Comments on Robert Langland's Lecture: "Endoscopy and Beyond"

by

Peter Sarnak*

Letter to Robert Langland's - April, 2001

Dear Bob,

Below I have put into writing my thoughts in connection with your lecture "Endoscopy and Beyond"[†] and especially the Section called "a pipe dream". Also I have tried to digest your more recent letters and numerical experiment and I address at least some of the points that you raise in the last section. There are no doubt a number of inaccuracies below but I think most of them are good approximations to the correct statements.

First, a comment about numerical testing your key problems T_1 and P_2 (the latter in general concerns the computation of the Hasse-Weil zeta function for a variety V over a number field while the former is functoriality for higher symmetric powers). The test that Booker has been using[‡] is as follows:

Let

$$L(s,\pi) = \sum_{n=1}^{\infty} \lambda_\pi(n) n^{-s}\,. \tag{1}$$

We normalize so that $\lambda_\pi(1) = 1$ and $|\lambda_\pi(n)|$ is on average of size 1. Now if $L(s,\pi)$ satisfies the expected analytic properties then one can show that if N_π is the usual conductor of π and C_π the analytic conductor (as defined in Iwaniec-Sarnak "Perspectives on the analytic theory of L-functions" [I-S]) then the following smooth sums $S(X,\pi)$ have a certain behavior. For $\psi \in C_0^\infty(0,\infty)$ set

$$S(X,\pi) := \sum_{(n,N_\pi)=1} \lambda_\pi(n)\psi\left(\frac{n}{X}\right) \tag{2}$$

For $X \le C_\pi$ these sums are of size $\sqrt{X}$ but as X increases beyond C_π these sums drop abruptly to zero. The latter can be used as a robust test of the analytic continuation of $L(s,\pi)$ to

*Department of Mathematics, Princeton University and the Courant Institute of Mathematical Sciences.

†The complete version of this paper will appear in the *American Journal of Math.*

‡Added in 2004: These appear in his Princeton thesis 2003

the complex plane. For example, this test quickly uncovered the fake modular form proposed by Katz (see Booker "A test for identifying Fourier coefficients of automorphic forms and applications to Kloosterman sums", *Exp. Math.*, **9**, (2000), 571-581).

I understand from Booker that if $\pi = \mathrm{sym}^{10}(\Delta)$, Δ being the Ramanujan cusp form of weight 12 for $SL(2,\mathbb{Z})$, then the sums in (2) pass the test convincingly. We expect for example, that low genus curves (with perhaps large conductor) will pass the test but the fact that we can test such things at all, compels us to do so.

The rest of this letter is concerned with the pipe dream section of your lecture. The idea of using the trace formula plus some analytic theory (ie sums over primes to compute poles of L-functions) to count Galois representations is tempting. As you remark, the sums in question only involve quadratic extensions. I stick to the simplest setting, that is $GL_2/\mathbb{Q}$ and consider only π which are everywhere unramified. That is $\pi = \phi$ a Maass cusp for $\Gamma = SL(2,\mathbb{Z})$ on $\mathbb{H}^2$. Let $\lambda_\phi = \frac{1}{4} + t_\phi^2$, $t_\phi > 0$ be its Laplace eigenvalue. Denote by $L(s,\phi,\mathrm{sym}^k)$ the symmetric k^{th} power L-function corresponding to the $k+1$ dimensional representation ρ_k of ${}^LG = GL_2(\mathbb{C})$. We write the finite part of corresponding L- function as

$$L(s,\phi,\mathrm{sym}^k) := \sum_{n=1}^{\infty} \lambda_{\phi,\mathrm{sym}^k}(n) n^{-s} \,. \tag{3}$$

For $k = 1$ write simply

$$L(s,\phi) \;=\; \sum_{n=1}^{\infty} \lambda_\phi(n) n^{-s} \,. \tag{4}$$

I note for later, as you do, that

$$\lambda_{\phi,\mathrm{sym}^k}(p) \;=\; \lambda_\phi(p^k) \tag{5}$$

when p is prime.

Define $m(\phi,\mathrm{sym}^k)$ to be the order of the pole of $L(s,\phi,\mathrm{sym}^k)$ at $s = 1$. For our choice of ϕ (ie everywhere unramified), we expect that $m(\phi,\mathrm{sym}^k) = 0$ for all $k \geq 1$. The goal then is to try to establish this without appeal to the theory of higher symmetric power L-functions but rather directly from the trace formula.

To this end, fix a smooth test function h on R which decays fast enough and also extends to an entire function. Then one wants to compute ($k \geq 1$ fixed)

$$D_k(h) \;=\; \sum_{\phi} h(t_\phi) m(\phi,\mathrm{sym}^k) \tag{6}$$

as a distribution in h.

