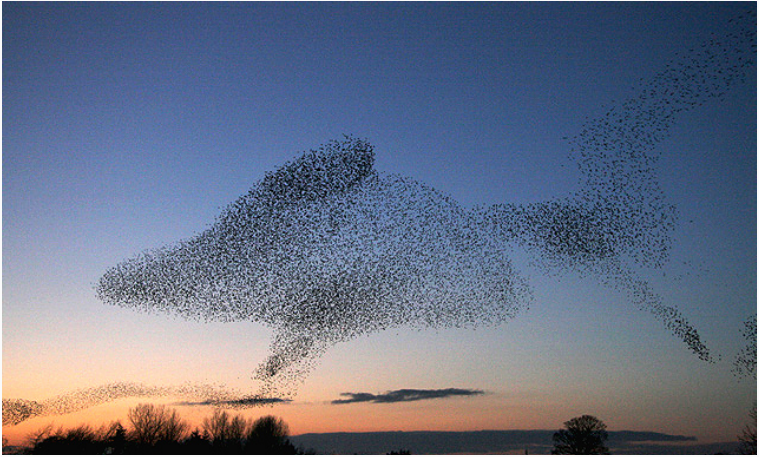
# Murmuration Phenomena

## Patterns in Number Theory

David Lowry-Duda (ICERM & Brown University)

November 2024

LSU Number Theory Seminar



*Starling shapes in the evening sky*, Walter Baxter,
`https://www.geograph.org.uk/photo/1065181`.

The murmurations phenomenon describes certain biases in averages of coefficients of $L$-functions, unnoticed but for a sequence of unlikely observations (and deep attentiveness from young mathematicians).

# Elliptic Curves

An elliptic curve (defined over $\mathbb{Q}$) is an algebraic curve of the form

$$E : Y^2 = X^3 + aX + b$$

for integers $a, b \in \mathbb{Z}$.

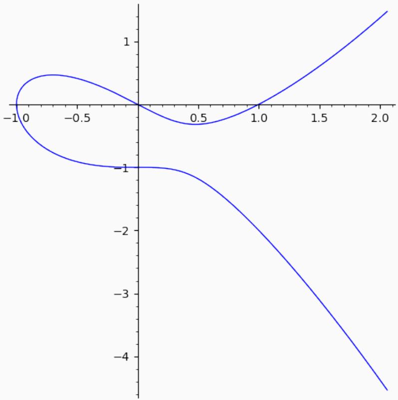
# Elliptic Curves

An elliptic curve (defined over $\mathbb{Q}$) is an algebraic curve of the form

$$E : Y^2 = X^3 + aX + b$$

for integers $a, b \in \mathbb{Z}$.

$$Y^2 = X^3 - 675X + 13662$$



# Elliptic Curves

An elliptic curve (defined over $\mathbb{Q}$) is an algebraic curve of the form

$$E : Y^2 = X^3 + aX + b$$

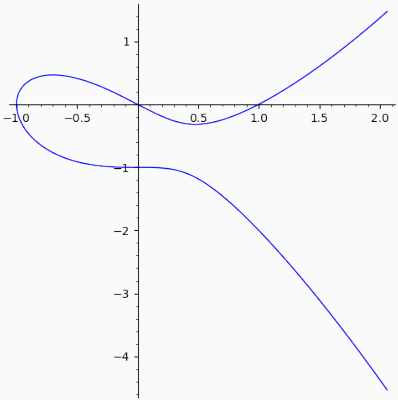
for integers $a, b \in \mathbb{Z}$.

To an elliptic curve $E$, we can associate an $L$-function

$$L(s, E) := \prod_{p \geq 2} L_p(s)^{-1} = \sum_{n \geq 1} \frac{a_n(E)}{n^s}.$$

The coefficients $a_n(E)$ satisfy many properties, but can be understood explicitly.

$$Y^2 = X^3 - 675X + 13662$$



When $p$ is a prime, the coefficient carries local point-counting information,

$$a_p(E) = p + 1 - \#E(\mathbb{F}_p),$$

where $\#E(\mathbb{F}_p)$ is the number of points on $E$ mod $p$

When $p$ is a prime, the coefficient carries local point-counting information,

$$a_p(E) = p + 1 - \#E(\mathbb{F}_p),$$

where $\#E(\mathbb{F}_p)$ is the number of *projective*[1] points on $E$ mod $p$, possibly including "points at $\infty$".

[1] projective points on the homogenized curve $Y^2 Z = X^3 + aXZ^2 + bZ^3$.

When $p$ is a prime, the coefficient carries local point-counting information,

$$a_p(E) = p + 1 - \#E(\mathbb{F}_p),$$

where $\#E(\mathbb{F}_p)$ is the number of *projective*[1] points on $E$ mod $p$, possibly including "points at $\infty$".

For example, consider $Y^2 = X^3 - 675X + 13662$ over $\mathbb{F}_5$. These are points with $Y^2 \equiv X^3 + 2 \bmod 5$, and there are $5 + 1$ of these points

$$(2,0), (3,2), (3,3), (4,1), (4,4), \qquad \text{and} \quad (0:1:0).$$

[1]projective points on the homogenized curve $Y^2Z = X^3 + aXZ^2 + bZ^3$.

When $p$ is a prime, the coefficient carries local point-counting information,

$$a_p(E) = p + 1 - \#E(\mathbb{F}_p),$$

where $\#E(\mathbb{F}_p)$ is the number of *projective*[1] points on $E$ mod $p$, possibly including "points at $\infty$".

For example, consider $Y^2 = X^3 - 675X + 13662$ over $\mathbb{F}_5$. These are points with $Y^2 \equiv X^3 + 2 \bmod 5$, and there are $5 + 1$ of these points

$$(2,0),(3,2),(3,3),(4,1),(4,4), \qquad \text{and} \quad (0:1:0).$$

We extend $a(n)$ to composite $n$ via

$$\prod_p (1 - a_p(E)p^{-s} + p^{-2s+1}) = \sum_{n\geq 1} \frac{a_n(E)}{n^s}.$$

[1] projective points on the homogenized curve $Y^2Z = X^3 + aXZ^2 + bZ^3$.

It's easy to compute these points slowly. It's much harder to compute them *quickly*. But we've gotten very good, and now we make huge tables and databases where we try to understand their properties.

# Elliptic curve with LMFDB label 14.a5 (Cremona label 14a4)

Show commands: Magma / Oscar / PariGP / SageMath

This is a model for the modular curve $X_1(14)$.

## Simplified equation

$y^2 = x^3 - 675x + 13662$ (homogenize, minimize)

## Mordell-Weil group structure

$\mathbb{Z}/6\mathbb{Z}$

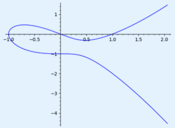
## Torsion generators

$(1, 0)$

## Integral points

**Properties**

**Label** 14.a5



| | |
|---|---|
| **Conductor** | $14$ |
| **Discriminant** | $-28$ |
| **j-invariant** | $-\frac{15625}{28}$ |
| **CM** | no |
| **Rank** | $0$ |
| **Torsion structure** | $\mathbb{Z}/6\mathbb{Z}$ |

# A series of unlikely events

About two years ago, Yang-Hui He, Kyu-Hwan Lee, Thomas Oliver, and Alexey Pozdnyakov experimented with using techniques from machine learning to study elliptic curves. They used data from the LMFDB.

