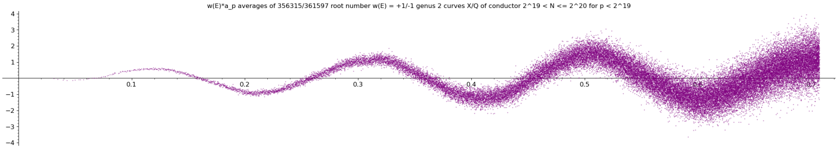
Genus 2 curves over $\mathbb{Q}$ of small conductor

Andrew V. Sutherland

Massachusetts Institute of Technology



The Simons Collaboration in Arithmetic Geometry, Number Theory, and Computation

(joint work with Andrew R. Booker)

Reasonable projects for the near future? (circa 1996)

Poonen's list of proposed projects for genus 2 curves X presented at ANTS II:

- Implement a polynomial-time algorithm to compute $\#X(\mathbb{F}_q)$. ✓ (Gaudry-Harley ANTS IV, Gaudry-Schost 2012)
- Devise and implement an algorithm to compute $\mathrm{End}(\mathrm{Jac}(X))$ for $X/\mathbb{Q}$. ✓ (Costa-Mascot-Sijsling-Voight 2019)
- Devise and implement an algorithm to compute $\mathrm{Jac}(X)_{\mathrm{tor}}$ for $X/\mathbb{Q}$. ✓ (Stoll 1999, . . ., Müller-Stoll 2024)
- Devise and implement algorithms to compute bounds on $\mathrm{rk}\,\mathrm{Jac}(X)(\mathbb{Q})$. ✓ (Stoll 2001, . . .)
- Automate the method of Chabauty–Coleman to compute $X(\mathbb{Q})$. ✓ (Balakrishnan 2006, . . .)

Reasonable projects for the near future?

Poonen's proposed projects for genus 2 curves X (continued):

- Extend Liu's conductor/reduction-type algorithm for $X/\mathbb{Q}$ to $p = 2$. ✓
  (Rüth-Wewers 2015, Bouw-Wewers 2017, Dokchitser-Doris 2018) ⚠️
- Given $X/\mathbb{Q}$, enumerate $X'/\mathbb{Q}$ with $\mathrm{Jac}(X') \sim \mathrm{Jac}(C)$. ✓
  (van Bommel-Chidambaram-Costa-Kieffer 2023 and work in progress)
- List all $X/\mathbb{Q}$ for which $\mathrm{Jac}(X)$ has good reduction away from 2.
  (not yet, but recent progress by Visser 2024)

and finally...

- Assemble a list of genus 2 curves over $\mathbb{Q}$ of small conductor analogous to ellitpc curve tables compiled by Birch, Swinnerton-Dyer, and Cremona.

Reasonable questions before embarking on such a project

Q: Why conductors?

A: The conductor is the fundamental invariant of the L-function $L(X,s)$; it measures its complexity and is the key parameter in its (conjectured) functional equation.

Q: Why L-functions?

A: Riemann, Birch and Swinnerton-Dyer, Sato-Tate, Lang-Trotter, Brumer-Stark, Brumer-Kramer, Langlands, murmurations, ..., these are all about L-functions.

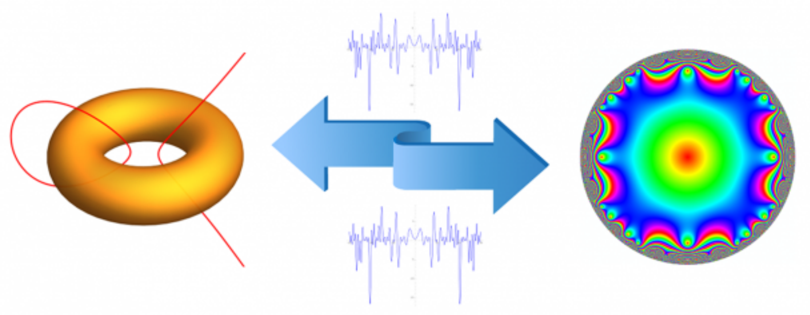
Q: Why *small conductors*?

A: Only L-functions of small conductor are computationally accessible.

Q: Doesn't the LMFDB already have a database of genus 2 curves of small conductor?

A: Only those with *small discriminant* (Booker-Sijsling-S-Yasaki-Voight ANTS XII).

Elliptic curves and their L-functions



Theorem (Eichler-Shimura, Langlands-Tunnell, Serre, Ribet, Wiles, Taylor-Wiles, Breuil-Conrad-Diamond-Taylor)

For each positive integer N, the set of L-functions $L(E, s)$ of elliptic curves $E/\mathbb{Q}$ of conductor N is equal to the set of L-functions $L(f, s)$ of newforms $f \in S_2^{\mathrm{new}}(\Gamma_0(N))$ of weight 2 and level N with rational q-expansions.

