quatrième série - tome 49 fascicule 5 septembre-octobre 2016

a*NNALES SCIEN*n*IFIQUES SUPÉRIEU*k*^E de L ÉCOLE* $NORMALE$

Byungchul CHA & Daniel FIORILLI & Florent JOUVE *Prime number races for elliptic curves over function fields*

SOCIÉTÉ MATHÉMATIQUE DE FRANCE

Annales Scientifiques de l'École Normale Supérieure

Publiées avec le concours du Centre National de la Recherche Scientifique

Responsable du comité de rédaction / *Editor-in-chief*

Antoine CHAMBERT-LOIR

Rédaction / *Editor*

Annales Scientifiques de l'École Normale Supérieure, 45, rue d'Ulm, 75230 Paris Cedex 05, France. Tél. : (33) 1 44 32 20 88. Fax : (33) 1 44 32 20 80. annales@ens.fr

Édition / *Publication* **Abonnements /** *Subscriptions*

Société Mathématique de France Maison de la SMF Institut Henri Poincaré Case 916 - Luminy 11, rue Pierre et Marie Curie 13288 Marseille Cedex 09 Fax : (33) 01 40 46 90 96

75231 Paris Cedex 05 Fax : (33) 04 91 41 17 51 Tél. : (33) 01 44 27 67 99 email : smf@smf.univ-mrs.fr

Tarifs

Europe : 519 \in . Hors Europe : 548 \in . Vente au numéro : 77 \in .

© 2016 Société Mathématique de France, Paris

En application de la loi du 1er juillet 1992, il est interdit de reproduire, même partiellement, la présente publication sans l'autorisation de l'éditeur ou du Centre français d'exploitation du droit de copie (20, rue des Grands-Augustins, 75006 Paris). *All rights reserved. No part of this publication may be translated, reproduced, stored in a retrieval system or transmitted in any form or by any other means, electronic, mechanical, photocopying, recording or otherwise, without prior permission of the publisher.*

PRIME NUMBER RACES FOR ELLIPTIC CURVES OVER FUNCTION FIELDS

BY BYUNGCHUL CHA, DANIEL FIORILLI AND FLORENT JOUVE

ABSTRACT. – We study the prime number race for elliptic curves over the function field of a proper, smooth and geometrically connected curve over a finite field. This constitutes a function field analogue of prior work by Mazur, Sarnak and the second author. In this geometric setting, we can prove unconditional results whose counterparts in the number field case are conditional on a Riemann Hypothesis and a linear independence hypothesis on the zeros of the implied L-functions. Notably we show that in certain natural families of elliptic curves, the bias generically dissipates as the conductor grows. This is achieved by proving a central limit theorem and combining it with generic linear independence results that will appear in a separate paper. Also we study in detail a particular family of elliptic curves that have been considered by Ulmer. In contrast to the generic case we show that the race exhibits very diverse outcomes, some of which are believed to be impossible in the number field setting. Such behaviors are possible in the function field case because the zeros of Hasse-Weil L-functions for those elliptic curves can be proven to be highly dependent among themselves, which is a very non generic situation.

R. – Nous étudions une version du biais de Chebyshev pour les courbes elliptiques sur le corps de fonctions d'une courbe lisse, propre, et géométriquement irréductible. Il s'agit de l'analogue, dans le cas des corps de fonctions, de travaux de Mazur, Sarnak, et Fiorilli. Le cadre géométrique dans lequel on se place permet d'établir inconditionnellement des résultats qui, sur les corps de nombres, nécessitent de supposer l'hypothèse de Riemann ou la conjecture de simplicité généralisée pour les zéros des fonctions L intervenant. On démontre notamment que, dans certaines familles naturelles de courbes elliptiques, il y a dissipation générique du biais lorsque le conducteur tend vers l'infini. La preuve s'appuie sur des résultats d'indépendance linéaire de zéros de fonctions L qui font l'objet d'un autre article des mêmes auteurs. Nous étudions par ailleurs la famille de courbes elliptiques d'Ulmer, et nous montrons qu'elle se comporte de manière pathologique, comparée au cas générique. Divers biais sont mis en évidence pour cette famille, dont certains sont conjecturalement impossibles à réaliser dans le cas des corps de nombres.

1240 B. CHA, D. FIORILLI AND F. JOUVE

1. Introduction and statement of the main results

1.1. Background

It was first noticed by Chebyshev that primes are biased in their distribution modulo 4, in that there seems to be more primes of the form $4n + 3$ than of the form $4n + 1$ in initial intervals of the integers. A number of papers have been written on this phenomenon and its generalizations, and it is now known that such a bias appears in many number theoretical contexts, such as primes in [arit](#page-39-0)hmetic progressions, Frobenius elements in conjugacy classes of the Galois group of extensions of number fields, Fourier coefficients of modular forms, prime polynomials in residue classes over $\mathbb{F}_q(t)$, and so on.

In their seminal paper [18], Rubinstein and Sarnak have given a framework to study questions of this type. One of the features of their work is the quantification of the so-called *Chebyshev bias*in te[rm](#page-3-0)s of an associated measure which is expressed using an explicit formula as a function of the nontrivial zeros of the involved L-functions.

In the case of Chebyshev's original question, Rubinstein and Sarnak determined that the logarithmic density⁽¹⁾ of the set of x for which $\pi(x; 4, 3) > \pi(x; 4, 1)$ exists and is given by $\delta(4; 3, 1) \approx 0.9959$. (Here, $\pi(x; q, a)$ is the count of primes $\leq x$ that are congruent to a modulo q .) Their results are conditional on the Generalized Riemann Hypothesis (GRH), and on the assumption (the Linear Independence hypothesis, or LI in short) that the multiset of (the ordinates of) all nontrivial zeros of the involved L-functions is linearly independent over $\mathbb Q$. One might think that the modulus 4 is not exceptional here, and that there should exist other moduli q and residue classes a and b modulo q such that $\delta(q; a, b)$, the logarithmic density of the set of $x \ge 1$ for which $\pi(x; q, a) > \pi(x; q, b)$ $\pi(x; q, a) > \pi(x; q, b)$ $\pi(x; q, a) > \pi(x; q, b)$, is also very close to 1. It turns out that as Rubinstein and Sarnak have shown, $\delta(q; a, b)$ approaches $\frac{1}{2}$ as $q \to \infty$, hence races of large [modu](#page-39-2)li are very moderately biased. One can also quantify the rate of convergence here, showing for example as in [11] that whenever $\delta(q; a, b) \neq \frac{1}{2}$, we have $|\delta(q;a,b)-1/2|=q^{-\frac{1}{2}+o(1)}.$

In the recent paper [10], the second author considered the more general race between two subsets A and B of the invertible reduced residues modulo q. It turns out that when studying the inequality $\pi(x; q, A) > \pi(x; q, B)$ with

$$
\pi(x;q,A):=\sum_{a\in A}\pi(x;q,a),
$$

things can become dramatically different from the previous case where only two residue classes were involved. Indeed, one can show under GRH and a multiplicity assumption on the zeros of $L(s, \chi)$ that there exist sequences of moduli $\{q_k\}$ and subsets $\{A_k\}$ and $\{B_k\}$ such that the associated lower and upper densities $\overline{\delta}(q_k; A_k, B_k)$ get arbitrarily close to 1. (Note also that it is known $\overline{\delta}(q; A, B) < 1$ for any q, A, B.) In other words, there exist 'highly biased prime number races'. Under the additional assumption that LI holds, it is also proven that in order to obtain highly biased prime number races, the moduli q_k need to have many prime factors, and hence highly biased prime number races are very rare. Most races are very moderately biased, in the sense that $\delta(q; A, B)$ is usually very close to $\frac{1}{2}$.

(1) The logarithmic density of a set $S \subset \mathbb{N}$ is defined by $\delta(S) := \lim_{N \to \infty} \frac{1}{\log N} \sum_{n \in S} \sum_{n \in S}$ $\frac{1}{n}$, if this limit exists.

In the context of elliptic curves, Mazur [16] introduced the race between the primes for which $a_p(E)$, the trace of the Frobenius at a prime p, is positive, against those for which $a_p(E)$ is negative. Sarnak's framework ⁽²⁾ in [19] to study this question turned out to be very effective, and explained this race very well in terms of the zeros (and potential poles) of $L(SymⁿE, s)$, the symmetric power L-functions attached to E, conditional on a Riemann Hypothesis and LI. Sarnak also remarked that one can stud[y a](#page-39-3) related race by focusing on the sign of the summatory function of $a_p(E)/\sqrt{p}$ using the zeros of $L(E, s)$ alone. For this race, Sarnak uncovered the influence of the analytic rank of E on the bias.

Building on Sarnak's work, the second author studied in [9] the following question: is it possible to find highly biased prime number races in the context of elliptic curves over Q, or are all races of this type only moderately biased? It turns out that conditionally on a Riemann Hypothesis and the assumption that the multiplicity of nonreal zeros of $L(E, s)$ $L(E, s)$ $L(E, s)$ i[s un](#page-39-5)iformly bounded (which is referred to as a bounded multiplicity assumption), the key to finding such races is to find curves E whose analytic rank is significantly larger than $\sqrt{\log N_E}$, where N_E is the conductor of E. Interestingly, the two existing conjectures (stated in $[8]$ and $[20]$) on the growth of the rank of elliptic curves over $\mathbb Q$ bo[th i](#page-39-6)mply the existence of the aforementioned curves. Note also that elliptic curves of large rank are extremely rare. It is widely believed that 100% of the elliptic curves over Q have rank either 0 or 1, depending on the root number of $L(E, s)$. One c[an](#page-39-7) show, as is explained in [19], that the bias for such curves dissipates as $N_E \rightarrow \infty$. Hence, highly biased elliptic curve prime number races over Q are very rare.

Coming back to the original Chebyshev bias, but in the function field setting, the first author showed in [4] that the framework of Rubinstein and Sarnak can be replicated in the one-variable polynomial ring $\mathbb{F}_q[t]$ over a finite field \mathbb{F}_q of q elements. One of the advantages in this setting is that much more is known about the (inverse[\) z](#page-39-7)eros of L-functions. [Firs](#page-0-0)t of all, the Riemann Hypothesis is known to be true in this context. Also, one can explicitly calculate the zeros of relevant L-functions in some specific cases and prove the function field version of LI (see the definition of Grand [Sim](#page-39-7)plicity Hypothesis in [4] and also Definition 4.1 below).

The goal of the current paper is to present an unconditional analysis of the Chebyshev bias for elliptic curves E over a function field K (in particular it can be seen as a partial 2-dimensional generalizati[on](#page-11-0) of [4]). More precisely we study the sign of the [sum](#page-39-6)matory function of the (normalized) trace $a_v(E)/q^{\deg v/2}$ of the Frobenius at v, where v runs over the places of the function field of a smooth proper geometrically connected curve over a finite field \mathbb{F}_q . In Section 2 we present the analogue of the work of Sarnak [19] in our geometric setting; as in *loc. cit.*, our analysis is mor[e g](#page-23-0)eneral than what is needed to study bias phenomena in the distribution of the traces of Frobenius. Indeed it involves an arbitrary smooth function of the angles θ_v of the local Frobenius traces, and as such zeros of higher symmetric power L-functions come into play. Section 3 is devoted to the study of Chebyshev's bias for Ulmer's family of elliptic curves E_d over the rational function field $\mathbb{F}_q(t)$ given by the Weierstrass equation

$$
E_d: y^2 + xy = x^3 - t^d.
$$

(2) Granville independently worked out the link between these types of prime number races and the distribution of the zeros of Hasse-Weil L-functions, and explained it to the second author.

For this particular family, the Hasse–Weil L-function is complet[ely](#page-39-8) explicit and can be described in an elementary fashion (involving, e.g., multiplicative orders modulo the divisors of the p[aram](#page-0-0)eter d). Finally in Section 4 we prove a central limit theorem which, in conjunction with the generic linear independence results proven in [5], allows us to deduce that most elliptic curve prime number races over function fields are very moderately biased (see Theorem 1.3).

Notations

Throughout the paper p denotes a prime number that does not divide 6 and ℓ is a prime number different from p. We fix a finite field \mathbb{F}_q , where q is a power of p, and a proper, smooth and geometrically connected curve C/\mathbb{F}_q , with function field $K = \mathbb{F}_q(C)$. At each place v of K we have the residue field k_v which is the unique extension of \mathbb{F}_q (in a fixed algebraic closure) of degree $\deg(v)$. We fix a separable closure K_s of K and we let $G_K := \text{Gal}(K_s/K)$ be the absolute Galois group of K. Finally, E/K is an elliptic curve with nonconstant j-invariant and its analytic rank will be denoted by $rank(E/K)$.

1.2. Main results

For a closed point v of C at which E/K has good reduction, we let a_v be the trace of Frobenius at v. We study the average behavior of $a_v/q^{\deg(v)/2} = 2\cos\theta_v$. Note that the results of Section 2 apply to the more general context of any smooth function of θ_v , but we will focus on $2 \cos \theta_v$ for now. We are interested in the limiting distribution arising from

(1)
$$
T_E(X) := -\frac{X}{q^{X/2}} \sum_{\substack{\deg(v) \le X \\ v \text{ good}}} 2 \cos \theta_v = -\frac{X}{q^{X/2}} \sum_{\substack{\deg(v) \le X \\ v \text{ good}}} \frac{a_v}{q^{\deg(v)/2}}.
$$

The quantity $T_E(X)$ oscillates, and usually takes both positive and negative values. To measure how long $T_E(X)$ stays positive or negative, that is, to measure its bias, we define

$$
\overline{\underline{\delta}}(E):=\overline{\lim}_{M\to\infty}\frac{1}{M}\sum_{\substack{X\leq M\\ T_E(X)>0}}1.
$$

If $\delta(E)$ and $\overline{\delta}(E)$ coincide we denote by $\delta(E)$ the common limit.

Since E/K has nonconstant *j*-invariant, we write its Hasse-Weil *L*-function as

$$
L(E/K, T) = \prod_{j=1}^{N_{E/K}} (1 - q e^{i\theta_j} T).
$$

Here, $N_{E/K}$ is given by the formula

$$
N_{E/K} = 4(g_C - 1) + \deg(\mathfrak{n}_{E/K}),
$$

where g_C is the genus of C and $\mathfrak{n}_{E/K}$ is the conductor of E/K (see, e.g., [22, Th. 9.3]).

Our first theorem, which will be obtained by combining Corollary 2.9, Corollary 2.10, and Theorem 4.5, provides a general description of the prime number race for elliptic curves in function fields.

DEFINITION 1.1. – We say that a function $S : \mathbb{Z}_{\geq 1} \to \mathbb{R}$ has a *limiting distribution* if there exists a Borel measure μ on $\mathbb R$ such that for any bounded Lipschitz continuous function $f : \mathbb{R} \longrightarrow \mathbb{R}$ we have

(2)
$$
\lim_{M \to \infty} \frac{1}{M} \sum_{X=1}^{M} f(S(X)) = \int_{\mathbb{R}} f(t) d\mu(t).
$$

In case S is chosen to be the function T_E of (1) we prove the following result.

THEOREM 1.2. – We keep the notation as above.

(i) *The function* $T_E(X)$ *has a limiting distribution. Denoting by* X_E *the associated random variable, its mean and variance are given by*

$$
\mathbb{E}[X_E] = \frac{\sqrt{q}}{\sqrt{q}-1} \left(\text{rank}(E/K) - \frac{1}{2} \right),
$$

and

$$
\mathbb{V}[X_E] = \frac{1}{4} \left(\frac{\sqrt{q}}{\sqrt{q} + 1} \right)^2 + \sum_{\theta_j \neq 0}^* \frac{m(\theta_j)^2}{|1 - q^{-1/2} e^{-i\theta_j}|^2}.
$$

 $Here, m(\theta_j)$ [is](#page-0-0) the multiplicity of θ_j , and the starred-summation (\sum^*) means that it runs *over all* $\theta_i \neq 0$ *counted without multiplicity.*

(ii) *Let* {E/K} *be a family of elliptic curves of unbounded conductor satisfying LI (see Definition* 4.1) and such that $\text{rank}(E/K) = o(\sqrt{N_{E/K}})$ as $N_{E/K} \to \infty$. Then, the *random variable*

$$
\sqrt{\frac{q-1}{q}}X_E/\sqrt{N_{E/K}}
$$

converges in distribution to the standard Gaussian as $N_{E/K} \rightarrow \infty$ *, and as a consequence we have that* $\delta(E) \rightarrow \frac{1}{2}$ *.*

This theorem says that the prime number race for elliptic curves can be generally described by its rank and the multiplicities of the zeros of $L(E/K, T)$. Indeed, up to a constant the variance is give[n by](#page-39-6) the [squ](#page-39-3)are of the 2-norm of the vector of multiplicities of the $\theta_i \neq 0$. Further, if we assume LI, then the bias in the race dissipates as $N_{E/K}$ gets large, unless the rank grows faster than $\sqrt{N_{E/K}}$. These results are in line with corresponding number field counterparts in [19] and [9]. However, our results are much more unconditional than in the number field setting. This is mainly because the ne[ce](#page-39-8)ssary analytic properties of $L(E/K, T)$ are established in the function field setting. We also note that the confirmation of LI is completely conjectural in the number field case. However over function fields, it is possible to prove LI in some cases. In fact, in a separate paper [5], we prove LI among certain families of elliptic curves generically by establishing quantitive bounds for the number of elliptic curves in the families satisfying LI. Let us briefly recall the construction of one of the main families studied in *loc. cit.*

[Fi](#page-39-9)x an elliptic curve E/K . The family we consider is a family of *quadratic twists* of E/K . For ease of exposition let us recall the necessary definitions only in the case where $C = \mathbb{P}^1$, in which case K is simply the rational function field $\mathbb{F}_q(t)$. The setting we use is the same as in [2, §2.1, 2.3] and we only recall the main definition and notation, referring the reader to *loc. cit.* for a more detailed exposition.

