SPECTRAL ASPECT SUBCONVEXITY FOR PGL_2: A COUPLE APPROACHES

Contents

Abstract. We discuss and compare two ways to estimate a short second moment of L-functions on PGL_2 in the spectral aspect: that of Iwaniec [7], using the approximate functional equation and Kuznetsov formula, and that of [12] using period integrals.

Remark 0.1. This is a lightly edited version of "log-2022-03-08-a.tex". It needs a lot of polishing, but might be useful to someone even in its current state.

1. Goal

Take

$$G := \mathrm{PGL}_2(\mathbb{R}),$$
$$\Gamma := \mathrm{PGL}_2(\mathbb{Z}) \hookrightarrow G,$$
$$H := \mathrm{GL}_1(\mathbb{R}) \cong \begin{pmatrix} * & 0 \\ 0 & 1 \end{pmatrix} \hookrightarrow G,$$
$$\Gamma_H := \mathrm{GL}_1(\mathbb{Z}) = \{\pm 1\} \hookrightarrow H.$$

Let $\pi \subset L^2_{\mathrm{cusp}}(\Gamma \backslash G)$ be a cuspidal automorphic representation corresponding to a Maass form of eigenvalue $1/4 + T^2$. We aim to show that

$$L(\pi, \tfrac{1}{2}) \ll T^{1/2-\delta} \tag{1.1}$$

for some fixed $\delta > 0$.

Such an estimate was first shown by Iwaniec [7] (initially under an additional hypothesis, which Iwaniec later observed could be removed using identity $\lambda(p)^2 - \lambda(p^2) = 1$). Iwaniec's method consists of estimating an amplified second moment over T in an interval of length a bit more than 1. This method was adapted to

PGL_3 by Blomer–Buttcane [3, 4], using an amplified fourth moment, and then to GL_n in [12] (direct link), using an amplified $2(n-1)$th moment. In the works of Iwaniec and Blomer—Buttcane, the moment is estimated using the approximate functional equation, Kuznetsov formula, and several applications of Poisson summation. In [12], the moment is estimated implicitly by direct analysis of an integral representation for the L-function, with vectors chosen as in [14] and the averaging implemented via the pretrace formula like in [8].

Remark 1.1. The best known bound for (1.1) is to due to Ivic [6] , who showed that (1.1) holds for any fixed $\delta < 1/6$. Ivic's method consists of estimating an unamplified fourth moment over T in an interval of length a bit more than $T^{1/3}$ (compare with [10, 11, 5, 2]). Such approaches using "higher moments" have no known extension beyond GL_2. It would be significant to identify such an extension.

2. Automorphic background

2.1. **Cuspidal automorphic representations.** The space $L^2_{\mathrm{cusp}}(\Gamma\backslash G)$ consists of square-integrable functions on $\Gamma\backslash G$ whose constant term vanishes. It is simultaneously a representation for the group G, acting via right translation, and the Hecke algebra, acting via left translation with respect to double cosets in $\Gamma\backslash \mathrm{PGL}_2(\mathbb{Q})/\Gamma$. Under these actions, it decomposes as a direct sum of irreducible representations, each occurring with multiplicity one. We denote by π the space of smooth vectors inside one such representation.

2.2. **Whittaker expansion.** The Hecke eigenvalues of π are described by a multiplicative function $\lambda_\pi : \mathbb{N} \to \mathbb{C}$, which specifies the eigenvalues for a spanning set of double cosets. For $\varphi \in \pi$, one defines the Whittaker function $W_\varphi : G \to \mathbb{C}$ by

$$W_\varphi(g) := \int_{x\in\mathbb{R}/\mathbb{Z}} \varphi\left(\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} g\right) e(-x)\, dx,$$

where we employ the standard abbreviation $e(t) := e^{2\pi i t}$. Conversely, φ may be recovered from λ_π and W_φ via the formula

$$\varphi(g) = \sum_{n\neq 0} \frac{\lambda_\pi(|n|)}{|n|^{1/2}} W_\varphi\left(\begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix} g\right), \tag{2.1}$$

with the sum taken over all nonzero integers n.

2.3. **Kirillov model.** The Whittaker model $\mathcal{W}(\pi)$ for π is defined to consist of all functions of the form W_φ. The theory of the Kirillov model implies that the restriction map

$$\mathcal{W}(\pi) \to \{\text{functions } H \to \mathbb{C}\}$$

is injective, and that its image contains $C_c^\infty(H)$. Combining this fact with (2.1) tells us that

- any $\varphi \in \pi$ is determined by $W_\varphi : H \to \mathbb{C}$, and
- every smooth compactly-supported function $H \to \mathbb{C}$ determines a unique $\varphi \in \pi$.