As you note, one would like to proceed directly from the limit formula (definition)

$$-m(\phi, \mathrm{sym}^k) = \lim_{M\to\infty} \frac{1}{M} \sum_{p\le M} \log p \, \lambda_{\phi,\mathrm{sym}^k}(p)$$
$$= \lim_{M\to\infty} \frac{1}{M} \sum_{p\le M} \log p \, \lambda_{\phi}(p^k) \,. \tag{7}$$

While (7) is structurally simple, being a limit of a sum over primes it is quite difficult to handle unconditionally. In fact, by forming similar (necessarily weighted) sums over integers it is shown below that by a direct use of a trace-like formula that $m(\phi, \mathrm{sym}^k) = 0$ for $k = 1$ and 2. I believe that this is roughly what Lapid was doing; though, I proceed using the Petersson-Kuznietzov formula rather than the trace formula. For $k \ge 3$, the analysis in Section B highlights where the difficulties lie and isolates the critical sums (46).

§ A. Sums over Primes

We recognize that (7) is equivalent to the non-vanishing of $L(s, \phi, \mathrm{sym}^k)$ for $\Re(s) = 1$. That is, (7) is a statement which is at the level of the prime number theorem (if we were dealing with $\zeta(s)$) and a proof of (7) without appeal to L-functions would be an achievement by itself. With (7), (6) becomes a question of evaluating

$$\lim_{M\to\infty} \frac{1}{M} \sum_{p\le M} \log p \sum_{\phi} h(t_\phi) \, \lambda_{\phi}(p^k) \,. \tag{8}$$

The idea then is to apply the trace formula to the inner sum in (8). Experience with sums that result from applying the trace formula suggests that if one can in fact estimate these sums, then one obtains a power saving. That is, we might shoot for an estimate of the form

$$S_M(h, k) := \sum_{p\le M} \log p \sum_{\phi} h(t_\phi) \, \lambda_{\phi}(p^k) \ll M^{1-\delta} \qquad \text{with } \delta > 0 \,. \tag{9}$$

This already is a very strong statement since it is essentially equivalent to a quasi-Riemann Hypothesis:

$$L(s, \phi, \mathrm{sym}^k) \neq 0 \qquad \text{for } \Re(s) > 1 - \delta \,. \tag{10}$$

while this may seem to be too ambitious, it is interesting to see what the trace formula offers. As mentioned before, rather than using the usual trace formula which requires the removal of the trivial representation and introduces class numbers, we apply the Kuznietzov formula which gets to the oscillatory terms directly. The formula is as follows: For $m \ge 1$ and $n \ge 1$

$$
\begin{aligned}
&\sum_{\phi} \omega_\phi \, h(t_\phi) \, \lambda_\phi(m) \, \overline{\lambda_\phi(n)} \\
&= \frac{\rho_{m,n}}{\pi} \int_{-\infty}^{\infty} t \, \tan h(\pi t) \, h(t) \, dt \\
&\qquad + \frac{2i}{\pi} \sum_{c=1}^{\infty} \frac{S(m,n,c)}{c} \int_{-\infty}^{\infty} J_{2ir} \left(\frac{4\pi\sqrt{mn}}{c} \right) \frac{h(r)r}{\cosh \pi r} \, dt \\
&\qquad\qquad + \text{continuous spectrum} \,.
\end{aligned} \tag{11}
$$

Here

$$\omega_\phi \;=\; \frac{1}{L(1, \mathrm{sym}^2 \phi)} \;>\; 0 \tag{12}$$

though we don't need to know this and the weight is harmless for our purposes (even when we allow ϕ to vary as we do later, since it is known that $t_\phi^\epsilon \ll \omega_\phi \ll t_\phi^\epsilon$ for any $\epsilon > 0$). Finally,

$$S(m, nc) \;=\; \sideset{}{^*}\sum_{x \bmod \mathrm{c}} e \left(\frac{mx + n\bar{x}}{c} \right) \tag{13}$$

is the Kloosterman sum.

For what follows the contributions in (11) from the continuous spectrum can be handled similarly and so I omit these terms.

For h fixed, one can think of (11) as giving an expression of the type

$$\lambda_\phi(n) \;=\; \sum_{c=1}^{\infty} \frac{S(n,1,c)}{c} \, V \left(\frac{\sqrt{n}}{c} \right) \tag{14}$$

$V \in C_0^\infty(0, \infty)$, for $\lambda_\phi(n)$. This expression is even useful for estimating $\lambda_\phi(n)$. If we use the Weil bound $S(m,n,c) \ll c^{1/2}$ we get (we have $c \approx \sqrt{n}$ because of the factor $V\left(\frac{\sqrt{n}}{c}\right)$)

$$|\lambda_\phi(n)| \ll n^{1/4} \,. \tag{15}$$

However, for our discussion the expression (14) in the form

$$\lambda_\phi(n) \;=\; \sum_{c=1}^{\infty} \frac{1}{c} \sideset{}{^*}\sum_{x \bmod \mathrm{c}} e\left(\frac{\bar{x}}{c}\right) \, e\left(\frac{xn}{c}\right) \, V \left(\frac{\sqrt{n}}{c} \right) \,. \tag{16}$$

is well suited for executing sums over n. Thus, when $k = 1$, (9) becomes

$$S_M(h,1) \;=\; \sum_{c=1}^{\infty} \frac{1}{c} \sum_{x \bmod c}^{*} e\left(\frac{\bar{x}}{c}\right) \sum_{p\le M} \log p \, e\left(\frac{px}{c}\right) V\left(\frac{\sqrt{p}}{c}\right). \tag{17}$$

