PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 143, Number 4, April 2015, Pages 1365–1375
S 0002-9939(2014)12218-0
Article electronically published on December 15, 2014

THE DISTRIBUTION OF POINTS ON SUPERELLIPTIC CURVES OVER FINITE FIELDS

GILYOUNG CHEONG, MELANIE MATCHETT WOOD, AND AZEEM ZAMAN

(Communicated by Matthew A. Papanikolas)

Abstract. We give the distribution of points on smooth superelliptic curves over a fixed finite field, as their degree goes to infinity. We also give the distribution of points on smooth m-fold cyclic covers of the line, for any m, as the degree of their superelliptic model goes to infinity. This builds on the previous work of Kurlberg, Rudnick, Bucur, David, Feigon, and Lalín for p-fold cyclic covers, but the limits taken differ slightly and the resulting distributions are interestingly different.

1. Introduction

For a family of smooth curves over a fixed finite field $\mathbb{F}_q$, it is natural to ask how many $\mathbb{F}_q$-points those curves have on average, or more precisely, what distribution of points one obtains from a random curve in the family. This question has been studied as the genus, degree, or related invariants go to infinity in many cases including: hyperelliptic curves [KR09], cyclic trigonal curves [BDFL10b], cyclic p-fold curves [BDFL11], plane curves [BDFL10a], complete intersections in projective space [BK11], trigonal curves [Woo12], and curves in Hirzebruch surfaces [EW12].

In this paper, we give the distribution of points on smooth superelliptic (affine) curves, C_f, given by the equation

$$y^m = f(x)$$

for varying $f(x) \in \mathbb{F}_q[x]$, with q and m fixed, as the degree of f goes to infinity. We also give the distribution for the normalizations of (possibly singular) superelliptic curves, which for $q \equiv 1 \pmod{m}$ is exactly the case of m-fold cyclic covers of the line (c.f. with the work on prime degree cyclic covers cited above). One important advance in our work is that the normalizations may have different numbers of points from their singular models, and it is, in general, difficult to write down the smooth curves explicitly. The above-cited works have all counted points on curves with explicitly given equations. (See also [Xio10] which gives the distribution on certain families of explicitly given superelliptic curves that include singular curves.) In Section 2 of this paper, we relate the number of points on a smooth curve to its explicitly given superelliptic model.

In contrast to studying a similar question when a family of fixed genus is chosen and $q \to \infty$, the philosophy of Katz-Sarnak [KS99] suggests that the limit distributions should be predicted by a certain group of random matrices. In the large

Received by the editors October 1, 2012 and, in revised form, February 14, 2013.
2010 *Mathematics Subject Classification.* Primary 11G20, 11R45, 11R58, 11T55, 14H25; Secondary 11G25, 11R20, 11T06.

genus limit for fixed q, there is no general conjectural picture of what one should expect. Thus it is important to have many examples of different families exhibiting different phenomena. In the cases we study, we obtain a particularly interesting contrast to [BDFL10b, BDFL11], as we count the same cyclic covers of the line, but with a different invariant going to infinity, and obtain different distributions.

Throughout, m will be fixed and $(q, m) = 1$. We write $\mathbb{F}_q[x]_d$ to denote the degree d polynomials in $\mathbb{F}_q[x]$. Note that C_f is smooth if and only if $f(x)$ is square-free, for which curves we have the following result.

Theorem 1.1. *Let $(q, m) = 1$. Then*

$$\lim_{d\to\infty} \frac{\#\{f \in \mathbb{F}_q[x]_d \mid \text{the curve } C_f \text{ is smooth with } k\ \mathbb{F}_q\text{-points}\}}{\#\{f \in \mathbb{F}_q[x]_d \mid \text{the curve } C_f \text{ smooth}\}} = \mathrm{Prob}(\sum_{i=1}^{q} X_i = k),$$

where the X_j are independent and identically distributed random variables with

$$X_j = \begin{cases} 0 \text{ with probability } \left(1 - \frac{1}{(m,q-1)}\right)\frac{1}{1+q^{-1}}, \\ 1 \text{ with probability } \frac{q^{-1}}{1+q^{-1}}, \\ (m, q-1) \text{ with probability } \frac{1}{(m,q-1)}\frac{1}{1+q^{-1}}, \end{cases}$$

where when $(m, q-1) = 1$ the last two probabilities are added to give $\mathrm{Prob}(X_j = 1)$. *We have the same results if we restrict to f such that C_f is irreducible or geometrically irreducible.*

In the case $m = 2$, Theorem 1.1 reduces to [KR09, Theorem 1] giving the distribution of $\mathbb{F}_q$-points on hyperelliptic curves.

