Large values of modular forms

NICOLAS TEMPLIER[*]

We show that there are primitive holomorphic modular forms f of arbitrary large level N such that $|f(z)| \gg N^{\frac{1}{4}}$ for some $z \in \mathfrak{H}$. Thereby we disprove a folklore conjecture that the L^∞-norm of such forms would be as small as $N^{o(1)}$.

AMS 2000 SUBJECT CLASSIFICATIONS: 11F70, 11F41.

KEYWORDS AND PHRASES: Whittaker functions, quantum chaos, automorphic forms, sup-norm, L-functions, mean value estimates.

1. Introduction

Let $N \geq 1$ and $\Gamma_0(N) \subset SL_2(\mathbb{Z})$ be the Hecke congruence subgroup of level N. Let χ be a Dirichlet character modulo N and $S_2(N, \chi)$ the vector space of holomorphic cusp forms of weight two, level N and nebentypus χ. Let $S_2^*(N, \chi)$ be the set of primitive forms

$$(1.1) \qquad f(z) = y \sum_{n=1}^{\infty} a_n n^{\frac{1}{2}} e^{2i\pi nz}, \quad z = x + iy \in \mathfrak{H}.$$

Recall that primitive forms satisfy $a_1 = 1$ and are eigenvalues of all Hecke operators T_n, $n \geq 1$, see e.g [26, §4.6]. The L-function $L(s, f) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ is entire and admits a functional equation and an Euler product.

The absolute value $|f(z)|$ is $\Gamma_0(N)$-invariant and bounded on $\mathfrak{H}$. It is of great interest to estimate the sup-norm $\|f\|_\infty$. Such estimates are closely related to quantum ergodicity and entropy bounds [6], they occur in the subconvexity problem for L-functions [12], in the distribution of zeros of modular forms [10, 11] and in the study of Arakelov invariants of $X_0(N)$ [1, 21, 24].

In [1, Thm. A] for N square-free and [20, Cor. 3.2] in general, it is proven that $\|f\|_\infty \ll_\epsilon N^{\frac{1}{2}+\epsilon}$ for all $\epsilon > 0$. In [3] it is shown that $\|f\|_\infty \ll N^{\frac{1}{2}-\delta}$ for

N square-free and $\delta < 25/914$. Further improvements in [37] and [14, 36] give $\delta < 1/6$, thus

$$(1.2) \qquad 1 \ll \|f\|_\infty \ll_\epsilon N^{\frac{1}{3}+\epsilon}.$$

1.1. A folklore conjecture. Conversely, how large can $|f(z)|$ be? A folklore conjecture is that

$$(A) \qquad \|f\|_\infty \overset{?}{\ll_\epsilon} N^\epsilon, \quad \text{for all } \epsilon > 0.$$

See e.g. [3, Eq. (5)][1] and [20, Rem. 3.3]. There have been partial evidences in favor of the estimate (A): it is true on average over an orthonormal basis of $S_2(N, \chi)$ by [24]; it would imply the Lindelöf hypothesis in the level aspect; it directly applies to old forms; and it is consistent with square-root cancellation of sums of Fourier coefficients.

1.2. Main result. We show that the estimate (A) is false:

Theorem 1.1. *There are primitive forms $f \in S_2^*(N, \chi)$ of arbitrary large level N such that $\|f\|_\infty \gg N^{\frac{1}{4}}$.*

Remarks.   (i) The multiplicative constant is absolute and indeed we shall produce an explicit value such as $\|f\|_\infty \geq (2\pi e)^{-1} N^{\frac{1}{4}}$, where $\frac{1}{2\pi e} = 0.0585498\ldots$

  (ii) Our method is specific to levels N which are not square-free and to the non-compact case. Although the lower bound seems to appear for purely analytic reason it remains an open problem whether it holds similarly for square-free levels or for compact surfaces.

(iii) One can see a similar phenomenon for Hecke eigenfunctions of the quantized cat map as shown by Olofsson [27]. In fact the same exponent $N^{\frac{1}{4}}$ occurs where N denotes the underlying quantum multiplicity. Also it is interesting to compare Proposition 2.3 below with [28] and Theorem 1.2 below with [27, Thm. 3.2].

Conditional under the GRH it is shown by Lau [23, Thm. 1] that there are primitive forms f of arbitrary large level N such that

$$\|f\|_\infty \geq \exp\left((\frac{1}{2} + o(1)) \sqrt{\frac{\log N}{\log \log N}} \right).$$

---

[1][3] record this conjecture with the two reservations that it is optimistic and that it is necessary that f be an Hecke eigenform since it fails for holomorphic Poincaré series.

The approach of [23] to this lower bound is via the Hecke integral (see (2.8) below) and Soundararajan's resonance method [35] to produce sharp omega results for central values of L-functions. There are analogous results of Milićević [25] in the eigenvalue aspect.

Other results on lower bounds for the sup-norm of automorphic forms in general may be found in Rudnick–Sarnak [29] and Lapid–Offen [22]. The mechanism there is quite different since large values arise from functorial lift from other groups. In our GL(2) situation dihedral forms could play a similar role but we don't know whether they exhibit special behavior with respect to sup-norms.

Here we shall establish the following which surprisingly shows that the primitive forms $f \in S_2^*(N, \chi)$ with large sup-norm are actually abundant:

Theorem 1.2. *For any primitive character χ (mod N) and any form $f \in S_2^*(N, \chi)$,*

$$\|f\|_\infty \gg \prod_{p^c \| N} p^{\frac{1}{2} \lfloor \frac{c}{2} \rfloor}$$

where $p^c \| N$ for a prime p and an integer $c \geq 0$ means that $p^c \mid N$ and $p^{c+1} \nmid N$.

Since χ is primitive, each form $f \in S_2^*(N, \chi)$ is twist-minimal in the sense that for any Dirichlet character η the level of the form $f \otimes \eta$ is at least N (see Lemma 3.3).

Examples. (i) If N is an even power of a prime or more generally a perfect square, then the right-hand side is equal to $N^{\frac{1}{4}}$. Thus we have $\|f\|_\infty \gg N^{\frac{1}{4}}$ for *all* primitive χ (mod N) and *all* $f \in S_2^*(N, \chi)$ if N is a square.

  (ii) If N is the third power of a square-free integer, then the right-hand side is equal to $N^{\frac{1}{6}}$. If N is square-full in the sense that $p \mid N$ implies $p^2 \mid N$ for all primes p, then the right-hand side is between $N^{\frac{1}{6}}$ and $N^{\frac{1}{4}}$.

One may now wonder what the true size of $\|f\|_\infty$ should be. A possible answer is to expect a purity phenomenon [30]: the accumulation points of the ratio $\frac{\log\|f\|_\infty}{\log N}$ could be restricted to a certain set E of exponents. Purity conjectures are put forward by Clozel and Sarnak [32] in analogy with Deligne's purity theorem in algebraic geometry. Purity can for example accommodate the exceptions to Ramanujan type bounds. The question is to determine what these accumulation points should be. What is known so far

is that $E \subset [0, \frac{1}{3}]$ by (1.2) and by Theorem 1.1 that the intersection $E \cap [\frac{1}{4}, \frac{1}{2}]$ is non-empty.

Since in Theorem 1.2 we have that $\frac{1}{2c}\lfloor\frac{c}{2}\rfloor$ takes infinitely many values below $\frac{1}{4}$, this suggests caution when speculating on the purity of $\frac{\log\|f\|_\infty}{\log N}$. For example it could be that the accumulation set E is infinite although our argument is not yet conclusive since a corresponding upper bound for $\|f\|_\infty$ is not known.

1.3. Hilbert modular forms. Our results and methods of proof generalize to automorphic forms on GL(2). Let F be a totally real number field of degree d. We show that there are Hilbert modular forms f of weight $(k_1, \ldots, k_d)$ and level $\mathfrak{N}$ such that

$$\|f\|_\infty \gg (k_1 \cdots k_d)^{\frac{1}{4}} \mathbf{N}_{F/\mathbb{Q}}(\mathfrak{N})^{\frac{1}{4}} \|f\|_2 .$$

The general result is stated below in Theorem 1.3 and includes Maass forms. Let $\pi = \otimes_v \pi_v$ be an unitary cuspidal automorphic representation of GL$(2, \mathbb{A}_F)$. If the place v is non-archimedean, then π_v has level $\mathfrak{N}_v$. If the place v is archimedean, then either π_v is a discrete series (resp. limit of discrete series) of some weight $k \geq 2$ (resp. of weight $k = 1$) and we denote by h the number of such places; or π_v is a principal or complementary series of spectral parameter $r > 0$ and we denote by m the number of such places. Thus $h + m = d$. The representation π is generated by a unique Hilbert–Maass newform $f \in \pi$. Hilbert–Maass forms belong to $L^2_\omega(\mathrm{GL}(2, F)\backslash \mathrm{GL}(2, \mathbb{A}_F))$ where ω is the central character.

