MAY 1993

ARITHMETIC QUANTUM CHAOS*

PETER SARNAK

SECTION 1. BACKGROUND

These lectures are aimed at a general audience. A detailed and up to date account of Quantum Chaos (Q.C.) can be found in the proceedings of the recent conference "Chaos and Quantum Physics" [G-V-Z] and in Gutzwiller's article [GU]. The mathematical problems associated with Q.C. are very difficult and one must be satisfied with proving partial results and gaining insight into the truth. Our main focus will be on problems of Q.C. associated with arithmetic hyperbolic manifolds. For these we can bring to bear techniques from analytic number theory and arithmetic. This allows us to resolve some of the problems and also gain understanding of the key issues. It turns out that some of these key issues are intimately connected with well known problems in the analytic theory of L-functions. In much of theoretical physics the theory revolves around certain well understood examples (harmonic oscillator, hydrogen atom...). With time as more is established about the spectra of arithmetic hyperbolic manifolds we hope that they become central models for Q.C. The new results (see Section 3) described in these lectures have been obtained in separate collaborations with H. Iwaniec, W. Luo and Z. Rudnick.

The basic problem of Q.C. is to understand the quantization of a classical Hamiltonian system whose dynamics are "chaotic." By this we mean that the flow when restricted to a fixed energy surface (H = constant) is ergodic and has (at least) almost everywhere positive Liapunov exponents [SI1]. In particular one wants to understand the semi-classical limit $\hbar \to 0$ of the quantized system. The correspondence principle is concerned with the relation between the quantum mechanics and

---

*Expanded version of the Schur lectures Tel-Aviv 1992 and Blythe lectures Toronto 1993.
Supported in part by NSF Grant DMS-9102082.

its classical counterpart in the semi-classical limit. In what follows we will primarily restrict ourselves to Hamiltonians which are geodesic flows on compact Riemannian manifolds. Precisely let X be such a manifold with metric

$$ds^2 = g_{ij}dx^i dx^j \ . \tag{1.1}$$

The Hamiltonian $H(x,\xi)$ is defined on the cotangent bundle $T^*(X)$ by

$$H(x,\xi) = g^{ij}\xi_i\xi_j \ . \tag{1.2}$$

It gives rise to the geodesic flow on $T^*(X)$. Let Δ denote the Laplace-Beltrami operator on functions on X.

$$\Delta = \frac{1}{\sqrt{g}}\frac{\partial}{\partial x^i}\left(\sqrt{g}g^{ij}\frac{\partial}{\partial x^j}\right) \tag{1.3}$$

where $g = \det(g_{ij})$. By the standard quantization procedure, see for example [L-L], $-\hbar^2\Delta$ corresponds to a quantization of H. The stationary eigenstate (or eigenfunctions, modes or wave functions) equation reads

$$-\hbar^2\Delta\psi_k = \lambda_k\psi_k \ . \tag{1.4}$$

It is well known that the eigenvalues $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \ldots$ are discrete and that the ψ_k's can be chosen to be an orthonormal basis (o.n.b.). From (1.4) we see that for the geodesic flow the semi-classical limit $\hbar \to 0$ is in this case the same as the large eigenvalue limit $\lambda_k \to \infty$. Hence the basic problem in this case is to understand the behavior of λ_k and ψ_k as $k \to \infty$ and especially their relation to the geodesic flow (correspondence principle). It is this form of the problem that we will examine.

If the classical flow is completely integrable (see [A-A] for definitions) then the $\hbar \to 0$ limit for eigenvalues and eigenfunctions is quite well understood, see Colin-de-Verdiere [CD1]. In this case the quantization condition was given by Bohr-Sommerfeld and Einstein. If p, q are generalized coordinates, $H(p,q)$ the Hamiltonian and T an invariant torus for the full set of integrals of the motion then the quantization condition is (see [CD1])

$$\int_{\gamma_\ell} p.dq = 2\pi\hbar(n_\ell + \mu_\ell/4) \tag{1.5}$$

where $n_\ell \in \mathbb{Z}$ and $\mu_\ell \in \mathbb{Z}$ is the Keller-Maslov index. The γ_ℓ's are a $\mathbb{Z}$ basis for the homology cycles in $H_1(T, \mathbb{Z})$. The corresponding approximate eigenfunctions or quasimodes, which are constructed in deducing (1.5), localize onto T as $\hbar \to 0$. Moreover their eigenvalues are

$$E_n = H(\hbar(n + \mu/4)) \tag{1.6}$$

where $I = \hbar(n + \mu/4)$ are the action variables for the quantized torus.

We illustrate the above with the well known whispering gallery effect. Let X be an ellipse in the plane. This is a manifold with boundary, a case which we also allow especially in connection with the numerical investigations in Section 2. The classical mechanics corresponds to a billiard ball in the ellipse which undergoes linear motion in the interior and angle of incidence equals angle of reflection at the boundary. The quantum system is the eigenvalue problem for the Laplacian with (say) Dirichlet boundary conditions. The billiard in this example is integrable. Indeed if C is a confocal ellipse (or hyperbola) as shown in Figure 1.1 then a billiard tangent to C will remain so forever.

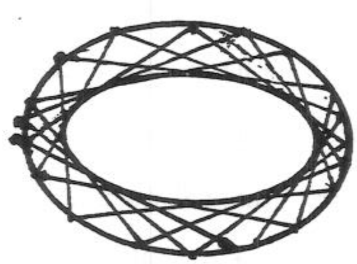

Figure 1.1

In phase space these vectors tangent to C give us an invariant torus T. For a quantized such torus the quasimodes localize onto T and hence their projection onto X are localized as is apparent in Figure 1.2, which is an eigenstate for the circle taken from McDonald and Kaufman [M-K].

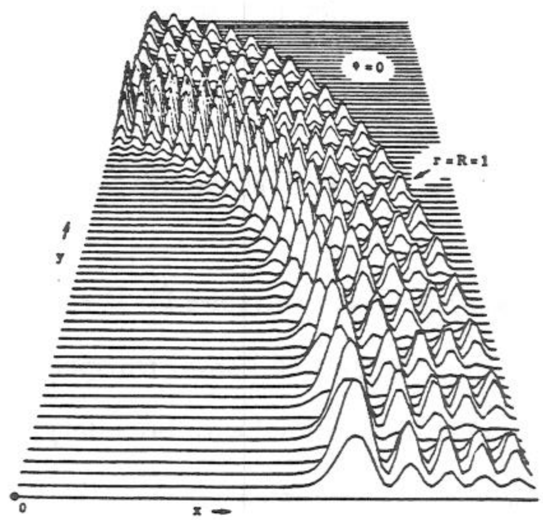

FIG 1.2: [redacted]. Intensity distribution $|\psi_{m,n}|^2(x,y)$ in positive quadrant $(x,y>0)$ of the $(\gamma=0)$ circle. This mode is $\psi_{m,n}=\psi_{40,5}=J_{40}(k_{40,5}r)\sin 40\theta$ with eigenvalue $l_{40,5}=65.012$.

The existence of such localized modes near the boundary of a smooth convex region explains the whispering gallery effect: that sonorous vibrations have a tendency to cling to the boundary of a convex surface. For details see Rayleigh [RAY] and Keller-Rubinow [K-R].

It was recognized early on (see Einstein [EI]) that if H is not integrable then such quantization conditions (as $\hbar \to 0$) are not available and perhaps do not exist. This in fact was the key stumbling block of the old quantum theory [PA, Chapter 10]. This applies especially if the classical mechanics is chaotic. For the case at hand viz the geodesic flow on $T^*(X)$ it is well known [SI1, A-A] that if X has negative sectional curvature then the classical flow is ergodic, Anosov… and in particular chaotic. Almost every orbit is dense but on the other hand there are a countably infinite number of unstable periodic orbits as well as orbits of arbitrary complexity.

For comparison, we recall the strong influence that a stable periodic orbit has on the spectrum. Suppose that $\dim X = 2$ and that γ is a stable closed geodesic of length L. That is the linearized Poincaré map about γ [A-A] has eigenvalues $e^{i\theta}, e^{-i\theta}$ with $0 < \theta < \pi$. The following theorem is due to Babich–Lazutkin [B-L] and Ralston [R].

Theorem 1.1. *Let γ be as above then there are eigenvalues γ_m of the Laplacian satisfying*

$$\sqrt{\lambda_m} = \frac{2\pi m + \theta/2 + \pi p_0}{L} + O(m^{-1/2})$$

where $p_0 = 0, 1$ depending on the behavior of the Jacobi fields along γ.

Thus γ accounts for an arithmetic progression of $\sqrt{\lambda_j}$'s. To see what part of the spectrum this comprises recall Weyl's Law with remainder due to Hörmander [HO]

$$\#\{j \mid \sqrt{\lambda_j} \leq T\} = C_n \ \mathrm{Vol}(X) T^n + O(T^{n-1}) \quad \text{as } T \to \infty \,. \tag{1.7}$$

where $c_n^{-1} = 2^n \pi^{n/2} \Gamma(\frac{n}{2} + 1)$ and $n = \dim X$. Hence we see that for $n = 2$ such a stable closed geodesic accounts for about T of the T^2 eigenvalues. As with the integrable case the method of proof of Theorem 1.1 gives at the same time quasimodes $u_m(x)$ which localize onto γ as $m \to \infty$. We normalize $u_m(x)$ to have L^2-norm equal to one, so that

$$d\mu_m = |u_m(x)|^2 dv(x), dv(x) = \sqrt{g} dx \tag{1.8}$$

are probability measures. Then $\mu_m \to \nu$ the singular arc-length measure as $m \to \infty$. This is an instance of the correspondence principle. On the other hand if γ is an unstable periodic orbit as is the case for all closed geodesics when the curvature is negative, then its influence on the spectrum (even as $\hbar \to 0$) is unclear. This is one of the central problems of Q.C.

In general we may define probability densities as in (1.8) with ψ_k replacing u_m. We denote by μ_k the corresponding probability measure on X. The quantum mechanical interpretation that μ_k is the probability density of a particle in the state ψ_k, is well known [L-L]. The question of whether subsequences of μ_k can localize and what are the possible weak star limits of these measures has been raised on many occasions. In contrast to the integrable case we have the following beautiful result of Shnirelman [SHN], Zeldtich [Z] and Colin-de Verdeire [CD2].

Theorem 1.2 (Quantum Ergodicity). *Assume that the geodesic flow is ergodic then for a subsequence λ_{k_j} of λ_k of full density*

$$d\mu_{k_j} \to dV/VOL(X) \ .$$

The meaning of full density is that

$$\#\{k_j \mid \lambda_{k_j} \leq T\} \sim \#\{k \mid \lambda_k \leq T\}$$

as $T \to \infty$. Thus the Theorem asserts that almost all of the square moduli of the eigenfunctions become equidistributed (hence the name Quantum Ergodicity). In fact their Theorem is stronger. It is concerned with measures on $T_1^*(X)$ the unit cotangent bundle to X. They show how to associate with ψ_k a measure $\tilde{\mu}_k$ on $T_1^*(X)$ such that for all $f(x,\xi) = f(x)$ i.e. ξ independent functions

$$\int_{T_1^*(X)} f(x,\xi)d\tilde{\mu}_k(x,\xi) = \int_X f(x)d\mu_k(x) \ . \tag{1.9}$$

Their theorem asserts that $\tilde{\mu}_k$ converge to μ the Liouville measure for almost all λ_k's. Along the way it is shown using Egorov's Theorem [E] that any limit of the $\tilde{\mu}_k$ must be invariant under the geodesic flow. Besides the Liouville measure, the arc length measure on a closed geodesic or volume form on an invariant torus are the simplest examples and as we have seen they can occur as such limits. We will call any limit point of the μ_k's a quantum limit. Whether in the case of a negatively manifold a closed geodesic can be the support of a quantum limit is not answered by Theorem 1.2. This question has been raised by Colin-de-Verdiere on a number of occasions [CD2-CD3]. It is closely related to scarring which will be defined below.

The central question then in the chaotic case is whether individiual eigenfunctions localize or scar or show any relation to the classical dynamics or do they as has been suggested by Berry [B1] behave like random waves. See Heller [HEL2] and Berry [B1] for discussions on the models for random waves. As a test of the random wave model note that the analogue of the central limit theorem [S-Z] would require that the distribution function of the ψ_k's tend to a Gaussian. Also the analogue of the Law of the Iterated Logarithm [S-Z] would require that

$$\|\psi_k\|_\infty \approx \sqrt{\log \lambda_k}\ . \tag{1.10}$$

We will look again at questions concerning the behavior of eigenfunctions after presenting the numerical experiments. We turn now to the fine structure of the distribution of the eigenvalues. Assume $n = 2$ and Weyl's Law is normalized to read

$$N(T) := \#\{j \mid \lambda_j \leq T\} \sim T\ . \tag{1.11}$$

So on average the eigenvalues occur with unit spacing. In Figure 1.3 some examples of 50 such spectral lines (high in the spectrum) are given. The Figure except for column (b) is taken from Bohigus-Giannoni [B-G]

All data have been scaled to have mean level spacing one. The Poisson Column corresponds to random numbers i.e. a Poisson process. Notice they have clusters and near degeneracies and gaps. The $SL_2(\mathbb{Z})$ spectrum comes from an arithmetic chaotic system and is described in detail in Sections 2 and 3. Note it has similar features to the Poisson. The next column are the experimental levels of the heavy nuclei [B-G]. These are a lot more regular or rigid in the sense that the number in a small interval is close to the expected number. The same is true for the chaotic

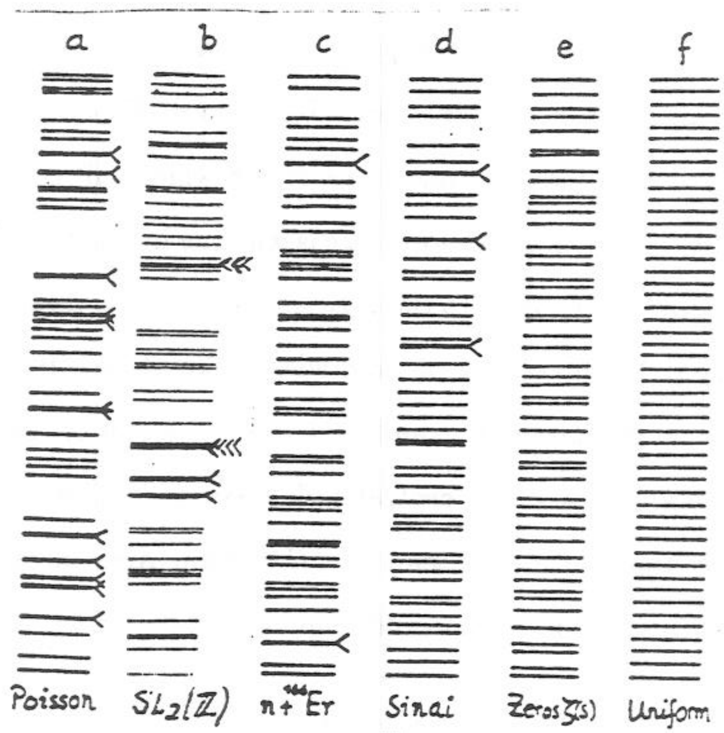

Figure 1.3

Sinai billiard, see Figure 2.5 in Section 2 for the description. The imaginary parts of the zeros of the Riemann zeta function (Section 4) are very rigid as is clear from the second to last column.

There are many statistics that we may use in order to study the finer aspects of the distribution of the levels. The statistic we examine is the Number variance $\Sigma^2(\lambda, L)$ (see Dyson-Mehta [D-M] where this and the related spectral rigidity are introduced in connection with the study of the spectral lines of heavy nuclei). The number variance measures the average deviation from the expected number, of the number of energy levels in intervals of length L. For the local averaging we use

$$\Sigma^2(\lambda, L) := \frac{1}{\lambda} \int_{\lambda}^{2\lambda} (N(x+L) - N(x) - L)^2 dx \ . \tag{1.12}$$

In order to explain the expected behavior of Σ^2 we recall some standard models for the λ's.

(i) Poisson: The λ_j follow a Poisson process that is $\mu_j = \lambda_{j+1} - \lambda_j$ has a Poisson distribution and the μ_j's are independent of each other. The left hand column in Figure 3.1 is a sample of such a sequence.

(ii) G.O.E. (Gaussian Orthogonal Ensemble) The λ_j's are the eigenvalues of a random symmetric $N \times N$ matrix ($N \to \infty$) which have been rescaled to satisfy (1.11). That is on the space of real symmetric matrices whose elements we denote by $B = (B_{ij})$ we put the Gaussian Orthogonal measure

$$d\mu(B) = C_N \exp\{-\frac{1}{2}\sum_j B_{jj} - \sum_{i<k} B_{jk}^2\} \prod_{j\leq k} dB_{jk} \tag{1.13}$$

μ is invariant under the action of the orthogonal group $B \to PBP^t$. From (1.13) one gets the joint distribution of the eigenvalues. From the latter, though it is by no means straightforward, one can compute the expectation of the number variance as well as other statistics such as the level spacing distribution. See Mehta's book [ME] for details.

(iii) G.U.E. (Gaussian Unitary Ensemble) Same as in (ii) except that real symmetric matrices are replaced by complex Hermitian ones.

For each of the above models one computes the expectation for $\Sigma^2(L)$:

(i) $$\Sigma^2_{\text{POISSON}}(L) = L$$

(ii) For L large

$$\Sigma^2_{GOE}(L) = \frac{2}{\pi^2}(\log 2\pi L + \gamma + 1 - \pi^2/8) + O(L^{-1}) .$$

(iii) For L large

$$\Sigma^2_{\text{GUE}}(L) = \frac{1}{\pi^2}(\log 2\pi L + \gamma + 1) + O(L^{-1}) .$$

Here γ is Euler's constant. Note that the number variance is much smaller for (ii) and (iii) than it is for (i). We express this by saying that the spectrum in (ii) and (iii) is rigid.

After Wigner, Dyson, Mehta, Bohigas, Berry, Tabor and others it is now believed that for integrable systems the eigenvalues follow the Poisson behavior while for chaotic systems they follow the GOE distribution. Put another way the spectrum of an integrable system is random while that of a chaotic system is rigid! Berry [B2] has given convincing heuristic arguments using trace formulae for the following universal behaviors of $\Sigma^2(\lambda, L)$.

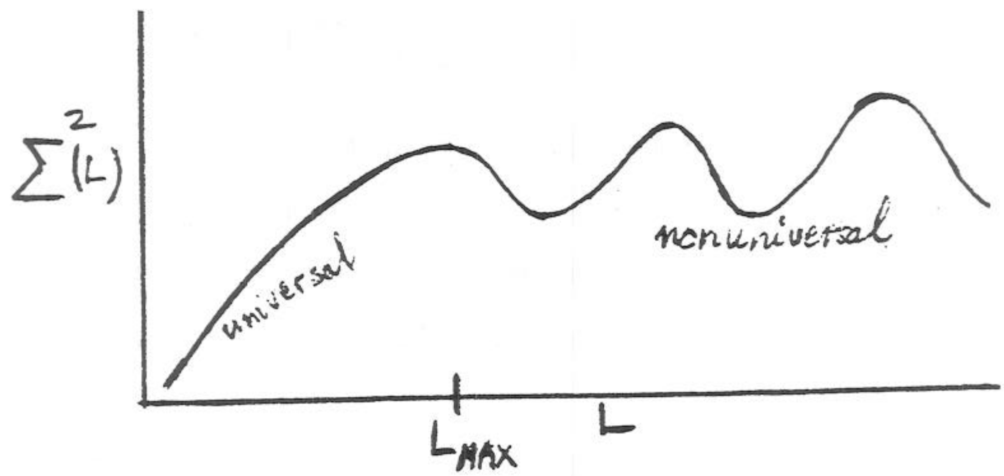

Figure 1.4

For $1 << L << L_{\mathrm{MAX}} \simeq \sqrt{\lambda}$ (recall we are assuming $n = 2$ for simplicity) the universal behavior is

(i) $\Sigma^2(L) = \Sigma^2_{\mathrm{POISSON}}(L)$ if the geodesic flow is integrable.

(ii) $\Sigma^2(L) = \Sigma^2_{\mathrm{GOE}}(L)$ if the geodesic flow is chaotic.

For $L > L_{\mathrm{MAX}}$ nonuniversal oscillations take over as indicated in Figure 1.4. There are similar conjectures for the closely related spectral rigidity [B2].

This completes our discussion of general background, we turn now to numerical experiments on the behaviors of eigenfunctions and eigenvalues of chaotic systems. In Section 3 we describe results concerning the eigenvalues and eigenfunctions of arithmetic hyperbolic manifolds. In Section 4 we review the theory of L-functions and relate them to the problems of Q.C. In Section 5 we outline some proofs.

Section 2. Report on Numerical Experiements

Our report on numerical experiments concerning the spectra of chaotic systems have been done by various groups: Heller, Bohigas Schmit, Steiner Aurich Steil, Hejhal and Rackner. They have used different methods to compute the eigenvalues and eigenfunctions. Indeed ingeneous methods are needed in order to compute such large eigenvalues for these problems. The reader should consult the original articles cited below for details.

