#### Murmurations of Arithmetic L-functions

#### Andrew V. Sutherland

#### Massachusetts Institute of Technology Simons Collaboration in Arithmetic Geometry, Number Theory, and Computation

#### September 14, 2023



Joint work with Yang-Hui He, Kyu-Hwan Lee, Thomas Oliver, and Alexey Pozdnyakov, with thanks to Nina Zubrilina and Peter Sarnak, and to Eran Assaf, and also to Jonathan Bober, Andrew Booker, Min Lee, and David Lowry-Duda.

#### Elliptic curves and their L-functions

Let  $E/\mathbb{Q}$  be an elliptic curve, say  $E: y^2 = x^3 + Ax + B$  with  $A, B \in \mathbb{Z}$ . For primes  $p \nmid \Delta(E) := -16(4A^3 + 27B^2)$  this equation defines an elliptic curve  $E/\mathbb{F}_p$ . For all such primes p we have the trace of Frobenius  $a_p(E) := p + 1 - \#E(\mathbb{F}_p) \in \mathbb{Z}$ .

One can also define  $a_p(E)$  for  $p|\Delta(E)$ , and then construct the *L*-function

$$L(E,s) \coloneqq \prod_{p} (1 - a_p p^{-s} + \chi(p) p^{1-2s})^{-1} = \sum_{p} a_n n^{-s}$$

where  $\chi(p) = 0$  for p|N(E) and  $\chi(p) = 1$  otherwise and  $N(E)|\Delta(E)$  is the conductor.

But in fact the  $a_p$  for  $p \nmid \Delta(E)$  determine L(E, s) (via strong multiplicity one), and also the conductor and root number  $w(E) = \pm 1$ , which appear in the functional equation

$$\Lambda(E,s) = w(E)N(E)^{1-s}\Lambda(E,2-s)$$

where  $\Lambda(s) := \Gamma_{\mathbb{C}}(s)L(E,s)$ . The *L*-function L(E,s) determines the isogeny class of *E*.

Arithmetic statistics of Frobenius traces of elliptic curves over  $\mathbb{Q}$ 

Three conjectures from the 1960s and 1970s (the first is now a theorem):

- 1. **Sato-Tate**: The sequence  $x_p := a_p(E)/\sqrt{p}$  is equidistributed with respect to the pushforward of the Haar measure of the Sato-Tate group of *E* (typically SU(2)).
- 2. Birch and Swinnerton-Dyer:

$$\lim_{x\to\infty}\frac{\log x}{2\sqrt{x}}\sum_{p\leq x}\frac{a_p(E)}{\sqrt{p}}=\frac{1}{2}-r(E).$$

3. Lang-Trotter: For every nonzero  $t \in \mathbb{Z}$  there is a real number  $C_{E,t}$  for which

$$\#\{p \leq x : a_p(E) = t\} \sim C_{E,t} \frac{\sqrt{x}}{\log x}$$

These conjectures depend only on L(E, s) and generalize to other L-functions.

## Example: Elkies' curve of rank $\geq 28$ (= 28 under GRH).

al histogram of y^2 + xy + y = x^3 - x^2 - 20067762415575526585033208209338542750930230312178956502x + 34481611795030556467032985690390720374855944359319180361266008296291939448732243429 for p <= 2^10 172 data points in 13 buckets, z1 = 0.023, out of range data has area 0.250



Moments: 1 1.034 1.716 2.532 4.446 7.203 13.024 22.220 40.854 72.100 133.961

## Example: Elkies' curve of rank $\geq 28$ (= 28 under GRH).

al histogram of y^2 + xy + y = x^3 - x^2 - 20067762415575526585033208209338542750930230312178956502x + 34481611795030556467032985690390720374855944359319180361266008296291939448732243429 for p <= 2^40 41203088796 data points in 202985 buckets



### How rank effects trace distributions

One formulation of the BSD conjecture implies that

$$\lim_{x \to \infty} \frac{1}{\log x} \sum_{p \le x} \frac{a_p(E) \log p}{p} = -r + \frac{1}{2},\tag{1}$$

and sums of this form (Mestre-Nagao sums) are often used as a tool when searching for elliptic curves of large rank (which necessarily have large conductor N).<sup>1</sup><sup>2</sup>

#### Theorem (Kim-Murty 2023)

If the limit on the LHS of (1) exists then it equals the RHS with r the analytic rank, and the L-function of E satisfies the Riemann hypothesis.