When $f(x)$ is not square-free, we can consider the normalization $\tilde{C}_f$ of C_f, which is a smooth curve. Note that for $g \in \mathbb{F}_q[x]$, the curves C_f and C_{fg^m} have the same normalization. So it is natural to consider the smooth normalizations $\tilde{C}_f$ for f that are mth power-free.

Theorem 1.2. *Let $(q, m) = 1$. Then*

$$\lim_{d\to\infty} \frac{\#\{f \in \mathbb{F}_q[x]_d \mid f \text{ is } m\text{th power-free and } \tilde{C}_f \text{ has } k\ \mathbb{F}_q\text{-points}\}}{\#\{f \in \mathbb{F}_q[x]_d \mid f \text{ is } m\text{th power-free}\}} = \mathrm{Prob}(\sum_{i=1}^{q} X_i = k),$$

where the X_j are independent and identically distributed random variables with

$$X_j = \begin{cases} 0 \text{ with probability } \displaystyle\sum_{0\le s\le m-1} q^{-s}\left(1 - \frac{1}{(m, s, q-1)}\right)\frac{1-q^{-1}}{1-q^{-m}}, \\ N \text{ with probability } \displaystyle\sum_{\substack{0\le s\le m-1 \\ (m,s,q-1)=N}} \frac{q^{-s}(1-q^{-1})}{N(1-q^{-m})}. \end{cases}$$

We have the same results if we restrict to f such that $\tilde{C}_f$ is irreducible or geometrically irreducible.

When $q \equiv 1 \pmod{m}$, by Kummer theory, the cyclic m-fold covers of the line are exactly the irreducible $\tilde{C}_f$ of Theorem 1.2. Both Theorems 1.1 and 1.2 will follow from Theorem 3.1, which will give the distribution of points on superelliptic curves and their normalizations with $f(x)$ that are nth power-free, further refined by x values. We prove Theorem 3.1 by first relating the number of points on the normalizations to the explicit affine model in Section 2, and then using counting

methods similar to those of [KR09] in Section 3 (with the difference that our setup allows us to avoid counting polynomials interpolated to take zero values).

Note that the average number of points in both theorems above is q, which agrees with the average in [BDFL10b, BDFL11, KR09] for prime degree cyclic covers of the (affine) line. However, our distributions differ from those in [BDFL10b, BDFL11], which is not a contradiction as we take different invariants going to infinity. In our case, we are letting the degree of f, or C_f, the affine model of the curve, go to infinity. Cyclic p-fold curves have a signature $d_1, \dots, d_{p-1}$, and in [BDFL10b, BDFL11] one has $\min_i(d_i) \to \infty$. It would be very interesting to understand why and how this difference has an impact on the resulting distribution of points, not only in this case, but more generally. For this reason, we particularly highlight the first contrast, in the case of cyclic trigonal curves, in which we have Theorem 1.2 when $q \equiv 1 \pmod 3$ with

$$X_j = \begin{cases} 0 \text{ with probability } \dfrac{2}{3(1+q^{-1}+q^{-2})}, \\ 1 \text{ with probability } \dfrac{q^{-1}+q^{-2}}{1+q^{-1}+q^{-2}}, \text{ and} \\ 3 \text{ with probability } \dfrac{1}{3(1+q^{-1}+q^{-2})}, \end{cases}$$

and when instead of $\deg(f) \to \infty$, we have $\min_i(d_i) \to \infty$, [BDFL10b] gives an analogous theorem with

$$X_j = \begin{cases} 0 \text{ with probability } \dfrac{2}{3(1+2q^{-1})}, \\ 1 \text{ with probability } \dfrac{2q^{-1}}{1+2q^{-1}}, \text{ and} \\ 3 \text{ with probability } \dfrac{1}{3(1+2q^{-1})}. \end{cases}$$

The difference seems very suggestive but the reason for it is not immediately clear.

2. Points on the normalization

In this section, we determine the number of points on the normalization $\tilde{C_f}$ in terms of f. Note that we have a map $\tilde{C_f} \to C_f \to \mathbb{A}^1 = \operatorname{Spec} \mathbb{F}_q[x]$, and so a degree 1 point of C_f maps to a degree 1 point of $\mathbb{A}^1 = \operatorname{Spec} \mathbb{F}_q[x]$, and thus we can talk about the x-value of such a point.

We say $f \in \mathbb{F}_q[x]$ has *power r over a field F* if r is the greatest integer such that $f(x) = g(x)^r$ for some $g \in F[x]$. The following lemma deals with when our curve C_f is irreducible or geometrically irreducible.