2.1. The Stadium.

A chaotic system in the plane is the billiard in a Bunimovich Stadium which is shown with a perodic orbit in Figure 2.1.

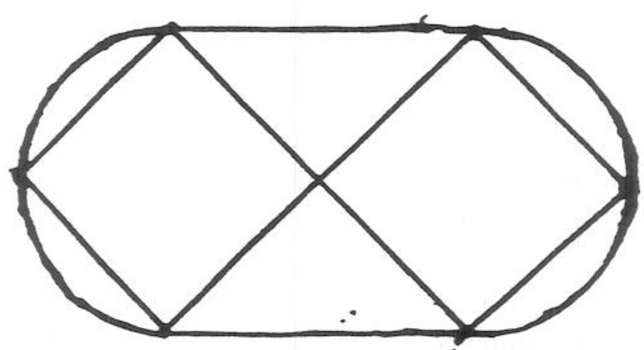

Figure 2.1

Bunimovich [BUN] has shown that this billiard dynamics is both ergodic and chaotic. It has a certain mixed aspect to it in that there are non isolated families of periodic orbits – the bouncing balls in Figure 2.2. While these are of measure zero they do have a strong influence on the spectrum (see Proposition 3.1). In [HEL1, HEL2] (see also [M-K]) Heller has numerically investigated the eigenfunctions in detail. He finds numerous behaviors as shown in Figure 2.3. Gerard and Leichtnam [G-L] have extended Theorem 1.2 to this case of a nonsmooth boundary. So almost all of the μ_k's must be equidistributed. However there are modes which are clearly associated with the bouncing balls in Figure 2.2. The most intriguing discovery of Heller is the scarring on unstable periodic orbits see Figure 2.4. His definition of a state scarring on a closed orbit is that $|\phi_k(x)|^2$ deviate significantly from its expected value, along the periodic orbit. Usually he has in mind that there is an enhanced probability around the orbit. Such scarring however is not expected to survive as $\hbar \to 0$.

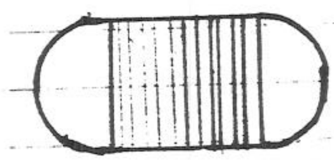

Figure 2.2

We are more interested in effects which persist as $\lambda_k \to 0$ and so we define the following stronger form of scarring:

Definiton 2.1. *A subsequence μ_{k_j} is said to strongly scar on $\Lambda \subset X$ if μ_{k_j} has limit μ and sing supp(μ) $= \Lambda$.*

By singular support we mean in the measure theoretical sense, i.e. write $\mu = \mu_1 + \mu_2$ with $d\mu_1 << dV$ and $d\mu_2 \perp dV$. Then sing supp(μ) is by definition the support of μ_2. The singular support is created by $|\phi_k(x)|^2$ behaving singularly on Λ. We will return in Section 3 to the question of strong scarring for certain chaotic systems.

A domain with similar properties to those of the stadium is the Sinai domain in Figure 2.5. Its levels as computed by Bohigas-Giannoni [B-G] were illustrated in Figure 1.3.

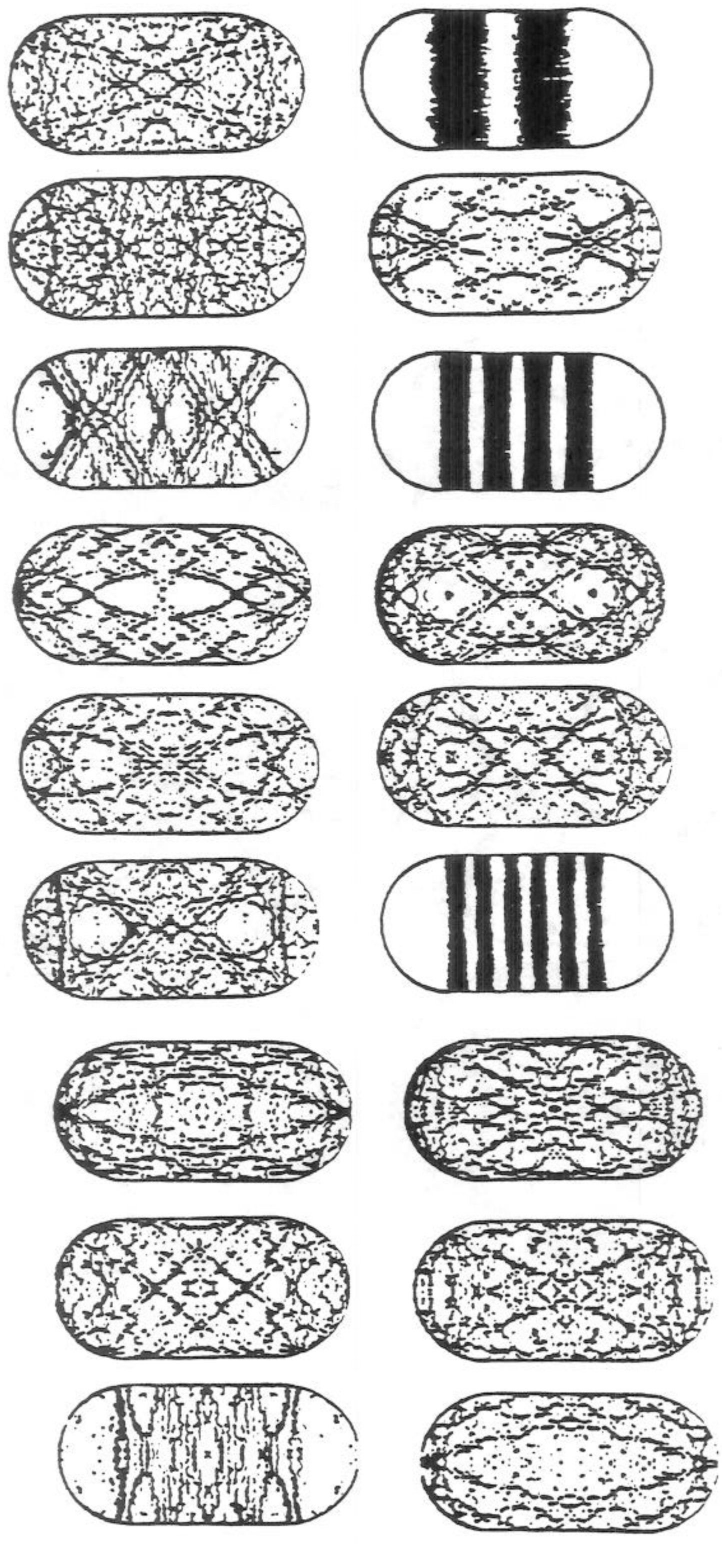

Figure 2.3. Density plot of $|\Phi(x)|^2$ for eigenstates of the stadium (Black signifies high density) for eigenvalues $\sqrt{\lambda} = k$, where going from top to bottom, $k = 110.389$, 119.413, 119.417, 119.451, 119.499, 119.512, 119.512, 119.525, 119.547, 119.587, 119.637, 119.672, 119.691, 119.701, 119.740, 119.802, 119.809, 119.839.

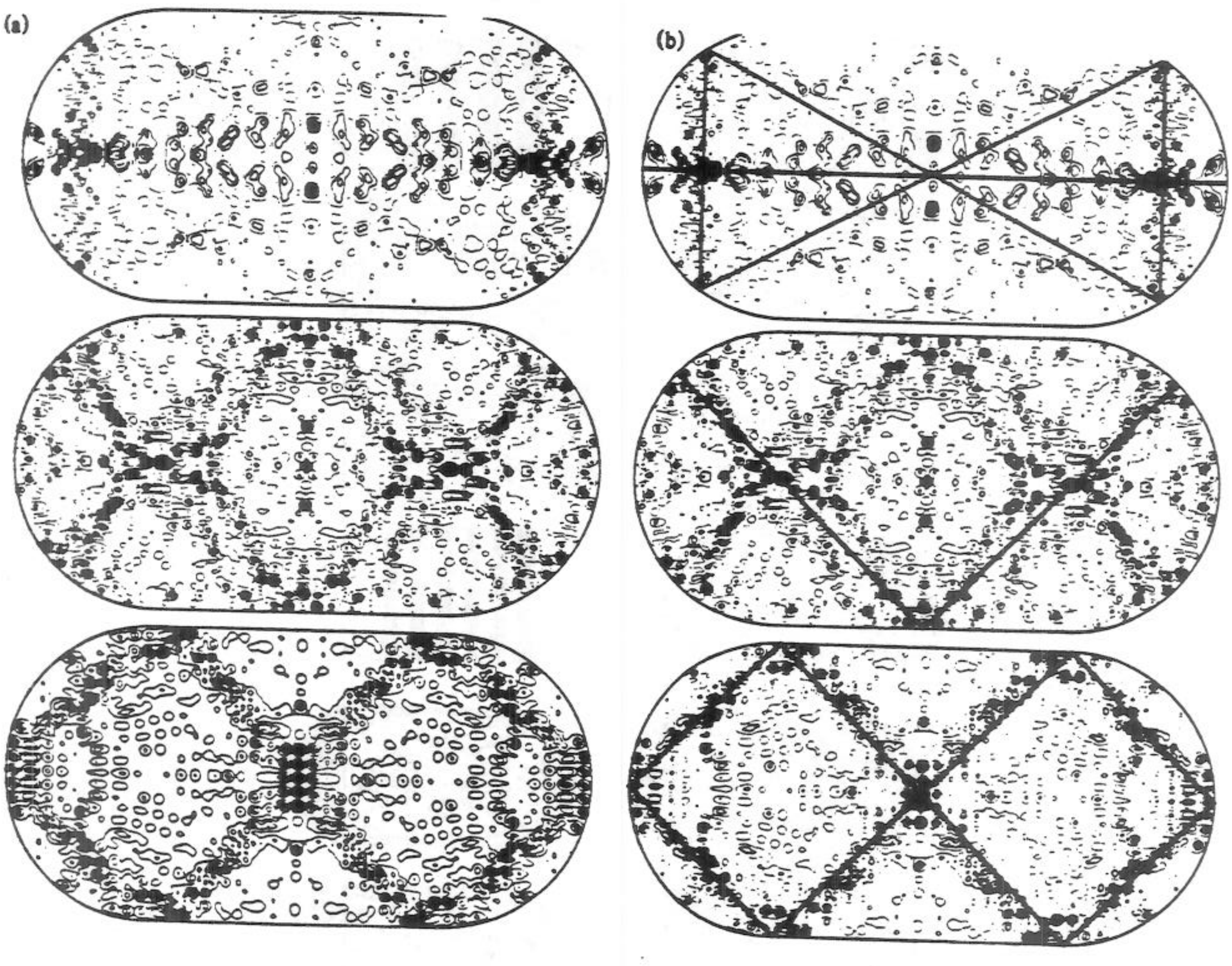

Figure 2.4. Left column, three scarred states of the stadium; right column, the isolated, unstable periodic orbits corresponding to the scars.

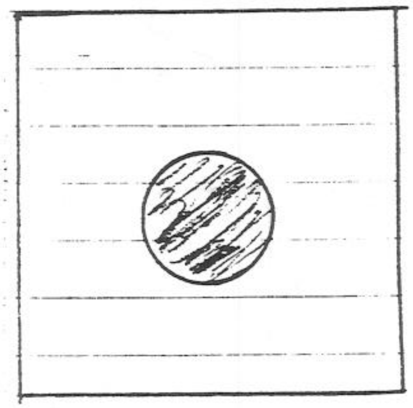

Figure 2.5

Section 2.2. Hyperbolic Surfaces.

As mentioned in Section 1 surfaces of negative Gaussian curvature K give rise to chaotic dynamical systems. Among these the ones most extensively studied are surfaces with constant curvature. These may be realized as $\Gamma\backslash\mathbb{H}^2$, where $\Gamma \leq \mathrm{SL}_2(\mathbb{R})$ is a discrete subgroup and $\mathbb{H}^2$ is the non-Euclidean upper half plane with line element $ds = |dz|/y$. We can also use the disk model for $\mathbb{H}^2$ viz $U = \{z \mid |z| < 1\}$ with $ds = 2(1-|z|^2)^{-1}|dz|$. In this case the discrete group is a subgroup of $\mathrm{SU}(1,1)$. In Figure 2.6 a fundamental domain, for the discrete group Γ generated by γ_1, γ_2, γ_3, γ_4, may be taken to be the intersection of the exteriors of the semi circles (see Katok [K] for details). The quotient is a so called hyperbolic orbifold (because Γ acts with fixed points) of genus 1. Another example is the family of Hecke groups whose fundamental domains are given in Figure 2.7. These are hyperbolic triangles in $\mathbb{H}^2$ with angles $\{\pi/m, \pi/m, 0\}$, $m \geq 3$. The group generated by the reflections R_1, R_2, R_3 as shown, is discrete. It has the Hecke group Γ_m generated by $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2\cos\pi/m \\ 0 & 1 \end{bmatrix}$, as an index two subgroup. The case $m = 3$ is the classical modular group $\mathrm{SL}_2(\mathbb{Z})$ and is of special interest. Note that $\Gamma_m\backslash\mathbb{H}^2$ is not compact.

Aurich and Steiner [A-S] have computed the eigenstates for the group whose fundamental domain in U is the octagon in Figure 2.8. With the sides identified as indicated $X = \Gamma\backslash\mathbb{H}^2$ is a genus 2 hyperbolic surface. Figures 2.9, 2.10, and 2.11 give plots of the density $|\phi_j(z)|^2$ for states with the energy indicated. They found

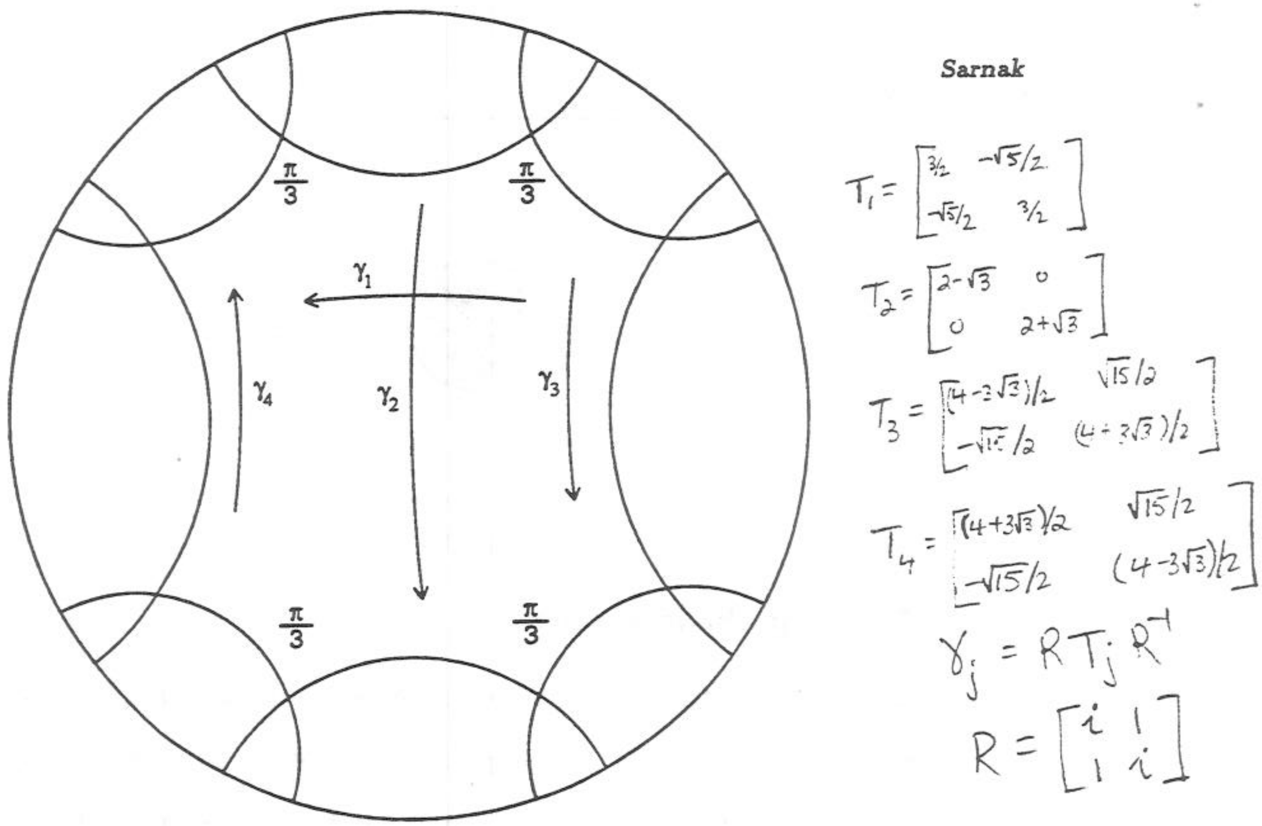

Figure 2.6

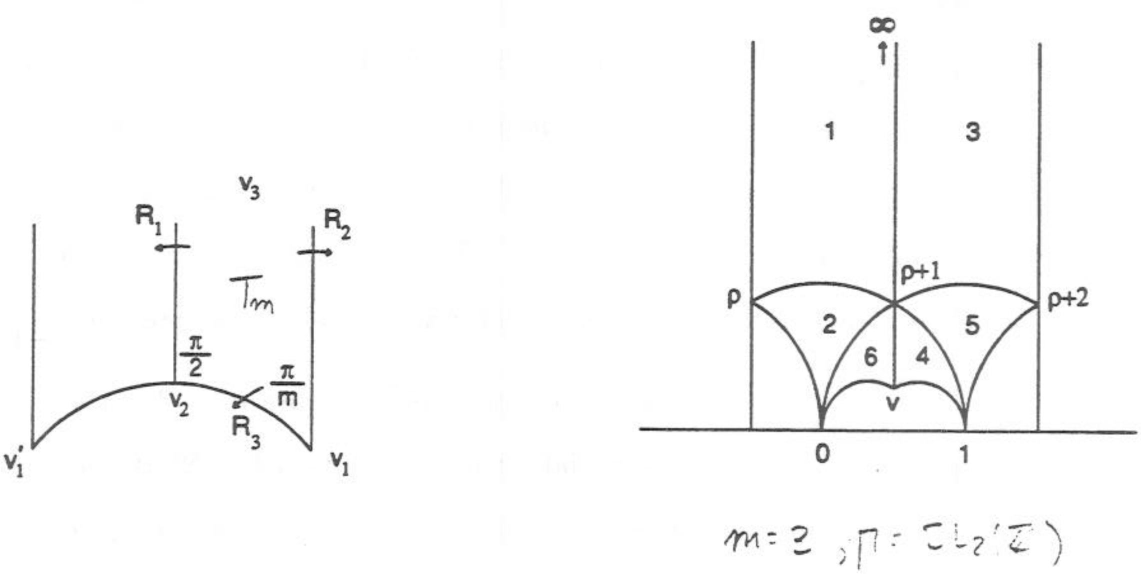

Figure 2.7

similar behaviors for all the eigenstates – that is there is no scarring or localization. In Figure 2.12 the cummulative distribution function for these eigenstates are displayed. Clearly these are Gaussian as predicted by the random wave model.

