

ON FOURIER COEFFICIENTS OF MAASS WAVEFORMS FOR $\mathrm{PSL}(2, \mathbb{Z})$

D. A. HEJHAL AND S. ARNO

Dedicated to the memory of D. H. Lehmer

ABSTRACT. In this paper, we use machine experiments to test the validity of the Sato-Tate conjecture for Maass waveforms on $\mathrm{PSL}(2, \mathbb{Z}) \backslash H$. We also elaborate on Stark's iterative method for calculating the Fourier coefficients of such forms.

1. INTRODUCTORY REMARKS

Around 20 years ago, D. H. Lehmer [19] empirically investigated the extent to which the numbers $\xi_p \equiv \tau(p)p^{-11/2}$ obey Sato-Tate statistics as $p \rightarrow \infty$. Here, $\tau(n)$ is the usual Ramanujan tau-function¹ and the proposed statistics assert that

$$(1.1) \quad \lim_{x \rightarrow \infty} \frac{N[p \leq x : \theta_p \in E]}{\pi(x)} = \frac{2}{\pi} \int_E \sin^2 \theta \, d\theta$$

for $\xi_p = 2 \cos(\theta_p)$ and any Jordan measurable $E \subseteq [0, \pi]$. [$\pi(x)$ is the usual counting function for the primes.]

The corresponding assertion for ξ_p itself will then read:

$$(1.2) \quad \lim_{x \rightarrow \infty} \frac{N[p \leq x : \xi_p \in E]}{\pi(x)} = \frac{1}{2\pi} \int_E \sqrt{4 - \xi^2} \, d\xi$$

for $E \subseteq [-2, 2]$. In this form, the proposed distribution coincides with the so-called Wigner semicircle law familiar from the study of spectra of random Hermitian matrices (cf. [1, 2, 21, 22, 29, 36] for the proper perspective).

Lehmer looked at ξ_p for $p < 10000$ and found fairly good agreement with (1.1)–(1.2). Such agreement is important partly because of Serre's observation that (1.1) would follow, not only for Δ , but also for more general automorphic representations π on $\mathrm{GL}_2(\mathbb{A})$, anytime the Langlands L -functions $L(s, \mathrm{Sym}^k)$ of the symmetric powers of π have good holomorphy properties on $\{\mathrm{Re}(s) \geq 1\}$ (cf. [33, 37, 18] and, more recently, [4, 7, 25, 27, 28]). (Any exceptional cases should reduce to [10].)

Received by the editor March 12, 1992 and, in revised form, January 26, 1993.

1991 *Mathematics Subject Classification.* Primary 11F03, 11F30.

The first author was partially supported by NSF Grant DMS 89-10744 with computer time from the Minnesota Supercomputer Institute (CRAY2), Pittsburgh Supercomputing Center (YMP832), and Supercomputing Research Center (CRAY2).

¹Associated with $\Delta(z) = e^{2\pi iz} \prod_{n=1}^{\infty} (1 - e^{2\pi inz})^{24}$.

An *alternative* way of suggesting that (1.2) typically holds was recently found by Sarnak [31]. His method, which is *unconditional*, rests on taking a kind of average over π and exploits the Selberg trace formalism for $\mathrm{PSL}(2, \mathbb{Z})$ à la [11, 12, 32]. One finds an interesting interplay between p and π (cf. [3] for some related ideas).

Our primary objective in this paper is to extend [19] to the case of Maass waveforms. We will do this by making an empirical test of (1.1)–(1.2) [and equation (1.5) below] in a variety of cases wherein π corresponds to a run-of-the-mill Maass waveform of weight zero on $\mathrm{PSL}(2, \mathbb{Z}) \backslash H$. (The observation of Serre will thus apply.)

In classical terms, we propose to study functions

$$(1.3) \quad \varphi(z) = \sum_{n=1}^{\infty} c_n y^{1/2} K_{iR}(2\pi n y) \left\{ \frac{\cos(2\pi n x)}{\sin(2\pi n x)} \right\}$$

on $\mathrm{PSL}(2, \mathbb{Z}) \backslash H$ such that [12, 20, 34]

$$(1.4_a) \quad \varphi(Tz) = \varphi(z) \quad \text{for } T \in \mathrm{PSL}(2, \mathbb{Z}),$$

$$(1.4_b) \quad y^2(\varphi_{xx} + \varphi_{yy}) + (\tfrac{1}{4} + R^2)\varphi = 0 \quad \text{for } z \in H,$$

$$(1.4_c) \quad \sum_{n=1}^{\infty} \frac{c_n}{n^s} = \prod_p \frac{1}{1 - c_p p^{-s} + p^{-2s}} \quad \text{for } s \gg 1,$$

and shall seek to prepare extensive lists of c_p -values, whose statistics we shall then proceed to analyze directly.

For φ of this type, the analog of the Ramanujan-Petersson conjecture asserts that

$$(1.5) \quad |c_p| \leq 2.$$

This estimate remains unproved. The best (unconditional) estimate presently available (1991) says that

$$(1.6) \quad |c_p| \leq p^{1/5} + p^{-1/5}$$

(cf. [4, 7, 24, 26]). (See [27, Theorem 4] and [25, pp. 436–438, 433 (Theorem 4)] for several improvements of a probabilistic nature.)

Before continuing, it should be stressed that implementation of the proposed empirical test of (1.5) and (1.1) is by no means trivial.

In fact, until fairly recently, there were *major* difficulties associated with computing the eigenvalues $\frac{1}{4} + R_k^2$ themselves, particularly when R_k exceeded 25 or so. Even in smaller cases, computing c_p for $p > 7$ (!) seemed very problematic.

Part of these difficulties were successfully addressed in [14, 15]. Indeed, it is now possible to reliably determine² not only the eigenvalue, but also the first few Fourier coefficients c_n in (1.3) for R -values ranging out to 500 or so. The number of “good” coefficients scales something like $2 + \frac{1}{4}R$. (See [12, Appendix C] for a glimpse of the situation prior to [14, 15].)

Having reached this stage, the next step is obviously one of “refining things” so as to allow the systematic computation of significantly larger numbers of c_p -values (for each R_k).

²To 5 or more places.

Currently, the only published work in this direction is that of Stark [35]. Though very suggestive, Stark's technique suffers from three drawbacks:

(A) the method was tested only for relatively small R -values (compare [12, p. 729]);

(B) Stark's subroutine for $K_{iR}(X)$ needs to be applied outside its demonstrated region of validity [35, p. 266, paragraph 3];

(C) the method does not seem to include sufficient internal checks on the accuracy of the computed c_p -values, particularly for larger p (cf. [35, pp. 266 (bottom), 268 (middle)]).

In order to develop a "firmer" c_p -algorithm, we therefore decided to make our own investigation of Stark's main idea (focussing, in particular, on any subtleties that occurred).

Sections 4 and 5 will contain a brief report on the outcome of our experiments—plus a general description of the (production) algorithm that we ultimately implemented.

Prior to that time, we will simply proceed as if there were no major problems expanding Stark's original idea into *some* type of algorithm capable of computing both R_k and the first $\pi(10^4)$ c_p 's out to 9–10 places. This possibility was alluded to in [35, pp. 268 (middle), 266 (top)].

Incidentally: to help get the main iteration started, it is understood that one may need to use the data in [14] or [15, Chapter 2] as an 0-order approximation (for each R_k that we consider). Compare [35, p. 266 (lines 5–8)].

2. CONCERNING HECKE OPERATORS

Recall that φ satisfies (1.4_a)–(1.4_c). Condition (c) is just another way of saying that

$$(2.1) \quad T_n[\varphi] = c_n \varphi,$$

where T_n is the familiar Hecke operator

$$(2.2) \quad T_n[f] \equiv \frac{1}{\sqrt{n}} \sum_{\substack{ad=n \\ d>0}} \sum_{\substack{0 \leq b < d}} f\left(\frac{az+b}{d}\right) \quad \text{on } \text{PSL}(2, \mathbb{Z}) \backslash H$$

(cf. [34] or [23] modulo a slight change in normalization). This fact is largely a reflection of the basic multiplicative property of T_n ; i.e.,

$$(2.3) \quad T_n T_m = \sum_{d|(m,n)} T_{nm/d^2}.$$

In particular, one checks that

$$(2.4) \quad c_n c_m = \sum_{d|(m,n)} c_{nm/d^2}.$$

The identities implicit in (2.4) are very important and need to be constantly kept in mind in what follows. Note, for instance, that

$$\begin{aligned} c(mn) &= c(m)c(n) \quad \text{for } (m, n) = 1, \\ c(p^2) &= c(p)^2 - 1, \\ c(p^3) &= c(p)^3 - 2c(p), \\ c(p^4) &= c(p)^4 - 3c(p)^2 + 1, \\ c(p^5) &= c(p)^5 - 4c(p)^3 + 3c(p). \end{aligned}$$

When one is “sitting” exactly on φ , equation (2.4) holds perfectly. In numerical work, however, there are always errors in R and c_n to contend with.

To pave the way for our later discussion in §§4–5, it is helpful to use a simple (but loose) example to illustrate how the relations embodied in (2.4) can be turned around to help one “zero in” on the correct data for φ .

To start with, we assume that R is known at least approximately (say from [14]). In the Fourier expansion (1.3), we suppose that the “harmonics” with $n > N$ contribute terms of (relative) order 10^{-14} at most, at least for $y \geq \frac{1}{2}\sqrt{3}$.

[As usual in this business, it is helpful to renormalize any $K_{iR}(X)$ terms by multiplying through by $\exp(\frac{\pi}{2}R)$. We tacitly assume this has been done. We also let \mathcal{F} represent the standard fundamental polygon for $\text{PSL}(2, \mathbb{Z}) \backslash H$.]

Fundamental to Stark’s method is a certain nonlinear system of equations in $\{c_2, c_3, \dots, c_N\}$ ³, which one seeks to solve by iteration. (This is where an approximate knowledge of c_2, c_3, \dots, c_N comes in handy.) The system (\mathcal{S}) has R as an input parameter and is (essentially) obtained from (2.1) by neglecting those terms in (1.3) deemed to have order less than 10^{-14} on \mathcal{F} (cf. [35, pp. 265 (line 22), 266 (lines 5–8)]).

To the extent there are errors in R , there will be corresponding uncertainties in c_2, c_3, \dots, c_N .

For the sake of argument, suppose that $17 \leq N \leq 32$ (say).

