Maass forms, nearing completion

David Lowry-Duda

December 2022

ICERM

December 2022 Meeting of Simons AGNTC

This is an update on my project to compute Maass forms. I last talked about this *at our November 2020* meeting. Since then, the broad plan to compute Maass forms rigorously has been unchanged, but the details have changed significantly.

This work has been heavily informed by Andy Booker and his soon-to-be PhD student Andrei Seymour-Howell at the University of Bristol. I should also thank Drew and Brendan for frequent support.

Our goal is to rigorously compute Maass forms. The problem is that everything associated to a generic Maass form is transcendental and one will never *exactly* compute a Maass form.

Each Maass form discussed today has an expansion

$$f(z) = \sum_{n \geq 1} \frac{a(m)}{\sqrt{m}} W_{ir}(2\pi m y) \, \mathrm{cs}(2\pi m x),$$

where $\mathrm{cs}(\cdot)$ is either $\cos(\cdot)$ or $\sin(\cdot)$, depending on the symmetry type of the Maass form. This is a (real) analytic function on $\Gamma_0(N)\backslash\mathcal{H}$ for some squarefree N, and it is an eigenfunction of a Laplacian with eigenvalue $\lambda = \frac{1}{4} + r^2$.

By "compute a Maass form", we mean to rigorously estimate the eigenvalue λ (or equivalently the spectral parameter r) to to rigorously estimate the coefficients $a(m)$.

Heuristic Evaluation

Last time I talked about Maass forms, I talked about *heuristic* computation of Maass forms using Hejhal's algorithm. This uses the modularity of $f(z)$ to construct an approximate homogenous linear system that can be solved to (heuristically) approximate a Maass form.[1]

To obtain better heuristics, one can iteratively apply Hejhal's algorithm.

Broadly, today I describe a way to make Hejhal's algorithm rigorous. To do that, we first briefly review Hejhal's algorithm.

[1] The details, including how to use Atkin-Lehner operators to reduce the dimension of the system and how to incorporate the various cusps, were the focus of my previous talk.

Hejhal's Algorithm

Let $f(z) = \sum_{n \geq 1} \frac{a(m)}{\sqrt{m}} W_{ir}(2\pi my)\,\mathrm{cs}(2\pi mx)$ denote a Maass form. For a fixed $H > 0$ and $Y > 0$, define the $2H$ points

$$z_m = x_m + iY = \frac{m - \frac{1}{2}}{2H} + it, \quad 1 - H \leq m \leq H.$$

These are equispaced points on a horocycle, and a form of Fourier inversion shows that

$$\frac{2}{H} \sum_{m=1}^{H} f(z_m)\,\mathrm{cs}(2\pi k x_m) =$$

$$\sum_{\epsilon \in \{-1,1\}} \sum_{\substack{n \geq 1 \\ n \equiv \epsilon k \bmod 2H}} (-1)^{\epsilon+1+(n-\epsilon k)/2H} \frac{a(n)}{\sqrt{n}} W_{ir}(2\pi nY).$$

Note that if we truncate the Fourier expansion of f at some L for some $L < H$, then the only term appearing on the RHS is $\frac{a(k)}{\sqrt{a(k)}} W_{ir}(2\pi kY)$.

This truncation introduces an error. We return to this error later.

To construct a nontrivial linear system, we choose Y small and use the modularity of f. If Y is sufficiently small (and we pretend that there is only one cusp), then $z_m = x_m + iY$ will be outside "the" fundamental domain for f. Let $z_m^* = x_m^* + iy_m^*$ denote the pullback of z_m to the fundamental domain, so that $f(z_m) = f(z_m^*)$.

With this choice of horocycle and the above computation, we find that

$$\frac{a(k)}{\sqrt{k}} W_{ir}(2\pi kY) = \frac{2}{H} \sum_{m=1}^{H} f(z_m^*) \operatorname{cs}(2\pi k x_m) + (\text{truncation error}).$$

We do this for each k with $1 \leq k \leq L$. Note that if we again substitute the (truncated) Fourier expansions for $f(z_m^*)$, this becomes a noisy homogenous linear system in terms of the L unknowns $\{a(k)\}_{1 \leq k \leq L}$.

Abstractly, we think of this linear system as being of the form

$$T(r)\vec{a} \approx 0,$$

where $\vec{a} = \big(a(k)\big)_{1 \le k \le L}^{T}$. For Maass newforms, $a(1) = 1$. Using this, we can remove the first column of the system of equations (corresponding to $a(1)$) and separate the first row as an auxiliary equation. Explicitly (and abusing notation), we have

$$T(r)\vec{a} \approx b(r),$$

where now $\vec{a} = \big(a(k)\big)_{{}_{2 \le k \le L}^{T}}$ and $b(r)$ are explicit in terms of the coefficients of $a(1) = 1$.