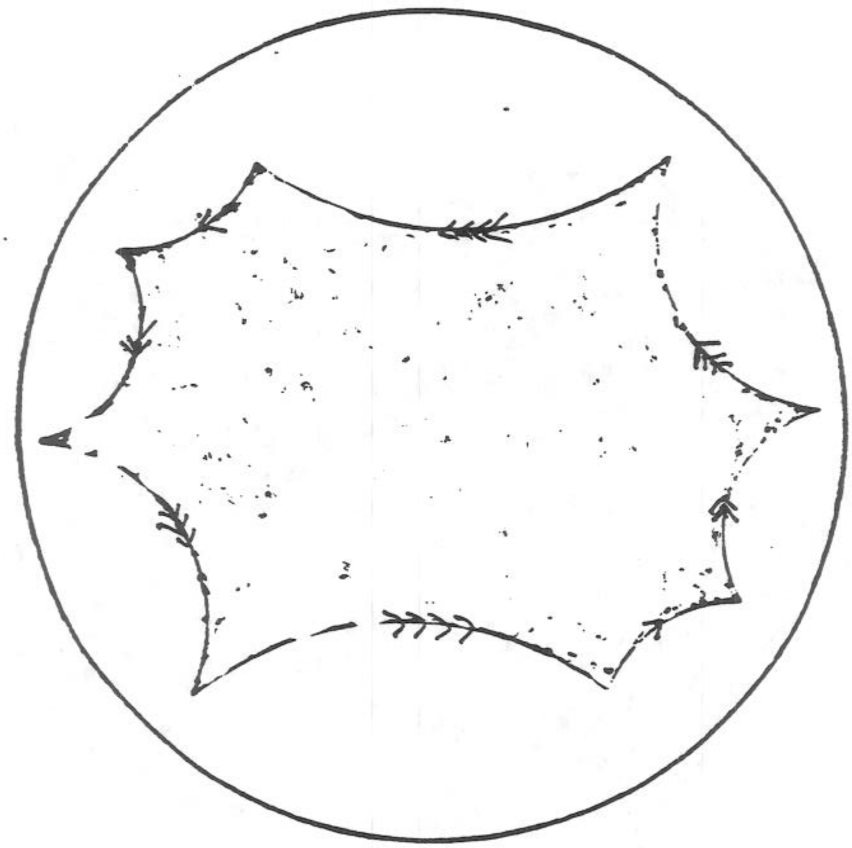

Figure 2.8

The final set of results were done by Hejhal's group (se [H-R]). They concern the eigenfunctions for the modular group $\Gamma = SL_2(\mathbb{Z})$. Its fundamental domain was shown in Figure 2.7 with $m = 3$. Since $\Gamma\backslash\mathbb{H}^2$ is not compact in this case, the existence of discrete spectrum (that is L^2 eigenfunctions) is not obvious. There is a continuous spectrum given by Eisenstein series (see Section 4). For this Γ, Selberg [SEL-2] showed that there are also an infinite number of embedded eigenvalues; in fact that in suitable sense most the spectrum consists of bound states (or cusp forms as they are called in this setting). Figures 2.13 shows the nodal lines for the cusp form with $r = \sqrt{\lambda - 1/4} = 125.13840$. Figures 2.14 and 2.15 give density plots

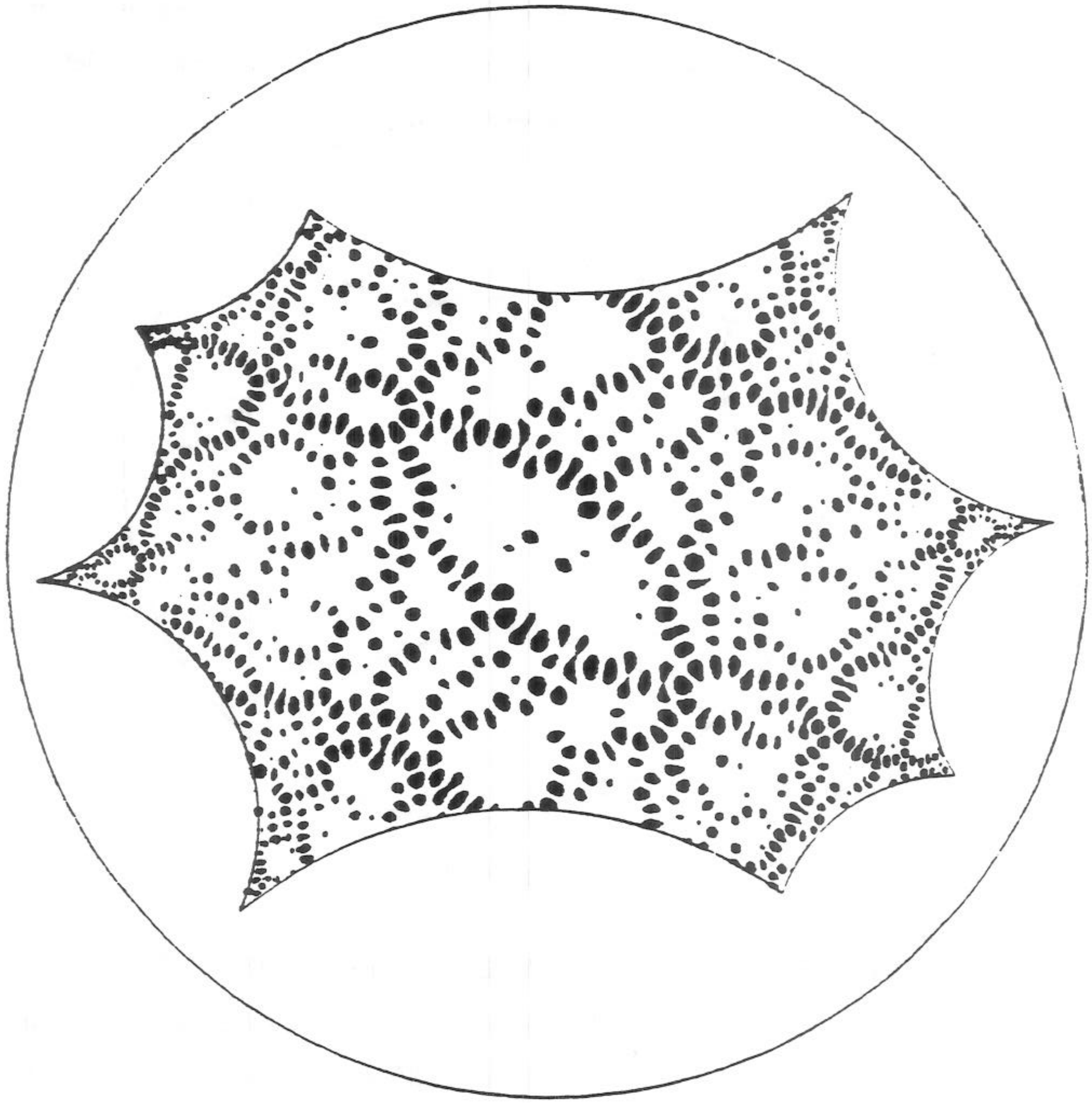

Figure 2.9. The intensity $\psi_n(z)^2$ of the eigenstate at energy $\lambda = 2000.695$ is shown in the Poincaré disc, whose boundary $z = 1$ is presented by the circle. The intensity is plotted in black, if it is above the threshold value $-ln(c\sqrt{2})/2\pi$ with $c = 0.6$.

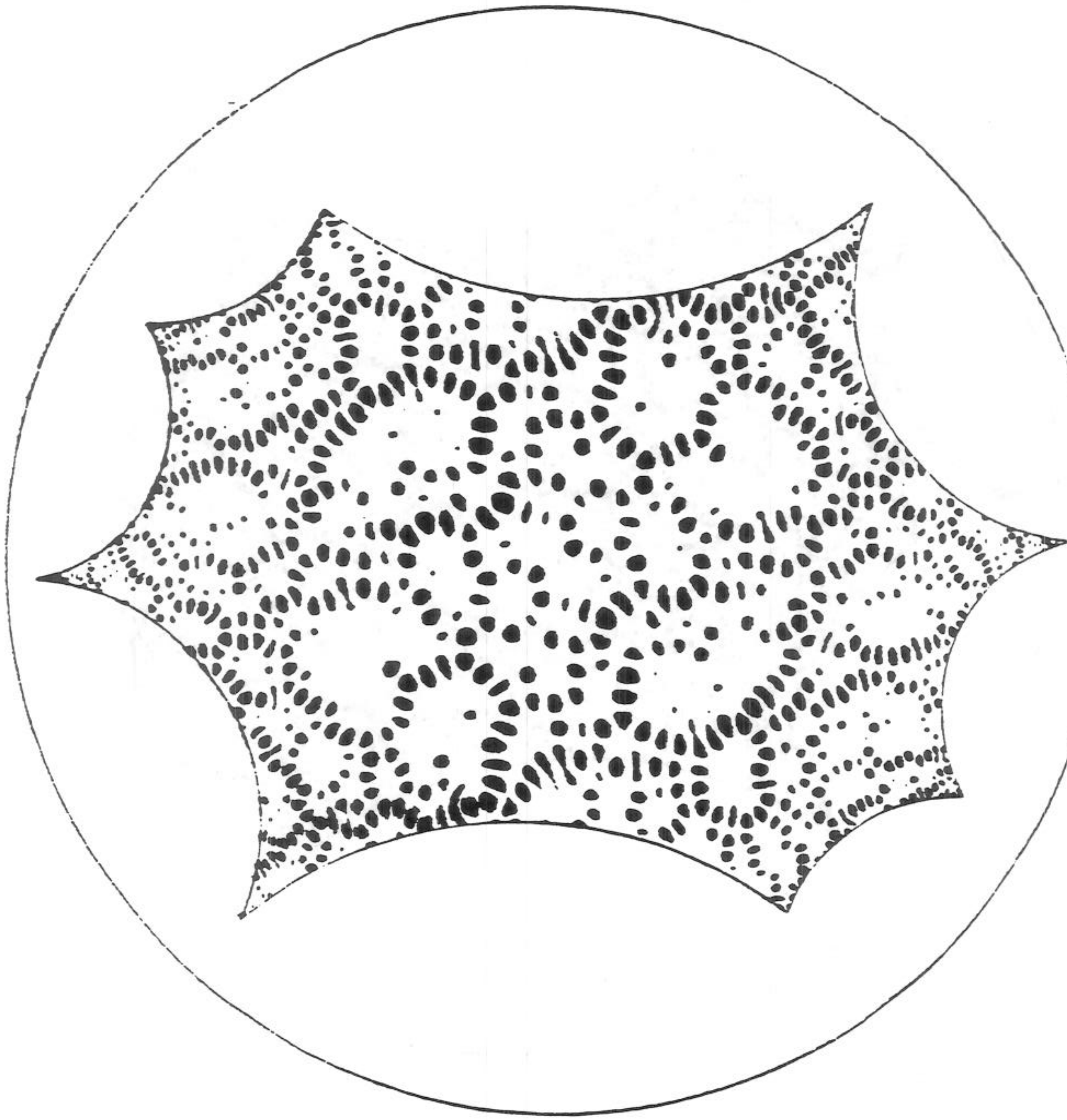

Figure 2.10. The same as in Figure 2.9 for the eigenstate at energy $\lambda = 2003.117$.

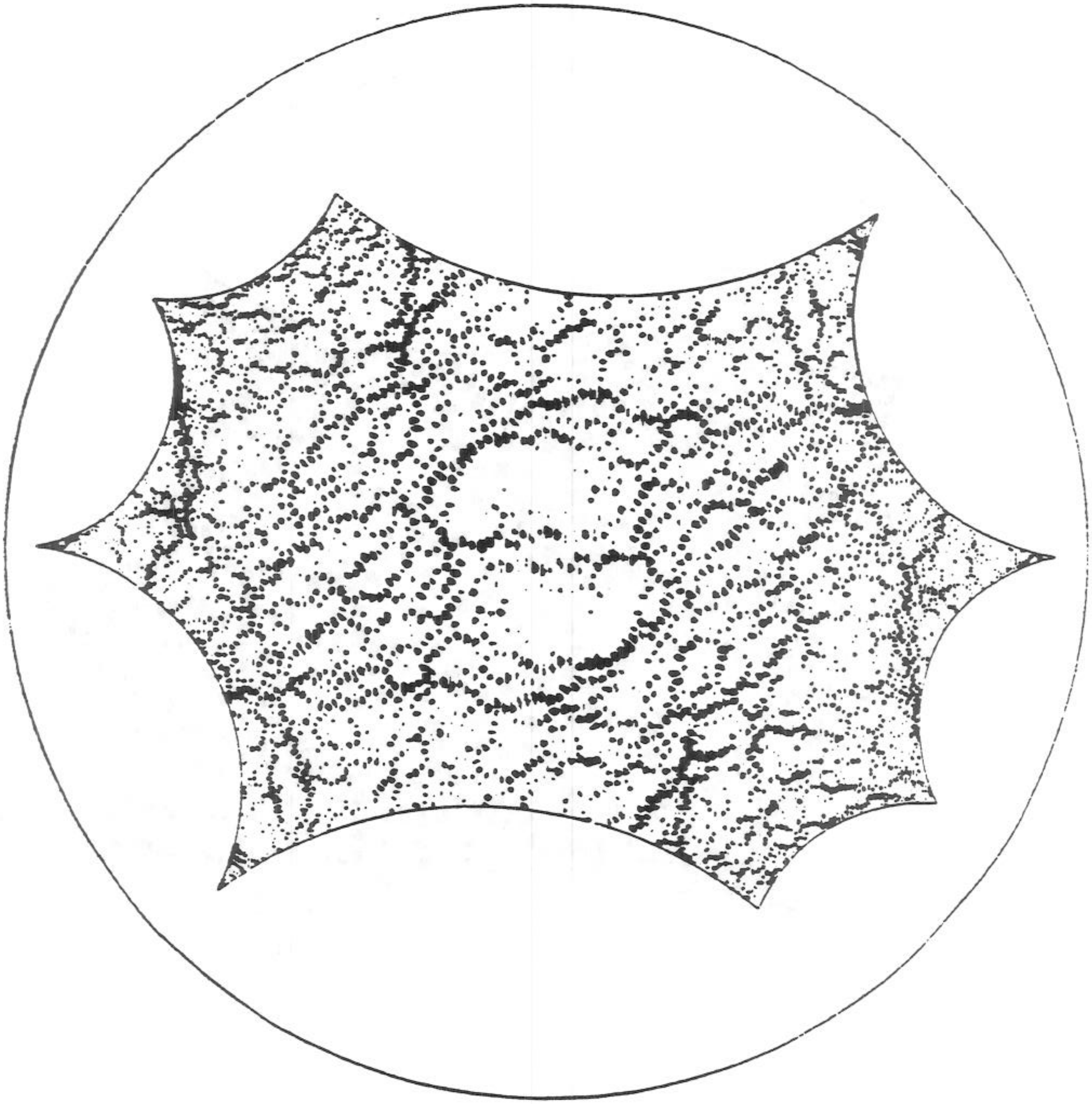

Figure 2.11. The same as in Figure 2.9 for the eigenstates at energy $\lambda = 10001.092$.

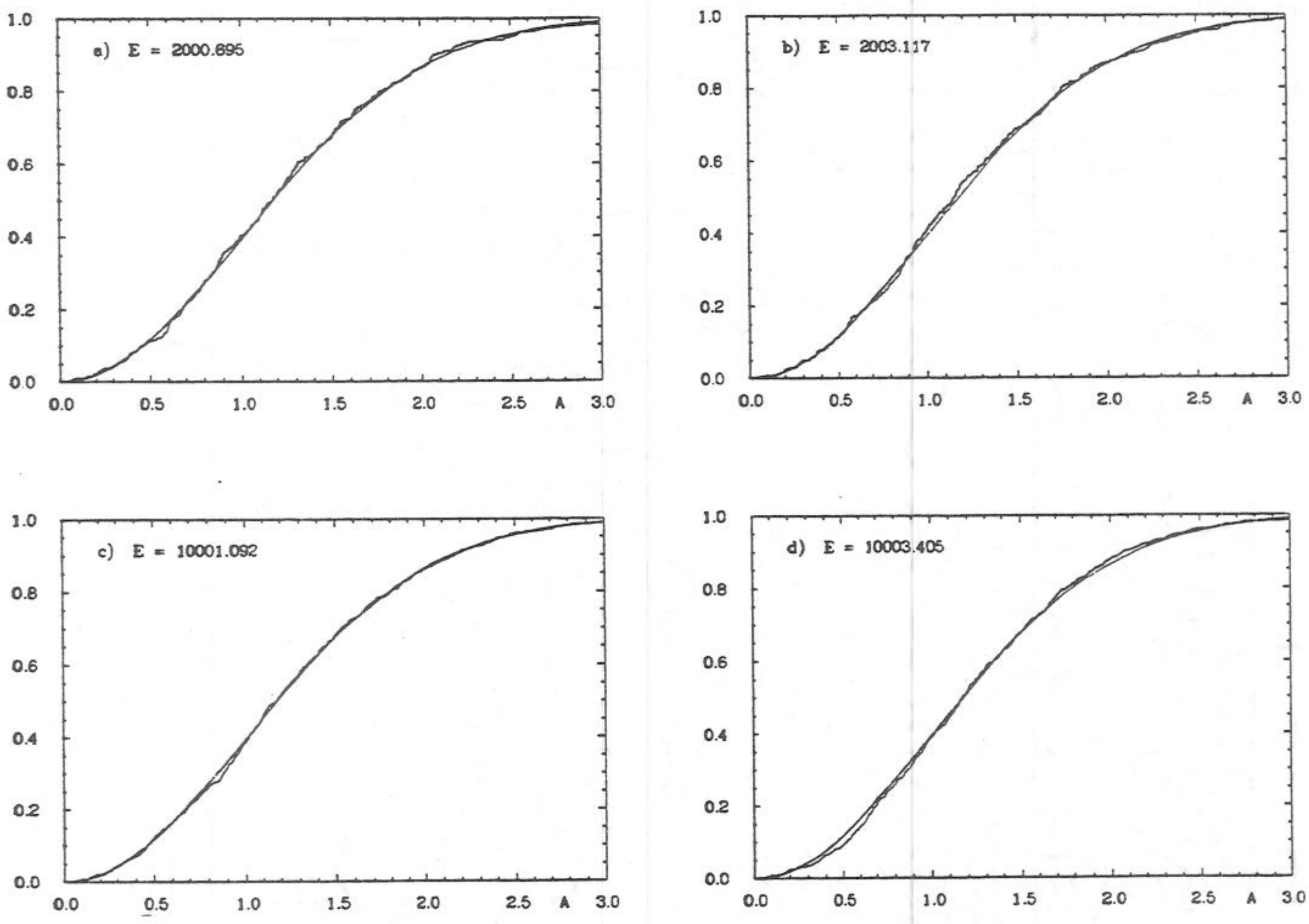

Figure 2.12. The cumulative distribution of the amplitudes A_n is shown in comparison with the theoretical expectation $I(A) = 1 - \exp(-A^2/2)$ (full line).

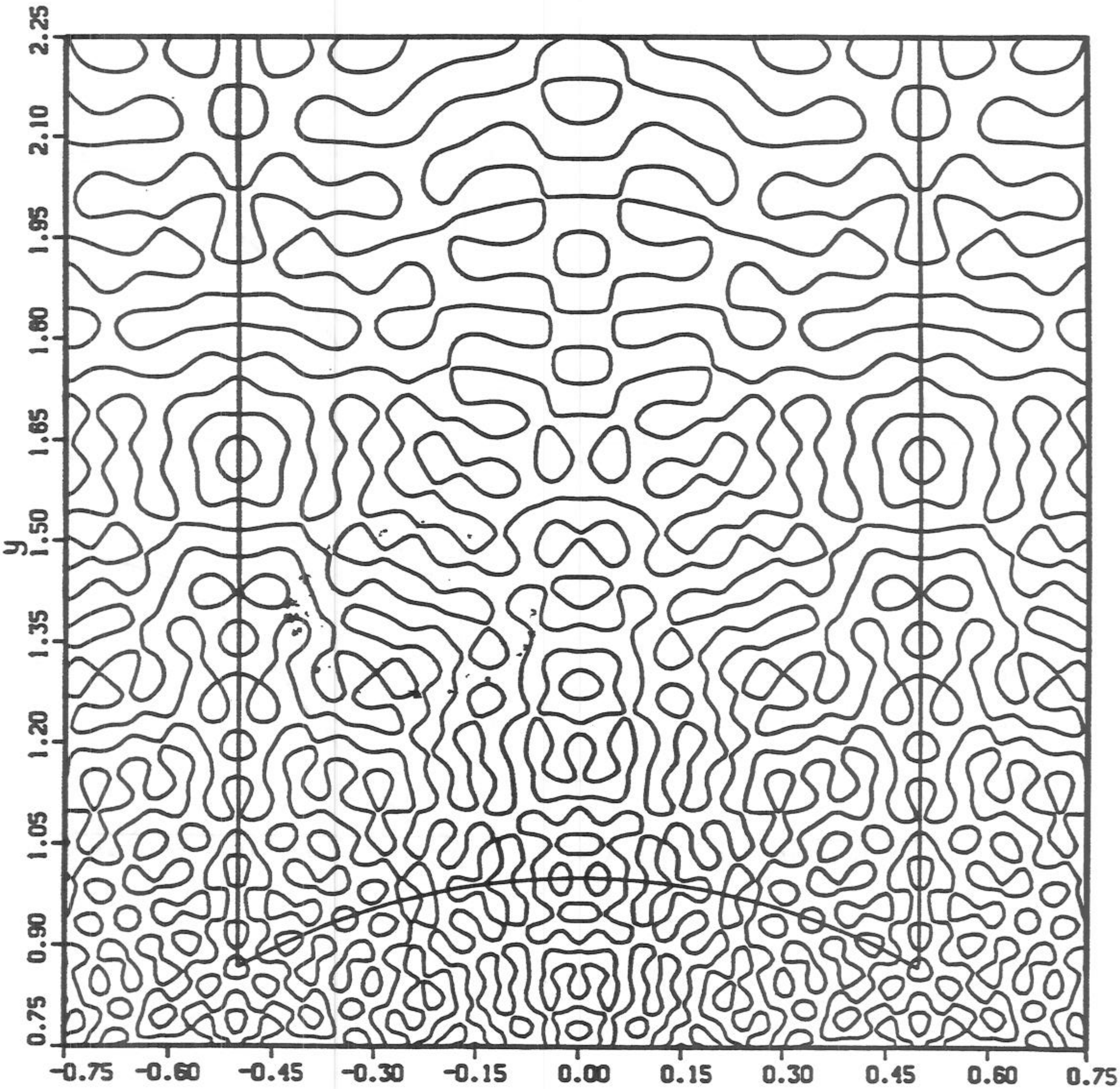

Figure 2.13. Nodal lines for the cusp form with $R = 125.13840$ for $SL_2(\mathbb{Z})$.

with $r = 125.313840$ and $r = 500.283548$. The results are similar to the previous compact quotient examples. There is no localization and the random wave model seems to apply. The distribution functions of ϕ_λ for various λ are given in Figure 2.16. They exhibit a clear Gaussian behavior. Hejhal and Rackner give convincing heuristic arguments for a local Gaussian behavior and they are led to

Conjecture 2.1 ([H-R]). *The distribution functions of the eigenfunctions $\phi_k(z)$ of $X = \Gamma\backslash\mathbb{H}^2$ tend to a Gaussian with mean 0 and standard deviation $Vol(X)^{-1/2}$, as $k \to \infty$.*

Section 2.3. Fine structure of the spectrum.

