Arithmetic Conjectures Suggested by the Statistical Behavior of Modular Symbols

Barry Mazur\textsuperscript{a} and Karl Rubin\textsuperscript{b}

\textsuperscript{a}Department of Mathematics, Harvard University, Cambridge, MA, USA; \textsuperscript{b}Department of Mathematics, UC Irvine, Irvine, CA, USA

\begin{abstract}
Suppose $E$ is an elliptic curve over $\mathbb{Q}$ and $\chi$ is a Dirichlet character. We use statistical properties of modular symbols to estimate heuristically the probability that $L(E, \chi, 1) = 0$. Via the Birch and Swinnerton-Dyer conjecture, this gives a heuristic estimate of the probability that the Mordell–Weil rank grows in abelian extensions of $\mathbb{Q}$. Using this heuristic, we find a large class of infinite abelian extensions $F$ where we expect $E(F)$ to be finitely generated. Our work was inspired by earlier conjectures (based on random matrix heuristics) due to David, Fearnley, and Kisilevsky. Where our predictions and theirs overlap, the predictions are consistent.
\end{abstract}

\begin{subject}
Primary: 11G05, Secondary: 11F67, 11G40, 14G05
\end{subject}

1. Introduction

By a number field, we will mean a field of finite degree over $\mathbb{Q}$, and any field denoted “$K$” below will be assumed to be a number field.

\textbf{Definition.} A variety $V$ over $K$ is said to be diophantine stable for the extension $L/K$ if $V(L) = V(K)$; i.e., if $V$ acquires no “new” rational points when extended from $K$ to $L$. Depending on emphasis intended, we will also sometimes say $L/K$ is diophantine stable for $V$.

The “minimalist philosophy” leads us to the following question.

\textbf{Question A.} If $A$ is an abelian variety over $K$, $\ell$ is an odd prime number, and $m$ is a positive integer, is it the case that the cyclic Galois extensions of $K$ of degree $\ell^m$ that are diophantine stable for $A$ are of density 1 (among all cyclic Galois extensions of $K$ of degree $\ell^m$, ordered by conductor)?

The following result of [20] is an embarrassingly weak theorem in the direction of answering this question.

\textbf{Theorem B.} If $A/K$ is a simple abelian variety such that $\text{End}_K(A) = \text{End}_K(A)$, then there is a set $S$ of prime numbers of positive density such that for all $\ell \in S$ and for all positive integers $m$ there are infinitely many cyclic Galois extensions of $K$ of degree $\ell^m$ that are diophantine stable for $A$.

Unfortunately, we cannot replace the phrase “infinitely many” in the statement of Theorem B by “a positive proportion” (where “proportion” is defined by organizing these cyclic extensions by size of conductor). If we order the extensions that way and consider what we have proved for a given $\ell^m$, among the first $X$ of these we get at least $X/\log^2 X$ as $X \to \infty$ (for a small, but positive, $\alpha$).

Note that, for example, when $K/\mathbb{Q}$ is quadratic, the cyclic $\ell^m$ extensions of $K$ that are Galois dihedral extensions of $\mathbb{Q}$ are potentially a source of systematic diophantine instability but are of density 0 in all cyclic $\ell^m$ extensions of $K$.

The aim of this article is to study the distributions of values of modular symbols and of what we call $\theta$-coefficients, in order to develop a heuristic prediction of the probability of diophantine stability (or instability) in the specific case where $V = E$ is an elliptic curve, $K = \mathbb{Q}$, and $L/\mathbb{Q}$ is abelian.

For example, our heuristic leads to the following conjecture.

\textbf{Conjecture 11.2.} Let $E$ be an elliptic curve over $\mathbb{Q}$ and $F/\mathbb{Q}$ any real abelian extension such that $F$ contains only finitely many subfields of degree 2, 3 or 5 over $\mathbb{Q}$. Then the group of $F$-rational points $E(F)$ is finitely generated.

For example, we can take the field $F$ in Conjecture 11.2 to be

- the cyclotomic $\mathbb{Z}_\ell$-extension of $\mathbb{Q}$ for any prime $\ell$, in which case Conjecture 11.2 is known to be true by work of Kato [16] and Rohrlich [25],

\begin{contact}
Karl Rubin krubin@uci.edu Department of Mathematics, UC Irvine, Irvine, CA 92697, USA.
\end{contact}
the compositum of all the above.

Question A was inspired by conjectures of David, Fearnley and Kisilevsky [10], who deal specifically with the case of elliptic curves over \( \mathbb{Q} \). They conjecture (cf. Conjecture 10.5) that for a fixed elliptic curve \( E \) over \( \mathbb{Q} \) and a fixed prime \( \ell \geq 7 \), there are only finitely many Dirichlet characters of order \( \ell \) for which \( L(E, \chi, 1) = 0 \). Consequently, (assuming the Birch–Swinnerton-Dyer Conjecture) only finitely many cyclic extensions of \( \mathbb{Q} \) of order \( \ell \) are diophantine unstable for a fixed elliptic curve \( E \) over \( \mathbb{Q} \). Their conjectures are based on computations and random matrix heuristics.

In this article, we will consider a more naive heuristic, that leads us to make conjectures first regarding the distributions of \( \theta \)-coefficients (see Sections 4 and 7, and Conjecture 7.2). These conjectures, supported by numerical evidence, lead us to Conjecture 11.2.

In the first part of the article, we introduce modular symbols, recall the properties we need, and define the \( \theta \)-coefficients (which are sums of modular symbols). We use the known distribution properties of modular symbols, along with numerical calculations, to make conjectures about the (more mysterious) distributions of \( \theta \)-coefficients.

We say "more mysterious" because (as suggested by random matrix heuristics) there may well be no useful limiting distribution, no matter how one normalizes the \( \theta \)-coefficients. Nevertheless, our heuristic needs only quite weak estimates, that we label "upper bounds for heuristic likelihood." These estimates seem to be in accord with the computations that we have made—and (it seems to us) by a comfortable margin.

In the second part of the article, we use these conjectures about \( \theta \)-coefficients to develop a heuristic for the probability of vanishing of twisted \( L \)-values \( L(E, \chi, 1) \) where \( \chi \) is a Dirichlet character of order at least 3.

### Part I

#### Modular symbols

2. Basic properties of modular symbols

Fix once and for all an elliptic curve \( E \) defined over \( \mathbb{Q} \). Since \( E \) is fixed we will usually suppress it from the notation. Let \( N \) be the conductor of \( E \),

\[
\phi_E : \mathbf{H} \to X_0(N) (\mathbb{C}) \to E(\mathbb{C})
\]

the (optimal) modular uniformization of \( E \) by \( \mathbf{H} := \mathbb{H} \cup \mathbb{P}^1(\mathbb{Q}) \), the completion of the upper half-plane \( \mathbf{H} := \{ z \in \mathbb{C} | \text{Im}(z) > 0 \} \),

\[
f_E(z) = \sum_n a_n e^{2\pi i nz}
\]

the newform associated to \( E \), and

\[
\Omega_E^\pm = \Omega^\pm = \int_{\gamma^\pm} \omega_E
\]

the real and imaginary periods of \( E \), where \( \omega_E \) is the Néron differential of (the Néron model of) \( E \), \( \gamma^+ \) is the appropriately oriented connected component of the real locus of \( E \), and \( \gamma^- \) is, similarly, the appropriately oriented cycle in \( E(\mathbb{C}) \) stabilized, but not fixed, by complex conjugation.

Note that we have two maps (one with "+", one with "−") that we can evaluate on closed curves \( \gamma \) in \( E(\mathbb{C}) \), namely

\[
\gamma \mapsto \frac{1}{\Omega^\pm} \left( \int_{\gamma} \omega_E \pm \int_{-\gamma} \omega_E \right) \in \mathbb{Z}
\]

These induce linear maps

\[
H_1(E(\mathbb{C}); \mathbb{Z}) \to \mathbb{Z}.
\]

The relationship of \( \omega_E \) to \( f_E \) is given (cf. [1]) by

\[
\phi_E^* \omega_E = c_E \cdot 2\pi i f_E(z) dz,
\]

where \( c_E \) is a nonzero integer ("Manin’s Constant").