The auxiliary equation from the first row can be written

$$c(r) := \vec{a} \cdot v(r) + w(r) \approx 0.$$

Iteration

$$T(r)\vec{a} \approx b(r), \tag{1}$$

$$c(r) := \vec{a} \cdot v(r) + w(r) \approx 0. \tag{2}$$

Note that the matrix $T(r)$ and components $b(r), v(r)$, and $w(r)$ depend on the (a priori unknown) parameter r. One form of Hejhal's algorithm is to guess an initial r, solve (1) to get approximations to the coefficients $a(k)$, and then iterate while trying to minimize the error term in the auxiliary equation (2).

One part of making this rigorous is to find precise error bounds in the linear system (1) and the first coefficient auxiliary equation (2), including the error coming from truncation.

Additional Ingredient

The key difference from my original efforts to find Maass forms has been the unexpectedly strong success of a method to find low-precision estimates for the eigenvalues λ using an explicit, computational form of Selberg's trace formula. This is the topic of Andrei Seymour-Howell's PhD thesis, and he talked about this at ANTS earlier this year.

The broad strategy is now to first compute several intervals $[r - \epsilon, r + \epsilon]$ that are known to contain a unique eigenvalue parameter r^* (with explicit ϵ bound); we then use a rigorous version of Hejhal's algorithm to refine the intervals.

Aside

Andrei's work currently applies only to squarefree level and trivial inner character. In a closely related project, we are working to develop an explicit, computational trace formula for squarefull level and nontrivial character.

Removing $\approx$ signs

We now fix a single interval $[r - \epsilon, r + \epsilon]$ known to contain a unique (but unknown) eigenvalue r^*.

Write $r^* = r + \delta$.

To make things rigorous, we must remove $\approx$ signs from our Hejhal system

$$T(r)\vec{a} \approx b(r),$$
$$c(r) := \vec{a} \cdot v(r) + w(r) \approx 0.$$

To do that, let $b^\natural(r)$ denote the vector we get for r by truncating all Fourier series at L, setting up the system, and ignoring all truncation error terms. Then define

$$e = T(r^*)\vec{a} - b^\natural(r^*), \qquad b(r) := b^\natural(r) + e.$$

Now $b(r)$ is precisely defined (though we are ignorant of its exact value).

Let $\vec{a}(r)$ denote the computed solution for $\vec{a}$ at r in the dehomogenized, now well-defined Hejhal system

$$T(r)\vec{a} = b(r).$$

In practice, we prove bounds using this idealized form but compute in interval arithmetic. Though we don't know $b(r)$ exactly (because we don't know the error e exactly), we can bound the error e. With these definitions, it follows that $\vec{a}(r^*)$ is an *exact* solution for the coefficients of the desired Maass form, and thus the error e comes entirely from truncation error.

Bounding the sizes of the tails of the Fourier expansions for f allows us to compute interval estimates for $b(r)$, and thus interval estimates for $\vec{a}(r)$.

Core Idea

We are now ready to state the core idea of how to refine the error. Recall that the true eigenvalue parameter $r^* \in [r - \epsilon, r + \epsilon]$ and $r^* = r + \delta$. We compute (interval estimates) for $\vec{a}(r)$ and look at the auxiliary system

$$c(r) := \vec{a}(r) \cdot v(r) + w(r).$$

Near r^*, $c(r)$ has the expansion $c(r^*) = c(r) + c'(r)\delta + c''(\widetilde{r})\delta^2/2$ for some $\widetilde{r}$ between r and r^*. Rearranging, we find that

$$|\delta| = \frac{|c(r) - c(r^*)|}{|c'(r) + c''(\widetilde{r})\delta/2|}.$$

The core idea is to find tight (interval) approximations for $c(r)$ and $c'(r)$ and rigorously bound $c(r^*)$ and c''. As $|\delta| \leq \epsilon$, we find for example that if $\epsilon|c''(\widetilde{r})| < |c'(r)|$, then

$$|\delta| \leq 2\left|\frac{c(r) - c(r^*)}{c'(r)}\right| \approx \frac{\text{extremely small}}{c'(r^*)}.$$

If ϵ is small enough in terms of $c'(r^*)$ and $c''(r^*)$, then this strategy succeeds.

But this doesn't always succeed. In the worst-case scenario, we might have $c'(r^*) = 0$. I have never observed this to occur, but if it did then we would have to construct a different auxiliary equation and work from there.

More often, the problem is that our initial ϵ-interval estimate from the trace formula isn't quite strong enough. It is possible to work harder to produce better initial intervals, but there is more work to be done. This is a *later* project.

For initial upload into the LMFDB, I intend to indicate which Maass forms are rigorously certified.

For the remainder of this talk, I'll comment on estimating and bounding the components of the auxiliary expansion

$$c(r^*) = c(r) + c'(r)\delta + c''(\widetilde{r})\delta^2/2.$$

$c(r)$ and $c(r^*)$

Estimating $c(r)$ is "easy" as long as we can compute $T(r), b^\natural(r), v(r)$, and $w(r)$ very precisely. As previously noted, we can bound the truncation error e arbitrarily well by using enough coefficients, giving an arbitrarily good interval bound for $b(r)$. As long as $T(r)$ is invertible,[2] we compute interval bounds for $\vec{a}(r)$. These directly give interval bounds for $c(r)$.