Since $p \sim c^2$ (by which we mean that $1 \ll \frac{p}{c^2} \ll 1$) we break the sum over primes into progressions mod c and apply the Riemann Hypothesis for the corresponding Dirichlet L=functions $L(s,\chi)$. Keeping track of the remainder terms yields

$$S_M(h,1) \ll \sum_{c \ll \sqrt{M}} \frac{1}{c} M^{\frac{1}{2}} c^{1+\epsilon} \ll M^{1+\epsilon}.$$

That is, we recover the trivial bound for S. This indicates that some aspect of the distribution of the zeroes of $L(s,\chi)$ would do better and yield a quasi-Riemann Hypothesis (9) for $L(s,\phi)$! The exact issue concerning the distribution of the zeroes of $L(s,\chi)$ appears difficult to isolate in this setting. However, if one proceeds in a somewhat different way as Iwaniec-Luo and I do in our paper ("Low Lying zeros of families of L-functions," *Publ. IHES*, (2001), one can isolate this phenomenon. That is, we allow h to vary with M. Consider

$$I \;=\; \sum_{T \le t_\phi \le 2T} \omega_\phi \left| \sum_{p \le M} \log p \, \lambda_\phi(p) \right|^2 \tag{18}$$

(note that we really mean the interval $[T, 2T]$ and not $[0, 2T]$). From the Riemann Hypothesis for $L(s,\phi)$, we expect

$$I \;\ll\; \#\{T \le t_\phi \;\le\; 2T\} \,\cdot\, M^{1+\epsilon} \;\ll\; T^2 M^{1+\epsilon}. \tag{19}$$

Thus, if we can establish (19) with $T \le M^{\frac{1}{2}-\delta}$ for some $\delta > 0$, we would obtain (9) at least in the form

$$\sum_{p \sim M} \log p \, \lambda_\phi(p) \;\ll\; M^{1-\delta} \tag{20}$$

for $M \sim t_\phi^{2-\delta}$.

Since in (20) M is limited, it gives us a quasi-Riemann Hypothesis for zeroes of $L(s,\phi)$ near the real axis.

To analyze (18) we square out and apply Kuznietzov's formula (11) with an appropriate test function $h_T(t)$ depending on T. Analyzing the resulting integral of $h_T(t)$ against the Bessel function leads to

$$I \;=\; T^2 M \log M \;+\; T \sum_{\substack{p \le M \\ q \le M \\ c \le M/T^2}}^{\text{smooth}} \frac{\log p \log q \, S(p,q,c)}{c} \, e\left(\frac{2\sqrt{pq}}{c}\right) \;+\; \text{smaller terms} \tag{21}$$

(by a superscript smooth we mean that the sum is weighted by a smooth function of p, q and c and the ranges are as indicated). The first term on the right hand side of (21) comes from the diagonal term and it is of the order of magnitude that we expect. If we choose $M = T^{2+\delta}, \delta > 0$ (small) then the second term has c limited to being at most M^δ - ie c will be small. We are thus facing classical "GL_1" analytic sums

$$\sum_{p,q} e(2\sqrt{pq}) \log p \log q$$

that were considered by Vinogradov.

What we show in the I-L-S paper (for holomorphic ϕ's) is that estimates on such classical sums over primes imply a quasi-Riemann Hypothesis (9). Precisely, let H_∞ be the Hypothesis:

$$H_\infty : \sum_{\substack{p \le M \\ p \equiv a(c)}} \log p \, e\left(\frac{2\sqrt{p}}{c}\right) \ll M^\alpha |c|^A \text{ for some fixed } A \text{ and any } c \ge 1, M \ge 1. \tag{22}$$

Vinogradov established H_α for $\alpha = 7/8$. In I-L-S it is shown that if H_α is true for any $\alpha < 3/4$ (we expect H_α holds for any $\alpha > 1/2$) and if $L(s, \phi)$ satisfies the Riemann Hypothesis except for real zeroes then (9) is true. That is to say that standard questions about $GL(1)$ analytic sums over primes imply a quasi-Riemann Hypothesis for $L(s, \phi)$!