For every \( r \in \mathbb{P}^1(\mathbb{Q}) \) the image of the "vertical line" in the upper half-plane

\[
\{ z = r + iy \mid 0 \leq y \leq \infty \} \subset \mathbf{H}
\]

in \( E \) is an oriented compact curve that "begins" at \( i\infty \) and terminates at some cusp in \( E \) (and any cusp is of finite order by the Manin–Drinfeld theorem).
Definition 2.1. For every $r \in \mathbb{Q}$ define the (raw) modular symbols
\[
[r] := 2\pi i \int_{\infty}^{z} f_E(z) dz \in \mathbb{C}
\]
and the plus/minus normalized modular symbols
\[
[r]^{\pm} := \frac{[r] \pm \{-r\}}{2\Omega^{\pm}}.
\]
We will also write simply $\Omega$ and $[r]$ for $\Omega^{+}$ and $[r]^{+}$, respectively.

The modular symbols have the following well-known properties.

Lemma 2.2. For every $r \in \mathbb{Q}$ we have:

(i) $[r]^{\pm} \in \delta_E^{-1}\mathbb{Z}$ for some positive integer $\delta_E$ independent of $r$
(ii) $[r + 1]^{\pm} = [r]^{\pm} = \pm [-r]^{\pm}$
(iii) If $A \in \Gamma_0(N) \subset SL_2(\mathbb{Z})$ then, viewing $A$ as a linear fractional transformation we have $[r]^{\pm} = [A(r)]^{\pm} - [A(\infty)]^{\pm}$. If further $A$ has a complex (quadratic) fixed point, then $[A(\infty)]^{\pm} = 0$, and therefore
\[
[A(r)]^{\pm} = [r]^{\pm}
\]
for all $r \in \mathbb{Q} \cup \{\infty\}$.
(iv) Atkin–Lehner relation: Let $m \geq 1$ and write $N = cf$ where $f := \gcd(m, N)$. Assume that $c$ and $f$ are relatively prime, and let $W_c$ be the Atkin–Lehner Hecke operator $[3]$. Denote by $w_c$ the eigenvalue of $W_c$ on $f_E$. If $a, d \in \mathbb{Z}$ and $ade \equiv -1 \pmod{m}$, then
\[
[d/m]^{\pm} = -w_c \cdot [a/m]^{\pm}.
\]
(v) Hecke relations: Suppose $\ell$ is a prime, and $a_\ell$ is the $\ell$-th Fourier coefficient of $f_E$. Then
(a) If $\ell \nmid N$, then $a_\ell \cdot [r]^{\pm} = [\ell r]^{\pm} + \sum_{i=0}^{\ell-1} [(r + i)/\ell]^{\pm}$.
(b) If $\ell \mid N$, then $a_\ell \cdot [r]^{\pm} = \sum_{i=0}^{\ell-1} [(r + i)/\ell]^{\pm}$.

Proof. For (i), we can take $\delta_E$ to be (any positive multiple of) $c_E \cdot \lambda_E$ where $\lambda_E$ is the l.c.m. of the orders of the image of the cusps of $X_0(N)$ in $E$. Assertion (ii) follows directly from the definition.

The first part of assertion (iii) is evident. For the second part, let $z = A(z)$ be the fixed point of $A$, and $\gamma$ a geodesic from $\infty$ to $z$. By invariance under $A$ we get
\[
\int_{\infty}^{z} f_E(z) dz = \int_{A(\infty)}^{A(z)} f_E(z) dz = \int_{A(\infty)}^{z} f_E(z) dz.
\]
Thus $\int_{\infty}^{A(\infty)} f_E(z) dz = 0$, and it follows that $[A(\infty)]^{\pm} = [\infty]^{\pm} = 0$.

For (iv), here is a construction of the Atkin–Lehner operator $W_c$. Let $f := \gcd(m, N)$ and $e := N/f$. The $W_c$ operator is given by (any) matrix of the following form:
\[
W_c := \begin{pmatrix} -ae & b \\ cN & de \end{pmatrix},
\]
with $a, b, c, d \in \mathbb{Z}$ and $\det(W_c) = e$.

Let $c := -m/f$ and $b := (ade + 1)/m \in \mathbb{Z}$. With these choices the matrix $W_e$ of (2.3) has determinant $e$,
\[
W_e(\infty) = -\frac{ae}{cN} = -\frac{a}{cf} = \frac{a}{m},
\]
and (computing)
\[
W_e(d/m) = \infty.
\]
Thus $W_e$ takes the path $[\infty, d/m]$ to the path $[a/m, \infty]$. It follows that $[d/m] = -w_e[a/m]$ where $w_e$ is the eigenvalue of $W_e$ acting on $f_E$. This is (iv), and the proof of (v) is straightforward. □
3. Modular symbols and $L$-values

**Definition 3.1.** Suppose $\chi$ is a primitive Dirichlet character of conductor $m$. Define the Gauss sum

$$\tau(\chi) := \sum_{a=1}^{m} \chi(a)e^{2\pi ia/m}$$

and, if $L(E, s) = \sum a_n n^{-s}$, the twisted $L$-function

$$L(E, \chi, s) := \sum_{n=1}^{\infty} \chi(n)a_n n^{-s}.$$  

If $F/\mathbb{Q}$ is a finite abelian extension of conductor $m$, we will identify characters of $\text{Gal}(F/\mathbb{Q})$ with primitive Dirichlet characters of conductor dividing $m$ in the usual way.

**Proposition 3.2.** If $F/\mathbb{Q}$ is a finite abelian extension, and the Birch and Swinnerton-Dyer conjecture holds for $E_{/\mathbb{Q}}$ and $E_{/F}$, then

$$\text{rank}(E(F)) = \text{rank}(E(\mathbb{Q})) + \sum_{\chi: \text{Gal}(F/\mathbb{Q}) \to \mathbb{C}^\times} \text{ord}_{s=1} L(E, \chi, s).$$

**Proof.** This follows from the identity

$$L(E/F, s, \chi) = \prod_{\chi: \text{Gal}(F/\mathbb{Q}) \to \mathbb{C}^\times} L(E, \chi, s).$$

\[\square\]

**Theorem 3.3 (Birch–Stevens).** If $\chi$ is a primitive Dirichlet character of conductor $m$, then

$$\sum_{a=1}^{m} \chi(a)[a/m]^{e} = \frac{\tau(\chi)L(E, \chi, 1)}{\Omega^e}.$$  

where the sign $e := \chi(-1)$ is the sign of the character $\chi$.

4. $\theta$-elements and $\theta$-coefficients

**Definition 4.1.** Suppose $m \geq 1$, and let $\Gamma_m = \text{Gal}(\mathbb{Q}(\mu_m)/\mathbb{Q})$. Identify $\Gamma_m$ with $(\mathbb{Z}/m\mathbb{Z})^\times$ in the usual way, and let $\sigma_{a,m} \in \Gamma_m$ be the Galois automorphism corresponding to $a \in (\mathbb{Z}/m\mathbb{Z})^\times$ (i.e., $\sigma_{a,m}(\zeta) = \zeta^a$ for $\zeta \in \mu_m$). Define

$$\theta_{m}^{\pm} := \delta_{E} \sum_{a \in (\mathbb{Z}/m\mathbb{Z})^\times} [a/m]^{\pm} \sigma_{a,m} \in \mathbb{Z}[\Gamma_m]$$

where $\delta_{E}$ is as in Lemma 2.2(i).

If $F/\mathbb{Q}$ is a finite abelian extension of conductor $m$, so $F \subset \mathbb{Q}(\mu_m)$, define the $\theta$-element (over $F$, associated to $E$) to be:

$$\theta_{F}^{\pm} := \theta_{m}^{\pm}|_{F} \in \mathbb{Z}[\text{Gal}(F/\mathbb{Q})]$$

where $\theta_{m}^{\pm}|_{F}$ is the image of $\theta_{m}^{\pm}$ under the natural restriction homomorphism

$$\mathbb{Z}[\text{Gal}(\mathbb{Q}(\mu_m)/\mathbb{Q})] \to \mathbb{Z}[\text{Gal}(F/\mathbb{Q})].$$

By Lemma 2.2(i) we have

$$\theta_{F}^{\pm} = \sum_{\gamma \in \text{Gal}(F/\mathbb{Q})} c_{F,\gamma}^{\pm} \gamma \in \mathbb{Z}[\text{Gal}(F/\mathbb{Q})]$$

where

$$c_{F,\gamma}^{\pm} := \delta_{E} \sum_{a \in (\mathbb{Z}/m\mathbb{Z})^\times} [a/m]^{\pm} \sigma_{a,m} \in \mathbb{Z}[\text{Gal}(F/\mathbb{Q})].$$

We will refer to the $c_{F,\gamma}^{\pm} \in \mathbb{Z}$ as $\theta$-coefficients. Since we will most often be dealing with the "plus"-$\theta$-elements, we will simplify notation by letting $\theta_{F} := \theta_{F}^{+}$, $c_{F,\gamma} := c_{F,\gamma}^{+}$, and $\Omega := \Omega^{+}$. If $F$ is a real field, then $\sigma_{-1,m}|_{F} = 1$, and $[a/m] = [-a/m]$, so

$$c_{F,\gamma} = 2\delta_{E} \sum_{a \in (\mathbb{Z}/m\mathbb{Z})^\times /\{1\}} [a/m].$$

With this notation, Proposition 3.3 can be rephrased as follows:
Corollary 4.4. Suppose $F/\Q$ is a finite real cyclic extension of conductor $m$ and $\chi : (\Z/m\Z)^{\times} \to \Gal(F/\Q) \hookrightarrow \C^{\times}$ is a character that cuts out $F$. Then
\[
\tilde{\chi}(\theta_F) = \delta_{E}(\sum_{a|m \chi} \tau(\chi)L(E,\chi,1)) / \Omega.
\]
In particular $\tilde{\chi}(\theta_F)$ vanishes if and only if $L(E,\chi,1)$ does.