Our method for “zeroing in” on the correct φ -data basically consists of little more than minimizing some functional like

$$(2.5) \quad \max[|c(33) - c(3)c(11)|, |c(34) - c(2)c(17)|, |c(35) - c(5)c(7)|]$$

with respect to R . Here, since $N < 33$, one needs to explain precisely what is meant by $c(33), c(34), c(35)$. These numbers are simply auxiliary functions of the solution data $\{c_2, c_3, \dots, c_N\}$, which are obtained by rewriting (2.1) as

$$(2.6) \quad c_n = \frac{T_n[\varphi](z_0)}{T_1[\varphi](z_0)}$$

and then substituting a suitable value for z_0 . In each case, φ is understood to be truncated after $n = N$, and the relevant points in (2.2) “tossed back over” to within \mathcal{F} .

To the extent that (2.5) can be “brought down” to something like 10^{-14} , one is inclined to suspect that R_{critical} must be very close to the true eigenvalue (i.e., R_k). [At a random R , there is simply no a priori reason for $c(33)$ to be exactly $c(3) \cdot c(11)$. Accidents *do* occur, but a simultaneous accident at 33, 34, 35 is well nigh impossible!]

Once a successful “minimization” has been achieved, the c_p -values for $p > N$ are then determined by repeating the procedure used for $c(33), c(34), c(35)$.

This is the story in a nutshell.

As N and R grow, one naturally needs to employ different composite n ’s in (2.5). It is in the use of *such* n that our approach differs (most significantly) from that of Stark [35].

Incidentally: in (2.6), note that varying z_0 furnishes an additional check on any proposed φ -data.

³Call it (\mathcal{S}).

For later use, we remark that

$$(2.7) \quad L(s) = \sum_{n=1}^{\infty} \frac{c_n}{n^s} = \prod_p \frac{1}{1 - c_p p^{-s} + p^{-2s}}$$

is actually an entire function satisfying

$$(2.8) \quad G(s)L(s) = \eta G(1-s)L(1-s),$$

where

$$\begin{cases} \eta = +1, & G(s) = \pi^{-s} \Gamma\left(\frac{s+iR}{2}\right) \Gamma\left(\frac{s-iR}{2}\right) & \text{for even } \varphi, \\ \eta = -1, & G(s) = \pi^{-s} \Gamma\left(\frac{s+1+iR}{2}\right) \Gamma\left(\frac{s+1-iR}{2}\right) & \text{for odd } \varphi. \end{cases}$$

The proof is a standard application of Mellin transforms to (1.3) and (1.4). One exploits [38, p. 388(8)] along the way.

3. STATEMENT OF RESULTS

Our primary goal was to study the first five even and odd Maass waveforms on $\text{PSL}(2, \mathbb{Z}) \backslash H$, as accurately as possible, for p ranging out to 10000.

Since the K -Bessel routine in [14] has an intrinsic accuracy of about 11–12 places (for $R \lesssim 100$), one would naively expect problems to occur if we tried to obtain *more* than this number of places in R and c_p . In fact, since c_p is determined using (2.6), it is clear that the accuracy must gradually begin to taper off as p grows (due to the increasing amount of roundoff error in the numerator). In line with this, we made a very slight change in the procedure outlined in §2; viz., we replaced 10^{-14} by 10^{-12} .

In addition to testing the first five even and odd R_k 's, we thought it would be interesting to see if our method could still be applied successfully in a couple of cases with somewhat larger R . For this purpose, we chose two *even* values; one near $R = 50$, the other near 125.

The total number of production runs was thus 12.

In this connection, however, it should be stressed that the production stage is reached only *after* c_2, c_3, \dots, c_N, R are computed to sufficient accuracy. Achieving this, basically requires a *number* of test runs, each aimed at trying to “successfully” minimize (2.5) through a sequence of steps reminiscent of regula falsi. [Ultimately achieving the desired accuracy is not a foregone conclusion, since system (\mathcal{S}) tends to become rather sensitive as R increases. Suffice it to say, though, that the only place a little persistence was really needed was with the two larger R . Further details can be found in §4.]

The R -values we used to get the whole process started were taken from [14] and read as in Table 1 (next page). To speed things up in solving (\mathcal{S}), we did not hesitate to use the approximate values of the first few c_n found in [14] or [12, Appendix C]. We mentioned this point earlier in §1.

For the production runs, we then used Table 2 (next page).

In each case, our tests with the analog of (2.5) suggested that the indicated

TABLE 1

| | | | |
|------|------------|-----|-----------|
| A | 13.779751 | H | 9.533695 |
| B | 17.738563 | I | 12.173008 |
| C | 19.423481 | J | 14.358509 |
| D | 21.315796 | K | 16.138073 |
| E | 22.785909 | L | 16.644259 |
| F | 47.926558 | | |
| G | 125.313840 | | |
| even | | odd | |

TABLE 2

| case | R | N | $p < x$ | $\pi(X)$ |
|------|------------------|-----|-------------|----------|
| A | 13.7797513518907 | 9 | $p < 10000$ | 1229 |
| B | 17.7385633810580 | 9 | $p < 10000$ | 1229 |
| C | 19.423481470828 | 10 | $p < 10000$ | 1229 |
| D | 21.3157959402041 | 10 | $p < 10000$ | 1229 |
| E | 22.785908494190 | 10 | $p < 10000$ | 1229 |
| F | 47.926558330595 | 20 | $p < 10000$ | 1229 |
| G | 125.3138401770 | 32 | $p < 5000$ | 669 |
| H | 9.5336952613536 | 9 | $p < 10000$ | 1229 |
| I | 12.1730083246801 | 9 | $p < 10000$ | 1229 |
| J | 14.358509518259 | 9 | $p < 10000$ | 1229 |
| K | 16.138073171521 | 9 | $p < 10000$ | 1229 |
| L | 16.6442592018999 | 9 | $p < 10000$ | 1229 |

R -value should be accurate to within

$$\begin{cases} 3 \times 10^{-13}, & \text{if 13 places are shown,} \\ 2 \times 10^{-12}, & \text{if 12 places are shown and } R < 23, \\ 3 \times 10^{-12}, & \text{in case F,} \\ 2 \times 10^{-10}, & \text{in case G.} \end{cases}$$

Tables of c_p -values were then prepared. Our spot checks (using both composite n and variation of z_0) suggest that the accuracy we obtained is quite good: 9–10 places initially, dropping off to 8–9 for p around X . (Incidentally: there was generally 9–10 place agreement between [35, Table 1] and case A.)

In *no* instance were any counterexamples to Ramanujan-Petersson detected.

Readers interested in scanning the actual c_p -values should refer to the Supplemental Listing available directly from the authors.

On the matter of “extreme” values, Table 3 holds. These values should be compared with the numbers found by D. H. Lehmer for $\tau(p)$:

$$\max_{p \leq 10^4} |\xi_p| = 1.961, \quad \min_{p \leq 10^4} |\xi_p| = .00062.$$

Finally, to check (1.1) and (1.2), we made some histograms [using the values contained in the Supplemental Listing]. These are included as Figures 1–6 in the Supplement section at the end of the issue. We are grateful to Andrew Odlyzko for his assistance with this.

TABLE 3

| | | | |
|---|---|---|-----------------|
| A | $c(8317) = 1.984$ $c(7687) = -1.987$ | $c(9539) = .000540$ $c(97) = -.003257$ | |
| B | $c(4217) = 1.984$ $c(7451) = -1.962$ | $c(7583) = .000637$ $c(8329) = -.000357$ | |
| C | $c(2699) = 1.969$ $c(8543) = -1.980$ | $c(4703) = .000576$ $c(6277) = -.004243$ | |
| D | $c(1201) = 1.973$ $c(8233) = -1.951$ | $c(1741) = .001437$ $c(9539) = -.000048$ | |
| E | $c(6599) = 1.997$ $c(7451) = -1.978$ | $c(6043) = .005829$ $c(727) = -.002484$ | |
| F | $c(8629) = 1.990$ $c(4217) = -1.991$ | $c(2903) = .002581$ $c(2953) = -.001902$ | |
| G | $c(131) = 1.944$ $c(3347) = -1.958$ | $c(1381) = .001109$ $c(1451) = -.006469$ | $p < 5000$ only |
| H | $c(3457) = 1.998$ $c(1277) = -1.987$ | $c(5953) = .003227$ $c(3121) = -.000615$ | |
| I | $c(7517) = 1.993$ $c(5501) = -1.988$ | $c(709) = .001142$ $c(601) = -.000066$ | |
| J | $c(7309) = 1.963$ $c(4561) = -1.945$ | $c(5009) = .006958$ $c(7411) = -.000487$ | |
| K | $c(8837) = 1.993$ $c(9539) = -1.914$ | $c(2311) = .002891$ $c(7583) = -.002180$ | |
| L | $c(1259) = 1.977$ $c(6211) = -1.994$ | $c(8933) = .000010$ $c(647) = -.002424$ | |

The agreement with (1.1), (1.2) certainly seems reasonable.

Out of curiosity, we also tested the *individual* R -values (but on somewhat coarser scales). This gave the results illustrated in Figures 7–18 of the Supplement.

In all cases, the agreement (still) seems fairly good.

The obvious question now is why can't we do $p < 10^5$, or maybe even $p < 10^6$?

To address this point, we need to first look at the CPU times for our production runs (Table 4 on next page). Since these jobs typically included spot checks at about a half a dozen composite n as well,⁴ the *actual* time (for obtaining just the c_p) will be about 10%–15% less.

CPU times in the range of $1\frac{1}{2}$ – $6\frac{1}{2}$ hours are, of course, nontrivial.

The key issue then is how things scale with respect to X (and R).

The *asymptotics* of $\exp(\frac{\pi}{2}R)K_{iR}(v)$ are known from [5], say. These asymptotics and the shape of \mathcal{F} make it clear that N must scale something like (constant) R . In determining c_p by means of (2.6), approximately $N(p+1)$ [independent] K -Bessel calls will be needed. To tabulate c_p for *all* $p \leq X$ will

⁴The form of these n being either $5p$ or $13p$.

TABLE 4

| case | machine | time (in secs.) | X |
|------|---------|-----------------|-------|
| A | Cray2 | 10103 | 10000 |
| B | YMP | 8151 | 10000 |
| C | Cray2 | 12631 | 10000 |
| D | Cray2 | 13454 | 10000 |
| E | Cray2 | 14039 | 10000 |
| F | YMP | 25718 | 10000 |
| G | YMP | 11791 | 5000 |
| H | YMP | 5812 | 10000 |
| I | YMP | 6861 | 10000 |
| J | YMP | 7441 | 10000 |
| K | YMP | 7820 | 10000 |
| L | YMP | 7887 | 10000 |

therefore require about

$$N \sum_{p \leq X} (p+1) \approx \frac{N}{2} \frac{X^2}{\log X} \approx (\text{constant}) R \frac{X^2}{\log X}$$

calls.⁵ The essential point is the occurrence of $X^2/(\log X)$. Unless the K -Bessel calls can somehow be speeded up, raising X by a factor of 10 is going to cause the CPU time to go up by something closer to 100. (Expressed differently: two hours of CRAY time is one thing, 200 is another!)