Lemma 2.1. *If the power of $f(x)$ over $\bar{\mathbb{F}_q}$ is relatively prime to m, then $y^m - f(x)$ is irreducible in $\bar{\mathbb{F}_q}[x, y]$ (i.e. C_f is geometrically irreducible).*

Proof. If the power of f over $\bar{\mathbb{F}_q}$ is relatively prime to m, then f has order m in $\bar{\mathbb{F}_q}(x)^*/(\bar{\mathbb{F}_q}(x)^*)^m$. Then, by Kummer theory, we have that $y^m - f(x)$ is irreducible in $\bar{\mathbb{F}_q}[x, y]$. □

Lemma 2.2. *Let $f \in \mathbb{F}_q[x]$ have power over $\bar{\mathbb{F}_q}$ relatively prime to m and for some $x_0 \in \mathbb{F}_q$, we write $f(x) = (x - x_0)^s g(x)$ with $s \geq 0$, and $g \in \mathbb{F}_q[x]$, and $g(x_0) \neq 0$.*

Then, the number of degree 1 *points of* $\tilde{C}_f$ *with* $x = x_0$ *is equal to*

$$\begin{cases} 0 & \text{if } g(x_0) \text{ is not an } (m,s) \text{ power in } \mathbb{F}_q, \\ (m, s, q-1) & \text{if } g(x_0) \text{ is an } (m,s) \text{ power in } \mathbb{F}_q. \end{cases}$$

Proof. The number of degree 1 points of $\tilde{C}_f$ with $x = x_0$ for some $x_0 \in \mathbb{F}_q$ is equal to the number of degree 1 primes over $(x - x_0)$ in $K = \mathbb{F}_q(x)[y]/(y^m - f(x))$ by the usual correspondence between curves and their function fields. (We have that K is a field by Lemma 2.1.) Without loss of generality, we can consider the case $x_0 = 0$. We have $f(x) = x^s g(x)$, where $g(x) \in \mathbb{F}_q[x]$ and $g(0) \neq 0$.

We let $z = y^{m/(m,s)} x^{-s/(m,s)} \in K$ and $z^{(m,s)} = g(x)$, so z is integral over $\mathbb{F}_q[x]$. Thus at any prime $\wp$ over (x), modulo $\wp$ we have $z^{(m,s)} \equiv g(0)$. Thus if $g(0)$ is not an (m,s) power in $\mathbb{F}_q$, then the inertia degree at $\wp$ is > 1, and there are no degree 1 primes of K over (x).

Note that K contains $L = \mathbb{F}_q(x)[z]/(z^{(m,s)} - g(x))$ (since the power of f over $\bar{\mathbb{F}}_q$ is relatively prime to m, the power of g over $\bar{\mathbb{F}}_q$ is relatively prime to (m,s) and thus this polynomial is irreducible by Lemma 2.1). We have that L is unramified over (x). Since $\mathbb{F}_q[x,z]/(z^{(m,s)} - g(x))$ is maximal at (x), we can compute the splitting of (x) in L by computing the splitting of $z^{(m,s)} - g(x)$ modulo (x). If $g(0)$ is an (m,s) power in $\mathbb{F}_q$, then $z^{(m,s)} - g(x)$ has exactly $(m, s, q-1)$ distinct degree 1 factors modulo (x), and there are exactly $(m, s, q-1)$ degree 1 primes in L over (x).

We write $(m,s) = mi + sj$ for some $i, j \in \mathbb{Z}$, and then $(y^j x^i)^m = x^{mi+sj} g(x)^j$ has valuation (m,s) in $\mathbb{F}_q(x)$ (with respect to x). So at any place of K above (x) there is an element of valuation $\frac{(m,s)}{m}$, and thus any prime over (x) has ramification degree $e \geq \frac{m}{(m,s)}$. This means every prime of L over (x) must be completely ramified from L to K, and thus there are exactly $(m, s, q-1)$ degree 1 primes in K over (x). The lemma follows. □

3. Main theorem

Throughout this section, we fix an integer $n \geq 2$. We say a polynomial $f \in \mathbb{F}_q[x]$ is *nth power free* if f is not a multiple of the nth power of any positive degree polynomial in $\mathbb{F}_q[x]$. Let

$$\mathcal{F} := \{f \in \mathbb{F}_q[x] \mid f \text{ is } n\text{th power-free}\}$$

and $\mathcal{F}_d$ be the elements of $\mathcal{F}$ of degree d. We write $x_1, \ldots, x_q$ for the elements of $\mathbb{F}_q$.

