STATISTICS OF MODULAR FORMS WITH SMALL RATIONALITY FIELDS

ALEX COWAN AND KIMBALL MARTIN

Abstract. We describe a database of weight 2 newforms in prime level less than 2 million with rationality field of degree $d \leq 6$. We use this to study the distribution of degrees and rationality fields of weight 2 newforms. In particular, we give heuristic upper bounds on how often degree d rationality fields occur in squarefree levels, and predict finiteness if $d \geq 7$. When $d = 2$, we make predictions about how frequently specific quadratic fields occurs, prove lower bounds, and conjecture that $\mathbb{Q}(\sqrt{5})$ is the most common quadratic rationality field. We also discuss how often non-conjugate degree d newforms occur in the same level, distributions of Atkin–Lehner signs, and Lang–Trotter-type questions.

Contents

Date: January 4, 2023.

AC was supported by the Simons Foundation (Collaboration Grant 550031).

KM was supported by the Simons Foundation (Collaboration Grant 512927), the Japan Society for the Promotion of Science (Invitational Fellowship L22540), and the Osaka Central Advanced Mathematical Institute (MEXT Joint Usage/Research Center on Mathematics and Theoretical Physics JPMXP0619217849).

1. Introduction

In this paper, we present a new dataset consisting of weight 2 newforms with trivial nebentypus, prime level between 10^4 and $2 \cdot 10^6$, and rationality field degree at most 6. With this dataset, we explore a variety of statistics of modular forms.

Our dataset builds on several others. All weight 2 modular forms with level at most 10^4 were computed in [3]. All weight 2 classical eigenforms with rational coefficients and level at most $5 \cdot 10^5$ were computed by Cremona [9]. Weierstrass equations for all elliptic curves of prime conductor less than $2 \cdot 10^9$ were computed by Bennett, Gherga, and Rechnitzer [2]. These datasets are available in the LMFDB [26]. Our dataset was computed using an algorithm of the first author [7]. The q-expansion of each form was computed up to at least the Sturm bound.

For a newform f, denote by K_f its rationality field. By the (rationality) *degree* of f, we mean the degree of $K_f/\mathbb{Q}$, and by the *discriminant* of f we mean the discriminant of K_f. Each number field which appears as K_f for some form in our dataset is the only totally real number field of degree 6 or less with discriminant $\mathrm{Disc}(K_f)$, and in that sense is uniquely determined by its discriminant. Unless mentioned otherwise, we always consider forms up to Galois conjugacy.

Table 1.1 summarizes our dataset. We tabulate the number of forms of prime level in the ranges 1 to 10^4, 10^4 to 10^6, and 10^6 to $2 \cdot 10^6$, grouped by discriminant and Atkin–Lehner eigenvalue. We omit discriminants which do not appear in our dataset, even when there is a prime level newform with that discriminant and level less than 10^4. Throughout the paper we adopt the convention that a blank entry in a table means that the entry is 0.

Deg	Disc	$\mathrm{Gal}(K_f/\mathbb{Q})$	Total	$1 — 10^4$		$10^4 — 10^6$		$10^6 — 2 \cdot 10^6$	
				$+$	$-$	$+$	$-$	$+$	$-$
1	1	C_1	15578	140	189	4364	4479	3206	3200
2	5	C_2	3044	93	65	938	962	508	478
	8	C_2	379	18	19	115	127	54	46
	13	C_2	59	4	9	21	19	1	5
	12	C_2	18		1	8	6	1	2
	21	C_2	5		1	1	2		1
	17	C_2	1			1			
3	49	C_3	154	19	15	40	50	20	10
	229	S_3	29	6	2	13	7		1
	148	S_3	18	7	5	3	3		
	81	C_3	16	2	1	2	11		
	257	S_3	16	3	6	4	2		1
	169	C_3	11	1	1	2	4	1	2
	321	S_3	3		2		1		
4	725	D_4	22	10	6	2	3		1
	1957	S_4	6	2	2	1	1		
	2777	S_4	5	2	1		2		
	8768	D_4	1			1			
5	70601	S_5	3	2			1		
	11^4	C_5	1				1		
6	13^5	C_6	1				1		

Table 1.1. Number of prime level newforms by degree, discriminant, and Atkin–Lehner sign

The first statistic we investigate is the distribution of degrees of newforms. In [28], the second author made conjectures about the distribution of Galois orbits of newforms for squarefree level, and computed the rationality fields of weight 2 newforms for prime levels up to $6 \cdot 10^4$. The main conjecture asserts that, on average, each Atkin–Lehner eigenspace consists of a single Galois orbit. In particular, 100% of the time there are no small degree newforms. Our data is in strong agreement with this conjecture. The heuristics in [28] also suggest higher degree newforms occur less frequently than lower degree newforms, but no precise conjecture was formulated. In Section 2, we refine earlier heuristics to conjecture asymptotic upper bounds for the number of degree d newforms of squarefree level. In particular, this number is finite if $d \geq 7$. Our data, which we present in Figure 2.9, supports this conjecture and suggests the true growth rate may be smaller.

We remark that the restriction to squarefree level is primarily to avoid complications from considering more local inertial types in general levels. However, philosophically there should not be too much difference between such counting problems for squarefree levels and general levels, except that for general levels one should avoid simple constructions—namely quadratic twists and CM forms. Thus we expect counting quadratic twist classes (by minimal level) of non-CM newforms should be roughly similar to counting newforms of squarefree level.

In Section 3, we study the distribution of rationality fields of weight 2 newforms in small degree. We present heuristics based on counting rational points on Hilbert modular varieties, and use this to suggest possible lower bounds on counts of newforms with rationality field $\mathbb{Q}(\sqrt{D})$ for $d = 5, 8, 12, 13, 17$. Moreover, we prove weaker lower bounds for $D = 5, 8$ and conjecture that 100% of the time the rationality field is quadratic, it is $\mathbb{Q}(\sqrt{5})$. Figure 3.8 and Figure 3.11 visualize the corresponding data.

Section 4 examines how often multiple forms of a prescribed discriminant have the same level. One natural hypothesis is that the number of forms in a given level follows a Poisson distribution. We construct a statistical test and with it find that forms of discriminant 1 in our dataset are inconsistent with this hypothesis. Forms of discriminants 5 and 8 in our dataset appear to have a similar type of bias. Table 4.1 summarizes our findings. There are known constructions for degree 1 which would lead to a distribution different than that hypothesized, but no analogous constructions are known for other degrees.

If one considers all newforms, not up to Galois conjugacy, there is a bias towards Atkin–Lehner sign -1 [27]. In Section 5, we investigate whether a similar bias appears for forms of fixed degree. Figure 5.1 visualizes our data. We find that biases towards each sign, as well as no bias at all, are nearly equiprobable given our dataset. We give a probabilistic heuristic that in fact, unless any underlying bias is very large, it would be impossible to discern even with datasets of arbitrary size.

In Section 6, we consider three Lang–Trotter-type questions. The first two of these were formulated by Murty [32, §3]. The first question is simply the classic Lang–Trotter conjecture [22] in higher degree. The second asks how often the Fourier coefficients of a newform f lie in a proper subfield of K_f. The third is about how many algebraic integers are the value of exactly k Fourier coefficients in a given range. This statistic appears to depend on the size of the image of the associated Galois representation—see Figure 6.7 for a picture in degree 2.

Acknowledgements. We are especially grateful to Noam Elkies for many insights and suggestions. We have also benefited from conversations with Armand Brumer and Ari Shnidman. The computations for our dataset were performed at the OU Supercomputing Center for Education & Research (OSCER) at the University of Oklahoma (OU).

2. Counts by degree

In this section we discuss estimates of the number of newforms of fixed degree. Write $\mathrm{Orb}(S_2^{\mathrm{new}}(N))$ for the set of newforms of $S_2(N)$ up to Galois conjugacy. For a given newform orbit f, let $\deg f$ and Δ_f denote the degree and discriminant of the number field K_f respectively, and let $w_p(f)$ denote its Atkin–Lehner eigenvalue. Define

$$\mathcal{F}(d,\Delta,w,N) \coloneqq \{f \in \mathrm{Orb}(S_2(N)) \,:\, \deg f = d,\, \Delta_f = \Delta,\, w_p(f) = w\} \tag{1}$$

and

$$C_{d,\Delta,w}(X) \coloneqq \sum_{\substack{p<X \\ p \text{ prime}}} \#\mathcal{F}(d,\Delta,w,p). \tag{2}$$

We also adopt the convention that if any of d, Δ, or w are not present, then no restriction on the degree, discriminant, or sign respectively is imposed. For example,

$$C_d(X) \coloneqq \#\{f \in \mathrm{Orb}(S_2(p)) \,:\, p < X, p \text{ prime}, \deg f = d\}.$$

Moreover, when N or X is omitted in these definitions, we will mean the union or sum over primes between 10^4 and $2 \cdot 10^6$. Since our data is for prime level, we restrict to prime level in this statistic, but many of our heuristics and results are formulated in terms of squarefree or general level.

Watkins [43], building on heuristics of Brumer and McGuinness [6], formulates heuristics that suggest

$$\mathrm{li}(X^{\frac{5}{6}}) \ll C_1(X) \ll \mathrm{li}(X^{\frac{5}{6}}). \tag{3}$$

These heuristics come from analysis of Weierstrass models of elliptic curves, and (3) agrees very well with the data available in the large database of [2]. The best known result is $C_1(X) \ll X^{1+\varepsilon}$ [11].

For $d > 1$, the situation is much more mysterious. In [38], Serre proves that $C_d(X) = o(\mathrm{li}(X^2))$ for any d (in fact, he proves, for any prime ℓ, that the number of forms of degree d and level N coprime to ℓ is $o(\dim(S_2(N)))$ as $N \to \infty$). Serre's theorem was made effective by Murty and Sinha [31], and an improved bound in the case of N prime was given by Sarnak and Zubrilina [37]. One cannot yet prove that $C_d(X) \ll X^{\beta_d}$ for any $d > 1$ and any $\beta_d < 2$.

Let us consider statements of the form

$$\mathrm{li}(X^{\alpha_d}) \ll C_d(X) \ll \mathrm{li}(X^{\beta_d}). \tag{4}$$

The Brumer–McGuinness–Watkins conjecture (3) is the assertion that (4) holds for $\alpha_1 = \beta_1 = \frac{5}{6}$. We know of no other conjectures about statements of the form (4), and it seems very hard even to put forward a reasonable guess about what might be true when $d > 1$. This situation is more difficult to reason about heuristically, in part because there's no clear analogue of an elliptic curve's Weierstrass model, and the data previously available was over too small of a range to allow for guesses based on empirics. Recently, based on computations up to $X = 6 \cdot 10^4$, the second author observed [28, Question 3.1] that perhaps (4) holds with $\beta_d = 0$ for sufficiently large d.

Based on a combination of inspection of our dataset and heuristics, we propose a number of statements of the form (4). In Section 2.1, we formulate Conjecture 2.4, that (4) holds with $\beta_d = 1 - \frac{d}{6} + \varepsilon$ for $d \leq 6$, and $\beta_d = 0$ for $d > 7$. Then, in Corollary 3.6, based on constructions of genus 2 curves, we show that (4) holds with $\alpha_2 = \frac{1}{3} - \varepsilon$ if one assumes the Bateman–Horn conjecture [1]. In Section 2.2, we present empirical results from our dataset.

2.1. **Random Hecke polynomial model.** Here we present a random model to estimate the distribution of degree d newforms. This is based on ideas for heuristics suggested in [36] and [28].

Consider a newspace $S_{2k}^{\mathrm{new}}(N)$. One can further decompose this space into $2^{\omega(N)}$ joint eigenspaces of the Atkin–Lehner operators W_p $(p \mid N)$, which we call the Atkin–Lehner eigenspaces. Each Atkin–Lehner eigenspace is Galois invariant. In general, one can further decompose each Atkin–Lehner eigenspace into smaller Galois invariant subspaces according to local inertia types of non-CM forms (see [10]) and the subspace of CM forms.

For simplicity, assume N is squarefree. Then there are no CM forms of trivial nebentypus and only one local inertial type. Let S be an Atkin–Lehner eigenspace in $S_{2k}^{\mathrm{new}}(N)$. For a newform $f \in S$, the single Fourier coefficient $a_p(f)$ generates K_f for 100% of p [21], and it is conjectured that this is true for all but finitely many p if $[K_f : \mathbb{Q}] > 4$ [32]. Hence, for fixed $p \nmid N$, the factorization type of the characteristic polynomial $c_{T_p}(x) \in \mathbb{Z}[x]$ will usually tell us the degrees of the newforms in S. In fact, it will always give us lower bounds.

Let $n = \dim S$. As in [36] and [28], we can model $c_{T_p}(x)$ as a random polynomial in the set H_n of degree n monic integral polynomials whose roots α satisfy $|\alpha| \le 2p^{k-1/2}$. Alternatively, one could look at the set of Weil q-polynomials of degree $2n$, where $q = p^k$, or isogeny classes of n-dimensional abelian varieties over $\mathbb{F}_{p^k}$.

Set $h(n) = \#H_n$. Since $h(n)$ grows relatively quickly with n, we expect that a random element of $h(n)$ is irreducible with high probability (approaching 1 as $n \to \infty$), and the probability that a random element factors is roughly the probability that is has a degree 1 factor, which is about $\frac{h(1)h(n-1)}{h(n)} \approx \frac{h(n-1)}{h(n)}$. (See [28] for more details.) Similarly, for $d \ll n$, the probability of a degree d factor is approximately

$$\frac{h(d)h(n-d)}{h(n)} \approx \frac{h(n-d)}{h(n)} \approx \frac{h(n-d)}{h(n-d+1)} \cdot \frac{h(n-d+1)}{h(n-d+2)} \cdots \frac{h(n-1)}{h(n)} \approx \left(\frac{h(n-1)}{h(n)}\right)^d.$$

Studying the precise asymptotics of $h(n)$ for fixed p, k as $n \to \infty$ appears to be a very hard problem. But it at least suggests the following way to relate asymptotics for degree d newforms to rational newforms.

Lemma 2.1. *Let $\nu_{2k}(X)$ be the number of Atkin–Lehner eigenspaces in $\bigcup_N S_{2k}^{\mathrm{new}}(N)$, where N ranges over squarefree levels, having dimension less than X. Then $X \ll \nu_{2k}(X) \ll X^{1+\varepsilon}$, for any $\varepsilon > 0$.*

Proof. Since $\dim S_{2k}^{\mathrm{new}}(N) \ll N$, the lower bound is obvious.