Figure 2.17 displays the first 250 even (i.e. invariant by $z \to -\overline{z}$) eigenvalues of $SL_2(\mathbb{Z})$ as computed recently by Steil [SL]. Previously, in his seminal computations Hejhal [H3] had determined the first 50 eigenvalues. Perhaps the most important point to note is that the spectrum is simple. In fact Steil has now checked this up to the 3000^{th} eigenvalue. The conjecture that the spectrum is simple has some important applications (see Theorem 3.15 for example). This spectrum is perhaps one of the most fundamental in number theory.

The number variance has been computed for the various examples discussed in Sections 2.1 and 2.2. In Figure 2.18, $\Sigma^2(L)$ is depicted for the Bunimovich Stadium and the semicircle billiard table [B-G]. The latter is integrable and as the results show $\Sigma^2(L)$ follows the expected Poisson. For the stadium it follows the G.O.E. as expected. Note however that in the nonuniversal range the number variance for this example must increase substantially because of the family of bouncing ball billiards – see Proposition 3.1.*

To describe the numerical results for the spectral statistics for $\Gamma\backslash\mathbb{H}^2$ we need to introduce the notion of an arithmetic group. These are the main objects of these

*Indeed this effect for the stadium has been confirmed recently in a physical experiment, see [G-H-L-L-R-R-S-W].

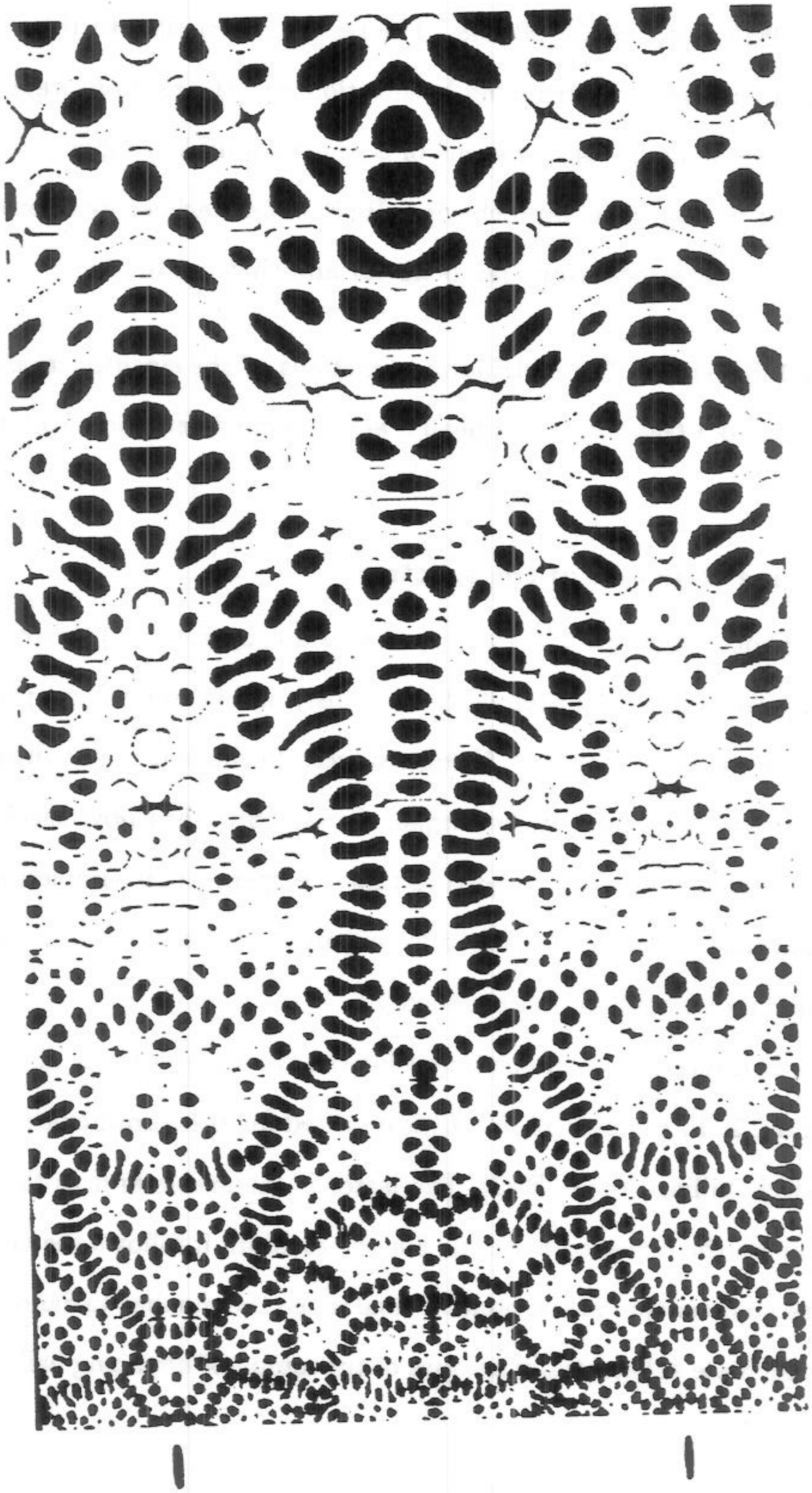

Figure 2.14. Density plot of the bound state (cusp form) for $SL_2(\mathbb{Z})\backslash\mathbb{H}$ with $R = 125.313840$.

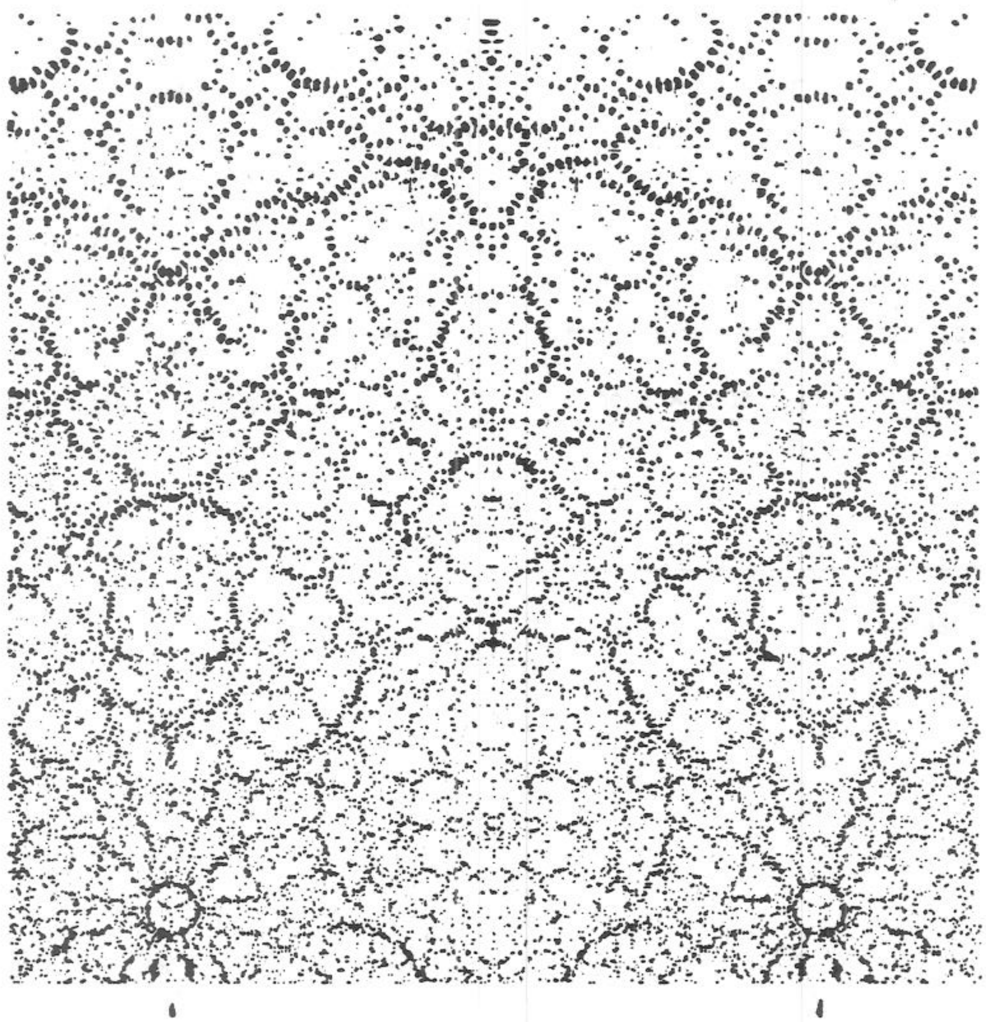

Figure 2.15. Density of the bound state of $SL_2(\mathbb{Z})\backslash\mathbb{H}^2$ with $R = 500.283548$. The points $e^{i\pi 3}$ and $e^{2\pi/3}$ which are the corners of the fundamental domain are encircled with mass.

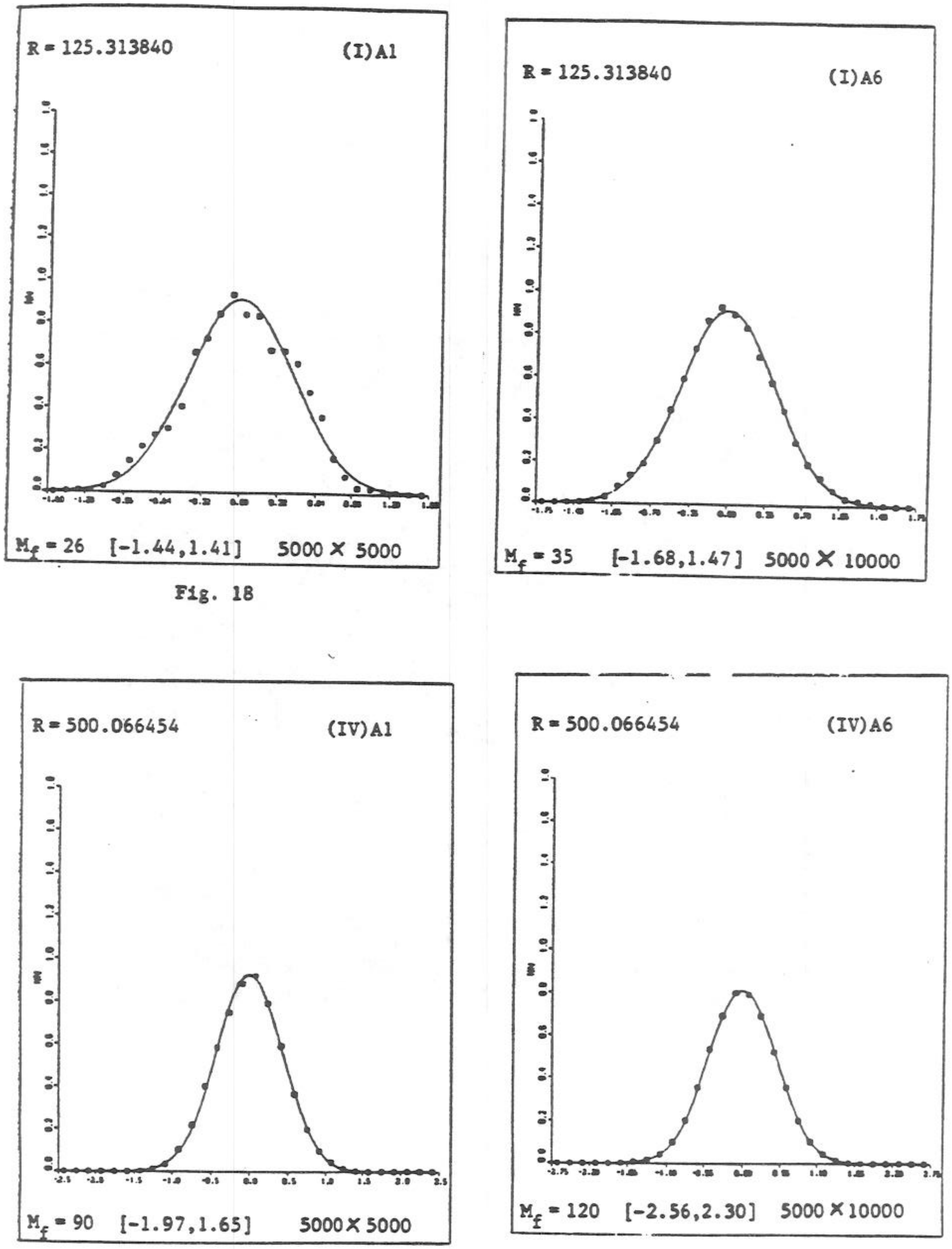

Figure 2.16. Local distribution of eigenfunctions for $SL_2(\mathbb{Z})\backslash\mathbb{H}$.
$A1 = [0, 0.2] \times [1, 1.2]$, $A6 = [-0.25, 0.25] \times [0.75, 3.75]$.

+	0	50	100	150	200
1	13.77975135189	50.25749138336	65.96209734837	77.98611704475	87.73649308039
2	17.73856338106	50.97145393295	66.23574908504	78.09252647365	87.76149544734
3	19.42348147083	51.29345538344	66.49100365587	78.20534347109	88.08154565610
4	21.31579594020	51.95599451481	67.04422392500	78.35339464386	88.16357318690
5	22.78590849419	51.97022883779	67.07029649539	78.37255228557	88.25453029914
6	24.11235272984	52.37358029018	67.34625908379	78.77823131287	88.51022821226
7	25.82624371271	52.62862052152	67.47742355856	78.95781318301	88.67026672751
8	26.15208544922	53.30057348875	68.13572078797	79.32224355438	88.96066800616
9	27.33270808315	53.36849756883	68.19151768676	79.33458819255	88.97436581691
10	28.53074769292	53.70758614195	68.27566473746	79.61334421892	89.43768617697
11	28.86339435392	54.45304383733	68.29637718449	79.89490171062	89.50582750141
12	30.41067880465	54.46777801657	68.88149627021	80.10361509406	89.54555368712
13	31.52658219679	54.47270843632	69.08836116230	80.26065640791	89.73343982086
14	31.56627541175	55.23229925600	69.25485308280	80.39368178281	89.80719411036
15	32.50811775991	55.37409277935	69.42696454672	80.87049505509	90.11362514641
16	32.89117021351	55.61692633822	69.92125375716	80.93955181528	90.41193016587
17	34.02788420010	56.23748384737	70.24290169136	81.10357062734	90.57861477668
18	34.45627153303	56.65882031164	70.32378254992	81.29439559126	90.70318734469
19	35.50234977137	56.89369689292	70.61572950727	81.32100656490	90.71234986439
20	35.84167643258	57.06541326358	71.01078372428	81.85612290670	90.95544578325
21	36.67755299315	57.54010430420	71.15704698684	82.02948035834	91.36718953829
22	36.85634949592	57.79524385295	71.37166757678	82.10306654087	91.39575276806
23	37.82507229059	57.99643152172	71.48995730929	82.14287108044	91.50982677824
24	38.30327615249	58.65093981625	71.70131027275	82.28298775627	91.63598281491
25	39.16808496793	58.78091445622	72.16797776854	82.70770750481	91.72529956741
26	39.40753186152	58.82020155454	72.23058701293	82.83179772306	91.84395089639
27	39.77362261904	59.11324417177	72.24874572552	83.03064561177	92.22025421815
28	40.54335121045	59.70160912149	72.38643028876	83.29192431362	92.36721794299
29	40.68866644493	59.93123655675	72.96286750763	83.56890078312	92.40899986691
30	41.55557767358	60.06709808489	73.12415395515	83.64002134805	92.78138740376
31	41.88300328542	60.46488819309	73.41907786242	83.91873587258	92.80189710505
32	42.64348841466	60.79740986912	73.49242967097	83.99881647469	93.07650975800
33	42.92222778356	60.82214006249	73.97080822746	84.24704551101	93.17186123241
34	43.26718203879	61.22103389976	74.10428030133	84.39590126839	93.46660516698
35	44.07740476167	61.80212021708	74.20416356027	84.77616371669	93.52238846484
36	44.42634811862	62.07168153918	74.43067661438	84.95024943562	93.54648940473
37	45.28743844249	62.20759315100	74.65128108423	84.95302975945	93.95631302682
38	45.36161360215	62.62501774403	74.93071791065	85.18848146530	94.18768303829
39	45.39846953131	62.97730391627	75.05855945306	85.29248644734	94.34175882653
40	46.10145632159	63.04910103639	75.45071957389	85.86252245475	94.39891249484
41	46.48140241232	63.19843639284	75.69209661168	85.86681912405	94.53732604252
42	46.65331835999	63.80322442707	76.12814986165	85.89707240986	94.60616740451
43	47.42289589850	64.10919880514	76.16884742696	85.92906122834	95.00492646164
44	47.92655833060	64.12517778026	76.27239358714	85.97613707223	95.18552860102
45	48.03933090509	64.21323387188	76.38882231717	86.57446409475	95.20790727963
46	48.74166634764	64.70981984493	76.61676668703	86.64937508745	95.30834305580
47	48.99830765408	65.00356195228	77.06049358257	86.85380347476	95.51585265716
48	49.68352007526	65.40390680728	77.21088762760	86.93039800832	95.69239631435
49	49.96169629050	65.40735293304	77.24214580553	86.95561068340	95.97201251932
50	50.08970547300	65.77666511886	77.61781292814	87.44469721193	96.04840199713

Figure 2.17. The first 200 even eigenvalues r for $SL_2(\mathbb{Z})$, $r^2 + 1/4 = \lambda$ [from Steil [SL]).

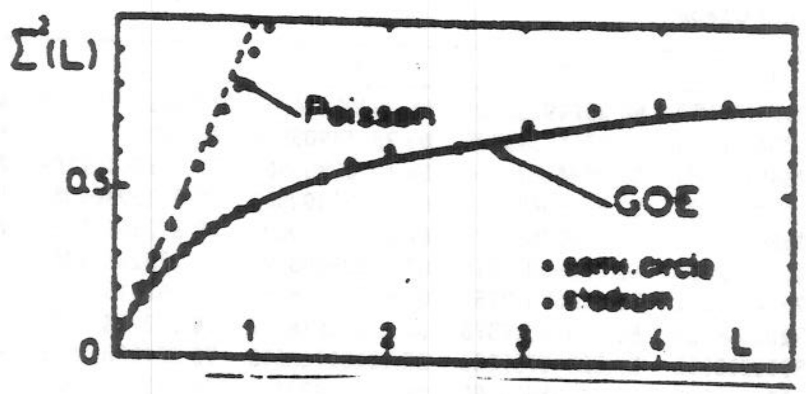

Figure 2.18

notes and the reason for the choice of the title. Firstly $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ is the most important example of an arithmetic lattice Γ in $\mathrm{SL}_2(\mathbb{R})$. Rather than give the usual definition [BO] we give an equivalent one, which is a characterization in terms of the traces of elements of Γ, due to Takeuchi [TA]. Let

$$\mathcal{L}(\Gamma) = \{\mathrm{tr}(\gamma) \mid \gamma \in \Gamma\} . \tag{2.1}$$

(A) Γ is arithmetic iff

(a) $K = Q(\mathcal{L}(\Gamma))$ is a finite extension of Q and $\mathcal{L}(\Gamma) \subset \mathcal{O}_K =$ the ring of integers of K.

(b) If $\phi\colon K \hookrightarrow \mathbb{C}$ is an embedding such that $\phi(t^2) \neq t^2$ for some $t \in \mathcal{L}(\Gamma)$ than $\phi(\mathcal{L}(\Gamma))$ is bounded in $\mathbb{C}$.

(B) Γ is derived from a quaternion algebra iff $\mathcal{L}(\Gamma)$ satisfies A(a) and (b). If ϕ is an embedding of K in $\mathbb{C}$, $\phi \neq$ identity, then $\phi(\mathcal{L}(\Gamma))$ is bounded.

For example using this, one can check that the Hecke groups Γ_m defined at the beginning of Section 2.2 are arithmetic only when $m = 3, 4, 6$. The group defined in Figure 2.6 is also arithmetic.

