On the Topography of Maass Waveforms for PSL(2, Z)

Dennis A. Hejhal and Barry N. Rackner

CONTENTS

Hejhal was supported in part by NSF Grant DMS 89-10744 and by the Minnesota Supercomputer Institute (with CPU time on the Cray-2 and Cray-XMP).

This paper is an expanded version of Hejhal's lecture at the Workshop on Discrete Groups, Number Theory and Ergodic Theory held at MSRI in November 1991.

This article provides a glimpse into "arithmetical quantum chaos" through a study of the topography and statistical properties of the eigenfunctions of the Laplacian for the modular surface $\mathrm{PSL}(2,\mathbf{Z})\backslash H$.

1. INTRODUCTION

The eigenfunctions of the Laplacian are central objects of study in the harmonic analysis of manifolds. In particular, they figure prominently in the Selberg trace formalism [Hejhal 1976, 1983; Selberg 1956]. For curiosity's sake alone, it would be interesting to try to produce pictures of such eigenfunctions for representative examples, particularly as the eigenvalue, and presumably the complexity, increase.

This problem was first actively considered by theoretical physicists working with quantum chaos, the principal motivation there having been to seek manifestations of quantum chaos in individual eigenstates of classically chaotic (or ergodic) systems.

Compared to the multitude of papers on eigenvalue statistics, works devoted to *eigenfunctions* are still rather sparse. For a sampling, see [Berry 1977, 1989; Bogomolny 1988; Heller 1984; Heller et al. 1989; McDonald and Kaufman 1988; Ozorio de Almeida 1988, pp. 210–213, 217–220; Gutzwiller 1990, Ch. 15].

Of particular interest to us is the speculation by Berry [1977] that the eigenfunctions Ψ_n of a classically ergodic system should tend to exhibit Gaussian random behavior as the wavenumber tends to infinity. Compare [Longuet-Higgins 1957a,b, 1962].

1058-6458/92 $0.50 per page

Certain aspects of this problem are treated in [McDonald and Kaufman 1988; Shapiro and Goelman 1984; Shapiro et al. 1988] for the case of a stadium domain in $\mathbf{R}^2$, where the ambient geometry is Euclidean.

The purpose of the present paper is to report on a similar series of experiments, but for a surface of constant negative curvature.

To a lesser extent, we are also interested in commenting on the analog of the "ridges" or "scars" discussed in [Berry 1989; Bogomolny 1988; Heller 1984; Heller et al. 1989; Ozorio de Almeida 1988]. Readers unfamiliar with this topic will find a brief description of it near the end of Section 3. The results we obtain will serve to amplify and extend an earlier series of experiments by Aurich and Steiner [1991].

Throughout our discussion, it is important to keep in mind that one of the oldest examples of ergodicity is given by the geodesic flow associated with a Fuchsian group $\Gamma \subseteq \mathrm{PSL}(2,\mathbf{R})$ whose quotient space $\Gamma\backslash H$ has finite hyperbolic area [Hedlund 1937, 1939; Hopf 1937, pp. 29–30, 69–80; Sinai 1977, pp. 74–80].

(Here H is the Poincaré upper half-plane. Recall that H has the metric $ds = y^{-1}\,|dz|$ of constant negative curvature, and that $\mathrm{PSL}(2,\mathbf{R})$ acts on H by isometries, the action of $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PSL}(2,\mathbf{R})$ being given by

$$z \mapsto \frac{az+b}{cz+d}, \quad \text{for } z \in \mathbf{C}.$$

For our purposes, $\Gamma \subset \mathrm{PSL}(2,\mathbf{R})$ is *Fuchsian* when it is discrete, non-cyclic, and finitely generated.)

In this context, the quantal eigenstates are nothing other than Γ-invariant eigenfunctions of the hyperbolic Laplacian $\Delta u = y^2(u_{xx}+u_{yy})$ [Gutzwiller 1990, p. 358]. We impose a boundary condition at infinity by requiring that u be square-integrable over $\Gamma\backslash H$.

One of the main tools now available in the study of these eigenfunctions is the Selberg trace formalism [Hejhal 1976, 1983; Selberg 1956]. In order to apply the computational techniques developed in [Hejhal 1991, 1992b; Hejhal and Arno 1992], we restrict ourselves to groups having at least one cusp (conjugacy class of parabolic elements). This means that the quotient $\Gamma\backslash H$ of constant negative curvature is compact *except* for a finite number of punctures (always assuming that $\Gamma\backslash H$ has finite area).

Far and away the most important example of this kind is the *modular group*

$$\mathrm{PSL}(2,\mathbf{Z}) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PSL}(2,\mathbf{R}) : a,b,c,d \in \mathbf{Z} \right\}.$$

Figure 1 shows the standard fundamental polygon $\mathcal{F}$ for $\mathrm{PSL}(2,\mathbf{Z})\backslash H$; observe that there is just one cusp, $i\infty$.

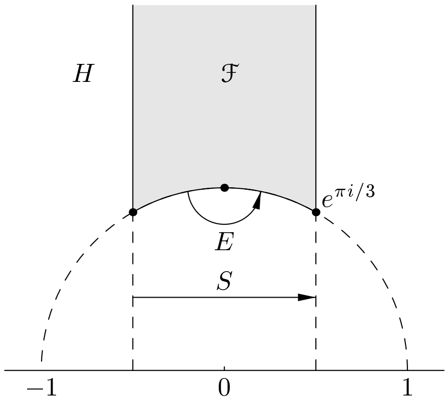


FIGURE 1. Fundamental domain $\mathcal{F}$ for the modular group $\mathrm{PSL}(2,\mathbf{Z})$, corresponding to the generators $S : z \mapsto z+1$ and $E : z \mapsto -1/z$.

The presence of punctures means [Hejhal 1983] that there is a continuous spectrum as well as a discrete one. From this standpoint, it would have been preferable for $\Gamma\backslash H$ to have no cusps, in which case the spectrum would be purely discrete. See [Hejhal 1976, p. 303; Takeuchi 1977a,b; 1983] for some natural examples of this type of Γ.

The importance of $\mathrm{PSL}(2,\mathbf{Z})$ is beyond dispute, however. Moveover, the situation with $\mathrm{PSL}(2,\mathbf{Z})$ and its congruence subgroups is almost as nice as in the case of no punctures. In fact, the continuous spectrum is here completely characterized, and well-controlled, by means of Epstein zeta functions, which are very familiar in analytic number theory. See [Hejhal 1983, Ch. 6 and 11; Maass 1949; Petersson 1982, pp. 286–294; Siegel 1977] for full details.

When $\Gamma = \mathrm{PSL}(2,\mathbf{Z})$ or one of its congruence subgroups, it is customary to refer to the quantal eigenstates as *Maass waveforms*.

2. ADDITIONAL BACKGROUND

To make our subsequent discussion more intelligible, we say a few words about the general format of a Maass waveform on $\mathrm{PSL}(2,\mathbf{Z})$.

Suppose for a moment that Γ is *any* Fuchsian group with just one cusp, which we place at $i\infty$ and then take to have width 1 as in Figure 1. Let $\mathcal{F}$ again denote the fundamental polygon. We assume, unless we say otherwise, that the hyperbolic area $\mu(\mathcal{F})$ of $\mathcal{F}$ is finite.

Exactly as in [Hejhal 1983, pp. 22–26], one finds that any square-integrable eigenfunction of Δ has a simple expansion in terms of the modified Bessel function $K_\nu(u)$. Specifically, if Ψ satisfies

$$\Delta\Psi + \lambda\Psi = 0$$

and we set $R = \sqrt{\lambda - \frac{1}{4}}$, we then have

$$\Psi(x+iy) = by^{\frac{1}{2}-\sqrt{\frac{1}{4}-\lambda}} + \sum_{n\neq 0} c_n y^{1/2} K_{iR}(2\pi\,|n|\,y) e^{2\pi i n x}$$

with $b = 0$ unless $0 \le \lambda < \frac{1}{4}$. Since we're interested in large λ, the case $0 \le \lambda < \frac{1}{4}$ is irrelevant, so we just have

$$\Psi(x+iy) = \sum_{n\neq 0} c_n y^{1/2} K_{iR}(2\pi\,|n|\,y) e^{2\pi i n x}, \quad (2.1)$$

and $R \ge 0$. At the same time,

$$\Psi(Tz) = \Psi(z) \quad \text{for all } T \in \Gamma. \quad (2.2)$$

The function Ψ can, of course, be taken to be real-valued.

The asymptotics of $K_{iR}(u)$ with respect to u (see (2.3) or [Hejhal 1983, p. 22]) immediately show that $\Psi(x+iy) = O(e^{-2\pi y})$ for large y. By virtue of [Hejhal 1983, p. 585 (middle)], we also have

$$|c_n| = O(|n|^{1/2}).$$

The method used in [Hejhal 1991, 1992b; Hejhal and Arno 1992] for determining R basically consists of forcing (2.2) to hold at sufficiently many $z \in \mathcal{F}$, for T ranging over a set of generators of Γ (say $T \in \{E, S\}$ in Figure 1).

In practical terms, this procedure really only involves the first few c_n. This stems from the exponential decay in $K_{iR}(2\pi\,|n|\,y)$. To be more specific, for $u > R$ we have [Erdélyi et al. 1953, vol. 2, pp. 87–88]

$$e^{\pi R/2} K_{iR}(u) \approx \frac{\sqrt{\pi/2}}{\sqrt[4]{u^2 - R^2}} \exp\Bigl(-Rg\Bigl(\frac{u}{R}\Bigr)\Bigr), \quad (2.3)$$

where

$$g(t) = \sqrt{t^2-1} + \arcsin(t^{-1}) - \pi/2 \quad \text{for } t > 1.$$

For $u < R$ the behavior is

$$e^{\pi R/2} K_{iR}(u) \approx \frac{\sqrt{2\pi}}{\sqrt[4]{R^2 - u^2}} \sin\Bigl(\frac{\pi}{4} + Rh\Bigl(\frac{u}{R}\Bigr)\Bigr), \quad (2.4)$$

where

$$h(t) = \operatorname{argcosh}(t^{-1}) - \sqrt{1-t^2} \quad \text{for } t < 1.$$

In both cases, it is understood that the right-hand side "cuts off" at a value like constant $\times R^{-1/3}$ once $|u - R|$ drops significantly below $R^{1/3}$. (Cf. [Balogh 1967].)

Note now that $g'(t) = \sqrt{t^2-1}/t$, so that

$$g(t) \simeq \frac{2\sqrt{2}}{3}(t-1)^{3/2}$$

near $t = 1$. As a consequence, one easily sees that $e^{\pi R/2} K_{iR}(2\pi\,|n|\,y)$ is already less than 10^{-15} when

$$|n| > \frac{R + 12R^{1/3}}{2\pi y} \quad (2.5)$$

and $R > 100$, say.

In imposing condition (2.2), we use values of z well away from $i\infty$. The ordinates of the associated Tz are therefore bounded away from both 0 and ∞ (since T was one of only a finite number of generators). We are also free to premultiply K_{iR} by $e^{\pi R/2}$ in (2.1). The process of determining R to a modest number of decimal places therefore hinges only on those c_n's with $1 \le |n| \le M$, with M as in the right-hand side of (2.5).

Obtaining good graphics of Ψ will generally require more c_n's than that. This creates problems (algorithmically). Here the type of Γ begins to matter. We distinguish two cases, depending on whether or not Γ is an arithmetic group. An example of Γ arithmetic is $\mathrm{PSL}(2, \mathbf{Z})$; examples of Γ nonarithmetic are the Hecke triangle groups with $N \neq 3, 4, 6$. (Noncongruence subgroups of $\mathrm{PSL}(2, \mathbf{Z})$ are best lumped together with the nonarithmetic case; compare [Maass 1983, pp. 66, 68, 72, 108].)

Typically, for Γ arithmetic, there are available certain number-theoretical symmetries, known as Hecke operators. Iterative techniques, like those in [Hejhal and Arno 1992], can then be employed to obtain large numbers of c_n with high accuracy.

In the nonarithmetic case, things are less clear, and we are still working on developing good methods for calculating additional c_n. This work will be reported on in a future paper.

There is, of course, another major difficulty associated with the nonarithmetic case. Namely, in line with the Sarnak–Phillips philosophy [Phillips and Sarnak 1985], it is unlikely that more than a finite number of Ψ can ever exist, unless Γ admits some type of algebraic symmetry (or group-theoretic inclusion).

In Hecke triangle groups, for instance, there is an obvious symmetry with respect to the y-axis. This means that our eigenfunction Ψ can be rewritten as

$$\Psi = \sum_{n=1}^{\infty} c_n y^{1/2} K_{iR}(2\pi n y) \left\{ \begin{array}{c} \cos(2\pi n x) \\ - - - - \\ \sin(2\pi n x) \end{array} \right\}, \qquad (2.6)$$

depending on whether Ψ is even or odd. In the cosine case, for nonarithmetic Γ, one expects *no* such Ψ to exist. In the sine case, however, things remain purely discrete and one gets good Weyl asymptotics, much as in the case of no cusps [Hejhal 1992b; Venkov 1983, §§ 6.5 and 6.7; Hejhal 1983, pp. 91–108].

To the extent that "discrete" Ψ do exist in either the arithmetic or the nonarithmetic case, determining the finer asymptotic properties of their Fourier coefficients c_n represents a very important problem. An analogous point is made in [Balazs and Voros 1986, pp. 168, 193].

For the modular group $\mathrm{PSL}(2, \mathbf{Z})$, Hecke operators exist and show that, in representation (2.6), one can take

$$c_n c_m = \sum_{d \mid (n,m)} c_{nm/d^2} \quad \text{and} \quad c_1 = 1 \qquad (2.7)$$

without loss of generality. Relation (2.7) is equivalent to asserting that

$$\sum_{n=1}^{\infty} \frac{c_n}{n^s} = \prod_p \frac{1}{1 - c_p p^{-s} + p^{-2s}},$$

where p runs over the primes and $\mathrm{Re}(s) > \frac{3}{2}$. As far as the results of [Hejhal and Arno 1992] go, there is strong support for both the (generalized) Ramanujan–Petersson conjecture $|c_p| \le 2$ and the Sato–Tate conjecture

$$\lim_{X \to \infty} \frac{N[\,p \le X : c_p \in E]}{N[p \le X]} = \frac{1}{2\pi} \int_E \sqrt{4 - \xi^2}\, d\xi,$$

where E is a Jordan measurable set. (Take $\xi = 2\cos\theta$ to get the usual form of Sato–Tate!) This equality is a variant of the more familiar Wigner semicircle law describing the distribution of eigenvalues of random Hermitian matrices [Mehta 1967].

It is clear from (2.7) that, for $\mathrm{PSL}(2, \mathbf{Z})$, the c_n are *not* statistically independent. There are dependencies corresponding to the arithmetic "symmetries" of $\mathrm{PSL}(2, \mathbf{Z})\backslash H$.

Although, for p prime, the coefficients c_p would appear to be statistically independent, it must be borne in mind that, even in studying fairly simple correlation functions of the c_n (not to mention the c_p), one immediately encounters major open problems lying at the frontier of modern number theory: see, for instance, [Bump 1989, pp. 54–59, 62–66; Gelbart and Shahidi 1988, pp. 2(L), 65–67, 84–85, 94–97, 113; Linnik 1963, Ch. 3; Moreno and Shahidi 1985; Selberg 1965, 1991].