Let us show the upper bound. First, it follows from the dimension formulas for Atkin–Lehner eigenspaces from [27] that any Atkin–Lehner eigenspace in $S_{2k}^{\mathrm{new}}(N)$ has dimension $\frac{(k-1)\phi(N)}{12 \cdot 2^{\omega(N)}} + O(N^{1/2+\varepsilon})$. (The necessary argument, though not the statement, is given in the proof of [29, Proposition 3.10].) This dimension is $\gg N^{1-\varepsilon}$. Hence if an Atkin–Lehner eigenspace has dimension less than X, it occurs in a level $N \ll X^{1+\varepsilon}$. Now the number of Atkin–Lehner eigenspaces in levels less than t is bounded by $\sum_{N \le t} 2^{\omega(N)}$. It is known that this latter sum is $\frac{6}{\pi^2} t \log t + O(t)$. $\square$

Heuristic 2.2. *Suppose the number rational newforms of weight $2k$ and squarefree level $N < X$ is $O(X^{1-\alpha})$ for some $\alpha < 1$. Then, for any $\varepsilon > 0$, the number of degree d weight $2k$ newforms of squarefree level $N \le X$ is $O(X^{1-\alpha d+\varepsilon})$ as $X \to \infty$.*

Our reasoning for this heuristic is as follows. Under the hypothetical bound $O(X^{1-\alpha})$, the lemma indicates that the probability of an Atkin–Lehner space of dimension n having size 1 Galois orbit is approximately $n^{-\alpha}$. Our above estimates then suggest a probability roughly $n^{-d\alpha}$ for a size d Galois orbit. Applying the lemma again leads to the stated heuristic.

Remark 2.3. Roberts' conjecture [36] implies that there are only finitely weight k rational non CM-newforms up to quadratic twists if $2k \ge 6$. Hence this heuristic suggests that there are only finitely many newforms of squarefree level of fixed weight $2k$ and degree d. One might similarly expect that for arbitrary level, there are finitely quadratic twist classes of non-CM newforms of fixed degree d and weight $2k$.

There are two significant caveats to our heuristic. First, ignores any potential arithmetic reasons for nontrivial decompositions of the Atkin–Lehner eigenspaces as Hecke modules beyond the existence of rational newforms. We will discuss one aspect of this later: how often small degree forms arise in the same level in weight 2? However, such questions seem to be somewhat different than of the problem of counting averages over all levels, which is what Heuristic 2.2 addresses.

More seriously, the probabilistic argument assumes that the characteristic polynomials of T_n on these spaces are equidistributed among possible Hecke polynomials in H_n. This does not seem to be actually true. E.g., our heuristic reasoning would lead us to expect that counts of the following three type of objects have the same asymptotic growth rates (up to constants): (i) pairs of rational newforms in the same Atkin–Lehner eigenspace, (ii) degree 2 forms with rationality field $\mathbb{Q}(\sqrt{5})$, and (iii) degree 2 forms with rationality field $\mathbb{Q}(\sqrt{21})$. In fact what appears to be true, at least in weight 2, is that (i) is more common. (Also, (ii) appears much more common than (iii). We discuss this below.) That is, the mostly likely way to get a degree d factor of an element in H_n is to d degree 1 factors. Since we have only stated Heuristic 2.2 as predicting an upper bound, this only suggests that the upper bound in Heuristic 2.2 may not be sharp.

With this in mind, comparing our heuristic to those of Brumer–McGuinness and Watkins suggests following.

Conjecture 2.4. *Let $\varepsilon > 0$. The number of degree d weight 2 newforms of squarefree level $N \leq X$ is $O(X^{1-d/6+\varepsilon})$ as $X \to \infty$. In particular, this number is finite if $d \geq 7$.*

Similar to Remark 2.3, one can ask if this conjecture holds for arbitrary levels if one counts quadratic twist classes of non-CM newforms.

One can also ask whether or not a version of Conjecture 2.4 which is uniform in d holds.

Question 2.5. *Do there exist integers d^* and N^* such that, for all primes $p > N^*$,*

$$\#\{f \in \mathrm{Orb}(S_2^{\pm}(p)) \,:\, \deg f \geq d^*\} \leq 1?$$

2.2. **Data for counts by degree.** In generating this dataset, we determined the number of eigenform orbits of prime level less than $2\cdot 10^6$ and degree at most 6, as well as the decomposition of all newspaces of prime level less than 10^6. Table 2.7 shows the number of eigenform orbits of each degree, and Table 2.8 gives the decomposition of all levels in this range which have two or more forms of degree at least 7 in the same Atkin–Lehner eigenspace. Since Table 2.8 is quite short, we are prompted to ask (the negation of) Question 2.5 for the choice $N^* = 4751$ and $d^* = 7$:

Question 2.6. *Is there a prime $p > 4751$ and a sign $\pm$ such that*

$$\#\{f \in \mathrm{Orb}(S_2^{\pm}(p)) \,:\, \deg f \geq 7\} > 1?$$

Degree	$[1, 10^4]$	$[10^4, 10^6]$	$[10^6, 2\cdot 10^6]$	Total
1	329	8843	6406	15578
2	212	2200	1096	3508
3	76	142	35	253
4	28	10	1	39
5	20	2		22
6	11	1		12

Table 2.7. Counts of Galois orbits by degree and level

Level	$S_2^+(p)$	$S_2^+(p)$
607	$5+7+7$	31
911	$9+14$	53
1223	34	$9+59$
1249	$7+37$	59
4751	153	$18+225$

Table 2.8. Atkin–Lehner eigenspaces with multiple orbits of size ≥ 7

A least-squares fit to the data

$$\{(\log X, \log C_d(X)) : 10^4 < X < 2\cdot 10^6,\ X \text{ prime}\} \tag{5}$$

by functions of the form $y = \log(a\mathrm{li}(\exp(x)^b)))$ yields, for $d = 1, 2, 3, 4$, the fits depicted in Figure 2.9. A least-squares fit to the data $(X, C_d(X))$ yields similar values; the exponents are instead 0.841, 0.622, 0.329, and 0.106.

For degree 1, our best fit values $a = 0.97$ and $b = 0.832$ are in agreement with the elliptic curve database [2], which, for $X = 2\cdot 10^9$, finds $a \approx 0.97$ and $b \approx 0.833$, as the Brumer–McGuinness–Watkins heuristic (3) predicts.

Despite the visual disagreement between the model and the data for degree 4 in Figure 2.9, the value of the exponent b is rather insensitive to the starting count of 28 at $X = 10^4$; reducing this count from 28 to 10 changes the best-fit value of b from 0.107 to 0.105. The best-fit value of a changes substantially however, going from 12.0 to 6.4.

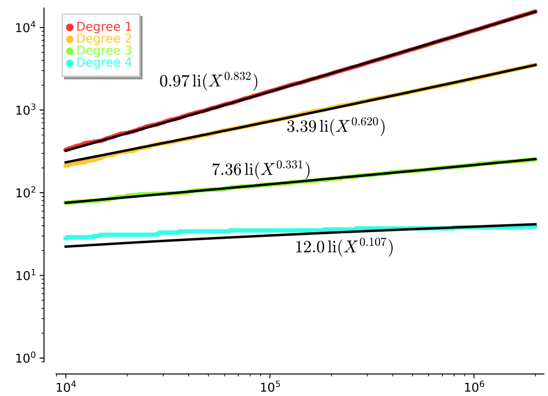

Figure 2.9. Number of forms with prime level less than X by degree, with least squares fits to the log-log data (5)

Conjecture 2.4 is in agreement with our prime level data. In Figure 2.9, the growth rate $O(X^{1-d/6})$ appears to be an upper bound in prime level for $d \le 4$. In this range the best fit has a notably lower exponent for $d = 3, 4$, but there may exist logarithmic factors in the main asymptotics for $d \ge 2$. For example, for $d = 2$, the Mestre obstruction to rationality of genus 2 curves may introduce a logarithmic factor in the denominator; there is no analogous obstruction for elliptic curves. In this range, $\log X > X^{\frac{1}{6}}$.

3. Hecke fields

In the previous section we considered the question of how often degree d newforms occur, and presented a random Hecke polynomial model, which at least for prime levels, appears to give asymptotic upper bounds. Here we consider the refined question of how often a specific degree d rationality field K should occur, and relate this question to rational points on Hilbert modular varieties. We discuss lower bound heuristics for quadratic fields, and predict that $\mathbb{Q}(\sqrt{5})$ is the most common quadratic rationality field. We also as examine the prime level data for $d \le 6$.

3.1. **Modular varieties.** First recall the connection between weight 2 modular forms and abelian varieties.

Let $N \ge 1$. To a newform $f \in S_2(N)$ with $[K_f : \mathbb{Q}] = d$, Shimura constructed a d-dimensional simple abelian variety $A_f/\mathbb{Q}$ satisfying the following properties. First, A_f is defined up to isogeny as a factor of $J_0(N)$. Moreover if A_f is isogenous to A_g if and only if f and g are Galois conjugates. The endomorphism algebra is $\mathrm{End}^0(A_f) = \mathrm{End}(A_f) \otimes \mathbb{Q} \simeq K_f$. The conductor of A_f is N^d. Finally, $L(s, A_f) = \prod_\sigma L(s, f^\sigma)$, where f^σ ranges over Galois conjugates of f.

In general, the center of the endomorphism algebra of a d-dimensional abelian variety A has degree $\le d$. Thus the assertion that $\mathrm{End}^0(A_f) \simeq K_f$ means that A_f has maximal real multiplication (RM). Conversely, if $A/\mathbb{Q}$ is a simple abelian variety with maximal RM, it is of GL(2) type, and so by Ribet [35] together and the work of Khare–Winterberger [20] on Serre's conjecture, $A = A_f$ for some cusp form $f \in S_k(N, \varepsilon)$ with some nebentypus ε. However, since K_f is totally real, either $\varepsilon = 1$ or f has CM by ε (e.g., see [34]). The latter can only happen in odd weight, so in fact $f \in S_k(N)$.

Hence the correspondence $f \mapsto A_f$ yields a bijection between degree d newforms f of weight 2 and isogeny classes of d-dimensional simple abelian varieties $A/\mathbb{Q}$ with maximal RM.

Now we suggest a heuristic approach to predicting coarse asymptotics for counts of such objects.

Let K be a totally real number field of degree d, and $\mathfrak{a}$ be an ideal in $\mathcal{O}_K$. The quotient $\mathfrak{H}^d/\mathrm{SL}(\mathcal{O}_K \oplus \mathfrak{a})$ parametrizes d-dimensional complex abelian varieties with RM by $\mathcal{O}_K$ together with a polarization structure corresponding to $\mathfrak{a}$ (see [16] for a precise statement). Compactifying this quotient and taking a projective resolution gives a Hilbert modular variety $Y(\mathcal{O}_K \oplus \mathfrak{a})$.

Now consider a newform $f \in S_2(N)$ with $K_f = K$. The abelian variety A_f has endomorphism ring an order in $\mathcal{O}_K$. Typically we expect it is all of $\mathcal{O}_K$, but if not, one can replace A_f by an isogenous variety with RM by $\mathcal{O}_K$. Thus f corresponds to a rational point y on $Y(\mathcal{O}_K \oplus \mathfrak{a})$ for some $\mathfrak{a}$, which we can take to be in a given set of representatives for $\mathrm{Cl}^+(K)$ This correspondence is far from one-to-one. First, replacing A_f by an isogenous variety, or modifying the polarization structure, may give a different point y on $Y(\mathcal{O}_K \oplus \mathfrak{a})$. Second, if g is another weight 2 newform and A_g is $\mathbb{C}$-isogenous to A_f, then both f and g correspond to the same rational points. Third, this is not a fine moduli space, so not all rational points on $Y(\mathcal{O}_K \oplus \mathfrak{a})$ will correspond to abelian varieties defined over $\mathbb{Q}$, and of those that do, some will correspond to non-simple abelian varieties.

That said, it seems reasonable to expect that quadratic twist classes of Galois orbits of weight 2 newforms generically correspond to finite sets of rational points on $Y = \bigcup_{a \in \mathrm{Cl}^+(K)} Y(\mathcal{O}_K \oplus \mathfrak{a})$. Thus one can attempt estimate the number of quadratic twist classes by estimating counts of rational points on Y. A priori, it is not clear how different orderings of classes of newforms (e.g., by minimal level) will correlate with different orderings of sets of rational points (e.g., by minimal height, for some choice of height function), and we will speculate more on this for $d = 2$ anon.

Note that typically (each component of) Y will be of general type, and one might expect that it has finitely many (and often no) rational points. Hence, for a given d, to estimate counts of degree d newforms as in Conjecture 2.4, it should in principle suffice to consider finitely many Y. Moreover, this philosophy suggests that some totally real degree d rationality fields will be more common than others, roughly according to whether the moduli spaces Y have many or few rational points. Of course this is not the only consideration, due to various complications of the correspondence between newforms and rational points mentioned above.

At least this philosophy is in line with Coleman's conjecture (e.g., see [4]), which predicts there are only finitely many isomorphism classes of endomorphism algebras for d-dimensional abelian varieties over $\mathbb{Q}$. Hence Coleman's conjecture implies that, for a fixed d, only finitely many degree d rationality fields K_f occur as f varies over weight 2 newforms.

In the next section, we will make these ideas more precise for $d = 2$.

3.2. **Counting rational points on Hilbert modular surfaces.** Now we estimate point counts on certain Hilbert modular surfaces, based on the ideas of the previous section for $d = 2$.

Let $D > 0$ be a fundamental discriminant and $\mathcal{O}_D$ be the ring of integers of $\mathbb{Q}(\sqrt{D})$. Let $Y_-(D)$ be the Hilbert modular surface constructed from the quotient $\mathfrak{H}^2/\mathrm{SL}(\mathcal{O}_D \oplus \sqrt{D}\mathcal{O}_D)$. This parametrizes principally polarized abelian surfaces with RM by $\mathcal{O}_D$ (together with a polarization structure). See [41] for details. For brevity, we will write RM D for RM by $\mathcal{O}_D$.

For our heuristic point counts, we will use explicit models for Hilbert modular surfaces. For $D < 100$, Elkies and Kumar [14] computed models for $Y_-(D)$. By work of Hirzebruch and Zagier [18], $Y_-(D)$ is rational (i.e., birational to $\mathbb{P}^2$) if and only if $D \in \{5, 8, 12, 13, 17\}$.

We expect that 100% of degree 2 weight 2 newforms correspond to rational points on Hilbert modular surfaces with the most rational points, i.e., the rational surfaces. While we do not know if any Hilbert modular surfaces parametrizing non-principally polarized surfaces with RM D are rational, we at least expect that the 5 rational $Y_-(D)$'s should account for a positive proportion of degree 2 weight 2 newforms, and this is supported by data.

Namely, the polarization classes of abelian surfaces with RM D are in bijection with the narrow ideal classes $\mathrm{Cl}^+(\mathbb{Q}(\sqrt{D}))$. In particular, if $D = 5, 8, 13, 17$, the abelian surface is automatically principally polarizable. Second, our data suggest that, at least for prime level, most degree 2 weight 2 newforms have rationality field $\mathbb{Q}(\sqrt{D})$ with $D = 5$ or 8, and thus correspond to points on $Y_-(5)$ and $Y_-(8)$. For other D, e.g., $D = 12$, it is not clear whether a positive proportion of abelian surfaces with RM D are principally polarizable.