<sup>&</sup>lt;sup>1</sup>See Sarnak's 2007 letter to Mazur.

<sup>&</sup>lt;sup>2</sup>See the recent paper of Kazalicki-Vlah for some recent machine-learning work on this topic.





#### Murmurations of elliptic curves

In their 2022 preprint *Murmurations of elliptic curves*, Yang-Hui He, Kyu-Hwan Lee, Thomas Oliver, and Alexey Pozdnyakov observed a curious fluctuation in average Frobenius traces of elliptic curves in a given conductor interval depending on the rank.



#### Murmurations of elliptic curves

Elliptic curve *L*-functions of conductor  $N \in (M, 2M]$  for  $M = 2^{11}, 2^{12}, \ldots, 2^{17}, 250000$ . The *x*-axis range is [0, 2M]. A blue/red or purple dot at  $(p, \bar{a}_p \text{ or } \bar{m}_p)$  shows the average of  $a_p$  or  $m_p := w(E)a_p(E)$  over even/odd or all  $E/\mathbb{Q}$  with  $N_E \in (M, 2M]$ .



#### Murmurations of elliptic curves

Elliptic curve *L*-functions of conductor  $N \in (M, 2M]$  for  $M = 2^{11}, 2^{12}, \ldots, 2^{17}, 250000$ . The *x*-axis range is [0, 2M]. A blue/red or purple dot at  $(p, \bar{a}_p \text{ or } \bar{m}_p)$  shows the average of  $a_p$  or  $m_p := w(E)a_p(E)$  over even/odd or all  $E/\mathbb{Q}$  with  $N_E \in (M, 2M]$ .



#### **Bias cancellation**

There is a negative bias in  $\bar{a}_p$  that is parity-independent and disappears in  $\bar{m}_p$ . This is especially noticeable at primes  $p \equiv 1 \mod 24$  and  $p \equiv \Box \mod 5, 7$ .



#### **Bias cancellation**

There is a negative bias in  $\bar{a}_p$  that is parity-independent and disappears in  $\bar{m}_p$ . This is especially noticeable at primes  $p \equiv 1 \mod 24$  and  $p \equiv \Box \mod 5, 7$ .



#### Moving averages

Moving average line plots of  $\bar{m}_{p}$  for 8 individual and all  $E/\mathbb{Q}$  with  $N_{E} \in (M, 2M]$ , using subintervals of size  $\sqrt{M}$  for p < 2M, with  $M = 2^{17}$ . 30 20 -10 -20 -30147455.b2, 163839.a1, 180222.be2, 196606.b1, 212990.11, 229374.a1, 245758.a1, 262143.d1

, 10000014

## Ordering by naive height

Elliptic curves with ht(E) := max(4|A|<sup>3</sup>, 27B<sup>2</sup>) in (M, 2M] for  $M = 2^{16}, \ldots, 2^{25}$ . The x-axis range is [0, 2M]. A blue/red or purple dot at (p,  $\bar{a}_p$  or  $\bar{m}_p$ ) shows the average of  $a_p$  or  $m_p$  over even/odd or all  $E/\mathbb{Q}$  with  $N_E \in (M, 2M]$ .



## Ordering by naive height

Elliptic curves with ht(E) := max(4|A|<sup>3</sup>, 27B<sup>2</sup>) in (M, 2M] for  $M = 2^{16}, \ldots, 2^{25}$ . The x-axis range is [0, 2M]. A blue/red or purple dot at (p,  $\bar{a}_p$  or  $\bar{m}_p$ ) shows the average of  $a_p$  or  $m_p$  over even/odd or all  $E/\mathbb{Q}$  with  $N_E \in (M, 2M]$ .

a\_p averages of 631953/630995 root number +1/-1 elliptic curves E/Q of naive height  $2^26 < ht(E) <= 2^{27}$  for p <  $2^{27}$ 





## Ordering by *j*-invariant



## Ordering by *j*-invariant











## Ordering by naive height (redux)

Elliptic curves with ht(E) := max(4|A|<sup>3</sup>, 27B<sup>2</sup>) in (M, 2M] for  $M = 2^{16}, \ldots, 2^{25}$ . The x-axis range is [0, 2M]. A blue/red or purple dot at (p,  $\bar{a}_p$  or  $\bar{m}_p$ ) shows the average of  $a_p$  or  $m_p$  over even/odd or all  $E/\mathbb{Q}$  with  $N_E \in (M, 2M]$ .