Theorem 1.3. *Fix $\epsilon > 0$ and a totally real number field F of degree d. For all but finitely many tuples of integers $k_1, \ldots, k_h, t_1, \ldots t_m \in \mathbb{N}_{\geq 1}$ and square integral ideals $\mathfrak{N}$, there are Hilbert-Maass newforms f of level $\mathfrak{N}$, archimedean type (h, m) with weights $(k_1, \ldots, k_h)$ and spectral parameters in the respective intervals $[t_1 - 1, t_1], \ldots, [t_m - 1, t_m]$ such that*

$$\|f\|_\infty \gg (t_1 \cdots t_m)^{\frac{1}{6}-\epsilon}(k_1 \cdots k_h)^{\frac{1}{4}-\epsilon}\mathbf{N}_{F/\mathbb{Q}}(\mathfrak{N})^{\frac{1}{4}-\epsilon} \|f\|_2 .$$

The multiplicative constant may depend only on F and $\epsilon > 0$.

1.4. Whittaker models and newvector theory. We will introduce a local invariant to study large values of modular forms, the key novelty being Proposition 1.4 below. It is shown by Iwaniec–Sarnak in the corrigendum in [30] of [17, Lem. A.1] that a Hecke–Maass form f on $SL(2, \mathbb{Z})\backslash \mathfrak{H}$

of large spectral parameter r satisfies $\|f\|_\infty \gg_\varepsilon r^{\frac{1}{6}-\varepsilon} \|f\|_2$. This is related to the asymptotic behavior of the K-Bessel function in the transition range.

This paper shows that large values are a general phenomenon that occurs not only in the eigenvalue aspect but also in the level aspect. In the level aspect one might hope that exponential sums will play the role of special functions. Thus our starting point has been to draw in the analogy between archimedean and non-archimedean places and to look for a non-archimedean analogue of the transition range of the K-Bessel function. Such analogy isn't obvious but indeed exists as we now explain.

We first isolate the following as the key of the lower bound [30] for Hecke–Maass forms:

$$(1.3) \qquad \max_{y>0} y^{\frac{1}{2}} |K_{ir}(y)| \asymp r^{\frac{1}{6}} e^{-\frac{\pi r}{2}}, \quad \text{as } r \to \infty.$$

See §1.5 and §5.2 for details.

Now $y^{\frac{1}{2}} K_{ir}(y)$ is the spherical vector in the Kirillov model of a principal series representation of $\mathrm{GL}(2,\mathbb{R})$. So in analogy it is natural to investigate the newvector in the Kirillov model of a ramified representation π of $\mathrm{GL}(2,\mathbb{Q}_p)$. Explicit formulas are known [7,8], namely we find

$$\eta(y) |y|^{\frac{1}{2}} \mathbb{1}_{\mathbb{Z}_p}(y) \quad \text{or} \quad \mathbb{1}_{\mathbb{Z}_p^\times}(y)$$

depending on the representation (here η is an unramified character of $\mathbb{Q}_p^\times$). The analogy with the archimedean $y^{\frac{1}{2}} K_{ir}(y)$ is good; we observe the same oscillatory behavior as $y \to 0$ and the same rapid decay as $y \to \infty$.

However there is a difference which is that the Kirillov newvector in the non-archimedean case doesn't exhibit any transition range as a function in the y-variable. In particular there is no analogue of (1.3), the Kirillov function sharply goes from one to zero between $|y| = 1$ and $|y| > 1$.

Our solution is to introduce a local invariant $h(\pi)$ of generic representations of $\mathrm{GL}(2,\mathbb{Q}_p)$. It is defined as the maximum value of a newvector $W_\circ$ in the Whittaker model. We have that $W_\circ \left(\left(\begin{smallmatrix} y & 0 \\ 0 & 1 \end{smallmatrix} \right) \right)$ is the newvector in the Kirillov model as above. Our main point is that $W_\circ$ achieves large values outside of the diagonal. Thus the correct analogue of (1.3) and the transition region in the non-archimedean case is to be found in the Whittaker function as opposed to the Kirillov function. We establish the following result in Section 4 which implies $h(\pi) = p^{\frac{1}{2}}$ for certain type of representations:

Proposition 1.4. *For a twist-minimal unitary principal series representation of* $\mathrm{GL}(2, \mathbb{Q}_p)$ *of conductor* p^2 *the newvector* $W_\circ$ *in the Whittaker model satisfies*

$$\max_{g \in \mathrm{GL}(2, \mathbb{Q}_p)} |W_\circ(g)| = p^{\frac{1}{2}} |W_\circ(\mathbf{e})|.$$

1.5. Ideas of proof. Let $f \in S_2^*(N, \chi)$ be a primitive newform of level N as in Theorem 1.1. For each cusp $\mathfrak{a}$ of $\Gamma_0(N) \backslash \mathfrak{H}$ we can consider the Fourier expansion at $\mathfrak{a}$ and integrate f against closed horocycles. The number of inequivalent cusps of $\Gamma_0(N) \backslash \mathfrak{H}$ is $\sum_{ab=N} \varphi((a, b))$ where φ is Euler function and (a, b) is the g.c.d of a and b. If N is square-free then there are $2^{\omega(N)}$ cusps where $\omega(N)$ is the number of prime factors of N.

Next we take into account the group of Atkin–Lehner involutions [2] because if a cusp $\mathfrak{a}$ is conjugate to $i\infty$ by an Atkin–Lehner involution, then the Fourier expansions of f at $\mathfrak{a}$ and $i\infty$ are directly related. The cusp $\mathfrak{a}$ does carry the same information as the cusp $i\infty$ so that the same asymptotics arise when investigating periods against closed horocycles at such cusps. The number of cusps is greater than the number $2^{\omega(N)}$ of Atkin–Lehner involutions if and only if N is not square-free.

The conclusion so far is that one should investigate the case when N is *not* square-free and look at closed horocycles around cusps of $\Gamma_0(N) \backslash \mathfrak{H}$ which are *not* conjugate to $i\infty$ by the group of Atkin–Lehner involutions. Fortunately this move to a somehow inextricable situation will actually succeed!

At this stage we bring into play representation theory, viewing f as generating a cuspidal automorphic representation $\pi \simeq \otimes_p \pi_p$. Our key point is to work with the invariants $h(\pi_p)$ attached to the representations π_p. Among other things it captures the local harmonic analysis underlying the solution of the problem, namely the asymptotic of p-adic Whittaker functions in the transition range. The idea should apply more generally to several kind of periods of automorphic forms as it primarily relies on factorization of period integrals in local functionals.

The method will be formalized in Section 3. We start with a simple inequality in Lemma 3.2 which is a variant of Hecke classical bound [15, Satz 8 p. 484]. This reduces lower bounds for $\|f\|_\infty$ to lower bounds for $h(\pi_p)$. Then we are in position to apply the new asymptotic in Proposition 1.4.

Note that the global argument saying that $\mathfrak{a}$ is not conjugate to $i\infty$ translates into the local fact explained in the previous §1.4 that the newvector $W_\circ$ achieves its maximum at a matrix $g \in \mathrm{GL}(2, \mathbb{Q}_p)$ that is neither diagonal nor anti-diagonal.

The value $h(\pi_p) = p^{\frac{1}{2}}$ equals the fourth-root of the conductor p^2 of π_p, which explains the exponent $\frac{1}{4}$ in Theorem 1.1. It seems to be the largest among all representations of $\mathrm{GL}(2, \mathbb{Q}_p)$ in which case Theorem 1.1 would provide the best lower bound that can be attained by our method.

The proof of Theorem 1.3 is based on the same ideas by including the uniformity in the parameters at infinity, see Section 5. The introduction of the local invariant $h(\pi)$ allows to make precise the analogy between archimedean and non-archimedean places. Indeed we can explain the respective exponents $\frac{1}{4}$ in the level aspect, $\frac{1}{6}$ in the eigenvalue aspect [30] (Maass forms), $\frac{1}{4}$ in the weight aspect [40] (holomorphic forms), by showing that $h(\pi) = N^{\frac{1}{4}}$ for certain non-archimedean π of conductor $N = p^2$, that $h(\pi) \asymp r^{\frac{1}{6}}$ for π an archimedean principal series of parameter r and that $h(\pi) \asymp k^{\frac{1}{4}}$ for π a discrete series of weight k.

2. Complements to the main results

We explore variants and corollaries as well as relation to other questions. The goal is to revisit some previous estimates in light of our results. The content of this section is summarized as follows:

(i) For a given square level N and a primitive character χ (mod N) the Theorem 1.1 says that $\|f\|_\infty \gg N^{\frac{1}{4}}$. Surprisingly there is a single point $z_\chi \in \mathfrak{H}$ depending only on χ at which these large values are achieved. Namely we have

$$\|f\|_\infty \geq |f(z_\chi)| \gg N^{\frac{1}{4}}$$

for all $f \in S_2^*(N, \chi)$, see Theorem 2.2.

(ii) For any $4 < r \leq \infty$ we have that $\|f\|_r \to \infty$ as $N \to \infty$ (again assuming that N is a square and χ is primitive, see Proposition 2.3).

(iii) We establish upper and lower estimates for Wilton partial sums

$$\sum_{m \leq M} a_m e(mx)$$

where $x \in (0, 1)$ and the Fourier coefficients a_m are given by (1.1).

