J. Inst. Math. Jussieu (2018), 1–69 1
doi:10.1017/S147474801800004X

AN EQUIDISTRIBUTION THEOREM FOR HOLOMORPHIC SIEGEL MODULAR FORMS FOR GSP_4 AND ITS APPLICATIONS

HENRY H. KIM[1,2], SATOSHI WAKATSUKI[3] AND TAKUYA YAMAUCHI[4]

[1]*Department of Mathematics, University of Toronto, Toronto, Ontario M5S 2E4, Canada* (henrykim@math.toronto.edu)

[2]*Korea Institute for Advanced Study, Seoul, Korea*

[3]*Faculty of Mathematics and Physics, Institute of Science and Engineering, Kanazawa University, Kakumamachi, Kanazawa, Ishikawa, 920-1192, Japan* (wakatsuk@staff.kanazawa-u.ac.jp)

[4]*Mathematical Inst. Tohoku Univ. 6-3, Aoba, Aramaki, Aoba-Ku, Sendai 980-8578, Japan* (tyamauchi@m.tohoku.ac.jp)

(Received 1 July 2016; revised 21 November 2017; accepted 22 November 2017)

Abstract We prove an equidistribution theorem for a family of holomorphic Siegel cusp forms for $GSp_4/\mathbb{Q}$ in various aspects. A main tool is Arthur's invariant trace formula. While Shin [Automorphic Plancherel density theorem, *Israel J. Math.* **192**(1) (2012), 83–120] and Shin–Templier [Sato–Tate theorem for families and low-lying zeros of automorphic L-functions, *Invent. Math.* **203**(1) (2016) 1–177] used Euler–Poincaré functions at infinity in the formula, we use a pseudo-coefficient of a holomorphic discrete series to extract holomorphic Siegel cusp forms. Then the non-semisimple contributions arise from the geometric side, and this provides new second main terms A, B_1 in Theorem 1.1 which have not been studied and a mysterious second term B_2 also appears in the second main term coming from the semisimple elements. Furthermore our explicit study enables us to treat more general aspects in the weight. We also give several applications including the vertical Sato–Tate theorem, the unboundedness of Hecke fields and low-lying zeros for degree 4 spinor L-functions and degree 5 standard L-functions of holomorphic Siegel cusp forms.

Keywords: trace formula; Hecke operators; Siegel modular forms

2010 *Mathematics subject classification*: 11F46; 11F70; 22E55; 11R45

Contents

The first author is partially supported by NSERC. The second author is partially supported by JSPS Grant-in-Aid for Scientific Research (nos. 26800006, 25247001, 15K04795). The third author is partially supported by JSPS Grant-in-Aid for Scientific Research (C) no. 15K04787.

1. Introduction

Recently equidistribution theorems for a family of automorphic forms or automorphic representations of a reductive group G over a number field have been studied in various aspects. The basic conceptual studies have been proposed by Sarnak and Serre in the case $G = GL_2$ (cf. [57, 61]) though the case when G has the compact symmetric space has been studied thoroughly. Since then, the trace formula for G has become one of most powerful tools to analyze equidistribution theorems related to a distribution of eigenvalues for a fixed operator acting on a family of automorphic forms for G. After Selberg's celebrated work, the trace formula for Hecke operators on the space of automorphic representations for G whose symmetric space is non-compact, has been developed by many people and it took much time and several stages to reach the current form which is invariant under conjugation (see [2] and the references therein), which is called Arthur's invariant trace formula.

In [62], Shin made good use of Sauvageot's important results [59] to show that the limit of an automorphic counting measure is the Plancherel measure. It implies the equidistribution of Hecke eigenvalues of automorphic forms on G. His key idea is to relate the automorphic counting measure with the spectral side of Arthur's invariant trace formula and then estimate its geometric side. After that, in [63], he and Templier tackled a difficult problem of making it explicit to obtain a power-saving error term for the purpose of an application to low-lying zeros of a family of automorphic L-functions. To do that, as in [62], they used the geometric expansion of the error term mainly consisting of global coefficients, invariant distributions, and orbital integrals. Then they estimated each invariant in a uniform way. One of main difficulties seems to be a uniform boundedness of the orbital integrals. However, since they chose the Euler–Poincaré function at infinity, they had only to consider the semisimple contributions in the geometric side. This would be a usual way to get around the non-semisimple contributions. Their results are fully general as much as possible within current knowledge under certain hypotheses such as Langlands functoriality conjecture but they work on all automorphic forms or automorphic representations which exhaust L-packets of the discrete series representations at infinity.

It is quite natural to consider the automorphic counting measure on automorphic representations with a fixed discrete series representation at infinity. In [62] Shin did not address this problem but he claimed that it may be possible to do so. In this paper we carry it out, namely, we study equidistributions of Hecke eigenvalues of holomorphic Siegel modular forms of degree 2, i.e., cuspidal representations which have a holomorphic discrete series at infinity. The strategy is similar to Shin [62] and Shin–Templier [63], but in addition to the semisimple contributions, we also have to estimate the non-semisimple contributions which have not been understood well in general. This makes the situation more difficult but as a payoff we will be able to observe the meaning of the second main terms A, B_1 coming from the non-semisimple contribution in comparison with the spectral side and a mysterious contribution B_2 also in the second main term from the semisimple part. Furthermore our results are unconditional in contrary to [63]. To explain our main results, we fix our notation.

Let $G = GSp_4$ and S' be a finite set of rational primes. Note that the symbol S will be used for a finite set of places including ∞ in §§ 4 and 5. Let $\mathbb{A}$ (respectively $\mathbb{A}_f$) be the ring of (respectively finite) adeles of $\mathbb{Q}$, $\mathbb{Q}_{S'} = \prod_{p\in S'} \mathbb{Q}_p$, and $\mathbb{A}^{S',\infty} = \prod'_{p\notin S'\cup\{\infty\}} \mathbb{Q}_p$. Let $\widehat{\mathbb{Z}}$ be the profinite completion of $\mathbb{Z}$. We denote by $\widehat{G(\mathbb{Q}_{S'})}$ the unitary dual of $G(\mathbb{Q}_{S'}) = \prod_{p\in S'} G(\mathbb{Q}_p)$ equipped with Fell topology. Put $A_{G,\infty} = Z_G(\mathbb{R})^\circ \simeq \mathbb{R}_{>0}$. Fix a Haar measure $\mu^{S',\infty}$ of $G(\mathbb{A}^{S',\infty})$ so that $\mu^{S',\infty}(G(\widehat{\mathbb{Z}}^{S'})) = 1$, and let U be a compact open subgroup of $G(\mathbb{A}^{S',\infty})$.

Consider the algebraic representation $\xi = \xi_{\underline{k}}$ for $\underline{k} = (k_1, k_2)$, $k_1 \geqslant k_2 \geqslant 3$ as in (3.5), and let $D^{\mathrm{hol}}_{l_1,l_2}$ be the holomorphic discrete series of $G(\mathbb{R})$ with the Harish–Chandra parameter $(l_1, l_2) = (k_1 - 1, k_2 - 2)$, and whose central character equal to $\chi_{\xi^\vee}$ on $A_{G,\infty}$. We choose the test function $f_{S'} = f_\xi f_U$ such that f_ξ is a pseudo-coefficient of $D^{\mathrm{hol}}_{l_1,l_2}$. Then we define a measure on $\widehat{G(\mathbb{Q}_{S'})}$ by

$$\widehat{\mu}_{U,\xi_{\underline{k}},D^{\mathrm{hol}}_{l_1,l_2}} := \frac{1}{\mathrm{vol}(G(\mathbb{Q})A_{G,\infty}\backslash G(\mathbb{A}))\cdot \dim \xi_{\underline{k}}} \times \sum_{\pi^0_{S'}\in \widehat{G(\mathbb{Q}_{S'})}} \mu^{S',\infty}(U) m_{\mathrm{cusp}}(\pi^0_{S'}; U, \xi_{\underline{k}}, D^{\mathrm{hol}}_{l_1,l_2}) \delta_{\pi^0_{S'},\xi}, \tag{1.1}$$

where $\delta_{\pi^0_{S'},\xi}$ is a normalized Dirac delta measure supported on $\pi^0_{S'}$ with respect to the Plancherel measure $\widehat{\mu}^{\mathrm{pl}}_{S'}$ on $\widehat{G(\mathbb{Q}_{S'})}$ (see (3.12)), and for a given unitary representation $\pi^0_{S'}$ of $G(\mathbb{Q}_{S'})$,

$$m_{\mathrm{cusp}}(\pi^0_{S'}; U, \xi_{\underline{k}}, D^{\mathrm{hol}}_{l_1,l_2}) = \sum_{\substack{\pi\in\Pi(G(\mathbb{A})) \\ \pi_{S'}\simeq\pi^0_{S'},\ \pi_\infty\simeq D^{\mathrm{hol}}_{l_1,l_2}}} m_{\mathrm{cusp}}(\pi)\mathrm{tr}(\pi^{S',\infty}(f_U))\cdot \mathrm{tr}(\pi_\infty(f_\xi)). \tag{1.2}$$

To state the equidistribution theorem, we need to introduce the Hecke algebra $C^\infty_c(G(\mathbb{Q}_{S'}))$ which is dense under the map $f_{S'} \mapsto \widehat{f}_{S'} = [\widehat{f}_{S'} : \pi_{S'} \mapsto \mathrm{tr}\pi_{S'}(f_{S'})]$ in a reasonable space $\mathcal{F}(\widehat{G(\mathbb{Q}_{S'})})$ consisting of suitable $\widehat{\mu}^{\mathrm{pl}}_{S'}$-measurable functions on $\widehat{G(\mathbb{Q}_{S'})}$ (see [62, § 2.3] for that space). Put $K_p = G(\mathbb{Z}_p)$. The spherical Hecke algebra $H^{\mathrm{ur}}(G(\mathbb{Q}_p))$ is defined by the subalgebra consisting of K_p-bi-invariant functions in $C^\infty_c(G(\mathbb{Q}_p))$. It is well known that

$$H^{\mathrm{ur}}(G(\mathbb{Q}_p)) = \mathbb{C}[h_{a_1,a_2,a_3} \mid a_1, a_2, a_3 \in \mathbb{Z},\quad a_3 \geqslant a_1 \geqslant a_2 \geqslant 0]$$

where h_{a_1,a_2,a_3} is the characteristic function of $K_p\mathrm{diag}(p^{-a_1}, p^{-a_2}, p^{a_1-a_3}, p^{a_2-a_3})K_p$. For any $\kappa \in \mathbb{Z}_{\geqslant 0}$ we denote by $H^{\mathrm{ur}}(G(\mathbb{Q}_p))^\kappa$ the $\mathbb{C}$-span of the functions h_{a_1,a_2,a_3} satisfying $a_3 \leqslant \kappa$.

In this paper, we restrict ourselves to $U = K(N)$ for $N \in \mathbb{Z}_{>0}$ which is the kernel of the natural quotient map from $G(\widehat{\mathbb{Z}})$ to $G(\widehat{\mathbb{Z}}/N\widehat{\mathbb{Z}})$. Put $\Gamma(N) = Sp_4(\mathbb{Q}) \cap K(N)$. Let $S' = \{p\}$ for $p \nmid N$. Let $\widehat{\mu}^{\mathrm{pl}}_p = \widehat{\mu}^{\mathrm{pl}}_{\{p\}}$. Then our main theorem is as follows:

Theorem 1.1. *Let $\{(K(N), \xi_{\underline{k}})\}$ be a family of weights and levels so that $p \nmid N$, $k_1 \geqslant k_2 \geqslant 3$ and $N + k_1 + k_2 \longrightarrow \infty$. For any $f \in H^{\mathrm{ur}}(G(\mathbb{Q}_p))^\kappa$,*

$$\lim_{k_1+k_2+N\to\infty} \widehat{\mu}_{K(N),\xi_{\underline{k}},D^{\mathrm{hol}}_{l_1,l_2}}(\widehat{f}) = \widehat{\mu}^{\mathrm{pl}}_p(\widehat{f}).$$

More precisely, there exist constants a, b, a' and b' depending only on G such that

(1) *(level aspect) Fix k_1, k_2. Then for $N \gg p^{10\kappa}$,*

$$\widehat{\mu}_{K(N),\xi_{\underline{k}},D^{\mathrm{hol}}_{l_1,l_2}}(\widehat{f}) = \widehat{\mu}^{\mathrm{pl}}_p(\widehat{f}) + A + O(p^{a\kappa+b}\varphi(N)N^{-3}), \quad A = O(p^{\kappa}\varphi(N)N^{-2}),$$

where φ stands for Euler's phi function:

(2) *(weight aspect) Fix N. Then as $k_1 + k_2 \to \infty$,*

$$\widehat{\mu}_{K(N),\xi_{\underline{k}},D^{\mathrm{hol}}_{l_1,l_2}}(\widehat{f}) = \widehat{\mu}^{\mathrm{pl}}_p(\widehat{f}) + B_1 + B_2 + O\left(\frac{p^{a'\kappa+b'}}{(k_1-k_2+1)(k_1-1)(k_2-2)}\right),$$

$$B_1 = O\left(\frac{p^{\kappa}}{(k_1-1)(k_2-2)}\right), \quad B_2 = O\left(\frac{p^{\kappa}}{(k_1-k_2+1)(k_1+k_2-3)}\right).$$

Remark 1.2. Here the second main terms A, B_1 come from non-semisimple contributions ($I_2(f)$ in Propositions 5.2 and 5.3), while B_2 comes from semisimple contributions ($I_3(f)$ in Proposition 5.3).

We explain several applications of Theorem 1.1. First, we consider the vertical Sato–Tate theorem which is formulated in terms of the classical setting. Let $S_{\underline{k}}(\Gamma(N), \chi)$ be the space of classical holomorphic Siegel cusp forms of level $\Gamma(N)$ with a central character $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \longrightarrow \mathbb{C}^\times$ and weight $\underline{k} = (k_1, k_2), k_1 \geqslant k_2 \geqslant 3$ (see (2.12) of §2.1). For a prime $p \nmid N$, let $T(p^n)$ be the Hecke operator with the similitude p^n (see §2.2). Any eigenform with respect to $T(p^n)$ for any non-negative integer n and any prime $p \nmid N$ is called a Hecke eigen cusp form. Let $HE_{\underline{k}}(\Gamma(N), \chi)$ be a basis of $S_{\underline{k}}(\Gamma(N), \chi)$ consisting of Hecke eigen forms outside N. Let $S_{\underline{k}}(\Gamma(N), \chi)^{\mathrm{tm}}$ be the subspace of $S_{\underline{k}}(\Gamma(N), \chi)$ generated by any Hecke eigen form F outside N so that $\pi_{F,p}$ is tempered for any $p \nmid N$. Put $HE_{\underline{k}}(\Gamma(N), \chi)^{\mathrm{tm}} = S_{\underline{k}}(\Gamma(N), \chi)^{\mathrm{tm}} \cap HE_{\underline{k}}(\Gamma(N), \chi)$. For each $p \nmid N$, we fix a square root $\chi(p)^{1/2}$ in $\mathbb{C}$ and we write $\chi(p)^{-1/2}$ for $(\chi(p)^{1/2})^{-1}$. For a Hecke eigen form $F \in S_{\underline{k}}(\Gamma(N), \chi)^{\mathrm{tm}}$, the Satake parameter of $\pi_{F,p}$ at $p \nmid N$ is given by $\{\alpha_{0p}, \alpha_{0p}\alpha_{1p}, \alpha_{0p}\alpha_{1p}, \alpha_{0p}\alpha_{1p}\alpha_{2p}\}$. Since $\alpha_{0p}^2\alpha_{1p}\alpha_{2p} = \chi(p)^2$, if we let $\alpha_{F,p} = \alpha_{0p}$ and $\beta_{F,p} = \alpha_{0p}\alpha_{1p}$, it can be written as $\{\alpha^{\pm}_{F,p}, \beta^{\pm}_{F,p}\}$. Then it follows from the temperedness that if we set

$$a_{F,p} := \alpha_{F,p}\chi(p)^{-1/2} + \alpha^{-1}_{F,p}\chi(p)^{1/2}, \quad b_{F,p} := \beta_{F,p}\chi(p)^{-1/2} + \beta^{-1}_{F,p}\chi(p)^{1/2},$$

then

$$a_{F,p}, b_{F,p} \in [-2, 2].$$

We introduce a suitable measure

$$\mu_p = f_p(x, y)g^+_p(x, y)g^-_p(x, y) \cdot \mu^{\mathrm{ST}}_\infty$$

on $\Omega := [-2, 2]^2/\mathfrak{S}_2$ (the non-trivial element in $\mathfrak{S}_2$ acts by $(x, y) \mapsto (y, x)$ on $[-2, 2]^2$), where

$$f_p(x, y) = \frac{(p+1)^2}{\left(\left(\sqrt{p}+\frac{1}{\sqrt{p}}\right)^2 - x^2\right)\left(\left(\sqrt{p}+\frac{1}{\sqrt{p}}\right)^2 - y^2\right)},$$

$$\mu_{\infty}^{\rm ST} = \frac{(x-y)^2}{\pi^2}\sqrt{1-\frac{x^2}{4}}\sqrt{1-\frac{y^2}{4}},$$
$$g_p^{\pm}(x,y) = \frac{p+1}{\left(\sqrt{p}+\frac{1}{\sqrt{p}}\right)^2 - 2\left(1+\frac{xy}{4}\pm\sqrt{1-\frac{x^2}{4}}\sqrt{1-\frac{y^2}{4}}\right)}.$$

Note that the denominator of $g^+(x,y)g^-(x,y)$ is $x^2+y^2-xy(p+p^{-1})-4+(p+p^{-1})^2$. Let $C^0(\Omega,\mathbb{R})$ be the space of $\mathbb{R}$-valued continuous functions on Ω. To control non-tempered part of $S_{\underline{k}}(\Gamma(N))$ we need to assume that $(N, 11!) = 1$ which might be unnecessary. Then we have

Theorem 1.3. *Let $p \nmid N$, $k_1 \geqslant k_2 \geqslant 3$, and $N+k_1+k_2 \longrightarrow \infty$ satisfying $(N, 11!) = 1$. Put $d_{\underline{k},N}^{\rm tm}(\chi) = \dim S_{\underline{k}}(\Gamma(N),\chi)^{\rm tm}$. Then the set*

$$\{(a_{F,p}, b_{F,p}) \in \Omega \mid F \in HE_{\underline{k}}(\Gamma(N),\chi)^{\rm tm}\}$$

is μ_p-equidistributed in Ω, namely, for any $f \in C^0(\Omega,\mathbb{R})$,

$$\lim_{\substack{N+k_1+k_2\to\infty \\ p\nmid N,\ (N,11!)=1}} \frac{1}{d_{\underline{k},N}^{\rm tm}(\chi)} \sum_{F\in HE_{\underline{k}}(\Gamma(N),\chi)^{\rm tm}} f(a_{F,p}, b_{F,p}) = \int_{\Omega} f(x,y)\mu_p.$$

It follows immediately from this that we obtain a special case of [64, Theorem 6.1].

Corollary 1.4. *Let the notations be in Theorem 1.3. Fix a weight $\underline{k} = (k_1, k_2)$ with $k_1 \geqslant k_2 \geqslant 3$ and a prime p. Then*

$$\limsup_{\substack{N\to\infty \\ p\nmid N,\ (N,11!)=1}} \{[\mathbb{Q}_F : \mathbb{Q}] \mid F \in HE_{\underline{k}}(\Gamma(N),\chi)^{\rm tm}\} = \infty$$

where $\mathbb{Q}_F$ is the Hecke field of F in Definition 2.14.

The space $S_{\underline{k}}(\Gamma(N),\chi)^{\rm tm}$ contains endoscopic lifts from elliptic modular forms which are called Yoshida lifts. The above corollary is still true even if we restrict F to be a non-endoscopic lift but as mentioned before we set a condition on the level so that N is coprime to 11! to control the conductor under the functoriality:

Corollary 1.5. *Let the notations be as in Theorem 1.3. Then*

$$\limsup_{\substack{N\to\infty \\ p\nmid N,\ (N,11!)=1}} \{[\mathbb{Q}_F : \mathbb{Q}] \mid F \in HE_{\underline{k}}(\Gamma(N),\chi)^{\rm tm} : \textit{non-endoscopic}\} = \infty.$$

In the above corollary it is also interesting to remove among $HE_{\underline{k}}(\Gamma(N),\chi)^{\rm tm}$ the base change lift from GL_2 over a quadratic field of $\mathbb{Q}$ and the symmetric cubic lift other than endoscopic lifts. We treat it in a forthcoming paper.

Next, we consider the distribution of the low-lying zeros of either spinor or standard L-functions for our family. For simplicity, denote $S_{\underline{k}}(\Gamma(N),1)$, $HE_{\underline{k}}(\Gamma(N),1)$ by $S_{\underline{k}}(N)$, $HE_{\underline{k}}(N)$, respectively. For $F \in S_{\underline{k}}(N)$ we denote the non-trivial zeros of $L(s,\pi_F,*)$, $* \in \{\text{Spin}, \text{St}\}$ by $\sigma_F = \frac{1}{2}+\sqrt{-1}\gamma_F$. We do not assume GRH, and hence γ_F can be a complex number. Let ϕ be a Schwartz function which is even and whose transform has a compact support (so ϕ extends to an entire function). Define

$$D(\pi_F, \phi, *) = \sum_{\gamma_F} \phi\left(\frac{\gamma_F}{2\pi} \log c_{\underline{k},N}\right),$$

where $\log c_{\underline{k},N} = \dfrac{1}{\dim S_{\underline{k}}(N)} \displaystyle\sum_{F \in HE_{\underline{k}}(\Gamma(N))} \log c(F, *)$ and $c(F, *)$ stands for the analytic conductor (see § 8). Then we prove

Theorem 1.6. *Let the notations be as in Theorem 1.3. Put $d_{\underline{k},N} = \dim S_{\underline{k}}(N)$. Let ϕ be a Schwartz function which is even and whose Fourier transform has a support sufficiently smaller than $(-1, 1)$. Then*

(1) $$\lim_{\substack{N+k_1+k_2\to\infty \\ (N,11!)=1}} \frac{1}{d_{\underline{k},N}} \sum_{F \in HE_{\underline{k}}(N)} D(\pi_F, \phi, \mathrm{Spin}) = \hat{\phi}(0) + \frac{1}{2}\phi(0) = \int_{\mathbb{R}} \phi(x) W(G)(x)\, dx,$$

where $G = SO(\mathrm{even})$, $SO(\mathrm{odd})$, or O type.

(2) $$\lim_{\substack{N+k_1+k_2\to\infty \\ (N,11!)=1}} \frac{1}{d_{\underline{k},N}} \sum_{F \in HE_{\underline{k}}(N)} D(\pi_F, \phi, \mathrm{St}) = \hat{\phi}(0) - \frac{1}{2}\phi(0) = \int_{\mathbb{R}} \phi(x) W(Sp)(x)\, dx,$$

where the corresponding density functions $W(G)$ are

$$W(SO(\mathrm{even}))(x) = 1 + \frac{\sin 2\pi x}{2\pi x}, \quad W(SO(\mathrm{odd}))(x) = 1 - \frac{\sin 2\pi x}{2\pi x} + \delta_0(x),$$

$$W(O)(x) = 1 + \frac{1}{2}\delta_0(x), \text{ and } W(Sp)(x) = 1 - \frac{\sin 2\pi x}{2\pi x}.$$

Remark 1.7. We stated our result only for $\Gamma(N)$ for simplicity. In weight aspect we expect that our result holds for other congruence subgroups such as $\Gamma_0(N) = \left\{\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \middle| C \equiv 0\ (N)\right\}$.

Remark 1.8. Kowalski–Saha–Tsimerman [36] (in level one case) and Dickson [18] considered weighted one-level density of spinor L-functions of scalar-valued Siegel cusp forms, namely, let $\mathcal{F}_k(N)$ be a basis of the space of Siegel eigen cusp forms of weight k with respect to $\Gamma_0(N)$. Then

$$\lim_{N+k\to\infty} \frac{1}{\displaystyle\sum_{F \in \mathcal{F}_k(N)} \omega_{F,N,k}} \sum_{F \in \mathcal{F}_k(N)} \omega_{F,N,k} D(\pi_F, \phi, \mathrm{Spin}) = \hat{\phi}(0) - \frac{1}{2}\phi(0)$$

$$= \int_{\mathbb{R}} \phi(x) W(Sp)(x)\, dx,$$

where

$$\omega_{F,N,k} = \frac{\sqrt{\pi}(4\pi)^{3-2k}\Gamma(k-\frac{3}{2})\Gamma(k-2)|A(F;E_2)|^2}{\mathrm{vol}(\Gamma_0(N)\backslash\mathbb{H}_2)4\langle F, F\rangle},$$

and $F(Z) = \sum_{T>0} A(F;T)e^{2\pi i Tr(TZ)}$. So the symmetry type is Sp. Notice that the symmetry type is changed due to the weighted sum.