Examples of lattices derived from quaternion algebras are the following: Fix $a > 0$, $b \in \mathbb{Z}$, and let

$$\Gamma_{a,b} = \left\{ \begin{bmatrix} x_0 + x_1\sqrt{a} & x_2 + x_3\sqrt{a} \\ b(x_2 - x_3\sqrt{a}) & x_0 - x_1\sqrt{a} \end{bmatrix}, x_0^2 = ax_1^2 - bx_2^2 + abx_3^2 = 1 \right\} \tag{2.2}$$

(see S. Katok [K]).

Schmit [SCH] has computed $\Sigma^2(\lambda, L)$ for both arithmetic and nonarithmetic triangle groups (see also [B-G-G-S], [B-S-S]). In Figure 2.19 the results for an arithmetic triangle group are shown. Surprisingly Σ^2 is close to Poisson in the universal range. Other arithmetic triangles have been studied and they all have the same behavior. This includes the modular group $\Gamma = SL_2(\mathbb{Z})$ for which one can see the Poisson like behavior from the energy level plot in Figure 1.3. On the other hand in Figure 2.20, the number variance, computed by Schmit for a nonarithmetic triangle, is shown. Unlike the arithmetic case this as expected follows the G.O.E. distribution in the universal range. A similar behavior has been found for other nonarithmetic lattices.

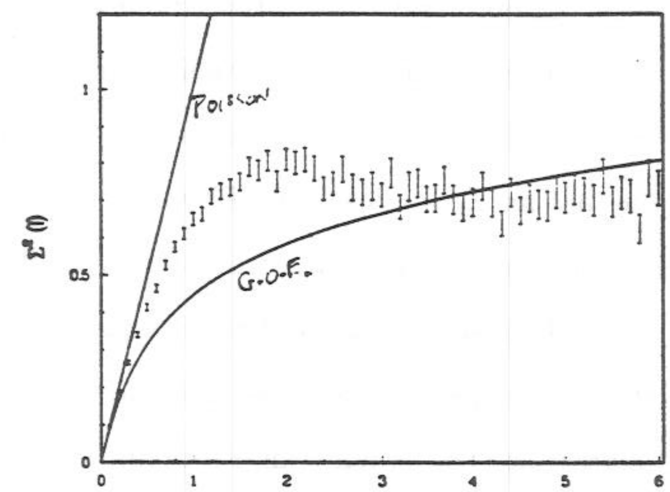

Figure 2.19: $\Sigma^2(L)$ for an Arithmetic Triangle

This completes our brief report on the numerical experiments. Summarizing the findings we have: At least for $n = 2$, the random wave model for the eigenfunctions appears to apply for both arithmetic and nonarithmetic hyperbolic surfaces. However only for the nonarithmetic surfaces do the eigenvalues behave according to random matrix theory. For the arithmetic surfaces the spectrum appears to be Poissonian. In the next section we state a series of results which explain these behaviors for arithmetic surfaces.

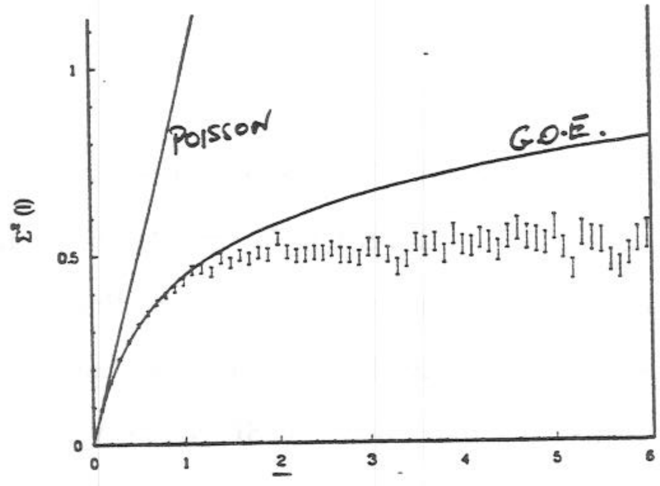

Figure 2.20: Σ^2 for a Nonarithmetic Triangle

Section 3. Arithmetic Manifolds

Before describing our results for arithmetic manifolds we make some remarks concerning the number variance for the stadium and similar manifolds. From the numerical results, viz Figure 2.18 we see that $\Sigma^2(\lambda, L)$ follows G.O.E. However we cannot expect that for $L >> L_{\mathrm{MAX}}, \Sigma^2(L)$ will oscillate in a bounded way about $\Sigma^2(\lambda, L_{\mathrm{MAX}})$. The reason is that the stadium has a family of "bouncing balls" (see Figure 2.2).These form a 1-parameter family of periodic billiards with a common period. Even though these points are of measure zero in phase space they affect the spectrum according to the following Proposition. It concerns the case without boundary but the effect in the case of the stadium should be similar.*

Proposition 3.1. *Let X be a compact 2-dimensional Riemannian manifold (without boundary). Assume that for some $T > 0$ the geodesic flow $\phi_t \colon S_1^*(X) \to S_1^*(X)$ satisfies* $\dim\{v \mid \phi_T v = v\} = 2$. *We also assume that this set satisfies some further standard technical conditions (see the proof in Section 5). Then for λ large,*

$$\frac{1}{\lambda}\int_{\lambda}^{2\lambda} \left|N(y) - \frac{\mathrm{Area}(X)}{4\pi}y\right|dy >> \lambda^{1/4} .$$

*It is, see footnote page 23.

In particular

$$\Sigma^2(\lambda, L) >> \lambda^{1/2} \quad \textit{for } L \sim \lambda \ .$$

The Sinai billiard (Figure 2.5) and Buminovich stadium have such families. So does a flat torus or more generally any case with integrable geodesic flow. For the case of the flat torus the lower bound implied by Proposition 3.1 gives a lower bound for the counting of lattice points in a circle radius $\sqrt{\lambda}$. In this case this lower bound is the well known Hardy-Landau Theorem [H-L]. Concerning the statistics of the level spacings for integrable Hamiltonians some interesting progress has been by Sinai [SI2] and Blecher [BL] for surfaces of revolution and by Uribe and Zelditch [U-Z] for Zoll surfaces.

We turn to the number variance for arithmetic surfaces. The following lower bounds show that these are highly nonrigid. Let $\overline{\Sigma}^2$ denote the average of Σ^2 over L, i.e.

$$\overline{\Sigma}(\Gamma, \lambda, L) = \frac{1}{L} \int_0^L \Sigma^2(\Gamma, \lambda, \xi) d\xi \ . \tag{3.1}$$

The behavior of $\overline{\Sigma}^2$ is similar to Σ^2 except for being a little smoother. In particular lower bounds for $\overline{\Sigma}^2$ imply ones for Σ^2.

Theorem 3.2 (Luo-Sarnak [L-S]). *Let $\Gamma \leq SL_2(\mathbb{R})$ be arithmetic then*

$$\overline{\Sigma}^2(\Gamma, \lambda, L) >> \frac{\sqrt{\lambda}}{(\log \lambda)^2}, \quad \textit{for } \frac{\sqrt{\lambda}}{\log \lambda} << L << L_{\text{MAX}} \simeq \sqrt{\lambda} \ .$$

Theorem 3.3 ([L-S]). *If Γ is derived from a quaternion algebra then*

$$\overline{\Sigma}^2(\Gamma, \lambda, L) >> L^2 \lambda^{-1/2}$$

$$\textit{for} \qquad \lambda^{1/4} << L << \lambda^{1/2} / \log \lambda \ .$$

Remarks 3.4.

(1) For $L \approx \sqrt{\lambda}/\log\lambda$ which is just inside the universal range, Theorem 3.2 asserts that

$$\overline{\Sigma}^2(\Gamma,\lambda,L) >> \frac{L}{\log L}\,. \tag{3.2}$$

This is close to establishing and consistent with the Poisson behavior which was found numerically.

(2) Theorem 3.3 gives in the more restricted case of Γ derived from a quaternion algebra, an effective lower bound for Σ^2 in the range $L \in [\lambda^{1/4}, \lambda^{1/2}/\log\lambda]$. It shows that in this range the spectrum is nonrigid and hence not the conjectured G.O.E.

The feature of arithmetic groups that is at the root of the nonrigidity of the spectrum is the high multiplicity of the length spectrum. The high multiplicity was observed by Selberg and Hejhal [H1] in their derivation of a lower bound for $|N(\lambda)-\lambda|$ for certain groups derived from quaternion algebras. This has been pointed out again in Bogomolny et al [B-G-G-S] and Steiner etal [A-S-S]. For us the key property of arithmetic groups in this connection is the following bounded clustering property (B-C) which may be easily deduced from Takeuchi's characterization given in Section 2. For $\Gamma \leq \mathrm{SL}_2(\mathbb{R})$ an arithmetic lattice there is $C(\Gamma) < \infty$ such that

$$|\mathcal{L}(\Gamma)\cap[m,m+1]| \leq C(\Gamma) \quad \text{for } m \in \mathbb{Z}\,. \tag{3.3}$$

This (B-C) property is technically difficult to work with, which is the reason for Theorem 3.2 being weaker than Theorem 3.3.* We will outline a proof of Theorem 3.2 in Section 5. While Theorem 3.2 explains the nonrigidity of the spectrum for the arithmetic quotients we emphasize that we have no upper bounds for the number variance.

We turn now to the behavior of the eigenfunctions for arithmetic quotients. Firstly we will also allow higher dimensional hyperbolic manifolds and specifically $n = 3$. Examples of arithmetic hyperbolic 3-folds are gotten from quadratic

*The (B-C) property may well be a geometric charcterization of arithmeticity.

forms as follows: Let $F(x) = F(x_1, x_2, x_3, x_4)$ be an integral quadratic form in 4 variables, which over $\mathbb{R}$ has signiture $(3, 1)$. For $\varepsilon = \pm 1$ chosen appropriately the hyperboloid $V = \{x \mid F(x) = \varepsilon\}$ consists of two sheets. The line element $ds^2 = F(dx_1, dx_2, dx_3, dx_4)$ restricted to $V(\mathbb{R})$ gives a model of hyperbolic 3-space. The orthogonal group of F denoted G_F acts on V as isometries and the quotient $X_F = G_F(\mathbb{Z})\backslash V(\mathbb{R})$, where $G_F(\mathbb{Z})$ consists of all 4×4 integral matrices preserving F, gives a hyperbolic 3-manifold. If F is anistropic over Q (that is, $F(x) = 0$ for $x \in Q^4 \Rightarrow x = 0$) then it is known that X_F is compact [CA]. An example is $F_0(x) = x_1^2 + x_2^2 + x_3^2 - 7x_4^2$. It is anisotropic, as may be seen by considering F_0 modulo 8.

In general when studying eigenfunctions by trace like or spectral expansions, one is limited to investigating sums of these quantities over quite large energy ranges. This makes the investigation of individual eigenfunctions very difficult. For general manifolds we don't know how to go beyond this natural barrier but in the arithmetic quotient case this can be done. The key point here is the analytic use of certain arithmetically defined operators, called Hecke operators. These arise from correspondences on the manifold X. By a correspondence C of order $r \geq 1$ on X we mean a map $z \xrightarrow{C} \{z_1, \dots, z_r\}$ of $X \to X \times X \ldots \times X/\Sigma$ where Σ is the permutation group of r-letters and where $z_j(z)$ are locally isometries of z (The z_j's are not individually globally defined but as a set they are). In particular a correspondence of order 1 is an isometry of X. If $X = \Gamma\backslash S$ with S a Riemannian space and Γ a subgroup of the isometry group G of S, then we may obtain such a correspondence as follows: Given $\delta \in G$ for which $\delta^{-1}\Gamma\delta \cap \Gamma = B$ is of finite index in both Γ and $\delta^{-1}\Gamma\delta$, let C_δ be given by

$$\Gamma x \overset{C_\delta}{\rightsquigarrow} \{\Gamma\delta\alpha_1 x, \dots, \gamma\delta\alpha_r x\} \tag{3.4}$$

where $\Gamma = \bigcup_{j=1}^{r} B\alpha_j$ is a coset decomposition.

C_δ is then a correspondence on X. To such a C_δ we associate a Hecke operator

$$T_\delta \colon L^2(X) \to L^2(X)$$
$$T_\delta f(x) = \sum_{j=1}^{r} f(\delta \alpha_j x) . \tag{3.5}$$

It is easily checked that T_δ is well defined and moreover since $\delta\alpha_j \in G$, T_δ commutes with the Laplacian Δ. The set of such T's generate an algebra which is of most significance when the commensurator $\mathrm{COM}(\Gamma) = \{\delta \in G \mid \delta^{-1}\Gamma\delta \cap \Gamma$ is finite index in both$\}$, is large. For $G = \mathrm{SL}_2$, or $\mathrm{SO}(n,1)$ (or any other semisimple Lie group for that matter) Margulis [M] has shown that $\mathrm{COM}(\Gamma)$ is dense in $G(\mathbb{R})$ iff Γ is arithmetic. In fact if Γ is not arithmetic then $\mathrm{COM}(\Gamma)/\Gamma$ is a finite group. Hence being arithmetic is equivalent to having an infinite family of correspondences.

Examples 3.5. *(i) For $\Gamma = PSL_2(\mathbb{Z})$ in $PGL_2(\mathbb{R})$, $COM(\Gamma) = PGL_2(Q)$. If $n \geq 1$, $\delta_n = \begin{bmatrix} n & 0 \\ 0 & 1 \end{bmatrix} \in COM(\Gamma)$ and it gives rise to the classical Hecke operators $T_n \colon L^2(\Gamma\backslash\mathbb{H}^2) \to L^2(\Gamma\backslash\mathbb{H}^2)$ defined by (see [SER])*

$$T_n f(z) = \sum_{\delta \in \Gamma\backslash R(n)} f(\delta z) = \sum_{\substack{b \bmod d \\ ad=n \\ d>0}} f\Big(\frac{az+b}{d}\Big) . \tag{3.6}$$

Here $R(n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix}, ad - bc = n, a, b, c, d \in \mathbb{Z} \right\}$.

The T_n's satisfy

$$T_n T_m = \sum_{d \mid (n,m)} d T_{nm/d^2} . \tag{3.6$'$}$$

Moreover they commute with Δ and they are self adjoint on $L^2(\Gamma\backslash\mathbb{H}^2)$. In choosing the eigenfunctions ϕ_k of Δ we may therefore assume that these are also eigenfunctions of the Hecke algebra. According to the numerical results discussed at the beginning of Section 2.3 this assumption about ϕ_k being a Hecke eigenform is probably automatic since the spectrum is in all liklihood simple. In any case it is this basis of eigenstates that is of interest in arithmetic.

(ii) The examples $\Gamma_{a,b}$ of (2.2) are called quaternion groups. Let A be the quaternion algebra over Q (one can of course work over more general number fields as well) generated linearly by 1, ω, Ω, $\omega\Omega$, where $\omega^2 = a$, $\Omega^2 = b$, $\omega\Omega + \Omega\omega = 0$. Then $\Gamma_{a,b}$ corresponds to $R(1) = \{\alpha \in A(\mathbb{Z}) \mid N(\alpha) = 1\}$, $N(\alpha)$ being equal to $\alpha\overline{\alpha}$ where $\overline{\alpha} = x_0 - x,\omega - x_2\Omega - x_3\omega\Omega$ if $\alpha = x_0 + x_1\omega + x_2\Omega + x_3\omega\Omega$. Just as with example (i) above we can use the set $R(1)\backslash R(n)$, where $R(n) = \{\alpha \in A(\mathbb{Z}) \mid N(\alpha) = n\}$, to define a correspondence $C_n : \Gamma_{a,b}\backslash\mathbb{H} \to \Gamma_{a,b}\backslash\mathbb{H}$. One can also form the Hecke operators T_n (see Eichler [EI]).

We say Γ is a congruence group if Γ is a subgroup the unit group of a quaternion algebra and Γ contains $\Gamma(q)$ for some q, where $\Gamma(q) = \{\alpha \in \Gamma \mid \alpha \equiv 1 (\bmod\ q)\}$. For these we have a commutative self adjoint Hecke algebra. In this case we will always assume that the basis of eigenfunctions ϕ_k are also Hecke eigenfunctions.

In connection with the behavior of the eigenfunctions our first result is that strong scarring on closed geodesics does not occur.

Theorem 3.6 (Rudnick-Sarnak [R-S]).

Let X be a congruence arithmetic manifold of dimension 2 or 3 then there is no strong scarring on closed geodesics or more generally proper geodesic submanifolds. Precisely, if the singular support of a quantum limit ν is contained in a finite union of points, or closed geodesics or closed totally geodesic surfaces (in the case $n = 3$) then the singular support must be the empty set (that is ν is absolutely continuous).

In particular this answers the question of Colin de Verdiere (see page 16) at least for congruence groups and Hecke bases. The above Theorem is a first step towards the following conjecture which rules out localization.

Conjecture 3.7 ([R-S]) Quantum Unique Ergodicity. *If $K < 0$ and $n = 2$ or 3 then $\mu_k \to dV/Vol(X)$ as $k \to \infty$. That is the μ_k's become (individually) equidistributed.*

Further evidence for this conjecture in special cases will be given in Section 4 where its relation to the Lindelhof-Hypothesis in the theory of L-functions, will be explained. This relation leads to the following quantitative form of the conjecture for hyperbolic surfaces: For $\varepsilon > 0$,

$$\int_X |\phi_k(z)|^2 f(z)dv(z) = \frac{1}{\mathrm{Vol}(X)} \int_X f(z)dv(z) + O_{\varepsilon,f}(\lambda_k^{-1/4+\varepsilon}) \,. \tag{3.7}$$

One can show that the exponent $-1/4$ in (3.7) cannot be replaced by anything smaller (c.f. the question in [CD-3, pp.]). Conjecture 3.7 if true is remarkable since it asserts that at the quantum level one has unique ergodicity (the conjecture extends to the $\tilde{\mu}_k$'s in (1.9) and asserts the uniqueness of such quantum limits) while classically unique ergodicity (i.e. uniqueness of invariant measures for the Hamiltonian flow) is never satisfied by chaotic systems. The proof of Theorem 3.6 is based on X having a large family of correspondences and the ϕ's being Hecke eigenforms, see Section 5.

All of the above is consistent with the experiments and the random wave model. To examine this further we investigate the sizes of the eigenfunctions. For a general compact X we define

Definiton 3.8. *For $2 < p \leq \infty$ let*

$$M_p(\lambda) = \begin{cases} \max\limits_{\Delta\phi+\lambda\phi=0} \frac{\|\phi\|_p}{\|\phi\|_2} & \text{for } \lambda \in \ Spec(\Delta) \\ 0 & otherwise \end{cases} .$$

There are general bounds for $M_p(\lambda)$ which are derived from local considerations (and in particular do not depend on the geometry of X) due to Seger and Sogge [S-S]:

Theorem 3.9. *Let $\delta(p) = \max\{n|\frac{1}{p} - \frac{1}{2}| - \frac{1}{2}, 0\}$. Then*

$$M_q(\lambda) = \begin{cases} 0(\lambda^{\delta(q)/2}), & \frac{2(n+1)}{n-1} \leq q \leq \infty \\ 0(\lambda^{(n-1)(2-q')/q'}) & 2 \leq q \leq \frac{2(n+1)}{n-1} \end{cases}$$

where $\frac{1}{q} + \frac{1}{q'} = 1$.

In particular

$$M_\infty(\lambda) = O(\lambda^{\frac{n-1}{4}}) . \tag{3.8}$$

The point in Theorem 3.9 is that for general $2 < q < \infty$ it gives a better bound than what one would get by simply interpolating between (3.8) and $M_2(\lambda) = O(1)$. The improvement is closely related to the restriction theorem, see Stein [ST]. As mentioned above Theorem 3.9 is local and while it is sharp for the n-sphere it presumably is far from the truth when $K < 0$ (i.e. the Q.C. regime). Even for the n-torus $X = \mathbb{R}^n/\mathbb{Z}^n$ where the eigenfunctions are exponentials, $M_p(\lambda)$ is not completely understood because of the multiplicities of the eigenvalues. For $X = \mathbb{R}^2/\mathbb{Z}^2$ we have $M_\infty(\lambda) = \sqrt{\mu(\lambda)}$ where $\mu(\lambda)$ is the multiplicity of λ. From elementary properties of sums of two squares it follows that for this X

$$\mu(\lambda) = O_\varepsilon(\lambda^\varepsilon) \quad \text{for all } \varepsilon > 0$$

$$M_\infty(\lambda) = O_\varepsilon(\lambda^\varepsilon) \quad \text{for all } \varepsilon > 0 . \tag{3.9}$$

Also $\mu(\lambda)$ is unbounded. It is less obvious that

$$M_4(\lambda) \leq \sqrt[4]{5} \tag{3.10}$$

which is due to Zygmund [Z]. This last estimate shows that on $\mathbb{R}^2/\mathbb{Z}^2$ every quantum limit is absolutely continuous. For $n \geq 3$ a bound like (3.10) for some $p > 2$ is not known. However Bourgain [BO] has recently shown that for $n \geq 4$ and $p \geq \frac{2(n+1)}{n-3}$, $M_p(\lambda)) \underset{\varepsilon}{<<} \lambda^{\frac{n-2}{4} - \frac{n}{2p} + \varepsilon}$. Moreover in this range these estimates are essentially sharp.