2.1. Mean value estimates. Let

(2.1)
$$M(z) := \sum_{g \ o.b. \ S_2(N,1)} |g(z)|^2, \quad z \in \mathfrak{H}$$

where the sum is over an orthonormal basis of $S_2(N,1)$ for the Petersson inner-product. Note that $\dim S_2(N,1) = N^{1+o(1)}$ and by definition

$$(2.2) \qquad \frac{1}{\mathrm{vol}(\Gamma_0(N)\backslash\mathfrak{H})} \int_{\Gamma_0(N)\backslash\mathfrak{H}} M(z)\frac{dxdy}{y^2} = \frac{1}{4\pi} + O(N^{-1+o(1)}).$$

It is shown by Jorgenson–Kramer [20] that

$$(2.3) \qquad\qquad\qquad \|M\|_\infty \ll 1.$$

The estimate is quite robust since it applies more generally to finite coverings of a given non-compact Riemann surface.

It follows from (2.3) that the upper bound $\|f\|_\infty \ll N^{\frac{1}{2}+\epsilon}$ holds for all primitive forms $f \in S_2^*(N,1)$. Indeed $g := \frac{f}{(f,f)^{\frac{1}{2}}}$ is an orthonormal vector for the Petersson inner product and $(f,f) = N^{1+o(1)}$ by results of Iwaniec and Hoffstein–Lockhart.

Another deep estimate for $M(z)$ is in the work of Michel–Ullmo [24, Thm. 1.6]:

$$(2.4) \qquad \frac{1}{\mathrm{vol}(\Gamma_0(N)\backslash\mathfrak{H})} \int_{\Gamma_0(N)\backslash\mathfrak{H}} M(z)^2\frac{dxdy}{y^2} = \frac{1}{(4\pi)^2} + O(N^{-\delta}),$$

for N square-free and some $\delta > 0$.

The previously conjectured estimate (A) is compatible with (2.3) and (2.4), e.g. it would have followed from (2.3) *if the mass $|f(z)|^2$ were* equally distributed among all primitive forms in $S_2^*(N,1)$, see [20, Rem. 3.3].

We now know from Theorem 1.1 that (A) doesn't hold in general. The conclusion one may draw is that the values $f(z)$ for different $f \in S_2^*(N,\chi)$ fluctuate much more than one might originally expect. In fact we shall derive a precise statement by considering the mean value estimate (2.3) for general nebentypus. Thus for χ an even Dirichlet character, let

$$(2.5) \qquad\qquad M_\chi(z) := \sum_{g \ o.b. \ S_2(N,\chi)} |g(z)|^2, \quad z \in \mathfrak{H}.$$

The sum is over an orthonormal basis of $S_2(N,\chi)$ and the asymptotic (2.2) still holds. If $\chi = \mathbb{1}$, then $M_\chi(z) = M(z)$. Consider also the average over characters

$$(2.6) \qquad\qquad M_{\Gamma_1(N)}(z) := \frac{2}{\phi(N)} \sum_{\chi \ (N)} M_\chi(z), \quad z \in \mathfrak{H}.$$

The number $\frac{\phi(N)}{2}$ of even Dirichlet characters modulo N equals the index $[\Gamma_0(N) : \Gamma_1(N)]$.

The function $M_{\Gamma_1(N)}(z)$ (resp. $M(z)$) can be interpreted as the ratio of the Arakelov metric and the hyperbolic metric for $\Gamma_1(N)\backslash\mathfrak{H}$ (resp. $\Gamma_0(N)\backslash\mathfrak{H}$). It is also shown in [20] that $\left\|M_{\Gamma_1(N)}\right\|_\infty \ll 1$ (their result applies to any finite cover of a given Riemann surface). This implies that for all $\epsilon > 0$,

$$\frac{1}{\mathrm{vol}(\Gamma_1(N)\backslash\mathfrak{H})} \sup_{z\in\mathfrak{H}} \sum_{\chi\ (N)} \sum_{f\in S_2^*(N,\chi)} |f(z)|^2 \ll_\epsilon N^\epsilon.$$

On the other hand we have seen in Theorem 1.1 that there are primitive forms $f \in S_2^*(N,\chi)$ of arbitrary large level such that $\|f\|_\infty \gg N^{\frac{1}{4}}$. This might be interpreted by saying that for certain $z \in \mathfrak{H}$ the absolute value $|f(z)|$ fluctuate much with f.

Now it is interesting to ask about the size of $\|M_\chi\|_\infty$ itself, for χ primitive. Our first observation is that the proof in [20] that $\|M\|_\infty \ll 1$ doesn't extend to M_χ. Indeed the argument on p.1275 of [20] uses the positivity of the heat kernel $K(t; z, \gamma z)$ which for non-trivial nebentypus should be replaced by $\chi(\gamma)K(t; z, \gamma z)$, destroying the positivity. Our second observation is that the proof in [24] doesn't apply either, one reason is that it is restricted to square-free levels.

In fact we have the following result which clarifies the situation because it shows that $\|M_\chi\|_\infty$ can become surprisingly large for non-trivial nebentypus.

Proposition 2.1. *For all even primitive Dirichlet character χ modulo $N = p^2$ with p prime, we have $\|M_\chi\|_\infty \gg_\epsilon N^{\frac{1}{2}-\epsilon}$.*

Not only the individual forms may assume large values but also their entire average $M_\chi(z)$ in (2.5)! Thus the respective arguments in [24] and [20, p. 1275] don't generalize to M_χ for a good reason.

Proof of Proposition 2.1. It follows from Theorem 2.2 below that $g(z_\chi) \gg N^{-\epsilon}$ where $g := \frac{f}{(f,f)^{\frac{1}{2}}}$ and $f \in S_2^*(N,\chi)$. Since $z_\chi \in \mathfrak{H}$ depends only on χ (and not on g), we have $M_\chi(z_\chi) \gg N^{\frac{1}{2}-\epsilon}$ which concludes the proof. $\square$

A comparison may also be drawn with (2.6): since $\left\|M_{\Gamma_1(N)}\right\|_\infty \ll 1$, the absolute value $|M_\chi(z)|$ fluctuate much with χ for certain $z \in \mathfrak{H}$. This is consistent with (2.7) below in which z_χ shall vary with χ.

2.2. A specific CM-point. We refine the statement of Theorem 1.1 by finding a specific point $z \in \mathfrak{H}$ such that $|f(z)| \gg N^{\frac{1}{4}}$.

Theorem 2.2. *Let χ be an even primitive Dirichlet character modulo p^2. There is a unique $b \in (\mathbb{Z}/p\mathbb{Z})^\times$ such that $\chi(1 - pz) = e^{2\pi i \frac{bz}{p}}$ for all $z \in \mathbb{Z}$. Let $a \in (\mathbb{Z}/p\mathbb{Z})^\times$ be the multiplicative inverse: $ab \equiv 1 \pmod{p}$.*

For all primitive forms $f \in S_2^*(p^2, \chi)$, we have $|f(z_\chi)| \gg p^{\frac{1}{2}}$ where

$$(2.7) \qquad\qquad z_\chi := \frac{a}{p} + \frac{i}{p^3} \in \mathfrak{H}.$$

Remarks.   (i) When establishing $\|f\|_\infty \ll N^{\frac{1}{3}}$ for N square-free in [3, 13, 14, 37], the critical region is $N^{-1} \ll \Im(z) \ll N^{-\frac{2}{3}}$. The lower bound $N^{-1} \ll \Im(z)$ follows by an application of Atkin–Lehner operators and the fact that N is square-free. In comparison, $N^{-\frac{3}{2}} = \mathrm{Im}(z_\chi)$ in (2.7) above is much smaller, where $N := p^2$.

  (ii) One has $\mathrm{Im}(z_\chi) \geq \mathrm{Im}(\gamma z_\chi)$ for all $\gamma \in \Gamma_0(p^2)$. Also if we let $z_\chi' := \frac{-1}{p^2 z_\chi}$, then $\frac{1}{p^3} > \mathrm{Im}(\gamma z_\chi')$. Actually it is possible to verify the following:

$$\frac{1}{p^7} \ll \frac{1}{p^3} - \max_{\gamma \in \Gamma_0(p^2)} \mathrm{Im}(\gamma z_\chi') \ll \frac{1}{p^5}.$$

 (iii) Since z_χ belongs to $\mathbb{Q}(i)$ it is a CM-point. It would be interesting to carry out the analysis of the local zeta integral at p attached to Waldspurger formula for $L(\frac{1}{2}, f \times \theta)$, where θ is a Grössencharacter of finite order of $\mathbb{Q}(i)$ ramified at p.

The proof of Theorem 2.2 will be given in §5.1. With more work we also establish lower bounds for other L^r-norms. If $2 \leq r \leq \infty$, let

$$\|f\|_r^r := \frac{1}{\mathrm{vol}(\Gamma_0(N)\backslash\mathfrak{H})} \int_{\Gamma_0(N)\backslash\mathfrak{H}} |f|^r \frac{dxdy}{y^2}.$$

Proposition 2.3. *(i) For any fixed $0 < \delta < \frac{1}{2}$ we have $|f| \gg p^\delta$ on a neighborhood of z_χ of volume at least $p^{\frac{1}{2}-\delta}$, with z_χ as in Theorem 2.2.*

(ii) Thus for all $r > 4$, we have that $\frac{\|f\|_r}{\|f\|_2} \to \infty$ as $p \to \infty$ uniformly for all $f \in S_2^*(p^2, \chi)$ and primitive character χ modulo p^2.