Suppose that E/K is given by the Weierstrass equation $y^2 = x^3 + ax + b$, where we can assume (up to a change of coordinates) that $a, b \in \mathbb{F}_q[t]$ and that the discriminant $\Delta = 4a^3 + 27b^2$ has minimal degree as a polynomial in t. For each $f \in K^\times$ we consider

 E_f : $y^2 = x^3 + f^2 ax + f^3 b$

which is a Weierstrass equation for an elli[pt](#page-39-9)ic curve over K . A quadratic twist of E/K is an elliptic curve E_f/K such that f is not a square in K. Note that E_f is isomorphic to E_g over K if and only if there exists $c \in K^{\times}$ such that $f = gc^2$.

A suitable change of variable (see, e.g., [2, §2.1]) yields a model of E over $\mathbb{F}_q[1/t]$ such that gluing the two models over the annulus $\mathbb{P}^1 \setminus \{0, \infty\}$ we obtain the identity [com](#page-39-10)ponent of the the Néron model of E w[hic](#page-39-8)h we denote $\mathcal{E} \to \mathbb{P}^1$.

Let us assume that $\mathcal{E} \to \mathbb{P}^1$ has at least one fiber of multiplicative reduction away from ∞ . (This assumption is crucial to ensure big monodromy statements such as [13, Th. 6.3] which in turn play a key role in $[5]$ to study linear independence properties of zeros of L -functions of E_f/K , where f runs over the parameter family we now define.)

The "twisting family" we consider was first introduced by Ka[tz.](#page-7-0) It is the $(d + 1)$ -dimension[al a](#page-39-9)ffine variety for which the F-rational points are:

(3)
$$
\mathcal{F}_d(\mathbb{F}) = \{ f \in \mathbb{F}[t] : f \text{ squarefree, } \deg f = d, \gcd(f, \Delta) = 1 \},
$$

for any algebraic extension $\mathbb{F} \supseteq \mathbb{F}_q$, and where $d \geq 8$ is an integer⁽³⁾. A remarkable fact (see, e.g., [2, §2.3] for a proof) is that if $f \in \mathcal{F}_d(\mathbb{F}_{q^n})$ then the conductor of E_f only depends on d and q. In particular we can let $n \to \infty$ without affecting the value of the common conductor of the twists E_f .

As was already mentioned, the above construction of quadratic twists can be done over any function field $K = \mathbb{F}_q(C)$ $K = \mathbb{F}_q(C)$ $K = \mathbb{F}_q(C)$, where C/\mathbb{F}_q is any smo[oth g](#page-0-0)eometrically connected proper curve. In particular (3) c[an b](#page-0-0)e defined in this more general context; it will then consist of elements of Θ , the ring of integers of the compositum $\mathbb{F}K$.

The conjunction of the results of [5] with Theorem 4.5, which is a stronger (quantitative) version of Theorem 1.2, gives the following result which holds for any function field $K = \mathbb{F}_q(C)$, with C as above.

T 1.3. – *With notation as above, there exists an absolute constant* c *such that the proportion of parameters* $f \in \mathcal{F}_d(\mathbb{F}_{q^n})$ *for which* $\delta(E_f)$ *exists and the inequality*

$$
\left|\delta(E_f) - \frac{1}{2}\right| \le \frac{c}{\sqrt{d}}
$$

 $holds$ is at least $1 - O_{d,E/\mathbb{F}_q(C)}(n \log q/q^{n c_E d^{-2}})$ $1 - O_{d,E/\mathbb{F}_q(C)}(n \log q/q^{n c_E d^{-2}})$ $1 - O_{d,E/\mathbb{F}_q(C)}(n \log q/q^{n c_E d^{-2}})$, where the positive constant c_E depends only *on the base curve* E*.*

Since the statement of Theorem 1.3 involves many different parameters, let us make several comments on the way we think one should interpret it. First one fixes the piece of data E/K (in particular the field [of](#page-37-0) constants \mathbb{F}_q of $K = \mathbb{F}_q(C)$ has fixed cardinality q). One should then pick d large so that the first i[neq](#page-0-0)uality means that $\delta(E_f)$ is very close to 1/2.

(3) The condition $d \geq 8$ ensures that $N(E_f/K) \geq 8$ by (63). This is needed to guarantee the existence of the density $\delta(E_f)$, as we will see in the proof of Theorem 1.3.

Then we use the remark preceding the statement of Theorem 1.3 to choose n large so that the proportion of curves excluded by the second inequality gets very close to 0. This way Theorem 1.3 can be seen as a result asserting that "generically" prime number races for elliptic curves over function fields are very moderately biased.

REMARK 1.4. – One may ask for the detailed description of a concrete example (i.e., an example where one starts with a concrete base elliptic curve E/K) where the unspecified constants appearing in Theorem 1.3 can be made more explicit. Let us consider the case where K is the rational function field $\mathbb{F}_q(t)$ and E/K is the Legendre elliptic curve given by:

$$
y^2 = x(x-1)(x-t).
$$

The curve E/K has multiplicative reduction at t and t – 1 and has discriminant $\Delta = 16t^2(t-1)^2$ so that elements $f \in \mathcal{F}_d$ are precisely monic squarefree polynomials of degree d that do not vanish at 0, 1.

Fix $d \geq 2$ and $\tilde{f} \in \mathcal{F}_{d-1}(\mathbb{F}_q)$. Let $n \geq 1$ be an integer. We restrict to twists of E by the \mathbb{F}_{q^n} -points of the open affine curve $U_{\tilde{f}}$ with geometric points:

$$
U_{\tilde{f}}(\overline{\mathbb{F}_q}) = \{c \in \overline{\mathbb{F}_q} \colon (c-t)\tilde{f}(t) \in \mathcal{F}_d(\overline{\mathbb{F}_q})\} = \{c \in \overline{\mathbb{F}_q} \setminus \{0,1\} \colon \tilde{f}(c) \neq 0\}.
$$

If $c \in U_{\tilde{f}}(\mathbb{F}_{q^n})$ $c \in U_{\tilde{f}}(\mathbb{F}_{q^n})$ $c \in U_{\tilde{f}}(\mathbb{F}_{q^n})$ we denote by E_c [the](#page-0-0) quadratic t[wis](#page-31-0)t of E by f where $f(t) = (c - t)\tilde{f}(t)$. For $f \in \mathcal{F}_d(\mathbb{F}_{q^n})$ the conductor of E_f/K is 2d (resp. 2d – 1) if d is even (resp. if d is odd). (See [5, Cor. 2.2] and the references therein.) Combining the arguments we develop in the proof of Theorem 1.3 (see Section 4) with [5, Cor. 2.2] we deduce that there exists an absolute constant c_0 such that $|\delta(E_c) - 1/2| \leq c_0 d^{-1/2}$ except for a proportion $\ll_{\tilde{f}} d^2 n \log q / q^{n/(24d^2)}$ (where the implied constant depends only on \tilde{f} and thus is independent of n) of exceptions in \mathbb{F}_{q^n} . Thus for big enough d and q and for the choice $n = d^2$ we get densities $\delta(E_c)$ [ver](#page-0-0)y close to 1/2 up to a proportion of exceptions $c \in \mathbb{F}_{q^n}$ very close to 0.

One can say that Theorem 1.3 presents an orderly picture regarding the prime number races for elliptic curves in general[—th](#page-39-5)eir bias dissipates as the conductor gets large. In contrast, our next finding shows that the races can exhibit very diverse outcomes when we look into the behaviors of $\delta(E)$ as E varies in a specific family of elliptic curves. We specialize to the family Ulmer considered in [20] and uncover many different and surprising prime number races[. In](#page-39-5)terestingly, many of the outcomes we discover are believed to be impossible in the number field case. The reason why Ulmer's family shows such diverse results is that LI is proven to be strongly violated in this family.

Following [20], we let E_d be the elliptic curve over $\mathbb{F}_q[t]$ given by the Weierstrass equation

(4)
$$
y^2 + xy = x^3 - t^d,
$$

where d and p are chosen so that $d \mid p^n + 1$ for some $n \geq 1$. We will be interested in the associated quantity defined by (1) which we will denote by $T_d(X)$. Our main tool for studying $T_d(X)$ will be the explicit formula given in Proposition 3.2.

Let us first state a result asserting that an extreme bias may occur $(T_d(X))$ either taking mostly positive or negative values) for suitable choices of parameters.

THEOREM 1.5. – *For the family* ${E_d/\mathbb{F}_q(t)}$ *(where we recall that the integer d and the characteristic* p of \mathbb{F}_q are linked by the relation $d \mid p^n + 1$ for some $n \geq 1$), one has the following *cases of extreme bias.*

(i) *Suppose that* $q \geq 3$ *, and assume that either* **–** d *is divisible by* 2 *and* q ≡ 1 mod 4*, or* **–** d *is divisible by* 3*.*

Then, $T_d(X) > 0$ *for all large enough* X, and thus $\underline{\delta}(E_d) = \overline{\delta}(E_d) = 1$.

- (ii) If $q = p^k$ with p large enough and $d = p^n + 1$ for some $1 \le n \le e^{q^{\frac{1}{2}}/2}$ with $n \equiv 0 \bmod k$, *then* $T_d(X) > 0$ *for all large enough* X, and thus $\underline{\delta}(E_d) = \overline{\delta}(E_d) = 1$.
- (iii) *Fix* $\epsilon > 0$ *. There exists primes* $d \geq 3$ *and* p *such that* p *is a primitive root modulo d,* and such that if we pick $q = p^{\frac{d-1}{2}+1}$, then the associated curve E_d has analytic rank 1 *(resp.* 2*) if* $(d-1)/2$ *is even (resp. odd) and*

$$
0 < \underline{\delta}(E_d) \leq \overline{\delta}(E_d) < \epsilon \, .
$$

REMARK 1.6. – One might wonder whether it is possible to have $\delta(E_d) = 0$. We will show in Corollary 3.4 that for $d \ge 7$ this is impossible, [since](#page-0-0) we always have $\delta(E_d) \ge 1/2n$, where *n* is the least positive integer such that $d | p^n + 1$.

The first ph[eno](#page-39-0)menon we uncover in Theorem 1.5 (i) is the existence of elliptic curves E_d for which $\delta(E_d) = 1$. This is quite surprising since one [can s](#page-0-0)how for an elliptic curve E over Q that under the Riemann Hypothesis for $L(E, s)$, we always have $\delta(E) < 1$. (This follows from the analysis in [18, Th. 1.2].) In the second point of the statement (i), the integer d is divisible by 3, which creates 'extra rank' for E_d (see Proposition 3.1). However (ii) shows the existence of infinitely many d not necessarily divisible by 3 for which $\delta(E_d) = 1$.

Part (iii) of the theorem highlights a remarkable feature of Ulmer's family. Indeed there are curves within the family $E_d/\mathbb{F}_q(t)$ of analytic rank ≥ 1 for which $\delta(E_d) < \frac{1}{2}$. This is surprising since in the case of elliptic curves over $\mathbb Q$, Sarnak showed ⁽⁴⁾ under a Riemann Hypothesis and a Linear Independence hypothesis that whenever the analytic rank of E is greater [than](#page-39-6) or equal to 1, we have $\delta(E) > \frac{1}{2}$. By (iii) we can find prime number races which are arbitrarily biase[d](#page-39-3) towards negative values. Interestingly, the involved curves E_d have rank at most 2, and a high bias is quite unexpected for such curves. Indeed as was remarked by Sarnak [19], one can show under GRH and LI that for elliptic curves of rank at most 2, the density $\delta(E)$ approaches $\frac{1}{2}$ as E runs over a family satisfying $N_E \to \infty$ (see [9, Proof of Th. 1.5] for a similar result with a weaker hypothesis).

Next we turn to subfa[milie](#page-39-6)s of $E_d/\mathbb{F}_q(t)$ with behavior very different to the above examples. Precisely the following statement shows the existence of curves for which there is no bias at all, in other words $\delta(E_d) = \frac{1}{2}$. Note that this is believed to be impossible for an elliptic curve E/\mathbb{Q} , since Sarnak [19] has shown under GRH and LI that $\delta(E) \neq \frac{1}{2}$.

⁽⁴⁾ When comparing our results with those of Sarnak, one should keep in mind that we are considering the race of opposite sign, that is we are considering the summatory function of $-a_v/q^{\text{deg }v/2}$, and Sarnak is considering the summatory function of a_p/\sqrt{p} .

THEOREM 1.7. – *For the family* ${E_d/\mathbb{F}_q(t)}$ *, one has the following cases where* $T_d(X)$ *is completely unbiased. Fix* $p \equiv 3 \mod 4$ *and let* $d \geq 5$ *be a divisor of* $p^2 + 1$ *. Pick* $q = p^{4k+1}$ *with* $k \geq 1$ *. Then the analytic rank of* E_d *is either* $(d-1)/4$ *or* $(d-2)/4$ *depending on whether* d *is congruent to* 1 *or* 2 *modulo* 4*, [and](#page-0-0) we have*

$$
\underline{\delta}(E_d) = \overline{\delta}(E_d) = \frac{1}{2}.
$$

Another reason why Theorem 1.7 is surprising is that for elliptic cur[ves o](#page-0-0)ver \mathbb{Q} , the key to producing highly biased races is to find elliptic curves E for w[hich](#page-10-0) the analytic rank is consid[erab](#page-0-0)ly larger than $\sqrt{\log N_E}$, where N_E is the conductor of E (see [9, Th. 1.2]). However, if we pick $d = p^2 + 1$, then many of the curves in Theorem 1.7 have very high rank[, qu](#page-39-5)ite close to the Brumer-Mestre bound when k is not too large (5) . Indeed in this case Th[eo](#page-39-3)rem 1.7 states that the rank of E_d equals $(p^2 - 1)/4$ $(p^2 - 1)/4$ $(p^2 - 1)/4$. This is considerably larger than $\sqrt{\deg(\mathfrak{n}_d)}$ (which is the analog of $\sqrt{\log N_E}$), as this last quantity is given by $\sqrt{p^2 + O(1)}$ (see [20, § 10.2]). In [su](#page-10-1)ch a situation one should expect $\delta(E)$ to be very close to 1, in light of [9, Theorem 1.2]. However Hypothesis BM of [9], which states that the multiplicities of the nonreal zeros of the L -functions associated to elliptic curves over $\mathbb Q$ are uniformly bounded, is strongly violated ⁽⁶⁾ for the elliptic curves E_d . This explains why no such extreme bias occurs.

Our final result shows that for any fixed $m \geq 1$, there are many curves for which $\delta(E_d)$ is very close to $(2m)^{-1}$. Those are races whose bias is moderate, but does not dissipate as the conductor grows. This result is motivated by the second part of Theorem 1.1 of [10] where the author shows under GRH and LI that, in the context of primes in arithmetic progressions, the set of all densities $\delta(q; NR, R)$ is dense in $[\frac{1}{2}, 1]$.

THEOREM 1.8. – *Define the set of all possible densities coming from Ulmer curves:*

$$
S := \{ \underline{\delta}(E_d) : d \mid p^n + 1, p \ge 3, n \ge 1; q = p^k, k \ge 1 \}.
$$

Then for every $m \geq 1$ *there exists elements of* S *that are arbitrarily close to* $1/(2m)$ *, that is:*

$$
\{0,1\} \cup \{1/(2m) : m \geq 1\} \subset S,
$$

where S *is the closure of* S*. The same holds if one replaces lower densities by upper densities.*

The corresponding statement for elliptic curves over $\mathbb Q$ is plainly false (under GRH and LI), as the only curves E/\mathbb{Q} having $\delta(E) < \frac{1}{2}$ are curves of rank 0, and for these curves $\delta(E)$ $\delta(E)$ $\delta(E)$ approaches $\frac{1}{2}$ as N_E tends to infinity. Moreover, it is unclear whether one should expect to have any limit points in $(\frac{1}{2}, 1)$, given our limited knowledge on ranks. Indeed, to obtain a limit point $\eta \in (\frac{1}{2}, 1)$ for the set of all $\delta(E)$ with E running over the curves over \mathbb{Q} , one would need (7) an infinite sequence of curves of analytic rank equal to $(\kappa + o(1))\sqrt{\log N_E}$, where κ is the unique real solution to the equation

$$
\eta = \frac{1}{\sqrt{2\pi}} \int_{-\kappa}^{\infty} e^{-x^2/2} \, \mathrm{d}x.
$$

⁽⁵⁾ For elliptic curves over the rational function field $\mathbb{F}_q(t)$, Brumer's analogue of Mestre's bound [3, Proposition 6.9] states that rank $(E/K) \ll \deg(\mathfrak{n}_d)/\log_q \deg(\mathfrak{n}_d)$.

⁽⁶⁾ This fact can actually be checked directly using Proposition 3.1.

⁽⁷⁾ Under the Riemann Hypothesis and LI for the functions $L(E, s)$, this is an equivalence.

1248 B. CHA, D. FIORILLI AND F. JOUVE

2. Limiting distributions associated with elliptic curves over function fields

2.1. Recollection on L**-functions**

We keep the notation as in §1.2. For a closed point v of C of good reduction for E, we let $a_v(E)$ be the integer defined by

$$
a_v(E) := q_v + 1 - \#E_v(k_v).
$$

Here, $\#E_{v}(k_{v})$ is the number of k_{v} -rational points on the reduction E_{v} of E at v. Also, since E has a good reduction at v, it is well-known that $a_v(E)$ is the trace of the k_v -Frobenius map on the ℓ -adic Tate module of E_v/k_v . Moreover, if we let α_v and β_v be its eigenvalues, then

$$
|\iota(\alpha_v)| = |\iota(\beta_v)| = q_v^{1/2} = q^{\deg(v)/2},
$$

for any embedding ι of $\overline{\mathbb{Q}}_{\ell}$ into the field of complex numbers (for simplicity we will omit ι from now on; its use, where needed, will be implicit). Therefore, after we fix one such embedding, there exists a unique θ_v in $[0, \pi]$ for each v of good reduction such that

$$
\alpha_v = \overline{\beta_v} = q^{\deg(v)/2} e^{i\theta_v}
$$

.

