J. Ramanujan Math. Soc. **20,** No.4 (2005) 283–339

Numerical tests of modularity

Andrew R. Booker

Mathematics Department, 530 Church Street, University of Michigan, Ann Arbor, MI 48109
e-mail: arbooker@umich.edu

Communicated by: V. Kumar Murty

Received: July 22, 2005

Abstract. We propose some numerical tests for identifying L-functions of automorphic representations of $\mathrm{GL}(r)$ over a number field. We then apply the tests to various conjectured automorphic L-functions, providing evidence for their modularity and the associated Riemann hypotheses. Our chief examples are the Hasse–Weil L-functions attached to curves of genus 2 over $\mathbb{Q}$ and to elliptic curves over $\mathbb{Q}(\sqrt{-1})$. We discuss also three miscellaneous applications. The first two include the L-functions of high symmetric powers of Ramanujan's Δ and the modular form in $S_2(\Gamma_0(11))$. The third application is an even 2-dimensional icosahedral Galois representation over $\mathbb{Q}$, which conjecturally corresponds to a Maass form of eigenvalue $\frac{1}{4}$.

1. Introduction

In 1955, at the International Symposium on Algebraic Number Theory in Tokyo, Yutaka Taniyama hinted at a link between the coefficients of certain Hasse–Weil zeta functions of elliptic curves and the Fourier coefficients of certain modular forms. That there could be such a connection between such broadly separated areas of mathematics seemed dubious; but over time and after much work by Shimura [41] and a paper of Weil [48] making the link more plausible, it became known as the Taniyama–Shimura conjecture. Now, more than forty years later, it is the celebrated

Theorem [Wiles, Taylor, Breuil, Conrad, Diamond]. *Let E be an elliptic curve defined over $\mathbb{Q}$ of conductor N_E, and put $\lambda_E(p) = \frac{p+1-\#E(\mathbb{F}_p)}{\sqrt{p}}$. Then*

$$L(s,E) = \prod_{p\nmid N_E} \frac{1}{1-\lambda_E(p)p^{-s}+p^{-2s}} \cdot \prod_{p\mid N_E} \frac{1}{1-\lambda_E(p)p^{-s}} = L(s,f),$$

for some $f \in S_2(\Gamma_0(N_E))$.

(Here and throughout we use the analyst's normalization of L-functions, so that the functional equation relates s to $1-s$.) Expectations have since been vastly broadened, with the zeta functions of algebraic varieties conjecturally related to automorphic representations of reductive groups. Despite the recent success resolving Taniyama–Shimura, much more remains conjecture than is known.

In this paper we propose some numerical tests for identifying L-functions of automorphic representations of $\mathrm{GL}(r)$ over a number field K. We then apply the tests to various conjectured automorphic L-functions, providing evidence for their modularity and associated Riemann hypotheses. While it is pure speculation to say so, if the computing technology of today had been available in 1955, it is possible that with these experiments mathematicians of the day might have been more ready to accept Taniyama–Shimura. It is our hope that the techniques developed here, more than the actual examples, will continue to be of use in providing evidence for related conjectures. Therefore, our emphasis will be on tests that are easy to perform and work in the greatest possible generality.

Our chief examples, for which we carry out all details, are the L-functions attached to curves of genus 2 over $\mathbb{Q}$ and to elliptic curves over $\mathbb{Q}(\sqrt{-1})$. In the final section we also consider some conjectural cases of Langlands' functoriality, including L-functions of high symmetric powers of forms on $\mathrm{GL}(2)$, and that of an even icosahedral Galois representation. First, we recall some properties of the objects of interest.

1.1 Automorphic L-functions

Let K be a number field, $\mathbb{A}_K$ its ring of adeles, and π an automorphic representation of $\mathrm{GL}_r(\mathbb{A}_K)$. Associated to π is an L-function given by an Euler product (over primes $\mathfrak{p}$ of K),

$$L(s,\pi)=\prod_{\mathfrak{p}\notin S}\prod_{i=1}^{r}\frac{1}{1-\alpha_\pi^{(i)}(\mathfrak{p})(N_{K/\mathbb{Q}}\mathfrak{p})^{-s}}\cdot\prod_{\mathfrak{p}\in S}L_{\mathfrak{p}}(s,\pi). \tag{1.1}$$

Here S is a finite set of primes. In words, for $\mathfrak{p}\notin S$, the local factor at $\mathfrak{p}$ is the reciprocal of a polynomial of degree r in $(N_{K/\mathbb{Q}}\mathfrak{p})^{-s}$, with reciprocal roots $\alpha_\pi^{(i)}(\mathfrak{p})\in\mathbb{C}$ called the *Satake parameters* of π. The product of these terms alone (or over any set of all but finitely many primes, including those in S) is called a *partial L-function*, denoted $L_S(s,\pi)$. For $\mathfrak{p}\in S$, the local factors $L_{\mathfrak{p}}(s,\pi)$ are again given by the reciprocal of a polynomial, but with degree less than r.

It is known that the product in (1.1) converges for $\Re s$ sufficiently large. Furthermore, π has a complete L-function of the form

$$\Lambda(s,\pi) := L(s,\pi)\gamma(s,\pi) := L(s,\pi) \prod_{i=1}^{r[K:\mathbb{Q}]} \Gamma_{\mathbb{R}}(s+\mu_\pi^{(i)}), \tag{1.2}$$

where $\Gamma_{\mathbb{R}}(s) = \pi^{-s/2}\Gamma(s/2)$, and the $\mu_\pi^{(i)}$ are numbers associated with π, called *archimedean parameters*. (One may think of the terms of the product in (1.2) as local factors corresponding to the archimedean places of K.) The complete L-function then has meromorphic continuation to the complex plane, with at most finitely many poles that are well understood, and satisfies a functional equation relating π to the contragredient representation $\widetilde{\pi}$:

$$\Lambda(s,\pi) = \epsilon_\pi N_\pi^{1/2-s}\Lambda(1-s,\widetilde{\pi}), \tag{1.3}$$

where ϵ_π is a complex number of magnitude 1, and N_π is a positive integer called the *conductor* of π.

1.2 L-functions associated to projective curves

Now let C be a projective curve of genus g defined over K which is non-singular over the algebraic closure $\overline{K}$. For each prime $\mathfrak{p}$ of K, one may consider the reduction $C_\mathfrak{p}$ of C modulo $\mathfrak{p}$. This is a curve whose defining equations have coefficients in the finite field $k_\mathfrak{p} = \mathcal{O}_K/\mathfrak{p}$. Furthermore, there is a non-zero ideal $\mathfrak{f}_C$ called the *conductor* of C such that for all $\mathfrak{p}$ not dividing $\mathfrak{f}_C$ (i.e. for all but finitely many primes), $C_\mathfrak{p}$ is non-singular over all finite extensions of $k_\mathfrak{p}$.

Then, one associates to $C_\mathfrak{p}$ the zeta function

$$Z_{C_\mathfrak{p}}(t) = \exp\left(\sum_{k=1}^{\infty} \frac{N_k t^k}{k}\right), \tag{1.4}$$

where N_k is the number of points of $C_\mathfrak{p}$ over an extension of degree k of $k_\mathfrak{p}$. It is known that this function is rational in t, of the form

$$Z_{C_\mathfrak{p}}(t) = \frac{P_{C_\mathfrak{p}}(t)}{(1-t)(1-qt)}, \tag{1.5}$$

where $q = N_{K/\mathbb{Q}}\mathfrak{p}$. Moreover the polynomial $P_{C_\mathfrak{p}}(t)$ has integer coefficients, and satisfies

(1) $\deg P_{C_\mathfrak{p}} = 2g$.
(2) $P_{C_\mathfrak{p}}$ factors over $\mathbb{C}$ as $P_{C_\mathfrak{p}}(t) = (1-a_1t)\cdots(1-a_{2g}t)$.
(3) The map $a \mapsto q/a$ is a bijection of the reciprocal roots $a_1,\ldots,a_{2g}$.
(4) $|a_i| = q^{1/2}$.

Note that property 4 is a special case of the function field analogue of the Riemann hypothesis.

To define the Hasse–Weil L-function of C, one combines the local zeta functions into the Euler product

$$L(s, C) = \prod_{\mathfrak{p} \nmid \mathfrak{f}_C} \frac{1}{P_{C_\mathfrak{p}}((N_{K/\mathbb{Q}}\mathfrak{p})^{-(s+1/2)})}, \tag{1.6}$$

where $P_{C_\mathfrak{p}}$ is the non-trivial part of the zeta function of $C_\mathfrak{p}$, as above. (Note again that this definition differs from the usual normalization by the shift of 1/2.)

Then it is conjectured that $L(s, C)$ agrees with the partial L-function $L_S(s, \pi)$ (where S is the set of primes dividing $\mathfrak{f}_C$) of an automorphic representation π of $\mathrm{GL}(2g)$ over K. Furthermore, the data N_π and $\mu_\pi^{(i)}$ of the complete L-function $\Lambda(s, \pi)$ of the conjectured representation have explicit descriptions in terms of the curve C [36]; in particular, the conductors are related by $N_\pi = (N_{K/\mathbb{Q}}\mathfrak{f}_C)|d_{K/\mathbb{Q}}|^{2g}$, where $d_{K/\mathbb{Q}}$ is the discriminant of K.

As discussed above, this is now known for elliptic curves over $\mathbb{Q}$ [53,47,8] and in a few other special cases, including some elliptic curves over totally real fields [44,45]. The cases we consider are perhaps the simplest ones where there is so far little theoretical evidence in support of these conjectures. There is already much numerical evidence collected by Cremona et al. [14] for elliptic curves of low conductor over imaginary quadratic fields of class number 1. One hopes eventually to be able to prove modularity of specific examples in this way using results of Harris, Soudry and Taylor [20,46] and a comparison of ℓ-adic Galois representations. Our work differs from theirs, however, in that they start with a form on a hyperbolic 3-manifold, computed using modular symbols, and look for a corresponding elliptic curve with matching Fourier coefficients (see below); our method goes in the opposite direction, testing the modularity of an arbitrary elliptic curve, without having to find the corresponding form. In practical terms, this allows us to consider curves of much larger conductor.

1.3 The distribution of Fourier coefficients

Note that the Euler products (1.1) and (1.6) may be multiplied out to Euler products over rational primes,

$$\prod_p (1 + \lambda(p)p^{-s} + \lambda(p^2)p^{-2s} + \cdots), \tag{1.7}$$

and further to Dirichlet series

$$\sum_{n=1}^{\infty} \lambda(n)n^{-s}. \tag{1.8}$$

The numbers $\lambda(n)$ in (1.7) and (1.8) are called *Fourier coefficients*. Now, as a first step to determining numerically if a given L-function is modular, one should check the distribution of its Fourier coefficients $\lambda(p)$ at primes p, since the coefficients of automorphic L-functions have expected equidistribution properties. For example, Dirichlet's theorem on primes in arithmetic progressions was the first example of such a law. For generic elliptic curve L-functions, this is the famous conjecture of Sato–Tate. Serre [38] has formulated the conjecture as follows.

Let E be an elliptic curve over K. Then the denominator of (1.6) (with E in place of C) factors as

$$(1 - e^{i\theta(\mathfrak{p})}(N_{K/\mathbb{Q}}\mathfrak{p})^{-s})(1 - e^{-i\theta(\mathfrak{p})}(N_{K/\mathbb{Q}}\mathfrak{p})^{-s}), \tag{1.9}$$

for some angle $\theta(\mathfrak{p}) \in [0, \pi]$. (This is equivalent to the factorization given above for $P_{E_\mathfrak{p}}$, but with the reciprocal roots normalized to be of magnitude 1.) To the angle $\theta(\mathfrak{p})$ one associates the conjugacy class in $\mathrm{SU}(2)$ of the element

$$\begin{pmatrix} e^{i\theta(\mathfrak{p})} & 0 \\ 0 & e^{-i\theta(\mathfrak{p})} \end{pmatrix}. \tag{1.10}$$

Serre then shows that the angles $\theta(\mathfrak{p})$ are equidistributed with respect to the Sato–Tate measure $\frac{2}{\pi}\sin^2\theta\, d\theta$ (which is the Haar measure $\mu(\mathrm{SU}(2)^\#)$ on the space of conjugacy classes in $\mathrm{SU}(2)$) if and only if the symmetric power L-functions $L(s, \mathrm{Sym}^k(E))$, $k \geq 1$ are holomorphic and non-vanishing for $\Re s \geq 1$. For most curves (those without complex multiplication), this is expected to be true.

Moreover, one has an ℓ-adic representation on the Tate module of E,

$$\phi_\ell : \mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}\, T_\ell(E). \tag{1.11}$$

The reciprocal roots a_i defined above arise naturally as the eigenvalues of the Frobenius endomorphism $\mathrm{Frob}_\mathfrak{p}$ under this map. Serre proves that the image of ϕ_ℓ is open in the ℓ-adic topology.

More generally, to a projective curve C of genus g over K, one may associate an ℓ-adic representation on the Tate module of the Jacobian variety $\mathrm{Jac}\, C$. For almost all primes $\mathfrak{p}$ of K, the reciprocal roots in (1.6) (or eigenvalues of Frobenius) define a conjugacy class in $\mathrm{USp}(2g)$. As generalized by Langlands [32], the equidistribution of these classes with respect to the Haar measure $\mu(\mathrm{USp}(2g)^\#)$ is equivalent to the holomorphicity and non-vanishing for $\Re s \geq 1$ of the L-functions $L(s, \rho(C))$ associated to each non-trivial irreducible representation ρ of $\mathrm{USp}(2g)$. In accordance with the Langlands philosophy, this should be true (for most curves) if the L-function $L(s, C)$ is modular.

Now, assuming that the reciprocal roots corresponding to different primes of K lying above a rational prime p are independent, this in particular implies a distribution for the Fourier coefficients $\lambda_C(p)$. In Section 2 we compute histograms of the prime Fourier coefficients of many curves of genus 2 over $\mathbb{Q}$

and elliptic curves over $\mathbb{Q}(\sqrt{-1})$, and compare them with these expectations. In this way, we will be able to distinguish these two types of L-functions, both of which are given by Euler products of degree 4 over $\mathbb{Q}$.

1.4 Converse theorems

Once we have established the expected distribution of coefficients for the L-functions under consideration, the question remains of how to distinguish the coefficients of automorphic L-functions from "typical" numbers with the same statistics. Hecke [21] pointed the way with his converse theorem

Theorem [Hecke]. *Let $k \in \mathbb{Z}$ and let $L(s) = \sum_{n=1}^{\infty} \lambda(n)n^{-s}$ be absolutely convergent in a right half-plane. Suppose $\Lambda(s) = (2\pi)^{-s}\Gamma(s+(k-1)/2)L(s)$ has analytic continuation to the entire complex plane, is bounded in vertical strips, and satisfies the functional equation*

$$\Lambda(s) = (-1)^{k/2}\Lambda(1-s). \tag{1.12}$$

Then $L(s) = L(s, f)$ for some $f \in S_k(\mathrm{SL}_2(\mathbb{Z}))$.

Thus, the modularity of a form $f \in S_k(\mathrm{SL}_2(\mathbb{Z}))$ is equivalent to the analytic continuation and precise functional equation of its L-function.

Hecke's result, which is for modular forms of full level, is atypical in that only one L-function is involved. Subsequent generalizations, by Weil [48] to congruence subgroups, Jacquet, Piatetski-Shapiro and Shalika [23,24] to automorphic forms on $\mathrm{GL}(3)$, and Cogdell and Piatetski-Shapiro [13] to $\mathrm{GL}(r)$, all require analytic continuation and functional equation of a family of twisted L-functions (by $\mathrm{GL}(r-2)$ in the most general case). However, it is believed that one twisted functional equation should be sufficient to imply modularity, this essentially being the modularity conjecture for the Selberg class [35]. Razar [33] made progress toward that end, giving a version of Weil's theorem for congruence subgroups requiring only finitely many twists. In [5] we generalized Weil's result in a different way, allowing one to relax the conditions on the twisted L-functions in certain cases by allowing poles. Thus, some infinite families of examples are now known to require only one twisted functional equation. Although this result does not apply to our examples of L-functions associated to curves, it does apply to the Galois representation example of Section 6, and we will in general regard evidence obtained from a single twist as a strong indication of modularity.

Moreover, by the strong multiplicity one theorem, automorphic representations of $\mathrm{GL}(r)$ are determined by all but finitely many of the Euler factors of their associated L-functions. With that in mind, we aim to design tests that can be used with knowledge of only a partial L-function. In fact the test of

Section 5 gives an algorithm that can recover any missing Euler factors given only a partial L-function. This is useful because in practice there are a few Euler factors that are either unknown or difficult to compute. For example, we do not compute the factors of Hasse–Weil L-functions at primes dividing the conductor $\mathfrak{f}_C$.

Note that our test of the distribution of Fourier coefficients above, as well as subsequent tests below, use the coefficients in the Dirichlet series over rational primes. Since we carry out each test for only one twist in general, we thus make no distinction between the L-function of an automorphic representation over a number field and the corresponding L-function over $\mathbb{Q}$. However, by Langlands' functoriality, the latter L-function should also come from a (non-cuspidal) automorphic representation, over $\mathbb{Q}$, so we expect this not be an issue. Likewise, automorphic representations on other groups are conjectured to have functorial transfers to $\mathrm{GL}_r(\mathbb{A}_{\mathbb{Q}})$, so the tests that we describe may be tailored to work in such settings as well.

1.5 Smooth sums and analytic continuation

We turn now to the problem of numerically measuring the analytic continuation of an L-function. Let

$$L(s) = \sum_{n=1}^{\infty} \lambda(n) n^{-s} \tag{1.13}$$

be a Dirichlet series, absolutely convergent in some right half-plane. By a smooth sum of the coefficients $\lambda(n)$ we mean a sum of the form

$$S(X) = \frac{1}{\sqrt{X}} \sum_{n=1}^{\infty} \lambda(n) F(n/X), \tag{1.14}$$

for $X > 0$ and F a Schwartz function on $(0, \infty)$, meaning F is smooth and, together with its derivatives, is of rapid decay at 0 and ∞. One may think of F, for example, as a bump function of compact support, so that $S(X)$ measures the sum of $\lambda(n)$ for n of size about X.

Now, (1.13) and (1.14) are related by a Mellin transform. Indeed, if $\widetilde{F}$ is the Mellin transform of F, we have

$$\begin{aligned} S(X) &= \frac{1}{\sqrt{X}} \sum_{n=1}^{\infty} \lambda(n) \frac{1}{2\pi i} \int_{\Re s=\sigma} \left(\frac{n}{X}\right)^{-s} \widetilde{F}(s)\, ds \\ &= \frac{1}{2\pi i} \int_{\Re s=\sigma} L(s) X^{s-1/2} \widetilde{F}(s)\, ds, \end{aligned} \tag{1.15}$$

where to justify the exchange of sum and integral we assume σ is large. Similarly, we have the inverse identity

$$L(s)\widetilde{F}(s) = \int_0^\infty S(X)X^{1/2-s}\frac{dX}{X}. \tag{1.16}$$

From (1.15) it is clear that if $L(s)$ extends to an entire function, with at most polynomial growth in vertical strips, then $S(X)$ is Schwartz. Conversely, if $S(X) = O(X^\sigma)$ then $L(s)\widetilde{F}(s)$ is analytic in $\Re s > 1/2+\sigma$. Thus, the analytic continuation of $L(s)$ is measured by the rate of growth or decay of the smooth sums $S(X)$.

Now, suppose that the numbers $\lambda(n)$ are selected randomly and independently from a fixed distribution of mean 0 on $[-1, 1]$, say. Then one would expect (1.14) typically to be of size 1; in fact the central limit theorem then implies that with probability 1, the random Dirichlet series (1.13) converges for $\Re s > 1/2$, where it has a natural boundary [27].

It is in this sense that the Fourier coefficients $\lambda_\pi(n)$ of an automorphic representation π differ from typical random numbers; the analytic continuation and functional equation of the L-function $L(s, \pi)$ may be viewed as a certain regularity governing the coefficients. More precisely, the associated smooth sums $S(X)$ are rapidly decreasing once X is much larger than the *analytic conductor* of π

$$C_\pi := N_\pi \prod_{i=1}^{r[K:\mathbb{Q}]} \frac{1+|\mu_\pi^{(i)}|}{2\pi}. \tag{1.17}$$

(Note that our definition differs from that first given by Iwaniec and Sarnak [22] by the factors of 2π.) This is analogous to the fact that the zeta functions in the function field setting are polynomials, i.e. their terms vanish beyond the degree, which plays the role of conductor.