2.3. Square-root cancellation and uniform Wilton estimates. Integral representations of special values of L-functions yield relations between the sup-norm of forms and the subconvexity problem. Let θ be a fixed dihedral GL(2) form induced from a Hecke character of finite order on a quadratic

field. The Lindelöf hypothesis states that $L(\frac{1}{2}, f \times \theta) \ll_{\epsilon,\theta} N^\epsilon$, which is of course very reliable because it would follow from the GRH for $L(s, f \times \theta)$.

The conjectured estimate (A) would have implied the Lindelöf hypothesis for $L(\frac{1}{2}, f \times \theta)$ as follows from the Waldspurger period formula, see [3] and [4, §6]. Conversely the Lindelöf hypothesis for $L(\frac{1}{2}, f \times \theta)$ implies a bound $|f(z)| \ll_\epsilon N^\epsilon$ for *fixed* CM points $z \in \mathfrak{H}$. For sup-norm bounds in the eigenvalue aspect the relation to the subconvexity problem is discussed in [17, Rem. D] and [31, §4].

Similarly the Hecke integral gives yet another relation to the subconvexity problem. It is not difficult to verify that

$$(2.8) \qquad \int_{1/N}^1 f(z) \frac{dy}{y} = (2\pi)^{-1} L(\tfrac{1}{2}, f) + O(1).$$

Thus (A) would have implied the Lindelöf hypothesis in the level aspect for $L(\frac{1}{2}, f)$. Since we know that the estimate (A) doesn't hold in general and since we believe in the Lindelöf hypothesis we see that the Hecke integral (2.8) has to carry a lot of cancellation.

We now turn to square-root cancellation heuristics based on Fourier expansion. For $z = x + iy$, the tail of the Fourier expansion (1.1) is negligible when ny becomes large. Thus, setting $M := 1/y$, we have an approximation

$$(2.9) \qquad f(x + iy) \approx M^{-\frac{1}{2}} \sum_{n \sim M} a_n e^{2i\pi nx}.$$

The normalization (1.1) is such that the Deligne bound reads $|a_n| \le \tau(n)$ for all $n \ge 1$. In many aspects the coefficients a_n behave at random (cf. the Sato-Tate distribution of a_p, the sign changes, the bounds for L-values and character twists). A basic heuristic is to compare (2.9) to a random trigonometric polynomial of degree M. With high probability the sup-norm over $x \in [0,1]$ of a random trigonometric polynomial is $\ll (M \log M)^{\frac{1}{2}}$, see e.g. [5] and the references herein.

For a *fixed* form f, the right-hand side of (2.9) is in fact $\ll_{f,\epsilon} M^\epsilon$ by the classical Wilton estimate. The estimate (A) was equivalent to the stronger estimate $\ll_\epsilon (MN)^\epsilon$ that would be uniform in the level N of f.

The following is a uniform version of the classical Wilton estimate (2.9).

Proposition 2.4. *Let* $f \in S_2^*(N, \chi)$ *be a primitive form with normalized coefficients* $(a_m)_{m \ge 1}$ *as in* (1.1)*. For all integer* $M \ge 1$ *and* $\epsilon > 0$,

$$\sum_{m \le M} a_m e(mx) \ll_\epsilon M^{\frac{1}{2}+\epsilon} \|f\|_\infty.$$

Proof. We consider the Dirichlet kernel

$$D_M(x+iy) := y^{\epsilon-1} \sum_{m \leq M} m^\epsilon e(-mx),$$

whose L^1-norm satisfies $\int_0^1 |D_M(x+iy)| \ll y^{\epsilon-1} M^\epsilon$. Then we compute

$$\int_0^\infty \int_0^1 f(\alpha+x+iy) D_M(x+iy) dx dy = \int_0^\infty y \sum_{m \leq M} a_m e(m\alpha) m^{\frac{1}{2}} e^{-2\pi my} (my)^\epsilon \frac{dy}{y}$$

$$= \sum_{m \leq M} \frac{a_m e(m\alpha)}{m^{\frac{1}{2}}} \int_0^\infty y^{1+\epsilon} e^{-2\pi y} \frac{dy}{y}.$$

By the triangle inequality this implies

$$\sum_{m \leq M} \frac{a_m e(m\alpha)}{m^{\frac{1}{2}}} \ll_\epsilon M^\epsilon \int_0^\infty \left(\sup_{\Im z = y} |f(z)| \right) y^{\epsilon-1} dy$$

$$\ll_\epsilon M^\epsilon \|f\|_\infty$$

The desired estimate follows by integration by parts. □

Conversely the Theorem 1.1 yields the following lower bound for Wilton sums.

Proposition 2.5. *There exist forms* $f \in S_2^*(N, \chi)$ *of arbitrary large level* N, *and* $M \geq 1$, $x \in [0,1]$ *such that*

$$\sum_{m \leq M} a_m e(mx) \gg_\epsilon M^{\frac{1}{2}} N^{\frac{1}{4}-\epsilon}.$$

Remark. Thus we have constructed arithmetic sequences $m \mapsto a_m e(mx)$ whose partial sums do not satisfy square-root cancellation by a large $N^{\frac{1}{4}}$ margin. This is despite the fact that the sequence is random in many aspects, which is reminiscent of a similar phenomenon for $m \mapsto a_m e(-2\sqrt{m})$ observed by Iwaniec–Luo–Sarnak [18, App. C].

Example. The proof below shows that one can choose $M \leq y^{-1}$ if $f(x+iy) \gg N^{\frac{1}{4}}$. From Theorem 2.2 we see that for a level $N = p^2$ and χ primitive the bound holds with $x = \frac{a}{p}$ and some integer $M \leq p^3$. It would be interesting to have a direct approach in this particular case.

Proof. Let $\alpha > 0$ be such that $|S(M)| \leq \alpha M^{\frac{1}{2}+\epsilon}$ for all integer $M \geq 1$, where

$$S(M) := \sum_{m \leq M} a_m e(mx).$$

Using partial summation we find that

$$f(x + iy) = y \sum_{n=1}^{\infty} \left(S(n) - S(n-1) \right) n^{\frac{1}{2}} e^{-2\pi ny}$$

$$\ll y \sum_{n=1}^{\infty} |S(n)| \, n^{-\frac{1}{2}} e^{-2\pi ny} \ll_\epsilon \alpha y^{-\epsilon}.$$

Let $f \in S_2^*(N, \chi)$ be such that $\|f\|_\infty \gg N^{\frac{1}{4}}$ (Theorem 1.1). Thus $\alpha \gg N^{\frac{1}{4}} y^\epsilon$ and the claim follows. $\qquad\square$

3. Whittaker periods and special forms

3.1. A local invariant. We introduce the following normalized local invariant attached to generic representations.

Definition 3.1. *For a unitary generic representation π of $\mathrm{GL}(2, \mathbb{Q}_p)$, let*

$$h(\pi) := \max_{g \in \mathrm{GL}(2, \mathbb{Q}_p)} |W_\circ(g)| \,,$$

where $W_\circ$ is the newvector in the Whittaker model, normalized by $W_\circ(\mathbf{e}) = 1$.

The definition is licit because [7] a nonzero newvector in the Whittaker model is unique up to scalar and doesn't vanish at the identity.

Example. If π is unramified (that is $p^{c(\pi)} = 1$), then $h(\pi) = 1$. The invariant $h(\pi)$ is a measure of the ramification of π in the sense that it behaves somehow similarly to the conductor $p^{c(\pi)}$. (They are different because $h(\pi) = 1$ if $p^{c(\pi)} = p$).

3.2. A lower bound. The introduction of the invariant $h(\pi)$ is motivated by the following estimate.

Lemma 3.2. *Let $f \in S_2^*(N, \chi)$ be a primitive form and for all primes p, let π_p be its local component at p. The following holds:*

$$\|f\|_\infty \geq (2\pi e)^{-1} \prod_p h(\pi_p).$$

Proof. The proof is better achieved in the adelic framework. Thus we work with the corresponding automorphic form φ on $\mathrm{GL}_2(\mathbb{Q})\backslash \mathrm{GL}_2(\mathbb{A})$. The image (inside $\mathbb{R}_+$) of $z \mapsto |f(z)|$ is identical to the image of $g \mapsto |\varphi(g)|$, thus $\|f\|_\infty = \|\varphi\|_\infty$. Let ψ be the standard additive character of $\mathbb{Q}\backslash\mathbb{A}$ which is unramified at all finite places and let dx be the self-dual Haar measure on $\mathbb{A}$. Let

$$(3.1) \qquad W(g) := \int_{\mathbb{Q}\backslash\mathbb{A}} \varphi(n(x)g)\overline{\psi(x)}dx, \quad g \in G(\mathbb{A})$$

be the corresponding Whittaker function. Since f is primitive, we find that $W(\mathbf{e}) = e^{-2\pi}$.

