IMRN International Mathematics Research Notices
2005, No. 34

On the Asymptotic Distribution of Zeros of Modular Forms

Zeév Rudnick

1 Introduction

1.1 Preface

Our purpose in this paper is to study the limiting distribution of zeros of modular forms. We review some definitions: a modular form of weight k for $SL_2(\mathbb{Z})$ is a holomorphic function on the upper half-plane $\mathbb{H}$, transforming as $f((az+b)/(cz+d)) = (cz+d)^k f(z)$, for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$ (this forces k to be even), and "holomorphic at the cusp" (see Section 2.1). A form is *cuspidal* if it vanishes at the cusp. For a modular form of weight k, let $\nu(f)$ be the number of inequivalent zeros of f in $\mathbb{H}$, with the convention that a zero at z is counted with weight $w(z)$ inverse to the number of elements of $SL_2(\mathbb{Z})/\{\pm I\}$ fixing z. Then $\nu(f) \le k/12$.

For a sequence of modular forms, where we assume that the number of inequivalent zeros tends to infinity, we would like to examine the manner in which the resulting configuration of zeros is distributed in the modular domain $SL_2(\mathbb{Z})\backslash\mathbb{H}$.

1.2 Example: Eisenstein series

The Eisenstein series are noncuspidal modular forms of weight $k > 2$, given by the sum

$$E_k(z) = \frac{1}{2} \sum_{(c,d)=1} \frac{1}{(cz+d)^k}. \tag{1.1}$$

Received 15 June 2005.
Communicated by Peter Sarnak.

Rankin and Swinnerton-Dyer [13] showed that all zeros (in the fundamental domain) of E_k lie on boundary arc $\{|z| = 1\}$. Moreover, as $k \to \infty$, the zeros become uniformly distributed on the unit arc. Rankin [14] gave a similar result for the (cuspidal) Poincaré series

$$G_k(z, m) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \frac{\exp(2\pi i m \gamma z)}{(cz + d)^k} \tag{1.2}$$

with $m \geq 1$ fixed, $k \gg m$. See [1] for yet another such example.

1.3 Example: higher Weierstrass points

An example in a slightly different setting is the higher Weierstrass points on a (compact) Riemann surface C. Let $S_k(C)$ be the space of holomorphic forms of weight k on C, $d := \dim S_k(C)$, and let $f_1, \ldots, f_d$ be a basis for $S_k(C)$. The Wronskian is defined as

$$\mathrm{Wr}_k\left(\{f_j\}; z\right) := \det\left(f_j^{(i)}\right)_{1 \leq i, j \leq d}. \tag{1.3}$$

It is a modular form of weight $d(k + d - 1)$.

A k*th-order Weierstrass point* is a zero of Wr_k. Equivalently, these are points of C where there is a nonzero form which vanishes to order $\geq \dim S_k(C)$.

It was conjectured by Bers and proved by Olsen [10] that as $k \to \infty$, these become dense in C. Mumford (see [7, page 11]) and Neeman [8] showed that in fact these become equidistributed with respect to the Arakelov (or Bergmann) measure on C, which arises from the metric on C gotten by embedding the curve in its Jacobian and pulling back the flat metric.

1.4 Cuspidal Hecke eigenforms

In contrast to these examples, we consider the case of cuspidal Hecke eigenforms. Recall that the Hecke operators

$$T_n f(z) := \frac{1}{n} \sum_{ad = n} a^k \sum_{b \bmod d} f\left(\frac{az + b}{d}\right) \tag{1.4}$$

act on the space of cusp forms of weight k, commute with each other, and are selfadjoint with respect to the Petersson inner product

$$\langle f, g\rangle := \int_{SL_2(\mathbb{Z})\backslash\mathbb{H}} f(z)\overline{g(z)}y^k \frac{dx\,dy}{y^2}. \tag{1.5}$$

Thus the space of cusp forms admits a basis consisting of joint eigenfunctions of all Hecke operators (the Eisenstein series E_k is also an eigenfunction). The cuspidal Hecke eigenfunctions have a simple zero at the cusp. For these Hecke eigenforms, we have the following.

Theorem 1.1. Assume the generalized Riemann hypothesis (GRH). Let $\{f_k\}$ be a sequence of cuspidal Hecke eigenforms, then as $k \to \infty$, their zeros are equidistributed in $SL_2(\mathbb{Z})\backslash\mathbb{H}$ with respect to the normalized hyperbolic measure $dV(z) = (3/\pi)(dxdy/y^2)$. □

Equidistribution here means that for any nice *compact* subset $\Omega \subset SL_2(\mathbb{Z})\backslash\mathbb{H}$, the proportion of (inequivalent) zeros of such a Hecke eigenform which lie in Ω is asymptotically the relative area of Ω:

$$\frac{1}{\nu(f)} \sum_{z\in\Omega} w(z)\,\mathrm{ord}_z(f) \sim \int_\Omega dV(z). \tag{1.6}$$

1.5 Equidistribution of masses

Theorem 1.1 is the outgrowth of some ideas from the theory of "quantum chaos." The mechanism is that equidistribution of the zeros of a sequence of forms f_k is implied by equidistribution of the "masses" $y^k|f_k(z)|^2 dV(z)$ of the forms. This idea was discovered by Nonnenmacher and Voros [9] in the context of quantum maps. Around the same time a very general result of this sort was obtained by Shiffman and Zelditch [18] for zeros of sections of high powers of a positive holomorphic Hermitian line bundle over any compact complex manifold ([9] deals with curves of genus one). All these are in a compact setting. The analogous result in our case is the following.