As we will see in Section 3, using the functional equation for $L(s, \pi)$, this decay can be explicitly predicted. Thus, computing the sum (1.14) for $X \gg C_\pi$ yields a test for identifying the coefficients of automorphic L-functions. The method is robust in the sense that if applied to numbers that are not coefficients of an automorphic L-function, the test bears this out quickly; an example is given in Section 3. However, this test has limitations, and gives weak bounds when there are unknown Euler factors at large primes. In such situations, as is the case with many of our main examples, the approximate functional equation tests described below are preferable.

1.6 Sums over primes and the Riemann hypothesis

The alert reader will note that we have argued on the one hand that the Fourier coefficients $\lambda_\pi(p)$ should act randomly (independently, moreover), and on the other that there is a regularity to them. This is not a contradiction; rather it is the difference between $\lambda_\pi(p)$ at primes and $\lambda_\pi(n)$ at integers. More explicitly, if we form a smooth sum out of only the coefficients $\lambda(p)$ at primes,

$$S^{\mathrm{pr}}(X) = \frac{1}{\sqrt{X}} \sum_{p \text{ prime}} (\log p)\lambda(p)F(p/X) \tag{1.18}$$

(it is customary in analytic number theory to weight the primes by $\log p$), we are measuring the analytic continuation not of $L(s)$, but its logarithmic derivative

$$-\frac{L'}{L}(s) = \sum_{p \text{ prime}} ((\log p)\lambda(p)p^{-s} + (\text{terms of order } p^{-2s})). \tag{1.19}$$

With bounded coefficients, the error term converges absolutely for $\Re s > 1/2$.

For automorphic L-functions, by computing $S^{\mathrm{pr}}(X)$, we are thus testing the Riemann hypothesis! More precisely, by a calculation including the second order terms that we carry out in Section 3, Riemann is equivalent to the bound $S^{\mathrm{pr}}(X) = O(1)$. As above, if the $\lambda(p)$ are chosen independently randomly then $L(s)$ will typically be zero-free in $\Re s > 1/2$, which is best possible; the Riemann hypothesis may thus be viewed as a statement about the independence of the numbers $\lambda_\pi(p)$. Also in Section 3, we apply this test to our main examples.

1.7 Seeing the analytic conductor

The numerical test described above is moreover a prediction that the sums (1.14) for coefficients of an automorphic L-function are of size 1 up to about the analytic conductor C_π, then decay rapidly thereafter. In particular, we would like to know that $S(X)$ is not too small for X of size C_π^δ, with $\delta > 0$. In Section 4 we establish a result to that effect for some infinite families of hyperelliptic curves of arbitrary genus over $\mathbb{Q}$.

1.8 Approximate functional equations

In our discussion of smooth sums above, all smoothing functions F were placed on an equal footing. However, for automorphic L-functions $L(s,\pi)$ there is a natural choice of smoothing function, namely the inverse Mellin transform

$$F_\pi(x) = \frac{1}{2\pi i} \int_{\Re s = \sigma} \gamma(s,\pi)x^{-s}\, ds. \tag{1.20}$$

Note that F_π is not rapidly decreasing near 0, reflecting the fact that $\gamma(s,\pi)$ has poles. However, $L(s,\pi)\gamma(s,\pi)$ is analytic, so that (1.15) is still rapidly decreasing as $X \to \infty$. As we will see in Section 5, with F_π as smoothing function, the functional equation (1.3) is manifested not only as a precise bound for $S(X)$, but as the symmetry

$$S(X) = \epsilon_\pi \overline{S(N_\pi/X)}, \tag{1.21}$$

called an *approximate functional equation*.

Since (1.21) is an exact equality, this yields a more precise test of analytic continuation and the functional equation than that of Section 3. It has the advantage that (1.21) is verifiable for X as small as $\sqrt{N_\pi}$, so that this test requires fewer coefficients of the L-series. However, verifying equation (1.21) requires knowledge of *all* local factors. In Section 5 we show how to circumvent this problem at the finite places; we consider $S(X) - \epsilon_\pi \overline{S(N_\pi/X)}$ as a quantity to be minimized, and thus determine the unknown Euler factors by solving a non-linear regression problem. This technique complements results of Dokchitser [18], who has considered similar problems.

1.9 Miscellaneous applications

Finally, in Section 6 we show results of three miscellaneous applications of the tests developed here. The first two include the L-functions of high symmetric powers of Ramanujan's Δ and the modular form in $S_2(\Gamma_0(11))$. One knows already [30] the meromorphic continuation and functional equation of these L-functions up to the ninth symmetric power, but not beyond; we test in particular their tenth symmetric powers. The third application is an even 2-dimensional icosahedral Galois representation over $\mathbb{Q}$. This conjecturally corresponds to a Maass form of eigenvalue $\frac{1}{4}$.

1.10 Summary of results and future work

Using the techniques developed in this paper, approximately 800 examples were tested for modularity and the Riemann hypothesis. The complete list of examples tested is available electronically [7]. No apparent counterexamples were found. Note, however, that the tests described here give evidence but not a certificate of modularity. An interesting question is whether it is possible in general to *prove* modularity of a candidate L-function with a finite computation. A good starting case would be the L-function of an even icosahedral Galois representation, for which the result of [5] may be useful; see also [6] for the related computational aspects. (Note that there is such an algorithm for odd icosahedral representations, using the theorem of Deligne and Serre [17].) In this case, the problem can be phrased as follows: Is the Artin conjecture for a given Galois representation decidable?

2. The distribution of Fourier coefficients

In this section we specialize the discussion of Fourier coefficients given in the introduction to elliptic curves over $\mathbb{Q}(\sqrt{-1})$ and genus 2 curves over $\mathbb{Q}$.

2.1 Elliptic curves

For an elliptic curve E over a number field K, the L-function defined in the introduction takes the form

$$L(s, E) = \prod_{\mathfrak{p} \nmid \mathfrak{f}_E} \frac{1}{1 - a_E(\mathfrak{p})(N\mathfrak{p})^{-s} + (N\mathfrak{p})^{-2s}}, \tag{2.1}$$

where, unraveling the definitions, one finds

$$a_E(\mathfrak{p}) = \frac{N\mathfrak{p} + 1 - \#E(k_{\mathfrak{p}})}{\sqrt{N\mathfrak{p}}}. \tag{2.2}$$

(Here we have omitted the factors at primes dividing the conductor; since they are finite in number, they are irrelevant to determining the distribution of coefficients.) The number $a_E(\mathfrak{p})$ is also the sum of the normalized reciprocal roots $e^{\pm i\theta(\mathfrak{p})}$, as defined in (1.9), i.e. $a_E(\mathfrak{p}) = 2\cos\theta(\mathfrak{p})$. Thus, we have the bound $|a_E(\mathfrak{p})| \leq 2$, due originally to Hasse.

For rational primes p not dividing $N\mathfrak{f}_E$, one computes the Fourier coefficients $\lambda_E(p)$ by collecting the Euler factors for $\mathfrak{p}$ dividing p. When $K = \mathbb{Q}(\sqrt{-1})$, there are three cases, depending on the factorization of p in K:

- inert primes, $p \equiv 3 \pmod 4$; then $\lambda_E(p) = 0$.
- split primes, $p \equiv 1 \pmod 4$; then $\lambda_E(p) = a_E(\mathfrak{p}) + a_E(\overline{\mathfrak{p}})$, for $\mathfrak{p}$ a prime dividing p.
- the ramified prime $p = 2$; then $\lambda_E(p) = a_E(\mathfrak{p})$, for $\mathfrak{p}$ the prime dividing 2.

We see immediately that the distribution of $\lambda_E(p)$ will have mass $1/2$ at 0. The more interesting part of the distribution is that of $\lambda_E(p)$ for $p \equiv 1 \pmod 4$. As explained in the introduction, it is conjectured that as $\mathfrak{p}$ varies, the angles $\theta(\mathfrak{p})$ vary according to the Sato–Tate measure $d\mu(\mathrm{SU}(2)^{\#}) = \frac{2}{\pi}\sin^2\theta\, d\theta$. This implies that $a_E(\mathfrak{p}) = 2\cos\theta(\mathfrak{p})$ follows the 'semi-circle' distribution $\frac{1}{2\pi}\sqrt{4 - x^2}\, dx$. Moreover, one expects $a_E(\mathfrak{p})$ and $a_E(\overline{\mathfrak{p}})$ to be independent, meaning that the $\lambda_E(p)$ should have density function given by the convolution

$$\frac{1}{4\pi^2} \int_{|x|-2}^{2} \sqrt{(4 - t^2)(4 - (|x| - t)^2)}\, dt, \tag{2.3}$$

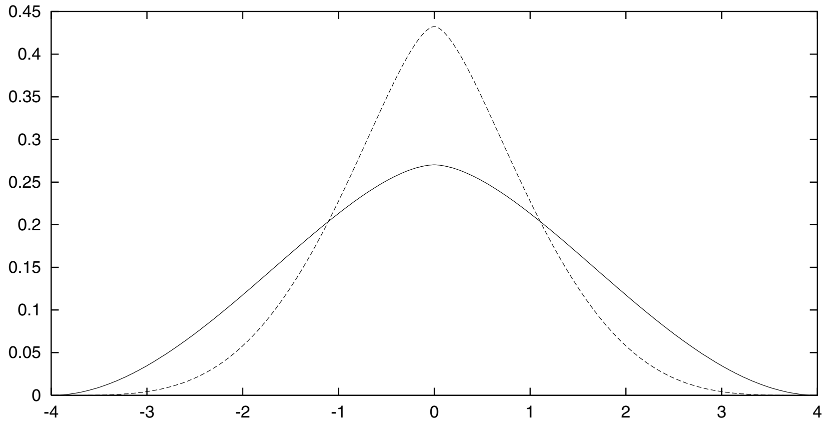

Figure 2.1. Expected distributions of Fourier coefficients $\lambda_E(p)$ for $p \equiv 1 \pmod 4$ (solid line), and $\lambda_C(p)$ for all primes (dashed line)

supported on $[-4, 4]$. By standard techniques [50], one reduces this elliptic integral to the Legendre normal form

$$\frac{4+|x|}{24\pi^2}\left[(x^2+16)E\left(\frac{4-|x|}{4+|x|}\right)-8|x|K\left(\frac{4-|x|}{4+|x|}\right)\right]. \qquad (2.4)$$

The graph of this function is shown as the solid line in Figure 2.1.

2.2 Curves of genus 2

Similarly, a curve C of genus 2 over a number field has the L-function

$$L(s, C) = \prod_{\mathfrak{p}\nmid\mathfrak{f}_C} \frac{1}{1-a_C(\mathfrak{p})(N\mathfrak{p})^{-s}+b_C(\mathfrak{p})(N\mathfrak{p})^{-2s}-a_C(\mathfrak{p})(N\mathfrak{p})^{-3s}+(N\mathfrak{p})^{-4s}}. \qquad (2.5)$$

Here, we have

$$a_C(\mathfrak{p}) = \frac{N\mathfrak{p}+1-\#C(k_\mathfrak{p})}{\sqrt{N\mathfrak{p}}} \quad \text{and} \quad b_C(\mathfrak{p}) = a_C(\mathfrak{p})^2 - \frac{(N\mathfrak{p})^2+1-\#C(\ell_\mathfrak{p})}{N\mathfrak{p}}, \qquad (2.6)$$

where $\ell_\mathfrak{p}$ is a quadratic extension of $k_\mathfrak{p}$. The coefficient $a_C(\mathfrak{p})$ is again the sum of the reciprocal roots $e^{\pm i\theta_1(\mathfrak{p})}$ and $e^{\pm i\theta_2(\mathfrak{p})}$, so the bound in this case, due to Weil, becomes $|a_C(\mathfrak{p})| \le 4$. Similar formulas hold for general genus, with the coefficients related to the numbers of points over extensions of $k_\mathfrak{p}$ via Newton's formulas.

Now, for $K = \mathbb{Q}$ we have simply $\lambda_C(p) = a_C(p)$. To compute its expected distribution, we first need the distribution of the space of conjugacy classes in

USp(4), indexed by θ_1 and θ_2, with $0 \le \theta_1 \le \theta_2 \le \pi$. This may be found, using the Weyl integration formula [29,51], to be

$$d\mu(\mathrm{USp}(4)^{\#}) = \frac{2}{\pi^2}(2\cos\theta_1 - 2\cos\theta_2)^2 \sin^2\theta_1 \sin^2\theta_2\, d\theta_1 d\theta_2. \tag{2.7}$$

A similar formula holds for general genus.

Then, the distribution of $a_C(p)$ is that of the trace:

$$\begin{aligned} &\frac{d}{dx}\frac{2}{\pi^2}\int_{\substack{0 \le \theta_1 \le \theta_2 \le \pi \\ 2\cos\theta_1 + 2\cos\theta_2 \le x}} (2\cos\theta_1 - 2\cos\theta_2)^2 \sin^2\theta_1 \sin^2\theta_2\, d\theta_1 d\theta_2 \\ &\quad = \frac{d}{dx}\frac{1}{8\pi^2}\int_{\substack{2 \ge t_1 \ge t_2 \ge -2 \\ t_1 + t_2 \le x}} (t_1 - t_2)^2 \sqrt{(4 - t_1^2)(4 - t_2^2)}\, dt_1 dt_2 \\ &\quad = \frac{1}{16\pi^2}\int_{|x|-2}^{2} (2t - |x|)^2 \sqrt{(4 - t^2)(4 - (|x| - t)^2)}\, dt. \end{aligned} \tag{2.8}$$

The result is again an elliptic integral, this time with Legendre normal form

$$\frac{4+|x|}{240\pi^2}\left[(x^4+224x^2+256)E\left(\frac{4-|x|}{4+|x|}\right) - 8|x|(x^2+24|x|+16)K\left(\frac{4-|x|}{4+|x|}\right)\right]. \tag{2.9}$$

The graph of this function is shown as the dashed line in Figure 2.1.

2.3 Notes on computation

2.3.1 Finding curves to test

In order to carry out the tests in this and subsequent sections, we first computed lists of curves of low conductor, as described below. The complete lists are available electronically [7].

Over a number field, an elliptic curve always has an affine model of the form (in the standard notation [31])

$$y^2 = x^3 - 27c_4 x - 54c_6, \tag{2.10}$$

where c_4 and c_6 are algebraic integers. This model is most convenient for computation, but has the disadvantage of being singular at primes dividing 2 and 3. Since we are most interested in curves of low conductor, it is better to consider the model

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6. \tag{2.11}$$

Table 2.1. Elliptic curves over $\mathbb{Q}(\sqrt{-1})$ (bottom) and genus 2 curves over $\mathbb{Q}$ (top).

Name	Δ	Curve
277a	277	$y^2 + y = x^5 - 2x^3 + 2x^2 - x$
1757a	1757	$y^2 + y = x^5 - x^4 - x^3 + 2x^2 - x$
9136a	-9136	$y^2 + xy = x^5 - 2x^4 - 2x^3 + 2x^2 + 3x + 1$
18690a	18690	$y^2 + xy = x^5 + x^4 - 4x^3 - 3x^2 + 3x + 3$
233a	$13 - 8i$	$y^2 + iy = x^3 + (1+i)x^2 + ix$
3577a	$21 - 56i$	$y^2 + iy = x^3 - ix^2 - x$
9657a	$21 - 96i$	$y^2 + y = x^3 - (1+i)x^2 - (1-i)x$
16970a	$-53 - 119i$	$y^2 + xy + iy = x^3 - (1+i)x^2 - x$

The coefficients of (2.10) may then be computed by a change of variables from (2.11).

To find curves of low conductor, it is most efficient to search through models of the form (2.10), then try to pull them back to the form (2.11). This is possible when c_4 and c_6 satisfy certain congruence conditions modulo 1728. The resulting curve (2.11) has discriminant given by the indefinite form

$$\Delta = \frac{c_4^3 - c_6^2}{1728} \tag{2.12}$$

and j-invariant

$$j = \frac{c_4^3}{\Delta}. \tag{2.13}$$

In general, the conductor $\mathfrak{f}_E$ divides Δ, although possibly properly. If desired, the conductor may be computed exactly by Tate's algorithm [42]; however, it will be determined as a consequence of our test from Section 5. If one believes in the truth of Hall's conjecture [43] then this method quickly finds "most" (isomorphism classes of) curves of low conductor. Applying the method, we found 390 non-isomorphic curves over $\mathbb{Q}(\sqrt{-1})$ of discriminant with norm less than 20000, after removing those with rational j-invariant. A few examples, indexed by the norm of discriminant, are listed in Table 2.1. The letters used in naming the curves are to distinguish those with the same discriminant in the extended lists of [7]; note that no attempt was made to find the "strong Weil curve" in these cases, so the notation is not standardized, and there are some isogenies.

Similarly, any curve of genus 2 over a number field is hyperelliptic, and thus has a model

$$y^2 = f(x), \tag{2.14}$$

where $f(x)$ is a monic square-free polynomial of degree 5 with algebraic integer coefficients. Again, to find curves of low conductor we search for models of the form

$$y^2 + Axy + By = f(x). \tag{2.15}$$

Warning. Unlike the case of elliptic curves, not every hyperelliptic curve over the ring of integers $\mathcal{O}_K$ takes this form. Thus, the curves we find in this way represent only a fraction of all curves of low conductor. Also, to find models (2.15), we simply performed an exhaustive search through all such equations with bounded coefficients; this is less elegant than the approach taken for elliptic curves, and not as efficient for finding curves of low conductor. Despite these restrictions, we found with this method 399 curves with absolute discriminant less than 20000. A few examples, indexed by absolute discriminant, are listed in Table 2.1.

2.3.2 Computing Fourier coefficients

Now, for a given curve we compute the Fourier coefficients via formulas (2.2) and (2.6). To perform the test of Section 3, we will need not only the coefficients at primes, but at all integers n up to some M proportional to (the absolute norm of) the discriminant. This can be done by computing all coefficients of the Euler factors in (2.1) and (2.5), and extending multiplicatively.

For elliptic curves over $\mathbb{Q}(\sqrt{-1})$, that amounts to computing $a_E(\mathfrak{p})$ for $\mathfrak{p}$ lying over inert primes p, in addition to the split and ramified ones. For inert primes, the residue field $k_\mathfrak{p}$ is isomorphic to $\mathbb{F}_{p^2}$, and thus the point count computations take longer than for split primes. Fortunately, we only need such counts for $p \leq \sqrt{M}$, so the inert primes contribute little to the total running time. On the other hand, for all split primes $p \leq M$, we count over the residue fields $k_\mathfrak{p}$ and $k_{\bar{\mathfrak{p}}}$, isomorphic to $\mathbb{F}_p$.

To perform these point counts, we used Shanks' 'baby step-giant step' algorithm [40] for computing the order of the group associated to an elliptic curve. For a count over $\mathbb{F}_p$, the algorithm has running time (probabilistically) $O(p^{1/4} \log p)$. Hence, the overall time is $O(M^{5/4})$. While there exist asymptotically faster algorithms, this is certainly the best in practice for numbers of the size that we consider.

For curves of genus 2 over $\mathbb{Q}$, computing the Euler factors in (2.5) requires one point count over $\mathbb{F}_p$ for $p \leq M$ and one over $\mathbb{F}_{p^2}$ for $p \leq \sqrt{M}$. Thus, asymptotically the same number of point counts are required for genus 2 curves over $\mathbb{Q}$ as for elliptic curves over $\mathbb{Q}(\sqrt{-1})$. However, although the Jacobian of a genus 2 curve has a group law, its dimension is greater than 1, and does not yield a fast algorithm for counting points; we are forced to use the direct $O(p)$ algorithm. That leads to an overall complexity of $O(M^2/\log M)$, and

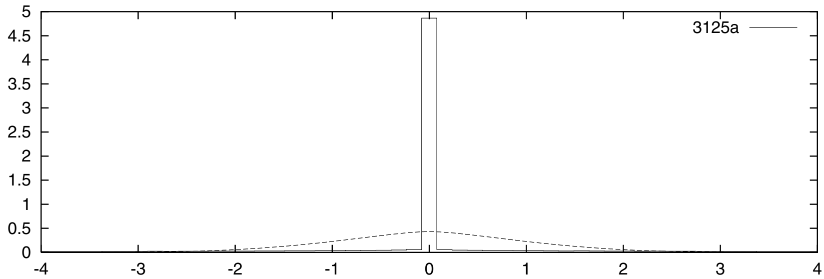

Figure 2.2. Distribution of Fourier coefficients for genus 2 curve 3125a

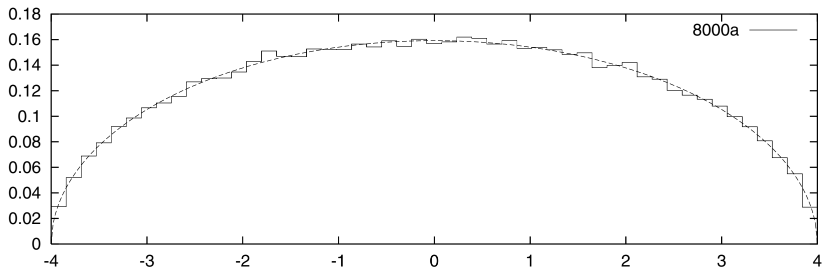

Figure 2.3. Distribution of Fourier coefficients for elliptic curve 8000a, compared with a stretched Sato–Tate measure

a dramatic difference in the computing time required; in our case, the elliptic curve computations took less than a day to complete, compared to more than a year for the genus 2 curves.