It is here, with the K -Bessel calls, that some major improvement (or restructuring) seems essential for further progress.

The routine that we used was simply that of [14, 15]. (We wanted to use something well tested!) Though highly stable, this routine requires Gaussian quadrature and a deformation of path. [Typical CPU time: .00012 secs.]

Iterative techniques, or asymptotic expansions, may offer some better possibilities (at least in certain ranges). Another idea worth exploring would be interpolation.

4. THE BASIC ALGORITHM IN GREATER DETAIL

Let $g(t) = \sqrt{t^2 - 1} + \arcsin(1/t)$ for $t \geq 1$. In line with [5], one knows that $\exp(\frac{\pi}{2}R)K_{iR}(v)$ behaves something like $e^{-R[g(v/R) - \pi/2]}$ for $v > R$. It is therefore possible to determine relatively easily how big v must be in order to guarantee that $\exp(\frac{\pi}{2}R)K_{iR}(v)$ will be less than 10^{-14} (say). When $v = 2\pi ny$, as in (1.3), this typically translates into a condition on n of the form

$$n > \frac{R + \mu R^{1/3}}{2\pi y},$$

where μ is some modest constant. [For large R , the constant μ is about 12.]

If y is known to be at least $\frac{1}{2}\sqrt{3}$, taking

$$(4.1) \quad n > \frac{R + \mu R^{1/3}}{\pi\sqrt{3}}$$

⁵To be completely correct, these calls should be subdivided into cases $v \ll R$, $v \approx R$, $v \gg R$. For fixed R , the relative proportion of each type remains relatively constant as $X \rightarrow \infty$. (Calls with $v \ll R$ take the most time ...)

will then accomplish exactly what was asked for in §2 with regard to N . (Similarly for 10^{-12} , as used in §3.)

As a temporary notation, let \hat{z} denote the point in \mathcal{F} which is equivalent to z under the action of $\mathrm{PSL}(2, \mathbb{Z})$.⁶

In our earlier discussion, system (\mathcal{S}) was intentionally kept a bit vague. This was done partly to avoid mentioning that several variants are, in fact, possible.

To see this, suppose for a moment that $R = R_k$ is known.

One way of getting a system of equations for $\{c_2, c_3, \dots, c_N\}$ is obvious. Simply rewrite (2.1) as (2.6) and then insist that, in forming $T_l[\varphi](z_0)$ according to (2.2), the points $(az_0 + b)/d$ are all given hats. Follow this by replacing φ with its truncation after N terms. The resulting system will then make sense for *any* R (and will clearly depend on both R and z_0). We will refer to z_0 as the test point.

To the extent that the denominator remains bounded away from 0, one would expect to see (at $R = R_k$) a final accuracy roughly commensurate with the choice of N (cf. (1.5) and (1.6)). Letting the denominator go to 0, however, tends to reduce this accuracy by causing *catastrophic* cancellation in the numerator. [Bear in mind here that the early K -Bessel terms are roughly $O(1)$ in size and only accurate to about 11–12 places.]⁷ For this reason, some care in the choice of z_0 is definitely necessary.

An *alternative* system can be obtained by using (2.4) to re-express any c_l with composite $l \leq N$ as a polynomial in the c_p 's (with $p \leq N$). One then considers (2.6) for *prime* n ($\leq N$) only. This is basically what Stark does (cf. [35, p. 266 (paragraph 1)]).⁸

This “smaller” system seems especially appropriate on $\mathrm{PSL}(2, \mathbb{Z}) \backslash H$.

In either case, the final format of our system will be

$$(4.2) \quad \mathbf{c} = T(\mathbf{c}),$$

where \mathbf{c} is a point in \mathbb{R}^M ,

$$(4.3) \quad M = \left\{ \frac{N-1}{\pi(N)} \right\},$$

and the component functions T_j of T are all rational.

Anytime one tries to solve a nonlinear equation like (4.2) by iteration, strange things can happen. One need only examine some of the recent literature on dynamical systems to be convinced of this.

In the absence of any kind of “hard” information about T , one simply has to expect that:

- (a) there may be multiple solutions;
- (b) the limiting behavior of $T^n(\mathbf{c}_0)$ as $n \rightarrow \infty$ may depend strongly on the choice of (initial point) \mathbf{c}_0 ;
- (c) to get convergence, one may have to take \mathbf{c}_0 rather close to the proposed fixed point;

⁶The algorithm for finding \hat{z} is familiar: one simply does the obvious flip-flop between translation and inversion. Cf. [30, p. 51 (lines 19–27)].

⁷Readers unfamiliar with catastrophic cancellation may wish to quickly consult §5, Example 5.

⁸In Stark's approach, the denominator $T_l[\varphi](z_0)$ is left as $\varphi(z_0)$ and is not replaced by $\varphi(\hat{z}_0)$. This is important only if $z_0 \notin \mathcal{F}$. Since the (modular) correspondence $T_l\{z\}$ “lives” on $\mathrm{PSL}(2, \mathbb{Z}) \backslash H$, our version has the effect of requiring that $z_0 \in \mathcal{F}$.

(d) the limiting behavior of $T^n(c_0)$ may be significantly affected by small perturbations in T (e.g., movement in z_0).

Though the “bigger” system is not without interest, we decided to follow Stark [35] and concentrate solely on the case of $\pi(N)$ -variables (cf. (4.3)).

One of the advantages to using the CRAY was that its speed allowed us to plot the “trajectories” $\{T^n(c_0): n \geq 1\}$ virtually instantaneously.

In the next section, we will present several examples which illustrate points (a)–(d).

The upshot will be that the global situation is not nearly so simple as was suggested in [35].

In cases where $T^n(c_0)$ converges, we like to think of the point as having made its way into some type of “attracting basin”. *The problem is, that as the dimension M grows, it becomes easier for $T^n(c_0)$ to get “lost”.*

In a certain sense: one wants to “help” $T^n(c_0)$ find its way (particularly when $R \approx R_k$).

Before explaining how we do this, it should be noted that, in defining the “smaller” system, declaring $n = \text{prime}$ at the very last step was a bit arbitrary. Indeed, one can legitimately maintain that

$$(4.4) \quad \begin{aligned} c_2 &= \frac{T_2[\varphi](z_0)}{T_1[\varphi](z_0)}, & c_2 &= \frac{T_6[\varphi](z_0)}{c_3 T_1[\varphi](z_0)} \\ c_2 &= \frac{T_{10}[\varphi](z_0)}{c_5 T_1[\varphi](z_0)}, & c_2 &= \frac{T_{14}[\varphi](z_0)}{c_7 T_1[\varphi](z_0)}, \text{ etc.} \end{aligned}$$

should all be considered on an equal footing.

This adds still another variant to (\mathcal{S}) , and simultaneously supplies us with an important hint.

The point is this. In numerous cases, one knows (from [14]!) the first few c_p -values for $\varphi(z)$ out to, say, six decimal places. [We will temporarily refer to these approximate values as $\hat{c}_2, \hat{c}_3, \dots, \hat{c}_q$.] Philosophically speaking, one is inclined to believe that there simply *must* be *some* way of exploiting this knowledge to help keep $T^n(c_0)$ ’s “nose” pointed in the right direction. (One seeks to impose a kind of potential field.)

In this context, one naturally assumes that any c_0 we start with will have components at least as good as $\hat{c}_2, \hat{c}_3, \dots, \hat{c}_q$. With this in mind, one way of introducing a kind of “restraining force” [to prevent divergence] would be to replace

$$c_2 = \frac{T_2[\varphi](z_0)}{T_1[\varphi](z_0)}$$

in (2.6) by

$$c_2 = \text{that quotient in (4.4) which lies closest to } \hat{c}_2.$$

This choice suffers, however, from the defect of wanting to pull c_2 too much in the direction of \hat{c}_2 (which is, after all, *not* the exact answer). To rectify this, it seems better to “sit on the particle itself” and go with:

$$(4.5) \quad c_2 = \text{that quotient in (4.4) which lies closest to } c_2.$$

Note that the right-hand side is, in fact, a legitimate function of c . (This requires a moment’s thought!)

To prevent “runaways”, one can erect a kind of “fence” by taking

$$(4.6) \quad c_2 = \begin{cases} \text{RHS of (4.5),} & \text{if } \textit{that} \text{ lies within } \delta \text{ units of } \hat{c}_2, \\ \hat{c}_2, & \text{otherwise.} \end{cases}$$

Similarly for the other c_p (with $p \leq q$).

To avoid confusion in what follows, we will let (\mathcal{S}) denote the *original* “smaller” system. The variants associated with (4.5) and (4.6) are then denoted by (\mathcal{S}^*) and (\mathcal{S}^{**}) .

Once a system has been decided upon, the basic procedure is essentially that outlined in §2.

In practice, one minor irritation with both (\mathcal{S}^*) and (\mathcal{S}^{**}) is that T is no longer real-analytic [especially with respect to R]. This allows the limiting trajectory to exhibit occasional *nonsmooth* behavior, which, in turn, can adversely affect the minimization procedure being used in connection with (2.5) (cf. §3, paragraph 5). The problem is basically one of trying to get c_0 as close as possible to the true coefficient vector (cf. point (c)).

Even when the “pump is primed” with data from [14], the process of computing $\{R_k, c_2, c_3, \dots, c_N\}$ to sufficient accuracy can be time-consuming. As R and M grew, it was found to be increasingly difficult to get by with just (\mathcal{S}) . [One would typically try many z_0 and have almost no luck finding the right attracting basin!] To a lesser extent, the same thing happened with (\mathcal{S}^*) . In order to treat $R = 47.926558$ and 125.313840 , passing to (\mathcal{S}^{**}) seems to be the key.

Even then, depending on the quality of the initial data, great sensitivity with respect to z_0 can still occur. At $R = 125.313840$, one needed to try something like 20 different test points (before finally achieving success).

To accelerate the process of determining $\{R_k, c_2, c_3, \dots, c_N\}$, we made a slight modification in the approach outlined in §2. Namely: once an approximation better than c_0 became available (from one of our runs), we did not hesitate to use it as the “new” c_0 . These updates were then repeated, as necessary. (We also found it helpful to run a variety of test points simultaneously.)

At some point, it might be useful to implement these ideas in a separate code. For purposes of §3, however, we were content to do everything “by hand”.

For simplicity, we decided to use \mathcal{S}^{**} throughout.