Returning to the chaotic case the following is our main conjecture.

Conjecture 3.10. *For $n = 2$ and $K < 0$*

$$M_\infty(\lambda) = O_\varepsilon(\lambda^\varepsilon) \quad \textit{for all } \varepsilon > 0 .$$

Some comments about this conjecture are in order. Firstly the conjecture is consistent with the random wave model (see 1.10). In fact it emerges from random wave considerations in much the same way as the Ramanujan conjecture (see 4.25) emerges from random matrix theory (see the comments at the end of this section). Thus we may view Conjecture 3.10 as the eigenfunction analogue of the Ramanujan conjectures. In the same way Conjecture 2.1 may be viewed as the eigenfunction analogue of the Sato-Tate conjectures (see [SA3]). There is further basis for Conjecture 3.10 in the arithmetic quotient case. We show in Section 4 that it implies the Lindelof Hypothesis for the Riemann zeta function as well as certain other L-functions. Thus if true the conjecture is very deep. It also puts these analytic questions about L-functions in a new context. The next theorem gives a partial result towards Conjecture 3.10. It gives the first improvement over the "convexity bound" (3.8) (recall that when dealing with congruence groups we are assuming ϕ's are Hecke eigenforms).

Theorem 3.11 (Iwaniec-Sarnak [I-S]). *If $\Gamma\backslash\mathbb{H}^2$ is a congruence quotient then*

(a) $M_\infty(\lambda) = O_\varepsilon(\lambda^{5/24+\varepsilon})$ for $\varepsilon > 0$.

(b) $M_\infty(\lambda_j) \geq c\sqrt{\log\log\lambda_j}$ for $c > 0$ and infinitely many j's.

Remarks 3.12.

(i) Part (b) of Theorem 3.11 shows in particular that the ϕ_j's are not uniformly bounded which is by no means obvious. It is consistent with the large deviations that occur for random waves – viz (1.10). The proof of (b) proceeds by using the eigenvalues of the Hecke operators as weights in sums over the spectrum (see Section 5). In this way one is able to investigate the individual eigenfunctions beyond what the usual trace methods yield. In fact the proof shows that the eigenfunctions are large at a certain dense set of arithmetical points of X. For $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ these points are the "C.M." points $\{z \in \mathbb{H}^2 \mid az^2 + bz + c = 0, a, b, c \in \mathbb{Z}\}$. The point

$\rho = e^{\pi i/3}$ is such a point and it clearly an enhanced point in Figure 2.15.

(ii) The proof of the upper bound (a) employs polynomials in the Hecke eigenvalues as weights in certain spectral sums (see Section 5). Thus by arithmetical means we are able to go beyond the local bound (3.8). It would be much more satisfactory to bring in the chaotic dynamics as the global ingredient (rather than arithmetic) to establish a result like (a). However at present no technique is known to do so. Part (a) actually leads to non-trivial bounds for the zeta function (see Section 4) on the line $\mathrm{Re}(s) = 1/2$ (by the same method used to show that Conjecture 3.10 implies the Lindelof-Hypothesis). While better results are known for the zeta function this does place Theorem 3.11 (a) in context.

(iii) The proof of Theorem 3.11 applied for Hecke eigenbases only. In general it is easy to see that the multiplicity $\mu(\lambda)$ of eigenvalue λ satisfies

$$\mu(\lambda) \leq (M_\infty(\lambda))^2 \; . \tag{3.11}$$

Hence the main Conjecture 3.10 implies that

$$\mu(\lambda) = O_\epsilon(\lambda^\epsilon) \; . \tag{3.12}$$

Unfortunately for arithmetic hyperbolic surfaces we know of no better bound for $\mu(\lambda)$ than

$$\mu(\lambda) = O(\lambda^{1/2}/\log\lambda) \tag{3.13}$$

which follows from Berard [BD]. This is a sad state of affairs since recall that for $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, $\mu(\lambda) \leq 1$ is most likely true (see the beginning of Section 2.3).

All of the above results are consistent with the random wave model for eigenfunctions. However for $n = 3$ and Γ arithmetic the eigenfunctions can display much more drastic variations. At least a certain subset of the eigenfunctions do not behave like random waves. Let X_F be the hyperbolic 3-fold associated to the quadratic form F as on page 32.

Theorem 3.13 (Rudnick-Sarnak [R-S]).

Let X_F be as above, then for $6 < p \leq \infty$

$$M_p(\lambda_j) \geq c_p \lambda_j^{1/4-3/2p}$$

for c_p a positive constant and infinitely many j's.

The proof of Theorem 3.13 uses Siegel Theta Functions [SHI] and will not be discussed further here. The subsequence of eigenfunctions exhibiting this singular behavior of being very large on a dense set of arithmetical points, are "theta lifts" from SL_2 [R-S]. That theta lifts can be singular in other aspects and for other groups is not a new phenomenon. For example the failure of the Ramanujan Conjectures for GSP_4 [KU, H-PS] or $\mathrm{SO}(n,1)$ $n \geq 4$ [B-L-S]. However the fact that these ϕ_j's are so singular is quite unexpected. The lower bound in (3.13) is not compatible with (1.10) and so these theta lifts do not behave like random waves. The reason for the behavior in Theorem 3.13 is beautifully illustrated in the following special example.

Let F_0 be the form defined on page 32 and let $V_0 = \{x \mid F_0(x) = -1\}$, $\Gamma = G_F(\mathbb{Z})$, $Y = \Gamma \backslash V_0(\mathbb{R})$.

Theorem 3.14 ([R-S]).

Let $P = (2,1,1,1) \in Y$. Of the $N(\lambda)$ $(= c\lambda^{3/2} + O(\lambda))$ eigenfunctions with $\lambda_j \leq \lambda$ at most $O(\lambda)$ do not vanish at P.

The fact that so many eigenfunctions vanish at P is what forces those that do not, to be very large at P. Once the L^∞-norm is shown to be as large as indicated, one can deduce from the fact that ϕ_j is an eigenfunction, that its L^p norm is large for the range of p indicated. The special feature of P, is that $V_0(\mathbb{Z})$ (that is the integral points of V_0) consists of a single Γ orbit. This is very special. It is an interesting problem to explain the geometric source (rather than arithmetical) for these singular eigenfunctions as well as the vanishing at P. Recall that this is an example of a chaotic Hamiltonian with 3 degrees of freedom and there are no caustics or focusing.

Before ending this Section we report briefly on some recent developments concerning existence of bound states (that is cusp forms) for noncompact but finite area surfaces $\Gamma\backslash\mathbb{H}^2$. Some years ago [SA2] we made the conjecture that the existence of infinitely many such forms is intimately tied to the arithmeticity of Γ. For example consider the triangles T_m in Figure 2.7. We seek nonconstant L^2-solutions to $\Delta u + \lambda u = 0$, $\partial_n u = 0$ on ∂T_m (that is Neumann boundary conditions). According to the arithmeticity remarks on page 27, our Conjecture concerning the existence of such solutions asserts that unless $m = 3, 4, 6$ there should be no such solutions. For $m = 3, 4, 6$ Selberg [SEL2] proved that such cusp forms exist in abundance. Numerical experiments by Hejhal [H3] have confirmed that for $m \neq 3, 4, 6$ no cusp forms exist for λ up to about 3600. The above conjecture emerged from the theory developed in Phillips-Sarnak [P-S1] concerning the behavior of cusp forms under deformation of Γ. It was shown that the nonvanishing of certain L-functions (precisely, Rankin-Selberg L-functions – see Section 4 for definitions) at special points ensures that such a form is dissolved generically in the moduli space. This nonvanishing condition, which may be viewed as a form of Fermi's Golden Rule [PS2, PS3], is especially Golden here since one can show by techniques from number theory that many of these special values are not zero. Recent developments by LUO [LU] give the strongest results to date on nonvanishing. Wolpert [W] has made further significant progress by developing the above theory for singular deformations of Γ's degenerating to the boundary of moduli space. For example one striking application of his theory, due to Judge [JU], is the following:

Theorem 3.15 (Judge). *Assume that the spectrum of $\Gamma_0(D)\backslash\mathbb{H}^2$, $D = 1, 2, 4$ is simple on new-forms, then for all but a countable number of $m \in (2, \infty)$ the triangle T_m has a finite number of bound states.*

In the above, $\Gamma_0(D) = \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) \mid \gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with $D|C\}$. $L^2(\Gamma_0(D)\backslash\mathbb{H}^2)$ splits into a space of new-forms and old forms [A-L]. The assumption about the simplicity of the new form spectrum for $\Gamma_0(1) = \mathrm{SL}_2(\mathbb{Z})$ is the same one that we have mentioned previously in Section 2.3. Again we see the importance of this multiplicity question. In Phillips and Sarnak [PS-4] a related result is proven under the much milder assumption that a positive proposition of the eigenvalues are of bounded multiplicity. Such information is a consequence of the λ_j's having a Poisson level spacing. What is needed here are upper bounds on the number variance!

To end this Section we make some remarks about the reversal of the use of random matrix theory. While the random matrix model does not apply to the Laplace eigenvalues for arithmetic surfaces it is appropriate for the spectrum of the Hecke operators (page 36). In [SA3] we showed that the Wigner semicircle level density corresponds to the Sato-Tate conjecture for the Hecke eigenvalues. Moreover Friedman [FR] has shown that for random regular graphs, which are the appropriate models for these Hecke operators, the eigenvalues satisfy bounds which correspond to the Ramanujan Conjectures [SA1]. Given this it is perhaps not too surprising that one can reverse the philosophy and use these Hecke correspondences to give explicit constructions of graphs which mimick random ones in many respects. This was done some time ago by Lubotzky-Phillips-Sarnak [L-P-S] and independently by Margulis [M2]. These explicit graphs are known as Ramanujan graphs and have found many practical applications [SA1].

SECTION 4: L-FUNCTIONS

We begin with Riemann's zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_{p} (1 - p^{-s})^{-1} . \tag{4.1}$$

The infinite "Euler" product is over the primes and the above definition and identity holds for $\mathrm{Re}(s) > 1$. Besides the Euler product $\zeta(s)$ has an analytic continuation and functional equation. Set

$$\xi(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s) .$$

Then

$$\xi(s) = \xi(1-s) . \tag{4.2}$$

The Riemann hypothesis, R-H, asserts that all the nontrivial zeros (i.e. those with $0 < \mathrm{Re}(s) < 1$) lie on $\mathrm{Re}(s) = 1/2$. This remains one of the major unsolved problems. A problem of importance in number theory is to estimate $\zeta(s)$ along the line $\mathrm{Re}(s) = 1/2$. The Lindelof hypothesis, L-H, asserts that

$$\zeta(\frac{1}{2} + it) = O_\varepsilon(1 + |t|)^\varepsilon \ , \varepsilon > 0 . \tag{4.3}$$

One can easily deduce the bound $\zeta(\frac{1}{2} + it) = O(|t|^{1/4})$ using convexity arguments (in particular the Phragmen-Lindelof principle [T]) together with the functional equation (4.2). Weyl (or at least his method) was the first to go beyond the convexity bound by using exponential sums. He showed that $\zeta(\frac{1}{2} + it) = O(|t|^{1/6})$. More recently Bombieri and Iwaniec [B-I] obtained $\zeta(1/2 + it) = O_\varepsilon(|t|^{9/56+\varepsilon})$. One reason to believe the L-H is that it follows from R-H by the convexity argument applied to $\log \zeta(s)$ [T].

More generally one can consider Euler products of higher degree. By an Euler product of degree k we mean an L-function (this is the name given to a generalized zeta function) of the form

$$L(s)=\sum_{n=1}^{\infty}\frac{a_n}{n^s}=\prod_p\prod_{j=1}^{k}\left(1-\alpha_{j,p}p^{-s}\right)^{-1} . \tag{4.4}$$

We assume that $a_1=1$ and that

$$a_n=O_\varepsilon(n^\varepsilon)\ ,\varepsilon>0 . \tag{4.5}$$

As with the zeta function we require that $L(s)$ have an analytic continuation and functional equation of the form

$$\xi(1-\overline{s})=\overline{\xi(s)} \tag{4.6}$$

where

$$\xi(s)=Q^s\Big(\prod_{j=1}^{k}\Gamma(\frac{s}{2}+ir_j)\Big)L(s) \tag{4.7}$$

$Q>0$, $r_j\in\mathbb{C}$.

Such L-functions include Dirichlet L-functions (see [DA]) as well as L-functions associated with automorphic forms [LA]. It is in the latter form that we will need them below. They have also been considered recently by Selberg [SEL3] and Piatetsky-Shapiro [P]. An interesting problem in connection with these Euler products is to show that except for simple change of variables and deformations [C-G] there are only a countable number of such Euler products. Such would have to be the case if they all come from automorphic forms. As with the zeta function these L-functions have a R-H. That is all their nontrivial zeros should lie on $\mathrm{Re}(s)=\frac{1}{2}$. Also, as before, the R-H implies the L-H in the t, Q and r aspects. That is:

Lindelof-Hypothesis 4.1. *For* $\varepsilon>0$

(i) $$L(\frac{1}{2}+it)=O_{\varepsilon,Q,r}((|t|+1)^\varepsilon$$

(ii) $$L(\frac{1}{2}+it_o)=O_{\varepsilon,t_0,r}(Q^\varepsilon)$$

(iii) $$L(\frac{1}{2}+it_0) = O_{\varepsilon,t_o,Q}(N(r)^{\varepsilon}) \text{ where } N(r) = \prod_{j=1}^{k}(1+|r_j|)\ .$$

The standard convexity argument which relies on the functional equation but not on the Euler product leads to

Proposition 4.2. *For* $\varepsilon > 0$,

(i) $$L(\frac{1}{2}+it) = O_{\varepsilon,r,Q}((|t|+1)^{k/4+\varepsilon})$$

(ii) $$L(\frac{1}{2}+it_o) = O_{\varepsilon,r,t_0}(Q^{1/2+\varepsilon})$$

(iii) $$L(\frac{1}{2}+it_0) = O_{\varepsilon,t_0,Q}(N(r)^{1/4+\varepsilon})\ .$$

The recent results of Iwaniec [I2] and Duke-Friedlander-Iwaniec [D-F-I] establish improvements (in the exponent) over the convexity bound in all three aspects, for all Euler products of degree at most 2 (coming from automorphic forms).

We now discuss the relation between these L functions and the previous lectures and in particular Conjectures 3.7 and 3.10. The relation emerges on considering $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ and its congruence subgroups. The continuous spectrum for $\Gamma\backslash\mathbb{H}^2$ is furnished by the Eisenstein series $E(z,s)$ which are defined by:

$$E(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} (y(\gamma z))^s \tag{4.8}$$

where

$$\Gamma_\infty = \left\{ \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \mid m \in \mathbb{Z} \right\}.$$

This converges for $\mathrm{Re}(s) > 1$. Since y^s satisfies

$$\Delta(y^s) + s(1-s)y^s = 0$$

it follows that

$$\Delta E(z,s) + s(1-s)E(z,s) = 0\ . \tag{4.9}$$

Thus $E(z,s)$ are eigenfunctions of Δ, however they are not in $L^2(\Gamma\backslash\mathbb{H})$. For $\mathrm{Re}(s) = 1/2$ they are almost in $L^2(\Gamma\backslash\mathbb{H})$ and furnish the extended states (i.e. continuous spectrum). We may see this from the Fourier expansion; since $E(z,s) = E(z+1,s)$ we develop E in such an expansion in $e^{2\pi inx}$. A straightforward calculation (see [SA1]) yields

$$E(z,s) = y^s + \phi(s)y^{1-s} + \frac{2y^{1/2}}{\xi(2s)}\sum_{n=1}^{\infty} n^{s-1/2}\sigma_{1-2s}(n)K_{s-1/2}(2\pi|n|y)\cos(2\pi nx) \tag{4.10}$$

where $\xi(s)$ is defined in (4.2) and

$$\begin{aligned}\phi(s) &= \frac{\xi(2s-1)}{\xi(2s)}\\ \sigma_\nu(n) &= \sum_{d|n} d^\nu\\ K_\nu(t) &= \int_0^\infty e^{-t\cosh u}\cosh(\nu u)du\ .\end{aligned} \tag{4.11}$$

The series (4.10) is rapidly convergent and from it we can easily deduce its meromorhic continuation in s. Also we see that $E(z,\frac{1}{2}+it)$ is almost square integrable on $\Gamma\backslash\mathbb{H}$. The orthogonal compliment, of the span $E(z,\frac{1}{2}+it)$ together with the constant function, consists of the cuspidal space [SEL2, HE2]

$$L_0^2(\Gamma\backslash\mathbb{H}) = \left\{f \in L^2(\Gamma\backslash\mathbb{H}) \mid \int_0^1 f(x,y)dx = 0 \quad \text{for } a.a.\ y\right\}\ .$$

The spectrum of Δ on $L_0^2(\Gamma\backslash\mathbb{H})$ is discrete and we may choose an o.n.b. $\phi_j(z)$ of such eigenfunctions. These are the cusp forms for $\mathrm{SL}_2(\mathbb{Z})$ mentioned in the previous sections. We may choose them to be eigenfunctions of the Hecke operators

T_n, $n \geq 1$. Write

$$\Delta\phi_j + \lambda_j\phi_j = 0$$
$$T_n\phi_j = \lambda_j(n)\phi_j$$
$$\eta_j(n) = \lambda_j(n)/\sqrt{n}\ . \tag{4.12}$$

The ϕ_j's and $E(z, \frac{1}{2}+it)$ play the same role and we may test our conjectures on the Eisenstein series (whatever can be proven for ϕ_j can also be proven for $E(\cdot, 1/2+it)$). Consider the $M_\infty(\lambda)$ problem for $E(z, 1/2+it)$. Since $E(\cdot, 1/2+it)$ is not square integrable we cannot L^2-normalize it. One way to get around this, which we choose to do here, is to consider the local analogue. That is we normalize E on a compact subset of $\Gamma\backslash\mathbb{H}^2$. This is equivalent to cutting off the zeroth coefficient of $E(z, 1/2+it)$ in the cusp. Define $E_A(z,s)$ where A is a large constant by

$$E_A(z,s) = \begin{cases} E(z,s) & \text{if } y \leq A, z \in \Gamma\backslash\mathbb{H} \\ E(z,s) - y^s - \phi(s)y^{1-s} & \text{if } y > A\ . \end{cases} \tag{4.13}$$

Then $E_A \in L^2(\Gamma\backslash\mathbb{H})$. An important calculation [SEL2] yields

$$\int_{\Gamma\backslash\mathbb{H}} |E_A(z, \frac{1}{2}+it)|^2 \frac{dxdx}{y^2} = 2\log A - \frac{\phi'}{\phi}(\frac{1}{2}+it) + \frac{\overline{\phi(\frac{1}{2}+it)}A^{2it} - \phi(\frac{1}{2}+it)A^{-2it}}{2it}\ . \tag{4.14}$$

Now $\frac{\phi'}{\phi}(\frac{1}{2}+it)$ involves the Gamma function and the zeta function at $1+2it$. Hence well known estimates for zeta [T] lead to

$$1 << \int_{\Gamma\backslash\mathbb{H}} |E_A(z, \frac{1}{2}+it)|^2 \frac{dxdy}{y^2} \underset{\varepsilon}{<<} (1+|t|)^\varepsilon\ . \tag{4.15}$$

That is E_A is essentially L^2-normalized. Consider now

$$M_\infty(\lambda) = M_\infty(1/4+t^2) = \max_z |E_A(z, \frac{1}{2}+it)|\ . \tag{4.16}$$

For $z = \sqrt{-1} := i$

$$E(i,s) = \sum_{(c,d)=1} (c^2+d^2)^{-s} = \frac{4\zeta_k(s)}{\zeta(s)} \tag{4.17}$$

where $k = Q(\sqrt{-1})$ and $\zeta_k(s)$ is its (Dedekind) zeta function. In fact $\zeta_k(s) = L(s,\chi_4)\zeta(s)$ with

$$L(s,\chi_4) = \sum_{n=1}^{\infty} \frac{\chi_4(s)}{n^s} = \prod_p \left(1 - \chi_4(p)p^{-s}\right)^{-1}$$

$$\chi_4(p) = \begin{cases} 1 & \text{if } p \equiv 1(4) \\ -1 & \text{if } p \equiv 3(4) \\ 0 & \text{if } p = 2\ . \end{cases}$$

Thus $\zeta_k(s)$ is an Euler product of degree 2 in the sense defined at the beginning of this Section. Again we need well known estimates for $\zeta(s)$ on $\mathrm{Re}(s) = 1$.

$$(1+|t|)^{-\varepsilon} \underset{\varepsilon}{<<} \zeta(1+it) \underset{\varepsilon}{<<} (1+|t|)^{\varepsilon}\ . \tag{4.18}$$

With this and (4.17) we see that the main Conjecture 3.10 implies

$$\zeta_k(\frac{1}{2}+it) = O_\varepsilon(1+|t|)^\varepsilon\ . \tag{4.19}$$

That is it implies the L-H for this classical zeta function. We note also that the convexity bound for $\zeta_k(s)$ via Proposition (4.2)(ii) reads

$$E(i,\frac{1}{2}+it) = O_\varepsilon\left((|t|+1)^{1/2+\varepsilon}\right)\ .$$

This is essentially the same as the general bound (3.8). By considering E evaluated at other CM points (see page 55) one can show that Conjecture (3.10) implies L-H for $\zeta_k(s)$ where k is any imaginary quadratic number field from which the L-H for $\zeta(s)$ may also be deduced.