## Ordering by naive height (redux)

Elliptic curves with ht(E) := max(4|A|<sup>3</sup>, 27B<sup>2</sup>) in (M, 2M] for  $M = 2^{16}, \ldots, 2^{25}$ . The x-axis range is [0, 2M]. A blue/red or purple dot at (p,  $\bar{a}_p$  or  $\bar{m}_p$ ) shows the average of  $a_p$  or  $m_p$  over even/odd or all  $E/\mathbb{Q}$  with  $N_E \in (M, 2M]$ .

a\_p averages of 351546/351348 root number +1/-1 elliptic curves E/Q of naive height  $2^25 < ht(E) <= 2^26$  for p <  $2^26$ 



#### Ordering by conductor in the Stein-Watkins database (SWDB)



#### Ordering by conductor in the Stein-Watkins database (SWDB)



### Arithmetic L-functions

We call an *L*-function is analytic if it has the properties every good *L*-function should: analytic continuation, functional equation, Euler product, temperedness, central character; see FPRS18; it is analytically normalized if its central value is at s = 1/2.

An analytically normalized *L*-function  $L_{an}(s) = \sum a_n n^{-s}$  is arithmetic if  $a_n n^{\omega/2} \in \mathcal{O}_K$  for some number field *K* and  $\omega \in \mathbb{Z}_{\geq 0}$ . The least such  $\omega$  is the motivic weight. Its arithmetic normalization  $L(s) := L_{an}(s + \omega/2)$  has coefficients in  $\mathcal{O}_K$  and satisfies

$$\Lambda(s) = N^{1-s} w \overline{\Lambda}(1+\omega-s).$$

*L*-functions of abelian varieties have motivic weight  $\omega = 1$ . *L*-functions of weight-*k* holomorphic cuspforms have motivic weight  $\omega = k - 1$ .

We consider Galois-closed families of self-dual arithmetically normalized *L*-functions. In any such family the values of  $a_p$  and  $m_p$  are integers and  $w = \pm 1$ .

When averaging  $a_p$ 's in motivic weight  $\omega > 1$  we normalize them via  $a_p \mapsto a_p/p^{(\omega-1)/2}$ . This ensures that we always have  $|a_p| = O(\sqrt{p})$ , as with elliptic curves.

## Newforms for $\Gamma_0(N)$ of weight k = 2, 4, 6 with rational coefficients.

w(E)\*a p averages of 1691/1772 root number w(E) = +1/-1 weight 2 newforms for Gamma 0(N) of level 2^10 < N <= 2^11 and dimension g <= 1 for p < 2^11







## Newforms for $\Gamma_0(N)$ of weight k = 2, 4, 6 with rational coefficients.

-1 -2

-3



w(E)\*a p/p^2 averages of 259/304 root number w(E) = +1/-1 weight 6 newforms for Gamma 0(N) of level 2^10 < N <= 2^11 and dimension g <= 1 for p < 2^11



# Newforms for $\Gamma_0(N)$ of weight k = 2, 4, 6, 8.



## Newforms for $\Gamma_0(N)$ of weight k = 2, 4, 6, 8.



## Newforms for $\Gamma_0(N)$ of weight k = 2.



## Newforms for $\Gamma_0(N)$ of weight k = 4.



## Newforms for $\Gamma_0(N)$ of weight k = 6.



## Newforms for $\Gamma_0(N)$ of weight k = 8.

a p/p^3 averages of 66427462/65502558 root number +1/-1 weight 8 newforms for Gamma 0(N) of level 2^14 < N <= 2^15 for p < 2^16



