MODELLING λ-INVARIANTS BY p-ADIC RANDOM MATRICES.

JORDAN S. ELLENBERG, SONAL JAIN, AKSHAY VENKATESH

Abstract. How often is the p-adic λ-invariant of an imaginary quadratic field equal to m? We model this by the statistics of random p-adic matrices, and test these predictions numerically.

Contents

1. Introduction

1.1. Let p be an odd prime number and χ a quadratic Dirichlet character. The function

$$s \mapsto L^{(p)}(s,\chi) := (1-\chi(p)p^{-s})L(s,\chi) \tag{1}$$

takes values in $\mathbf{Q}^\times$ for every negative integer s. Moreover, it is p-adically continuous on the set $\{s \in \mathbb{Z} : s < 0,\ s \equiv k_0 \bmod p-1\}$ for any $k_0 \in \mathbf{Z}$; by interpolation, one constructs the p-adic L-function $L_p(s,\chi)$ attached to χ. We recall the details in §2.

Our goal is to model features of this p-adic L-function by means of random p-adic matrices, in parallel with the well-known story for complex analytic L-functions [8, 9]. If one applies this idea to model the largest power of p dividing $L(0,\chi)$, one arrives at (a special case of) the Cohen-Lenstra heuristics for the p-part of quadratic class groups. In this paper we focus on the somewhat subtler (but in some sense more natural) λ-invariant; as we explain in §2.4, this is somewhat analogous to certain questions about "low-lying zeroes" studied in the literature on complex analytic L-functions.

1.2. The λ-invariant is an invariant of an imaginary quadratic field K which measures the number of zeroes of an associated p-adic L-function: the function that interpolates values of $L^{(p)}(s, \chi_K)$ at integers $s \equiv 0$ modulo $p-1$. Equivalently, the λ-invariant measures the growth of class numbers in cyclotomic towers over K; see §2.2 for details. In particular, if p is inert in K, then $\lambda(K) > 0$ if and only if p divides the class number of K.

For example, the 3-adic λ-invariant attached to the quadratic field $\mathbf{Q}(\sqrt{-41})$ of discriminant -164 equals 3. This is related to various unusual 3-divisibilities, such as:

$$L(-2, \chi_{-41}) = -8 \cdot 3^6, \quad \mathrm{Cl}(\mathbf{Q}(\sqrt{-41}, \zeta_9)) \cong \mathbf{Z}/72\mathbf{Z} \oplus \mathbf{Z}/3\mathbf{Z},$$

Unlike the p-part of ideal class groups, the Iwasawa invariant apparently takes rather large values even for discriminants of moderate size; e.g., the 3-adic λ-invariant of $\mathbf{Q}(\sqrt{-53301})$ is 11.[1] The largest λ-invariant we found in our computations is 14, which is the 3-adic λ-invariant of $\mathbf{Q}(\sqrt{-956238})$. Computations with p-adic random matrices that we carry out in §4 suggest the following conjecture:

1.3. Conjecture. *Amongst imaginary quadratic fields K in which p does not split, the proportion with $\lambda(K) = r$ is*

$$p^{-r} \prod_{t>r} (1 - p^{-t}).$$

Here the $\lambda(K)$ denotes the λ-invariant of the cyclotomic $\mathbf{Z}_p$-extension.

In particular, one expects the λ-invariant to be unbounded as χ varies over quadratic characters.

Contrast Conjecture 1.3 with the analogous Cohen-Lenstra conjectures on p-parts of class groups: the probability that the class group of $\mathbf{Q}(\sqrt{-d})$ contains a subgroup isomorphic to $(\mathbf{Z}/p\mathbf{Z})^r$ is expected to be of order p^{-r^2}. The large quantitative difference between the two conjectures "explains" the fact that large λ-invariants are easier to find than class groups with large p-rank.

A related question: *can one give a reasonable upper bound for λ in terms of* $\mathrm{disc}(K)$? A logarithmic bound $\lambda \leq C \cdot \log(\mathrm{disc}(K))$ seems likely; note that this is false for $p = 2$, as shown by Ferrero [4] and Kida [10], but this appears to be a phenomenon special to the fact that complex conjugation has order 2.

Kraft and Washington formulate in [11] a conjecture along the lines of ours in the cases $r = 0, 1$ and $p = 3$; it agrees with Conjecture 1.3, and indeed one of our motivations was to generalize the work of [11].

[1] In this case $\mathrm{Cl}(\mathbf{Q}(\sqrt{-53301}, \zeta_9) \otimes \mathbf{Z}_3 \cong (\mathbf{Z}/3\mathbf{Z})^4$; it is striking that not merely the 3-part but also the rank over $\mathbb{F}_3$ is large.

1.4. The following table records the results of computing the 3-adic λ-invariants of the first 10^6 admissible imaginary quadratic fields, i.e., those in which 3 does not split:

λ	0	1	2	3	4	5	6
observed	0.589887	0.2680	0.0936	0.0322	0.0109	0.0035	0.0012
predicted	0.5601	0.2800	0.1050	0.0363	0.0122	0.0041	0.0013

From this one may see that the conjecture consistently *overestimates* the fraction of K with large λ-invariants. Further data for short ranges at higher discriminant indeed show the observed λ-probabilities moving towards the predicted ones. Similar observations have been made when testing the Cohen-Lenstra heuristics – in words, there are *unusually few quadratic fields with p dividing the order of the class group* compared to the number predicted by the heuristics, even in cases where the heuristics have been proven correct. In the Cohen-Lenstra case with $p = 3$, this discrepancy between asymptotic and observed behavior is explained by a secondary main term which is negative [1, 13, 17]. It is interesting to wonder whether the same is the case in the present context.

More ambitiously, one can speculate that the distribution of more refined features of the p-adic L-function can also be predicted by random matrix models; we discuss one example, concerning the spacing of zeroes. Here the numerical evidence is less compelling but not wholly discouraging (see §3.2).

1.5. Computation. Our method of computation differs from [11]: we interpolate the p-adic L-function from values of the classical L-function. The primary advantage of this method arises at very large discriminant d: one may compute the classical L-function by transcendental methods in time $O(d^{1/2+\varepsilon})$, whereas using exact arithmetic it appears to require time $O(d)$.

Note that one can compute by integer arithmetic a *block of values* of $L(-k, \chi_d)$, with $d \in [X, 2X]$, in time $O(X^{1+\epsilon})$. Although this would presumably speed up some of our computations, we did not try to carry it out.

It seems like a very interesting question to give a genuinely p-adic algorithm for computing the p-adic L-function that has performance comparable to the methods of rapidly computing classical L-functions.

1.6. Acknowledgements. We are very grateful to Brian Conrey and Larry Washington for their helpful comments on a first draft of this paper.

We thank the sponsoring agencies that have supported our work: A.V. was supported by a Packard Foundation fellowship, a Sloan fellowship and NSF grant DMS-0903110.

2. Review of p-adic L-functions.

We review the definitions of the p-adic L-function attached to a Dirichlet character. We shall always suppose $p > 2$.

We fix an algebraic closure $\overline{\mathbf{Q}}_p$ of $\mathbf{Q}_p$ and denote by $\overline{\mathbf{Z}}_p$ the integral closure of $\mathbf{Z}_p$ in $\overline{\mathbf{Q}}_p$.