Theorem 3.1. *Let* $k_1, \ldots, k_q$ *be integers and* $(q, m) = 1$. *Then*

$$\lim_{d \to \infty} \frac{\#\{f \in \mathcal{F}_d \mid \text{for each } i,\ C_f \text{ has } k_i \text{ points with } x = x_i\}}{\#\mathcal{F}_d} = \operatorname{Prob}(X_i = k_i \text{ for all } i),$$

where the X_j *are independent and identically distributed random variables with*

$$X_j = \begin{cases} 0 \text{ with probability } \left(1 - \frac{1}{(m,q-1)}\right) \frac{1-q^{-1}}{1-q^{-n}}, \\ 1 \text{ with probability } \frac{q^{-1}-q^{-n}}{1-q^{-n}}, \\ (m, q-1) \text{ with probability } \frac{1}{(m,q-1)} \frac{1-q^{-1}}{1-q^{-n}}, \end{cases}$$

where when $(m, q-1) = 1$ *the last two probabilities are added to give* $\mathrm{Prob}(X_j = 1)$. *Also,*

$$\lim_{d\to\infty} \frac{\#\{f \in \mathcal{F}_d \mid \textit{for each } i,\ \tilde{C}_f \textit{ has } k_i \textit{ points with } x = x_i \textit{ for all } i\}}{\#\mathcal{F}_d} = \mathrm{Prob}(X_i = k_i \textit{ for all } i),$$

where the X_j *are independent and identically distributed random variables with*

$$X_j = \begin{cases} 0 \textit{ with probability } \displaystyle\sum_{0\le s\le n-1} q^{-s}\left(1 - \frac{1}{(m, s, q-1)}\right)\frac{1-q^{-1}}{1-q^{-n}}, \\ N \textit{ with probability } \displaystyle\sum_{\substack{0\le s\le n-1 \\ (m,s,q-1)=N}} \frac{q^{-s}(1-q^{-1})}{N(1-q^{-n})}. \end{cases}$$

We have the same results if we restrict to f *such that* C_f *(or* $\tilde{C}_f$*) is irreducible or geometrically irreducible.*

When $q \equiv 1 \pmod m$ and $m = n$, Theorem 3.1 reduces to [Xio10, Theorem 1.1].

3.1. **Notation.** For a prime $h \in \mathbb{F}_q[x]$, we write $h^s || f$ if $h^s | f$ and $h^{s+1} \nmid f$. For $s = (s_1, \dots, s_q) \in \mathbb{N}^q$, let

$$\mathcal{F}^{\mathbf{s}} := \{f \in \mathcal{F} \mid \ (x - x_i)^{s_i} || f \text{ for all } i\},$$

and $\mathcal{F}_d^{\mathbf{s}}$ be the elements of $\mathcal{F}^{\mathbf{s}}$ of degree d. Note that if some $_i \ge n$, then $\mathcal{F}^{\mathbf{s}}$ is empty. We use V to denote the set of monic polynomials in $\mathbb{F}_q[x]$, and $V_d \subset V$ to denote those of degree d.

We define the zeta function

$$\zeta(s) := \prod_{\substack{P\in\mathbb{F}_q[x] \\ \text{P is monic irreducible}}} (1 - q^{-s\deg(P)}) = \frac{1}{1-q^{1-s}}$$

and the Möbius function for $f \in \mathbb{F}_q[x]$

$$\mu(f) := \begin{cases} 0 & \text{if } f \text{ is not square-free,} \\ (-1)^k & \text{if } f \text{ is the product of } k \text{ distinct irreducible factors.} \end{cases}$$

3.2. **Lemmas.** The following is an analog of [KR09, Lemma 5] (which is the special case $n = 2$). We count the number of nth power-free polynomials interpolating given values.

Lemma 3.2. *Let* $x_1, \cdots, x_l \in \mathbb{F}_q$ *be distinct elements, and fix any* $a_1, \cdots, a_l \in \mathbb{F}_q^*$. *Then*

$$\#\{f \in \mathcal{F}_d \mid f(x_1) = a_1, \cdots, f(x_l) = a_l\} = \frac{q^{d-l}(q-1)}{\zeta(n)(1-q^{-n})^l} + O(q^{d/n+1}),$$

where the constant in the O *is absolute.*

Proof. By inclusion-exclusion, we have

$$\begin{aligned}
&\#\{f \in \mathcal{F}_d \mid f(x_1) = a_1, \cdots, f(x_l) = a_l\} \\
&= \sum_{\substack{f_2 \in V \\ 0 \le \deg(f_2) \le d/n}} \mu(f_2) \#\{f_1 \in \mathbb{F}_q[x]_{d-n\deg(f_2)} \mid f_2(x_j)^n f_1(x_j) = a_j \text{ for } 1 \le j \le l\} \\
&= \sum_{\substack{f_2 \in V \\ 0 \le \deg(f_2) \le d/n}} \mu(f_2) \#\{f_1 \in \mathbb{F}_q[x]_{d-n\deg(f_2)} \mid f_1(x_j) = a_j/f_2(x_j)^n \text{ for } 1 \le j \le l\}.
\end{aligned}$$

If $f_2(x_j) = 0$, then by convention, there are no f_1 such that $f_1(x_j) = a_j/f_2(x_j)^n$.