Let us define precisely what are the L-functions that naturally come into play in our study. We will follow the definition of [21, § 3.1.7] to define the L-function $L(\rho, K, T)$ for any continuous, absolutely irreducible ℓ -adic representation

$$
\rho\colon G_K\longrightarrow\operatorname{GL}(V)
$$

of the absolute Galois group G_K in some finite dimensional \mathbb{Q}_ℓ -vector space V. For each v, we choose a decomposition group $D_v \subset G(K)$ and we let I_v and Frob_v be the corresponding inertia group and the geometric Frobenius conjugacy class. Then, the L-function $L(\rho, K, T)$ is defined by the formal product

(5)
$$
L(\rho, K, T) = \prod_v \det \left(1 - \rho(\text{Frob}_v) T^{\deg v} | V^{\rho(I_v)} \right)^{-1},
$$

where $V^{\rho(I_v)}$ is the subspace of inertia invariants of V.

Of interest to us is the continuous ℓ -adic representation

$$
\rho_{\ell,E/K}\colon G_K\longrightarrow \operatorname{Aut}(V_{\ell}(E)),
$$

arising from the Galois action on $V_{\ell}(E) := T_{\ell}(E) \otimes \mathbb{Q}_{\ell}$, where $T_{\ell}(E)$ is the ℓ -adic Tate mod[ule](#page-40-1) of E/K . Because of a well-known independence of ℓ property, (namely $(\rho_{E/K})_{\ell}$) forms a compatible system of representations), the L-function $L(\rho_{\ell,E/K}, K, T)$ will be denoted simply $L(E/K, T)$ in the sequel. Its local factors are given explicitly as follows (see, e.g., [22, Lecture 1]):

(6)
$$
L(E/K,T) = \prod_{v \text{ good}} (1 - a_v(E) T^{\deg(v)} + q_v T^{2 \deg(v)})^{-1} \cdot \prod_{v \text{ bad}} (1 - a_v(E) T^{\deg(v)})^{-1}.
$$

Here, when v is a prime of bad reduction, we set $a_v(E) := 1, -1$ or 0, depending on the reduction type of E at v being split multiplicative, nonsplit multiplicative or additive, respectively. Also, for each $m \geq 1$, we form

$$
\text{Sym}^m(\rho_{\ell,E/K}): G_K \longrightarrow \text{Aut}(\text{Sym}^m(V_{\ell}(E))),
$$

by taking the m-th symmetric power of $\rho_{\ell,E/K}$. Again, by independence of ℓ , we can and we will write $L((\text{Sym}^m E)/K, T)$ for the L-function associated with $\text{Sym}^m(\rho_{\ell,E/K})$. The local factors of $L((Sym^m E)/K, T)$ can be described as follows. If E/K has good reduction at v, then its local factor at v is

(7)
$$
\prod_{j=0}^{m} (1 - \alpha_v^{m-j} \beta_v^j T^{\deg(v)})^{-1},
$$

whereas, for a ramified prime v , its local factor is

(8)
$$
(1 - a_v(E)^m T^{\deg(v)})^{-1},
$$

wit[h, a](#page-39-11)gain, $a_v(E) = 1, -1$, or 0, depending on the reduction type of E/K at v as before.

Recall that E/K is assumed to have nonconstant j-invariant. As a result of this, $L((\text{Sym}^m E)/K, T)$ is a polynomial in T (see [21, §3.1.7], as well as the introduction of [15]). More precisely, $L((Sym^m E)/K, T) \in 1 + T\mathbb{Z}[T]$. We define ν_m to be the degree of $L((\text{Sym}^m E)/K, T)$ and write

(9)
$$
L((\text{Sym}^m E)/K,T) = \prod_{j=1}^{\nu_m} (1 - \gamma_{m,j} T),
$$

for some complex numbers $\gamma_{m,j}$. For $m=1$, we also use the notation $N_{E/K} := \nu_1 = \deg(L(E/K, T)).$ Deligne's purity result [6, § 3.2.3] implies that $\gamma_{m,j}$ is of absolute value $q^{(m+1)/2}$ under any complex embedding of $\overline{\mathbb{Q}_\ell}$. Therefore, we can define the angles $\theta_{m,j}$ by the equation

$$
\gamma_{m,j} = q^{(m+1)/2} \mathrm{e}^{\mathrm{i} \theta_{m,j}},
$$

for all $j = 1, \ldots, \nu_m$ and for each $m \ge 1$. Note that ν_m can be given explicitly by the formula (see [21, § 3.1.7])

(11)
$$
\nu_m = (2g_C - 2)(m+1) + \deg(\mathfrak{n}_m).
$$

Here, g_C is the genus of C/k and \mathfrak{n}_m is the global Artin conductor of $\text{Sym}^m(\rho_{\ell,E/K})$. We will need the following lemma, which says that ν_m grows at most linearly with m.

LEMMA 2.1. – *There exists a positive constant* $C_{E/K}$ *which depends only on* E/K *such that*

$$
\nu_m \leq C_{E/K} \cdot m
$$

for all $m \geq 1$ *.*

Proof. – Let $m \geq 1$ be a fixed integer. Suppose that E/K has bad reduction at v and let K_v be the completion of K with respect to v. Thanks to the Euler characteristic Formula (11), it is enough to show that the exponent $f(\mathrm{Sym}^m \rho_{\ell,E/K_v})$ of the (local) Artin conductor of Sym^m $\rho_{\ell,E/K_v}$ is bounded by dim Sym^m $\rho_{\ell,E/K_v} = m + 1$, up to a multiplicative constant depending only on E/K_v .

Fix a separable closure K_v^{sep} of K_v and, for simplicity, set $\rho_\ell := \rho_{\ell,E/K_v}$ and $\rho_{\ell,m} := \text{Sym}^m \rho_{\ell,E/K_v}$. Let us further denote by V_ℓ and $V_{\ell,m}$ the representation spaces respectively attached to these representations. For an arbitrary ℓ -adic representation

 τ : Gal(K_v^{sep}/K_v) \rightarrow GL(W) we use the following expression (see, e.g., [23, §4, §5, and (9.1)] and the references therein):

$$
f(\tau) := \int_{-1}^{\infty} \operatorname{codim} W^{G^s} \mathrm{d} s \,,
$$

where $(G^s)_{s\geq -1}$ is the (decreasing) ramification filtation on $Gal(K_v^{\text{sep}}/K_v)$ in the upper numbering. Let $s_0(E/K_v) > 0$ be the smallest positive integer such that codim $V_\ell^{G^s} = 0$ for all $s \geq s_0(E/K_v)$. The integer $s_0(E/K_v)$ only depends on E/K_v and the fact that codim $V_t^{G^s} = 0$ for all $s \geq s_0(E/K_v)$ can be rephrased by saying that G^s acts trivially on V_t for all $s \geq s_0(E/K_v)$. In particular G^s acts trivially on $V_{\ell,m}$ for all $s \geq s_0(E/K_v)$ and we deduce

$$
f(\rho_{\ell,m})=\int_{-1}^{s_0(E/K_v)} \operatorname{codim} V_{\ell,m}^{G^s}{\rm d} s\,.
$$

A trivial upper bound for the right hand side is $(s_0(E/K_v)+1)$ dim $V_{\ell,m}$. This completes the proof, since this last quantity equals $(s_0(E/K_v) + 1)(m + 1)$. \Box

R 2.2. – A somewhat more explicit argument was indicated to us by Henniart. Let us briefly sketch it, using the same notation as in the proof above.

We use the fact that the representation we start with is 2-dimensional since it corresponds to the ℓ -adic Tate module $V_{\ell}(E)$ attached to E/K . Its conductor can be related to the conductor of a Weil-Deligne representation of the Weil group W_{K_v} (see, e.g., [23, Th. 1]). According to the classification of the 2-dimensional Weil-Deligne representations, only three cases may occur:

1. $\rho_{\ell,E/K_v} = \chi_1 \oplus \chi_2$, where χ_1, χ_2 are characters. Thus

$$
\text{Sym}^m \, \rho_{\ell, E/K_v} = \sum_{k=0}^m \chi_1^k \chi_2^{m-k} \,,
$$

and the statement follows by linearity of the Swan conductor (again, see, e.g., [23] and the references therein for definitions).

2. $\rho_{\ell,E/K_v}$ = $\chi \otimes$ St, i.e., $\rho_{\ell,E/K_v}$ is a twist of the Steinberg representation by an unramified character χ . Thus

$$
\operatorname{Sym}^m \rho_{\ell,E/K_v} = \chi^m \otimes \operatorname{St}_m,
$$

and the statement follows by properties of the Steinberg representation.

3. $\rho_{\ell,E/K_v}$ is irreducible; it is the so-called *supersingular* case. In this case, one can use the inequality of slo[pes](#page-40-2)

$$
\lambda(\operatorname{Sym}^m \rho_{\ell, E/K_v}) \leq \lambda(\rho_{\ell, E/K_v}),
$$

and invoke, e.g., [23, Prop. 2] to conclude.

2.2. An explicit form[ula](#page-39-6)

Throughout this section we fix an elliptic curve E over K . For readability we do not indicate the dependency on E of the objects we introduce whenever it is clear from context.

Following Sarnak [19], we will consider a function $V : [0, \pi] \longrightarrow \mathbb{R}$ and study the limiting distribution arising from

$$
T_V(X) := \frac{X}{q^{X/2}} \sum_{\substack{\deg(v) \leq X \\ v \text{ good}}} V(\theta_v)
$$

as $X \to \infty$. This is done by computing the Fourier expansion of V using the functions

$$
U_m(\theta) := \frac{\sin((m+1)\theta)}{\sin \theta},
$$

for $m\geq 0.$ Indeed the family $\{U_m\}_{m=0}^\infty$ forms an orthonormal basis of $L^2([0,\pi])$ with respect to the inner product

$$
\langle V_1, V_2 \rangle := \frac{2}{\pi} \int_0^{\pi} V_1(\theta) V_2(\theta) \sin^2 \theta \, d\theta.
$$

Obviously computing quantities $T_{U_m}(X)$ is crucial and suffices to understand more generally $T_V(X)$ for functions V that coincide with their Fourier expansion with respect to the family $\{U_m\}_{m\geq 0}$. We now proceed to comp[uting](#page-5-1) explicitly $T_{U_m}(X)$ for $m \geq 1$.

The following result can be seen as a so-called "explicit formula" that relates a summation over primes to a summation over zeros of a certain L-function. It will be of crucial importance in the proof of the theorems stated in § 1.2.

THEOREM 2.3. – Let $m \ge 1$ and $N \ge 1$ be integers. Let $\mathcal{E}(N) := 1$ if N is even, and $\mathcal{E}(N) := 0$ if N is odd. Then one has:

$$
\frac{N}{q^{N/2}} \sum_{\deg(v)=N} U_m(\theta_v) = (-1)^{m+1} \mathcal{E}(N) - \sum_{j=1}^{\nu_m} e^{iN \theta_{m,j}} + O_E(m^2 q^{-N/6}).
$$

To prove the theorem, we need some preliminary results.

PROPOSITION 2.4. – *With notation as in Theorem 2.3, one has:*

(12)
$$
-\sum_{j=1}^{\nu_m} e^{i N \theta_{m,j}} = q^{-N/2} \sum_{d|N} d \sum_{\substack{\deg(v) = d \\ v \text{ good}}} U_m(\frac{N}{d} \theta_v) + O_E(q^{-(m+1)N/2}).
$$

Proof. – We take the log derivative of $L((Sym^m E)/K, T)$ using the Euler factors (7) and (8) and obtain

(13)
$$
T\frac{L'}{L}((\text{Sym}^m E)/K,T) = \sum_{d=1}^{\infty} \sum_{\substack{\text{deg}(v)=d \\ v \text{ good}}} \sum_{j=0}^{m} \frac{d \alpha_v^{m-j} \beta_v^j T^d}{1 - \alpha_v^{m-j} \beta_v^j T^d} + \sum_{d=1}^{\infty} \sum_{\substack{\text{deg}(v)=d \\ v \text{ bad}}} \frac{d \alpha_v^m T^d}{1 - \alpha_v^m T^d}.
$$

.

To simplify the summation over good primes v , we rewrite U_m as

(14)
$$
U_m(\theta) = \sum_{j=0}^m e^{i(m-2j)\theta}
$$

This gives, for any positive integer $k \geq 1$,

$$
U_m(k\theta_v) = \sum_{j=0}^m e^{ik(m-2j)\theta_v} = q_v^{-km/2} \sum_{j=0}^m \left(q_v^{k(m-j)/2} e^{ik(m-j)\theta_v} \right) \left(q_v^{kj/2} e^{-ikj\theta_v} \right)
$$

= $(q^{m/2})^{-dk} \sum_{j=0}^m \alpha_v^{k(m-j)} \beta_v^{kj}$.

Thus

(15)
$$
(q^{m/2})^{dk} U_m(k\theta_v) = \sum_{j=0}^m \alpha_v^{k(m-j)} \beta_v^{kj}.
$$

The summation over good primes v in (13) becomes

(16)
\n
$$
\sum_{d=1}^{\infty} \sum_{\substack{\deg(v)=d}} \sum_{j=0}^{m} \frac{d \alpha_v^{m-j} \beta_v^j T^d}{1 - \alpha_v^{m-j} \beta_v^j T^d} = \sum_{d,k=1}^{\infty} \sum_{\substack{\deg(v)=d \\ v \text{ good}}} d \sum_{j=0}^{m} \alpha_v^{k(m-j)} \beta_v^{kj} T^{dk}
$$
\n
$$
= \sum_{N=1}^{\infty} \left((q^{m/2})^N \sum_{\substack{d|N \\ d|N}} d \sum_{\substack{\deg(v)=d \\ v \text{ good}}} U_m(\frac{N}{d} \theta_v) \right) T^N,
$$

where the last equality follows from (15). Similarly the contribution of bad primes v to (13) is $\sqrt{2}$

(17)
$$
\sum_{d=1}^{\infty} \sum_{\substack{\deg(v)=d\\v \text{ bad}}} \frac{d a_v^m T^d}{1 - a_v^m T^d} = \sum_{N=1}^{\infty} \left(\sum_{d|N} d \sum_{\substack{\deg(v)=d\\v \text{ bad}}} a_v^{mN/d} \right) T^N.
$$

Since $a_v = 0$ or ± 1 and there are only finitely many bad primes v we see that the coefficient of T^N above is bounded by a constant depending on E/K . Using this and (16), we simplify (13) as follows:

(18)
$$
T\frac{L'}{L}((\text{Sym}^m E)/K,T) = \sum_{N=1}^{\infty} \left((q^{m/2})^N \sum_{\substack{d|N \ d \text{eg}(v)=d}} d \sum_{\substack{deg(v)=d \ v \text{ good}}} U_m(\frac{N}{d} \theta_v) + O_E(1) \right) T^N.
$$

On the other hand, taking the log derivative of (9) and comparing the N-th coefficient of T^N with (18), we conclude that

$$
-\sum_{j=1}^{\nu_m} \gamma_{m,j}^N = (q^{m/2})^N \sum_{d|N} d \sum_{\substack{\text{deg}(v)=d\\v \text{ good}}} U_m(\frac{N}{d} \theta_v) + O_E(1).
$$

Dividing out both sides by $q^{(m+1)N/2}$ this finishes the proof of Proposition 2.4.

 \Box

C 2.5. – *With notation as in Theorem 2.3, one has:*

$$
\sum_{\deg(v)=N} U_m(\theta_v) \ll_E m \frac{q^{N/2}}{N}.
$$

 $\frac{1}{2}$

Proof. – In (12), we split out the term $d = N$ to get (19)

$$
\frac{N}{q^{N/2}} \sum_{\deg(v)=N} U_m(\theta_v) = -\sum_{j=1}^{\nu_m} e^{i N \theta_{m,j}} - q^{-N/2} \sum_{\substack{d|N \ d\leq N/2}} d \sum_{\substack{\deg(v)=d \deg(v)=d}} U_m(\frac{N}{d} \theta_v) + O_E(q^{-N/2}).
$$

Then, it is enough to show that the right side is bounded by $(m + 1)$ times an absolute constant. From Lemma 2.1, we have

(20)
$$
\left| - \sum_{j=1}^{\nu_m} e^{i N \theta_{m,j}} \right| \leq C_{E/K} m.
$$

To bound the rest of th[e su](#page-14-0)mmation in the right side of (19), we use the trivial bound

$$
|U_m(\theta)| \le m + 1,
$$

which is immediate from (14), and the count of places of fixed degree in $K = \mathbb{F}_q(C)$ (see, e.g., [3, Prop. 6.3]) which yields the estimate

(22)
$$
\left| \left(\sum_{\deg(v)=d} 1 \right) - \frac{q^d}{d} \right| \le \frac{2g_C + 1}{1 - q^{-1}} q^{d/2}.
$$

The second term on the right hand side of (19) is then bounded above in absolute value by

$$
(m+1)q^{-N/2} \sum_{\substack{d|N\\d\le N/2}} d \sum_{\substack{\deg(v)=d\\v \text{ good}}} 1 \ll_C mq^{-N/2} \sum_{\substack{d|N\\d\le N/2}} q^d \ll m.
$$

 \Box

The proof follows.

We are now read[y to p](#page-0-0)rove Theorem 2.3.