The Whittaker function factors into $W = W_\infty \prod_p W_p$. Here for all prime p, $W_p(g)$ is a newvector in the Whittaker model $\mathcal{W}(\pi_p, \psi_p)$ and $W_\infty \in \mathcal{W}(\pi_\infty, \psi_\infty)$ is proportional to the lowest weight vector. We see that we may arrange so that

$$W_\infty(\left(\begin{smallmatrix} y & 0 \\ 0 & 1 \end{smallmatrix}\right)) = ye^{-2\pi y}, \text{ and } W_p(\mathbf{e}) = 1.$$

Thus W_p is the normalized newvector denoted $W_\circ$ in the previous section.

Since $\mathbb{Q}\backslash\mathbb{A}$ has volume one, we have the inequality

$$\|f\|_\infty \geq \sup_{g\in G(\mathbb{A})} |W(g)|$$
$$= \sup_{g\in G(\mathbb{R})} |W_\infty(g)| \prod_p \sup_{g\in G(\mathbb{Q}_p)} |W_p(g)|.$$

The claim now follows from Definition 3.1 and an elementary calculation at infinity which shows that the maximum is attained at $y = (2\pi)^{-1}$. $\qquad\square$

The proof of Lemma 3.2 is similar to that of Hecke bound. In fact we can recover Hecke bound since the coefficients of f satisfy the inequality

$$|a_n| \leq n^{\frac{1}{2}} \prod_{p|n} h(\pi_p), \quad n \geq 1.$$

3.3. Modular forms with prescribed ramification.

Lemma 3.3. *Let π be an irreducible admissible representation of $\mathrm{GL}(n, \mathbb{Q}_p)$ whose central character ω_π has the same conductor. Then π is a twist-minimal principal series representation, that is: π of the form $\chi_1 \boxplus \chi_2 \cdots \boxplus \chi_n$ with χ_i unramified for all $1 \leq i \leq n-1$.*

Proof. Let (V, N) be the representation of the Weil-Deligne group attached to π by the local Langlands correspondence. We want to show that $N = 0$ and V is a direct sum of one-dimensional characters.

Let I be the inertia group of $\mathbb{Q}_p$ and $(I^u)_{u \geq 0}$ be the upper numbering filtration. Then the conductor of π is equal to the Artin conductor of (V, N), thus [33, Chap. VI]:

$$c(\pi) = \log_p f(V, N) = \operatorname{codim} \left(V^I \right)^{N=0} + \int_0^\infty \operatorname{codim} V^{I^u} du.$$

By assumption $c(\pi) = c(\omega_\pi)$. The corresponding formula also holds for $c(\omega_\pi) = \log_p f(\det V)$, which implies $N = 0$ on V^I and

$$\operatorname{codim} V^{I^u} = \operatorname{codim} (\det V)^{I^u} \quad \text{for all } u \geq 0.$$

Since $\det V$ is one-dimensional, we deduce that $\operatorname{codim} V^I \in \{0, 1\}$ from which the claim follows. $\square$

Corollary 3.4. *Let χ be an even primitive Dirichlet character of conductor N. Then all primitive forms in $S_2^*(N, \chi)$ are twist-minimal. Moreover the components at any prime p are twist-minimal principal series. The p-th Fourier coefficients in the expansion* (1.1) *are units:* $|a(p)| = 1$.

Remarks.   (i) A direct proof of the last assertion that $|a(p)| = 1$ may be found in [26, Prop. 4.6.17].
  (ii) The condition that χ be *primitive* is essential. For example if N is square-free and $\chi = \mathbb{1}$ then the components at any prime $p \mid N$ are Steinberg. Thus there are local conditions on N and on the conductor of χ that need to be satisfied.
 (iii) Conversely if the desired local conditions are satisfied (e.g. $N \mid f(\chi)^2$ would be sufficient), then a positive proportion of forms in $S_2^*(N, \chi)$ have a component at every prime that is a principal series representation; this follows from [34, 39].

The Corollary 3.4 immediately follows from Lemma 3.3 with $n = 2$. We provide below two alternative proofs which are of independent interest. It is instructive to see how each argument naturally points towards the same conclusion that π is a principal series.

Alternative proof of Lemma 3.3 when $n = 2$. We first prove that π cannot be a twist $\eta \operatorname{St}$ of a Steinberg representation. Indeed if η were unramified then $c(\eta \operatorname{St}) = 1$, a contradiction. If η were ramified, $c(\eta \operatorname{St}) = 2c(\eta)$ which is strictly larger than $c(\omega_\pi) = c(\eta^2)$, again a contradiction.

We next prove that π cannot be a dihedral supercuspidal representation. Indeed otherwise $c(\pi) \geq 2$ and π would be induced from a quasi-character η of a quadratic extension E. Its central character ω_π would be equal to $\eta|_{\mathbb{Q}_p^\times} \chi_E$ where χ_E is the quadratic character attached to E. If E were unramified then $c(\pi) = 2c(\eta)$ which is strictly larger than $c(\omega_\pi) = c(\eta|_{\mathbb{Q}_p^\times})$, a contradiction. If E were tamely ramified then $c(\pi) = c(\eta) + 1$ which is strictly larger than $c(\eta|_{\mathbb{Q}_p^\times}) \geq c(\omega_\pi)$, again a contradiction. If E were widely ramified then $p = 2$ and $c(\pi) = c(\eta) + 2$ which is strictly larger than $c(\eta|_{\mathbb{Q}_p^\times} \chi_E)$, again a contradiction. The case of non-dihedral supercuspidal representations (when $p = 2$) follows from [38].

Thus π is a principal series representation $\chi_1 \boxplus \chi_2$. We have $c(\pi) = c(\chi_1) + c(\chi_2)$. On the other hand $\omega_\pi = \chi_1 \chi_2$ which implies

$$c(\omega_\pi) \leq \max(c(\chi_1), c(\chi_2))$$

and thus $c(\chi_1)$ or $c(\chi_2)$ is equal to zero. The claim follows. $\square$

Second alternative proof of Lemma 3.3 when $n = 2$. Let $r = c(\pi) = c(\omega_\pi)$. Recall [7] that the newvector in the representation of π is stabilized by the congruence subgroup I_r consisting of matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right)$ with $p^r | c$ and transforms via the character $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) \mapsto \omega_\pi(d)$.[2]

Let τ be the representation of $\mathrm{GL}(2, \mathbb{Z}_p)$ induced from this character of I_r. By assumption $c(\omega_\pi) = r$, which implies that τ is irreducible. By Frobenius reciprocity the restriction $\pi|_{\mathrm{GL}(2,\mathbb{Z}_p)}$ contains τ.

By a result of Henniart [16] the representation τ is a type for the Bernstein component of twist-minimal principal series with central character equal to ω_π on $\mathbb{Z}_p^\times$. (Since ω_π is of conductor p^r we are away from the exceptional cases in [16, §A.1.5] where a Bernstein component does not possess a type). This implies that $\pi|_{\mathrm{GL}(2,\mathbb{Z}_p)}$ contains τ if and only if $\pi = \chi_1 \boxplus \chi_2$ with χ_1 unramified and $\chi_2|_{\mathbb{Z}_p^\times} = \omega_\pi|_{\mathbb{Z}_p^\times}$, as claimed. $\square$

3.4. The case of oldforms. The space of oldforms of level N is spanned by functions $g(z) = f(dz)$ where f is a primitive form of level R strictly dividing N and $d \mid \frac{N}{R}$. Since $\|g\|_\infty = \|f\|_\infty$, the bound (A) for the primitive form f would imply the same bound for the oldform g. For example the oldforms of level N that come from a level 1 form are trivially uniformly bounded. Also character twists would be compatible with (A) since for all Dirichlet characters η, we have $\|f \otimes \eta\|_\infty = \|f\|_\infty$.

---

[2]There is a typo in [7, Eq. (1.3)] where "a'" should read "d'".

4. Ramified Whittaker functions

4.1. A formula for the Whittaker newvector. Let χ be a unitary character of $\mathbb{Q}_p^\times$ of conductor p^2. Consider the principal series representation $\pi = \mathbb{1} \boxplus \chi$ of $G = \mathrm{GL}(2, \mathbb{Q}_p)$. Let ψ be an unramified additive character and $W_\circ$ be the newvector in the Whittaker model of π, normalized by $W_\circ(\mathbf{e}) = 1$. It is stable under the action of the congruence subgroup I_2 of matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}(2, \mathbb{Z}_p)$ such that $p^2 \mid c$. For $i \in \{0, 1, 2\}$, let $\mathbf{k}_i := \left(\begin{smallmatrix} p^i & 0 \\ p^i & 1 \end{smallmatrix}\right)$.

