Rigorous computation of Maass cusp forms of squarefree level

Andrei Seymour-Howell

University of Bristol

ANTS XV, 8th August 2022

Maass cusp forms

Let $\mathbb{H} = \{z = x + iy \mid y > 0\}$ denote the upper half-plane. We define the Hecke congruence subgroup $\Gamma_0(N) < \mathrm{SL}(2, \mathbb{Z})$ by

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z}) \mid c \equiv 0 \pmod{N} \right\}$$

for $N > 0$.

Maass cusp forms

Let $\mathbb{H} = \{z = x + iy \mid y > 0\}$ denote the upper half-plane. We define the Hecke congruence subgroup $\Gamma_0(N) < \mathrm{SL}(2, \mathbb{Z})$ by

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z}) \mid c \equiv 0 \pmod{N} \right\}$$

for $N > 0$. This group acts on $\mathbb{H}$ by linear fractional transformations, i.e

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{az + b}{cz + d} \qquad \forall \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N), z \in \mathbb{H}.$$

Maass cusp forms

Let $\mathbb{H} = \{z = x + iy \mid y > 0\}$ denote the upper half-plane. We define the Hecke congruence subgroup $\Gamma_0(N) < \mathrm{SL}(2, \mathbb{Z})$ by

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z}) \mid c \equiv 0 \pmod{N} \right\}$$

for $N > 0$. This group acts on $\mathbb{H}$ by linear fractional transformations, i.e

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{az + b}{cz + d} \qquad \forall \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N), z \in \mathbb{H}.$$

The modular surface $X = \Gamma_0(N) \backslash \mathbb{H}$ is a finite volume non-compact surface with Laplacian

$$\Delta = -y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right).$$

We also have the measure

$$\frac{dx\, dy}{y^2}.$$

Maass cusp forms

We call a function $f : \mathbb{H} \to \mathbb{C}$ a **Maass cusp form** of level N (trivial character) if

Maass cusp forms

We call a function $f : \mathbb{H} \to \mathbb{C}$ a **Maass cusp form** of level N (trivial character) if

1) f is an eigenfunction of the Laplacian, $\Delta f = \lambda f, \lambda \geq 0$,

Maass cusp forms

We call a function $f : \mathbb{H} \to \mathbb{C}$ a **Maass cusp form** of level N (trivial character) if

1) f is an eigenfunction of the Laplacian, $\Delta f = \lambda f, \lambda \geq 0$,
2) f is automorphic, $f(\gamma z) = f(z)$ for all $\gamma \in \Gamma_0(N)$,

Maass cusp forms

We call a function $f : \mathbb{H} \to \mathbb{C}$ a **Maass cusp form** of level N (trivial character) if

1) f is an eigenfunction of the Laplacian, $\Delta f = \lambda f, \lambda \geq 0$,
2) f is automorphic, $f(\gamma z) = f(z)$ for all $\gamma \in \Gamma_0(N)$,
3) $f \in L^2(\Gamma_0(N)\backslash\mathbb{H})$, i.e f is square-integrable,

Maass cusp forms

We call a function $f : \mathbb{H} \to \mathbb{C}$ a **Maass cusp form** of level N (trivial character) if

1) f is an eigenfunction of the Laplacian, $\Delta f = \lambda f, \lambda \geq 0$,
2) f is automorphic, $f(\gamma z) = f(z)$ for all $\gamma \in \Gamma_0(N)$,
3) $f \in L^2(\Gamma_0(N)\backslash\mathbb{H})$, i.e f is square-integrable,
4) f vanishes at all of the cusps of $\Gamma_0(N)\backslash\mathbb{H}$.

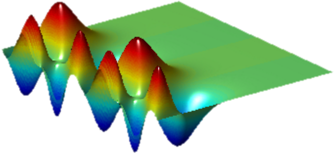
Maass cusp forms

We call a function $f : \mathbb{H} \to \mathbb{C}$ a **Maass cusp form** of level N (trivial character) if

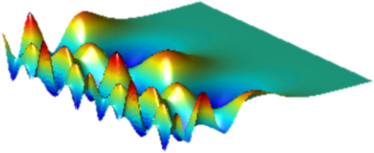
1) f is an eigenfunction of the Laplacian, $\Delta f = \lambda f, \lambda \geq 0$,
2) f is automorphic, $f(\gamma z) = f(z)$ for all $\gamma \in \Gamma_0(N)$,
3) $f \in L^2(\Gamma_0(N)\backslash\mathbb{H})$, i.e f is square-integrable,
4) f vanishes at all of the cusps of $\Gamma_0(N)\backslash\mathbb{H}$.