Theorem 1.2. Suppose that $\{f_k\}$ is a sequence of L^2-normalized cusp forms (f_k of weight k) for which the bulk of zeros lies in the fundamental domain: $\nu(f_k) \sim k/12$. Assume that for some $c > 0$,

$$y^k\left|f_k(z)\right|^2 dV(z) \xrightarrow{w} c \cdot dV(z). \tag{1.7}$$

Then the zeros of f_k are equidistributed with respect to $dV(z)$. □

Here $\xrightarrow{w}$ denotes weak convergence when we test against compactly supported functions.

The proof of Theorem 1.2, given in Section 3, follows closely the argument of [9, 18], with care taken to handle the complications due to the lack of compactness of the modular domain. We require the hypothesis (1.7), which is weaker than equidistribution of the masses as it allows "leakage" of some of the mass at the cusp (which cannot happen in a compact setting). This slightly weaker version of equidistribution of masses is what Lindenstrauss [5] proved unconditionally for the analogous case of Maass forms, though he cannot exclude $c = 0$. In the holomorphic case, we do not have an unconditional proof of (1.7). However, it has been known for some time to follow (with $c = 1$) from GRH [16, 21]. This is sketched in Section 4.

2 Potential theory on $\Gamma\backslash\mathbb{H}$

2.1 Preliminaries

We review some definitions: a modular form of weight k for $SL_2(\mathbb{Z})$ is a holomorphic function on the upper half-plane $\mathbb{H}$, transforming as

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z), \quad \forall \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \tag{2.1}$$

(this forces k to be even), and *holomorphic at the cusp*. This means the following: since $f(z)$ is periodic, it can be expressed as a holomorphic function $\tilde{f}(q)$ of $q = e^{2\pi i z}$ in the punctured disk $0 < |q| < 1$. The requirement to be holomorphic at the cusp means that it extends to a holomorphic function at $q = 0$. A form is *cuspidal* if it vanishes at the cusp, that is, $\tilde{f}(0) = 0$. The order of vanishing at the cusp $\mathrm{ord}_\infty(f)$ is defined as the order of vanishing at $q = 0$ of $\tilde{f}(q)$.

We denote by $\Gamma = SL_2(\mathbb{Z})/\{\pm I\}$ and speak interchangeably about modular forms for Γ and for $SL_2(\mathbb{Z})$. For each $z \in \mathbb{H}$, let $\Gamma_z = \{\gamma \in \gamma : \gamma z = z\}$ be the stabilizer in Γ of z, and set

$$w(z) := \frac{1}{\#\Gamma_z} = \begin{cases} \frac{1}{2}, & z \in \Gamma i, \\ \frac{1}{3}, & z \in \Gamma e^{2\pi i/3}, \\ 1, & \text{otherwise.} \end{cases} \tag{2.2}$$

For a modular form of weight k, let

$$\nu(f) = \sum_{z \in \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}} w(z)\,\mathrm{ord}_z(f) \tag{2.3}$$

be the (weighted) number of Γ-inequivalent zeros of f in $\mathbb{H}$. Then $\nu(f) \le k/12$, in fact, $\nu(f) + \mathrm{ord}_\infty(f) = k/12$. Note that $\nu(f) = 0$ only for powers of the modular discriminant.

2.2 A fundamental identity

For a smooth compactly supported function $\phi \in C_c^\infty(\mathbb{H})$ on the upper half-plane $\mathbb{H}$, set

$$F_\phi(z) := \sum_{\gamma\in\Gamma} \phi(\gamma z) \in C_c^\infty(\Gamma\backslash\mathbb{H}). \tag{2.4}$$

Note that

$$\int_{\Gamma\backslash\mathbb{H}} F_\phi(z)\frac{dx\,dy}{y^2} = \int_{\mathbb{H}} \phi(z)dx\,dy. \tag{2.5}$$

Let $dV(z) = (1/\mathrm{vol}(\Gamma\backslash\mathbb{H}))(dx\,dy/y^2)$ be the normalized hyperbolic measure on the quotient $\Gamma\backslash\mathbb{H}$.