Prior to making the final production run at $R = R_k$, we generally liked to run two or three mini-versions (out to $p = 103$ or so) using significantly different test points. These runs always included six composite n (à la footnote 4) plus a direct calculation of $T(c) - c$. The case displaying the “best” behavior on these two items was the one we then chose for our full-scale run.

[The mini-runs also provided a further form of “insurance” in that it was always very satisfying to be able to watch the c_p -values (for $p > N$) consistently agree to 9–10 places, no matter how disparate the z_0 were!]

5. SOME SUBTLETIES IN THE ITERATION

The task of setting up the production runs for §3 was very instructive. By working interactively through the various steps, one was able to gain important insight into the (dynamical) behavior of \mathcal{S} , \mathcal{S}^* , \mathcal{S}^{**} .

In several cases, prompted by what was seen, it was natural to push things a bit further just for their own sake.

Our aim in this section will be to report on *some* of the more interesting things we found (see [16] for a fuller list).

Before doing this, however, we need to preface things with a number of preliminary remarks.

First and foremost: we have not really sought to make any kind of rigorous, all-inclusive analysis of \mathcal{S} , \mathcal{S}^* , \mathcal{S}^{**} . Our use of these systems has basically been only as a means to an end. On the matter of rigor, it is the usual story; the computer may suggest (to 12 decimal places!) that T has a fixed point at \mathbf{p} , but that does not constitute a proof. Accordingly, any phenomena depicted in our examples must⁹ technically be regarded as conjectural. We hasten to add, however, that T is basically rational with respect to \mathbf{c} (cf. (4.3) and (4.4)). This fact, combined with the existence of a variety of very similar examples, leads us to believe (in each instance) that the properties highlighted are, in fact, quite rigorous.

Secondly: it has been more or less implicit in our earlier discussion that, when $R = R_k$, the transformation T for case (\mathcal{S}) *does* have an actual fixed point someplace very close to the c_p -vector for $\varphi_k(z)$. In what follows, we will simply *assume* that this is so.

[In practice, the only cases which give one pause are those where the prevailing flow pattern seems to be repellent. Here, however, a little numerical differentiation with respect to \mathbf{c} will (in all likelihood) suffice to settle the matter. The problem is essentially one of computing the Jacobian matrix of $T(\mathbf{x}) - \mathbf{x}$.]

For convenience, we will generally engage in a slight abuse of language and make *no* distinction between the fixed point and the actual φ -data.

Our third point concerns \mathcal{S}^* and \mathcal{S}^{**} . The physical analogies used earlier were intended mainly as motivation. To a large extent, the expressions in (4.5) and (4.6) can only be regarded as ad hoc. The thing to keep in mind, though, at (say) $R \approx 125$ is that the problem confronting us was basically one of “cautiously groping through the fog” so as to be able to successively improve the accuracy in our c_p -values (after starting with the partial result found in [14]). We thus tend to see \mathcal{S}^{**} more as a kind of *device* than anything else [a device which is applied only near φ_k !].

Though successful at $R = 125.313840$, it was not a case of running it once and then, presto, there was the answer. What was needed was a willingness to test a variety of z_0 and to make updates in \mathbf{c}_0 as one moved along. In this connection we often found it helpful to have the machine display the results of the *best* iteration [as measured by (2.5)] and not simply those of the last. (If \mathcal{S}^* had been used instead, similar remarks would still apply.)

The fourth, and final preliminary, is simply a reminder that the nature of a fixed point \mathbf{p} is, to a large extent, characterized by the eigenvalues of the Jacobian matrix $J_T(\mathbf{p})$. For good convergence, one wants these eigenvalues to have modulus less than 1.

[In specific cases, if worst comes to worst, one can of course seek to calculate $J_T(\mathbf{c})$ numerically. This is not unreasonable when T is rational.]

To avoid ambiguities below, let T_0 be the T -transformation associated with system (\mathcal{S}) .¹⁰

⁹At least for now.

¹⁰Note too that any T_0 -fixed point is automatically one for \mathcal{S}^{**} (but not conversely)!

Example 1. The following data came from a job of type (\mathcal{S}^{**}) with parameters as shown. The vector $(c_0(p))$ is accurate to about eight places. On the next level, the *first* $c(p)$ -vector is the answer after 400 iterations (of T). The middle column shows what one then gets after an application of T_0 . The final column lists the difference between the two. (The *largest* difference, i.e., $1.7E-10$, will be called the T_0 -indicator.) Incidentally: in this example, the size of the analog of (2.5) was about $4.2E-11$.

Parameters for (4.3) and (4.6)

$$N = 32, \quad M = \pi(N) = 11, \quad \delta = 10^{-4}, \quad q = 13$$

```
R= 125.313840177000   ICOUNT:  400   MAX:  400
FOR X=  0.05700000000   Y=  0.92300000000
and initial values:
C0( 1)=  1.000000000000
C0( 2)= -0.332695680000
C0( 3)=  1.056573770000
C0( 5)= -0.946096480000
C0( 7)=  0.324132110000
C0(11)=  0.040097140000
C0(13)=  0.479453980000
C0(17)=  1.254988810000
C0(19)=  0.893140170000
C0(23)=  1.605561140000
C0(29)=  0.856710520000
C0(31)= -0.250878130000
```

| | | |
|------------------------|-----------------|-------------|
| C(1)= 1.000000000000 | 1.000000000000 | 0.0000E+00 |
| C(2)= -0.332695683096 | -0.332695683078 | -0.1845E-10 |
| C(3)= 1.056573773111 | 1.056573773281 | -0.1700E-09 |
| C(5)= -0.946096480439 | -0.946096480532 | 0.9242E-10 |
| C(7)= 0.324132112930 | 0.324132112918 | 0.1231E-10 |
| C(11)= 0.040097144250 | 0.040097144250 | 0.2220E-15 |
| C(13)= 0.479453986303 | 0.479453986287 | 0.1646E-10 |
| C(17)= 1.254988815396 | 1.254988815396 | 0.1421E-13 |
| C(19)= 0.893140180104 | 0.893140180104 | 0.4974E-13 |
| C(23)= 1.605561152662 | 1.605561152662 | 0.7105E-14 |
| C(29)= 0.856710527884 | 0.856710527884 | 0.4619E-13 |
| C(31)= -0.250878136751 | -0.250878136751 | -0.3197E-13 |

the (2.6) denominator was: .6331E00

With $(c_0(p))$ accurate to 8 places, one might think that the same job run with (\mathcal{S}) would converge. This did not happen, however. The fact is, after $n = 24$ or so, the point $T_0^n(c_0)$ basically became “lost” (and began to wander aimlessly).

Switching to 11-place c_0 -data made very little difference; 24 just became 29.

This shows how sensitive things can be (and provides some indication of the utility of \mathcal{S}^{**}).

In both \mathcal{S} -jobs, there was a steady growth in the size of (2.5)¹¹ prior to things getting completely lost. We are inclined to speculate that the vector

¹¹And in the T_0 -indicator.

TABLE 5

| ITER. | T_0 -indicator | (2.5) analog | T_0 -indicator | (2.5) analog |
|---------------------------|------------------|------------------------|------------------|--------------|
| 100 | 9.8E-8 | 7.7E-8 | 3.7E-3 | 4.6E-3 |
| 400 | 1.7E-10 | 4.2E-11 | 2.1E-10 | 2.3E-10 |
| 800 | 1.7E-10 | 4.1E-11 | 2.1E-10 | 2.3E-10 |
| system \mathcal{S}^{**} | | system \mathcal{S}^* | | |
| with eight-place $c_0(p)$ | | | | |

TABLE 6

| $R = 125.313840$ |
|-------------------------|
| $c_{00}(2) = -0.332696$ |
| $c_{00}(3) = 1.056574$ |
| $c_{00}(5) = -0.946097$ |
| $c_{00}(7) = 0.324132$ |
| $c_{00}(11) = 0.040097$ |
| $c_{00}(13) = 0.479454$ |
| $c_{00}(17) = 1.254990$ |
| $c_{00}(19) = 0.893140$ |
| $c_{00}(23) = 1.605562$ |
| $c_{00}(29) = 0.85671$ |
| $c_{00}(31) = -0.25088$ |

$(c(p))$ must simply represent some type of *repellent* fixed point for T_0 . [We did not compute $J_{T_0}(c)$, however.]

For the original (eight-place) $c_0(p)$, system (\mathcal{S}^*) yields results very similar to those of (\mathcal{S}^{**}) , but with less accuracy.

In this connection, Table 5 tells an interesting story.

To give some indication of how much additional accuracy (\mathcal{S}^{**}) ultimately provided, we remark that the *original* data taken from [14] was simply Table 6. [The eight-place vector $(c_0(p))$ used in Table 5 is update #2.]

Finally, let:

$$\mathcal{A} = \{z_0 \in H : T_0 \text{ has an attractive fixed point at the } \varphi\text{-data for } R = R_k = 125.3138401770\}.$$

Since none of the many \mathcal{S} -jobs that we ran with $(c_{00}(p))$, or with 8–11 place $(c_0(p))$, ever managed to converge properly,¹² we suspect that \mathcal{A} is rather sparse. (It should also be borne in mind here that for 19 earlier choices of z_0 , \mathcal{S}^{**} failed to converge properly as well.)

One might also ask, for *given* z_0 (and, say, $R = 125.3138401770$), how big a set

$$\mathcal{N}_{z_0} \equiv \left\{ \mathbf{p} \in \mathbb{R}^m : \lim_{n \rightarrow \infty} T_0^n(\mathbf{p}) \text{ does not exist} \right\}$$

is (cf. our earlier remarks about wandering aimlessly for $n \geq 24, 29$).

¹²I.e., to the correct φ -data.

Example 2. Passing to (\mathcal{S}^*) or (\mathcal{S}^{**}) does *not* always yield improved dynamics.

Type (\mathcal{S}) $N = 20, M = \pi(N) = 8$

```
R=      47.926558330590  ICOUNT:  265  MAX:  400
FOR X=   0.1280000000  Y=   0.7195000000
and initial values:
C0( 1)=  1.000000000000
C0( 2)=  0.511058000000
C0( 3)=  1.700681000000
C0( 5)= -1.358583000000
C0( 7)=  0.028926000000
C0(11)= -0.295543000000
C0(13)= -0.087480000000
C0(17)=  0.000000000000
C0(19)=  0.000000000000
```

crude data
from [14]

(with *attractive* behavior throughout
as measured by the successive differences)

| | | | |
|--------|-----------------|-----------------|-------------|
| C(1)= | 1.000000000000 | 1.000000000000 | 0.0000E+00 |
| C(2)= | 0.511058433834 | 0.511058433833 | 0.1208E-12 |
| C(3)= | 1.700681262519 | 1.700681262519 | 0.7816E-13 |
| C(5)= | -1.358582751399 | -1.358582751399 | 0.6537E-12 |
| C(7)= | 0.028925945761 | 0.028925945761 | -0.2591E-12 |
| C(11)= | -0.295544126038 | -0.295544126039 | 0.5276E-12 |
| C(13)= | -0.087366865379 | -0.087366865379 | 0.3979E-12 |
| C(17)= | 0.937817000442 | 0.937817000442 | 0.2878E-12 |
| C(19)= | 1.795503232584 | 1.795503232584 | 0.2061E-12 |

(2.6) denominator was: $-.4314E00$
final value of the (2.5) analog: $5.6E-11$
 T_0 -indicator: $6.5E-13$

Even when the iteration count was only about 100, one was finding T_0 -indicators and (2.5) values of about 10^{-8} . There is little doubt that, for this test point, the φ -data is T_0 -attractive.