We will denote the space of Maass cusp forms of level N and Laplace eigenvalue λ by $\mathcal{S}_\lambda(N)$.
The set of functions that just satisfy points (2), (3) and (4) we shall denote as $L^2_{\text{cusp}}(\Gamma_0(N)\backslash\mathbb{H})$.

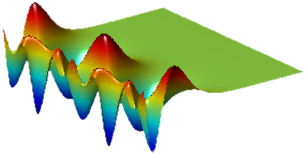
Pictures of Maass forms



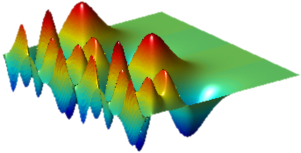
(a) Level 1, $\lambda = 91.141345\ldots$



(b) Level 1, $\lambda = 190.131547\ldots$



(c) Level 2, $\lambda = 79.867724\ldots$



(d) Level 3, $\lambda = 182.713668\ldots$

Figure: Images of Maass forms from the LMFDB.

Fourier expansion

For any $f \in \mathcal{S}_\lambda(N)$ and any non-zero integer n coprime to N, we have the **Hecke operator** T_n that maps $\mathcal{S}_\lambda(N) \to \mathcal{S}_\lambda(N)$. Further f has a Fourier expansion of the form

$$f(z) = f(x+iy) = \sum_{n \neq 0} a(n)\sqrt{y}K_{ir}(2\pi|n|y)\exp(2\pi inx)$$

where $K_\nu(u)$ is the K-Bessel function and $\lambda = \frac{1}{4} + r^2$.

Fourier expansion

For any $f \in \mathcal{S}_\lambda(N)$ and any non-zero integer n coprime to N, we have the **Hecke operator** T_n that maps $\mathcal{S}_\lambda(N) \to \mathcal{S}_\lambda(N)$. Further f has a Fourier expansion of the form

$$f(z) = f(x+iy) = \sum_{n \neq 0} a(n)\sqrt{y}K_{ir}(2\pi|n|y)\exp(2\pi inx)$$

where $K_\nu(u)$ is the K-Bessel function and $\lambda = \frac{1}{4} + r^2$.
If f is also a Hecke eigenfunction for all Hecke operators T_n with $(n, N) = 1$, i.e $T_n f = \lambda(n) f$, then we can normalise such that $a(1) = 1$ and we have

$$a(n) = \lambda(n),$$
$$a(-n) = \varepsilon\lambda(n),$$

where $\varepsilon = 1$ if $a(n) = a(-n)$ and -1 if $a(n) = -a(-n)$.

Fourier expansion

For any $f \in \mathcal{S}_\lambda(N)$ and any non-zero integer n coprime to N, we have the **Hecke operator** T_n that maps $\mathcal{S}_\lambda(N) \to \mathcal{S}_\lambda(N)$. Further f has a Fourier expansion of the form

$$f(z) = f(x+iy) = \sum_{n \neq 0} a(n)\sqrt{y}K_{ir}(2\pi|n|y)\exp(2\pi inx)$$

where $K_\nu(u)$ is the K-Bessel function and $\lambda = \frac{1}{4} + r^2$.
If f is also a Hecke eigenfunction for all Hecke operators T_n with $(n, N) = 1$, i.e $T_n f = \lambda(n)f$, then we can normalise such that $a(1) = 1$ and we have

$$a(n) = \lambda(n),$$
$$a(-n) = \varepsilon\lambda(n),$$

where $\varepsilon = 1$ if $a(n) = a(-n)$ and -1 if $a(n) = -a(-n)$.
From this, we know there exists an orthogonal basis $\{f_j\}$ in $L^2_{\text{cusp}}(\Gamma_0(N)\backslash\mathbb{H})$ consisting of eigenfunctions to all Hecke operators T_n with $(n, N) = 1$.