Proof of Theorem 2.3. – We begin with (19) and separate out the term $d = N/2$, which exists only when N is even. (Recall, by definition, $\mathcal{E}(N) = 1$ if N is even and 0 otherwise.)

$$
(23)
$$

$$
\frac{N}{q^{N/2}} \sum_{\deg(v)=N} U_m(\theta_v) = -\sum_{j=1}^{\nu_m} e^{i N \theta_{m,j}} - \frac{1}{q^{N/2}} \sum_{\substack{d|N \ d \le N/2}} d \sum_{\deg(v)=d} U_m(\frac{N}{d} \theta_v) + O_E(q^{-N/2})
$$

$$
= -\sum_{j=1}^{\nu_m} e^{i N \theta_{m,j}} - \mathcal{E}(N) \frac{N/2}{q^{N/2}} \sum_{\substack{d \in S(v)=N/2}} U_m(2\theta_v)
$$

$$
- \frac{1}{q^{N/2}} \sum_{\substack{d|N \ d \le N/3}} d \sum_{\substack{v \text{ good}}} U_m(\frac{N}{d} \theta_v) + O_E(q^{-N/2}).
$$

First, we handle the terms with $d \le N/3$. From (21) and (22) again,

(24)
$$
\left| \frac{1}{q^{N/2}} \sum_{\substack{d|N \\ d \le N/3}} d \sum_{\substack{\deg(v) = d \\ v \text{ good}}} U_m(\frac{N}{d} \theta_v) \right| \le \frac{m+1}{q^{N/2}} \sum_{\substack{d|N \\ d \le N/3}} d\left(\frac{q^d}{d} + O_C\left(\frac{q^{d/2}}{d}\right)\right)
$$

Next, to estimate $\sum_{\deg(v)=N/2} U_m(2\theta_v)$, we consider the following Fourier expansion of $V(\theta) := U_m(2\theta)$ with respect to the orthonormal basis $\{U_k\}_{k=0}^{\infty}$

$$
V(\theta) = \sum_{k=0}^{\infty} \langle V, U_k \rangle U_k(\theta).
$$

One has:

$$
\sum_{\deg(v)=N/2} V(\theta_v) = \sum_{k=0}^{\infty} \langle V, U_k \rangle \sum_{\deg(v)=N/2} U_k(\theta_v)
$$

= $\langle V, U_0 \rangle \sum_{\deg(v)=N/2} 1 + \sum_{k=1}^{\infty} \langle V, U_k \rangle \sum_{\deg(v)=N/2} U_k(\theta_v)$
= $\langle V, U_0 \rangle \left(\frac{q^{N/2}}{N/2} + O_C(q^{N/4}/N) \right) + \sum_{k=1}^{\infty} \langle V, U_k \rangle \sum_{\deg(v)=N/2} U_k(\theta_v).$

Note that Lemma 2.6 b[elow](#page-0-0) provides the Fourier coefficients of V explicitly. In particular,

(25) $\langle V, U_0 \rangle = (-1)^m$.

Moreover, Corollary 2.5 gives

$$
\sum_{\deg(v)=N/2} U_k(\theta_v) \ll_E k \frac{q^{N/4}}{N}.
$$

Again, by Lemma 2.6, we can determine $\langle V, U_k \rangle$ for all $k \ge 1$. We deduce

(26)
$$
\sum_{k=1}^{\infty} \langle V, U_k \rangle \sum_{\deg(v) = N/2} U_k(\theta_v) \ll_E \frac{q^{N/4}}{N} \sum_{k=1}^{\infty} |\langle V, U_k \rangle| k
$$

$$
\ll_E \frac{q^{N/4}}{N} m^2.
$$

From (25) and (26), we obtain

(27)
$$
\frac{N/2}{q^{N/2}} \sum_{\deg(v)=N/2} U_m(2\theta_v) = (-1)^m + O_E(m^2q^{-N/4}).
$$

Collect (24) and (27) and put them into (23), to obtain

$$
\frac{N}{q^{N/2}} \sum_{\deg(v)=N} U_m(\theta_v) = -\sum_{j=1}^{\nu_m} e^{i\,N\theta_{m,j}} - \mathcal{E}(N)(-1)^m + O_E(m^2q^{-N/6}).
$$

This finishes the proof of Theorem 2.3.

4 ^e SÉRIE – TOME 49 – 2016 – N^o 5

 \Box

LEMMA 2.6. – *[For](#page-14-0)* $m > 0$,

$$
U_m(2\theta) = U_{2m}(\theta) - U_{2m-2}(\theta) + \cdots + (-1)^{m+1}U_2(\theta) + (-1)^m.
$$

Proof. – Using (14), one can prove that

$$
U_{m+1}(2\theta) + U_m(2\theta) = U_{2(m+1)}(\theta).
$$

 \Box

Now, an easy induction on m co[mple](#page-0-0)tes the proof.

2.3. Limiting distribution arising from smooth functions

We now derive from Theorem 2.3 a decomposition of $T_V(X)$ (under suitable hypotheses on the function V) from which we deduce the existence of a limiting distribution for $T_V(X)$.

Let $V : [0, \pi] \longrightarrow \mathbb{R}$ be a function and let $V_m := \langle V, U_m \rangle$ be the m-th Fourier coefficient of V. To ensure convergence, we will assume that $V_m \ll m^{-3-\eta}$ for some $\eta > 0$. In particular this will guarantee that the Fourier expansion

$$
V(\theta) = \sum_{m=1}^{\infty} V_m U_m(\theta)
$$

converges (uniformly and absolutely) for $\theta \in [0, \pi]$, by the tri[vial](#page-0-0) bound $|U_m(\theta)| \le m + 1$. The reason why we require such a strong decay rate for V_m is that we also need to ensure the convergence of the error term in (28).

Assuming further that $\langle V, U_0 \rangle = 0$ we may apply Theorem 2.3 to get:

$$
(28) \quad T_V(X) = \frac{X}{q^{X/2}} \sum_{N=1}^{X} \sum_{\substack{\text{deg } v = N \\ v \text{ good}}} V(\theta_v) = \frac{X}{q^{X/2}} \sum_{N=1}^{X} \sum_{m=1}^{\infty} V_m \sum_{\substack{\text{deg } v = N \\ v \text{ good}}} U_m(\theta_v)
$$

$$
= \frac{X}{q^{X/2}} \sum_{N=1}^{X} \frac{q^{N/2}}{N} \left(\mathcal{E}(N) \sum_{m=1}^{\infty} V_m(-1)^{m+1} - \sum_{m=1}^{\infty} V_m \sum_{j=1}^{\nu_m} e^{iN\theta_{m,j}} + O\left(\sum_{m=1}^{\infty} m^2 V_m q^{-N/6}\right) \right).
$$

We break the last line into:

$$
T_V(X) = T_V^{(1)}(X) + T_V^{(II)}(X) + O(X(\log X)^{-1}q^{-X/6}),
$$

where

$$
T_V^{(I)}(X) := \frac{X}{q^{X/2}} \sum_{N=1}^X \frac{q^{N/2}}{N} \mathcal{E}(N) \left(\sum_{m=1}^\infty V_m(-1)^{m+1} \right),
$$

$$
T_V^{(II)}(X) := -\frac{X}{q^{X/2}} \sum_{N=1}^X \frac{q^{N/2}}{N} \left(\sum_{m=1}^\infty V_m \sum_{j=1}^{\nu_m} e^{i N \theta_{m,j}} \right).
$$

To simplify $T_V^{(1)}$ $V_V^{(\rm I)}(X)$ and $T_V^{(\rm II)}$ $V_V^{(11)}(X)$ further, we use [4, Cor. 2.3 and Cor. 2.4]. Since the statements are quite short we recall them without proof in the following lemma.

LEMMA 2.7. – Let $N \geq 1$ *be an integer and let* $\mathcal{E}(N)$ *be as above. The following holds.*

(i) *Let*

(29)
$$
c_{\pm}(X) := \begin{cases} q/(q-1) & \text{for even } X, \\ \sqrt{q}/(q-1) & \text{for odd } X. \end{cases}
$$

Then

$$
\frac{X}{q^{X/2}} \sum_{N=1}^{X} \mathcal{E}(N) \frac{q^{N/2}}{N} = c_{\pm}(X) + o(1) ,
$$

 $as X \rightarrow \infty$ *.*

(ii) Let $\gamma := \sqrt{q}e^{i\theta}$ *be a complex number of argument* $\theta \in [0, 2\pi]$ *. Then*⁽⁸⁾

$$
\frac{X}{q^{X/2}}\sum_{N=1}^X\frac{\gamma^N}{N}=\frac{\gamma\mathrm{e}^{\mathrm{i}\theta X}}{\gamma-1}+o(1)\,,
$$

 $as X \to \infty$ *.*

Applying (i) of the lemma we obtain

(30)
$$
T_V^{(1)}(X) = c_{\pm}(X) \sum_{m=1}^{\infty} (-1)^{m+1} V_m + o(1).
$$

For $T_V^{({\rm II})}$ $V_V^{(11)}(X)$, we apply (ii) of the lemma with $\gamma = \gamma_{m,j}/q^{m/2}$ (see (10)). Then, as $X \to \infty$,

$$
\frac{X}{q^{X/2}} \sum_{N=1}^{X} \frac{q^{N/2}}{N} e^{iN\theta_{m,j}} = \frac{X}{q^{X/2}} \sum_{N=1}^{X} \frac{(\gamma_{m,j}/q^{m/2})^N}{N} = \frac{\gamma_{m,j}}{\gamma_{m,j} - q^{m/2}} e^{i\theta_{m,j}X} + o(1).
$$

Thus,

(31)
$$
T_V^{(\text{II})}(X) = \sum_{m=1}^{\infty} V_m \left(\sum_{j=1}^{\nu_m} \frac{\gamma_{m,j}}{\gamma_{m,j} - q^{m/2}} e^{i \theta_{m,j} X} \right) + o(1),
$$

which converges absolutely by our assumption that $V_m \ll m^{-3-\eta}$.

Next we separate out the terms with $\gamma_{m,j} = q^{(m+1)/2}$, or equivalently $\theta_{m,j} = 0$. To do so, define

(32)
$$
M_m(1) := \#\{j \mid \theta_{m,j} = 0 \text{ with } j = 1, ..., \nu_m\}.
$$

In other words, $M_m(1)$ is the multiplicity of the zero $T = q^{-(m+1)/2}$ in $L((\text{Sym}^m E)/K, T)$. Then, from (31),

$$
(33) T_V^{(II)}(X) = -\sum_{m=1}^{\infty} V_m \frac{\sqrt{q}}{\sqrt{q}-1} M_m(1) - \sum_{m=1}^{\infty} V_m \sum_{\substack{j=1,\dots,\nu_m \\ \theta_{m,j} \neq 0}} \frac{\gamma_{m,j}}{\gamma_{m,j} - q^{m/2}} e^{i\theta_{m,j}X} + o(1).
$$

Combining (30) and (33) we obtain the following result.

(8) In [4] this is stated under the assumption $\theta \in [0, \pi]$ but the conclusion obviously holds more generally for any $\theta \in [0, 2\pi]$ simply by applying complex conjugation.

PROPOSITION 2.8. – *For a function* $V : [0, \pi] \longrightarrow \mathbb{R}$ *with* $\langle V, U_0 \rangle = 0$ *and* $V_m \ll m^{-3-\eta}$ *for some* $\eta > 0$ *, we have that*

$$
T_V(X) = Q_V(X) + R_V(X) + o_{X \to \infty}(1)
$$

where

$$
Q_V(X) := \sum_{m=1}^{\infty} \left((-1)^{m+1} c_{\pm}(X) - \frac{\sqrt{q}}{\sqrt{q}-1} M_m(1) \right) V_m,
$$

$$
R_V(X) := -\sum_{m=1}^{\infty} V_m \sum_{\substack{j=1,\ldots,\nu_m \\ \theta_{m,j} \neq 0}} \frac{\gamma_{m,j}}{\gamma_{m,j} - q^{m/2}} e^{i\theta_{m,j}X}.
$$

Here $c_{\pm}(X)$ *and* $M_m(1)$ *are defined as in* [\(29](#page-0-0)) *and* (32).

COROLLARY 2.9. – Let V be as in Proposition 2.8. The quantity $T_V(X)$ has a limiting *distribution* μ_V *in the sense of Definition 1.1. Moreover, for* $k \geq 1$ *,*

(34)
$$
\lim_{M \to \infty} \frac{1}{M} \sum_{X=1}^{M} T_V(X)^k = \int_{\mathbb{R}} t^k d\mu_V(t).
$$

Proof. – This is the function field analogue of [9, Lemmas 2.3 and 2.5]. We just need to make a slight adaption to handle the difference [aris](#page-39-0)ing fro[m th](#page-39-12)e fact that our definition of limiting di[str](#page-39-12)ibution (Definition 1.1) uses a (discrete) summation, rather than an integral.

First, we outline the proof of existence of a limiting distribution. This is well-known in the number field setting, originally described in [18] and in [1] for more general situations. Following [1], we say that a real-valued function $\phi(X)$ defined for all positive integers X is a B^2 -almost periodic function if for any $\epsilon > 0$, there exists a real-valued trigonometric polynomial

(35)
$$
P_{N(\epsilon)}(X) = \sum_{n=1}^{N(\epsilon)} r_n(\epsilon) e^{i\lambda_n(\epsilon) X}
$$

such that

(36)
$$
\limsup_{Y \to \infty} \frac{1}{Y} \sum_{X=1}^{Y} |\phi(X) - P_{N(\epsilon)}(X)|^2 < \epsilon^2.
$$

Here, $\lambda_n(\epsilon)$ [is r](#page-20-0)eal and $r_n(\epsilon)$ is complex. It can [be](#page-39-12) shown that any trigonometric polynomial of the form (35) has a limiting distribution (in the sense of Definition 1.1.) This is due to a discrete version of th[e Kr](#page-0-0)onecker-Weyl Theorem, whose proof can be found in [14, Append. A]. ⁽⁹⁾ Also, an obvious adaptation of [1, Th. 2.9] shows that any B^2 -almost periodic function has a limiting distribution. Next, for any (large) positive number M , we define (using the notation of Proposition 2.8)

$$
\mathcal{E}_V(X,M) = -\sum_{m \ge M} V_m \sum_{\substack{j=1,\ldots,\nu_m \\ \theta_{m,j} \neq 0}} \frac{\gamma_{m,j}}{\gamma_{m,j} - q^{m/2}} e^{i\theta_{m,j}X} + o_{X \to \infty}(1),
$$

(9) For a complete statement and proof for a "continuous" version, see http://mathoverflow.net/questions/ 162875/reference-for-kronecker-weyl-theorem-in-full-generality. For the discrete version, the statement and its proof are similar.

so that $T_V(X)$ is a sum of a trigonometric polynomial and $\mathcal{E}_V(X, M)$. Using a trivial bound we easily obtain

(37)
$$
|\mathcal{E}_V(X,M)| \leq \frac{\sqrt{q}}{\sqrt{q}-1} \sum_{m \geq M} |V_m| \nu_m + o_{X \to \infty}(1).
$$

The decay condition on V_m now implies its mean square can be [ma](#page-20-1)de arbi[tra](#page-39-3)rily small by choosing M large and we deduce that $T_V(X)$ is B^2 -almost periodic, thus has a limiting di[str](#page-6-0)ibution, say, μ_V . In fact, a straightforward calculation shows that $T_V(X)$ is bounded for all X, therefore, μ_V is supported on a bounded set in R. To prove (34), as in [9, Lemma 2.5], we choose a Lipschitz continuous function f equal to t^k on a set containing the support of μ_V and to zero outside and apply (2). We omit the details. \Box

2.4. Limiting distributi[on](#page-5-0) arising from $T_E(X)$

In this subsection, we look into the special case $T_E(X) = T_V(X)$, i.e., $V(\theta) = -U_1(\theta) =$ $-2\cos\theta$, as defined in (1). Thus we have $V_m = 0$ for all $m \geq 2$. Recall from §2.1 that $L(E/K, T)$ is a polynomial in T of degree $\nu_1 = N_{E/K}$. We will write its inverse zeros as $\gamma_j = q \mathrm{e}^{\mathrm{i} \theta_j}$ for $j = 1, \ldots, N_{E/K},$ so that

(38)
$$
L(E/K,T) = \prod_{j=1}^{N_{E/K}} (1 - q e^{i\theta_j} T).
$$

We note from the definition (32) that $M_1(1) = \text{rank}(E/K)$, the analytic rank of E/K . Proposition 2.8 gives the following statement in the case $V = -U_1$.

COROLLARY 2.10. – *With notation as above we have*

(39)
$$
T_E(X) = Q_E(X) + R_E(X) + o_{X \to \infty}(1),
$$

where

$$
Q_E(X) := \frac{\sqrt{q}}{\sqrt{q} - 1} \text{rank}(E/K) - c_{\pm}(X),
$$

\n
$$
R_E(X) := \sum_{\substack{j=1,\ldots,N_{E/K}\\ \theta_j \neq 0}} \frac{1}{1 - q^{-1/2} e^{-i\theta_j}} e^{i\theta_j X} = \sum_{\substack{j=1,\ldots,N_{E/K}\\ \theta_j \neq 0}} \frac{\gamma_j}{\gamma_j - \sqrt{q}} e^{i\theta_j X}.
$$

Proof. – [This](#page-0-0) is obvious from Proposition 2.8 because $V_1 = -1$ $V_1 = -1$ $V_1 = -1$ and $V_m = 0$ for all $m \geq 2$. \Box

Corollary 2.9 applied to the case $V = -U_1$ enables us to study the random variable X_E associated to the li[miting](#page-0-0) distribution of $T_E(X)$. Theorem 1.2(i) shows that we can obtain simple closed formulæ for the mean and variance of X_E .