5. Distribution of modular symbols

From now on, we assume that our elliptic curve $E$ is semistable, so its conductor $N$ is squarefree.

Theorem 5.2 (Petridis & Risager [24, see also [23]). As $X$ goes to infinity the values
\[
\left\{ \frac{[a/m]}{\sqrt{\log(m)}} : m \leq X, a \in (\Z/m\Z)^{\times} \right\}
\]
approach a normal distribution with variance $C_E$.

Numerical experiments led to the following conjecture. Denote by $\text{Var}(m)$ the variance
\[
\text{Var}(m) := \frac{1}{\varphi(m)} \sum_{a \in (\Z/m\Z)^{\times}} (\frac{[a/m]}{\sqrt{\log(m)}})^2
\]

Conjecture 5.3. (i) As $m$ goes to infinity, the distribution of the sets
\[
\left\{ \frac{[a/m]}{\sqrt{\log(m)}} : a \in (\Z/m\Z)^{\times} \right\}
\]
converges to a normal distribution with mean zero and variance $C_E$.

(ii) For every divisor $\kappa$ of the conductor $N$, there is a constant $D_{E,\kappa} \in \R$ such that
\[
\lim_{m \to \infty} \sum_{(m,N) = \kappa} \text{Var}(m) - C_E \log(m) = D_{E,\kappa}.
\]

Remark 5.4. We expect that a version of Conjecture 5.3 holds also for non-semistable elliptic curves, with a slightly more complicated formula for $C_E$.

Note that Theorem 5.2 is an “averaged” version of Conjecture 5.3(i). Inspired by Conjecture 5.3, Petridis and Risager [24, Theorem 1.6] obtained the following result, which identifies the constant $D_{E,\kappa}$ and proves an averaged version of Conjecture 5.3(ii).

Theorem 5.5 (Petridis & Risager [24]). For every divisor $\kappa$ of $N$, there is an explicit (see [24, (8.12)]) constant $D_{E,\kappa} \in \R$ such that
\[
\lim_{X \to \infty} \sum_{\substack{m \leq X \\text{prime}}} \frac{1}{\varphi(m)} \sum_{(m,N) = \kappa} \varphi(m) \text{Var}(m) - C_E \log(m) = D_{E,\kappa}.
\]

Remark 5.6. Petridis & Risager compute $D_{E,\kappa}$ in terms of the values $L(\Sym^2(E),1)$ and $L'(\Sym^2(E),1)$. They deal with non-holomorphic Eisenstein series twisted by modular symbols.

Lee and Sun [19] more recently have proven the same result (for arbitrary $N$, averaged over $m$, but without explicit determination of the constants $C_E$ and $D_{E,\kappa}$) by considering dynamics of continued fractions.

A version of Conjecture 5.3(ii) with $m$ ranging over primes was proved by Blomer, Fouvry, Kowalski, Michel, Miličević, and Sawin in [5, Theorem 9.2].

For related results regarding higher weight modular eigenforms see [4]. See also [8] and [22] for other related results.

Remark 5.7. The modular symbols are not completely “random” subject to Conjecture 5.3. Specifically, partial sums $\sum_{a=m}^\beta [a/m]$ behave in a somewhat orderly way. Numerical experiments led the authors and William Stein to conjecture the following result, which was then proved by Diamantis, Hoffstein, Kiral, and Lee (for arbitrary $N$, with an explicit rate of convergence) [12, Theorem 1.2].
Theorem 5.8 ([12]). If $0 < x < 1$ then
\[
\lim_{m \to \infty} \frac{1}{m} \sum_{a=1}^{m} \frac{\sin(\pi nx)}{n^2 \Omega} = \sum_{n=1}^{\infty} \alpha_n q^n
\]
where $\sum_n \alpha_n q^n$ is the modular form $f_E$ corresponding to $E$.

Another proof in the case of squarefree $N$ was given by Sun [26, Theorem 1.1].

6. The involution $t_F$

Definition 6.1. Suppose $F$ is a finite real cyclic extension of $\mathbb{Q}$. Let $m$ be the conductor of $F$, $f := \gcd(m,N)$, and $e := N/f$. Since $N$ is assumed squarefree, $e$ is prime to $m$. Let $\gamma_F$ be the image of $e$ under the map $(\mathbb{Z}/m\mathbb{Z})^\times \to \text{Gal}(F/\mathbb{Q})$. Define an involution $t_F$ of the set $\text{Gal}(F/\mathbb{Q})$ by
\[
t_F(\gamma) = (\gamma_F)^{-1}
\]
Let $w_F = -w_e$ where $w_e$ is the eigenvalue of the Atkin–Lehner operator $W_e$ (see Lemma 2.2(iv)) acting on $f_E$.

Recall from (4.2) that $\theta_F = \sum_{\gamma \in \text{Gal}(F/\mathbb{Q})} c_{F,\gamma} \gamma$.

Lemma 6.2. Suppose $F$ is a finite real cyclic extension of $\mathbb{Q}$.

(i) We have $c_{F,\gamma} = w_F c_{F,\gamma'}$ where $\gamma' = t_F(\gamma)$.

(ii) The fixed points of $t_F$ are the square roots of $\gamma_F^{-1}$ in $\text{Gal}(F/\mathbb{Q})$, so the number of fixed points is:
- one if $|F: \mathbb{Q}|$ is odd,
- zero if $\gamma_F$ is not a square in $\text{Gal}(F/\mathbb{Q})$,
- two if $|F: \mathbb{Q}|$ is even and $\gamma_F$ is a square in $\text{Gal}(F/\mathbb{Q})$.

(iii) If $\gamma = t_F(\gamma)$ and $w_F = -1$, then $c_{F,\gamma} = 0$.

Proof. Assertion (i) follows from the Atkin–Lehner relations satisfied by the modular symbols (Lemma 2.2(iv)). Assertion (ii) is immediate from the definition, and (iii) follows directly from (i). \qed

Definition 6.3. If $F/\mathbb{Q}$ is a real cyclic extension, we say that $\gamma \in \text{Gal}(F/\mathbb{Q})$ is generic, (resp., special$^+$, resp., special$^-$) if $\gamma \neq t_F(\gamma)$ (resp., $\gamma = t_F(\gamma)$ and $w_F = 1$, resp., $\gamma = t_F(\gamma)$ and $w_F = -1$). By Lemma 6.2(iii), if $\gamma$ is special$^-$ then $c_{F,\gamma} = 0$.

7. Distributions of $\theta$-coefficients

Consider real cyclic extensions $F/\mathbb{Q}$ of fixed degree $d \geq 3$ and varying conductor $m$. By (4.3), if $\gamma$ is generic (resp., special$^+$) then the $\theta$-coefficient $c_{F,\gamma}$ is $2\delta_E$ times a sum of $\varphi(m)/(2d)$ modular symbols (resp., $4\delta_E$ times a sum of $\varphi(m)/(4d)$ modular symbols). If these were randomly chosen modular symbols $[a/m]$, one would expect from Conjecture 5.3(i) that the collection of data
\[
\Sigma_d := \left\{ \frac{c_{F,\gamma} \sqrt{d}}{\sqrt{\varphi(m) \log(m)}} : F/\mathbb{Q} \text{ real, cyclic of degree } d, m = \text{cond}(F), \gamma \in \text{Gal}(F/\mathbb{Q}) \text{ generic} \right\}
\]
(ordered by $m$) would converge to a normal distribution with variance $2\delta_E^2 \mathbb{C}_E$ as $m$ tends to $\infty$. Similarly the data defined in the same way except with $\gamma$ special$^+$ instead of generic, would converge to a normal distribution with variance $4\delta_E^2 \mathbb{C}_E$.