History of Computations

- In the 1990's Hejhal developed an algorithm to compute Maass cusp forms, that was generalised by Strömberg in 2006 to work for general congruence (and non-congruence) subgroups. This algorithm works very well in practice, however it relies on a heuristic argument and it has not been proven yet to rigorously converge to genuine Maass cusp forms.

History of Computations

- In the 1990's Hejhal developed an algorithm to compute Maass cusp forms, that was generalised by Strömberg in 2006 to work for general congruence (and non-congruence) subgroups. This algorithm works very well in practice, however it relies on a heuristic argument and it has not been proven yet to rigorously converge to genuine Maass cusp forms.
- In 2006 Booker, Strömbergsson and Venkatesh developed an algorithm to verify the computations of Maass cusps forms. This was originally only derived for level 1, however recent work of Child has generalised this to general level and character.

History of Computations

- In the 1990's Hejhal developed an algorithm to compute Maass cusp forms, that was generalised by Strömberg in 2006 to work for general congruence (and non-congruence) subgroups. This algorithm works very well in practice, however it relies on a heuristic argument and it has not been proven yet to rigorously converge to genuine Maass cusp forms.
- In 2006 Booker, Strömbergsson and Venkatesh developed an algorithm to verify the computations of Maass cusps forms. This was originally only derived for level 1, however recent work of Child has generalised this to general level and character.
- Also in 2006, Booker and Strömbergsson used the Selberg trace formula for computations of Maass cusp forms. However they were mainly focused on proving the non-existence of Maass forms in an interval to help prove the Selberg eigenvalue conjecture.

Selberg trace formula

The Selberg trace formula allows one to consider the whole spectrum of Maass cusp forms for a fixed level N. Selberg derived this in the 1950's to prove the existence of Maass cusp forms.

Selberg trace formula

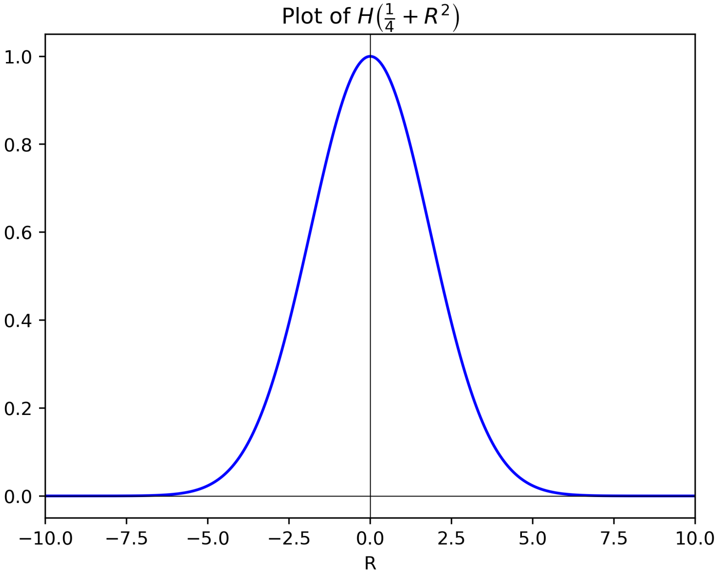
The Selberg trace formula allows one to consider the whole spectrum of Maass cusp forms for a fixed level N. Selberg derived this in the 1950's to prove the existence of Maass cusp forms.

In our case, if we have a Hecke eigenbasis $\{f_j\}$ of $L^2_{\text{cusp}}(\Gamma_0(N)\backslash\mathbb{H})$ with respective Laplace eigenvalues λ_j and Hecke eigenvalues $a_j(n)$, the Selberg trace formula allows us to compare

$$\text{(Spectral side)} \sum_{j=1}^{\infty} a_j(n)H(\lambda_j) = \text{(Geometric side)}$$

for some nice (analytic) test function H and $n \neq 0$.