Surely, then, running (\mathcal{S}^*) should produce the same answer. Not so!

Type (\mathcal{S}^*)

```
R=      47.926558330590  ICOUNT:  800  MAX:  800
FOR X=   0.1280000000  Y=   0.7195000000
and initial values:
C0( 1)=  1.000000000000
C0( 2)=  0.511058000000
C0( 3)=  1.700681000000
C0( 5)= -1.358583000000
C0( 7)=  0.028926000000
C0(11)= -0.295543000000
C0(13)= -0.087480000000
C0(17)=  0.000000000000
C0(19)=  0.000000000000
```

(with repellent behavior before $I \approx 50$
 and attractive behavior after $I \approx 120$
 as measured by the successive differences)

| | | | |
|---------|-----------------|-----------------|----------|
| C(1)= | 1.000000000000 | 1.000000000000 | .000E+00 |
| C(2)= | 0.376415454703 | 0.885343665274 | .533E-14 |
| C(3)= | 1.188036394793 | 1.337190554675 | .000E+00 |
| C(5)= | -0.820559895739 | -1.797243672303 | .355E-14 |
| C(7)= | 0.215736965622 | 0.215736965622 | .799E-14 |
| C(11)= | -1.168200588376 | -1.168200588376 | .711E-14 |
| C(13)= | -0.360588849565 | -0.360588849565 | .355E-14 |
| C(17)= | 0.994237766499 | 0.994237766499 | .142E-13 |
| C(19)= | 1.247090592260 | 1.247090592260 | .711E-14 |

(2.6) denominator was: $-.3023E00$
 largest successive difference: about 10^{-7} at $I = 200$
 about 10^{-12} at $I = 300$

In this batch, we have changed the format so that the third column displays the absolute value of the last *successive* difference. The middle column continues to be the T_0 -image of the first. As the numbers show, things have clearly stabilized (with respect to T) but *not* at a T_0 -fixed point.

So how did \mathcal{S}^* escape from φ ? The answer would seem to be that, for at least one choice among the various quotients in (4.4), etc., the φ -data is a *repellent* fixed point of rather strong intensity.

Going over to \mathcal{S}^{**} [with $q = 13$ and $\delta = 10^{-3}$] did not help. This yielded only a random wandering within a small neighborhood of the true φ -data; hence a divergence. (Note, too, the relevance of the *third* preliminary remark.)

We should also mention that, once $c_0(p)$ started to differ appreciably from the true φ -data for at least one small p , we were *unable* to come up with a single instance of T_0 -convergence (after 10^4 iterations) in over 50 test cases. This illustrates point (c) in §4.¹³

By switching over to $X = .234$, $Y = .7195$ and modifying the vector $(c_0(p))$ with appropriate \pm signs, it is possible to produce at least two fixed points for T_0 *in addition to* the fixed point at the φ -data (cf. [16]). The former are presumably attractive; the latter repellent.

In one of the attractive cases, Ramanujan-Petersson is clearly violated. (The behavior reported in [35, pp. 266 (bottom), 268] thus seems to be *atypical*.)

Example 3. The following \mathcal{S} -data came from some jobs we ran to investigate Stark's assertion [35, p. 266 (lines 9–11)] that, for small R , it is possible to start literally from nothing.

¹³Each test run took about 28 seconds on the CRAY-2.

$$N = 20, M = \pi(N) = 8$$

```

R=      13.7797513518907  ICOUNT:   11  MAX:   100
FOR X=   0.10000000000  Y=   1.00500000000
and initial values:
C0( 1)=   1.000000000000
C0( 2)=   0.000000000000
C0( 3)=   0.000000000000
C0( 5)=   0.000000000000
C0( 7)=   0.000000000000
C0(11)=   0.000000000000
C0(13)=   0.000000000000
C0(17)=   0.000000000000
C0(19)=   0.000000000000

```

(with rapidly attractive behavior throughout)

| | | | |
|--------|-----------------|-----------------|------------|
| C(1)= | 1.000000000000 | 1.000000000000 | 0.0000E+00 |
| C(2)= | 1.549304477941 | 1.549304477941 | 0.0000E+00 |
| C(3)= | 0.246899772454 | 0.246899772454 | 0.0000E+00 |
| C(5)= | 0.737060385348 | 0.737060385348 | 0.0000E+00 |
| C(7)= | -0.261420075765 | -0.261420075765 | 0.0000E+00 |
| C(11)= | -0.953564652617 | -0.953564652617 | 0.0000E+00 |
| C(13)= | 0.278827029162 | 0.278827029162 | 0.0000E+00 |
| C(17)= | 1.307341714533 | 1.307341714533 | 0.0000E+00 |
| C(19)= | 0.092558582508 | 0.092558582508 | 0.0000E+00 |

(2.6) denominator was: .9699E00
final value of the (2.5) analog: 0

```

R=      13.7797513518907  ICOUNT:    9  MAX:   100
FOR X=   0.15000000000  Y=   1.00500000000
and initial values:
C0( 1)=   1.000000000000
C0( 2)=   0.000000000000
C0( 3)=   0.000000000000
C0( 5)=   0.000000000000
C0( 7)=   0.000000000000
C0(11)=   0.000000000000
C0(13)=   0.000000000000
C0(17)=   0.000000000000
C0(19)=   0.000000000000

```

(with rapidly attractive behavior throughout)

| | | | |
|--------|-----------------|-----------------|------------|
| C(1)= | 1.000000000000 | 1.000000000000 | 0.0000E+00 |
| C(2)= | -1.122947942450 | -1.122947942450 | 0.0000E+00 |
| C(3)= | -0.389949016161 | -0.389949016161 | 0.0000E+00 |
| C(5)= | -0.523189739313 | -0.523189739313 | 0.0000E+00 |
| C(7)= | 0.507684269216 | 0.507684269216 | 0.0000E+00 |
| C(11)= | 0.609969290410 | 0.609969290410 | 0.0000E+00 |
| C(13)= | 0.044620901213 | 0.044620901213 | 0.0000E+00 |
| C(17)= | -0.049027146897 | -0.049027146897 | 0.0000E+00 |
| C(19)= | 0.637093574481 | 0.637093574481 | 0.0000E+00 |

(2.6) denominator was: .7151E00

```

R=      13.7797513518907  ICOUNT:  800  MAX:  800
FOR X=   0.1200000000  Y=   0.6000000000
and initial values:
C0( 1)=  1.000000000000
C0( 2)=  0.000000000000
C0( 3)=  0.000000000000
C0( 5)=  0.000000000000
C0( 7)=  0.000000000000
C0(11)=  0.000000000000
C0(13)=  0.000000000000
C0(17)=  0.000000000000
C0(19)=  0.000000000000

```

(with repellent behavior before $I \approx 20$)
then steadily attractive afterwards)

| | | | |
|--------|------------------|------------------|-------------|
| C(1)= | 1.000000000000 | 1.000000000000 | 0.0000E+00 |
| C(2)= | -35.836533862512 | -35.836533862514 | 0.1819E-11 |
| C(3)= | -36.625899101444 | -36.625899101446 | 0.2046E-11 |
| C(5)= | -73.310502980020 | -73.310502980024 | 0.4093E-11 |
| C(7)= | -31.181855497617 | -31.181855497619 | 0.1705E-11 |
| C(11)= | -1.099116370964 | -1.099116370963 | -0.2629E-12 |
| C(13)= | -42.913663228011 | -42.913663228013 | 0.2728E-11 |
| C(17)= | 77.526863693281 | 77.526863693285 | -0.4547E-11 |
| C(19)= | 18.904232881022 | 18.904232881023 | -0.9095E-12 |

(2.6) denominator was: .1690E00

```

R=      13.7797513518907  ICOUNT:  800  MAX:  800
FOR X=   0.1250000000  Y=   0.6000000000
and initial values:
C0( 1)=  1.000000000000
C0( 2)=  0.000000000000
C0( 3)=  0.000000000000
C0( 5)=  0.000000000000
C0( 7)=  0.000000000000
C0(11)=  0.000000000000
C0(13)=  0.000000000000
C0(17)=  0.000000000000
C0(19)=  0.000000000000

```

(with attractive behavior throughout)

| | | | |
|--------|-----------------|-----------------|-------------|
| C(1)= | 1.000000000000 | 1.000000000000 | 0.0000E+00 |
| C(2)= | 1.549304477941 | 1.549304477941 | -0.6963E-12 |
| C(3)= | 0.246899772454 | 0.246899772455 | -0.9255E-12 |
| C(5)= | 0.737060385349 | 0.737060385351 | -0.1577E-11 |
| C(7)= | -0.261420075764 | -0.261420075763 | -0.8278E-12 |
| C(11)= | -0.953564652617 | -0.953564652617 | 0.2132E-13 |
| C(13)= | 0.278827029166 | 0.278827029167 | -0.9592E-12 |
| C(17)= | 1.307341714536 | 1.307341714534 | 0.1734E-11 |
| C(19)= | 0.092558582508 | 0.092558582508 | 0.4245E-12 |

(2.6) denominator was: -.1467E00
final value of the (2.5) analog: 4E-12

In batches 3 and 4, note that the test points are rather close together. Taking $X = .124$ gave something which, after 800 iterations, was trying to converge to the φ -data but was still off by up to $5E-2$. [Not surprisingly perhaps, the convergence in batch 4 was a bit on the slow side: the (2.5) analog had values $4.8E-2$ and $6.8E-8$ at $I = 100, 400$. Something like $E-11$ was achieved only around $I = 600$.]

To get a better handle on this type of situation, it is tempting to make a series of further runs (with, say, $ICOUNT = 25000$) at a large number of intermediate X .

Let case 1 refer to $Y = .6000$ and case 2 to $Y = 1.005$. In *both* cases, one easily checks that $\varphi(x + iY)$ has constant sign on the given x -interval. This is important because it shows that whatever subtleties are present are not merely caused by passing through a zero of φ .