Let

$$\chi : (\mathbf{Z}/f\mathbf{Z})^\times \to \overline{\mathbf{Q}}_p^\times, \quad \omega : (\mathbf{Z}/p\mathbf{Z})^\times \to \mathbf{Q}_p^\times,$$

be, respectively, a Dirichlet character of primitive conductor f and the Teichmüller character. (Thus ω is the unique character so that $\omega(i) \equiv i$ modulo p for each i relatively prime to p; the reader may wish to look now at §2.3 to get a sense of the role ω will play).

Although χ is not valued in $\mathbf{C}$, one can nonetheless define a value $L(-k, \chi) \in \overline{\mathbf{Q}}_p$ for the Dirichlet L-function attached to χ, evaluated at a negative integer. There is a unique $L \in \overline{\mathbf{Q}}_p$ so that $\iota(L) = L(-k, \iota \circ \chi)$ for every field embedding $\iota : \overline{\mathbf{Q}}_p \hookrightarrow \mathbf{C}$, and we define accordingly $L(-k, \chi) := L$.

2.1. Proposition. (Iwasawa, [18, Theorem 7.10]) *There is a unique $f_\chi \in \overline{\mathbf{Z}}_p[[T]]$ so that*

$$f_\chi((1+p)^{-s} - 1) = L^{(p)}(s, \chi\omega^s) \tag{2}$$

for all negative integers s.

Accordingly, we define the p-adic L-function attached to χ as:

$$L_p(s, \chi) := f_\chi((1+p)^{-s} - 1), \ (s \in \overline{\mathbf{Q}}_p, |s| \leq 1)$$

which interpolates the values of the usual L-function at the points $s \equiv 0 \bmod p-1$.

Warning. Usually, the definition of the p-adic L-function associated to χ is $s \mapsto L^{(p)}(s, \chi\omega^{s-1})$, and f_χ is defined accordingly. This does not change the mathematics: the two definitions are mapped to one another under the reparameterization $\chi \leftarrow \chi\omega$. We have chosen the present normalization because it matches particularly conveniently with the function field case. We apologize for any confusion this causes.

2.2. Distinguished polynomials. A *distinguished polynomial* $P \in \overline{\mathbf{Z}}_p[T]$ is one for which $P \equiv T^{\deg\ P}$ modulo the maximal ideal of $\overline{\mathbf{Z}}_p$. For example, $3 + 9T + 2T^2 \in \mathbf{Z}_3[T]$ is distinguished; or, if $N \in \mathrm{GL}_n(\mathbf{Z}_p)$ has the property that its reduction $\bar{N} \in \mathrm{GL}_n(\mathbf{F}_p)$ is nilpotent, then the characteristic polynomial $\det(T \cdot \mathrm{id} - N)$ is distinguished. The latter example will play an important role in §4.

The importance of this notion stems, in part, from the p-adic Weierstrass preparation theorem [18, Theorem 7.3]: if $f \in \overline{\mathbf{Z}}_p[[T]]$ is a power series, there is a unique expression

$$f = p^\mu \cdot U \cdot P,$$

where $U \in \overline{\mathbf{Z}}_p[[T]]$ is a unit, and P is distinguished. It follows, in particular, that f has precisely $\deg(P)$ zeros in the disc $\{T \in \overline{\mathbf{Q}}_p : |T| < 1\}$, for U has no zeroes in this region, and it is easy to check that P has all its zeroes there.

A fundamental theorem of Ferrero and Washington [4] asserts that, if $f = f_\chi$ is as in Prop. 2.1, then $\mu = 0$, i.e. there is a unique factorization

$$f_\chi = P_\chi U_\chi, \quad U_\chi \in \mathbf{Z}_p[[T]]^\times, \quad P_\chi \text{ distinguished.}$$

We define $\lambda_\chi = \deg(P_\chi)$ to be the *λ-invariant* of χ, and refer to P_χ as the *Iwasawa polynomial* of χ. The λ-invariant is easy to extract directly from f_χ: if we write $f_\chi = \sum a_i T^i$, then $\lambda = \min\{i : a_i \text{ is a unit in } \overline{\mathbf{Z}}_p.\}$

The λ-invariant controls the growth of class numbers in cyclotomic $\mathbf{Z}_p$-towers. Suppose, for example, that χ is the quadratic character of discriminant f; let p^{e_n} be the size of the Sylow subgroup of the relative class group of $\mathbf{Q}_n(\sqrt{f}, \zeta_{p^n})/\mathbf{Q}_n$, where $\mathbf{Q}_n$ denotes the cyclotomic $\mathbf{Z}/p^n\mathbf{Z}$-extension of $\mathbf{Q}$.

Then

$$e_n = \lambda n + \nu,$$

for n sufficiently large. Thus λ_χ is exactly the λ-invariant of the cyclotomic $\mathbf{Z}_p$-extension of the quadratic field attached to χ.

2.3. An example of conductor -31, with $p = 3$. Take χ to be the quadratic character of conductor -31. Let χ' be the quadratic character of conductor 93.

The values of $L(s, \chi)$ vanish at negative odd integers; the values of $L(s, \chi')$ vanish at negative even integers. Part of the content of Proposition 2.1 is that one may make a p-adic analytic function on the negative integers by *intertwining the values of* $L(s, \chi)$ *and* $L(s, \chi')$. This illustrates the role of twisting by powers of the Teichmüller character ω: for $p = 3$, ω is the quadratic character of conductor -3, and so $\chi' = \chi\omega$.

Then, for instance,

$$\begin{aligned} L_p(-4, \chi) &= (1 - \chi(3)3^4) \cdot L(-4, \chi) &&= 82 \times 25920 &&= 2 \cdot 3^4 + 3^5 + \dots \\ L_p(-5, \chi) &= (1 - \chi'(3)3^5) \cdot L(-5, \chi') &&= -257650476 &&= 3 + 2 \cdot 3^2 + \dots \end{aligned}$$

By computing more values of L_p and interpolating via Proposition 5.2, we can compute f_χ to any desired degree of precision. In this case, the Weierstrass factorization is

$$f_\chi = u \cdot (T - t_0),$$

where $u \in \mathbf{Z}_p[[T]]^\times$and $t_0 = -(2 \cdot 3 + 3^2 + 2 \cdot 3^3 + 3^4 + O(3^5))$. In particular, the λ-invariant $\lambda(\mathbf{Q}(\sqrt{-31})) = 1$; as commented earlier, the fact that the λ-invariant is nonzero is equivalent to the fact that the class number of $\mathbf{Q}(\sqrt{-31})$ is divisible by 3.

Thus f_χ has a zero at $T = t_0$; correspondingly, the L-function has a zero at the solution s_0 to

$$4^{-s_0} - 1 = t_0,$$

which we compute to be $2 + 3 + 2 \cdot 3^2 + 3^3 + 2 \cdot 3^4 + O(3^5)$. In particular $s_0 \equiv -4$ modulo 27 (which accounts for the high 3-divisibility of $L_p(-4, \chi)$

noted above), and $s_0 \equiv -31$ modulo 81; indeed one checks that $L_p(-31, \chi)$ is divisible by 3^6.

Remark. As observed by Wagstaff, not every zero of f_χ necessarily arises from a zero of $L_p(s, \chi)$, i.e., not every zero of f_χ lies in the image of the map $s \mapsto (1+p)^{-s} - 1$. The "extra zeroes" do appear as zeroes of p-adic L-functions if we pass from $\mathbf{Q}$ to a higher layer of the cyclotomic $\mathbf{Z}_p$-extension, as shown by Childress and Gold. In the present paper we shall take the zeros of f_χ as the objects of primary interest.