# A series of unlikely events

About two years ago, Yang-Hui He, Kyu-Hwan Lee, Thomas Oliver, and Alexey Pozdnyakov experimented with using techniques from machine learning to study elliptic curves. They used data from the LMFDB.

$$\begin{pmatrix} a_2(E_0) & a_3(E_0) & a_5(E_0) & a_7(E_0) & a_{11}(E_0) & \cdots \\ a_2(E_1) & a_3(E_1) & a_5(E_1) & a_7(E_1) & a_{11}(E_1) & \cdots \\ a_2(E_2) & a_3(E_2) & a_5(E_2) & a_7(E_2) & a_{11}(E_2) & \cdots \\ a_2(E_3) & a_3(E_3) & a_5(E_3) & a_7(E_3) & a_{11}(E_3) & \cdots \\ a_2(E_4) & a_3(E_4) & a_5(E_4) & a_7(E_4) & a_{11}(E_4) & \cdots \\ a_2(E_5) & a_3(E_5) & a_5(E_5) & a_7(E_5) & a_{11}(E_5) & \cdots \\ \vdots & & & & & \end{pmatrix}$$

They looked at $a_p(\cdot)$ values of elliptic curves in the same order the LMFDB uses (by conductor).

# A series of unlikely events

About two years ago, Yang-Hui He, Kyu-Hwan Lee, Thomas Oliver, and Alexey Pozdnyakov experimented with using techniques from machine learning to study elliptic curves. They used data from the LMFDB.

$$\begin{pmatrix} a_2(E_0) & a_3(E_0) & a_5(E_0) & a_7(E_0) & a_{11}(E_0) & \cdots \\ a_2(E_1) & a_3(E_1) & a_5(E_1) & a_7(E_1) & a_{11}(E_1) & \cdots \\ a_2(E_2) & a_3(E_2) & a_5(E_2) & a_7(E_2) & a_{11}(E_2) & \cdots \\ a_2(E_3) & a_3(E_3) & a_5(E_3) & a_7(E_3) & a_{11}(E_3) & \cdots \\ a_2(E_4) & a_3(E_4) & a_5(E_4) & a_7(E_4) & a_{11}(E_4) & \cdots \\ a_2(E_5) & a_3(E_5) & a_5(E_5) & a_7(E_5) & a_{11}(E_5) & \cdots \\ \vdots & & & & & \end{pmatrix}$$

While applying various unsupervised learning techniques, they decided to use a dimension reduction and principal component analysis (PCA).

# A series of unlikely events

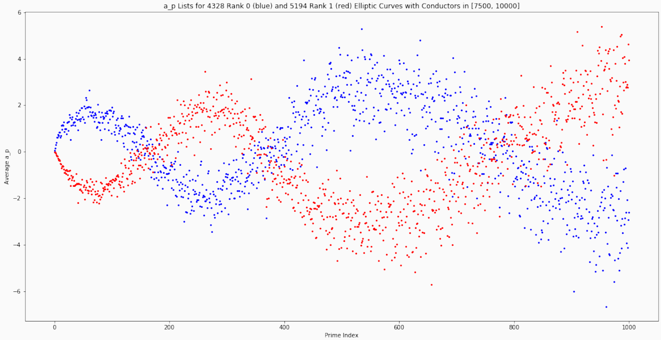
About two years ago, Yang-Hui He, Kyu-Hwan Lee, Thomas Oliver, and Alexey Pozdnyakov experimented with using techniques from machine learning to study elliptic curves. They used data from the LMFDB.

$$\begin{pmatrix} a_2(E_0) & a_3(E_0) & a_5(E_0) & a_7(E_0) & a_{11}(E_0) & \cdots \\ a_2(E_1) & a_3(E_1) & a_5(E_1) & a_7(E_1) & a_{11}(E_1) & \cdots \\ a_2(E_2) & a_3(E_2) & a_5(E_2) & a_7(E_2) & a_{11}(E_2) & \cdots \\ a_2(E_3) & a_3(E_3) & a_5(E_3) & a_7(E_3) & a_{11}(E_3) & \cdots \\ a_2(E_4) & a_3(E_4) & a_5(E_4) & a_7(E_4) & a_{11}(E_4) & \cdots \\ a_2(E_5) & a_3(E_5) & a_5(E_5) & a_7(E_5) & a_{11}(E_5) & \cdots \\ \vdots & & & & & \end{pmatrix}$$

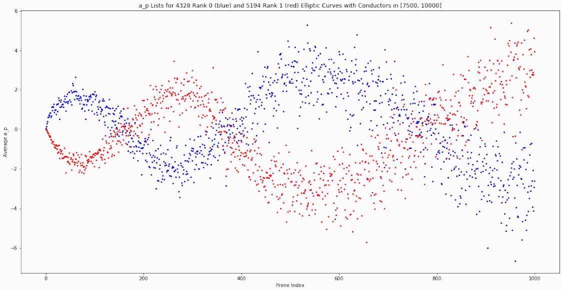
While applying various unsupervised learning techniques, they decided to use a dimension reduction and principal component analysis (PCA). Alexey decided to look at the weights in the PCA. But there were thousands and thousands, so he decided to plot them.

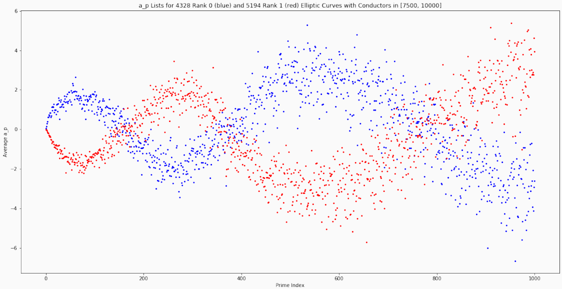
# A series of unlikely events

About two years ago, Yang-Hui He, Kyu-Hwan Lee, Thomas Oliver, and Alexey Pozdnyakov experimented with using techniques from machine learning to study elliptic curves. They used data from the LMFDB.




This shows for each prime $p \in [2, 7919]$ the average of $a_p(\cdot)$ where $E$ ranges over curves of conductor $N \in [7500, 10000]$. Points in blue are from rank 0 curves, and points in red are from rank 1 curves.

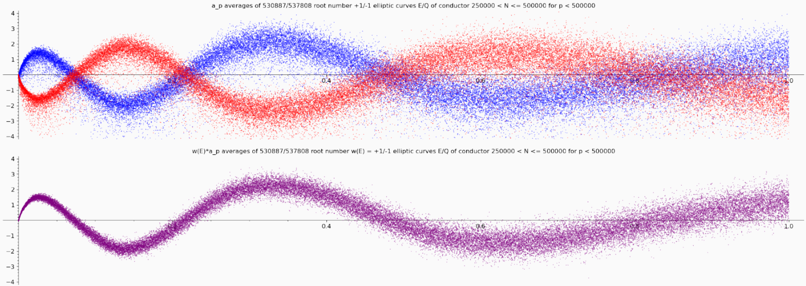

a_p Lists for 4328 Rank 0 (blue) and 5194 Rank 1 (red) Elliptic Curves with Conductors in [7500, 10000]
Average a_p
Prime Index
6
4
2
0
−2
−4
−6
0
200
400
600
800
1000



a_p Lists for 4328 Rank 0 (blue) and 5194 Rank 1 (red) Elliptic Curves with Conductors in [7500, 10000]
Average a_p
Prime Index
6
4
2
0
-2
-4
-6
0
200
400
600
800
1000


They emailed Drew Sutherland, who happily began to collaborate (and who has given several excellent talks on this). He computed much further out.

They emailed Drew Sutherland, who happily began to collaborate (and who has given several excellent talks on this). He computed much further out.




There is a small normalization issue here. The $x$-axis label has changed. It turns out that normalizing the $x$-axis to consider ratios $p/N$, where $N$ is the conductor of the elliptic curves, seems to make these plots uniform.

Can we try to explain this phenomenon?