Still, number-theoretical techniques do provide the only way presently known of attaining any kind of rigorous probabilistic control on the c_n, either for $\mathrm{PSL}(2, \mathbf{Z})$ or for nonarithmetic Γ. Although the latter case is less common in the literature, note in particular the results of A. Good [1981, 1983] concerning the Rankin–Selberg method. Compare [Selberg 1965, §§ 2 and 4] and [Hejhal 1983, Appendix E].

The whole situation is further complicated by the fact that one ultimately needs to let $R \to \infty$.

Given this state of affairs, it makes sense to resort to some exploratory experiments.

3. SOME KNOWN RESULTS

Prior to discussing the experiments, we need to draw attention to several additional facts.

Let $\mathcal{M}$ denote for now any C^∞ Riemannian manifold of dimension 2. Let Δ be the usual Laplace–Beltrami operator.

For $\mathcal{M}$ compact, it is known that, generically, the nodal lines of the Δ-eigenfunctions don't cross and the eigenvalue multiplicities are all 1 [Uhlenbeck 1976; Courant and Hilbert 1953, p. 395 (7)]. Uhlenbeck's result is presumably still valid when the metric is hyperbolic, even if $\mathcal{M}$ is allowed to have a finite number of cusps. (In the latter case, it is understood that we are referring only to the discrete spectrum.)

For $\mathcal{M}$ compact, it is also known that, within any geodesic ball of radius $c/\sqrt{\lambda}$ (where c is a universal constant), any eigenfunction Ψ necessarily has a change of sign [Courant and Hilbert 1953, pp. 451–452; Brüning 1978, p. 18].

A small amount of further work shows that the same result holds not only for any hyperbolic $\mathcal{M}$ of finite area but also for any real-valued (hyperbolic) eigenfunction *on* H. The simplest proof is to pass to mean values as in [Hejhal 1983, p. 570 (7.6)] and then quote [Hobson 1931, § 237]. Compare [Courant and Hilbert 1953, p. 455; Courant and Hilbert 1962, p. 289]. Also see [Donnelly and Fefferman 1988, 1990].

When $\mathcal{M}$ is hyperbolic of finite area, classical ergodicity holds, and the result about balls of radius $c/\sqrt{\lambda}$ strongly suggests that the nodal lines tend to become increasingly complicated (chaotic) as the $\lambda \to \infty$. See [Gutzwiller 1990, p. 237] or [McDonald and Kaufman 1979, Fig. 1] for a related Euclidean example.

Ergodicity also plays a decisive role in the work of Shnirelman [1974], Colin de Verdiere [1985] and Zelditch [1987]. In discussing this work, we assume from the outset that $\mathcal{M} = \Gamma\backslash H$ is a hyperbolic manifold. As usual, we denote by $\mathcal{F}$ some fundamental domain of $\mathcal{M}$, and by μ the hyperbolic area on H.

Suppose first that $\mathcal{M}$ is compact. The orthonormal basis $\{\varphi_n\}_{n=0}^{\infty}$ formed from the eigenfunctions of Δ is then controlled, at least in part, by asymptotic estimates associated with the names of Weyl, Minakshisundaram and Pleijel [Hörmander 1968; Weyl 1950, § 5]. In particular, one knows that

$$N[\lambda_n \le X] \sim \frac{\mu(\mathcal{F})}{4\pi} X,$$

$$\sum_{\lambda_n \le X} \varphi_n(P)^2 \sim \frac{1}{4\pi} X,$$

where $P \in \mathcal{M}$.

The presence of ergodicity allows one to go further. Specifically, the formalism of [Shnirelman 1974; Colin de Verdiere 1985; Zelditch 1987] will now ensure that, after the possible exclusion of a set of λ_n of density 0, we have

$$\lim_{n\to\infty} \frac{1}{\mu(A)} \int_A \varphi_n(z)^2 d\mu(z) = \frac{1}{\mu(\mathcal{F})} \qquad (3.1)$$

for *every* Jordan region A in $\mathcal{F}$. (Density 0 simply means that the number of bad λ_n up to height X is at most $o(X)$.) We stress that the rate of convergence in (3.1) may depend strongly on A, particularly as the area of A shrinks.

Relation (3.1) is customarily regarded as a kind of equidistribution statement. It clearly implies, for instance, that, for nonexceptional λ_n, the mass of $\varphi_n(z)$ can never localize to, say, just a finite number of closed geodesics on $\Gamma\backslash H$.

Zelditch [1991] proved that (3.1) continues to hold for $\Gamma = \mathrm{PSL}(2, \mathbf{Z})$ and its congruence subgroups.

The presence of an exceptional set is clearly a bit troubling. So long as one is present, the door remains wide open to a variety of unusual behaviors. We mention this principally because of the "scarring effect" that has been observed on stadium domains (in $\mathbf{R}^2$). What happens there is that, for numerous n, the topography of φ_n is found to contain clear "ridges of mass", or "scars", situated roughly along what would appear to be closed geodesics. The location of these scars changes with n. See [Heller 1984; Heller et al. 1989; Gutzwiller 1990, p. 251].

Heuristic explanations for these ridges have been provided by Bogomolny [1988], Berry [1989] and Ozorio de Almeida [1988, pp. 210–213, 217–220] on the basis of semiclassical asymptotic expansions of what is now commonly referred to as the pre-trace formula [Gutzwiller 1990, pp. 188, 261 (top), 291–295, 297 and 206 (middle)]. The results in those three papers thus refer mainly to "packets" like

$$\sum_{|R_n - X| \le \delta} \varphi_n(P)^2,$$

as opposed to individual eigenstates.

The resulting expressions contain contributions from *every* periodic orbit. The magnitude of these contributions is essentially a function of $X \pm \delta$, the length of the orbit and the relative position of P [Berry 1989, Eqs. (38), (41), (45)]. To ensure convergence of the overall sum, one must take δ about $O(1)$ in size. This, needless to say, compares unfavorably with the mean R_n-gap of c/X. Even in the best of cases, it is doubtful that δ can ever be taken significantly less than $1/\sqrt{X}$. There are thus major problems in drawing conclusions about individual eigenstates. Compare [Berry 1989, p. 229(x); Delande 1991, pp. 284–285].

Further difficulties arise from the fact that, in handling the pre-trace formula, only semiclassical *approximations* were used. For now at least, supplying precise error estimates, even in the case of hyperbolic $\mathcal{M}$, seems largely out of the question, particularly if $\delta \to 0$.

The upshot of all this is that the "scarring phenomenon" is not yet set in stone. In particular, based on our experience [Hejhal 1976, pp. 131–139, 280–315], it seems entirely possible that the semi-classical (ripple) effects described in [Berry 1989; Bogomolny 1988; Heller 1984; Ozorio de Almeida 1988] may ultimately blend into the rigorous error term for the pre-trace formula and no longer be explicitly discernible (along geodesics) for large X. Compare [Berry 1991, § 4.1] and [Selberg 1991, Theorems 1 and 2].

Rudnick and Sarnak [1992] have begun to address some of these more subtle questions (such as accurate error terms) for surfaces of the form $\Gamma\backslash H$, and have found, for instance, that on congruence subgroups of $\mathrm{PSL}(2, \mathbf{Z})$, *no finite collection of closed geodesics can ever serve as the limiting support of a subsequence of φ_n with $n \to \infty$.* In fact, a stronger statement holds: no subsequence of $\varphi_n^2 \, d\mu$ can ever converge to a measure having *singular* support restricted to a finite number of closed geodesics. The proof uses Hecke operators.

The same paper also includes a valuable discussion of [Bogomolny 1988, Eqs. (2), (9), (10)] from the standpoint of the classical Selberg trace formalism.

4. PICTURES OF WAVEFORMS

Our experiments were carried out in two stages. In the first stage, we made a variety of plots depicting the topography of about a dozen Maass waveforms on $\mathrm{PSL}(2, \mathbf{Z})\backslash H$. The second stage was devoted to statistical analyses (see Section 5). In all cases, we took the values of R from [Hejhal 1991], with slight improvements in accuracy. Table 1 summarizes the waveforms investigated.

As mentioned earlier, the Fourier coefficients c_k are readily computed to high accuracy, using the techniques of [Hejhal and Arno 1992].

Convention. From this point on, it is understood that all our K-Bessel functions are premultiplied by $e^{\pi R/2}$.

Figures 2–5 use color to depict representative waveforms. See also Figure 9, and the cover of this issue. For comparison, Figure 10 shows a "mock waveform" obtained by summing the cosine branch of (2.6) with $R = 500$, $c_1 = 1$, and c_n chosen randomly, with uniform distribution, in the interval $(-1, 1)$. (Note that these random coefficients do not satisfy (2.7).) Mock waveforms will be discussed more extensively in Sections 5.2 and 6.

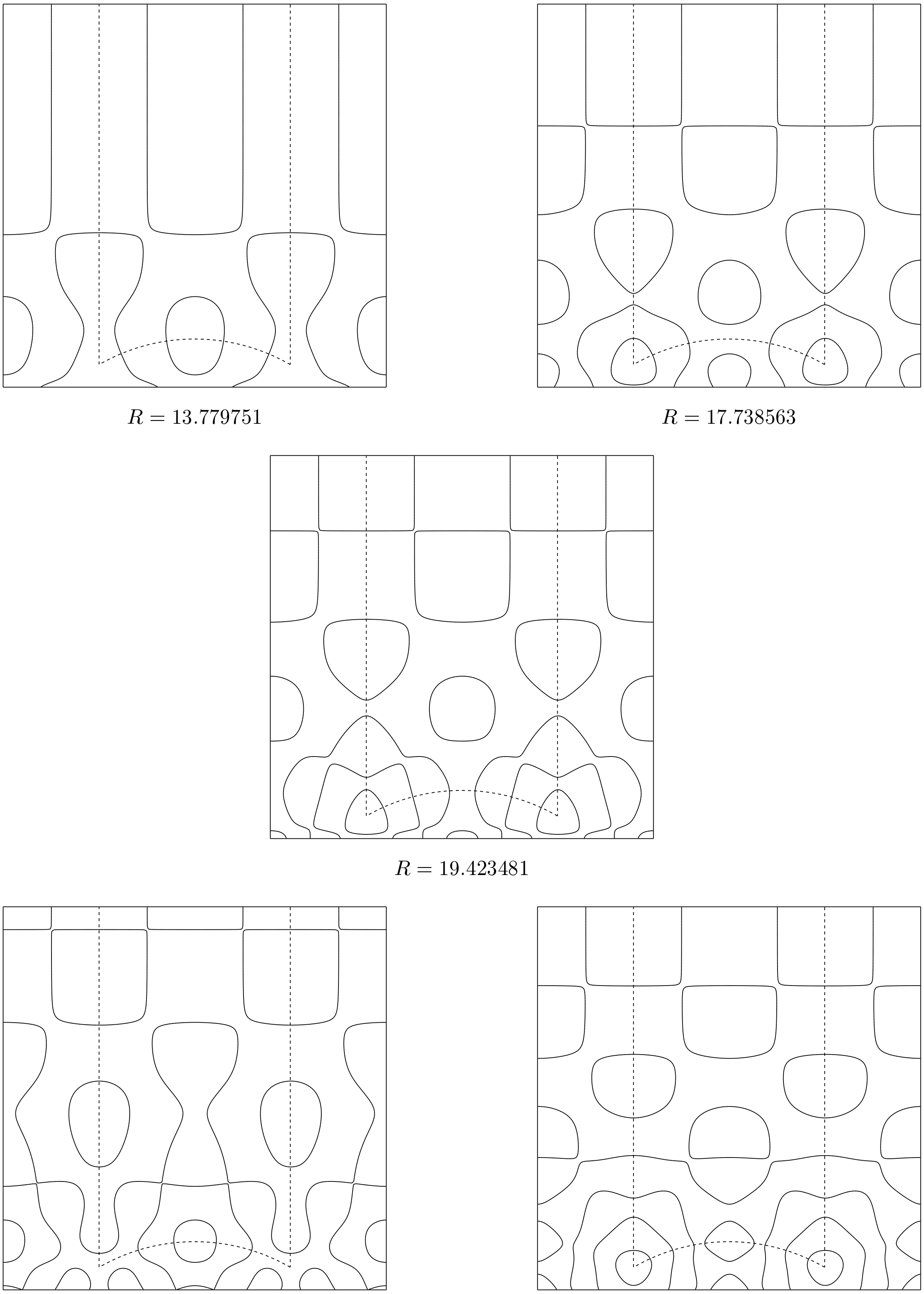
R	parity	M_0	Figures
13.779751	even	6.06	6 (top left)
17.738563	even	7.09	6 (top right)
19.423481	even	7.52	6 (middle)
21.315796	even	7.99	6 (bottom left)
22.785908	even	8.36	6 (bottom right)
9.533695	odd	4.87	7 (top left)
12.173008	odd	5.62	7 (top right)
14.358510	odd	6.21	7 (middle)
16.138073	odd	6.68	7 (bottom left)
16.644259	odd	6.81	7 (bottom right)
47.926558	even	14.15	4, 8
125.313840	even	30.39	2, 8, cover
125.347558	even	30.39	3, 8
125.523988	even	30.43	8
500.066454	even	103.57	9
500.283548	even	103.61	5

TABLE 1. Summary of the waveforms investigated. The number $M_0 = (R + 8R^{1/3})/(\pi\sqrt{3})$ is the approximate threshold for five-place accuracy at the point $z = e^{\pi i/3}$: see (2.5) and the discussion leading to it.

Figures 6–8 show the nodal lines of the waveforms Ψ corresponding to the first 14 values of R in Table 1. (Nodal lines are simply the curves where $\Psi = 0$.) We omit the plots for $R \approx 500$, which are similar to those for $R \approx 125$, only four times finer. When Ψ is even (Figures 6 and 8), no nodal lines cross; the dashed lines indicate the boundary of the fundamental domain $\mathcal{F}$. When Ψ is odd (Figure 7), the crossings are all consistent with the Dirichlet boundary condition. Figures 6 and 7 show good agreement with some earlier, rougher plots made by Huntebrinker [1991], using finite element methods. Huntebrinker also computed the first few waveforms for several congruence subgroups of $\mathrm{PSL}(2, \mathbf{Z})$.

In studying these graphics, we can make several comments more or less immediately.

(Continued on p. 287)



$R = 13.779751$ $R = 17.738563$

$R = 19.423481$

$R = 21.315796$ $R = 22.785908$

FIGURE 6. Nodal lines of even waveforms Ψ for small R. The illustrated region is $[-1, 1] \times [.75, 2.75]$. The dashed lines indicate the boundary of $\mathcal{F}$.