Calculations do not support this expectation, at least not for small values of $d$. See Examples 7.7 below. In addition, random matrix heuristics (see [10, 17]) predict that the data above for $d = 3$ will not converge to a non-zero distribution.

However, calculations do support the following much weaker Conjecture 7.2 below, which is strong enough for our purposes.

Definition 7.1. For every $F$ of degree $d$ and conductor $m$, and every $\gamma \in \text{Gal}(F/\mathbb{Q})$ define the normalized $\theta$-coefficient
\[
\bar{c}_{F,\gamma} := \frac{c_{F,\gamma} \sqrt{d}}{\sqrt{\varphi(m) \log(m)}}.
\]

For $d \geq 3$, $\alpha, \beta \in \mathbb{R}$, and $X \in \mathbb{R}_{>0}$, let $\Sigma_{d,\alpha,\beta}(X)$ be the collection of data (counted with multiplicity)
\[
\Sigma_{d,\alpha,\beta}(X) := \left\{ \bar{c}_{F,\gamma} m^\alpha \log(m)^\beta : F/\mathbb{Q} \text{ real, cyclic of degree } d, m = \text{cond}(F) < X, \gamma \in \text{Gal}(F/\mathbb{Q}) \text{ generic or special$^+$} \right\}.
\]
**Conjecture 7.2.** There is a $B_E > 0$ and for every $d \geq 3$ there are $\alpha_d, \beta_d \in \mathbb{R}$ such that

(i) for every real open interval $(a, b)$,

$$\limsup_{X \to \infty} \frac{\# \{ \Sigma_{d, \alpha_d, \beta_d}(X) \cap (a, b) \}}{\# \Sigma_{d, \alpha_d, \beta_d}(X)} < B_E(b - a),$$

(ii) $\{ \alpha_d \varphi(d) : d \geq 3 \}$ is bounded, and $\lim_{d \to \infty} \beta_d = 0$.

**Remark 7.3.** Random matrix theory heuristics (see [10, 11, 17]) suggest that for $d = 3$ and every real open interval $(a, b)$,

$$\limsup_{X \to \infty} \frac{\# \{ \Sigma_{d, \alpha_d, \beta_d}(X) \cap (a, b) \}}{\# \Sigma_{d, \alpha_d, \beta_d}(X)} < B_d(b - a),$$

(7.4)

with $\alpha_3 = 0, \beta_3 = 3/4$ and a sufficiently large $B_3$. Empirical data (see Examples 7.7) suggest (7.4) holds for all $d$, with $\alpha_d = 0, \beta_d$ converging to zero for large $d$ and $B_d$ bounded for large $d$. Taking $B_E$ to be the maximum of the $B_d$ leads to the statement of Conjecture 7.2.

**Remark 7.5.** Here is a heuristic shorthand description of Conjecture 7.2: There is a $B_E > 0$ such that for

- every cyclic extension $F/\mathbb{Q}$ of degree $d \geq 3$,
- every generic or special $\gamma \in \text{Gal}(F/\mathbb{Q})$, and
- every real interval $(a, b)$,

the “heuristic likelihood” that $c_{F, \gamma}$ lies in $(a, b)$, which we will denote by $\mathcal{P}[c_{F, \gamma} \in (a, b)]$, is bounded above by

$$B_E(b - a)m^{\alpha_d} \log(m)^{\beta_d},$$

where $m$ is the conductor of $F$ (and $\alpha_d, \beta_d$ are as in Conjecture 7.2).

By heuristic likelihood $\mathcal{P}[X]$ of an outcome $X$ occurring we mean a real number that gives a (merely heuristically supported guess for an) upper bound for the number of times that the outcome $X$ occurs within a specific range of events. The main use we will make of such (nonproved, but only heuristically suggested) estimates is that in various natural contexts we produce computationally supported upper bounds of heuristic likelihoods $\mathcal{P}[X_i]$ (for $i$ in some index set $I$) where the $\{X_i\}_{i \in I}$ is a collection of possible outcomes for some specific range of events. If $\sum_{i \in I} \mathcal{P}[X_i]$ is finite, we simply interpret our heuristic as suggesting that the totality of outcomes of the form $X_i$ (for $i \in I$) that actually occur—within the specific range of events—is finite.

**Remark 7.6.** The smaller those exponents $\alpha_d, \beta_d$ are—and the faster they approach 0 as $d$ grows—the stronger is the claim in Conjecture 7.2.

It may well be that $\alpha_d$ may be taken to be 0 for all $d \geq 3$; the empirical data we have computed (see 7.7 below) which has led us to frame Conjecture 7.2, might almost suggest this. But we don't believe we have computed far enough yet to say anything that precise, and—more importantly—the eventual qualitative arithmetic conjecture that we make (Conjecture 11.2) doesn’t need such precision for its heuristic support; we will only rely on the formulation of Conjecture 7.2.

**Examples 7.7.** For each of the three elliptic curves 11A1, 37A1, and 32A1 (in the notation of Cremona’s tables [9]), and five (prime) values of $d$, we computed the first (approximately) 50,000 normalized $\theta$-coefficients $c_{F, \gamma}$, with $F$ ordered by conductor and $\gamma$ generic. The resulting distributions are shown in Figures 1 through 3. As $d$ grows the distributions approach the “expected” normal distribution, shown as the dashed line in each figure.

**Part II**

**Heuristics**

8. The heuristic for cyclic extensions of prime degree

Recall that we have assumed that $E$ is semistable.

Our heuristic is based on two assumptions:

- (H1) the $c_{F, \gamma}$ are distributed randomly subject to the constraints imposed by Conjecture 7.2 and Remark 7.5,
- (H2) the only correlations among the $c_{F, \gamma}$ for fixed $F$ and varying $\gamma$ are the Atkin–Lehner relations of Lemma 6.2. See Remark 9.3 below for more about the issue of “intra-correlation”.
For simplicity, we first consider the case of extensions of prime degree. Fix for this section a character $\chi$ of prime order $p \geq 3$. As above we write $m$ for the conductor of $\chi$, $F$ for the corresponding real cyclic extension of $\mathbb{Q}$, and $G := \text{Gal}(F/\mathbb{Q})$. Let $B_E, \alpha_p, \beta_p \in \mathbb{R}$ be as in Conjecture 7.2.

Since $p$ is odd, Lemma 6.2(ii) shows that the involution $\iota_E$ has a unique fixed element $\sigma \in G$. Choose a subset $S \subset G$ consisting of one element from each pair $\{\gamma, \iota_E(\gamma)\}$ of generic elements.

**Lemma 8.1.** We have

$$\sum_{\gamma \in G} \chi(\gamma)c_{F,\gamma} = 0 \iff c_{F,\gamma} = c_{F,\sigma} \text{ for all } \gamma \in S.$$ 

**Proof.** The only $\mathbb{Q}$-linear relation among the values of $\chi$ (i.e., the $p$-th roots of unity) is that their sum is zero. It follows that the sum over $\gamma$ is zero if and only if all $c_{F,\gamma}$ are equal. The lemma now follows from Lemma 6.2(i) (note that $a_0 = 0$ if $w_E = -1$). \qed
Figure 3. Distribution of normalized $\theta$-coefficients for $E = 32A1$ and varying $d$.

Heuristic 8.2. With notation as above,

$$\mathcal{P}[L(E, \chi, 1) = 0] < \left( B_E^2 \frac{pm^{2\alpha_p}}{\phi(m) \log(m)^{1-2\beta_p}} \right)^{(p-1)/4}.$$ 

Remark 8.3. The notation $\mathcal{P}[L(E, \chi, 1) = 0]$ is meant to stand for (an upper bound for) the heuristic likelihood that the value $L(E, \chi, 1)$ vanishes (see Remark 7.5).