Lemma 2.1. Let f be a (weakly holomorphic) modular form of weight k for Γ and let $\{z_j\}$ be a set of Γ-inequivalent zeros of f in $\mathbb{H}$. Then

$$\begin{aligned}\sum_j w(z_j)F_\phi(z_j) &= k\frac{\mathrm{vol}(\Gamma\backslash\mathbb{H})}{4\pi}\int_{\Gamma\backslash\mathbb{H}} F_\phi(z)dV(z)\\ &\quad + \frac{1}{2\pi}\int_{\mathbb{H}} \log\left(y^{k/2}|f(z)|\right)\Delta\phi(z)dx\,dy.\end{aligned} \tag{2.6}$$

□

Proof. Let $S = \{z \in \mathbb{H} : f(z) = 0\}$ be the set of zeros of the form f. It is a discrete set of points, which is stable under Γ ($\gamma S = S$ for all $\gamma \in \Gamma$). For any such S, we have

$$\sum_{z\in\Gamma\backslash S} w(z)F_\phi(z) = \sum_{s\in S}\phi(s). \tag{2.7}$$

Indeed,

$$\begin{aligned}\sum_{s\in S}\phi(s) &= \sum_{z\in\Gamma\backslash S}\sum_{s\in\Gamma z}\phi(s) = \sum_{z\in\Gamma\backslash S}\sum_{\gamma\in\Gamma/\Gamma_z}\phi(\gamma z)\\ &= \sum_{z\in\Gamma\backslash S}\frac{1}{\#\Gamma_z}\sum_{\gamma\in\Gamma}\phi(\gamma z) = \sum_{z\in\Gamma\backslash S}\frac{1}{\#\Gamma_z}F_\phi(z)\end{aligned} \tag{2.8}$$

as required.

We recall that the electrostatic potential for a point charge in the plane is $(1/2\pi)\log|z|$, that is, for $\phi \in C_c^\infty(\mathbb{C})$, we have

$$\int_{\mathbb{C}} \frac{1}{2\pi} \log|z| \Delta\phi(z)\, dx\, dy = \phi(0). \tag{2.9}$$

Consequently, for the holomorphic function $f(z)$ on $\mathbb{H}$, we have

$$\int_{\mathbb{H}} \frac{1}{2\pi} \log\left|f(z)\right| \Delta\phi(z)\, dx\, dy = \sum_{s:f(s)=0} \phi(s) \tag{2.10}$$

for all $\phi \in C_c^\infty(\mathbb{H})$ (with multiple zeros repeated). Thus

$$\begin{aligned}\sum_j w(z_j) F_\phi(z_j) &= \frac{1}{2\pi}\int_{\mathbb{H}} \log\left|f(z)\right| \Delta\phi(z)\, dx\, dy \\ &= -\frac{1}{2\pi}\int_{\mathbb{H}} \log y^{k/2} \Delta\phi(z)\, dx\, dy \\ &\quad + \frac{1}{2\pi}\int_{\mathbb{H}} \log\left(y^{k/2}\left|f(z)\right|\right) \Delta\phi(z)\, dx\, dy.\end{aligned} \tag{2.11}$$

The first term above is transformed via integration by parts into

$$-\frac{1}{2\pi}\int_{\mathbb{H}} \Delta\left(\log y^{k/2}\right)\phi(z)\, dx\, dy = \frac{k}{4\pi}\int_{\mathbb{H}} \phi(z)\frac{dx\, dy}{y^2}. \tag{2.12}$$

Note that

$$\int_{\mathbb{H}} \phi(z)\frac{dx\, dy}{y^2} = \int_{\Gamma\backslash\mathbb{H}} F_\phi(z)\frac{dx\, dy}{y^2} = \mathrm{vol}(\Gamma\backslash\mathbb{H})\int_{\Gamma\backslash\mathbb{H}} F_\phi(z)\, dV(z) \tag{2.13}$$

so that we get

$$-\frac{1}{2\pi}\int_{\mathbb{H}} \Delta\left(\log y^{k/2}\right)\phi(z)\, dx\, dy = k\frac{\mathrm{vol}(\Gamma\backslash\mathbb{H})}{4\pi}\int_{\Gamma\backslash\mathbb{H}} F_\phi(z)\, dV(z) \tag{2.14}$$

as required. ■

Remark 2.2. We may reformulate this in Γ-invariant form as follows: instead of the Euclidean Laplacian $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2$, we use the hyperbolic Laplacian

$$\mathcal{L} = y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) \tag{2.15}$$

which is Γ-invariant, and instead of $|f(z)|$, we use $y^k|f(z)|^2$ which is Γ-invariant. Then

$$\begin{aligned}\sum_j w(z_j)F_\phi(z_j) = k\frac{\mathrm{vol}(\Gamma\backslash\mathbb{H})}{4\pi}\int_{\Gamma\backslash\mathbb{H}} F_\phi(z)dV(z) \\ + \frac{\mathrm{vol}(\Gamma\backslash\mathbb{H})}{2\pi}\int_{\Gamma\backslash\mathbb{H}} \log\left(y^{k/2}\left|f(z)\right|\right)\mathcal{L}F_\phi(z)\frac{dx\,dy}{y^2}.\end{aligned} \tag{2.16}$$