If $D \ge l$, there are $(q-1)q^{D-\ell}$ degree D elements of $\mathbb{F}_q[x]$ that interpolate l given values because each interpolation point imposes a linearly independent condition on the coefficients of a monic polynomial of degree D (by the Vandermonde determinant). If $l \ge D+1$, then $D+1$ of these conditions are still linearly independent on the set of (not necessarily monic) polynomials of degree at most D, and thus there is at most 1 polynomial of degree D interpolating l given values.

We split the above expression into two sums depending on the degree of f_2, and thus the above is

$$= \sum_{\substack{f_2 \in V \\ 0 \le \deg(f_2) \le (d-l)/n \\ f_2(x_j) \ne 0 \text{ for } 1 \le j \le l}} \mu(f_2)(q-1)q^{d-n\deg(f_2)-l} + O\left(\sum_{\substack{f_2 \in V \\ (d-l)/n < \deg(f_2) \le d/n \\ f_2(x_j) \ne 0 \text{ for } 1 \le j \le l}} 1 \right).$$

We bound the error term by counting all monic polynomials of degree at most d/n to see that the above is

$$\begin{aligned}
&= \sum_{\substack{f_2 \in V \\ 0 \le \deg(f_2) \le (d-l)/n \\ f_2(x_j) \ne 0 \text{ for } 1 \le j \le l}} \mu(f_2)(q-1)q^{d-n\deg(f_2)-l} + O(q^{d/n}) \\
&= q^{d-l} \sum_{\substack{f_2 \in V \\ f_2(x_j) \ne 0 \text{ for } 1 \le j \le l}} \mu(f_2)(q-1)q^{-n\deg(f_2)} - q^{d-l} \\
&\quad \times \sum_{\substack{f_2 \in V \\ (d-l)/n < \deg(f_2) \\ f_2(x_j) \ne 0 \text{ for } 1 \le j \le l}} \mu(f_2)(q-1)q^{-n\deg(f_2)} + O(q^{d/n}).
\end{aligned}$$

We bound the middle term by counting that there are at most q^i monic polynomials of degree i, and summing the geometric series to obtain that the above is

$$= q^{d-l} \sum_{\substack{f_2 \in V \\ f_2(x_j) \ne 0 \text{ for } 1 \le j \le l}} \mu(f_2)(q-1)q^{-n\deg(f_2)} + O(q^{d/n+1}).$$

We have

$$\sum_{\substack{f_2\in V\\ f_2(x_j)\neq 0 \text{ for } 1\le j\le l}} \mu(f_2)(q-1)q^{-n\deg(f_2)} = (q-1)\prod_{\substack{P\in\mathbb{F}_q[x]\\ \text{P is monic irreducible}\\ P(x_j)\neq 0 \text{ for } 1\le j\le l}} (1-q^{-n\deg(P)})$$
$$= (q-1)\zeta(n)^{-1}(1-q^{-n})^{-l}.$$

Thus

$$\#\{f\in\mathcal{F}_d \| \ f(x_1)=a_1,\cdots,f(x_l)=a_l\} = \frac{q^{d-l}(q-1)}{\zeta(n)(1-q^{-n})^l} + O(q^{d/n+1}).$$

□

Now we refine our interpolation count to $\mathcal{F}_d^{\mathbf{s}}$. For $g\in\mathbb{F}_q[x]$, we write $g|_{x=x_i}$ for the evaluation of g at $x=x_i$.

Lemma 3.3. *Let $a_1,\dots,a_q\in\mathbb{F}_q^*$ and $0\le s_1,\dots,s_q\le n-1$. Then*

$$\#\{f\in\mathcal{F}_d^{\mathbf{s}} \mid \left.\frac{f}{(x-x_i)^{s_i}}\right|_{x=x_i} = a_i \text{ for all } i\} = \frac{q^{d-\sum s_i-q}(q-1)}{\zeta(n)(1-q^{-n})^q} + O(q^{(d-\sum s_i)/n+1}),$$

where the constant in the O is absolute.

Proof. For any $f\in\mathcal{F}^{\mathbf{s}}$, we write $f = g\prod_{1\le i\le q}(x-x_i)^{s_i}$. Then,

$$\left.\frac{f}{(x-x_i)^{s_i}}\right|_{x=x_i} = a_i \text{ if and only if } g(x_i) = \frac{a_i}{\prod_{j\neq i}(x_i-x_j)^{s_j}}.$$

Furthermore, if $h\in\mathbb{F}_q[x]$ is nth power free with no linear factors, then $h\prod_{1\le i\le q}(x-x_i)^{s_i}\in\mathcal{F}^{\mathbf{s}}$ since the $s_i\le n-1$.