One can remove the assumption about the only violations of **RH** for $L(s, \phi)$ being real as follows: Consider

$$\Pi = \sum_{T \le t_\phi \le 2T} \omega_\phi \left| \sum_{m \le M} \mu(m) \lambda_\phi(m) \right|^2 \tag{23}$$

Here $\mu(m)$ is the Möbius function. Again,

$$\sum_{m \le M} \lambda_\phi(m) \, \mu(m) \ll M^{1-\delta} \tag{24}$$

is equivalent to the quasi-Riemann Hypothesis (9). Proceeding as above, one finds that

$$\Pi = T^2 M + \sum_{\substack{n \le M \\ m \le M \\ c \le M/T^2}}^{(\text{smooth})} \mu(m) \, \mu(n) \, S(m, n, c) \, e\left(\frac{2\sqrt{mn}}{c}\right) + \text{ lower order terms} \tag{25}$$

Again, we are facing GL_1 sums. Let K_α denote the hypothesis: K_α: There is an A s.t. for $c \ge 1$, $M \ge 1$

$$\sum_{n \le M, m \le M} \mu(n) \mu(m) \, e\left(\frac{2\sqrt{mn}}{c}\right) \ll (N + M)^\alpha \, (NM)^{\frac{1}{2}} |c|^A \tag{26}$$

One can determine a sharp estimate for the norm of the matrix

$$\left(e\left(\frac{2\sqrt{mn}}{c}\right)\right)_{\substack{1\le m\le M\\ 1\le n\le n}}\cdot \tag{27}$$

This is done in lecture notes of Iwaniec (Rutgers 1990) I have enclosed a copy of the relevant pages. It asserts that for any sequences $a(m)\,m\le M$, $b(n)\,n\le N$,

$$\sum_{\substack{m\le M\\ n\le N}} a(m)\overline{b(n)}\,e\left(\frac{2\sqrt{mn}}{c}\right) \ll (M+N)^{1/2}\,|\,c\,|^{1/2}\ \|\,a\,\|\cdot\|\,b\,\|\ . \tag{28}$$

Note that (28) is sharp for $c=1$ and $a(m)=b(n)=1$ for $m,n\le N$.

In particular, K_α is true for $\alpha=\frac{1}{2}$. From (25) we conclude that if K_α is true for any $\alpha<\frac{1}{2}$ (we expect it holds for any $\alpha>0$) then there is a $\delta>0$ such that for $M\sim t_\phi^{2+\delta}$

$$\sum_{m\le M}\mu(m)\,\lambda_\phi(m)\ \ll\ M^{1-\delta}\,. \tag{29}$$

The point of this discussion is that it shows that the passage via the trace formula of the sums over primes in (9), is to reduce their estimation to that of classical exponential sums as H_α and K_α. Rather than entering into an analysis of such sums over primes when $k>1$, which of course is much more problematic, we examine next the sums over integers.

§B. Sums over Integers

One can also capture the pole at $s=1$ of $L(s,\phi,\mathrm{sym}^k)$ by considering the sums

$$\sum_{n\le M}\lambda_{\phi,\mathrm{sym}^k}(n)\,. \tag{30}$$

Better still, consider the smooth sums

$$\sum_{k,\phi}(M)\ :=\ \sum_n F\left(\frac{n}{M}\right)\,\lambda_{\phi,\mathrm{sym}^k}(n) \tag{31}$$

where $F\ge 0$ smooth and supported in $(1,2)$ say.

Indeed, if $m(\phi,\mathrm{sym^k})\ge 1$ then

$$\sum_{k,\phi}(M)\ \sim\ \left(\int_0^\infty F(x)dx\right)R_\phi\,M(\log M)^{m(\phi,\mathrm{sym^k})-1},\ \text{as } M\to\infty\,. \tag{32}$$

Here $R_\phi > 0$ and is the coefficient of $(s-1)^{-m}$ of $L(s,\phi,\mathrm{sym}^k)$. If $m(\phi,\mathrm{sym}^k)=0$ then $\sum\limits_{k,\phi}(M)$ will be very small as $M \to \infty$.

For h a fixed test function as before, set

$$\sum_{h,k}(M) \;=\; \sum_{\phi} h(t_\phi)\,\omega_\phi \sum_n F\left(\frac{n}{M}\right)\,\lambda_{\phi,\mathrm{sym}^k}(n) \tag{33}$$

Consider the case $k=1$ and apply (11).