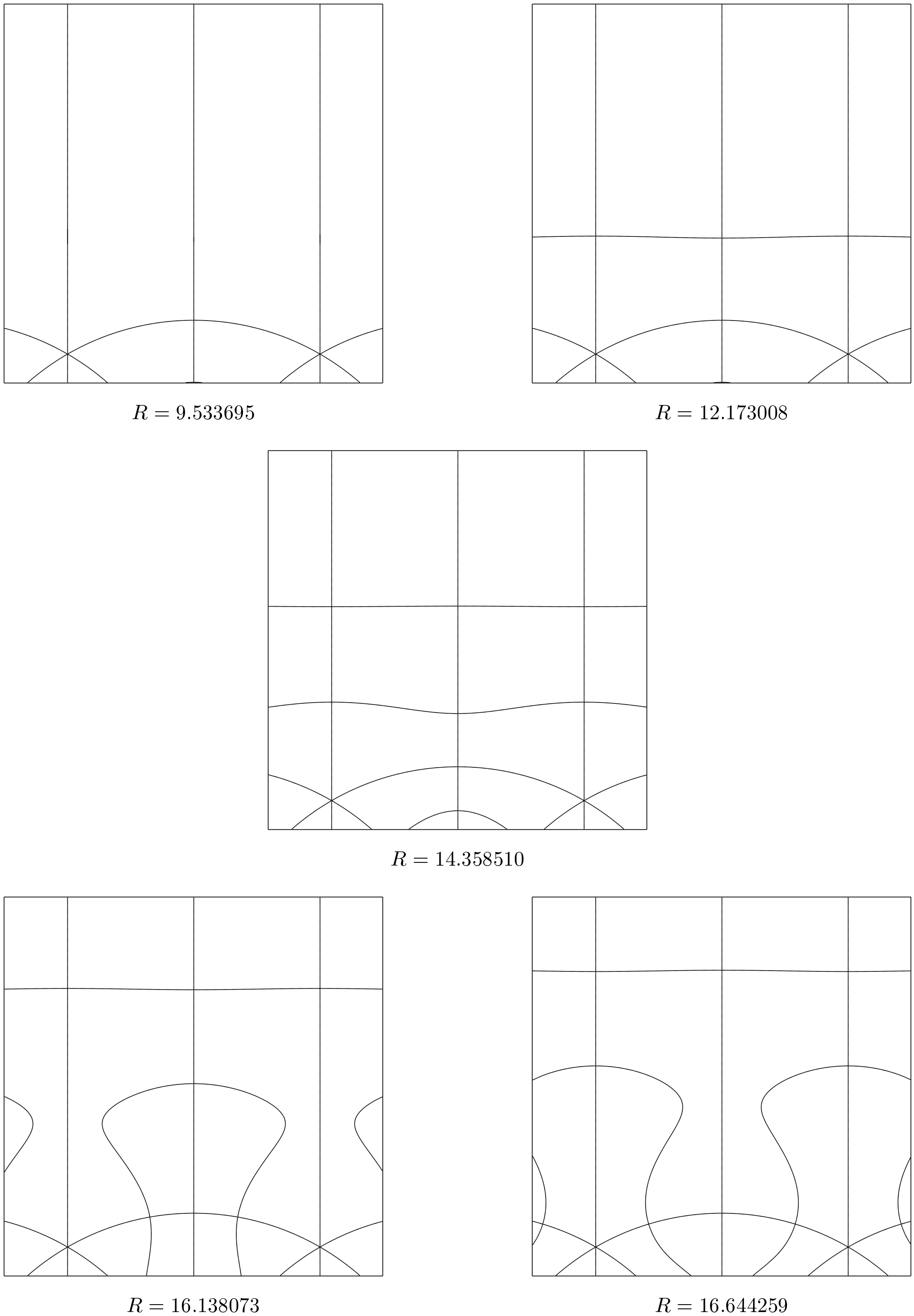
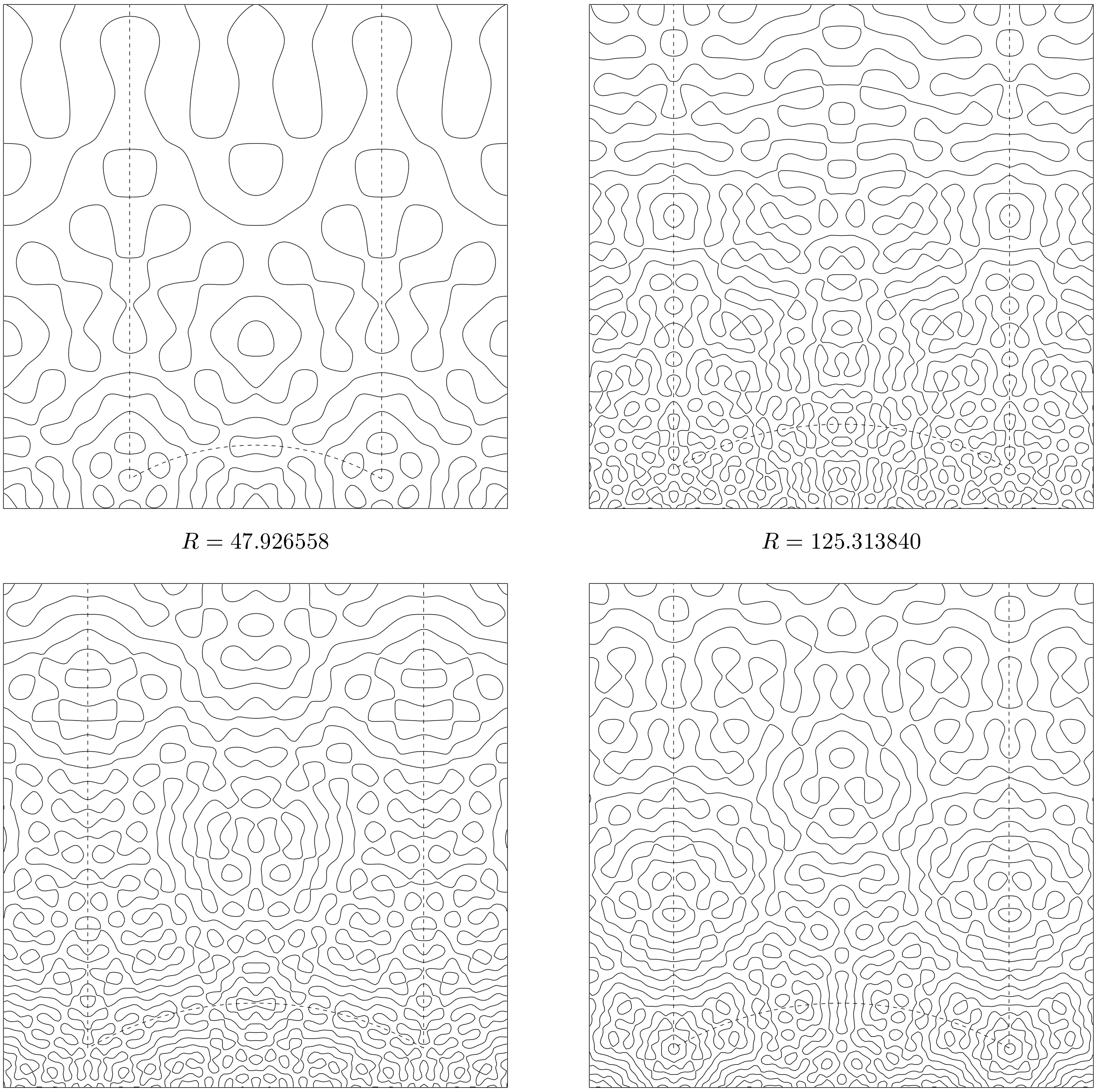


FIGURE 7. Nodal lines of odd waveforms Ψ for small R. The illustrated region is $[-.75, .75] \times [.75, 2.25]$.



$R = 47.926558$ $R = 125.313840$

$R = 125.347558$ $R = 125.523988$

FIGURE 8. More nodal lines of even waveforms. The illustrated region is $[-1, 1] \times [.75, 2.75]$ for $R = 47.926558$ and $[-.75, .75] \times [.75, 2.25]$ for the remaining graphs.

- Although ridges are clearly visible as soon as R is moderately high (Figures 2, 3, 5 and 9), they do not seem to lie along closed geodesics. (We recall, incidentally, that $\mathrm{PSL}(2, \mathbf{Z})\backslash H$ has no periodic orbits passing through $i\infty$.)
- Texturally, Figures 5, 9 and 10 are very similar, even though Figure 10 is random.
- In Figures 2, 5 and 9, there are roughly circular scars surrounding the elliptic fixpoint at $e^{\pi i/3}$. By contrast, Figure 3 does not show this phenomenon.
- As shown in [Hejhal 1992b, p. 93], the function Ψ will typically have a positive local maximum or a negative local minimum at $e^{\pi i/3}$. A look at the numerics shows that in general this is not a global maximum or minimum (see Figure 3). The point $z = i$ is also a critical point, but its type appears to be variable.

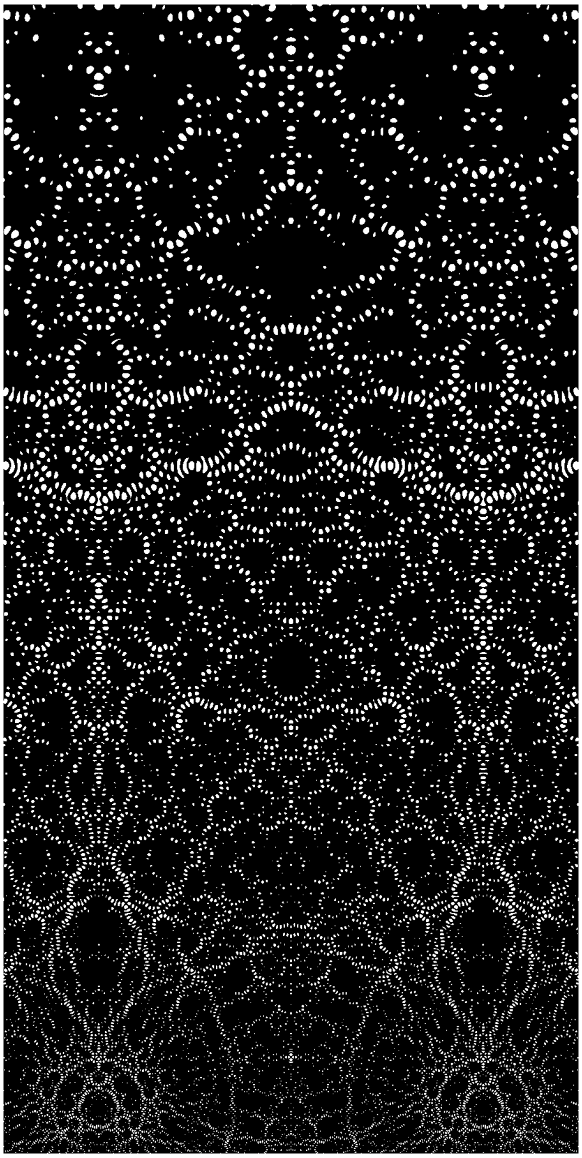


FIGURE 9. Waveform for $R = 500.066454$, in the region $[-.75, .75] \times [.75, 3.75]$. The maximum is 2.300, at $(.430, 1.391)$, and the minimum is -2.558, at $(.259, 2.539)$. Black represents values of Ψ in the middle third of the interval $[-\max|\Psi|, \max|\Psi|]$.

- Figures 5, 9 and 10 are very reminiscent of Figures 1a and 7 in [Heller et al. 1989]. Note, however, that our superpositions consist solely of waves coming in from $i\infty$. See also [Longuet-Higgins 1957a,b, 1962].
- The geometric patterns formed by the alternating "hills" and "holes", best visible in the figures with the black-white-red color scheme, may have some significance; see [Heller et al. 1989]. It is also interesting that, at least in certain cases, these hills and holes have oblong shapes,

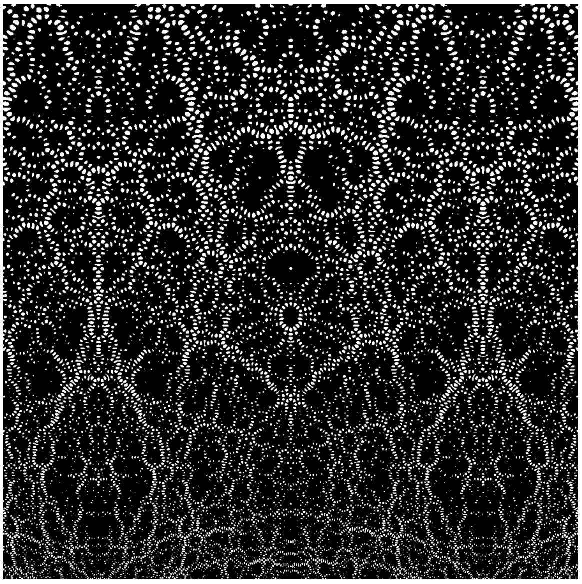


FIGURE 10. Mock waveform (p. 280) for $R = 500$, in the region $[-.75, .75] \times [.75, 2.25]$. The maximum is 1.514, at $(\frac{1}{2}, 1.726)$, and the minimum is -1.654, at $(0, 1.793)$. The color coding is as in Figure 9.

in rough agreement with [Berry 1989, p. 228 (iii)] and [Bogomolny 1988, p. 174, Eq. (15) and p. 176, l. 14]. (In these references $\hbar$ is the analog of $c/\sqrt{\lambda}$ in Section 3.)
- Figure 4 nicely illustrates how the successive K-Bessel functions "kick in" as y decreases, giving rise to increasing levels of "chaos" along the way. One can also see, for instance, that the first two K-Bessels must have zeros close to $y = 2.79$ and 3.18; likewise for $y \approx 2.14$ and $n \leq 3$.

Finally, some technical remarks. In making these plots, it is essential to use a sufficiently fine grid, consistent with $c/\sqrt{\lambda}$ and the color graduation. This point is easily addressed by retaining $\|\text{grad}\,\Psi\|$ as a "control value" in the machine output. To this end, we initially used a 5000×10000 grid for the rectangular regions (Figures 2–5, 9), and a 5000×5000 one for the square regions (Figures 5 and 10). Data preparation for Figure 10, say, took about 7 minutes of CPU time on the Cray-XMP. To optimize the subsequent color separation, we then switched over to 900×1800 and 900×1800 grids, after confirming that this entailed no significant loss of graphical accuracy.

Figures 2–5 and 9–10 were then produced on a Silicon Graphics 4D 310VGX in tiff format, using tools from the Utah Raster Toolkit and using the PBMplus Toolkit. The computer graphics

environment created by Joel Neisen at the Minnesota Supercomputer Center allowed us to produce these images with the greatest of ease. We also acknowledge the expert advice of Wes Barris of MSC and of Silvio Levy of the Geometry Center (University of Minnesota), Editor of *Experimental Mathematics*, whose many suggestions greatly improved our initial pictures.

5. STATISTICAL MATTERS

5.1. Introduction

Recall from the end of Section 3 that (ignoring a relatively small set of exceptional eigenvalues) we have an equidistribution relation (3.1) that holds for all Jordan regions A in $\mathcal{F}$. The most natural way of explaining why this should be so would be for the probability measures

$$\nu_{nA}(E) \equiv \frac{\mu\{z \in A : \varphi_n(z) \in E\}}{\mu(A)} \tag{5.1}$$

to converge nicely to some probability distribution G, independent of A and having mean 0 and standard deviation $\mu(\mathcal{F})^{-1/2}$ [Billingsley 1986, pp. 344 (ii), 348 and 408 (top)]. The optimal situation, especially from the standpoint of chaos [Moran 1968, p. 243], would clearly be for G to be Gaussian.

In (3.1), the φ_n were orthonormal. Basing the normalization on (2.6) and (2.7), as we do, leads only to the insertion of a modest scaling factor before $\mu(\mathcal{F})^{-1/2}$, in the expected value of the standard deviation. More specifically, as in [Iwaniec 1984, § 5; Iwaniec 1990, §§ 2–3; Kuznecov 1981, Theorem 6; Smith 1981], one has

$$\varphi_n(z) = B_n \Psi_n(z), \tag{5.2}$$

where $B_n \equiv 2\rho_n(1)e^{-\pi R_n/2}$ satisfies

$$R_n^{-\varepsilon} \ll |B_n| \ll \sqrt{R_n}\,, \tag{5.3}$$

$$\sum_{R_n \le X} |B_n|^2 \sim \frac{2X^2}{\pi^2}. \tag{5.4}$$

Here $\rho_n(j)$ is the obvious Fourier coefficient of φ_n, and $e^{-\pi R_n/2}$ reflects our K-Bessel convention. It was recently shown that $\sqrt{R_n}$ can be replaced by R_n^{ε} in (5.3) [Hoffstein and Lockhart 1992].