One form of the Birch and Swinnerton-Dyer conjecture implies

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

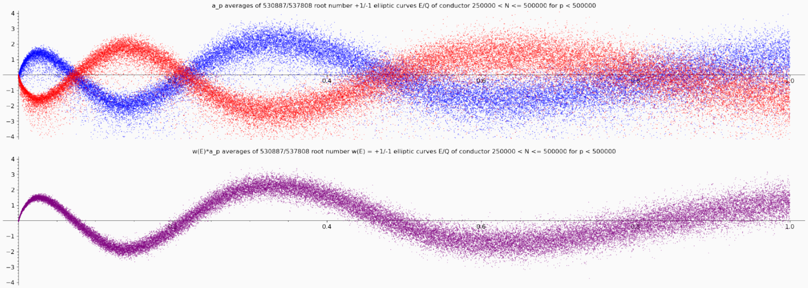
A folklore conjecture (often attributed to Goldfeld) asserts that 50% of elliptic curves have rank 0 and 50% of elliptic curves have rank 1.

# This murmuration is not surprising

Conjecturally, half the time $r = 1$ and BSD suggests there should be a negative correlation between the $a_p(E)$ and $p$, and half the time $r = 0$ and there should be a positive correlation.

# This murmuration is ~~not~~ VERY surprising

Conjecturally, half the time $r = 1$ and BSD suggests there should be a negative correlation between the $a_p(E)$ and $p$, and half the time $r = 0$ and there should be a positive correlation.



# Talk Outline

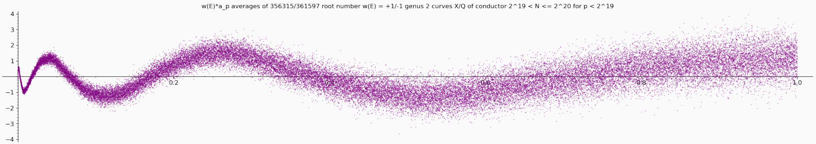
1. Lots of murmurations
2. Connections to one-level density
3. Things we can prove
4. Is machine learning useful?
5. Archimedean murmurations

# More murmurations!

He, Lee, Oliver, Pozdnyakov, and Sutherland have carried out more experiments and found many more, analogous, murmuration phenomena. Even more experiments were carried out in July of 2023 at ICERM.

# More murmurations!

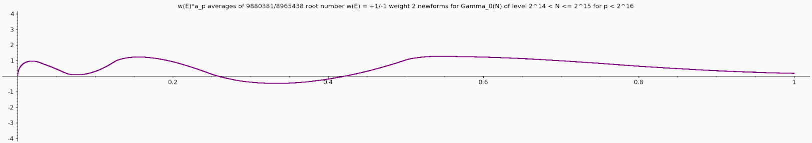
He, Lee, Oliver, Pozdnyakov, and Sutherland have carried out more experiments and found many more, analogous, murmuration phenomena. Even more experiments were carried out in July of 2023 at ICERM.




abelian surfaces with generic Sato–Tate group

# More murmurations!

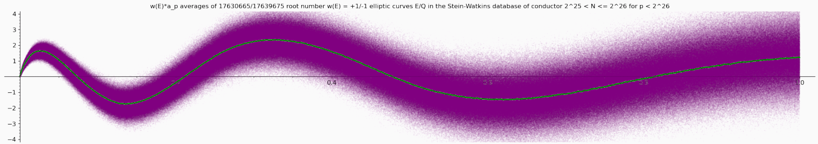
He, Lee, Oliver, Pozdnyakov, and Sutherland have carried out more experiments and found many more, analogous, murmuration phenomena. Even more experiments were carried out in July of 2023 at ICERM.




self-dual holomorphic newforms of weight 2

# More murmurations!

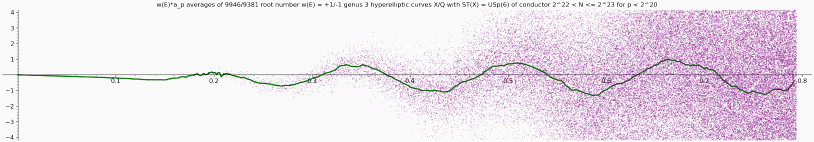
He, Lee, Oliver, Pozdnyakov, and Sutherland have carried out more experiments and found many more, analogous, murmuration phenomena. Even more experiments were carried out in July of 2023 at ICERM.




elliptic curves from the Stein–Watkins database

# More murmurations!

He, Lee, Oliver, Pozdnyakov, and Sutherland have carried out more experiments and found many more, analogous, murmuration phenomena. Even more experiments were carried out in July of 2023 at ICERM.




genus 3 curves with generic Sato–Tate group

This seems to really be about correlations between coefficients of $L$-functions and the root numbers of the functional equations.

A general (self-dual) $L$-function has the shape

$$L(s, \pi) = \sum_{n \geq 1} \frac{a(n)}{n^s}$$

and satisfies a functional equation

$$\Lambda(s, \pi) := N^s G(s) L(s, f) = \epsilon N^{1-s} G(1-s) L(1-s, f).$$

For functions associated to $\pi$ in some family $\mathcal{F}$ (ordered by some height function, typically analytic conductor) murmuration phenomena are correlations between the coefficients $a(n)$ and the root numbers $\epsilon$.

Given a smooth nonnegative weight function $\Phi : (0, \infty) \longrightarrow \mathbb{R}$ of compact support and a complex-valued function $f$ defined on a family $\mathcal{F}$ of $L$-functions ordered with respect to a height function $h$, we look at

$$A_{\Phi}^{\mathcal{F}}(f, X) = A(f, X) := \sum_{\pi \in \mathcal{F}} \Phi(h(\pi)/X) f(\pi),$$

Given a smooth nonnegative weight function $\Phi : (0, \infty) \longrightarrow \mathbb{R}$ of compact support and a complex-valued function $f$ defined on a family $\mathcal{F}$ of $L$-functions ordered with respect to a height function $h$, we look at

$$A_{\Phi}^{\mathcal{F}}(f, X) = A(f, X) := \sum_{\pi \in \mathcal{F}} \Phi(h(\pi)/X) f(\pi),$$

and more generally expected values

$$\mathop{\mathbb{E}}_{\pi \in \mathcal{F}}[f; X] := \frac{A(f, X)}{A(1, X)}.$$

Given a smooth nonnegative weight function $\Phi : (0, \infty) \longrightarrow \mathbb{R}$ of compact support and a complex-valued function $f$ defined on a family $\mathcal{F}$ of $L$-functions ordered with respect to a height function $h$, we look at

$$A_{\Phi}^{\mathcal{F}}(f, X) = A(f, X) := \sum_{\pi \in \mathcal{F}} \Phi(h(\pi)/X) f(\pi),$$

and more generally expected values

$$\mathop{\mathbb{E}}_{\pi \in \mathcal{F}}[f; X] := \frac{A(f, X)}{A(1, X)}.$$

For example, take $\mathcal{F} = \mathcal{E}^{\pm}$, elliptic curves ordered by conductor with root number $\epsilon = \pm 1$; and take $\Phi$ to be the indicator function on $[X, 2X]$. Then the murmuration plots shown before are exactly plots of

$$\mathbb{E}_{\mathcal{E}^{\pm}}[a_E(p); X] = \frac{\sum_{\substack{E \in \mathcal{E}^{\pm} \\ X \leq \mathrm{cond}(E) \leq 2X}} a_E(p)}{\sum_{\substack{E \in \mathcal{E}^{\pm} \\ X \leq \mathrm{cond}(E) \leq 2X}} 1}$$

for various primes $p$ and in various dyadic ranges $[X, 2X]$.