We have

$$\sum_{h,1}(M) \;=\; \sum_n F\left(\frac{n}{M}\right)\sum_{c=1}^{\infty}\frac{s(n,1,c)}{c}\,V\left(\frac{\sqrt{n}}{c}\right) \tag{34}$$

$$= \sum_{c\sim\sqrt{M}}\frac{1}{c}\sideset{}{^*}\sum_{x \bmod c} e\left(\frac{\bar{x}}{c}\right)\sum_n F\left(\frac{n}{M}\right)\,e\left(\frac{nx}{c}\right)\,V\left(\frac{\sqrt{n}}{c}\right) \tag{35}$$

Now that we are summing over integers we can apply Poisson Summation to the n-sum:

$$\sum_n F\left(\frac{n}{M}\right)\,e\left(\frac{nx}{c}\right)\,V\left(\frac{\sqrt{n}}{c}\right) \;=\; M\sum_{\nu=-\infty}^{\infty}\int_{-\infty}^{\infty}F(t)\,V\left(\frac{\sqrt{M}\sqrt{t}}{c}\right)\,e\left(tm\left(\frac{x}{c}-\nu\right)\right)c \tag{36}$$

Now $(x,c)=1$ and hence for $\nu \in \mathbb{Z}$

$$\left|\frac{x}{c}-\nu\right| \;\geq\; \frac{1}{c}\,.$$

Hence, integration by parts in (36) shows that the sum over ν is $O(M^{-B})$ for any B. Hence for any $B>0$

$$\sum_{h,1}(M) \;=\; O(M^{-B})\,. \tag{37}$$

This is much stronger than the statement that $m(\phi,\mathrm{sym}^1)=0$ for each ϕ. It essentially gives the analytic continuation of $L(s,\phi)$ to the complex plane.

One can proceed in a similar way for $k=2$ though I haven't analyzed this in complete detail. We have

$$\sum_{h,2}(M) \;=\; \sum_{c\sim M}\frac{1}{c}\sideset{}{^*}\sum_{x \bmod c} e\left(\frac{\bar{x}}{c}\right)\sum_{n=1}^{\infty}e\left(\frac{n^2x}{c}\right)\,F\left(\frac{n}{M}\right)\,V,\left(\frac{n}{x}\right)\,. \tag{38}$$

It still pays to break the n sum into progressions mod c even though they are of the same size.

$$\sum_{n} e\left(\frac{n^2 x}{c}\right) F\left(\frac{n}{M}\right) V\left(\frac{n}{c}\right) = \sum_{a(c)} e\left(\frac{a^2 x}{c}\right) \sum_{\lambda} F\left(\frac{a+\lambda c}{M}\right) V\left(\frac{a+\lambda c}{c}\right)$$

$$= \sum_{m=-\infty}^{\infty} \hat{G}_c(m) \sum_{a(c)} e\left(\frac{a^2 x}{c}\right) e\left(\frac{ma}{c}\right) \tag{39}$$

where

$$\hat{G}_c(m) = \int_{-\infty}^{\infty} F\left(\frac{c}{M}x\right) V(x)\, e(-mx)\, dx\,.$$

In particular, since in this case $c \sim M$

$$\hat{G}_c(m)\, 1\, (|m|+1)^{-A} \text{ for any } A\,. \tag{40}$$

Hence,

$$\sum_{h,2}(M) = \sum_{c\sim M}^{\text{smooth}} \frac{1}{c} \sum_{m=-\infty}^{\infty} \hat{G}_c(m) \left(\sideset{}{^*}\sum_{x(c)} \sum_{a(c)} e\left(\frac{a^2 x + ma + \bar{x}}{c}\right)\right) \tag{41}$$

Set

$$\beta_m(c) = \sideset{}{^*}\sum_{x(c)} \sum_{a(c)} e\left(\frac{a^2 x + 2ma + \bar{x}}{c}\right) \tag{42}$$

These Gauss sums may be calculated explicitly by completing the square. $\beta_m(c)$ is multiplicative in c and for c an odd prime

$$\beta_m(c) = \left(\frac{1-m^2}{c}\right) \tau_c^2 \tag{43}$$

where

$$\tau_c = \sum_{y(c)} \left(\frac{y}{c}\right) e\left(\frac{y}{c}\right)$$

is the usual Gauss sum. For each m (recall that in view of (40) m is small) $\hat{G}_c(m)$ is smoothly and slowly varying with c. Hence, the asymptotic behavior of

$$\sum_{c\sim M} \frac{1}{c} \hat{G}_c(m)\, \beta_m(c) \tag{44}$$

can explicitly be determined from (43). Thus, we can determine the asymptotic behavior as $M \to \infty$ of $\sum_{h,2}(M)$ in (41). As I said earlier, I haven't worked this out[§], but it appears that it will pick up the desired dihedral representations when it is supposed to (ie through $\beta_m(c)$ explicitly not having cancellations under certain circumstances).

[§] Added in 2004: This and much more is proven in the 2002 Princeton thesis of A. Venkatesh.