To derive (2.16) from **Lemma 2.1**, we transform the second term in (2.6) by noting that since $dxdy/y^2$ is Γ-invariant, we have

$$\begin{aligned}&\frac{1}{2\pi}\int_{\mathbb{H}} \log\left(y^{k/2}\left|f(z)\right|\right)\mathcal{L}\phi(z)\frac{dx\,dy}{y^2} \\ &\quad = \frac{1}{2\pi}\int_{\Gamma\backslash\mathbb{H}} \log\left(y^{k/2}\left|f(z)\right|\right)\sum_{\gamma\in\Gamma}(\mathcal{L}\phi)(\gamma z)\frac{dx\,dy}{y^2} \\ &\quad = \frac{1}{2\pi}\int_{\Gamma\backslash\mathbb{H}} \log\left(y^{k/2}\left|f(z)\right|\right)F_{\mathcal{L}\phi}(z)\frac{dx\,dy}{y^2} \\ &\quad = \frac{\mathrm{vol}(\Gamma\backslash\mathbb{H})}{2\pi}\int_{\Gamma\backslash\mathbb{H}} \log\left(y^{k/2}\left|f(z)\right|\right)F_{\mathcal{L}\phi}(z)dV(z).\end{aligned} \tag{2.17}$$

This proves (2.16) once we note that $F_{\mathcal{L}\phi} = \mathcal{L}F_\phi$ since $\mathcal{L}$ is Γ-invariant.

2.3 Background on subharmonic functions

Let $\Omega \subset \mathbb{C} = \mathbb{R}^2$ be an open connected set. An upper semicontinuous[1] function $u : \Omega \to [-\infty, \infty)$ is *subharmonic* if u is not identically $-\infty$ and for all $z \in \Omega$,

$$u(z) \le \frac{1}{2\pi}\int_0^\infty u\left(z + re^{i\theta}\right)d\theta \tag{2.18}$$

whenever the disc $\{w : |z - w| \le r\}$ is contained in Ω.

Note. u is subharmonic $\Leftrightarrow \Delta u \ge 0$ as a distribution.

[1] $u : \Omega \to [-\infty, \infty)$ is upper semicontinuous if for all $c \in \mathbb{R}$, the set $\{z : u(z) < c\}$ is open.

A fundamental example of a subharmonic function is $\log|f(z)|$ where $f(z)$ is holomorphic in Ω. Moreover, we have

$$\Delta \log\left|f(z)\right| = 2\pi \sum_{f(z_j)=0} \delta\left(z - z_j\right). \tag{2.19}$$

We denote the space of subharmonic functions on Ω by $\mathcal{SH}(\Omega)$.

A basic compactness property of subharmonic functions is the following (cf. [2, Theorem 4.1.9]).

Lemma 2.3 (compactness property)**.** If $\{u_j\} \subset \mathcal{SH}(\Omega)$ are locally uniformly upper bounded (i.e., for all compact $K \subset \Omega$, there is c_K so that $u_j(z) \leq c_K$ for all $z \in K$ and all j), then either

(1) $u_j \to -\infty$ uniformly on compacta, or
(2) there is a subsequence u_{j_k} which converges weakly to some $u \in \mathcal{SH}(\Omega)$.
(3) Moreover, in this case, $\limsup u_j \leq u$ and $\limsup u_j = u$ almost everywhere.

□

The following is known as "Hartog's lemma" (see [15, Theorem 3.4.3]).

Lemma 2.4. If $\{u_j\} \subset \mathcal{SH}(\Omega)$ are locally uniformly bounded above and there is a continuous function $\phi : \Omega \to \mathbb{R}$ such that $\limsup u_j \leq \phi$, then $\max(u_j, \phi) \to \phi$ locally uniformly on Ω as $j \to \infty$. □

3 Ergodicity of eigenfunctions implies equidistribution of zeros

3.1 Proof of Theorem 1.2

Assume that there is some $c > 0$ so that $y^k|f_k|^2 dV \xrightarrow{w} c\,dV$ for a sequence of $k \to \infty$. We will show that necessarily the zeros $\{z_j\}$ of f_k become equidistributed relative to dV as $k \to \infty$. We need to show that for all $F \in C_c^\infty(\Gamma\backslash\mathbb{H})$, we have

$$\frac{1}{\nu(f_k)} \sum_j w(z_j)F(z_j) \sim \int_{\Gamma\backslash\mathbb{H}} F(z)dV(z), \quad k \longrightarrow \infty. \tag{3.1}$$

It suffices to show this for F of the form $F(z) = F_\phi(z) = \sum_{\gamma\in\Gamma} \phi(\gamma z)$, $\phi \in C_c^\infty(\mathbb{H})$, since these are dense in $C_c^\infty(\Gamma\backslash\mathbb{H})$ with respect to the uniform topology.[2] For these we have the

[2]To see this it suffices to show that for fixed ball $B \subset \mathbb{H}$, the space $\{F_\phi : \operatorname{supp}\phi \subset B\}$ is dense in $C(\pi(B))$, $\pi : \mathbb{H} \to \Gamma\backslash\mathbb{H}$ the projection, which follows by applying the Stone-Weierstrass theorem. Note that the uniform closure of the space $\{F_\phi : \phi \in C_c^\infty(\mathbb{H})\}$ is $C_0(\Gamma\backslash\mathbb{H})$, the continuous functions vanishing at infinity.