2.4 Results

Figures 2.4 and 2.5 show histograms of the computed Fourier coefficients of the curves listed in Table 2.1, over all primes for genus 2 curves over $\mathbb{Q}$, and primes $\equiv 1 \pmod 4$ for elliptic curves over $\mathbb{Q}(\sqrt{-1})$, against their expected distributions. Note that since proportionally more Fourier coefficients were computed for curves of high discriminant, the agreement is better for those curves. In all but two cases the agreement is good and clearly distinguishes the nature of these two types of Euler products. (If desired, this may be quantified by a χ^2 or Kolmogorov–Smirnov statistical test; we content ourselves with the pictures.) It has the added benefit of showing that there are likely no systematic errors in our computation of the Fourier coefficients at primes.

The exceptional cases are shown in Figures 2.2 and 2.3. The first came from the genus 2 curve

$$y^2 + y = x^5 \tag{2.16}$$

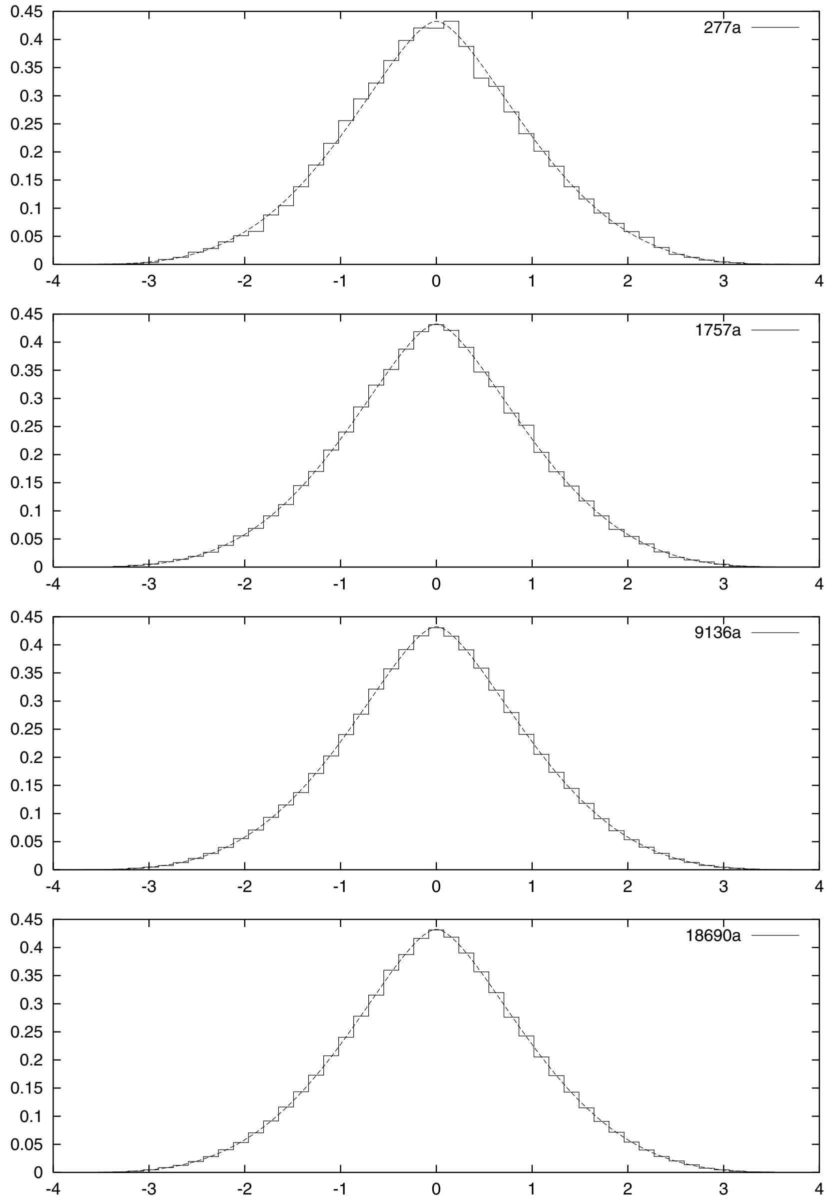

Figure 2.4. Distribution of Fourier coefficients for genus 2 curves over $\mathbb{Q}$

of discriminant 5^4. The curve displays a symmetry akin to complex multiplication for elliptic curves, and after completing the square, is of the type considered by Weil [49]. Its L-function is that of a Hecke character over $\mathbb{Q}(\zeta_{10})$, where ζ_{10} is a primitive tenth root of unity; thus, the analytic continuation and

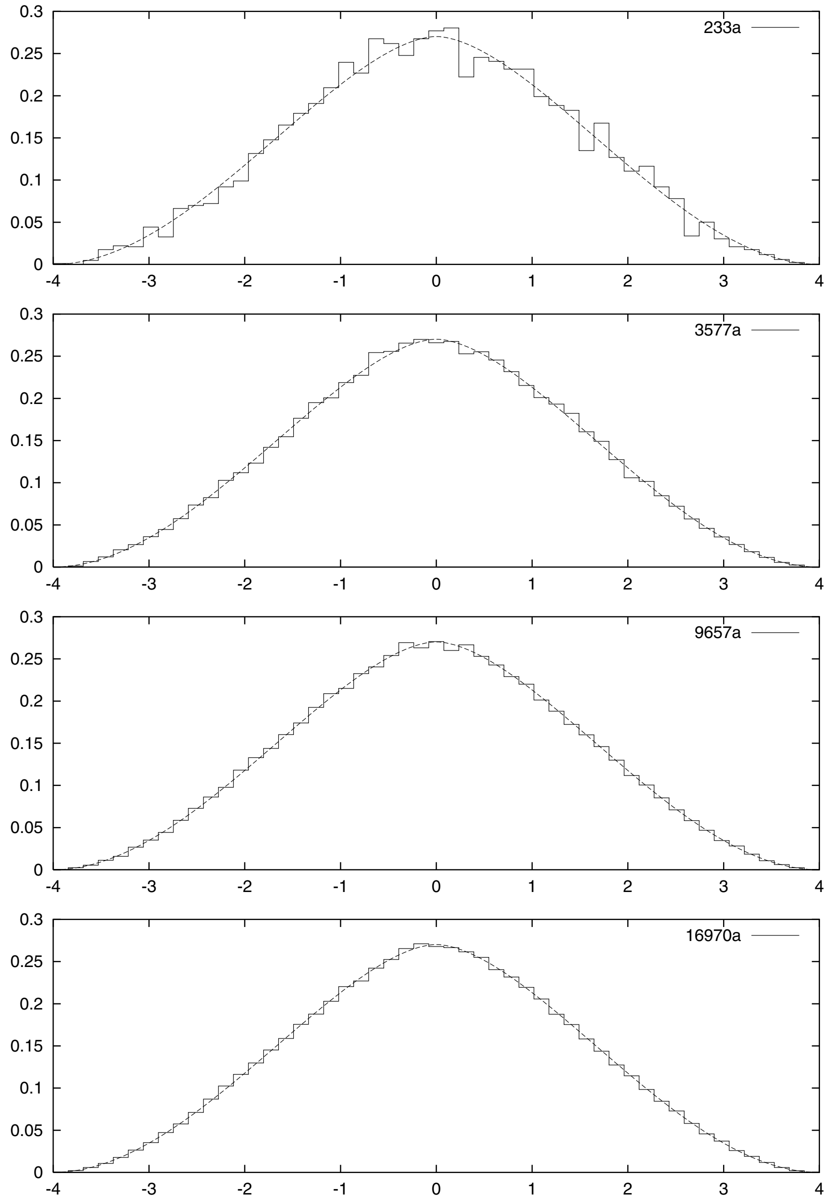

Figure 2.5. Distribution of Fourier coefficients for elliptic curves over $\mathbb{Q}(\sqrt{-1})$

functional equation that we test in Sections 3 and 5 are known for this example. The second exceptional case was that of the elliptic curve

$$y^2 + (1+i)xy = x^3 + (1-i)x^2 - (13+27i)x - (3+69i), \qquad (2.17)$$

of discriminant $40 + 80i$. It is a $\mathbb{Q}$-curve, meaning that, while not definable over $\mathbb{Q}$, it is isogenous to its complex conjugate. Thus, $a(\mathfrak{p}) = a(\overline{\mathfrak{p}})$ for every prime $\mathfrak{p}$, and the expected distribution in this case is the Sato–Tate measure stretched over the interval $[-4, 4]$, to which Figure 2.3 shows good agreement. Modularity is known in this case as well by extensions of Wiles' work due to Ellenberg and Skinner [19].

3. Smooth sums

This section will expand our discussion of smooth sums from the introduction, and give examples of the sort of numerical results that one may obtain with them.

3.1 The prototypical test

Theorem 3.1. *Let π be an automorphic representation of $\mathrm{GL}_r(\mathbb{A}_{\mathbb{Q}})$ whose L-function*

$$L(s, \pi) = \sum_{n=1}^{\infty} \lambda_\pi(n) n^{-s} \tag{3.1}$$

extends to an entire function. Define

$$S^*(X) = \frac{1}{\sqrt{X}} \sum_{n=1}^{\infty} \lambda_\pi(n) F(n/X), \tag{3.2}$$

for F a Schwartz function on $(0, \infty)$. Then for any positive integer k,

$$S^*(X) \ll_{F,r,k} \left(\frac{C_\pi}{X}\right)^k. \tag{3.3}$$

Proof. Recalling equation (1.15), we have

$$S^*(X) = \frac{1}{2\pi i} \int L(s, \pi) \widetilde{F}(s) X^{s-1/2}\, ds, \tag{3.4}$$

taken initially along a vertical line far to the right. Using the functional equation (1.3) for $L(s, \pi)$, this becomes

$$S^*(X) = \frac{\epsilon_\pi}{2\pi i} \int L(1 - s, \widetilde{\pi}) \frac{\gamma(1 - s, \widetilde{\pi}) \widetilde{F}(s)}{\gamma(s, \pi)} \left(\frac{X}{N_\pi}\right)^{s-1/2} ds. \tag{3.5}$$

Now we shift the line of integration to $\Re s = 1/2 - k$, writing $s = 1/2 - k + it$. Recall that $\gamma(s, \pi)$ is a product of the terms $\Gamma_{\mathbb{R}}(s + \mu_\pi^{(i)})$. Corresponding to the archimedean parameter $\mu_\pi^{(i)}$ of π, $\widetilde{\pi}$ has parameter $\mu_{\widetilde{\pi}}^{(i)} = \overline{\mu_\pi^{(i)}}$. Thus, the ratio of γ-factors in (3.5) is a product of terms of the form

$$\frac{\pi^{-(1/2+k-it+\bar{\mu})/2}\Gamma\left(\frac{1/2+k-it+\bar{\mu}}{2}\right)}{\pi^{-(1/2-k+it+\mu)/2}\Gamma\left(\frac{1/2-k+it+\mu}{2}\right)}$$

$$= \left(\pi^{i(t+\Im\mu)}\frac{\Gamma\left(\frac{1/2+k-it+\bar{\mu}}{2}\right)}{\Gamma\left(\frac{1/2+k+it+\mu}{2}\right)}\right)\cdot \pi^{-k}\frac{\Gamma\left(\frac{1/2+k+it+\mu}{2}\right)}{\Gamma\left(\frac{1/2-k+it+\mu}{2}\right)}. \tag{3.6}$$

The first factor above, in parentheses, is of absolute value 1, while the second is the polynomial

$$P(\mu, t) = \prod_{n=1}^{k} \frac{\mu + it + 1/2 + (k-2n)}{2\pi}. \tag{3.7}$$

Substituting this into (3.5) we obtain

$$|S^*(X)| \le \frac{1}{2\pi}\left(\frac{N_\pi}{X}\right)^k \int_{-\infty}^{\infty} \left| L(1/2+k-it, \widetilde{\pi})\widetilde{F}(1/2-k+it) \prod_{i=1}^{r} P(\mu_\pi^{(i)}, t)\right| dt. \tag{3.8}$$

Now, assuming the Ramanujan conjecture for π (and thereby $\widetilde{\pi}$), $L(1/2+k-it, \widetilde{\pi})$ is bounded by $\zeta^r(1/2+k)$. There is no loss in assuming Ramanujan, since it is known in the examples we consider, and anyway we first verify it numerically using the techniques of Section 2. However, note that it is not strictly necessary, since here an average result obtained from the Rankin–Selberg method would suffice. In any case, it is clear that the integral is dominated by

$$\prod_{i=1}^{r}\left(\frac{1+|\mu_\pi^{(i)}|}{2\pi}\right)^k, \tag{3.9}$$

independently of π. □

Remarks.

(1) The factors of 2π in the denominator of (3.9) may seem arbitrary since the implied constant in (3.3) depends on r and k. However, stated as is, the result is asymptotically correct in the sense that if all $|\mu_\pi^{(i)}|$ are assumed very large (depending in a precise way on r and k), then the constant may be taken to be $\zeta^r(1/2+k)$ times a number dependent only on F.

(2) It is not valid to take $k = 0$ in the theorem, since then we would require a bound for $L(s, \pi)$ in the critical strip, which would necessarily depend on π. However, the Lindelöf hypothesis for π predicts the bound $O_\varepsilon(N_\pi^\varepsilon)$, meaning we expect $S^*(X)$ to be bounded *almost* independently of π. This agrees with the philosophy stated in the introduction that $S^*(X)$ should be "random" of size 1 for X up to the analytic conductor, then rapidly decaying thereafter. The Ω-result of Section 4 also gives supporting evidence for this.

(3) We assumed that F is of rapid decay near 0. If not then $\widetilde{F}$ will not be analytic far to the left, and (3.3) will only be valid for $k \leq$ some k_0. The results below with explicit constants use smoothing functions that are only $O(x^2)$ near 0, with k taken appropriately small.

A quick glance at some sample graphs of $S^*(X)$ (see below) reveals that it is most natural to consider X on a logarithmic scale, i.e. we put $X = N_\pi^\sigma$. Then in terms of σ, Theorem 3.1 says that S^* is of arbitrarily fast exponential decay, which should start to take effect near $X = C_\pi$. Thus, from a graph of S^* not only is it easy to see the analytic continuation of $L(s, \pi)$, but also the analytic conductor. Moreover, it gives indirect evidence of the functional equation, through explicit forms of the bound (3.3), which was derived assuming it.

3.2 Modified tests

The problem with using Theorem 3.1 as a practical test is that it requires knowledge of all Dirichlet coefficients $\lambda_\pi(n)$. In accordance with our basic philosophy that it should be possible to verify modularity without knowledge of the Euler factors at finitely many places, we consider the following modification of Theorem 3.1.

Theorem 3.2. *Let π, $L(s, \pi)$, and F be as in Theorem 3.1 and define*

$$S(X) = \frac{1}{\sqrt{X}} \sum_{(n, N_\pi)=1} \lambda_\pi(n) F(n/X). \tag{3.10}$$

Then

$$S(X) \ll_{F,r} \left(\frac{C_\pi (\log(1 + C_\pi))^{2r}}{X} \right)^{1/2}. \tag{3.11}$$

Proof. In place of (3.4) we have

$$S(X) = \frac{1}{2\pi i} \int L(s, \pi) \prod_{p | N_\pi} L_p(s, \pi)^{-1} \widetilde{F}(s) X^{s-1/2} \, ds, \tag{3.12}$$

where $L_p(s,\pi)^{-1}$ is the local factor polynomial at the prime p. This yields

$$S(X)=\frac{\epsilon_\pi}{2\pi i}\int L(1-s,\widetilde{\pi})\prod_{p|N_\pi}L_p(s,\pi)^{-1}\frac{\gamma(1-s,\widetilde{\pi})\widetilde{F}(s)}{\gamma(s,\pi)}\left(\frac{X}{N_\pi}\right)^{s-1/2}ds. \tag{3.13}$$

We proceed as in the proof of Theorem 3.1, except that now shifting the contour far to the left introduces large powers of N_π from the local factors. Instead we shift to $\Re s=-1/\log(1+C_\pi)$. Then, again assuming the Ramanujan conjecture for π, we have

$$\left|\prod_{p|N_\pi}L_p(s,\pi)^{-1}\right|\le\prod_{p|N_\pi}(1+p^{1/\log(1+C_\pi)})^r\ll_r 1. \tag{3.14}$$

Next we apply (3.6) with $k=1/2+1/\log(1+C_\pi)$, and use Stirling's formula, since k is no longer an integer. We arrive at

$$\begin{aligned}|S(X)|&\ll_{F,r}\zeta^r\left(1+\frac{1}{\log(1+C_\pi)}\right)\left(\frac{C_\pi}{X}\right)^{1/2+1/\log(1+C_\pi)}\\&\ll(\log(1+C_\pi))^r\left(\frac{C_\pi}{X}\right)^{1/2},\end{aligned} \tag{3.15}$$

assuming $X\gg 1$. □

Theorem 3.2 again shows exponential decay in σ for X of size $C_\pi^{1+\varepsilon}$. Unfortunately, unlike the case of Theorem 3.1, it is not possible in general to improve upon the exponent $1/2$ in (3.11). Thus, Theorem 3.2 gives evidence for the analytic continuation only down to $\Re s=0$. A more serious problem is that without very fast decay this test cannot detect poles with large imaginary part. That is because of the fast decay of $\widetilde{F}(s)$ in vertical strips; a pole at $s=1/2+it$, say, with residue of size 1, would introduce a term proportional to $\widetilde{F}(1/2+it)$ when the contour of (3.12) is shifted to the left. So if the given function has meromorphic continuation without being analytic, that fact is not apparent with Theorem 3.2. This problem will be resolved with the more precise tests of Section 5.

However, there are some situations in which the tests of this section are more useful than those of Section 5. These will be discussed in the next subsection. First, we compute smooth sums from our main examples, which will help to understand why the exponent $1/2$ in (3.11) is best possible in the general setting. As mentioned in the introduction, if Δ denotes the discriminant, we expect the L-functions of Section 2 to have conductor $D=|\Delta|$ in the case of genus 2 curves, and $D=16|N_{\mathbb{Q}(\sqrt{-1})/\mathbb{Q}}\Delta|$ in the case of elliptic curves, unless

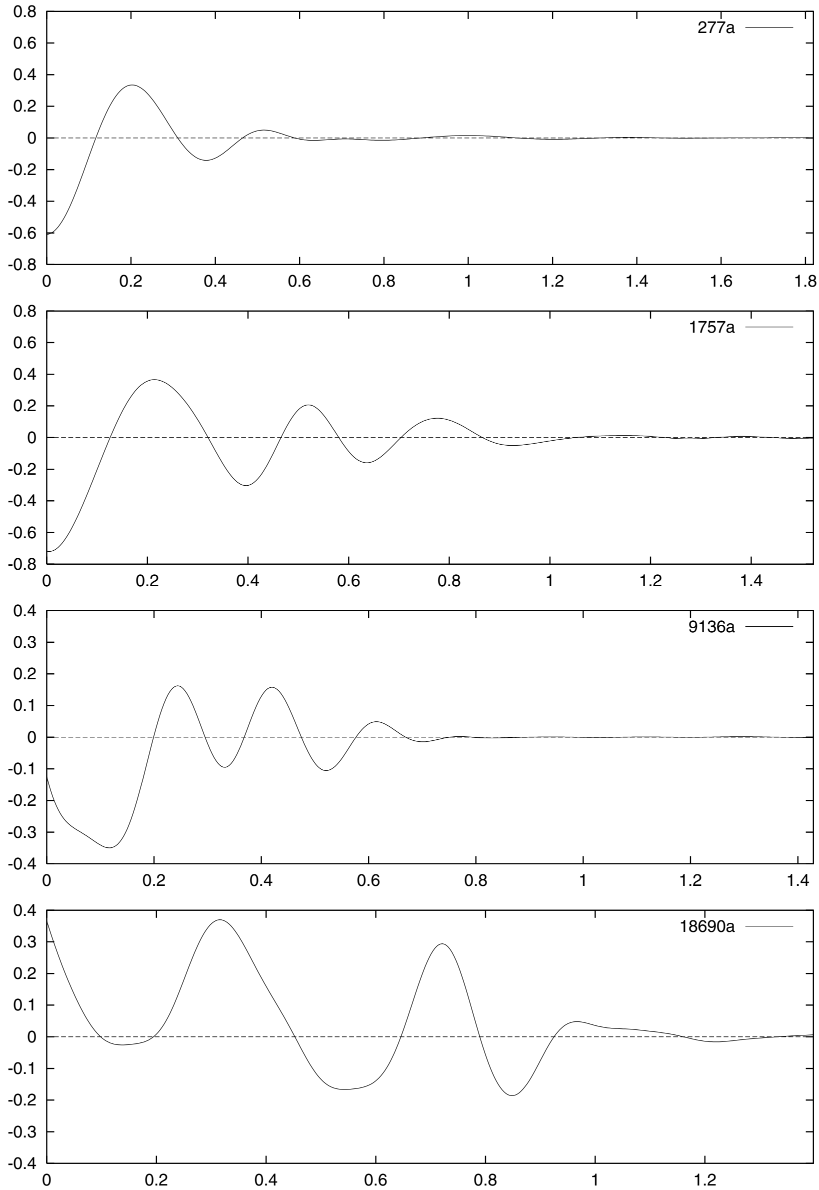

Figure 3.1. $S(D^\sigma)$ for genus 2 curves over $\mathbb{Q}$

the discriminant properly divides the conductor. Figures 3.1 and 3.2 below show the graphs of $S(D^\sigma)$ against σ, with smoothing function $F(x) = x^2e^{-x}$, applied to the example curves from Section 2. Note that each graph starts to decay at or before $\sigma = 1$, as expected. (The archimedean parameters are constant among each class of curves, so the conductor and analytic conductor differ by a constant factor, which is not significant on the logarithmic scale.)