Justification for Heuristic 8.2. By Corollary 4.4 and Lemma 8.1,

$$L(E, \chi, 1) = 0 \iff \sum_{\gamma \in G} \chi(\gamma)c_{F, \gamma} = 0 \iff c_{F, \gamma} = c_{F, \sigma} \text{ for all } \gamma \in S.$$ 

Let

$$r_\chi := \frac{\sqrt{p}}{\sqrt{\psi(m) \log(m)}},$$

so the normalized $\theta$-coefficients $\tilde{c}_{F, \gamma}$ of Definition 7.1 are given by $\tilde{c}_{F, \gamma} = c_{F, \gamma}r_\chi$. Since the $c_{F, \gamma}$ are integers, we have

$$c_{F, \gamma} = c_{F, \sigma} \iff c_{F, \gamma} \in \left( \frac{c_{F, \sigma}}{2} \right)^2, c_{F, \sigma} + \frac{1}{2} \iff \tilde{c}_{F, \gamma} \in \left( \tilde{c}_{F, \sigma} - \frac{r_\chi}{2}, \tilde{c}_{F, \sigma} + \frac{r_\chi}{2} \right).$$

Using Conjecture 7.2, Remark 7.5 and our heuristic assumptions (H1) and (H2) above, we have

$$\mathcal{P}[c_{F, \gamma} = c_{F, \sigma} \forall \gamma \in S] = \prod_{\gamma \in S} \mathcal{P}[\tilde{c}_{F, \gamma} \in \left( \tilde{c}_{F, \sigma} - \frac{r_\chi}{2}, \tilde{c}_{F, \sigma} + \frac{r_\chi}{2} \right)] \leq (B_E \left( \frac{pm^{2\alpha_p} \log(m)^{2\beta_p}}{\phi(m) \log(m)} \right)^{(p-1)/2}$$

as desired.

9. The heuristic for general cyclic extensions

Fix for this section an even character $\chi$ of arbitrary degree $d \geq 3$. As above we write $m$ for the conductor of $\chi$, $F$ for the corresponding real cyclic extension of $\mathbb{Q}$, and $G := \text{Gal}(F/\mathbb{Q})$. Let $B_E, \alpha_d$ and $\beta_d$ be the constants in Conjecture 7.2.

Heuristic 9.1. With notation as above,

$$\mathcal{P}[L(E, \chi, 1) = 0] < \left( B_E^2 \frac{dm^{2\alpha_d}}{\phi(m) \log(m)^{1-2\beta_d}} \right)^{\phi(d)/4}.$$
The rest of this section is devoted to justifying Heuristic 9.1, using heuristic assumptions (H1) and (H2) of Section 8. Recall the sign \( w_F = \pm 1 \), the element \( y_F \in G \), and the involution \( t_F : G \to G \) of Definition 6.1. We extend \( t_F \) to a \( \mathbb{Z} \)-linear involution of \( \mathbb{Z}[G] \) (also denoted by \( t_F \)).

Fix a generator \( g \) of \( G \). Let \( \zeta \in \mu_d \) denote the primitive \( d \)-th root of unity \( \chi(g) \). Define a \( \mathbb{Z} \)-linear involution \( i_\chi \) of the cyclotomic ring \( \mathbb{Z}[\zeta] \) by

\[
i_\chi(\rho) = \chi(y_F)^{-1} \rho
\]

where \( \rho \mapsto \hat{\rho} \) is complex conjugation. Then we have a commutative diagram

\[
\begin{array}{ccc}
\mathbb{Z}[G]^{i_F=\rho_F} & \xrightarrow{\chi} & \mathbb{Z}[G] \\
\downarrow & & \downarrow \\
\mathbb{Z}[\zeta]^{i_\chi=\rho_F} & \xrightarrow{i_\chi} & \mathbb{Z}[\zeta]
\end{array}
\]

where, for example, \( \mathbb{Z}[G]^{i_F=\rho_F} \) denotes \( \{ \rho \in \mathbb{Z}[G] : t_F(\rho) = w_F \rho \} \).

**Definition 9.2.** For \( \gamma \in G \), define \( v_\gamma \in \mathbb{Z}[G]^{i_F=\rho_F} \) by

\[
v_\gamma := \begin{cases} 
\gamma + w_F t_F(\gamma) & \text{if } \gamma \text{ is generic}, \\
\gamma & \text{if } \gamma \text{ is special}^+, \\
0 & \text{if } \gamma \text{ is special}^-.
\end{cases}
\]

Fix a subset \( G_0 \subset G \) consisting of all special\(^+\) elements of \( G \) and one element from each pair \( \{ \gamma, t_F(\gamma) \} \) of generic elements. Then the set \( \{ v_\gamma : \gamma \in G_0 \} \) is a \( \mathbb{Z} \)-basis of \( \mathbb{Z}[G]^{i_F=\rho_F} \).

Since \( \chi : \mathbb{Z}[G]^{i_F=\rho_F} \to \mathbb{Z}[\zeta]^{i_\chi=\rho_F} \) is surjective, we can choose a subset \( S \subset G_0 \) such that, writing \( A \subset \mathbb{Q}[G]^{i_F=\rho_F} \) for the \( \mathbb{Q} \)-vector space spanned by \( \{ v_\gamma : \gamma \in S \} \), \( \chi \) maps \( A \) isomorphically to \( \mathbb{Q}(\zeta)^{i_\chi=\rho_F} \). Let \( \lambda : \mathbb{Q}(\zeta)^{i_\chi=\rho_F} \to A \) be the inverse isomorphism. In particular we have

\[
#S = \dim_\mathbb{Q} \mathbb{Q}(\zeta)^{i_\chi=\rho_F} = \varphi(d)/2.
\]

**Justification for Heuristic 9.1.** Define

\[
\rho := \sum_{\gamma \in G} c_{F,\gamma} v_\gamma.
\]

Then \( \rho \in \mathbb{Z}[G]^{i_F=\rho_F} \), so \( \rho = \sum_{\gamma \in G_0} c_{F,\gamma} v_\gamma \). Let

\[
\rho_1 := \sum_{\gamma \in S} c_{F,\gamma} v_\gamma, \quad \rho_2 := \sum_{\gamma \in G_0 - S} c_{F,\gamma} v_\gamma,
\]

so \( \rho = \rho_1 + \rho_2 \) and \( \rho_1, \rho_2 \in A \). We have

\[
\sum_{\gamma \in G} \chi(\gamma) c_{F,\gamma} = 0 \iff \chi(\rho) = 0 \iff \chi(\rho_1) = -\chi(\rho_2) \iff \rho_1 = -\lambda(\chi(\rho_2)).
\]

Since \( \lambda(\chi(\rho_2)) \in A \) we can write \( -\lambda(\chi(\rho_2)) = \sum_{\gamma \in S} b_\gamma v_\gamma \) with \( b_\gamma \in \mathbb{Q} \). Hence, writing \( r_\chi := \sqrt{d} / \sqrt[\varphi(m) \log(m)} \) as in Section 8,

\[
\sum_{\gamma \in G} \chi(\gamma) c_{F,\gamma} = 0 \iff c_{F,\gamma} = b_\gamma \quad \text{for every } \gamma \in S
\]

\[
\iff c_{F,\gamma} \in (b_\gamma - \frac{1}{2}, b_\gamma + \frac{1}{2}) \quad \text{for every } \gamma \in S
\]

\[
\iff \tilde{c}_{F,\gamma} \in (b_\gamma r_\chi - \frac{d}{2}, b_\gamma r_\chi + \frac{d}{2}) \quad \text{for every } \gamma \in S
\]

so using Conjecture 7.2, Remark 7.5 and the heuristic assumptions (H1) and (H2) of Section 8 we have

\[
\mathbb{P} \left[ \sum_{\gamma \in G} \chi(\gamma) c_{F,\gamma} = 0 \right] = \prod_{\gamma \in S} \mathbb{P} \left[ \tilde{c}_{F,\gamma} \in (b_\gamma r_\chi - \frac{d}{2}, b_\gamma r_\chi + \frac{d}{2}) \right] \ll (B \delta r_\chi m^{6d} \log(m)^{6d})^{#S/2}
\]

(note that the \( b_\gamma \) for \( \gamma \in S \) depend only on the \( c_{F,\gamma} \) for \( \gamma \in G_0 - S \)). Since \( #S = \varphi(d)/2 \), and \( L(E, \chi, 1) = 0 \) if and only if \( \sum_{\gamma \in G} \chi(\gamma) c_{F,\gamma} = 0 \) by Corollary 4.4, this gives the desired heuristic. \( \square \)
Remark 9.3. The upper bound of Heuristic 9.1, specifically the exponent $\psi(d)/4$, depends on the assumption (H2) of no intra-correlations, i.e., no statistical correlations connecting $\psi(d)/2$ different $\theta$-coefficients of a given $\theta$-element, these being chosen to have the property that no two are brought into one another by the involution $\sigma$ (see Section 6). We might refer to the exponent $\psi(d)/4$ as the intra-correlation exponent. We will see that a qualitative version of our heuristic will still seem viable even if we assume significantly smaller intra-correlation exponents, e.g., if we have an upper bound of the form:

$$ P[L(E, \chi, 1) = 0] < \left( B_E^2 \frac{dm^{2\alpha_d}}{\psi(m) \log(m)^{1-2\beta_d}} \right)^{t(d)} $$

where $t(d) \gg \log(d)$ (see Proposition 12.15 below).