Proof of Theorem 1.2 (i). – First, we compute

$$
\mathbb{E}[X_E] = \lim_{M \to \infty} \frac{1}{M} \sum_{X=1}^{M} T_E(X).
$$

From Corollary 2.10, we have

(40)
$$
T_E(X) = \frac{\sqrt{q}}{\sqrt{q}-1} \text{rank}(E/K) - c_{\pm}(X) + \sum_{\theta_j \neq 0} \frac{1}{1 - q^{-1/2} e^{-i\theta_j}} e^{i\theta_j X} + o_{X \to \infty}(1).
$$

It is easy to show that

$$
\lim_{M \to \infty} \frac{1}{M} \sum_{X=1}^{M} \left(\frac{\sqrt{q}}{\sqrt{q}-1} \text{rank}(E/K) - c_{\pm}(X) \right) = \frac{\sqrt{q}}{\sqrt{q}-1} \left(\text{rank}(E/K) - \frac{1}{2} \right).
$$

In addition, for any $\theta \in (0, 2\pi)$,

(41)
$$
\sum_{X=1}^{M} e^{i\theta X} = O\left(\frac{1}{\|\theta/2\pi\|}\right),
$$

where $\lVert \cdot \rVert$ denotes the distance to the nearest integer. The formula for $\mathbb{E}[X_E]$ follows immediately from (40). For $\mathbb{V}[X_E]$, we must compute

$$
\mathbb{V}[X_E] = \lim_{M \to \infty} \frac{1}{M} \sum_{X=1}^{M} (T_E(X) - \mathbb{E}[X_E])^2.
$$

Note that $\sqrt{q}/(2(\sqrt{q}-1)) - c_{\pm}(X) = \pm \sqrt{q}/(2(\sqrt{q}+1))$, with a + (resp. –) sign if X is odd (resp. even). Thus

(42)
$$
\left(\frac{1}{2}\frac{\sqrt{q}}{\sqrt{q}-1} - c_{\pm}(X) + \sum_{\theta_j \neq 0} \frac{1}{1 - q^{-1/2}e^{-i\theta_j}} e^{i\theta_j X}\right)^2 = \frac{1}{4} \left(\frac{\sqrt{q}}{\sqrt{q}+1}\right)^2 \pm \frac{\sqrt{q}}{\sqrt{q}+1} \sum_{\theta_j \neq 0} \frac{1}{1 - q^{-1/2}e^{-i\theta_j}} e^{i\theta_j X} + \left(\sum_{\theta_j \neq 0} \frac{1}{1 - q^{-1/2}e^{-i\theta_j}} e^{i\theta_j X}\right)^2.
$$

Using this and (40), we obtain

$$
\mathbb{V}[X_E] = \lim_{M \to \infty} \frac{1}{M} \sum_{X=1}^{M} \left(\frac{1}{2} \frac{\sqrt{q}}{\sqrt{q} - 1} - c_{\pm}(X) + \sum_{\theta_j \neq 0} \frac{1}{1 - q^{-1/2} e^{-i\theta_j}} e^{i\theta_j X} \right)^2
$$

= $\frac{1}{4} \left(\frac{\sqrt{q}}{\sqrt{q} + 1} \right)^2 + \lim_{M \to \infty} \frac{1}{M} \sum_{X=1}^{M} \left(\sum_{\theta_j \neq 0} \frac{1}{1 - q^{-1/2} e^{-i\theta_j}} e^{i\theta_j X} \right)^2$,

where the last line follows from (42) and (41). Denoting by \sum^* a sum over θ_j counted without multiplicity, we compute

$$
\sum_{X=1}^{M} \left(\sum_{\theta_j \neq 0} \frac{1}{1 - q^{-1/2} e^{-i\theta_j}} e^{i\theta_j X} \right)^2 = \sum_{X=1}^{M} \left(\sum_{\theta_j \neq 0}^* \frac{m(\theta_j) e^{i\theta_j X}}{1 - q^{-1/2} e^{-i\theta_j}} \right)^2
$$

=
$$
\sum_{X=1}^{M} \left(\sum_{\theta_k \neq 0}^* \frac{m(\theta_k) e^{i\theta_k X}}{1 - q^{-1/2} e^{-i\theta_k}} \right) \left(\sum_{\theta_l \neq 0}^* \frac{m(\theta_l) e^{-i\theta_l X}}{1 - q^{-1/2} e^{i\theta_l}} \right).
$$

Splitting out the diagonal term, the right hand side equals

$$
M\sum_{\theta_j\neq 0}^* \frac{m(\theta_j)^2}{|1-q^{-1/2}\mathrm{e}^{-\mathrm{i}\theta_j}|^2} + \sum_{\substack{\theta_k,\theta_l\neq 0\\k\neq l}}^* \frac{m(\theta_k)m(\theta_l)}{(1-q^{-1/2}\mathrm{e}^{-\mathrm{i}\theta_k})(1-q^{-1/2}\mathrm{e}^{\mathrm{i}\theta_l})}\sum_{X=1}^M \mathrm{e}^{\mathrm{i}(\theta_k-\theta_l)X}\,.
$$

We divide the last line by M, let $M \to \infty$, and use (41) one more time to obtain

$$
\mathbb{V}[X_E] = \frac{1}{4} \left(\frac{\sqrt{q}}{\sqrt{q} + 1} \right)^2 + \sum_{\theta_j \neq 0}^* \frac{m(\theta_j)^2}{|1 - q^{-1/2} e^{-i\theta_j}|^2}.
$$

This concludes the proof.

3. Ulmer's fami[ly](#page-39-5)

The goal of this section is to prove the results stated in § 1.2 regarding the function $T_E(X)$ associated to the elliptic curves of Ulmer's family in [20]. Let $\mathbb{F}_q(t)$ be the rational function field over \mathbb{F}_q . Following [20], we define $E_d/\mathbb{F}_q(t)$ to be the elliptic curve over $\mathbb{F}_q(t)$ given by the Weierstrass equation

$$
E_d: y^2 + xy = x^3 - t^d,
$$

and we write $T_d(X)$ for the function $T_E(X)$ that arises from $E_d/\mathbb{F}_q(t)$. Essential to us is the following explicit description of their Hasse-Weil L-function.

PROPOSITION 3.1. – *Suppose that d divides* $p^{n} + 1$ *for some n*, *and let* $L(E_d/\mathbb{F}_q(t), T)$ *be the Hasse-Weil L-function of* E_d *over* $\mathbb{F}_q(t)$ *. Then,*

$$
L(E_d/\mathbb{F}_q(t),T) = (1-qT)^{\epsilon_d} \prod_{\substack{e|d\\e\nmid 6}} \left(1 - (qT)^{o_e(q)}\right)^{\phi(e)/o_e(q)}.
$$

Here, $\phi(e) = \#(\mathbb{Z}/e\mathbb{Z})^*$ *is the Euler-phi function and* $o_e(q)$ *is the (multiplicative) order of* q $\sin(\mathbb{Z}/e\mathbb{Z})^*$ *. Further,* ϵ_d *is defined as*

$$
\epsilon_d := \begin{cases} 0 & \text{if } 2 \nmid d \text{ or } 4 \nmid q - 1 \\ 1 & \text{if } 2 \mid d \text{ and } 4 \mid q - 1 \end{cases} + \begin{cases} 0 & \text{if } 3 \nmid d \\ 1 & \text{if } 3 \mid d \text{ and } 3 \nmid q - 1 \\ 2 & \text{if } 3 \mid d \text{ and } 3 \mid q - 1 \end{cases}.
$$

In particular we have an explicit formula for the analytic rank of $E_d/\mathbb{F}_q(t)$ *:*

$$
\operatorname{rank}(E/\mathbb{F}_q(t)) = \epsilon_d + \sum_{\substack{e \mid d \\ e \nmid 6}} \frac{\phi(e)}{o_e(q)}.
$$

Proof. – This is essentially a corollary of the main results from [20]. To briefly highlight the main ingredients, [20, Cor. 7.7 and Prop. 8.1] computes the characteristic polynomial, under the action of Frobenius, of the (degree 2 component of the) étale cohomohology group for (a certain quotient of) the Fermat surface of degree d, whose image under blow-ups provides a smooth proper model for E_d over $\mathbb{F}_q(t)$. Under the assumption that d divides $p^{n} + 1$ for some *n*, this characteristic polynomial is given precisely by the expression in the statement of the present proposition.

4 ^e SÉRIE – TOME 49 – 2016 – N^o 5

 \Box

The statement about the rank is then a straightf[orw](#page-19-2)ard consequence of the formula for the L-function of $E_d/\mathbb{F}_q(t)$. \Box

PROPOSITION 3.2. – Let $c_{+}(X)$ be defined as in (29). Then

$$
T_d(X) = -c_{\pm}(X) + \frac{\epsilon_d}{1 - q^{-\frac{1}{2}}} + \sum_{\substack{e \mid d \\ e \nmid 6}} \phi(e) \frac{q^{-(X \mod o_e(q))/2}}{1 - q^{-o_e(q)/2}} + o_{X \to \infty}(1)
$$

for X large enough, where $0 \leq (X \mod \ell) \leq \ell - 1$ $0 \leq (X \mod \ell) \leq \ell - 1$ $0 \leq (X \mod \ell) \leq \ell - 1$ *is the [rema](#page-0-0)inder in the Euclidean division of* X *by* ℓ *.*

Proof. – We combine Corollary 2.10 and Proposition 3.1 to obtain

$$
(43) T_d(X) = -c_{\pm}(X) + \frac{\epsilon_d}{1 - q^{-\frac{1}{2}}} + \sum_{\substack{e|d\\e \neq 6}} \frac{\phi(e)}{o_e(q)} \sum_{k=0}^{o_e(q)-1} \frac{e^{2\pi i k X / o_e(q)}}{1 - q^{-1/2} e^{-2\pi i k / o_e(q)}} + o_{X \to \infty}(1).
$$

To simplify this, define

$$
f_{\ell}(X) := \sum_{k=0}^{\ell-1} \frac{e^{2\pi i kX/\ell}}{1 - q^{-1/2} e^{-2\pi i k/\ell}}
$$

for any positive integer ℓ . Then,

$$
f_{\ell}(X) = \sum_{k=0}^{\ell-1} e^{2\pi i kX/\ell} \sum_{m=0}^{\infty} q^{-m/2} e^{-2\pi i k m/\ell} = \sum_{m=0}^{\infty} q^{-m/2} \sum_{k=0}^{\ell-1} e^{2\pi i k(X-m)/\ell}
$$

=
$$
\sum_{m=0}^{\infty} q^{-m/2} \begin{cases} \ell & \text{if } X \equiv m \bmod \ell \\ 0 & \text{otherwise} \end{cases} = \ell \sum_{j=0}^{\infty} q^{-((X \bmod \ell) + j\ell)/2} = \frac{\ell q^{-(X \bmod \ell)/2}}{1 - q^{-\ell/2}}.
$$

Now, we use t[h](#page-0-0)is with $\ell = o_e(q)$ in (43) and finish the proof.

As a first consequence we deduce the existence and a nontrivial lower bound for $\delta(E_d)$ which is valid in general. We also get a conditional upper bound for $\overline{\delta}(E_d)$ that will be used to prove Proposition 3.8. It will be useful to consider the periodic part of $T_d(X)$, so we define

 \Box

(44)
$$
T_d^{\text{per}}(X) := -c_{\pm}(X) + \frac{\epsilon_d}{1-q^{-\frac{1}{2}}} + \sum_{\substack{e|d\\e\nmid 6}} \phi(e) \frac{q^{-(X \mod o_e(q))/2}}{1-q^{-o_e(q)/2}}.
$$

REMARK 3.3. – Proposition 3.2 asserts that $T_d^{\text{per}}(X) = T_d(X) + o_{X \to \infty}(1)$. In particular if $T_d^{\text{per}}(X)$ does not vanish, then we can conclude that the density $\delta(E_d)$ exists. Otherwise the $o(1)$ term contributes to the sign of $T_d(X)$ in a way we do not control and we cannot deduce that $\delta(E_d) = \overline{\delta}(E_d)$.

COROLLARY 3.4. – *Assuming that* $d \geq 7$ *divides* $p^{n} + 1$ *for some n, the function* $T_d^{\text{per}}(X)$ *is* 2n*-periodic. Furthermore, one has*

$$
\underline{\delta}(E_d) \ge \frac{1}{2n},
$$

and, provided there exists some $X_0 \in \mathbb{N}$ such that $T_d^{\text{per}}(X_0) < 0$, one also has

$$
\overline{\delta}(E_d)\leq 1-\frac{1}{2n}.
$$

Proof. – Let us prove the second part of the statement. Let $k \ge 1$ be the integer such that $q = p^k$. Note first that the expression

$$
\sum_{\substack{e|d\\e\nmid 6}} \phi(e) \frac{q^{-(X \bmod o_e(q))/2}}{1-q^{-o_e(q)/2}}
$$

is 2n-periodic as a function of X. This follows since $p^{2n} \equiv 1 \mod d$, which implies that $q^{2n} = p^{2kn} \equiv 1 \mod e$ for every $e \mid d$, and hence $o_e(q) \mid 2n$.

If $X \equiv 0 \mod 2n$, then Proposition 3.2 implies that

$$
T_d(X) = -c_{\pm}(X) + \frac{\epsilon_d}{1 - q^{-\frac{1}{2}}} + \sum_{\substack{e \mid d \\ e \nmid 6}} \phi(e) \frac{1}{1 - q^{-o_e(q)/2}} + o_{X \to \infty}(1),
$$

a quantity which is positive for X large enough, since $\phi(d)$ is even for $d > 3$ and thus

$$
\frac{2}{1 - q^{-o_e(q)/2}} > c_{\pm}(X).
$$

Finally the upper bound is a trivial consequence of the 2n-periodicity of $T_d^{\text{per}}(X)$ and of the existence of some X_0 such that $T_d^{\text{per}}(X_0) < 0$, co[mbin](#page-0-0)ed with the fact that $T_d(X) = T_d^{\text{per}}(X) + o_{X \to \infty}(1).$ $T_d(X) = T_d^{\text{per}}(X) + o_{X \to \infty}(1).$ $T_d(X) = T_d^{\text{per}}(X) + o_{X \to \infty}(1).$ \Box

3.1. Cases of extreme bias for Ulmer's family: proof of Theorem 1.5.

Proof of Theorem 1.5 (i). – If d and q satisfy one of the two stated assumptions, then $\epsilon_d \geq 1$. Then, the statement easily follows because

$$
-c_{\pm}(X) + \frac{\epsilon_d}{1-q^{-1/2}} \ge -\frac{1}{1-q^{-1}} + \frac{1}{1-q^{-\frac{1}{2}}} = \frac{q^{-\frac{1}{2}}-q^{-1}}{(1-q^{-1})(1-q^{-\frac{1}{2}})} > 0,
$$

 \Box

thus using Proposition 3.2 we see that $T_d(X) > 0$ for all large enough X.

Proof of Theorem 1.5 (ii). – By Proposition 3.2 and by positivity, we have (recall that p is large enough, and therefore so is d)

$$
T_d(X) = -c_{\pm}(X) + \frac{\epsilon_d}{1 - q^{-\frac{1}{2}}} + \sum_{\substack{e|d\\e\nmid b}} \phi(e) \frac{q^{-(X \mod o_e(q))/2}}{1 - q^{-o_e(q)/2}} + o_{X \to \infty}(1)
$$

$$
\geq \phi(d)q^{-(X \mod o_d(q))/2} - 1 + o_{p \to \infty}(1) + o_{X \to \infty}(1)
$$

$$
\geq \phi(d)q^{-(o_d(q)-1)/2} - 1 + o_{p \to \infty}(1) + o_{X \to \infty}(1).
$$

However, we have that $q^{2n/k} \equiv 1 \mod d$, that is $o_d(q) | 2n/k$. We conclude that

$$
T_d(X) \ge \phi(d)q^{\frac{1}{2}}q^{-\frac{n}{k}} - 1 + o_{p \to \infty}(1) + o_{X \to \infty}(1)
$$

= $\phi(d)q^{\frac{1}{2}}(d-1)^{-1} - 1 + o_{p \to \infty}(1) + o_{X \to \infty}(1).$

This quantity is positive for large enough X since, for d large enough, we have the inequality $\phi(d)/(d-1) \geq (e^{-\gamma} + o(1))/\log \log d$ and the condition on n implies that $\log \log (p^n + 1) \leq q^{\frac{1}{2}}/2 + \log \log p + 1.$ \Box

Our aim is now to prove Theorem 1.5 (iii). As a preliminary result we give examples of very biased races, for which the lower bound of Corollary 3.4 is essentially sharp. We will then show using a result of Goldfeld that there exist arbitrarily large values of p and d satisfying the conditions of the statement (i) of the proposition.

PROPOSITION 3.5. – (i) Let $d \ge 7$ and $p \ge 3$ be two primes such that p is a primitive root *modulo d. Selecting* $q = p^{\frac{d-1}{2}+1}$, we have that $T_d(X)$ is quite biased towards negative *values; precisely*

$$
\frac{1}{d-1} \le \underline{\delta}(E_d) \le \overline{\delta}(E_d) \le \frac{4}{d-1}.
$$

(ii) *Assume that* $d = p^n + 1 = 2\ell$ *with* $\ell \ge 7$ *a prime,* $p \equiv 3 \mod 4$ *and* $n \ge 4$ *an even number. Pick* $q = p^{n-1}$ *. Then* $T_d(X)$ *is quite biased towards negative values; precisely*

$$
\frac{1}{2n} \le \underline{\delta}(E_d) \le \overline{\delta}(E_d) \le \frac{2}{n}.
$$

Proof. – (i) We first see that the conditions of Ulmer's construction are s[atisfi](#page-0-0)ed. This is clear since $p^{\frac{d-1}{2}} \equiv -1 \mod d$, hence $d \mid p^n + 1$ with $n = (d-1)/2$.

Note also that since $d \geq 5$ is prime, $\epsilon_d = 0$. Moreover, $o_d(q) \in \{(d-1)/2, d-1\}$. This is clear since $q \equiv -p \mod{d}$, and p is a primitive root modulo d. Proposition 3.2 then takes the form

$$
T_d(X) = -c_{\pm}(X) + \phi(d) \frac{q^{-(X \mod o_d(q))/2}}{1 - q^{-o_d(q)/2}} + o_{X \to \infty}(1).
$$

If $X \equiv j \mod (d-1)$ with $j \notin \{0, 1, (d-1)/2, (d-1)/2 + 1\}$, then

$$
T_d(X) \le -q^{-\frac{1}{2}} + 2(d-1)q^{-2/2} + o_{X \to \infty}(1) = -q^{-\frac{1}{2}} \left(1 - \frac{2(d-1)}{p^{\frac{d+1}{4}}} \right) + o_{X \to \infty}(1),
$$

a quantity which is negative for X large enough since for $d \ge 7$ and $p \ge 3$ we have $4(d-1) < p^{\frac{d+1}{4}}$. We have thus shown that $T_d(X)$ is negative for most of the values of X mod $(d-1)$.