2.4. We now discuss the theory over function fields. This both motivates our choice of model in §4, and will clarify our remark that *zeros of the p-adic L-function are analogous to low-lying zeros of a complex-analytic L-function.*

Let X be a curve over a finite field $\mathbb{F}_\ell$, and let K be the field of rational functions on X. We assume that $\ell \neq p$, and, for simplicity, that $\ell^{p-1} - 1$ is not divisible by p^2; equivalently, the subgroup of $\mathbf{Z}_p^*$ generated by ℓ contains $1 + p\mathbf{Z}_p$. Let $\bar{K}$ be a separable closure of K.

Let χ be a nontrivial quadratic "Dirichlet character" of K, that is to say, a nontrivial one-dimensional representation of $\mathrm{Gal}(\bar{K}/K)$ with image ± 1. Associated to it is an L-function, defined as

$$L(s, \chi) = \prod_x (1 - \chi(\mathrm{F}_x) q_x^{-s})^{-1},$$

where the product is taken over closed points x of X that are unramified in the quadratic extension corresponding to χ, and q_x denotes the cardinality of the residue field at x. Note that, in this case, there is no distinction between L and $L^{(p)}$: there is no analogue of "removing the p-Euler factor" because K has no primes of characteristic p.

The L-function is a polynomial in ℓ^{-s}; it is explicitly given as

$$L(s, \chi) = P(\ell^{-s})$$

where $P(t) = \det(1 - tF|V) \in \mathbf{Q}_p[t]$ is the characteristic polynomial of the Frobenius F acting on a certain $\mathbf{Q}_p$-vector space V: if Y is the (projective model of the) double cover of X determined by χ, we may take V to be the quotient of $H^1(Y \times \overline{\mathbb{F}}_\ell, \mathbf{Q}_p)$ by its subspace $H^1(X \times \overline{\mathbb{F}}_\ell, \mathbf{Q}_p)$.

We shall proceed by closely following (2). Set

$$f_\chi(T) = P((1+T)^\beta),$$

where β is the unique element of $\mathbf{Z}_p$ such that $(1+p)^{(p-1)\beta} = \ell^{p-1}$; such a β exists by our hypothesis on ℓ. Then $f_\chi \in \overline{\mathbf{Z}}_p[[T]]$ and

$$f_\chi((1+p)^{-s} - 1) = \det(1 - \ell^{-s} F|V)$$

whenever $s \in \mathbf{Z}$ is divisible by $p - 1$; more generally, when $s \equiv k$ modulo $p - 1$,

$$\begin{aligned} f_\chi((1+p)^{-s} - 1) = &\quad \det(1 - \ell^{-s}\langle \ell \rangle^s F|V) \\ = &\quad L(s, \chi\omega^s) \end{aligned} \tag{3}$$

where $\langle \ell \rangle = \ell^{-1}(1+p)^{\beta}$ is the Teichmüller lift of ℓ modulo p and $\omega : \mathrm{Gal}(\bar{K}/K) \to \overline{\mathbf{Q}_p}^{\times}$ is the composition of the cyclotomic character valued in $(\mathbf{Z}/p\mathbf{Z})^{\times}$ with the Teichmüller lift.

Let Z be the subspace of V consisting of vectors v such that $(F-1)^n v \to 0$ as $n \to \infty$. Then Z is stable under the action of F. So we have a factorization

$$\det(1 - tF|V) = Q(t)R(t),$$

where $Q(t) := \det(1 - tF|_Z)$. One checks that the factor of f_χ corresponding to Q is a distinguished polynomial and the one corresponding to R is a unit; so the λ-invariant of χ equals $\dim Z$ – that is,

$$\lambda = \#\{\text{eigenvalues } \sigma \text{ of } F \text{ such that } |\sigma - 1| < 1\},$$

where the count is taken with multiplicity.

Note that any zero of $L(s, \chi)$ with $s \equiv 0$ corresponds to an eigenvalue ℓ^s of F on V that is congruent to 1; therefore, we may indeed regard λ as counting a p-adic analogue of the set of "low lying zeroes" of the L-function: it tracks zeros that are in a small p-adic neighbourhood of zero, rather than a small archimedean neighbourhood of the critical point.

Example. Take $\ell = 11, p = 7, K = \mathbf{F}_{11}(T)$; let χ be the character that corresponds to the quadratic extension $K(\sqrt{T^3 + T + 1})$. Then V is a 2-dimensional vector space, the 7-adic Tate module of the elliptic curve $y^2 = x^3 + x + 1$ in characteristic 11. Then $\det(1 - tF|V) = 1 + 2t + 11t^2 \in \mathbf{Q}_7[t]$. Exactly one of the two roots of this polynomial is congruent to 1 modulo 7. Consequently, the λ-invariant equals 1.

3. Data

3.1. Distribution of the Iwasawa invariant. We describe here our data testing Conjecture 1.3.

We have already given the data for $p = 3$.

λ	0	1	2	3	4	5	6
predicted	0.5601	0.2800	0.1050	0.0363	0.0122	0.0041	0.0013
observed – (a)	0.589887	0.2680	0.0936	0.0322	0.010932	0.003497	0.001237
observed – (b)	0.5783	0.2715	0.09874	0.03369	0.01161	0.003933	0.001284
observed – (c)	0.5769	0.2736	0.09802	0.03304	0.01137	0.0041	0.0015

Here are the origin of the three data sets.

(a) The value of λ for $\mathbf{Q}(\sqrt{-d})$ for the first 10^6 values of d for which the prime 3 does not split.
(b) The value of λ for $\sim 10^5$ fields $\mathbf{Q}(\sqrt{-d})$ so that $\left(\frac{-d}{3}\right) = -1$ taken from a range of intervals in $[10^7, 2 \times 10^7]$.
(c) The value of $\lambda - 1$ for $\sim 10^5$ fields $\mathbf{Q}(\sqrt{-d})$ within $[10^7, 2 \times 10^7]$, conditioned on the p-adic L-function having a trivial zero.

This in particular gives credence to the idea that the distribution of $\lambda - 1$ among L-functions with a trivial zero agrees, as one might expect, with the

distribution of λ among L-functions without a trivial zero. In fact, further numerical investigations are consistent with the more refined hypothesis that the distribution of f_χ/T, conditional on the existence of a trivial zero, agrees with the distribution of f_χ, conditional on the absence of a trivial zero.

For $p = 5$ and the first 10^6 values of d for which the prime 5 does not split in $\mathbb{Q}(\sqrt{-d})$, the corresponding table reads:

λ	0	1	2	3	4	5	6
predicted	0.7603	0.1901	0.03960	0.007984	0.001599	0.0003200	0.00006400
observed	0.7801	0.1755	0.03561	0.007103	0.001383	0.0003300	0.00001400

3.2. The discriminant of the Iwasawa polynomial. Another interesting invariant, beyond the λ-invariant, is the discriminant of P_χ: this measures the spacing between zeros of the p-adic L-function, and, as such, is analogous to the well-studied question of spacing between zeroes of *complex* L-functions. The random matrix heuristics detailed in the next section describe a measure on the space of distinguished polynomials, not only on their degrees, and consequently allow us to make predictions about the statistics of this discriminant. See §4.7.

For the first 7500 admissible discriminants with $\lambda = 2$, we have:

$\mathrm{ord}_3\mathrm{disc}(P_\chi)$	1	2	3	4	5	6	7
predicted	0.5926	0.2716	0.0878	0.0320	0.0106	0.0036	0.001204
observed	.65280	0.22867	0.083467	0.025733	0.008933	0.000267	0.000133

The observed data in this range is not clearly in line with the predictions, and we do not know whether slow convergence or the prediction itself is to blame.