Example of a test function



Aside to modular forms

Explicit versions of trace formulas for modular forms have been used to compute basis elements of the spaces of modular forms, for example in the pari command `mfeigenbasis`. Since these spaces are finite dimensional, one can use the Hecke operators and linear algebra to extract the Fourier coefficients of the basis elements.

Aside to modular forms

Explicit versions of trace formulas for modular forms have been used to compute basis elements of the spaces of modular forms, for example in the pari command `mfeigenbasis`. Since these spaces are finite dimensional, one can use the Hecke operators and linear algebra to extract the Fourier coefficients of the basis elements.

In the Maass form case, the spaces we will be working with are infinite dimensional, hence the requirement for the test function in the trace formula. Our idea is to choose a test function such that the contribution from the larger eigenvalues is negligible and we can then treat the problem as a finite linear algebra one.

Verification using the Selberg trace formula

Fix a level N. Let $\{f_j\}$ be a Hecke eigenbasis of $L^2_{\text{cusp}}(\Gamma_0(N)\backslash\mathbb{H})$ with respective Laplace eigenvalues λ_j such that $\lambda_1 \leq \lambda_2 \leq \ldots$. Let $a_j(n)$ be the Hecke eigenvalues of f_j, that is $T_n f_j = a_j(n) f_j$ for $(n, N) = 1$.

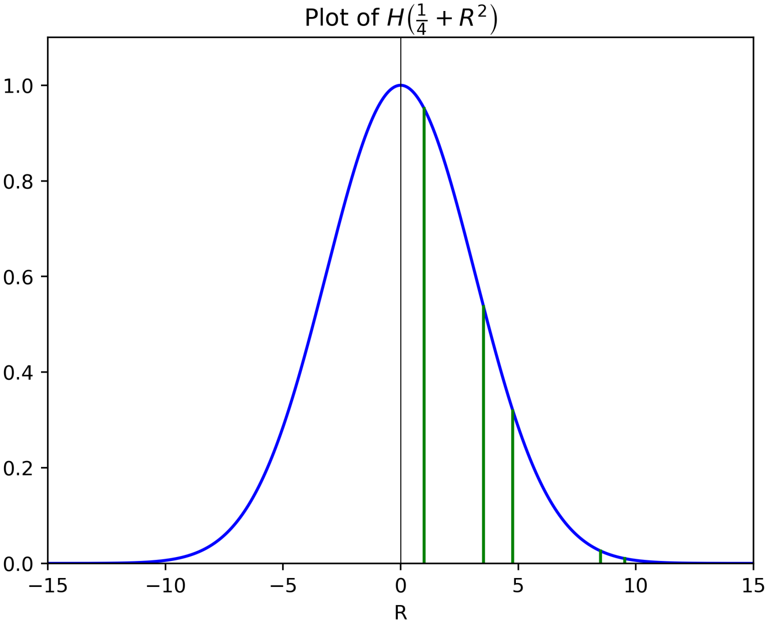
Verification using the Selberg trace formula

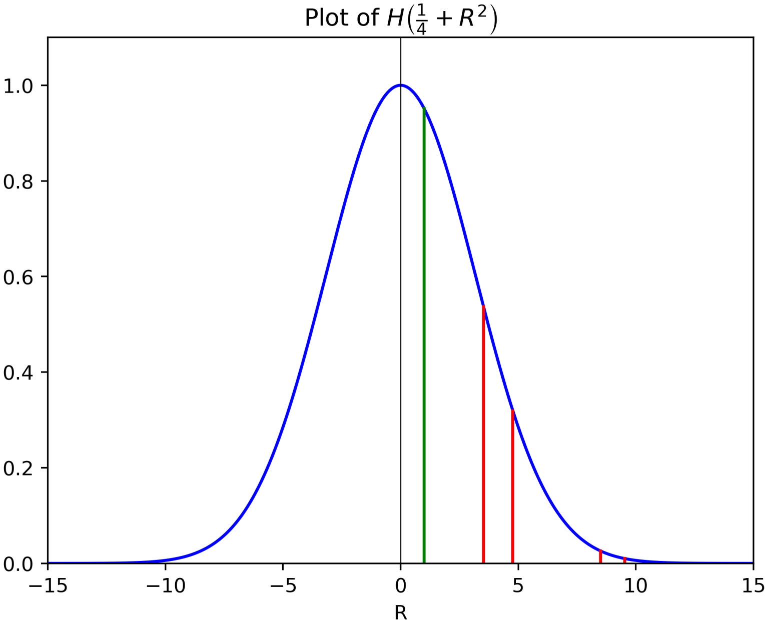
Fix a level N. Let $\{f_j\}$ be a Hecke eigenbasis of $L^2_{\text{cusp}}(\Gamma_0(N)\backslash\mathbb{H})$ with respective Laplace eigenvalues λ_j such that $\lambda_1 \leq \lambda_2 \leq \ldots$. Let $a_j(n)$ be the Hecke eigenvalues of f_j, that is $T_n f_j = a_j(n) f_j$ for $(n, N) = 1$.