Remark 1.9. If we apply the main results in Shin–Templier [63] for $G = GSp_4$, one-level density is for a family of all cuspidal representations whose infinity types are in the local L-packet consisting of both holomorphic discrete series $D^{\mathrm{hol}}_{l_1,l_2}$ and the large discrete series $D^{\mathrm{large}}_{l_1,l_2}$. Hence their family consists of both holomorphic Siegel cusp forms and non-holomorphic forms. The global L-packet of a holomorphic Siegel cusp form always contains a non-holomorphic form. However, there are cuspidal representations with the large discrete series at the infinity whose global L-packet does not contain any cuspidal representation with a holomorphic discrete series at infinity.

Let $\widetilde{S}_{\underline{k}}(N)$ be the set of isomorphic classes of cuspidal representations with the given infinity type in the local L-packet $\{D^{\mathrm{hol}}_{l_1,l_2}, D^{\mathrm{large}}_{l_1,l_2}\}$. Then Shin–Templier showed that

$$\lim_{k_1+k_2+N\to\infty} \frac{1}{|\widetilde{S}_{\underline{k}}(N)|} \sum_{\pi\in\widetilde{S}_{\underline{k}}(N)} D(\pi,\phi,*) = \hat{\phi}(0) \pm \frac{1}{2}\phi(0) = \int_{\mathbb{R}} \phi(x)W(G)(x)\,dx,$$

where $\pm$ is according to $* =$ Spin or St, respectively and G is as in Theorem 1.6. The symmetry type is the same because the L-functions of representations in the same L-packet remain the same. Also we can see that the contribution from endoscopic non-holomorphic forms is negligible, and in fact, we expect that it matches with non-semisimple contributions from the geometric side. We elaborate it more in § 10.

Some experts in the trace formula may be wondering why we did not use the stable trace formula for proving the main theorem. A reason is that even if we use the stable trace formula for $G = GSp_4$, we would have to take care of the fundamental lemma for Hecke elements under several transfers which might be feasible, but this gives rise to a conditional result in contrast to our main theorem. However, the method of the stabilization is still important in our proof. In fact, our argument of estimating non-semisimple terms is essentially the same as the stabilization.

This paper is organized as follows. In § 2, we recall the correspondence between classical holomorphic Siegel cusp forms and their adelic forms and their associated cuspidal automorphic representations of GSp_4. We also recall various subspaces and their dimensions.

In § 3, we give a precise description of the spectral decomposition and residual spectrum and classification of CAP forms. Even though we consider a pseudo-coefficient of holomorphic discrete series, there are non-trivial contributions from the residual spectrum and CAP forms. In § 3.2, we quickly recall the classification of the algebraic representations of $GSp_4(\mathbb{R})$ and adjust a central character to match with the holomorphic discrete series in question. We also define a normalized (automorphic) counting measures which will be related to the Plancherel measure with error terms. In § 4, we estimate the global coefficients, invariant distributions which are (limit) character formulas of holomorphic discrete series, and orbital integrals for spherical elements. The geometric side will be decomposed into seven main terms according to the shape of conjugacy classes. In § 5, by using results in previous section we estimate those seven terms. In § 6, we give a proof of the main theorems. Their interpretation in terms of classical Siegel modular forms will be given in § 7. In § 8, we review the spinor L-function and the standard L-function of automorphic representations for GSp_4 whose infinite component is the holomorphic discrete series and estimate the conductor which shows up in the functional equation. Then using the main theorem, we can estimate the sum of the coefficients of automorphic L-functions. In § 9, we apply results of § 8 to obtain the one-level density result for degree 4 spin L-functions and degree 5 standard L-functions.

In § 10, we compare Shin's work with ours to explain the meaning of the second main terms in terms of automorphic representations. An analytic property of an infinite sum which is necessary in § 9 will be proved in the appendix.

2. Preliminaries for holomorphic Siegel modular forms

In this section, we recall holomorphic Siegel modular forms of genus 2. We refer to [1, 25, 68] for the classical setting and [8] for the adelic setting. First we fix our notations. For any commutative ring R with a unit, let $M_n(R)$ be the algebra consisting of all square matrices over R with size n and denote by E_n (respectively 0_n) the identity matrix (respectively zero matrix). We write $\mathrm{diag}(a_1, \dots, a_n) \in M_n(R)$ for the diagonal matrix whose entries are $a_1, \dots, a_n \in R$. We also write ${}^tX$ for the transpose of $X \in M_n(R)$.

We define the generalized symplectic group by

$$G = GSp_4 = \left\{ X \in GL_4 \;\middle|\; {}^tX \begin{pmatrix} 0_2 & E_2 \\ -E_2 & 0_2 \end{pmatrix} X = \nu(X) \begin{pmatrix} 0_2 & E_2 \\ -E_2 & 0_2 \end{pmatrix}, \;\; \nu(X) \in GL_1 \right\}$$

which is a smooth algebraic group scheme over $\mathbb{Z}$. The similitude ν defines the character $\nu : G \longrightarrow GL_1$ and defines the symplectic group $Sp_4 := \mathrm{Ker}(\nu)$ which is of rank 2.

Let B be the standard Borel subgroup consisting of upper triangular matrices and T be the split diagonal torus in B. For $i = 1, 2$, let $P_i = M_i N_i$ be the parabolic subgroup of $G = GSp_4$ defined by

$$M_1 = \left\{ \begin{pmatrix} A & 0 \\ 0 & u\,{}^tA^{-1} \end{pmatrix} \;\middle|\; A \in GL_2, u \in GL_1 \right\} \simeq GL_2 \times GL_1,\; N_1 = \left\{ \begin{pmatrix} E_2 & S \\ 0 & E_2 \end{pmatrix} \;\middle|\; S = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \right\},$$

$$N_2 = \left\{ \begin{pmatrix} E_2 & A \\ 0 & E_2 \end{pmatrix} \begin{pmatrix} C & 0 \\ 0 & C' \end{pmatrix} \;\middle|\; A = \begin{pmatrix} a & b \\ b & 0 \end{pmatrix}, C = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}, C' = \begin{pmatrix} 1 & 0 \\ -d & 1 \end{pmatrix} \right\},$$

$$M_2 = \left\{ \begin{pmatrix} t & & & \\ & a & & b \\ & & \det(g)/t & \\ & c & & d \end{pmatrix} \;\middle|\; t \in GL_1, \;\; g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2 \right\}.$$

We sometimes use the theory of elliptic modular forms on the upper half-plane $\mathbb{H}_1$ with respect to the following congruence subgroups of $SL_2(\mathbb{Z})$:

$$\Gamma^1(N) = \{g \in SL_2(\mathbb{Z}) \mid g \equiv E_2 \bmod N\},$$

$$\Gamma_1^1(N) = \left\{ g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \;\middle|\; a - 1 \equiv c \equiv 0 \bmod N \right\},$$

$$\Gamma_0^1(N) = \left\{ g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \;\middle|\; c \equiv 0 \bmod N \right\}.$$

2.1. Classical Siegel modular forms

Let $\mathcal{H}_2 = \{Z \in M_2(\mathbb{C}) |\ {}^tZ = Z,\ \mathrm{Im}(Z) > 0\}$ be the Siegel upper half-plane of degree 2. For a pair of non-negative integers $\underline{k} = (k_1, k_2)$, $k_1 \geqslant k_2 \geqslant 0$, we define the algebraic representation $\lambda_{\underline{k}}$ of GL_2 with the highest weight $\underline{k}$ by

$$V_{\underline{k}} = \mathrm{Sym}^{k_1-k_2}\mathrm{St}_2 \otimes \det{}^{k_2}\mathrm{St}_2,$$

where St_2 is the standard representation of dimension 2 with the basis $\{e_1, e_2\}$. More explicitly, if R is any ring, then $V_{\underline{k}}(R) = \bigoplus_{i=0}^{k_1-k_2} Re_1^{k_1-k_2-i} \cdot e_2^i$ and for $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(R)$, $\lambda_{\underline{k}}(g)$ acts on $V_{\underline{k}}(R)$ by

$$g \cdot e_1^{k_1-k_2-i} \cdot e_2^i := \det(g)^{k_2}(ae_1 + ce_2)^{k_1-k_2-i} \cdot (be_1 + de_2)^i.$$

We identify $V_{\underline{k}}(R)$ with $R^{\oplus(k_1-k_2+1)}$, and $\lambda_{\underline{k}}(g)$ with the representation matrix of $\lambda_{\underline{k}}(g)$ with respect to the above basis.

We have the action and the automorphy factor $J(\gamma, Z)$ by

$$\gamma Z = (AZ + B)(CZ + D)^{-1}, \quad J(\gamma, Z) = CZ + D \in GL_2(\mathbb{C}), \tag{2.1}$$

for $\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in GSp_4(\mathbb{R})^+$ and $Z \in \mathcal{H}_2$.

For an integer $N \geqslant 1$, we define a principal congruence subgroup $\Gamma(N)$ to be the group consisting of the elements $g \in Sp_4(\mathbb{Z})$ such that $g \equiv 1 \bmod N$. For a $V_{\underline{k}}(\mathbb{C})$-valued function F on $\mathcal{H}_2$, the action of $\gamma \in GSp_4(\mathbb{R})^\circ$ is defined by

$$F(Z)|[\gamma]_{\underline{k}} := \lambda_{\underline{k}}(\nu(\gamma)J(\gamma, z)^{-1})F(\gamma Z). \tag{2.2}$$

The algebra of all $Sp_4(\mathbb{R})$-invariant differential operators on $\mathcal{H}_2$ is isomorphic to $\mathbb{C}[\Omega, \Delta]$, the commutative polynomial ring of two variables, where Ω is the degree 2 Casimir element, and Δ is the degree 4 element. (See [34, § 5] for the details and $\Omega = \Delta_1$, $\Delta = \Delta_2$ in the notation there.) It is easy to see that

$$\Omega F = \tfrac{1}{12}((k_1 - 1)^2 + (k_2 - 2)^2 - 5)F, \quad \Delta F = ((k_1 - 1)(k_2 - 2))^2 F, \tag{2.3}$$

for all $V_{\underline{k}}(\mathbb{C})$-valued holomorphic function F on $\mathcal{H}_2$ when $k_1 \geqslant k_2 \geqslant 0$. When $(k_1, k_2) = (3, 3)$, $(3, 1)$, or $(0, 0)$, the two eigenvalues are 0 and 4. It is similar for $(k_1, 3)$ and $(k_1, 1)$. However, if $k_2 > 3$, then there is no weight other than (k_1, k_2) itself with the same eigenvalues. This observation would be related to the sets in (4.8).

Suppose that $k_2 \geqslant 3$ and let us introduce the Harish–Chandra parameter (l_1, l_2) for the holomorphic discrete series generated by (non-zero) F of weight (k_1, k_2). It has the relation $(k_1, k_2) = (l_1 + 1, l_2 + 2)$. Then in terms of the Harish–Chandra parameter we rewrite (2.3) as follows:

$$\Omega F = \tfrac{1}{12}(l_1^2 + l_2^2 - 5)F, \quad \Delta F = (l_1 l_2)^2 F. \tag{2.4}$$

Let us turn to the general case of $k_1 \geqslant k_2 \geqslant 0$. For a principal congruence subgroup $\Gamma(N)$, $N \geqslant 1$, we say that a holomorphic function $F : \mathcal{H}_2 \longrightarrow V_{\underline{k}}(\mathbb{C})$ is a holomorphic

Siegel modular form of weight (k_1, k_2) with respect to $\Gamma(N)$ if it satisfies $F|[\gamma]_{\underline{k}} = F$ for any $\gamma \in \Gamma(N)$. If we further impose the following condition, we call F a holomorphic Siegel cusp form:

$$\lim_{Z \to \partial \mathcal{H}_2} F|[\gamma]_{\underline{k}}(Z) = 0 \quad \text{for any } \gamma \in Sp_4(\mathbb{Q})$$

where $\partial\mathcal{H}_2$ stands for the boundary of the Satake compactification of $\mathcal{H}_2$. We denote by $M_{\underline{k}}(\Gamma(N))$ (respectively $S_{\underline{k}}(\Gamma(N))$) the space of such holomorphic Siegel modular (respectively cusp) forms. The space $E_{\underline{k}}(\Gamma(N))$ of Eisenstein series is defined by the orthogonal complement of $S_{\underline{k}}(\Gamma(N))$ in $M_{\underline{k}}(\Gamma(N))$ with respect to Petersson inner product. Hence we have

$$M_{\underline{k}}(\Gamma(N)) = S_{\underline{k}}(\Gamma(N)) \oplus E_{\underline{k}}(\Gamma(N)).$$

By Harish–Chandra (see [8, Theorem 1.7]) the space $M_{\underline{k}}(\Gamma(N))$ is finite-dimensional. We give an estimation of the dimension of $S_{\underline{k}}(\Gamma(N))$ and of its specific subspace later on.

The group $\Gamma(N)$ contains the subgroup consisting of $\begin{pmatrix} E_2 & NS \\ 0 & E_2 \end{pmatrix}$, $S = {}^tS \in M_2(\mathbb{Z})$. Hence for a given $F \in M_{\underline{k}}(\Gamma(N))$, we have the Fourier expansion

$$F(Z) = \sum_{T \in \mathrm{Sym}^2(\mathbb{Z})_{\geqslant 0}} A_F(T) e^{(2\pi\sqrt{-1}/N)\mathrm{tr}(TZ)}, \quad Z \in \mathcal{H}_2, \tag{2.5}$$

where $\mathrm{Sym}^2(\mathbb{Z})_{\geqslant 0}$ is the subset of $M_2(\mathbb{Q})$ consisting of all symmetric matrices $\begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}$, $a, b, c \in \mathbb{Z}$, which are semi-positive definite.

2.2. Hecke operators

We define the Hecke operators on $M_{\underline{k}}(\Gamma(N))$ as in [19]: For any positive integer n coprime to N, let

$$\Delta_n(N) := \left\{ g \in M_4(\mathbb{Z}) \cap GSp_4(\mathbb{Q}) \;\middle|\; g \equiv \begin{pmatrix} E_2 & 0 \\ 0 & \nu(g)E_2 \end{pmatrix} \bmod N,\ \nu(g) = n \right\}.$$

For $m \in \Delta_n(N)$, we introduce the action of the Hecke operators on $M_{\underline{k}}(\Gamma(N))$ by

$$T_m F(Z) := \nu(m)^{(k_1+k_2)/2-3} \sum_{\alpha \in \Gamma(N)\backslash\Gamma(N)m\Gamma(N)} F(Z)|[(\nu(m)^{-1/2}\alpha]_{\underline{k}}, \tag{2.6}$$

and for any positive integer n, put

$$T(n) := \sum_{m \in \Gamma(N)\backslash\Delta_n(N)} T_m.$$

These actions preserve the space $S_{\underline{k}}(\Gamma(N))$. For $t_1 = \mathrm{diag}(1, 1, p, p)$, $t_2 = \mathrm{diag}(1, p, p^2, p)$ for a prime p, put $T_{j,p^k} := T_{t_j^k}$ $j = 1, 2$, $k \in \mathbb{Z}_{>0}$ and fix $\widetilde{S}_{p^i,1}, \widetilde{S}_{p^i,p^i} \in Sp_4(\mathbb{Z})$ so that $\widetilde{S}_{p^i,1} \equiv \mathrm{diag}(p^{-i}, 1, 1, p^i) \bmod N$ and $\widetilde{S}_{p^i,p^i} \equiv \mathrm{diag}(p^{-i}, p^{-i}, p^i, p^i) \bmod N$ for each $i \in \mathbb{Z}_{>0}$. Put $R_{p^i} := \widetilde{S}_{p^i,p^i} T_{p^i E_4} = p^{i(k_1+k_2-6)} \widetilde{S}_{p^i,p^i}$ and note that it commutes with any Hecke operator. Then we see that

$$\begin{aligned} &T(p)=T_{1,p}, \quad T(p^2)=T_{1,p^2}+T_{2,p}+R_p,\\ &T_{1,p}^2-T(p^2)-p^2R_p=p(T_{2,p}+(1+p^2)R_p),\\ &T_{2,p}^2=T_{\mathrm{diag}(1,p^2,p^4,p^2)}+(p+1)T_{\mathrm{diag}(p,p,p^3,p^3)}\\ &\qquad +(p^2-1)T_{\mathrm{diag}(p,p^2,p^3,p^2)}+(p^4+p^3+p+1)R_{p^2}. \end{aligned} \tag{2.7}$$

The last relation is obtained as follows: First, we note that $M \in T_{2,p}$ if and only if $r_p(M)=1$, where $r_p(M)$ is the rank of M mod p. Then

$$T_{2,p}^2=\sum_M t(KMK)KMK,$$

where M runs over all double coset representatives of $K\backslash\Delta_{p^4}(N)/K$, and $t(KMK)$ is the number of left coset representatives A of $T_{2,p}/K$ such that $A^{-1}M \in T_{2,p}$, i.e., $r_p(A^{-1}M)=1$. Now M runs over the following elements;

$$\begin{aligned} &\mathrm{diag}(1,1,p^4,p^4),\ \mathrm{diag}(1,p,p^4,p^3),\ \mathrm{diag}(1,p^2,p^4,p^2),\\ &\mathrm{diag}(p,p,p^3,p^3),\ \mathrm{diag}(p,p^2,p^3,p^2),\ \mathrm{diag}(p^2,p^2,p^2,p^2). \end{aligned}$$

Then we may use the explicit left coset representatives in [51, pp. 189–190].

By [19, (1.15)], we have the following relation:

$$\begin{aligned} &\sum_{n=0}^{\infty}T(p^n)t^n=\frac{P_p(t)}{Q_p(t)}, \quad P_p(t)=1-p^2R_pt^2,\\ &Q_p(t)=1-T(p)t+\{T(p)^2-T(p^2)-p^2R_p\}t^2-p^3R_pT(p)t^3+p^6R_{p^2}t^4. \end{aligned} \tag{2.8}$$

The finite group $Sp_4(\mathbb{Z}/N\mathbb{Z})\simeq Sp_4(\mathbb{Z})/\Gamma(N)$ acts on $M_{\underline{k}}(\Gamma(N))$ by $F\mapsto F|[\tilde{\gamma}]_{\underline{k}}$ if we fix a lift $\tilde{\gamma}$ of $\gamma\in Sp_4(\mathbb{Z}/N\mathbb{Z})$. We denote this action by the same notation $F|[\gamma]_{\underline{k}}$. This action does not depend on the choice of lifts of γ. The diagonal subgroup of $Sp_4(\mathbb{Z}/N\mathbb{Z})$ is isomorphic to $(\mathbb{Z}/N\mathbb{Z})^\times\times(\mathbb{Z}/N\mathbb{Z})^\times$ by sending $S_{a,b}:=\mathrm{diag}(a^{-1},b^{-1},a,b)$ to (a,b) and it also acts on $M_{\underline{k}}(\Gamma(N))$, factoring through the action of $Sp_4(\mathbb{Z}/N\mathbb{Z})$. Then we have the character decomposition

$$M_{\underline{k}}(\Gamma(N))=\bigoplus_{\chi_1,\chi_2:(\mathbb{Z}/N\mathbb{Z})^\times\longrightarrow\mathbb{C}^\times}M_{\underline{k}}(\Gamma(N),\chi_1,\chi_2), \tag{2.9}$$

where $M_{\underline{k}}(\Gamma(N),\chi_1,\chi_2)=\{F\in M_{\underline{k}}(\Gamma(N)) \mid F|[S_{a,1}]_{\underline{k}}=\chi_1(a)F \text{ and } F|[S_{a,a}]_{\underline{k}}=\chi_2(a)F\}$. It is easy to see that the Hecke operators preserve $M_{\underline{k}}(\Gamma(N),\chi_1,\chi_2)$ (cf. [55]). We should remark that in order that $M_{\underline{k}}(\Gamma(N),\chi_1,\chi_2)\neq 0$, the weight (k_1,k_2) has to satisfy the parity condition

$$\chi_2(-1)=(-1)^{k_1+k_2}. \tag{2.10}$$

In particular, if $N=1$ or 2, then χ_2 has to be trivial and therefore

$$k_1+k_2\equiv 0 \bmod 2. \tag{2.11}$$

Put

$$M_{\underline{k}}(\Gamma(N),\chi) := \bigoplus_{\chi_1:(\mathbb{Z}/N\mathbb{Z})^\times \longrightarrow \mathbb{C}^\times} M_{\underline{k}}(\Gamma(N),\chi_1,\chi)$$
$$S_{\underline{k}}(\Gamma(N),\chi) := M_{\underline{k}}(\Gamma(N),\chi) \cap S_{\underline{k}}(\Gamma(N)). \tag{2.12}$$

Throughout this paper, we assume this parity condition (2.10) for F. Let

$$F(Z) = \sum_{T\in \mathrm{Sym}^2(\mathbb{Z})_{\geqslant 0}} A_F(T) e^{(2\pi\sqrt{-1}/N)\mathrm{tr}(TZ)} \in M_{\underline{k}}(\Gamma(N),\chi),$$

be an eigenform for all $T(p^i)$, $p \nmid N$, $i \in \mathbb{N}$ with eigenvalues $\lambda_F(p^i)$, i.e.,

$$T(p^i)F = \lambda_F(p^i)F.$$

By (2.8) we have the following relation:

$$\begin{aligned}
&\sum_{n=0}^{\infty} \lambda_F(p^n)t^n = \frac{P_{F,p}(t)}{Q_{F,p}(t)}, \quad P_{F,p}(t) = 1 - p^{\mu-1}\chi(p)t^2,\\
&Q_{F,p}(t) = 1 - \lambda_F(p)T + \{\lambda_F(p)^2 - \lambda_F(p^2) - p^{\mu-1}\chi(p)\}t^2\\
&\qquad - \chi(p)p^{\mu}\lambda_F(p)t^3 + \chi(p)^2 p^{2\mu} t^4
\end{aligned} \tag{2.13}$$

where $\mu = k_1 + k_2 - 3$. Then we define the partial spinor L-function of F by

$$L^N(s,\mathrm{spin},F) := \prod_{p\nmid N} Q_{F,p}(p^{-s})^{-1}.$$

Definition 2.14. We define the Hecke field $\mathbb{Q}_F$ of F by

$$\mathbb{Q}_F = \mathbb{Q}(\lambda_F(p^i),\ \chi_j(p),\ j = 1,2 \text{ for } p \nmid N \text{ and } i \geqslant 0).$$

It is well known (cf. [68]) that $\mathbb{Q}_F$ has a finite degree over $\mathbb{Q}$ since $M_{\underline{k}}(\Gamma(N))$ has a $\mathbb{Q}$-structure $L_{\mathbb{Q}}$ which is preserved by Hecke actions and the Hecke algebra inside $\mathrm{End}_{\mathbb{Q}}(L_{\mathbb{Q}})$ is finitely generated.

2.3. Adelic forms

For a positive integer N, let $K(N)$ be the group consisting of the elements $g \in GSp_4(\widehat{\mathbb{Z}})$ such that $g \equiv E_4 \bmod N$. Then we see that $\Gamma(N) = Sp_4(\mathbb{Q}) \cap K(N)$ and $\nu(K(N)) = 1 + N\widehat{\mathbb{Z}}$. Then it follows from the strong approximation theorem for Sp_4 that

$$G(\mathbb{A}) = \coprod_{\substack{1\leqslant a<N\\(a,N)=1}} G(\mathbb{Q})G(\mathbb{R})^+ d_a K(N) = \coprod_{\substack{1\leqslant a<N\\(a,N)=1}} G(\mathbb{Q})Z_G(\mathbb{R})^+ Sp_4(\mathbb{R}) d_a K(N) \tag{2.15}$$

where d_a is the diagonal matrix such that $(d_a)_p = \mathrm{diag}(a,a,1,1)$ if $p|N$, $(d_a)_p = E_4$ otherwise.