From a practical standpoint, the main drawback to the current form of (3.1) is the lack of effective bounds on its error terms. In a certain sense, the existence of a possible exceptional set is but one manifestation of this. Moreover, as already mentioned, the rate of convergence in (3.1) or (5.1) may well depend strongly on A.

To elaborate on this a bit, recall that, in the physics literature, $c/\sqrt{\lambda_n}$ is commonly referred to as the de Broglie wavelength. At length scales below $c/\sqrt{\lambda_n}$, one expects the topography of Ψ_n to look "essentially sinusoidal", that is, regular. It is only when A is substantially bigger than the de Broglie wavelength that one stands any chance of seeing any type of Gaussian distribution.

(A similar situation holds for the logarithms of various number-theoretical L-functions along the line $\mathrm{Re}(s) = \frac{1}{2}$; compare [Bombieri and Hejhal 1987, § 3; Hejhal 1989; Selberg 1991, § 2].)

This graininess basically implies that one should not expect too much uniformity in $\nu_{nA}(E)$ under slight variations in A (and, to a lesser extent, E) for relatively modest R.

Be these things as they may, there are three questions about which one would very much like to gather further, even if only very sketchy, information:

Question 1. Is G, in fact, Gaussian?

Question 2. Are the restrictions of Ψ_n to well-separated subregions of $\mathcal{F}_+ \equiv \mathcal{F} \cap \{\mathrm{Re}(z) > 0\}$ in any sense statistically independent as $n \to \infty$?

Question 3. Do the functions Ψ_n and Ψ_{n-q} tend to become statistically independent on arbitrary $A \subseteq \mathcal{F}_+$ for $n \to \infty$ and $q \ge 1$?

Compare [McDonald and Kaufman 1988; Shapiro and Goelman 1984; Shapiro et al. 1988].

It is to these questions that we addressed the second stage of our experiments.

If Berry's conjecture (page 275) is correct, Question 1 should have a positive answer. Since (3.1) is already known for PSL(2, **Z**), the key issue is simply whether the measures ν_{nA} actually look like Gaussians with mean 0. (The standard deviations should, by all rights, take care of themselves, at least if the exceptional set is empty, as seems to be likely.)

Some caution needs to be exercised in dealing with all three questions, because of the special format implicit in (2.6) and (2.7). The philosophy of [Berry 1977] and [Longuet-Higgins 1957a,b, 1962] is predicated on the "wave-vectors" being able to come in from all directions. Compare Figure 4 (left) and the asymptotics of (2.4) for $u \ll R$, where $u = 2\pi ny$. Or, equivalently, see (6.3).

In any event, note that if the measures ν_{nA} go Gaussian with any kind of uniformity in A (at scales bigger than the de Broglie wavelength), a simple argument based on conditional probabilities will immediately yield the plausibility of a positive answer to Question 3 whenever $(n-q)/n \to 0$. The worst case is when $q = o(n)$. (The granularity of any test sets is understood here to be appropriately large compared to $c/\sqrt{\lambda_{n-q}}$.)

5.2. Description of the Statistical Experiments

Table 2 and Figures 11–12 summarize our explorations in trying to answer Question 1. We calculated (2.6) for the waveforms Ψ corresponding to the last five values of R in Table 1, over a variety of rectangular windows, as listed in the top part of Table 2. Note that some of the windows are not contained in $\mathcal{F}_+$; we chose them in this way in order to allow for a wider class of tests.

We also made histograms by throwing hyperbolic areas into thirty buckets, according to the size of the local Ψ-average over a 5000×5000 or

waveform	window	$E(\Psi)$	SD	grid size	M	range	Figure
$R = 125.313840$	$[0,.2] \times [1,1.2]$	.010	.439	5000×5000	26	$[-1.44, 1.41]$	11 (top left)
	$[.2,.4] \times [1,1.2]$	$-.024$	.500				
	$[0,.2] \times [\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}+.2]$	$-.005$	.426				
	$[.3,.5] \times [\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}+.2]$	.003	.488				
	$[-.75,.75] \times [.75,3.75]$	.000	.435	5000×10000	35	$[-1.68, 1.47]$	11 (top right)
$R = 125.347558$	$[0,.2] \times [1,1.2]$	.004	.322				
	$[.2,.4] \times [1,1.2]$	.005	.371				
	$[0,.2] \times [\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}+.2]$	$-.002$	.324				
	$[.3,.5] \times [\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}+.2]$	.003	.345				
	$[-.75,.75] \times [.75,3.75]$	.000	.389				
$R = 125.523988$	$[0,.2] \times [1,1.2]$	$-.009$	.567				
	$[.2,.4] \times [1,1.2]$	$-.009$	.592				
	$[0,.2] \times [\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}+.2]$	.010	.727				
$R = 500.066454$	$[0,.2] \times [1,1.2]$	$-.001$	.429	5000×5000	90	$[-1.97, 1.65]$	11 (middle left)
	$[.2,.4] \times [1,1.2]$	.000	.451				
	$[0,.2] \times [\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}+.2]$	$-.002$	.521				
	$[-.75,.75] \times [.75,3.75]$	.000	.490	5000×10000	120	$[-2.56, 2.30]$	11 (middle right)
$R = 500.283548$	$[0,.2] \times [1,1.2]$	$-.003$	.521	5000×5000	90	$[-1.93, 1.97]$	11 (bottom left)
	$[.2,.4] \times [1,1.2]$	.004	.582				
	$[0,.2] \times [\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}+.2]$	.000	.611				
	$[-.75,.75] \times [.75,3.75]$	.000	.571	5000×10000	120	$[-2.82, 2.40]$	11 (bottom right)
mock (a), $R \approx 500$	$[0,.2] \times [1,1.2]$	$-.002$	.564				
mock (b), $R = 500$	$[0,.2] \times [1,1.2]$	$-.002$	.339	5000×5000	90	$[-1.21, 1.30]$	12 (top)
	$[.2,.4] \times [1,1.2]$	.001	.355				
	$[-.75,.75] \times [.75,2.25]$	.000	.359	5000×5000	120	$[-1.65, 1.51]$	12 (middle)
mock (c), $R = 1000$	$[0,.2] \times [1,1.2]$	.000	.349				
	$[.2,.4] \times [1,1.2]$	.000	.334				
mock (d), $R = 1000$	$[0,.2] \times [1,1.2]$	.000	.339	5000×5000	172	$[-1.54, 1.50]$	12 (bottom)
	$[.2,.4] \times [1,1.2]$	.000	.364				

TABLE 2. Statistics of individual waveforms and mock waveforms Ψ (see page 291) corresponding to diverse values of R and rectangular windows A. For all experiments we show the mean and the standard deviation SD of Ψ. We also show, in those cases for which a histogram is included in Figures 11 and 12, the size of the grid used, the approximate value of $M = (R + 8R^{1/3})/(2\pi y_{\min})$, as in Table 1, and the range of values of Ψ occurring within the window.

5000×10000 grid. Some of these histograms are shown in Figure 11, and the corresponding data are given in the long rows of Table 2.

The bottom part of Table 2 and Figure 12 report the same experiments for several *mock* waveforms:

(a) $R = 500.283548$, with the coefficients of $R = 13.779751$;
(b) $R = 500$, with $c_1 = 1$ and uniform random $c_n \in (-1, 1)$;
(c) $R = 1000$, with $c_1 = 1$ and uniform random $c_n \in (-1, 1)$;
(d) as in (c), but with a different batch of coefficients.

We estimate the error level in Table 2 and subsequent ones to be no more than a few thousandths, given the controls we exercised on $\|\mathrm{grad}\,\Psi\|$ when passing to Riemann sums.

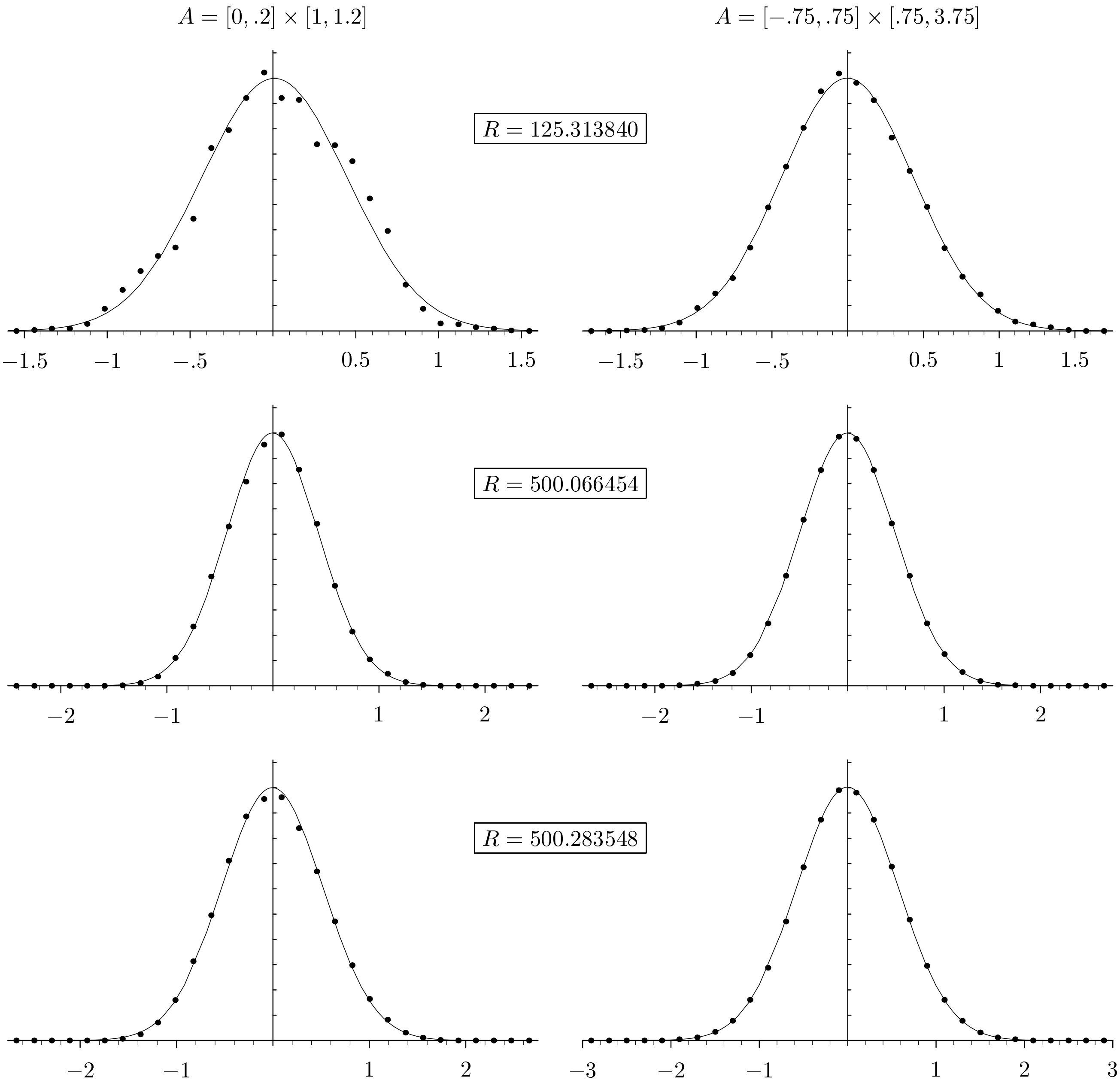


FIGURE 11. Histograms of the value distribution of automorphic waveforms Ψ in the given window A. See Table 2 (top) for the statistics.

We obtained very good support throughout for the conjecture that the distribution G is Gaussian. The bigger the R, the better the fit. The textures for automorphic waveforms with $R \approx 500$ are entirely comparable to those for the mock waveforms.

Note, however, that the standard deviations for automorphic Ψ are having a bit of trouble stabilizing to something independent of the window. This is mildly disturbing but not wholly unexpected, given our earlier discussion. See the end of Section 6, items (a)–(d), for more on this point.

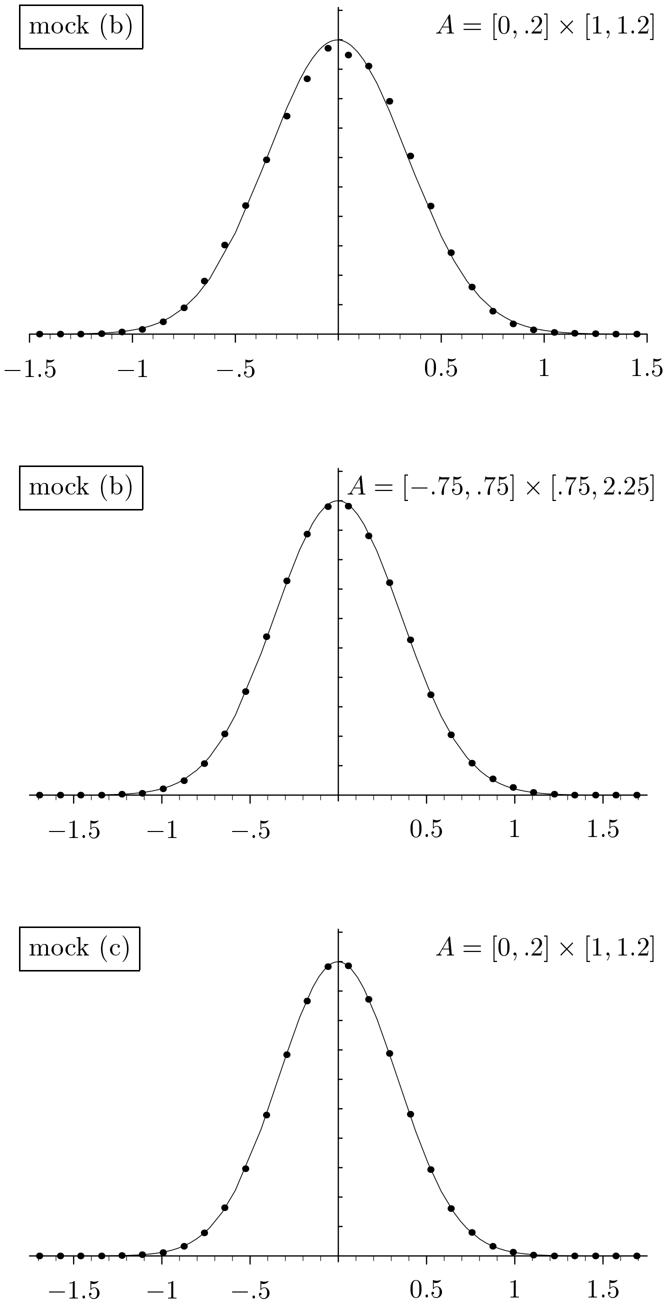


FIGURE 12. Histograms of the value distribution of mock waveforms Ψ. See Table 2 (bottom).

To round things out, we ran similar tests on a number of vertical and horizontal cross-sections of the rectangles used before. Representative results are shown in Table 3 and Figures 13 and 14. Here we found nothing nearly as striking as in the earlier figures. In all cases tested, both vertical and horizontal, the means and standard deviations tended to exhibit relatively high levels of fluctuation. Any evidence for a *one*-dimensional analog of (3.1), and corresponding Gaussian limit, must therefore be regarded as sketchy at best.

We next turned our attention to Questions 2 and 3, which have to do with statistical independence. Here, rather than make detailed comparisons of joint probability distributions, it seemed much easier to compute a variety of correlation coefficients; see [Billingsley 1986, p. 417 (7)] and the theorem in [Feller 1971, vol. 2, p. 136].