10. Expectations

In this section we suppose that Conjecture 7.2 holds, and we fix data $B_E, \alpha_d, \beta_d$ for all $d \geq 3$ as in that conjecture.

Definition 10.1. For every $d$ let $C_d$ denote the set of even primitive Dirichlet characters of order $d$. Define

$$ F_d(m) := \#\{\chi \in C_d : \text{conductor}(\chi) = m\}, $$

and for every $X$

$$ C_d(X) := \{\chi \in C_d : \text{conductor}(\chi) \leq X\}, $$

$$ E_d(X) := \sum_{m \leq X} F_d(m) \left( B_E^2 \frac{dm^{2\alpha_d}}{\psi(m) \log(m)^{1-2\beta_d}} \right)^{\psi(d)/4}. \quad (10.2) $$

Our Heuristic 9.1 suggests the following.

Heuristic 10.3. We expect

$$ \#\{\chi \in C_d(X) : L(E, \chi, 1) = 0\} \leq E_d(X). $$

In particular if $E_d(\infty)$ is finite, then the number of $\chi$ of order $d$ with $L(E, \chi, 1) = 0$ should be finite.

More generally, if $D \subseteq \mathbb{Z}_{\geq 3}$, then we expect

$$ \#\{\chi \in \bigcup_{d \in D} C_d(X) : L(E, \chi, 1) = 0\} \leq \sum_{d \in D} E_d(X). $$

Proposition 10.4. Suppose $p$ is prime, and let

$$ \sigma := (1 - 2\alpha_p)(p - 1)/4, \quad \tau := \beta_p(p - 1)/2. $$

As $X \in \mathbb{R}$ goes to infinity we have the following upper bounds on $E_p(X)$.

(i) If $\sigma > 1$, or $\sigma = 1$ and $\tau < -1$, then $E_p(\infty)$ is finite.

(ii) If $\sigma = 1$ and $\tau = -1$, then $E_p(X) \ll_{E,p} \log(X)$.

(iii) If $\sigma = 1$ and $\tau > -1$, then $E_p(X) \ll_{E,p} \log(X)^{\tau+1}$.

(iv) If $0 < \sigma < 1$, then $E_p(X) \ll_{E,p} X^{1-\sigma} \log(X)^{\tau}$.

Proof. This follows from Lemma 12.7 below. Precisely, since $m \ll \psi(m) \log(m)$ (see for example [15, Theorem 328]), we have

$$ E_p(X) \ll_{E,p} \sum_{m \leq X} F_p(m) \frac{\log(m)^{\tau}}{m^{\sigma}}. $$

Let $g(x) := \log(x)^{\tau} / x^\sigma$. Then on a suitable interval $[r, \infty)$, $g$ satisfies the hypotheses of Lemma 12.7, so that lemma shows that

$$ E_p(X) \ll_{E,p} \int_{r}^{X} g(t) dt. $$

Now the proposition follows.  

The following conjectures and predictions are suggested by the random matrix heuristic. This involves work of Conrey et al. [7] for $p = 2$, and David et al. [10, 11] and Fearnley et al. [14] for $p \geq 3$. For more precise references see Remark 10.6.

If $p(X), q(X) : \mathbb{R}_{\geq r} \rightarrow \mathbb{R}_{>0}$ are functions for some real number $r$, then we write $p(X) \sim q(X)$ to mean that $\lim_{X \to \infty} p(X)/q(X)$ exists and is nonzero.

If $p(X), q(X) : \mathbb{R}_{\geq r} \rightarrow \mathbb{R}_{>0}$ are functions for some real number $r$, then we write $p(X) \sim q(X)$ to mean that $\lim_{X \to \infty} p(X)/q(X)$ exists and is nonzero.
Conjecture 10.5 ([10, 11, 14]). For a prime $p \geq 3$ let

$$n_p(X) := \#\{\chi \in C_p(X) : L(E, \chi, 1) = 0\}.$$

Then there are nonzero constants $c_{E,p}$ such that

(i) $n_3(X) \sim \sqrt{X} \log(X)^{c_{E,3}},$
(ii) $n_5(X) \sim \log(X)^{c_{E,5}},$
(iii) $n_p(X)$ is bounded independently of $X$ if $p \geq 7.$

Remark 10.6. Assertion (iii) is part of [11, Conjecture 1.2].

The statements (i) and (ii) are implicit in [11, p. 256] but not explicitly stated as conjectures. The authors of [11] comment:

*The exact power of log $X$ that is obtained with the random matrix approach depends subtly on the discretization, and is difficult to predict.*

They state a weaker conjecture:

Conjecture 10.7 (Conjecture 1.2 of [11]).
(i) $\lim_{X \to \infty} \log n_3(X) / \log(X) = 1/2.$
(ii) $n_5(X)$ is unbounded and $\ll X^\epsilon$ as $X$ tends to infinity, for any $\epsilon > 0.$

Remark 10.8. See [14] for further interesting results related to the above conjecture for cubic characters. Fearnley and Kisilevsky [13] exhibit one example of $L(E, \chi, 1) = 0$ with $\chi$ a character of order $\ell = 11$ and $E$ the elliptic curve $5906B1$ (using Cremona’s classification [9]). The character $\chi$ is of conductor 23 (i.e., $\chi$ has the smallest possible conductor for characters of its order). In this example $\text{rank}(E(Q)) = 2,$ and $\text{rank}(E(Q(\mu_{23}))^+) = 12.$

Remark 10.9. Suppose $p = 3,$ $\alpha_3 = 0,$ and $\beta_3 = c_{E,3}.$ Then Proposition 10.4(iv) gives

$$E_3(X) \ll X^{1/2} \log(X)^{c_{E,3}}.$$

Now suppose $p = 5,$ $\alpha_5 = 0,$ and $\beta_5 = (c_{E,5} - 1)/2.$ Then Proposition 10.4(iii) gives

$$E_5(X) \ll \log(X)^{c_{E,5}}.$$

Now suppose $p \geq 7$ and $\alpha_p < (p - 5)/(2(p - 1)).$ Then Proposition 10.4(i) shows that $E_p(\infty)$ is finite.

In other words, for appropriate values of $\alpha$ and $\beta$ (that are consistent with computational data), Heuristic 10.3 is consistent with Conjecture 10.5.

Proposition 10.10. We have

$$\sum_{d: \psi(d)(1 - 2\alpha_d) > 4} E_d(\infty) < \infty.$$

Proof. This follows from Proposition 12.15 below (and comparing equations (10.2) and (10.16)) taking

$$t(d) := \psi(d)/4, \quad \sigma_d := 2\alpha_d, \quad \tau_d := 2\beta_d, \quad M := B_E^2.$$

11. Conjectures

Proposition 10.10, Heuristic 10.3, and Conjecture 7.2 lead to the following “analytic” and “arithmetic” conjectures.

Conjecture 11.1. Let

$$C = \bigcup_{\psi(d) > 4} C_d$$

(the set of all even Dirichlet characters of order at least 7 and different from 8, 10, or 12). Then the set $\{\chi \in C : L(E, \chi, 1) = 0\}$ is finite.

Conjecture 11.2. Suppose $F/\mathbb{Q}$ is an (infinite) real abelian extension that has only finitely many subfields of degree 2, 3, or 5. Then $E(F)$ is finitely generated.
Conjecture 11.2 would be a consequence of Conjecture 11.1 and the Birch and Swinnerton-Dyer conjecture. As in Proposition 3.2, if the Birch and Swinnerton-Dyer conjecture holds then it is enough to show that there are only finitely many characters $\chi$ of $\text{Gal}(F/\mathbb{Q})$ such that $L(E, \chi, 1) = 0$. But the hypotheses in Conjecture 11.2 imply that $\text{Gal}(F/\mathbb{Q})$ has only finitely many characters of order 2, 3, or 5, and consequently it has only finitely many characters of order 6, 8, 10, or 12 as well. Now the conjecture follows from Conjecture 11.1.

Remark 11.3. As mentioned in Remark 9.3, our heuristic—following Proposition 12.15—would suggest qualitatively similar conjectures (possibly with a larger set of exceptional degrees $d$) even if there were significant “intra-correlation”.