Similarly, let $K_2(N)$ be the groups consisting of the elements $g \in GSp_4(\widehat{\mathbb{Z}})$ such that $g \equiv \mathrm{diag}(1,1,a,a) \bmod N$ for some $a \in \widehat{\mathbb{Z}}^\times$. Then $K(N) \subset K_2(N)$, $\Gamma(N) = GSp_4(\mathbb{Q}) \cap K_2(N)$, and the similitude map $\mu : K_2(N) \to \widehat{\mathbb{Z}}^\times$ is surjective. Then we have,

$$G(\mathbb{A}) = G(\mathbb{Q})G(\mathbb{R})^+ K_2(N).$$

Let $I := E_2\sqrt{-1} \in \mathcal{H}_2$ and $U(2) = \mathrm{Stab}_{Sp_4(\mathbb{R})}(I)$. For any open compact subgroup U of $GSp_4(\widehat{\mathbb{Z}})$ we let $\mathcal{A}_{\underline{k}}(U)^\circ$ denote the subspace of functions $\phi : GSp_4(\mathbb{Q})\backslash GSp_4(\mathbb{A}) \longrightarrow V_{\underline{k}}(\mathbb{C})$ such that

(1) $\phi(guu_\infty) = \lambda_{\underline{k}}(J(u_\infty, I)^{-1})\phi(g)$ for all $g \in G(\mathbb{A})$, $u \in U$, and $u_\infty \in U(2)A_{G,\infty}$;

(2) for $h \in G(\mathbb{A}_f)$, the function

$$\phi_h : \mathcal{H}_2 \longrightarrow V_{\underline{k}}(\mathbb{C}),\ \ \phi_h(Z) = \phi_h(g_\infty I) := \lambda_{\underline{k}}(J(g_\infty, I))\phi(hg_\infty)$$

is a holomorphic function where $Z = g_\infty I,\ \ g_\infty \in Sp_4(\mathbb{R})$ (note that this definition is independent of the choice of g_∞);

(3) for $g \in G(\mathbb{A})$, $\displaystyle\int_{N_R(\mathbb{Q})\backslash N_R(\mathbb{A})} \phi(ng)\, dn = 0$ for any parabolic $\mathbb{Q}$-subgroup R with the Levi decomposition $R = M_R N_R$ and dn is the Haar measure on $N_R(\mathbb{Q})\backslash N_R(\mathbb{A})$.

We define similarly $\mathcal{A}_{\underline{k}}(U)$ by omitting the last condition (3).

Let $\Gamma(N)_a := Sp_4(\mathbb{Q}) \cap d_a^{-1} K(N) d_a$. Note that $\Gamma(N)_a = \Gamma(N)$ for each a. Then we have the isomorphism

$$\mathcal{A}_{\underline{k}}(K(N)) \stackrel{\sim}{\longrightarrow} \bigoplus_{\substack{1\leqslant a<N \\ (a,N)=1}} M_{\underline{k}}(\Gamma(N)_a), \quad \phi \mapsto (\phi_{d_a})_a. \tag{2.16}$$

We also have the isomorphism

$$\mathcal{A}_{\underline{k}}(K(N))^\circ \simeq \bigoplus_{\substack{1\leqslant a<N \\ (a,N)=1}} S_{\underline{k}}(\Gamma(N)_a) \simeq \bigoplus_{\substack{1\leqslant a<N \\ (a,N)=1}} S_{\underline{k}}(\Gamma(N)), \tag{2.17}$$

as well (cf. [8] for checking the cuspidality). We should note that it follows from the condition (1) that

$$\phi(gz_\infty) = z_\infty^{-(k_1+k_2)}\phi(g), \quad g \in G(\mathbb{A}),\ z_\infty \in A_{G,\infty}. \tag{2.18}$$

Similarly, we have the isomorphism

$$\mathcal{A}_{\underline{k}}(K_2(N)) \stackrel{\sim}{\longrightarrow} M_{\underline{k}}(\Gamma(N)). \tag{2.19}$$

Now given a Siegel modular form $F \in M_{\underline{k}}(\Gamma(N))$, we define the adelization of F (with respect to $K_2(N)$) to be the unique automorphic form $\phi_F \in \mathcal{A}_{\underline{k}}(K_2(N))$[1] such that

$$\phi_F(g_\infty) = \lambda_{\underline{k}}(J(g_\infty, I))^{-1} F(g_\infty I). \tag{2.20}$$

Given a cusp form $F \in M_{\underline{k}}(\Gamma(N))$, we obtain $\phi_F \in \mathcal{A}_{\underline{k}}(K_2(N))$, and ϕ_F gives rise to a cuspidal representation π_F. Conversely, given a cuspidal representation π of $GSp_4(\mathbb{A})$, every admissible representation of $GSp_4(\mathbb{A}_f)$ has a vector fixed by $K_2(N)$ for some N [52, 56]. Hence one can pick a cusp form $F \in M_{\underline{k}}(\Gamma(N))$ using the isomorphism (2.17). Hence the definition of the adelic form (2.20) is an appropriate one.

[1]Thanks are due to the referee who suggested this definition and the commutative diagram.

Then we have the following commutative diagram:

$$\begin{array}{ccc} \mathcal{A}_{\underline{k}}(K(N)) & \xrightarrow{\sim} & \bigoplus_{\substack{1\leqslant a<N\\(a,N)=1}} M_{\underline{k}}(\Gamma(N)_a) \\ \uparrow \iota & & \uparrow d \\ \mathcal{A}_{\underline{k}}(K_2(N)) & \xrightarrow{\sim} & M_{\underline{k}}(\Gamma(N)) \end{array}$$

where ι is the inclusion, and d is the diagonal embedding. We denote the image of ϕ_F under ι by ϕ_F again. Then $F(g_\infty I)=\lambda_{\underline{k}}(J(g_\infty,I))\phi_F(d_a g_\infty)$ for all a, $(a,N)=1$.

Remark 2.1. The principal congruence subgroup $K(N)$ is required to study $S_k(\Gamma(N))$ by the trace formula on $G(\mathbb{A})$. (See the proof of Proposition 5.2; Shin and Templier works with $K(N)$, not $K_2(N)$ in [63, Lemma 8.4]. An analogous result for $K_2(N)$ is not available at this time.) Therefore, we work with $K(N)$ in adelic situations. However, as the referee pointed it out, one needs $K_2(N)$ to associate a Siegel cusp form to an automorphic cuspidal representation. Namely, the results in [52] and [56] ensure that every automorphic cuspidal representation of $G(\mathbb{A})$ includes a vector in the image of the map d in the commutative diagram.

Now we restrict the isomorphism (2.16) to specific subspaces, using the character decomposition (2.9). Given two Dirichlet characters $\chi_i:(\mathbb{Z}/N\mathbb{Z})^\times\longrightarrow\mathbb{C}^\times$, $i=1,2$, associate the characters $\chi_i':\mathbb{A}_f^\times\longrightarrow\mathbb{C}^\times$ via the natural map $\mathbb{A}_f^\times\longrightarrow\widehat{\mathbb{A}}_f^\times/\mathbb{Q}_{>0}=\widehat{\mathbb{Z}}^\times\longrightarrow(\mathbb{Z}/N\mathbb{Z})^\times$. Define $\tilde{\chi}:T(\mathbb{A}_f)\longrightarrow\mathbb{C}^\times$ by

$$\tilde{\chi}(\mathrm{diag}(*,*,c,d))=\chi_1'(d^{-1}c)^{-1}\chi_2'(d)^{-1}.$$

Choose $F=(F_a)$ from the RHS of (2.16) which satisfies $F|[S_{z,z}]_{\underline{k}}=(F_a|[S_{z,z}]_{\underline{k}})=(\chi_2(z)F_a)=\chi_2(z)F$ and $F|[S_{z,1}]_{\underline{k}}=\chi_1(z)F$. If we write $g\in G(\mathbb{A})$ as $g=rz_\infty d_a g_\infty k\in G(\mathbb{A})$ and take $z_f\in T(\mathbb{A}_f)$, then

$$\phi_F(gz_f)=\lambda_{\underline{k}}(J(g,I))^{-1}F_a(g_\infty I)\tilde{\chi}(z_f). \tag{2.21}$$

Then this gives rise to the isomorphism of the subspaces

$$\mathcal{A}_{\underline{k}}(K(N),\tilde{\chi})\xrightarrow{\sim}\bigoplus_{\substack{1\leqslant a<N\\(a,N)=1}} M_{\underline{k}}(\Gamma(N)_a,\chi_1,\chi_2).$$

The central character of any element in $\mathcal{A}_{\underline{k}}(K(N),\tilde{\chi})$ is given by χ_2 and hence $\mathcal{A}_{\underline{k}}(K(N),1)\simeq\bigoplus_{\chi_1\in\widehat{(\mathbb{Z}/N\mathbb{Z})^\times}}\bigoplus_{\substack{1\leqslant a<N\\(a,N)=1}} M_{\underline{k}}(\Gamma(N)_a,\chi_1,1)$.

We now study the Hecke operators on $\mathcal{A}_{\underline{k}}(K(N))$ and its relation to classical Hecke operators. Let ϕ be an element of $\mathcal{A}_{\underline{k}}(K(N))$ and $F=(F_a)$ be the corresponding element of the RHS via the isomorphism (2.16). For any prime $p\nmid N$ and $\alpha\in G(\mathbb{Q})\cap T(\mathbb{Q}_p)$, define the Hecke action with respect to α

$$\widetilde{T}_\alpha\phi(g):=\int_{G(\mathbb{A}_f)}([K(N)_p\alpha K(N)_p]\otimes 1_{K(N)^p})(g_f)\phi(gg_f)\,dg_f, \tag{2.22}$$

where dg_f is the Haar measure on $G(\mathbb{A}_f)$ so that $\mathrm{vol}(K) = 1$. Here $K(N)_p$ is the p-component of $K(N)$ and $K(N)^p$ is the group consisting of all elements of $K(N)$ with trivial p-component. Here $[K(N)_p\alpha K(N)_p]$ stands for the characteristic function of $K(N)_p\alpha K(N)_p$. Then by using (2.15) and (2.22), we can easily see that

$$\nu(\alpha)^{-(k_1+k_2-3)/2} T_\alpha F(Z) = \nu(\alpha)^{-3/2} \widetilde{T}_{\alpha^{-1}} \phi(g), \tag{2.23}$$

where $g = rz_\infty g_a g_\infty k$ as above and $Z = g_\infty I$ (cf. [45, § 8]). The same formula also holds for $K_2(N)$. From this relation, up to the factor of $\nu(\alpha)^{(k_1+k_2)/2-3}$, the isomorphism (2.16) preserves Hecke eigenforms in both sides.

To end this subsection we give an estimation of the dimension of each space which immediately follows from [69, 70]:

Proposition 2.2. *For any $\underline{k} = (k_1, k_2), k_1 \geqslant k_2 \geqslant 3$ and N, as $N + k_1 + k_2 \to \infty$,*

(1) $\dim S_{\underline{k}}(\Gamma(N)) \sim C \cdot N^{10}(k_1 - 1)(k_2 - 2)(k_1 - k_2 + 1)(k_1 + k_2 - 3)$;

(2) $\dim \mathcal{A}_{\underline{k}}(K(N))^\circ \sim C \cdot \varphi(N) N^{10}(k_1 - 1)(k_2 - 2)(k_1 - k_2 + 1)(k_1 + k_2 - 3)$;

(3) $\dim S_{\underline{k}}(\Gamma(N), \chi) \sim C \cdot \frac{N^{10}}{\varphi(N)}(k_1 - 1)(k_2 - 2)(k_1 - k_2 + 1)(k_1 + k_2 - 3)$;

(4) $\dim \mathcal{A}_{\underline{k}}(K(N), \chi)^\circ \sim C \cdot N^{10}(k_1 - 1)(k_2 - 2)(k_1 - k_2 + 1)(k_1 + k_2 - 3)$;

where C is a positive constant which is independent of $\underline{k}$ and N.

Proof. First we assume $k_2 \geqslant 5$ and treat the cases (1) and (2). Then one can apply [69, Theorem 3, 4] in case $N = 1, 2$, and [70, Theorem 7.3] in case $N \geqslant 3$ for $(j, k) = (k_1 - k_2, k_2)$. If $k_2 = 3$ or 4, the argument in [30, § 5] shows the above dimension formulas is still validity for $k_2 \geqslant 3$ up to the difference comes from $\dim M_{k_1-1,1}(\Gamma(N))$ which occurs in case $k_2 = 3$. By Proposition 3.7 and 3.1, we see that $\dim M_{k_1-1,1}(\Gamma(N)) = o(\dim S_{\underline{k}}(\Gamma(N)))$ as $k_1 + k_2 + N \to \infty$. Hence we have the claim. The second claim follows from (2.17).

For (3), (4), consider a normal subgroup of $Sp_4(\mathbb{Z})$ defined by

$$\Gamma'(N) = \{\gamma \in Sp_4(\mathbb{Z}) \mid (\gamma \bmod N) \in (\mathbb{Z}/N\mathbb{Z})^\times E_4\}.$$

By [70, Theorem 3.2] the main term contributing to the dimension is

$$[Sp_4(\mathbb{Z}) : \Gamma'(N)](k_1 - 1)(k_2 - 2)(k_1 - k_2 + 1)(k_1 + k_2 - 3)$$

up to an absolute constant. Hence we have the assertion since

$$[Sp_4(\mathbb{Z}) : \Gamma'(N)] = \varphi(N)^{-1}[Sp_4(\mathbb{Z}) : \Gamma(N)]. \qquad \square$$

2.4. The infinity component of π_F

Let F be a Hecke eigenform in $S_{\underline{k}}(\Gamma(N))$ with the associated cuspidal representation $\pi_F = \pi_{F,\infty} \otimes \otimes'_p \pi_{F,p}$ of $GSp_4(\mathbb{A})$. We assume that $k_1 \geqslant k_2 \geqslant 3$. Then $\pi_{F,\infty}$ is a unitary tempered representation in the discrete spectrum and its minimal K-type is $(k_1, k_2) = (l_1 + 1, l_2 + 2)$, where (l_1, l_2) isthe Harish–Chandra parameter in the holomorphic discrete series. Under the condition $l_1 > l_2 > 0$, the discrete series with the Harish–Chandra parameter (l_1, l_2) is unique up to isomorphism as $(\mathfrak{g}, K)$-cohomology where $\mathfrak{g}$ stands for the complexification of $\mathrm{Lie}(Sp_4(\mathbb{R}))$ and $K = U(2)$. Let $D^{\mathrm{hol}}_{l_1,l_2}$ be the holomorphic discrete series of $GSp_4(\mathbb{R})$ with Harish–Chandra parameter (l_1, l_2) as above and its central character is given by $z \mapsto z^{-k_1-k_2} = z^{-l_1-l_2-3}$ on $A_{G,\infty} \simeq \mathbb{R}_{>0}$. If we want to insist on the minimal K-type (k_1, k_2) instead of the Harish–Chandra parameter (l_1, l_2), the holomorphic discrete series in question would be denoted by $D^{\mathrm{hol}}_{k_1,k_2}$. Hence we have

$$\pi_{F,\infty} \simeq D^{\mathrm{hol}}_{l_1,l_2}.$$

Note that both of central characters coincide on $A_{G,\infty}$ because of (2.18).

3. Spectral decomposition and automorphic counting measures

In this section as Shin did in [62], we introduce measures related to several automorphic forms to look carefully inside the spectral side. This is necessary to extract only holomorphic forms from Arthur's trace formula. It will be clear later on. However, in order to do that, we have to use a single pseudo-coefficient in the trace formula and it causes defects which never appear in the setting of [62]. Namely, non-semisimple orbit contributions arise from the geometric side.

Let us first fix a measure on $G(\mathbb{A})$. For any finite prime p, let μ_p be the Haar measure on $G(\mathbb{Q}_p)$ so that $\mu_p(G(\mathbb{Z}_p)) = 1$. Let μ_∞ be the Euler–Poincaré measure (see [62, § 2]). Then the product measure $\mu = \prod_{p \leqslant \infty} \mu_p$ on $G(\mathbb{A})$ is compatible with the point counting measure on $G(\mathbb{Q})$ and the Lebesgue measure on A_G and therefore it defines the quotient measure $\overline{\mu}$ on $G(\mathbb{Q})A_{G,\infty}\backslash G(\mathbb{A}) = G(\mathbb{Q})\backslash G(\mathbb{A})^1$ where $G(\mathbb{A})^1 = \{g \in G(\mathbb{A}) \mid |\nu(g)|_{\mathbb{A}} = 1\}$. It follows that

$$G(\mathbb{A}) \simeq G(\mathbb{A})^1 \times A_{G,\infty}, \quad g = (g_f, g_\infty) \mapsto ((g_f, g_\infty|\nu(g)|_{\mathbb{A}}^{-1}), |\nu(g_\infty)|) \tag{3.1}$$

and clearly $G(\mathbb{A})^1 \supset G(\mathbb{Q})$.

3.1. Spectral decomposition

For a quasi-character χ (which is not necessarily unitary) on $A_{G,\infty}$ we define $L^2 := L^2(G(\mathbb{Q})\backslash G(\mathbb{A}), \chi)$ as the space of $\mathbb{C}$-valued functions on $G(\mathbb{A})$ which are square integrable modulo $A_{G,\infty}$ with respect to the measure χ, left $G(\mathbb{Q})$-invariant, and transform under $A_{G,\infty}$ by χ. Here 'square integrable modulo $A_{G,\infty}$' means that the integral over $G(\mathbb{Q})\backslash G(\mathbb{A})^1$ is square integrable which makes sense because of (3.1). The regular action of $G(\mathbb{A})$ decomposes the Hilbert space $L^2(G(\mathbb{Q})\backslash G(\mathbb{A}), \chi)$ into the discrete spectrum and the continuous one. Then

$$L^2_{\mathrm{disc}}(G(\mathbb{Q}\backslash G(\mathbb{A}), \chi) = L^2_{\mathrm{cusp}}(G(\mathbb{Q})\backslash G(\mathbb{A}), \chi) \oplus L^2_{\mathrm{res}}(G(\mathbb{Q})\backslash G(\mathbb{A}), \chi).$$

For $* \in \{\mathrm{disc}, \mathrm{cusp}, \mathrm{res}\}$, let

$$L^2_* := L^2_*(G(\mathbb{Q}\backslash G(\mathbb{A}), \chi) = \bigoplus_{\pi} m_*(\pi)\, \pi,$$

where $m_*(\pi)$ is the multiplicity of π in $L^2_*(G(\mathbb{Q}\backslash G(\mathbb{A}), \chi)$.

We review the residual spectrum of GSp_4 in [33]:

$$L^2_{\text{res}}(G(\mathbb{Q}\backslash G(\mathbb{A}), \chi) = L^2(B) \oplus L^2(P_1) \oplus L^2(P_2).$$

Then

$$L^2(P_1) = \bigoplus_{(\pi,\eta)} J\left(\frac{1}{2}, \pi \otimes \eta\right),$$

where π runs over cuspidal representations of $GL_2(\mathbb{A}_{\mathbb{Q}})$ with the trivial central character such that $L(\frac{1}{2}, \pi) \neq 0$, and η runs over grössencharacters of $\mathbb{A}_{\mathbb{Q}}^\times$ such that $\eta_\infty^2 = \chi$, and $J(\frac{1}{2}, \pi \otimes \eta)$ is the unique quotient of $\text{Ind}_{P_1}^G \pi |\det|^{1/2} \otimes \eta$.

Similarly,

$$L^2(P_2) = \bigoplus_{(\eta,\pi)} J(1, \eta \otimes \pi), \tag{3.2}$$

where η runs over non-trivial quadratic characters of $\mathbb{A}_{\mathbb{Q}}^\times$, and π runs over monomial cuspidal representations of $GL_2(\mathbb{A}_{\mathbb{Q}})$ such that $\pi \simeq \pi \otimes \eta$, and $\eta\omega_\pi = \chi$, and $J(1, \eta \otimes \pi)$ is the unique quotient of $\text{Ind}_{P_2}^G \eta|\cdot| \otimes \pi$.

Finally,

$$L^2(B) = J(\rho_B, \chi(1, 1, \mu)) \oplus \bigoplus_{\nu} J(e_1, \chi(\nu, \nu, \mu)),$$

where ν runs over non-trivial quadratic characters and $\mu^2 = \chi$. Here $J(\rho_B, \chi(1, 1, \mu))$ is the unique quotient of $\text{Ind}_B^G \chi(1, 1, \mu) \otimes \exp(\langle \rho_B + \rho_B, H_B(\,)\rangle)$, and $J(e_1, \chi(\nu, \nu, \mu))$ is the unique quotient of $\text{Ind}_B^G \chi(\nu, \nu, \mu) \otimes \exp(\langle e_1 + \rho_B, H_B(\,)\rangle)$, which is the Langlands quotient of $\text{Ind}_{P_2}^G |\cdot| \otimes \sigma$, where $\sigma = \text{Ind}_B^{M_2} \chi(\nu, \nu, \mu)$.

Any local component of a representation in the residue spectrum is known to be non-tempered (cf. [72]). Recall that $D^{\text{hol}}_{k_1,k_2}$ is tempered for $k_1 \geqslant k_2 \geqslant 3$. Therefore,

$$\text{Hom}_{G(\mathbb{R})}(D^{\text{hol}}_{k_1,k_2}, L^2_{\text{res}}) = 0.$$

3.2. Algebraic representations of $GSp_4(\mathbb{R})$

In this section we quickly recall algebraic representations of $GSp_4(\mathbb{R})$. This is necessary to fix the central characters of the representations associated to classical Siegel modular forms.

Recall $G = GSp_4$ and put $G_0 = Sp_4$. Let T_0 be the split maximal diagonal torus of G_0 and K_0 be the maximal compact subgroup of $G_0(\mathbb{R})$. Let $\xi = (\xi, V)$ be an irreducible algebraic representation of $GSp_4(\mathbb{R})$. We have a decomposition $\text{Lie}(G(\mathbb{R})) = \mathfrak{z} \oplus \text{Lie}(G_0(\mathbb{R}))$ where $\mathfrak{z} \simeq \mathbb{R}$ is the center of $\text{Lie}(G(\mathbb{R}))$. The infinitesimal action of $\text{Lie}(G(\mathbb{R}))$ on V uniquely determines ξ since there exists a sufficiently small, open neighborhood of $GSp_4(\mathbb{R})$ at the origin in Euclidean topology whose Zariski closure is $GSp_4(\mathbb{R})$. It follows from this and the classification of all algebraic representations for Sp_4 (cf. [21, § 16.2]) that

$$\xi \simeq \nu^c \otimes \rho_{a,b}, \ \ c \in \mathbb{Z} \tag{3.3}$$

where $\nu : GSp_4 \longrightarrow GL_1$ is the similitude and $\rho_{a,b}$ is a unique irreducible algebraic representation of $G(\mathbb{R})$ with highest weight $(a, b) \in \mathbb{Z}^2, a \geqslant b \geqslant 0$. It is clear that the compact torus $U(1)\times U(1)$ in $U(2)$ is Zariski dense in $T_0(\mathbb{C})$ via $U(2)\simeq K$. Hence the algebraic character of T_0 is completely determined by the action of the diagonal compact torus $U(1)\times U(1)$. It follows from this that $\rho_{a,b}|_{K_0}$ contains $V_{(a,b)}$ (see § 2.1 for the notation). The central character of ξ is given by $\mathbf{diag}(z,z,z,z)\mapsto z^{2c+(a+b)}$ by (3.3). By Weyl's dimension formula (see [21, (24.19) in p. 406]),

$$\dim\xi=\dim\rho_{a,b}=\frac{(a-b+1)(a+b+3)(a+2)(b+1)}{6}. \tag{3.4}$$

Put $\xi_{c,a,b}=\nu^c\otimes\rho_{a,b}$. We usually consider

$$\xi_{\underline{k}}=\xi_{3,k_1-3,k_2-3},\quad \underline{k}=(k_1,k_2),k_1\geqslant k_2\geqslant 3 \tag{3.5}$$

and then the central character of its dual $\xi_{\underline{k}}^\vee$ coincides on $A_{G,\infty}$ with one of $D^{\mathrm{hol}}_{k_1,k_2}$. Note that

$$\dim\xi_{\underline{k}}=\frac{(k_1-k_2+1)(k_1+k_2-3)(k_1-1)(k_2-2)}{6}=\frac{(\ell_1-\ell_2)(\ell_1+\ell_2)\ell_1\ell_2}{6}=\frac{1}{6}d(D^{\mathrm{hol}}_{k_1,k_2}),$$

where $d(D^{\mathrm{hol}}_{k_1,k_2})=(\ell_1-\ell_2)(\ell_1+\ell_2)\ell_1\ell_2$ is the formal degree of $D^{\mathrm{hol}}_{k_1,k_2}$.