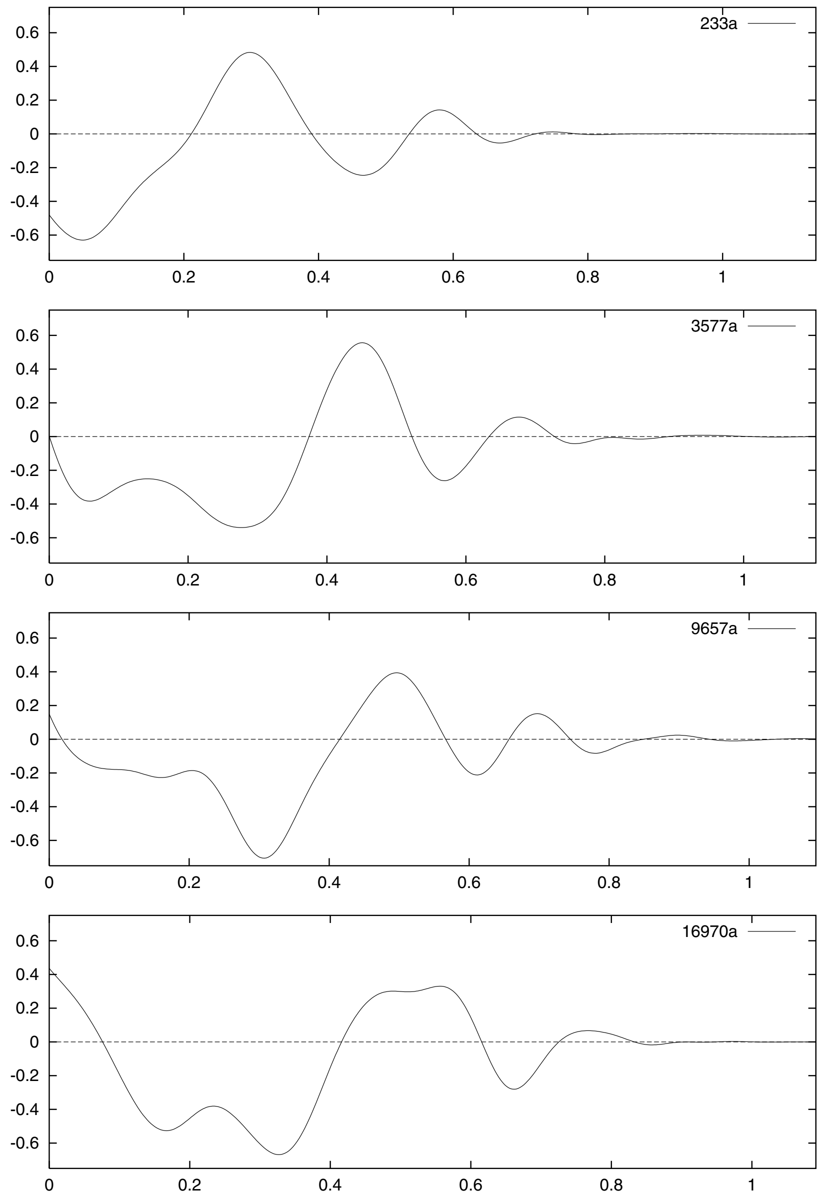

Figure 3.2. $S(D^\sigma)$ for elliptic curves over $\mathbb{Q}(\sqrt{-1})$

Now, consider equation (3.12) above. We may expand the product of Euler factors to a finite Dirichlet series $\sum_n c_n n^{-s}$. Then (3.12) takes the form

$$S(X) = \sum_n \frac{c_n}{\sqrt{n}} S^*(X/n), \tag{3.16}$$

where $S^*(X)$ is the smooth sum over all integers, i.e. with no missing Euler factors. By Theorem 3.1, $S^*(X)$ is of rapid decay once X is significantly beyond C_π. We see from (3.16) that $S(X)$ is the same as $S^*(X)$ (from the main term $n = 1$), followed by several "echoes" $S^*(X/n)$, of intensity decaying as $1/n^{1/2}$; that is the square root decay observed in (3.11). This phenomenon may be seen in some of the graphs below; for example, the small peak near $\sigma = 1$ in the graph for genus 2 curve 277a is an echo of the one near $\sigma = 0$.

The echoes persist to the largest value of n in (3.16). For example, in the case of prime conductor N, the first echo starts near $X = N$, just as the main term decays. The last echo occurs near $X = N^{r-1}$, after which we expect rapid decay for X beyond N^r. (In practice this is usually too large to observe.) Of course, for composite conductors there are many more terms, and one may not be able to see clearly where one term in (3.16) ends and the next begins; this also helps to explain why the graph for the example genus 2 curve of discriminant $18690 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 89$ does not decay as quickly as the others. A natural question is whether it is possible to guess values for c_n, in an attempt to cancel the echoes. The answer is yes, and we shall take this up rigorously in Section 5 using regression and approximate functional equations.

3.3 Improvements

Before leaving this test, we mention some cases in which (3.11) can be improved upon. First, the inequality (3.14) was derived assuming the worst case of the Ramanujan bound for each missing Euler factor. Sometimes one has available better known bounds for the coefficients. For example, for a self-dual modular form for $\Gamma_0(N)$, the coefficients $\lambda(p)$ for p dividing N are, by Atkin–Lehner theory [1], either 0 or $\pm 1/\sqrt{p}$. (This is notably false for non-self-dual forms; consider e.g. forms that arise from complex Galois representations, which always have reciprocal roots of absolute value 1.) Using this, in [3] we derive the precise bound

Theorem 3.3. *Let f be a Maass or holomorphic modular form and Hecke eigenform for $\Gamma_0(N)$, and put $Q = \prod_{p \| N}(p+1)$. Let $S(X)$ be the smooth sum of Fourier coefficients*

$$S(X) = \sum_{(n,N)=1} \lambda_f(n) F(n/X), \tag{3.17}$$

with smoothing function $F(x) = x^2 e^{-x}$. Then

$$|S(X)| < Q \cdot \left(\frac{N(\lambda + 3)}{42.88X} \right)^2 \tag{3.18}$$

when f is a Maass form of eigenvalue λ, and

$$|S(X)| < Q \cdot \left(\frac{N(k^2+2k+9)}{171.5X}\right)^2 \tag{3.19}$$

when f is a holomorphic form of weight k.

Here the analytic conductor is expressed as a function of the weight or eigenvalue. This "Atkin–Lehner effect" persists in higher dimensional cases as well. For example, for genus 2 curves over $\mathbb{Q}$, the Euler factors for primes p dividing the conductor typically have one or more reciprocal roots of size $1/\sqrt{p}$ (see Section 5).

Second, if N_π has many repeated factors then we may shift the contour in (3.13) farther to the left without a large penalty; e.g. in the proof of Theorem 3.3 the contour is shifted to $\Re s = -3/2$. This can be useful in showing that a given function is *not* modular. For example, consider the classical Kloosterman sums,

$$S(\alpha, p) = \frac{1}{\sqrt{p}} \sum_{x\bar{x}\equiv 1\,(p)} e\left(\frac{x+\alpha\bar{x}}{p}\right), \tag{3.20}$$

for prime p not dividing α. In [28], Katz proved that the collection of all $S(\alpha, p)$ is distributed according to the Sato–Tate measure as $p \to \infty$, and conjectured the same for fixed α. This claim is supported by numerical evidence [3], as in Section 2. This led him to question whether the numbers $\lambda(p) = \epsilon S(1, p)$, for fixed $\epsilon = \pm 1$, could be the prime coefficients of a Maass eigenform for some $\Gamma_0(N)$, presumably with $N = 1$ or possibly a power of 2. (They cannot be the coefficients of a holomorphic modular form, since the latter, when properly normalized, lie in a fixed number field.) Using Theorem 3.3, in [3] we establish

Theorem 3.4. *If a Katz form of level $N = 2^\nu$ and eigenvalue λ exists, then*

$$N(\lambda + 3) > 18.3 \times 10^6. \tag{3.21}$$

Therefore, it is unlikely that any such modular form exists. (Interestingly, W. Li [12] has shown that a function field analogue of Katz's question, with $\epsilon = 1$, does hold.) Note that it is essential for this result that our test not require knowledge of the eigenvalue λ of the purported form. Thus, the techniques of Section 5 would not apply here.

3.4 A test of GRH

In this subsection we let π be as in Theorem 3.1 and study the sums

$$S^{\mathrm{pr}}(X) = \frac{1}{\sqrt{X}} \sum_{p \nmid N_\pi} (\log p)\lambda_\pi(p)F(p/X), \tag{3.22}$$

over prime numbers p. Note that we leave out the terms for p dividing the conductor, although unlike the case of sums over integers, here it makes little difference.

To simplify the discussion we will assume that π is self-dual and satisfies the Ramanujan conjecture, although these restrictions may likely be removed in what follows. Also, it suffices to consider π cuspidal, as the L-functions of interest to us are products of cuspidal ones; note that $S^{\mathrm{pr}}(X)$ for a product L-function is the sum of $S^{\mathrm{pr}}(X)$ for the individual factors.

Proceeding, let $L_S(s,\pi) = \prod_{p \nmid N_\pi} L_p(s,\pi)$ be the partial L-function with S the set of primes dividing N_π. Each local factor for $p \notin S$ may be written in terms of Satake parameters

$$L_p(s,\pi) = \prod_{i=1}^{r} (1 - \alpha_\pi^{(i)}(p) p^{-s})^{-1}, \tag{3.23}$$

so that

$$-\frac{L_p'}{L_p}(s,\pi) = \sum_{i=1}^{r} \frac{\alpha_\pi^{(i)}(p)(\log p) p^{-s}}{1 - \alpha_\pi^{(i)}(p) p^{-s}} = (\log p) \sum_{k=1}^{\infty} \sum_{i=1}^{r} \alpha_\pi^{(i)}(p)^k p^{-ks}. \tag{3.24}$$

Now the first term is $\sum_{i=1}^{r} \alpha_\pi^{(i)}(p) p^{-s} = \lambda_\pi(p) p^{-s}$. Also,

$$\mathrm{Tr}\,\mathrm{Sym}^2\,\mathrm{diag}(\alpha_\pi^{(1)}(p), \ldots, \alpha_\pi^{(r)}(p)) = \sum_{i=1}^{r} \alpha_\pi^{(i)}(p)^2 + \sum_{1 \le i < j \le r} \alpha_\pi^{(i)}(p) \alpha_\pi^{(j)}(p) \tag{3.25}$$

and

$$\mathrm{Tr} \wedge^2 \mathrm{diag}(\alpha_\pi^{(1)}(p), \ldots, \alpha_\pi^{(r)}(p)) = \sum_{1 \le i < j \le r} \alpha_\pi^{(i)}(p) \alpha_\pi^{(j)}(p) \tag{3.26}$$

so that

$$\begin{aligned} \sum_{i=1}^{r} \alpha_\pi^{(i)}(p)^2 = {} & \mathrm{Tr}\,\mathrm{Sym}^2\,\mathrm{diag}(\alpha_\pi^{(1)}(p), \ldots, \alpha_\pi^{(r)}(p)) \\ & - \mathrm{Tr} \wedge^2 \mathrm{diag}(\alpha_\pi^{(1)}(p), \ldots, \alpha_\pi^{(r)}(p)), \end{aligned} \tag{3.27}$$

which we write in the more compact form $\lambda_{\mathrm{Sym}^2 \pi}(p) - \lambda_{\wedge^2 \pi}(p)$. Thus, (3.24) is

$$\begin{aligned} -\frac{L_p'}{L_p}(s,\pi) = {} & (\log p)(\lambda_\pi(p) p^{-s} + (\lambda_{\mathrm{Sym}^2 \pi}(p) - \lambda_{\wedge^2 \pi}(p)) p^{-2s} \\ & + (\text{terms of order } p^{-3s})). \end{aligned} \tag{3.28}$$

Summing over p, we get

$$\sum_{p\notin S}(\log p)\lambda_\pi(p)p^{-s}$$
$$= -\frac{L_S'}{L_S}(s,\pi) + \frac{L_S'}{L_S}(2s,\operatorname{Sym}^2\pi) - \frac{L_S'}{L_S}(2s,\wedge^2\pi) + (\text{error}), \qquad (3.29)$$

with error term analytic in $\Re s > 1/3$.

Next, by the work of Bump and Friedberg [10] and Bump and Ginzburg [11], the partial L-functions $L_S(2s,\wedge^2\pi)$ and $L_S(2s,\operatorname{Sym}^2\pi)$ are analytic, except that one of them has a simple pole at $s = 1/2$. Further, Shahidi [39] showed that their product $L_S(2s,\pi\times\pi)$ does not vanish on the line $\Re s = 1/2$. Therefore, the corresponding terms of (3.29) are analytic in $\Re s \ge 1/2$, again except for a simple pole at $s = 1/2$. Integrating over s and applying the Wiener-Ikehara theorem, we thus have the following explicit formula.

Theorem 3.5. *There is an integer m (depending on π) such that*

$$S^{\mathrm{pr}}(X) = \frac{m}{2}\widetilde{F}(1/2) - \sum_\rho \widetilde{F}(\rho)X^{\rho-1/2} + o(1), \qquad (3.30)$$

where the sum runs over all non-trivial zeros ρ of $L(s,\pi)$, with multiplicity.

Corollary 3.6. *Suppose that $L(s,\pi)$ satisfies the Riemann hypothesis, i.e. all of its non-trivial zeros lie on the line $\Re s = 1/2$. Then*

$$S^{\mathrm{pr}}(X) = O(1). \qquad (3.31)$$

Remarks.

(1) While there are known effective zero-free regions for $L(s,\pi\times\pi)$, they yield very weak estimates for the error in Theorem 3.5 that are likely far from the truth. In fact, one expects that $L(s,\pi\times\pi)$ itself satisfies a Riemann hypothesis, so that the error term decays as a power of X and is small when X is larger than the conductor.

(2) The integer m is an artifact that comes from summing over primes rather than all prime powers. Instead of estimating the higher order terms, one could include them in the definition of $S^{\mathrm{pr}}(X)$.

(3) When F is a Schwartz function, the sum over zeros ρ is very small unless $L(s,\pi)$ happens to have a zero of low height. Therefore, to use Corollary 3.6 as a practical test of the Riemann hypothesis, in this situation (and only this situation, we hasten to add) it makes sense to use a function F which is not smooth so that the decay of $\widetilde{F}$ is not too great. For example, we may take

$$S^{\mathrm{pr}}(X) = \frac{1}{\sqrt{X}}\sum_{p\nmid N_\pi,\ p\le X}(\log p)\lambda_\pi(p)(1-p/X), \qquad (3.32)$$

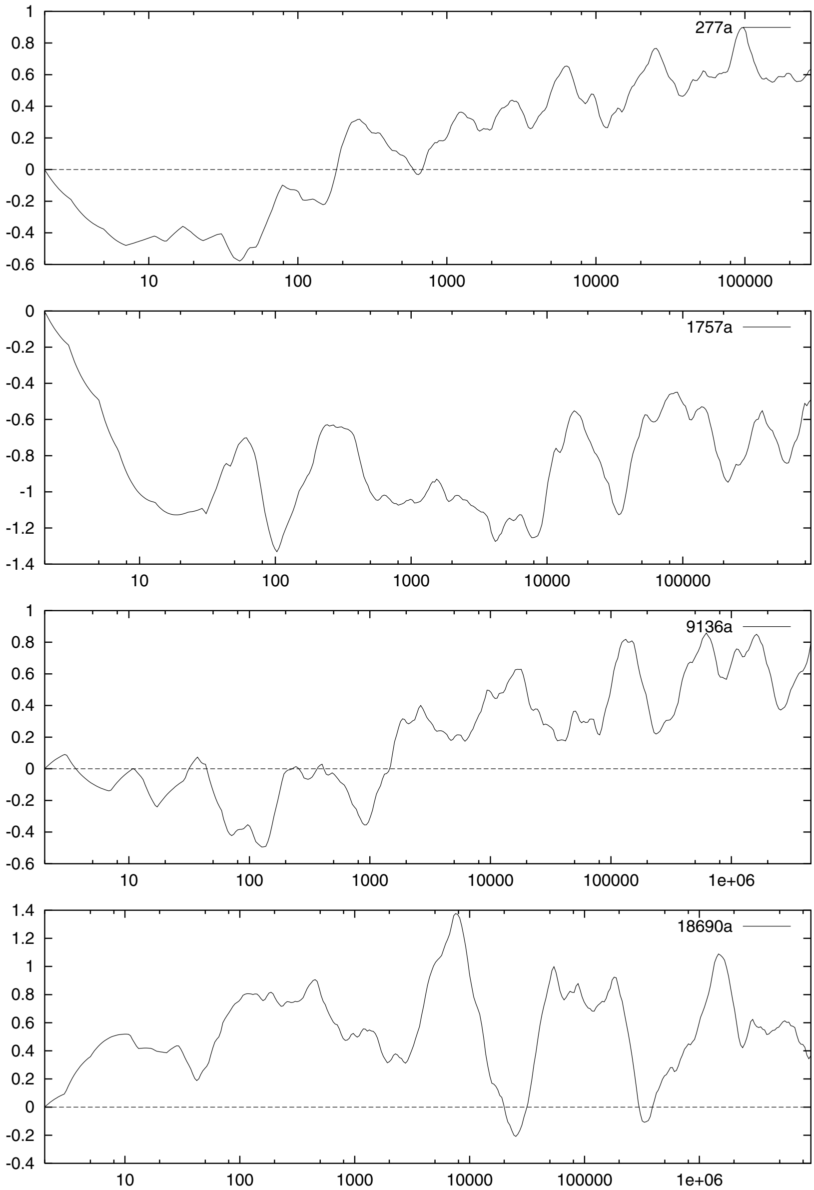

Figure 3.3. $S^{\mathrm{pr}}(X)$ for genus 2 curves over $\mathbb{Q}$

which corresponds to the choice $\widetilde{F}(s) = 1/s(s+1)$. Corollary 3.6 remains valid in this case.