We turn to (33) when $k \geq 3$. Now we face a fundamental new difficulty. We will have to consider the sums

$$\sum_{\phi} h(t_\phi)\, \omega_\phi \sum_{n \sim M}^{\text{smooth}} \lambda_\phi(n^k) \tag{45}$$

By the analysis above, these get transformed into sums (smooth) of the shape

$$M^{-k/2} \sum_{m \sim M} \sum_{c \sim M^{k/2}} \sum_{x \bmod c}^{*} e\left(\frac{m^k x + \bar{x}}{c}\right) \tag{46}$$

The goal is to establish an estimate of the form $O(M^{1-\delta})$, $\delta > 0$ for this sum. The difficulty with treating (46) as above is that for $k \geq 3$, c is substantially larger than M. So switching the order and doing the m sum first is problematic. For example, summing for m in progressions mode is perhaps now counter-productive though for small k is worth trying. The sums

$$W \;=\; \sum_{m \sim M}^{\text{smooth}} e\left(\frac{m^k \bar{x}}{c}\right) \tag{47}$$

are classical Weyl sums and as he showed they do have cancellation

$$W \;\ll\; MC^{-1/K}\,,\, K \;=\; 2^{k-1} \tag{48}$$

However, (48) by itself is of little help here.

Returning to the sums in (46) we have to face the fact that the m sum is short compared with c. One expects that the three variable sum has "square root" of the number of terms cancellation and this would yield a bound for (46) of

$$M^{-k/2}\left(M^{k+1}\right)^{\frac{1}{2}+\epsilon} \;=\; M^{\frac{1}{2}+\epsilon} \tag{49}$$

This would do the job. One can also imagine what might happen if we allow the original ϕ's to have ramification (so as to capture the cases where $m(\phi, \mathrm{sym}^k) \geq 1$). Clearly, then the square root philosophy fails just like several variable exponential sums over finite fields. Intermediate cohomology groups (not in the middle dimension) can contribute terms larger than square root. In this case, the existence of the Artin representations with $m(\phi, \mathrm{sym}^k) \geq 1$ must be reflected in the sums (46) by some such mechanism. Still the feature that the variables c and x are much larger than m is a serious problem. Anyway, it appears that the sums in (46) are at the heart of the issue.

If we are more modest in our goal and we allow the test function h to vary so as to gain non-trivial upper bounds for the number of π in a family for which $m(\pi, \mathrm{sym}^k) \geq 1$, then progress can be made. Roughly speaking, this is what Duke is doing (via the Petersson formula) in

[*IMRN*, **2**, (1995), 99-109.] He gives non-trivial upper bounds for the number of holomorphic cusp forms of weight 1 for $\Gamma_0(N)$, with $N \to \infty$. His upper bound of $N^{11/12}$ (the trivial bound being N) is based on these forms corresponding to 2-dimensional Galios representations.

§C. Trace Formula

Since you have approached these questions with the usual trace formula, we should compare this with the use of the Petersson formula as in Section A. With the trace formula, one has to remove the contribution of the trivial representation directly (which is a non-trivial task) and one is facing class numbers rather than Kloosterman sums. I believe that in the end it doesn't matter which formula one uses. Eventually one will arrive at the same critical sums in (46). I follow your analysis (22) and (23) of Endoscopy and Beyond and (J) and (N) of M... 2000 letters. These give expressions for the quantity $\sum_\phi h(t_\phi)\lambda_\phi(N)$ in the form of a sum over n

$$\sum_{n \le \sqrt{N}} T(n,N)\,, \tag{50}$$

where

$$T(n,N) \;=\; \frac{1}{n} \sum_{r^2+s^2D=N} \frac{1}{s} \left(\frac{-D}{n}\right) \psi\left(\frac{r}{\sqrt{N}}\right) \tag{51}$$

Here I have taken only one of the two contributions to your formula (J) (in which I set $A = 1$). The second term with E_1 is similar. In this way, n in (51) is at most $N^{1/2}$ (or more precisely $N^{1/2+\epsilon}$ for any $\epsilon > 0$). We will be summing $N \sim N_0$ say (again smoothly) and restricting N to be $4p$ with p prime. Also $\psi(x)$ is of the form $\sqrt{1-x^2}\phi(x)$ with ϕ smooth of compact support in $(-1,1)$.

In sub-sums of

$$\sum_{N \sim N_0} T(n,N) \tag{52}$$

answers of order $N_0^{1-\delta}$, $\delta \ge 0$ may be neglected (as far as picking up $m(\phi, \mathrm{sym}^k) \ge 1$). So the first thing to note is that we can assume s is quite small in (51). Indeed, if we set

$$T_A(n,N) \;=\; \frac{1}{n} \sum_{\substack{r^2+s^2D=N \\ s \ge A}} \frac{1}{s} \left(\frac{-D}{n}\right) \psi\left(\frac{r}{\sqrt{N}}\right) \tag{53}$$

then

$$\sum_{N\sim N_0} |T_A(n,N)| \leq \frac{1}{n}\sum_{s\geq A}\frac{1}{s}\sum_{r^2+s^2D\ll N_0} 1$$

$$\leq \frac{1}{n}\sum_{s\geq A}\frac{1}{s}\sum_{r\ll\sqrt{N_0}}\left[\frac{N_0-r^2}{s^2}+1\right]$$