fundamental identity of Lemma 2.1:

$$\frac{1}{\nu(f)}\sum_j w(z_j)F_\phi(z_j) = \frac{\frac{k}{12}}{\nu(f_k)}\int_{\Gamma\backslash\mathbb{H}} F_\phi(z)dV(z) + \mathcal{E}, \tag{3.2}$$

where

$$\mathcal{E} = \frac{1}{2\pi\nu(f_k)}\int_{\mathbb{H}} \log\left(y^k|f_k(z)|^2\right)\Delta\phi(z)dx\,dy. \tag{3.3}$$

Since we assume that $\nu(f_k) \sim k/12$, the main term is the desired one (the mean value of F_ϕ) and we need to show that $\mathcal{E} \to 0$ as $k \to \infty$, that is, that for all $\phi \in C_c^\infty(\mathbb{H})$,

$$\int_{\mathbb{H}} \frac{1}{k}\log\left(y^k|f_k(z)|^2\right)\Delta\phi(z)dx\,dy \longrightarrow 0, \tag{3.4}$$

or that

$$\int_{\mathbb{H}} \frac{1}{k}\log\left(|f_k(z)|^2\right)\Delta\phi(z)dx\,dy \longrightarrow \int_{\mathbb{H}} -\log y\Delta\phi(z)dx\,dy. \tag{3.5}$$

Assume that this is false, that is, there are some test function $\phi_0 \in C_c^\infty(\mathbb{H}/\Gamma)$ and a subsequence of the f_k (which we will continue to call f_k for notational ease) for which (3.5) fails. We will proceed to derive a contradiction.

We begin by listing some properties of $\nu_k := (1/k)\log|f_k^2|$. These are subharmonic in $\mathbb{H}/\Gamma$. By Proposition A.1, we have $y^k|f_k(z)|^2 \ll k$ uniformly on compacta and so

$$\nu_k \le -\log y + \frac{\log k + O(1)}{k}, \tag{3.6}$$

thus

$$\limsup \nu_k \le -\log y \tag{3.7}$$

locally uniformly. Thus the family $\{\nu_k\} \subset \mathcal{SH}$ is locally bounded above. By the compactness property (Lemma 2.3), this gives us two possibilities:

(i) $\nu_k \to -\infty$ locally uniformly, or

(ii) $\{\nu_k\}$ has a weakly convergent subsequence.

We will dispose of both possibilities.

Option (i). On the support of the test function ϕ_0, we have $\nu_k \to -\infty$ uniformly, so that there is some K so that for all $k \ge K$ and all $z \in \operatorname{supp}\phi_0$ we have

$$\nu_k(z) \le -2H, \tag{3.8}$$

where $H = \max\{\operatorname{Im} z : z \in \operatorname{supp} \phi_0\}$. But then exponentiating, we find

$$\left|f_k\right|^2 \leq e^{-2kH} \tag{3.9}$$

so that for all ϕ supported inside $\operatorname{supp}\phi_0$ (and so that $\operatorname{supp}\phi$ is contained in a single fundamental domain)

$$\int_{\Gamma\backslash\mathbb{H}} \phi(z)\left|f_k(z)\right|^2 y^k \frac{dx\,dy}{y^2} \longrightarrow 0 \tag{3.10}$$

contradicting the assumption that $y^k|f_k|^2 dV \xrightarrow{w} c\,dV$.

Option (ii). We assume $\{\nu_k\}$ has a weakly convergent subsequence, which for notational convenience we continue calling $\{\nu_k\}$, which then converges to some $\nu \in \mathcal{SH}$, and moreover $\limsup \nu_k \leq \nu$, and these are equal almost everywhere. Then by (3.7), $\nu + \log y \leq 0$ almost everywhere.

Since $\nu_k \xrightarrow{w} \nu$ but $\int_{\mathbb{H}} \nu(z)\Delta\phi_0(z)dz \neq \int_{\mathbb{H}} -\log y \Delta\phi_0(z)dz$, we know that $\nu \neq -\log y$ on a set of positive measure. Since ν is upper semicontinuous, there is some $\delta > 0$ so that $\nu < -\log y - \delta$ on some nonempty open relatively compact set U.

Since $\nu = \limsup \nu_k < -\log y - \delta$, by Lemma 2.4, there is $K = K(\delta, U)$ so that for all $k \geq K$, $\nu_k \leq -\log y - \delta/2$ on U. That is to say,

$$\frac{1}{k}\log\left|f_k\right|^2 < -\log y - \frac{\delta}{2} \tag{3.11}$$

or

$$y^k\left|f_k\right|^2 \leq e^{-k\delta/2} \tag{3.12}$$

on U, which as in the first option contradicts $y^k|f_k|^2 dV \xrightarrow{w} c\,dV$.