Automorphic forms associated to abelian surfaces

Type	Conductor	Curve Equation	Motive	Modular form
$\mathbf{A}[C_1]_{(s)}$	$277 = 277^1$	$y^2+(x^3+x^2+x+1)y = -x^2-x$	typical surface	paramodular form
$\mathbf{B}[C_1]_s$	$529 = 23^2$	$y^2+(x^3+x+1)y = -x^5$	surface with RM by $\mathbb{Q}(\sqrt{5})$ over $\mathbb{Q}$	CMF 23.2.1.a
$\mathbf{B}[C_1]_{ns}$	$294 = 2^1 3^1 7^2$	$y^2+(x^3+1) = x^4+x^2$	product of ECs 14a4 and 21a4 over $\mathbb{Q}$	CMFs 14.2.1.a and 21.2.1.a
$\mathbf{B}[C_2]_s$	$10368 = 2^7 3^4$	$y^2+x^2y = 3x^5-4x^4+6x^3-3x^2+1$	surface with RM by $\mathbb{Q}(\sqrt{2})$ over $\mathbb{Q}(\sqrt{2})$	HMF 162.1-a over $\mathbb{Q}(\sqrt{2})$
$\mathbf{B}[C_2]_{ngs}$	$1088 = 2^6 17^1$	$y^2+(x^3+x^2+x+1)y = x^4+x^3+2x^2+x+1$	Weil restriction of 17.1-a1 over $\mathbb{Q}(\sqrt{2})$	HMF 17.1-a over $\mathbb{Q}(\sqrt{2})$
$\mathbf{C}[C_2]_{(ns)}$	$448 = 2^6 7^1$	$y^2+(x^3+x)y = x^4-7$	product of PCM EC 32a3 and EC 14a6 over $\mathbb{Q}$	CMFs 32.2.1.a and 14.2.1.a
$\mathbf{D}[C_4]_{(s)}$	$3125 = 5^5$	$y^2+y = x^5$	surface with CM by $\mathbb{Q}(\zeta_5)$ over $\mathbb{Q}(\zeta_5)$	CM HMF 125.1-a over $\mathbb{Q}(\sqrt{5})$
$\mathbf{D}[D_2]_{(ns)}$	$8192 = 2^{13}$	$y^2 = x^6-9x^4+16x^2-8$	product of PCM ECs 32a3 and 256d1 over $\mathbb{Q}$	CMFs 32.2.1.a and 256.2.1.d
$\mathbf{E}[C_1]_{(ns)}$	$196 = 2^2 7^2$	$y^2+(x^2+x)y = x^6+3x^5+6x^4+7x^3+6x^2+3x+1$	square of EC 14a1 over $\mathbb{Q}$	CMF 14.2.1.a
$\mathbf{E}[C_2,\mathbb{C}]_{(ngs)}$	$576 = 2^6 3^2$	$y^2+(x^3+x^2+x+1)y = -x^3-x$	square of EC 9.1-a3 over $\mathbb{Q}(\sqrt{2})$	CMF 24.2.13.a
$\mathbf{E}[C_3]_{(ngs)}$	$324 = 2^2 3^4$	$y^2+(x^3+x+1)y = x^5+2x^4+2x^3+x^2$	square of EC 8.1-a1 over 3.3.81.1	CMF 18.2.13.a
$\mathbf{E}[C_4]_{(ngs)}$	$256 = 2^8$	$y^2+y = 2x^5-3x^4+x^3+x^2-x$	square of EC 1.1-a5 over 4.4.2048.1	CMF 16.2.5.a
$\mathbf{E}[C_6]_{(ngs)}$	$169 = 13^2$	$y^2+(x^3+x+1)y = x^5+x^4$	square of EC 1.1-a3 over 6.6.371293.1	CMF 13.2.4.a
$\mathbf{E}[C_2,\mathbb{R}\times\mathbb{R}]_s$	$455625 = 3^6 5^4$	$y^2+(x^3+x^2+x+1)y = x^5-3x^4-2x-1$	surface with QM $(D=6)$ over 2.0.3.1	BMF over 2.0.3.1 of level 50625
$\mathbf{E}[C_2,\mathbb{R}\times\mathbb{R}]_{ngs}$	$3969 = 3^4 7^2$	$y^2+(x^2+x+1)y = -3x^5+5x^4-4x^3+x$	Weil restriction of 441.2-a over 2.0.3.1	BMF 2.0.3.1-441.2-a
$\mathbf{E}[C_2,\mathbb{R}\times\mathbb{R}]_{ns}$	$675 = 3^3 5^2$	$y^2 = -x^6-14x^5-44x^4+28x^3-44x^2-14x-1$	product of ECs 15a2 and 45a2 over $\mathbb{Q}$	CMFs 15.2.1.a and 45.2.1.a
$\mathbf{E}[D_2]_s$	$20736 = 2^8 3^4$	$y^2 = -27x^6-54x^5-27x^4+18x^3+18x^2-2$	surface with QM $(D=6)$ over 4.0.576.2	HMF 324.1-b over $\mathbb{Q}(\sqrt{2})$
$\mathbf{E}[D_3]_s$	$34992 = 2^4 3^7$	$y^2 = -2x^6-6x^5+10x^3+9x^2-18x+6$	surface with QM $(D=6)$ over 6.0.2834352.2	BMF over 2.0.3.1 of level 3888
$\mathbf{E}[D_4]_s$	$20736 = 2^8 3^4$	$y^2+y = 6x^5+9x^4-x^3-3x^2$	surface with QM $(D=6)$ over 8.0.339738624.10	BMF over 2.0.3.1 of level 2304
$\mathbf{E}[D_6]_s$	$8100 = 2^2 3^4 5^2$	$y^2+x^3y = x^6+3x^5-42x^4+43x^3+21x^2-60x-28$	surface with QM $(D=6)$ over degree 12 field	BMF over 2.0.3.1 of level 900
$\mathbf{E}[D_2]_{ngs}$	$6400 = 2^8 5^2$	$y^2 = 2x^5+5x^4+8x^3+7x^2+6x+2$	square of EC 256.1-a1 over $\mathbb{Q}(\sqrt{5})$	HMF 2.2.5.1-256.1-a
$\mathbf{E}[D_3]_{ngs}$	$2187 = 3^7$	$y^2+(x^3+1)y = -1$	square of EC over 6.0.177147.2	BMF over 2..0.3.1 of level 243
$\mathbf{E}[D_4]_{ngs}$	$3600 = 2^4 3^2 5^2$	$y^2+x^2y = x^5-3x^4+11x^2-16x$	square of EC over 4.0.13500.2	BMF over $\mathbb{Q}(i)$ of level 225
$\mathbf{E}[D_6]_{ngs}$	$3600 = 2^4 3^2 5^2$	$y^2+x^3y = 14x^3-20$	square of EC over 6.0.7200000.1	BMF over 2.0.3.1 of level 400
$\mathbf{F}[D_2,C_2,\mathcal{H}]_{ngs}$	$576 = 2^6 3^2$	$y^2+x^3y = 5x^3-2$	square of PCM EC 1.1-a2 over $\mathbb{Q}(\sqrt{6})$	CM HMF 1.1-a over $\mathbb{Q}(\sqrt{6})$
$\mathbf{F}[C_2,C_1,\mathrm{M}_2(\mathbb{R})]_{ns}$	$729 = 3^6$	$y^2+y = -48x^6+15x^3-1$	square of PCM EC 27.a4 over $\mathbb{Q}$	CM CMF 27.2.1.a