4. Random p-adic matrices and their characteristic polynomials.

We now compute the statistics of random p-adic matrices which motivate the conjectures numerically investigated in the first part of the paper. Let us describe the main thrust of the section in words:

We choose matrices from $\mathrm{GL}_n(\mathbf{Z}_p)$ – or the related group $\mathrm{GSp}(\mathbf{Z}_p)$ – according to Haar measure. In this case, the Haar measure is simply the limit of the normalized counting measures on each $\mathrm{GL}_n(\mathbf{Z}/p^r\mathbf{Z})$, as $r \to \infty$.

The characteristic polynomial of $A \in \mathrm{GL}_n(\mathbf{Z}_p)$ of primary interest is not the statistic of primary interest in the present contrast; rather, we consider a "piece" of of the characteristic polynomial, defined as follows. We discard all roots of the characteristic polynomial that are not p-adically close to 1 (i.e., those roots α for which $\alpha - 1$ is a unit). The resulting polynomial is, after shifting $T \leftarrow T - 1$, distinguished in the sense of §2.2; it is denoted by P_A below.

We compute the statistics of P_A in Theorem 4.3. The proof goes as follows: partitioning the eigenvalues of A into those close to 1 and those far from

1 induces a corresponding direct sum decomposition $\mathbf{Z}_p^n = X \oplus Y$, where $A-1$ acts nilpotently on one summand and invertibly on the other; we then count the number of such decompositions, and, for each such decomposition, analyze the set of $A \in \mathrm{GL}_n(\mathbf{Z}_p)$ that give rise to it.

4.1. Let V be a free finite rank $\mathbf{Z}_p$-module, and $A \in \mathrm{GL}(V)$. We set:

$$Z = \{v \in V : (A-1)^N v \to 0\}, P_A = \det(((1+T)\cdot \mathrm{id} - A)\,|_Z) \ \in \mathbf{Z}_p[T].$$

We refer to P_A as the *associated polynomial* to A; it is distinguished in the sense of §2.2. Note that if A is an element of $\mathrm{GL}(V/p^mV)$, we may still define P_A as a polynomial modulo p^m.

We may equivalently describe Z as the sum of all generalized eigenspaces V_λ for eigenvalues satisfying $|\lambda - 1|_p < 1$; consequently, $P_A = \prod_\lambda (T+1-\lambda)$, where the product is taken (with multiplicity) over generalized eigenvalues of A satisfying $|\lambda - 1|_p < 1$. In words, P_A is that part of the characteristic polynomial corresponding to roots that are p-adically close to 1.

Our goal in this section is to investigate the statistics of P_A where A is chosen randomly from $\mathrm{GL}(V)$ or a coset of the symplectic group. (It will transpire that the statistics are the same.) The analogy between number fields and function fields (see §2.4) suggests that the resulting distribution on distinguished polynomials can be viewed as a model for the distribution of f_χ as χ ranges over quadratic imaginary Dirichlet characters.

4.2. For simplicity of notation we choose a basis for V and consider random matrices in the group $\mathrm{GL}(n, \mathbf{Z}_p)$ of $n \times n$ matrices with entries in $\mathbf{Z}_p$. Let ω be the standard nondegenerate symplectic form on $\mathbf{Z}_p^{2n}$:

$$\omega(x_i, y_i) = \sum_{1 \le i \le n} (x_i y_{i+n} - x_{i+n} y_i) \tag{4}$$

and define

$$\mathrm{GSp}_\alpha(2n, \mathbf{Z}_p) = \{g \in \mathrm{GL}_n(\mathbf{Z}_p) : \langle gx, gy\rangle = \alpha\langle x, y\rangle\}.$$

Then the generalized symplectic group $\mathrm{GSp}(2n)$ is the union $\bigcup_{\alpha \in \mathbf{Z}_p^*} \mathrm{GSp}_\alpha(2n, \mathbf{Z}_p)$, and each $\mathrm{GSp}_\alpha(2n, \mathbf{Z}_p)$ is a coset of the symplectic group in the generalized symplectic group.

We assume for the remainder of the paper that

$$\alpha \in \mathbf{Z}_p^* \text{ does not reduce to 1 mod } p.$$

Absent this assumption, the distribution of P_A as A ranges over $\mathrm{GSp}_\alpha(2n, \mathbf{Z}_p)$ will be different from that described here; see [7] for a discussion of a related phenomenon.

When we speak of a "random" element of a compact group like $GL(n, \mathbf{Z}_p)$, we shall always will mean an element that is random with respect to the Haar probability measure. A random element in $\mathrm{GSp}_\alpha(2n, \mathbf{Z}_p)$ is drawn from the unique probability measure invariant under $\mathrm{Sp}(2n, \mathbf{Z}_p)$. By abuse of notation we refer to both measures as μ.

We denote by ρ the infinite product

$$\rho := \prod_{i \geq 1} (1 - p^{-i})$$

This is the large n limit of the probability that a random $n \times n$ matrix over $\mathbf{F}_p$ is nonsingular. It plays a central role in the Cohen-Lenstra distribution on finite abelian p-groups: ρ is the probability that a random such group (drawn from the Cohen-Lenstra distribution) is trivial.

The main result of this section is the following.

4.3. Theorem. *Let S be an open subset of the set of distinguished polynomials of degree r. Denote by e_S the measure of the subset of $\mathrm{GL}_r(\mathbf{Z}_p)$ consisting of matrices A for which $P_A \in S$.*

For each n, let $P(n, S)$ be the set of elements in $\mathrm{GL}_n(\mathbf{Z}_p)$ such that P_A lies in S; similarly, let $Q(n, S)$ be the set of elements in $\mathrm{GSp}_\alpha(2n, \mathbf{Z}_p)$ such that P_A lies in S. Then

$$\lim_{n \to \infty} \mu(P(n, S)) = \lim_{n \to \infty} \mu(Q(n, S)) = \rho e_S.$$

For example, take $r = 0$; the set of degree r distinguished polynomials is a singleton $\{1\}$. Take $S = \{1\}$. The theorem asserts that the fractions $P(n, S)$ and $Q(n, S)$ both approach ρ as $n \to \infty$; this recovers the prediction that the proportion of χ with λ-invariant 0 is ρ (in agreement with the Cohen-Lenstra heuristics).

More generally we draw the following corollary, which supports Conjecture 1.3:

Corollary. *The fraction of elements $A \in \mathrm{GL}_n(\mathbf{Z}_p)$ or $A \in \mathrm{GSp}_\alpha(2n, \mathbf{Z}_p)$ with $\deg(P_A) = r$ approaches*

$$p^{-r} \prod_{t > r} (1 - p^{-t})$$

as $n \to \infty$.

Proof. We apply Theorem 4.3, taking S to be the set of *all* distinguished polynomials of degree r. The probability that $\deg P_A = r$ is then the measure of the set of $A \in \mathrm{GL}_r(\mathbf{Z}_p)$ for which the characteristic polynomial of $A - 1$ is distinguished. The characteristic polynomial of A is distinguished if and only if the reduction of A mod p is unipotent. [15] We now use the fact that the number of unipotent elements in $\mathrm{GL}_r(\mathbf{F}_p)$ equals $p^{r^2 - r}$ together with the fact that the cardinality of $\mathrm{GL}_r(\mathbf{F}_p)$ is $p^{r^2} \prod_{i \leq r} (1 - p^{-i})$ It follows that

$$e_S = p^{-r} \prod_{i \leq r} (1 - p^{-i})^{-1}$$

and the corollary follows immediately. □

We now prove Theorem 4.3. The first step is to verify the assertion of the theorem when $r = 0$, i.e. we need to compute the probability that $P_A \equiv 1$, or, equivalently, that $A - 1$ is invertible.