4 Sketch of the proof of Theorem 1.1

As is well understood by now [16, 21], the equidistribution of masses (1.7) (with $c = 1$!) follows from the generalized Riemann hypothesis (for certain automorphic L-functions), in fact, from a "subconvexity estimate" for the central value of certain L-functions. To relate the equidistribution of the masses for Hecke eigenforms to GRH, one has to examine the behavior of the "periods"

$$\int_{SL_2(\mathbb{Z})\backslash\mathbb{H}} g(z)\left|f_k(z)\right|^2 y^k dV(z), \tag{4.1}$$

where g is fixed and $k \to \infty$. We need to show that whenever g is orthogonal to the constants, then the period vanishes as $k \to \infty$.

In the case where g is an Eisenstein series, the period is the (completed) Rankin-Selberg L-function $L^*(1/2 + it, f_k \times f_k)$, the standard convexity estimates give a bound of k^ϵ (for all $\epsilon > 0$) and vanishing as $k \to \infty$ follows from GRH (see [**16**, Section 4]).

To handle the case when g is a cuspidal Hecke-Maass form, one exploits Watson's formula [**21**] which relates the period with the central value of the triple product L-function

$$L\left(\frac{1}{2}, g \times f_k \times f_k\right). \tag{4.2}$$

To get the decay as $k \to \infty$, it transpires that, again, one has to beat the "convexity bound" on the central value, for which one uses GRH.

In the case of CM-forms,[3] Sarnak [**17**] proved that their masses are equidistributed. Thus the reasoning above implies (unconditionally) that the zeros of CM-forms are equidistributed with respect to dV as $k \to \infty$.

Appendix

A An L^∞-bound for cusp forms

A holomorphic form of weight k for $SL_2(\mathbb{Z})$ is bounded, since we require that the q-expansion contains no negative powers of q. We will need a bound on the supremum with an explicit dependence on the weight k. For our purposes, any bound on $y^{k/2}|f(z)|$ which is subexponential in k for fixed z suffices. Below we derive a bound of size $k^{1/2}$; since the dimension of the space of cusp forms grows linearly in k, this bound is optimal (for arbitrary forms) as far as the k-dependence is. The proof is adapted from [**3**, Lemma A.1] which treats the case of Maass forms, though unlike [**3**], we do not need to control the z-aspect.

A.1 The incomplete Gamma function

We first recall some properties of the incomplete Gamma function $\Gamma(a, x)$ defined for $x > 0$ by

$$\Gamma(a, x) := \int_x^\infty e^{-t} t^a \frac{dt}{t}. \tag{A.1}$$

[3]These only exist for certain congruence subgroups of the modular group.

From the definition, it is clear that $\Gamma(a,x)$ is decreasing in x. For integer k, we have

$$\Gamma(k-1,x) = (k-2)!e^{-x}\sum_{m=0}^{k-2}\frac{x^m}{m!}. \tag{A.2}$$

What is not completely straightforward is the asymptotic behavior of $\Gamma(a,x)$ when a, x tend to infinity. We will need the asymptotic

$$\Gamma(k-1,k) \sim \frac{1}{2}\Gamma(k-1), \quad k \longrightarrow \infty. \tag{A.3}$$

By (A.2), relation (A.3) is equivalent to

$$e^{-k}\sum_{m=0}^{k-2}\frac{k^m}{m!} \sim \frac{1}{2}. \tag{A.4}$$

This is close to a conjecture of Ramanujan [11, 12, 19, 20] that

$$e^{-k}\left(\sum_{m=0}^{k-1}\frac{k^m}{m!} + \theta\frac{k^k}{k!}\right) = \frac{1}{2} \tag{A.5}$$

for some $1/3 \le \theta = \theta(k) \le 1/2$; he showed [12] that $\theta(\infty) = 1/3$. We need the weaker asymptotic (A.4) or, equivalently (by Stirling's formula), that

$$e^{-k}\sum_{m=0}^{k}\frac{k^m}{m!} \sim \frac{1}{2}. \tag{A.6}$$

Here is a quick "proof" using the central limit theorem (this kind of argument goes back to Kac [4]): let $X_1, \dots, X_k$ be independent Poisson variables with parameter 1 (so having mean and variance 1). Let $S_k = X_1 + \cdots + X_k$ be their sum. By the central limit theorem, $(S_k - k)/\sqrt{k}$ is asymptotically normal, in particular, as $k \to \infty$,

$$\mathrm{Prob}\left(\frac{S_k - k}{\sqrt{k}} \le 0\right) \longrightarrow \frac{1}{2}, \tag{A.7}$$

that is,

$$\mathrm{Prob}\left(S_k \le k\right) \longrightarrow \frac{1}{2}. \tag{A.8}$$

However, since S_k is Poisson with parameter k, we have

$$\mathrm{Prob}\left(S_k \le k\right) = e^{-k} \sum_{m=0}^{k} \frac{k^m}{m!} \tag{A.9}$$

which gives (A.6).