3. Integral representation

The L-function $L(\pi, s)$ may be defined for $\Re(s) > 1$ by the absolutely convergent Dirichlet series $\sum_{n\in\mathbb{N}} \lambda_\pi(n)n^{-s}$ and in general by meromorphic continuation. By unfolding (2.1), one obtains the Hecke/Jacquet–Langlands integral representation

$$\int_{\Gamma_H\backslash H} \varphi = L(\pi, \tfrac{1}{2}) \int_H W_\varphi. \tag{3.1}$$

4. Coadjoint orbits

We denote by $\mathfrak{g}$ the Lie algebra of G and by $\mathfrak{g}^*$ its linear dual. Using the trace pairing, we may identify $\mathfrak{g}^*$ with $\mathfrak{g}$, which we identify further with the space $\mathfrak{sl}_2(\mathbb{R})$ of traceless 2×2 matrices ξ. The coadjoint orbit

$$\mathcal{O}_\pi \subseteq \mathfrak{g}^*$$

attached to π is the one-sheeted hyperboloid cut out by the equation

$$\det(\xi) = -T^2.$$

(See for instance §6.2-§6.3 of Quantum variance III.) It comes with a natural symplectic volume form that describes the character of π via the Kirillov formula (see [14, §6]). We note that in the coordinates

$$\xi = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} = \begin{pmatrix} x & y-z \\ y+z & -x \end{pmatrix},$$

we have

$$\det(\xi) = -a^2 - bc = z^2 - y^2 - x^2.$$

In particular, $\mathcal{O}_\pi$ contains the circle $\{(x, y, 0) : x^2 + y^2 = T^2\}$. The circular strip $\{(x, y, z) \in \mathcal{O}_\pi : |z| \leq 1/2\}$ has symplectic volume one; in the orbit method heuristic described in [14, §1.7], it corresponds to the weight zero vector in π.

The integrals (3.1) may be understood as describing how π oscillates against the trivial character of H. The orbit method suggests [14, §1.9] that such oscillation may be understood in terms of the intersection

$$\mathcal{O}_\pi \cap \mathfrak{h}^\perp,$$

i.e., the preimage in $\mathcal{O}_\pi$ of the trivial element 0 of $\mathfrak{h}^*$. That intersection is given in (x, y, z) coordinates by

$$\{(0, y, z) : y^2 - z^2 = T^2\}$$

and in (a, b, c) coordinates by

$$\{(0, b, c) : bc = T^2\}.$$

It is a closed H-orbit, with trivial stabilizer.

We pick a point $\tau \in \mathcal{O}_\pi \cap \mathfrak{h}^\perp$ of size comparable to T. For concreteness, let us take

$$\tau = \begin{pmatrix} 0 & T \\ T & 0 \end{pmatrix}.$$

Remark 4.1. For the variant problem concerning π attached to a holomorphic form of weight $2k$, we would take $T := k - 1/2$, we would take $\mathcal{O}_\pi$ cut out by $\det(\xi) = T^2$, and we would take

$$\tau = \begin{pmatrix} 0 & -T \\ T & 0 \end{pmatrix}.$$

See for instance §8.4 of Quantum variance III.

5. Choice of vector

We seek a unit vector $\varphi \in \pi$ that is "localized at τ." Informally, this means that for each fixed Lie algebra element $X \in \mathfrak{g}$, we have

$$X\varphi = i\langle X, \tau\rangle\varphi + \mathrm{O}(T^{1/2}). \tag{5.1}$$

For further informal discussion of this concept, see [14, §1.7] and [13, §2.5] . For some precise definitions, see §7.1 of Quantum variance III, or [12, §14] (direct link: §14).