One page of the “giant table” [Booker-Sijsling-S-Voight-Yasaki 2024?]

Enumerating elliptic curves by conductor

To enumerate $E/\mathbb{Q}$ by conductor we may proceed as follows:

1. Prove modularity (this step is optional and may be (was) deferred).
2. Enumerate rational newforms $f \in S_2^{\mathrm{new}}(\Gamma_0(N))$ for $N = 1, 2, 3, \ldots$
3. Use Eichler-Shimura to get an isogeny class representative A_f for each f.
4. Fill out isogeny classes by finding all the elliptic curves $E/\mathbb{Q}$ isogenous to E_f.

For $N \leq 500000$ this yields 3064705 elliptic curves with 2164260 distinct L-functions.

Each one of these steps is substantially more difficult for $g > 1$, even for $g = 2$.

Lots of recent progress on steps 1 (BGCP) and 4 (vBCCK), we seem to be stuck on step 2. And even if we were to get unstuck, there is no step 3 (not even in principle).

Alternatively, one can use fast Thué-Mahler solvers (BGR, GS) to enumerate all elliptic curves with discriminant supported on a given set of primes: $N \leq 10^6$ coming soon!

But this approach is particular to equations of degree 3 and 4, and even if we could extend them to degrees 5 and 6, enumerating curves by discriminant won't work.

Challenges in dimension two

We currently have nothing close to the abelian surface equivalent of even the 1972 Antwerp tables of elliptic curves. We know only the first 36 modular abelian surface L-functions unconditionally, of which 5 are typical (the 1972 Antwerp tables had 749).

- Enumerating paramodular forms of a given level is very difficult; even counting them is hard, due to the lack of dimension formulae. We have provably complete lists of paramodular forms only up to level 353 (five of them).
- Computing the L-function of a given paramodular form is very difficult; it is typically only feasible to compute a handful of Hecke eigenvalues (not enough!).
- **There is no analog of Eichler-Shimura for paramodular forms**.
- Not all abelian surfaces over $\mathbb{Q}$ are Jacobians of genus 2 curves over $\mathbb{Q}$. One can generically represent an abelian surface as a projective variety in $\mathbb{P}^{15}$ defined by 72 quadratic forms, but this is not a particular pleasant thing to do.
- There is no algorithm known to enumerate all genus 2 curves over $\mathbb{Q}$ of a given conductor. Even computing the conductor of a given curve is hard!