A.2 The supremum of cusp forms

Proposition A.1. Let f be a cusp form of weight k for $SL_2(\mathbb{Z})$. Then uniformly for z in compact subsets of $\mathbb{H}$,

$$\frac{y^k \left|f(z)\right|^2}{\langle f, f\rangle} \ll k, \tag{A.10}$$

where $\langle f, f\rangle$ is the Petersson inner product. $\square$

Proof. Consider the integral of $y^k|f(z)|^2$ over the "Siegel set"

$$\mathfrak{S} := \left\{ z = x + iy : 0 \le x \le 1,\ y > \frac{1}{4\pi} \right\}. \tag{A.11}$$

This integral converges since f is exponentially decreasing in the cusp. Since the Siegel set $\mathfrak{S}$ is contained in a fixed number of translates of the standard fundamental domain $\mathcal{F} = \{z = x + iy : |z| \ge 1,\ |x| \le 1/2\}$, we have

$$\int_{\mathfrak{S}} \left|f(z)\right|^2 y^k \frac{dx\, dy}{y^2} \ll \langle f, f\rangle. \tag{A.12}$$

On the other hand, using the Fourier expansion $f(z) = \sum_{n\ge 1} a(n) e^{2\pi i n z}$ and Parseval, we find

$$\begin{aligned} \int_{\mathfrak{S}} \left|f(z)\right|^2 y^k \frac{dx\, dy}{y^2} &= \sum_{n\ge 1} \left|a(n)\right|^2 \int_{1/4\pi}^{\infty} e^{-4\pi n y} y^{k-2} dy \\ &= \sum_{n\ge 1} \frac{\left|a(n)\right|^2}{(4\pi n)^{k-1}} \Gamma(k-1, n). \end{aligned} \tag{A.13}$$

Thus we find that

$$\sum_{n\ge 1} \frac{\left|a(n)\right|^2}{(4\pi n)^{k-1}} \Gamma(k-1, n) \ll \langle f, f\rangle. \tag{A.14}$$

To estimate $y^k|f(z)|^2$, use the Fourier expansion and Cauchy-Schwartz:

$$\begin{aligned} y^k\left|f(z)\right|^2 &= y^k\left|\sum_{n\geq 1} a(n)e^{-2\pi n y}e(nx)\right|^2 \\ &\leq \sum_{n\geq 1}\frac{\left|a(n)\right|^2}{(4\pi n)^{k-1}}\Gamma(k-1,n)\times y^k\sum_{n\geq 1}e^{-4\pi n y}\frac{(4\pi n)^{k-1}}{\Gamma(k-1,n)}. \end{aligned} \tag{A.15}$$

For the first sum, use (A.14) to bound it by $\langle f, f\rangle$. Thus we find

$$\frac{y^k\left|f(z)\right|^2}{\langle f,f\rangle} \ll y^k\sum_{n\geq 1}e^{-4\pi n y}\frac{(4\pi n)^{k-1}}{\Gamma(k-1,n)}. \tag{A.16}$$

To bound the sum, we split it up into "small" and "large" n's. For "small" n, that is, those satisfying $n \leq k$, we can give an upper bound as follows: first use $\Gamma(k-1,n) \geq \Gamma(k-1,k)$, for $n \leq k$. This gives

$$\begin{aligned} y^k\sum_{n\leq k}e^{-4\pi n y}\frac{(4\pi n)^{k-1}}{\Gamma(k-1,n)} &\leq \frac{y^k}{\Gamma(k-1,k)}\sum_{n\leq k}e^{-4\pi n y}(4\pi n)^{k-1} \\ &\leq \frac{y}{\Gamma(k-1,k)}\sum_{n=1}^{\infty}e^{-4\pi n y}(4\pi n y)^{k-1}. \end{aligned} \tag{A.17}$$

We have[4]

$$\begin{aligned} \sum_{n=1}^{\infty}e^{-4\pi n y}(4\pi n y)^{k-1} &\leq \int_0^{\infty}e^{-4\pi x y}(4\pi x y)^{k-1}\,dx + 2e^{-(k-1)}(k-1)^{k-1} \\ &\ll \Gamma(k)\left(\frac{1}{y}+\frac{1}{\sqrt{k}}\right). \end{aligned} \tag{A.18}$$

Thus we find

$$y^k\sum_{n\leq k}e^{-4\pi n y}\frac{(4\pi n)^{k-1}}{\Gamma(k-1,n)} \ll \frac{y\Gamma(k)}{\Gamma(k-1,k)}\left(\frac{1}{y}+\frac{1}{\sqrt{k}}\right), \tag{A.19}$$

which by (A.3) is

$$\ll k\left(1+\frac{y}{\sqrt{k}}\right). \tag{A.20}$$

[4]If $f(t)$ is increasing for $t < t_0$ and decreasing for $t > t_0$, then $\sum_{n=1}^{\infty} f(n) \leq \int_0^{\infty} f(t)\,dt + 2f(t_0)$.