Proposition 4.1. *For all $y \in \mathbb{Q}_p^\times$, $W_\circ(\left(\begin{smallmatrix} y & 0 \\ 0 & 1 \end{smallmatrix}\right)) = |y|^{\frac{1}{2}} \mathbb{1}_{\mathbb{Z}_p}(y)$ and*

$$W_\circ(\left(\begin{smallmatrix} y & 0 \\ 0 & 1 \end{smallmatrix}\right) \mathbf{k}_0) = p^{-1} |y|^{\frac{1}{2}} \chi(-y)\psi(y)\epsilon(\tfrac{1}{2}, \chi, \psi)\mathbb{1}_{p^{-2}\mathbb{Z}_p}(y).$$

Proof. The first identity is well-known [7,8]. The second identity follows from the Jacquet–Langlands functional equation and may also be established in the same way as Proposition 4.2 below. We omit the details which are not directly relevant to the proof of Theorem 1.1. $\qquad\square$

Proposition 4.2. *If $y \in p^{-2}\mathbb{Z}_p^\times$ is such that*

$$\chi(1 - z) = \psi(yz), \quad \forall z \in p\mathbb{Z}_p,$$

then $W_\circ(\left(\begin{smallmatrix} y & 0 \\ 0 & 1 \end{smallmatrix}\right) \mathbf{k}_1) = p^{\frac{1}{2}}$. Otherwise $W_\circ(\left(\begin{smallmatrix} y & 0 \\ 0 & 1 \end{smallmatrix}\right) \mathbf{k}_1) = 0$.

Remark. Let B be the Borel subgroup of upper-triangular matrices. Then $\{\mathbf{k}_0, \mathbf{k}_1, \mathbf{k}_2\}$ are representatives for the double quotient $B\backslash G/I_2$. Thus we have determined all the values of $W_\circ$ because any element $g \in G$ can be written as

$$g = z \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \mathbf{k}_i \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with $x \in \mathbb{Q}_p$, $y, z \in \mathbb{Q}_p^\times$, $i \in \{0, 1, 2\}$ and $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in I_2$, in which case

$$W_\circ(g) = \psi(x)\chi(dz)W_\circ(\left(\begin{smallmatrix} y & 0 \\ 0 & 1 \end{smallmatrix}\right) \mathbf{k}_i).$$

Proof. We first recall the result of Casselman [7] who shows that the newform $f_\circ$ in the induced model of $\mathbb{1} \boxplus \chi$ is supported on $B \cdot I_2$. This determines $f_\circ : G \to \mathbb{C}$ entirely up to a multiplicative constant, which we normalize by the condition $f_\circ(\mathbf{e}) = 1$,

Then we shall use the fact[3] that

$$f_\circ(\mathbf{w}\,(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix})\,\mathbf{k}_1) = p^{-\frac{1}{2}}\,|x|^{-1}\,\chi(x), \quad \text{if } x \in -1 + p\mathbb{Z}_p,$$

and is zero otherwise, and also

$$f_\circ(\mathbf{w}\,(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix})) = |x|^{-1}\,\chi(x), \quad \text{if } v(x) \le -2,$$

and is zero otherwise.

The Jacquet integral gives an intertwinning from the induced model to the Whittaker model $\mathcal{W}(\pi, \psi)$. Thus letting

$$W(g) := \int_F f_\circ(\mathbf{w}\,(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix})\,g)\overline{\psi(x)}dx, \quad g \in G,$$

it follows that $W_\circ(g) = W(g)/W(\mathbf{e})$.

We find that

$$W(\mathbf{e}) = \int_{v(x) \le -2} \chi(x)\overline{\psi(x)}\frac{dx}{|x|} = \epsilon(1, \chi^{-1}, \overline{\psi}),$$

and on the other hand,

$$W((\begin{smallmatrix} y & 0 \\ 0 & 1 \end{smallmatrix})\,\mathbf{k}_1) = \chi(y)\,|y|^{\frac{1}{2}} \int_F f_\circ(\mathbf{w}\,(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix})\,\mathbf{k}_1)\overline{\psi(xy)}dx$$

$$= p^{-\frac{1}{2}}\,|y|^{\frac{1}{2}}\,\chi(y) \int_{-1+p\mathbb{Z}_p} \chi(x)\overline{\psi(xy)}dx.$$

The proposition follows using known identities on Gauss sums and epsilon factors. $\qquad\square$

The following two early observations served to indicate that the values of the Whittaker function outside of the diagonal could play a role in the context of disproving the conjecture (A). First observation is that we need to go beyond the relation between the sup-norm problem and the subconvexity problem for L-functions: in the theory of L-functions one integrates the Whittaker function on the diagonal $(\begin{smallmatrix} * & 0 \\ 0 & * \end{smallmatrix})$ and on the antidiagonal $(\begin{smallmatrix} 0 & * \\ * & 0 \end{smallmatrix})$ (cf. the Jacquet–Langlands functional equation and the Hecke integral (2.8)); the diagonal is included in BI_2 while the antidiagonal is included in $B\mathbf{k}_0 I_2$. Second observation (already seen in §1.5) is that we need to investigate cusps

---

[3]This follows by writing $\mathbf{w}\,(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix})\,\mathbf{k}_1 = \left(\begin{smallmatrix} p & 1 \\ p(1+x) & x \end{smallmatrix}\right) = \left(\begin{smallmatrix} -x^{-1}p & 1 \\ 0 & x \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ p(1+x^{-1}) & 1 \end{smallmatrix}\right).$

which are not conjugate to $i\infty$ by the group of Atkin–Lehner involutions; the cusps conjugate to $i\infty$ correspond to the double cosets BI_2 and $B\mathbf{k}_0I_2$ again. Thus we were led to study the Whittaker function on $B\mathbf{k}_1I_2$ which is the complement of $BI_2 \cup B\mathbf{k}_0I_2$ in G. This was the underlying motivation of the Proposition 4.2.

Corollary 4.3. *If χ_1 is unramified and χ_2 has conductor p^2, then $h(\chi_1 \boxplus \chi_2) = p^{\frac{1}{2}}$.*

Proof. Let $\chi := \chi_1^{-1}\chi_2$ which has conductor p^2. We have $\chi_1 \boxplus \chi_2 \simeq \chi_1(\mathbb{1} \boxplus \chi)$. Let $W_\circ$ be the normalized newvector in the Whittaker model of $\mathbb{1} \boxplus \chi$ as above. The function $W_\circ^{\chi_1}(g) := \chi_1(\det g)W_\circ(g)$ is the newvector in the Whittaker model of $\chi_1 \boxplus \chi_2$. Indeed this follows from the fact that

$$W_\circ^{\chi_1}(gk) = \chi_1\chi_2(d)W_\circ^{\chi_1}(g),$$

for all elements $g \in G$ and $k = \left(\begin{smallmatrix} * & * \\ * & d \end{smallmatrix}\right)$ in I_2. Thus $h(\pi_1 \boxplus \chi_2) = h(\mathbb{1} \boxplus \chi)$.

The claim follows since Proposition 4.1 and Proposition 4.2 together imply that $h(\mathbb{1} \boxplus \chi) = p^{\frac{1}{2}}$. $\square$

In fact we can generalize the above results to characters χ_2 of arbitrary conductor. The computation is similar thus we omit the proof.

Proposition 4.4. *If χ_1 is unramified and χ_2 has conductor p^c, then*

$$h(\chi_1 \boxplus \chi_2) = p^{\frac{1}{2}\lfloor \frac{c}{2} \rfloor}.$$

This determines $h(\pi)$ for all twist-minimal principal series. More generally if both χ_1 and χ_2 are ramified and if π is supercuspidal we shall present elsewhere a complete formula for $W_\circ(g)$. The formula will involve $_2F_1$ hypergeometric sums which are the non-archimedean analogue of the classical Whittaker and K-Bessel functions.

4.2. Proof of Theorem 1.1. Let $N = p^2$ with p a large enough prime. Let χ be an even primitive Dirichlet character of conductor p^2 and $f \in S_2^*(N, \chi)$. From Corollary 3.4 the component at p is a principal series representation $\chi_1 \boxplus \chi_2$ with χ_1 unramified and χ_2 of conductor p^2.

The Lemma 3.2 implies that $\|f\|_\infty \gg h(\chi_1 \boxplus \chi_2)$. Indeed the remaining local invariants are 1 because π is unramified outside p. The Corollary 4.3 says that $h(\chi_1 \boxplus \chi_2) = p^{\frac{1}{2}} = N^{\frac{1}{4}}$. This concludes the proof. $\square$

Remark. Interestingly, if χ_1 and χ_2 have conductor p, then $h(\chi_1 \boxplus \chi_2) \asymp 1$. This shows that the condition in Theorem 1.1 that the central character χ be primitive modulo N is necessary in our proof.