An axiomatic approach to arithmetic L-functions (FPRS)

An arithmetic L-function of motivic weight $w \in \mathbb{Z}_{\geq 0}$ with field of coefficients K is a Dirichlet series $L(s) = \sum_{n\geq 1} a_n n^{-s}$ with $a_1 = 1$, $a_n \in \mathcal{O}_K$, $\mathbb{Q}(a_n) = K$ such that:

- Analytic continuation: $L_{\mathrm{an}}(s) := L(s + w/2)$ converges absolutely on $\mathrm{Re}(s) > 1$ and has a meromorphic continuation with finitely many poles, all on $\mathrm{Re}(s) = 1$.
- Functional equation: For some $N \in \mathbb{Z}_{<0}$, $\varepsilon \in \mathbb{C}$ and $\mu_i, \nu_j \in \mathbb{Z}$ or $\mu_i, \nu_j \in \frac{1}{2} + \mathbb{Z}$,

  $$\Lambda_{\mathrm{an}}(s) := \Gamma_{\mathbb{R}}(s+\mu_1)\cdots\Gamma_{\mathbb{R}}(s+\mu_{d_1})\Gamma_{\mathbb{C}}(s+\nu_1)\cdots\Gamma_{\mathbb{C}}(s+\nu_{d_2})L_{\mathrm{an}}(s)$$

  is bounded in vertical strips away from $\mathrm{Re}(s) = 1$ with $\boldsymbol{\Lambda(s) = \varepsilon N^{1-s}\bar{\Lambda}(1-s)}$. Here ε is the root number, N is the conductor, and $d = d_1 + 2d_2$ is the degree.
- Euler product: $L_{\mathrm{an}}(s) = \prod_p F_p(p^{-s})^{-1}$ where $F_p(z) = (1 - \alpha_{1,p}z)\cdots(1-\alpha_{d_p,p}z)$ with $d_p \leq d$ ($d_p = p$ if $p \nmid N$) and $|\alpha_{j,p}| = p^{-m_j/2}$ with $m_j \in \mathbb{Z}_{\geq 0}$, $\sum m_j \leq d - d_p$.
- Central character: There is a Dirichlet character χ of modulus N for which $F_p(z) = 1 - a_p Z + \cdots + (-1)^d \chi(p) z^d$ and $\chi(-1) = (-1)^{\sum \mu_j + \sum (2\nu_k + 1)}$.

An axiomatic approach to L-functions of abelian varieties over $\mathbb{Q}$

Fix a positive integer g. We shall consider arithmetic L-functions of degree $2g$, motivic weight 1, field of coefficients $\mathbb{Q}$, defined by an Euler product

$$L(s) := \sum_n a_n n^{-s} = \prod_p L_p(p^{-s})^{-1},$$

with $L_p \in \mathbb{Z}[T]$. We further assume that

- $\Lambda(s) := \Gamma_{\mathbb{C}}(s)^g L(s)$ is holomorphic on $\mathbb{C}$ and satisfies the functional equation

  $$\Lambda(s) = \varepsilon N^{1-s}\Lambda(2-s)$$

  with root number $\varepsilon = \pm 1$ and conductor N.

- the a_n are integers that satisfy $|a_n| \leq d_{2g}(n)\sqrt{n}$, where $d_r(n) = \sum_{n_1\cdots n_r = n} 1$.

Under the Hasse–Weil conjecture, every $A/\mathbb{Q}$ of dimension g has such an L-function.

Conductor bounds for abelian varieties over $\mathbb{Q}$

The Brumer–Kramer formula gives explicit bounds on the conductor exponents of abelian varieties $A/\mathbb{Q}$ as a function of the dimension g:

$$v_p(N) \le 2g + pd + (p-1)\lambda_p(d),$$

where $d = \lfloor \frac{2g}{p-1} \rfloor$ and $\lambda_p(d) = \sum i d_i p^i$, with $d = \sum d_i p^i$ with $0 \le d_i < p$.

g	$p=2$	$p=3$	$p=5$	$p=7$	$p>7$
1	8	5	2	2	2
2	20	10	9	4	4
3	28	21	11	13	6

For $g \le 2$ these bounds are tight (see www.lmfdb.org for examples).