We will fix a sufficiently nice test function H that is positive and monotonically decreasing for $\lambda > 0$. The Selberg trace formula allows us to compute

$$t(n, H) := \sum_{j=1}^{\infty} a_j(n) H(\lambda_j)$$

for any $n \neq 0$ and $(n, N) = 1$.


Plot of $H\left(\frac{1}{4}+R^2\right)$
1.0
0.8
0.6
0.4
0.2
0.0
−15
−10
−5
0
5
10
15
R


Plot of $H\left(\frac{1}{4}+R^2\right)$
1.0
0.8
0.6
0.4
0.2
0.0
−15
−10
−5
0
5
10
15
R

The details

Using the Hecke relations, we have for any real sequence of numbers $\{c(m)\}_{m=1}^{M}$ satisfying $c(m) = 0$ if $(m, N) > 1$, that

$$\left(\sum_{m=1}^{M} c(m) a_j(m) \right)^2 = \sum_{m_1=1}^{M} \sum_{m_2=1}^{M} c(m_1) c(m_2) \sum_{d | (m_1, m_2)} a_j \left(\frac{m_1 m_2}{d^2} \right).$$

The details

Using the Hecke relations, we have for any real sequence of numbers $\{c(m)\}_{m=1}^{M}$ satisfying $c(m) = 0$ if $(m, N) > 1$, that

$$\left(\sum_{m=1}^{M} c(m) a_j(m)\right)^2 = \sum_{m_1=1}^{M} \sum_{m_2=1}^{M} c(m_1) c(m_2) \sum_{d|(m_1,m_2)} a_j\left(\frac{m_1 m_2}{d^2}\right).$$

We define

$$\begin{aligned} Q(c, H) &:= \sum_{j=1}^{\infty} \left(\sum_{m=1}^{M} c(m) a_j(m)\right)^2 H(\lambda_j) \\ &= \sum_{m_1=1}^{M} \sum_{m_2=1}^{M} c(m_1) c(m_2) \sum_{d|(m_1,m_2)} \sum_{j=1}^{\infty} a_j\left(\frac{m_1 m_2}{d^2}\right) H(\lambda_j) \\ &= \sum_{m_1=1}^{M} \sum_{m_2=1}^{M} c(m_1) c(m_2) \sum_{d|(m_1,m_2)} t\left(\frac{m_1 m_2}{d^2}, H\right). \end{aligned}$$

Intuition

Suppose that we have putative numerical approximations $\tilde{\lambda}_j, \tilde{a}_j(m)$ to $\lambda_j, a_j(m)$. Then we want to choose numbers $c_i(m)$ such that $c_i(m) = 0$ if $(m, N) > 1$ and

$$\sum_{m=1}^{M} c_i(m)\tilde{a}_j(m) = \begin{cases} 1 & \text{if } j = i, \\ 0 & \text{otherwise.} \end{cases}$$

Intuition

Suppose that we have putative numerical approximations $\tilde{\lambda}_j, \tilde{a}_j(m)$ to $\lambda_j, a_j(m)$. Then we want to choose numbers $c_i(m)$ such that $c_i(m) = 0$ if $(m, N) > 1$ and

$$\sum_{m=1}^{M} c_i(m)\tilde{a}_j(m) = \begin{cases} 1 & \text{if } j = i, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\tilde{H}_i(\lambda) = H(\lambda)(\lambda - \tilde{\lambda}_i)^2$. For the verification we shall prove that there exists a Laplace eigenvalue near $\tilde{\lambda}_i$. For this, we use the definition of Q to compute

$$\varepsilon_i := \sqrt{\frac{Q(c_i, \tilde{H}_i)}{Q(c_i, H)}}.$$

Then there exists a cuspidal eigenvalue $\lambda \in [\tilde{\lambda}_i - \varepsilon_i, \tilde{\lambda}_i + \varepsilon_i]$.