We computed $S^{\mathrm{pr}}(X)$ using the definition (3.32) for the example curves from Section 2; the results are shown in Figures 3.3 and 3.4 below. Note that

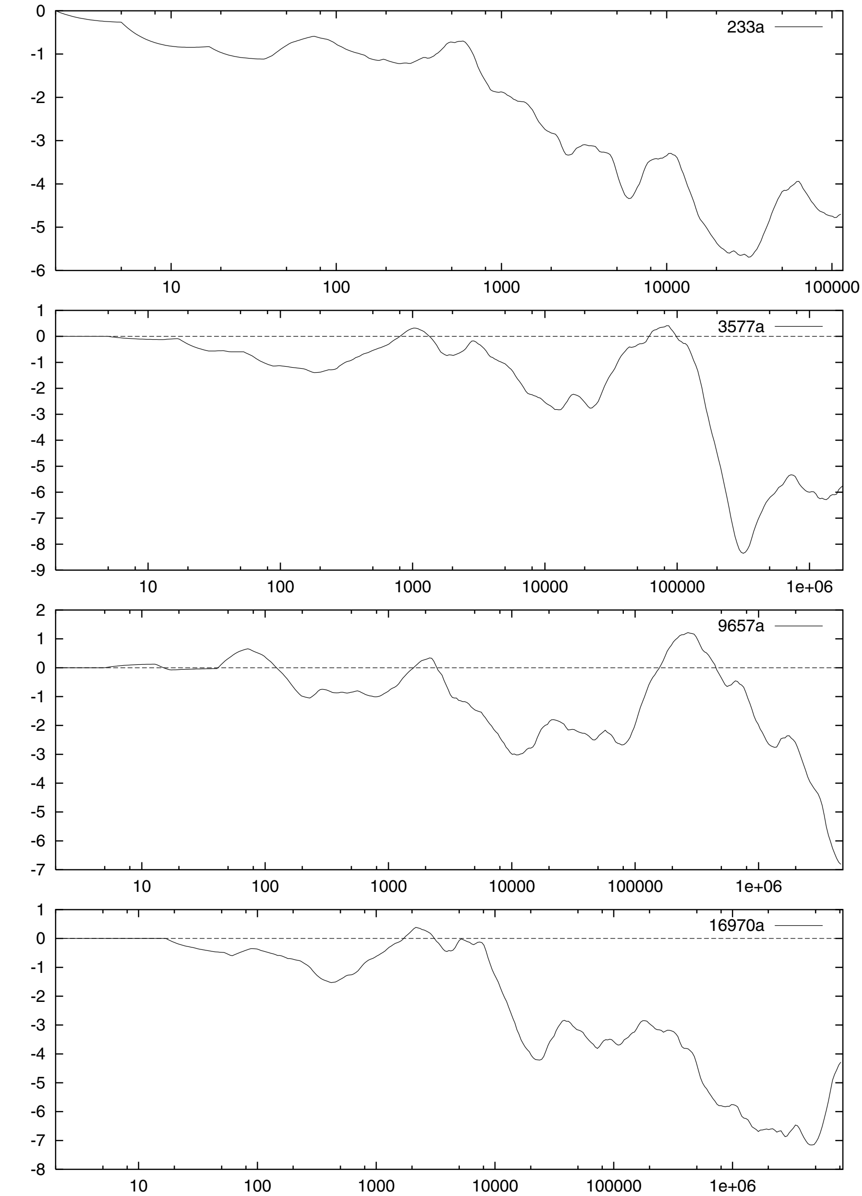

Figure 3.4. $S^{\mathrm{pr}}(X)$ for elliptic curves over $\mathbb{Q}(\sqrt{-1})$

each function oscillates around a non-zero value for large X, although those of Figure 3.4 do not reach their true size until X is large, due to the extra factor of 16 in the conductor. If the L-function does not vanish at $1/2$ then that reflects the term $\frac{m}{2}\widetilde{F}(1/2) = \frac{2}{3}m$; e.g. the example 18690a appears to oscillate around $2/3$, so we may guess that $m = 1$. (We will verify in Section 5 that $L(1/2) > 0$

in this case.) None of the graphs appears to grow significantly, and thus it is likely that the L-functions have no low-lying counterexamples to the Riemann hypothesis. On the other hand, their oscillations do not decay either, which indicates the presence of zeros on the line $\Re s = 1/2$.

4. The analytic conductor

Let $\mathcal{F}(M)$ be the family of polynomials

$$\mathcal{F}(M) = \{f_0(x)(x-a) \mid a = 1 \dots M\}, \tag{4.1}$$

where $f_0(x)$ is a fixed square-free polynomial of degree at least 2, with integer coefficients. For $f \in \mathcal{F}(M)$, let

$$L(s, f) = \sum_{n=1}^{\infty} \lambda_f(n) n^{-s} \tag{4.2}$$

be the partial Hasse–Weil L-function associated to the curve $y^2 = f(x)$, without the Euler factors at primes of bad reduction, i.e. we take $\lambda_f(n) = 0$ when $(n, 2\Delta(f)) \neq 1$. Further, let $S_f(X)$ be the smooth sum

$$S_f(X) = \frac{1}{\sqrt{X}} \sum_{n=1}^{\infty} \lambda_f(n) F(n/X), \tag{4.3}$$

where F is a given Schwartz function on $(0, \infty)$.

Theorem 3.2 of the last section shows that if $L(s, f)$ is modular then $S_f(|\Delta(f)|^\sigma)$ decays when σ is larger than 1. (The precise point of decay depends on the archimedean parameters, but these are constant over the family.) Conversely, we expect, and the numerical examples of Section 3 demonstrate, that the decay typically does not start when σ is much smaller than 1. While such a statement is likely out of reach, we obtain in this section a lower bound $\delta > 0$ such that many of the S_f from our chosen family do not start decaying until $\sigma \geq \delta$. Precisely, we have

Theorem 4.1. *There are positive constants c and δ such that*

$$\#\{f \in \mathcal{F}(M) \mid |S_f(X)| > c \text{ for some } X \geq |\Delta(f)|^\delta\} \gg \frac{M}{\log M}. \tag{4.4}$$

The constants c and δ, and the implied constant in the theorem all depend on the choice of f_0. Such dependence is assumed implicitly throughout.

The proof goes as follows. We study the second and fourth moments of $S_f(X)$, summed over the family. As it turns out, there is a small gain to be had in the fourth moment calculation if we assume that the smoothing function $F(x)$ is *balanced*, in the sense that

$$\int_0^\infty \frac{F(x)}{\sqrt{x}}\,dx = 0. \tag{4.5}$$

There is no loss of generality since we may replace our given F by $F(x) - \sqrt{2}F(2x)$, which satisfies (4.5); this amounts to studying $S_f(X) - S_f(X/2)$ in place of $S_f(X)$. Note that whenever this quantity is large at X, the original S_f must be large at either X or $X/2$.

For balanced F, we obtain a lower bound of $c_1 M$ for the second moment and an upper bound $c_2 M \log X$ for the fourth. Of course, these estimates will not hold unconditionally; we will require $1 \ll X \ll M^\alpha$ for some positive constant α. Since $\Delta(f)$ is a polynomial in the parameter a, we may thus let X range up to a power $|\Delta(f)|^\delta$.

Assuming the estimates for now, for a fixed X, let T be the set

$$\{f \in \mathcal{F}(M) \mid |S_f(X)| > \sqrt{c_1/2}\}. \tag{4.6}$$

Then, removing the contribution to the second moment from the terms outside of T and applying the Schwarz inequality, we have

$$\frac{c_1}{2}M \le \sum_{f\in T} S_f(X)^2 \le \sqrt{\sum_{f\in T} S_f(X)^4}\sqrt{\sum_{f\in T} 1} \le \sqrt{c_2 M(\log X)\#T}. \tag{4.7}$$

Thus,

$$\#T \ge \frac{c_1^2}{4c_2}\frac{M}{\log X} \gg \frac{M}{\log M}. \tag{4.8}$$

To complete the proof, we turn to the evaluation of the moments.

4.1 A general moment formula

The rth moment of $S_f(X)$ is

$$\sum_{f\in\mathcal{F}(M)} S_f(X)^r = \sum_{f\in\mathcal{F}(M)} \left(\frac{1}{\sqrt{X}}\sum_{n=1}^\infty \lambda_f(n)F(n/X)\right)^r. \tag{4.9}$$

Writing the inner sum as an integral of the Mellin transform $\widetilde{F}$ of F, this is

$$\sum_{f\in\mathcal{F}(M)} \left(\int_{\Re s\gg 1} X^{s-1/2} L(s,f)\widetilde{F}(s)\frac{ds}{2\pi i}\right)^r$$
$$= \int_{\Re s_1\gg 1}\cdots\int_{\Re s_r\gg 1} X^{s_1+\cdots+s_r-r/2}$$
$$\times \sum_{f\in\mathcal{F}(M)} L(s_1,f)\cdots L(s_r,f)\widetilde{F}(s_1)\cdots\widetilde{F}(s_r)\frac{ds_1}{2\pi i}\cdots\frac{ds_r}{2\pi i}. \tag{4.10}$$

Now,

$$\sum_{f\in\mathcal{F}(M)} L(s_1, f)\cdots L(s_r, f) = \sum_{n_1,\dots,n_r}\sum_{a=1}^{M}\lambda_{f_a}(n_1)\cdots\lambda_{f_a}(n_r)n_1^{-s_1}\cdots n_r^{-s_r}, \tag{4.11}$$

where $f_a(x) = f_0(x)(x-a)$. Note that $\lambda_{f_a}(n)$ depends only on a modulo n. Thus, we may break the inner sum of (4.11) into periods modulo $n_1\cdots n_r$:

$$\sum_{n_1,\dots,n_r}\left(\frac{M}{n_1\cdots n_r}\sum_{a\ (n_1\cdots n_r)}\lambda_{f_a}(n_1)\cdots\lambda_{f_a}(n_r)+R_M(n_1,\dots,n_r)\right)n_1^{-s_1}\cdots n_r^{-s_r}. \tag{4.12}$$

Here the error term $R_M(n_1,\dots,n_r)$ is bounded as $O_\varepsilon((n_1\cdots n_r)^{1+\varepsilon})$, independently of M. Thus, the contribution of this term to (4.10) is at most $O_{r,\varepsilon}(X^{3r/2+\varepsilon})$. (One expects that it is in fact $O_{r,\varepsilon}(X^{r+\varepsilon})$, but we will not need this.)

The main term of (4.12) is

$$\begin{aligned} M\sum_{n=1}^{\infty}\sum_{n_1\cdots n_r=n}\sum_{a\ (n)}&\lambda_{f_a}(n_1)\cdots\lambda_{f_a}(n_r)n_1^{-(s_1+1)}\cdots n_r^{-(s_r+1)} \\ &= M\sum_{n=1}^{\infty}A(n;s_1,\dots,s_r), \end{aligned} \tag{4.13}$$

where $A(n;s_1,\dots,s_r)$ is defined by this equation. By the Chinese Remainder Theorem, A is multiplicative in n. Therefore, (4.13) may be written as the Euler product

$$M\prod_p\left(1+\sum_{k=1}^{\infty}A(p^k;s_1,\dots,s_r)\right). \tag{4.14}$$

Let $L_p(s, f)$ denote the local factor of the L-function $L(s, f)$ at the prime p. Then the local factors of (4.14) are related to the $L_p(s, f)$ by

$$1+\sum_{k=1}^{\infty}A(p^k;s_1,\dots,s_r) = \frac{1}{p}\sum_{a\ (p)}L_p(s_1, f_a)\cdots L_p(s_r, f_a). \tag{4.15}$$

Now, we would like to meromorphically continue (4.14) and evaluate (4.10) by shifting contours to the left. As it turns out, the main term will come from poles near $s_i = 1/2$. In view of (4.15), the only obstructions to the product converging for $\Re s_i > 1/3$ are the terms

$$A(p;s_1,\dots,s_r) = \left(\frac{1}{p}\sum_{a\ (p)}\lambda_{f_a}(p)\right)(p^{-s_1}+\cdots+p^{-s_r}) \tag{4.16}$$

and

$$A(p^2; s_1, \dots, s_r) = \left(\frac{1}{p}\sum_{a\ (p)} \lambda_{f_a}(p^2)\right)(p^{-2s_1} + \cdots + p^{-2s_r})$$
$$+ \left(\frac{1}{p}\sum_{a\ (p)} \lambda_{f_a}(p)^2\right) \sum_{1\le i<j\le r} p^{-(s_i+s_j)}. \tag{4.17}$$

For these terms we have

Lemma 4.2.

i. $\sum_{a\ (p)} \lambda_{f_a}(p)^2 = p + O(1)$,
ii. $\sum_{a\ (p)} \lambda_{f_a}(p) = O(1)$,
iii. $\sum_{a\ (p)} \lambda_{f_a}(p^2) = O(1)$.

Proof. Note first that if p divides $2\Delta(f_0)$ then all sums are trivially 0. There are only finitely many such primes, and we may take the implied constant in i large enough to cover these cases. Hence, assume that p does not divide $2\Delta(f_0)$.

i. If f_a is not divisible by a square modulo p, we have

$$\lambda_{f_a}(p)^2 = \frac{1}{p}\sum_{x,y\ (p)} \left(\frac{f_0(x)f_0(y)}{p}\right)\left(\frac{(x-a)(y-a)}{p}\right). \tag{4.18}$$

If f_a is divisible by a square modulo p, then $\lambda_{f_a}(p) = 0$. This happens precisely when $a \in Z_p(f_0)$, the set of roots of f_0 modulo p. Thus, the sum in i is

$$\frac{1}{p}\sum_{x,y\ (p)} \left(\frac{f_0(x)f_0(y)}{p}\right)\sum_{a\ (p)}\left(\frac{(x-a)(y-a)}{p}\right)$$
$$- \sum_{a\in Z_p(f_0)} \left(\frac{1}{\sqrt{p}}\sum_{x\ (p)}\left(\frac{f_0(x)(x-a)}{p}\right)\right)^2. \tag{4.19}$$

Now, as we see by counting points on the conic $Y^2 = (x-X)(y-X)$,

$$\sum_{a\ (p)} \left(\frac{(x-a)(y-a)}{p}\right) = -1 + p\delta_{x=y}, \tag{4.20}$$

where $\delta_{x=y} = 1$ if $x = y$, and 0 otherwise. Thus, (4.19) is

$$p - \left[\lambda_{f_0}(p)^2 + \#Z_p(f_0) + \sum_{a\in Z_p(f_0)} \left(\frac{1}{\sqrt{p}}\sum_{x\ (p)}\left(\frac{f_0(x)(x-a)}{p}\right)\right)^2\right]. \tag{4.21}$$

The quantity in brackets is bounded. Interestingly, we also see that the error term in i is always ≤ 0.

ii. We have similarly

$$\sum_{a\ (p)} \lambda_{f_a}(p) = \frac{-1}{\sqrt{p}} \sum_{x\ (p)} \left(\frac{f_0(x)}{p}\right) \sum_{a\ (p)} \left(\frac{x-a}{p}\right) + \sum_{a\in Z_p(f_0)} \frac{1}{\sqrt{p}} \sum_{x\ (p)} \left(\frac{f_0(x)(x-a)}{p}\right). \tag{4.22}$$

The first term above is 0, and the second is evidently bounded.

iii. When $a \notin Z_p(f_0)$,

$$\lambda_{f_a}(p^2) = \frac{1}{2}\left(\lambda_{f_a}(p)^2 - \frac{1}{p} \sum_{x\in\mathbb{F}_{p^2}} \chi(f_a(x))\right), \tag{4.23}$$

where $\chi : \mathbb{F}_{p^2}^* \to \{\pm 1\}$ is the quadratic character $x \mapsto x^{(p^2-1)/2}$, extended to $\mathbb{F}_{p^2}$ by setting $\chi(0) = 0$. Summing over a and applying part *i*, this yields

$$\sum_{a\ (p)} \lambda_{f_a}(p^2) = O(1) + \frac{1}{2}\left(p - \frac{1}{p} \sum_{x\in\mathbb{F}_{p^2}} \chi(f_0(x)) \sum_{a\in\mathbb{F}_p\setminus Z_p(f_0)} \chi(x-a)\right). \tag{4.24}$$

Now, we extend the sum over a to all of $\mathbb{F}_p$, introducing an error of $O(1)$. The complete sum $\sum_{a\in\mathbb{F}_p} \chi(x-a)$ is clearly $p-1$ when $x \in \mathbb{F}_p$. Moreover, for any $c \in \mathbb{F}_p^*$ and $d \in \mathbb{F}_p$,

$$\sum_{a\in\mathbb{F}_p} \chi((cx+d)-a) = \sum_{a\in\mathbb{F}_p} \chi(x-a), \tag{4.25}$$

and thus the sum is constant for x outside of $\mathbb{F}_p$. Summing over x, we see that it equals -1 there. Therefore, (4.24) is

$$O(1) + \frac{1}{2}\left(p - \frac{1}{p} \sum_{x\in\mathbb{F}_{p^2}} \chi(f_0(x))(-1 + p\delta_{x\in\mathbb{F}_p})\right) \tag{4.26}$$

$$= O(1) + \frac{1}{2}\left(\frac{1}{p} \sum_{x\in\mathbb{F}_{p^2}} \chi(f_0(x)) + p - \sum_{x\in\mathbb{F}_p} \chi(f_0(x))\right) = O(1).$$

□

Now, the estimates given in Lemma 4.2 say roughly that (4.16) is small and that (4.17) is close to

$$\sum_{1\le i<j\le r} p^{-(s_i+s_j)}. \tag{4.27}$$

More precisely, (4.14) may be written as

$$M \prod_{1 \le i < j \le r} \zeta(s_i + s_j) \cdot \prod_p \left(\prod_{1 \le i < j \le r} (1 - p^{-(s_i+s_j)}) \right) \left(1 + \sum_{k=1}^{\infty} A(p^k; s_1, \dots, s_r) \right), \tag{4.28}$$

with the product over p convergent for $\Re s_i > 1/3$. Thus, this expression meromorphically continues to $\Re s_i > 1/3$, with simple poles at $s_i + s_j = 1$ for $i \ne j$.

Remarks.

(1) We have been assuming that $\lambda_f(p) = 0$ when p divides $2\Delta(f)$. The proof of Lemma 4.2 shows that this is not necessary, i.e. the conclusion of Theorem 4.1 is not changed if we have outside knowledge of some or all of the "correct" values of $\lambda_f(p)$ at primes of bad reduction.

(2) We do not need the full strength of Lemma 4.2 for the subsequent argument to work. Theorem 4.1 will hold more generally for any family for which the Euler product in (4.28) analytically continues a small amount to the left of $\Re s_i = 1/2$.

Next, we evaluate the main term of (4.10) for $r = 2$ and $r = 4$.

4.2 The second moment

The results of the previous section show that the second moment of $S_f(X)$ is

$$O_\varepsilon(X^{3+\varepsilon}) + M \int_{\Re s_1 \gg 0} \int_{\Re s_2 \gg 0} X^{s_1+s_2-1} \zeta(s_1+s_2) P(s_1, s_2) \widetilde{F}(s_1) \widetilde{F}(s_2) \frac{ds_1}{2\pi i} \frac{ds_2}{2\pi i}, \tag{4.29}$$

where $P(s_1, s_2)$ is the Euler product

$$\prod_p (1 - p^{-(s_1+s_2)}) \left(1 + \sum_{k=1}^{\infty} A(p^k; s_1, s_2) \right), \tag{4.30}$$

convergent for $\Re s_i > 1/3$.

Thanks to the rapid decay of $\widetilde{F}$ along vertical lines, we may shift the contours of (4.29) as we please, keeping track of the residues from any poles. Assume that the contours are arranged so that $\Re s_1$ and $\Re s_2$ are a small distance to the right of $1/2$. This is far enough to the right to not pick up any poles. Then, shifting the contour of s_1 to the left, we get a term from the pole at $s_1 = 1 - s_2$,

$$M \int P(1 - s_2, s_2) \widetilde{F}(1 - s_2) \widetilde{F}(s_2) \frac{ds_2}{2\pi i}, \tag{4.31}$$

and an error term of the same form as (4.29). We may estimate the error as $O_\varepsilon(MX^{-1/3+\varepsilon})$ by shifting the contours of s_1 and s_2 appropriately close to 1/3. As for the main term (4.31), it seems most natural to shift the contour to $\Re s_2 = 1/2$, since the integrand there is real. Further, then

$$\widetilde{F}(1-s_2)\widetilde{F}(s_2) = |\widetilde{F}(s_2)|^2 \tag{4.32}$$

and by equation (4.15) the local factors of $P(1-s_2, s_2)$ are

$$(1-p^{-1})\frac{1}{p}\sum_{a\ (p)} |L_p(s_2, f_a)|^2 > 0. \tag{4.33}$$

Thus, the integral in (4.31) is a strictly positive constant c.

Altogether, we have the asymptotic formula

$$\sum_{f\in\mathcal{F}(M)} S_f(X)^2 = cM + O_\varepsilon(MX^{-1/3+\varepsilon} + X^{3+\varepsilon}). \tag{4.34}$$

Fixing $\varepsilon < 1/3$, we get a lower bound of the desired type when $1 \ll X \ll M^{1/(3+\varepsilon)}$.