An integral converse theorem for GL_2

Theorem (Dimitrov 2023)

Let K be a number field, $k, q \in \mathbb{Z}_{>0}$, $L(s) = \sum_{n\geq 1} a_n n^{-s}$ be an L-function with $a_1 = 1$, $qa_n \in \mathbb{Z}$ for $n \geq 1$, $a_n = O(n^{k-1})$, and $\widetilde{L}(s)$ any L-function. Suppose $L(s)$ and $\widetilde{L}(s)$ admit a holomorphic continuation to $\mathbb{C}$ that is bounded on vertical strips such that

$$\Lambda(s) = i^k N^{k/2-s} \widetilde{\Lambda}(k-s)$$

for some $N \in \mathbb{Z}_{>0}$, with $\Lambda(s) := \Gamma_{\mathbb{C}}(s)L(s)$ and $\widetilde{\Lambda}(s) := \Gamma_{\mathbb{C}}(s)\widetilde{L}(s)$.
Then $L(s) = L(f, s)$ and $\widetilde{L}(s) = L(f|_{W_N}, s)$ for some $f \in S_k(\Gamma_0(N))$.

Corollary

Every rational L-function of degree 2, conductor N, and motivic weight w with $L_\infty(s) = \Gamma_{\mathbb{C}}(s)$ is the L-function of a newform in $S_k^{\mathrm{new}}(\Gamma_0(N))$ with $k = w + 1$.
If $w = 1$, it is also the L-function of an elliptic curve of conductor N.

Remark: The analogue for degree 4 L-functions with $w = 1$ is false (but almost true).

A finite problem

Let $\mathcal{S}(g, N, \varepsilon)$ denote the set of L-functions $L(s)$ that satisfy our axioms for a particular choice of $g, N \in \mathbb{Z}_{>0}$ and $\varepsilon = \pm 1$.

We expect every $L \in \mathcal{S}(g, N, \varepsilon)$ to be the L-function of a g-dimensional $A/\mathbb{Q}$ (this is far beyond anything we can currently hope to prove, but we don't need to).

Shafarevich's conjecture (proved by Faltings), then implies that $\mathcal{S}(g, N, \varepsilon)$ is finite. Moreover there is an effectively computable $n_0 = O(\sqrt{N})$ for which the coefficients $a_1, \ldots, a_{n_0}$ uniquely determine each $L \in \mathcal{S}(g, N, \varepsilon)$ (and $n_0 = O(\log^2 N)$ under GRH).

We seek an algorithm that takes inputs g, N, ε, determines a suitable n_0, and then outputs a list of distinct tuples $(a_1, \ldots, a_{n_0})$, one for each $L \in \mathcal{S}(g, N, \varepsilon)$. See Booker and Farmer–Koutsoliotas–Lemurell for prior work in this direction.

Our plan: Compute $\mathcal{S}(g, N, \varepsilon)$ via linear algebra, then search for corresponding $A/\mathbb{Q}$.

Our plan depends crucially on being able to compute $\mathcal{S}(g, N, \varepsilon)$ exactly. This not only tells us when to stop searching, knowing $a_1, \ldots, a_{n_0}$ helps us search.

The approximate functional equation

Fix g, N, ε. For each nonnegative integer k we define $S_k(x) := \sum_n f_k(n/x) a_n/n$, where

$$f_k(x) := \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} (s-1)^k \Gamma_{\mathbb{C}}(s)^g x^{1-s}\, ds.$$

The functional equation then implies the identity

$$S_k(x) = \varepsilon(-1)^k S_k(N/x),$$

valid for all $x > 0$; this is the approximate functional equation.If we choose k so that $(-1)^k = -\varepsilon$ and put $x = \sqrt{N}$ we obtain a nontrivial linear constraint on the a_n:

$$\sum_n \frac{a_n}{n} f_k(n/\sqrt{N}) = 0. \tag{1}$$

The $O(\sqrt{n})$ bounds on a_n and rapid decay of $f_k(x)$ allows us to compute an interval $I_{k,m}$ containing the truncated sum in (1) for $n \le m$ that does not depend on the a_n.

A system of linear constraints

For each $k \geq 0$ of the correct parity (meaning $(-1)^k = -\varepsilon$), we have linear constraints

$$\sum_{n \leq m} f_k(n/\sqrt{N}) a_n/n \in I_{k,m}.$$

These become less useful as k grows, so we restrict to $k = O(N^{1/4})$.
We also have the constraints $|a_n| \leq d_{2g}(n)\sqrt{n}$ for $n \geq 1$.

If we further assume that the $L \in \mathcal{S}(g, N, \varepsilon)$ are automorphic (which we do), we can obtain additional constraints by twisting $L(s)$ by a Dirichlet character $\chi\colon \mathbb{Z} \to \mathbb{C}$, equivalently, taking the Rankin-Selberg convolution of $L(s)$ with $L(\chi, s)$.

This generally increases the conductor and widens the corresponding interval $I_{\chi,k,m}$, but for χ of small conductor q and small k we obtain useful constraints

$$\sum_{n \leq m} f_k(n/\sqrt{q^4 N}) \chi(n) a_n/n \in I_{\chi,k,m}.$$

Solving the system rigorously using the simplex method

The Euler product for $L(s)$ implies that the a_n are determined by the a_q for prime powers $q = p^e$ with $e \leq 2g$. In order to take advantage of this, and to obtain rigorous results using off-the-shelf simplex solvers with fixed precision, we proceed as follows.