So, by writing $f = g\prod_{1\le i\le q}(x-x_i)^{s_i}$, we have

$$\#\{f\in\mathcal{F}_d^{\mathbf{s}} \mid \left.\frac{f}{(x-x_i)^{s_i}}\right|_{x=x_i} = a_i \text{ for all } i\}$$
$$= \#\{g\in\mathcal{F}_{d-\sum s_i} \mid g(x_i) = \frac{a_i}{\prod_{j\neq i}(x_i-x_j)^{s_j}} \text{ for all } i\}$$

and then we apply Lemma 3.2. □

Lemma 3.4. *We have that $\#\mathcal{F}_d = (q-1)(q^d-q^{d-n+1}) = q^d(q-1)/\zeta(n)$ for $d\ge n$.*

Proof. By counting all monic polynomials of degree d, we have

$$\frac{1}{1-qt} = \sum_{d\ge 0}(qt)^d = \prod_{\substack{P\in\mathbb{F}_q[x]\\ \text{P is monic irreducible}}} \frac{1}{1-t^{\deg(P)}}.$$

By counting nth power-free monic polynomials, we then have

$$\sum_d \frac{\#\mathcal{F}_d}{q-1}t^d = \prod_{\substack{P\in\mathbb{F}_q[x]\\ \text{P is monic irreducible}}} (1+t^{\deg(P)}+\cdots+t^{(n-1)\deg(P)})$$
$$= \prod_{\substack{P\in\mathbb{F}_q[x]\\ \text{P is monic irreducible}}} \frac{1-t^{n\deg(P)}}{1-t^{\deg(P)}} = \frac{1-qt^n}{1-qt}.$$

We conclude that $\#\mathcal{F}_d = (q-1)(q^d-q^{d-n+1}) = q^d(q-1)/\zeta(n)$ for $d\ge n$. □

See also [VW12, Proposition 5.9(a)], which gives a much more general proof of this kind of identity (with the line replaced by any variety), and without using the Euler product.

3.3. **Proof of Theorem 3.1.** We are now ready to prove our main theorem, which we do by reducing it to the interpolation counts of Lemma 3.3.

Proof of Theorem 3.1. We will first see that for $f \in \mathcal{F}_d$ with $(x-x_i)^s || f$ and $\frac{f}{(x-x_i)^s}\Big|_{x=x_i} = a$, the number of points on C_f with $x = x_i$ only depends on s and a. If $s = 0$, then if a is one of the $\frac{q-1}{(m,q-1)}$ mth powers in $\mathbb{F}_q^*$, then there are $(m, q-1)$ points on C_f with $x = x_i$ (the number of mth roots of an mth power in $\mathbb{F}_q^*$, i.e. choices for y given $x = x_i$). If $s = 0$ and a is one of the $(q-1)(1-\frac{1}{(m,q-1)})$ non-mth powers in $\mathbb{F}_q^*$, then there are 0 points on C_f with $x = x_i$. If $s \geq 1$, then all $q-1$ values of a give 1 point with $x = x_i$ (as necessarily $y = 0$).

Let $\phi(s,k)$ be the set of $a \in \mathbb{F}_q^*$ such that for any $f \in \mathcal{F}$ with $(x-x_i)^s || f$, we have that $\frac{f}{(x-x_i)^s}\Big|_{x=x_i} = a$ implies that C_f has k points with $x = x_i$. From the above we see that

$$\#\phi(s,k) = \begin{cases} (q-1)(1-\frac{1}{(m,q-1)}) & \text{if } s = k = 0, \\ \frac{q-1}{(m,q-1)} & \text{if } s = 0 \text{ and } k = (m, q-1), \\ q-1 & \text{if } s \geq 1 \text{ and } k = 1, \text{ and} \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Applying Lemma 3.3, we have

$$\begin{aligned}
&\#\{f \in \mathcal{F}_d |\ C_f \text{ has } k_i \text{ points with } x = x_i \text{ for all } i\} \\
&= \sum_{0 \leq s_1, \ldots, s_q \leq n-1} \#\{f \in \mathcal{F}_d^{\mathbf{s}} |\ C_f \text{ has } k_i \text{ points with } x = x_i \text{ for all } i\} \\
&= \sum_{0 \leq s_1, \ldots, s_q \leq n-1} \sum_{(a_1, \ldots, a_q) \in \prod_i \phi(s_i, k_i)} \#\{f \in \mathcal{F}_d^{\mathbf{s}} |\ \frac{f}{(x-x_i)^{s_i}}\Big|_{x=x_i} = a_i \text{ for all } i\} \\
&= \sum_{0 \leq s_1, \ldots, s_q \leq n-1} \sum_{(a_1, \ldots, a_q) \in \prod_i \phi(s_i, k_i)} \left(\frac{q^{d - \sum s_i - q}(q-1)}{\zeta(n)(1-q^{-n})^q} + O(q^{(d-\sum s_i)/n+1}) \right) \\
&= \sum_{0 \leq s_1, \ldots, s_q \leq n-1} \left(\frac{q^{d - \sum s_i - q}(q-1)}{\zeta(n)(1-q^{-n})^q} + O(q^{(d-\sum s_i)/n+1}) \right) \sum_{(a_1, \ldots, a_q) \in \prod_i \phi(s_i, k_i)} 1 \\
&= \sum_{0 \leq s_1, \ldots, s_q \leq n-1} \left(\frac{q^{d - \sum s_i - q}(q-1)}{\zeta(n)(1-q^{-n})^q} + O(q^{(d-\sum s_i)/n+1}) \right) \left(\prod_{i=1}^{q} \#\phi(s_i, k_i) \right).
\end{aligned}$$