Computing the Laplace eigenvalues

With the same H as before, we define $\tilde{H}(\lambda) = \lambda H(\lambda)$. Let Q and $\tilde{Q}$ denote the respective matrices of the quadratic forms $Q(c, H)$ and $Q(c, \tilde{H})$. To find the eigenvalues λ_j, we seek solutions to the generalised symmetric eigenvalue problem

$$\tilde{Q}x = \lambda Qx.$$

This gives us numerical approximations to the Laplace eigenvalues.

Computing the Laplace eigenvalues

With the same H as before, we define $\tilde{H}(\lambda) = \lambda H(\lambda)$. Let Q and $\tilde{Q}$ denote the respective matrices of the quadratic forms $Q(c, H)$ and $Q(c, \tilde{H})$. To find the eigenvalues λ_j, we seek solutions to the generalised symmetric eigenvalue problem

$$\tilde{Q}x = \lambda Qx.$$

This gives us numerical approximations to the Laplace eigenvalues. The corresponding eigenvectors we then use as the c_i. This way means we only have to do two matrix diagonalisations per level.

In practice

Suppose that we have putative numerical approximations $\tilde{\lambda}_j$. Define

$$\varepsilon_i := \sqrt{\frac{Q(c_i, \tilde{H}_i)}{Q(c_i, H)}},$$

for the c_i from the eigenvectors of the previous computation. Then, similar to before, there exists a cuspidal eigenvalue $\lambda \in [\tilde{\lambda}_i - \varepsilon_i, \tilde{\lambda}_i + \varepsilon_i]$ for each i.

In practice

Suppose that we have putative numerical approximations $\tilde{\lambda}_j$. Define

$$\varepsilon_i := \sqrt{\frac{Q(c_i, \tilde{H}_i)}{Q(c_i, H)}},$$

for the c_i from the eigenvectors of the previous computation. Then, similar to before, there exists a cuspidal eigenvalue $\lambda \in [\tilde{\lambda}_i - \varepsilon_i, \tilde{\lambda}_i + \varepsilon_i]$ for each i.

We can further prove that we did not miss any by essentially using our approximations to approximate the trace formula and measure the difference between the two values.

In practice

Suppose that we have putative numerical approximations $\tilde{\lambda}_j$. Define

$$\varepsilon_i := \sqrt{\frac{Q(c_i, \tilde{H}_i)}{Q(c_i, H)}},$$

for the c_i from the eigenvectors of the previous computation. Then, similar to before, there exists a cuspidal eigenvalue $\lambda \in [\tilde{\lambda}_i - \varepsilon_i, \tilde{\lambda}_i + \varepsilon_i]$ for each i.

We can further prove that we did not miss any by essentially using our approximations to approximate the trace formula and measure the difference between the two values.

In addition, we can also get rigorous error bounds on the Fourier coefficients $a_j(n)$.

Computational remarks

- The computations were done in C using a library called ARB, which does rigorous real and complex arithmetic with arbitrary precision. This is done using ball arithmetic, a form of interval arithmetic, which represents its numbers by a midpoint and radius.

Computational remarks

- The computations were done in C using a library called ARB, which does rigorous real and complex arithmetic with arbitrary precision. This is done using ball arithmetic, a form of interval arithmetic, which represents its numbers by a midpoint and radius.
- Currently this method has only been implemented for square-free level N due to availability of explicit forms of the Selberg trace formulas with Hecke operators.