For the remainder of the section, assume $D \in \{5, 8, 12, 13, 17\}$. Then $Y_-(D)$ is birational to $\mathbb{P}^2_{m,n}$. Let $\mathcal{A}_2$ be the moduli space for principally polarized abelian surfaces. Forgetting the RM action yields a map $Y_-(D) \to \mathcal{A}_2$.

Let $\mathcal{M}_2$ be the moduli space of genus 2 curves. To a genus 2 curve $C : y^2 = h(x)$, one associates Igusa–Clebsch invariants $I_{2j}(C)$ for $j = 1, 2, 3, 5$. Here $I_{2j}(C)$ can be regarded as a degree $2j$-polynomial in the coefficients of $h(x)$, and $I_{10}(C)$ is the discriminant of h. One can realize $\mathcal{A}_2$ as weighted projective space $\mathbb{P}^3_{1,2,3,5}$ with coordinates $(I_2 : I_4 : I_6 : I_{10})$. The Torelli map $\mathcal{M}_2 \to \mathcal{A}_2$ sends the moduli of C to $(I_2(C) : I_4(C) : I_6(C) : I_{10}(C))$, and the image is the complement of the hyperplane $I_{10} = 0$. We note that Igusa–Clebsch invariants are only isomorphism invariants of C up to weighted projective scaling.

For generic affine coordinates $(m, n) \in \mathbb{A}^2$, one has an associated point on $Y_-(D)$ and thus weighted projective coordinates $(I_2(m, n) : I_4(m, n) : I_6(m, n) : I_{10}(m, n)) \in \mathcal{A}_2$, where the $I_{2j}(m, n)$'s are rational functions in m, n. Elkies and Kumar [14] gave birational models for $Y_-(D)$ together with an explicit description of the rational functions $I_{2j} = I_{2j}(m, n)$.

Now we will attempt to estimate the number of rational points (m, n) with bounded Igusa–Clebsch invariants. First we want to scale Igusa–Clebsch invariants (in $\mathbb{P}^3_{1,2,3,5}$) to be integral, as will be the case for the $I_{2j}(C)$'s given a curve C over $\mathbb{Z}$. Let us write $(m, n) = (a/c, b/c)$ for $a, b, c \in \mathbb{Z}$ with $\gcd(a, b, c) = 1$. Regarding a, b, c as variables, we scale the $I_{2j}(m, n)$'s to get polynomials $I_{2j}(a, b, c) \in \mathbb{Z}[a, b, c]$'s which are minimal integral over $\mathbb{Z}[a, b, c]$. That is, we scale out denominators, and also any factors of the numerators π so that $\pi^j \mid I_{2j}(a, b, c)$ for all $j \in \{1, 2, 3, 5\}$ implies π is a unit in $\mathbb{Z}[a, b, c]$. The resulting $I_{10}(a, b, c)$'s (which are uniquely determined up to ± 1) are given in Table 3.1.

D	$I_{10}(a,b,c)$
5	$8(a^5 - 10a^3b^2 + 25ab^4 + 5a^4c - 50a^2b^2c + 125b^4c - 5a^3c^2 + 25ab^2c^2 - 45a^2c^3 + 225b^2c^3 + 108c^5)^2$
8	$8c^3(a-c)^3(a+c)^6(-16a^2b^2 + 32b^4 + a^3c - 56ab^2c + 9a^2c^2 - 72b^2c^2 + 27ac^3 + 27c^4)^2$
12	$(a+c)^3(a-c)^9(-27a^2 + b^2 + 27c^2)^2(a^2b + 9a^2c - 8c^3)^3$
13	$2^3 \cdot 3^{11} \cdot (-267a^3 + 72a^2b - ab^2 - 3552a^2c + 1440abc - 128b^2c + 768ac^2)^2 \cdot (-12a^3 + 3a^2c + b^2c)^4(-a^3 - 150a^2c + 6abc - 264ac^2 + 120bc^2 + 64c^3)^4$
17	$2^{15} \cdot 3^{11} \cdot (-132a + b + 3c)^3(4608a^3 - 1728a^2c + b^2c + 216ac^2 - 9c^3)^2 \cdot (456a^2 + ab + 723ac - 8bc + 24c^2)^3(-256a^3 - 1200a^2c + 18abc - 6006ac^2 + 99bc^2 + 41c^3)^5$

Table 3.1. I_{10} polynomials for $Y_-(D)$

Specializing a, b, c to integers, we denote by $I_{2j}^{\min}(a,b,c) \in \mathbb{Z}$ scalings which are minimal integral over $\mathbb{Z}$. Note that $I_{2j}^{\min}(a,b,c) \mid I_{2j}(a,b,c)$ but they are often not equal. E.g., when $D = 5$, then the invariants $I_{2j}(1,3,2)$'s are $(-2^4 \cdot 5^3, 2^8 \cdot 5^4, -2^{15} \cdot 5 \cdot 599, 2^{21} \cdot 3^8)$, whereas the $\mathbb{Z}$-minimal invariants $I_{2j}^{\min}(a,b,c)$ are obtained by scaling out a factor of 2^4, i.e., they are $(-5^3, 5^4, -2^3 \cdot 5 \cdot 599, 2 \cdot 3^8)$.

First we want to estimate, in terms of a real parameter T, the growth of the cardinality of

$$Z_D(T) := \{(a,b,c) \in \mathbb{Z}^3 - U_D : \gcd(a,b,c) = 1 \text{ and } |I_{2j}^{\min}(a,b,c)| < T^{2j} \text{ for } j \in 1,2,3,5\},$$

where U_D consists of (a,b,c) such that the map $(a/c, b/c) \to \mathcal{A}_2$ is either undefined (e.g., $c = 0$) or is not finite-to-one (e.g., for $D = 8$, all points with $m = a/c = -1$ map to $(1:0:0:0) \in \mathcal{A}_2$). Really our interest is just in bounding $I_{10}^{\min}$, but we impose bounds on the other $I_{2j}^{\min}$'s to guarantee finiteness of $Z_D(T)$.

Precise estimates are difficult, so we make two simplifications which are sufficient to get lower bounds: (1) We impose the stronger bound $|I_{2j}^{\min}(a,b,c)| \le |I_{2j}(a,b,c)| < T^{2j}$. (2) We will suppose each monomial in $I_{2j}(a,b,c)$ is bounded by T^{2j}. Note that (1) and (2) can respectively be thought of as non-archimedean and archimedean simplifications to monomials.

Proposition 3.2. *We have $\#Z_D(T) \gg T^{r_D}$ as $T \to \infty$, where respectively $r_D = 3, \frac{3}{2}, 2, 1, 1$ for $D = 5, 8, 12, 13, 17$.*

Proof. Each $I_{2j}(a,b,c)$ is a homogeneous polynomial in a, b, c, say of degree d. Taking a, b, c independently up to size $T^{2j/d}$ shows there are $\gg T^{6j/d}$ tuples (a,b,c) with $|I_{2j}(a,b,c)| < T^{2j}$. Moreover, the ratio j/d is independent of the choice of j. Since that the conditions $\gcd(a,b,c) = 1$ and $(a,b,c) \notin U_D$ are satisfied for a positive proportion of (a,b,c), we get the asymptotic lower bound $\#Z_D(T) \gg T^{30/d}$, where $d = \deg I_{2j}(a,b,c)$. We respectively have $d = 10, 20, 25, 30, 30$ for $D = 5, 8, 12, 13, 17$, which gives the asserted lower bounds for $D = 5, 8, 13, 17$.

For $D = 12$, one can get better lower bounds using a $\mathbb{P}^1 \times \mathbb{P}^1$ parametrization for (m, n). Namely, write $(m,n) = (r/s, t/u)$ for $r, s, t, u \in \mathbb{Z}$. Let $I_{2j}(r,s,t,u)$ be the minimal invariants over $\mathbb{Z}[r,s,t,u]$. These have degrees $6, 12, 18, 30$ for $j = 1, 2, 3, 5$. Using same argument as above gives a lower bound of $\#Z_{12}(T) \gg T^{4/3}$. However, if we regard $I_{2j}(r,s,t,u)$ as polynomials only in t and u, the respective degrees are $2, 4, 6, 10$ for $j = 1, 2, 3, 5$. Thus by taking r, s uniformly bounded and $|t|, |u| \ll T$ yields $\#Z_{12}(T) \gg T^2$ as claimed. □

The lower bounds in the proposition are the optimal ones we could find using the $\mathbb{P}^2$ or $\mathbb{P}^1 \times \mathbb{P}^1$ parametrizations for (m,n) by allowing either each of a, b, c or r, s, t, u to vary independently up to some power of T (not necessarily the same power for each variable). We remark that the optimal exponents for lower bounds using the $\mathbb{P}^1 \times \mathbb{P}^1$ for $D = 5, 8, 13, 17$ are respectively $2, 4/3, 1, 1$. To get these exponents, for $D = 5$ one can take each of $|r|, |s|, |t|, |u| \ll T$. For $D = 8$, one takes $|r|, |s| \ll T^{2/3}$ and $|t|, |u| \ll 1$. For both $D = 13$ and $D = 17$, one takes $|r|, |s| \ll 1$ and $|t|, |u| \ll 1/2$.

Question 3.3. *For $D \in \{5, 8, 12, 13, 17\}$, is $\#Z_D(T) \ll T^{r_D + \varepsilon}$ for any $\varepsilon > 0$?*

It is not clear if the simplifications to monomials affect the exponents in our estimates, but if the I_{2j} polynomials are sufficiently general type, one might expect they only account for a multiplicative factor of size $O(1 + T^{\varepsilon})$. Note that in the Brumer–McGuinness heuristics, it is believed the analogous archimedean simplification (2) only affects counts by an $O(1)$ factor.

A more serious reason to doubt the exponents in these lower bounds are optimal is that there may exist (i) other rational parametrizations of $Y_-(D)$ where the I_{2j} degrees are smaller, or (ii) special curves on $Y_-(D)$ which intersect Z_D in an especially large number of rational points. Indeed, the proof makes clear that different types of parametrizations may yield better counts for different surfaces.

Now we explain how these estimates for $\#Z_D(T)$ are related to counting quadratic twist classes of weight 2 newforms f with rationality field $\mathbb{Q}(\sqrt{D})$. As explained above, a positive proportion of f (all, if $D \neq 12$) should correspond to rational points on $Y_-(D)$. Conversely, a rational point on $Y_-(D)$ may not come from a simple abelian surface with RM defined over $\mathbb{Q}$—one needs that the Mestre conic has a point and the Jacobian is nonsplit. However, we expect that these obstructions will only contribute logarithmic factors to asymptotics. (See [8] for details about when the Mestre obstruction vanishes.)

Consider a point $(a, b, c) \in Z_D(T)$ which corresponds to a $\bar{\mathbb{Q}}$-isomorphism class of simple abelian surfaces $A/\mathbb{Q}$. Generically this $\bar{\mathbb{Q}}$-isomorphism class is just the family of Jacobians of quadratic twists of a genus 2 curve $C/\mathbb{Q}$ with RM D. Suppose this, and assume C is a minimal quadratic twist of conductor N_C. Then (a, b, c) corresponds to the quadratic twist class of some minimal newform f of level N_f where $N_C = N_f^2$. One can write down a minimal integral model for C, and the polynomially-defined Igusa–Clebsch invariants $I_{2j}(C)$ are necessarily divisible by $I_{2j}^{\min}(a, b, c)$. Thus $I_{10}^{\min}(a, b, c)$ divides the minimal discriminant $\Delta_C = 2^{-12} I_{10}(C)$ of C. One also knows that the conductor $N_C \mid \Delta_C$.

Now we would like to understand how N_C relates to $I_{10}^{\min}(a, b, c)$ or $I_{10}(a, b, c)$. There are no general upper or lower bounds, and in fact there are competing issues in opposite directions. One is that Δ_C may be much larger than either $I_{10}^{\min}(a, b, c)$ or $I_{10}(a, b, c)$, and numerically this is quite typical. E.g., if $p^m \parallel I_{10}^{\min}(a, b, c)$ it often happens that $p^{m+10} \mid \Delta_C$ (see [25] for local results). On the other hand, the prime powers occurring in N_C are often smaller than those in Δ_C. E.g., if $p^3 \parallel \Delta_C$ then necessarily $p^2 \parallel N_C$ or $p \nmid N_C$. Some preliminary investigations suggest that the latter issue has more impact, and asymptotic counts of curves by $I_{10}^{\min}$ may essentially be lower bounds (up to logarithmic factors) of counts by conductor. This leads us to ask:

Question 3.4. *Let $C_2^{\mathrm{tw}}(D; X)$ be the number of quadratic twist classes of non-CM weight 2 newforms of minimal level $N < X$ with rationality field $\mathbb{Q}(\sqrt{D})$. Is $C_2^{\mathrm{tw}}(D; X) \gg X^{\alpha-\varepsilon}$ for α such that $\#Z_D(T) \gg T^{\alpha/5}$? Note that can take $\alpha = \frac{3}{5}, \frac{3}{10}, \frac{2}{5}, \frac{1}{5}, \frac{1}{5}$ for $D = 5, 8, 12, 13, 17$.*

Note that all of these exponents are less than the exponent of $2/3$ from the $d = 2$ case of Conjecture 2.4. We will compare these questionable lower bounds with our prime level data below.

Even if these lower bounds hold, there are several reasons why they may not be sharp. For one, there is the issue of Question 3.3. Perhaps most serious, there is the issue of how prime powers in N_C relate to prime powers in Δ_C mentioned above.

Another potential issue comes from the way we defined $Z_D(T)$: for the comparison with conductors, we are only interested in bounds on $I_{10}^{\min}$, and there may be many points with $I_{10}^{\min}$ small relative to $I_2^{\min}, I_4^{\min}, I_6^{\min}$. In particular, for $D = 12$, if one views $I_{2j}(a, b, c)$ as a polynomial in b with $a = 0$ and c fixed, then the degrees are $2, 2, 4, 4$, so one gets at least $T^{5/2}$ points with $I_{10}^{\min} \ll T^{10}$, which is a better than the lower bound $\gg T^2$ in Proposition 3.2. These rational points correspond to the curve $m = 0$ on $Y_-(12)$, and numerical investigations suggests this is a Shimura curve parametrizing abelian surfaces with geometric endomorphism algebra the quaternion algebra of discriminant 6. Consequently, one might be able to take $\alpha = \frac{1}{2}$ in Question 3.4 when $D = 12$. However, it is not clear that there is a family of genus 2 curves with RM 12 over $\mathbb{Q}$ that would achieve $\alpha = \frac{1}{2}$.

3.3. **Lower bounds for quadratic fields.** There are several known families of genus 2 curves with RM 5 and RM 8. These can be used to give lower bounds on counting such curves with bounded discriminant, and therefore conductor. For a genus 2 curve $C/\mathbb{Q}$, let Δ_C denote the minimal integral discriminant.

