_1 &= (-2)^{2s} \frac{(2s)!}{2^s s!} \sum_{\substack{v_1, \ldots, v_s \in \mathcal{S}_k \\ \text{distinct}}} \prod_{i=1}^{s} \left\{ c(\deg v_i) |v_i|^{-\frac{1}{2}} (\deg v_i) \right\}^2 \\ &\quad \times \frac{1}{\# E(k,g)} \sum_{\substack{\mathbb{K} \in E(k,g) \\ v_1, \ldots, v_s \in \mathcal{S}_k}} f_{\mathbb{K}}(v_1)^2 \cdots f_{\mathbb{K}}(v_s)^2\,. \end{aligned}$$

The error in applying Zhao's theorem is again $\ll 1$, and the main term is

$$\frac{(-2)^{2s}(2s)!}{2^s\, s!} \sum_{\substack{v_1, \ldots, v_s \in \mathcal{S}_k \\ \text{distinct}}} \prod_{i=1}^{s} \left\{ c(\deg v_i)^2 |v_i|^{-1} (\deg v_i)^2 \left(1 + |v_i|^{-1} + |v_i|^{-2} \right)^{-1} \right\}.$$

If we remove the restraint that $v_1, \ldots, v_s$ are distinct, this becomes

$$\frac{2^s(2s)!}{s!}\left(\sum_{v\in\mathcal{S}_k} c(\deg v)^2|v|^{-1}(\deg v)^2\left(1+|v|^{-1}+|v|^{-2}\right)^{-1}\right)^s. \tag{13}$$

Simplifying, and using the prime number theorem in the form $\#\left\{v\in\mathcal{S}_k : \deg v = n\right\} = q^n/n + O\left(q^{n/2}\right)$, we obtain

$$\begin{aligned}&\frac{2^s(2s)!}{s!}\left(\sum_{n\le l} c(n)^2n^2q^{-n}\left(1+q^{-n}+q^{-2n}\right)^{-1}\left\{\frac{q^n}{n}+O(q^{n/2})\right\}\right)^s\\ &=\frac{2^s(2s)!}{s!}\left(\sum_{n\le l} c(n)^2n+O(1)\right)^s = \frac{2^s(2s)!}{s!}\left(\frac{1}{2\pi^2}\log l\beta+O(1)\right)^s\\ &=\frac{(2s)!}{\pi^{2s}s!}\log^s(l\beta)+O\left(\log^{s-1}(l\beta)\right).\end{aligned}$$

From this we also see that removing the restraint that the $v_1, \ldots, v_s$ are distinct resulted in an error bounded by $O\left(\log^{s-2}(l\beta)\right)$, and that the terms for which $\lambda_i \ge 2$ for each i and $\lambda_j \ge 3$ for some j contribute $O\left(\log^{(r-3)/2}(l\beta)\right)$. This completes the proof. □

Proposition 7. *We have, for each r,*

$$\langle(\Delta_2)^r\rangle \ll_r 1. \tag{14}$$

Proof. We again start by expanding out the rth power and extracting the error term from Zhao's theorem, which is $\ll 1$. The remaining main term is

$$\left(-2\sum_{v\in\mathcal{S}_k} c(2\deg v)|v|^{-1}(\deg v)\widetilde{g}(v)\right)^r, \tag{15}$$

where $\widetilde{g}(v) = \frac{1+|v|^{-1}}{1+|v|^{-1}+|v|^{-2}} = 1 + O\left(|v|^{-1}\right)$ is the main term in the average of $\widetilde{g}_{\mathbb{K}}(v)$ over $\mathbb{K}$, coming again from Zhao's theorem. By the prime number theorem and property (v) of the $c(n)$, (15) is the rth power of

$$-2\sum_{n\le l} c(2n)q^{-n}n\left(1+O\left(q^{-n}\right)\right)\left(\frac{q^n}{n}+O(q^{n/2})\right) = \sum_{n\le l} c(2n)+O(1) \ll 1\,.$$