Let $q \leq n_0 < m$ be a prime power. Assume we have recursively fixed values for $a_1, \ldots, a_{q-1}$ that we cannot rule out this sequence as a prefix of a feasible solution.

We now apply the simplex method to a system of linear constraints on variables $a_{q'}$, with q' ranging over prime powers $q \leq q' \leq m$, using the objective functions $\pm a_q$.

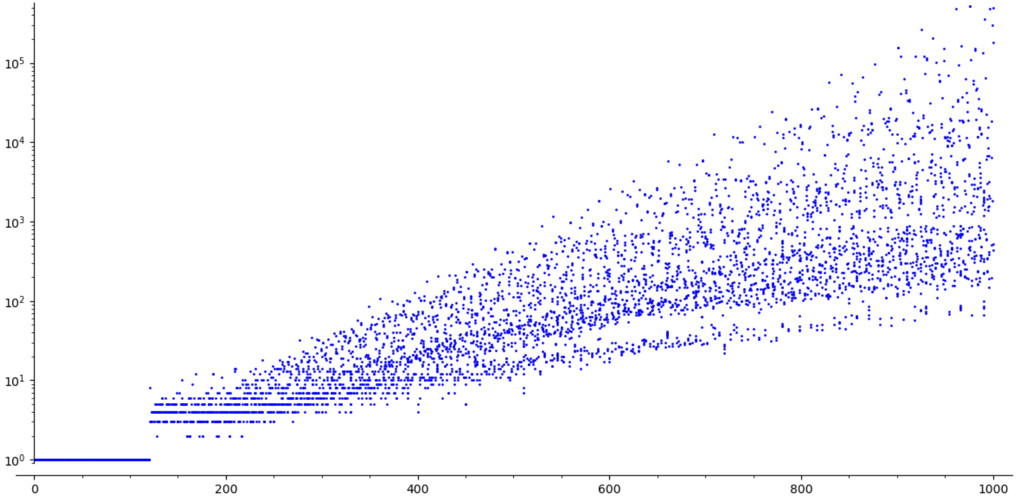
The dual solution yields a linear combination of constraints we can compute using interval arithmetic. Plugging in bounds on $a_{q'}$ yields an interval I_q containing a_q.

If $I_q \cap \mathbb{Z}$ is empty, then $a_1, \ldots, a_{q-1}$ is not the prefix of any $L \in \mathcal{S}(g, N, \varepsilon)$. Otherwise, for each $a \in I_q$ we add the tuple $(a_1, \ldots, a_{q-1}, a)$ to our list of feasible tuples.

We continue in this fashion until we run out of feasible prefixes or reach $q = n_0$.

Show, don't tell.

Timings



Proving completeness

If our algorithm outputs a nonempty list of feasible tuples $(a_1, \ldots, a_{n_0})$, the next step is to show there is at most one L-function in $\mathcal{S}(g, N, \varepsilon)$ for each prefix.

For this step, if we suppose that $(a_1, \ldots, a_{n_0})$ is the prefix of two distinct L-functions $L(s, \pi_1)$ and $L(s, \pi_2)$ of isobaric cuspidal automorphic representations of $\mathrm{GL}_{2g}(\mathbb{A}_{\mathbb{Q}})$ whose L-functions lie in $\mathcal{S}(g, N, \varepsilon)$. Using the Rankin–Selberg convolution L-function $L(s, \pi_1 \boxtimes \pi_2)$ we construct an inequality which will be violated if n_0 is sufficiently large.

If it is not violated, we increase n_0, extend our tuples, and try again.

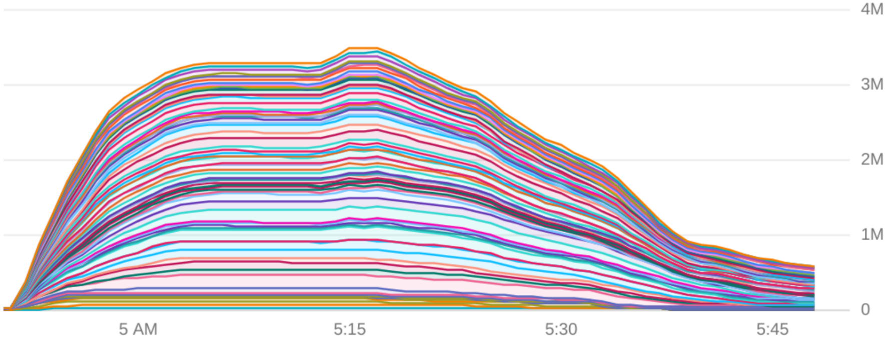
Eventually we obtain a list of distinct tuples $(a_1, \ldots, a_{n_0})$, each of which is provably the prefix of at most one automorphic L-function in $\mathcal{S}(g, N, \varepsilon)$.

This gives us an upper bound for our search that we expect to be tight.
Finding an abelian variety for each prefix proves completeness subject to modularity.