By running jobs with a variety of input vectors, one quickly confirms that the φ -data goes from being T_0 -repellent at one endpoint to T_0 -attractive at the other.

The transition points occur at about $.1238$ and $.12877$, respectively.

The iterative behavior near these X should prove interesting (if only for the fact that one of the eigenvalues of $J_{T_0}(\varphi)$ now has modulus very close to 1).

Taking $X = .1238$ and $c_0(p) \equiv 0$ in case 1 produces output which shows no sign of convergence even after 500000 iterations. Yet, in taking $X = .1238 + (.0001)n$, $n \neq 0$, we find output which is already quite stable at $I = 25000$. The resulting limit vectors continue to display the same lopsided behavior seen in batches 3 and 4. Given the size of the "gap", there is little doubt that *someplace* near $X = .1238$, a point of true divergence must occur (for this choice of $c_0(p)$).

The dynamics in case 2 were significantly different. As X approached $.12877$, the matrix $J_{T_0}(\varphi) - I$ seemed to become singular, or else very nearly so (in marked contrast to case 1). In line with some standard structure theorems in several complex variables, the *zero-set* of $T_0(c) - c$ must (therefore) be expected to exhibit a kind of branching behavior as X passes through $.12877$. Our experiments with $I = 25000$ tended to confirm this—in that, what was found, was a kind of *criss-cross* in the attractive and repellent fixed points (as opposed to an actual jump).

Example 4. Another job of type (\mathcal{S}) . In this case, R is (obviously) not an eigenvalue.

$$N = 20, \quad M = \pi(N) = 8$$

```

R=          5.000000000000  ICOUNT:   19  MAX:   400
FOR X=      0.1890000000  Y=      0.8650000000
and initial values:
C0( 1)=     1.000000000000
C0( 2)=     0.000000000000
C0( 3)=     0.000000000000
C0( 5)=     0.000000000000
C0( 7)=     0.000000000000
C0(11)=     0.000000000000
C0(13)=     0.000000000000
C0(17)=     0.000000000000
C0(19)=     0.000000000000

```

(with rapidly attractive behavior throughout)

| | | | |
|---------|-----------------|-----------------|------------|
| C(1)= | 1.000000000000 | 1.000000000000 | 0.0000E+00 |
| C(2)= | 3.980104536067 | 3.980104536067 | 0.0000E+00 |
| C(3)= | -4.353479845878 | -4.353479845878 | 0.0000E+00 |
| C(5)= | 3.747186760026 | 3.747186760026 | 0.0000E+00 |
| C(7)= | -2.495431559510 | -2.495431559510 | 0.0000E+00 |
| C(11)= | 0.424977267960 | 0.424977267960 | 0.0000E+00 |
| C(13)= | 3.269324551156 | 3.269324551156 | 0.0000E+00 |
| C(17)= | -4.379743415131 | -4.379743415131 | 0.0000E+00 |
| C(19)= | -8.749979373183 | -8.749979373183 | 0.0000E+00 |

the (2.6) denominator was: .1060E-1

Example 5 [illustrating catastrophic cancellation]. In the following \mathcal{S} -job, the 4th column displays the effect of placing *absolute value* signs on all the various harmonics [from (1.3)] which enter into the (2.6) numerator.

The input data is shown in column 1; the result of *one* iteration of (2.6) in column 2; the difference 1–2 in column 3. Note that the input data is accurate to about 12 places.

$$N = 9, \quad M = \pi(N) = 4$$

| | | | | |
|---------|------------------|-----------------|--------------|------------|
| R= | 13.7797513518907 | ICOUNT: 1 | | |
| FOR X= | 0.0000000000 | Y= | 0.2000000000 | |
| C(1)= | 1.000000000000 | 1.000000000000 | 0.0000E+00 | |
| C(2)= | 1.549304477942 | 1.549304494482 | -0.1654E-07 | 0.5010E+00 |
| C(3)= | 0.246899772454 | 0.246899868708 | -0.9625E-07 | 0.7784E+00 |
| C(5)= | 0.737060385348 | 0.737061249696 | -0.8643E-06 | 0.2914E+01 |
| C(7)= | -0.261420075765 | -0.261419800439 | -0.2753E-06 | 0.1692E+01 |
| C(11)= | | -0.953563663742 | | 0.2037E+01 |
| C(13)= | | 0.278826741784 | | 0.1941E+01 |
| C(17)= | | 1.307341713868 | | 0.1805E+01 |
| C(19)= | | 0.092558532888 | | 0.1778E+01 |
| C(23)= | | 1.138070636669 | | 0.1629E+01 |
| C(29)= | | 0.752113320290 | | 0.1859E+01 |

(2.6) denominator was: .1385E-5

typical c_p -accuracy in output: about 10^{-6}

The reason for the (accuracy) degradation lies in the size of the denominator. In terms of floating-point arithmetic in (2.6), the earliest [i.e., “lowest”] harmonics have magnitude about $O(1)$ and a precision of (say) 11 digits. When the denominator is about 10^{-6} (as here), in order for c_p to be sensibly bounded, this makes the *numerator* in (2.6) have to have value like $O(1)10^{-6}$. But, then, the first *five* digits in the earliest harmonics must cancel out exactly (under addition). The number of significant figures in c_p should therefore be about $11 - 5 = 6$. This of course agrees with what we found.

Taking $z_0 = i(.125)$ produces a denominator of $2.8\mathrm{E}-14$, and essentially *no* significant figures in c_p . Similar behavior would ensue anytime z_0 were accidentally positioned along a nodal line of φ .

6. CONCLUDING REMARKS

Any number of worthwhile problems have been raised (either directly or indirectly) in the preceding sections.

Before closing, we only want to draw attention to one idea which may have already occurred to some readers.

Namely, recall (2.7)–(2.8). One would very much like to know where the *zeros* of $G(s)L(s)$ lie; in particular, are they all located along $\mathrm{Re}(s) = \frac{1}{2}$? In this connection, it would also be nice to know their spacing distribution (cf. [13]).

Having the ability to calculate many c_p -values is an important prerequisite to being able to explore problems of this kind experimentally. (Indeed, to being able to simply compute $L(s)$!)

In matters of this type, cases with large $\mathrm{Im}(s)$ will typically be most important. To reach the necessary heights, one definitely needs to know many c_p [probably out to $p \approx (\text{constant})\mathrm{Im}(s)$] and *know* them to *good* accuracy.

From this standpoint, then, the problem raised at the end of §3 takes on a new significance.

Whether it will ultimately prove to be possible to obtain enough c_p to study $L(s)$ at heights like 10^6 is by no means clear. [Similar remarks apply to the more general functions $L(s, \mathrm{Sym}^k)$ mentioned in §1 (and to an even greater degree!).] See [6, 8, 9, 13, 17] for the proper perspective.

BIBLIOGRAPHY

1. L. Arnold, *On the asymptotic distribution of the eigenvalues of random matrices*, J. Math. Anal. Appl. **20** (1967), 262–268.
2. —, *Deterministic version of Wigner's semicircle law for the distribution of matrix eigenvalues*, Linear Algebra Appl. **13** (1976), 185–199.
3. B. J. Birch, *How the number of points of an elliptic curve over a fixed prime field varies*, J. London Math. Soc. **43** (1968), 57–60.
4. D. Bump, *The Rankin-Selberg method: a survey*, Number Theory, Trace Formulas, and Discrete Groups (K. Aubert, E. Bombieri, and D. Goldfeld, eds.), Academic Press, New York, 1989, pp. 49–109; especially pp. 57 (middle)–59 (top).
5. A. Erdélyi et al., *Higher transcendental functions*, vol. 2, McGraw-Hill, New York, 1953, pp. 86–88; especially pp. 86(7), 87(18), 88(19)(20).
6. H. R. P. Ferguson et al., *On zeros of Mellin transforms of $\mathrm{SL}_2(\mathbb{Z})$ cusp forms*, Math. Comp. **42** (1984), 241–255.
7. S. Gelbart and F. Shahidi, *Analytic properties of automorphic L-functions*, Academic Press, New York, 1988; especially pp. 65–67, 85, 107–109.
8. A. Good, *Approximative Funktionalgleichungen und Mittelwertsätze für Dirichletreihen, die Spitzenformen assoziiert sind*, Comment. Math. Helv. **50** (1975), 327–361.
9. G. Guthmann, *Asymptotische Entwicklungen für unvollständige Gammafunktionen*, Forum Math. **3** (1991), 105–141.
10. E. Hecke, *Mathematische Werke*, Vandenhoeck and Ruprecht, Göttingen, 1959; especially pp. 268–275, 220, 228, 230.
11. D. A. Hejhal, *The Selberg trace formula for $\mathrm{PSL}(2, \mathbb{R})$* , vol. 1, Lecture Notes in Math., vol. 548, Springer-Verlag, New York, 1976.