For the cusp forms ϕ_j the $M_\infty(\lambda)$ conjecture is also related to the L-H but for another Euler product of degree 2. Let ϕ_j be as in (4.11) then in view of (3.6′), $L(s,\phi_j)$ is an Euler product of degree two, where

$$L(s,\phi_j) = \sum_{m=1}^{\infty} \frac{\eta_j(m)}{m^s} = \prod_p \left(1 - \eta_j(p)p^{-s} + p^{-2s}\right)^{-1}\ . \tag{4.20}$$

The Gamma Factor that goes with the functional equation is (see [BU])

$$\Gamma(\frac{s}{2} + \frac{ir_j}{2})\Gamma(\frac{s}{2} - \frac{ir_j}{2}) , \quad \text{where } \frac{1}{2} + r_j^2 = \lambda_j .$$

This time Conjecture (3.10) implies the L-H in the r_j aspect for $L(s, \phi_j)$, with $s = 1/2$. The analogue of (4.17) with E replaced by ϕ_j was recently derived in S. Katok and Sarnak [K-SK]: Let ϕ_j be an even Hecke-Maass cusp form for $SL_2(\mathbb{Z})$ then

$$|a_j(1)|^2 \phi_j(i) = 48\sqrt{2}\pi \rho_j(-4)\overline{\rho_j(1)} \tag{4.21}$$

where $a_j(1)$ is the first Fourier coefficient of $\phi_j(z)$, $\rho_j(m)$ is the j-th Fourier coefficient of the Shimura correspondent F_j of ϕ_j. F_j is a form of weight 1/2. In his thesis K. Khuri-Madkisi [MA] has shown that $|\rho_j(d)|^2$ is essentially equal to $L(\phi_j \otimes \chi_d, 1/2)$ (see Waldspurger [WA], Kohn-Zagier [K-Z], Shimura [SH] for the corresponding statement for holomorphic modular forms). Morever combining the results in Iwaniec [I1] and Hoffstein-Lockhart [HN-L] we have

$$|r_j|^{-\varepsilon} << \frac{|a_j(1)|^2}{\cosh(\pi r_j)} \underset{\varepsilon}{<<} |r_j|^{\varepsilon} . \tag{4.22}$$

The upshot is that the M_∞ conjecture for $z = i$ is equivalent to the L-H in the r_j aspect for $L(\phi_j, 1/2)$.

Conjecture 3.7 when restricted to $\Gamma = SL_2(\mathbb{Z})$ is also related to questions about L functions, but this time of degree 4. Associated to each ϕ_j as above is an Euler product of degree 4 known as the Rankin-Selberg L-function. It is defined by

$$L(\phi_j \otimes \phi_j, s) = \rho(2s) \sum_{m=1}^{\infty} \frac{|\eta_j(m)|^2}{m^s} . \tag{4.23}$$

Again (3.6′) may be used to show that $L(\phi_j \otimes \phi_j, s)$ is an Euler product of degree 4. It has an integral representation (see [BU])

$$\begin{aligned}\xi(s) &:= |a_j(1)|^2 \pi^{-s}\Gamma^2(s/2)\Gamma(\frac{s}{2} + ir)\Gamma(\frac{s}{2} - ir)L(s, \phi_j \otimes \phi_j) \\ &= 2\pi^{-s}\Gamma(s)\zeta(2s)\int_{\Gamma\backslash\mathbb{H}} |\phi_j(z)|^2 E(z,s)\frac{dx\,dy}{y^2} .\end{aligned} \tag{4.24}$$

In as much as $\pi^{-s}\Gamma(s)\zeta(2s)E(z,s)$ is analytic in $\mathbb{C}$ (except for simple poles at $s=0$ and $s=1$) and that it satisfies a functional equation $s \to 1-s$, we see that $\xi(s)$ satisfies the same properties. Thus $L(s,\phi_j \otimes \phi_j)$ is an Euler product of the type described at the beginning of this section. We should mention the issue of (4.5) being satisfied for $L(s,\phi_j \otimes \phi_j)$ or $L(s,\phi_j)$. This is the statement

$$\eta_j(n) = O_\varepsilon(n^\varepsilon) \tag{4.25}$$

and is known as the Ramanujan Conjecture [SA1]. The best estimate to date towards this widely believed conjecture, is due Bump-Duke-Hoffstein-Iwaniec [B-D-H-I] and reads $|\eta_j(p)| \leq 2p^{5/28}$ for p a prime. In any event the L-H in the r-aspect for $L(s,\phi_j \otimes \phi_j)$ together with (4.21), (4.23) and Stirling's series for $\Gamma(s)$ implies that for a fixed t

$$\int_{\Gamma\backslash\mathbb{H}} |\phi_j(z)|^2 E(z, \frac{1}{2}+it)\frac{dxdy}{y^2} = O_\varepsilon\big(\lambda_j^{-1/4+\varepsilon}\big)\ . \tag{4.26}$$

This in turn implies that for a fixed $h \in C_0^\infty(\Gamma\backslash\mathbb{H})$ which is in the space of Eisenstein series (i.e. in the span of $E(z,1/2+it), t \in \mathbb{R}$)

$$\int_{\Gamma\backslash\mathbb{H}} |\phi_j(z)|^2 h(z)\frac{dxdy}{y^2} = O_{\varepsilon,h}\big(\lambda_j^{-1/4+\varepsilon}\big)\ . \tag{4.27}$$

We expect that (4.27) holds equally well with the Eisenstein series $E(z,1/2+it)$ replaced by a cusp form $\phi_\ell(z)$– but in this case we don't have a direct relation to L-functions. However we can prove that (4.26) is valid in the mean with either the Eisenstein series or cusp forms. More precisely let $h \in C_0^\infty(\Gamma\backslash\mathbb{H})$ satisfying $\int_{\Gamma\backslash\mathbb{H}} h(z)dv(z) = 0$, then for $\varepsilon > 0$

$$\sum_{\lambda_j \leq \lambda} \left| \int_{\Gamma\backslash\mathbb{H}} h(z) d\mu_j(z)\right|^2 \underset{\varepsilon,h}{<<} \lambda^{1/2+\varepsilon}\ . \tag{4.28}$$

This should be compared with Theorem (1.2) (Quantum ergodicity) which is equivalent to the statement that for any smooth h of mean zero

$$\sum_{\lambda_j \leq \lambda} \left| \int_X h(x) d\mu_j(x)\right| = o(N(\lambda))\ . \tag{4.29}$$

For $X = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$, (4.29) was established by Zelditch [Z2]. He has recently [Z3] extended this as follows: If X is a compact manifold of (possibly variable) negative curvature then for $m \geq 1$ and h as above

$$\sum_{\lambda_j \leq \lambda} \Big| \int_X h(x) d\mu_j(x) \Big|^m = O\Big(\frac{N(\lambda)}{(\log \lambda)^{m/2}}\Big) .$$

The proof of (4.28) uses in a crucial way that the ϕ_j's are Hecke eigenforms and also estimates on Kloosterman sums via use of the Kuznietzov trace formula [KU]. In any event it is the above line of reasoning with L-functions and the average result (4.28) that leads one to Conjecture 3.7 and (3.7). Note that the individual equidistribution (against the continuous spectrum) would already follow from a bound of the form $L(\frac{1}{2} + it_o, \phi_j \otimes \phi_j) = O(|\lambda_j|^{1/4-\delta})$ for some $\delta > 0$. That is we need to beat convexity for this L-function in the r aspect.

The connections between the analytic theory of L-functions and the problems of the first three Sections show the importance of developing the analytic theory of general L-functions. For many purposes improving the convexity bound is enough. Needless to say a more geometric or dynamical approach to the problems described in the first three Sections could have a profound effect on the theory of L-functions.

Section 5. Outline of some proofs

In this Section we outline proofs of some of the theorems mentioned in Section 3. Details can be found in the papers cited. We begin with a proof of Proposition 3.1 which is a simple consequence of Duistermaat-Guillemin trace formula [D-G]. The technical condition that we are assuming is that the fixed point set for the geodesic flow at time T, which we denote by $Z_T \subset S^*(X)$, is a disjoint union $Z_{1,T} \sqcup Z_{2,T} \ldots \sqcup Z_{k,T}$ where $Z_{j,T}$ is a clean intersection [D-G]. Also we are assuming that (say) $\dim Z_{1,T} = 2$. According to the formula of Duistermaat and Guillemin if ϕ is equal to 1 in a neighborhood of T_0 and supported in a sufficiently small neighborhood of T_0 and smooth then

$$\sum_j \hat{\phi}(x - \sqrt{\lambda_j}) \sim c_0 x^{1/2} + c_1 + c_2 x^{-1/2} \ldots \text{ as } x \to \infty . \tag{5.1}$$

We will also assume that $c_0 \neq 0$, a property which can be checked for our given X. Let

$$N_1(t) := \sum_{\sqrt{\lambda_j} \leq t} 1 := \frac{\text{Vol}(X)}{4\pi} t^2 + S(t) . \tag{5.2}$$

Then we can write (5.1) as

$$\int_{-\infty}^{\infty} \hat{\phi}(x - t)\Big(\frac{\text{Vol(X)}}{2\pi} t dt + dS(t)\Big) \sim c_0 x^{1/2} . \tag{5.3}$$

Since $T_0 \neq 0$ it follows that

$$\int_{-\infty}^{\infty} \hat{\phi}(x - t) dS(t) \sim c_0 x^{1/2}$$

or

$$\int_{-\infty}^{\infty} \hat{\phi}'(x - t) S(t) dt \sim -c_0 x^{1/2} . \tag{5.4}$$

The following calculus lemma is easily established.

Lemma 5.1. *Let $S(t)$ be a locally integrable function on $\mathbb{R}$ satisfying $S(t) = O(|t|)$. Assume that for some Schwartz function ψ we have*

$$\Big| \int_{-\infty}^{\infty} S(t) \psi(x - t) dt \Big| >> x^{1/2}$$

then

$$\frac{1}{X}\int_X^{2X} |S(t)|dt >> X^{1/2} .$$

The condition that $S(t) = O(|t|)$ in our case is Weyl's Law (1.7). It follows that if $N(\lambda) = \sum_{\lambda_j \leq \lambda} 1$ and $E(\lambda) = N(\lambda) - \frac{\mathrm{Vol}(X)}{4\pi}\lambda$ then

$$\frac{1}{\lambda}\int_\lambda^{2\lambda} |E(u)|du >> \lambda^{1/4}$$

and

$$\frac{1}{\lambda}\int_\lambda^{2\lambda} |E(u)|^2 du = \sum{}^2(\lambda,\lambda) >> \lambda^{1/2} .$$

This proves Proposition 3.1 and the comments following it.

We turn next to Theorems 3.1 and 3.2. The starting point is the exact form of the trace formula for $X = \Gamma\backslash\mathbb{H}^2$, that is the Selberg Trace Formula [SEL1, H1]. It reads

$$\sum_j h(r_j) - \frac{\mathrm{Vol}(\Gamma\backslash\mathbb{H})}{4\pi}\int_{-\infty}^{\infty} h(r) r \tan h(\pi r) dr$$
$$= \sum_{\{\gamma_0\}}\sum_{k=1}^{\infty} \frac{\log N(\gamma_0) g(k\log N(\gamma_0))}{N(\gamma_0)^{k/2} - N(\gamma_0)^{-k/2}} \quad (5.5)$$

where $\lambda_j = \frac{1}{4} + r_j^2$, g is a smooth even function of compact support and h is its Fourier transform. The sum on the right hand side extends over all primitive hyperbolic conjugacy classes of Γ and $N(\gamma)$ is related to trace$(\gamma) = t(\gamma)$ by $N(\gamma)^{1/2} + N(\gamma)^{-1/2} = t(\gamma)$. Geometrically $\log N(\gamma)$ is the length of the closed geodesic on $\Gamma\backslash\mathbb{H}$ created by γ.

For $t > 0$ define $S(t)$ by

$$S(t) = \#\{0 \leq r_j \leq t\} - \frac{\mathrm{Vol}(\Gamma\backslash\mathbb{H})}{4\pi} t^2 . \quad (5.6)$$

Theorem 3.1 (and similarly Theorem 3.2) follows by a change of variables from

$$\frac{1}{TU}\int_U^{2U}\int_T^{2T} (S(t+u) - S(t))^2 dt du >> \frac{T}{(\log T)^2} \quad (5.7)$$

where $U \sim 1/\log T$.

Now the left hand side of the trace formula is, except for the finite number of terms with $r_j \notin \mathbb{R}$, essentially equal to

$$\int_0^\infty h(r)dS(r) \ . \tag{5.8}$$

For $u > 0$ we have

$$\begin{aligned}\int_0^\infty (S(v+u) - S(v))h(v)dv &= \int_0^\infty \int_v^{v+u} \frac{dS}{dx} dx h(v)dv \\ &= \int_{-\infty}^\infty \frac{dS}{dx} \int_x^{x+u} h(v)dvdx \ .\end{aligned}$$

In view of (5.8) the last can be expressed in terms of the right hand side of 5.5. In this way one can derive the following approximate formula, which is a variant of Proposition 18.11 in Hejhal [H1]: For $10 \le \beta \le 2\log t$, $0 < u \le t$, g a fixed test function,

$$\begin{aligned}&\int_{|t-v|\le t/2} e^{i\beta(v-t)} h(v-t)(S(v+u) - S(v))dv \\ &= -i \sum_{\substack{e^{\beta-1} \le N_\ell \le e^{\beta+1} \\ \ell \in \mathcal{L}(\Gamma)}} \frac{\mu(\ell) g(\beta - \log N_\ell)}{\sqrt{N_\ell}} \left(e^{it\log N_\ell} e^{i(t+u)\log N_\ell}\right) + O_\varepsilon(t^\varepsilon) \end{aligned} \tag{5.9}$$

where

$$\mu(\ell) = \sum_{\substack{\{\gamma\} \\ N(\gamma) = N_\ell}} 1 \ .$$

Note that (5.9) yields an approximate formula for a smoothed form of $S(t)$. From it we can deduce lower bounds for Σ^2. Without such strong smoothing we have no useful approximation by Dirichlet polynomials which is what makes upper bounds for S or Σ^2 so illusive. In the case that Γ is arithmetic we can use the B-C property (3.3) to study the mean modulus squared of the right hand side of (5.9) With this knowledge applied to the left hand side of (5.9) one gets the lower bounds in (5.7) on using Cauchy's inequality. To see how the B-C property is used first note that if the traces in Γ are well spaced (that is $|t(\gamma)-t(\gamma')| \geq \delta$ whenever $t(\gamma) \neq t(\gamma')$) as is the case if Γ is derived from a quaternion algebra, then we can appeal to the sharp form of the Hilbert inequality [M-V] to study the mean square. For the general arithmetic Γ the following Lemma from [L-S] is used:

Lemma 5.2. *Let $0 < t_1 < t_2 < t_3 \ldots$ satisfy $|\{t_j\} \cap [n^2, (n+1)^2]| \leq C$ for all $n \in \mathbb{N}$. Let $a_j \geq 0$, then there is $D > 0$ such that for $R \geq 1$*

$$\int_R^{2R} \Big| \sum_{t_j \leq N} a_j t_j^{-ir} \Big|^2 \Big(\frac{1}{2} - \frac{|r - \frac{3R}{2}|}{R}\Big) dr \geq DR \sum_{t_j \leq N} a_j^2 + O(\sqrt{N} \sum_{t_j \leq N} a_j^2) \,.$$

With this we get a lower bound for 5.7 from the diagonal sum in Lemma 5.2. Appealing a second time to the B-C property for arithmetic groups we have

$$\sum_{|t_0| \leq X} \mu(t_0) << X^{1/2} \Big(\sum_{|t_0| \leq X} \mu^2(t_0) \Big)^{1/2} \tag{5.10}$$

where t_0 runs through numbers which are the traces of primitive conjugacy classes and $\mu(t_0)$ is the number of distinct such classes with trace t_0. The left hand side is asymptotic to $X^2/\log X$ by the Prime Geodesic Theorem [SEL1, H1] and hence

$$\sum_{|t_0| \leq X} \mu(t_0)^2 >> X^3/(\log X)^2 \,. \tag{5.11}$$

One can show that the order of magnitude of the left hand side of (5.11) is in fact $X^3/(\log X)^2$. The lower bound (5.11) leads eventually by the above reasoning to the lower bound in Theorem 3.1. In the case that Γ is nonarithmetic we have no control of the off diagonal contributions in Lemma (5.2) (i.e. no B-C property) nor a sufficient understanding of the left hand side of (5.11) which presumably is of order $X/\log X$.

Next we outline the proof of Theorem 3.6. The idea is to exploit the nonlocal aspect of the Hecke correspondences (3.5). Let μ be a quantum limit for a congruence quotient $X = \Gamma\backslash\mathbb{H}^n$, $n = 2, 3$, that is μ is a limit of a subsequence of the

$|\phi_j(z)|^2 dV(z)$. Recall that we are assuming that ϕ_j is an eigenform of a suitable subalgebra of the Hecke algebra. Let $\Lambda = \text{sing supp}(\mu)$. The key is the following separation lemma.

Lemma 5.3. *If Λ is non empty and is contained in a finite union of geodesic subspaces as described in Theorem 3.6 then there is a correspondence C of X and $\hat{z} \in X \backslash \Lambda$ such that $C\hat{z} \cap \Lambda$ consists of exactly one point.*

The proof of this Lemma will not be given here. It may be found in Rudnick-Sarnak [R-S]. The proof is algebraic and uses the structure of the Hecke operators and the theory of representation of numbers by binary quadratic and Hermitian forms. To see how Lemma 5.3 is used to establish the Theorem, let T_C be the Hecke operator associated to the correspondence C. (The Lemma actually also ensures that T_C is in the suitable Hecke subalgebra). ϕ_j is an eigenfunction of T_C so we may write

$$\lambda_j(C)\phi_j(z) = \phi_j(z_1(z)) + \sum_{k=2}^{r} \phi_j(z_k(z)) \tag{5.12}$$

where C is degree r. We have separated out $z_1(z)$ so that $C\hat{z} \cap \Lambda = z_1(\hat{z}) := w$. The eigenvalues $\lambda_j(C)$ satisfy

$$|\lambda_j(C)| \leq r \ . \tag{5.13}$$

Let U be a small neighborhood of w in $\mathcal{F} :=$ union of the finite number of geodesic subspaces in which Λ lies. Let $B(U,\varepsilon) = \{\xi \mid d(\xi, U) < \varepsilon\}$. By Lemma 5.3 we can find U_1 a neighborhood of $\hat{z}$ and $\varepsilon_0 > 0$ such that for $0 < \varepsilon < \varepsilon_0$

$$\begin{gathered} B(U_1,\varepsilon) \cap \Lambda = \phi \\ z_j(B(U_1,\varepsilon)) \cap \Lambda = \phi \quad \text{for } j = 2, \dots, r \\ z_1(B(U_1,\varepsilon)) = B(U,\varepsilon) \ . \end{gathered} \tag{5.14}$$

From (5.12) and Cauchy's inequality

$$\int_{B(U_1,\varepsilon)} |\phi_j(z_1(z))|^2 dv(z) = \int_{B(U,\varepsilon)} |\phi_j(\xi)|^2 dv(\xi)$$
$$\leq r\Big(|\lambda_j(C)|^2 \int_{B(U_1,\varepsilon)} |\phi_j(\xi)|^2 dv(\xi) + \sum_{k=2}^{r} \int_{z_k(U_1,\varepsilon)} |\phi_j(\xi)|^2 dv(\xi)\Big) .$$

It follows from this and (5.13) on letting $j \to \infty$ that for all $\varepsilon < \varepsilon_0$

$$\mu(B(U,\varepsilon)) \leq r\big(r\mu(B(U_1,\varepsilon) + \mu(z_2(B(U_1,\varepsilon)) + \ldots \mu(z_r(B(U_1),\varepsilon))\big) .$$

If μ is absolutely continuous w.r.t. dv outside Λ then the r.h.s. of the last inequality tends to 0 as $\varepsilon \to 0$. That is we have $\mu(U) = 0$. This however contradicts the fact that $w \in \Lambda = \Lambda \cap \mathcal{F}$ and so Theorem 3.6 follows.