Proposition 3.5. *The number of $\bar{\mathbb{Q}}$-isomorphism classes of genus 2 curves over $\mathbb{Q}$ with RM 5 (resp. RM 8) with $|\Delta_C| < X$ is $\gg X^{1/6}$ (resp. $\gg X^{1/7}$).*

Proof. First consider RM 5. Brumer exhibited a 3-parameter family of curves $C_{b,c,d}$ with RM 5 over $\mathbb{Q}$ (see [5] for an announcement and [17] for a proof), however it is not clear when two such curves are isomorphic, either over $\mathbb{Q}$ or $\bar{\mathbb{Q}}$. We consider the 1-parameter subfamily C_d with $b = c = 0$, which is given by

$$C_d : y^2 + (x^3 + x + 1)y = -dx^3 + x^2 + x.$$

This defines a genus 2 curve with RM 5 for all $d \in \mathbb{Z}$ (I_{10} is never 0 for $d \in \mathbb{Z}$), and the discriminant of this model, $(27d^3 - 81d^2 - 34d - 103)^2$, is degree 6 in d. Thus to complete the RM 5 case of the proposition, it suffices to show that the number of $C_{d'}$ isomorphic to a given C_d over $\bar{\mathbb{Q}}$ is finite and uniformly bounded.

Let I_{2j} (resp. I'_{2j}) denote the polynomial Igusa–Clebsch invariants for C_d (resp. $C_{d'}$) for $j = 1, 2, 3, 5$. Then I_4/I_2^2 and is an absolute invariant for C_d. Thus if C_d and $C_{d'}$ are $\bar{Q}$-isomorphic, one must have $F = (I'_2)^2 I_4 - I_2^2 I'_4 = 0$. Now $F = 0$ defines a union of 3 curves in the (d, d')-plane, none of which are of the form $d = d_0$. So the number of $C_{d'}$ which are $\bar{\mathbb{Q}}$-isomorphic to a fixed C_d is bounded by $\deg F$.

Now consider RM 8. Here we use Mestre's 2-parameter family $C'_{a,b}$ of genus 2 RM 8 curves over $\mathbb{Q}$ from [30]. Consider the subfamily $C'_b = C'_{2,b}$, which is given by

$$C'_b : y^2 = 7500x^5 + (-75b + 3400)x^4 + (-34b + 2283)x^3 + (-3b + 1111)x^2 + 177x + 9.$$

This is a genus 2 curve with RM 8 for $b \in \mathbb{Z} - \{-88, 112\}$, and the discriminant has degree 7 in b. One can complete the argument just as in the RM 5 case. □

Corollary 3.6. *The number of quadratic twist classes of weight 2 newforms with rationality field $\mathbb{Q}(\sqrt{5})$ (resp. $\mathbb{Q}(\sqrt{2})$) and minimal level $N < X$ is $\gg X^{1/3}$ (resp. $\gg X^{2/7}$).*

Proof. If C is a genus 2 curve with RM D over $\mathbb{Q}$ with nonsplit Jacobian, then modularity tells us that C corresponds to a weight 2 newform f of level $N \leq \sqrt{|D_C|}$ and rationality field $\mathbb{Q}(\sqrt{D})$. So it suffices to show that a positive proportion of the curves in the proofs have nonsplit Jacobian. For the RM 5 C'_d family above, one can check that the above model has discriminant coprime to 5 when $d \equiv 1 \bmod 5$. Moreover computing the L-polynomial shows the mod 5 Jacobian is nonsplit, whence the Jacobian of C_d over $\mathbb{Q}$ is nonsplit. Similarly, for the RM 8 family, the curve C'_b has nonsplit Jacobian when $b \equiv 1 \bmod 7$. □

Note that these lower bounds are significantly smaller than those in Question 3.4.

Remark 3.7.

(1) Under the Bateman–Horn conjecture [1], $27d^3 - 81d^2 - 34d - 103$ is prime for $\gg X/\log X$ integers $d \equiv 1 \bmod 5$. For such d, C_d has prime-squared discriminant and the associated newform f has prime conductor. Consequently, subject to this conjecture, the above argument shows there are $\gg X^{1/3}/\log X$ weight 2 newforms f with rationality field $\mathbb{Q}(\sqrt{5})$ and prime level $< X$. The analogous argument does not work for RM 8 as the discriminant of C'_b splits into linear factors over $\mathbb{Z}$.

(2) Elkies recently gave similar lower bounds for RM 5 curves with an Eisenstein congruence, but with the exponent $\frac{1}{3}$ replaced by $\frac{1}{4}$ [13].

(3) One should similarly be able to give explicit lower bounds for $D = 12, 13, 17$ (and some other D). Namely, [14] give infinite families of RM D genus 2 curves for various D, though without explicit models. One can use Mestre's algorithm to construct models, and then follow the above argument. However, this approach typically yields families

with discriminants of very large degree, so one likely needs some care and/or luck to get decent lower bounds this way.

3.4. **Data for counts by discriminant.** Here we present some data on how often newforms of small degree d have a given rationality field. For $d \le 6$ and prime level N between 10^4 and $2 \cdot 10^6$, the rationality fields K_f can be specified by their discriminant D. That is, for each such $D = \mathrm{Disc}(K_f)$ appearing, K_f is the unique totally real number field (up to isomorphism) of degree at most 6 and discriminant D. Total counts by discriminant are summarized in Table 1.1.

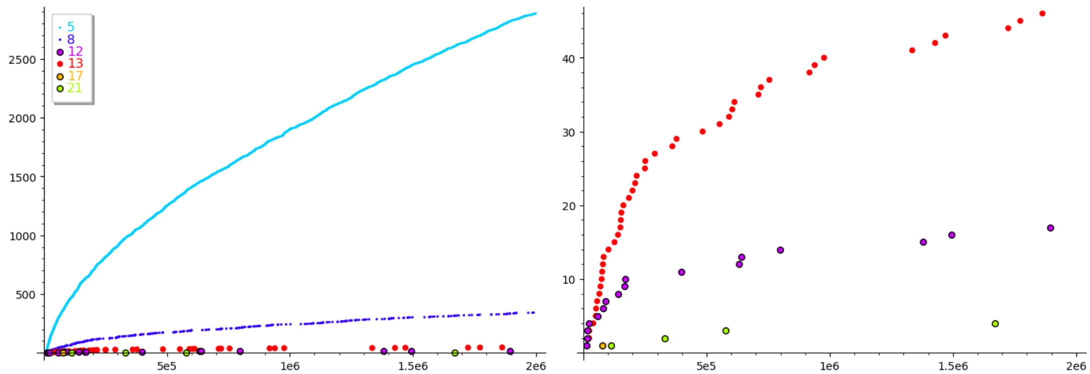

Figure 3.8. Counts of degree 2 forms by discriminant. The graph on the right excludes discriminants 5 and 8.

In Figure 3.8, we plot counts of degree 2 forms (up to Galois conjugacy) by their Hecke field discriminant. Note that, the growth rate of these count appears largest for $D = 5$, and then $D = 8$ appears the next largest. In particular, this suggests:

Conjecture 3.9. *Among prime levels $N \to \infty$, 100% of the degree 2 newforms in $S_2(N)$ have rationality field $\mathbb{Q}(\sqrt{5})$.*

The best fit to the data presented in Figure 3.8 by the model given in Section 2.2 yields exponents $b \approx 0.65$, 0.55, 0.47, and 0.42 respectively for $D = 5$, 8, 12, and 13. These are all higher than the lower bounds Question 3.4 asks about, those for discriminants 8 and 13 substantially so. There is only one form of discriminant 17 and prime level less than $2 \cdot 10^6$, at level 75653.

The heuristics from Question 3.4 as well as Corollary 3.6 support Conjecture 3.9. Namely, the suggested and proven lower bounds for counts by quadratic rationality field are largest for $\mathbb{Q}(\sqrt{5})$. Since these heuristics do not rely on a restriction to prime level, one is led to ask:

Question 3.10. *Do 100% of quadratic twist classes (ordered by minimal level) of weight 2 degree 2 non-CM newforms have rationality field $\mathbb{Q}(\sqrt{5})$?*

There is an arithmetic reason to expect a relative scarcity of certain quadratic fields in prime level compared to the lower bounds for arbitrary levels suggested in Question 3.4. Namely, if C is a genus 2 curve with RM D, then we typically expect odd primes dividing $I_{10}^{\min}$ to divide the conductor N_C. Based on the factorizations of I_{10}'s in Table 3.1, we expect $I_{10}^{\min}$ to be a 2-power times a prime power very infrequently for $D = 12, 13, 17$. Indeed, the LMFDB [26] lists 5485, 2189, 1230, and 1643 forms in all levels $N \le 10000$ for $D = 5$, 12, 13, and 17 respectively. Restricting to squarefree level, these numbers are 1820, 445, 319, and 461.

In Figure 3.11, we plot counts of degree 3 forms (up to Galois conjugacy) by their Hecke field discriminant. It appears that 100% of prime level degree 3 forms have rationality field $\mathbb{Q}(\zeta_7)^+$, the cubic field of discriminant 49. The cubic fields with the next highest growth rates appear to be those of discriminant 81 and 229, but it is not clear if one of these has a larger growth rate than the other. Note the field of discriminant 81 is $\mathbb{Q}(\zeta_9)^+$.

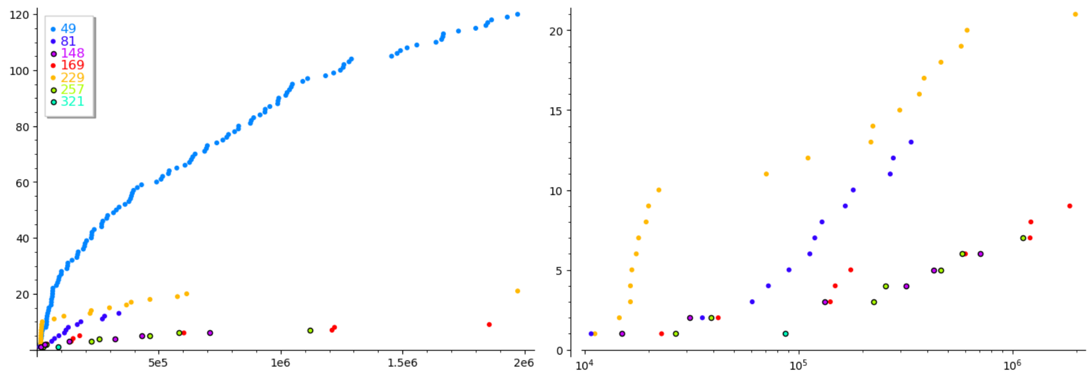

Figure 3.11. Counts of degree 3 forms by discriminant. The graph on the right excludes discriminant 49 and is on a log scale.

3.5. **Poisson model.** Fitting the data in Figure 3.8 to the model from Section 2.2 is natural, but we are also motivated to propose an alternative model for the following two reasons.

First, we're not sure how appropriate of a choice a least-squares fit of this type is. Least-squares fits maximize the probability that the best-fit parameters generated the observed data under the assumptions that the differences between the fit and the data are independent and normally distributed, with the same mean and variance for each data point. How reasonable in our setting all three of these assumptions are: that the deviations are independent, identically distributed, and normally distributed? With these assumptions, we did find in Section 2.2 best-fit parameters $(a, b) = (0.97, 0.833)$ for degree 1 forms that match what one obtains using the database [2], which is 1000 times larger.

Second, it's hard to get a reasonable estimate of the precision of these fits. For instance, we found above that forms of discriminant 12 and 13 had best-fit exponents 0.47 and 0.42 respectively, but Figure 3.8 leads one to expect that the exponent for discriminant 12 is smaller than that for discriminant 13. Are the numbers 0.47 and 0.42 given by this model given with enough precision to say with confidence that one is larger than the other? For comparison, fitting the data to $a\mathrm{li}(X^b)$, without taking logs as in (5), gives exponents of 0.40 and 0.38 instead.

We now present an alternative model which avoids the two concerns above. Let $Z_{a,b}(p)$ be a Poisson random variable with mean ap^b, and let $\mathcal{F}(d, \Delta, p)$ as in (1). Define

$$\mathrm{Prob}(\mathrm{data}_\Delta \mid a, b) := \prod_{\substack{10^4 < p < 2\cdot 10^6 \\ p \text{ prime}}} \mathrm{Prob}\Big(Z_{a,b}(p) = \#\mathcal{F}(d, \Delta, p)\Big). \tag{6}$$

For a given Δ, the *maximum likelihood estimate* $(\hat{a}, \hat{b})$ is the choice of (a, b) which maximizes $\mathrm{Prob}(\mathrm{data}_\Delta \mid a, b)$. For example, based on the Brumer–McGuinness–Watkins heuristic (3) we would expect $\hat{b} \approx -\frac{1}{6}$ for $\Delta = 1$. In Figure 3.12, we show, for $\Delta = 1$, 5, 8, and 49, the maximum likelihood estimates, along with, for $k = 1, 2, 3$, the regions

$$\left\{ (a, b) \, : \, \frac{\mathrm{Prob}(\mathrm{data}_\Delta \mid a, b)}{\mathrm{Prob}(\mathrm{data}_\Delta \mid \hat{a}, \hat{b})} > 10^{-k} \right\}. \tag{7}$$

By Bayes' theorem, these ratios of probabilities can inform one about the relative plausibility of the various choices of (a, b). They also form the basis of the *likelihood-ratio test*, also known as Wilks' test [44, 24, 23], which is widely used in statistics.

For discriminants 12 and 13, the maximum likelihood estimates are $(6.0, -0.85)$ and $(2.6, -0.70)$ respectively, but even the $k = 1$ regions from (7) are enormous, e.g. containing both $(0.1, -0.5)$ and $(200, -1.1)$ for discriminant 12.