We then use Faltings-Serre, Boxer-Calegari-Gee-Pilloni, Calegari-Chidambaram-Ghitza, or other methods to prove modularity for individual abelian varieties.

Searching for genus 2 curves

Over the past several years we have conducted several searches for genus 2 curves of small conductor (including one last week!). Below is CPU histogram from a computation we ran in 2022 that enumerated more than 10^{19} genus 2 curves using a large parallel computation running on Google cloud platform.



We used a total of 4,034,560 Intel/AMD vCPUs in 73 data centers across the globe.

How much carbon does a 300 vCPU-year computation emit?

This is a question http://www.green-algorithms.org/ can help answer.

300 vCPU-years is about 1 314 900 core-hours (2 vCPUs per core).

CPU	cores	platform	location	energy	carbon
i9-9900K (64GB)	1	desktop	Massachusetts	46.99 MWh	19 750 Kg
i9-9900K (64GB)	16	desktop	Massachusetts	17,61 MWh	7 400 Kg
Ryzen 3990X (256GB)	64	desktop	Massachusetts	7.44 MWh	3 260 Kg
Ryzen 3990X (256GB)	64	cloud	Virginia	8.60 MWh	2 650 Kg
Ryzen 3990X (256GB)	64	cloud	Montreal	8.60 MWh	13 Kg

Searching for genus 2 curves

We found millions of genus 2 curves of small conductor, including the curve

$$C_{903} : y^2 + (x^2 + 1)y = x^5 + 3x^4 - 13x^3 - 25x^2 + 61x - 28$$

of conductor 903 whose L-function coefficients match those of the paramodular form of level 903 computed by Poor–Yuen that had not previously been matched.

We also found curves of conductor 657, 760, 775, 924 not previously known to occur, and many new genus-2 L-functions of small conductor:

conductor bound	1000	10000	100000	1000000
curves in LMFDB	159	3069	20265	66158
curves found	807	25438	447507	5151208
L-functions in LMFDB	109	2807	19775	65534
L-functions found	200	9409	212890	2426708

A provisional result

Provisional Theorem (proof in progress)

Assume the paramodular conjecture.
There are 456 L-functions of abelian surfaces $A/\mathbb{Q}$ *with conductor* $N \leq 1000$*, of which*

- *360 arise from products of elliptic curves over* $\mathbb{Q}$*;*
- *17 arise from weight-2 newforms with quadratic coefficients;*
- *2 arise from the Weil restriction of an elliptic curve over a quadratic field;*
- *77 arise from generic abelian surfaces, of which at least 67 are Jacobians.*

It may be feasible to remove the paramodular hypothesis (but that will depend largely on work by others, it is not a problem we are working on).

In addition to proving this theorem, we hope to extend it well beyond $N \leq 1000$.

Exploiting Galois representations

Let $A/\mathbb{Q}$ be an abelian surface of conductor N. For each $m \in \mathbb{Z}_{>1}$ we have a mod-m Galois representation

$$\rho_{A,m} \colon \operatorname{Gal}(\mathbb{Q}(A[m])/\mathbb{Q}) \to \operatorname{GSp}_4(\mathbb{Z}/m\mathbb{Z}).$$

For $p \nmid mN$ the charpoly $\chi_p \in (\mathbb{Z}/m\mathbb{Z})[T]$ of $\rho_{A,m}(\operatorname{Frob}_p) \in \operatorname{GSp}_4(\mathbb{Z}/m\mathbb{Z})$ satisfies

$$\chi_p(T) \equiv T^{2g} L_p(T^{-1}) \bmod \ell.$$

The m-torsion field $\mathbb{Q}(A[m])$ is unramified away from $p|mN$ and of degree at most $\#\operatorname{GSp}_4(\mathbb{Z}/m\mathbb{Z})$. For small m and N it is feasible to enumerate all such fields K and their associated mod-m GSp_4-representations, especially $m = 2$ and N a prime power.

Each representation yields mod-m congruence constraints on $L_p(T)$ for primes $p \nmid mN$. This dramatically reduces the amount of branching in our algorithm.

What I did over (the first few weeks of) my summer vacation

Last week we ran another search using completely new (128-bit) code that uses our L-functions-from-nothing approach to efficiently compute/bound conductors.