3.3. Special cohomological representation of $GSp_4(\mathbb{R})$

As we will see in § 3.6, the trace of a pseudo-coefficient of a holomorphic discrete series $D^{\mathrm{hol}}_{l_1,l_2}$ may be non-vanishing for $D^{\mathrm{hol}}_{l_1,l_2}$, ω_{l_1}, and $\mathbf{1}$, where $\mathbf{1}$ is the trivial representation of $G(\mathbb{R})$, and ω_l is a certain unitary representation defined as follows: For $l\geqslant 2$, it is the induced representation $\mathrm{Ind}_{Sp_4(\mathbb{R})}^{G(\mathbb{R})}\omega_l'$, where ω_l' is the Langlands quotient of $\mathrm{Ind}_{P_2(\mathbb{R})\cap Sp_4(\mathbb{R})}^{Sp_4(\mathbb{R})}|\cdot|\mathrm{sgn}\rtimes D_l^+$, where D_l^+ is the holomorphic discrete series of $SL_2(\mathbb{R})$ with minimal K-type l.

By the classification of the residual spectrum, $\mathbf{1}$ occurs as an infinity component of residual spectrum from the Borel subgroup:

$$\mathrm{Hom}_{G(\mathbb{R})}(\mathbf{1},L^2)^{K(N)}=\mathrm{Hom}_{G(\mathbb{R})}(\mathbf{1},L^2(B))^{K(N)}=\bigoplus_{\substack{\chi\in\widehat{(\mathbb{Z}/N\mathbb{Z})^\times}\\ \chi^2=1}}\mathbb{C}. \tag{3.6}$$

One can see easily that there are no cuspidal representations with the infinity component $\mathbf{1}$: If F is such a form, then clearly, $\Omega F=\Delta F=0$. This is clearly not possible. (Or one can see from strong approximation theorem that $G(\mathbb{Q})G(\mathbb{R})$ is dense in $G(\mathbb{A})$. Hence if $\mathbf{1}$ is the infinity component of an automorphic representation, the automorphic representation has to be $\mathbf{1}$.)

The special unitary representation ω_l occurs as the infinity component of residual spectrum of Klingen parabolic subgroup, and CAP forms of weight $(l+1,1)$ from the Klingen parabolic subgroup (Proposition 3.4). It is non-tempered. Hence by (3.2)

$$\mathrm{Hom}_{G(\mathbb{R})}(\omega_{l_1},L^2_{\mathrm{res}})^{K(N)}=\mathrm{Hom}_{G(\mathbb{R})}(\omega_{l_1},L^2(P_2))^{K(N)}=\bigoplus_{(\eta,\pi)}J(1,\eta\otimes\pi)^{K(N)}$$

which is related to the space of Klingen Eisenstein series of weight $(k_1,1)$ with $l_1+1=k_1$. Here η runs over the quadratic characters of imaginary quadratic fields, $\pi_\infty|_{SL_2(\mathbb{R})}$ is isomorphic to $D^+_{l_1}\oplus\overline{D^+_{l_1}}$, and π corresponds to a CM modular form in $S_{k_1}(\Gamma^1(N))$. Since $\dim S_{k_1}(\Gamma^1(N))=O(k_1N^3)$ as $k_1+N\to\infty$, we have a rough estimation:

Proposition 3.1. *For any $k_1 \geqslant k_2 \geqslant 1$,*

$$\dim \bigoplus_{(\eta,\pi)} J(1, \eta \otimes \pi)^{K(N)} = O(k_1 \varphi(N) N^3) = O(k_1 N^4) \text{ as } k_1 + N \to \infty.$$

3.4. Classification of endoscopic forms

In this section we study endoscopic forms in $S_{\underline{k}}(\Gamma(N))$ and give an estimation of the dimension of the space $S_{\underline{k}}^{\text{en}}(\Gamma(N))$ generated by such forms. For simplicity we work on adelic forms instead of classical forms. We also write $\mathcal{A}_{\underline{k}}^{\text{en}}(K(N))^{\circ}$ for the space generated by Hecke eigen automorphic forms in $\mathcal{A}_{\underline{k}}(K(N))^{\circ}$ whose automorphic representation is endoscopic via the isomorphism (2.17). We freely use the results in [50]. The endoscopic lift in our situation is a functorial lift from the endoscopic group $H := GSO(2,2)$ to GSp_4. Note that $GSO(2,2) \simeq (GL_2 \times GL_2)/\{(z, z^{-1}) : z \in GL_1\}$, and $GSO(4) \simeq (D^{\times} \times D^{\times})/\{(z, z^{-1}) : z \in GL_1\}$, where D is a quaternion division algebra. Hence a cuspidal representation of $GSO(2,2)$ (respectively $GSO(4)$) can be written as (π_1, π_2), where π_1, π_2 are cuspidal representations of GL_2 (respectively $D^{\times}$) with the same central characters. When we lift a pair of elliptic new forms of weights $\geqslant 2$, it is known that

(1) if the lift factors through $GSO(4)$ via Jacquet–Langlands correspondence, then it has to be a holomorphic discrete series at infinity; and
(2) if the lift does not factor through $GSO(4)$, then the lift cannot be a holomorphic discrete series at infinity (it is a large discrete series).

Since we are interested in holomorphic Siegel forms, only the first case can happen.

For any reductive group G over $\mathbb{Q}$, we denote by $\Pi(G(\mathbb{A}))$ the set consisting of the isomorphism classes of cuspidal automorphic representations of $G(\mathbb{A})$.

Let $\Pi(\tau)$ be the global packet for GSp_4 constructed from a cuspidal representation τ of $H(\mathbb{A})$ via the theta lift due to Roberts [50]. Put $r_1 = k_1 + k_2 - 2$ and $r_2 = k_1 - k_2 + 2$. By adjusting the central character, we may assume that any element of $\Pi(\tau)$ is realized in the space $L^2_{\text{cusp}}(G(\mathbb{Q})\backslash G(\mathbb{A}), \chi_{\xi^\vee})$ for $\xi = \xi_{\underline{k}}$. Then we see that

$$\dim \mathcal{A}_{\underline{k}}^{\text{en}}(K(N))^{\circ} = \text{vol}(K(N))^{-1} \sum_{\substack{\tau \in \Pi(H(\mathbb{A})) \\ \tau_\infty \simeq D_{r_1} \otimes D_{r_2}}} \sum_{\substack{\Pi \in \Pi(\tau) \\ \Pi_\infty \simeq D^{\text{hol}}_{k_1,k_2}}} m(\Pi) \text{tr}(\Pi_f(1_{K(N)})), \ \ m(\Pi) \in \{0, 1\} \tag{3.7}$$

where $1_{K(N)} = \text{char}_{K(N)}$. Note that $s \cdot \tau \not\simeq \tau$ (s interchanges the two components) because of the infinite type and $r_1 \neq r_2$. We shall describe the RHS more precisely. For $\Pi \in \Pi(\tau)$, let T_Π be the set of every places v of $\mathbb{Q}$ so that $\Pi_v \simeq \theta_{GSO(4)}(\tau_v^{\text{JL}})$ where $GSO(4) \simeq (D_v^{\times} \times D_v^{\times})/\mathbb{Q}_v^{\times}$ for a unique quaternion division field over $\mathbb{Q}_v$ and θ is the local theta lift from $GSO(4)$ to GSp_4. Note that τ_v is necessarily square integrable. We write $\tau = (\pi_1, \pi_2)$ for $\tau \in \Pi(H)$ where each π_i is a cuspidal automorphic representation of $GL_2(\mathbb{A})$ such that $\omega_{\pi_1} = \omega_{\pi_2}$. For such a τ, let T_τ be the set of all places of $\mathbb{Q}$ so that both of $\pi_{1,v}$ and $\pi_{2,v}$ are square integrable.

By [50, Theorem 8.5] and [24, Theorems 8.1 and 8.2], if $m(\Pi)=1$, then for each $\Pi\in\Pi(\tau)$, $\tau=(\pi_1,\pi_2)$, there exists a definite quaternion algebra D over $\mathbb{Q}$ such that τ is in the image of Jacquet–Langlands correspondence from $GSO(D)(\mathbb{A})$ and the set of ramified places of D is given exactly by $T_\Pi=T_\tau$ and

(1) for each $v\in T_\tau$,
   (a) if $\sigma^{JL}=\pi_{1,v}=\pi_{2,v}$, then $\Theta(\sigma\boxtimes\sigma)=\theta(\sigma\boxtimes\sigma)$ is the unique non-generic direct factor of $I_{Q(Z)}(1,\sigma^{JL})$ where $Q(Z)$ stands for the Klingen parabolic subgroup;
   (b) if $\pi_{1,v}\neq\pi_{2,v}$, then $\Theta(\pi_{1,v}^{\rm JL}\boxtimes\pi_{2,v}^{\rm JL})=\theta(\pi_{1,v}^{\rm JL}\boxtimes\pi_{2,v}^{\rm JL})$ is a non-generic supercuspidal representation.

(2) for each $v\notin T_\tau$, one of $\pi_{1,v}$ and $\pi_{2,v}$ has to be a principal series representation and
   (a) if $\pi_{1,v}$ is square integrable and $\pi_{2,v}=\pi(\chi,\chi')$ for unitary characters χ,χ' so that $\omega_{\pi_{1,v}}=\chi\chi'$, then $\theta(\pi_{1,v}\boxtimes\pi_{2,v})=J_{P(Y)}(\pi_{1,v}\otimes\chi^{-1},\chi)$ is a non-generic Langlands quotient of $I_{P(Y)}(\pi_{1,v}\otimes\chi^{-1},\chi)$ where $P(Y)$ stands for the Siegel parabolic subgroup;
   (b) if $\pi_{1,v}=\pi(\chi_1,\chi_1')$ and $\pi_{2,v}=\pi(\chi_2,\chi_2')$ for unitary characters χ_i,χ_i' so that $\chi_1\chi_1'=\chi_2\chi_2'$, then $\theta(\pi_{1,v}\boxtimes\pi_{2,v})={\rm Ind}_B^G(\chi_2'/\chi_1,\chi_2/\chi_1;\chi_1)$.

In each case the local L-packets are given as follows (see [13, § 8] for (1) and [13, § 6.6] for (2)):

(1) (a) $\{\pi^+,\pi^-\}$, where π^+ (respectively π^-) is the unique generic (respectively non-generic) direct summand of $I_{Q(Z)}(1,\sigma^{JL})$;
   (b) $\{\pi^+:=\theta(\pi_{1,v}\boxtimes\pi_{2,v}),\pi^-:=\theta(\pi_{1,v}^{\rm JL}\boxtimes\pi_{2,v}^{\rm JL})\}$.

(2) (a) $\{\pi^+,\pi^-:=J_{P(Y)}(\pi_{1,v}\otimes\chi^{-1},\chi)\}$, where π^+ is the unique generic subrepresentation of $I_{P(Y)}(\pi_{1,v}\otimes\chi^{-1},\chi)$;
   (b) $\{{\rm Ind}_B^G(\chi_2'/\chi_1,\chi_2/\chi_1;\chi_1)\}$ (the L-packet is a singleton).

For a principal congruence subgroup $K(N)$ we put $K^H(N)=H(\mathbb{A}_f)\cap K(N)$. For a finite place $v=p_v$ of $\mathbb{Q}$, put $N_v=p_v^{{\rm ord}_v(N)}$. From now suppose that $(N,11!)=1$. Then by [20, Théorèm 3.2.3], $(1_{K(N)_v},N_v^{-2}1_{K^H(N)_v})$ is a transfer pair for each finite place v. We now apply the (endoscopic) local character identities ([13, Proposition 6.9] for the L-packets in (2) and Proposition 8.2 for the L-packets in (1)) with the transfer pair $(1_{K(N)_v},N_v^{-2}1_{K^H(N)_v})$ for each finite place$v$, the RHS of (3.7) is bounded by

$$\begin{aligned}
{\rm RHS} &\leqslant {\rm vol}(K(N))^{-1}\sum_{\substack{\tau\in\Pi(H(\mathbb{A}))\\ \tau_\infty\simeq D_{r_1}\otimes D_{r_2}}}\ \sum_{\substack{\Pi\in\Pi(\tau)\\ \Pi_\infty\in\{D^{\rm hol}_{k_1,k_2},D^{\rm large}_{k_1,k_2}\}}} m(\Pi){\rm tr}(\Pi_f(1_{K(N)}))\\
&\leqslant {\rm vol}(K(N))^{-1}\sum_{\substack{\tau\in\Pi(H(\mathbb{A}))\\ \tau_\infty\simeq D_{r_1}\otimes D_{r_2}}} 2^{|T_\tau|}N^{-2}{\rm tr}(\tau_f(1_{K^H(N)}))\\
&\ll {\rm vol}(K(N))^{-1}N^{-2+\varepsilon}{\rm vol}(K^H(N))\dim S_{r_1}(\Gamma^1(N))\times\dim S_{r_2}(\Gamma^1(N))\\
&\ll (r_1-1)(r_2-1)N^{9+\varepsilon}.
\end{aligned}$$

Since $\dim S_{\underline{k}}^{\mathrm{en}}(\Gamma(N)) = \varphi(N)^{-1}\dim \mathcal{A}_{\underline{k}}^{\mathrm{en}}(K(N))^\circ$, we have

Theorem 3.2. *Let* $(N, 11!) = 1$*. Then* $\dim S_{\underline{k}}^{\mathrm{en}}(\Gamma(N)) = O((k_1 - k_2 + 1)(k_1 + k_2 - 3)N^{8+\varepsilon})$ *as* $N + k_1 + k_2 \to \infty$*.*

3.5. Classification of CAP forms

In this section, we classify CAP forms. We say F is a CAP form associated to a parabolic subgroup P if π_F is a CAP representation associated to P in the sense of [22]. Such a representation is completely classified by [48, 60, 66].

For holomorphic Siegel modular CAP forms of weight $\underline{k} = (k_1, k_2)$, $k_1 \geqslant k_2 \geqslant 1$, it turns out that

(1) in the case when (k_1, k_2), $k_2 > 2$, it has to be a CAP form associated to Siegel parabolic subgroup and $k_1 = k_2$;

(2) in the case when $k_2 = 2$, we must have $k_1 = k_2 = 2$ and it can be a CAP form associated to any parabolic subgroup;

(3) in the case when $(k_1, k_2) = (1, 1)$, it can be a CAP form associated to Borel or Klingen parabolic subgroup, but not Siegel parabolic subgroup;

(4) in the case when $k_1 > k_2 = 1$, it has to be a CAP form associated to Klingen parabolic subgroup.

In [55], one can see several examples regarding to the second case. In the third case Weissauer proved that any (Hecke eigen) Siegel modular forms of weight one with respect to $\Gamma_0(N)$ is CAP (see [75]), but it is still open whether it is also the case for $\Gamma(N)$. In the case (4), it will be proved in Proposition 3.4 that any (Hecke eigen) Siegel modular forms of weight $(k_1, 1)$ $k_1 \geqslant 3$ is always a CAP form associated to Klingen. We expect that it holds even if $k_1 = 2$ but in this case we cannot use the geometric argument and thus we might have to rely on the classification of representations for GSp_4. Once Arthur's conjectural classification in [6] is completed, then our expectation would be true since it is known to be non-tempered at infinity by [49].

Consider the first case $k_1 \geqslant k_2 \geqslant 2$. We first observe that $k_1 = k_2$ by the argument in [60, p. 225] and hence put $k := k_1 = k_2$. In this case the central character of π_F should be a square of a character and by twisting we may assume that it is trivial. Schmidt completely characterized any holomorphic CAP form associated to P_1 (see [60, Theorem 3.1]) by constructing a lift from a cuspidal automorphic representation π of $\mathrm{PGL}_2(\mathbb{A})$. Schmidt's construction can be a functorial lift by the local Langlands conjecture established by [23] (see [60, Remark 3.2-(a)]). To characterize π_F by using the completed L-function (product over all places), we have to carefully look at the behavior at bad places since weak equivalence does not characterize F except for the case of level one.

Let $S_k^{P_1}(\Gamma(N))$ be the space generated by Hecke eigen cusp forms of parallel weight k which are CAP associated to P_1. We now try to estimate the dimension of this space. Let $\mathcal{A}_k^{P_1}(K(N))^\circ$ be the space of all Hecke eigen automorphic forms in $\mathcal{A}_k(K(N))^\circ$ whose automorphic representations are CAP associated to P_1. To do this we carefully check the behavior of the levels under the functorial lift constructed by Schmidt. As it is done for endoscopic lifts we work with adelic forms.

Since any CAP representation associated to the Siegel parabolic subgroup can be regarded as a cuspidal representation of $PGSp_4(\mathbb{A})$ by [48, Theorem 2.1], we first assume that any representation in question has the trivial central character.

Let $\mathcal{A}(\pi) := \mathcal{A}(\pi, 1)$ be the global A-packet for $PGSp_4$ constructed from a cuspidal representation π of $PGL_2(\mathbb{A})$ by Schmidt [60] and Gan [22, § 4.3]; they studied the same A-packets in conjunction with Waldspurger's local packers (cf. [60, § 3]). We use $\mathcal{A}(\pi)$ as a tool to describe the dimension of $\mathcal{A}_k^{P_1}(K(N), 1)^\circ := \mathcal{A}_k^{P_1}(K(N))^\circ \cap \mathcal{A}_k(K(N), 1)^\circ$. We do not care whether this packet satisfies desirable properties. Then we see that

$$\dim \mathcal{A}_k^{P_1}(K(N), 1)^\circ = \mathrm{vol}(K(N))^{-1} \sum_{\substack{\pi \in \Pi(PGL_2(\mathbb{A})) \\ \tau_\infty \simeq D_{2k-2}}} \sum_{\substack{\Pi \in \mathcal{A}(\pi) \\ \Pi_\infty \simeq D^{\mathrm{hol}}_{k_1,k_2}}} m(\Pi) \mathrm{tr}(\Pi_f(1_{K(N)})), \tag{3.8}$$

$m(\Pi) \in \{0, 1\}$. We shall study the RHS more precisely. For $\pi \in \Pi(PGL_2(\mathbb{A}))$ let S_π be the set of all places v of $\mathbb{Q}$ so that π_v is square integrable. By [60, Theorem 3.1], for each $\Pi \in \Pi(\tau)$ with $m(\Pi) = 1$, there exists a subset $S \subset S_\pi$ such that $\Pi = \Pi(\pi \otimes \pi_S)$ and $(-1)^{|S|} = \varepsilon(1/2, \pi)$ and

(1) if $v \notin S$, then Π_v is the unique, non-tempered irreducible quotient $Q(|\cdot|^{1/2}\pi_v, |\cdot|^{-1/2})$ (see [51, Proposition 5.5.1]) of $\mathrm{Ind}_{P_1}^G(|\cdot|^{1/2}\pi_v \rtimes |\cdot|^{-1/2})$ where P_1 is the Siegel parabolic subgroup;

(2) if $v \in S \setminus \{\infty\}$, then $\Pi_v = SK(\pi_v^{\mathrm{JL}}) = \theta((\pi_v^{\mathrm{JL}} \boxtimes \mathrm{St}^{\mathrm{JL}})^+)$ is a non-generic (tempered) supercuspidal representation; and

(3) if $v = \infty$, then $\Pi_\infty = D^{\mathrm{hol}}_{k,k}$.

Put $U = K(N)$ (respectively $U^1 = K^1(N)$, the group consisting of elements $g \in GL_2(\widehat{\mathbb{Z}})$ such that $g \equiv E_2 \mod N$) and let U_v (respectively U_v^1) be the v-component of U (respectively U^1). Let $N_v = p_v^{\mathrm{ord}_v(N)}$ for the prime p_v corresponding to v. In the first case since it is a quotient of $\mathrm{Ind}_{P_1}^G(|\cdot|^{1/2}\pi \rtimes |\cdot|^{-1/2})$ by [11, Corollary 1, p. 16], we have a rough estimation

$$\dim(\Pi_v)^{U_v} \leqslant \dim \left(\mathrm{Ind}_{P_1}^G(|\cdot|^{1/2}\pi_v \rtimes |\cdot|^{-1/2})\right)^{U_v} \leqslant t_v \cdot \mathrm{tr}(\pi_v(1_{U_v^1}))$$

where $t_v := \#(P_1(\mathbb{Q}_v)\backslash G(\mathbb{Q}_v)/U_v) = O(N_v^3)$.

In the second case Π_v forms a local L-packet $\{\pi^- := \Pi_v, \pi^+ := \theta((\pi_v \boxtimes \mathrm{St})^+)\}$. Therefore, one can apply the previous argument in the case of endoscopic lifts. To bound (3.8) we additionally count admissible representations Π' obtained from Π in $\Pi(\tau)$ by switching the non-generic representation π^- with the generic representation π^+ at the finite places v in the case (2). Therefore, we have

$$\begin{aligned} \mathrm{RHS} &= \mathrm{vol}(K(N))^{-1} \sum_{\substack{\pi \in \Pi(PGL_2(\mathbb{A})) \\ \tau_\infty \simeq D_{2k-2}}} \sum_{\substack{\Pi \in \mathcal{A}(\pi) \\ \Pi_\infty \simeq D^{\mathrm{hol}}_{k_1,k_2}}} \sum_{\substack{S \subset S_\pi \\ \varepsilon(1/2,\pi) = (-1)^{|S|}}} \mathrm{tr}(\Pi_f(1_{K(N)})) \\ &\leqslant \mathrm{vol}(K(N))^{-1} \sum_{\substack{\pi \in \Pi(PGL_2(\mathbb{A})) \\ \tau_\infty \simeq D_{2k-2}}} \sum_{\substack{\Pi \in \mathcal{A}(\pi) \\ \Pi_\infty \simeq D^{\mathrm{hol}}_{k_1,k_2}}} \sum_{\substack{S \subset S_\pi \\ \varepsilon(1/2,\pi) = (-1)^{|S|}}} \{\mathrm{tr}(\Pi_f(1_{K(N)})) + \mathrm{tr}(\Pi'_f(1_{K(N)}))\} \end{aligned}$$

$$\leqslant \mathrm{vol}(K(N))^{-1} \sum_{\substack{\pi\in\Pi(PGL_2(\mathbb{A}))\\ \tau_\infty\simeq D_{2k-2}}} \sum_{\substack{S\subset S_\pi\\ \varepsilon(1/2,\pi)=(-1)^{|S|}}} \prod_{v\notin S} t_v \cdot \mathrm{tr}(\pi_v(1_{U_v^1}))$$
$$\times \prod_{v\in S\setminus\{\infty\}} 2N_v^{-2}\mathrm{tr}((\pi_v\boxtimes \mathrm{St})(1_{K^H(N)_v}))$$
$$\ll \mathrm{vol}(K(N))^{-1}2^{|\{p|N\}|}N^{-2}\mathrm{vol}(K^H(N))\varphi(N)^{-1}N^3\dim S_{2k-2}(\Gamma^1(N))$$
$$\ll kN^{8+\varepsilon}. \tag{3.9}$$

Summing up we have

Theorem 3.3. *Let $(N, 11!) = 1$. Then as $N + k \to \infty$,*

$$\dim S_{\underline{k}}^{P_1}(\Gamma(N)) = O(kN^{8+\varepsilon}).$$

Proof. It follows from

$$\dim S_{\underline{k}}^{P_1}(\Gamma(N)) = O(\dim S_{\underline{k}}^{P_1}(\Gamma(N),1)\varphi(N)) = O(\dim \mathcal{A}_k^{P_1}(K(N),1)^\circ) = O(kN^{8+\varepsilon}). \ \square$$

Next we consider the case $k_1 \geqslant 3$ and $k_2 = 1$. First of all we prove the following:

Proposition 3.4. *Let $k \geqslant 3$ and $N \geqslant 1$ be integers. Then any Siegel cusp form F of weight $(k, 1)$ with respect to $\Gamma(N)$ which is a Hecke eigenform is a CAP form associated to the Klingen parabolic subgroup, and its associated cuspidal representation has the infinity type ω_{k-1}.*

Proof. By [49], $\pi_{F,\infty}|_{Sp_4(\mathbb{R})}$ has the component which is isomorphic to ω_{k-1}. (In [49], ω_{k-1} is denoted as $L(k-1, 1)$.) Since it is the Langlands quotient of $\mathrm{Ind}_{P_2(\mathbb{R})\cap Sp_4(\mathbb{R})}^{Sp_4(\mathbb{R})} |\cdot|\mathrm{sgn} \rtimes D_k^+$, $H^2((\mathrm{Lie}Sp_4(\mathbb{R}), K), \omega_{k-1}\otimes \xi_{(k,3)}^\vee) \neq 0$ by a direct calculation with [9, Proposition 3.1 of p. 36].