Thus we have, by Lemma 3.4, for $d \geq n$,

$$\begin{aligned}
&\frac{\#\{f \in \mathcal{F}_d |\ C_f \text{ has } k_i \text{ points with } x = x_i \text{ for all } i\}}{\#\mathcal{F}_d} \\
&= \sum_{0 \leq s_1, \ldots, s_q \leq n-1} \left(\frac{q^{-\sum s_i - q}}{(1-q^{-n})q} + O(q^{-d+1+(d-\sum s_i)/n+1}) \right) \left(\prod_{i=1}^{q} \#\phi(s_i, k_i) \right) \\
&= \prod_{i=1}^{q} \left(\sum_{s_i=0}^{n-1} \frac{\#\phi(s_i, k_i)}{q^{s_i+1}(1-q^{-n})} \right) + O(q^{-d+1+(d-\sum s_i)/n+1+q+n}).
\end{aligned}$$

So,

(2)
$$\lim_{d \to \infty} \frac{\#\{f \in \mathcal{F}_d |\ C_f \text{ has } k_i \text{ points with } x = x_i \text{ for all } i\}}{\#\mathcal{F}_d} = \prod_{i=1}^{q} \left(\sum_{s_i=0}^{n-1} \frac{\#\phi(s_i, k_i)}{q^{s_i+1}(1-q^{-n})} \right).$$

From equation (1) we have that

$$\sum_{s_i=0}^{n-1} \frac{\#\phi(s_i, k_i)}{q^{s_i+1}(1-q^{-n})} = \begin{cases} \left(1 - \frac{1}{(m, q-1)}\right) \frac{1-q^{-1}}{1-q^{-n}} & \text{if } k_i = 0, \\ \frac{q^{-1}-q^{-n}}{1-q^{-n}} & \text{if } k_i = 1 \text{ and } k_i \neq (m, q-1), \text{ and} \\ \frac{1}{(m,q-1)} \frac{1-q^{-1}}{1-q^{-n}} & \text{if } k_i = (m, q-1) \text{ and } k_i \neq 1, \text{ and} \\ \frac{q^{-1}-q^{-n}}{1-q^{-n}} + \frac{1}{(m,q-1)} \frac{1-q^{-1}}{1-q^{-n}} & \text{if } k_i = 1 = (m, q-1), \end{cases}$$

which, plugged into equation (2), proves the first statement of Theorem 3.1.

If a degree d polynomial f in $\mathbb{F}_q[x]$ is an rth power in $\bar{\mathbb{F}_q}[x]$ for some $r > 1$, then $f = cg^r$ for some $c \in \mathbb{F}_q$ and monic $g \in \mathbb{F}_q[x]$, and so there are at most $q^{d/r+1}$ of these polynomials f of degree d. So, in $\mathcal{F}_d$, we have at most

$$\sum_{\substack{i|m \\ i \geq 2}} q^{d/i+1} = O(mq^{d/2})$$

polynomials whose power over $\bar{\mathbb{F}_q}$ is not relatively prime to m. Using Lemma 3.4, we see that in the limit as $d \to \infty$, these polynomials will not contribute to the total proportion. By Lemma 2.1, this gives that the geometrically reducible C_f do not contribute in the $d \to \infty$ limit, giving the last statement of the theorem.

Let $\mathcal{F}_d^* = \{f \in \mathcal{F}_d \mid \text{the power over } \bar{\mathbb{F}_q} \text{ is relatively prime to } m\}$. Lemma 2.2 shows that for $f \in \mathcal{F}_d^*$ such that $(x - x_i)^s || f$ and $\frac{f}{(x-x_i)^s}\Big|_{x=x_i} = a$, that the number of points on $\tilde{C_f}$ with $x = x_i$ only depends on s and a. More explicitly, it shows that $\frac{q-1}{(m,s,q-1)}$ of the $q-1$ possible values of a_i give $(m, s, q-1)$ points with $x = x_i$ and the other $(q-1)(1 - \frac{1}{(m,s,q-1)})$ values of a_i give 0 points with $x = x_i$. Let $\tilde{\phi}(s, k)$ be the set of $a \in \mathbb{F}_q^*$ such that for any $f \in \mathcal{F}_d^*$ with $(x - x_i)^s || f$, we have that $\frac{f}{(x-x_i)^s}\Big|_{x=x_i} = a$ implies that $\tilde{C_f}$ has k points with $x = x_i$. From the above we see that