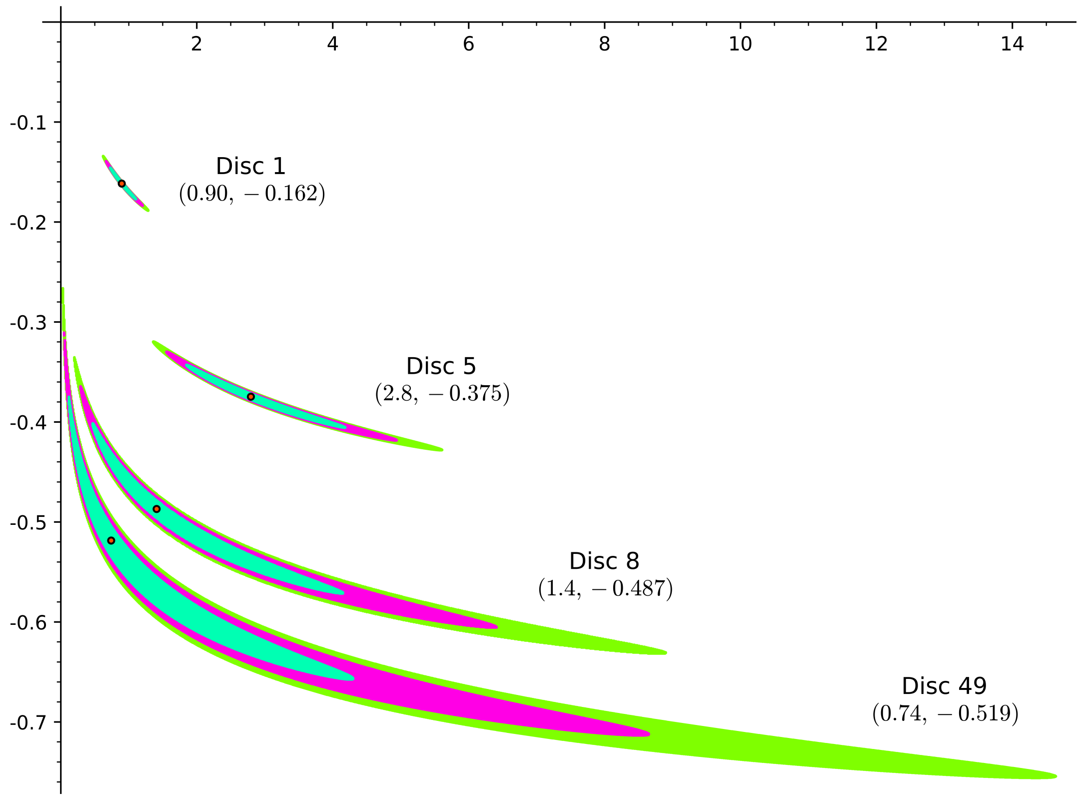

Figure 3.12. Poisson model (6) maximum likelihood estimates, and the regions (7)

4. Number of forms of same level

In this section we investigate how often multiple eigenform orbits appear at the same level. In our dataset, forms of discriminant 1 tend to congregate much more than they would if they behaved independently of one another, and there are known constructions which explain this behaviour. The forms of discriminants 5 and 8, for which no analogous constructions are known, tend to congregate an anomalous amount. More precisely, this section will give a test of the model we presented in Section 3.5, which was that $\#\mathcal{F}(d, \Delta, N)$ is Poisson distributed with mean aN^b for some a, b which depends only on Δ. Recall that $\mathcal{F}(d, \Delta, N)$, which was defined in (1), is the set of Galois orbits of degree d, discriminant Δ, and level N.

We begin by establishing some notation. Define

$$Q(\Delta, k) := \sum_{d=1}^{6} \#\{N : 10^4 < N < 2 \cdot N^6,\ N \text{ prime},\ \#\mathcal{F}(d, \Delta, N) = k\}. \tag{8}$$

The sum over d in the definition of $Q(\Delta, k)$ is only to allow us the notational convenience of omitting d; for all of the discriminants Δ we are considering, there is exactly one totally real number field with that discriminant and degree 6 or less.

For given $k \in \mathbb{Z}$ and $\lambda \in \mathbb{R}$, let $X(k, \lambda)$ be the Bernoulli random variable defined by

$$\begin{aligned}\operatorname{Prob}\big(X(k,\lambda) = 1\big) &= \operatorname{Prob}\big(\operatorname{Pois}(\lambda) = k\big)\\ \operatorname{Prob}\big(X(k,\lambda) = 0\big) &= 1 - \operatorname{Prob}\big(X(k,\lambda) = 1\big),\end{aligned}$$

where $\operatorname{Pois}(\lambda)$ denotes a Poisson random variable with mean λ.

The random variable

$$S_{a,b}(k) := \sum_{\substack{10^4 < p < 2\cdot 10^6 \\ p \text{ prime}}} X(k, ap^b)$$

follows a “Poisson binomial distribution”. By Le Cam’s theorem [12, §3.6], $S_{a,b}(k)$ is approximately Poisson distributed with mean

$$\mathbb{E}\big[S_{a,b}(k)\big] = \sum_{\substack{10^4 < p < 2\cdot 10^6 \\ p \text{ prime}}} \operatorname{Prob}\big(\operatorname{Pois}(ap^b) = k\big), \tag{9}$$

and the error term in this approximation is effective:

$$\begin{aligned}\sum_{j=0}^{\infty} \Big| \operatorname{Prob}\big(S_{a,b}(k) = j\big) - \operatorname{Prob}\Big(\operatorname{Pois}\big(\mathbb{E}\big[S_{a,b}(k)\big]\big) = j\Big) \Big| \\ < 2 \min\Big\{1, \mathbb{E}\big[S_{a,b}(k)\big]^{-1}\Big\} \sum_{\substack{10^4 < p < 2\cdot 10^6 \\ p \text{ prime}}} \operatorname{Prob}\big(\operatorname{Pois}(ap^b) = k\big)^2.\end{aligned} \tag{10}$$

Define $R_{a,b}(k)$ to be the quantity on the right hand side of the inequality above.

Let F_λ denote the cumulative distribution function of the Poisson distribution with mean λ, i.e.

$$F_\lambda(x) := \operatorname{Prob}\big(\operatorname{Pois}(\lambda) \le x\big).$$

For a given $x \in \mathbb{R}$, define

$$\rho(\lambda\,; x) := \begin{cases} F_\lambda(x) & x \le \lambda, \\ 1 - F_\lambda(x-1) & x > \lambda. \end{cases} \tag{11}$$

The quantity $\rho(\lambda\,; x)$ can be thought of as the probability that a sample of $\operatorname{Pois}(\lambda)$ will be “at least as extreme” as x.

Having established the notation above, we can now give a test for whether or not $\#\mathcal{F}(d, \Delta, N)$ appears to be Poisson distributed with mean aN^b. Supposing that it is, and additionally supposing that every level N behaves independently, the quantity $Q(\Delta, k)$ should be well modeled by $S_{a,b}(k)$. In Figure 3.12, we gave the maximum likelihood estimates $(\hat{a}, \hat{b})$ for discriminants 1, 5, 8, and 49. Using Le Cam’s theorem, the random variable $S_{\hat{a},\hat{b}}(k)$ is approximately poisson distributed with mean $\mathbb{E}[S_{\hat{a},\hat{b}}(k)]$, and the approximation error $R_{\hat{a},\hat{b}}(k)$ is given by (10). The quantity $\rho\big(\mathbb{E}\big[S_{\hat{a},\hat{b}}(k)\big]\,; Q(\Delta,k)\big)$ is a measure of how unlikely it would be to observe the data if the Poisson model was correct.

In Table 4.1, we list the values of

$$\begin{aligned} Q &:= Q(\Delta, k) & \mathbb{E} &:= \mathbb{E}\big[S_{\hat{a},\hat{b}}(k)\big] \\ \rho &:= \rho\big(\mathbb{E}\big[S_{\hat{a},\hat{b}}(k)\big]\,; Q(\Delta,k)\big) & R &:= R_{\hat{a},\hat{b}}(k). \end{aligned}$$

	$\Delta = 1$				$\Delta = 5$			
k	Q	$\mathbb{E}$	ρ	R	Q	$\mathbb{E}$	ρ	R
0	134363	133240.0	$1.1 \cdot 10^{-3}$	1.8	144876	144852.8	0.48	1.96
1	11922	13708.9	$3.3 \cdot 10^{-55}$	0.19	2772	2816.5	0.20	$4.7 \cdot 10^{-2}$
2	1038	727.5	$1.7 \cdot 10^{-27}$	$1.1 \cdot 10^{-2}$	54	34.3	$1.1 \cdot 10^{-3}$	$1.4 \cdot 10^{-3}$
3	302	26.8	$1.1 \cdot 10^{-200}$	$5.3 \cdot 10^{-4}$	2	0.38	$5.6 \cdot 10^{-2}$	$1.9 \cdot 10^{-5}$
4	60	0.77	$1.1 \cdot 10^{-89}$	$1.8 \cdot 10^{-5}$	0	$4.2 \cdot 10^{-3}$	0.99	$5.8 \cdot 10^{-9}$
5	15	$1.9 \cdot 10^{-2}$	$1.0 \cdot 10^{-38}$	$1.7 \cdot 10^{-8}$	0	$4.7 \cdot 10^{-5}$	0.9999	$1.3 \cdot 10^{-12}$
6	1	$4.1 \cdot 10^{-4}$	$4.1 \cdot 10^{-4}$	$1.3 \cdot 10^{-11}$				
7	2	$8.0 \cdot 10^{-6}$	$3.2 \cdot 10^{-11}$	$8.4 \cdot 10^{-15}$				
8	0	$1.5 \cdot 10^{-7}$	0.999999	$4.3 \cdot 10^{-18}$				
9	0	$2.5 \cdot 10^{-9}$	0.999999	$1.8 \cdot 10^{-21}$				
10	1	$4.0 \cdot 10^{-11}$	$4.0 \cdot 10^{-11}$	$6.4 \cdot 10^{-25}$				

	$\Delta = 8$				$\Delta = 49$			
k	Q	$\mathbb{E}$	ρ	R	Q	$\mathbb{E}$	ρ	R
0	147365	147362.3	0.50	2.0	147584	147584.0	0.50	2.0
1	336	341.1	0.40	$7.1 \cdot 10^{-3}$	120	119.9	0.51	$2.7 \cdot 10^{-3}$
2	3	0.61	$2.4 \cdot 10^{-2}$	$3.2 \cdot 10^{-5}$	0	$8.0 \cdot 10^{-2}$	0.92	$6.9 \cdot 10^{-7}$
3	0	$1.2 \cdot 10^{-3}$	0.999	$4.6 \cdot 10^{-10}$	0	$6.5 \cdot 10^{-5}$	0.999999	$1.6 \cdot 10^{-12}$

Table 4.1. $Q(\Delta, k)$ (8), $\mathbb{E}[S_{\hat{a},\hat{b}}(k)]$ (9), $\rho(\mathbb{E}\,; Q)$ (11), and $R_{\hat{a},\hat{b}}(k)$ (10) by discriminant Δ

Viewing the quantity ρ as one would a "p-value", i.e. a probability of a "null hypothesis" generating observed data which, when small, inclines one to reject the plausibility of that null hypothesis, Table 4.1 soundly rejects the hypothesis that counts of forms of discriminant 1 of a given level N are independently Poisson distributed with mean of the form aN^b. There were reasons a priori to expect this, however, stemming from the many constructions which produce multiple elliptic curves of the same conductor (e.g. [15]). For example, Elkies points out to us that in fact one expects $\gg X^{\frac{2}{3}-\varepsilon}$ prime levels up to X which contain 3 or more elliptic curves, contrasting with the prediction $\ll X^{\frac{1}{2}+\varepsilon}$ from the Poisson model being considered.

Table 4.1 also casts doubt on the same hypothesis for discriminants 5 and 8, where no constructions analogous to those for discriminant 1 are known. For $\Delta = 5$, we see that the most likely Poisson model produces on average 34.3 levels containing exactly two forms of discriminant 5, and in the dataset 54 such levels were observed. Le Cam's theorem leads to the approximation that, if the hypothesis was correct, one would observe 54 or more levels like this only 0.11% of the time, and approximating in this way induces an error of at most 0.14%. This is not absurdly infrequent, especially given that it is being considered along side many other similar table entries, yet it is conceivable that phenomena reminiscent of those known for degree 1 would exist here. Similarly, the $\Delta = 8$, $k = 3$ entry may be noteworthy in context, but on its own does not constitute sufficient evidence to confidently reject the hypothesis under consideration.

5. Distribution of Atkin–Lehner eigenvalues

In this section we investigate the statistics of the Atkin–Lehner signs of the eigenform orbits in our dataset. Define

$$\Delta_d(X) \coloneqq C_{d,+}(X) - C_{d,-}(X),$$

where $C_{d,\pm}(X)$ are as defined in (2). It is unclear how $\Delta_d(X)$ behaves as $X \to \infty$. It is widely expected that the very coarse bound $\Delta_d(X) = o(C_{d,\pm}(X))$ holds, and beyond this there seems to be no discussion about $\Delta_d(X)$ or similar in the literature. We ask about both the magnitude of $\Delta_d(X)$ and the sign. Table 5.2 lists the number of forms of a given degree and sign in our dataset. Figure 5.1 shows $\Delta_d(X) - \Delta_d(10^4)$, as well as $\Delta_1(X) - \Delta_1(10^4)$ after omitting all forms corresponding to Setzer–Neumann elliptic curves [33, 39]; we discuss these in more detail below.

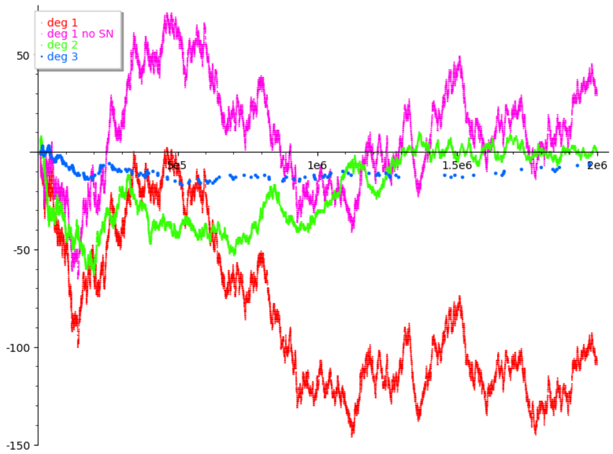

Figure 5.1. $\Delta_d(X) - \Delta_d(10^4)$

Degree	Sign	Count
1	$+1$	7570
	-1	7679
2	$+1$	1648
	-1	1648
3	$+1$	85
	-1	92
4	$+1$	4
	-1	7
5	$+1$	0
	-1	2
6	$+1$	0
	-1	1

Table 5.2. $C_{d,\pm}$ (2)

The approach we take in this section is to introduce a parameterized family of heuristic models for $\Delta_d(X)$, and then see how likely each of them would be to generate the data we observe. We will highlight three particular members of this family which seem quite natural yet lead to opposite predictions. We give a heuristic which predicts that there is a fundamental and rather small limit to how much models in our family can be distinguished, and that our dataset is close to this limit. Ultimately, we find that a bias towards Atkin–Lehner eigenvalue -1 is most likely in our dataset, but the opposite bias is only very slightly less likely.

For any $f \in \mathrm{Orb}(S_2(p))$ and $\beta \in \mathbb{R}$, let $Z_\beta(f)$ denote the random variable defined by

$$\mathrm{Prob}(Z_\beta(f) = 1) = \frac{\left(\dim S_2^+(p)\right)^\beta}{\left(\dim S_2^+(p)\right)^\beta + \left(\dim S_2^-(p)\right)^\beta},$$
$$\mathrm{Prob}(Z_\beta(f) = -1) = 1 - \mathrm{Prob}(Z_\beta(f) = 1).$$

Our model for $\Delta_d(X) - \Delta_d(10^4)$ is the random variable $\sum_{f \in \mathcal{F}(d)} Z_\beta(f)$, with $\mathcal{F}(d)$ defined by (1).