For $p \nmid N$, let $\mathrm{Sat}_p(\pi_F)$ be the set of all Satake parameters of $\pi_{F,p}$ which take the values in ${}^L GSp_4 = GSp_4(\mathbb{C})$. By [39, Theorem 7.5-(1)] (actually this case is for $k_1 = 3$ but it also holds from [40, Theorem 24.1] for any $k_1 \geqslant 3$), π_F is weakly equivalent to an induced representation for the Klingen parabolic P_2 from a cuspidal representation σ of $GL_2(\mathbb{A}_{\mathbb{Q}})$ with $\sigma_\infty|_{SL_2(\mathbb{R})} \simeq D_{l_1}^+ \oplus \overline{D_{l_1}^+}$. Hence σ is corresponding to an elliptic new form f of weight k with some level N' such that

$$\mathrm{Sat}_p(\pi_F) = \{p^{1/2}\alpha_p, p^{1/2}\beta_p, p^{-1/2}\alpha_p^{-1}, p^{-1/2}\beta_p^{-1}\}$$

for all $p \nmid N'$, where $\{\alpha_p, \beta_p\}$ is the Satake parameters of $\pi_{f,p}$. This means that π_F is weakly equivalent to $\mathrm{Ind}_{P_2(\mathbb{A})\cap Sp_4(\mathbb{A})}^{Sp_4(\mathbb{A})} |\cdot|\mathrm{sgn} \rtimes \pi_f$ and hence F is a CAP form associated to the Klingen parabolic subgroup. $\square$

By Proposition 3.4, if F is any Siegel cusp form of weight $(k, 1)$, $k \geqslant 3$ with respect to $\Gamma(N)$ which is a Hecke eigenform, then π_F is a CAP representation associated to P_2 and then it follows from [66] that π_F can be obtained by a theta lifting from a Hecke character for an imaginary quadratic field. In [74], Weissauer computed the dimension of $S_{(3,1)}(\Gamma(N))$ but his calculation works for any weight $(k, 1)$ with $k \geqslant 3$. (Just replace

$\delta(K,2,\mathfrak{a}N_1)$ with $\delta(K,k-1,\mathfrak{a}N_1)$ in his notation.) Let us use the notations in [74]. The proof is almost identical with that of [74]. Therefore, we focus only on key points.

For an imaginary quadratic field K, let $\mathcal{O}_K$ be the ring of integers, $w(K)$ the cardinality of the units in $\mathcal{O}_K$, D_K the fundamental discriminant of K, and θ_K the different of $K/\mathbb{Q}$. Let $T=T_K$ be a positive definite binary quadratic form over $\mathbb{Q}$ given by $(K,N_{K/\mathbb{Q}})$. We denote by $O(T)$ (respectively $SO(T)$) the (respectively special) orthogonal group associated to T. Fix an additive character $\psi:\mathbb{Q}\backslash\mathbb{A}_{\mathbb{Q}}\longrightarrow\mathbb{C}^\times$. For K and $t\in\mathbb{Q}_{>0}$ as in [74, § 4] we define the Weil representation $\omega_{T,t}=\omega^{\psi}_{T,t}$ of $\mathrm{Sp}_4(\mathbb{A}_{\mathbb{Q}})$ acting on the Schwartz space $\mathcal{S}(K^2_{\mathbb{A}})$ where $K_{\mathbb{A}}=K\otimes_{\mathbb{Q}}\mathbb{A}_{\mathbb{Q}}$. This action commutes with the action of $O(T)(\mathbb{A}_{\mathbb{Q}})$ on $\mathcal{S}(K^2_{\mathbb{A}})$ given by

$$h\cdot\phi(X):=\phi(h^{-1}X),\ \ h\in O(T)(\mathbb{A}_{\mathbb{Q}}),\ \ \phi\in\mathcal{S}(K^2_{\mathbb{A}}),\ \ X\in K^2_{\mathbb{A}}.$$

For $\phi\in\mathcal{S}(K^2_{\mathbb{A}})$ we define the theta kernel by

$$\theta^{\phi}_{T,t}(g,h):=\sum_{X\in K^2}\omega_{T,t}(g)\phi(h^{-1}X).$$

Let $\widehat{\sigma}$ be a non-trivial automorphic representation σ of $O(T)(\mathbb{A})$ which can be realized as a subrepresentation of the 2-dimensional induced representation $\mathrm{Ind}^{O(T)(\mathbb{A})}_{SO(T)(\mathbb{A})}\sigma$ from a character σ on $SO(T)(\mathbb{A}_{\mathbb{Q}})$. Therefore, we view an element of $\widehat{\sigma}$ a function on $O(T)(\mathbb{A}_{\mathbb{Q}})$. Then for an automorphic form $f\in\widehat{\sigma}$, a left $Sp_4(\mathbb{Q})$-invariant function

$$\theta^{\phi}_{T,t}(g;f):=\int_{O(T)(\mathbb{Q})\backslash O(T)(\mathbb{A}_{\mathbb{Q}})}\theta^{\phi}_{T,t}(g,h)\overline{f(h)}\,dh,\quad g\in Sp_4(\mathbb{A}_{\mathbb{Q}})$$

is called the theta lift of f. When $f\in\widehat{\sigma}$ and $\phi\in\mathcal{S}(K^2_{\mathbb{A}})$ vary, the space of all $\theta^{\phi}_{T,t}(*;f)$ makes up a $Sp_4(\mathbb{A}_{\mathbb{Q}})$ invariant subspace in the space of automorphic forms on $Sp_4(\mathbb{A}_{\mathbb{Q}})$. We denote it by $\theta(T,t,\sigma)$.

Lemma 3.5. *Assume that $\sigma_\infty(z)=\left(\frac{z}{|z|}\right)^{k-1}$, $k>1$. It holds that*

1. *$\theta(T,t,\widehat{\sigma})$ is irreducible.*
2. *$\theta(T,t,\widehat{\sigma})\simeq\theta(T',t',\widehat{\sigma}')$ if and only if tT and $t'T'$ are isomorphic as a quadratic form over $\mathbb{Q}$ and $\widehat{\sigma}\simeq\widehat{\sigma}'$.*
3. *When p splits in K, then*

$$\theta(T,t,\widehat{\sigma})_p\simeq\mathrm{Ind}^{Sp_4(\mathbb{Q}_p)}_{P_2(\mathbb{Q}_p)\cap Sp_4(\mathbb{Q}_p)}\sigma_p\rtimes 1_{SL_2(\mathbb{Q}_p)}.$$

4. *When p is inert in K, then*

$$\theta(T,t,\widehat{\sigma})_p\simeq\left(\mathcal{S}(K^2_p)\otimes\widehat{\sigma}^*\right)^{O(T)(\mathbb{Q}_p)}$$

where $K_p=K\otimes_{\mathbb{Q}}\mathbb{Q}_p$ and $\widehat{\sigma}^*$ is the contragredient representation of $\widehat{\sigma}$. Here the left hand side is a representation of $Sp_4(\mathbb{Q}_p)$ defined through $\omega_{T,t}$ commuting with the action of $O(T)(\mathbb{Q}_p)$.

(5) *if* ϕ_∞ *is given in Definition 4.12.-i) for* $\sigma_\infty(z) = \left(\frac{z}{|z|}\right)^{k-1}$, $k > 1$ *in [47], then*

$$\theta(T, t, \widehat{\sigma})_\infty \simeq \mathrm{Ind}_{P_2(\mathbb{R})\cap Sp_4(\mathbb{R})}^{Sp_4(\mathbb{R})} |\cdot|\mathrm{sgn} \rtimes D_k^+.$$

Proof. The claim follows from [74, Lemma 4.1]. For the last claim see p. 623 of [47] and it claims that $\theta(T, t, \widehat{\sigma})_\infty$ corresponds to a holomorphic Siegel modular form of weight $(k, 1)$. The representation of $Sp_4(\mathbb{R})$ generated by such a form is isomorphic to the RHS of (5) as clarified before. □

For an ideal $\mathfrak{f} \subset \mathcal{O}_K$, let $E_{K,\mathfrak{f}}$ be the units in $\mathcal{O}_K$ congruent to 1 modulo $\mathfrak{f}$. Clearly $|E_{K,f}|$ divides 6. For an integer $k \geqslant 2$ and an ideal $\mathfrak{f} \subset \mathcal{O}_K$, define $\delta(K, k, \mathfrak{f})$ to be 1 if $(E_{K,\mathfrak{f}})^k = \{1\}$ and 0 otherwise. For a fixed K and a positive integer N, put

$$N_1(K) = \prod_{\substack{p|N \\ \text{split in } K}} p^{\mathrm{ord}_p(N)}.$$

For an ideal $\mathfrak{a} \subset \mathcal{O}_K$ put $N(\mathfrak{a}) = |\mathcal{O}_K/\mathfrak{a}|$. For K, N, and $\mathfrak{a}$, put

$$p(\mathfrak{a}, N, K) := N(\mathfrak{a})^2 \prod_{\mathfrak{p}|\mathfrak{a}} (1 - N(\mathfrak{p})^{-2}) N_1(K)^4 \prod_{p|N_1(K)} (1 - p^{-4}).$$

Then we have the following

Proposition 3.6. *For* $k \geqslant 3$*, the following equality hold:*

$$\dim M_{(k,1)}(\Gamma(N)) = \sum_K \frac{h(K)}{w(K)} \sum_M \sum_{\mathfrak{a}} \delta(K, k-1, \mathfrak{a}N_1(K)) |E_{K,\mathfrak{a}N_1(K)}| p(\mathfrak{a}, N, K) \quad (3.10)$$

where K *runs over all imaginary quadratic field so that* $\theta_K^2 | N$ *and* M *all divisor of* N *whose all prime divisors are inert. The last summation runs over all ideal* $\mathfrak{a} \subset \mathcal{O}$ *dividing* $\theta(K, M, N)$ *where*

$$\theta(K, M, N) := \prod_{\substack{\mathfrak{p}\subset K \\ \text{non-split}}} \mathfrak{p}^{2[\mathrm{ord}_{\mathfrak{p}}(NM)/2] - \mathrm{ord}_{\mathfrak{p}}(M)}.$$

Proof. Put $K(N)' = K(N) \cap Sp_4(\widehat{\mathbb{Z}})$. The space $M_{(k,1)}(\Gamma(N)) = S_{(k,1)}(\Gamma(N)) \oplus M_{(k,1)}(\Gamma(N))^{\mathrm{Kl}}$ can be realized by $\mathrm{Hom}(\omega_l, L^2(Sp_4(\mathbb{Q})\backslash Sp_4(\mathbb{A}_{\mathbb{Q}}))^{K(N)'}$ as seen in § 4.9 where the latter space stands for the space of Klingen Eisenstein series of weight $(k, 1)$ with respect to $\Gamma(N)$. By the classification for the CAP forms it coincides with

$$\bigoplus_{(K,T,\widehat{\sigma})} \theta(T, t, \widehat{\sigma})^{K(N)'},$$

where $(K, T, \widehat{\sigma})$ runs over all triple defining $\theta(T, t, \widehat{\sigma})$ up to isomorphism according to Lemma -(2) and σ runs over all Hecke characters of weight $k-1$ at infinity. By Lemma -(3), (4), the fixed vectors at each finite place can be computed by using explicit realization given there (this is done in [74, § 5]). Then the argument in [74, § 6] yields the claim without any change. □

By using this we have

Proposition 3.7. *The notation being as above. Then for any $\varepsilon > 0$,*

$$\dim S_{(k,1)}(\Gamma(N)) = O(N^{6+\varepsilon}) \quad \text{as } N \to \infty.$$

Proof. The trivial bound is $w(K) \leqslant 6$, $|E_{K,\mathfrak{f}}| \leqslant 6$. By definition, $N_1(K) \leqslant \frac{N}{D_K}$ and then

$$\prod_{p|N_1(K)} (1 - p^{-4}) \leqslant \prod_{p} (1 - p^{-4}) = \zeta(4)^{-1}.$$

If $\mathfrak{p}$ is inert, $N(\mathfrak{p}) = p^2$, where $\mathfrak{p} = p$ is a rational prime. If $\mathfrak{p}$ is ramified, $N(\mathfrak{p}) = p$, where $p = \mathfrak{p}^2$. Since $1 - p^{-2} \leqslant 1 - p^{-4}$, one has

$$\prod_{\mathfrak{p}|\mathfrak{a}} (1 - N(\mathfrak{a})^{-2}) \leqslant \prod_{p} (1 - p^{-4}) = \zeta(4)^{-1}.$$

Hence

$$p(\mathfrak{a}, N, K) \leqslant \zeta(4)^{-2} N^4 D_K^{-4} N(\mathfrak{a})^2.$$

If $\mathfrak{p}$ is inert and $\mathfrak{p}|N$, then $2[\mathrm{ord}_{\mathfrak{p}}(NM)/2] - \mathrm{ord}_{\mathfrak{p}}(M) \leqslant \mathrm{ord}_{\mathfrak{p}}(N)$. If $\mathfrak{p}$ is ramified, then $\mathfrak{p}|N$ and $\mathrm{ord}_{\mathfrak{p}}(M) = 0$ by definition. Hence

$$\sum_{\mathfrak{a}} \delta\left(K, k, \frac{a}{N_1}(K)\right) |E_{K,\mathfrak{a}N_1(K)}|\ p(\mathfrak{a}, N, K) \ll N^4 D_K^{-4} \left(\sum_{d|D_K^2} d^2\right)\left(\sum_{d|N/D_K} d^2\right)$$

$$= N^4 D_K^{-4} \sigma_2(D_K^2) \sigma_2\left(\frac{N}{D_K}\right) \ll N^{6+\epsilon} D_K^{-2+\epsilon}.$$

The sum over M in (3.10) is majorized by $d(N)$. Since $h(K) \ll D_K^{1/2+\epsilon}$,

$$\dim S_{(k,1)}(\Gamma(N)) \leqslant \dim M_{(k,1)}(\Gamma(N)) \ll N^{6+\epsilon} d(N) \sum_{D_K|N} D_K^{-3/2+\epsilon} \ll N^{6+\epsilon}. \qquad \square$$

3.6. Automorphic counting measures

Let S' be a finite set of rational primes. Let $\xi = \xi_{\underline{k}}$ be an irreducible algebraic representation of $G(\mathbb{R})$ with the highest weight (k_1, k_2) satisfying $k_1 \geqslant k_2 \geqslant 3$ as in (3.5), and $D^{\mathrm{hol}}_{l_1,l_2}$ be the holomorphic discrete series of $G(\mathbb{R})$ with the Harish–Chandra parameter $(l_1, l_2) = (k_1 - 1, k_2 - 2)$ and whose central character equals $\chi_{\xi^\vee}$ on $A_{G,\infty}$. Here $\chi_{\xi^\vee}$ is the central character of $\xi^\vee$. We fix a pseudo-coefficient $f_\xi \in C_c^\infty(G(\mathbb{R}), \xi^\vee)$ (cf. [17], [2, p. 266], or [3, p. 161]), so that

$$\mathrm{tr}(\pi_\infty(f_\xi)) = (-1)^{q(G(\mathbb{R}))} = -1, \ \ \pi_\infty = D^{\mathrm{hol}}_{l_1,l_2},$$

where $q(G(\mathbb{R})) = \dim G(\mathbb{R})/A_{G,\infty}U(2) = 3$. By [27], we have,

$$\text{if } l_2 > 1, \quad \mathrm{tr}(\pi_\infty(f_\xi)) = \begin{cases} -1, & \text{if } \pi_\infty = D^{\mathrm{hol}}_{l_1,l_2}, \\ 0, & \text{otherwise}; \end{cases} \tag{3.11}$$

$$\text{if } l_1 > 2 \text{ and } l_2 = 1, \quad \operatorname{tr}(\pi_\infty(f_\xi)) = \begin{cases} -1, & \text{if } \pi_\infty = D^{\mathrm{hol}}_{l_1,l_2}, \\ 1, & \text{if } \pi_\infty = \omega_{l_1}, \\ 0, & \text{otherwise}; \end{cases}$$

$$\text{if } (l_1, l_2) = (2, 1), \quad \operatorname{tr}(\pi_\infty(f_\xi)) = \begin{cases} -1, & \text{if } \pi_\infty = D^{\mathrm{hol}}_{l_1,l_2}, \\ 1, & \text{if } \pi_\infty = \omega_{l_1} \text{ or } \mathbf{1}, \\ 0, & \text{otherwise}. \end{cases}$$

Let $\pi^0_{S'}$ be a given unitary representation of $G(\mathbb{Q}_{S'})$. Let $\delta_{\pi^0_{S'}}$ be the Dirac delta measure supported on $\pi^0_{S'}$ with respect to the Plancherel measure $\widehat{\mu}^{\mathrm{pl}}_{S'}$ on $\widehat{G(\mathbb{Q}_{S'})}$. Then we define a normalized Dirac delta measure supported on $\pi^0_{S'}$ by

$$\delta_{\pi^0_{S'},\xi}(\widehat{f_{S'}}) := \sum_i a_i |\nu(\alpha_i)|_{S'}^{-(k_1+k_2)/2} \varepsilon_{S'}(\alpha_i)^{-1} \delta_{\pi^0_{S'}}(\widehat{f_{S'}}) \tag{3.12}$$

for $f_{S'} = \sum_i a_i [G(\mathbb{Z}_{S'})\alpha_i G(\mathbb{Z}_{S'})] \in C_c^\infty(G(\mathbb{Q}_{S'}))$ with respect to the normalization of (2.23). The factor $\varepsilon_{S'}(\alpha_i)$ is defined by

$$\varepsilon_{S'}(\alpha_i) = \begin{cases} 2, & \text{if } \nu(\alpha_i) \in (\mathbb{Q}_{S'}^\times)^2, -1 \in \nu(U), \text{ and } Sp_4(\mathbb{Z}) \cap U \text{ has non-trivial center}, \\ 1, & \text{otherwise}. \end{cases} \tag{3.13}$$

We do not need the property of the Plancherel measure except for the following Plancherel formula by Harish–Chandra:

$$\widehat{\mu}^{\mathrm{pl}}_{S'}(\widehat{f_{S'}}) = f_{S'}(1). \tag{3.14}$$

For the above ξ and any compact open subgroup U of $G(\mathbb{A}^{S',\infty})$, we define a counting measure on $\widehat{G(\mathbb{Q}_{S'})}$ by

$$\begin{aligned} \widehat{\mu}_{U,\xi_{\underline{k}},D^{\mathrm{hol}}_{l_1,l_2}} := {} & \frac{1}{\operatorname{vol}(G(\mathbb{Q})A_{G,\infty}\backslash G(\mathbb{A})) \cdot \dim \xi_{\underline{k}}} \\ & \times \sum_{\pi^0_{S'} \in \widehat{G(\mathbb{Q}_{S'})}} \mu^{S',\infty}(U) m_{\mathrm{cusp}}(\pi^0_{S'}; U, \xi_{\underline{k}}, D^{\mathrm{hol}}_{l_1,l_2}) \delta_{\pi^0_{S'},\xi} \end{aligned} \tag{3.15}$$

where for a given unitary representation $\pi^0_{S'}$ of $G(\mathbb{Q}_{S'})$, the normalized multiplicity $m_{\mathrm{cusp}}(\pi^0_{S'}; U, \xi_{\underline{k}}, D^{\mathrm{hol}}_{l_1,l_2})$ is given by

$$m_{\mathrm{cusp}}(\pi^0_{S'}; U, \xi_{\underline{k}}, D^{\mathrm{hol}}_{l_1,l_2}) = \sum_{\substack{\pi \in \Pi(G(\mathbb{A})) \\ \pi_{S'} \simeq \pi^0_{S'},\, \pi_\infty \simeq D^{\mathrm{hol}}_{l_1,l_2}}} m_{\mathrm{cusp}}(\pi) \operatorname{tr}(\pi^{S',\infty}(f_U)) \cdot \operatorname{tr}(\pi_\infty(f_\xi)), \tag{3.16}$$

where $\Pi(G(\mathbb{A}))$ stands for the set of all isomorphism classes of automorphic representations of $G(\mathbb{A})$.

Since ω_{l_1} occurs in both cuspidal spectrum and residual spectrum, we define, for $* \in \{\mathrm{cusp}, \mathrm{res}\}$,

$$\widehat{\mu}_{U,\xi_{\underline{k}},\omega_{l_1},*} := \frac{1}{\mathrm{vol}(G(\mathbb{Q})A_{G,\infty}\backslash G(\mathbb{A}))\cdot \dim \xi_{\underline{k}}} \sum_{\pi_S^0\in \widehat{G(\mathbb{Q}_S)}} \mu^{S,\infty}(U) m_*(\pi_S^0; U, \xi_{\underline{k}}, \omega_{l_1})\delta_{\pi_S^0,\xi}$$

where

$$m_*(\pi_S^0; U, \xi_{\underline{k}}, \omega_{l_1}) = \sum_{\substack{\pi\in\Pi(G(\mathbb{A}))\\ \pi_{S'}\simeq\pi_{S'}^0,\ \pi_\infty\simeq\omega_{l_1}}} m_*(\pi)\mathrm{tr}(\pi^{S',\infty}(f_U))\cdot \mathrm{tr}(\pi_\infty(f_\xi)),\ \ * \in \{\mathrm{cusp}, \mathrm{res}\}.$$

Since $\mathbf{1}$ occurs only in the residual spectrum, we define

$$\widehat{\mu}_{U,\xi_{\underline{k}},\mathbf{1}} := \frac{1}{\mathrm{vol}(G(\mathbb{Q})A_{G,\infty}\backslash G(\mathbb{A}))\cdot \dim \xi_{\underline{k}}} \sum_{\pi_{S'}^0\in \widehat{G(\mathbb{Q}_{S'})}} \mu^{S',\infty}(U) m_{\mathrm{res}}(\pi_{S'}^0; U, \xi_{\underline{k}}, \mathbf{1})\delta_{\pi_{S'}^0,\xi}$$

where

$$m_{\mathrm{res}}(\pi_{S'}^0; U, \xi_{\underline{k}}, \mathbf{1}) = \sum_{\substack{\pi\in\Pi(G(\mathbb{A}))\\ \pi_{S'}\simeq\pi_{S'}^0,\ \pi_\infty\simeq\mathbf{1}}} m_{\mathrm{res}}(\pi)\mathrm{tr}(\pi^{S',\infty}(f_U))\cdot \mathrm{tr}(\pi_\infty(f_\xi)).$$

If $U\cap Sp_4(\mathbb{Z})$ has a non-trivial center, then it contains $-E_4$. In the case when U contains $-E_4$ and $-1\in\nu(U)$, there are two ways to extend a classical form F to an adelic form ϕ of level U. One way is explained in (2.21). Since $d = \mathrm{diag}(-1, 1, 1, -1) \in U,\ \nu(d) = -1$, we have a holomorphic form F_d defined by

$$F_d(Z) := \overline{F(dZ)} = \overline{F\left(\begin{pmatrix} -z_1 & z_2 \\ z_2 & -z_3 \end{pmatrix}\right)}.$$

Then F_d can be also extended to an adelic form of level U. Both of them generate the same automorphic representation. This explains the meaning of (3.13).

4. Arthur's invariant trace formula and some calculations

In this section we make use of Arthur's invariant trace formula, and as in [62, 63], we relate the Plancherel measure with the spectral expansion of the trace formula. Then this leads to the calculation of the geometric side. In our setting the pseudo-coefficient is chosen in a single discrete series. This causes the contribution from unipotent elements. This contribution should be understood in terms of endoscopic representations which appear in the spectral side. In general we do not know how to control the unipotent contribution, but in our case we compute every terms very explicitly. We give several estimates for invariants which appear in the trace formula.