12. Analytic results

Definition 12.1. For every $d \geq 1$, recall (Definition 10.1) that $F_d(m)$ denotes the number of even primitive Dirichlet characters of order $d$ and conductor $m$. We are interested in sums of the form

$$S_d(g; X) := \sum_{m=r}^{X} F_d(m)g(m)$$

where $g : \mathbb{R}_{\geq 1} \to \mathbb{R}_{\geq 0}$ is a function, and we let $S_d(g) = S_d(g; \infty)$.

Let $\tilde{F}_d(m)$ denote the number of all primitive Dirichlet characters of order dividing $d$ and conductor $m$. The following theorem is essentially due to Kubota [18].

Theorem 12.2 (Kubota). For every $d \geq 1$, define Dirichlet series

$$L_d(s) := \sum_{m=1}^{\infty} \tilde{F}_d(m)m^{-s}, \quad \prod_{t|d, t > 1} \zeta_{\mathbb{Q}(\mu_t)}(s)$$

where $\zeta_{\mathbb{Q}(\mu_t)}$ denotes the Dedekind $\zeta$-function of the cyclotomic field $\mathbb{Q}(\mu_t)$. Then $L_d(s)/Z_d(s)$ has an analytic continuation to the half plane $\Re(s) > 1/2$ and is nonzero at $s = 1$.

Proof. The key ideas of the proof are in [18, Theorem 6], although Kubota does not state a result this strong. See also [10, Proposition 5.2] for a proof in the case $d = 3$ that works for all primes $d$. We prove the theorem in general by showing that $L_d(s)/Z_d(s)$ is given by an Euler product that converges on the half plane $\Re(s) > 1/2$.

Suppose $m \in \mathbb{Z}_{>0}$ factors into powers of distinct primes as $m = \prod_t q_t$. Using the isomorphism $(\mathbb{Z}/m\mathbb{Z})^* \cong \prod_t (\mathbb{Z}/q_t\mathbb{Z})^*$, we see that every Dirichlet character of conductor $m$ and order dividing $d$ can be written uniquely as a product of Dirichlet characters of conductor $q_t$ and order dividing $d$. It follows that the function $\tilde{F}_d(m)$ is multiplicative in $m$, and thus $L_d(s)$ can be written as an Euler product

$$L_d(s) = \prod_p L_{d,p}(s) \quad \text{where} \quad L_{d,p}(s) := \sum_{k=0}^{\infty} \tilde{F}_d(p^k)p^{-ks}.$$

It is straightforward to verify that:

- $L_{d,p}(s)$ is a (finite degree) polynomial in $p^{-s}$,
- $\tilde{F}_d(p) = \gcd(d, p-1) - 1$,
- if $p \nmid d$ and $k \geq 2$ then $\tilde{F}_d(p^k) = 0$.

In particular

$$L_{d,p}(s) = 1 + (\gcd(d, p-1) - 1)p^{-s} \quad \text{if } p \nmid d. \quad (12.3)$$

Let $L_{d,p}(s)$ denote the Euler factor at $p$ of the Dirichlet series $L_d(s)^{-1}$. If $t \in \mathbb{Z}_{>0}$ and $p \nmid t$ let $\nu(p, t)$ denote the order of $p$ in $(\mathbb{Z}/t\mathbb{Z})^*$. Then if $p \nmid d$ we have

$$L_{d,p}(s) = (1 - p^{-s})L_{d,p}(s) = \prod_{t|d} (1 - p^{-\nu(p, t)})^{\psi(t)/\nu(p, t)} = \prod_{n|d} (1 - p^{-ns})^{\sum_{d|\gcd(d, p-1)} \psi(t)/n}.$$  \hspace{1cm} (12.4)

The exponent of the $n = 1$ term is $\sum_{t|d, t(p-1)} \psi(t) = \gcd(d, p-1)$, so

$$L_{d,p}(s) = (1 - p^{-s})^{\gcd(d, p-1) - 1} \prod_{n|d, n > 1} (1 - p^{-ns})^{\sum_{d|\gcd(d, p-1)} \psi(t)/n}. \quad (12.4)$$

It follows from (12.3) and (12.4) that the product $\prod_p L_{d,p}(s)L_{d,p}(s)$ converges on the half plane $\Re(s) > 1/2$ and is not zero at $s = 1$. This completes the proof of the theorem. \qed
Corollary 12.5. The Dirichlet series $L_d(s)$ of Theorem 12.2 converges on the half plane $\Re(s) > 1$, and has a pole of order $\tau(d) - 1$ at $s = 1$, where $\tau(d)$ is the number of divisors of $d$.

Proof. This is immediate from Theorem 12.2, since each $\zeta_{\mathbb{Q}(\mu_i)}$ has a simple pole at $s = 1$. \hfill \Box

Since $F_d(m) \leq F'_d(m)$, the following is an immediate consequence of Corollary 12.5.

Corollary 12.6. If $g(X) \ll X^{-s}$ for some $s > 1$, then $S_d(g)$ is finite.

Recall that if $p(X), q(X) : \mathbb{R}_{>r} \to \mathbb{R}_{>0}$ are functions for some real number $r$, we write $p(X) \sim q(X)$ to mean that $\lim_{X \to \infty} p(X)/q(X)$ exists and is nonzero.

Lemma 12.7. Suppose $r \in \mathbb{R}$ and $g : \mathbb{R}_{>r} \to \mathbb{R}_{>0}$ is decreasing, continuously differentiable, and satisfies $-g'(X) \ll g(X)/X$. Then for every prime $p$, we have

$$S_p(g; X) \ll_{p,g} \sum_{m=r}^{X} g(m) \sim \int_{r}^{X} g(t)dt.$$ \hfill (12.8)

Proof. Fix a prime $p$, and let $A(X) := \sum_{m=r}^{X} F_p(m)$. Since $p$ is prime, Corollary 12.5 shows that the Dirichlet series $L_p(s)$ of Theorem 12.2 has a simple pole at $s = 1$. Hence, a standard Tauberian theorem (e.g., [21, Theorem I, Appendix 2]) shows that

$$A(X) = cX + o(X)$$ \hfill (12.8)

where $c$ is the residue at $s = 1$ of $L_p(s)$.

Let

$$\tilde{S}_p(g; X) := \sum_{m=r}^{X} F_p(m)g(m).$$

Abel summation (see for example [2, Theorem 4.2]) gives

$$\tilde{S}_p(g; X) = A(X)g(X) - A(r)g(r) - \int_{r}^{X} A(t)g'(t)dt.$$ \hfill (12.9)

Integration by parts gives

$$\int_{r}^{X} g(t)dt = Xg(X) - rg(r) - \int_{r}^{X} tg'(t)dt.$$ \hfill (12.10)

Let $B(X) := A(X) - cX$ with $c$ as in (12.8). Subtracting $c$ times (12.10) from (12.9) gives

$$\tilde{S}_p(g; X) - c\int_{r}^{X} g(t)dt = B(X)g(X) - B(r)g(r) - \int_{r}^{X} B(t)g'(t)dt.$$ \hfill (12.11)

Since $g$ is decreasing and $-g'(X) \ll g(X)/X$, we have

$$\left|\int_{r}^{X} B(t)g'(t)dt\right| \leq -\int_{r}^{X} |B(t)|g'(t)dt \ll \int_{r}^{X} \frac{|B(t)|}{t}g(t)dt.$$ \hfill (12.12)

In addition, since $B(X) = o(X)$ and $g$ is decreasing,

$$B(X)g(X) = o(Xg(X)) \quad \text{and} \quad Xg(X) \leq \int_{r}^{X} g(t)dt + rg(r).$$ \hfill (12.13)

Combining (12.11) with (12.12) and (12.13) shows that

$$\tilde{S}_p(g; X) - c\int_{r}^{X} g(t)dt = \begin{cases} o(\int_{r}^{X} g(t)dt) & \text{if } \int_{r}^{X} g(t)dt \text{ is unbounded,} \\ O(1) & \text{if } \int_{r}^{X} g(t)dt \text{ is bounded.} \end{cases}$$

Since $S_p(g, X) \leq \tilde{S}_p(g, X)$, this proves the lemma. \hfill \Box

Lemma 12.14. There is an absolute constant $C$ such that for every $m \geq 3$ we have $\varphi(m) \log \log(m) > m/C$.