Computational remarks

- The computations were done in C using a library called ARB, which does rigorous real and complex arithmetic with arbitrary precision. This is done using ball arithmetic, a form of interval arithmetic, which represents its numbers by a midpoint and radius.
- Currently this method has only been implemented for square-free level N due to availability of explicit forms of the Selberg trace formulas with Hecke operators.
- We computed 33214 Laplace eigenvalues of Maass forms with squarefree levels between $2 \leq N \leq 105$. The range of the ε_i's computed is between 10^{-15} and 10^{-2}.

Computational remarks

- The computations were done in C using a library called ARB, which does rigorous real and complex arithmetic with arbitrary precision. This is done using ball arithmetic, a form of interval arithmetic, which represents its numbers by a midpoint and radius.
- Currently this method has only been implemented for square-free level N due to availability of explicit forms of the Selberg trace formulas with Hecke operators.
- We computed 33214 Laplace eigenvalues of Maass forms with squarefree levels between $2 \leq N \leq 105$. The range of the ε_i's computed is between 10^{-15} and 10^{-2}.
- We also verified the Ramanujan-Petersson conjecture for prime $p \leq 2000$ for 13271 forms.

Sato-Tate

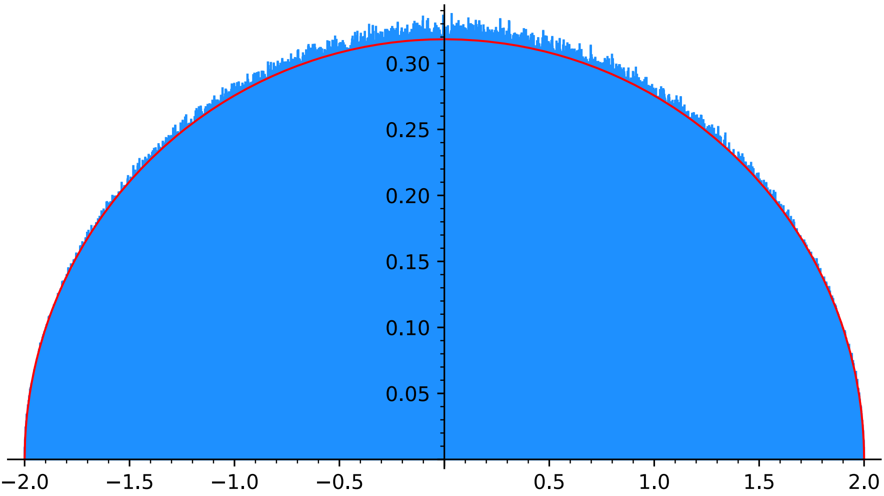


Figure: Distribution of the classical Sato-Tate distribution of the $a_j(p)$ for prime p.

Thanks for listening!

Example of trace formula, composite square-free level

$$\frac{\mu(N)\sigma_1(|n|)}{\sqrt{|n|}} h\left(\frac{i}{2}\right) + \sum_{j=1}^{\infty} h(R_j) a_j(n)$$

$$= \sum_{\substack{t\in\mathbb{Z} \\ \sqrt{D}=\sqrt{t^2-4n}\notin\mathbb{Q} \\ D>0}} c_N(D)\cdot g\left(\log\left(\frac{(|t|+\sqrt{D})^2}{4|n|}\right)\right)$$

$$+ \sum_{\substack{t\in\mathbb{Z} \\ \sqrt{D}=\sqrt{t^2-4n}\notin\mathbb{Q} \\ D<0}} c_N(D)\cdot \frac{\sqrt{|D/4n|}}{2\pi} \int_{-\infty}^{\infty} \frac{g(u)\cosh(u/2)}{\sinh^2(u/2)+|D/4n|} du$$

$$+ \left[\text{ if } \sqrt{n}\in\mathbb{Z} : \frac{\prod_{p|N}(p-1)}{12\sqrt{n}} \int_{-\infty}^{\infty} \frac{g'(u)}{\sinh\left(\frac{u}{2}\right)} du \right].$$

$$c_N(D) = \frac{L(1,\psi_d)}{l} \prod_{p|N} (\psi_d(p)-1) \prod_{p|l} \left[1 + (p-\psi_d(p))\frac{(l,p^\infty)-1}{p-1}\right].$$