There are infinitely many models in our family which predict $\mathbb{E}[\Delta_d(X)] > \lambda C_d(X)^{\frac{1}{2}}$ for arbitrary $\lambda \in \mathbb{R}$, but our family does not contain any models which predict $\mathbb{E}[\Delta_d(X)] \gg C_d(X)^{\frac{1}{2}+\varepsilon}$. However, when $p = u^2 + 64$ for some $u \in \mathbb{Z}$, there is degree 1 form associated to the Setzer–Neumann elliptic curves [33, 39] with conductor p, and it is known that these curves always have rank 0 (and thus Atkin–Lehner eigenvalue -1). There are 138 forms corresponding to Setzer–Neumann curves in our dataset. We will call these *Setzer–Neumann forms* (SN). Because Setzer–Neumann forms always have sign -1, one might expect that there exists a $c > 0$ such that

$$\Delta_1(X) < -cX^{\frac{1}{2}-\varepsilon} \tag{12}$$

for any $\varepsilon > 0$ and sufficiently large X depending on ε. It is conceivable that there exists a family akin to that of Setzer–Neumann but with a preference for sign $+1$, which is at least as large as the Setzer–Neumann family. It is also conceivable that, even with SN forms omitted, Δ_1 has fluctuations of order large than $X^{\frac{1}{2}-\varepsilon}$ for all $\varepsilon > 0$. Either of these cases could lead to (12) not holding. We would be quite surprised if SN forms did not drive ultimately the growth of Δ_1 however.

Define

$$\mathrm{Prob}(\mathrm{data}_d \mid \beta) := \prod_{f \in \mathcal{F}(d)} \mathrm{Prob}\Big(Z_\beta(f) = w_p(f)\Big),$$

where $w_p(f)$ is the Atkin–Lehner eigenvalue of f, and $\mathcal{F}(d)$ is as defined in (1).

In Figure 5.4, we plot

$$\mathrm{Prob}(\mathrm{data}_d \mid \beta) \cdot 2^{\#\mathcal{F}(d)}$$

for $d = 1, 2$, as well as $d = 1$ after having omitted all Setzer–Neumann forms. We also show the maximum likelihood estimate for the parameter β and the associated likelihood. In Figure 5.5, we show a zoomed out version which includes $d = 3$. For $d = 3$, the maximum likelihood estimate $\hat{\beta}$ is 4.155, and $\mathrm{Prob}(\mathrm{data}_3 \mid \hat{\beta}) = 2.09 \cdot 2^{-177}$. A least-squares fit by functions of the form $\exp(-a\beta^2 + b\beta)$ yields the parameters in Table 5.3. These fits overlap completely with the data for all but $d = 3$, where the fit is roughly a factor of 2 too small at $\beta = 15$. Just as was the case in Section 3.4, one can make use of these likelihoods through Bayes' theorem and the likelihood-ratio test.

Degree	$\hat{a}$	$\hat{b}$
1	1.422	2.136
1 (no SN)	1.434	0.715
2	0.556	1.084
3	0.041	0.348

Table 5.3. Best fit of $\exp(-a\beta^2 + b\beta)$ to $\mathrm{Prob}(\mathrm{data}_d \mid \beta) \cdot 2^{\#\mathcal{F}(d)}$

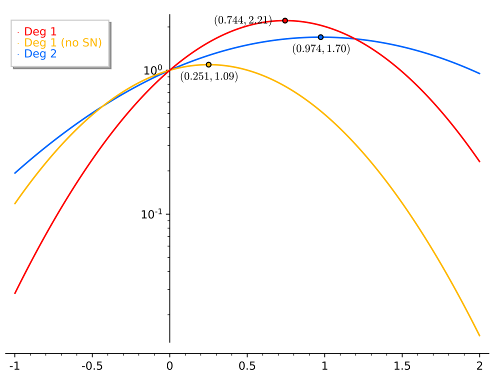

Figure 5.4. $\mathrm{Prob}(\mathrm{data}_d \mid \beta) \cdot 2^{\#\mathcal{F}(d)}$ and $\hat{\beta}$ for $d = 1$, 1 no SN, and 2

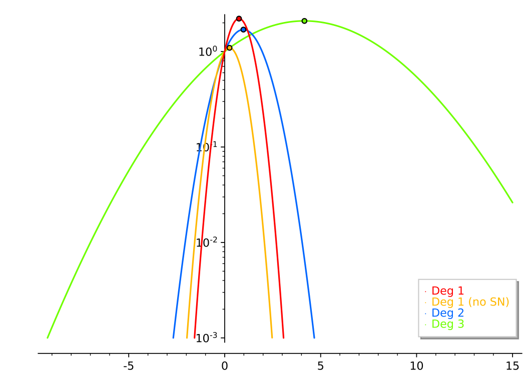

Figure 5.5. $\mathrm{Prob}(\mathrm{data}_d \mid \beta) \cdot 2^{\#\mathcal{F}(d)}$ and $\hat{\beta}$ for $d = 1$, 1 no SN, 2, and 3

Define

$$Z_\alpha^{(d)} := \Big(Z_\alpha(f) \;:\; f \in \mathcal{F}(d)\Big).$$

By the central limit theorem [12, Thm. 3.4.10], the distribution of the "log-likelihood" $\log(\mathrm{Prob}(Z_\alpha^{(d)} \mid \beta))$ converges to a normal distribution, and effective bounds on the rate of the convergence are known by the Berry–Esseen theorem [40, 12]. A straightforward calculation yields the following estimates for the mean and variance of this normal distribution:

$$\mathbb{E}\left[\log\left(\operatorname{Prob}(Z_\alpha^{(d)} \mid \beta)\right)\right] = \sum_{f\in\mathcal{F}(d)} \left(144\alpha\beta - 72\beta^2\right) \frac{\left(\dim S_2^-(N_f) - \dim S_2^+(N_f)\right)^2}{N_f^2} - \log 2 + \mathcal{O}\left(N_f^{-\frac{3}{2}}\right)$$

and

$$\operatorname{Var}\left[\log\left(\operatorname{Prob}(Z_\alpha^{(d)} \mid \beta)\right)\right] = \sum_{f\in\mathcal{F}(d)} 144\beta^2 \frac{\left(\dim S_2^-(N_f) - \dim S_2^+(N_f)\right)^2}{N_f^2} + \mathcal{O}\left(N_f^{-\frac{3}{2}}\right),$$

where N_f is the level of f.

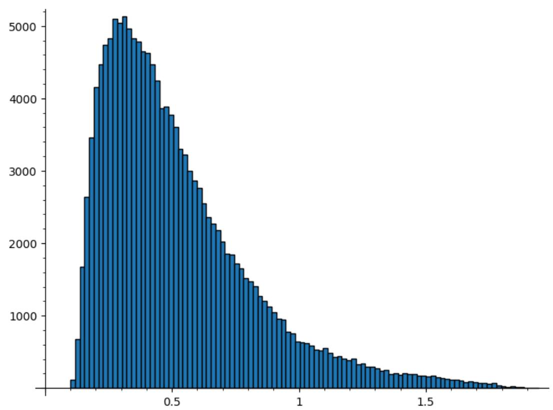

Figure 5.6. For $10^4 < p < 2\cdot 10^6$, histogram of $\dfrac{\dim S_2^-(p) - \dim S_2^+(p)}{\sqrt{p}}$

In [27], the second author gives formulas for $\dim S_2^\pm(p)$ in terms of class numbers of orders in imaginary quadratic fields. In particular,

$$\dim S_2^-(p) > \dim S_2^+(p), \text{ and}$$

$$\dim S_2^\pm(p) - \frac{1}{2}\dim S_2(p) \ll (\dim S_2(p))^{\frac{1}{2}+\varepsilon}.$$

In Figure 5.6, we plot the values of

$$\frac{\dim S_2^-(p) - \dim S_2^+(p)}{\sqrt{p}}$$

for p between 10^4 and $2\cdot 10^6$. The average value of this ratio in this range is about 0.52. For degrees d between 1 and 6, the values of

$$\sum_{f\in\mathcal{F}(d)} \frac{\left(\dim S_2^-(N_f) - \dim S_2^+(N_f)\right)^2}{N_f^2}$$

are given in Table 5.7.

One comfortably expects that $C_d(X) \ll X^{1-\delta}$ for some $\delta > 0$, causing the sums above to converge even when summed over all prime levels. Consequently, the distribution of log-likelihoods for fixed α and varying β will overlap for every α and β, i.e. there is a limit to how confidently one is able to reject the model Z_α in favour of Z_β as the potential source of the data, regardless of how much Atkin–Lehner eigenvalue data one compiles.

1	2	3	4	5	6
$1.99\cdot 10^{-2}$	$7.74\cdot 10^{-3}$	$6.08\cdot 10^{-4}$	$2.04\cdot 10^{-4}$	$5.24\cdot 10^{-6}$	$1.48\cdot 10^{-6}$

Table 5.7. $\displaystyle\sum_{f\in\mathcal{F}(d)} \frac{\left(\dim S_2^-(N_f) - \dim S_2^+(N_f)\right)^2}{N_f^2}$ by degree d

Without the Setzer–Neumann forms, the sum for degree 1 is $1.97\cdot 10^{-2}$. The sum over all primes between 10^4 and $2\cdot 10^6$ is 0.1617. For degrees 1 through 4, the median partial sum is more than 90% of the totals above.

We highlight three particular choices of the parameter β associated to plausible hypotheses for how Atkin–Lehner eigenvalues might be distributed:

The choice $\beta = 0$ models the situation where the sign of each form is ± 1 uniformly at random.

The choice $\beta = 1$ models the situation where the probability of a form having a given sign is proportional to the dimension of the associated eigenspace. If one chooses normalized eigenforms (not up to Galois conjugacy) uniformly at random, then it is tautologically true that $\beta = 1$ is the most appropriate choice. Thus, if $\beta = 1$ does not best model $\Delta_d(X)$, then the eigenforms of degree d have different sign statistics than eigenforms of degree not equal to d.

As discussed in Section 2.2, in our dataset we found that the number of forms of degree d with level less than X was, to good approximation, proportional to $\mathrm{li}(X^{\beta_d+1})$ with

$$\beta_1 \coloneqq -0.17, \quad \beta_2 \coloneqq -0.38, \quad \beta_3 \coloneqq -0.67.$$

If these exponents are a function only of Atkin–Lehner eigenspace dimension, then the appropriate choice of β in this model are these β_d's.

For each of $d = 1, 2, 3$, the model associated to $\beta = 1$ is the most likely among the three above to have generated the data. If one omits the Setzer–Neumann forms, then $\beta = 0$ is most likely to have generated the data, followed by $\beta = -0.17$. One does not know of analogues of the Setzer–Neumann forms in higher degrees, but it is conceivable that these would exist. However, for each degree, there is a large range of β for which the associated models could all easily have generated the observed data.

6. Lang–Trotter

In [32, §3], Kumar Murty outlines a generalization of the Lang–Trotter conjecture [22] to Fourier coefficients of newforms. For any algebraic integer $a \in K_f$, define

$$\pi_{f,a}(X) \coloneqq \#\{p < X \,:\, p \text{ prime}, a_f(p) = a\}.$$

The number of algebraic integers in K_f which satisfy the Weil bound for the p^{th} coefficient is on the order of $p^{\frac{d}{2}}$, where d is the degree of $K_f/\mathbb{Q}$. Thus, heuristically, one would expect that

$$\sum_{\substack{p < X \\ p \text{ prime}}} p^{-\frac{d}{2}} \ll \pi_{f,a}(X) \ll \sum_{\substack{p < X \\ p \text{ prime}}} p^{-\frac{d}{2}}$$

for non-CM newforms, so long as there are no local obstructions preventing $a_f(p)$ from taking the value a. Murty also considers the question of how often the values $a_f(p)$ lie in a fixed proper M subfield of K_f. Define

$$\pi_{f,M}(X) \coloneqq \#\{p < X \,:\, p \text{ prime}, a_f(p) \in M\}.$$

One can again give plausible heuristics for this quantity using the Weil bound. Murty conjectures that, if f is a normalized eigenform without CM, then there exist constants $c_{f,a}$ and $c_{f,M}$, possibly 0, such that

$$\pi_{f,a}(X) = (c_{f,a} + o(1)) \begin{cases} \frac{\sqrt{X}}{\log X}, & K_f = \mathbb{Q} \\ \log\log X, & [K_f : \mathbb{Q}] = 2 \\ 1 & \text{otherwise,} \end{cases}$$

and

$$\pi_{f,M}(X) = (c_{f,M} + o(1)) \begin{cases} \frac{X}{\log X}, & M = K_f \\ \frac{\sqrt{X}}{\log X}, & [K_f : \mathbb{Q}] = 2, M = \mathbb{Q} \\ \log\log X, & [K_f : \mathbb{Q}] = 3, M = \mathbb{Q} \\ \log\log X, & [K_f : \mathbb{Q}] = 4, [M : \mathbb{Q}] = 2 \\ 1 & \text{otherwise.} \end{cases}$$

The case $K_f = \mathbb{Q}$ above is the classical Lang–Trotter conjecture. Murty does not give explicit coefficients, but in the classical Lang–Trotter conjecture, they're given by an Euler product that essentially corresponds to the non-archimedean analogues of the Weil bound consideration alluded to above. In [42], Van Hirtum gives these infinite products explicitly for degree 2 forms, and investigates the conjecture computationally.

Given a form f of level N_f, define

$$X_f \coloneqq \max\left\{8196,\, 30\sqrt{N_f},\, \frac{N_f+1}{6}\right\}.$$

We computed $a_f(p)$ for all $p < X_f$. In this section, we use this data to investigate the behaviour of $\pi_{f,a}$, $\pi_{f,M}$, and

$$C_f(X;k) \coloneqq \#\{a \in K_f \,:\, \pi_{f,a}(X) = k\}.$$

We introduce $C_f(X;k)$ so that we can look at $\pi_{f,a}$ in more detail.

6.1. $\pi_{f,a}$. Murty conjectures that $\pi_{f,a}(X)$ is bounded for fixed f and a provided f is degree 3 or more. Among the forms in our dataset, we found $\pi_{f,a}(X_f) \le 3$ uniformly in f and a when $\deg f \ge 3$, and $\pi_{f,a}(X_f) \le 1$ uniformly for the 14 forms of degree 4, 5, and 6. We give exact counts in Table 6.1.

k	Deg 2	Deg 3	Deg 4	Deg 5	Deg 6
1		30	11	2	1
2	99	142			
3	1971	5			
4	1060				
5	142				
6	17				
7	5				
8	2				

Table 6.1. $\#\left\{f \,:\, \max_a \pi_{f,a}(X_f) = k\right\}$

The levels of the five forms of degree 3 with $\max_a \pi_{f,a}(X_f) = 3$ are 60209, 70843, 463963, 581657, and 1637617. The largest level among the 99 forms of degree 2 with $\max_a \pi_{f,a}(X_f) = 2$ is 313669, and they clearly cluster towards smaller levels. This is consistent with the conjecture that $\max_a \pi_{f,a}(X)$ is of order $\log\log X$, since we computed fewer Fourier coefficients for forms of smaller level.