□

3.3. Proof of Theorem 2 for $E(k, g)$.

Since

$$N_l(\mathbb{K}) \;=\; 2g\beta + T_l(\mathbb{K}) + \triangle_1(\mathbb{K}) + \triangle_2(\mathbb{K}) + O\left(\frac{g}{l}\right),$$

by combining our previous estimates we obtain

$$\left\langle\left(\frac{N_l(\bullet)-2g\beta}{\sqrt{\frac{2}{\pi^2}\log(l\beta)}}\right)^r\right\rangle = \frac{\delta(r/2)r!}{2^{r/2}\,(r/2)!} + O\left(\log^{-\frac{1}{4}}(l\beta)\right). \tag{16}$$

We claim that (16) holds true if the term $N_l(\bullet) := N_l^{\pm}(\bullet)$ is replaced by $N_{\mathbf{I}}(\bullet)$. Granting this, since for a standard Gaussian distribution the odd moments vanish and the even moments are

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^{2r}e^{-x^2/2}\mathrm{d}x = \frac{(2r)!}{2^r r!}\,,$$

we shall have concluded that as $\mathbb{K}$ runs through $E(k,g)$ with $g \to \infty$, the term

$$\frac{N_{\mathbf{I}}(\mathbb{K}) - 2g\beta}{\sqrt{\frac{2}{\pi^2}\log(l\beta)}}$$

converges weakly to a standard Gaussian distribution, hence proving Theorem 2 for $E(k,g)$.

It remains to prove the claim. We define

$$W_l(\mathbb{K}) \;\; := \;\; N_l^{+}(\mathbb{K}) - N_l^{-}(\mathbb{K})\,,$$

for which we see, for each r, that

$$\langle (W_l)^r \rangle \;\; \ll \log^{r/4}(l\beta)\,. \tag{17}$$

This is proved in the same manner as our previous bounds. The difference is that we replace expressions of the form $c(j\deg v)$ (with $j = 1,2,3$) with expressions of the form $c^{+}(j\deg v) - c^{-}(j\deg v)$, for which we have $|c^{+}(n) - c^{-}(n)| \leq \frac{1}{2(l+1)}$ for each n. In all cases the resulting expressions are bounded above by the previous expressions; moreover, we now obtain a bound of $O(1)$ in (13). We also apply the upper bound $g/l \ll \log^{r/4}(l\beta)$, and thereby deduce (17).

With this estimate, and defining $V_l(\mathbb{K})$ by

$$N_{\mathbf{I}}(\mathbb{K}) - 2g\beta = N_l^{-}(\mathbb{K}) - 2g\beta + V_l(\mathbb{K})\,,$$

we obtain

$$(N_{\mathbf{I}}(\mathbb{K}) - 2g\beta)^r = \left(N_l^{-}(\mathbb{K}) - 2g\beta\right)^r + E(\mathbb{K})\,,$$

where

$$|E(\mathbb{K})| \;\; \ll \;\; \sum_{\substack{u+v=r \\ v\geq 1}} \left|\left(N_l^{-}(\mathbb{K}) - 2g\beta\right)^u W_l(\mathbb{K})^v\right|\,.$$

Using the Cauchy-Schwartz inequality and (16) and (17), we have

$$\begin{aligned}\langle |E(\bullet)|\rangle &\ll \sum_{\substack{u+v=r \\ v\geq 1}} \left\langle \left(N_l^{-}(\bullet) - 2g\beta\right)^{2u}\right\rangle^{1/2} \left\langle (W_l)^{2v}\right\rangle^{1/2} \\ &\ll \sum_{\substack{u+v=r \\ v\geq 1}} (\log l\beta)^{u/2}(\log l\beta)^{v/4} \ll (\log l\beta)^{\frac{r}{2}-\frac{1}{4}}\,.\end{aligned}$$

This implies that

$$\left\langle \left(\frac{N_{\mathbf{I}}(\bullet) - 2g\beta}{\sqrt{\frac{2}{\pi^2}\log(l\beta)}}\right)^r \right\rangle = \frac{\delta(r/2)(r)!}{2^{r/2}\,(r/2)!} + O\left(\log^{-\frac{1}{4}}(l\beta)\right)\,. \tag{18}$$