# Newforms for $\Gamma_0(N)$ of weight k = 10.



## Newforms for $\Gamma_0(N)$ of weight k = 12.



#### Zubrilina's theorem

#### Theorem (Zubrilina 2023)

Let  $f = \sum_{n \ge 1} a_n q^n \in S_k^{\text{new}}(N)$  denote a newform of weight  $k \ge 2$  for  $\Gamma_0(N)$  with root number  $\varepsilon(f)$ . Let  $X, Y, P \to \infty$  with P prime,  $Y = (1 + o(1))X^{1-\delta_Y}$ ,  $P \ll X^{1+\delta_P}$ , for some  $\delta_Y, \delta_P > 0$  with  $2\delta_P < \delta_Y < 1$ . As a function of  $y \coloneqq P/X$  we have

$$\frac{\zeta(2)\pi}{XY} \sum_{\substack{N \in [X,X+Y]\\N \perp P \ \Box-\text{free}}} \sum_{f} \frac{\varepsilon(f) a_P(f)}{P^{(k/2-1)}} = A\sqrt{y} + (-1)^{k/2-1} \sum_{1 \le r \le 2\sqrt{y}} c(r)\sqrt{4y - r^2} U_{k-2}\left(\frac{r}{2\sqrt{y}}\right)$$
$$-\pi y \delta_{k=2} + O_{\varepsilon}(X^{-\delta'+\varepsilon})$$

where  $U_n$  is the Chebyshev polynomial defined by  $U_n(\cos x) = \sin(nx + x)/\sin x$ ,  $\delta' := \max(\delta_Y/2 - \delta_p, (\delta_Y + 1)/9 - \delta_P), c(r) := B \prod_{p|r} (1 + p^2/(p^4 - 2p^2 - p + 1)),$ where A = 1.450032... and B = 0.731311... are explicit constants.

For every  $\delta_P < 2/9$  one can choose  $\delta_Y$  so that  $\delta' > 0$ .

## Zubrilina's theorem



## Newforms for $\Gamma_0(N)$ of weight k = 2 with square root normalization.



## Newforms for $\Gamma_0(N)$ of weight k = 4 with square root normalization.



#### Murmurations of elliptic curves with squareroot normalization

Elliptic curve *L*-functions of conductor  $N \in (M, 2M]$  for  $M = 2^{11}, 2^{12}, \ldots, 2^{17}, 250000$ . The *x*-axis range is [0, 2M]. A blue/red or purple dot at  $(\sqrt{p}, \bar{a}_p \text{ or } \bar{m}_p)$  shows the average of  $a_p$  or  $m_p := w(E)a_p(E)$  over even/odd or all  $E/\mathbb{Q}$  with  $N_E \in (M, 2M]$ .



#### Murmurations of elliptic curves with squareroot normalization

Elliptic curve *L*-functions of conductor  $N \in (M, 2M]$  for  $M = 2^{11}, 2^{12}, \ldots, 2^{17}, 250000$ . The *x*-axis range is [0, 2M]. A blue/red or purple dot at  $(\sqrt{p}, \bar{a}_p \text{ or } \bar{m}_p)$  shows the average of  $a_p$  or  $m_p := w(E)a_p(E)$  over even/odd or all  $E/\mathbb{Q}$  with  $N_E \in (M, 2M]$ .



#### Murmurations in the weight aspect

Theorem (Bober, Booker, Lee, Lowry–Duda 2023)

Assume GRH for the L-functions of Dirichlet characters and modular forms. Let  $E \subset \mathbb{R}_{>0}$  be a compact interval with |E| > 0, let  $\delta \in \{0, 1\}$  and  $K, H, \varepsilon \in \mathbb{R}_{>0}$ , with  $K^{\frac{5}{6}+\varepsilon} < H < K^{1-\varepsilon}$  as  $K \to \infty$ , and put  $N := \left(\frac{\exp \psi(K/2)}{2\pi}\right)^2$ . We have

$$\frac{\sum_{\substack{k\equiv 2\delta \text{ mod } 4 \sum_{f} \sum_{p/N \in E} \frac{a_{p}(f) \log p}{p^{k/2-1}}}}{\sum_{\substack{k\equiv 2\delta \text{ mod } 4 \sum_{f} \sum_{p/N \in E} \log p}} = (-1)^{\delta} \left(\frac{\nu(E)}{|E|} + o_{E,\varepsilon}(1)\right), \quad \text{where}$$

$$\nu(E) \coloneqq \frac{1}{\zeta(2)} \sum_{\substack{r,q \in \mathbb{Z}_{>0} \\ \gcd(r,q)=1 \\ (r/q)^{-2} \in E}}^{*} \frac{\mu(q)^2}{\varphi(q)^2 \sigma(q)} \left(\frac{q}{r}\right)^4 = \frac{1}{2} \sum_{t \in \mathbb{Z}} \prod_{p \nmid t} \frac{p^2 - p - 1}{p^2 - p} \cdot \int_E \sqrt{x} \cos\left(\frac{2\pi t}{\sqrt{x}}\right) dx,$$

and the \* indicates that values of  $(r/q)^{-2}$  at endpoints of E have weight  $\frac{1}{2}$ .