Such a vector φ may be described readily in the Kirillov model for π. Recall (from §2.3) that this model consists of the restrictions to H of elements W of the Whittaker model $\mathcal{W}(\pi)$. It will be convenient to think of such restricted elements W as functions on $\mathbb{R}^\times$ via the abbreviation

$$W(y) := W\left(\begin{pmatrix} y & 0 \\ 0 & 1 \end{pmatrix}\right).$$

We consider the following basis elements for $\mathfrak{g}$:

$$\partial_a = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \partial_b = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \partial_c = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

The action of the first two of these elements on the Kirillov model is given very simply:

$$\partial_a W(y) = 2yW'(y), \quad \partial_b W(y) = 2\pi i y W(y).$$

From now on, we will be very vague and informal about asymptotics notation. For a precise discussion, we refer to Exercise 6 of this note. Since $\langle \partial_a, \tau\rangle = 0$ and $\langle \partial_b, \tau\rangle = T$, the condition (5.1) says in particular that

$$\partial_a W_\varphi = \mathrm{O}(T^{1/2}), \quad \partial_b W_\varphi = iTW_\varphi + \mathrm{O}(T^{1/2}).$$

These formulas suggest taking the following smoothened L^2-normalized characteristic function:

$$W_\varphi(y) = T^{1/4} 1^{\mathrm{smooth}}_{T/2\pi + \mathrm{O}(T^{1/2})}(y). \tag{5.2}$$

Using that W_φ is an eigenfunction under the Casimir operator (see [12, §12], direct link: §12), one can check also that

$$\partial_c W_\varphi = iTW_\varphi + \mathrm{O}(T^{1/2}).$$

Precise forms of all these estimates are established in a general setting in [12, Part 3] (direct link: Part 3) and in this specific setting in Exercise 6 of this note.

Remark 5.1. This choice of φ is closely related to the "microlocal lift" of Zelditch [17] et al., see [9, 16, 1]. More precisely, it is (asymptotically) a G-translate of the usual definition; the limit invariance for L^2-masses will be with respect to the stabilizer G_τ of τ, namely

$$G_\tau = \left\{ \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix} \right\}, \tag{5.3}$$

rather than with respect to the diagonal subgroup.

By choosing the smooth bump function implicit in (5.2) to be nonnegative, we may arrange that

$$\int_H W_\varphi \asymp T^{-1/4},$$

so that (3.1) reads

$$T^{-1/4} L(\pi, \tfrac{1}{2}) \asymp \int_{\Gamma_H \backslash H} \varphi. \tag{5.4}$$

Our task is to show that the right hand side is $\ll T^{1/4-\delta}$.

6. Truncation

We now indicate why the integral on the right hand side of (5.4) of may be effectively truncated to a fixed compact set. For a detailed discussion of this point, see [12, §5.3] (direct link: 5.3) or [10, §5.1.4] or [15, §3] (or §3 of this informal note).

Consider, for a fixed even "truncation" function $\mathcal{T} \in C_c^\infty(H)$, the map

$$I : H \to \mathbb{C}$$

$$I(Y) := \int_{y \in \Gamma_H \backslash H} \mathcal{T}(y/Y) \varphi(y) \, d^\times y$$

assigning to a parameter Y the smoothened integral of $\varphi|_H$ over the corresponding dyadic range. We have

$$\int_{\Gamma_H \backslash H} \varphi |.|^{s-1/2} = L(\pi, s) Z(W_\varphi, s),$$

where $Z(W_\varphi, s)$ denotes the local zeta integral

$$Z(W_\varphi, s) := \int_H W_\varphi |.|^{s-1/2}.$$

For $s = \mathrm{O}(1)$, we see by explicit calculation that

$$Z(W_\varphi, s) \approx T^{s-3/4}.$$

If moreover $|\Re(s)| \leq 1/2$, then the convexity bound reads

$$L(\pi, s) \ll T^{1-s}.$$

These bounds become only polynomially worse if we relax the condition $s = \mathrm{O}(1)$ to $\Re(s) = \mathrm{O}(1)$. Multiplying them together and convolving against the rapidly-decaying Mellin transform of the fixed test function $\mathcal{T}$, we deduce the Mellin transform estimate

$$|\Re(s)| \leq 1/2 \implies \tilde{I}(s) \ll T^{1/4} (1 + |s|)^{-\infty}.$$

It follows readily that for fixed $\kappa > 0$, we incur the acceptable error $\mathrm{O}(T^{1/4-\kappa/2})$ by smoothly truncating the integral $\int_H \varphi$ to the range $\{T^{-\kappa} < |y| < T^\kappa\}$. If we seek only a qualitative subconvex bound, then we can take κ as small as we like, so there is no harm in truncating to $\{|y| = T^{o(1)}\}$. Note that the relevant truncation is specific to our choice of vector φ.