Proof. This follows from [15, Theorem 328]. \hfill \Box
**Proposition 12.15.** Suppose $M \in \mathbb{R}_{>0}$, $t : \mathbb{Z}_{\geq 3} \to \mathbb{R}$ is a function satisfying $t(d) \gg \log(d)$, and $\sigma_d$, $\tau_d$ are sequences of real numbers such that $\{t(d)\sigma_d : d \geq 3\}$ is bounded and $\lim_{d \to \infty} \tau_d = 0$. Then

$$
\sum_{d: t(d) (1 - \sigma_d) > 1} \sum_{m=1}^{\infty} F_d(m) \left( \frac{M d^{\sigma_d} \log(m)^{\tau_d}}{\varphi(m) \log(m)} \right)^{t(d)}
$$

converges.

**Proof.** Let

$$T_d := \sum_{m=1}^{\infty} F_d(m) \left( \frac{M d^{\sigma_d} \log(m)^{\tau_d}}{\varphi(m) \log(m)} \right)^{t(d)}.
$$

We will show below that $T_d$ converges if $t(d) (1 - \sigma_d) > 1$. We want to show that in fact $\sum_{d: t(d) (1 - \sigma_d) > 1} T_d$ converges. Note that $F_d(m) = 0$ unless $m > d$. Fix $k$ such that if $d \geq k$, then

- $t(d) (1 - \sigma_d) > 3$,
- $\tau_d < 1/2$,
- $\log \log(x)/ \log(x)^{1/2}$ is decreasing on $[d, \infty)$.

Let $C$ be the positive constant from Lemma 12.14. Using Lemma 12.14, we have for every $d$

$$T_d = (M d)^{t(d)} \sum_{m > d} F_d(m) \left( \frac{m^{\sigma_d}}{\varphi(m) \log(m)^{(1 - \tau_d)}} \right)^{t(d)} < (M d)^{t(d)} \sum_{m > d} F_d(m) \left( \frac{C \log(m)}{m^{1 - \sigma_d} \log(m)^{1 - \tau_d}} \right)^{t(d)}.
$$

We have $F_d(m) \leq |\text{Hom}((\mathbb{Z}/m\mathbb{Z})^\times, \mu_d)| < m$, and using (12.17) we get

$$
\sum_{d \geq k} T_d \leq \sum_{d \geq k} (M d)^{t(d)} \sum_{m > d} m \left( \frac{C \log(m)}{m^{1 - \sigma_d} \log(m)^{1/2}} \right)^{t(d)}
\leq \sum_{d \geq k} \left( \frac{M C d \log(d)}{\log(d)^{1/2}} \right)^{t(d)} \sum_{m > d} m^{1 - t(d) (1 - \sigma_d)}
\leq \sum_{d \geq k} \left( \frac{M C d \log(d)}{\log(d)^{1/2}} \right)^{t(d)} \frac{d^2 - t(d) (1 - \sigma_d) - 2}{t(d) (1 - \sigma_d) - 2}
\leq \sum_{d \geq k} \left( \frac{M C \log(d) \log(d)^{1/2}}{\log(d)^{1/2}} \right)^{t(d)} d^{2 + b}
$$

where $b = \sup\{t(d)\sigma_d : d \geq 3\}$.

Fix an integer $n \geq b + 4$. Since $t(d) \gg \log(d)$, there is a positive constant $C'$ such that for all large $d$ we have

$$
\left( \frac{\log(d)^{1/2}}{M C \log(d)} \right)^{t(d)} \geq \log(d)^{C' \log(d)} = e^{C' \log(d) \log(d)} = d^{C' \log(d)} > d^n.
$$

Thus the sum (12.19) converges.

For the finitely many $d < k$ with $t(d) (1 - \sigma_d) > 1$, (12.18) and Corollary 12.6 applied with

$$
g(x) = \left( \frac{\log \log(x)}{x^{1 - \sigma_d} \log(x)^{1 - \tau_d}} \right)^{t(d)}
$$

shows that $T_d$ converges. This completes the proof of the proposition. \[\square\]

**Acknowledgments**

We thank Jon Keating and Asbjørn Nordencroft for helpful conversations about the random matrix theory conjectures.

**Declaration of Interest**

No potential conflict of interest was reported by the authors.
Funding

This material is based upon work supported by the National Science Foundation under grants DMS-1302409 and DMS-1500316.

References

[1] Agashe, A., Ribet, K., Stein, W. A. (2006). The Manin constant. Pure Appl. Math. Q., 2, 617–636.
[2] Apostol, T. (1976). Introduction to Analytic Number Theory. Undergraduate Texts in Mathematics, New York: Springer.
[3] Atkin, A. O. L., Lehner, J. (1970). Hecke operators on $\Gamma_0(m)$. Math. Annalen, 185, 134–160.
[4] Bettin, S., Drappeau, S. Limit laws for rational continued fractions and value distribution of quantum modular forms. http://sary-aurelien.drappeau.perso.luminy.univ-amu.fr/documents/LL-cfrac.pdf
[5] Blomer, V., Fouvry, E., Kowalski, E., Michel, P., Miličević, D., Sawin, W. The second moment theory of families of L-functions. https://arxiv.org/abs/1804.01450, to appear in Memoirs of the AMS.
[6] Coates, J. (2012). The enigmatic Tate-Shafarevich group. AMS/IP Stud. Adv. Math. 51, 43–50.
[7] Conrey, B., Keating, J., Rubinstein, M., Snaith, N. (2002). On the frequency of vanishing of quadratic twists of modular L-Functions. In: Number Theory for the Millennium I. Natick, MA: A K Peters, Ltd., 301–315.
[8] Constantinescu, P. Distribution of modular symbols in $\mathbb{H}^3$. Int. Math. Res. Not., https://doi.org/10.1093/imrn/rnaa241.
[9] Cremona, J. E. (1992). Algorithms for Modular Elliptic Curves. Cambridge: Cambridge University Press.
[10] David, C., Fearnley, J., Kisilevsky, H. (2004). On the vanishing of twisted L-functions of elliptic curves. Exp. Math. 13: 185–198.
[11] David, C., Fearnley, J., Kisilevsky, H. (2007). Vanishing of L-functions of elliptic curves over number fields. In: Ranks of Elliptic Curves and Random Matrix Theory, London Math. Soc. Lecture Note Ser., Vol. 341. Cambridge: Cambridge Univ. Press, 247–259.
[12] Diamantis, N., Hoffstein, J., Kiral, E., Lee, M. (2020). Additive twists and a conjecture by Mazur, Rubin, and Stein. J. Number Theory 209: 1–36.
[13] Fearnley, J., Kisilevsky, H. (2010). Critical values of derivatives of twisted elliptic L-functions, Exp. Math. 19: 149–160.
[14] Fearnley, J., Kisilevsky, H., Kuo, T. (2012). Vanishing and non-vanishing Dirichlet twists of L-functions of elliptic curves. J. London Math. Soc. (2) 86: 539–557.
[15] Hardy, G. H., Wright, E. M. (1960). An Introduction to the Theory of Numbers, 4th ed. London: Oxford University Press.
[16] Kato, K. (2004). p-adic Hodge theory and values of zeta functions and modular forms. Cohomologies $p$-adiques et applications arithmétiques III, Astérisque 295, 117–290.
[17] Keating, J. P., Snaith, N. C. (2000). Random Matrix Theory and $\zeta(1/2 + it)$, Comm. Math. Phys. 214: 57–89.
[18] Kubota, T. (1956). Density in a family of abelian extensions, Proceedings of the international symposium on algebraic number theory, Tokyo & Nikko (1955) 77–91, Tokyo: Science Council of Japan.
[19] Lee, J., Sun, H.-S. Dynamics of continued fractions and distribution of modular symbols. https://arxiv.org/abs/1902.06277.
[20] Mazur, B., Rubin, K., with an appendix by M. Larsen (2018). Diophantine stability. American J. Math. 140: 571–616.
[21] Narkiewicz, W. (2004). Elementary and Analytic Theory of Algebraic Numbers, 3rd ed. Springer Monographs in Mathematics, Springer.
[22] Nördell, A. (2021). Central values of additive twists of cuspidal L-functions. J. Reine Angew. Math. 776: 255–293.
[23] Petridis, Y. N., Risager, M. S. (2004). Modular symbols have a normal distribution. Geom. Funct. Anal. 14: 1013–1043.
[24] Petridis, Y. N., Risager, M. S. (2018). Arithmetic statistics of modular symbols. Invent. Math. 212: 997–1053.
[25] Rohrlich, D. (1984). On L-functions of elliptic curves and cyclotomic towers. Invent. Math. 75: 404–423.
[26] Sun, H.-S. (2021). A proof of the conjecture of Mazur-Rubin-Stein. Bull. Korean Math. Soc. 58: 163–170.