We calculated such coefficients for many pairs of cases in Table 2. To keep things simple, we restricted ourselves to cases where the two windows were isometric. (Proper attention was also paid to the necessary disjointness of the windows on $\mathrm{PSL}(2, \mathbf{Z})\backslash H$.)

Tables 4–6 are representative of the results we obtained. Table 4 shows self-correlation coefficients (Question 2), while the other two involve comparisons between different waveforms (Question 3).

In Table 5, the correlations tend to be biggest when the two values of R are close together. Similar behavior was seen in every other case tested, but tended to diminish as R grew: see Table 6.

These tables certainly support the statistical independence properties formulated in Questions 2 and 3. *All in all, then, our experiments, as far as they go, tend to confirm the basic thrust of Berry's hypothesis, at least for* $\Gamma = \mathrm{PSL}(2, \mathbf{Z})$. The obvious expectation, of course, is that Questions 1–3 will continue have an affirmative answer for any quotient $\Gamma\backslash H$ of finite area, especially if compact.

Lots of additional experiments are possible, but, on the whole, a natural picture of quantum chaos for waveforms on surfaces of constant negative curvature seems to be emerging. The properties stated in Questions 1–3 form its centerpiece.

waveform	segment	$E(\Psi)$	SD	range	Figure
$R = 500.066454$	$[-.75, .75] \times \{.8208\}$	.023	.535	$[-1.53, 1.70]$	13 (top)
	$[-.75, .75] \times \{1.5003\}$	.049	.498	$[-1.09, 1.22]$	similar to 13 (top)
mock (b)	$[-.75, .75] \times \{.8208\}$	.026	.381	$[-0.97, 1.09]$	13 (middle)
	$[-.75, .75] \times \{1.5003\}$	.025	.393	$[-1.02, 1.04]$	13 (bottom)
mock (b)	$\{.3603\} \times [.75, 2.25]$	.007	.388	$[-1.14, 1.14]$	14 (top)
	$\{.0003\} \times [.75, 2.25]$	.005	.472	$[-1.65, 1.24]$	14 (middle)
$R = 500.283548$	$\{.3603\} \times [.75, 3.75]$	$-.071$	.522	$[-1.91, 1.47]$	14 (bottom)
	$\{.0003\} \times [.75, 3.75]$	.083	.761	$[-2.01, 2.22]$	similar to 14 (bottom)

TABLE 3. Statistics for waveforms Ψ sampled along segments, rather than rectangles.

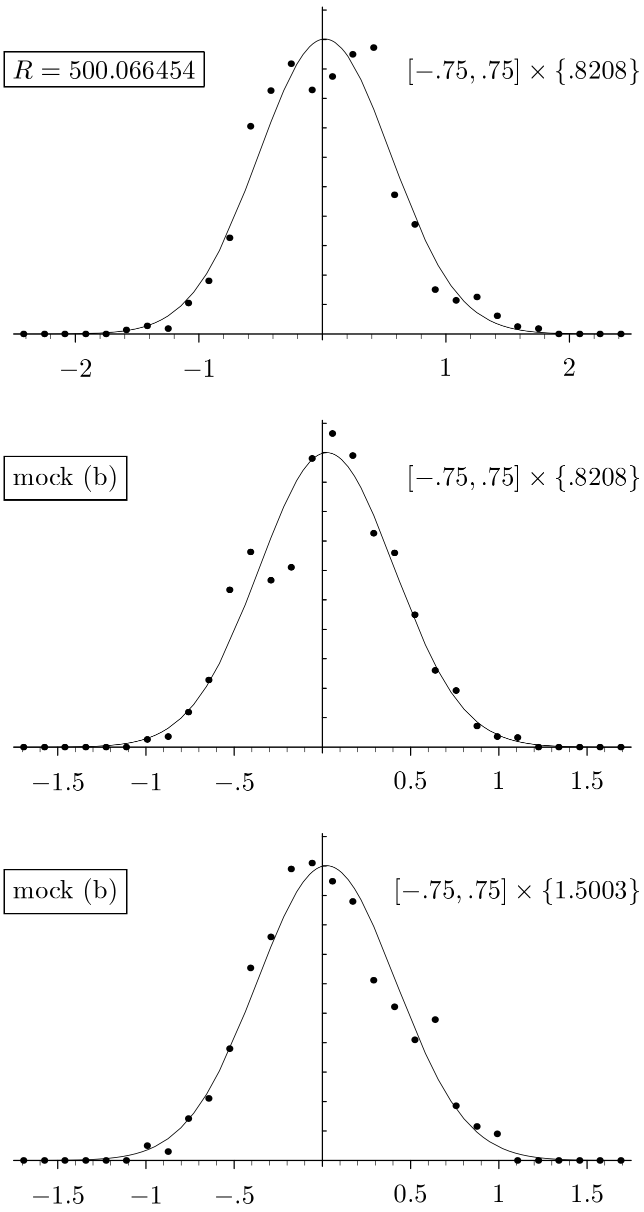


FIGURE 13. Histograms of Ψ along horizontal segments. See the top part of Table 3.

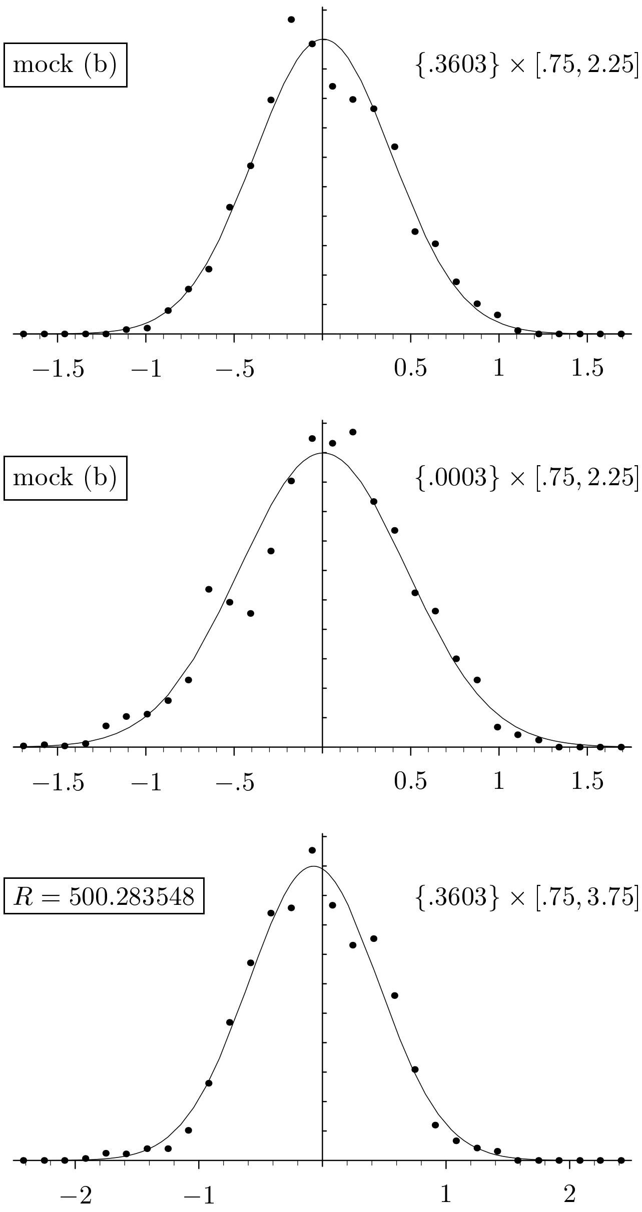


FIGURE 14. Histograms of Ψ along vertical segments. See the bottom part of Table 3.

waveform	window 1	window 2	m.s.c.	$\operatorname{cor}(f,g)$	$\operatorname{cor}(f^2,g^2)$
$R = 125.313840$	$[0,.2]\times[1,1.2]$	$[.2,.4]\times[1,1.2]$	.096	−.060	−.094
	$[0,.2]\times[\frac{\sqrt3}{2},\frac{\sqrt3}{2}+.2]$	$[.3,.5]\times[\frac{\sqrt3}{2},\frac{\sqrt3}{2}+.2]$	−.394	−.394	.070
$R = 125.347558$	$[0,.2]\times[1,1.2]$	$[.2,.4]\times[1,1.2]$	−.241	−.241	.015
	$[0,.2]\times[\frac{\sqrt3}{2},\frac{\sqrt3}{2}+.2]$	$[.3,.5]\times[\frac{\sqrt3}{2},\frac{\sqrt3}{2}+.2]$	−.295	−.295	−.005
$R = 125.523988$	$[0,.2]\times[1,1.2]$	$[.2,.4]\times[1,1.2]$	.255	.255	−.024
$R = 500.066454$	$[0,.2]\times[1,1.2]$	$[.2,.4]\times[1,1.2]$	−.166	−.166	.024
$R = 500.283548$	$[0,.2]\times[1,1.2]$	$[.2,.4]\times[1,1.2]$	−.048	−.048	−.024
mock (b)	$[0,.2]\times[1,1.2]$	$[.2,.4]\times[1,1.2]$	−.018	−.018	.007
mock (c)	$[0,.2]\times[1,1.2]$	$[.2,.4]\times[1,1.2]$	−.022	.014	−.021
mock (d)	$[0,.2]\times[1,1.2]$	$[.2,.4]\times[1,1.2]$	−.029	−.029	.001

TABLE 4. Correlative behavior of a *single* waveform on disjoint subregions. The functions f and g are the restrictions of Ψ to the two specified windows. The entry "m.s.c." (most significant correlation) indicates the correlation of largest absolute value among $\operatorname{cor}(f,g)$, $\operatorname{cor}(f^2,g^2)$, $\operatorname{cor}(|f|,|g|)$, $\operatorname{cor}(|f|,g)$, $\operatorname{cor}(f,|g|)$, $\operatorname{cor}(|f|^{1/2},|g|^{1/2})$ and $\operatorname{cor}(\operatorname{sgn} f,\operatorname{sgn} g)$.

waveform 2	window 1	window 2	m.s.c.	$\operatorname{cor}(f,g)$	$\operatorname{cor}(f^2,g^2)$
$R = 125.313840$	$[0,.2]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	−.024	−.002	−.024
	$[.2,.4]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	.024	.005	.024
$R = 125.347558$	$[0,.2]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	.010	.003	−.005
	$[.2,.4]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	−.039	.003	−.021
$R = 125.523988$	$[0,.2]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	.026	.007	.026
	$[.2,.4]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	−.006	−.001	.000
$R = 500.066454$	$[0,.2]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	.169	.169	.054
	$[.2,.4]\times[1,1.2]$	$[.2,.4]\times[1,1.2]$	.090	.090	−.037
	$[.3,.5]\times[\frac{\sqrt3}{2},\frac{\sqrt3}{2}+.2]$	$[.3,.5]\times[\frac{\sqrt3}{2},\frac{\sqrt3}{2}+.2]$	−.193	−.193	−.029
mock (b)	$[0,.2]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	.116	.116	.006
mock (c)	$[0,.2]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	.036	−.000	−.016
mock (d)	$[0,.2]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	.013	−.000	.000

TABLE 5. Correlative behavior of the waveform with $R = 500.283548$ versus other waveforms. (Window 1 refers to $R = 500.283548$.)

waveform 2	window 1	window 2	m.s.c.	$\operatorname{cor}(f,g)$	$\operatorname{cor}(f^2,g^2)$
$R = 125.313840$	$[0,.2]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	−.015	.000	−.002
$R = 125.523988$	$[0,.2]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	.013	−.001	.001
$R = 500.066454$	$[0,.2]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	−.013	.002	.003
mock (b)	$[0,.2]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	−.022	.001	.002
mock (c)	$[0,.2]\times[1,1.2]$	$[0,.2]\times[1,1.2]$	.080	.080	.005

TABLE 6. Correlative behavior of mock waveform (d) (at $R = 1000$) versus other waveforms.

As examples of directions for further experimentation, we cite:

- performing similar tests for Γ nonarithmetic, or cocompact;
- testing more waveforms, with larger values of R;
- studying the analog of the correlation function $C(X;q)$ of [Berry 1977, p. 2089, eq. (21) and last two lines];
- testing for correlative behavior over more general regions;
- testing analogs of the various geometric properties mentioned in [Longuet-Higgins 1957a,b, 1962].

We hope to treat several of these topics in a subsequent publication.

6. SOME HEURISTICS

The experiments of Section 5 have provided us with a tantalizing glimpse of what the deeper strata of the Selberg trace formalism may contain. The real challenge will come, of course, when one seeks to place things on a rigorous footing.

Given the paucity of our present data, it is probably wise to refrain from making any precise speculations as to the type of techniques that will ultimately come into play. One is reminded here of H. Weyl's famous quote [Weyl 1950, p. 131]:

> I feel that these informations about the proper oscillations of a membrane, valuable as they are, are still very incomplete. I have certain conjectures on what a complete analysis of their asymptotic behavior should aim at; but since for more than 35 years I have made no serious attempt to prove them, I think I had better keep them to myself.

Still, the histograms in Figures 11 and 12 are rather striking. Something is certainly going on there! Under the circumstances, offering some remarks of a largely heuristic nature may not be totally out of place.

Briefly put, our main idea will be to combine ideas of Rankin–Selberg type with some very suggestive results of Salem and Zygmund [1954] on partial sums of random(ized) Fourier series. Important motivation is provided by an earlier, closely related, discussion of Rice [1944, §§ 3.1, 2.8, 1.7].

To set the stage, it is best to begin with the case of a mock waveform

$$\Psi(x+iy)=\sum_{n=1}^{\infty} c_n y^{1/2} K_{iR}(2\pi ny)\cos(2\pi nx), \quad (6.1)$$

where R is arbitrary (but large!) and the c_n are chosen randomly in $[-1,1]$ with, say, uniform distribution. Since the coefficients c_n will later be viewed as independent random variables, it might have been better to write $c_n(\omega)$ in place of c_n.

Choose any $a \geq 12$ and keep y bounded away from 0 and ∞. Let

$$M=\frac{R+aR^{1/3}}{2\pi y}.$$

As in the discussion leading to (2.5), we see that it is not too far wrong to limit the summation in (6.1) to $n \leq M$. In view of (2.4), and taking into account the Convention in Section 4, we can further approximate $\Psi(x+iy)$ by

$$\sum_{n\leq M} c_n \frac{\sqrt{2\pi y}}{\sqrt[4]{R^2-(2\pi ny)^2}} \sin\Bigl(\frac{\pi}{4}+Rh\Bigl(\frac{2\pi ny}{R}\Bigr)\Bigr)\cos(2\pi nx) \quad (6.2)$$

with the convention that $\sqrt[4]{R^2-(2\pi ny)^2}$ freezes at $\sqrt[4]{2a}\,R^{1/3}$ in the range $|2\pi ny-R|\leq aR^{1/3}$, and that the sine term is then modified as appropriate.

For n significantly less than $R/2\pi y$, a quick calculation shows that the general term in (6.2) basically reduces to

$$c_n\frac{\sqrt{2\pi y}}{\sqrt{R}}\sin\Bigl(\frac{\pi}{4}+R\log\Bigl(\frac{R}{\pi nye}\Bigr)\Bigr)\cos(2\pi nx), \quad (6.3)$$

in good agreement with, say, Proposition 1.5 of [Hejhal 1990]. This shows that we are not simply dealing with something essentially equivalent to a double Fourier series.