For the sum over the remaining n's, use (A.2) to get

$$\Gamma(k-1,n) \geq e^{-n} n^{k-2}, \tag{A.21}$$

which gives

$$\begin{aligned} y^k \sum_{n>k} e^{-4\pi n y} \frac{(4\pi n)^{k-1}}{\Gamma(k-1,n)} &\leq y^k \sum_{n>k} e^{-4\pi n y} \frac{(4\pi n)^{k-1}}{e^{-n} n^{k-2}} \\ &\leq (4\pi y)^k \sum_{n>k} e^{-n(4\pi y-1)} n \\ &\ll (4\pi y)^k \frac{(k+1) e^{-(k+1)(4\pi y-1)}}{\left(1-e^{-(4\pi y-1)}\right)^2} \\ &\ll k e^{-4\pi y} \left(\frac{4\pi y e}{e^{4\pi y}}\right)^k, \end{aligned} \tag{A.22}$$

which, for fixed $y \geq \sqrt{3}/2$, is negligible. ■

Acknowledgments

This work was partially supported by Grant no. 2002088 from the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel. This paper is based on notes written in October 1999, while enjoying the hospitality of the Institute for Advanced Study, Princeton. I have benefited from several discussions with Peter Sarnak, Misha Sodin, and Steve Zelditch.

References

[1] J. A. Getz, *A generalization of a theorem of Rankin and Swinnerton-Dyer on zeros of modular forms*, Proc. Amer. Math. Soc. **132** (2004), no. 8, 2221–2231.

[2] L. Hörmander, *The Analysis of Linear Partial Differential Operators. I. Distribution Theory and Fourier Analysis*, 2nd ed., Grundlehren der Mathematischen Wissenschaften, vol. 256, Springer, Berlin, 1990.

[3] H. Iwaniec and P. Sarnak, L^∞ *norms of eigenfunctions of arithmetic surfaces*, Ann. of Math. (2) **141** (1995), no. 2, 301–320.

[4] M. Kac, *Note on the partial sums of the exponential series*, Univ. Nac. Tucumán. Revista A. **3** (1942), 151–153, reprinted in *Mark Kac: Probability, Number Theory, and Statistical Physics* (K. Baclawski and M. D. Donsker, eds.), Mathematicians of Our Time, vol. 14, MIT Press, Massachusetts, 1979.

[5] E. Lindenstrauss, *Invariant measures and arithmetic quantum unique ergodicity*, to appear in Ann. of Math. (2).

[6] W. Z. Luo and P. Sarnak, *Quantum ergodicity of eigenfunctions on* $\mathrm{PSL}_2(\mathbf{Z})\backslash\mathbf{H}^2$, Inst. Hautes Études Sci. Publ. Math. **81** (1995), 207–237.
[7] D. Mumford, *Curves and Their Jacobians*, The University of Michigan Press, Michigan, 1975.
[8] A. Neeman, *The distribution of Weierstrass points on a compact Riemann surface*, Ann. of Math. (2) **120** (1984), no. 2, 317–328.
[9] S. Nonnenmacher and A. Voros, *Chaotic eigenfunctions in phase space*, J. Statist. Phys. **92** (1998), no. 3-4, 431–518.
[10] B. A. Olsen, *On higher order Weierstrass points*, Ann. of Math. (2) **95** (1972), 357–364.
[11] S. Ramanujan, *Question 294*, J. Indian Math. Soc. **3** (1911), 128.
[12] ———, *On question 294*, J. Indian Math. Soc. **4** (1912), 151–152.
[13] F. K. C. Rankin and H. P. F. Swinnerton-Dyer, *On the zeros of Eisenstein series*, Bull. London Math. Soc. **2** (1970), 169–170.
[14] R. A. Rankin, *The zeros of certain Poincaré series*, Compositio Math. **46** (1982), no. 3, 255–272.
[15] T. Ransford, *Potential Theory in the Complex Plane*, London Mathematical Society Student Texts, vol. 28, Cambridge University Press, Cambridge, 1995.
[16] P. Sarnak, *Arithmetic quantum chaos*, The Schur Lectures (1992) (Tel Aviv), Israel Math. Conf. Proc., vol. 8, Bar-Ilan University, Ramat Gan, 1995, pp. 183–236.
[17] ———, *Estimates for Rankin-Selberg* L*-functions and quantum unique ergodicity*, J. Funct. Anal. **184** (2001), no. 2, 419–453.
[18] B. Shiffman and S. Zelditch, *Distribution of zeros of random and quantum chaotic sections of positive line bundles*, Comm. Math. Phys. **200** (1999), no. 3, 661–683.
[19] G. Szegö, *Über einige von S. Ramanujan gestellte Aufgaben*, J. London Math. Soc. **3** (1928), 225–232 (German).
[20] G. N. Watson, *Theorems stated by Ramanujan. V: approximations connected with* e^x, Proc. London Math. Soc. (2) **29** (1929), 293–308.
[21] T. Watson, *Rankin triple products and quantum chaos*, to appear in Ann. of Math. (2).

Zeév Rudnick: Raymond and Beverly Sackler School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel
E-mail address: rudnick@math.tau.ac.il