4.3 The fourth moment

As for the second moment, it is possible to derive an asymptotic for any of the moments. For simplicity, we derive only an upper bound for the fourth.

We have so far that the fourth moment is

$$O_\varepsilon(X^{6+\varepsilon}) + M\int_{\Re s_1\gg 0}\cdots\int_{\Re s_4\gg 0} \times \prod_{1\le i<j\le 4}\frac{1}{s_i+s_j-1}\cdot h(s_1,\dots,s_4)\frac{ds_1}{2\pi i}\cdots\frac{ds_4}{2\pi i}, \tag{4.35}$$

where $h(s_1,\dots,s_4)$ is $X^{s_1+\cdots+s_4-2}$ times a symmetric function, holomorphic in $\Re s_i > 1/3$ and of rapid decay along vertical lines in each variable.

Now we proceed to shift the contours of (4.35). We first arrange the lines of integration so that $1/2 < \Re s_1 < \cdots < \Re s_4 < 1/2+\delta$, with δ small. For brevity of notation, put $g(s) = \frac{1}{s-1}$. In what follows there will be many terms involving g and h, e.g. $g(2s_3)g(s_3+s_4)h(s_3, 1-s_3, s_3, s_4)$. (At each stage, we write only the integrand.) In all cases, the arguments to h will all be close in real part to 1/2. As soon as the sum of the arguments to h of any one term has real part ≤ 2, we get a bound of $O(1)$ for that term by integrating the absolute value. In the final stages of the argument, there will be terms involving derivatives of h that arise from second order poles. Differentiating h gives rise to the factor of $\log X$ that appears in the final answer.

In full detail, the function to be integrated is

$$g(s_1+s_2)g(s_1+s_3)g(s_1+s_4)g(s_2+s_3)g(s_2+s_4)g(s_3+s_4)h(s_1,s_2,s_3,s_4). \tag{4.36}$$

We start by shifting the contour of s_1 to the left of $1/2$. We get three terms from the residues of the poles at $1-s_2$, $1-s_3$, and $1-s_4$:

pole	residue
$s_1=1-s_2$	$g(s_2+s_3)g(s_2+s_4)g(s_3+s_4)g(1-s_2+s_3)g(1-s_2+s_4)$ $h(1-s_2,s_2,s_3,s_4)$
$s_1=1-s_3$	$g(s_2+s_3)g(s_2+s_4)g(s_3+s_4)g(1-s_3+s_2)g(1-s_3+s_4)$ $h(1-s_3,s_2,s_3,s_4)$
$s_1=1-s_4$	$g(s_2+s_3)g(s_2+s_4)g(s_3+s_4)g(1-s_4+s_2)g(1-s_4+s_3)$ $h(1-s_4,s_2,s_3,s_4)$

(4.37)

Once s_1 is far enough to the left of $1/2$, we bound the error term, as described above.

Next, we handle the terms of (4.37) separately. For the first term, we shift the line of s_3 to bring the argument sum to the left of 2. For the second and third, we shift s_2. Note that for the third term we need only shift s_2 past the pole at $1-s_3$ to pass to the left of 2. In total, we have six terms, with an error of $O(1)$:

pole	residue
$s_3=s_2$	$g(2s_2)g(s_2+s_4)^2g(1-s_2+s_4)h(1-s_2,s_2,s_2,s_4)$
$s_3=1-s_2$	$g(s_2+s_4)g(1-s_2+s_4)^2g(2-2s_2)h(1-s_2,s_2,1-s_2,s_4)$
$s_3=1-s_4$	$g(1-s_4+s_2)g(s_2+s_4)g(2-s_2-s_4)g(1-s_2+s_4)$ $h(1-s_2,s_2,1-s_4,s_4)$
$s_2=1-s_3$	$g(1-s_3+s_4)^2g(s_3+s_4)g(2-2s_3)h(1-s_3,1-s_3,s_3,s_4)$
$s_2=1-s_4$	$g(1-s_4+s_3)g(s_3+s_4)g(2-s_3-s_4)g(1-s_3+s_4)$ $h(1-s_3,1-s_4,s_3,s_4)$
$s_2=1-s_3$	$g(1-s_3+s_4)g(s_3+s_4)g(2-s_3-s_4)g(1-s_4+s_3)$ $h(1-s_4,1-s_3,s_3,s_4)$

(4.38)

We bound the terms of (4.38) with argument sum 2. Further, two of the remaining three terms are equal (by changing s_3 into s_2). Thus, we are down to two terms:

$$g(2s_2)g(s_2+s_4)^2g(1-s_2+s_4)h(1-s_2,s_2,s_2,s_4) \tag{4.39}$$

and

$$2g(s_2+s_4)g(1-s_2+s_4)^2g(2-2s_2)h(1-s_2,s_2,1-s_2,s_4). \tag{4.40}$$

For term (4.40), we shift s_4 to left of the pole at s_2, with residue

$$\frac{\partial}{\partial s_4} 2g(s_2 + s_4)g(2 - 2s_2)h(1 - s_2, s_2, 1 - s_2, s_4)\Big|_{s_4 = s_2}. \tag{4.41}$$

This residue involves a first-order derivative of h and has argument sum 2, and thus is $O(\log X)$.

As for (4.39), we shift s_2 to the left, past the poles at $1/2$ and $1 - s_4$:

$$\frac{1}{2} g\left(\frac{1}{2} + s_4\right)^3 h\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, s_4\right) + \frac{\partial}{\partial s_2} g(2s_2)g(1 - s_2 + s_4)h(1 - s_2, s_2, s_2, s_4)\Bigg|_{s_2 = 1 - s_4}. \tag{4.42}$$

The second of these terms is again $O(\log X)$.

Now, the first term would ordinarily have a triple pole at $s = 1/2$, and would thus be of size $(\log X)^2$. However, we saw that we may assume that the smoothing function F is balanced, i.e. $\widetilde{F}(1/2) = 0$. Then the first term vanishes.

This completes the proof that the fourth moment is $O_\varepsilon(M \log X + X^{6+\varepsilon})$. For fixed ε, this gives our upper bound when $X \ll M^{1/(6+\varepsilon)}$.

5. Approximate functional equations

In this section we see how the approximate functional equation discussed in the introduction leads naturally to a non-linear regression problem and a more precise test of modularity than that of Section 3.

Our starting point is the proof of Theorem 3.1; there π was an automorphic representation of $\mathrm{GL}_r(\mathbb{A}_\mathbb{Q})$, and to complete the proof we estimated (3.5). Here instead we reinterpret the integral as a smooth sum of Fourier coefficients. In order to make explicit the dependence on the smoothing function F, we write $S_F^*(X)$. Next, put $\widetilde{G}(s) = \widetilde{F}(1-s)\gamma(s, \widetilde{\pi})/\gamma(1-s, \pi)$ and define $G(x)$ to have Mellin transform $\widetilde{G}(s)$. Changing s into $1 - s$, (3.5) becomes then

$$S_F^*(X) = \frac{\epsilon_\pi}{2\pi i} \int L(s, \widetilde{\pi})\widetilde{G}(s) \left(\frac{N_\pi}{X}\right)^{s-1/2} ds = \frac{\epsilon_\pi}{\sqrt{N_\pi / X}} \sum_{n=1}^{\infty} \lambda_{\widetilde{\pi}}(n) G(nX/N_\pi). \tag{5.1}$$

Thus, the functional equation yields a relation between smooth sums S_F^* and S_G^*, with smoothing functions related by the generalized Bessel transformation $F \mapsto G$ defined above. This transformation is simplest when $\widetilde{F}(s) = \gamma(s, \pi)$, i.e. $F(x) = F_\pi(x)$, for which we have $G(x) = \overline{F_\pi(x)}$. Since $\lambda_{\tilde{\pi}}(n) = \overline{\lambda_\pi(n)}$, (5.1) thus becomes

$$S^*_{F_\pi}(X) = \epsilon_\pi \overline{S^*_{F_\pi}(N_\pi/X)}. \tag{5.2}$$

Hence, if we write $X = N_\pi^\sigma$ as in Section 3, then $S^*_{F_\pi}$ satisfies a functional equation relating σ to $1-\sigma$.

Throughout the remainder of this section we will use the smoothing function F_π and again suppress it from the notation.

5.1 An optimization problem

Now, (5.2) is an exact equality and can be used with smaller values of X than the tests of Section 3. However, it depends crucially on knowledge of all Euler factors. To handle partial L-functions, recall the setting of Section 3.2. There we defined the smooth sum $S(X)$ for a partial L-function, and saw the relation

$$S(X) = \sum_n \frac{c_n}{\sqrt{n}} S^*(X/n), \tag{5.3}$$

where $\sum_n c_n n^{-s} = \prod_{p|N_\pi} L_p(s,\pi)^{-1}$ is the finite Dirichlet series equal to the product of missing local factor polynomials. There is an inverse identity to (5.3). Let Σ_π denote the set of positive integers whose prime factors divide N_π. Then,

$$\sum_{n\in\Sigma_\pi} \lambda_\pi(n) n^{-s} = \prod_{p|N_\pi} L_p(s,\pi) = \frac{1}{\sum_n c_n n^{-s}}, \tag{5.4}$$

so that

$$S^*(X) = \sum_{n\in\Sigma_\pi} \frac{\lambda_\pi(n)}{\sqrt{n}} S(X/n). \tag{5.5}$$

Substituting this expression into (5.2), we have

$$\sum_{n\in\Sigma_\pi} \frac{\lambda_\pi(n)}{\sqrt{n}} S(X/n) = \epsilon_\pi \overline{\sum_{n\in\Sigma_\pi} \frac{\lambda_\pi(n)}{\sqrt{n}} S(N_\pi/nX)}. \tag{5.6}$$

Next, recall that the local factors at primes $p|N_\pi$ are of degree at most $r-1$, so they take the general form

$$L_p(s,\pi) = (1 + x_p^{(1)} p^{-s} + \cdots + x_p^{(r-1)} p^{-(r-1)s})^{-1}. \tag{5.7}$$

Thus, while there are an infinite number of unknown coefficients $\lambda_\pi(n)$ for $n \in \Sigma_\pi$, they are determined by the finitely many numbers $x_p^{(i)}$. Those numbers, together with ϵ_π, may be viewed as the unique parameters that minimize

$$\sum_X \left| \sum_{n\in\Sigma_\pi} \frac{1}{\sqrt{n}} (\lambda_\pi(n) S(X/n) - \epsilon_\pi \overline{\lambda_\pi(n) S(N_\pi/nX)}) \right|^2, \qquad (5.8)$$

where X ranges over a collection of sample points. Further, the parameters are constrained by the Ramanujan bound $|x_p^{(i)}| \leq \binom{r-1}{i}$, and $|\epsilon_\pi| = 1$. Thus, (5.8) gives a bounded (non-linear) optimization problem, whose solution may be found to high accuracy in finite time, e.g. by Newton's method. (Note that the sum over Σ_π is essentially finite, since $S(X/n)$ is very rapidly decreasing in n.)

The existence of a solution which makes (5.8) small gives evidence for the analytic continuation and functional equation of the L-function. The strength of that evidence depends on how over-determined the system is, in terms of the number of significant coefficients $\lambda_\pi(n)$ versus the range of sample points X. In practical terms, any function can be well-modeled as a linear combination of translates of another if enough coefficients are allowed. Thus, if there are many missing Euler factors for small primes (and thereby many choices for $\lambda_\pi(n)$) then such a solution would perhaps not be surprising even in the absence of a functional equation. Moreover, in that situation the solution space is large and the minimum not very stable, so that it can take a long time to locate it.

The answer to both problems, to make the evidence more convincing and to increase the system stability, is to add more data. That can be done by calculating more Fourier coefficients, increasing the range of possible sample points X. However, a more efficient use of the data is possible by appealing to twisted functional equations. That is, we add data from the approximate functional equations of $L(s, \pi \times \chi)$ for Dirichlet characters χ of conductor q relatively prime to N_π. Precisely, for those L-functions we consider in place of (5.8)

$$\sum_X \left| \sum_{n\in\Sigma_\pi} \frac{1}{\sqrt{n}} (\lambda_\pi(n)\chi(n) S_\chi(X/n) - \epsilon_{\pi\times\chi} \overline{\lambda_\pi(n)\chi(n) S_\chi(N_\pi q^r/nX)}) \right|^2, \qquad (5.9)$$

where $S_\chi(X) = \frac{1}{\sqrt{X}} \sum_{(n,N_\pi)=1} \lambda_\pi(n)\chi(n) F_{\pi\times\chi}(n/X)$. For simplicity, we may choose χ unramified at ∞ so that $F_{\pi\times\chi} = F_\pi$; this is not an issue for the Hasse–Weil L-functions of genus 2 curves over $\mathbb{Q}$ or elliptic curves over $\mathbb{Q}(\sqrt{-1})$, since their Γ-factors can be expressed entirely in terms of $\Gamma_\mathbb{C}(s) = \Gamma_\mathbb{R}(s)\Gamma_\mathbb{R}(s+1)$ (see below), which is invariant under twisting by a character over $\mathbb{Q}$. Note that only finitely many twists yield new data, since by twisting the conductor is multiplied by a factor of q^r. However, if willing to change smoothing functions, one may also consider higher rank twists or functorial transfers; we will not pursue that possibility.

Although our basic philosophy is that modularity should be determined by a single twist, using many twists can be useful in this way computationally. The essential point that makes it work is that the coefficients of the twisted L-function are determined by those of the original, and thus the twisted functional equations do not introduce more variables. To that end, we also need to know the relationship between ϵ_π and $\epsilon_{\pi\times\chi}$, which is given in general in [25,16]. When π has trivial central character and χ is unramified at ∞, we have simply

$$\epsilon_{\pi\times\chi} = \epsilon_\chi^r \chi(N_\pi)\epsilon_\pi, \tag{5.10}$$

where $\epsilon_\chi = \tau(\chi)/\sqrt{q}$ is the ϵ-factor of $L(s,\chi)$.

5.2 More notes on computation

The problem of efficiently computing values of the K-Bessel function $K_\nu(x)$ is classic; Whittaker and Watson [52] list a dozen or so formulas related to it. They can be divided into four types: integral representations, power series for small x, asymptotics for large x, and differential equations and recursion relations. In this subsection we describe briefly how to generalize such results to the inverse Mellin transform

$$F(x) = \frac{1}{2\pi i}\gamma(1)^{-1}\int \gamma(s)x^{-s}\,ds \tag{5.11}$$

of a general Γ-factor $\gamma(s) = \prod_{i=1}^r \Gamma_{\mathbb{R}}(s+\mu^{(i)})$, which is needed to carry out the test described above. The factor $\gamma(1)^{-1}$ is a normalization that amounts to insisting $\int_0^\infty F(x)\,dx = 1$ and thus avoids any question of the proper normalization of $\gamma(s)$. As we will see, the latter three types of formulas correspond to the behaviour of $\gamma(s)$ in the left and right half-planes and its recursion relations. (See also [18] and [34] for more detailed information, and [2] for a discussion of such Mellin transforms in connection with high moments of $\zeta(s)$ and a formulation of the Lindelöf hypothesis in terms of smooth sums.)

First, for small x, consider the simplest of all such functions,

$$e^{-x} = \frac{1}{2\pi i}\int \Gamma(s)x^{-s}\,ds, \tag{5.12}$$

with the contour initially a vertical line far to the right. Shifting the contour to the left, we pick up contributions from the poles of $\Gamma(s)$ at integers $s=-k$, with residues $(-1)^k/k!$. Thus, we recover the MacLaurin series

$$e^{-x} = \sum_{k=0}^\infty \frac{(-1)^k}{k!}x^k. \tag{5.13}$$

More generally, from the poles of $\gamma(s)$, which may have multiplicity, we get a power and log series of the form

$$F(x) = \frac{1}{2\pi i}\gamma(1)^{-1}\int \gamma(s)x^{-s}\,ds = \sum_{\nu} P_\nu(\log x)x^{\nu}, \tag{5.14}$$

where $-\nu$ ranges over the discrete set of poles of $\gamma(s)$ and P_ν is a polynomial of degree at most $r-1$.

In order to evaluate the terms of (5.14), we need the Laurent series expansion of the Γ function about an arbitrary point in the complex plane. This is obtained in a straightforward manner from the Weierstrass expansion

$$-\frac{\Gamma'}{\Gamma}(s) = \frac{1}{s} + \gamma + \sum_{n=1}^{\infty}\left(\frac{1}{n+s} - \frac{1}{n}\right), \tag{5.15}$$

where $\gamma = 0.577\ldots$ is Euler's constant. In particular, for the Hasse–Weil L-function of a curve, the local representation at the infinite places is of Hodge type, meaning that, up to a shift, the archimedean parameters $\mu^{(i)}$ are integers. For both genus 2 curves over $\mathbb{Q}$ and elliptic curves over an imaginary quadratic field, we have simply $\gamma(s) = \Gamma_{\mathbb{C}}(s+1/2)^2$, where $\Gamma_{\mathbb{C}}(s) = 2(2\pi)^{-s}\Gamma(s)$, for which (5.15) leads to the Laurent series

$$\Gamma_{\mathbb{C}}(s) = \frac{(-2\pi)^k}{k!}\frac{2}{s+k}\exp\left[\left(\sum_{n=1}^{k}\frac{1}{n} - \gamma - \log 2\pi\right)(s+k) + \sum_{j=2}^{\infty}\frac{1}{j}\left((-1)^j\zeta(j) + \sum_{n=1}^{k}\frac{1}{n^j}\right)(s+k)^j\right]. \tag{5.16}$$

For later examples it will be helpful to handle any Γ-factor that is a product of $\Gamma_{\mathbb{R}}(s)$ and $\Gamma_{\mathbb{R}}(s+1)$. Laurent expansions of these can be obtained easily from (5.16) and the Legendre duplication formula

$$\Gamma\left(\frac{s}{2}\right)\Gamma\left(\frac{s+1}{2}\right) = 2^{-s}\Gamma(s). \tag{5.17}$$

Next, we consider x large. In the right half-plane, any Γ factor of degree r looks roughly (up to a power of s) like $(\pi r)^{-rs/2}\Gamma(rs/2)$. More precisely, Stirling's formula yields

$$\gamma(s) = \prod_{i=1}^{r}\Gamma_{\mathbb{R}}(s+\mu^{(i)}) = \sqrt{r2^{r-1}}(\pi r)^{-(rs+\mu)/2}\Gamma\left(\frac{rs+\mu}{2}\right)\left(1+O\left(\frac{1}{s}\right)\right), \tag{5.18}$$

where $\mu = 1 + \sum_{i=1}^{r}(\mu^{(i)} - 1)$. This leads to the asymptotic

$$\begin{aligned} F(x) &\sim \frac{1}{2\pi i}\int \sqrt{r2^{r-1}}(\pi r)^{-(rs+\mu)/2}\Gamma\left(\frac{rs+\mu}{2}\right)x^{-s}\,ds \\ &= \sqrt{\frac{2^{r+1}}{r}}x^{\mu/r}e^{-\pi r x^{2/r}}. \end{aligned} \tag{5.19}$$

It is not immediately clear from (5.18) and (5.19) that the error term is small, but that may be rigorously verified by the method of stationary phase.

Note that the decay of (5.19) is subexponential for $r > 2$. Because of this, and for later use in estimating error terms, it is natural to consider F in terms of the variable $y = \pi r x^{2/r}$; (5.19) then takes the simple form of a constant times $y^{\mu/2}e^{-y}$. By using more terms from Stirling's series in (5.18), we may also obtain more terms in the asymptotic expansion of F, in powers of y^{-1}.

Lastly, note that multiplication of $\gamma(s)$ by s amounts to applying $-x\frac{d}{dx}$ to $F(x)$. Thus, from the recursion formula for Γ, F satisfies the differential equation

$$\left[\prod_{i=1}^{r}\left(x\frac{d}{dx} - \mu^{(i)}\right)\right]F(x) = (-2\pi)^r x^2 F(x). \tag{5.20}$$

In terms of y, this becomes

$$\frac{d^r F}{dy^r} = (-1)^r F + \sum_{k=0}^{r-1} a_k y^{k-r}\frac{d^k F}{dy^k}, \tag{5.21}$$

with coefficients a_k defined recursively by

$$(k - r\mu^{(1)}/2)\cdots(k - r\mu^{(r)}/2) + a_0 + a_1 k + a_2 k(k-1) + \cdots + a_k k! = 0. \tag{5.22}$$

This may be viewed as a perturbation of the equation $\frac{d^r F}{dy^r} = (-1)^r F$, to which Ce^{-y} is the unique solution of fastest decay.