4.3. Proof of Theorem 1.2. Let χ be an even primitive Dirichlet character of conductor N. From Corollary 3.4, all forms $f \in S_2^*(N, \chi)$ are such that for any prime $p \mid N$, the component π_p at p is a twist-minimal principal series. The Lemma 3.2 together with Proposition 4.4 imply that

$$\|f\|_\infty \gg \prod_{p^c \| N} h(\pi_p) = \prod_{p^c \| N} p^{\frac{1}{2}\lfloor \frac{c}{2} \rfloor}.$$

5. Proof of the other results

5.1. Proof of Theorem 2.2 and Proposition 2.3. Let $f \in S_2^*(p^2, \chi)$ and let φ be the automorphic form attached to f with the same notation as in the proof of Lemma 3.2. We have the expansion

$$(5.1) \qquad \varphi(g) = \sum_{n \in \mathbb{Q}^\times} W\left(\left(\begin{smallmatrix} n & 0 \\ 0 & 1 \end{smallmatrix} \right) g \right), \quad g \in \mathrm{GL}_2(\mathbb{A}),$$

where $W = W_\infty \prod_p W_{\circ p}$, with $W(\mathbf{e}) = 1$ and $W_{\circ p}$ the normalized Whittaker newvector of π_p. We choose $g \in \mathrm{GL}_2(\mathbb{A})$ such that $g_\infty = \left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix} \right)$, $g_p = \left(\begin{smallmatrix} \frac{b}{p} & 0 \\ p & 1 \end{smallmatrix} \right)$ and $g_v = 1$ for all other places $p \neq v, \infty$. Recall that $b \in \mathbb{Z}_p^\times$ is such that $\chi(1 - z) = \psi_p\left(\frac{bz}{p^2}\right)$ for all $z \in p\mathbb{Z}_p$. Since

$$\begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix} g_p = \begin{pmatrix} \frac{bn}{p^2} & 0 \\ 0 & 1 \end{pmatrix} \mathbf{k}_1,$$

we see from Proposition 4.2 that the summand in (5.1) is zero unless $n \in \mathbb{Z}_{\geq 1}$ and $n \equiv 1(p)$. This yields

$$\varphi(g) = y p^{\frac{1}{2}} \sum_{n \equiv 1(p)} a_n n^{\frac{1}{2}} e^{2i\pi nz}, \quad z = x + iy,$$

where a_n are the normalized coefficients as in (1.1).

Choosing $x = 0$, $y = 1$, that is $g_\infty = \mathbf{e}$, we obtain

$$(5.2) \qquad \varphi(g) = p^{\frac{1}{2}} e^{-2\pi} + O(e^{-2\pi p}).$$

It remains to relate $\varphi(g)$ to the values of the classical form f on $\mathfrak{H}$, which we do via the strong approximation $\mathrm{GL}_2(\mathbb{A}) = \mathrm{GL}_2(\mathbb{Q}) \, \mathrm{GL}_2(\mathbb{R})^+ K_0(p^2)$.

We have the decomposition

$$g_p = \begin{pmatrix} \frac{b}{p} & 0 \\ p & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{p} & 0 \\ ap & 1 \end{pmatrix} k,$$

for some $k \in I_2$, where $a \in \mathbb{Z}_p^\times$ is such that $ab \equiv 1(p)$. Let $\gamma := \begin{pmatrix} \frac{1}{p} & 0 \\ ap & 1 \end{pmatrix}$ viewed as an element in $\mathrm{GL}_2(\mathbb{Q})$. Then we have

$$g = \gamma \cdot \gamma_\infty^{-1} g_\infty \cdot k', \quad k' \in K_0(p^2).$$

We compute that $\gamma_\infty^{-1} = \begin{pmatrix} p & 0 \\ -ap^2 & 1 \end{pmatrix}$, thus

$$\gamma_\infty^{-1} \cdot i = \frac{pi}{1 - ap^2 i} = \frac{-1}{p^2 z_\chi}.$$

Since $z \mapsto \frac{-1}{p^2 z}$ is the Atkin–Lehner involution we have that $|\varphi(g)| = |f(z_\chi)|$ which concludes the proof of Theorem 2.2. □

For the Proposition 2.3 we repeat the proof until (5.2). We note that more generally choosing $g_\infty = \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix}$, we have $\varphi(g) \sim y p^{\frac{1}{2}} e^{2i\pi z}$ as long as $y \geq p^{1-\epsilon}$ for some fixed $\epsilon > 0$. Thus $|\varphi(g)| \geq p^\delta$ as long as $p^{\delta - \frac{1}{2}} \leq y \leq 1$. This spans a set of hyperbolic area at least $p^{\frac{1}{2} - \delta}$ (for the measure $\frac{dx dy}{y^2}$ which is preserved by the action of γ_∞^{-1}). □

5.2. Proof of Theorem 1.3. First we need a variant of Lemma 3.2. The formula in Lemma 3.2 was simple because we were working purposely with primitive forms, normalized in such a way that their first coefficient $a_1 = 1$. We now work in the context of Hilbert–Maass forms with the L^2-normalization $\|f\|_2 = 1$. For a place v and $W_v \in \mathcal{W}(\pi_v, \psi_v)$ let

$$(W_v, W_v) := \int_{F_v^\times} \left| W_v\left(\begin{pmatrix} y & 0 \\ 0 & 1 \end{pmatrix}\right) \right|^2 \frac{dy}{y}.$$

Lemma 5.1. *Let f be a Hilbert–Maass newform with $\|f\|_2 = 1$ and let π be the automorphic representation it generates. Let $W = W_\infty \prod_{v \nmid \infty} W_{\circ v}$ the Whittaker function attached to f. Then*

$$\|f\|_\infty^2 \gg_F L(1, \pi, \mathrm{Ad})^{-1} \frac{\max |W_\infty|^2}{(W_\infty, W_\infty)} \prod_{v \nmid \infty} \frac{h(\pi_v)^2 L(1, \pi_v, \mathrm{Ad})}{(W_{\circ v}, W_{\circ v})}$$

Example. If π_v is unramified, then $h(\pi_v) = 1$ and $\zeta_v(2)(W_{\circ v}, W_{\circ v}) = L(1, \pi_v, \mathrm{Ad})$. Thus the product is absolutely convergent.

Proof. The same argument as before based on (3.1) yields

$$\|f\|_\infty \gg \max_{g \in \mathrm{GL}_2(\mathbb{A})} |W(g)| = \max |W_\infty| \prod_{v \nmid \infty} h(\pi_v).$$

From the theory of Rankin–Selberg integrals the following identity [19] holds for some constant $c_F > 0$:

$$\|f\|_2^2 = c_F L(1, \pi, \mathrm{Ad})(W_\infty, W_\infty) \prod_{v \nmid \infty} \frac{(W_{\circ v}, W_{\circ v})}{L(1, \pi_v, \mathrm{Ad})}.$$

Since $\|f\|_2 = 1$ the estimate follows. $\qquad\square$

For almost all integers $k_1, \ldots, k_h, t_1, \ldots t_m \in \mathbb{N}_{\geq 1}$ and integral ideals $\mathfrak{N}$, there are Hilbert-Maass newforms f of level $\mathfrak{N}$, archimedean type (m, h) with weights $(k_1, \ldots, k_h)$ and spectral parameters in the respective intervals $[t_1 - 1, t_1], \ldots, [t_m - 1, t_m]$ such that the local component π_v at every non-archimedean place v is a twist-minimal principal series representation $\chi_1 \boxplus \chi_2$ with χ_1 unramified.

By Iwaniec's convexity bound, we have

$$L(1, \pi, \mathrm{Ad}) \ll_\epsilon (t_1 \cdots t_m)^\epsilon (k_1 \cdots k_h)^\epsilon \mathbf{N}_{F/\mathbb{Q}}(\mathfrak{N})^\epsilon$$

for all $\epsilon > 0$. Assuming $\mathfrak{N}$ is a square ideal, the Proposition 4.4 shows that $h(\pi_v) = N_{F/\mathbb{Q}}(\mathfrak{N}_v)^{\frac{1}{4}}$ for all non-archimedean place v. We can now conclude the proof of Theorem 1.3 using the following two lemmas.

Lemma 5.2. *Let π_v be a twist-minimal principal series representation of $\mathrm{GL}(2, F_v)$. Then $(W_{\circ v}, W_{\circ v}) = 1$ and $L(1, \pi_v, \mathrm{Ad}) = \zeta_v(1)$.*

Proof. The first equality follows from the known formula for $W_{\circ v}(\left(\begin{smallmatrix} y & 0 \\ 0 & 1 \end{smallmatrix}\right))$ in Proposition 4.1. The second equality follows from the fact that $L(1, \pi_v \times \tilde{\pi}_v) = \zeta_v(1)^2$. $\qquad\square$

Lemma 5.3. *Let π be a generic unitary representation of $\mathrm{GL}(2, \mathbb{R})$ and $W \in \mathcal{W}(\pi, \psi)$ be a nonzero lowest weight vector. Then*

$$\frac{\max |W|}{(W, W)^{\frac{1}{2}}} \asymp \begin{cases} k^{\frac{1}{4}}, & \text{if } \pi \text{ is a discrete series of weight } k, \\ r^{\frac{1}{6}}, & \text{if } \pi \text{ is a principal series of spectral parameter } r. \end{cases}$$

Proof. Since W transforms by a unitary character under $SO(2)$, we have by the Iwasawa decomposition $\max |W| = \max_{y>0} \left| W\left(\left(\begin{smallmatrix} y & 0 \\ 0 & 1 \end{smallmatrix} \right) \right) \right|$. We distinguish three cases.