# One Level Density

For "nice" families $\mathcal{F}$ of $L$-functions $L(s, f)$, ordered by conductor $N_f$, let $\mathcal{F}(N) = \{f \in \mathcal{F} : N_f = N\}$.

Katz and Sarnak predict that for large $N$, the low-lying zeros of $L(s, f)$ for $f \in \mathcal{F}(N)$ act like eigenvalues of matrices drawn randomly from certain groups of matrices associated to $\mathcal{F}$. One measure of zero behavior is one level density, which is

# One Level Density

For "nice" families $\mathcal{F}$ of $L$-functions $L(s, f)$, ordered by conductor $N_f$, let $\mathcal{F}(N) = \{f \in \mathcal{F} : N_f = N\}$.

Katz and Sarnak predict that for large $N$, the low-lying zeros of $L(s, f)$ for $f \in \mathcal{F}(N)$ act like eigenvalues of matrices drawn randomly from certain groups of matrices associated to $\mathcal{F}$. One measure of zero behavior is one level density, which is

$$\mathrm{OLD}_\Phi(\mathcal{F}) := \lim_{N\to\infty} \frac{1}{\#\mathcal{F}(N)} \sum_{f\in\mathcal{F}(N)} \sum_{\gamma_f} \Phi\left(\frac{\gamma_f \log N}{2\pi}\right),$$

where $\gamma_f$ runs through nontrivial zeros of $L(s, f)$ and $\Phi$ is a "nice" test function. This measures the distribution of low-lying zeros *on average* over elements of large conductor.

Katz and Sarnak predict that for many families, there is a measure $W_{\mathcal{F}}$ coming from matrices such that

$$\mathrm{OLD}_{\Phi}(\mathcal{F}) = \int_{\mathbb{R}} \widehat{\Phi}(x)\widehat{W_{\mathcal{F}}}(x)dx$$

for all nice test functions $\Phi$.

Katz and Sarnak predict that for many families, there is a measure $W_{\mathcal{F}}$ coming from matrices such that

$$\mathrm{OLD}_{\Phi}(\mathcal{F}) = \int_{\mathbb{R}} \widehat{\Phi}(x)\widehat{W_{\mathcal{F}}}(x)dx$$

for all nice test functions $\Phi$.

**Theorem (Iwaniec–Luo–Sarnak 2000)**

*Assume GRH. Let $\Phi$ be a Schwarz function with $\mathrm{supp}(\widehat{\Phi}) \subset (-2, 2)$. Let $H_k^{\pm}(N)$ denote a Hecke eigenbasis of modular newforms of weight $k$ and root number $\epsilon_f = \pm 1$. Then*

$$E(H_k^{\pm}; \Phi) = \int_{\mathbb{R}} \widehat{\Phi}(x)\widehat{W_{\mathrm{SO}(\pm)}}(x)dx$$

*where $W_{\mathrm{SO}(+)} = 1 + \frac{\sin(2\pi x)}{2\pi x}$ and $W_{\mathrm{SO}(-)} = 1 - \frac{\sin(2\pi x)}{2\pi x} + \delta_0(x)$.*

# Aside on modular forms

A modular form comes with two pieces of data:

1. a weight $k$, and
2. a group $\Gamma_0(N) = \{\gamma \in \mathrm{SL}(2, \mathbb{Z}) : \gamma \equiv \left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right) \bmod N\}$ of level $N$.

Then we say a holomorphic function $f$, defined on the complex upper half-plane $\mathcal{H} = \{x + iy : y > 0\}$ is of level $N$ and weight $k$ if

$$f\left(\tfrac{az+b}{cz+d}\right) = (cz+d)^k f(z) \qquad \forall \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N),$$

(and if it satisfies a certain complicated notion of holomorphy at the boundary).

# Modular forms, pictorially

The condition that $f(\gamma z) = (cz + d)^k f(z)$ for all $\gamma \in \Gamma_0(N)$ gives an infinite family of symmetries. As $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_0(N)$, we get that $f(z+1) = f(z)$.

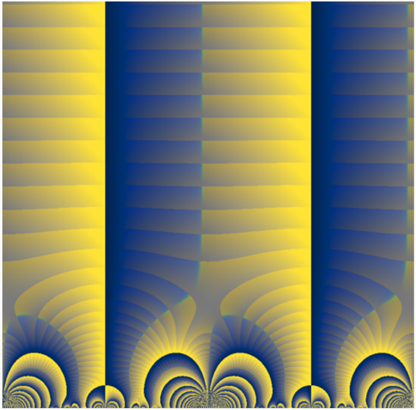
# Modular forms, pictorially

The condition that $f(\gamma z) = (cz + d)^k f(z)$ for all $\gamma \in \Gamma_0(N)$ gives an infinite family of symmetries. As $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_0(N)$, we get that $f(z + 1) = f(z)$.

The infinitely many other symmetries are hard to understand. But we can get a pretty good idea by looking at a picture.

# Modular forms, pictorially

The condition that $f(\gamma z) = (cz+d)^k f(z)$ for all $\gamma \in \Gamma_0(N)$ gives an infinite family of symmetries. As $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_0(N)$, we get that $f(z+1) = f(z)$.



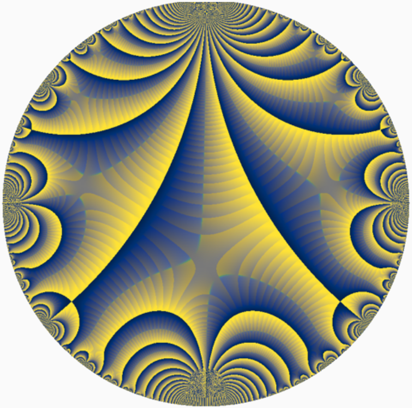
# Modular forms, pictorially

The condition that $f(\gamma z) = (cz+d)^k f(z)$ for all $\gamma \in \Gamma_0(N)$ gives an infinite family of symmetries. As $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_0(N)$, we get that $f(z+1) = f(z)$.

Or we can conformally map $\mathcal{H}$ to the unit disk, getting instead

# Modular forms, pictorially

The condition that $f(\gamma z) = (cz + d)^k f(z)$ for all $\gamma \in \Gamma_0(N)$ gives an infinite family of symmetries. As $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_0(N)$, we get that $f(z + 1) = f(z)$.



As noted above, a modular form $f(z)$ is periodic. The “nice” holomorphic properties guarantee that $f$ equals its Fourier expansion

$$f(z) = \sum_{n \geq 0} a_n(f) e^{2\pi i n z}.$$

As noted above, a modular form $f(z)$ is periodic. The "nice" holomorphic properties guarantee that $f$ equals its Fourier expansion

$$f(z) = \sum_{n \geq 0} a_n(f) e^{2\pi i n z}.$$

To a modular form $f$, we associate the Dirichlet series

$$L(s, f) = \sum_{n \geq 1} \frac{a_n(f)}{n^s}.$$

For it to be an *actual* $L$-function, $a_0(f)$ must be 0 and $f$ must be chosen from a distinguished basis $H_k(N)$ of "Hecke" forms. In this case, $L(s, f)$ satisfies a functional equation of the form $L(s, f)G(s) \mapsto \epsilon_f L(1-s, f)G(1-s)$ for some product of $\Gamma$ functions, $G(s)$, and where $\epsilon_f \in \{\pm 1\}$ is the root number or sign of the functional equation.