In the denominator we can also replace l by $2g$ because

$$\log(l\beta) = \log(2g\beta) + O(\log\log(2g\beta))\,.$$

This concludes the proof of Theorem 2 for $E(k,g)$ as $g \to \infty$. $\square$

4. Gaussian moments in Wood's theorem

We have the well known formula

$$\#C(\mathbb{F}_q) - q - 1 = -q^{1/2}\sum_{i=1}^{2g} e\left(\theta_{C,i}\right),$$

and using Proposition 3 with $n=1$ we write

$$W(\mathbb{K}) := \#C(\mathbb{F}_q) - q - 1 = \sum_{\substack{v\in\mathcal{S}_{111}\\ \deg v=1}} 2 + \sum_{\substack{v\in\mathcal{S}_{11^2}\\ \deg v=1}} 1 - \sum_{\substack{v\in\mathcal{S}_{3}\\ \deg v=1}} 1\,. \tag{19}$$

where $\mathbb{K}$ is the cubic extension corresponding to the trigonal curve C.

Proposition 8. *For each $\mathbb{K}\in E(k,g)$ and positive integer r, we have*

$$\frac{1}{\#E(k,g)}\sum_{\mathbb{K}\in E(k,g)} W(\mathbb{K})^r = \mathbb{E}(\widetilde{X}^r) + O\left(q^{-\delta g + Ar + 3}\right), \tag{20}$$

where $\widetilde{X} := \sum_{v\in\mathcal{S}_k}(X_v - 1)$ is a sum of $q+1$ independent identically distributed random variables $\{X_v - 1\}_{v\in\mathcal{S}_k}$, with each X_v is defined by (1).

Proof. By construction, the main term of (20) is precisely the main term originating from Zhao's theorem (Theorem 4), so it remains to evaluate the sum of the error terms.

We apply Zhao's theorem for all combinations of places $\{v_1,\dots,v_j\}$ with $j \le \min(q+1,r)$ and all choices of splitting types in $\{\mathcal{S}_{111},\mathcal{S}_{11^2},\mathcal{S}_3\}$ for each of the v_i, hence $\ll (3q)^r$ times. The error term in each application of Zhao's theorem is bounded by $q^{-\delta g+Ar}$, weighted by a factor bounded by 2^r (the maximum occurring when $v_i\in\mathcal{S}_{111}$ for each i). Therefore the total error is bounded above by $q^{-\delta g+Ar+1}(6q)^r$, and plugging in the crude bound $6^r < q^{2r}$ finishes the proof. $\square$

Taking $r=1$ gives Theorem 1.1 of Wood [24]. Moreover, since as $q\to\infty$,

$$\mathbb{E}(X_i) = q^{-1}(1+q^{-1}+q^{-2})^{-1} \sim q^{-1},$$

$$\mathbb{E}(X_i^2) = (1+q^{-1})^{-1}(1+q^{-1}+q^{-2})^{-1} \sim 1,$$

we find that as $q,g\to\infty$ and $\mathbb{K}$ ranges over $E(k,g)$, the limiting distribution of $\frac{W(\mathbb{K})-1}{\sqrt{q+1}}$ becomes the standard Gaussian with mean zero and variance one. This completes the proof of Theorem 1. $\square$

References

[1] B. Bektemirov, B. Mazur, W. Stein, and M. Watkins, *Average ranks of elliptic curves: tension between data and conjecture*, Bull. Amer. Math. Soc. (N.S.) **44** (2007), no. 2, 233–254.

[2] M. Bhargava, A. Shankar, and J. Tsimerman, *On the Davenport-Heilbronn theorem and second order terms*, Inv. Math., accepted for publication.

[3] W. Bosma, J. Cannon, and C. Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), 235–265.

[4] A. Bucur, C. David, B. Feigon, M. Lalín, *Statistics for traces of cyclic trigonal curves over finite fields*, Int. Math. Res. Not. IMRN (2010), No. 5, 932–967.