The model problem is thus to bound

$$\int_{y \in H, y \asymp 1} \varphi(y) \, d^\times y. \tag{6.1}$$

7. Symmetries

We use a convolution kernel $\omega \in C_c^\infty(G)$ to "remember" many of the symmetries satisfied by φ. (When amplifying, we really take $\omega \in C_c^\infty(\mathrm{PGL}_2(\mathbb{A}))$.) Roughly speaking, we take ω to be a character multiple of an approximate subgroup of G:

$$\omega := \mathrm{vol}(J)^{-1} 1_J^{\mathrm{smooth}} \chi_\tau^{-1},$$

where:

- J is a subset of G roughly of the shape
$$J = (1 + \mathrm{O}(T^{-\varepsilon})) \cap (G_\tau + \mathrm{O}(T^{-1/2-\varepsilon})),$$
with G_τ the stabilizer of τ, as described in (5.3).
- 1_J^{smooth} is a smoothened characteristic function of J.
- χ_τ is the "approximate character" of J attached to τ, given near the identity in exponential coordinates by
$$\chi_\tau(\exp(X)) = e^{i\langle X, \tau\rangle}.$$

We may also describe ω in terms of the function $\mathfrak{g}^* \to \mathbb{C}$ obtained by taking the Fourier transform of the pullback $\omega \circ \exp$. This function is roughly a smoothened characteristic function of a "coin-shaped" neighborhood of τ, of thickness T^ε (resp. $T^{1/2+\varepsilon}$) in directions transverse (resp. tangential) to the coadjoint orbit $\mathcal{O}_\pi$ at τ. The intersection of this neighborhood with $\mathcal{O}_\pi$ has symplectic volume $\approx T^{2\varepsilon} \approx 1$. The orbit method heuristic suggests that for an irreducible representation σ of G, we have $\sigma(\omega) \approx 0$ unless σ is a principal series representation of parameter $T + \mathrm{O}(T^\varepsilon)$, in which case $\sigma(\omega)$ is approximately a projection onto a rank $\approx T^{2\varepsilon} \approx 1$ "subspace" of vectors microlocalized at τ. In particular,

$$\pi(\omega)\varphi \approx \varphi. \tag{7.1}$$

These heuristics and definitions can be made precise, and the above approximation holds in an extremely strong sense (i.e., up to $\mathrm{O}(T^{-\infty})$ with respect to any fixed seminorm).

For further informal discussion concerning ω in a general setting, see [13, §2]. For a precise discussion in the current rank one example, see §7.1 of Quantum variance III.

8. The convexity threshold

Let's explain how to recover the convexity bound from here. Our task is to show that

$$\int_{y \in H, y \asymp 1} \pi(\omega)\varphi(y)\, d^\times y \ll T^{1/4}.$$

We view the square of the left hand side as one term arising from an integrated pretrace formula, like in the sup norm story [8]. Alternatively, we write the left hand side as the inner product of φ against a Poincaré series and apply Cauchy–Schwarz; see [13, §5.3]. Either way, we reduce to checking that

$$\int_{y_1, y_2 \asymp 1} \sum_{\gamma \in \Gamma} \omega(y_1^{-1}\gamma y_2) \ll T^{1/2}. \tag{8.1}$$

Since ω is supported on $1 + \mathrm{O}(T^{-\varepsilon})$, we see that the only γ that contribute are those in Γ_H. Combining the y_1 and y_2 integrals, it remains to check that

$$\int_{y \in H : y \asymp 1} \omega(y) \ll T^{1/2}. \tag{8.2}$$

To that end, we observe that

$$H \cap G_\tau = \{1\},$$

i.e., that no nontrivial matrices are simultaneously diagonal and of the form (5.3). (This is a baby case of the "stability" feature explained in [14, §1.9, §14].) It follows that (up to some ε's in the exponents)

$$H \cap J \subseteq \mathrm{O}(T^{-1/2}),$$

and so the volume of the integral in (8.2) is $\mathrm{O}(T^{-1/2})$.

On the other hand, the magnitude of the integrand is

$$\mathrm{vol}(J)^{-1} \approx T.$$

Indeed, J has dimensions roughly 1 along one direction and $T^{-1/2}$ along the remaining two directions.

These observations combine to give the required estimate (8.2).