Combining this with Corollary 3.4 we conclude that

$$
\frac{1}{d-1} \le \underline{\delta}(E_d) \le \overline{\delta}(E_d) \le \frac{4}{d-1}.
$$

(ii) First note that the given choice of parameters ensures that $\epsilon_d = 0$. Proposition 3.2 yields the formula

$$
T_d(X) = -c_{\pm}(X) + \phi(\ell) \frac{q^{-(X \mod o_{\ell}(q))/2}}{1 - q^{-o_{\ell}(q)/2}} + \phi(2\ell) \frac{q^{-(X \mod o_{2\ell}(q))/2}}{1 - q^{-o_{2\ell}(q)/2}} + o_{X \to \infty}(1).
$$

We have $q^n \equiv (-p^{-1})^n \equiv -1 \mod d$. We claim that n is the least positive integer such that this congruence holds. Indeed this minimality condition holds by definition for the congruence $p^n \equiv -1 \mod d$. Now $q \equiv -p^{-1} \mod d$ and n is even, thus the claim follows. Since $(\mathbb{Z}/d\mathbb{Z})^{\times}$ is cyclic (it is isomorphic to $(\mathbb{Z}/l\mathbb{Z})^{\times}$), this implies that $o_d(q) = o_{2\ell}(q) = 2n$. Now consider an integer j such that $4 \le j \le 2n - 1$. We have that

$$
\phi(2\ell) \frac{q^{-(j \bmod o_{2\ell}(q))/2}}{1-q^{-o_{2\ell}(q)/2}} \ll \ell q^{-4/2} \ll p^{n-2(n-1)} = p^{2-n}.
$$

Hence, since $-c_{\pm}(j) < 0$ and $c_{\pm}(j) \gg q^{-\frac{1}{2}} = p^{(1-n)/2}$, we have for p and X large enough and $n \geq 4$ that $T_d(X) < 0$ as soon as $X \equiv j \mod 2n$. Hen[ce,](#page-0-0)

$$
0 \leq \underline{\delta}(E_d) \leq \overline{\delta}(E_d) \leq \frac{2}{n}.
$$

As for the lower bound, it is given once more by Corollary 3.4.

We will now see that there are infinitely many choices of primes p and d in (i) of the proposition (in which both c[an](#page-39-13) be arbitrarily large) for which p is a primitive root modulo d . This is a corollary of the following result of Goldfeld.

THEOREM 3.6 (Goldfeld [12]). $-$ Let $x \geq 2$ be a real number and let $\text{Li}(x) := \int_2^x (\log t)^{-1} dt$. *Let* $1 < A \leq x$ *. Then for each* $D \geq 1$ *,*

(45)
$$
N_a(x) := \#\{p \le x : a \text{ is a primitive root mod } p\} = c\text{Li}(x) + O_D\left(\frac{x}{(\log x)^D}\right),
$$

for all $a \leq A$ *with at most* $c(D)A^{\frac{9}{10}}(5\log x + 1)^{D+2+\frac{\log x}{\log A}}$ exceptions, where c is Artin's *Constant:*

$$
c := \prod_{p} \left(1 - \frac{1}{p(p-1)} \right).
$$

COROLLARY 3.7. – *We have for any fixed* $C \geq 1$ *the estimate*

$$
\sum_{\substack{p,d \le x \text{ both prime} \\ p \text{ is a primitive root mod } d}} 1 = c \text{Li}(x)^2 \left(1 + O_C\left(\frac{1}{(\log x)^C} \right) \right),
$$

and as a consequence,

$$
\#\{p, d \in (x, 2x] \text{ both prime : } p \text{ is a primitive root mod } d\} \sim \frac{3cx^2}{(\log x)^2}
$$

.

 \Box

Proof. – Pick $A = x$ and $D = C$ in Goldfeld's Theorem, and let $\mathcal{C}(x)$ be the set of $a \leq x$ for which (45) does not hold, thus $|\mathcal{U}(x)| \ll_C x^{\frac{9}{10}}(\log x)^{C+3}$. We then write

$$
\sum_{\substack{p,d \leq x \text{ both prime} \\ p \text{ is a primitive root mod } d}} 1 = \sum_{p \leq x} N_p(x) = \sum_{p \leq x} c\text{Li}(x) + \sum_{p \leq x} (N_p(x) - c\text{Li}(x))
$$

$$
= c\text{Li}(x)^2 + O\left(\sum_{p \leq x} |N_p(x) - c\text{Li}(x)|\right) + O\left(\frac{x^2}{(\log x)^C}\right),
$$

by the Prime Number Theorem. The first error term is bounded as follows:

$$
\sum_{p \le x} |N_p(x) - c\text{Li}(x)| = \sum_{\substack{p \le x \\ p \in \mathcal{A}(x)}} |N_p(x) - c\text{Li}(x)| + \sum_{\substack{p \le x \\ p \notin \mathcal{A}(x)}} |N_p(x) - c\text{Li}(x)|
$$

$$
\ll_C |\mathcal{C}(x)|\pi(x) + \frac{x^2}{(\log x)^{C+1}} \ll_C \frac{x^2}{(\log x)^C}.
$$

Proof of Theorem 1.5 (iii). – This is a direct consequence of Proposition 3.5 (i) and Corollary 3.7. Note that the curves E_d in Proposition 3.5 have [rank](#page-0-0) either one or two (depending on the parity of $(d-1)/2$). Indeed the proof of Proposition 3.5 shows that the multiplicative order of q modulo d is $d-1$ ([resp](#page-0-0). $(d-1)/2$) if $(d-1)/2$ is even (resp. odd), and we get the corresponding value of the rank by applying Proposition 3.1. \Box

3.2. Cases of moderate bias for Ulmer's family: proof of Theorem 1.7.

We start this section by proving a first result about moderate bias. We will later specialize in order to obtain races which are not biased at all.

PROPOSITION 3.8. – *Fix* $n \geq 2$ *even,* $p \equiv 3 \mod 4$ *and* $k \geq 1$ *. Pick* $q = p^{kn+1}$ *and* $d = pⁿ + 1$. Then for n fixed and p large enough, $T_d(X)$ is m[oder](#page-0-0)ately biased, that is

$$
\frac{1}{2n} \le \underline{\delta}(E_d) \le \overline{\delta}(E_d) \le 1 - \frac{1}{2n}.
$$

REMARK 3.9. – Note that the bounds of Proposition 3.5 (ii) are much more precise. However, although we believe that there are infinitely many curves satisfying the hypotheses of Proposition 3.5 (ii), [this](#page-0-0) seems hard to prove given the restrictive arithmetic conditions that d has t[o sat](#page-0-0)isfy.

Proof of Proposition 3.8. – Since the given choice of parameters ensures that $\epsilon_d = 0$, Proposition 3.2 takes the form

$$
T_d(X) = -c_{\pm}(X) + \sum_{\substack{e|d\\e\nmid 6}} \phi(e) \frac{q^{-(X \mod o_e(q))/2}}{1 - q^{-o_e(q)/2}} + o_{X \to \infty}(1).
$$

We now show that for each e | d with e $\nmid 6$, we have $o_e(q) \geq 3$. Note first that $(p-1, e)$ and $(p+1, e)$ both divide 2. Indeed we have obviously $(p-1, e)$ | $(p-1, d)$ and $(p+1, e)$ | $(p+1, d)$. We compute

$$
(p-1, d) = (p-1, pn + 1) = (p-1, pn-1 + 1) = \dots = (p-1, p+1) = 2;
$$

\n
$$
(p+1, d) = (p+1, pn + 1) = (p+1, pn-1 - 1) = (p+1, pn-2 + 1) = \dots = (p+1, p-1) = 2,
$$

since n is even. Note that $e \nmid 6$ and $e \neq 4$, the latter being a consequence of $e \mid d$ and $d \equiv 2 \mod 4$. We conclude from the above computation that $e \nmid (p + 1)(p - 1)$, that is $o_e(p) \geq 3$. Since $q \equiv \pm p \mod e$ and $p \not\equiv \pm 1 \mod e$, we also have that $o_e(q) \geq 3$.

Using this fact, we have that if $X \equiv 2 \mod 2n$, then

$$
T_d(X) = -c_{\pm}(X) + \sum_{\substack{e|d\\e\nmid 6}} \phi(e) \frac{q^{-2/2}}{1 - q^{-o_e(q)/2}} + o_{X \to \infty}(1).
$$

This last quantity is negative for large enough $X \equiv 2 \mod 2n$, since $c_{\pm}(X) = 1 + o_{p \to \infty}(1)$, and

$$
\sum_{\substack{e|d\\e\nmid 6}} \phi(e) \frac{q^{-2/2}}{1 - q^{-o_e(q)/2}} \ll q^{-1}d \ll p^{n - (kn+1)} \ll p^{-1},
$$

which is negligible compared to $c_{\pm}(X)$. We conclude by invoking Corollary 3.4.

 \Box

Proof of Theorem [1.7](#page-0-0). – Firstly, $d | p^2 + 1 \equiv 2 \mod 3$ if $p \neq 3$ (otherwise $p^2 + 1 \equiv 1 \mod 3$) and $q \equiv p \equiv 3 \mod 4$, hence $\epsilon_d = 0$. Note also that if $e \mid d$ with $e \notin \{1, 2\}$, then $q^2 \equiv p^2 \equiv -1 \mod e$, hence $o_e(q) = 4$. The rank of E_d is then easily computed with the help of Proposition 3.1. Moreover, Proposition 3.2 becomes

$$
T_d(X) = -c_{\pm}(X) + \sum_{\substack{e|d\\e \neq 1,2}} \phi(e) \frac{q^{-(X \mod 4)/2}}{1 - q^{-2}} + o_{X \to \infty}(1),
$$

hence $T_d^{\text{per}}(X) = -c_{\pm}(X) + \sum_{e|d, e \neq 1,2} \phi(e) q^{-(X \mod 4)/2} (1 - q^{-2})^{-1}$ is 4-periodic. If $X \equiv 0 \mod 4$, then

$$
T_d(X) = -\frac{q}{q-1} + \sum_{\substack{e|d\\e \neq 1,2}} \frac{\phi(e)}{1-q^{-2}} + o_{X \to \infty}(1) \ge -2 + \frac{d-2}{1-q^{-2}} + o_{X \to \infty}(1),
$$

which is positive for X large enough.

If $X \equiv 1 \mod 4$, then

$$
T_d(X) = -\frac{q^{\frac{1}{2}}}{q-1} + \sum_{\substack{e|d\\e \neq 1,2}} \frac{\phi(e)q^{-\frac{1}{2}}}{1-q^{-2}} + o_{X \to \infty}(1) \ge q^{-\frac{1}{2}} \left(\frac{d-2}{1-q^{-2}} - 2\right) + o_{X \to \infty}(1),
$$

which is again positive for X large enough.

As for $X \equiv 2 \mod 4$, we have

$$
T_d(X) = -\frac{q}{q-1} + \sum_{\substack{e \mid d \\ e \neq 1,2}} \frac{\phi(e)q^{-1}}{1-q^{-2}} + o_{X \to \infty}(1) \le -1 + 2dq^{-1} + o_{X \to \infty}(1),
$$

which is negative for X large enough since $2d \leq 2(p^2 + 1) < p^5/2 \leq q/2$.

Finally, for $X \equiv 3 \mod 4$ we have

$$
T_d(X) = -\frac{q^{\frac{1}{2}}}{q-1} + \sum_{\substack{e|d\\e \neq 1,2}} \frac{\phi(e)q^{-\frac{3}{2}}}{1-q^{-2}} + o_{X \to \infty}(1) \leq q^{-\frac{1}{2}} \left(-1 + 2dq^{-1}\right) + o_{X \to \infty}(1),
$$

which is negative for X large enough. Since $T_d(X)$ is positive for asymptoticall[y ha](#page-0-0)lf of the values of X, we conclude that $\underline{\delta}(E_d) = \overline{\delta}(E_d) = \frac{1}{2}$. \Box

3.3. On the closure of the set of lower (resp. upper) densities: proof of Theorem 1.8

We first need the following preliminary result.

PROPOSITION 3.10. – Let $p \ge 17$ and $n \ge 3$ be primes such that $n \nmid p+1$, and set $d = (p^{n} + 1)/(p + 1)$ and $q = p^{k}$ with $k \ge 4$ even and coprime to n. Then the curve E_{d} *has rank exactly* (d − 1)/n*, and we have that*

$$
\frac{1}{k} - \frac{2}{nk} \le \underline{\delta}(E_d) \le \overline{\delta}(E_d) \le \frac{1}{k} + \frac{1}{2n}.
$$

Proof. – We first see that $\epsilon_d = 0$. Indeed we have that

(46)
$$
d = 1 - p + p^2 - p^3 + \dots + p^{n-1} \equiv n \equiv 1 \mod 2.
$$

Moreover, if $p \equiv 1 \mod 3$, then $d \equiv 2/2 \equiv 1 \mod 3$. If $p \equiv 2 \mod 3$, then we have

$$
d = 1 - p + p2 - p3 + \dots + pn-1 \equiv 1 - 2 + 1 - 2 + \dots + 1 \equiv n \not\equiv 0 \mod 3,
$$

hence $6 \nmid d$ and $\epsilon_d = 0$.

We will now show that for every $e \mid d$ with $e \nmid 6$, $o_e(q) = n$. First, $p^n \equiv -1 \mod e$, so $o_e(p) \in \{2, 2n\}$ (indeed $-1 \not\equiv 1 \mod e$ since $e \nmid 6$). Note that $(e, p + 1) = 1$; indeed

$$
(d, p+1) = (p^{n-1} - p^{n-2} + p^{n-3} - \dots + 1, p+1) = (-2p^{n-2} + p^{n-3} - \dots + 1, p+1)
$$

= $\dots = (- (n-1)p + 1, p+1) = (n, p+1) = 1,$

since *n* is prime and $n \nmid p + 1$. We also [hav](#page-30-0)e

$$
(pn + 1, p - 1) = (pn-1 + 1, p - 1) = \dots = (p + 1, p - 1) = 2,
$$

hence $(d, p - 1) = 1$ $(d, p - 1) = 1$ (since d is odd by (46)). We conclude that $(e, (p + 1)(p - 1)) = 1$, and so since $e > 1$ (recall that $e \nmid 6$), $p^2 \not\equiv 1 \mod e$ and thus $o_e(p) = 2n$. Since $(k, 2n) = 2$, it follows that $o_e(q) = n$. In particular the statement about the rank of E_d is easily deduced from Proposition 3.1.

We now turn to the study of $T_d(X)$. We have

$$
T_d(X) = -c_{\pm}(X) + \sum_{\substack{e|d\\e\notin 6}} \phi(e) \frac{q^{-(X \mod n)/2}}{1 - q^{-n/2}} + o_{X \to \infty}(1)
$$

=
$$
-c_{\pm}(X) + (d-1) \frac{q^{-(X \mod n)/2}}{1 - q^{-n/2}} + o_{X \to \infty}(1).
$$

If $X \equiv j \mod 2n$ with $0 \le j \le 2(n-2)/k$, then

$$
T_d(X) \ge -2 + (d-1)q^{-\frac{n-2}{k}} + o_{X \to \infty}(1) \ge -2 + \frac{p}{2} + o_{X \to \infty}(1),
$$

which is positive for large enough X .

If $X \equiv j \mod 2n$ with $(2n-1)/k + 1 \le j \le 2n - 1$, then

$$
T_d(X) \le -\frac{q^{-\frac{1}{2}}}{1-q^{-1}} + 2(d-1)q^{-\frac{n-1/2}{k} - \frac{1}{2}} + o_{X \to \infty}(1) \le -\frac{1}{q^{\frac{1}{2}}} + \frac{4}{(pq)^{\frac{1}{2}}} + o_{X \to \infty}(1),
$$

which is negative for large enough X, since we have assumed $p \geq 17$.

We conclude that

$$
\frac{1}{k} - \frac{2}{nk} \le \underline{\delta}(E_d) \le \overline{\delta}(E_d) \le \frac{1}{k} + \frac{1}{2n} - \frac{1}{2nk}.
$$

Proof of Theorem 1.8. – From Theorems 1.5 and 1.7, we already know that $\{0, \frac{1}{2}, 1\} \subset \overline{S}$. Fix $m \ge 2$ and $\epsilon > 0$. Let $p > \min(17, 2\epsilon^{-1})$, and pick a prime $n > p+1$, so that $n \nmid p+1$. Letting $k = 2m$ in Proposition 3.10, we find a curve E_d such that

$$
\frac{1}{2m} - \epsilon \leq \underline{\delta}(E_d) \leq \overline{\delta}(E_d) \leq \frac{1}{2m} + \epsilon.
$$

The theorem follows since ϵ is arbitrary.

The proof of Proposition 3.10 shows in fact that under the stated assumptions, $2n$ is the smallest period of the periodic part $T_d^{\text{per}}(X)$ of $T_d(X)$.

4. Central Limit Theorem

The goal of this section is to pro[ve a](#page-12-1) Cen[tral](#page-21-0) Limit Theorem, in particular we will prove Theorems 1.2 (ii) and 1.3 .

DEFINITION 4.1. – Recall from (10) and (38) that we denote the inverse zeros of $L(E/K,T)$ by $\{\gamma_j = q e^{i\theta_j}\}\$ with $j = 1, ..., N_{E/K}$. Let $\epsilon(E/K) = \pm 1$ be the sign of the functional equation f[or](#page-40-1) E/K , that is, the unique number satisfying

(47)
$$
L(E/K,T) = \epsilon(E/K)(qT)^{N_{E/K}} L(E/K, 1/(q^2T)).
$$

(See, e.g., $[22, Th. 2.2.1]$.) Define the set of *forced zeros* of E/K to be

(48)
$$
\text{FZ}(E/K) := \begin{cases} \{ \epsilon(E/K)q \} & \text{if } N_{E/K} \text{ is odd,} \\ \{q, -q \} & \text{if } N_{E/K} \text{ is even and } \epsilon(E/K) = -1, \\ \{ \} & \text{otherwise.} \end{cases}
$$

We will denote by $\{\theta_1, \ldots, \theta_k\}$ the multiset of angles in $[0, \pi]$ which do not come from forced zeros of $L(E/K, T)$. We will say that an elliptic curve E over K satisfies the *linear independence* (LI) hypothesis if the multiset

$$
\{\theta_j/\pi \colon 0 < \theta_j \le \pi, \, j \in \{1, \dots, k\} \} \cup \{1\}
$$

is linearly independent over Q.