4.4. Proposition. *The probability that an element A of $\mathrm{GL}_n(\mathbf{Z}_p)$ (or of $\mathrm{GSp}_\alpha(2n, \mathbf{Z}_p)$) has $P_A \equiv 1$ approaches ρ as $n \to \infty$.*

The first case, where A is drawn randomly from $\mathrm{GL}_n(\mathbf{Z}_p)$, was proved by Friedman and Washington in [6]. We present a uniform proof for both cases.

Proof. For the proof of the proposition, we replace α by its reduction modulo p, so α is an element of $\mathbf{F}_p^*$ not equal to 1.

For $1 \leq i \leq 3$ we define a random variable X_i as follows:

$i = 1$: the number of fixed vectors of a random matrix $A \in \mathrm{GL}_n(\mathbf{F}_p)$;
$i = 2$: the number of fixed vectors of a random matrix $A \in \mathrm{GSp}_\alpha(2n, \mathbf{F}_p)$.
$i = 3$: the number of vectors annihilated by a random $n \times n$ matrix with $\mathbf{F}_p$-entries;

In each case, "random" refers to counting measure; in each case, a "vector" is an element of $\mathbf{F}_p^n$.

We claim that, for all $t < n/2$, and for $j = 1, 2$, we have

$$\langle (X_j - 1)(X_j - p) \dots (X_j - p^{t-1}) \rangle = 1 \tag{5}$$

whereas for $j = 3$,

$$\langle (X_j - 1)(X_j - p) \dots (X_j - p^{t-1}) \rangle = (1 - p^{-n})(1 - p^{1-n}) \dots (1 - p^{t-n-1}). \tag{6}$$

Here $\langle \dots \rangle$ denotes "expected value." Note that, as $n \to \infty$, the right-hand side of (6) approaches 1; thus the results of all three computations agree as $n \to \infty$.

For $j = 1$, we have by Burnside's lemma [14, Theorem 3.22] that this expected value is the number of orbits of $\mathrm{GL}(\mathbf{F}_p^n)$ on injective maps $\mathbf{F}_p^t \hookrightarrow \mathbf{F}_p^n$, which is 1. For $j = 2$, one needs the corresponding statement[2] for orbits of $\mathrm{Sp}(\mathbf{F}_p^n)$ on injective maps $\mathbf{F}_p^t \hookrightarrow \mathbf{F}_p^n$. See [2, Lemma 8.8] for a proof of (5) in this case. For $j = 3$, the left hand side counts the expected number of such injections into the kernel of a random $n \times n$ matrix; this equals the number of injections $\mathbf{F}_p^t \hookrightarrow \mathbf{F}_p^n$, multiplied by p^{-tn}; that verifies (6).

The distribution of each X_j defines a probability measure on the non-negative integers $\mathbb{N}$ (indeed, on the set $\{1, p, p^2, \dots\} \subset \mathbb{N}$). For each j, let μ_j be any weak limit as $n \to \infty$ of this probability measure. Then μ_3 is easy to compute; in particular, since the fraction of $n \times n$ matrices which are *invertible* approaches ρ as $n \to \infty$, we see $\mu_3(\{0\}) = \rho$. It suffices, then, to check that $\mu_1 = \mu_2 = \mu_3$.

We claim that μ_i have the same moments for $1 \leq i \leq 3$, that is to say, for each k, the expected value $\langle X_i^k \rangle_{i \in \mathbb{N}}$ is the same whether we draw natural numbers from μ_1, μ_2 or μ_3. Indeed, this follows from (5) and (6), together with elementary convergence estimates to carry out switches of limits and summation; for instance, (5) and (6) imply that the probability that $X_i \geq p^k$ is $\leq Cp^{-k^2}$ when $i > 2k$, which is enough.

[2]We note that $\alpha \neq 1$ is used in a crucial way here; without it, the statement of the theorem would change.

It is not general true that a probability distribution is determined by its moments; but it has been checked by Fouvry and Klüners [5, Theorem 1,2] that it is so in the case at hand; the reader may also consult [2, §8.1, Lemma] for a different proof of a closely related result.

Thus $\mu_1 = \mu_2 = \mu_3$, as desired. □

4.5. We now prove the first part of Theorem 4.3, concerning GL_n. We may suppose that S is the preimage of a finite set $\bar{S}$ of degree-r polynomials mod p^m, for some $m \geq 1$. Therefore, e_S is precisely the proportion of $C \in \mathrm{GL}_r(\mathbf{Z}/p^m\mathbf{Z})$ which have $P_C \in \bar{S}$.

More precisely: let $\bar{V} = V/p^m V$. Let $\mathcal{Q}$ be the set of quadruples

$$(X \subset \bar{V}, Y \subset \bar{V}, A_X \in \mathrm{GL}(X), A_Y \in \mathrm{GL}(Y))$$

where X, Y are subgroups so that $X \oplus Y = V$, $A_Y - 1$ is invertible, and $A_X - 1$ nilpotent. Then we claim that the natural map

$$\mathcal{Q} \longrightarrow \mathrm{GL}(\bar{V}), \tag{7}$$

sending (X, Y, A_X, A_Y) to $A_X \oplus A_Y \in \mathrm{GL}(\bar{V})$, is a *bijection*.

Start with a matrix A in $\mathrm{GL}(\bar{V})$. The sequence $\ker(A-1) \subset \ker(A-1)^2 \subset \ldots$ must stabilize; let N be so large that $\ker(A-1)^N = \ker(A-1)^{N+1} = \ldots$. Let $X = \ker(A-1)^N$ and let $Y = \mathrm{image}(A-1)^N$. Then $A-1$ is nilpotent (resp. invertible) on X (resp. Y). It follows that $X \cap Y = \{0\}$, and clearly $|X| \cdot |Y| = |\bar{V}|$. So $\bar{V} = X \oplus Y$. The bijection is now clear.

Since X, Y are free $\mathbf{Z}/p^m\mathbf{Z}$-modules, we may speak of $P_{A|X}$. The bijection (7) implies immediately that the fraction of $A \in \mathrm{GL}(\bar{V})$ with $P_A \in \bar{S}$ is

$$\frac{\#\{A_Y \in \mathrm{GL}_{n-r}(\mathbf{Z}/p^m\mathbf{Z}) : A_Y - 1 \text{ invertible}\}}{\#\mathrm{GL}_{n-r}(\mathbf{Z}/p^m\mathbf{Z})} \cdot e_S,$$

using the fact that there are

$$\frac{\#\mathrm{GL}_n(\mathbf{Z}/p^m\mathbf{Z})}{\#\mathrm{GL}_r(\mathbf{Z}/p^m\mathbf{Z})\#\mathrm{GL}_{n-r}(\mathbf{Z}/p^m\mathbf{Z})}$$

splittings $\bar{V} = X \oplus Y$. Now, letting $n \to \infty$, the theorem follows from Proposition 4.4.