Moreover, using the differential equation it is possible to express all derivatives of F in terms of $F, \frac{dF}{dy}, \ldots, \frac{d^{r-1}F}{dy^{r-1}}$. Thus, the Taylor series for F around any point is determined by its first r terms. That suggests an algorithm for computing F by numerically solving the differential equation. Namely, starting with an initial set of values for the first r derivatives of F at some y_0, we obtain a Taylor series with which we compute the first r derivatives at $y_0 + h$ for a small increment h. Repeating this procedure we get Taylor expansions for F about a lattice of points, which is useful for later rapid evaluation.

It makes sense to work with the variable y rather than x, since the Taylor series for e^{-y} has the same rate of convergence around all points. That will be true for F as well when y is large, where its behaviour is controlled by the asymptotic. Since F may have a singularity at 0, the method becomes inefficient for small y, and it is better to use the power and log series expansion there. Also, because we are searching for the solution of most rapid decay to the differential equation, to avoid rampant error from other solutions it is best to start with a large value of y, with initial terms perhaps chosen from the asymptotic, and work downwards.

5.3 Results

We applied the regression analysis discussed above to the partial L-functions computed in Section 2. In all cases we found parameter values $x_p^{(i)}$ that made (5.8) zero to within the accuracy of the computation (about 12 decimals). The partially recovered local factors are listed in Tables 5.1 and 5.2, with coefficients $x_p^{(i)} p^{i/2}$ normalized to be integers. This extra arithmetic information was not used in the regression, so the fact that all coefficients are close to integers provides yet another verification of modularity. In Table 5.2 the entries at split primes $p = \mathfrak{p}\bar{\mathfrak{p}}$ are of smaller degree because in each case the curve has good reduction at one of $\mathfrak{p}$ or $\bar{\mathfrak{p}}$; only the part corresponding to the prime of bad reduction is listed. Note that the method only reproduces the Euler factorization over $\mathbb{Q}$, so that if a curve had bad reduction at both $\mathfrak{p}$ and $\bar{\mathfrak{p}}$ it would not be possible to distinguish the factors at those primes without more work.

Note also that not all coefficients could be recovered from the computed data; undetermined coefficients are identified in the tables with a ? mark. That is because the parameter $x_p^{(i)}$ only affects coefficients $\lambda_\pi(n)$ for n at least p^i. As a rule of thumb, we found that with the coefficients for $n \leq M$ we could

Table 5.1. Euler factors for genus 2 curves over $\mathbb{Q}$.

p	approximate local factor polynomial at p	factorization
	curve 277a, $N = 277$, $\epsilon = 1$:	
277	$1 - 7.0000000000t + 268.998t^2 + ?t^3$	$(1+t)(1-8t+277t^2)$
	curve 1757a, $N = 1757$, $\epsilon = -1$:	
7	$1 + 2.0000000000t + 4.0000000000t^2 - 7.0000000000t^3$	$(1-t)(1+3t+7t^2)$
251	$1 - 11.0000000000t + 239.00000003t^2 + ?t^3$	$(1+t)(1-12t+251t^2)$
	curve 9136a, $N = 1142$, $\epsilon = 1$:	
2	$1 - 1.0000000000t + 2.0000000000t^2 - 2.0000000000t^3$	$(1-t)(1+2t^2)$
571	$1 + 3.0000000000t - 567.0000007t^2 + ?t^3$	$(1-t)(1+4t+571t^2)$
	curve 18690a, $N = 18690$, $\epsilon = 1$:	
2	$1 - 1.0000000000t - 0.0000000000t^2 + 2.0000000000t^3$	$(1+t)(1-2t+2t^2)$
3	$1 - 2.0000000000t + 4.0000000000t^2 - 3.0000000000t^3$	$(1-t)(1-t+3t^2)$
5	$1 - 1.0000000000t + 5.0000000000t^2 - 5.0000000000t^3$	$(1-t)(1+5t^2)$
7	$1 + 0.0000000000t + 6.0000000000t^2 - 7.0000000000t^3$	$(1-t)(1+t+7t^2)$
89	$1 + 5.0000000000t + 83.000000000t^2 - 88.999999997t^3$	$(1-t)(1+6t+89t^2)$

Table 5.2. Euler factors for elliptic curves over $\mathbb{Q}(\sqrt{-1})$.

p	approximate local factor polynomial at p
	curve 233a, $N = 16 \cdot 233$, $\epsilon = -1$:
233	$1 - 1.0000000000t$
	curve 3577a, $N = 16 \cdot 3577$, $\epsilon = -1$:
7	$1 - 1.0000000000t^2$
73	$1 - 1.0000000000t$
	curve 9657a, $N = 16 \cdot 9657$, $\epsilon = -1$:
3	$1 + 1.0000000000t^2$
29	$1 - 1.0000000000t$
37	$1 - 1.0000000000t$
	curve 16970a, $N = 16 \cdot 16970$, $\epsilon = -1$:
2	$1 + 1.0000000000t$
5	$1 - 1.0000000000t$
1697	$1 - 1.0000000000t$

accurately recover all $x_p^{(i)}$ with $p^i \leq$ about $\frac{M}{3}$, and the accuracy trails off for larger p^i. This is more of an issue for curves of low conductor, for which proportionally fewer coefficients were computed. However, the polynomials that occur in Hasse–Weil zeta functions have certain limitations. The irreducible factors at primes of bad reduction for a curve of genus 2, for example, can only be $1 \pm t$, $1 + t^2$ or $1 + at + pt^2$ with $|a| < 2\sqrt{p}$ (and interestingly each of these types can be found in the extended tables in [7]). Thus, in our case the numbers $x_p^{(1)}$ and $x_p^{(2)}$ are enough to completely recover the local factors; Table 5.1 also shows the factorization of each polynomial, even for those that have $x_p^{(3)}$ missing.

For about 10% of the curves tested the conductor was found to properly divide the discriminant. One of these is that of discriminant 9136 listed in Table 5.1; that fact could perhaps have been guessed from the pictures of Section 3. The genus 2 curve of discriminant $18690 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 89$ is also interesting since it has fifteen parameters $x_p^{(i)}$, the most of any of the examples. For that case, when multiplied out in terms of the parameters, (5.8) has more than 20 million significant terms. Despite that, the algorithm was able to recover efficiently the complete local factors.

Next, Figures 5.1 and 5.2 show graphs of $S^*(N^\sigma)$, where N is the apparent conductor, computed using the reconstructed local factors. The symmetry present in each graph reflects the functional equation. Thus, in order to get any information about the functional equation, it is plainly necessary to compute at least $\sqrt{N}$ coefficients of the L-series. It should be stressed, however, that $\sqrt{N}$ terms are not sufficient to verify analytic continuation, nor is it enough to give a cursory glance at the graph; the functional equation must be verified to a high degree of accuracy. That is because it is possible that the L-function

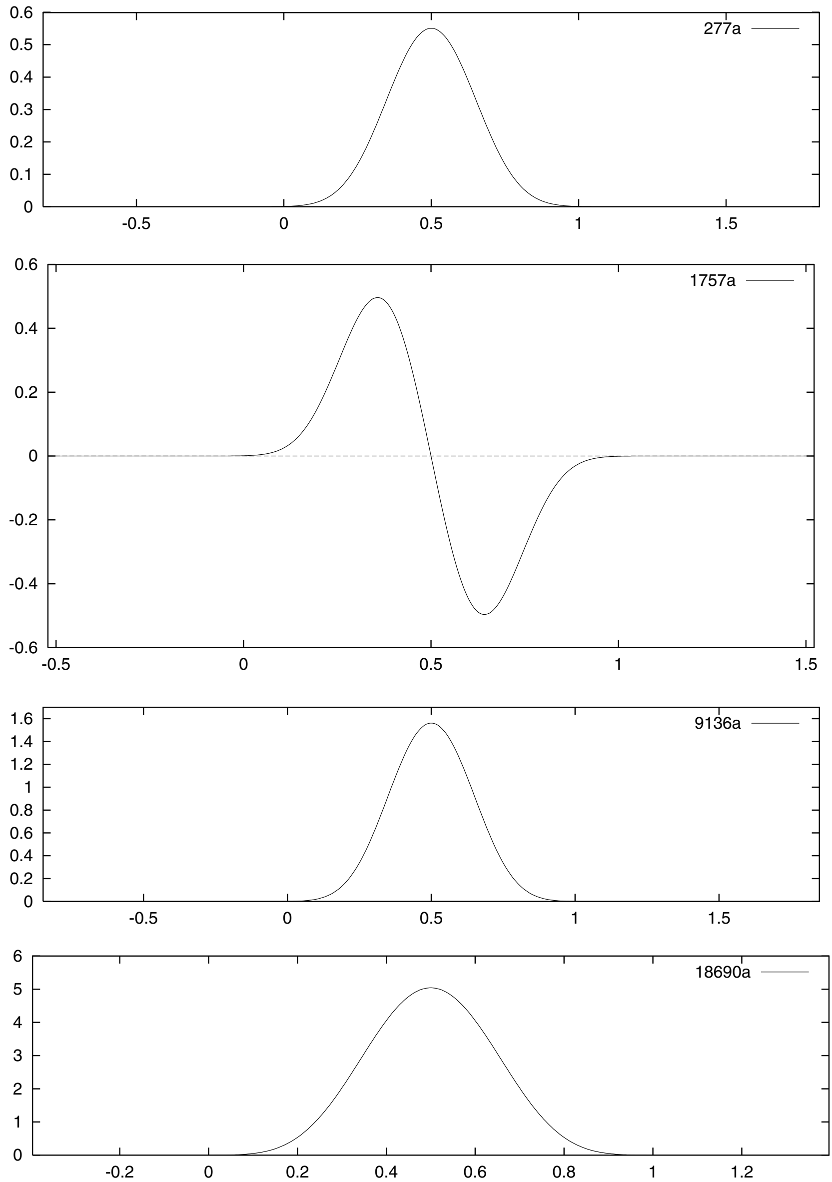

Figure 5.1. $S^*(N^\sigma)$ for curves of genus 2 over $\mathbb{Q}$

satisfies the appropriate functional equation, yet has only meromorphic continuation. Note that a pole of the L-function at $s = 1/2 + it$ with residue of size 1 would contribute an oscillation of amplitude roughly $e^{-\frac{\pi}{4}r|t|}$ to $S(X)$, due to the decay of the Γ-factor. On the other hand, cutting the L-series off at the nth term results in an error on the order of $e^{-\pi r (n/X)^{2/r}}$. Taking $X = \sqrt{N}$,

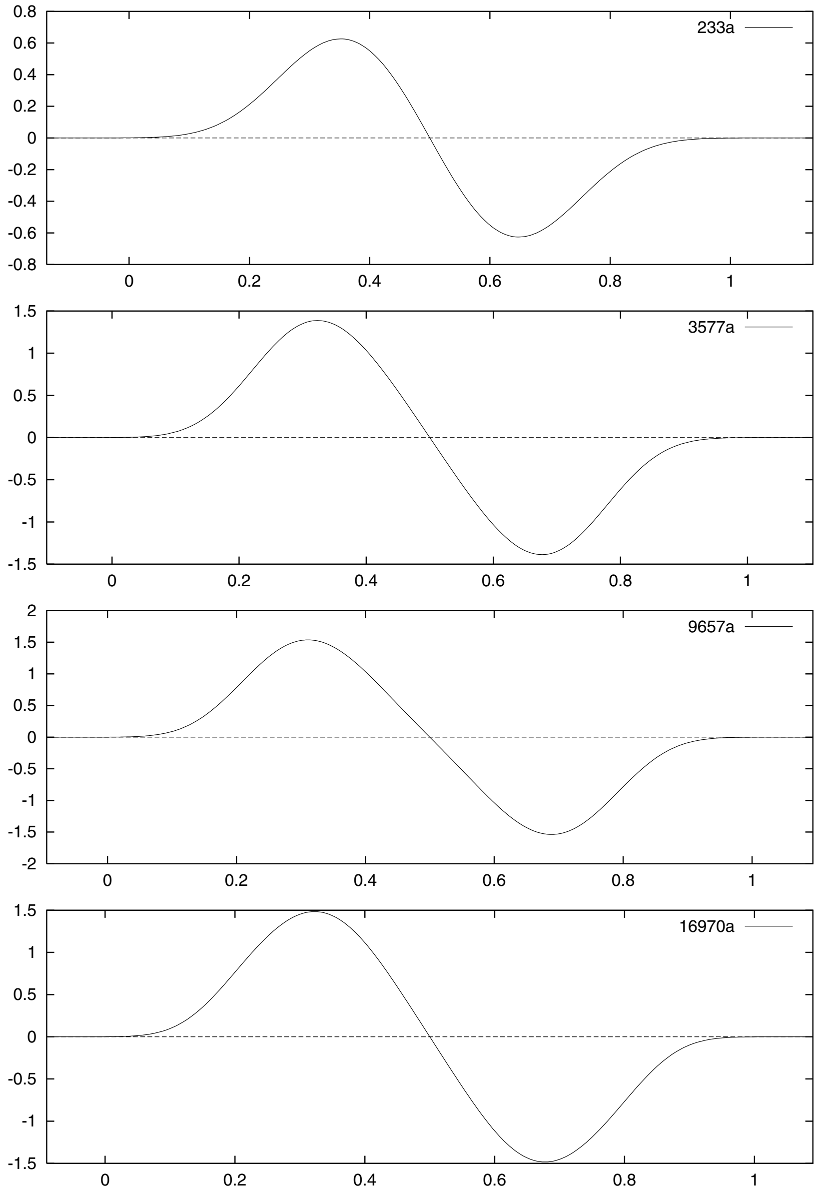

Figure 5.2. $S^*(N^\sigma)$ for elliptic curves over $\mathbb{Q}(\sqrt{-1})$

we see that to detect a pole with imaginary part t requires about $\sqrt{|t/4|^r N}$ terms. (This assumes the archimedean parameters are fixed; a finer analysis would include them as part of the analytic conductor.)

Other features can be read from the figures as well. For example, the area under each curve is proportional to $L(1/2)/\log N$, on the given scale. Thus, the Riemann hypothesis predicts that for those graphs that are even (so that $\epsilon = 1$) there should typically be a bias toward positive values; this is easily seen in the figures. The four examples in Figure 5.2 all happen to have $\epsilon = -1$. According to the Birch-Swinnerton-Dyer conjecture, the rank of the Mordell-Weil group in each case is therefore odd; see [14] for many more computations of this nature. Also, comparing to the graphs from Section 3, using the smoothing function F_π seems to take much of the "randomness" out of the picture. However, one should still think of $S^*(X)$ as random up to about $\sqrt{N}$, and indeed some examples from [7] have more peaks; see also Section 6 where we consider examples of much larger conductor.

6. Miscellaneous applications

6.1 A Galois representation

We consider in this section the L-function $L(s, \rho)$ of an even icosahedral Galois representation ρ, specifically the first such in the tables of Buhler [9]. It is a lift of minimal conductor $7947 = 3^2 \cdot 883$ of a projective representation of the Galois group of the polynomial

$$f(x) = x^5 + 5x^4 - 7x^3 - 11x^2 + 10x + 3. \tag{6.1}$$

That information alone is enough to determine the Dirichlet coefficients $\lambda_\rho(p)$ up to sign. Buhler shows how to determine the sign via computations in the idele class group of the splitting field of the sextic resolvent of f. Buhler's method is sufficient to find the first few hundred coefficients, but as he remarks, it seems to have exponential complexity and is impractical when many coefficients are needed.

Fortunately, the problem of finding a model for ρ itself was solved by Crespo [15] and Jehanne [26]. One such model is as follows. Let $\widetilde{\rho}$ denote the self-dual twist of ρ of minimal conductor $7017201 = 3^2 \cdot 883^2$. Then $\widetilde{\rho}$ is a representation of the Galois group of

$$\begin{aligned} g(x) = x^{24} &- 1719x^{22} + 1033803x^{20} \\ &- 247759929x^{18} + 18891275922x^{16} - 520186579740x^{14} \\ &+ 4132529834850x^{12} - 12489586256925x^{10} + 17037668338668x^8 \\ &- 10424803822722x^6 + 2281972718133x^4 - 481611663x^2 + 729, \end{aligned} \tag{6.2}$$

isomorphic to a cyclic central extension of A_5, and ρ is a twist of $\widetilde{\rho}$ by a cubic character of conductor 883. It turns out in this case that the coefficients

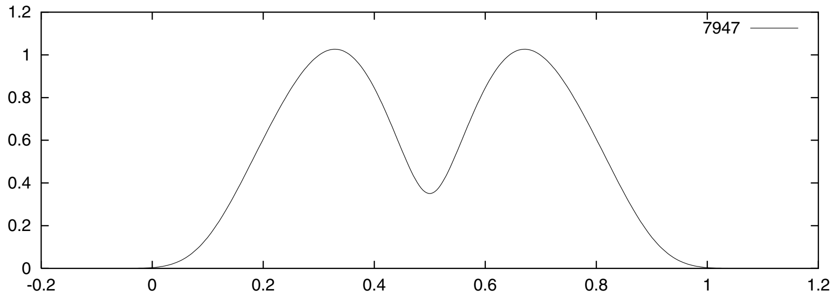

Figure 6.1. $|S^*(N^\sigma)|$ for the Galois representation ρ of conductor $N = 7947$

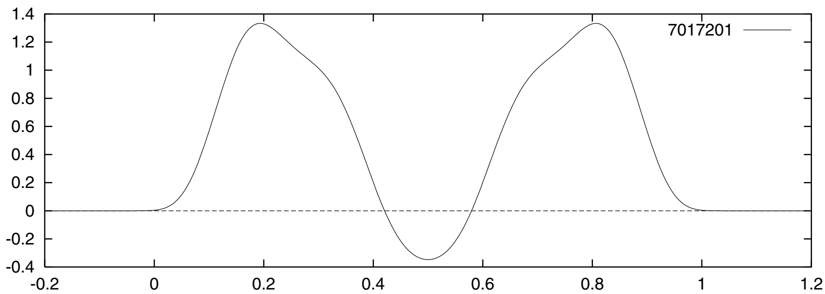

Figure 6.2. $S^*(N^\sigma)$ for the self-dual twist $\widetilde{\rho}$ of conductor $N = 7017201$

$\lambda_{\widetilde{\rho}}(p)$ are determined by the factorization of $g(x)$ modulo p, together with a trick that Buhler attributes to Serre for distinguishing the conjugacy classes of orders 5 and 10. Thus, no class group computations are needed, and there is an algorithm that computes $\lambda_\rho(p)$ in time $O(\log p)$. Hence, the coefficients $\lambda_\rho(n)$ for $n \le M$ may be computed in time $O(M)$, which up to a constant factor is best possible.

Using this method, we computed 10^9 coefficients $\lambda_\rho(n)$, more than enough to compute $S^*(X)$ for both ρ and $\widetilde{\rho}$; we are indebted to Jehanne for providing code that was needed for part of these computations. The results are shown in Figures 6.1 and 6.2; note since $S^*(X)$ is complex for ρ, we plot the absolute value. In both cases we see the symmetry $\sigma \to 1 - \sigma$. Note that if it were possible to *prove* that these symmetries exist then modularity of $L(s, \rho)$ would follow from the converse theorem of Section 6.

6.2 High symmetric powers

We consider now the symmetric power L-functions $L(s, \mathrm{Sym}^n \Delta)$ and $L(s, \mathrm{Sym}^n f)$ where $\Delta \in S_{12}(\Gamma(1))$ and $f \in S_2(\Gamma_0(11))$, normalized to have first Fourier coefficient 1. Both of these examples may be computed efficiently using identities of modular forms. For example, we take advantage of the fact that $S_{12}(\Gamma(1))$ is 1-dimensional, so that any two modular forms of weight

12 can be combined to yield Δ; in particular, by comparing initial Fourier coefficients we find that

$$\Delta = \frac{691}{2^6 3^5 7^2}(E_{12} - E_6^2), \tag{6.3}$$

where E_6 and E_{12} are the Eisenstein series [37]

$$E_6(z) = 1 - 504 \sum_{n=1}^{\infty} \sigma_5(n) e(nz) \quad \text{and} \quad E_{12}(z) = 1 + \frac{65520}{691} \sum_{n=1}^{\infty} \sigma_{11}(n) e(nz). \tag{6.4}$$

This involves only one squaring operation that can be done quickly using fast Fourier transform multiplication over finite fields. For M coefficients, one Fourier transform takes $O(M \log M)$ time, and we must perform $O(\log M)$ transforms over different fields for full precision because the weight is large. Thus M coefficients may be computed in time proportional to $M(\log M)^2$, only slightly worse than the Galois representation example.