(3)
$$\#\tilde{\phi}(s, k) = \begin{cases} (q-1)(1 - \frac{1}{(m,s,q-1)}) & \text{if } k = 0 \text{ and} \\ \frac{q-1}{(m,s,q-1)} & \text{if } k = (m, s, q-1). \end{cases}$$

As above, we have

$$\lim_{d\to\infty} \frac{\#\{f \in \mathcal{F}_d |\ \tilde{C}_f \text{ has } k_i \text{ points with } x = x_i \text{ for all } i\}}{\#\mathcal{F}_d} = \prod_{i=1}^{q} \left(\sum_{s_i=0}^{n-1} \frac{\#\tilde{\phi}(s_i, k_i)}{q^{s_i+1}(1-q^{-n})} \right).$$

From equation (3) we have that

$$\sum_{s_i=0}^{n-1} \frac{\#\tilde{\phi}(s_i, k_i)}{q^{s_i+1}(1-q^{-n})} = \begin{cases} \displaystyle\sum_{0\le s\le n-1} q^{-s} \left(1 - \frac{1}{(m, s, q-1)}\right) \frac{1-q^{-1}}{1-q^{-n}} & \text{if } k_i = 0, \text{ and} \\ \displaystyle\sum_{\substack{0\le s\le n-1 \\ (m,s,q-1)=N}} \frac{q^{-s}(1-q^{-1})}{N(1-q^{-n})} & \text{if } k_i = N, \end{cases}$$

which gives the second statement of the theorem. □

Acknowledgements

The authors would like to thank Alina Bucur for useful conversations and the referee for many comments that improved the exposition of the paper. Theorem 1.1 was given as a conjecture by E. Holzhausen and M. Willer as a result of work supported by NSF grant DMS-0838210. M. M. Wood was supported in this work by an American Institute of Mathematics Five-Year Fellowship and NSF grant DMS-1147782. G. Cheong and A. Zaman were supported by NSF grants DMS-1147782 and DMS-0838210.

References

[BDFL10a] Alina Bucur, Chantal David, Brooke Feigon, and Matilde Lalín, *Fluctuations in the number of points on smooth plane curves over finite fields*, J. Number Theory **130** (2010), no. 11, 2528–2541, DOI 10.1016/j.jnt.2010.05.009. MR2678860 (2011f:11076)

[BDFL10b] Alina Bucur, Chantal David, Brooke Feigon, and Matilde Lalín, *Statistics for traces of cyclic trigonal curves over finite fields*, Int. Math. Res. Not. IMRN **5** (2010), 932–967, DOI 10.1093/imrn/rnp162. MR2595014 (2011c:11100)

[BDFL11] Alina Bucur, Chantal David, Brooke Feigon, and Matilde Lalín, *Biased statistics for traces of cyclic p-fold covers over finite fields*, WIN—women in numbers, Fields Inst. Commun., vol. 60, Amer. Math. Soc., Providence, RI, 2011, pp. 121–143. MR2777802 (2012j:11130)

[BK11] Alina Bucur and Kiran S. Kedlaya, *The probability that a complete intersection is smooth* (English, with English and French summaries), J. Théor. Nombres Bordeaux **24** (2012), no. 3, 541–556. MR3010628

[EW12] Daniel Erman and Melanie Matchett Wood, *Semiample Bertini theorems over finite fields*, to appear in Duke Math. J. arXiv:1209.5266.

[KR09] Pär Kurlberg and Zeév 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, DOI 10.1016/j.jnt.2008.09.004. MR2488590 (2009m:14029)

[KS99] Nicholas M. Katz and Peter Sarnak. *Random matrices, Frobenius eigenvalues, and monodromy*, volume 45 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1999.

[VW12] Ravi Vakil and Melanie Matchett Wood, *Discriminants in the Grothendieck ring*, to appear in Duke Math. J. arXiv:1208.3166.

[Woo12] Melanie Matchett Wood, *The distribution of the number of points on trigonal curves over* $\mathbb{F}_q$, Int. Math. Res. Not. IMRN **23** (2012), 5444–5456. MR2999148

[Xio10] Maosheng Xiong, *The fluctuations in the number of points on a family of curves over a finite field* (English, with English and French summaries), J. Théor. Nombres Bordeaux **22** (2010), no. 3, 755–769. MR2769344 (2012a:11085)

Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Drive, Madison, Wisconsin 53705

Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Drive, Madison, Wisconsin 53705

E-mail address: mmwood@math.wisc.edu

American Institute of Mathematics, 360 Portage Ave, Palo Alto, California 94306-2244