$$\leq \frac{1}{n}\sum_{s\geq A}\left(\frac{N_0^{3/2}}{s^3}+\frac{N_0^{1/2}}{s}\right)$$

$$\leq \frac{N_0^{3/2}A^{-2}}{n} \tag{54}$$

Hence, if $A \geq N_0^{1/4+\delta}$ then T_A can be neglected. Thus, we can assume that

$$\delta \leq N_0^{1/4+\delta}\,. \tag{55}$$

The case that $n = q$ an odd prime is I believe the case that you are looking at numerically. So consider for $N = 4p$ large

$$\mathrm{III} \;=\; \frac{1}{qs}\sum_{r^2+s^2D=N}\left(\frac{-D}{q}\right)\psi\left(\frac{r}{\sqrt{N}}\right) \tag{56}$$

which is essentially

$$\frac{1}{qs}\sum_{r^2\equiv N(s^2)}\left(\frac{r^2-N}{q}\right)\psi\left(\frac{r}{\sqrt{N}}\right) = \frac{1}{qs}\sum_{\substack{a \bmod (qs^2)\\ a^2\equiv N(s^2)}}\left(\frac{a^2-N}{q}\right)\sum_{r\equiv a(qs^2)}\psi\left(\frac{r}{\sqrt{N}}\right).$$

Applying Poisson Summation we obtain

$$\mathrm{III} \;=\; \frac{\sqrt{N}}{q^2s^3}\sum_{\nu\in\mathbb{Z}}\hat{\psi}\left(\frac{\sqrt{N}\nu}{qs^2}\right)\sum_{\substack{a \bmod (qs^2)\\ a^2\equiv N(s^2)}}\left(\frac{a^2-N}{q}\right)e\left(\frac{a\nu}{qs^2}\right) \tag{57}$$

Now $qs^2 \ll N^{1/2+\delta}$ (from 55) so that the ν sum is restricted to $|\nu| \ll N^{\delta}$. For $\nu = 0$ we get a term that is part of the main term that cancels with the trivial representation. that is, the explicit term

$$\frac{\hat{\psi}(0)\sqrt{N}}{q^2s^3}\sum_{\substack{a \bmod qs^2\\ a^2\equiv N(s^2)}}\left(\frac{a^2-N}{q}\right). \tag{58}$$

If $(q,s)=1$ this is

$$-\frac{\hat{\psi}(0)\sqrt{N}}{q^2s^3}\sum_{b^2\equiv N(s^2)} 1\,. \tag{59}$$

The nonzero ν contribute only when ν is small and $s^2 \approx N^{1/2}$. If $(q,s)=1$ and $1=\lambda q+\mu s^2$ then

$$\sum_{\substack{a \bmod qs^2\\ a^2\equiv N(s^2)}} \left(\frac{a^2-N}{q}\right) e\left(\frac{a\nu}{s^2q}\right) = \sum_{a(q)} \left(\frac{a^2-N}{q}\right) e\left(\frac{a\mu\nu}{q}\right)\cdot \sum_{b^2\equiv N(s^2)} e\left(\frac{b\lambda\nu}{s^2}\right)\,. \tag{60}$$

We expect that the roots b of $b^2\equiv N(s^2)$ for $s\sim\sqrt{N}$ are equidistributed. In fact, the methods in Hooley [*Mathematatika* **11** (1964), 39-49] should apply to prove this, but I haven't checked. That is to say, there is non-trivial cancellation for $\nu\neq 0$ in the sum

$$\sum_{s\sim\sqrt{N}}\ \sum_{b^2\equiv N(s^2)} e\left(\frac{b\lambda\nu}{s^2}\right)\,. \tag{61}$$

If so, this confirms that the contribution to (57) of the terms with $\nu\neq 0$ is $0\left(\frac{1}{q^2}\right)$ and hence when summing over N these will be exceptionally small.

The critical range for the sums in (53) is when n is large rather than fixed, that is $n\sim\sqrt{N}$ which is its upper limit. One instance of this is s small say $s=1$. So consider

$$R(n,N) = \frac{1}{n}\sum_{r<\sqrt{N}} \left(\frac{r^2-N}{n}\right)\psi\left(\frac{r}{\sqrt{N}}\right)\,. \tag{62}$$

Now

$$\sum_{r\in\mathbb{Z}} \left(\frac{r^2-N}{n}\right)\psi\left(\frac{r}{\sqrt{N}}\right) = \sum_{x \bmod n}\left(\frac{x^2-N}{n}\right)\sum_{\lambda\in\mathbb{Z}}\psi\left(\frac{x+\lambda n}{\sqrt{N}}\right)$$

$$= \sum_{\nu\in\mathbb{Z}}\frac{\sqrt{N}}{n}\hat{\psi}\left(\frac{\nu\sqrt{N}}{n}\right)\sum_{x \bmod n}\left(\frac{x^2-N}{n}\right)e\left(\frac{x\nu}{n}\right)\,. \tag{63}$$