4.6. The proof of Theorem 4.3 for GSp_α proceeds along similar lines. We take $\bar{V} = (\mathbf{Z}/p^m\mathbf{Z})^{2g}$, endowed with the symplectic form (4) – by "symplectic form", we mean an alternating pairing $\bar{V} \times \bar{V} \to (\mathbf{Z}/p^m\mathbf{Z})$ with the property that the induced map $\bar{V} \longrightarrow \mathrm{Hom}(\bar{V}, \mathbf{Z}/p^m\mathbf{Z})$ is an isomorphism. We take $\mathcal{Q}$ to be the set of quadruples

$$(X = X \subset \bar{V}, Y \subset \bar{V}, A_1 \in \mathrm{GSp}_\alpha(X), A_Y \in \mathrm{GSp}_\alpha(Y))$$

so that $(A-1)(A-\alpha)$ is nilpotent on X, $(A-1)(A-\alpha)$ is invertible on Y, and so that $X \oplus Y$ is an orthogonal direct sum with respect to the symplectic form. As before, the claim is that the natural map

$$\mathcal{Q} \longrightarrow \mathrm{GSp}_\alpha(\bar{V})$$

sending (X, Y, A_X, A_Y) to $A_X \oplus A_Y$ is a bijection.

This is proved, as before, by constructing an inverse $\mathrm{GSp}_\alpha \to \mathcal{Q}$; this construction is similar, replacing $(A-1)$ by $B := (A-1)(A-\alpha)$. The argument that $\bar{V}$ splits canonically into a direct product $X \oplus Y$ proceeds exactly analogously. What remains is to show that X and Y are orthogonal with respect to the symplectic form, i.e. $\omega(x,y) = 0$ for all $x \in X, y \in Y$. Take $x \in X, y \in Y$; we can write $y = Bz$, and then easy computation yields

$$\omega(x,y) = \alpha\omega(A^{-2}Bx, z).$$

It follows that for each N we can write

$$\omega(x,y) = \alpha^N \omega(A^{-2N}B^N x, z_N)$$

for some $z_N \in Y$, and taking N large enough makes the right-hand side zero.

It is still true that $P_A \in \bar{S} \iff P_{A|X} \in \bar{S}$. Now $\deg(P_A) = r$ if and only if the rank of X over $\mathbf{Z}/p^m\mathbf{Z}$ is $2r$. Indeed, the spaces

$$X_1 = \bigcup_N \ker\,(A-1)^N, \quad X_2 = \bigcup_N \ker\,(A-\alpha)^N \tag{8}$$

satisfy $X_1 \cap X_2 = 0$ because $A - \alpha$ is invertible on X_1 and $A-1$ is invertible on X_2, and α is not congruent to 1 mod p. So, X breaks up as a (non-orthogonal) direct sum of X_1 and X_2, each of which is a free $\mathbf{Z}/p^m\mathbf{Z}$-module. Moreover, X_1 and X_2 are both isotropic; one argues iteratively as above, using the fact that

$$\omega(x_1, (A-1)x_2) = \omega((\alpha A^{-1} - 1)x_1, x_2).$$

Therefore, the map $X_i \to \mathrm{Hom}(X_j, \mathbf{Z}/p^m\mathbf{Z})$ given by ω is an isomorphism for $\{i,j\} = \{1,2\}$; so the two are in fact free $\mathbf{Z}/p^m\mathbf{Z}$-modules of the same rank, as claimed.

We shall prove in a moment that

Lemma. *The fraction of $A \in \mathrm{GSp}_\alpha(2r, \mathbf{Z}/p^m\mathbf{Z})$ for which $(A-1)(A-\alpha)$ is nilpotent and P_A lies in $\bar{S}$ equals e_S.*

Assume the Lemma for now. The fraction of $A \in \mathrm{GSp}_\alpha(\mathbf{Z}/p^m\mathbf{Z})$ with $P_A \in \bar{S}$ is thus given by:

$$\frac{\#\{A \in \mathrm{GSp}_\alpha(n-2r, \mathbf{Z}/p^m\mathbf{Z}) : A-1 \text{ invertible}\}}{\#\mathrm{Sp}(n-2r, \mathbf{Z}/p^m\mathbf{Z}) \cdot \#\mathrm{Sp}(2r, \mathbf{Z}/p^m\mathbf{Z})} \cdot (e_S.\#\mathrm{Sp}(2r, \mathbf{Z}/p^m\mathbf{Z})),$$

where we used the fact that the number of orthogonal splittings $(\mathbf{Z}/p^m\mathbf{Z})^{2n} = X \oplus Y$ equals $\frac{\mathrm{Sp}(n, \mathbf{Z}/p^m\mathbf{Z})}{\mathrm{Sp}(n-2r, \mathbf{Z}/p^m\mathbf{Z})\mathrm{Sp}(2r, \mathbf{Z}/p^m\mathbf{Z})}$, because $\mathrm{Sp}(n, \mathbf{Z}/p^m\mathbf{Z})$ acts transitively on such splittings. Taking into account Prop. 4.4, we are again done. □

Proof. (of the Lemma)

We write $X = (\mathbf{Z}/p^m\mathbf{Z})^{2r}$, and $\mathrm{GSp}_\alpha(X)$ instead of $\mathrm{GSp}_\alpha(2r, \mathbf{Z}/p^m\mathbf{Z})$.

The desired set can be written as a *disjoint union* over isotropic subgroups $X_1 \subset X$ that are isomorphic as groups to $(\mathbf{Z}/p^m\mathbf{Z})^r$:

$$\bigcup_{X_1} \{A \in \mathrm{GSp}_\alpha(X) \text{ such that } A \text{ preserves } X_1 \text{ and } P_{A|X_1} \in S.\}. \tag{9}$$

Indeed: if $(A-1)(A-\alpha)$ is nilpotent, then A determines, as in (8), a subgroup $X_1 \cong (\mathbf{Z}/p^m\mathbf{Z})^r \subset X = (\mathbf{Z}/p^m\mathbf{Z})^{2r}$ on which A acts unipotently, and P_A is the characteristic polynomial of $A|X_1$.

In the reverse direction, every element of the set defined by (9) has $(A-1)(A-\alpha)$ nilpotent: the assumption that $P_{A|X_1} \in S$ entails, in particular, that $A-1$ acts nilpotently on X_1; then $A-\alpha$ acts nilpotently on X/X_1 by duality.

The action of $\mathrm{GSp}(X)$ on isotropic subgroups isomorphic to $(\mathbf{Z}/p^m\mathbf{Z})^r$ is transitive – see for instance [12, Cor 2.8].[3] Thus, if we specify one such subgroup X_1, the size of the union is $\mathrm{GSp}(X)/\mathrm{stab}(X_1)$, where $\mathrm{stab}(X_1)$ denotes the stabilizer of X_1.

On the other hand, the action on X_1 and on the symplectic form yields a homomorphism

$$\mathrm{stab}(X_1) \to \mathrm{GL}(X_1) \times \mathbf{F}_p^\times$$

which is seen to be a *surjection*. Consequently, the set

$$\{A \in \mathrm{GSp}_\alpha(X) \text{ such that } A \text{ preserves } X_1 \text{ and } P_{A|X_1} \in S.\}$$

described in (9) has size $\frac{e_S}{p-1} \cdot |\mathrm{stab}(X_1)|$, because it is the preimage under a group homomorphism of a set of relative measure $\frac{e_S}{p-1}$.

Now $\frac{|\mathrm{GSp}(X)|}{|\mathrm{stab}(X_1)|} \cdot (\frac{e_S}{p-1}\mathrm{stab}(X_1)) = e_S \frac{|\mathrm{GSp}(X)|}{p-1} = e_S |\mathrm{GSp}_\alpha(X)|$, completing the proof. □

4.7. Distribution of discriminants.

Corollary. *The measure of* $P \in \mathrm{GL}_n(\mathbf{Z}_p)$ *or* $\mathrm{GSp}_\alpha(2n, \mathbf{Z}_p)$ *for which* $\deg(P_A) = 2$ *and* $|\mathrm{disc}(P_A)|_p = p^{-s}$ *approaches, as* $n \to \infty$,

$$\rho \frac{p^{-s-1}}{1-p^{-1}} \left(1 - \begin{cases} p^{-(s+1)/2}, & s \text{ odd.} \\ \frac{p^{-s/2-1}}{1+p^{-1}}, & s \text{ even.} \end{cases}\right). \tag{10}$$

We leave the proof to the reader.