# Connection to Murmurations

Unravelling, ILS shows that (under GRH)

$$\lim \frac{1}{\# H_k^{\pm}(N)} \sum_{f \in H_k^{\pm}(N)} \sum_{\gamma_f} \Phi\left(\frac{\gamma_f \log N}{2\pi}\right) = \int_{\mathbb{R}} \widehat{\Phi}(x) \widehat{W_{\mathrm{SO}(\pm)}}(x) dx.$$

The explicit formula relating zeros to sums over primes implies that

$$\sum_{\gamma_f} \Phi\left(\frac{\gamma_f \log N}{2\pi}\right) \approx \sum_{p} \frac{\lambda_f(p) \log p}{\sqrt{p}} \widehat{\Phi}\left(\frac{\log p}{\log N}\right).$$

Note that if $\widehat{\Phi}$ is supported on $[-\theta, \theta]$, then only primes $\leq N^\theta$ appear. Hence one level density behaves like

$$\mathop{\mathbb{E}}_{p \sim N^\theta} \mathop{\mathbb{E}}_{f \in H_k^{\pm}(N)} [\lambda_f(p) \log p / \sqrt{p}].$$

$$\mathop{\mathbb{E}}_{p\sim N^\theta}\mathop{\mathbb{E}}_{f\in H_k^\pm(N)}[\lambda_f(p)\log p/\sqrt{p}]$$

is *almost* an averaged murmuration behavior (with some prime scaling), except that murmurations concern $p \sim N$ instead of $p \sim N^\theta$. Specifically

$$\widehat{W_{\mathrm{SO}(+)}}(y) = \delta_0(y) + \frac{\mathbf{1}_{[-1,1]}(y)}{2}$$
$$\widehat{W_{\mathrm{SO}(-)}}(y) = \delta_0(y) + \frac{2-\mathbf{1}_{[-1,1]}(y)}{2}.$$

There is a discontinuity in behavior exactly when $\pm 1 \in \mathrm{supp}(\widehat{\Phi})$. Murmurations arise from the transition range for one-level density for known families, which is much more mysterious than $p \sim N^\theta$ for $\theta < 1$ or $\theta > 1$.

# Unpublished result

In some cases, it would be possible to *prove* murmuration-like phenomena from one-level density results. The proofs in ILS 2000 imply that

$$\frac{\sum_{f \in H_k^{\pm}(1)} \Phi(N_f/X) \frac{\lambda_f(p)\sqrt{p}}{L(1,\mathrm{sym}^2 f)}}{\sum_{f \in H_k^{\pm}(1)} \Phi(N_f/X) \frac{1}{L(1,\mathrm{sym}^2 f)}} = \pm \sum_{c \geq 1} \frac{\mu^2(c)}{c^2 \varphi(c)} \Phi(*),$$

which describes correlations between $\lambda_f(p)/L(1, \mathrm{sym}^2 f)$ with root numbers.

# Unpublished result

In some cases, it would be possible to *prove* murmuration-like phenomena from one-level density results. The proofs in ILS 2000 imply that

$$\frac{\sum_{f\in H_k^{\pm}(1)} \Phi(N_f/X)\frac{\lambda_f(p)\sqrt{p}}{L(1,\mathrm{sym}^2 f)}}{\sum_{f\in H_k^{\pm}(1)} \Phi(N_f/X)\frac{1}{L(1,\mathrm{sym}^2 f)}} = \pm\sum_{c\geq 1}\frac{\mu^2(c)}{c^2\varphi(c)}\Phi(*),$$

which describes correlations between $\lambda_f(p)/L(1,\mathrm{sym}^2 f)$ with root numbers.

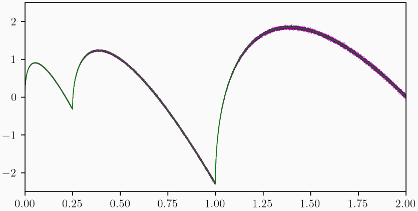
What can we learn from this?

# A first theorem!

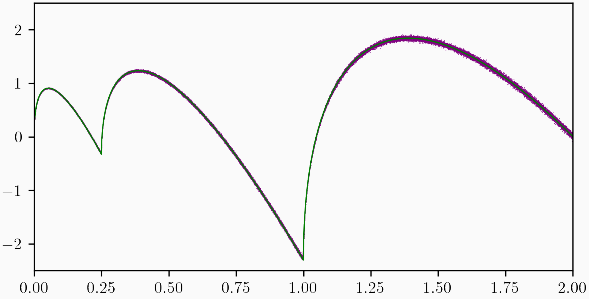
**Theorem (Zubrilina, arXiv:2310.07681)**

*Fix $k \in 2\mathbb{Z}_{>0}$. Then there is a continuous function $M_k : \mathbb{R}_{>0} \to \mathbb{R}$ such that, for any fixed $y \in \mathbb{R}_{>0}$ and $\delta \in (0,1)$,*

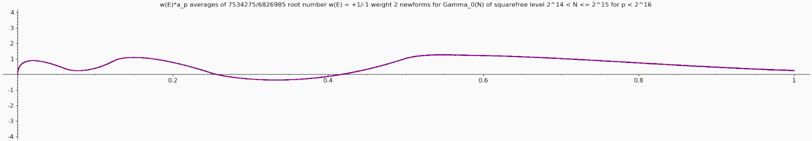
$$\lim_{\substack{p \text{ prime} \\ p \to \infty}} \frac{\sum_{\substack{N \in [p/y, p/y + p^\delta] \cap \mathbb{Z} \\ N \text{ squarefree}}} \sum_{f \in H_k(N)} \epsilon_f a_f(p) \sqrt{p}}{\sum_{\substack{N \in [p/y, p/y + p^\delta] \cap \mathbb{Z} \\ N \text{ squarefree}}} \sum_{f \in H_k(N)} 1} = M_k(y).$$




$M_2$ versus numerics for $|N - 2^{18}| < 2^{10}$ and $p < 2^{19}$




$M_2$ versus numerics for $|N - 2^{18}| < 2^{10}$ and $p < 2^{19}$



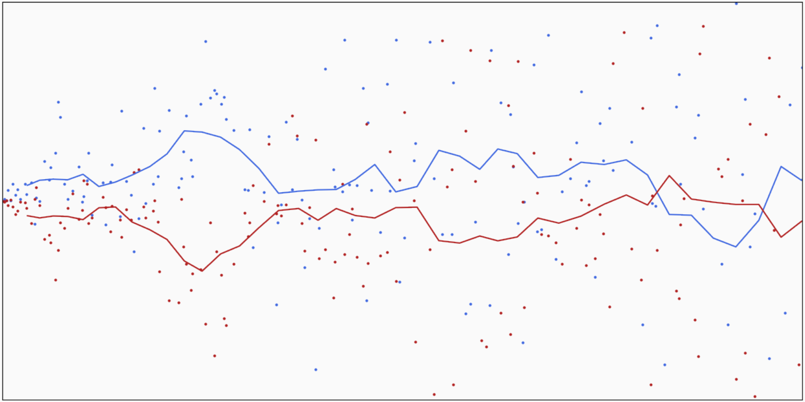

dyadic convolution of $M_2$ versus numerics for $N \in [2^{14}, 2^{15}]$ and $p < 2^{16}$

Last July, ICERM held a workshop on murmurations. Sarnak asked whether there are murmuration phenomena for families with varying Archimedean parameters (such as Maass forms with increasing eigenvalue, or modular forms of fixed level as the weight $k \to \infty$).