12. —, *The Selberg trace formula for $\mathrm{PSL}(2, \mathbb{R})$* , vol. 2, Lecture Notes in Math., vol. 1001, Springer-Verlag, New York, 1983.
13. —, *Zeros of Epstein zeta functions and supercomputers*, Proc. Internat. Congr. Math. (Berkeley 1986), pp. 1362–1384.
14. —, *Eigenvalues of the Laplacian for $\mathrm{PSL}(2, \mathbb{Z})$: some new results and computational techniques*, International Symposium in Memory of Hua Loo-Keng, vol. 1 (S. Gong, Q. Lu, Y. Wang, and L. Yang, eds.), Science Press and Springer-Verlag, New York, 1991, pp. 59–102.
15. —, *Eigenvalues of the Laplacian for Hecke triangle groups*, Mem. Amer. Math. Soc. **469** (1992).
16. D. A. Hejhal and S. Arno, *On Fourier coefficients of Maass waveforms for $\mathrm{PSL}(2, \mathbb{Z})$* , Minnesota Supercomputer Institute Research Report No. 1992/20, 72 pp.
17. J. Lagarias and A. M. Odlyzko, *On computing Artin L -functions in the critical strip*, Math. Comp. **33** (1979), 1081–1095.
18. R. P. Langlands, *Problems in the theory of automorphic functions*, Lecture Notes in Math., vol. 170, Springer-Verlag, New York, 1970, pp. 18–61; especially §8.
19. D. H. Lehmer, *Note on the distribution of Ramanujan's tau function*, Math. Comp. **24** (1970), 741–743.
20. H. Maass, *Über eine neue Art von nichtanalytischen automorphen Funktionen und die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen*, Math. Ann. **121** (1949), 141–183.
21. M. L. Mehta, *Random matrices*, Academic Press, New York, 1967; especially pp. 45, 59, 76, 240.
22. —, *Random matrices*, rev. ed., Academic Press, New York, 1991; especially pp. 9, 75, 78, 93, 135, 401.
23. T. Miyake, *Modular forms*, Springer-Verlag, New York, 1989; especially pp. 145–149, 142(26), 143(27).
24. C. Moreno and F. Shahidi, *The L -functions $L(s, \mathrm{Sym}^m(r), \pi)$* , Canad. Math. Bull. **28** (1985), 405–410.
25. M. R. Murty, *Oscillations of Fourier coefficients of modular forms*, Math. Ann. **262** (1983), 431–446.
26. —, *On the estimation of eigenvalues of Hecke operators*, Rocky Mountain J. Math. **15** (1985), 521–533.
27. —, *The Ramanujan τ -function*, Ramanujan Revisited (G. Andrews, R. Askey, B. Berndt, K. Ramanathan, and R. Rankin, eds.), Academic Press, New York, 1988, pp. 269–288.
28. V. K. Murty, *On the Sato-Tate conjecture*, Number Theory Related to Fermat's Last Theorem (N. Koblitz, ed.), Birkhäuser, Basel and New York, 1982, pp. 195–205.
29. C. Porter, ed., *Statistical theories of spectra: fluctuations*, Academic Press, New York, 1965; especially pp. 153, 154, 179, 197, 227.
30. R. A. Rankin, *Modular forms and functions*, Cambridge Univ. Press, London and New York, 1977.
31. P. Sarnak, *Statistical properties of eigenvalues of the Hecke operators*, Analytic Number Theory and Diophantine Problems (A. Adolphson, J. B. Conrey, A. Ghosh, and R. Yager, eds.), Birkhäuser, Basel and New York, 1987, pp. 321–331; especially page 324.
32. A. Selberg, *Collected papers*, vol. 1, Springer-Verlag, New York, 1989, pp. 626–674.
33. J.-P. Serre, *Abelian l -adic representations and elliptic curves*, Benjamin, New York, 1968, pp. 118–129.
34. G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Princeton Univ. Press, Princeton, NJ, 1971; especially pp. 79–80.
35. H. Stark, *Fourier coefficients of Maass waveforms*, Modular Forms (R. A. Rankin, ed.), Ellis Horwood, 1984, pp. 263–269.

36. G. Szegő, *Orthogonal polynomials*, rev. ed., Amer. Math. Soc. Colloq. Publ., No. 23, Amer. Math. Soc., Providence, RI, 1959; especially pp. 199(8.22.12), 139–143.
37. J. Tate, *Algebraic cycles and poles of zeta functions*, *Arithmetical Algebraic Geometry* (O. F. G. Schilling, ed.), Harper and Row, New York, 1965, pp. 93–110; especially pp. 106–107.
38. G. N. Watson, *A treatise on the theory of Bessel functions*, 2nd ed., Cambridge Univ. Press, London and New York, 1944.

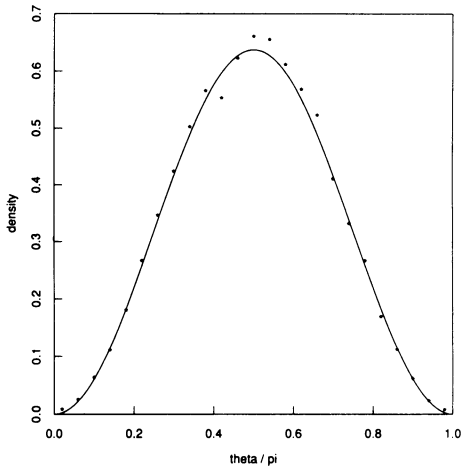
SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MINNESOTA 55455
E-mail address: hejhal@math.umn.edu

SUPERCOMPUTING RESEARCH CENTER, 17100 SCIENCE DRIVE, BOWIE, MARYLAND 20715
E-mail address: arno@super.org

**Supplement to
ON FOURIER COEFFICIENTS OF MAASS WAVEFORMS
FOR $PSL(2, \mathbb{Z})$**

D. A. HEJHAL AND S. ARNO

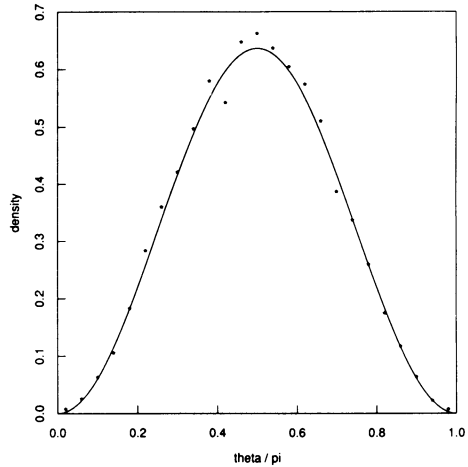
aggregation of all data (scatterplot) and Sato-Tate (curve)



CASES A - L (14188 coefficients)

Fig. 1

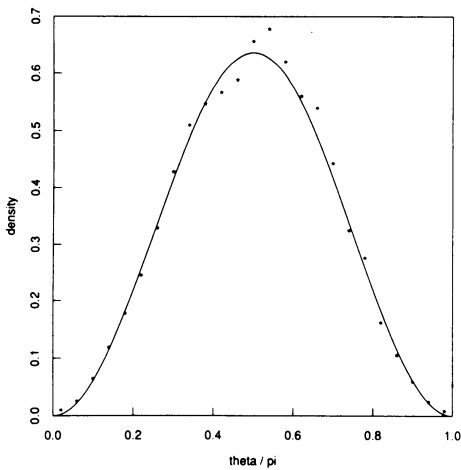
aggregation of all even data (scatterplot)
and Sato-Tate (curve)



CASES A - G (8043 coefficients)

Fig. 2

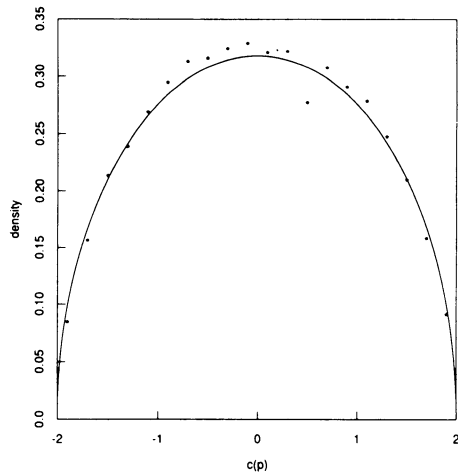
aggregation of all odd data (scatterplot)
and Sato-Tate (curve)



CASES H - L (6145 coefficients)

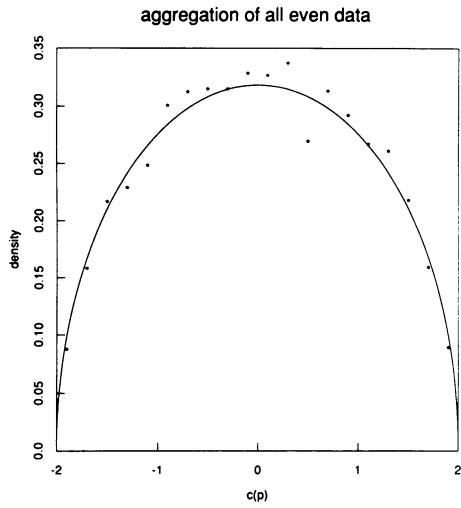
Fig. 3

aggregation of all data



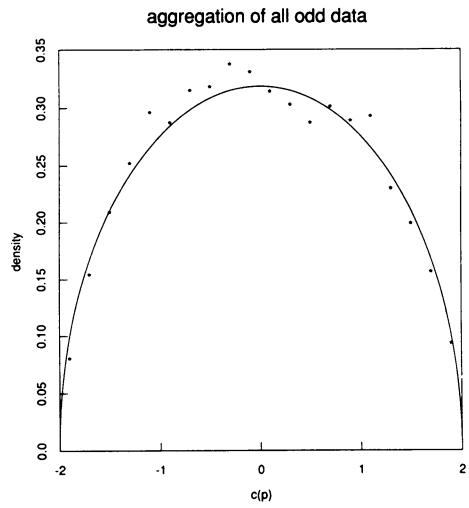
CASES A - L (14188 coefficients)

Fig. 4



CASES A – G (8043 coefficients)

Fig. 5



CASES H – L (6145 coefficients)

Fig. 6

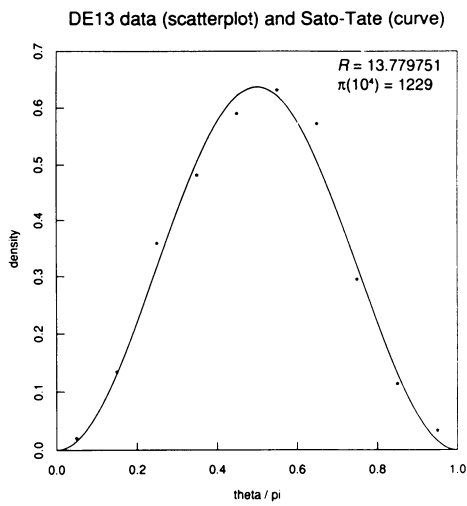


Fig. 7A

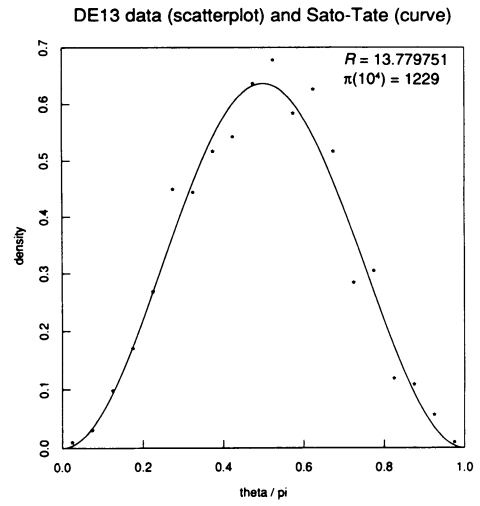


Fig. 7B

DE17 data (scatterplot) and Sato-Tate (curve)

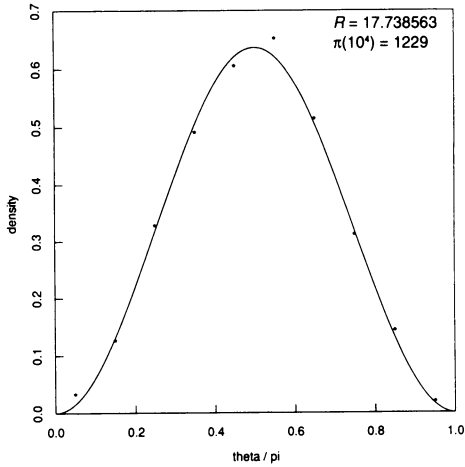


Fig. 8

DE19 data (scatterplot) and Sato-Tate (curve)