Recall the notations G, B, T, and M_j $(j = 0, 1, 2)$ given in § 2. Set $P_0 = B$ and $M_0 = T$. We denote by N_0 the unipotent radical of P_0. The Weyl group $W_0^G(= N_G(T)/T)$ for M_0 in G is generated by two elements s_0 and s_1 which satisfies the relations $s_0^2 = s_1^2 = 1$ and $s_0s_1s_0s_1 = s_1s_0s_1s_0$. We put $s_2 = s_0s_1s_0$. Then, we have

$$W_0^G = \{1,\ \ s_0,\ \ s_1,\ \ s_2,\ \ s_0s_1,\ \ s_0s_2,\ \ s_1s_2,\ \ s_0s_1s_2\}.$$

For s_0 and s_2 in W_0^G, their representatives w_{s_0} and w_{s_2} in $G(\mathbb{Q})\cap K$ can be chosen as

$$w_{s_0}=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix},\quad w_{s_2}=\begin{pmatrix}1&0&0&0\\0&0&0&1\\0&0&1&0\\0&-1&0&0\end{pmatrix}.$$

For all elements s in W_0^G, we fix their representatives w_s by w_{s_0}, w_{s_2}, and some products like $w_{s_1}=w_{s_0}w_{s_2}w_{s_0}$. We also find

$$W_0^{M_0}=1,\quad W_0^{M_1}=\{1,\ \ s_0\},\quad W_0^{M_2}=\{1,\ \ s_2\}.$$

For $s\in W_0^G$ and $H\subset G$, we set $sH=w_sHw_s^{-1}$. The set of all Levi subgroups containing M_0 is given by

$$\mathcal{L}=\{M_0,\ \ M_1,\ \ s_1M_1,\ \ M_2,\ \ s_0M_2,\ \ G\}.$$

Set

$$\mathbf{K}_\infty=\left\{\begin{pmatrix}A&B\\-B&A\end{pmatrix}\in G(\mathbb{R})\right\},\quad \mathbf{K}_v=G(\mathbb{Z}_v)\quad(\forall v<\infty).$$

Then $\mathbf{K}_v$ is a maximal compact subgroup of $G(\mathbb{Q}_v)$ and $\mathbf{K}=\prod_v\mathbf{K}_v$ is also a maximal compact subgroup of $G(\mathbb{A})$. We normalize Haar measures dk_v on $\mathbf{K}_v$ as $\int_{\mathbf{K}_v}dk_v=1$. A Haar measure dk on $\mathbf{K}$ is defined by $dk=\prod_v dk_v$. We also choose Haar measures dx_v on $\mathbb{Q}_v$ as $\int_{\mathbb{Z}_v}dx_v=1$ $(v<\infty)$ and the Lebesgue measure dx_∞ on $\mathbb{R}$. A Haar measure dx on $\mathbb{A}$ is defined by $dx=\prod_v dx_v$. For each M in $\mathcal{L}$, we fix Haar measures on $A_M(\mathbb{R})^0$ as in [29, Condition 5.1]. Moreover, we fix a Haar measure on $G(\mathbb{A})$. By the same manner as in [3, p. 32] we normalize Haar measures on $M(\mathbb{A})^1$.

4.1. Characters of holomorphic discrete series of $Sp_4(\mathbb{R})$

We recall character formulas for holomorphic discrete series of $Sp_4(\mathbb{R})$. These are necessary to control the geometric side $I_{\mathrm{geom}}(f)$ of Arthur's invariant trace formula.

Let

$$t_4(\theta_1,\theta_2)=\begin{pmatrix}\cos\theta_1&0&\sin\theta_1&0\\0&\cos\theta_2&0&\sin\theta_2\\-\sin\theta_1&0&\cos\theta_1&0\\0&-\sin\theta_2&0&\cos\theta_2\end{pmatrix}\quad(\theta_1,\theta_2\in\mathbb{R}).$$

We define a compact Cartan subgroup T_4 of $Sp_4(\mathbb{R})$ as

$$T_4=\{t_4(\theta_1,\theta_2)\mid\theta_1,\theta_2\in\mathbb{R}\}.$$

We write T_4^{reg} the subset of regular elements of T_4. For each (l_1,l_2) in $\mathbb{Z}\oplus\mathbb{Z}$, a function Θ_{l_1,l_2} on T_4^{reg} is defined by

$$\Theta_{l_1,l_2}(t_4(\theta_1,\theta_2))=\frac{-e^{il_1\theta_1+il_2\theta_2}+e^{il_2\theta_1+il_1\theta_2}}{(e^{i\theta_1}-e^{-i\theta_1})(e^{i\theta_2}-e^{-i\theta_2})(1-e^{i\theta_1+i\theta_2})(e^{-i\theta_1}-e^{-i\theta_2})}.\tag{4.1}$$

Assume that (l_1,l_2) satisfies $l_1>l_2>0$. Then, there exists a unique holomorphic discrete series D_{l_1,l_2} of $Sp_4(\mathbb{R})$ whose character equals Θ_{l_1,l_2} (cf. [35, Theorem 12.7]).

The parameter (l_1, l_2) is called the Harish–Chandra parameter for discrete series representations.

Throughout this section, we use the Harish–Chandra parameter (l_1, l_2) to describe holomorphic discrete series representation instead of the minimal $\mathbf{K}_\infty$-type (k_1, k_2). Since $(k_1, k_2) = (l_1 + 1, l_2 + 2)$, one can easily convert the results from one to the other.

For a, a_1, a_2, θ in $\mathbb{R}$, we set

$$t_0(a_1, a_2) = \mathrm{diag}(e^{a_1}, e^{a_2}, e^{-a_1}, e^{-a_2}),$$
$$t_1(a, \theta) = \begin{pmatrix} e^a \cos\theta & e^a \sin\theta & 0 & 0 \\ -e^a \sin\theta & e^a \cos\theta & 0 & 0 \\ 0 & 0 & e^{-a}\cos\theta & e^{-a}\sin\theta \\ 0 & 0 & -e^{-a}\sin\theta & e^{-a}\cos\theta \end{pmatrix},$$
$$t_2(a, \theta) = \begin{pmatrix} e^a & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & e^{-a} & 0 \\ 0 & -\sin\theta & 0 & \cos\theta \end{pmatrix}.$$

It is known that the group $Sp_4(\mathbb{R})$ has the four Cartan subgroups T_0, T_1, T_2, T_4 up to conjugation, where

$$T_0 = \{t_0(a_1, a_2) \mid a_1, a_2 \in \mathbb{R}_{>0}\} \subset M_0(\mathbb{R}),$$
$$T_j = \{t_j(a, \theta) \mid a \in \mathbb{R}_{>0},\quad \theta \in \mathbb{R}\} \subset M_j(\mathbb{R}) \quad (j = 1, 2).$$

Let T_j^{reg} denote the set of regular elements of T_j. A character formula of Θ_{l_1,l_2} on T_j^{reg} is known (cf. [26, 28, 43]). Here we summarize it briefly. For $t_0(a_1, a_2)$ in T_0^{reg}, we have

$$\Theta_{l_1,l_2}(t_0(a_1, a_2)) = \frac{\{-e^{-l_1|a_1|-l_2|a_2|} + e^{-l_2|a_1|-l_1|a_2|}\} \times \mathrm{sgn}(a_1 a_2)}{(e^{a_1} - e^{-a_1})(e^{a_2} - e^{-a_2})(1 - e^{a_1+a_2})(e^{-a_1} - e^{-a_2})} \tag{4.2}$$

if $|a_1| > |a_2| > 0$. For $t_0(a_1, a_2)$ in T_0^{reg} and $\delta_1 = \mathrm{diag}(1, -1, 1, -1)$, we get

$$\Theta_{l_1,l_2}(\delta_1 t_0(a_1, a_2)) = \frac{\{-(-1)^{l_2} e^{-l_1|a_1|-l_2|a_2|} + (-1)^{l_1} e^{-l_2|a_1|-l_1|a_2|}\} \times \mathrm{sgn}(a_1 a_2)}{(e^{a_1} - e^{-a_1})(-e^{a_2} + e^{-a_2})(1 + e^{a_1+a_2})(e^{-a_1} + e^{-a_2})}. \tag{4.3}$$

For $t_1(a, \theta)$ in T_1^{reg}, we have

$$\Theta_{l_1,l_2}(t_1(a, \theta)) = \frac{\{-e^{-l_1(|a|+i\theta)-l_2(|a|-i\theta)} + e^{-l_2(|a|+i\theta)-l_1(|a|-i\theta)}\} \times \mathrm{sgn}(a)}{(e^{a+i\theta} - e^{-a-i\theta})(e^{a-i\theta} - e^{-a+i\theta})(1 - e^{2a})(e^{-a-i\theta} - e^{-a+i\theta})}. \tag{4.4}$$

For $t_2(a, \theta)$ in T_2^{reg}, we have

$$\Theta_{l_1,l_2}(t_2(a, \theta)) = \frac{\{e^{-l_1|a|+il_2\theta} - e^{-l_2|a|+il_1\theta}\} \times \mathrm{sgn}(a)}{(e^{i\theta} - e^{-i\theta})(e^a - e^{-a})(1 - e^{i\theta+a})(e^{-a} - e^{-i\theta})}. \tag{4.5}$$

Since $\Theta_{l_1,l_2}(-\gamma) = (-1)^{l_1+l_2}\Theta_{l_1,l_2}(\gamma)$ and $\Theta_{l_1,l_2}(g^{-1}\gamma g) = \Theta_{l_1,l_2}(\gamma)$ $(g,\ \gamma \in Sp_4(\mathbb{R}))$, the formulas (4.1)–(4.5) cover all cases for regular semisimple elements in $Sp_4(\mathbb{R})$.

A closed subgroup $Sp_4^{\pm}(\mathbb{R})$ of $G(\mathbb{R}) = GSp_4(\mathbb{R})$ is defined by

$$Sp_4^{\pm}(\mathbb{R}) = \left\{ g \in GL_4(\mathbb{R}) \;\middle|\; g \begin{pmatrix} O_2 & E_2 \\ -E_2 & O_2 \end{pmatrix} {}^t g = \pm \begin{pmatrix} O_2 & E_2 \\ -E_2 & O_2 \end{pmatrix} \right\}.$$

Note that an isomorphism $G(\mathbb{R}) \cong A_{G,\infty} \times Sp_4^{\pm}(\mathbb{R})$ holds. We write $\Theta^{\mathrm{hol}}_{l_1,l_2}$ for the holomorphic discrete series $D^{\mathrm{hol}}_{l_1,l_2}$ of $G(\mathbb{R})$ $(l_1 > l_2 > 0)$. For the algebraic representation ξ of $G(\mathbb{R})$ corresponding to (l_1, l_2), we denote by χ_ξ the central character of ξ. For $\delta = \mathrm{diag}(1, 1, -1, -1) \in Sp_4^{\pm}(\mathbb{R})$, it is obvious that

$$Sp_4^{\pm}(\mathbb{R}) = Sp_4(\mathbb{R}) \sqcup Sp_4(\mathbb{R})\delta.$$

Hence, considering the action of δ, one finds

$$\Theta^{\mathrm{hol}}_{l_1,l_2}(zg) = \chi_\xi(z)^{-1} \times \{\Theta_{l_1,l_2}(g) + \overline{\Theta_{l_1,l_2}(g)}\} \quad (z \in A_{G,\infty}, \quad g \in Sp_4(\mathbb{R})). \tag{4.6}$$

Lemma 4.1. *We get $\Theta^{\mathrm{hol}}_{l_1,l_2}(\gamma) = 0$ for any regular semisimple element γ in $A_{G,\infty}Sp_4(\mathbb{R})\delta$.*

Proof. Let H_{l_1,l_2} denote a representation space of D_{l_1,l_2}. There exists an anti-holomorphic discrete series $\overline{D_{l_1,l_2}}$ with the same infinitesimal character as D_{l_1,l_2}. The Hilbert space H_{l_1,l_2} is also regarded as a representation space of $\overline{D_{l_1,l_2}}$, because $\overline{D_{l_1,l_2}}$ can be defined by $\overline{D_{l_1,l_2}}(g)v = D_{l_1,l_2}(\delta g\delta)v$ $(v \in H_{l_1,l_2})$. Therefore, the space $H_{l_1,l_2} \oplus H_{l_1,l_2}$ becomes a representation space of $D^{\mathrm{hol}}_{l_1,l_2}$. Namely, we have

$$D^{\mathrm{hol}}_{l_1,l_2}(g)(v_1, v_2) = (D_{l_1,l_2}(g)v_1, D_{l_1,l_2}(\delta g\delta)v_2), \quad D^{\mathrm{hol}}_{l_1,l_2}(\delta)(v_1, v_2) = (v_2, v_1)$$

for each vector (v_1, v_2) in $H_{l_1,l_2} \oplus H_{l_1,l_2}$ and each element $g \in Sp_4(\mathbb{R})$. By an orthonormal basis $\{v_j\}_{j=1}^{\infty}$ of H_{l_1,l_2}, we can choose an orthonormal basis $\{(v_j, 0), \ (0, v_k) \mid j, k = 1, 2, \dots\}$ of $H_{l_1,l_2} \oplus H_{l_1,l_2}$. Hence, for any function f in $C_c^\infty(GSp_4(\mathbb{R}))$ whose support is contained in $A_{G,\infty}Sp_4(\mathbb{R})\delta$, it follows that

$$\langle D^{\mathrm{hol}}_{l_1,l_2}(f)(v_j, 0), (v_j, 0)\rangle = \langle D^{\mathrm{hol}}_{l_1,l_2}(f)(0, v_k), (0, v_k)\rangle = 0 \quad (j, k = 1, 2, \dots).$$

This is obvious if one sees the action of $D^{\mathrm{hol}}_{l_1,l_2}(\delta)$. Thus, this lemma is proved. □

Let $D^M(\gamma)$ denote the Weyl denominator of γ in $M(\mathbb{R})$ and let $W(M, T)$ denote the Weyl group with respect to a torus T in M over $\mathbb{R}$. For each $0 \leqslant j \leqslant 2$, we set $M = M_j$ and T is a torus in M over $\mathbb{R}$ such that $T(\mathbb{R}) = T_j \sqcup (-T_j)$ when $j = 1$ or 2, and $T(\mathbb{R}) = M_0(\mathbb{R}) = T_0 \sqcup (-T_0) \sqcup \delta_1 T_0 \sqcup (-\delta_1)T_0$ when $j = 0$. In case of $M = G$, there exists a torus T over $\mathbb{R}$ such that $T(\mathbb{R}) = T_4$. We say that $\Theta^{\mathrm{hol}}_{l_1,l_2}$ is stable for M if $|D^M(\gamma)|^{-1/2}|D^G(\gamma)|^{1/2}\Theta^{\mathrm{hol}}_{l_1,l_2}(\gamma)$ on regular elements γ of $T(\mathbb{R})$ is a finite, $W(M, T)$-invariant linear combination of quasi-characters. This condition is the same as the assumption for $\Phi(\gamma)$ in [2, Lemma 4.1 in p. 271].

Lemma 4.2. *The character $\Theta^{\mathrm{hol}}_{l_1,l_2}$ is stable for M_0, M_1, M_2, and is not stable for G.*

Proof. This can be proved by using (4.1)–(4.5), (4.6) and Lemma 4.1. □

4.2. Spectral side

Let ξ be an irreducible algebraic representation of $G(\mathbb{R})$ with the highest weight (k_1, k_2) with $k_1 \geqslant k_2 \geqslant 3$, and $D^{\mathrm{hol}}_{l_1,l_2}$ denote the holomorphic discrete series representation of $G(\mathbb{R})$ with the Harish–Chandra parameter (l_1, l_2) so that the central character is same as $\xi^\vee$

on $A_{G,\infty}$, where $(l_1, l_2) = (k_1 - 1, k_2 - 2)$. Choose a test function h in $C_c^\infty(G(\mathbb{A}_{\mathrm{fin}}))$ and we write f_ξ for a pseudo-coefficient of $D^{\mathrm{hol}}_{l_1,l_2}$. If we set

$$f = f_\xi h \tag{4.7}$$

then f belongs to the Hecke algebra of $\mathbf{K}$-finite functions in $C_c^\infty(G(\mathbb{A})^1)$.

By (3.11), for $\underline{l} = (l_1, l_2)$, we set

$$\Pi(\underline{l}, \xi) = \begin{cases} \{\, D^{\mathrm{hol}}_{l_1,l_2} \,\} & \text{if } l_2 > 1, \\ \{\, D^{\mathrm{hol}}_{l_1,l_2},\ \omega_{l_1} \,\} & \text{if } l_1 > 2 \text{ and } l_2 = 1, \\ \{\, D^{\mathrm{hol}}_{l_1,l_2},\ \omega_2,\ \mathbf{1} \,\} & \text{if } (l_1, l_2) = (2, 1). \end{cases} \tag{4.8}$$

They correspond to (1) $k_1 \geqslant k_2 \geqslant 4$, (2) $k_1 > k_2 = 3$, and (3) $k_1 = k_2 = 3$, respectively in terms of the classical weight for Siegel modular forms. Then the spectral side of Arthur's invariant trace formula for f is

$$I_{\mathrm{spec}}(f) = \sum_{\pi=\pi_\infty\otimes\pi_{\mathrm{fin}}\in\widehat{G(\mathbb{A})},\ \pi_\infty\in\Pi(\underline{l},\xi)} m_{\mathrm{disc}}(\pi)\mathrm{tr}(\pi_\infty(f_\xi))\,\mathrm{tr}(\pi_{\mathrm{fin}}(h)) \tag{4.9}$$

where $m_{\mathrm{disc}}(\pi)$ denotes the multiplicity of π in the discrete spectrum of $L^2_{\mathrm{disc}}(G(\mathbb{Q})\backslash G(\mathbb{A}), \chi_{\xi^\vee})$ and the unramified Hecke action inside $\mathrm{tr}(\pi_{\mathrm{fin}}(h))$ is normalized as (2.22). For the proof of this expansion, we refer to [2, § 3]. It is fortunate that cohomological, non-holomorphic Saito–Kurokawa representations do not appear in this case. If f_ξ is a pseudo-coefficient of a large discrete series whose parameter satisfies $|l_1 - l_2| = 1$, then it appears on the spectral side.

We are now ready to relate the above measures to the spectral side $I_{\mathrm{spec}}(f)$ and also to the geometric side in Arthur's trace formula, as in [62, Proposition 4.2].

Proposition 4.3. *Let S' be a finite set of finite places of $\mathbb{Q}$. For any compact subgroup U of $G(\mathbb{A}^{S',\infty})$ and $f_{S'} = \sum_i a_i f_{S',\alpha_i} \in C_c^\infty(G(\mathbb{Q}_{S'}))$, $f_{S',\alpha_i} = [G(\mathbb{Z}_{S'})\alpha_i G(\mathbb{Z}_{S'})]$, $\alpha_i \in T(\mathbb{Q}_{S'})$,*

$$\sum_i a_i \frac{I_{\mathrm{geom}}(f_U f_{S',\alpha_i} f_\xi)}{\varepsilon_{S'}(\alpha_i)\overline{\mu}(G(\mathbb{Q})A_{G,\infty}\backslash G(\mathbb{A}))\dim\xi} = \sum_i a_i \frac{I_{\mathrm{spec}}(f_U f_{S',\alpha_i} f_\xi)}{\varepsilon_{S'}(\alpha_i)\overline{\mu}(G(\mathbb{Q})A_{G,\infty}\backslash G(\mathbb{A}))\dim\xi}$$

$$= \begin{cases} \widehat{\mu}_{U,\xi_{\underline{k}},D^{\mathrm{hol}}_{l_1,l_2}}(\widehat{f}_{S'})), & \text{if } l_2 > 1, \\ \widehat{\mu}_{U,\xi_{\underline{k}},D^{\mathrm{hol}}_{l_1,l_2}}(\widehat{f}_{S'})) + \sum_{*\in\{\mathrm{cusp},\mathrm{res}\}} \widehat{\mu}_{U,\xi_{\underline{k}},\omega_{l_1},*}(\widehat{f}_{S'}), & \text{if } l_1 > 2 \text{ and } l_2 = 1, \\ \widehat{\mu}_{U,\xi_{\underline{k}},D^{\mathrm{hol}}_{l_1,l_2}}(\widehat{f}_{S'})) + \widehat{\mu}_{U,\xi_{\underline{k}},\mathbf{1}}(\widehat{f}_{S'}) + \sum_{*\in\{\mathrm{cusp},\mathrm{res}\}} \widehat{\mu}_{U,\xi_{\underline{k}},\omega_{l_1},*}(\widehat{f}_{S'}), & \text{if } (l_1, l_2) = (2, 1), \end{cases}$$

where f_U is the characteristic function of U and the factor $\varepsilon_{S'}(\alpha_i)$ is defined by (3.13).

Proof. The claim follows from the definition and $\delta_{\pi^0_{S'}}(\widehat{f}_{S'}) = \int_{\widehat{G(\mathbb{Q}_{S'})}} \widehat{f}_{S'}\, d\delta_{\pi^0_{S'}} = \mathrm{tr}(\pi^0_{S'}(\widehat{f}_{S'}))$. □

4.3. Geometric side

Fix a finite set S' of finite places. Let S_0 be a finite set of finite places containing S' and put $S = S_0 \cup \{\infty\}$. We consider Hecke operators for $G(\mathbb{Q}_{S'})$ while we vary S_0 and hence S in the geometric side of the Arthur's trace formula. Put $q(G(\mathbb{R})) = \frac{1}{2}\dim G(\mathbb{R})/K_\infty A_{G,\infty} = 3$. Assume that S is sufficiently large. Then, the geometric side of Arthur's invariant trace formula is, for $f = f_\xi h$ as in (4.7),

$$I_{\text{geom}}(f) = \sum_{M\in\mathcal{L}} (-1)^{q(G(\mathbb{R}))+\dim(A_M/A_G)} \frac{|W_0^M|}{|W_0^G|} \sum_{\gamma\in(M(\mathbb{Q}))_{M,S}} a^M(S,\gamma)\, I_M^G(\gamma, f_\xi)\, J_M^M(\gamma, h_P), \tag{4.10}$$

where $(M(\mathbb{Q}))_{M,S}$ denotes the set of (M,S)-equivalence classes in $M(\mathbb{Q})$ (cf. [3, p. 113]) which turns out to be a finite set, for each M in $\mathcal{L}$ we choose a parabolic subgroup P such that M is a Levi subgroup of P, and we set

$$h_P(m) = \delta_P(m)^{1/2} \int_{\mathbf{K}_{S_0}} \int_{N_P(\mathbb{Q}_{S_0})} h(k^{-1}mnk)\, dn\, dk \quad (m \in M(\mathbb{Q}_{S_0})) \tag{4.11}$$

($\mathbf{K}_{S_0} = \prod_{v\in S_0} \mathbf{K}_v$ and δ_P is the modular function of $P(\mathbb{A})$). Regarding $J_M^M(\gamma, h_P)$ it follows from (4.11) that $h_P(m) = 0$ unless $k^{-1}m^{-1}\gamma mnk \in \mathrm{Supp}(h)$ where $m \in M(\mathbb{Q}_{S_0})$, $k \in K_{S_0}$, and $n \in N_P(\mathbb{Q}_{S_0})$. This implies

$$\nu(\gamma) \in \nu(\mathrm{Supp}(h)). \tag{4.12}$$

For the definitions of the invariant weighted orbital integral $I_M^G(\gamma, f_\xi)$ and the orbital integral $J_M^M(\gamma, h_P)$, we refer to [3, §§ 18 and 23]. The factor $a^M(S,\gamma)$ is called the global coefficient (see [3, § 19]). We later give their details for some cases which are necessary to explain our estimations.

Let Z_G denote the center of G. For each element γ in $G(\mathbb{A})$, we write $\{\gamma\}_G$ for the $G(\mathbb{Q})$-conjugacy class of γ. For convenience, we set

$$u_{\min}(x) = \begin{pmatrix} 1 & 0 & x & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \delta_1(x,y) = \begin{pmatrix} 1 & x & 0 & y \\ 0 & -1 & -y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & x & -1 \end{pmatrix},$$
$$u_{\min} = u_{\min}(1), \quad \delta_1 = \delta_1(0,0).$$

It is clear that all minimal unipotent elements belong to $\{u_{\min}\}_G$. Let γ be an element of $G(\mathbb{Q})$. If γ is a semisimple element whose diagonalization is δ_1, then γ is $G(\mathbb{Q})$-conjugate to δ_1. This fact can be proved by using the Galois cohomology.