[3] The special case needed here follows by a standard argument, which we include for self-containedness. Let X_1 be such a subgroup; choose a $\mathbf{Z}/p^m\mathbf{Z}$-basis $x_1, \dots, x_r$. Since ω gives a surjection $X \to \mathrm{Hom}(X, \mathbf{Z}/p^m\mathbf{Z})$, it follows that there exists $y_1, \dots, y_r$ so that $\omega(x_i, y_j) = \delta_{ij} \in \mathbf{Z}/p^m\mathbf{Z}$. Now we may choose y_i so that $\omega(y_i, y_j) = 0$: replace y_1 by $y_1 - \omega(y_1, y_1)x_1/2$, and so on. Now $y_1, \dots, y_r$ must span a free $\mathbf{Z}/p^m\mathbf{Z}$-submodule of X/X_1, and so $x_1, \dots, x_r, y_1, \dots, y_r$ span freely X. Given any other subgroup X_1', we may choose a similar basis $x_1', \dots, x_r', y_1', \dots, y_r'$, and then there is an element of $\mathrm{Sp}(X)$ sending x_i to x_i' and y_i to y_i'.

For instance, for $s = 1$, this measure equals $\frac{\rho}{p^2}$; on the other hand, the measure of A with $\deg(P_A) = 2$ equals $\rho\frac{p^{-2}}{(1-p^{-1})(1-p^{-2})}$. In particular, the probability that $\mathrm{disc}(P_A)$ has valuation 1 *given that* $\deg(P_A) = 2$ is $(1-p^{-1})(1-p^{-2})$. For $p = 3$ this yields the prediction of $16/27 \sim .5926$ appearing in section 3.2.

5. Computation: algorithms and implementation

All our computations were carried out by a `PARI/GP` script [16], available on request. We describe the main points of the computation below. We note that our computations rely on several inbuilt functions of `PARI`: computation of ζ-values, class numbers and units for quadratic fields, in particular.

We compute λ-invariants by, firstly, computing special values of L-functions using transcendental methods (§5.1) and, secondly, interpolating the Iwasawa power series (§5.2). From the power series we extract both the lambda invariant λ-invariant and the Iwasawa polynomial (§5.3).

We have tried to independently verify the correctness of our codes in a few different ways, detailed in §5.4.

5.1. Computation of L-functions. The p-adic L-function, as mentioned, is obtained by interpolation from special values $L_p(1), L_p(0), L_p(-1), \ldots$.

The special values $L_p(0)$ and $L_p(1)$ were obtained, when possible, through the class number formula and through the p-adic class number formula [18, Theorem 5.24]. As for the values $L_p(k)$ for integral $k < 0$, they were obtained either from the exact formula [18, Theorem 5.11] or via transcendental methods.

In practice, we have sometimes used the explicit formula, sometimes used the approximate functional equation, and sometimes used PARI's inbuilt function `zetak, zetakinit` and the formula $L(s, \chi_q) = \frac{L(\mathbf{Q}(\sqrt{q}),s)}{\zeta(s)}$, which in many cases ran faster than our code.

5.2. Interpolation. In this section, we explain how one can efficiently compute the p-adic L-function to some desired precision – in particular, how to provably compute the λ invariant – by interpolation of special values.

Proposition. (Interpolation of p-adic power series.) *Suppose $x_1, \ldots, x_N \in p\mathbf{Z}_p$ are distinct. Then there is a linear functional $S_n : \mathbf{Q}_p^N \to \mathbf{Q}_p$, depending only on $x_1, \ldots, x_N$, so that, for any power series $f = \sum c_s T^s \in \mathbf{Z}_p[[T]]$,*

$$c_n - S_n(f(x_1), f(x_2), \ldots, f(x_N)) \in p^{N-n}\mathbf{Z}_p.$$

If the x_i become p-adically close, the coefficients of S_N will be p-adically large, and it becomes essential that we know the values $f(x_j)$ to high p-adic accuracy.

Proof. Define, for any function $g : \mathbf{Z}_p \to \mathbf{Q}_p$ and any $n \geq 2$

$$\Lambda_n(g) = \det \begin{pmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-2} & x_2^{n-2} & \dots & x_n^{n-2} \\ g(x_1) & g(x_2) & \dots & g(x_n) \end{pmatrix}.$$

For $n = 1$, we define $\Lambda_1(g) := g(x_1)$.

For ease of notation, we denote $\Lambda_n((x \mapsto x^k)) \in \mathbf{Z}_p[x_1, \dots, x_n]$ by $\Lambda_n(x^k)$. Such determinants arise in the study of Schur polynomials. Note that $\Lambda_n(x^k) = 0$ for $k \leq n-2$. For larger k, we get an integrality condition: For any $x_1, \dots, x_n \in p\mathbf{Z}_p$, we have

$$\left| \frac{\Lambda_n(x^m)}{\Lambda_n(x^{n-1})} \right|_p \leq \frac{1}{p^{m+1-n}} \tag{11}$$

The quotient $\frac{\Lambda_n(x^m)}{\Lambda_n(x^{n-1})}$ is the sum of all monomials of degree $m+1-n$. In particular, it is a multiple of p^{m+1-n} when evaluated anywhere in $(p\mathbf{Z}_p)^n$. This proves (11).

Now write $f_N = \sum_{s=0}^{N-1} c_s T^s$, a truncated version of f. Then we have

$$\begin{pmatrix} \Lambda_1(1) & \Lambda_1(x) & \Lambda_1(x^2) & \Lambda_1(x^3) & \dots & \Lambda_1(x^{N-1}) \\ 0 & \Lambda_2(x) & \Lambda_2(x^2) & \Lambda_2(x^3) & \dots & \Lambda_2(x^{N-1}) \\ 0 & 0 & \Lambda_3(x^2) & \Lambda_3(x^3) & \dots & \Lambda_3(x^{N-1}) \\ 0 & 0 & 0 & \Lambda_4(x^3) & \dots & \Lambda_4(x^{N-1}) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \Lambda_N(x^{N-1}) \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_{N-1} \end{pmatrix} = \begin{pmatrix} \Lambda_1(f_N) \\ \Lambda_2(f_N) \\ \Lambda_3(f_N) \\ \Lambda_4(f_N) \\ \vdots \\ \Lambda_N(f_N) \end{pmatrix}.$$

We adopt the convention that the rows and columns of matrices are numbered starting from 0 instead of 1, so that c_i lies in row i in the above formula.