At the eigenvalue r^*, we should expect $c(r^*) \approx 0$. The only error is also due to truncation error. We can bound this by essentially the same techniques as bounding the truncation error e for $b^\natural(r)$.

[2] If $T(r)$ is not invertible, then Hejhal's algorithm fails and we choose a different horocycle.

$c''(r)$

Bounding $c''(r)$ is the most technically involved aspect of the analysis. A major problem is that naive derivatives here involve derivatives of the Whittaker functions W_{ir}, for which we don't have particularly accurate estimates.[3]

Generally, we can approximate W_{ir} very well, but we can only efficiently compute rigorous bounds for general $\frac{\partial}{\partial r} W_{ir}$ and $\frac{\partial^2}{\partial r^2} W_{ir}$.

We must be careful about which terms we can approximate and which terms we can merely bound.

[3] A surprisingly large amount of my time on Maass forms has been centered on producing various rigorous estimates for Bessel functions.

Taylor-type approximations

For any matrix valued function $M(t)$ that is twice differentiable in a neighborhood of r, we define

$$M_1(r,\delta) := \begin{cases} \frac{M(r+\delta)-M(r)}{\delta} & \delta \neq 0, \\ M'(r) & \delta = 0, \end{cases}$$

$$M_2(r,\delta) := \begin{cases} \frac{M(r+\delta)-M(r)-M'(r)\delta}{\delta^2} & \delta \neq 0, \\ \frac{M''(r)}{2} & \delta = 0. \end{cases}$$

This implies that

$$M(r+\delta) = M(r) + M_1(r,\delta)\delta = M(r) + M'(r)\delta + M_2(r,\delta)\delta^2$$

for sufficiently small δ.

To bound c'', we actually bound $c_2(r,\delta)$.

Second order Hejhal approximation

The Hejhal system $T(t) \cdot \vec{a} = b(t)$ is smooth as W_{it} is smooth, and has first order version

$$(T(r) + T_1(r,\delta)\delta)(\vec{a}(r) + \vec{a}_1(r,\delta)\delta) = b(r) + b_1(r,\delta)\delta,$$

which implies that

$$T(r)\vec{a}_1(r,\delta) = b_1(r,\delta) - T_1(r,\delta)\vec{a}(r+\delta).$$

Fixing again δ so that $r^* = r + \delta$ (and performing similar analysis on the second order version) shows that

$$\begin{aligned}
\vec{a}_1(r,\delta) &= T^{-1}(r)(b_1(r,\delta) - T_1(r,\delta)\vec{a}(r^*)) \\
\vec{a}_2(r,\delta) &= T^{-1}(r)(b_2(r,\delta) - T'(r)\vec{a}_1(r,\delta) - T_2(r,\delta)\vec{a}(r^*)).
\end{aligned}$$

With bounds for the first and partial r-derivatives of $W_{ir}(y)$, we can bound $T'(r)$, $T_1(r,\delta)$, $T_2(r,\delta)$, $b_1(r,\delta)$, and $b_2(r,\delta)$. We bound the exact coefficients $\vec{a}(r^*)$ using the Kim–Sarnak bound. Thus we can bound $\vec{a}_1$ and $\vec{a}_2$.

Similar analysis on the second order version, as well as the first and second order versions of the auxiliary equation, ultimately show that

$$c_2(r,\delta) = w_2(r,\delta) + \vec{a}_2(r,\delta) \cdot v(r) + \vec{a}_1(r,\delta) \cdot v'(r) + \vec{a}(r+\delta) \cdot v_2(r,\delta).$$

Using the bounds for $\vec{a}_1$ and $\vec{a}_2$ from before, we can bound $c_2(r,\delta)$.

$c'(r)$

Finally, we approximate $c'(r)$. In principle, we could do this through a first order approximation of the Hejhal system, as we did with c_2 above. But this would require a high accuracy algorithm to compute (not just bound) $\partial_r W_{ir}$.

Instead, we take a first order approximation to $c'(r)$ and apply our bound for c_2, via

$$c'(r) = \frac{c(r + \delta_0) - c(r)}{\delta_0} + c_2(r, \delta_0)\delta_0.$$

We precisely approximated $c(r)$ by solving Hejhal's linear system, and we do this again by solving another linear system for $c(r + \delta_0)$ for a small δ_0.

The final term $c_2(r, \delta_0)$ can be bounded as before, except that instead of using the Kim-Sarnak bound for idealized coefficients $\vec{a}(r^*)$, we use the fact that we now have explicitly computed coefficients $\vec{a}(r + \delta_0)$ and use their actual values.

Thank you very much.

There will be more details in a forthcoming preprint, and data on an LMFDB near you.