The idea behind this definition is quite simple. When we say that E/K satisfies LI, we are willing to [ig](#page-39-8)nore any trivial multiplicative relation among the zeros, namely, the relations coming from complex conjugation and forced zeros $\gamma = \pm q$, as well as those $\gamma = q$ arising from the vanishing of $L(E/K, T)$ at the central point, or, a positive analytic rank of E/K . In our work [5] we prove a stronger statement involving the possible multiplic[ati](#page-39-8)ve relations among inverse zeros of *reduced* L-functions, i.e., the Q-polynomial obtained by quotienting $L(E/K, T)$ $L(E/K, T)$ $L(E/K, T)$ by the product of linear factors corresponding to $FZ(E/K)$. (Note that any forced zero $\gamma = \pm q$ would cause linear dependence over \mathbb{Q} .) Indeed, our work [5] shows that generically not only does LI hold, but also the rank is at most one.

Before we prove a Central Limit Theorem which implies Theorem 1.2 (ii), we need several lemmas.

 $LEMMA 4.2. - We have that$

$$
\frac{L'}{L}(E/K, q^{-3/2}) = O_C(\log_q N_{E/K}).
$$

Proof. – This is a function field analogue of Littlewood's bound $(L'/L)(1, \chi) = O(\log \log q)$ for Dirichlet characters χ modulo q. Let $\mathcal D$ be an effective divisor on the curve C (recall that C/\mathbb{F}_q is the curve whose function field is by definition K), i.e., $\mathcal{D} = \sum_{v} n_v \cdot v$ where n_v are nonnegative integers, $n_v = 0$ for all but finitely many v's, and v runs over the set of places of K. Using the standard definition of the degree of a divisor (for $\mathcal D$ as above,

 $deg \mathcal{D} = \sum_{v} n_v deg v$, we let $|\mathcal{D}| := q^{\deg(\mathcal{D})}$. Also, we define the *von Mangoldt function on* C by

$$
\Lambda_C(\mathcal{D}) = \begin{cases} \deg v & \text{if } \mathcal{D} = n_v \cdot v \text{ for some } v, \\ 0 & \text{otherwise.} \end{cases}
$$

Define

$$
a_{n_v \cdot v}(E) = \alpha_v^{n_v} + \beta_v^{n_v},
$$

with notation as in §2.1, and further $\alpha_v = a_v(E)$, $\beta_v = 0$ if v is a place of bad reduction of E/K . We claim that:

(49)
$$
- \frac{L'}{L}(E/K, q^{-3/2}) = \sum_{\substack{\mathcal{D}: \ \mathcal{D}=n_v \cdot v \\ n_v \ge 0, v \text{ place of } K}} \frac{\Lambda_C(\mathcal{D}) a_{\mathcal{D}}(E)}{|\mathcal{D}|^{3/2}}.
$$

This follows from our previous calculation (13), (16) and (17) with $m = 1$.

We rewrite the right hand side of (49) as follows

(50)
$$
\sum_{n\geq 1} \sum_{v \text{ place of } K} \frac{\deg v}{q^{n \deg v}} \cdot \frac{a_{n \cdot v}(E)}{q^{(n \deg v)/2}}.
$$

The sum of the terms with $n \geq 2$ is

$$
\ll \sum_{n\geq 2}\sum_{v \text{ place of }K} \frac{\deg v}{q^{n\deg v}}=\sum_{d\geq 1}\sum_{\substack{v \text{ place of }K \\ \deg v=d}} \frac{d}{q^{2d}}\cdot \frac{1}{1-q^{-d}} \ll_C \sum_{d\geq 1} q^{-d}=\frac{1}{q-1}\,,
$$

by (22). Next we split the rest of the sum (50) as $S_1 + S_2$ where

$$
S_1 := \sum_{d \leq 2 \log_q N_{E/K}} \frac{d}{q^d} \sum_{\substack{v \text{ place of } K \\ \deg v = d}} \frac{a_v(E)}{q^{\deg v/2}}, \qquad S_2 := \sum_{d > 2 \log_q N_{E/K}} \frac{d}{q^d} \sum_{\substack{v \text{ place of } K \\ \deg v = d}} \frac{a_v(E)}{q^{\deg v/2}}.
$$

Applying again (22) w[e triv](#page-0-0)ially have

$$
S_1 \ll_C \log_q N_{E/K} .
$$

Finally, from Theorem 2.3 we deduce that

$$
S_2 \ll \sum_{d>2 \log_q N_{E/K}} \frac{d}{q^d} (1+\nu_1) \ll N_{E/K} \int_{2 \log_q N_{E/K}}^{\infty} \frac{t dt}{q^t} \ll N_{E/K} \frac{\log_q N_{E/K}}{q^{2 \log_q N_{E/K}}} \ll 1. \quad \Box
$$

The following sum over the angles θ_i which do not come from forced zeros gives an estimate for the variance of the random variable X_E .

LEMMA 4.3. – *We have the estimate*

$$
I_E := \sum_{j=1}^k \left| \frac{2}{1 - q^{-\frac{1}{2}} e^{-i\theta_j}} \right|^2 = \frac{2q}{q-1} N_{E/K} + O\left(\log N_{E/K} + \text{rank}(E/K)\right).
$$

Proof. – The quantity I_E can be computed by evaluating the log derivative of (56) at $T = q^{-1/2}$ (see Page 1370 in [4] for a similar calculation). We see that

(51)
$$
I_E = \frac{4q}{q-1} \left(q^{-1/2} \frac{L'}{L} (E/K, q^{-1/2}) - \frac{\sqrt{q}}{\sqrt{q}-1} \operatorname{rank}(E/K) - k \right).
$$

The functional equation (47) yields

(52)
$$
q^{-1/2} \frac{L'}{L} (E/K, q^{-1/2}) = N_{E/K} - q^{-3/2} \frac{L'}{L} (E/K, q^{-3/2}).
$$

This gives in turn

(53)
$$
I_E = \frac{2q}{q-1} N_{E/K} + \frac{4q}{q-1} \left(\frac{1}{2} - \frac{\sqrt{q}}{\sqrt{q}-1} \right) r(E/K) - \frac{4\sqrt{q}}{q-1} \frac{L'}{L} (E/K, q^{-3/2}).
$$

The desired estimate follows from an application of Lemma 4.2.

$$
\Box
$$

The following lemma is an application of the Berry-Esseen inequality.

LEMMA 4.4. – *Fix two parameters* $0 < \epsilon < 1$ and $1 \leq M \leq \epsilon^{-\frac{1}{2}}$, and assume that the *characteristic function* $\widehat{X}(\xi)$ *of the random variable* X *satisfies the following properties:*

- (i) $|\hat{X}(\xi)| \le 10\xi^{-4}$ *for* $|\xi| \ge \epsilon^{-\frac{1}{4}}$ *, and*
- (ii) $\log \hat{X}(\xi) + \frac{\xi^2}{2}$ $|\frac{\xi^2}{2}| \leq 10\epsilon(M\xi^2 + \xi^4)$ for $|\xi| \leq \epsilon^{-\frac{1}{4}}$.

Then, the distribution function $F_X(x)$ *of* X *satisfies*

$$
\sup_{x \in \mathbb{R}} |F_X(x) - G(x)| \ll M\epsilon,
$$

*where*G(x)*is the distribution function of the Gaussian an[d th](#page-39-14)e implied constant is absolute. (The constant* 10 *appearing both in* (i) *and* (ii) *can be replaced by any positive absolute constant.)*

Proof. – We will apply the Berry-Esseen inequality [7, Theorem 2a]

(54)
$$
\sup_{x \in \mathbb{R}} |F_X(x) - G(x)| \ll \int_{-T}^{T} \frac{\hat{X}(\xi) - e^{-\frac{\xi^2}{2}}}{\xi} d\xi + \frac{1}{T}.
$$

Taking $T = \epsilon^{-1}$ and applying the hypotheses on $\hat{X}(\xi)$, we note that the integral equals

$$
\int_{|\xi| \le \epsilon^{-\frac{1}{4}}} e^{-\xi^2/2} \frac{e^{O(\epsilon(M\xi^2 + \xi^4))} - 1}{\xi} d\xi + O\left(\int_{\epsilon^{-\frac{1}{4}} \le |\xi| \le \epsilon^{-1}} \left|\frac{\xi^{-4}}{\xi}\right| d\xi\right)
$$

\$\ll \int_{|\xi| \le \epsilon^{-\frac{1}{4}}} e^{-\xi^2/2} \frac{\epsilon(M\xi^2 + \xi^4)}{|\xi|} d\xi + \epsilon\$
\$\le \epsilon \int_{\mathbb{R}} |M\xi + \xi^3| e^{-\xi^2/2} d\xi + \epsilon \ll M\epsilon\$,

with absolute implied constants. The result follows.

THEOREM 4.5. – *Suppose that* ${E/K}$ *is a family of elliptic curves satisfying LI such that*

$$
\frac{\text{rank}(E/K)}{\sqrt{N_{E/K}}}\to 0
$$

as $N_{E/K}$ → ∞*. Let* X_E *be the random variable associated to the limiting distribution of* $T_E(X)$ *(see Corollary 2.9 in the case* $V = -U_1$ *), and define the normalized random variable* $Y_E := \sqrt{\frac{q-1}{q}} X_E/\sqrt{N_{E/K}}.$ Then, Y_E converges in distribution to the standard Gaussian. More *precisely, the distribution function* F_E *of* Y_E *satisfies*

$$
\sup_{x \in \mathbb{R}} |F_E(x) - G(x)| \ll \frac{\operatorname{rank}(E/K) + 1}{\sqrt{N_{E/K}}},
$$

where G *denotes the distribution function of the Gaussian.*

Proof. – Define $m(E/K, -q)$ to be the multiplicity of the inverse zero $\gamma = -q$ in $L(E/K, T)$. Then, we can write

(56)
$$
L(E/K,T) = (1+qT)^{m(E/K,-q)}(1-qT)^{\text{rank}(E/K)} \prod_{j=1}^{k} [(1-\gamma_j T)(1-\overline{\gamma}_j T)],
$$

with $\gamma_j = q e^{i\theta_j}$ $\gamma_j = q e^{i\theta_j}$ $\gamma_j = q e^{i\theta_j}$ and $0 < \theta_j < \pi$ for $j = 1, ..., k$. We have

(57)
$$
N_{E/K} = m(E/K, -q) + \text{rank}(E/K) + 2k.
$$

From (56) and Corollary 2.10, it is easy to deduce that

(58)
$$
T_E(X) = \frac{\sqrt{q}}{\sqrt{q} - 1} \text{rank}(E/K) - c_{\pm}(X)
$$

$$
-(-1)^X m(E/K, -q) \frac{\sqrt{q}}{\sqrt{q} + 1} + 2 \sum_{j=1}^k \Re\left(\frac{e^{j\theta_j X}}{1 - q^{-\frac{1}{2}} e^{-j\theta_j}}\right) + o_{X \to \infty}(1).
$$

From now on, we will assume that E/K satisfies LI. This implies in particular that

$$
m(E/K, -q) = \begin{cases} 1 & \text{if } -q \in \text{FZ}(E/K), \\ 0 & \text{if } -q \notin \text{FZ}(E/K). \end{cases}
$$

Define $\mathscr{B}_E(\xi)$ by

(59)
$$
\mathscr{B}_E(\xi) := \frac{1}{2} \left[\exp(iB_0 \xi) + \exp(iB_1 \xi) \right]
$$

with

$$
B_0 := \frac{\sqrt{q}}{\sqrt{q} - 1} \text{rank}(E/K) - \frac{q}{q - 1} + m(E/K, -q) \frac{\sqrt{q}}{\sqrt{q} + 1}
$$

and

$$
B_1 := \frac{\sqrt{q}}{\sqrt{q}-1} \text{rank}(E/K) - \frac{\sqrt{q}}{q-1} - m(E/K, -q) \frac{\sqrt{q}}{\sqrt{q}+1}.
$$

In order to prove Theorem 4.5, we will proceed as in [4, Th. 6.2]. First, we claim that

(60)
$$
\widehat{X}_E(\xi) = \mathcal{B}_E(\xi) \prod_{j=1}^k J_0 \left(\frac{2\xi}{|1 - q^{-\frac{1}{2}} e^{-i\theta_j}|} \right)
$$

This follows from (58) and LI (see [4, Th. 3.4].) Indeed, thanks to (the discrete version of) the Kronecker-Weyl Theorem, the set

.

$$
\{t\cdot (0,1/2,\theta_1/2\pi,\ldots,\theta_k/2\pi) \mid t=1,2,3,\ldots\}\subset \mathbb{R}^{k+2}/\mathbb{Z}^{k+2}
$$

is equidistributed in its topological closure, say, A . Moreover, LI implies that A is simply the union of two k-torii

$$
A = (\{0\} \times \{0\} \times \mathbb{R}^k / \mathbb{Z}^k) \cup (\{0\} \times \{1/2\} \times \mathbb{R}^k / \mathbb{Z}^k) \subset \mathbb{R}^{k+2} / \mathbb{Z}^{k+2}.
$$

This shows that, under LI, our random variable X_E can be written as a sum of certain *independent* random variables. More precisely, we can write $X_E = X_1 + X_2$, where X_1 is a Bernoulli random variable with values B_0 and B_1 , and $X_2 := 2\sum_{j=1}^k \Re(Z_j) / |1 - q^{-\frac{1}{2}} e^{-i\theta_j}|$ with Z_j uniformly distributed on the unit circle in $\mathbb C$ with the random variables $X_1, Z_1, Z_2, \ldots, Z_k$ being mutually independent. Therefore the characteristic function of X_E is given by the product of the characteristic functions of these random variables. In particular, we note

that $\mathbb{E}[\mathrm{e}^{\mathrm{i}tX_1}] = \frac{1}{2} \mathrm{e}^{\mathrm{i}tB_0} + \frac{1}{2} \mathrm{e}^{\mathrm{i}tB_1}$ and we use the identity $\frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{\mathrm{i}t\cos\theta} d\theta = J_0(t)$ to obtain $\mathbb{E}[\mathrm{e}^{\mathrm{i}t\Re(Z_j)}] = J_0(t)$. This proves (60).

It follows that the normalized random variable Y_E satisfies

$$
\log \widehat{Y}_E(\xi) = \log \mathscr{B}_E \left(\sqrt{\frac{q-1}{q}} \frac{\xi}{\sqrt{N_{E/K}}} \right) + \sum_{j=1}^k \log J_0 \left(\left| \frac{2}{1 - q^{-\frac{1}{2}} e^{-i\theta_j}} \right| \sqrt{\frac{q-1}{q}} \frac{\xi}{\sqrt{N_{E/K}}} \right).
$$

The Taylor series of $\log J_0(z)$ has radius of convergence slightly larger than 12/5, since $J_0(z)$ has no zero in this region⁽¹⁰⁾. From the bound

$$
\left|\frac{2}{1-q^{-\frac{1}{2}}{\rm e}^{-{\rm i}\theta_j}}\right|\sqrt{\frac{q-1}{q}}\frac{|\xi|}{\sqrt{N_{E/K}}}\leq \frac{2}{1-2^{-\frac{1}{2}}}\frac{|\xi|}{\sqrt{N_{E/K}}},
$$

we have that the argument of log J_0 in the last equation never exceeds 6.83 $|\xi|/\sqrt{N_{E/K}}$ in absolute value. We then have for $|\xi| \leq .35\sqrt{N_{E/K}}$ that

$$
\log \widehat{Y}_E(\xi) = O\left(|\xi| \frac{\text{rank}(E/K)}{\sqrt{N_{E/K}}}\right) - \frac{q-1}{4q} \frac{\xi^2}{N_{E/K}} \sum_{j=1}^k \left|\frac{2}{1-q^{-\frac{1}{2}}e^{-i\theta_j}}\right|^2 + O\left(\frac{\xi^4}{N_{E/K}}\right).
$$

Given the assumptions of the theorem and Lemma 4.3, it follows that $\log \widehat{Y}_E(\xi) \to -\frac{\xi^2}{2}$ 2 pointwise as $N_{E/K}$ tends to infinity. In light of Lév[y's t](#page-0-0)heorem [17, Section 4.2, Theorem 4], this establishes the first part of the theorem.

For the [seco](#page-0-0)nd part, we write $Y_i := (1 - q^{-1})^{-\frac{1}{2}} X_i / \sqrt{N_{E/K}}$ $(i = 1, 2)$, where the X_i were defined earlier in the proof. We will apply Lemma 4.4 to the random variable Y_2 , with the parameters $\epsilon := 1/N_{E/K}$, $M = c(\log N_{E/K} + \text{rank}(E/K))$, where c is the implied constant in Lemma 4.3. With the help of Lemma 4.3, we see that the characteristic function of Y_2

$$
\widehat{Y}_{2}(\xi) = \prod_{j=1}^{k} J_{0} \left(\left| \frac{2}{1 - q^{-\frac{1}{2}} e^{-i\theta_{j}}} \right| \sqrt{\frac{q-1}{q}} \frac{\xi}{\sqrt{N_{E/K}}} \right)
$$

satisfies the estimate

$$
\log \hat{Y}_2(\xi) = -\frac{\xi^2}{2} + O\left(\frac{\log N_{E/K} + \text{rank}(E/K)}{N_{E/K}}\xi^2 + \frac{\xi^4}{N_{E/K}}\right)
$$

in the range $|\xi| \leq 0.35\sqrt{N_{E/K}}$; hence (ii) of Lemma 4.4 holds for $N_{E/K}$ large enough. Moreover, the bound $\log J_0(x) \leq -x^2/2$ valid for $|x| \leq 12/5$ combined with Lemma 4.3 imply that

$$
\log \widehat{Y}_2(\xi) \le -(1 + o(1)) \frac{\xi^2}{2}
$$

for $|\xi| \leq 0.35\sqrt{N_{E/K}}$, and hence we clearly have $\widehat{Y}_2(\xi) \leq \xi^{-4}$ in this range. To show that (i) of Lemma 4.4 holds, it remains to show that $|\hat{Y}_2(\xi)| \ll \xi^{-4}$ in the range $|\xi| \ge 0.35 \sqrt{N_{E/K}}$.