We now set

$$Q=\sqrt{\frac{R^2}{R^2-(2\pi ny)^2}},$$

and rewrite (6.2) in the form

$$\sqrt{\frac{2\pi y}{R}}\sum_{n\leq M} c_n\sqrt{Q}\sin\Bigl(\frac{\pi}{4}+Rh\Bigl(\frac{2\pi ny}{R}\Bigr)\Bigr)\cos(2\pi nx), \quad (6.4)$$

regarding $\sqrt{Q}$ as a kind of reverse mollifier.

The quantities

$$c_n' = c_n \sin\Bigl(\frac{\pi}{4} + Rh\Bigl(\frac{2\pi n y}{R}\Bigr)\Bigr)$$

and

$$c_n'' = c_n' \sqrt{Q}$$

are presumably even more random than the c_n, because of the sine terms. A trivial calculation shows that

$$|c_n''| \le O(n^{1/6})|c_n'| \le O(n^{1/6})|c_n| \quad \text{for } n \le M. \tag{6.5}$$

Since the c_n are chosen randomly, one can safely assume that any moments or correlation-type sums built out of the c_n will be accurately estimable using results like the law of large numbers, Chebyshev's inequality, or the central limit theorem; and similarly for c_n' and c_n''.

At this point one is reminded of the central limit theorem for random(ized) Fourier series, as stated in [Salem and Zygmund 1954, Ch. III] or [Zygmund 1959, § 16.6]. Though Salem and Zygmund base their analysis on a "twisting" by Rademacher functions $\varphi_n(t)$, it is a familiar fact that this is formally equivalent to working with a special set of independent random variables [Billingsley 1986; Kac 1959, p. 6; Zygmund 1959, vol. I, p. 34 (6)].

To generalize matters, it's basically enough to replace any t-integrals in [Salem and Zygmund 1954] by their ω-counterparts (following some minor algebraic changes due to the dual role played by B_N^2). The resulting limit theorem will then hold for a relatively wide class of random Fourier coefficients $\mathring{c}_n$.

The formalism of [Salem and Zygmund 1954; Zygmund 1959] rests on the time-honored technique of characteristic functions. At least under some somewhat stronger hypotheses on $\mathring{c}_n$, it ought to be possible to recover exactly the same result using only successive moments, in line with [Billingsley 1986, p. 408]. For further insights on this, see [Rice 1944, § 4.5; Esseen 1945, pp. 32–38; Ghosh 1983, pp. 100–101].

To make things *effective*, it will then be necessary to find good estimates for various (higher) correlation-type sums involving the $\mathring{c}_n$. Doing so without first averaging over ω (that is, for specific ω) will generally be delicate.

Note that there is a strong analogy here with recent value-distribution theorems for logarithms of L-functions (and the associated Selberg moment formalism); see [Selberg 1991, § 2; Hejhal 1992a, § 3; Titchmarsh 1951, §§ 14.20–24; Tsang 1984].

In the present case, our need for an effective version of Salem–Zygmund stems mainly from the fact that (6.2) terminates at $n = M$. What one expects, of course, is that, for most ω, everything will turn out OK.

In fact, things should still be OK anytime the chosen c_n *mimic* the behavior of independent random variables sufficiently well, in terms of certain explicit higher correlation functions.

Having said this, we now go ahead and simply apply [Salem and Zygmund 1954, Ch. III] (or [Zygmund 1959, § 16.6]) for large R, with equation (6.5) in mind.

We are led to conclude heuristically that, on any segment $[x_1, x_2] \subseteq [0, \frac{1}{2}]$, the distribution of values of $\Psi(x + iy)$ should tend to look Gaussian with mean 0 and standard deviation approximately

$$\sqrt{\frac{2\pi y}{R}}\Bigl(\frac{1}{2}\sum_{n\le M}|c_n''|^2\Bigr)^{\frac{1}{2}}. \tag{6.6}$$

(In this connection, note that $E(\Psi)$ is trivially bounded by $O(1)R^{-1/2}\log(R/2\pi y)$, by virtue of (6.5).) It is understood here that suitable restrictions, depending on R, are placed on the granularity of any test sets.

Now set

$$M' = \frac{R - aR^{1/3}}{2\pi y}.$$

Then the sum in parentheses in (6.6) equals

$$\begin{aligned}\sum_{n\le M'} |c_n''|^2 &+ O(R^{2/3})\\ &= \sum_{n\le M'} |c_n|^2 Q \sin^2\Bigl(\frac{\pi}{4} + Rh\Bigr) + O(R^{2/3})\\ &= \frac{1}{2}\sum_{n\le M'} |c_n|^2 Q\,(1 + \sin(2Rh)) + O(R^{2/3}).\end{aligned}$$

At the same time, however,

$$\sum_{n\le X} |c_n|^2 \sim \Omega X, \tag{6.7}$$

with $\Omega = \frac{1}{2}\int_{-1}^{1} t^2\,dt$, by the law of large numbers. Therefore the preceding expression can also be written

$$\tfrac{1}{2}(1+O(R^{-1/3}))\sum_{n\le M'}|c_n|^2 Q + \frac{1}{2}\sum_{n\le M'}|c_n|^2 Q\sin(2Rh),$$

and the standard deviation satisfies

$$\mathrm{SD} \simeq \sqrt{T_1+T_2}, \tag{6.8}$$

where

$$T_1 = \frac{\pi y}{2R}(1+O(R^{-1/3}))\sum_{n\le M'}|c_n|^2 Q,$$
$$T_2 = \frac{\pi y}{2R}\sum_{n\le M'}|c_n|^2 Q\sin\Big(2Rh\Big(\frac{2\pi n y}{R}\Big)\Big).$$

Whenever (6.7) holds with remainder $o(X^{2/3})$ or less, we easily see using integration by parts that

$$T_1 \sim \frac{\pi\Omega}{8}. \tag{6.9}$$

By partitioning T_1 at an appropriate $(R-\mathfrak{L})/2\pi y$ and then using $|c_n| = O(1)$, we find that (6.9) actually holds unconditionally. (For $|c_n| = O(n^{\delta})$, a remainder term of $o(X^{1-2\delta})$ is needed.)

We now turn to T_2. The presence of the sine term makes one suspect that $|T_2|$ is typically much less than T_1. The simplest approach is to regard $|c_n|^2$ as a random variable and apply Chebyshev's inequality. The whole difficulty then comes down to showing that

$$\frac{\pi y}{2R}\sum_{n\le M'}\Omega Q\sin\Big(2Rh\Big(\frac{2\pi n y}{R}\Big)\Big)$$

is small. For this, the techniques of [Titchmarsh 1951, Ch. 4 and 5] will presumably suffice. (Bear in mind that the corresponding zeta function is just $\Omega\zeta(s)$; compare (6.3) for $n \ll R/2\pi y$. The typical "gain" will be a small power of R.)

In this connection, it is also reasonable to expect that sufficiently good control on

$$\sum_{n=1}^{\infty}|c_n|^2\, n^{-s}$$

would allow one to estimate T_2 directly. (Another approach that may be useful for *explicit* c_n would be to take moments of T_2 with respect to y.)

The upshot of all this is simple: for generic mock waveforms, one should expect to see

$$\mathrm{SD}(y) \approx \sqrt{\frac{\pi\Omega}{8}} \quad \text{as } R\to\infty.$$

By employing some type of simultaneous integration with respect to y, one should therefore be able to conclude that such mock waveforms will have Gaussian value distribution with mean 0 and standard deviation approximately $\sqrt{\pi\Omega/8}$ over any rectangle $[x_1,x_2]\times[y_1,y_2]$ as $R\to\infty$. Since we have $\sqrt{\pi\Omega/8} = .3618$ for $\Omega = \frac{1}{3}$, the agreement with the last seven rows of Table 2 is quite good.

Since varying y tends to make c_n' and c_n'' more random (inducing as it does something of a "multiple-shuffling" effect), it is natural to expect that the approach to normality over $[x_1,x_2]\times[y_1,y_2]$ will generally be somewhat faster and more robust than in cases where y is fixed; compare Figure 12 with Figure 13.

Our discussion of mock waveforms has been purposely phrased in such a way that the ingredients for a successful extension to true waveforms are readily discernible. The essential requirement, of course, is that the c_n need to mimic the behavior of independent random variables sufficiently well for $1\le n\le M$. The extent to which this occurs is measured by the size of certain explicit correlation-type sums involving c_n'', as mentioned earlier.

For a true (automorphic) waveform Ψ, relation (6.7) corresponds to the classical Rankin–Selberg estimate: see [Selberg 1965; Iwaniec 1984, § 5; Iwaniec 1990, §§ 2–3; Moreno 1977, § 2.4]. One gets

$$\Omega = \frac{4}{\pi}\,\frac{\cosh(\pi R)}{\mu(\mathcal{F})}\int_{\mathcal{F}}|2e^{-\pi R/2}\Psi(z)|^2\,d\mu(z) \tag{6.10}$$

and a remainder term of $O_R(X^{3/5})$, the subscript indicating that the implied constant may depend on R. Here

$$\Psi = \sum_{n=1}^{\infty} c_n y^{1/2} K_{iR}(2\pi n y)\cos(2\pi n x)$$

(subject to our usual K-Bessel convention). Sums like

$$\sum_{n\le X} c_n \quad\text{and}\quad \sum_{n\le X} c_{n+l}\,\bar{c}_n \quad \text{with} \quad l\ge 1 \tag{6.11}$$

can be treated using a variety of closely related techniques (of Rankin–Selberg type) that essentially yield $O_R(X^{1/2+\varepsilon})$ and $O_{Rl}(X^{2/3+\varepsilon})$, respectively [Epstein et al. 1985, § 3; Good 1983; Iwaniec 1985, §§ 9, 14.]

We also mention, in connection with calculating Ψ-moments over $0 \leq x \leq \alpha$, that analogs of these estimates can be successfully pushed through for cases in which n is restricted to lie in an arithmetic progression [Good 1983, p. 128; Shimura 1971, Prop. 3.64].

For fixed R, then, we thus see that the Fourier coefficients c_n are starting to simulate the behavior of independent random variables. This is true even though there are multiplicative relations corresponding to (2.7). In some sense, whatever dependencies are present tend to get "mixed out" as $X \to \infty$. The situation is somewhat analogous to Theorem 27.5 of [Billingsley 1986].

However, this much has basically been just the "zero-order" approximation. The *real* problems begin when our correlation-type sums involve products of three or more c''_n. In a nutshell, precious little is rigorously known here.

One can certainly experiment a bit with making connections to various "higher-order" Rankin–Selberg zeta functions, but this does not take care of everything. The fact that the analytic properties of such zeta functions are still largely conjectural does not help matters! See [Bump 1989, pp. 54–59, 62–66; Gelbart and Shahidi 1988, pp. 2, 65–67, 84–85, 94–97, 113; Moreno and Shahidi 1985; Shahidi 1990].

Connections of this type are further frustrated by the fact that any resulting error terms will generally possess an R dependence as well. Because of the restriction $n \leq M$, things are therefore going to have to be kept fairly explicit. (This difficulty is already visible in (6.11).)

The task at hand clearly seems daunting. One wonders if there isn't some simpler approach to the whole business! Whatever the answer, the accuracy manifested in Figures 11 and 12 serves as an important stimulus. In fact, in view of the values of M listed in Table 2, it would appear that the proposed mimicry tends to "kick in" relatively early, with respect to the size of R. The fact that y ranges over a comparatively long interval probably has something to do with this. (Cf. the definition of c''_n and the de Broglie wavelength.)

The issues of statistical independence raised in Questions 2 and 3 of Section 5 clearly lead to similar kinds of correlation-type sums built up out of mixtures of terms from the two given functions (at least when, in Question 2, the underlying windows are taken to be real translates of one another; without this restriction, things are much more complicated). The "ground-level" versions are simply the obvious analogs of (6.7) and (6.11). The expected estimates continue to hold. The subsequent difficulties appear to lie at about the same level as for the Gaussian question.

To *complete* the overall picture, it remains to forge a link between (6.10) and (3.1). We do so by starting with equations (5.2), (6.4) and (6.6). Assuming suitable randomness (and writing $R = R_k$), we thus have

$$\begin{aligned} \frac{1}{B-A}\int_A^B |\Psi_k(x+iy)|^2 dx \\ &= \frac{2\pi y}{R}\,\frac{1}{2}\sum_{n\leq M} |c''_n|^2 + E_1 \\ &= T_1 + T_2 + E_2 \\ &= \frac{\pi\Omega}{8} + E_3 \\ &= \frac{\pi}{8}\,\frac{4\cosh(\pi R)}{\pi\mu(\mathcal{F})}\,\frac{1}{\rho_k(1)^2} + E_3, \end{aligned}$$

where the E_j are obvious error terms. (Incidentally, an analogous equation holds for more general groups Γ.)

To switch this over to φ_k, we need to multiply both sides by $(2\rho_k(1)e^{-\pi R/2})^2$. This gives

$$\begin{aligned} \frac{1}{B-A}\int_A^B |\varphi_k(x+iy)|^2 dx \\ = \frac{1}{\mu(\mathcal{F})} + \frac{e^{-2\pi R}}{\mu(\mathcal{F})} + (2\rho_k(1)e^{-R\pi/2})^2 E_3. \end{aligned} \tag{6.12}$$

The limiting behavior of $(2\rho_k(1)e^{-R\pi/2})^2 E_3$ is thus the central issue.

In view of (5.3), the obvious hope is that $E_3 = O(R^{-\eta})$ for some positive η. This seems reasonable, and would clearly serve to round things out in a very natural manner. One needs to keep in mind, however, that the size of E_3 hinges on several factors, including:

(a) how random the c_n actually are, for $n \leq M$;

(b) what type of remainder term can be achieved in Salem–Zygmund;
(c) the extent of any R-dependence in the remainder term for (6.7);
(d) the size of $|T_2|$, and whether or not any averaging with respect to y was necessary to estimate it.

Of these four factors, (a) is clearly the most intractable, given current technology. In (d), though the need for averaging cannot be excluded, it seems somewhat more likely that the "texture" problems in Figure 13 are simply due to the relatively small number of summands in (6.2). (See also our comment on page 297 about multiple shuffling.)

For the modular group $\mathrm{PSL}(2, \mathbf{Z})$, we are inclined to wager that *both* (3.1) and Question 1 in Section 5 are true unconditionally (with respect to φ_k). Should this not be the case, the problem will almost certainly stem from (a)'s being true (i.e., accessible) only in some average sense as $R_k \to \infty$.

Though the situation for one-dimensional A's is less clear, we tend to think that things will still be OK along segments $y = \text{constant}$.

7. CONCLUDING REMARKS

The foregoing heuristics suggest any number of further ideas on both the theoretical and experimental fronts.

1. One idea would be to "increase" the randomness by looking at $\Psi_k(z)$ on some copy $T(A)$ instead of on A. This leads to a corresponding increase in the number of c_n entering into (6.4), which presumably helps the statistics.

Since perfect Gaussians are not ordinarily seen for bounded R, there has to be some catch! The point, of course, is that the correlation-type sums on which everything pivots are not merely functions of the c_n, but rather of the c''_n, which depend nontrivially on both y and R.

Since A and $T(A)$ are isometric, our earlier remarks about granularity and the de Broglie wavelength are also relevant in this connection.

The original idea does have a certain attractiveness, however. One wonders if there might not be some point to experimenting with, say, optimal choices of $T \in \mathrm{PSL}(2, \mathbf{Z})$.