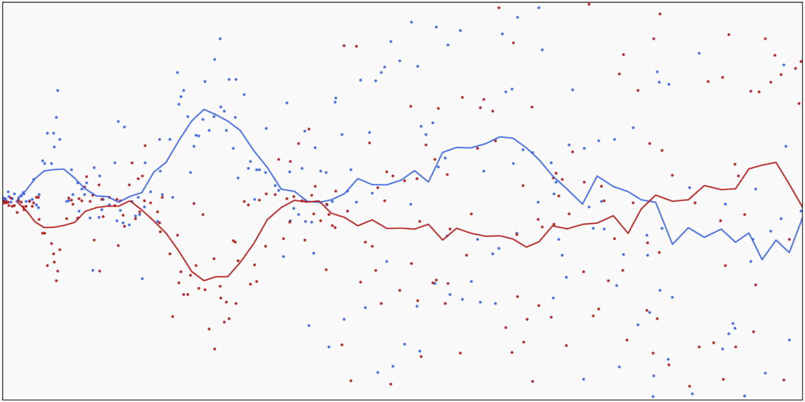
For both of these, we would need to use the *analytic conductor* instead of the standard conductor: the analytic conductor of a modular form of weight $k$ and fixed level behaves like $\mathcal{N}(k) = \left(\frac{k-1}{4\pi}\right)^2 + O(1)$, and the analytic conductor of a Maass form with eigenvalue $\lambda = \frac{1}{4} + R^2$ behaves like $\mathcal{N}(R) = \frac{R^2}{4\pi^2} + O(1)$. These are both quadratic in the Archimedean parameter, which leads to different behavior.

Andrei Seymour-Howell and I have been rigorously computing Maass forms. What does this data look like?

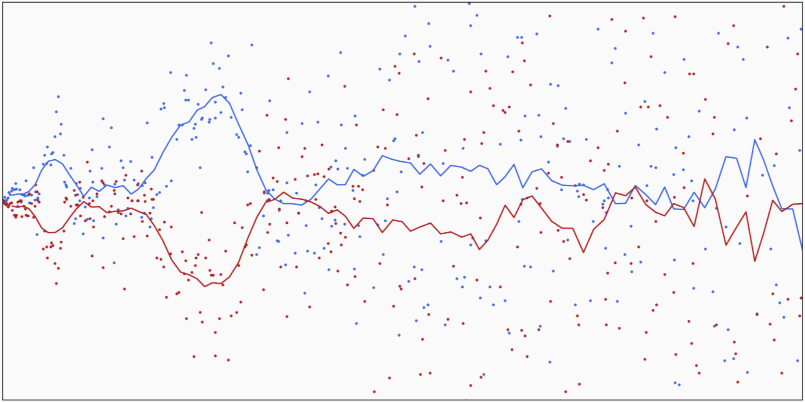
Andrei Seymour-Howell and I have been rigorously computing Maass forms. What does this data look like?



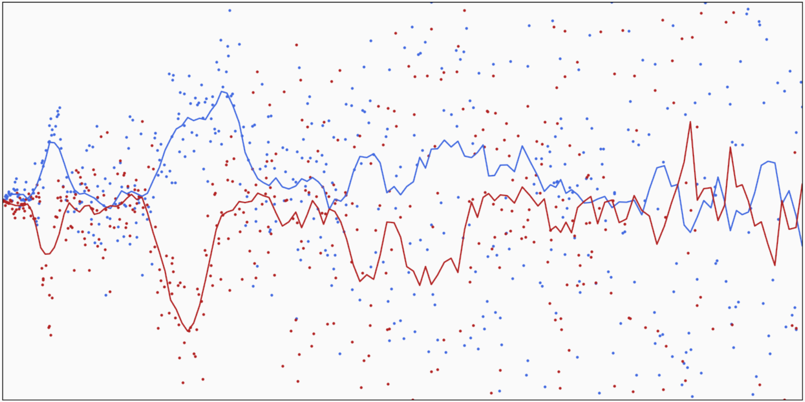
Andrei Seymour-Howell and I have been rigorously computing Maass forms. What does this data look like?



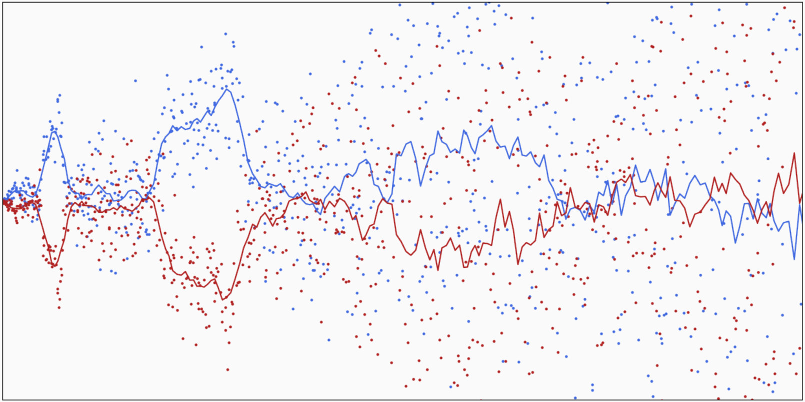
Andrei Seymour-Howell and I have been rigorously computing Maass forms. What does this data look like?



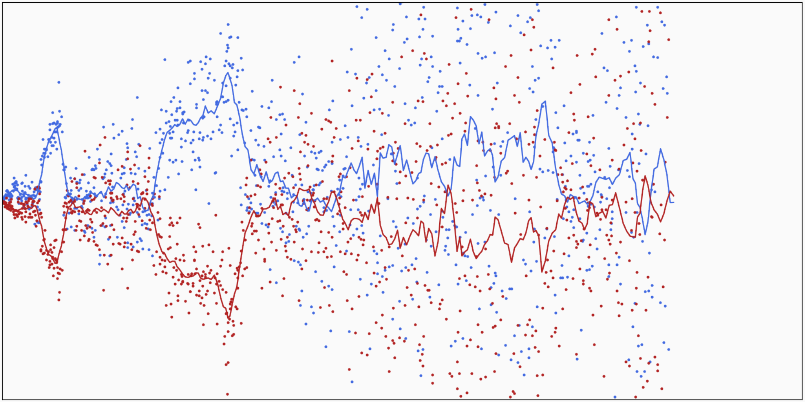
Andrei Seymour-Howell and I have been rigorously computing Maass forms. What does this data look like?



Andrei Seymour-Howell and I have been rigorously computing Maass forms. What does this data look like?



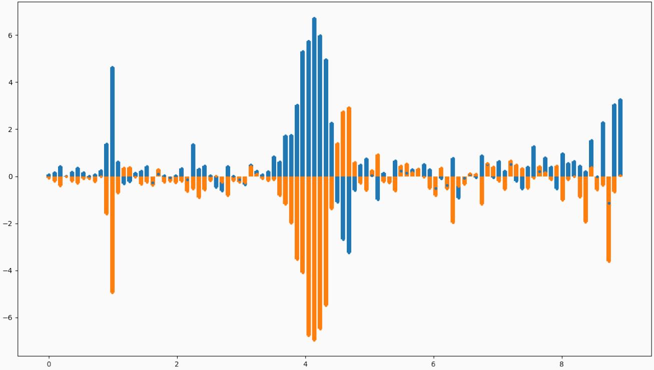
Andrei Seymour-Howell and I have been rigorously computing Maass forms. What does this data look like?



When generating data for classical modular forms in the LMFDB, Jonathan Bober implemented an (extremely fast) trace formula that can produce data for fixed level, rising weight modular forms.