Again we note that since $n\ll\sqrt{N}$ and $\hat{\psi}\in\rho(R)$ the ν sum is limited to $|\nu|\ll N^\delta$ for any $\delta>0$ (the totality of the rest being negligible). The main term comes from $\nu=0$ and this is again part of what cancels with the other main terms. Write

$$R(n,N) = R_0(n,N) + R_1(n,N) \tag{64}$$

with

$$R_0(n,N) = \frac{\sqrt{N}}{n^2}\hat{\psi}(0)\sum_{x \bmod n}\left(\frac{x^2-N}{n}\right)\,. \tag{65}$$

The inner sum mod n is multiplicative in n. Its value when $n = q$ an odd prime is easily calculated (as you do on page 17 of the 2000 letter)

$$\sum_{x(q)} \left[\left(\frac{x^2 - N}{q} \right) + 1 \right] \;=\; \sum_{x^2 - N \equiv y^2 (q)} 1 \;=\; q - 1 \text{ if } NTq\,.$$

Hence

$$\sum_{x(q)} \left(\frac{x^2 - N}{q} \right) \;=\; -1 \text{ if } NTq\,. \tag{66}$$

Thus for $n = q$ and NTq

$$T_0(n, N) = \frac{-\sqrt{N}}{n^2}\, \hat{\psi}(0)\,. \tag{67}$$

A similar expression is derived for general n and (67) is then ready for summation over $n \leq \sqrt{N}$ and $N = 4p \sim N_0$. These yield part of the term that cancels with the trivial representation.

Consider now the oscillatory terms

$$T_1(n, N) \;=\; \frac{\sqrt{N}}{n^2} \sum_{\nu \neq 0} \hat{\psi} \left(\frac{\nu \sqrt{N}}{n} \right) \left(\sum_{x \bmod n} \left(\frac{x^2 - N}{n} \right) e \left(\frac{\nu x}{n} \right) \right) \tag{68}$$

As mentioned above, the $\hat{\psi}\left(\frac{\nu\sqrt{N}}{n}\right)$ factor restricts

$$\left.\begin{array}{rl} & n \;\sim\; N^{1/2} \\ \text{and} & \\ & \nu \;\sim\; 1 \end{array}\right\} \tag{69}$$

The inner sum on n is multiplicative. If $n = q$ is an odd prime and $\nu \neq 0 \,(\bmod\, q)$ then

$$\sum_{x(q)} \left(\frac{x^2 - N}{q} \right) e \left(\frac{\nu x}{q} \right) \;=\; \sum_{x(q)} \left[1 \,+\, \left(\frac{x^2 - N}{q} \right) \right] e \left(\frac{\nu x}{q} \right)$$

$$=\; \sum_{x^2 - N \equiv y^2 (q)} e \left(\frac{\nu x}{q} \right) \;=\; \sum_{ab \equiv N (q)} e \left(\frac{\bar{2}\nu (a + b)}{q} \right)$$

$$=\; \sideset{}{^*}\sum_{a(q)} e \left(\frac{\bar{2}\nu a \,+\, \bar{2}\bar{a}\, \nu N}{q} \right) \;=\; S(\bar{2}\nu, \bar{2}\nu N, q) \tag{70}$$

the Kloosterman sum!

That is for $T_1(n, N)$ we are facing sums over $n, \nu \neq 0$, $n \sim \sqrt{N}$ and $\nu \ll N^\delta$, $\delta > 0$ of

$$\frac{1}{n} S(\bar{2}\nu, \bar{2}\nu N, n)\, \hat{\psi}\left(\frac{\nu\sqrt{n}}{n}\right) \tag{71}$$

This is pretty much the same expression that comes from the Petersson-Kuznietzov formula. Those were sums

$$\sum_{c \sim \sqrt{N}} \frac{1}{c} S(1, N, c)\, V\left(\frac{\sqrt{N}}{c}\right). \tag{72}$$

Thus we conclude that at least this part of the analysis via the trace formula coincides with the Petersson formula. I expect that with enough work one can prove the complete cancellation of all the main terms in the trace formula leaving the oscillatory terms as above. Thus, the key sums that need investigation are those in (46) or perhaps the related sums coming from (62):

$$\sum_{n \sim M^{k/2}} \frac{1}{n} \sum_{N \sim M} \sum_{x(n)} \left(\frac{x^2 - N^k}{n}\right). \tag{73}$$

I have no doubt that there is further structure to be uncovered with such sums.

Numerical experimentation, a more general investigation and a deeper understanding of these ideas would be welcome!

Best regards, Peter Sarnak

[I-S] H. Iwaniec and P. Sarnak, *GAFA* (2000), 705-741.

June 2, 2004:gpp.