#### Murmurations in the weight aspect



Also see [Iwaniec-Luo-Sarnak 2000] Low lying zeros of families of L-functions, where they consider  $a_p$  averages over p on the scale of  $\mathcal{N}^{\vartheta}$ , for  $\vartheta < 1$  and  $\vartheta > 1$ . They observe a phase transition at  $\vartheta = 1$ , which is the murmuration regime.

Recently constructed database of more than 5 million genus 2 curves  $X/\mathbb{Q}$  of conductor at most  $2^{20}$  includes 1,440,894 isogeny classes with Sato–Tate group USp(4). Conductor of L(X, s) in (M, 2M] for  $M = 2^{12}, \ldots, 2^{19}$  with x-axis range [0, M/2].



Coming soon to the LMFDB.

Recently constructed database of more than 5 million genus 2 curves  $X/\mathbb{Q}$  of conductor at most  $2^{20}$  includes 1,440,894 isogeny classes with Sato–Tate group USp(4). Conductor of L(X, s) in (M, 2M] for  $M = 2^{12}, \ldots, 2^{19}$  with x-axis range [0, M/2].



Recently constructed database of more than 5 million genus 2 curves  $X/\mathbb{Q}$  of conductor at most  $2^{20}$  includes 1,440,894 isogeny classes with Sato–Tate group USp(4). Conductor of L(X, s) in (M, 2M] for  $M = 2^{12}, \ldots, 2^{19}$  with x-axis range [0, M/2].



Coming soon to the LMFDB.

Recently constructed database of more than 5 million genus 2 curves  $X/\mathbb{Q}$  of conductor at most  $2^{20}$  includes 1,440,894 isogeny classes with Sato–Tate group USp(4). Conductor of L(X, s) in (M, 2M] for  $M = 2^{12}, \ldots, 2^{19}$  with x-axis range [0, M/2].



## Trace distributions of genus 2 curves





























WARDOW SHOL PARAMETERS

















A TRANSPORT OF AN ADDRESS MANY PR



## *L*-functions of genus 2 curves over $\mathbb{Q}$ , Sato-Tate group $N(SU(2) \times SU(2))$ .

These are primitive *L*-functions arising from Hilbert or Bianchi modular forms. Conductor of L(X, s) in (M, 2M] for  $M = 2^{12}, \ldots, 2^{19}$  with x-axis range [0, M/2].



## *L*-functions of genus 2 curves over $\mathbb{Q}$ , Sato-Tate group $N(SU(2) \times SU(2))$ .

These are primitive *L*-functions arising from Hilbert or Bianchi modular forms. Conductor of L(X, s) in (M, 2M] for  $M = 2^{12}, \ldots, 2^{19}$  with x-axis range [0, M/2].



# *L*-functions of products of $E/\mathbb{Q}$ , Sato-Tate group SU(2) × SU(2). Conductor of L(X, s) in (M, 2M] for $M = 2^{12}, \dots, 2^{17}$ with x-axis range [0, M/2].





# *L*-functions of products of $E/\mathbb{Q}$ , Sato-Tate group SU(2) × SU(2). Conductor of L(X, s) in (M, 2M] for $M = 2^{12}, \dots, 2^{17}$ with x-axis range [0, M/2].





# *L*-functions of genus 3 curves over $\mathbb{Q}$ with Sato-Tate group USp(6). Recently constructed database of genus 3 curves $X/\mathbb{Q}$ of conductor at most 10<sup>7</sup> includes 59,214 isogeny classes of hyperelliptic curves with ST group USp(6). Conductor of L(X, s) in (M, 2M] for $M = 2^{16}, \ldots, 2^{22}$ with x-axis range [0, M/2].



Coming soon to the LMFDB.

Recently constructed database of genus 3 curves  $X/\mathbb{Q}$  of conductor at most  $10^7$  includes 59,214 isogeny classes of hyperelliptic curves with ST group USp(6). Conductor of L(X, s) in (M, 2M] for  $M = 2^{16}, \ldots, 2^{22}$  with x-axis range [0, M/2].



Coming soon to the LMFDB.

## Thank you





Animations available at https://math.mit.edu/~drew/murmurations.html.