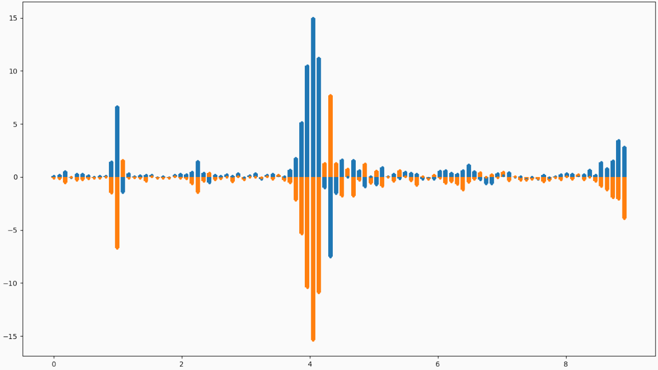
In collaboration with Min Lee, Bober, and Andy Booker, we began to look at experimental plots. We observed that normalizing points by $p/\mathcal{N}(k)$ is complicated to reason about — so we collect points into small bins (corresponding to small local averages in $p$).

When generating data for classical modular forms in the LMFDB, Jonathan Bober implemented an (extremely fast) trace formula that can produce data for fixed level, rising weight modular forms.



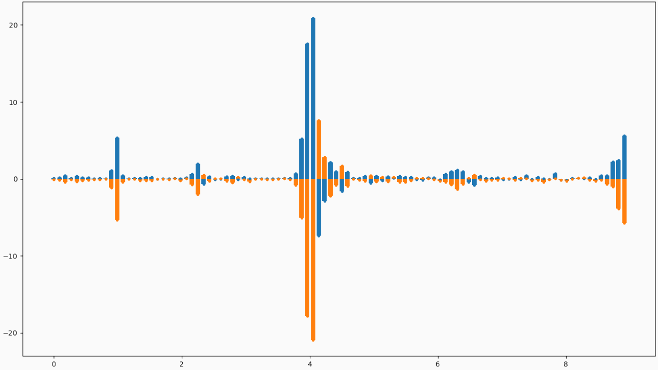

averages of $a_p(f)\sqrt{p}$ for f of weight $k \in [50, 250]$ and fixed $\epsilon_f$ (blue $= +1$, orange $= -1$), collated by $p/N$.

When generating data for classical modular forms in the LMFDB, Jonathan Bober implemented an (extremely fast) trace formula that can produce data for fixed level, rising weight modular forms.



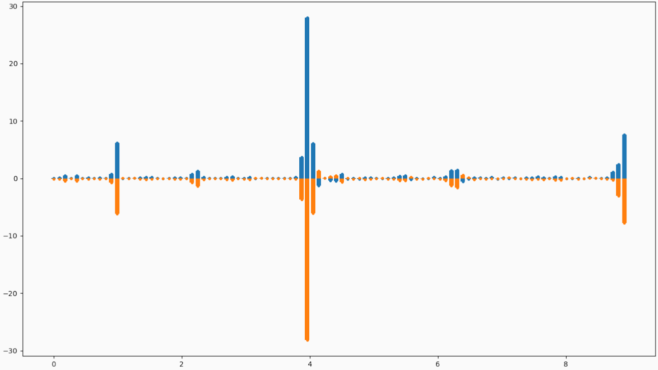

averages of $a_p(f)\sqrt{p}$ for f of weight $k \in [250, 600]$ and fixed $\epsilon_f$ (blue $= +1$, orange $= -1$), collated by $p/N$.

When generating data for classical modular forms in the LMFDB, Jonathan Bober implemented an (extremely fast) trace formula that can produce data for fixed level, rising weight modular forms.



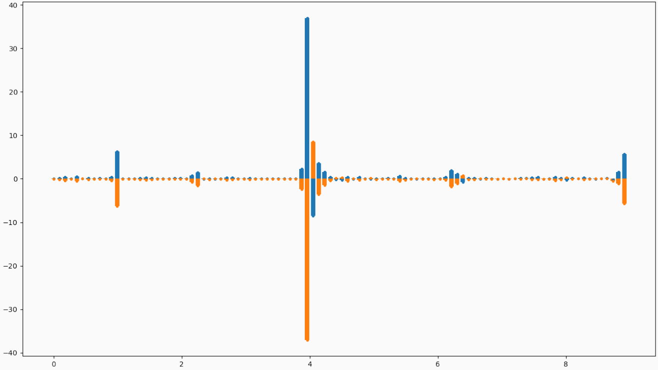

averages of $a_p(f)\sqrt{p}$ for f of weight $k \in [600, 1200]$ and fixed $\epsilon_f$ (blue $= +1$, orange $= -1$), collated by $p/N$.

When generating data for classical modular forms in the LMFDB, Jonathan Bober implemented an (extremely fast) trace formula that can produce data for fixed level, rising weight modular forms.




averages of $a_p(f)\sqrt{p}$ for f of weight $k \in [1200, 2400]$ and fixed $\epsilon_f$ (blue $= +1$, orange $= -1$), collated by $p/N$.

When generating data for classical modular forms in the LMFDB, Jonathan Bober implemented an (extremely fast) trace formula that can produce data for fixed level, rising weight modular forms.




averages of $a_p(f)\sqrt{p}$ for f of weight $k \in [2400, 3300]$ and fixed $\epsilon_f$ (blue $= +1$, orange $= -1$), collated by $p/N$.

It turns out that the murmurations tend towards a sequence of dirac delta functions. In arXiv:2310.07746, we prove the following.

**Theorem (Bober, Booker, Lee, Lowry-Duda)**

*Assume GRH. Fix $B \subset \mathbb{R}_{>0}$ compact with $|B| > 0$. Let $K, H \in \mathbb{R}_{>0}$ with $K^{\frac{5}{6}+\epsilon} < H < K^{1-\epsilon}$. Set $N$ to be the analytic conductor in $H_K(1)$. Then as $K \to \infty$,*

$$\frac{\sum_{p/N \in B} \log p \sum_{\substack{k \equiv 2^\delta \bmod 4 \\ |k-K| \leq H}} \sum_{f \in H_k} a_p(f)}{\sum_{p/N \in B} \log p \sum_{\substack{k \equiv 2^\delta \bmod 4 \\ |k-K| \leq H}} \sum_{f \in H_k} 1} = \frac{(-1)^\delta}{\sqrt{N}} \left( \frac{\nu(B)}{|B|} + o_{B,\epsilon}(1) \right),$$

*For an explicit measure $\nu(B)$, which looks more complicated to understand than it actually is.*

It turns out that the murmurations tend towards a sequence of dirac delta functions. In arXiv:2310.07746, we prove the following.

**Theorem (Bober, Booker, Lee, Lowry-Duda)**

*Assume GRH. Fix $B \subset \mathbb{R}_{>0}$ compact with $|B| > 0$. Let $K, H \in \mathbb{R}_{>0}$ with $K^{\frac{5}{6}+\epsilon} < H < K^{1-\epsilon}$. Set $N$ to be the analytic conductor in $H_K(1)$. Then as $K \to \infty$,*

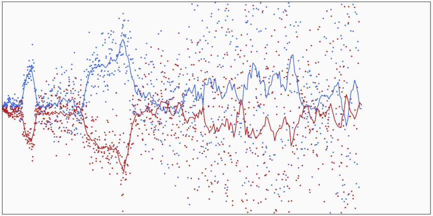
$$\frac{\sum_{p/N \in B} \log p \sum_{\substack{k \equiv 2^\delta \bmod 4 \\ |k-K| \leq H}} \sum_{f \in H_k} a_p(f)}{\sum_{p/N \in B} \log p \sum_{\substack{k \equiv 2^\delta \bmod 4 \\ |k-K| \leq H}} \sum_{f \in H_k} 1} = \frac{(-1)^\delta}{\sqrt{N}} \left( \frac{\nu(B)}{|B|} + o_{B,\epsilon}(1) \right),$$

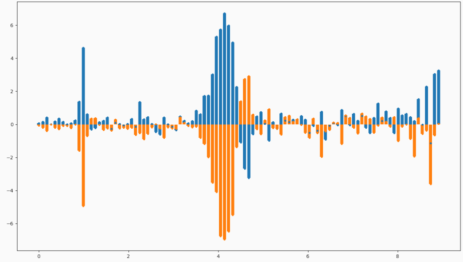
$$\text{where } \nu(B) = \frac{1}{\zeta(2)} \sideset{}{^*}\sum_{\substack{a,q \in \mathbb{Z}_{>0} \\ \gcd(a,q)=1 \\ (a/q)^{-2} \in B}} \frac{\mu(q)^2}{\varphi(q)^2 \sigma(q)} \left( \frac{q}{a} \right)^3$$

$$= \tfrac{1}{2} \sum_{t=-\infty}^{\infty} \prod_{p \nmid t} \frac{p^2 - p - 1}{p^2 - p} \cdot \int_B \cos\left( \frac{2\pi t}{\sqrt{y}} \right) dy.$$

We convinced ourselves that we would prove this in July of last year, and it was just a matter of ironing out details. But it turns out that we missed something that perhaps should have jumped out at us from numerics.