[5] A. Bucur, C. David, B. Feigon, M. Lalín, *Fluctuations in the number of points on smooth plane curves over finite fields*, J. Number Theory **130** (2010), no. 11, 2528–2541.

[6] A. Bucur, C. David, B. Feigon, M. Lalín, *Biased statistics for traces of cyclic p-fold covers over finite fields*, WINŮwomen in numbers, 121–143, Fields Inst. Commun., **60**, Amer. Math. Soc., Providence, RI, 2011.

[7] A. Bucur, C. David, B. Feigon, M. Lalín, K. Sinha, *Distribution of zeta zeroes of Artin-Schreier curves*, to appear in Math. Res. Lett.

[8] A. Bucur, K. Kedlaya, *The probability that a complete intersection is smooth*, J. Theor. Nombres Bordeaux **24** (2012), no. 3, 541–556.

[9] P. J. Cho and H. Kim, *Low lying zeros of Artin L-functions*, preprint.

[10] A. Entin, *On the distribution of zeroes of Artin-Schreier L-functions*, Geom. Funct. Anal. **22** (2012), 1322–1360.

[11] B. Datskovsky and D. Wright, *Density of discriminants of cubic extensions*, J. Reine Angew. Math. **386** (1988), 116–138.

[12] D. Faifman, Z. Rudnick, *Statistics of the zeros of zeta functions in families of hyperelliptic curves over a finite field*, Compos. Math. **146** (2010), no. 1, 81–101.

[13] B. Hough, *Equidistribution of Heegner points associated to the 3-part of the class group*, preprint; available at `http://arxiv.org/abs/1005.1458`.

[14] P. Kurlberg, Z. Rudnick, *the fluctuations in the number of points on a hyperelliptic curve over a finite field*, J. Number Theory **129** (2009), no. 3, 580–587.

[15] P. Kurlberg, I. Wigman, *Gaussian point count statistics for families of curves over a fixed finite field*, Int. Math. Res. Not. IMRN 2011, no. 10, 2217–2229.

[16] G. Martin and P. Pollack, *The average least character non-residue and further variations on a theme of Erdős*, J. Lond. Math. Soc. (2) **87** (2013), no. 1, 22–42.

[17] H. L. Montgomery, "Ten lectures on the interface between analytic number theory and harmonic analysis". CBMS Regional Conference Series in Mathematics, **84**. American Mathematical Society, Providence, RI, 1994.

[18] C. Moreno, "Algebraic curves over finite fields", Cambridge Tracts in Mathematics **97**, Cambridge University Press, 1991.

[19] M. Rosen, "Number theory in function fields". Graduate Texts in Mathematics, **210**. Springer-Verlag, New York, 2002.

[20] P. Rozenhart, M. Jacobson Jr., and R. Scheidler, *Tabulation of cubic function fields via polynomial binary cubic forms*, Math. Comp. **81** (2012), 2335–2359.

[21] T. Taniguchi and F. Thorne, *Secondary terms in counting functions for cubic fields*, Duke Math. J., to appear.

[22] A. Weil, *Sur les Courbes Algébriques et les Variétés qui s'en Déduisent*, Publ. Inst. Math. Univ. Strasbourg **7** (1945), Hermann et Cie., Paris, 1948.

[23] C. Weir, *Constructing and tabulating global function fields*, Ph.D. thesis, University of Calgary, 2013.

[24] M.M. Wood, *The distribution of the number of points on trigonal curves over $\mathbb{F}_q$*, Int. Math. Res. Not. IMRN 2012, doi: 10.1093/imrn/rnr256.

[25] D. Wright, *Distribution of discriminants of abelian extensions*, Proc. London Math. Soc. (3) **58** (1989), no. 1, 17–50.

[26] M. Xiong, *Statistics of the zeros of zeta functions in a family of curves over a finite field*, Int. Math. Res. Not. IMRN 2010, no. 18, 3489–3518.

[27] M. Xiong, *Distribution of zeta zeroes for abelian covers of algebraic curves over a finite field*, preprint.

[28] Y. Zhao, *to be added*.