(i) If π is a discrete series of weight k, then $W\left(\left(\begin{smallmatrix} y & 0 \\ 0 & 1 \end{smallmatrix} \right) \right) = y^{\frac{k}{2}} e^{-2\pi y}$, hence:

$$\max |W| = \left(\frac{k}{4\pi} \right)^{\frac{k}{2}} e^{-\frac{k}{2}}, \quad (W, W) = (4\pi)^{-k} \Gamma(k),$$

and the assertion follows by Stirling formula [40].

(ii) If π is a principal series representation with trivial central character then $W\left(\left(\begin{smallmatrix} y & 0 \\ 0 & 1 \end{smallmatrix} \right) \right) = y^{\frac{1}{2}} K_{ir}(2\pi y)$. Hence the assertion follows by the asymptotic behavior of the K-Bessel function in the transition range [30]:

$$\max |W| \asymp r^{\frac{1}{6}} e^{-\frac{\pi r}{2}}, \quad (W, W) = \frac{1}{4} \Gamma_{\mathbb{R}}(1 + 2ir) \Gamma_{\mathbb{R}}(1 - 2ir).$$

(iii) If π is a principal series representation with nontrivial central character then $W\left(\left(\begin{smallmatrix} y & 0 \\ 0 & 1 \end{smallmatrix} \right) \right) = W_{\frac{1}{2}, ir}(4\pi y)$. A uniform asymptotic behavior may be found in [9]:

$$W_{\kappa, ir}(y) \asymp r^{\frac{1}{6}+\kappa} e^{-\frac{\pi r}{2}} \operatorname{Ai}\left(r^{\frac{2}{3}} \eta\left(\frac{y}{r}, \frac{\kappa}{r} \right) \right),$$

for all fixed $\kappa \geq 0$ and uniformly in $y \asymp r$ away from the zeros of the Airy function $\operatorname{Ai}(r^{\frac{2}{3}} \eta)$. Here $\eta(.,.)$ is a continuous function on $\mathbb{R} \times [0, \infty)$ which is increasing in the first variable. We deduce

$$\max |W| \asymp r^{\frac{2}{3}} e^{-\frac{\pi r}{2}}, \quad (W, W) = \frac{\pi \operatorname{Im} \psi(ir)}{\sinh(2\pi r) \Gamma(ir) \Gamma(-ir)} \asymp r e^{-\pi r},$$

which concludes the proof of the lemma. $\qquad\square$

Acknowledgments

References

[1] A. Abbes and E. Ullmo, *Comparaison des métriques d'Arakelov et de Poincaré sur $X_0(N)$*, Duke Math. J. **80** (1995), no. 2, 295–307.

[2] A. O. L. Atkin and J. Lehner, *Hecke operators on $\Gamma_0(m)$*, Math. Ann. **185** (1970), 134–160.

[3] V. Blomer and R. Holowinsky, *Bounding sup-norms of cusp forms of large level*, Invent. Math. **179** (2010), no. 3, 645–681.

[4] V. Blomer and P. Michel, *Sup-norms of eigenfunctions on arithmetic ellipsoids*, Int. Math. Res. Not. IMRN **21** (2011), 4934–4966.

[5] P. Borwein and R. Lockhart, *The expected L_p norm of random polynomials*, Proc. Amer. Math. Soc. **129** (2001), no. 5, 1463–1472 (electronic).

[6] J. Bourgain and E. Lindenstrauss, *Entropy of quantum limits*, Comm. Math. Phys. **233** (2003), no. 1, 153–171.

[7] W. Casselman, *On some results of Atkin and Lehner*, Math. Ann. **201** (1973), 301–314.

[8] P. Deligne, *Formes modulaires et représentations de* GL(2), Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), 1972, pp. 55–105. Lecture Notes in Math., Vol. 349.

[9] T. M. Dunster, *Uniform asymptotic approximations for the Whittaker functions $M_{\kappa,i\mu}(z)$ and $W_{\kappa,i\mu}(z)$*, Anal. Appl. (Singap.) **1** (2003), no. 2, 199–212.

[10] A. Ghosh, A. Reznikov, and P. Sarnak, *Nodal domains of Maass forms I*, Geom. Funct. Anal. **23** (2013), no. 5, 1515–1568.

[11] A. Ghosh and P. Sarnak, *Real zeros of holomorphic Hecke cusp forms*, J. Eur. Math. Soc. (JEMS) **14** (2012), no. 2, 465–487.

[12] G. Harcos and Ph. Michel, *The subconvexity problem for Rankin-Selberg L-functions and equidistribution of Heegner points. II*, Invent. Math. **163** (2006), no. 3, 581–655.

[13] G. Harcos and N. Templier, *On the sup-norm of Maass cusp forms of large level. II*, Int. Math. Res. Not. (2011), Art. ID rnr202, 4764–4774.

[14] _____, *On the sup-norm of Maass cusp forms of large level. III*, Math. Ann. **356** (2013), no. 1, 209–216.

[15] E. Hecke, *Mathematische Werke*, Third Edition, Vandenhoeck & Ruprecht, Göttingen, 1983. With introductory material by B. Schoeneberg, C. L. Siegel and J. Nielsen.

[16] G. Henniart, *Sur l'unicité des types pour GL(2)*, Duke Math. J. **2** (2002), 298–310.

[17] H. Iwaniec and P. Sarnak, L^∞ *norms of eigenfunctions of arithmetic surfaces*, Ann. of Math. (2) **141** (1995), no. 2, 301–320.

[18] H. Iwaniec, W. Luo, and P. Sarnak, *Low lying zeros of families of L-functions*, Inst. Hautes Études Sci. Publ. Math. **91** (2000), 55–131 (2001).

[19] H. Jacquet, *Factorization of period integrals*, J. Number Theory **87** (2001), no. 1, 109–143.

[20] J. Jorgenson and J. Kramer, *Bounding the sup-norm of automorphic forms*, Geom. Funct. Anal. **14** (2004), no. 6, 1267–1277.

[21] J. Jorgenson and J. Kramer, *Bounds on Faltings's delta function through covers*, Ann. of Math. (2) **170** (2009), no. 1, 1–43.

[22] E. Lapid and O. Offen, *Compact unitary periods*, Compos. Math. **143** (2007), no. 2, 323–338.

[23] Y.-K. Lau, *An omega result for supremum norms of Hecke-eigenforms in the level aspect*, Sci. China Math. **53** (2010), no. 1, 223–242.

[24] Ph. Michel and E. Ullmo, *Points de petite hauteur sur les courbes modulaires $X_0(N)$*, Invent. Math. **131** (1998), no. 3, 645–674.

[25] D. Milićević, *Large values of eigenfunctions on arithmetic hyperbolic surfaces*, Duke Math. J. **155** (2010), no. 2, 365–401.

[26] T. Miyake, *Modular forms*, English, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2006. Translated from the 1976 Japanese original by Yoshitaka Maeda.

[27] R. Olofsson, *Large supremum norms and small Shannon entropy for Hecke eigenfunctions of quantized cat maps*, Comm. Math. Phys. **286** (2009), no. 3, 1051–1072.

[28] _____, *Large newforms of the quantized cat map revisited*, Ann. Henri Poincaré **11** (2010), no. 7, 1285–1302.

[29] Z. Rudnick and P. Sarnak, *The behaviour of eigenstates of arithmetic hyperbolic manifolds*, Comm. Math. Phys. **161** (1994), no. 1, 195–213.

[30] P. Sarnak, *Letter to Morawetz (2004); Letter to Reznikov on restrictions of eigenfunctions (2008).*

[31] ———, *Arithmetic quantum chaos*, The Schur lectures (1992) (Tel Aviv), 1995, pp. 183–236.

[32] ———, *Notes on the generalized Ramanujan conjectures*, Harmonic analysis, the trace formula, and Shimura varieties, 2005, pp. 659–685.

[33] J.-P. Serre, *Corps locaux*, Hermann, Paris.

[34] S.-W. Shin, *Automorphic Plancherel density theorem*, Israel J. Math. **192** (2012), no. 1, 83–120.

[35] K. Soundararajan, *Extreme values of zeta and L-functions*, Math. Ann. **342** (2008), no. 2, 467–486.

[36] N. Templier, *Hybrid sup-norm bounds for Hecke-Maass cusp forms*, To appear J. Eur. Math. Soc.

[37] ———, *On the sup-norm of Maass cusp forms of large level*, Selecta Math. (N.S.) **16** (2010), no. 3, 501–531.

[38] J. B. Tunnell, *On the local Langlands conjecture for $GL(2)$*, Invent. Math. **46** (1978), no. 2, 179–200.

[39] J. Weinstein, *Hilbert modular forms with prescribed ramification*, Int. Math. Res. Not. IMRN **8** (2009), 1388–1420.

[40] H. Xia, *On L^∞ norms of holomorphic cusp forms*, J. Number Theory **124** (2007), no. 2, 325–327.

NICOLAS TEMPLIER
DEPARTMENT OF MATHEMATICS
WASHINGTON ROAD
PRINCETON, NJ 08544
USA
E-mail address: templier@math.princeton.edu