To study concretely the geometric side, we separate the sum into the following seven types:

$$I_{\text{geom}}(f) = I_1(f) + I_2(f) + I_3(f) + I_4(f) + I_5(f) + I_6(f) + I_7(f)$$

where

- $I_1(f)$: $M = G$ and $\gamma \in Z_G(\mathbb{Q})$;
- $I_2(f)$: $M = G$ and $\gamma \in Z_G(\mathbb{Q})\{u_{\min}\}_G$;

- $I_3(f)$: $M=G$ and $\gamma\in Z_G(\mathbb{Q})\{\delta_1\}_G$;
- $I_4(f)$: $M=G$ and γ is semisimple and $\gamma\notin Z_G(\mathbb{Q})\sqcup Z_G(\mathbb{Q})\{\delta_1\}_G$;
- $I_5(f)$: $M=G$ and γ is not semisimple and $\gamma\notin Z_G(\mathbb{Q})\{u_{\min}\}_G$;
- $I_6(f)$: $M\neq G$ and γ is semisimple;
- $I_7(f)$: $M\neq G$ and γ is non-semisimple.

The main term will be $I_1(f)$ and the second main term will be $I_2(f)$ in general, but also $I_3(f)$ in weight aspect. The terms $I_4(f)$, $I_5(f)$, $I_6(f)$ and $I_7(f)$ never contribute to both the main term and the second main term in any aspect. We estimate each terms after the detailed studies of invariants $a^M(S,\gamma)$, $I_M^G(\gamma,f_\xi)$, and $J_M^M(\gamma,h_P)$. Since we clearly know

$$I_1(f)=(-1)^{q(G(\mathbb{R}))}\sum_{z\in Z_G(\mathbb{Q})}\mathrm{vol}(G(\mathbb{Q})\backslash G(\mathbb{A})^1)\,f(z),$$

we do not discuss it throughout this section.

If $M=G$, then $I_G^G(\gamma)=J_G^G(\gamma)$ and $h_Q=h$. For simplicity, we set

$$J_G(\gamma,f_\xi)=I_G^G(\gamma,f_\xi),\qquad J_G(\gamma,h)=J_G^G(\gamma,h).$$

4.4. Some measures concerning $I_2(f)$ and $I_3(f)$

We choose some measures on centralizers to calculate explicitly the orbital integrals $J_G(zu_{\min},f_\xi)$ and $J_G(z\delta_1,f_\xi)$. The centralizer $G_{u_{\min}}$ of $u_{\min}$ in G is given by

$$G_{u_{\min}}=\left\{z\begin{pmatrix}1&*&*&*\\0&1&*&0\\0&0&1&0\\0&0&*&1\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&*&0&*\\0&0&1&0\\0&*&0&*\end{pmatrix}\in G,\ z\in Z_G\right\}.$$

The centralizer G_{δ_1} of δ_1 satisfies

$$G_{\delta_1}=\left\{\begin{pmatrix}*&0&*&0\\0&*&0&*\\*&0&*&0\\0&*&0&*\end{pmatrix}\in G\right\}\cong\{(g_1,g_2)\in GL_2\times GL_2\mid \det(g_1)=\det(g_2)\}.$$

If we want to determine $J_G(zu_{\min},f_\xi)$ and $J_G(z\delta_1,f_\xi)$ precisely, we should choose measures on the centralizers. A Haar measure on $\mathbb{R}_{>0}$ is chosen by $x^{-1}dx$ and a Haar measure on $SL_2(\mathbb{R})$ is fixed by $(2\pi)^{-1}v^{-3}\,du\,dv\,d\theta$ for $\left(\begin{smallmatrix}1&u\\0&1\end{smallmatrix}\right)\left(\begin{smallmatrix}v&0\\0&v^{-1}\end{smallmatrix}\right)\left(\begin{smallmatrix}\cos\theta&\sin\theta\\-\sin\theta&\cos\theta\end{smallmatrix}\right)$. On the groups $\{\pm E_4\}$ and $\{E_4,\ \delta\}$, we take the counting measure. By the isomorphism $G_{u_{\min}}(\mathbb{R})\cong\{\pm E_4\}\times\mathbb{R}_{>0}\times(\mathbb{R}^3\ltimes SL_2(\mathbb{R}))$ (respectively $G_{\delta_1}(\mathbb{R})\cong\{E_4,\ \delta\}\rtimes(\mathbb{R}_{>0}\times SL_2(\mathbb{R})\times SL_2(\mathbb{R}))$), we obtain a Haar measure on $G_{u_{\min}}(\mathbb{R})$ (respectively $G_{\delta_1}(\mathbb{R})$).

To simplify the description for the global coefficient $a^G(S,u_{\min})$, we choose measures on the orbits as below. We define $J_G(zu_{\min},f_\xi)$ and $J_G(zu_{\min},h)$ as

$$J_G(zu_{\min},f_\xi)=\int_{\mathbb{R}}\int_{\mathbf{K}_\infty}f_\xi(zk^{-1}u_{\min}(x)k)\,|x|_\infty\,dk\,dx,\tag{4.13}$$

$$J_G(zu_{\min},h)=\prod_{p\in S_0}(1-p^{-1})^{-1}\times\int_{\mathbb{Q}_{S_0}}\int_{\mathbf{K}_{S_0}}h(zk^{-1}u_{\min}(x)k)\,|x|_{S_0}\,dk\,dx.\tag{4.14}$$

If we choose a suitable Haar measure on $G(\mathbb{R})$, then the integral $J_G(zu_{\min})$ coincides with the orbital integral of $zu_{\min}$ normalized by the above mentioned measure on $G_{u_{\min}}(\mathbb{R})$.

We may define $J_G(z\delta_1, h)$ as

$$J_G(z\delta_1, h) = \int_{\mathbb{Q}_{S_0}} \int_{\mathbb{Q}_{S_0}} \int_{\mathbf{K}_{S_0}} h(zk^{-1}\delta_1(x, y)k)\, dk\, dx\, dy. \tag{4.15}$$

This definition is useful to compute spherical Hecke algebras.

We determine the total contributions $I_2(f)$ and $I_3(f)$ up to constant multiples (cf. Lemmas 4.11, 4.12, 4.13 and 4.14). We do not explicitly calculate the constants, because it is unnecessary for our main purpose. However, if one wants to know their numerical values, one can explicitly calculate them by using Lemmas 4.6, 4.7, (4.17), and choosing some normalizations of measures.

4.5. Estimations and vanishings for $I_M^G(\gamma, f_\xi)$

By [4, 5] one knows

$$I_M^G(\gamma, f_\xi) = (-1)^{\dim \mathfrak{a}_M^G} |D^G(\gamma)|^{1/2}\, \Theta^{\mathrm{hol}}_{l_1,l_2}(\gamma)$$

for any G-regular semisimple element γ which is $\mathbb{R}$-elliptic in M. If γ is not $\mathbb{R}$-elliptic, then $I_M^G(\gamma, f_\xi)$ vanishes. Hence, our remaining work is to study its behaviors at singular elements.

Lemma 4.4. *Let M be a proper Levi subgroup in $\mathcal{L}$. Then, for any $M(\mathbb{R})$-conjugacy class γ in $M(\mathbb{Q})$, there exists a positive constant $c(\gamma)$ such that the absolute value of $I_M^G(\gamma, f_\xi)$ is bounded by $c(\gamma) \times \chi_\xi(\gamma)^{-1} \times \{l_1 + l_2\}$. Furthermore, we have $I_M^G(\gamma, f_\xi) = 0$ if the semisimple part of γ is not $\mathbb{R}$-elliptic in M. In addition, the term $I_7(f)$ vanishes.*

Proof. By Lemma 4.2 we can apply the same argument as in [2, Proof of Theorem 5.1] to $I_M^G(\gamma, f_\xi)$ related to M_0, M_1, s_1M_1, M_2 and s_0M_2, because the character satisfies the same assumption as in [2, Lemma 4.1]. Hence, one can explicitly compute $I_M^G(\gamma, f_\xi)$. In particular, $I_M^G(\gamma, f_\xi)$ vanishes for all non-semisimple conjugacy classes γ. Hence, we get $I_7(f) = 0$. □

A required estimation for $I_6(f)$ can be proved by this lemma (cf. § 5). Hence, it is enough to consider the terms related to $M = G$, i.e., $I_2(f)$, $I_3(f)$, $I_4(f)$, and $I_5(f)$.

For each non-semisimple conjugacy class γ in $G(\mathbb{R})$, the distribution $J_G(\gamma)$ on $C_c^\infty(G(\mathbb{R}))$ is expressed by a linear combination of limits of regular semisimple orbital integrals (cf. [2, Appendix]), but its coefficients are still unknown in general. Hence, we should consider them case by case.

Lemma 4.5. *Let γ be an element in $G(\mathbb{Q})$. We assume that the semisimple part of γ is not in $Z(\mathbb{Q})$ and γ does not belong to $Z_G(\mathbb{Q})\{\delta_1\}_G$. Then, there exists a positive constant $c(\gamma)$ such that the absolute value of $J_G(\gamma, f_\xi)$ is bounded by $c(\gamma) \times \chi_\xi(\gamma)^{-1} \times \{l_1 + l_2\}$. In particular, we have $J_G(\gamma, f_\xi) = 0$ if the semisimple part of γ is not $\mathbb{R}$-elliptic in G.*

Proof. As for semisimple singular elements, it was done by Langlands using Harish–Chandra's limit formula (cf. [37]). In case of $SL_2(\mathbb{R})$, one can study the limit formula for the unipotent elements in the book [35, § 3, Ch. XI]. Hence, one can easily calculate them (all such explicit calculations were done in [70]). □

Lemma 4.6. *If $\gamma = z\delta_1$ $(z \in Z_G(\mathbb{Q}))$, then there exists a positive constant $c(\gamma)$ such that*

$$J_G(\gamma, f_\xi) = c(\gamma) \times \chi_\xi(z)^{-1} \times (-1)^{l_2} l_1 l_2 \{1 + (-1)^{l_1 - l_2 - 1}\}.$$

If we choose the Haar measure given in § 4.4 on the centralizer, we have $c(\gamma) = 2^{-4}\pi^{-2}$.

Proof. This follows from the limit formula for $SL_2(\mathbb{R})$ (cf. [35]). □

Now, the remaining conjugacy classes are only unipotent orbits. The group G has the four unipotent classes; (1) regular, (2) subregular, (3) minimal, (4) unit.

Lemma 4.7. *If $\gamma = zu_{\min}$ $(z \in Z_G(\mathbb{Q}))$, then there exists a positive constant $c(\gamma)$ such that*

$$J_G(\gamma, f_\xi) = c(\gamma) \times \chi_\xi(z)^{-1} \times (l_1 - l_2)(l_1 + l_2).$$

If we choose the Haar measure given in § 4.4 on the centralizer, we have $c(\gamma) = -2^{-3}\pi^{-3}$.

Proof. This lemma can be proved by the limit formula of [53]. As for a suitable chamber and the constant $c(\gamma)$, we refer to [70, Lemmas 4.9 and 4.11] □

Lemma 4.8. *There is only one regular unipotent $G(\mathbb{R})$-conjugacy class u_{reg} in $G(\mathbb{Q})$. For $\gamma = zu_{\mathrm{reg}}$ $(z \in Z_G(\mathbb{Q}))$, we find $J_G(\gamma, f_\xi) = 0$.*

Proof. This obviously follows from the limit formulas of [10, 53]. □

Lemma 4.9. *There are two subregular unipotent $G(\mathbb{R})$-conjugacy classes $u_{\mathrm{sub},1}$ and $u_{\mathrm{sub},2}$ in $G(\mathbb{Q})$. For any z in $Z_G(\mathbb{Q})$, we have $J_G(zu_{\mathrm{sub},1}, f_\xi) = J_G(zu_{\mathrm{sub},2}, f_\xi) = 0$.*

Proof. For a real symmetric matrix S of degree 2, we set $u_{\mathrm{sub}}(S) = \begin{pmatrix} E_2 & S \\ O_2 & E_2 \end{pmatrix}$. Let $S_{++} = \mathrm{diag}(1, 1)$, $S_{+-} = \mathrm{diag}(1, -1)$, and $S_{--} = \mathrm{diag}(-1, -1)$. Then, $u(S_{++})$, $u(S_{+-})$, $u(S_{--})$ are representatives for subregular unipotent orbits of $Sp_4(\mathbb{R})$. But, the sum of the orbits of $u(S_{++})$ and $u(S_{--})$ forms a $G(\mathbb{R})$-conjugacy class. Then, we denote it by $u_{\mathrm{sub},1}$, and let $u_{\mathrm{sub},2}$ denote the $G(\mathbb{R})$-conjugacy class of $u(S_{+-})$.

Using the limit formulas [10, 53] and some associated constants [70, Lemma 4.11], one gets

$$J_{Sp_4}(zu(S_{++}), f_\xi) = -J_{Sp_4}(zu(S_{--}), f_\xi).$$

Hence, it follows that $J_G(zu_{\mathrm{sub},1}, f_\xi) = J_{Sp_4}(zu(S_{++}), f_\xi) + J_{Sp_4}(zu(S_{--}), f_\xi) = 0$. Since f_ξ is cuspidal, we deduce $\int_{\mathbf{K}_\infty} \int_{N_1(\mathbb{R})} f_\xi(zk^{-1}nk)\, dn\, dk = 0$ from the Plancherel formula for $M_1(\mathbb{R})$. By normalizing Haar measures on the centralizers, one can see that

$$J_G(zu_{\mathrm{sub},1}, f_\xi) + J_G(zu_{\mathrm{sub},2}, f_\xi) = \int_{\mathbf{K}_\infty} \int_{N_1(\mathbb{R})} f_\xi(zk^{-1}nk)\, dn\, dk = 0.$$

Hence, we get $J_G(zu_{\mathrm{sub},2}, f_\xi) = 0$. □

4.6. Global coefficients $a^G(S,\gamma)$

For details of global coefficients $a^G(S,\gamma)$, we refer to [3, § 19] and [29]. It is known that $a^G(S,1)=\mathrm{vol}(G(\mathbb{Q})\backslash G(\mathbb{A})^1)$ holds. For non-trivial unipotent orbits for G, they are explicitly calculated in [29]. By Lemmas 4.8 and 4.9, we need only an information for $a^G(S,u_{\min})$. Let γ be a (G,S)-conjugacy class in $G(\mathbb{Q})$. We shall consider the case γ is not unipotent. We may reduce to the centralizers of the semisimple part γ_s of γ. For each element γ_1 in $G(\mathbb{Q})$, we denote by $G_{\gamma_1,+}$ the centralizer of γ_1 in G over $\mathbb{Q}$ and by G_{γ_1} the connected component of 1 in $G_{\gamma_1,+}$. (Note that $G_{u_{\min}}=G_{u_{\min},+}$ and $G_{\delta_1}=G_{\delta_1,+}$.) We set $\iota(\gamma_s)=G_{\gamma_s,+}(\mathbb{Q})/G_{\gamma_s}(\mathbb{Q})$. If S is sufficiently large, then we have

$$a^G(S,\gamma)=\varepsilon^G(\gamma_s)\,|\iota^G(\gamma_s)|^{-1}\sum_{\{u:\gamma_s u\sim\gamma\}} a^{G_{\gamma_s}}(S,u)$$

where u runs over $G_{\gamma_s}(\mathbb{Q}_S)$-unipotent orbits in $G_{\gamma_s}(\mathbb{Q})$ such that $\gamma_s u$ are (G,S)-equivalent to γ and we set

$$\varepsilon^G(\gamma_s)=\begin{cases}1 & \text{if } \gamma_s \text{ is } \mathbb{Q}\text{-elliptic in } G,\\ 0 & \text{otherwise.}\end{cases}$$

Especially, if γ is semisimple, then we have

$$a^G(S,\gamma)=\varepsilon^G(\gamma)\,|\iota^G(\gamma)|^{-1}\,\mathrm{vol}(G_\gamma(\mathbb{Q})\backslash G_\gamma(\mathbb{A})^1). \tag{4.16}$$

Hence, from this and Lemma 4.5, one finds that a needed estimation for $I_4(f)$ is obviously reduced to some known results (cf. Proof of Proposition 5.3). Note that we carefully see the growth of $a^G(S,\gamma)$ with respect to S if γ is not semisimple. We also note that

$$a^G(S,z\gamma)=a^G(S,\gamma)$$

holds for any z in $Z_G(\mathbb{Q})$.

The following notations are necessary to describe $a^G(S,\gamma)$ explicitly. Let E be an algebraic number field and let $\chi=\prod_w\chi_w$ be a character on $E^\times\backslash\mathbb{A}_E^1\cong E^\times\mathbb{R}_{>0}\backslash\mathbb{A}_E^\times$. We set

$$S_E=\bigsqcup_{v\in S}\{w\mid w \text{ is a place of } E \text{ such that } w \text{ divides } v\},$$

$$L_E^S(s,\chi)=\prod_{w\notin S_E}L_{E,w}(s,\chi_w),$$

$$L_{E,w}(s,\chi_w)=\begin{cases}(1-\chi_w(\pi_w)\,q_w^{-s})^{-1} & \text{if } \chi_w \text{ is unramified,}\\ 1 & \text{if } \chi_w \text{ is ramified,}\end{cases}$$

where π_w is a prime element of E and q_w denotes the cardinality of the residue field of E_w. For the trivial representation $\mathbf{1}_E$ on $E^\times\backslash\mathbb{A}_E^1$, we set

$$\zeta_E^S(s)=L^S(s,\mathbf{1}_E).$$

We write c_E for the residue of $\zeta_E^{\Sigma_\infty}(s)$ where $\Sigma_\infty=\{w|\infty\}$. We denote $\mathfrak{c}_E(S)$ by the constant term of the Laurent expansion of $\zeta_E^S(s)$ at $s=1$, that is,

$$\zeta_E^S(s)=\frac{c_{E,S}^{-1}c_E}{s-1}+\mathfrak{c}_E(S)+(*)(s-1)+\cdots$$

where we set $c_{E,S} = \prod_{w \in S_E - \Sigma_\infty} (1 - q_w^{-1})^{-1}$. If $E = \mathbb{Q}$, then we set

$$L^S(s, \chi) = L^S_{\mathbb{Q}}(s, \chi), \quad \zeta^S(s) = \zeta^S_{\mathbb{Q}}(s), \quad \mathfrak{c}(S) = \mathfrak{c}_{\mathbb{Q}}(S)$$

for simplicity. We later use the following estimates.

Lemma 4.10. *Let $m \in \mathbb{N}$ be fixed. For any positive real number ε, there exists a positive constant $c(\varepsilon, m, E)$ such that*

$$\sum_{\chi} |\, L^S_E(1, \chi)\,|^m < c(\varepsilon, m, E) \times \prod_{p \in S_0} p^{\varepsilon}$$

where $\chi = \prod_w \chi_w$ runs over all non-trivial quadratic characters on $E^\times \backslash \mathbb{A}^1_E$ such that χ_w is unramified for any $w \notin S_E$.

Proof. Let $N(\chi)$ be the norm of the conductor of χ. Then by [73, Lemma 1.4],

$$N(\chi) \leqslant 2^{3n_E} \prod_{\substack{\mathfrak{p}_w \nmid 2 \\ w \in S_E - \Sigma_\infty}} N(\mathfrak{p}_w),$$

where $n_E = [E : \mathbb{Q}]$. Now $N(\mathfrak{p}_w) \leqslant p^{n_E}$. Hence $N(\chi) \leqslant 2^{2n_E} \prod_{p \in S_0} p^{n_E}$. By [41], $L_E(1, \chi) \ll_{E,\epsilon'} \exp\left(C \frac{\log N(\chi)}{\log\log N(\chi)}\right) \ll N(\chi)^{\epsilon'}$ for some constants C, ϵ'. Here

$$L^S_E(1, \chi) = L_E(1, \chi) \prod_{w \in S_E - \Sigma_\infty} L_{E,w}(1, \chi_w)^{-1}.$$

Hence

$$\left| \prod_{w \in S_E - \Sigma_\infty} L_{E,w}(1, \chi_w)^{-1} \right| \leqslant \left| \prod_{w \in S_E - \Sigma_\infty} (1 + \mathfrak{p}_w^{-1}) \right| \ll_E \prod_{p \in S_0} (1 + p^{-1})^{n_E}.$$

Now for $M = \prod_{p \in S_0} p$,

$$\log \prod_{p \in S_0} (1 + p^{-1}) = \sum_{p \in S_0} \log(1 + p^{-1}) \ll \sum_{p \in S_0} \frac{1}{p} \ll \sum_{p \leqslant M} \frac{1}{p} \ll \log\log M.$$

Since $N(\chi) \ll_E M^{n_E}$,

$$|L^S_E(1, \chi)| \ll_E \exp\left(C \frac{\log M}{\log\log M}\right) \log M \ll_{E,\epsilon'} M^{\epsilon'} \log M.$$

Hence for each $m \in \mathbb{N}$,

$$\sum_{\chi} |L^S_E(1, \chi)|^m \ll M^{\epsilon' m} (\log M)^m \sum_{d | M} 1 = M^{\epsilon' m} \phi(M) (\log M)^m \ll_{E,m,\epsilon} M^{\epsilon}. \qquad \square$$

For the $G(\mathbb{Q}_S)$-orbit of $u_{\min}$ and the chosen measures (4.13) and (4.14), we have

$$a^G(S, u_{\min}) = 2^{-1} \operatorname{vol}(M_2(\mathbb{Q}) \backslash M_2(\mathbb{A})^1)\, \zeta^S(2) \tag{4.17}$$

by [29, Theorem 6.1]. Next, we explain $a^{G_{\gamma_s}}(S,u)$ for non-semisimple and non-unipotent elements $\gamma=\gamma_s u$, which we call a mixed element. For all mixed elements of G, their global coefficients were studied in [29, § 3]. We mention only cases required for our estimation. Furthermore, we choose the same measures as in [29, § 3.4] on the unipotent orbits over $\mathbb{Q}_S$. However, we do not explain details for normalizations of measures, because they are unnecessary for estimations. Namely, the following equalities for the global coefficients hold under suitable normalizations of measures. By a classification of mixed elements in [29, § 5.3] it is enough to consider the following groups (as the centralizers of semisimple elements),

$$\begin{aligned} G_1 &= \{(g_1,g_2)\in GL_2\times GL_2 \mid \det(g_1)=\det(g_2)\}(\cong G_{\delta_1}),\\ G_2 &= \{g\in R_{E/\mathbb{Q}}(GL_2)\mid \det(g)\in GL_1\},\\ G_3 &= \{(x,g)\in R_{E/\mathbb{Q}}(GL_1)\times GL_2 \mid N_{E/\mathbb{Q}}(x)=\det(g)\} \end{aligned}$$

where E is a quadratic extension of $\mathbb{Q}$ and $R_{E/\mathbb{Q}}$ means the restriction of scalars. A unipotent element in $G_1(\mathbb{Q})$ can be written as

$$u_1(x,y)=\left(\begin{pmatrix}1&x\\0&1\end{pmatrix},\begin{pmatrix}1&y\\0&1\end{pmatrix}\right).$$

For the group G_1, representative elements of non-trivial unipotent orbits over $\mathbb{Q}_S$ are as follows:

$$u_1(1,0),\quad u_1(0,1),\quad u_1(\alpha,1)\ \ (\alpha\in\mathbb{Q}^\times/((\mathbb{Q}_S^\times)^2\cap\mathbb{Q}^\times)).$$

For α in $\mathbb{Q}^\times/((\mathbb{Q}_S^\times)^2\cap\mathbb{Q}^\times)$, there exist constants $c_{\min}(u_1)$ and $c_{\mathrm{reg}}(u_1)$ (which do not depend on S) such that

$$\begin{aligned} a^{G_1}(S,u_1(1,0)) &= a^{G_1}(S,u_1(0,1)) = c_{\min}(u_1)\times\mathfrak{c}(S),\\ a^{G_1}(S,u_1(\alpha,1)) &= c_{\mathrm{reg}}(u_1)\times\left\{\mathfrak{c}(S)^2+\sum_\chi \chi_S(\alpha)\,L^S(1,\chi)^2\right\}, \end{aligned} \tag{4.18}$$

where $\chi=\prod_v\chi_v$ runs over all non-trivial quadratic characters such that χ_v is unramified for any $v\notin S$, and we set $\chi_S=\prod_{v\in S}\chi_v$ (see [29, Example 3.9]). Representative elements of non-trivial unipotent $G_2(\mathbb{Q}_S)$-orbits in $G_2(\mathbb{Q})$ are

$$u_2(\alpha)=\begin{pmatrix}1&\alpha\\0&1\end{pmatrix}\quad (\alpha\in E^\times/(((E_{S_E}^\times)^2\mathbb{Q}_S^\times)\cap E^\times)).$$

Then, there exists a constant $c(u_2)$ (which does not depend on S) such that

$$a^{G_2}(S,u_2(\alpha))=c(u_2)\times\left\{\mathfrak{c}_E(S)+\sum_\chi \chi_{S_E}(\alpha)\,L_E^S(1,\chi)\right\} \tag{4.19}$$

where $\chi=\prod_w\chi_w$ runs over all non-trivial quadratic characters such that $\chi|_{\mathbb{A}_\mathbb{Q}^1}=1$ and χ_w is unramified for any $w\notin S_E$ (see [29, Example 3.8]). Representative elements of non-trivial unipotent $G_3(\mathbb{Q}_S)$-orbits in $G_3(\mathbb{Q})$ are

$$u_3(\alpha)=\left(1,\begin{pmatrix}1&\alpha\\0&1\end{pmatrix}\right)\quad (\alpha\in\mathbb{Q}^\times/(N_{E/\mathbb{Q}}(E_{S_E}^\times)\cap\mathbb{Q}^\times)),$$

where $N_{E/\mathbb{Q}}$ is the norm of $E/\mathbb{Q}$. Then, there exists a constant $c(u_3)$ (which does not depend on S) such that

$$a^{G_3}(S, u_3(\alpha)) = c(u_3) \times \left\{\mathfrak{c}(S) + \chi_S(\alpha)\, L^S(1, \chi)\right\} \tag{4.20}$$

where χ is the non-trivial quadratic character on $\mathbb{Q}^\times \backslash \mathbb{A}_{\mathbb{Q}}^1$ corresponding to E via the class field theory (see [29, § 3]).