The form f may be computed likewise from the identity [31]

$$f(z) = \eta^2(z)\eta^2(11z), \tag{6.5}$$

where

$$\eta(z) = e(z/24) \prod_{n=1}^{\infty} (1 - e(nz)) = e(z/24) \sum_{n=-\infty}^{\infty} (-1)^n e\left(\frac{3n(n+1)}{2} z\right). \tag{6.6}$$

Note that η^2 may be computed directly from (6.6) in time $O(M)$, so that this again involves only one multiplication. Since f has weight 2, precision is not a problem for numbers of the size that we consider, so this algorithm runs in time effectively $O(M \log M)$. However, in both cases the algorithm suffers from the fact that we must store all coefficients at once, which limits the size of M. To compute many coefficients of f, we may instead use the fact that $X_0(11)$ is an elliptic curve, and count points using Shanks' $O(M^{5/4})$ algorithm, which compares favorably in speed yet needs only $O(M^{1/4})$ memory. One model for $X_0(11)$ is [31]

$$y^2 + y = x^3 - x^2. \tag{6.7}$$

Using these ideas we computed $2^{26} \approx 67 \times 10^6$ coefficients for both Δ and f. With these numbers, we compute the symmetric power L-functions; precisely, if $L(s, \Delta)$ has local factors

$$L_p(s, \Delta) = \frac{1}{(1 - \alpha(p)p^{-s})(1 - \alpha(p)^{-1}p^{-s})} \tag{6.8}$$

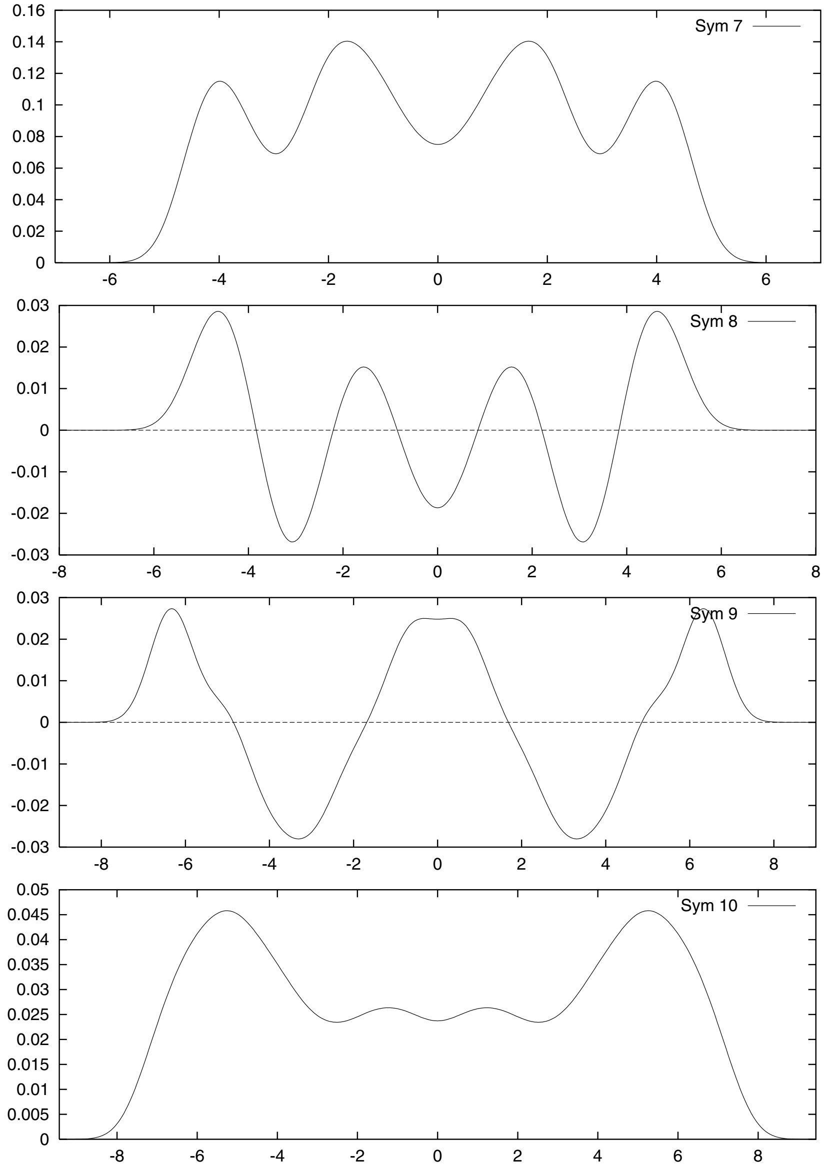

Figure 6.3. $S^*(e^\sigma)$ for $\mathrm{Sym}^n\,\Delta$

with $\alpha(p)$ and $\alpha(p)^{-1}$ the Satake parameters, then $L(s, \mathrm{Sym}^n\,\Delta)$ has local factors

$$L_p(s, \mathrm{Sym}^n\,\Delta) = \prod_{-n \le j \le n,\ j \equiv n\ (2)} (1 - \alpha(p)^j p^{-s})^{-1}. \tag{6.9}$$

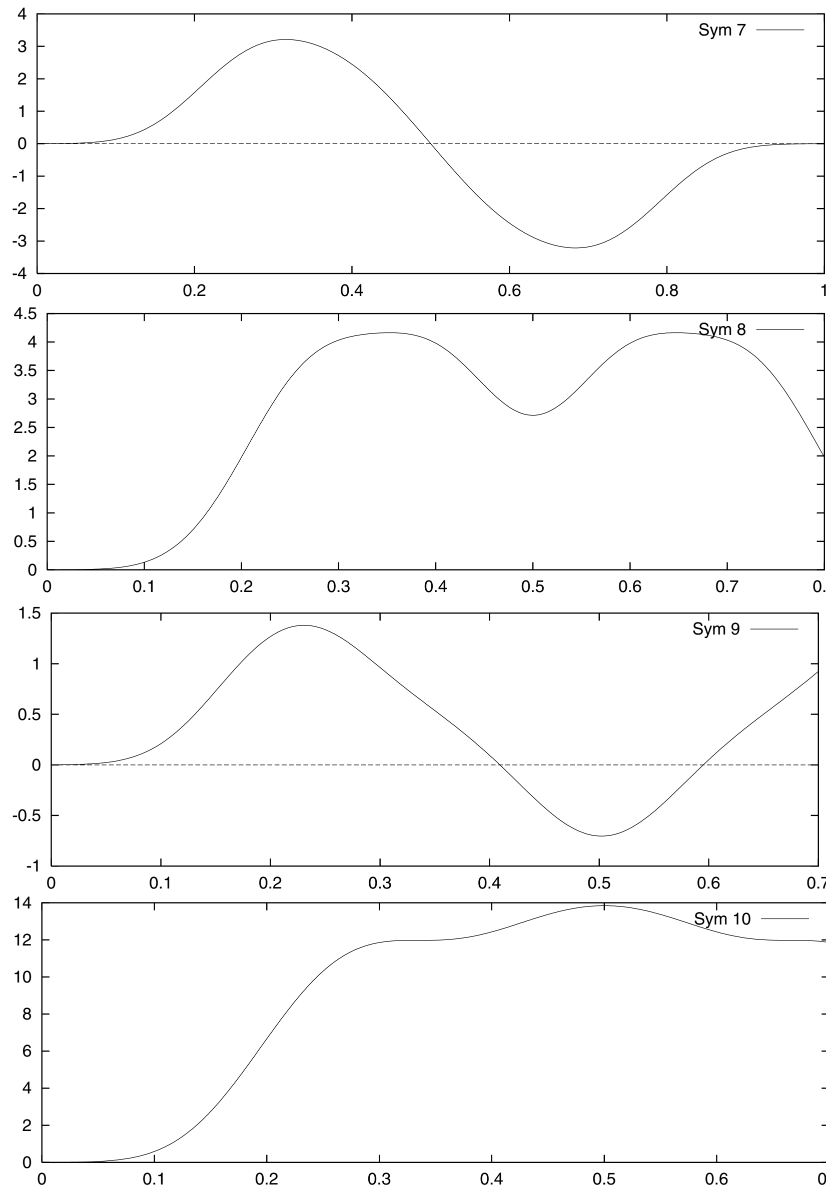

Figure 6.4. $S^*(N^\sigma)$ for $\mathrm{Sym}^n f$ of conductor $N = 11^n$

The formulation is similar for f except that the local representation at $p = 11$ is special for all powers. At this place we found the local factors to be

$$L_{11}(s, \mathrm{Sym}^n f) = (1 - 11^{-s-n/2})^{-1} \tag{6.10}$$

using the techniques of Section 5.

Finally, the Γ-factor in each case is [36]

$$\prod_{1\le j\le n,\ j\equiv n\ (2)} \Gamma_{\mathbb{C}}\left(s + j\cdot\frac{k-1}{2}\right)\cdot \prod_{a\in\{0,1\},\ 2a\equiv n\ (4)} \Gamma_{\mathbb{R}}(s+a), \qquad (6.11)$$

where k is the weight.

Figures 6.3 and 6.4 show $S^*(X)$ for the seventh through tenth symmetric powers of Δ and f. Note that since the conductor N is 1 for all powers of Δ, we cannot use our normal scale $X = N^\sigma$; in this case we put instead $X = e^\sigma$. The conductor of $\mathrm{Sym}^n f$ is $N = 11^n$.

Again we see a symmetry in each case. Note however that the coefficients we have computed are insufficient to see the decay of $S^*(X)$ for $\mathrm{Sym}^n f$ for $n \ge 8$. This could remedied for $n = 8$ and perhaps $n = 9$, but it quickly becomes impractical to go higher. (Note that in each graph 90% of the data are needed in only the last inch.) We remark that $L(s, \mathrm{Sym}^n \Delta)$ and $L(s, \mathrm{Sym}^n f)$ are meromorphic for $6 \le n \le 9$ [30], but this is not known for the tenth power.

Acknowledgements

This paper is an abbreviated account of my thesis [4]. I thank my advisor, Peter Sarnak, whose help during the preparation of the thesis was invaluable; many of the ideas contained here are due to him. I would also like to thank Robert Langlands for helpful comments and for inspiring much of this work.

References

[1] A. O. L. Atkin and J. Lehner, Hecke operators on $\Gamma_0(m)$, *Math. Ann.*, **185** (1970) 134–160. MR 42 #3022

[2] Richard Bellman, Wigert's approximate functional equation and the Riemann zeta-function, *Duke Math. J.*, **16** (1949) 547–552. MR 11,234c

[3] Andrew R. Booker, A test for identifying Fourier coefficients of automorphic forms and application to Kloosterman sums, *Experiment. Math.*, **9** (2000) no. 4, 571–581. MR 2001k:11077

[4] Andrew R. Booker, *Numerical tests of modularity*, Ph.D. thesis, Princeton University, 2003, available from `http://www.umich.edu/~arbooker/papers/`.

[5] Andrew R. Booker, Poles of Artin L-functions and the strong Artin conjecture, *Ann. of Math. (2)*, **158** (2003), no. 3, 1089–1098. MR MR2031863 (2004k:11082)

[6] Andrew R. Booker, *Artin's conjecture, Turing's method and the Riemann hypothesis*, preprint (2005).

[7] Booker, Andrew R, *web page containing data files and plots of extended examples*, 2005, available from `http://www.umich.edu/~arbooker/papers/thesis/`.

[8] Christophe Breuil, Brian Conrad, Fred Diamond and Richard Taylor, On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercises, *J. Amer. Math. Soc.*, **14** (2001) no. 4, 843–939 (electronic). MR 2002d:11058

[9] J. P. Buhler, *Icosahedral Galois representations*, Springer-Verlag, Berlin, 1978, Lecture Notes in Mathematics, vol. 654. MR 58 #22019

[10] Daniel Bump and Solomon Friedberg, *The exterior square automorphic L-functions on* GL(n), Festschrift in honor of I. I. Piatetski-Shapiro on the occasion of his sixtieth birthday, Part II (Ramat Aviv, 1989), Israel Math. Conf. Proc., vol. 3, Weizmann, Jerusalem, 1990, pp. 47–65. MR 93d:11050

[11] Daniel Bump and David Ginzburg, Symmetric square L-functions on GL(r), *Ann. of Math. (2)* **136** (1992) no. 1, 137–205. MR 93i:11058

[12] Ching-Li Chai and Wen-Ching Winnie Li, *Function fields: arithmetic and applications*, Applications of curves over finite fields (Seattle, WA, 1997), Contemp. Math., vol. 245, Amer. Math. Soc., Providence, RI, 1999, pp. 189–199. MR 2001f:11151

[13] J. W. Cogdell and I. I. Piatetski-Shapiro, Converse theorems for GL_n. II, *J. Reine Angew. Math.* **507** (1999) 165–188. MR 2000a:22029

[14] J. E. Cremona and E. Whitley, Periods of cusp forms and elliptic curves over imaginary quadratic fields, *Math. Comp.*, **62** (1994) no. 205, 407–429. MR 94c:11046

[15] Teresa Crespo, Explicit construction of $\tilde{A}_n$ type fields, *J. Algebra*, **127** (1989) no. 2, 452–461. MR 91a:12006

[16] P. Deligne, *Les constantes des équations fonctionnelles des fonctions L*, Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), Springer, Berlin, 1973, pp. 501–597. Lecture Notes in Math., vol. 349. MR 50 #2128

[17] Pierre Deligne and Jean-Pierre Serre, Formes modulaires de poids 1, *Ann. Sci. École Norm. Sup. (4)* **7** (1974) 507–530 (1975). MR 52 #284

[18] Tim Dokchitser, Computing special values of motivic L-functions, *Experiment. Math.*, **13** (2004) no. 2, 137–149. MR MR2068888 (2005f:11128)

[19] Jordan S. Ellenberg and Chris Skinner, On the modularity of $\mathbb{Q}$-curves, *Duke Math. J.*, **109** (2001) no. 1, 97–122.

[20] Michael Harris, David Soudry and Richard Taylor, l-adic representations associated to modular forms over imaginary quadratic fields. I. Lifting to $\mathrm{GSp}_4(\mathbf{Q})$, *Invent. Math.*, **112** (1993) no. 2, 377–411. MR 94d:11035

[21] Erich Hecke, *Mathematische Werke*, third ed., Vandenhoeck & Ruprecht, Göttingen, 1983, With introductory material by B. Schoeneberg, C. L. Siegel and J. Nielsen. MR 86a:01049

[22] H. Iwaniec and P. Sarnak, *Perspectives on the analytic theory of L-functions*, Geom. Funct. Anal. (2000), no. Special Volume, Part II, 705–741, GAFA 2000 (Tel Aviv, 1999). MR 2002b:11117

[23] H. Jacquet, I. Piatetski-Shapiro and J. Shalika, Automorphic forms on GL(3). I, *Ann. of Math. (2)* **109** (1979) no. 1, 169–212. MR 80i:10034a

[24] H. Jacquet, Automorphic forms on GL(3). II, *Ann. of Math. (2)*, **109** (1979) no. 2, 213–258. MR 80i:10034b

[25] Hervé Jacquet and Joseph Shalika, A lemma on highly ramified ϵ-factors, *Math. Ann.*, **271** (1985) no. 3, 319–332. MR 87i:22048

[26] A. Jehanne, Realization over $\mathbb{Q}$ of the groups $\tilde{A}_5$ and $\hat{A}_5$, *J. Number Theory*, **89** (2001) no. 2, 340–368. MR 2002f:12005

[27] Jean-Pierre Kahane, *Some random series of functions*, second ed., Cambridge Studies in Advanced Mathematics, vol. 5, Cambridge University Press, Cambridge, 1985. MR 87m:60119

[28] N. Katz, *Sommes exponentielles*, Société Mathématique de France, Paris, 1980, Course taught at the University of Paris, Orsay, Fall 1979. MR 82m:10059

[29] Nicholas M. Katz and Peter Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, American Mathematical Society Colloquium Publications, vol. 45, American Mathematical Society, Providence, RI, 1999. MR 2000b:11070

[30] Henry H. Kim and Freydoon Shahidi, Cuspidality of symmetric powers with applications, *Duke Math. J.*, **112** (2002) no. 1, 177–197. MR 2003a:11057

[31] A. Knapp, *Elliptic curves*, Princeton University Press, Princeton, NJ, 1992. MR 93j:11032

[32] R. P. Langlands, *Problems in the theory of automorphic forms*, Lectures in modern analysis and applications. III, Edited by C. T. Taam. Lecture Notes in Mathematics, vol. 170, Springer-Verlag, Berlin, 1970, pp. 18–61. MR 42 #4349

[33] Michael J. Razar, Modular forms for $G_0(N)$ and Dirichlet series, *Trans. Amer. Math. Soc.*, **231** (1977) no. 2, 489–495. MR MR0444576 (56 #2926)

[34] Michael Rubinstein, *Computational methods and experiments in analytic number theory*, preprint (2005).

[35] Atle Selberg, *Old and new conjectures and results about a class of Dirichlet series*, Proceedings of the Amalfi Conference on Analytic Number Theory (Maiori, 1989) (Salerno), Univ. Salerno, 1992, pp. 367–385. MR 94f:11085

[36] J.-P. Serre, *Facteurs locaux des fonctions zêta des variétés algébriques*, Séminaire Delange-Pisot-Poitou, 11e année (1969/70), Théorie des nombres, Fasc. 1, Exp. No. 19, Secrétariat Mathématique, Paris, 1970, p. 15.

[37] J.-P. Serre, *A course in arithmetic*, Springer-Verlag, New York, 1973, Translated from the French, Graduate Texts in Mathematics, No. 7. MR 49 #8956

[38] J.-P. Serre, *Abelian l-adic representations and elliptic curves*, A K Peters Ltd., Wellesley, MA, 1998, With the collaboration of Willem Kuyk and John Labute, Revised reprint of the 1968 original. MR 98g:11066

[39] Freydoon Shahidi, On nonvanishing of L-functions, *Bull. Amer. Math. Soc. (N.S.)*, **2** (1980) no. 3, 462–464. MR 81c:12020

[40] Daniel Shanks, *Class number, a theory of factorization, and genera*, 1969 Number Theory Institute (Proc. Sympos. Pure Math., Vol. XX, State Univ. New York, Stony Brook, N.Y., 1969), Amer. Math. Soc., Providence, R.I., 1971, pp. 415–440. MR 47 #4932

[41] Goro Shimura, On the zeta-functions of the algebraic curves uniformized by certain automorphic functions, *J. Math. Soc. Japan* **13** (1961) 275–331. MR 26 #84

[42] Joseph H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 151, Springer-Verlag, New York, 1994. MR 96b:11074

[43] Joseph H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 106, Springer-Verlag, New York, 1994, Corrected reprint of the 1986 original. MR 95m:11054

[44] C. M. Skinner and A. J. Wiles, Residually reducible representations and modular forms, *Inst. Hautes Études Sci. Publ. Math.*, (1999) no. 89, 5–126 (2000). MR 2002b:11072

[45] C. M. Skinner and Andrew J. Wiles, Nearly ordinary deformations of irreducible residual representations, *Ann. Fac. Sci. Toulouse Math. (6)*, **10** (2001), no. 1, 185–215. MR 1 928 993

[46] Richard Taylor, l-adic representations associated to modular forms over imaginary quadratic fields. II, *Invent. Math.*, **116** (1994) no. 1–3, 619–643. MR 95h:11050a

[47] Richard Taylor and Andrew Wiles, Ring-theoretic properties of certain Hecke algebras, *Ann. of Math. (2)*, **141** (1995) no. 3, 553–572. MR 96d:11072

[48] A. Weil, Über die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen, *Math. Ann.*, **168** (1967) 149–156. MR 34 #7473

[49] André Weil, Jacobi sums as "Grössencharaktere", *Trans. Amer. Math. Soc.*, **73** (1952) 487–495. MR 14,452d

[50] Eric W. Weisstein, *Crc concise encyclopedia of mathematics*, CRC Press, Boca Raton, Florida, 1999.

[51] Hermann Weyl, *The classical groups*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997, Their invariants and representations, Fifteenth printing, Princeton Paperbacks. MR 98k:01049

[52] E. T. Whittaker and G. N. Watson, *A course of modern analysis*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1996, An introduction to the general theory of infinite processes and of analytic functions; with an account of the principal transcendental functions, Reprint of the fourth (1927) edition. MR 97k:01072

[53] Andrew Wiles, Modular elliptic curves and Fermat's last theorem, *Ann. of Math. (2)* **141** (1995) no. 3, 443–551. MR 96d:11071