The upper triangular matrix above factors as the product DU, where D is the diagonal matrix with entries $\Lambda_1(1), \Lambda_2(x), \dots, \Lambda_{N+1}(x^N)$, and U is the upper triangular unipotent matrix whose (i,j)-entry is $\Lambda_{i+1}(x^j)/\Lambda_{i+1}(x^i)$ for $j > i$. Note that U lies in the group of matrices whose (i,j) entry is divisible by p^{j-i}; thus U^{-1} lies in this group as well. Writing u'_{ik} for the (i,k)-entry of U^{-1}, we have

$$c_i = \sum_{k=i}^{N-1} u'_{ik} \frac{\Lambda_{k+1}(f_N)}{\Lambda_{k+1}(x^k)}$$

This gives a formula for c_i in terms of f_N; it remains to bound the error when we replace f_N by f. Write $g = f_N - f$. Then

$$c_i - \sum_k u'_{ik} \frac{\Lambda_{k+1}(f)}{\Lambda_{k+1}(x^k)} = \sum_k u'_{ik} \frac{\Lambda_{k+1}(g)}{\Lambda_{k+1}(x^k)}.$$

Taking into account the bounds $|u'_{ik}|_p \leq p^{k-i}$ and $|\frac{\Lambda_{k+1}(g)}{\Lambda_{k+1}(x^k)}|_p \leq p^{N-k}$, one has

$$\left| c_i - \sum_k u'_{ik} \frac{\Lambda_{k+1}(f)}{\Lambda_{k+1}(x^k)} \right| \leq p^{N-i}.$$

This finishes the proof, taking

$$S_i(f(x_1), \ldots, f(x_N)) = \sum_{k=i}^{N} u'_{ik} \frac{\Lambda_{k+1}(f)}{\Lambda_{k+1}(x^k)}.$$

□

5.3. Computing the Iwasawa polynomial. Let

$$f(T) = \sum_{i=0}^{\infty} a_i T^i \in \mathbf{Z}_p[[T]]$$

be a power series whose p-adic Weierstrass factorization is

$$f(T) = u(T)P(T),$$

where $P(T) = T^\lambda + c_{\lambda-1}T^{\lambda-1} + \ldots c_0$ is a distinguished monic polynomial of degree λ and $u(T)$ is a unit. Note that λ is easy to obtain from f: it is the smallest positive integer i for which a_i is a unit.

Our entire setup presupposes that such an integer i exists; this is always the case in our application because of the vanishing of the μ-invariant [4].

Proposition. *Suppose one knows the coefficients a_i of the power series $f(T)$ to precision $O(p^{K+1-i})$. Then for each $k \leq \lfloor K/\lambda \rfloor$, one can compute the coefficients $c_0, c_1, \ldots, c_{K-\lambda k}$ of the polynomial part $P(T)$ to precision $O(p^k)$.*

Proof. Write $\frac{1}{u(T)} = \sum_{i=0}^{\infty} b_i T^i$. We compute $b_0, b_1, \ldots b_{K-\lambda k}$ to precision $O(p^k)$. Then, since $c_n = \sum_{i=0}^{n} a_i b_{n-i}$, the proposition follows.

We use the fact that P is a distinguished polynomial of degree λ to compute the b_i. Consider the system of equations below, where s ranges from 1 to $K - k\lambda$:

$$\sum_{i+j=\lambda} a_i b_j = 1 \tag{12}$$

$$\sum_{i+j=\lambda+s} a_i b_j = 0 \tag{13}$$

First suppose that $k = 1$. Because $a_i \equiv 0 \mod p$ for $i < \lambda$, we find using equation (12) above that $a_\lambda b_0 \equiv 1 \mod p$, and thus we determine b_0

mod p. Next, suppose we know $b_{s'} \mod p$ for all $s' < s$, then again noting that $a_i \equiv 0 \mod p$ for $i < \lambda$, we use equation (13) to determine $b_s \mod p$:

$$b_s \equiv -\frac{1}{a_\lambda} \sum_{i=1}^{s} a_{\lambda+i} b_{s-i} \mod p.$$

Thus we can determine b_s to precision $O(p)$ for $s \leq K - \lambda$.

Now suppose that for all $k' < k$ we can compute b_s to precision $O(p^{k'})$ for $s \leq K - k'\lambda$. We use the equation (12) and fact that $a_i \equiv 0 \mod p$. Since we know b_s to precision $O(p^{k-1})$ for $s \leq K - (k-1)\lambda$, we are able to determine b_0 to precision $O(p^k)$. Suppose we know $b_{s'}$ to precision $O(p^k)$ for all $s' < s$, where s is some number $\leq K - k\lambda$. Then we use equation (13) and the fact that $a_i \equiv 0 \mod p$ to determine b_s to precision $O(p^k)$. □

5.4. Verification. Our routine passed the following three tests of correctness:

- Checks against other tables of λ-invariants;
- Check against Ferrero's formula for the 2-adic λ-invariant.
- Checking against the computations of Ernvall-Metsalynka. (We note that our normalization of the Iwasawa polynomial does not agree with [3], and a change of variable is required.)

We did not implement these checks "systematically." We simply chose a few discriminants by hand in each case and verified matching.

References

[1] M. Bhargava, A. Shankar, and J. Tsimerman. On the Davenport-Heilbronn theorem and second order terms. *Arxiv preprint arXiv:1005.0672*, 2010.

[2] J.S. Ellenberg, A. Venkatesh, and C. Westerland. Homological stability for Hurwitz spaces and the Cohen-Lenstra conjecture over function fields. *Arxiv preprint arXiv:0912.0325*, 2009.

[3] R. Ernvall and T. Metsänkylä. Computation of the zeros of p-adic L-functions. *Math. Comp.*, 58(198):815–830, S37–S53, 1992.

[4] Bruce Ferrero and Lawrence C. Washington. The Iwasawa invariant μ_p vanishes for abelian number fields. *Ann. of Math. (2)*, 109(2):377–395, 1979.

[5] Étienne Fouvry and Jürgen Klüners. Cohen-Lenstra heuristics of quadratic number fields. In *Algorithmic number theory*, volume 4076 of *Lecture Notes in Comput. Sci.*, pages 40–55. Springer, Berlin, 2006.

[6] Eduardo Friedman and Lawrence C. Washington. On the distribution of divisor class groups of curves over a finite field. In *Théorie des nombres (Quebec, PQ, 1987)*, pages 227–239. de Gruyter, Berlin, 1989.

[7] Derek Garton. Random matrices and the Cohen-Lenstra statistics for global fields with roots of unity. 2010.

[8] N.M. Katz and P. Sarnak. Zeroes of zeta functions and symmetry. *Bulletin of the American Mathematical Society, volume=36, pages=1–26, year=1999, publisher=American Mathematical Society.*

[9] JP Keating and NC Snaith. Random matrix theory and L-functions at s= 1/2. *Communications in Mathematical Physics*, 214(1):91–100, 2000.

[10] Y. Kida. On cyclotomic Z2-extensions of imaginary quadratic fields. *Tôhoku Math. J.(2)*, 31:91–96, 1979.

[11] James S. Kraft and Lawrence C. Washington. Heuristics for class numbers and lambda invariants. *Math. Comp.*, 76(258):1005–1023 (electronic), 2007.
[12] AA Michael. Finite Abelian actions on surfaces. *Topology and its Applications*, 153(14):2591–2612, 2006.
[13] D.P. Roberts. Density of cubic field discriminants. *Mathematics of Computation*, 70(236):1699–1705, 2001.
[14] Joseph J. Rotman. *An introduction to the theory of groups*, volume 148 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, fourth edition, 1995.
[15] Robert Steinberg. *Endomorphisms of linear algebraic groups*. Memoirs of the American Mathematical Society, No. 80. American Mathematical Society, Providence, R.I., 1968.
[16] The PARI Group, Bordeaux. *PARI/GP, version* `2.3.4`, 2008. available from `http://pari.math.u-bordeaux.fr/`.
[17] F. Thorne. The secondary term in the counting function for cubic fields. *preprint*, 2010.
[18] Lawrence C. Washington. *Introduction to cyclotomic fields*, volume 83 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1997.