6.2. $\pi_{f,M}$. When M is a proper subfield of K_f, Murty conjectures that $\pi_{f,M}(X)$ is bounded if f has degree 5 or more, as well as if f has degree 4 and $M = \mathbb{Q}$. The two forms of degree 5 in our dataset had no rational values of $a_f(n)$ other than $n = 0$ or 1. The form of degree 5 of largest prime level in the LMFDB, at level 6277, does have $a_2 = -1$. The unique form f_{171713} of degree 6 in our dataset had, for $M \coloneqq \mathbb{Q}[x]/(x^3 - x^2 - 4x - 1)$, that

$$a_{f_{171713}}(n) \in M, \quad n = 2, 4, 8, 16, \ldots, 16384,$$

and that, for all other $n \in [2, 28623]$, the coefficient $a_{f_{171713}}(n)$ does not lie in a proper subfield of $K_{f_{171713}} = \mathbb{Q}(\zeta_{13^5})^+$. There are 15 algebraic integers in $\mathbb{Q}(\zeta_{13^5})^+$ which satisfy the Weil bound at 2. Of those, 5 are rational, 2 generate $\mathbb{Q}(\sqrt{13})$, 4 generate the cubic subfield, and 4 generate $\mathbb{Q}(\zeta_{13^5})^+$. The ratio of the Plancherel measures [37] of these sets of integers is about 45 : 12 : 19 : 24.

In table Table 6.2, we list all cases in our dataset where the Fourier coefficient of a degree 4 form lies in a proper subfield of K_f. Underlines indicate values of n which are prime.

Level	Δ	$\mathrm{Gal}(K_f/\mathbb{Q})$	Largest n	$a_n \in \mathbb{Q}$	a_n quadratic
10169	8768	D_8	8196	$\underline{3}, \underline{5}$	$\underline{2}, \underline{7}, \underline{67}, \underline{97}, 1027, \underline{1237}, 3667, 5767$
13681	725	D_8	8196		$\underline{3}, 58, 152, 1057, 2407, 2482, 5644$
14759	725	D_8	8196		$\underline{3}, \underline{7}, 169, 332, 415, 1300, \underline{1531}$
28057	2777	S_4	8196	$\underline{2}$	No quadratic subfield
28789	1957	S_4	8196	$\underline{17}$	No quadratic subfield
35977	725	D_8	8196		$\underline{2}, \underline{5}, 49, 259, \underline{1103}, 1369, 1509, \underline{5987}, 6721$
63607	2777	S_4	10607	$\underline{2}$	No quadratic subfield
185599	1957	S_4	30968	$\underline{2}$	No quadratic subfield
264919	725	D_8	44158		$\underline{3}, 185, 952, 1255, 5795, 6314, 13784, 21392, 37249$
794111	725	D_8	132357	21	$15, 45, 140, 230, 1009^*, 10850, \underline{49033}, 85490, 110009$
1716109	725	D_8	286024		$22, 56, 297, 637, 11768, 12964, 16205, 22116, 26075$ $27645, 55120, 125446, 130405, 209076, 261345$

Table 6.2. Fourier coefficients of degree 4 forms in proper subfields of K_f, primes underlined

In Figure 6.3 on the right, we show a histogram of the values of $\pi_{f,\mathbb{Q}}(X_f)$ for forms of degree 3 in our dataset. Murty conjectures that $\pi_{f,\mathbb{Q}}(X)$ should be of order $\log\log X$ for these forms. Since $\log\log X_f$ is a bit less than 2.6 for the largest forms in our dataset we can't test this conjecture in any meaningful way. We do see that forms of larger discriminant tend to have larger values of $\pi_{f,\mathbb{Q}}(X_f)$, even considering that these forms have a strong tendency to appear in smaller levels, therefore having a smaller X_f. Heuristically, it is expected that forms of larger discriminant will have more rational a_p's, since the number of algebraic integers a in K_f which satisfy $\iota(a) < B$ for some fixed B and all real embeddings ι is inversely proportional to $\sqrt{\Delta_f}$, essentially by definition of the discriminant. This explains some of the behaviour visible in Figure 6.3, but not all of it.

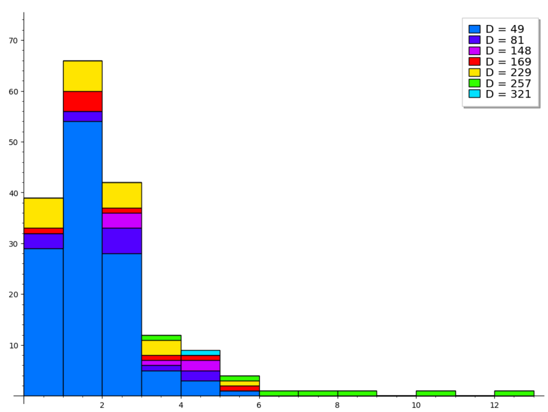

Figure 6.3. Histogram of $\pi_{f,\mathbb{Q}}(X_f)$ for degree 3 forms by discriminant

In the degree 2 case, the ratio $\pi_{f,\mathbb{Q}}(X)\pi(X)^{-\frac{1}{2}}$ is conjectured to approach a constant as X goes to infinity. Figure 6.4 shows that, for different f with X_f varying by roughly a factor of 20, these ratios at $X = X_f$ are of the same order of magnitude, supporting the conjecture.

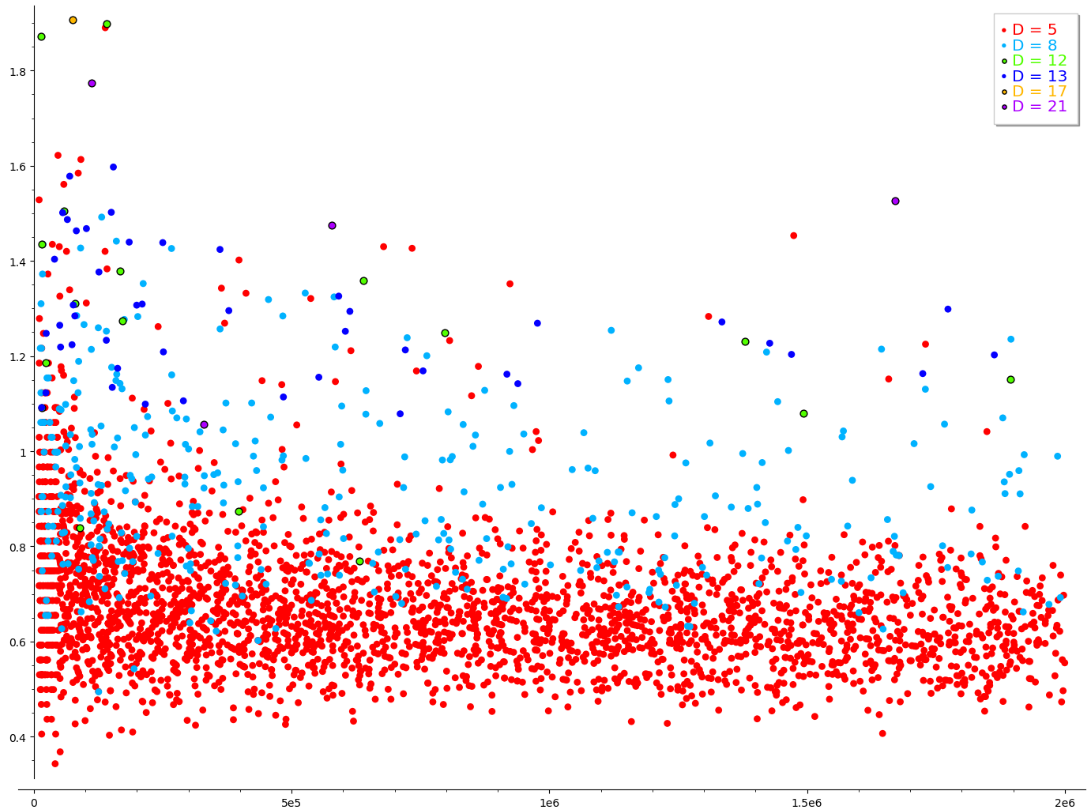

Figure 6.4. $\pi_{f,\mathbb{Q}}(X)\pi(X)^{-\frac{1}{2}}$ for degree 2 forms

6.3. $C_f(X;k)$**.** Recall that we defined

$$C_f(X;k) \coloneqq \#\{a \in K_f \,:\, \pi_{f,a}(X) = k\}.$$

We found in Section 6.1 that in our dataset all values of $\pi_{f,a}(X)$ were small for fixed f of degree at least 2, fixed a, and $X < X_f$. This motivates us to consider $C_f(X;k)$. After relaxing the condition that a must be fixed, we observe a wider range of behaviour. Fixing k comes at very little cost, as Table 6.1 shows. Lang–Trotter-type heuristics can also be applied directly to $C_f(X;k)$, producing questions in the same spirit.

It follows from the fact that $\pi_{f,a}(X_f) \leq 1$ for all a and for all f of degree 4 or more in our dataset that, for these forms, $C_f(X_f;1) = \pi(X_f)$ and $C_f(X_f;k) = 0$ for all $k \geq 2$. However, we also find it interesting to look at instances where $a_f(n) = a_f(m)$ without requiring m and n be prime. Many such instances will follow "trivially" from others. For example, if $a_f(n) = a_f(m)$, then $a_f(rn) = a_f(rm)$ when $\gcd(r, mn) = 1$. Cases where $a_f(n) = 0$ or 1 will also yield many trivial equalities. Below, for forms degree 5 and 6 in our dataset, we give a list of instances of $a_f(n) = a_f(m)$ from which all others can be easily inferred.

$$\begin{aligned}
a_{f_{26777}}(\underline{383}) &= a_{f_{26777}}(979), \quad & N(a_{f_{26777}}(\underline{383})) &= -3^2 \cdot 18211,\\
a_{f_{86161}}(122) &= a_{f_{86161}}(2535), \quad & N(a_{f_{86161}}(122)) &= 23^2,\\
a_{f_{171713}}(\underline{47}) &= a_{f_{171713}}(180), \quad & N(a_{f_{171713}}(\underline{47})) &= 13.
\end{aligned}$$

Underlines indicate primes, and subscripts the level of the form. The degree 6 form has level 171713.

For forms of degree 4, we find that $a_f(n) = a_f(m)$ for $m \neq n$ quite often. Below, for four forms of degree 4 of small level, we list values of $a_f(n)$ from which all equalities $a_f(n) = a_f(m)$ with $m, n < X_f$ can be deduced. We omit the subscript f, and instead mention the level of the form. This information for other forms of degree 4 in our dataset looks roughly similar to that of those presented here.

Level 10169 ($\Delta = 8768$):	Norm
$a(3) = 0$	
$a(5) = 1$	
$a(16) = a(2 \cdot 73) = 3$	
$a(28) = a(32)$	7^2
$a(50) = a(97)$	2^8
$a(125) = a(144) = a(2 \cdot 3^4 \cdot 7) = -9$	
$a(3^6) = a(2000) = -27$	
$a(17 \cdot 53) = a(1217)$	$-2^4 \cdot 7 \cdot 11^2$
$a(2089) = a(2 \cdot 7 \cdot 409)$	$2^4 \cdot 7 \cdot 353$

Level 13681 ($\Delta = 725$):	Norm
$a(2 \cdot 73) = a(2^2 \cdot 83)$	-2^4
$a(1163) = a(2^2 \cdot 3^3 \cdot 11)$	$5^2 \cdot 281$
$a(2^2 \cdot 3^3 \cdot 13) = a(2 \cdot 3 \cdot 17 \cdot 29)$	-5^2

Level 28057 ($\Delta = 2777$):	Norm
$a(2) = -1$	
$a(5 \cdot 23) = a(31 \cdot 37)$	$-2^3 \cdot 31$
$a(13 \cdot 67) = a(2 \cdot 5 \cdot 593)$	$-2^5 \cdot 23 \cdot 47$
$a(5^3 \cdot 23) = a(2 \cdot 5 \cdot 19 \cdot 31)$	$2^4 \cdot 31 \cdot 883$

Level 28789 ($\Delta = 1957$):	Norm
$a(17^2) = -1$	
$a(2^2 \cdot 3) = a(2 \cdot 7 \cdot 17^2)$	31
$a(2^3 \cdot 7) = a(2^2 \cdot 3 \cdot 5 \cdot 43)$	$3 \cdot 7 \cdot 31$
$a(7 \cdot 11) = a(2^5 \cdot 3 \cdot 43)$	$-3 \cdot 19 \cdot 31$
$a(109) = a(2^2 \cdot 5 \cdot 7)$	$-7 \cdot 31^2$
$a(2^2 \cdot 59) = a(2 \cdot 7 \cdot 37)$	$-31 \cdot 73$
$a(2 \cdot 5 \cdot 37) = a(3 \cdot 5^2 \cdot 79)$	$7 \cdot 73$
$a(2 \cdot 5 \cdot 43) = a(2^3 \cdot 17^2)$	$-3 \cdot 7$
$a(2 \cdot 3 \cdot 83) = a(2^3 \cdot 73)$	$3^2 \cdot 7 \cdot 43$

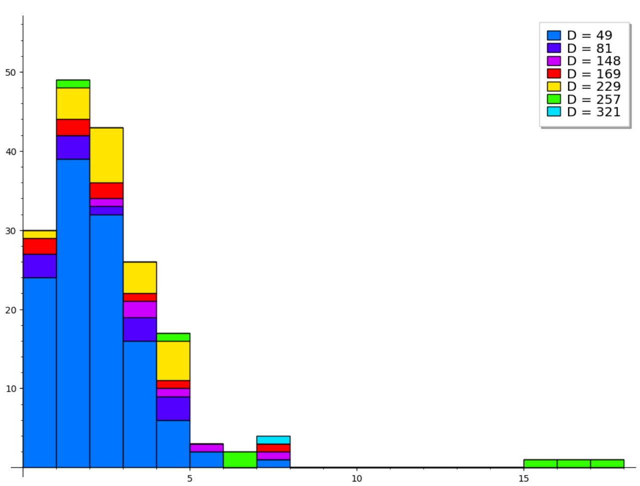

Figure 6.5. Histogram of $C_f(X_f; 2)$ for degree 3 forms by discriminant

n	Observed	Poisson
0	24	24.2
1	39	38.8
2	32	31.0
3	16	16.5
4	6	6.6
5	2	2.1
6	0	0.6
7	1	0.1
8	0	0.03

Table 6.6. Number of discriminant 49 forms with $C_f(X_f; 2) = n$, and expected number from best fit Poisson distribution

The histogram in Figure 6.5 shows the values of $C_f(X_f; 2)$ for the 177 forms of degree 3 in our dataset. The levels of the three forms of discriminant 257 with $C_f(X_f; 2) = 15$, 16, and 17 are 463249, 581657, and 1120969 respectively.