2. Independently of any Gaussian behavior, it is very natural to ask what one can say about the maximum and minimum of $\Psi_k(z)$ over an arbitrary A as $R_k \to \infty$. An effective remainder term in the Salem–Zygmund central limit theorem would allow one to make certain statements along these lines. To achieve better precision, however, one is tempted to simply apply the results of [Salem and Zygmund 1954, Ch. IV] (albeit very heuristically). This leads to the conclusion that, on any $[x_1, x_2] \subseteq [0, \frac{1}{2}]$, the maximum and minimum of $\Psi_k(x + iy)$ should generally have magnitude about

$$(\text{constant}) \sqrt{\Omega \log \frac{R_k}{2\pi y}}, \tag{7.1}$$

so long as $R_k \gg 2\pi y$. (Consult [Moreno and Shahidi 1983, 1985] for the fourth-power analog of (6.7).) In view of (6.10), the expression in (7.1) becomes

$$(\text{constant}) \frac{\sqrt{\log \dfrac{R_k}{2\pi y}}}{\rho_k(1) e^{-\pi R_k/2}}.$$

For $\varphi_k(x + iy)$, then, one gets

$$(\text{constant}) \sqrt{\log\left(\frac{R_k}{2\pi y}\right)}. \tag{7.2}$$

Entirely similar estimates follow from a heuristic application of [Kahane 1985, pp. 67–71].

(Whether these estimates have any basis in fact will naturally depend on factors like (a)–(d) at the end of Section 6. Indeed, because of relation (2.7) and the analogous issues for $\zeta(\frac{1}{2} + it)$—as in [Titchmarsh 1951, §§ 8.12 and 14.7]—it has to be expected that (7.2) will occasionally be much too small.)

See [Rudnick and Sarnak 1992] for the latest unconditional results in this area. (In a nutshell: using Hecke operators, Rudnick and Sarnak essentially show that the φ_k-extrema must have magnitude at least $\sqrt{\log \log R_k}$.)

3. Finally, what about the Eisenstein series? Concerning this, we offer the following (very rough) sketch.

The aim is to extend the heuristics of Section 6 to $E(z; s)$, with $s = \frac{1}{2} + iR$. (See [Hejhal 1983, Ch. 6 and 11] for the necessary background.) To do so, we use the identity

$$\sum_{n=1}^{\infty} \frac{|\sigma_{i\gamma}(n)|^2}{n^s} = \frac{\zeta(s)^2\zeta(s-i\gamma)\zeta(s+i\gamma)}{\zeta(2s)}, \quad (7.3)$$

with $\gamma = 2R$ [Titchmarsh 1951, (1.3.3)], and focus initially on

$$F(z; \tfrac{1}{2}+iR) = 4\sum_{n=1}^{\infty}(n^{-iR}\sigma_{2iR}(n))y^{1/2}K_{iR}(2\pi ny)\cos(2\pi nx).$$

Keeping the K-Bessel convention in mind, observe that

$$F(z; \tfrac{1}{2}+iR) \equiv A(\tfrac{1}{2}+iR)\tilde{E}(z; \tfrac{1}{2}+iR), \quad (7.4)$$

where

$$\tilde{E}(z;s) = E(z;s) - y^s - \varphi(s)y^{1-s},$$
$$A(s) = \pi^{-s}\Gamma(s)\exp(\pi|\operatorname{Im}(s)|/2)\zeta(2s).$$

Also note that the original coefficients $n^{-iR}\sigma_{2iR}(n)$ are real.

Because of (7.3), the new Ω is in some sense

$$\frac{16\,|\zeta(1+2iR)|^2}{\zeta(2)}\log\frac{R}{2\pi y}. \quad (7.5)$$

In this connection, see [Titchmarsh 1951, §§ 5.14, 5.17, (14.2.4)] and the various manipulations found in [Hejhal 1983, pp. 694–708]. The line $\operatorname{Re}(s) = \varepsilon$ is replaced by $\operatorname{Re}(s) = \frac{5}{6}$ (or by something closer to $\frac{1}{2}$ if the Riemann Hypothesis holds). For results akin to (6.11), consult [Goldfeld 1981; Hejhal 1982; Iwaniec 1985, §§ 9, 14; Deshouillers and Iwaniec 1982, pp. 230–231; Kuznecov 1982, 1985; Vinogradov and Takhtazhyan 1987].

For fixed y, one should therefore expect the value of $F/\sqrt{\log R}$ to go Gaussian with mean 0 and standard deviation

$$\mathrm{SD}(y) \approx \sqrt{\frac{2\pi|\zeta(1+2iR)|^2}{\zeta(2)}} \quad (7.6)$$

as $R \to \infty$.

If we now set

$$e^{i\theta(t)} = \frac{A(\frac{1}{2}+it)}{|A(\frac{1}{2}+it)|},$$

the numbers

$$e^{i\theta(R)}\frac{E(z;\frac{1}{2}+iR)}{\sqrt{\log R}} \text{ and } e^{i\theta(R)}\frac{\tilde{E}(z;\frac{1}{2}+iR)}{\sqrt{\log R}} \quad (7.7)$$

are real. In view of (7.4) and (7.6), these renormalized functions should then exhibit statistics that are asymptotically Gaussian with mean 0 and standard deviation

$$\mathrm{SD}(y) \sim \sqrt{\frac{6}{\pi}}. \quad (7.8)$$

In particular, one should have

$$\frac{1}{B-A}\int_A^B |\tilde{E}(x+iy;\tfrac{1}{2}+iR)|^2 dx \sim \frac{6}{\pi}\log R \quad (7.9)$$

for $\delta \le y \le \delta^{-1}$, as well as a two-dimensional analog thereof.

Note that a purely formal integration of (7.9) over $\mathcal{F}$ produces a result which agrees, as $R \to \infty$, with the principal term $-(\varphi'/\varphi)(\frac{1}{2}+iR)$ in the Maass–Selberg relation [Hejhal 1983, pp. 200, 201, 508 (2.4), 434 (2.7)]. Compare [Zelditch 1991, p. 38 (lines 14–20)] and (6.12).

Clearly, one should at least try to check these assertions empirically! Standing in the way, however, are two potentially serious problems, from the standpoint of computational time. The first is that, if $\sqrt{\log R}$ needs to be large, machine computations are likely to be out of the question. The second is that, in the analysis surrounding (7.5), one quickly sees that, even on the Riemann Hypothesis, one has to expect a fairly large (relative) error term in the analog of (6.7) for $X \approx R/2\pi y$, perhaps of the order of $O(1/\log R)$. This error then persists in both (7.8) and (7.9).

To try to palliate this second difficulty, it is natural to seek at least a heuristic refinement in the analog of (6.7) and (6.9) for $c_n = 4n^{-iR}\sigma_{2iR}(n)$. The situation is somewhat like that of the Dirichlet divisor problem [Titchmarsh 1951, § 12.1]. The upshot is that, in (6.9), the *effective* value of Ω turns out to be

$$\frac{16\,|\zeta(1+2iR)|^2}{\zeta(2)} \times \left(\log\frac{R}{2\pi y} - \log 2 + 2\gamma + 2\operatorname{Re}\frac{\zeta'}{\zeta}(1+2iR) - 2\frac{\zeta'}{\zeta}(2)\right),$$

not merely (7.5). (Here γ is the usual Euler constant.)

In line with this, we performed a number of tests using the modified function

$$e^{i\theta(R)}\frac{\tilde{E}(z;\frac{1}{2}+iR)}{\sqrt{\mathcal{P}}}, \tag{7.10}$$

where

$$\mathcal{P}=\log\frac{R}{2\pi y_{\mathrm{av}}}-\log 2+2\gamma+2\operatorname{Re}\frac{\zeta'}{\zeta}(1+2iR)-2\frac{\zeta'}{\zeta}(2),$$

y_{av} being the average y-value over the given rectangle. (Of course, $\mathcal{P}\sim\log R$.) To two decimal places, here is what we found:

R	region	mean	SD	max	|min|
500	$[0,.2]\times[1,1.2]$	.00	1.45	5.17	5.78
1000	$[0,.2]\times[1,1.2]$	.00	1.37	6.55	5.41
5000	$[0,.2]\times[1,1.2]$	.00	1.39	7.25	6.51
25,000	$[.1,.12]\times[1,1.004]$	.00	1.43	6.51	5.95
25,000	$[.1,.12]\times[1.004,1.008]$	.00	1.40	6.56	5.38
50,000	$[.1,.12]\times[1,1.004]$	.00	1.38	6.07	6.08
50,000	$[.1,.12]\times[1.004,1.008]$	.00	1.42	8.55	5.78
50,001	$[.1,.11]\times[1,1.004]$	.00	1.39	6.26	6.00
50,002	$[.1,.11]\times[1,1.004]$	.00	1.41	6.18	6.04
50,500	$[.1,.11]\times[1,1.004]$	.00	1.35	5.94	6.40
51,000	$[.1,.11]\times[1,1.004]$	.00	1.386	6.59	6.54

We remark that, in the last four rows, the typical CPU time per job was about 39 minutes on the Cray-XMP.

In interpreting the results of this table, bear in mind here that $\sqrt{6/\pi}\simeq 1.382$, and that (7.2) is correctly normalized for use with (7.10).

In each case, the histograms were nearly perfect Gaussians. Furthermore, note that the fluctuation level in $|\,\mathrm{SD}-\sqrt{6/\pi}|$ seems generally consistent with the presence of a pre-factor $1+O(R^{-1/3})$ in T_1; see (6.8).

For $R=51{,}000$, we also made histograms for a number of horizontal cross-sections. The results were as follows:

y	mean	SD	y	mean	SD
1.000000	.00	1.40	1.002500	−.05	1.32
1.000500	−.01	1.63	1.003000	−.05	1.46
1.001000	−.02	1.34	1.003500	−.05	1.32
1.001500	−.04	1.25	1.004000	−.05	1.36
1.002000	−.04	1.30	average	−.03	1.376

The histograms were surprisingly close to being Gaussian, considering the size of the window.

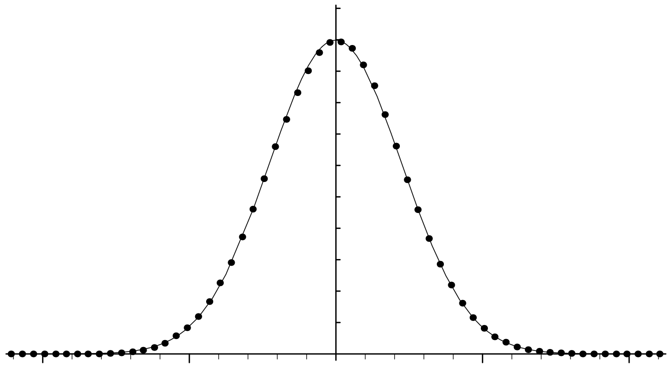


FIGURE 15. Histogram of Ψ, for $R=51{,}000$ (and $M=8164$), in the window $[.1,.11]\times[1,1.004]$ (for $F/4$), sampled on a 2000×800 grid. The minimum and maximum are -4.26 and 4.29.

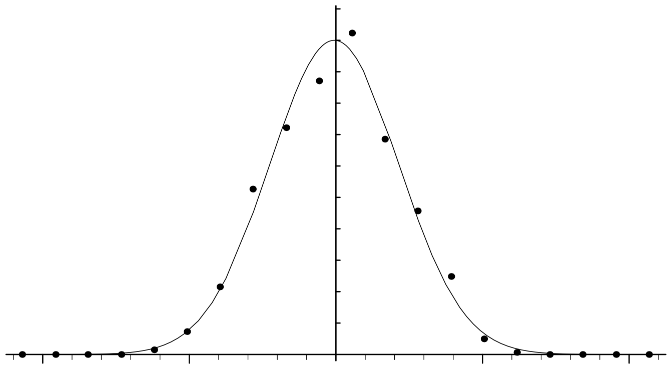


FIGURE 16. Histogram of Ψ, for $R=51{,}000$ (and $M=8156$), in the segment $[.1,.11]\times\{1.001\}$, with 2000 samples. The minimum and maximum are -2.41 and 2.34.

Figures 15 and 16 show representative examples.

On the basis of these results, we *conjecture* that letting $R\to\infty$ in (7.7) does in fact produce a limiting Gaussian with mean 0 and standard deviation $\sqrt{6/\pi}$ for every y. An analogous result should hold for any congruence subgroup. Compare [Zelditch 1991, p. 38].

This conjecture, coupled with (7.9), furnishes an interesting new perspective on the first few paragraphs of [Selberg 1991, § 5]. (These nascent links between the z- and R-aspects of E certainly deserve closer study in future years...)

In closing, we can't resist "summarizing" things with the following lines from [Kac 1959, p. 52]:

> That we are led here to the normal law... usually associated with random phenomena is perhaps an indication that the deterministic and probabilistic points of view are not as irreconcilable as

they may appear at first sight. To dwell further on this question would lead us too far afield, but it may be appropriate to quote a statement of Poincaré, who said (partly in jest no doubt) that there must be something mysterious about the normal law since mathematicians think it is a law of nature whereas physicists are convinced that it is a mathematical theorem.

REFERENCES

[Aurich and Steiner 1991] R. Aurich and F. Steiner, "Exact theory for the quantum eigenstates of a strongly chaotic system", *Physica* D**48** (1991), 445–470.

[Balazs and Voros 1986] N. L. Balazs and A. Voros, "Chaos on the pseudosphere", *Phys. Rep.* **143** (1986), 109–240.

[Balogh 1967] C. Balogh, "Asymptotic expansions of the modified Bessel function of the third kind of imaginary order", *SIAM J. Appl. Math.* **15** (1967), 1315–1323.

[Berry 1977] M. Berry, "Regular and irregular semiclassical wavefunctions", *J. Phys.* A**10** (1977), 2083–2091.

[Berry 1989] M. Berry, "Quantum scars of classical closed orbits in phase space", *Proc. Royal Soc. London* A**423** (1989), 219–231.

[Berry 1991] M. Berry, "Some quantum-to-classical asymptotics", pp. 251–303 in *Chaos et Physique Quantique, Les Houches 1989* (edited by M. J. Giannoni et al.), North-Holland, Amsterdam, 1991.

[Billingsley 1986] P. Billingsley, *Probability and Measure*, 2nd ed., Wiley, New York, 1986.

[Bogomolny 1988] E. B. Bogomolny, "Smoothed wave functions of chaotic quantum systems", *Physica* D**31** (1988), 169–189.

[Bombieri and Hejhal 1987] E. Bombieri and D. Hejhal, "Sur les zéros des fonctions zêta d'Epstein", *C. R. Acad. Sci. Paris* **304** (1987), 213–217.

[Brüning 1978] J. Brüning, "Über Knoten von Eigenfunktionen des Laplace–Beltrami–Operators", *Math. Zeit.* **158**, 15–21.

[Bump 1989] D. Bump, "The Rankin–Selberg method: a survey", pp. 49–109 in *Number Theory, Trace Formulas, and Discrete Groups* (edited by K. Aubert et al.), Academic Press, Boston, 1989.

[Colin de Verdiere 1985] Y. Colin de Verdiere, "Ergodicité et fonctions propres du laplacien", *Comm. Math. Phys.* **102** (1985), 497–502.

[Courant and Hilbert 1953] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, vol. 1, Interscience, New York, 1953.

[Courant and Hilbert 1962] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, vol. 2, Interscience, New York, 1962.

[Delande 1991] D. Delande, "Eigenstates of a chaotic system", *Comments on Atom. Molec. Phys.* **25** (1991), 281–290.

[Deshouillers and Iwaniec 1982] J.-M. Deshouillers and H. Iwaniec, "Kloosterman sums and Fourier coefficients of cusp forms", *Invent. Math.* **70** (1982), 219–288.