In the range $|\xi| \ge 5\sqrt{N_{E/K}}$, we use the bound $|J_0(x)| \le \sqrt{2/(\pi|x|)}$ to deduce that

$$
|\widehat{Y}_2(\xi)| \leq \prod_{j=1}^k \left(\frac{1 + q^{-\frac{1}{2}}}{(1 - q^{-1})^{\frac{1}{2}}} \frac{\sqrt{N_{E/K}}}{\pi \xi} \right) \leq \left(\frac{7\sqrt{N_{E/K}}}{\pi \xi} \right)^{k/2},
$$

(10) See for instance http://mathworld.wolfram.com/BesselFunctionZeros.html.

a quantity which is $\leq \xi^{-4}$ as soon as $|\xi| \geq 5\sqrt{N_{E/K}}$ (we used LI and the fact that $k = (\frac{1}{2} + o(1))N_{E/K}$, which follows from our assumption on the growth of the rank in terms of $N_{E/K}$).

The last range that we need to treat is $0.35\sqrt{N_{E/K}} \leq |\xi| \leq 5\sqrt{N_{E/K}}$; we need to show that $|\widehat{Y}_2(\xi)| \leq \xi^{-4}$. For these values of ξ we always have that

$$
0.28 \le \left| \frac{2}{1 - q^{-\frac{1}{2}} e^{-i\theta_j}} \right| \sqrt{\frac{q - 1}{q}} \frac{|\xi|}{\sqrt{N_{E/K}}} \le 35,
$$

and hence from the properties of the Bessel function J_0 ,

$$
|\widehat{Y}_2(\xi)| \leq 0.981^k.
$$

This last quantity is $\leq \xi^{-4}$, since $0.981^k \leq (5N_{E/K})^{-4}$ is clearly true for large enough $N_{E/K}$.

.

The conclusion of Lemma 4.4 is then that the distribution function of Y_2 satisfies

(61)
$$
\sup_{x \in \mathbb{R}} |F_{Y_2}(x) - G(x)| \ll \frac{\text{rank}(E/K) + \log N_{E/K}}{N_{E/K}}
$$

Now note that $Y_E = Y_1 + Y_2$ and that $|Y_1| \ll (\text{rank}(E/K) + 1)/\sqrt{N_{E/K}}$ with probability one. Hence, denoting by Z a standard Gaussian random variable, we can apply (61) to deduce the following about the distribution function of Y_E :

$$
F_E(x) = \mathbb{P}[Y_1 + Y_2 \le x] = \mathbb{P}[Y_1 + Y_2 \le x \mid |Y_1| \ll (\text{rank}(E/K) + 1) / \sqrt{N_{E/K}}]
$$

\n
$$
= \mathbb{P}[Y_2 \le x + O((\text{rank}(E/K) + 1) / \sqrt{N_{E/K}})]
$$

\n
$$
= \mathbb{P}[Z \le x + O((\text{rank}(E/K) + 1) / \sqrt{N_{E/K}})] + O\left(\frac{\text{rank}(E/K) + \log N_{E/K}}{N_{E/K}}\right)
$$

\n
$$
= G(x) + O\left(\frac{\text{rank}(E/K) + 1}{\sqrt{N_{E/K}}}\right),
$$

since $e^{-x^2/2}$ is boun[ded](#page-0-0) by 1. Noting that the above calculati[on is](#page-0-0) uniform in x, the result follows. \Box

Proof of Theorem 1.2 (ii). – It is a consequence of Theorem 4.5. \Box

The computation of the characteristic function of X_E performed in the proof and (60) allows us to deduce the existence of $\delta(E)$ as soon as $N(E/K)$ is not too small.

L 4.6. – *Suppose that* E/K *is an elliptic curve satisfying LI and such that* $N(E/K) - \text{rank}(E/K) \ge 7$ $N(E/K) - \text{rank}(E/K) \ge 7$ $N(E/K) - \text{rank}(E/K) \ge 7$, then $\delta(E)$ e[xists](#page-0-0).

Proof. – Since $N(E/K)$ – rank $(E/K) \ge 7$ and LI holds for E/K , we have $k \ge 3$ by (57). Now with notation as in Theorem 4.5 the characteristic function $\hat{X}_E(\xi)$ given by (60) is integrable since it is $\ll (1 + |\xi|)^{-3/2}$ as noted in the proof of Theorem 4.5. It follows that the associated measure μ_E is absolutely continuous, and in particular $\delta(E)$ exists. \Box

COROLLARY 4.7. – Let ${E/K}$ *be a family of elliptic curves satisfying the hypotheses of Theorem* 1.2(ii)*. Assume further that* $N(E/K)$ – rank $(E/K) \ge 7$ *for every member of the family. Then* δ(E) *exists and satisfies*

$$
\delta(E) = \frac{1}{2} + O\left(\frac{\text{rank}(E/K) + 1}{\sqrt{N_{E/K}}}\right).
$$

Proof. – Define the random variable $Y_E := \sqrt{\frac{q-1}{q}} X_E/\sqrt{N_{E/K}}$, and denote by $F_E(x)$ its distribution function. We have by Theorem 4.5 and Lemma 4.6 that $\delta(E)$ exists and satisfies

$$
1 - \delta(E) = F_E(0) = \frac{1}{2} + O\left(\frac{\text{rank}(E/K) + 1}{\sqrt{N_{E/K}}}\right),
$$

 \Box

since the Gaussian is symmetric around 0.

Proof of Theorem 1.3. – The proof is obtained by combining Corollary 4.7 and [5, Th. 2.3] (choosing $k = 1$ $k = 1$). Let us briefly recall what the latter result asserts. In the notation set before stating Theorem 1.3 let $f \in \mathcal{F}_d(\mathbb{F}_{q^n})$ and let $\gamma_i(f)$ (seen as complex numbers after fixing an embedding of $\overline{\mathbb{Q}_\ell}$ in C) be the set of inverse roots of $L(E_f/K, T)$ that are *not* forced zeros (see (48)) of the Q-polynomial $L(E_f/K;T)$. By Deligne's purity result the inverse r[oo](#page-39-8)ts $\gamma_i(f)$ all have the same modulus (equal to the cardinality of the subfield of constants in K) so we may divide the $\gamma_j(f)$'s by the common modulus and write $e^{i\theta_j(f)}$, $\theta_i(f) \in [0, 2\pi)$, for the complex numbers of modulus 1 thus obtained. The main object of study in [5] is the multiplicative \mathbb{Z} -module

$$
Rel((\gamma_j(f))) = \left\{ (n_j) \subseteq \mathbb{Z} \colon \prod_j e^{in_j \theta_j(f)} = 1 \right\},\,
$$

which is the set of multiplicative relations among the inverse roots $\gamma_i(f)$. Of course there are multiplicative relations among the $\gamma_j(f)'s$: the ones coming from the functional equation satisfied by $L(E_f/K,T)$. In other words these relations come from the invariance of the set of roots of $L(E_f/K,T)$ under inversion. In terms of the a[ngl](#page-39-8)es $\theta_i(f) \in [0,2\pi)$, this can be rephrased by saying that for each j, $2\pi - \theta_i(f)$ is again the angle of some inverse root of $L(E_f/K, T)$. In [5] these relations are called *trivial* and the Z-module Rel $((\gamma_i(f)))$ is called trivial if it only consists of trivial relations. From [5, Th. 2.3] we know that for all p bigger than a constant depending only on d, for all big enough p-power $q := p^m$ (precisely m has to be bigger than a constant depending only on $\overline{\mathcal{F}_d} := \mathcal{F}_d \times \overline{\mathbb{F}_p}$) and for all d bigger than an absolute constant,

(62)
$$
q^{-n(d+1)} \# \{ f \in \mathcal{F}_d(\mathbb{F}_{q^n}) : \text{Rel}\left((\gamma_{,j}(f))\right) \text{ is nontrivial } \} \ll nq^{-n\gamma^{-1}} \log q,
$$

where one can take $2\gamma = 4 + 7N_f(N_f - 1)$, the implied constant depends only on d and the base curve E/K , and N_f is the degree of the the Q-polynomial $L(E_f/K, T)$.

Now we claim that

$$
(63) \t\t d \le N_f \le 2d + C_{E/K} \,,
$$

where $C_{E/K}$ is a constant depending only on E/K . To see why this holds, first recall (see, e.g., [5, Th. 1.1(3)] and the references therein) that

$$
N_f = \deg M_f + 2 \deg A_f + 4(g-1).
$$

Recall also that if $\Delta \in K$ is the discriminant of the minimal Weierstrass model of E/K then the discriminant of the minimal Weierstrass model of E_f/K is $f^6\Delta$ (see [5, Section 2.1]). Thus the locus Sing E_f of bad reduction of E_f/K consists, besides the locus Sing E of bad reduction of the base curve E/K , of the irreducible factors of f. From the above formula for N_f , one deduces

$$
\deg(\operatorname{Sing} E_f) \le N_f \le 2 \deg(\operatorname{Sing} E_f) + 4g.
$$

We have

$$
\sum_{\pi \mid f} [\mathbb{F}_{q^{\deg \pi}} : \mathbb{F}_q] \leq \deg (\operatorname{Sing} E_f) \leq \deg (\operatorname{Sing} E) + \sum_{\pi \mid f} [\mathbb{F}_{q^{\deg \pi}} : \mathbb{F}_q]
$$

where the summation is over irreducible factors π of f. Obviously this sum equals deg $f = d$. The claim follows. Inserting the [upp](#page-37-1)er bound in (62) we obtain (64)

$$
q^{-n(d+1)} \# \{ f \in \mathcal{F}_d(\mathbb{F}_{q^n}) \colon \operatorname{Rel}\left((\gamma_{,j}(f))\right) \text{ is nontrivial } \} \ll nq^{-nd^{-2}(14+b_E/d)^{-1}} \log q,
$$

under the same conditions as in (62), and where b_E is a constant depending only on E.

Next we use Corollary 4.7. To do so we note that if the Z-module $Rel((\gamma_{,j}(f))_j)$ is trivial then E_f/K satisfies LI and rank $(E_f) \leq 1$. Indeed any non-trivial Q-linear relation⁽¹¹⁾ involving elements of $\{1\} \cup \{\theta_i(f)/\pi \in [0,1]\}$ immediately leads ([by c](#page-38-1)learing denominators and exponentiating) to a non-trivial relation in $\mathrm{Rel}(\left(\gamma_{,j}(f)\right)_{j}).$ Moreover if $\mathrm{rank}\, E_{f}\geq 2$ then one of the $\theta_i(f)$'s is zero (since the forced zeros contribute at most 1 to this count) which of course produces a non-trivial rel[atio](#page-38-1)n. This disc[ussi](#page-0-0)on shows that (64) implies

(65)
$$
\#\{f \in \mathcal{F}_d(\mathbb{F}_{q^n})\colon E_f/K \text{ violates LI or } \text{rank}(E_f) \geq 2\} \ll \frac{nq^{n(d+1)}\log q}{q^{nd^{-2}(14+b_E/d)^{-1}}},
$$

under the same conditions as in (64). Corollary 4.7 applies since rank(E_f) ≤ 1 and thus as soon as $d \geq 8$, the inequality $N(E_f)$ – rank $(E_f) \geq 7$ holds by (63). Theorem 1.3 now follows from combining this with (65) and the lower bound in (63). \Box

Acknowledgments. – We would like to thank Andrew Granville for introducing the second author to elliptic curve prime number races and for suggesting to work on the summatory function of the trace of Frobenius. We would also like to thank Chris Hall for clarifying some technical aspects of Katz's family of quadratic twists, Guy Henniart for enlightening explanations regarding Artin conductors, Zeev Rudnick and Douglas Ulmer for their feedback and comments, and the referee for a careful reading and valuable suggestions that helped us improve the manuscript. The first author is grateful to Lehigh University for their hospitality during his visit where part of this work was done. The second author was supported by an NSERC Postdoctoral Fellowship, NSF grant DMS-0635607 as well as a Postdoctoral

⁽¹¹⁾ Note that the tuple of coefficients $(r_1, 0, \ldots, 0), r_1 \in \mathbb{Q}^{\times}$, does not give a \mathbb{Q} -linear relation.

Fellowship from the Fondation Sciences Mathématiques de Paris. This work was accomplished while the second author was at the Institute for Advanced Study, the University of Michigan, and Université Paris Diderot.

BIBLIOGRAPHY

- [1] A. AKBARY, N. NG, M. SHAHABI, Limiting distributions of the classical error terms of prime number theory, *Q. J. Math.* **65** (2014), 743–780.
- [\[2\]](http://smf.emath.fr/Publications/AnnalesENS/4_49/html/ens_ann-sc_49_5.html#4) S. BAIG, C. HALL, Experimental data for Goldfeld's conjecture over function fields, *Exp. Math.* **21** (2012), 362–374.
- [\[3\]](http://smf.emath.fr/Publications/AnnalesENS/4_49/html/ens_ann-sc_49_5.html#5) A. BRUMER, The average rank of elliptic curves. I, *Invent. math.* **109** [\(1992\), 445–472.](doi:10.1093/imrn/rnw087)
- [\[4\]](http://smf.emath.fr/Publications/AnnalesENS/4_49/html/ens_ann-sc_49_5.html#6) B. CHA, Chebyshev's bias in function fields, *Compos. Math.* **144** (2008), 1351-1374.
- [\[5\]](http://smf.emath.fr/Publications/AnnalesENS/4_49/html/ens_ann-sc_49_5.html#7) B. CHA, D. FIORILLI, F. JOUVE, Independence of the zeros of elliptic curve L -functions over function fields, *Int. Math. Res. Not.* **2016** (2016), doi:10.1093/imrn/rnw087.
- [\[6\]](http://smf.emath.fr/Publications/AnnalesENS/4_49/html/ens_ann-sc_49_5.html#8) P. DELIGNE, La conjecture de Weil. II, *Publ. Math. IHÉS* 52 (1980), 137–252.
- [7] C.-G. ESSEEN, Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law, *Acta Math.* **77** (1945), 1–125.
- [8] D. W. FARMER, S. M. GONEK, C. P. HUGHES, The maximum size of L-functions, *J. reine angew. Math.* **609** (2007), 215–236.
- [9] D. FIORILLI, Elliptic curves of unbounded rank and Chebyshev's bias, *Int. Math. Res. Not. IMRN* **2014** (2014), 4997–5024.
- [10] D. FIORILLI, Highly biased prime number races, *Algebra Number Theory* 8 (2014), 1733–1767.
- [\[11\]](http://smf.emath.fr/Publications/AnnalesENS/4_49/html/ens_ann-sc_49_5.html#13) D. FIORILLI, G. MARTIN, Inequities in the Shanks-Rényi prime number race: an asymptotic formula for the densities, *J. reine angew. Math.* **676** (2013), 121–212.
- [\[12\]](http://smf.emath.fr/Publications/AnnalesENS/4_49/html/ens_ann-sc_49_5.html#14) M. G, Artin's conjecture on the average, *Mathematika* **15** (1968), 223–226.
- [13] C. H, Big symplectic or orthogonal monodromy modulo l, *Duke Math. J.* **141** (2008), 179–203.
- [14] P. HUMPHRIES, *The Mertens and Pólya conjectures in function fields*, Master's thesis, Australian National University, 2012.
- [15] N. M. K, *Twisted* L*-functions and monodromy*, Annals of Math. Studies **150**, Princeton Univ. Press, 2002.
- [16] B. M, Finding meaning in error terms, *Bull. Amer. Math. Soc. (N.S.)* **45** (2008), 185–228.
- [\[17\]](http://smf.emath.fr/Publications/AnnalesENS/4_49/html/ens_ann-sc_49_5.html#19) M. M. R, *Probability theory with applications*, Probability and Mathematical Statistics, Academic Press Inc., Harcourt Brace Jovanovich Publishers, 1984.
- [18] M. R, P. S, Chebyshev's bias, *Experiment. Math.* **3** (1994), 173–197.
- [19] P. SARNAK, letter to Barry Mazur on Chebyshev's bias for $\tau(p)$ (2007).
- [20] D. ULMER, Elliptic curves with large rank over function fields, *Ann. of Math.* **155** (2002), 295–315.

- [\[21\]](http://smf.emath.fr/Publications/AnnalesENS/4_49/html/ens_ann-sc_49_5.html#23) D. ULMER, Geometric non-vanishing, *Invent. math.* **159** (2005), 133-186.
- [22] D. U, Elliptic curves over function fields, in *Arithmetic of* L*-functions*, IAS/Park City Math. Ser. **18**, Amer. Math. Soc., 2011, 211–280.
- [23] D. ULMER, Conductors of ℓ -adic representations, *Proc. Amer. Math. Soc.* 144 (2016), 2291–2299.

(Manuscrit reçu le 6 mai 2015 ; accepté, après révision, le 17 décembre 2015.)

Byungchul C Department of Mathematics and Computer Science Muhlenberg College 2400 Chew st., Allentown, PA 18104, USA E-mail: cha@muhlenberg.edu

Daniel FIORILLI Département de mathématiques et de statistique Université d'Ottawa 585 King Edward Ottawa, Ontario, K1N 6N5, Canada E-mail: daniel.fiorilli@uottawa.ca

Florent J Laboratoire de Mathématiques d'Orsay Université Paris-Sud & CNRS Université Paris-Saclay 91405 Orsay, France E-mail: florent.jouve@math.u-psud.fr