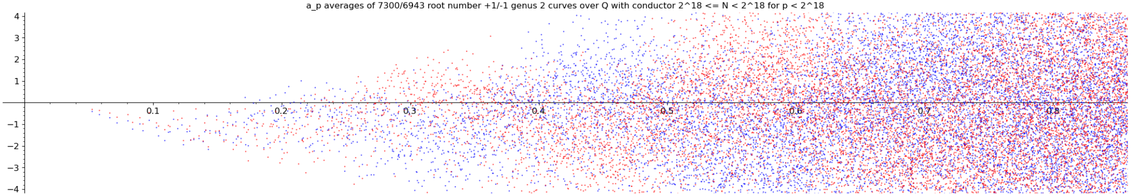
- We enumerated integral models $X\colon y^2 + h(x)y = f(x)$ with $h_i \in \{0,1\}$ and $\|f\| \le 90$ for which $\Delta_{\min}(X)$ is compatible with $\operatorname{cond}\operatorname{Jac}(X) \le 2^{20}$, ignoring prime-power factors of the form $p^{12a+10b}$ compatible with almost good reduction.
- Liu's `genus2red` algorithm (Pari/GP) to compute $\operatorname{odd}(N_{\min}) \le N_{\max} = 2^{20}$.
- Allombert's `lfungenus2` algorithm (Pari/GP) to compute degree-3 Euler factors with conductor exponent 1 and discriminant exponent at most 12.
- Maistret-S for Euler factors at primes of almost good reduction.
- Harvey-S average poly-time for Euler factors at good $p \le C\sqrt{N_{\max}} \approx 12{,}000$.
- Fast (milliseconds) heuristic L-function test iterating over $v_2(N_{\min})$.
- Slower (minutes) rigorous L-function test to rigorously compute $v_2(N_{\min})$ via `arb`.

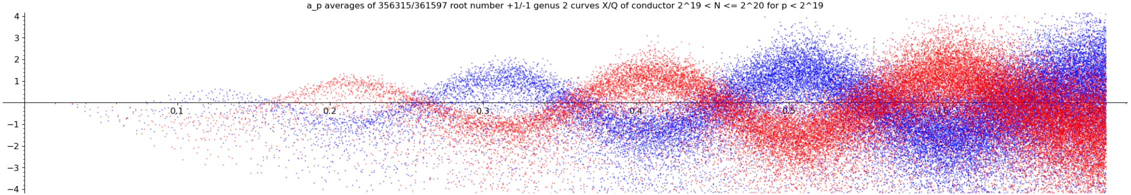
Some highlights

- About 250 nanoseconds per curve to enumerate $\approx 4 \times 10^{16}$ smooth curves (covering 10^{17}) and test their discriminants for compatibility with small conductor.
- Of these, roughly 5×10^9 (about $1/10^7$) have sufficiently smooth discriminants.
- Of these, roughly 7×10^8 (about $1/10$) have $\mathrm{odd}(N_{\mathrm{min}}) \leq 2^{20}$.
- Of these, roughly 6×10^7 (about $1/10$) have $N_{\mathrm{min}} \leq N_{\mathrm{max}}$.
- ≈ 7 million quadratic-twist-minimal curves (some are twists).
- ≈ 1.3 million twist-minimal isogeny classes.
- Twisting yields about 1.8 million isogeny classes, of which at least $200,000$ are new (smallest new conductor is 1343).
- Even using minimal twists, about 250,000 have a prime of almost good reduction that cannot be removed (this proportion will grow as we expand isogeny classes).

L-functions of genus 2 curves over $\mathbb{Q}$ with Sato-Tate group USp(4).

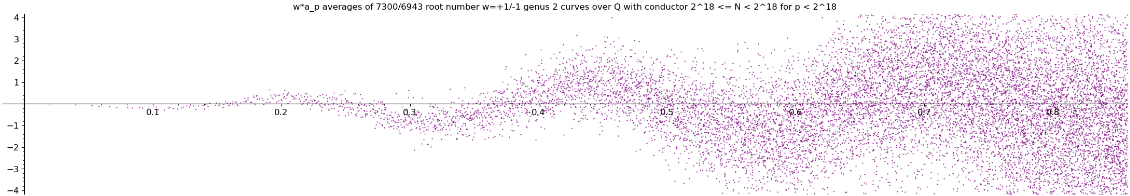
Before and after genus 2 murmuration plots (top LMFDB, bottom new dataset).

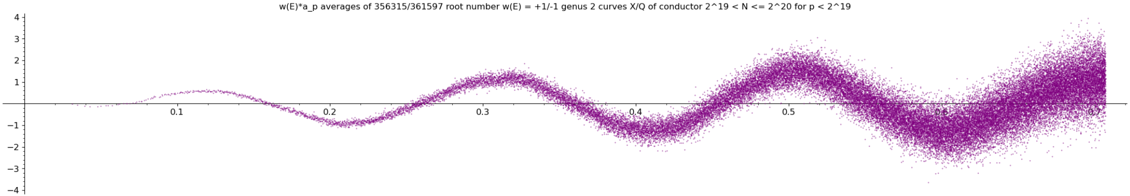




L-functions of genus 2 curves over $\mathbb{Q}$ with Sato-Tate group USp(4).

Before and after genus 2 murmuration plots (top LMFDB, bottom new dataset).





Sixteenth Algorithmic Number Theory Symposium

Massachusetts Institute of Technology
July 15–19, 2024

ANTS XVI

Also check out The Mordell conjecture 100 years later the week before, July 8–12.