Finally we discuss the spectral inequalities that lead to Theorem 3.11. To explain the ideas we consider only the case $X = \Gamma\backslash\mathbb{H}^2$ with $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, pretending X is compact. So we ignore the continuous spectrum. We use Selbergs theory of point pair invariants [SEL1]. Let

$$\begin{aligned} k(t) &= -\frac{1}{\pi}\int_t^\infty \frac{dQ(w)}{\sqrt{w-t}} \\ g(u) &= Q(e^u + e^{-u} - 2) \\ h(r) &= \int_{-\infty}^{\infty} e^{iru} g(u)du \end{aligned} \tag{5.15}$$

where we view h as the source function, h should be even and holomorphic in $|\mathrm{Im}(r)| < 1/2 + \delta_1$, for some $\delta_1 > 0$ and also $|h(r)| << (1+|r|)^{-\delta}$, $\delta > 2$. We form the spectral expansion

$$\sum_{\gamma\in\Gamma} k(u(\gamma z, w)) = \sum_{j=0}^{\infty} h(r_j)\phi_j(z)\overline{\phi_j(w)} \tag{5.16}$$

where

$$u(z,w) = \frac{|z-w|^2}{\mathrm{Im}(z)\mathrm{Im}(w)}$$

and ϕ_j and r_j as in (5.5). A useful choice of h is as follows; let $T \geq 1$ and

$$h(r) = 4\pi^2 \frac{\cosh \frac{\pi r}{2} \cosh \frac{\pi T}{2}}{\cosh \pi r + \cosh \pi T} . \tag{5.17}$$

then

$$g(\xi) = \frac{2\pi \cos(\xi T)}{\cosh(\pi \xi)} . \tag{5.18}$$

For this h we have

(i) $\quad h(r) \geq 0 \quad$ for $\quad r \in \mathbb{R} \cup i\mathbb{R}$ (5.19)

(ii) $\quad h(r) \geq 1, T \leq r \leq T+1$.

A simple analysis concerning the behavior of $k(z,z)$ as $T \to \infty$ shows that the left hand side of (5.16) is $\nu(z)T + O(T^{1/2})$, where $\nu(z) = |\{\gamma \in \Gamma \mid \gamma z = z\}|$. Hence from (5.19) (i) and (ii) we have

$$\sum_{T \leq r_j \leq T+1} |\phi_j(z)|^2 << T . \tag{5.20}$$

This yields $\|\phi_j\|_\infty << \lambda_j^{1/4}$ which is the local bound 3.8. To go beyond this would require shortening the sum in (5.20). This however renders the geometric side to be no longer local and to have an exponential number of terms. This is very hard to deal with in general and remains a natural barrier. However in the case that we have corresondences and in particular the Hecke operators T_n as for $\mathrm{SL}_2(\mathbb{Z})$ we can break this barrier. Firstly apply T_n to both sides of 5.16 (in the variable z).

Assuming as we are that the ϕ_j's are Hecke eigenforms as in (4.11) and using 3.6 we get

$$\sum_{\gamma \in R(n)} k(\gamma z, w) = \sum_{j=0}^{\infty} h(r_j)\lambda_j(n)\phi_j(z)\overline{\phi_j(w)} . \tag{5.21}$$

Let

$$M(z,n,\delta) = \#\{\gamma \in R(n) \mid u(\gamma z, z) < \delta\} . \tag{5.22}$$

In this (and the more general case of quaternion algebras) one can estimate $M(z,n,\delta)$ from above according to the following Lemma – see Iwaniec-Sarnak [I-S]:

Lemma 5.4. *For* $\varepsilon > 0$

$$M(z,n,\delta) \underset{\varepsilon}{<<} (\delta + \delta^{1/4})n^{1+\varepsilon} + n^{\varepsilon} \ .$$

The main reason one is able to do this is that the conditions defining $R(n)$ and M are a quadratic equation with linear inequalities. A. Zaharescu has recently improved the bound above by replacing $\delta^{1/4}$ by $\delta^{1/2}$. With Lemma 5.4 and a careful analysis of $k(z,w)$ one shows [I-S]

Lemma 5.5.

$$\sum_{\gamma \in R(n)} k(\gamma z, z) \underset{\varepsilon}{<<} (T + nT^{1/2})n^{\varepsilon} \ .$$

With this the relation $\eta_j(n)\eta_j(m) = \sum_{d|(n,m)} \eta_j(\frac{nm}{d^2})$ (see (4.11) and (3.6′)) and (5.21) one deduces the following weighted inequality.

Theorem 5.6. *Let* $\varepsilon > 0$, N, $T > 1$. *For any* $\alpha_n \in \mathbb{C}$ *we have*

$$\sum_{T \leq r_j \leq T+1} |\phi_j(z)|^2 \Big| \sum_{n \leq N} \alpha_n \eta_j(n) \Big|^2 \underset{\varepsilon}{<<} N^{\varepsilon} T \sum_{n \leq N} |\alpha_n|^2 + N^{1+\varepsilon} T^{1/2} \sum_{n \leq N} |\alpha_n|^2 \ .$$

One can then choose the weights α_n to highlight a particular $|\phi_j(z)|^2$. The obvious choice of $\alpha_n = \overline{\eta_{j_o}(n)}$ has the difficulty that we don't have a lower bound for $\sum_{n \leq N} |\eta_j(n)|^2$. One has to be more careful in choosing the weights, see [I-S]. In any event this can be done and leads to Theorem 3.11 (i).

To prove (ii) of the same Theorem we again appeal to (5.21) but with a different choice of h. Let g, h, k be chosen with $g \geq 0$, $h \geq 0$, $k \geq 0$ and g of compact support in $[-1,1]$. It is easily seen that such a choice is possible. For $T \geq 1$ set

$h_T(r) = h(r/T)$ and let g_T, k_T be the corresponding transforms. The idea is to choose z in the relation

$$\sum_j h(r_j/T)\lambda_j(m)|\phi_j(z)|^2 = \sum_{\gamma \in R(m)} k_T(\gamma z, z)$$

to be a fixed point of many correspondences. Let $S(X) = \{z \in \Gamma\backslash\mathbb{H} \mid az^2 + bz + c = 0, (a,b,c) = 1, a,b,c \in \mathbb{Z}\}$ be the set of "C.M." points in X. The following Theorem is proven in [I-S].

Theorem 5.7.

(a) *Let $z \in S(X)$, $D = b^2 - 4ac$ and $P = \{\text{primes } p \mid (\frac{4D}{p}) = 1\}$. If n has its prime factors in P and is divisible by a certain fixed integer depending only on D then*

$$\sum_{r_j \le T} |\phi_j(z)|^2 |\lambda_j(n)|^2 >> T^2 \sum_{d|n} d\tau(\frac{n^2}{d^2})$$

where $\tau(n)$ is the number of divisors of n.

(b) *For any $n < 2T$ we have*

$$\sum_{r_j \le T} |\lambda_j(n)|^2 \underset{\varepsilon}{<<} (\sum_{d|n} d)T^2 + n^{3+\varepsilon}T \,.$$

The point is that the coefficient of T^2 in (a) can be substantially larger than the one in (b). This can be seen by choosing $n = \prod_{p<\underline{Y}, p\in \underline{P}} p$. Doing so leads to Theorem 3.11 (ii) and shows that the ϕ_j's are at least that large on the C.M. points.

Section 6. Conclusion

To summarize, both the theoretical and experimental discussions have centered around the quantization of the geodesic flow of an arithmetic hyperbolic manifold. For these, at least in dimension two, the large energy limit (or equivalently $\hbar \to 0$) the quantum eigenstates appear to behave like random waves with little structure beyond what is minimally necessary. There is no localization of the eigenstates and few if any of the chaotic features of the classical Hamiltonian appear at the quantum level, even in the semiclassical limit. Strong scarring onto periodic orbits does not occur and quantum unique ergodicity is expected. A weaker type of persistent enhancement of the eigenstates occurs at certain arithmetical points of the manifold. The fine analysis of the size and distribution of the eigenfunctions leads to some basic conjectures which are closely related to the classical Lindelof Hypothesis for the Riemann zeta function and which are natural eigenfunction analogues of the classical Ramanujan and Sato-Tate conjectures which concern eigenvalues. It is interesting that these conjectures as well as the random wave model do not apply universally to three dimensional arithmetic hyperbolic manifolds. The reason being that a certain infinite subset of the eigenstates which are "theta lifts" exhibit a singular behavior. The level spacing distribution for the eigenvalues of all arithmetic manifolds also exhibit a singular behavior. They do not follow the expected G.O.E. statistics and in dimension two appear to be following a Poisson like behavior, which is what is associated with integrable systems.

We expect that the basic eigenfunction behavior found for the arithmetic surface cases is typical of the quantization of the general chaotic system. The techniques that we have described in these lectures, which go beyond the local theory and which make some progress towards the basic conjectures mentioned above, are very special to arithmetic manifolds relying heavily on the correspondences carried by such a manifold. It is highly desirable to find methods which bring in the dynamics, rather than arithmetic, into the analysis. There is clearly still much to be done and learned.

Acknowledgements:

My debt to the researchers mentioned in Section 2 in connection with the numerical experiments is clear. I also would like to acknowledge the many discussions I have had with Iwaniec, Hejhal, Luo, Rudnick and Phillips. These have been instrumental in shaping the views that have been outlined in these lectures. Finally I thank the mathematics departments at the Universities of Tel Aviv and Toronto for the opportunity to lecture on this material.

References

[A-L] Atkin, A.O. and J. Lehner. *Math. Ann.* **185** (1970), 134–160.

[A-A] Arnold, V. and Avez. *Ergodic Problems of Classical Mechanics.* W.A. Benjamin, New York, 1968.

[A-A] Aurich, R., F. Steiner. "Statistical properties of highly excited quantum eigenstates of a strongly chaotic system." preprint, 1992.

[B-I] Bombieri, E. and H. Iwaniec. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **13** (1986) 449–472.

[B-L] Babich, V., V. Latzutkin. In Topics in Math Phys. Vol 2 (M.S. Birman Ed.) Consultants Bureau, New York, 1968, 9–18.

[BD] Berard, P. *Math. Zeit* **155** (1977), 249–276.

[B1] Berry, M. *J. Phys. A.* **A10** (1977), 2083–2091.

[B2] ______. *Proc. Royal Soc. London A.* **413** (1987), 183–198.

[B-T] Berry, M. and M. Tabor. *Proc. Royal Soc. Lond. A.* **356**, (1977), 375–394.

[B-G-G-S] Bogomolny, E.B., B. Georgeot, M. Giannoni, and C. Schmit. "Chaotic billiards generated by arithmetic groups." *Phys. Rev. Lett* **69** (1992), 1477–1480.

[B-S-S] Bolte J., G. Steil and F. Steiner. *Phys. Rev. Lett.* Vol. 69, No. 15, (1992), 2188–2191. *BAMS* Vol. 26, No. 2, (1992), 253–257.

[B-L-S] Burger, M., J. S. Li, and P. Sarnak. *BAMS*, Vol. 26, No. 2, (1992), 253–257.

[B-G] Bohigas, S.O., and M. Giannoni. "Chaotic motions and random matrix theories." *L.N. in Physics* **209** (1984), 1–99, Springer Verlag.

[BO] Bourgain, J. *Int. Math. Research Notices of Duke Math Jnl.* Vol. 3, (1993), 61–66.

[BL] Bleher, P. "Distribution of Energy Levels of Quantum Free Particles on a Surface of Revolution." preprint, 1993.

[BO] Borel, A. *Ann. Sci. Norm. Pisa Cl. Sci.* **8** (1981), 1–33.

[BU] Bump, D. Proc. Symp. "Number Theory, Trace Formulas Discrete Groups." Academic Press, Ed. K. Aubert, E. Bombieri, D. Goldfeld, 1981, 49–111.

[BUN] Bunimovich, L.A. *Func. Anal. and Appl.* **8** (19874), 254.

[CA] Cassels, J.W.S. *Rational Quadratic Forms.* Academic Press, 1978.

[CD1] Colin de Verdiere, Y. *Invent. Math.* **43** (1977), 15–52.

[CD2] ———. *Comm. Math. Phys.* **102** (1985), 497–502.

[CD3] ———. In *Chaos and Quantum Phys.* Le Houches 1989, Ed. M. Giannoni, A. Voros and I. Inn-Justin, 307–329.

[C-G] Conrey, B. and A. Gohsh. "On the Selberg Class of Dirichlet Series, Small Weights." preprint, 1992.

[D] Davenport, H. *Multiplicative Number Theory.* Grad. Text Math. Vol. 74, Springer, 1980.

[D-M] Dyson, F and M.L. Mehta. *J. Math. Phys.* **4** (1963), 701–712.

[D-G] Duistermaat, J.J. and V. Guillemin. *Invent. Math.* **29** (1975), 39–79.

[D-F-1] Duke, W., J. Friedlander, and H. Iwaniec. "Bounds for Automorphic L-Functions." I and II, 1993.

[E] Egorov, Y.V. *Uspenkski Mat. Nauk* **25** (1969), 235–236.

[EI] Eichler, M. "Lectures on modular correspondences." *Tata Inst.* **9** (1955).

[EIN] Einstein, A. *Verh, Dt. Phys. Ges.* **19**, (1917), 82–92.

[FR] Friedman, J. "On the Second Eigenvalue and Random Walks in d-Regular Graphs." 1998.

[G-L] Gerard, P. and E. Leichtnam. "Ergodic Properties of Eigenfunctions for the Dirichlet Problem." preprint, 1992.

[G-V-Z] Giannoni, M. A. Voroz, and J. Zinn-Justin editors. *Chaos and Quantum Physics.* Les Houches, 1991, Elsevier Publ.

[G-H-L-L-R-R-S-W] Gräf, H. et al. "Distribution of Eigenmodes in a Superconducting Stadium Billiard with Chaotic Dynamics." preprint, 1992.

[GU] Gutzwiller, M.C. "Quantum Chaos." *Scientific American* January 1992, 78–84.

[H1] Hejhal, D. "The Selberg Trace Formula," Vol. I, S.L.N. Vol. 548, 1976.

[H2] ______. "The Selberg Trace Formula" Vol. II, S.L.N. Vol. 1001, 1980.

[H3] ______. "Eigenvalues for the Laplacian for Hecke triangle groups." *Memoirs AMS 469* (1992).

[H-R] Hejhal, and D., Rackner. "On the topography of Maas wave forms." to appear in *Jn. of Exp. Math.* 1993.

[HEL1] Heller, E.J. *Phys. Rev. Letters* **53** (1984), 1515.

[HEL2] ______. In "Chaos and Quantum Physics." Les Houches, 1989, 549–661.

[H-L] Hardy, G., and E. Landau. "The lattice points in a circle." Collected works of Hardy Vol. II, Oxford, 1967. 329–356.

[H-PS] Howe, R. and I. Piatetsky-Shapiro. *Duke Math Jnl.* **50** (1983), 55–106.

[HO] Hörmander, L. "The spectrum of a positive elliptic operator." *Acta Math.* **121** (1968), 193–218.

[HN-L] Hoffstein, J. and Lockhart. "Coefficients of Maass forms and the Siegel zero." To appear in *Annals of Math.* 1993.

[I1] Iwaniec, H. "Small eigenvalues for $\Gamma_0(N)$." *Acta Arith.* **LVI** (1990), 65–82.

[I2] ______. "The spectral growth of automorphic L-functions." *J. Reine Angew. Math.* **428** (1992), 139–159.

[I-S] Iwaniec, H., and P. Sarnak. "L^∞ norms of eigenfunctions of arithmetic surfaces." preprint, 1993.

[JU] Judge, C. Thesis, Univ. of Maryland, 1993.

[K] Katok, S. *Fuchsian Groups.* University of Chicago Press, 1992.

[K-SK] Katok, S. and P. Sarnak. "Heegner points, cycles and Maass forms." To appear in the *Israel Math. Jnl.* 1993.

[K-R] Keller, J. and S.I. Rubinow. *Annals of Physics* **9** (1960), 24–75.

[K-Z] Kohnen, N.W. and D. Zagier. *Math. Ann.* **271** (1985), 237–268.

[K] Kurokawa. N. *Invent. Math.* **49** (1978), 149–165.

[KU] Kuznietzov, N. *Mat. Sb. (N.S.).* **39** (1981), 299–342.

[LA] Langlands, R. "Euler Products." Yale Math Monographs.

[LU] Luo, W.J. "On the nonvanishing of Rankin Selberg L-functions." *Duke Math. Jnl.*, Vol. 69, No. 2, 411–427, (1993).

[L-S] Luo, W.J. and P. Sarnak. "Number variance for arithmetic surfaces." preprint, 1993.

[L-L] Landau, L.C. and E.M. Lifschitz. "Quantum Mechanics." Vol. 3, Pergammon Press, 1988.

[L-P-S] Lubotzky, A. R. Phillips and P. Sarnak. *Combinatorica* **8** (1988), 261–277.

[M1] Margulis. G. *Usp. Math. Nauk.* **29** (1974), 49–98.

[M2] Margulis, G. "Explicit Group Theoretic Constructions...." *Jnl. of Problems of Information Transmission*, (1988).

[M-K] McDonald, S.W. and N. Kaufman. *Phys. Rev. A.* **37**, (1988), 3067.

[MA] Khuri-Makdisi, K. Thesis, Princeton University, 1993.

[ME] Mehta, M.L. *Random matrices and the statistical theory of energy* *levels.* Acad. Press., 2nd edition, 1990.

[M] Montgomery, H. *B.A.M.S.* **84** (1978), 547–567.

[PA] Pais, A. *Niels Bohr's Times.* Clarendon Press, Oxford, 1991.

[P] Piatetsky-Shapiro, TI. Work in progress on *Converse Theorems.*

[PS-1] Phillips, R. and P. Sarnak. *Invent. Math.* **80** (1985), 339–364.

[PS-2] ______. *Jn. Am. Math. Soc.* Vol 5, No. 1, (1992), 1–32.

[PS-3] ______. *Jnl. D'Analyse Math.* Vol. 59 (1992), 179–187.

[PS-4] ______. "Cusp forms in character varieties." preprint, 1992.

[R] Ralston, J. *Jnl. of Diff. Geom.* **12** (1977), 87–100.

[RAY] Lord Rayleigh. *Phil. Mag.* [6] 20, 1001 (1910).

[R-S] Rudnick, Z. and P. Sarnak. "The behavior of eigenstates of arithmetic hyperbolic manifolds." preprint, 1993.

[SA1] Sarnak, P. "Some applications of modular forms." Cambridge Univ. Press, **99**, 1990.

[SA2] ______. *A.M.S. Cont. Math.***53** (1986).

[SA3] ______. In Proc. Conf. "Analytic Number Theory and Diophantine Problem." Birkhauser, 1987, 321–333.

[S-Z] Salem, R. and A. Zygmund. *Acta Math.* **91** (1954), 245–301.

[SCH] Schmit, C. in Ref [G-V-Z], pp. 333-369.

[SER] Serre, J.P. *A course in arithmetic.* Grad. Texts in Math. **7**, Springer Verlag, 1970.

[SEL1] Selberg, A. *Collected Works*, Vol. 1, Springer Verlag, 1989, pp. 423–464.

[SEL2] ______. *ibid.*, 626–675.

[SEL3] ______. *Collected Works Vol. II*, Springer 1991, 47–65.

[SI1] Sinai, Y. *Introduction to Ergodic Theory.* Princeton Univ. Press, Vol. 18.

[SI2] ______. "Mathematical problems in the theory of quantum chaos." preprint, 1991.

[SHN] Shnirelman, A.I. *Uspenski Math. Nauk.* **29/6**, 1974, 181–182.

[SH] Shimura, G. "On Fourier Coefficients of Hilbert Modular Forms of Half Integral Weight." to appear in *Duke Math. Jnl.*, 1993.

[SHI] Shintani, T. *Nagoya Math. Jnl.* **58** (1975), 83–126.

[S-S] Seger, A. and C. Sogge, *Indiana J. of Math.* **38** (1989), 669–682.

[SL] Steil, G. "Uber die eigenwerte des Laplace operators und de Hecke operatoren fur SL(2, $\mathbb{Z}$)." preprint, 1993.

[ST] Stein, E. In Lectures in Harmonic Analysis, Princeton Press, 1986, 307–356.

[T] Titchmarsh, E. *The theory of the zeta function.* Oxford Univ. Press, 1967.

[TA] Takeuchi, K. *J. Math. Soc. Japan* **29**, No. 4, (1975), 600–612.

[U-Z] Uribe, A. and S. Zelditch. "Spectral statistics on Zoll surfaces." preprint, 1992.

[WA] Waldsburger. J.L. *J. Math. Pures. Appl.* **60** (1981), 375–484.

[W] Wolpert, S. "Dissapearence of cusp forms in families." To appear in *The Annals of Math.*, 1993.

[Z] Zelditch, S., *Duke Math. Jnl.* **55** (1987), 919–941.

[Z2] ______. *Memoirs of A.M.S.* Vol. 96, No. 495, 1992.

[Z3] ______. "On the rate of Quantum Ergodicity." preprint, 1993.

[ZY] Zygmund, A. *Studia Math. T.L.*, (1974), 189–201.