We convinced ourselves that we would prove this in July of last year, and it was just a matter of ironing out details. But it turns out that we missed something that perhaps should have jumped out at us from numerics.





# Surprise!

These murmurations are exactly the same.

**Theorem (Booker, Lee, Lowry-Duda, Seymour-Howell, Zubrilina)**

*Assume GRH. Fix* $B \subset \mathbb{R}_{>0}$ *compact with* $|B| > 0$*. Let* $R, H \in \mathbb{R}_{>0}$ *with* $R^{\frac{5}{6}+\epsilon} < H < R^{1-\epsilon}$*. Set* $N$ *to be the analytic conductor for Maass forms of level* $1$ *with eigenvalue parameter* $R$*. Then as* $R \to \infty$*,*

$$\frac{\sum_{p/N \in B} \log p \sum_{|r_j - R| \leq H} \epsilon_f a_p(f)}{\sum_{p/N \in B} \log p \sum_{|r_j - R| \leq H} 1} = \frac{1}{\sqrt{N}} \left( \frac{\nu(B)}{|B|} + o_{B,\epsilon}(1) \right),$$

*where* $\nu(B)$ *is exactly as above.*

# Aside on ML and Number Theory

All the work I describe today was inspired by machine learning.

I've been working a lot recently on trying to get *real*, *actual* mathematics from machine learning. I think it's worth discussing further.

1. percent correct isn't percent understanding ($\mu(n)$)
2. pattern recognition is amazing (clustering)
3. one-sided information oracle ($a(p)$ and signs)
4. what's next?

# Sketch of proof ideas

Superficially, the proofs here and in Zubrilina's work are all similar: begin with an explicit trace formula and try to make sense through repeated averaging.

But the trace formulas (and whether they use holomorphic or nonholomorphic objects) are very different.

We used an Eichler–Selberg trace formula to prove the holomorpic modular form case as the weights $k \to \infty$, and a variant of the Selberg trace formula due to Strömbergsson to prove the Maass form case.

For fixed level, rising weight holomorphic modular forms, the Eichler–Selberg trace formula shows

$$\sum_{f \in H_k} a_p(f) \approx \frac{(-1)^{\frac{k}{2}}}{\pi} \sum_{\substack{t \in \mathbb{Z} \\ t^2 < 4p}} \cos\left((k-1)\phi_{t,p}\right) L(1, \psi_{t^2-4p}),$$

where $\phi_{t,p} = \arcsin\left(\frac{t}{2\sqrt{p}}\right)$ and $\psi$ is a Dirichlet character.

We average over primes $p$ and weights $k$ (in a particular congruence class $2^\delta \bmod 4$), giving sums of the shape

$$\sum_p \sum_k \sum_{f \in S_k} a_p(f) \approx \frac{(-1)^\delta}{\pi} \sum_p \sum_{\substack{t \in \mathbb{Z} \\ t^2 < 4p}} \sum_k \cos\big((k-1)\phi_{t,p}\big) L(1, \psi_{t^2-4p}).$$

Recalling that $\phi_{t,p} = \arcsin\left(\frac{t}{2\sqrt{p}}\right) \in (-\pi/2, \pi/2)$ and that $k$ is summed over a fixed congruence class mod 4, we observe that the sum over $k$ strongly concentrates over values of $\phi_{t,p}$ that are close to either 0 or $\pm\frac{\pi}{2}$. But as there aren't many integers $t$ with $t^2$ near $4p$, the mass of the sum actually concentrates around $\phi_{t,p} \approx 0$.

In particular, we can focus on $t$ with $|t| < T$ for some (smaller) parameter $T$, and replace $\phi_{t,p}$ by a linear approximation. We are led to

$$\sum_p \sum_k \sum_{f \in S_k} a_p(f) \approx \frac{(-1)^\delta}{\pi} \sum_k \sum_p \sum_{\substack{t \in \mathbb{Z} \\ |t| < T}} \cos\left(\frac{(k-1)t}{2\sqrt{p}}\right) L(1, \psi_{t^2-4p}).$$

$$\sum_{p}\sum_{k}\sum_{f\in S_k} a_p(f) \approx \frac{(-1)^\delta}{\pi}\sum_{k}\sum_{p}\sum_{\substack{t\in\mathbb{Z}\\|t|<T}} \cos\left(\frac{(k-1)t}{2\sqrt{p}}\right) L(1,\psi_{t^2-4p}).$$

To handle the $L$-function and sum over primes, we show that $L(1,\psi_{t^2-4p})$ can be understood through studying the local average $L(1,\overline{\psi}_t)$, where

$$\overline{\psi}_t(m) = \frac{1}{\varphi(m^2)}\sum_{\substack{n \bmod m^2\\(n,m)=1}} \psi_{t^2-4n}(m).$$

In particular, for sufficiently nice $\Phi$

$$\sum_{p\in[A,B]} L(1,\psi_{t^2-4p})\Phi(p)\log p \approx L(1,\overline{\psi}_t)\int_A^B \Phi(u)\,du.$$

This lets us transform the sum over primes into an integral, which (after a change of variables) shows that

$$\sum_{p}\sum_{k}\sum_{f\in S_k} a_p(f) \approx \frac{(-1)^\delta}{\pi}\sum_{k}\left(\frac{k-1}{4\pi}\right)^2\int\sum_{t} L(1,\overline{\psi}_t)\cos(2\pi\alpha t)\frac{d\alpha}{\alpha^3}.$$

$$\sum_p \sum_k \sum_{f \in S_k} a_p(f) \approx \frac{(-1)^\delta}{\pi} \sum_k \left(\frac{k-1}{4\pi}\right)^2 \int \sum_t L(1, \overline{\psi}_t) \cos(2\pi\alpha t) \frac{d\alpha}{\alpha^3}.$$

Main terms come from angles $\alpha \approx \frac{a}{q}$ for rational numbers $\frac{a}{q}$ — to quantify this, we use the circle method.

The arithmetic content of the main terms come from the arithmetic behavior of $L(1, \overline{\psi}_t) \cos(2\pi\alpha t)$.

For Maass forms, there is a particularly annoying initial difference: we require an analytic test function for the Selberg trace formula. It thus cannot be compactly supported.

Nonetheless, the main contribution again comes from hyperbolic terms in the trace formula, corresponding to a sum

$$\sum_t L(1, \psi_{t^2+4p})\Phi(\cdot),$$

where the test function $\Phi(\cdot)$ is more complicated but which also concentrates the sum around $|t| \leq T$.

Although all the error handling is different (and the presence of a noncompact weight complicates every step), the computation of the main term is almost identical.

**Thank you very much.**

**Please note that these slides are (or will soon be) available on my website (davidlowryduda.com).**