The distribution of the values of $C_f(X_f; 2)$ as f ranges across forms of degree 3 and discriminant 49 in our dataset is shown in Table 6.6. The mean value is 1.6, and the values are in very nearly Poisson-distributed with this mean (which is the maximum likelihood estimate), as we would expect from Le Cam's theorem [12, §3.6]. In contrast, the distribution for the forms of discriminant 257 is essentially impossible to explain in this way.

We found $C_f(X_f; 3) = 1$ for the five forms of degree 3 at the levels 60209, 70843, 463963, 581657, and 1637617, with discriminants 49, 229, 229, 257, and 49 respectively. As Table 6.1 implies, $C_f(X_f; 3) = 0$ for all other degree 3 forms in the dataset.

In Figure 6.7 we show, for all forms f of degree 2 in our dataset, the values $C_f(X_f; 2)$, plotted by level. The black curve is $\frac{1}{50}\pi(X_f)$.

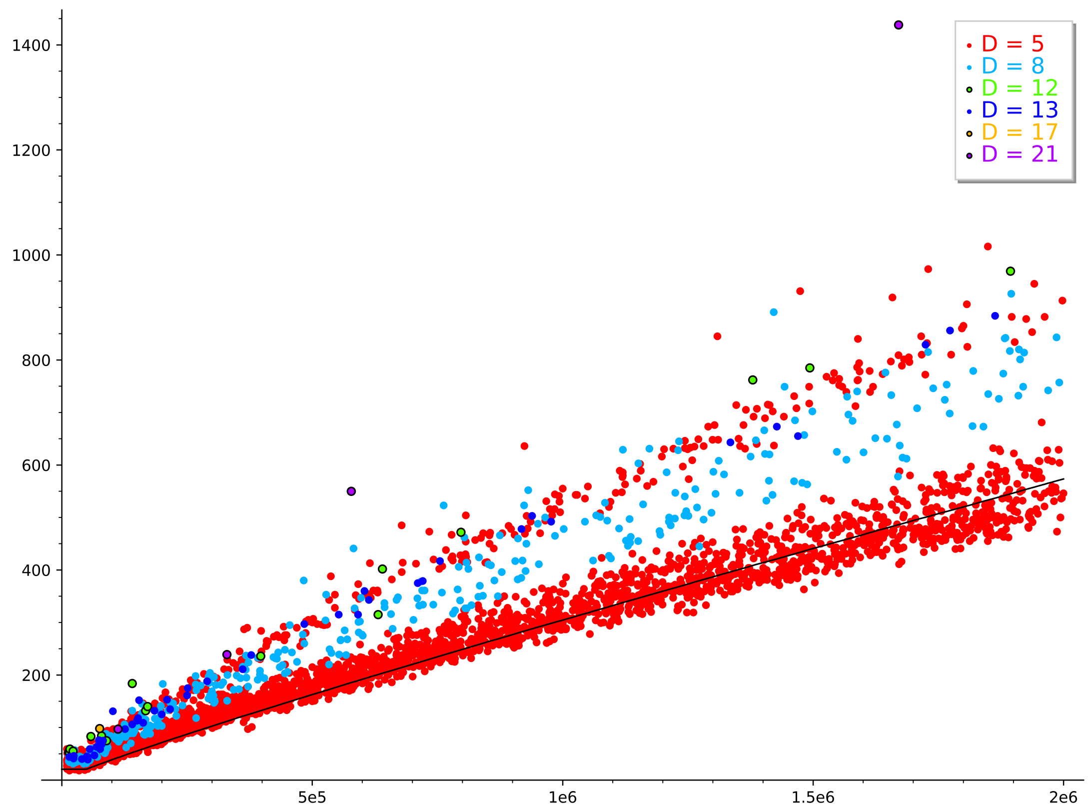

Figure 6.7. $C_f(X_f; 2)$ for degree 2 forms by discriminant

There is a clear separation between some of the forms of discriminant 5 and others. Broadly speaking, the Fourier coefficients of the forms in the upper group seem to always be constrained by a congruence of some sort, but these congruence conditions can come about in a variety of ways, and, for some of the forms, we were unable to find a convincing explanation for why they were in the upper group. We now give a few examples.

The form f of level 1997773 has the second largest level among all discriminant 5 forms in the dataset. It is in the upper group, and clearly visible in Figure 6.7. Noam Elkies suggested the following explanation. A Weierstrass model for the associated genus 2 curve is $C : y^2 = x^6 + 19x^5 + 44x^4 - 221x^3 + 25x^2 + 16x$. This clearly has a rational Weierstrass point. An RM 5 curve $C : y^2 = h(x)$, where $h(x)$ is sextic, has mod 2 Galois representation with image (isomorphic to $\mathrm{Gal}(h)$) contained in a transitive copy of A_5 in S_6. Having a rational Weierstrass point means the mod 2 Galois representation is contained in D_{10}, and that in turn implies that the residue class of a_p in $\mathbb{Z}[\frac{1+\sqrt{5}}{2}]/(2)$ is 1 if and only if $p = 2$. This restriction explains why $C_f(X_f; 2)$ is relatively large for this f. There are 301 discriminant 5 forms in our dataset for which $a_p = 1$ modulo 2 if and only if $p = 2$.

Not all of the forms in the upper group for discriminant 5 correspond to genus 2 curves with rational Weierstrass points. For instance, the forms at levels 35977, 49831, 86161, 1658417, and 1729823, some of which are quite special, have a variety of types of restrictions mod 2, and the form at level 1239943 has Fourier coefficients that are restricted mod 3.

There are 7 forms of degree 2 in our dataset which have an *Eisenstein congruence*, i.e. they are congruent mod ℓ to the unique weight 2 Eisenstein series of that level, for some prime ℓ. One form has discriminant 13, $\ell = \frac{1+\sqrt{13}}{2}$ (the prime above 3), and level 154333, and six have discriminant 5, $\ell = \sqrt{5}$, and levels 25951, 72931, 91331, 398011, 537241, and 923701. In Figure 6.7, the four of these with largest level are clearly visible above the bulk of the points. Eisenstein congruences correspond to reducible mod ℓ Galois representations, and satisfy

$$(p+1)^2 - (p+1)\mathrm{Tr}(a_p) + N(a_p) \equiv 0 \mod \ell \quad (p \neq N_f),$$

which is related to rational ℓ-torsion on the associated abelian variety [19]. Noam Elkies investigates Eisenstein congruences for genus 2 curves with RM 5 in [13].

References

[1] Paul T. Bateman and Roger A. Horn. A heuristic asymptotic formula concerning the distribution of prime numbers. *Math. Comp.*, 16:363–367, 1962. 2, 1

[2] Michael A. Bennett, Adela Gherga, and Andrew Rechnitzer. Computing elliptic curves over $\mathbb{Q}$. *Math. Comp.*, 88(317):1341–1390, 2019. 1, 2, 2.2, 3.5

[3] Alex J. Best, Jonathan Bober, Andrew R. Booker, Edgar Costa, John Cremona, Maarten Derickx, David Lowry-Duda, Min Lee, David Roe, Andrew V. Sutherland , and John Voight. Computing classical modular forms. *arXiv e-prints*, page arXiv:2002.04717, February 2020. 1

[4] Nils Bruin, E. Victor Flynn, Josep González, and Victor Rotger. On finiteness conjectures for endomorphism algebras of abelian surfaces. *Math. Proc. Cambridge Philos. Soc.*, 141(3):383–408, 2006. 3.1

[5] Armand Brumer. The rank of $J_0(N)$. Number 228, pages 3, 41–68. 1995. Columbia University Number Theory Seminar (New York, 1992). 3.3

[6] Armand Brumer and Oisín McGuinness. The behavior of the Mordell-Weil group of elliptic curves. *Bull. Amer. Math. Soc. (N.S.)*, 23(2):375–382, 1990. 2

[7] Alex Cowan. Computing newforms using supersingular isogeny graphs. *Res. Number Theory*, 8(4):Paper No. 96, 2022. 1

[8] Alex Cowan and Kimball Martin. Moduli for rational genus 2 curves with real multiplication for discriminant 5. arXiv 2206.05752. 3.2

[9] J. E. Cremona. *Algorithms for modular elliptic curves.* Cambridge University Press, Cambridge, 1992. 1

[10] Luis Victor Dieulefait, Ariel Pacetti, and Panagiotis Tsaknias. On the number of Galois orbits of newforms. *J. Eur. Math. Soc. (JEMS)*, 23(8):2833–2860, 2021. 2.1

[11] W. Duke and E. Kowalski. A problem of Linnik for elliptic curves and mean-value estimates for automorphic representations. *Invent. Math.*, 139(1):1–39, 2000. With an appendix by Dinakar Ramakrishnan. 2

[12] Rick Durrett. *Probability—theory and examples*, volume 49 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2019. Fifth edition of [MR1068527]. 4, 5, 6.3

[13] Noam Elkies. Genus-2 jacobians with a 5-torsion point, Preprint. 2, 6.3

[14] Noam Elkies and Abhinav Kumar. K3 surfaces and equations for Hilbert modular surfaces. *Algebra Number Theory*, 8(10):2297–2411, 2014. 3.2, 3

[15] Noam D. Elkies and Mark Watkins. Elliptic curves of large rank and small conductor. In *Algorithmic number theory*, volume 3076 of *Lecture Notes in Comput. Sci.*, pages 42–56. Springer, Berlin, 2004. 4

[16] Eyal Z. Goren. *Lectures on Hilbert modular varieties and modular forms*, volume 14 of *CRM Monograph Series*. American Mathematical Society, Providence, RI, 2002. With the assistance of Marc-Hubert Nicole. 3.1
[17] Ki-ichiro Hashimoto. On Brumer's family of RM-curves of genus two. *Tohoku Math. J. (2)*, 52(4):475–488, 2000. 3.3
[18] F. Hirzebruch and D. Zagier. Classification of Hilbert modular surfaces. In *Complex analysis and algebraic geometry*, pages 43–77. Iwanami Shoten, Tokyo, 1977. 3.2
[19] Nicholas M. Katz. Galois properties of torsion points on abelian varieties. *Invent. Math.*, 62(3):481–502, 1981. 6.3
[20] Chandrashekhar Khare and Jean-Pierre Wintenberger. Serre's modularity conjecture. I. *Invent. Math.*, 178(3):485–504, 2009. 3.1
[21] Koopa Tak-Lun Koo, William Stein, and Gabor Wiese. On the generation of the coefficient field of a newform by a single Hecke eigenvalue. *J. Théor. Nombres Bordeaux*, 20(2):373–384, 2008. 2.1
[22] Serge Lang and Hale Trotter. *Frobenius distributions in* GL_2*-extensions*. Lecture Notes in Mathematics, Vol. 504. Springer-Verlag, Berlin-New York, 1976. Distribution of Frobenius automorphisms in GL_2-extensions of the rational numbers. 1, 6
[23] Lucien Le Cam. *Asymptotic methods in statistical decision theory*. Springer Series in Statistics. Springer-Verlag, New York, 1986. 3.5
[24] Bing Li and G. Jogesh Babu. *A graduate course on statistical inference*. Springer Texts in Statistics. Springer, New York, 2019. 3.5
[25] Qing Liu. Conducteur et discriminant minimal de courbes de genre 2. *Compositio Math.*, 94(1):51–79, 1994. 3.2
[26] The LMFDB Collaboration. The L-functions and modular forms database. http://www.lmfdb.org, 2023. [Online; accessed 3 January 2023]. 1, 3.4
[27] Kimball Martin. Refined dimensions of cusp forms, and equidistribution and bias of signs. *J. Number Theory*, 188:1–17, 2018. 1, 2.1, ??
[28] Kimball Martin. An on-average Maeda-type conjecture in the level aspect. *Proc. Amer. Math. Soc.*, 149(4):1373–1386, 2021. 1, 2, 2.1
[29] Kimball Martin and Jordan Wiebe. Zeroes of quaternionic modular forms and central L-values. *J. Number Theory*, 217:460–494, 2020. 2.1
[30] J.-F. Mestre. Familles de courbes hyperelliptiques à multiplications réelles. In *Arithmetic algebraic geometry (Texel, 1989)*, volume 89 of *Progr. Math.*, pages 193–208. Birkhäuser Boston, Boston, MA, 1991. 3.3
[31] M. Ram Murty and Kaneenika Sinha. Effective equidistribution of eigenvalues of Hecke operators. *J. Number Theory*, 129(3):681–714, 2009. 2
[32] V. Kumar Murty. Frobenius distributions and Galois representations. In *Automorphic forms, automorphic representations, and arithmetic (Fort Worth, TX, 1996)*, volume 66 of *Proc. Sympos. Pure Math.*, pages 193–211. Amer. Math. Soc., Providence, RI, 1999. 1, 2.1, 6
[33] Olaf Neumann. Elliptische Kurven mit vorgeschriebenem Reduktionsverhalten. I. *Math. Nachr.*, 49:107–123, 1971. 5, 5
[34] Kenneth A. Ribet. Galois representations attached to eigenforms with Nebentypus. In *Modular functions of one variable, V (Proc. Second Internat. Conf., Univ. Bonn, Bonn, 1976)*, Lecture Notes in Math., Vol. 601, pages 17–51. Springer, Berlin, 1977. 3.1
[35] Kenneth A. Ribet. Abelian varieties over $\mathbf{Q}$ and modular forms. In *Modular curves and abelian varieties*, volume 224 of *Progr. Math.*, pages 241–261. Birkhäuser, Basel, 2004. 3.1
[36] David P. Roberts. Newforms with rational coefficients. *Ramanujan J.*, 46(3):835–862, 2018. 2.1, 2.3
[37] Peter Sarnak and Nina Zubrilina. Convergence to the plancherel measure of hecke eigenvalues, 2022. 2, 6.2
[38] Jean-Pierre Serre. Répartition asymptotique des valeurs propres de l'opérateur de Hecke T_p. *J. Amer. Math. Soc.*, 10(1):75–102, 1997. 2
[39] Bennett Setzer. Elliptic curves of prime conductor. *J. London Math. Soc. (2)*, 10:367–378, 1975. 5, 5
[40] I. G. Shevtsova. Refinement of estimates for the rate of convergence in Lyapunov's theorem. *Dokl. Akad. Nauk*, 435(1):26–28, 2010. 5
[41] Gerard van der Geer. *Hilbert modular surfaces*, volume 16 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1988. 3.2
[42] Jasper Van Hirtum. On the distribution of Frobenius of weight 2 eigenforms with quadratic coefficient field. *Exp. Math.*, 26(2):165–188, 2017. 6
[43] Mark Watkins. Some heuristics about elliptic curves. *Experiment. Math.*, 17(1):105–125, 2008. 2
[44] S. S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics*, 9(1):60 – 62, 1938. 3.5

Department of Mathematics, Harvard University, Cambridge, MA 02138 USA
Email address: `cowan@math.harvard.edu`

Department of Mathematics, University of Oklahoma, Norman, OK 73019 USA
Email address: `kimball.martin@ou.edu`