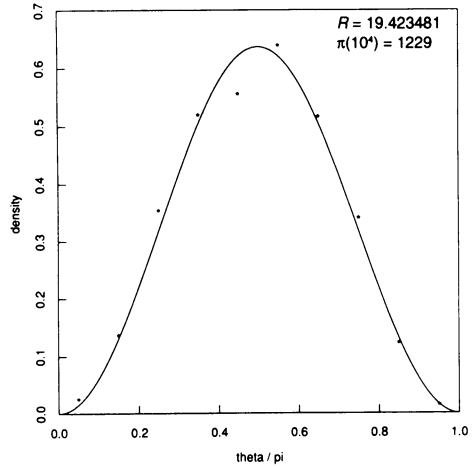


Fig. 9

DE21 data (scatterplot) and Sato-Tate (curve)

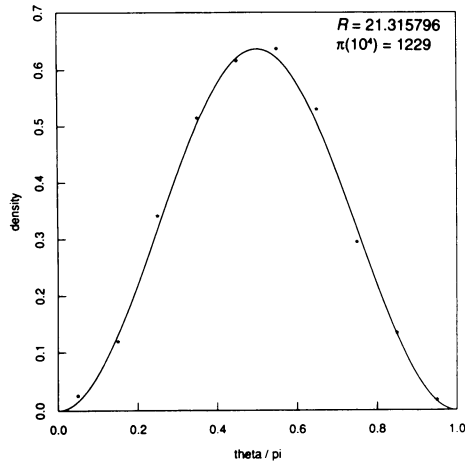


Fig. 10

DE22 data (scatterplot) and Sato-Tate (curve)

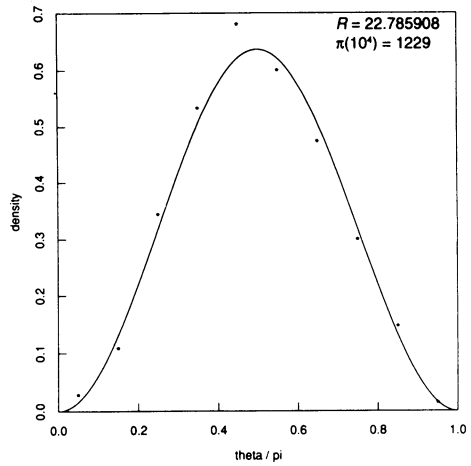


Fig. 11

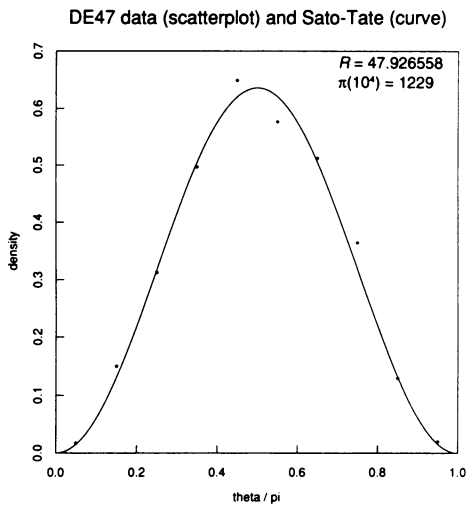


Fig. 12A

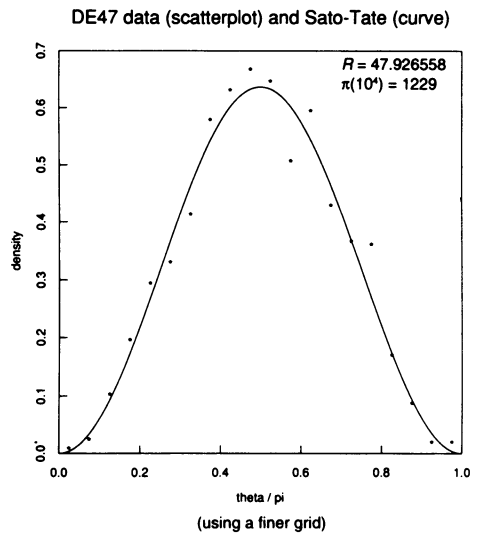


Fig. 12B

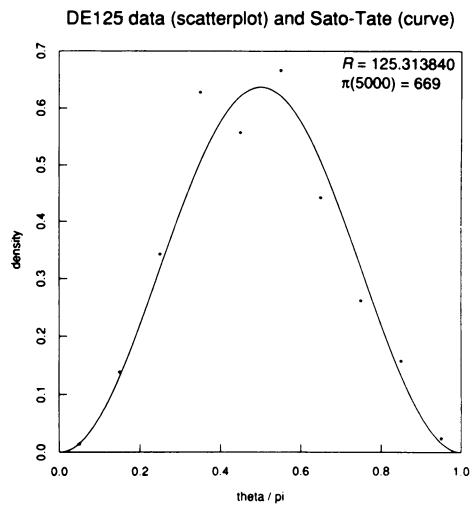


Fig. 13

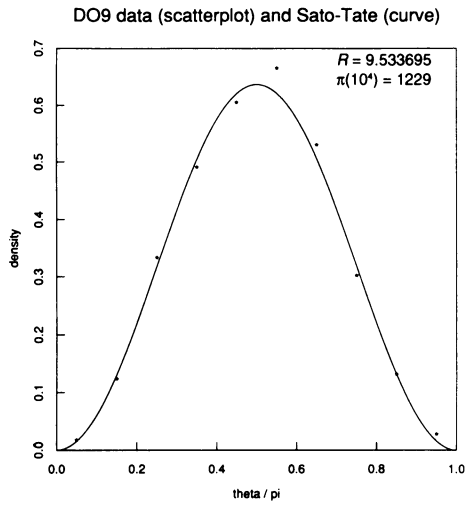


Fig. 14A

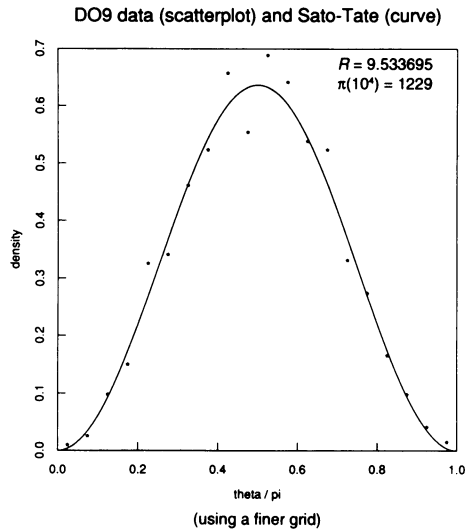


Fig. 14B

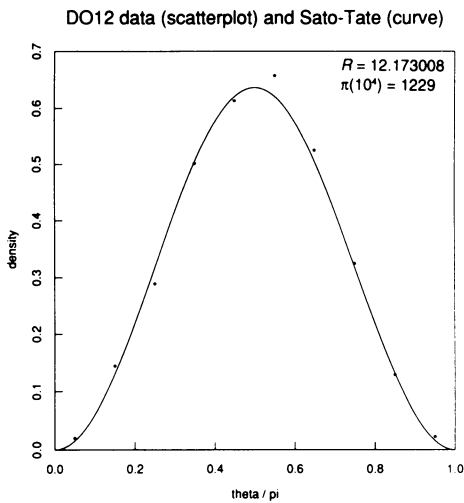


Fig. 15

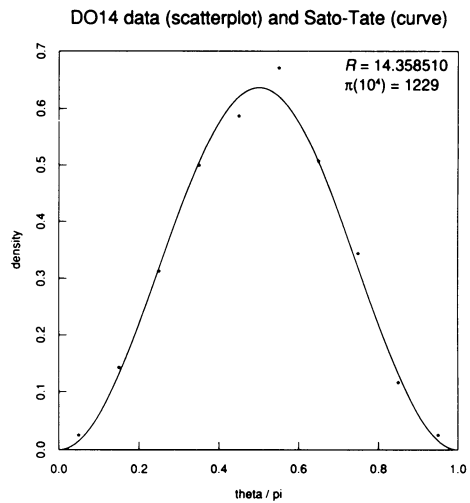


Fig. 16

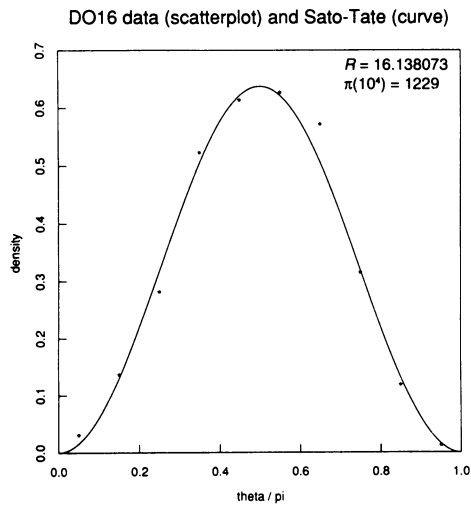


Fig. 17

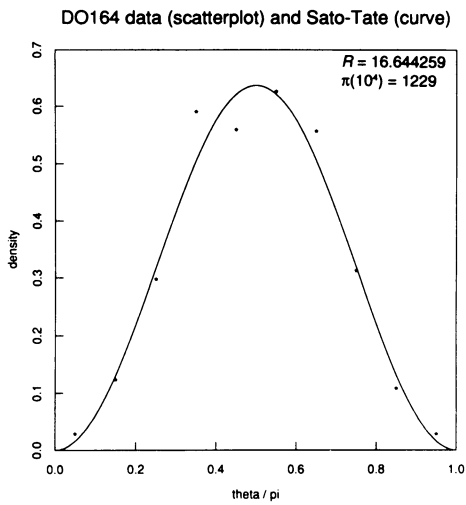


Fig. 18A

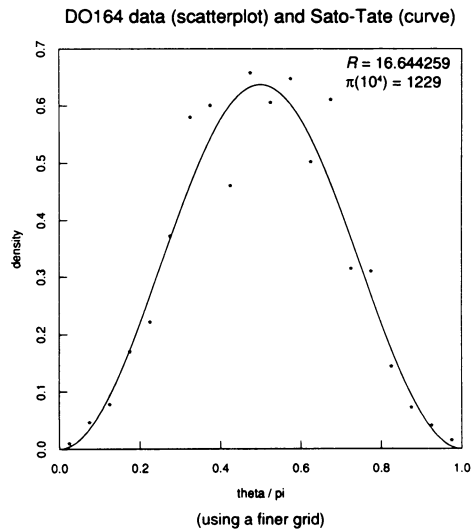


Fig. 18B