[Donnelly and Fefferman 1988] H. Donnelly and C. Fefferman, "Nodal sets of eigenfunctions on Riemannian manifolds", *Invent. Math.* **93** (1988), 161–183.

[Donnelly and Fefferman 1990] H. Donnelly and C. Fefferman, "Nodal sets for eigenfunctions of the Laplacian on surfaces", *J. Amer. Math. Soc.* **3** (1990), 333–353.

[Epstein et al. 1985] C. Epstein, J. Hafner and P. Sarnak, "Zeros of L-functions attached to Maass forms", *Math. Zeit.* **190** (1985), 113–128.

[Erdélyi et al. 1953] A. Erdélyi et al. *Higher Transcendental Functions*, McGraw-Hill, New York, 1953.

[Esseen 1945] C. G. Esseen, "Fourier analysis of distribution functions: a mathematical study of the Laplace–Gaussian law", *Acta Math.* **77** (1945), 1–125.

[Feller 1971] W. Feller, *An Introduction to Probabilility Theory and Its Applications*, 2nd ed., Wiley, New York, 1971.

[Gelbart and Shahidi 1988] S. Gelbart and F. Shahidi, *Analytic Properties of Automorphic L-functions*, Academic Press, Boston 1988.

[Ghosh 1983] A. Ghosh, "On the Riemann zeta function—mean value theorems and the distribution of $|S(T)|$", *J. Number Th.* **17** (1983), 93–102.

[Goldfeld 1981] D. Goldfeld, "On convolutions of nonholomorphic Eisenstein series", *Adv. in Math.* **39** (1981), 240–256.

[Good 1981] A. Good, "Cusp forms and eigenfunctions of the Laplacian", *Math. Ann.* **255** (1981), 523–548.

[Good 1983] A. Good, "On various means involving the Fourier coefficients of cusp forms", *Math. Zeit.* **183** (1983), 95–129.

[Gutzwiller 1990] M. C. Gutzwiller, *Chaos in Classical and Quantum Mechanics*, Springer-Verlag, New York, 1990.

[Hedlund 1937] G. A. Hedlund, "A metrically transitive group defined by the modular group", *Amer. J. Math.* **57** (1937), 668–678.

[Hedlund 1939] G. A. Hedlund, "The dynamics of geodesic flows", *Bull. Amer. Math. Soc.* **45** (1939), 241–260.

[Hejhal 1976] D. A. Hejhal, *The Selberg Trace Formula for* $\mathrm{PSL}(2,\mathbf{R})$, vol. 1, Lecture Notes in Math. **548**, Springer-Verlag, Berlin, 1976.

[Hejhal 1982] D. A. Hejhal, "Quelques exemples de séries de Dirichlet dont les pôles ont un rapport étroit avec les valeurs propres de l'opérateur de Laplace–Beltrami hyperbolique", *C. R. Acad. Sci. Paris* **294** (1982), 637–640.

[Hejhal 1983] D. A. Hejhal, *The Selberg Trace Formula for* $\mathrm{PSL}(2,\mathbf{R})$, vol. 2, Lecture Notes in Math. **1001**, Springer-Verlag, Berlin, 1983.

[Hejhal 1989] D. A. Hejhal, "On polynomial approximations to $Z(t)$", to appear in *Proc. of the 1989 Amalfi International Symposium on Analytic Number Theory.*

[Hejhal 1990] D. A. Hejhal, "On a result of G. Pólya concerning the Riemann ξ-function", *J. d'Analyse Math.* **55** (1990), 59–95.

[Hejhal 1991] D. A. Hejhal, "Eigenvalues of the Laplacian for $\mathrm{PSL}(2,\mathbf{Z})$: some new results and computational techniques", pp. 59–102 in *International Symposium in Memory of Hua Loo-Keng* (edited by S. Gong et al.), vol. 1, Science Press, Beijing, and Springer-Verlag, New York, 1991. Reprinted with [Hejhal 1992b].

[Hejhal 1992a] D. A. Hejhal, "On the distribution of zeros of a certain class of Dirichlet series", Internat. Math. Res. Notices, *Duke Math. J.* **66**(4) (1992), 83–91.

[Hejhal 1992b] D. A. Hejhal, *Eigenvalues of the Laplacian for Hecke triangle groups*, Memoirs Amer. Math. Soc. **469** (1992), 3–124.

[Hejhal and Arno 1992] D. A. Hejhal and S. Arno, *On Fourier coefficients of Maass waveforms for* $\mathrm{PSL}(2,\mathbf{Z})$, Research Report 1992/20, Univ. of Minnesota Supercomputer Institute, Minneapolis, 1992.

[Heller 1984] E. J. Heller, "Bound-state eigenfunctions of classically chaotic Hamiltonian systems: scars of periodic orbits", *Phys. Rev. Lett.* **53** (1984), 1515–1518.

[Heller et al. 1989] E. J. Heller, P. O'Connor and J. Gehlen, "The eigenfunctions of classically chaotic systems", *Physica Scripta* **40** (1989), 354–359. See also Heller's survey article in *Chaos et Physique Quantique, Les Houches 1989* (edited by M. J. Giannoni et al.), North-Holland, Amsterdam, 1991.

[Hobson 1931] E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics*, Cambridge University Press, Cambridge, 1931.

[Hoffstein and Lockhart 1992] J. Hoffstein and P. Lockhart, "Coefficients of Maass forms and the Siegel zero" (preprint), Brown University, 1992.

[Hopf 1937] E. Hopf, *Ergodentheorie*, Springer-Verlag, Berlin, 1937.

[Hörmander 1968] L. Hörmander, "The spectral function of an elliptic operator", *Acta Math.* **121** (1968), 193–218.

[Huntebrinker 1991] W. Huntebrinker, "Numerische Bestimmung von Eigenwerten des Laplace-Operators auf hyperbolischen Räumen mit adaptiven Finite-Element-Methoden", Ph.D. dissertation, Univ. Bonn, *Bonner Math. Schriften* **225** (1991).

[Iwaniec 1984] H. Iwaniec, "Nonholomorphic modular forms and their applications", pp. 157–196 in *Modular Forms* (edited by R. A. Rankin), E. Horwood, Chichester (England), and Halsted Press, New York, 1984. Note: In equation (5.7), replace 6 by 12.

[Iwaniec 1985] H. Iwaniec, "Promenade along modular forms and analytic number theory", pp. 221–303 in *Topics in Analytic Number Theory* (edited by S. Graham and J. Vaaler), University of Texas Press, Austin, 1985.

[Iwaniec 1990] H. Iwaniec, "Small eigenvalues of Laplacian for $\Gamma_0(N)$", *Acta Arith.* **56** (1990), 65–82.

[Kac 1959] M. Kac, *Statistical Independence in Probablility, Analysis, and Number Theory*, MAA Carus Mathematical Monographs **12**, distributed by Wiley, New York, 1959.

[Kahane 1985] J. P. Kahane, *Some Random Series of Functions*, 2nd ed., Cambridge University Press, Cambridge (UK), 1985.

[Kuznecov 1981] N. V. Kuznecov, "Petersson's conjecture for cusp forms of weight zero and Linnik's conjecture; sums of Klosterman sums", *Math. USSR Sbornik* **39** (1981), 299–342.

[Kuznecov 1982] "Spectral methods in arithmetic problems", N. V. Kuznecov, *J. Soviet Math.* **18** (1982), 398–404.

[Kuznecov 1985] N. V. Kuznecov, "Convolution of the Fourier coefficients of the Eisenstein–Maass series", *J. Soviet Math.* **29** (1985), 1131–1159.

[Linnik 1963] Ju. V. Linnik, *The Dispersion Method in Binary Additive Problems*, Transl. Math. Monographs **4**, Amer. Math. Soc., Providence, RI, 1963.

[Longuet-Higgins 1957a] M. S. Longuet-Higgins, "The statistical analysis of a random, moving surface", *Philos. Trans. Royal Soc. London* A**249** (1957), 321–387.

[Longuet-Higgins 1957b] M. S. Longuet-Higgins, "Statistical properties of an isotropic random surface", *Philos. Trans. Royal Soc. London* A**250** (1957), 157–174.

[Longuet-Higgins 1962] M. S. Longuet-Higgins, "The statistical geometry of random surfaces", *Proc. Symposia Applied Math.* **13** (1962), 105–143. Note especially the final paragraph on p. 142.

[Maass 1949] H. Maass, "Über eine neue Art von nichtanalytischen automorphen Funktionen und die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen", *Math. Ann.* **121** (1949), 141–183.

[Maass 1983] H. Maass, *Lectures on Modular Functions of One Complex Variable*, revised ed., Springer-Verlag, Berlin, 1983.

[McDonald and Kaufman 1979] S. McDonald and A. Kaufman, "Spectrum and eigenfunctions for a Hamiltonian with stochastic trajectories", *Phys. Rev. Lett.* **42** (1979), 1189–1191.

[McDonald and Kaufman 1988] S. McDonald and A. Kaufman, "Wave chaos in the stadium: statistical properties of short-wave solutions of the Helmholtz equation", *Phys. Rev.* A**37** (1988), 3067–3086.

[Mehta 1967] M. L. Mehta, *Random Matrices*, Academic Press, Boston, 1967. Revised ed., 1991.

[Moran 1968] P. Moran, "*An Introduction to Probability Theory*, Oxford University Press, Oxford, 1968.

[Moreno 1977] C. Moreno, "Explicit formulas in the theory of automorphic forms", pp. 73–216 in *Number Theory Day* (edited by M. B. Nathanson), Lecture Notes in Math. **626**, Springer-Verlag, Berlin, 1977. Note: On page 131, replace $3/\pi$ by $12/\pi$.

[Moreno and Shahidi 1983] C. Moreno and F. Shahidi, "The fourth moment of Ramanujan τ-function", *Math. Ann.* **266** (1983), 233–239.

[Moreno and Shahidi 1985] C. Moreno and F. Shahidi, "The L-functions $L(s, \mathrm{Sym}^m(r), \pi)$", *Canad. Math. Bull.* **28** (1985), 405–410.

[Ozorio de Almeida 1988] A. M. Ozorio de Almeida, *Hamiltonian Systems: Chaos and Quantization*, Cambridge University Press, Cambridge, 1988.

[Petersson 1982] H. Petersson, *Modulfunktionen und quadratische Formen*, Ergebnisse der Mathematik **100**, Springer-Verlag, Berlin, 1982.

[Phillips and Sarnak 1985] R. S. Phillips and P. Sarnak, "On cusp forms for cofinite subgroups of $\mathrm{PSL}(2, \mathbf{R})$", *Invent. Math.* **80** (1985), 339–364.

[Rice 1944] S. O. Rice, "The mathematical analysis of random noise", *Bell Sys. Tech. J.* **23** (1944), 282–332 and **24** (1945), 46–156. Reprinted as pp. 133–294 in *Selected Papers on Noise and Stochastic Processes* (edited by N. Wax), Dover, New York, 1954.

[Rudnick and Sarnak 1992] Z. Rudnick and P. Sarnak, "Notes on arithmetic quantum chaos" (preprint), Princeton University, 1992.

[Salem and Zygmund 1954] R. Salem and A. Zygmund, "Some properties of trigonometric series whose terms have random signs", *Acta Math.* **91** (1954), 245–301.

[Selberg 1956] A. Selberg, "Harmonic analysis and discontiuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series", *J. Indian Math. Soc.* **20** (1956), 47–87.

[Selberg 1965] A. Selberg, "On the estimation of Fourier coefficients of modular forms", *Proc. Symposia Pure Math.* **8** (1965), 1–15.

[Selberg 1991] A. Selberg, "Old and new conjectures and results about a class of Dirichlet series", pp. 47–63 in *Collected Papers*, vol. 2, Springer-Verlag, Berlin, 1991.

[Shahidi 1990] F. Shahidi, "Best estimates for Fourier coefficients of Maass forms", pp. 135–141 in *Automorphic Forms and Analytic Number Theory* (edited by Ram Murty), Publ. Centre de Reserches Mathématiques, Montréal, 1990.

[Shapiro and Goelman 1984] M. Shapiro and G. Goelman, "Onset of chaos in an isolated energy eigenstate", *Phys. Rev. Lett.* **53** (1984), 1714–1717.

[Shapiro et al. 1988] M. Shapiro, J. Ronkin and P. Brumer, "Scaling laws and correlation lengths of quantum and classical ergodic states", *Chem. Phys. Lett.* **148** (1988), 177–182.

[Shimura 1971] G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Functions*, Princeton University Press, Princeton, NJ, 1971.

[Shnirelman 1974] A. I. Shnirelman, "Ergodic properties of eigenfunctions", *Uspekhi Mat. Nauk.* **29**(6) (1974), 181–182 (in Russian).

[Siegel 1980] C. L. Siegel, *Advanced Analytic Number Theory*, Tata Institute, Bombay, 1980.

[Sinai 1977] Ya. G. Sinai, *Introduction to Ergodic Theory*, Princeton University Press, Princeton, NJ, 1977.

[Smith 1981] R. A. Smith, "The L_2-norm of Maass wave functions", *Proc. Amer. Math. Soc.* **82** (1981), 179–182.

[Takeuchi 1977a] K. Takeuchi, "Commensurability classes of arithmetic triangle groups", *J. Fac. Sci. Univ. Tokyo* **24** (1977), 201–212.

[Takeuchi 1977b] K. Takeuchi, "Arithmetic triangle groups", *J. Math. Soc. Japan* **29** (1977), 91–106.

[Takeuchi 1983] K. Takeuchi, "Arithmetic Fuchsian groups with signature $(1;e)$", *J. Math. Soc. Japan* **35** (1983), 381–407.

[Titchmarsh 1951] E. Titchmarsh, *The Theory of the Riemann Zeta-Function*, Oxford Universisty Press, Oxford, 1951.

[Tsang 1984] K. M. Tsang, *The Distribution of the Values of the Riemann Zeta-Function*, Ph.D. dissertation, Princeton University, Princeton, NJ, 1984.

[Uhlenbeck 1976] K. Uhlenbeck, "Generic properties of eigenfunctions", *Amer. J. Math.* **98** (1976), 1059–1078.

[Venkov 1983] A. B. Venkov, *Spectral Theory of Automorphic Functions*, Proc. Steklov Inst. Math. **153**(4), Amer. Math. Soc., Providence, RI, 1983.

[Vinogradov and Takhtazhyan 1987] A. I. Vinogradov and L. A. Takhtazhyan, "The zeta function of the additive divisor problem and spectral decomposition of the automorphic Laplacian", *J. Soviet Math.* **36** (1987), 57–78, with continuation in **46** (1989), 1734–1759 and **52** (1990), 3004–3016.

[Weyl 1950] H. Weyl, "Ramifications, old and new, of the eigenvalue problem", *Bull. Amer. Math. Soc.* **56** (1950), 115–139.

[Zelditch 1987] S. Zelditch, "Uniform distribution of eigenfunctions on compact hyperbolic surfaces", *Duke Math. J.* **55** (1987), 919–941.

[Zelditch 1991] S. Zelditch, "Mean Lindelöf hypothesis and equidistribution of cusp forms and Eisenstein series", *J. Funct. Anal.* **97** (1991), 1–49.

[Zygmund 1959] A. Zygmund, *Trigonometric Series*, vols. 1 and 2, 2nd ed., Cambridge University Press, Cambridge, 1959.

Dennis A. Hejhal, School of Mathematics, University of Minnesota, Minneapolis, MN 55455 (mf10402@uc.msc.edu)

Barry N. Rackner, Minnesota Supercomputer Center, 1200 Washington Ave. S., Minneapolis, MN 55415 (bnr@msc.edu)

Received August 26, 1992; accepted November 2