Remark 8.1. It's clear in retrospect that we should have obtained such an estimate: the orbit method heuristic applied to ω suggests that the left hand side of (8.1) is a proxy for the sum of $|L(\pi, \frac{1}{2})|^2$ over T in an interval of width roughly $\mathrm{O}(T^{\varepsilon})$ (see [13, §2.3]), which is of the appropriate size for an averaged Lindelöf estimate to recover convexity.

9. Amplification

This section is just a stub; see [13, §1.5, §2.7-2.10] for details and pictures relevant for this example. To carry out the amplification step, we basically need to know that the vector φ' obtained by averaging the translates of φ under elements of H of size $\asymp 1$ satisfies a matrix coefficient estimate $\langle g\varphi', \varphi' \rangle \ll T^{-\delta}$ except when g is very close to H. This is word-for-word what happens in the sup norm problem [8], where, if we replace H with $\mathrm{SO}(2)$, then φ' becomes something like the weight zero vector.

10. Comparison with Iwaniec

The "two arguments" are "equivalent." Recall that Iwaniec [7] starts with the approximate functional equation, applies Kuznetsov, and then applies Poisson summation to both variables. The reduction to (6.1) is essentially the approximate functional equation (TODO: explain this in some detail?). In the passage to (8.1), rather than applying the pretrace formula, we could have instead applied the Fourier expansion of φ and averaged each term in the resulting double sum using Kuznetsov. A couple applications of Poisson to the geometric side of Kuznetsov would then bring us right back to (8.1).

It's more interesting to compare the generalization of this argument to GL_3 with Blomer–Buttcane [3]. The arguments are again ultimately "equivalent," but the present approach seems less miraculous.

References

[1] Nalini Anantharaman and Steve Zelditch. Patterson-Sullivan distributions and quantum ergodicity. *Ann. Henri Poincaré*, 8(2):361–426, 2007.

[2] Olga Balkanova, Dmitry Frolenkov, and Han Wu. On weyl's subconvex bound for cube-free hecke characters: Totally real case, 2022.

[3] Valentin Blomer and Jack Buttcane. On the subconvexity problem for L-functions on GL(3). *Ann. Sci. Éc. Norm. Supér. (4)*, 53(6):1441–1500, 2020.

[4] Valentin Blomer and Jack Buttcane. Subconvexity for L-functions of non-spherical cusp forms on GL(3). *Acta Arith.*, 192(1):31–62, 2020.

[5] Valentin Blomer, Subhajit Jana, and Paul D. Nelson. The Weyl bound for triple product L-functions. *Duke Math. J.*, 172(6):1173–1234, 2023.

[6] Aleksandar Ivić. On sums of Hecke series in short intervals. *J. Théor. Nombres Bordeaux*, 13(2):453–468, 2001.

[7] Henryk Iwaniec. The spectral growth of automorphic l-functions. *Journal für die reine und angewandte Mathematik*, 428:139–160, 1992.

[8] Henryk Iwaniec and Peter Sarnak. L^∞ norms of eigenfunctions of arithmetic surfaces. *Ann. of Math. (2)*, 141(2):301–320, 1995.

[9] Elon Lindenstrauss. On quantum unique ergodicity for $\Gamma\backslash\mathbb{H}\times\mathbb{H}$. *Internat. Math. Res. Notices*, (17):913–933, 2001.

[10] Philippe Michel and Akshay Venkatesh. The subconvexity problem for GL_2. *Publ. Math. Inst. Hautes Études Sci.*, (111):171–271, 2010.

[11] Paul D. Nelson. Eisenstein series and the cubic moment for PGL_2. arxiv preprint, 2019.

[12] Paul D. Nelson. Bounds for standard L-functions. *arXiv e-prints*, page arXiv:2109.15230, September 2021.

[13] Paul D. Nelson. Spectral aspect subconvex bounds for $U_{n+1}\times U_n$. *Invent. Math.*, 232(3):1273–1438, 2023.

[14] Paul D. Nelson and A. Venkatesh. The orbit method and analysis of automorphic forms. *Acta Math.*, 226(1):1–209, 2021.

[15] Raphael Schumacher. Subconvexity for $GL_3(R)$ L-Functions via Integral Representations. *arXiv e-prints*, page arXiv:2004.06791, April 2020.

[16] Lior Silberman and Akshay Venkatesh. On quantum unique ergodicity for locally symmetric spaces. *Geom. Funct. Anal.*, 17(3):960–998, 2007.

[17] Steven Zelditch. Uniform distribution of eigenfunctions on compact hyperbolic surfaces. *Duke Math. J.*, 55(4):919–941, 1987.