4.7. Explicit calculations for $I_2(f)$

Recall $J_G(zu_{\min}, f_\xi)$ and $J_G(zu_{\min}, h)$ defined as (4.13) and (4.14).

Lemma 4.11. *Let $c(zu_{\min})$ denote the constant given in Lemma 4.7. Then, we have*

$$I_2(f) = \sum_{z \in Z(\mathbb{Q})} 2^{-1} \operatorname{vol}(M_2(\mathbb{Q}) \backslash M_2(\mathbb{A})^1)\, \zeta^S(2) \times c(zu_{\min}) \times \chi_\xi(z)^{-1} \times (l_1 - l_2)(l_1 + l_2) \times J_G(zu_{\min}, h).$$

In particular, $\zeta^S(2)$ is bounded by a positive constant for any S.

Proof. This lemma follows from Lemma 4.7 and (4.17). □

Lemma 4.12. *Assume that h_{a_1,a_2,a_3} is the characteristic function of the open compact set $\mathbf{K}_p \mathrm{diag}(p^{-a_1}, p^{-a_2}, p^{a_1-a_3}, p^{a_2-a_3})\mathbf{K}_p$ on $G(\mathbb{Q}_p)$ $(a_3 \geqslant a_1 \geqslant a_2 \geqslant 0)$. If a_3 is odd, then we have $J_G(zu_{\min}, h_{a_1,a_2,a_3}) = 0$. If a_3 is even, then we may set $a_3 = 2m$ and assume $m \geqslant a_1 \geqslant a_2$ by the action of W_0^G, and we get*

$$J_G(zu_{\min}, h_{a_1,a_2,a_3}) = \begin{cases} (1-p^{-2})^{-1} & \text{if } a_3 = 2m,\ a_1 = a_2 = m, \text{ and } |z|_p = p^m, \\ p^{a_3 - 2a_2} & \text{if } a_3 = 2m,\ m = a_1 > a_2, \text{ and } |z|_p = p^m, \\ 0 & \text{otherwise.} \end{cases}$$

If h_p is the characteristic function of $\{x \in \mathbf{K}_p \mid x \equiv E_4 \mod p^l \mathbb{Z}_p\}$ on $G(\mathbb{Q}_p)$, then

$$J_G(zu_{\min}, h_p) = \begin{cases} p^{-2l}(1-p^{-2})^{-1} & \text{if } z \equiv 1 \mod p^l \mathbb{Z}_p, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. The first assertion is stated in [7, Theorem 2.4.1]. The second assertion is trivial. He assumed that the residual characteristic is not 2 in the paper. However, since [7, Lemma 2.1.1] can be applied for $\mathbb{Q}_2$, one can easily compute the above integral. □

4.8. Explicit calculations for $I_3(f)$

Lemma 4.13. *Let $c(z\delta_1)$ denote the constant given in Lemma 4.6. Then, we have*

$$I_3(f) = \sum_{z \in Z(\mathbb{Q})} \operatorname{vol}(G_{\delta_1}(\mathbb{Q}) \backslash G_{\delta_1}(\mathbb{A})^1) \times c(z\delta_1) \times \chi_\xi(z)^{-1} \times (l_1 - l_2)(l_1 + l_2) \times J_G(z\delta_1, h).$$

Proof. This lemma follows from Lemma 4.6 and (4.16). □

Lemma 4.14. *Assume that $J_G(z\delta_1, h)$ is defined as (4.15) and h_{a_1,a_2,a_3} is the characteristic function of $\mathbf{K}_p\mathrm{diag}(p^{-a_1}, p^{-a_2}, p^{a_1-a_3}, p^{a_2-a_3})\mathbf{K}_p$ on $G(\mathbb{Q}_p)$ $(a_3 \geqslant a_1 \geqslant a_2 \geqslant 0)$. If a_3 is odd, then we have $J_G(z\delta_1, h_{a_1,a_2,a_3}) = 0$. If a_3 is even, then we may set $a_3 = 2m$ and assume $m \geqslant a_1 \geqslant a_2$ by the action of W_0^G, and we get*

$$J_G(z\delta_1, h_{a_1,a_2,a_3}) = \begin{cases} 1 & \text{if } a_3 = 2m,\ a_1 = a_2 = m, \text{ and } |z|_p = p^m, \\ p^{a_3-a_1-a_2}(1-p^{-2}) & \text{if } a_3 = 2m,\ m > a_1 = a_2, \text{ and } |z|_p = p^m, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. This lemma can be proved by [7, Lemma 2.1.1]. □

4.9. Estimations for $I_5(f)$

From now on, we assume that $\gamma \in G(\mathbb{Q})$ is not semisimple and $\gamma \notin Z_G(\mathbb{Q})\{u_{\min}\}_G$. By Lemmas 4.5, we may also assume that the γ_s is $\mathbb{R}$-elliptic in G.

Lemma 4.15. *If the centralizer of γ is not isomorphic to G_1, G_2, and G_3, then we get $J_G(\gamma, f_\xi) = 0$.*

Proof. For unipotent elements γ, it was proved in Lemmas 4.8 and 4.9. All mixed elements in $G(\mathbb{Q})$ are classified in [29, § 5.3]. Using the classification and the limit formula for $SL_2(\mathbb{R})$ (cf. [35]) one can show $J_G(\gamma, f_\xi) = 0$. □

Lemma 4.16. *Let $z \in Z_G(\mathbb{Q})$. Then, we get*

$$J_G(z\delta_1 u_1(1,0), f_\xi) = J_G(z\delta_1 u_1(0,1), f_\xi) = 0.$$

There exists a constant $c(z\delta_1 u_1)$ such that

$$J_G(z\delta_1 u_1(1,1), f_\xi) = -J_G(z\delta_1 u_1(1,-1), f_\xi) = c(z\delta_1 u_1) \times \chi_\xi(z)^{-1} \times \{(-1)^{l_2} - (-1)^{l_1}\}.$$

Proof. This can be proved by the limit formula for $SL_2(\mathbb{R})$ (cf. [35]). □

Lemma 4.17. *Let $z \in Z_G(\mathbb{Q})$. The contribution of $z\delta_1 u_1(1,0)$ and $z\delta_1 u_1(0,1)$ to $I_5(f)$ is zero. For any positive real number ε, there exists a constant $c(z\delta_1 u_1, \varepsilon) > 0$ such that the contribution of (G, S)-equivalence classes of elements $z\delta_1 u_1(1,\alpha)$ $(\alpha \in \mathbb{Q}^\times/(\mathbb{Q}_S^\times)^2 \cap \mathbb{Q}^\times)$ is bounded by*

$$c(z\delta_1 u_1, \varepsilon) \times \chi_\xi(z)^{-1} \times J_{M_0}^{M_0}(z\delta_1, |h_{P_0}|) \times \prod_{p \in S_0} p^\varepsilon \times \prod_{p \in S_0} (1-p^{-1})^{-2}.$$

Proof. The first assertion obviously follows from Lemma 4.16. By (4.18) and Lemma 4.16, the contribution equals

$$\sum_{\alpha \in \mathbb{Q}^\times/(\mathbb{Q}_S^\times)^2 \cap \mathbb{Q}^\times} a^G(S, z\delta_1(1,\alpha))\, J_G(z\delta_1 u_1(1,\alpha), f_\xi)\, J_G(z\delta_1 u_1(1,\alpha), h)$$
$$= 2c(z\delta_1 u_1)\{(-1)^{l_2} - (-1)^{l_1}\} \sum_\chi L^S(1,\chi)^2 \sum_{\alpha \in \mathbb{Q}^\times/(\mathbb{Q}_{S_0}^\times)^2 \cap \mathbb{Q}^\times} \chi_{S_0}(\alpha)\, J_G(z\delta_1 u_1(1,\alpha), h)$$

where $\chi = \prod_v \chi_v$ runs over all quadratic characters such that $\chi_\infty = \mathrm{sgn}$ and χ_v is unramified for any $v \notin S$. By calculating $G_{z\delta_1 u_1(1,\alpha)}$, we find that $G_{z\delta_1 u_1(1,\alpha)}$ is contained in $Z_G N_0$ and we may set

$$x = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a^{-1} & 0 & 0 \\ 0 & 0 & b & 0 \\ 0 & 0 & 0 & ab \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & y \\ 0 & 1 & y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & r & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -r & 1 \end{pmatrix}$$

for each element x in $G_{z\delta_1 u_1(1,\alpha)}(\mathbb{Q}_{S_0})\backslash P_0(\mathbb{Q}_{S_0})$. Then, we get

$$x^{-1} z\delta_1 u_1(1,\alpha) x = z\delta_1 \begin{pmatrix} \begin{matrix} 1 & 2r \\ 0 & 1 \end{matrix} & \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix}\begin{pmatrix} b & 2y \\ 2y & \alpha a^2 b \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -r & 1 \end{pmatrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \begin{matrix} 1 & 0 \\ -2r & 1 \end{matrix} \end{pmatrix}.$$

From this, we deduce

$$\begin{aligned} \Bigg| \sum_{\alpha \in \mathbb{Q}^\times/(\mathbb{Q}_{S_0}^\times)^2 \cap \mathbb{Q}^\times} \chi_{S_0}(\alpha)\, J_G(z\delta_1 u_1(1,\alpha), h) \Bigg| &\leqslant \sum_{\alpha \in \mathbb{Q}^\times/(\mathbb{Q}_{S_0}^\times)^2 \cap \mathbb{Q}^\times} J_G(z\delta_1 u_1(1,\alpha), |h|) \\ &= 4 \times \prod_{p \in S_0} (1-p^{-1})^{-2} \times \int_{\mathbf{K}_{S_0}} \int_{N_0(\mathbb{Q}_{S_0})} \left| h(k^{-1} z\delta_1 n\, k) \right| dn\, dk \\ &= 4 \times \prod_{p \in S_0} (1-p^{-1})^{-2} \times J_{M_0}^{M_0}(z\delta_1, |h_{P_0}|). \end{aligned}$$

Hence, the inequality follows from this estimation and Lemma 4.10. □

Lemma 4.18. *Let δ_2 be a semisimple element in $G(\mathbb{Q})$ such that $G_{\delta_2} \cong G_2$ for a quadratic extension $E/\mathbb{Q}$. Each unipotent element $u_2(\alpha)$ in $G_2(\mathbb{Q})$ is identified with an element in $G_{\delta_2}(\mathbb{Q}) \subset G(\mathbb{Q})$. The $G(\mathbb{Q})$-conjugacy class of $\delta_2 u_2(\alpha)$ has an intersection with $P_1(\mathbb{Q})$ and we may assume that δ_2 belongs to $M_1(\mathbb{Q})$ as a representative element of the conjugacy class. For any positive real number ε, there exists a constant $c(\delta_2 u_2, \varepsilon) > 0$ such that the contribution of (G, S)-equivalence classes of elements $\delta_2 u_2(\alpha)$ ($\alpha \in E^\times/((E_{S_E}^\times)^2 \mathbb{Q}_S^\times) \cap E^\times$) is bounded by*

$$c(\delta_2 u_2, \varepsilon) \times \chi_\xi(\delta_2)^{-1} \times J_{M_1}^{M_1}(\delta_2, |h_{P_1}|) \times \prod_{p \in S_0} p^\varepsilon \times \prod_{p \in S_0} (1-p^{-1})^{-2}.$$

Proof. By [29, § 5.3], we may choose the semisimple element δ_2 as

$$\delta_2 = z \begin{pmatrix} h_\beta & O_2 \\ O_2 & {}^t h_\beta \end{pmatrix}, \quad h_\beta = \begin{pmatrix} 0 & 1 \\ \beta & 0 \end{pmatrix}$$

for an element β in $\mathbb{Q}^\times - (\mathbb{Q}^\times)^2$ and an element z in $Z_G(\mathbb{Q})$. In particular, we have $E = \mathbb{Q}(\sqrt{\beta})$. If β is negative, then $\nu(\delta_2)$ is also negative. Hence, by Lemma 4.1, the contribution vanishes if β is negative. We may assume that β is positive, i.e., E is a real quadratic field. Then, there exists an element z' in $Z_G(\mathbb{R})$ such that δ_2 is $G(\mathbb{R})$-conjugate to $z'\delta_1$. Thus, it follows from (4.19) and Lemma 4.16 that the contribution equals

$$\sum_{\alpha\in E^\times/((E^\times_{S_E})^2\mathbb{Q}^\times_S)\cap E^\times)} a^G(S,\delta_2u_2(\alpha))\,J_G(\delta_2u_2(\alpha),f_\xi)\,J_G(\delta_2u_2(\alpha),h)$$
$$=2c(z\delta_1u_1)\{(-1)^{l_2}-(-1)^{l_1}\}\sum_\chi L^S_E(1,\chi)\sum_{\alpha\in E^\times/((E^\times_{(S_0)_E})^2\mathbb{Q}^\times_{S_0})\cap E^\times)}\chi_{(S_0)_E}(\alpha)\,J_G(\delta_2u_2(\alpha),h)$$

where $\chi=\prod_w\chi_w$ runs over all non-trivial quadratic characters such that $\chi|_{\mathbb{A}^1_{\mathbb{Q}}}=1$, $\chi_w=\mathrm{sgn}$ for any $w|\infty$, and χ_w is unramified for any $w\notin S_E$. Therefore, we finish the proof by using Lemma 4.10 and an argument similar to the proof of Lemma 4.17. □

Lemma 4.19. *Let δ_3 be a semisimple element in $G(\mathbb{Q})$ such that $G_{\delta_3}\cong G_3$ for a quadratic extension $E/\mathbb{Q}$. Each unipotent element $u_3(\alpha)$ in $G_3(\mathbb{Q})$ is identified with an element in $G_{\delta_3}(\mathbb{Q})\subset G(\mathbb{Q})$. The $G(\mathbb{Q})$-conjugacy class of $\delta_3u_3(\alpha)$ has an intersection with $P_2(\mathbb{Q})$ and we may assume that δ_3 belongs to $M_2(\mathbb{Q})$ as a representative element of the conjugacy class. There exists a positive constant $c(\delta_3u_3)$ such that the contribution of (G,S)-equivalence classes of elements $\delta_3u_3(\alpha)$ ($\alpha\in\mathbb{Q}^\times/(N_{E/\mathbb{Q}}(E^\times_{S_E})\cap\mathbb{Q}^\times)$ is bounded by*

$$c(\delta_3u_3)\times\chi_\xi(\delta_3)^{-1}\times J^{M_2}_{M_2}(\delta_3,|h_{P_2}|)\times\prod_{p\in S_0}(1-p^{-1})^{-1}.$$

Proof. By [29, § 5.3], the semisimple element δ_3 can be written as

$$\delta_3=z\begin{pmatrix}1&0&0&0\\0&x&0&y\\0&0&1&0\\0&\beta y&0&x\end{pmatrix}$$

for an element z in $Z_G(\mathbb{Q})$, an element β in $\mathbb{Q}^\times-(\mathbb{Q}^\times)^2$, and elements x, y in $\mathbb{Q}$ such that $x^2-\beta y^2=1$. Note that $E=\mathbb{Q}(\sqrt{\beta})$. If β is positive, then δ_3 is not $\mathbb{R}$-elliptic in G. Hence, we have $J_G(\delta_3u_3(\alpha))=0$ by Lemma 4.5. So, we assume that β is negative, i.e., E is an imaginary quadratic field. Since $u_3(\alpha)$ is $G(\mathbb{R})$-conjugate to $u_3(1)$ or $u_3(-1)$, using the limit formula for $SL_2(\mathbb{R})$, we have

$$J_G(\delta_3u_3(1),f_\xi)=-J_G(\delta_3u_3(-1),f_\xi)=c_3'\times(-e^{il_2\theta}+e^{il_1\theta})$$

for a positive constant c_3'. Hence, by (4.20), the contribution is equal to

$$\sum_{\alpha\in\mathbb{Q}^\times/(N_{E/\mathbb{Q}}(E^\times_{S_E})\cap\mathbb{Q}^\times)} a^G(S,\delta_3u_3(\alpha))\,J_G(\delta_3u_3(\alpha),f_\xi)\,J_G(\delta_3u_3(\alpha),h)$$
$$=2c_3'\times(-e^{il_2\theta}+e^{il_1\theta})L^S(1,\chi)\sum_{\alpha\in\mathbb{Q}^\times/(N_{E/\mathbb{Q}}(E^\times_{(S_0)_E})\cap\mathbb{Q}^\times)}\chi_{S_0}(\alpha)\,J_G(\delta_3u_3(\alpha),h)$$

where χ denotes the non-trivial quadratic character on $\mathbb{Q}^\times\backslash\mathbb{A}^1$ corresponding to E. We can derive this lemma from this equality and an argument similar to the proof of Lemma 4.17. □

5. An estimation of the geometric side

Let us recall our setting. Let S' be a (non-empty) finite set of finite primes and S_0 be a finite set of finite primes containing S'. We choose S_0 sufficiently large so that Arthur's geometric expansion works. Put $S = S_0 \cup \{\infty\}$. For a positive integer N whose prime divisors do not belong to S', put $f_{K(N)} = \mathrm{char}_{K(N)}$. When we study the level aspect, we always choose the level N for the fixed S' as above. We denote by $\underline{k} = (k_1, k_2)$ the highest weight of $\xi = \xi_{\underline{k}}$ for $k_1 \geqslant k_2 \geqslant 3$ as in § 3.2. Let us put

$$f_{S',\alpha} := [G(\mathbb{Z}_{S'})\alpha G(\mathbb{Z}_{S'})] \in C_c^\infty(G(\mathbb{Q}_{S'}))$$

for $\alpha \in T(\mathbb{Q})$. Note that usually we would choose α from $T(\mathbb{Q}_{S'})$, but due to the comparison with spectral side, intentionally we choose $\mathbb{Q}$-rational elements and clearly this never changes anything. If $\nu(\alpha) \in (\mathbb{Q}^\times)^2$, then such an α can be uniquely written as

$$\alpha = (z_\alpha \cdot E_4)k_\alpha, \;\; z_\alpha \cdot E_4 \in Z_G^+(\mathbb{Q}), \;\; k_\alpha \in (T \cap Sp_4)(\mathbb{Q}).$$

Put

$$z'_\alpha = \begin{cases} z_\alpha \cdot E_4 & \text{if } \nu(\alpha) \in (\mathbb{Q}^\times)^2 \\ E_4 & \text{otherwise.} \end{cases} \tag{5.1}$$

To prove Theorem 1.1 we have only to check it for

$$f = f_\xi h, \quad h = f_{z'_\alpha K(N)} f_{S',\alpha}\Bigg(\bigotimes_{p \in S\setminus(S' \cup \{v|N\infty\})} \mathrm{char}_{K_p}\Bigg) \in C_c^\infty(G(\mathbb{Q}_{S_0}))$$

where $f_{z'_\alpha K(N)}$ stands for the characteristic function of $z'_\alpha K(N)$.

Since $I_7(f) = 0$ (see Lemma 4.4), it is unnecessary to consider it. Fix a test function f as above. Let γ be a (M, S)-equivalence class whose contribution to $I_{\mathrm{geom}}(f)$ is not zero. By (4.12) we see that

$$\nu(\alpha) = \nu(\gamma).$$

Furthermore its semisimple part γ_s is $\mathbb{R}$-elliptic in M (cf. Lemmas 4.4 and 4.5). If we set $\gamma_s = z_\gamma \gamma_1$ where $z_\gamma \in \mathbb{R}^\times \simeq Z_G(\mathbb{R})$ and $\gamma_1 \in M(\mathbb{R})$ with $\nu(\gamma_1) \in \{\pm 1\}$. Therefore, we have

$$\begin{aligned}\chi_\xi(z_\gamma) &= z_\gamma^{k_1+k_2} = \mathrm{sgn}(z_\gamma^{k_1+k_2})|\nu(z_\gamma)|_{\mathbb{R}}^{(k_1+k_2)/2} \\ &= \mathrm{sgn}(z_\gamma^{k_1+k_2})|z_\gamma|_{\mathbb{R}}^{k_1+k_2} = \mathrm{sgn}(z_\gamma^{k_1+k_2})|\nu(\alpha)|_{S'}^{-(k_1+k_2)/2}.\end{aligned}$$

This observation will be implicitly used in the proof of Proposition 5.1 below.

Each orbital integral $J_M^M(\gamma, h_P)$ is bounded by that of $PGSp_4(\mathbb{Q}_S)$ using the projection $G \to PGSp_4$. Therefore, we can use the same arguments as in [62, 63] to estimate it for any semisimple element γ.

Proposition 5.1. *Fix a finite set* S' *and a function* $f_{S',\alpha} = [G(\mathbb{Z}_{S'})\alpha G(\mathbb{Z}_{S'})] \in C_c^\infty(G(\mathbb{Q}_{S'}))$. *Then, we have*

$$I_2(f) \times \mathrm{vol}(K(N))^{-1} \times |\nu(\alpha)|_{S'}^{-(k_1+k_2)/2} = O((k_1 - k_2 + 1)(k_1 + k_2 - 3)\varphi(N)N^8),$$

$$I_3(f)\times \mathrm{vol}(K(N))^{-1}\times|\nu(\alpha)|_{S'}^{-(k_1+k_2)/2}=O((k_1-1)(k_2-2)),$$
$$\{I_4(f)+I_5(f)\}\times \mathrm{vol}(K(N))^{-1}\times|\nu(\alpha)|_{S'}^{-(k_1+k_2)/2}=O(k_1+k_2-3),$$
$$I_6(f)\times \mathrm{vol}(K(N))^{-1}\times|\nu(\alpha)|_{S'}^{-(k_1+k_2)/2}=O((k_1+k_2-3)\varphi(N)N^7)$$

for any weight (k_1,k_2) $(k_1\geqslant k_2\geqslant 3)$ *and any level* $N>0$ *prime to* $\prod_{p\in S'}p$. *Note that* $\mathrm{vol}(K(N))^{-1}=[\Gamma(1):\Gamma(N)]=N^{10}\prod_{p|N}(1-p^{-2})(1-p^{-4})$.

Proof. If $zu_{\min}$ contributes to $I_{\mathrm{geom}}(f)$, then we have $z=z_\alpha$. Hence, only $z_\alpha u_{\min}$ contributes to $I_2(f)$ and the estimation for $I_2(f)$ obviously follows from Lemmas 4.11 and 4.12.

By Lemmas 4.5, 4.8, 4.9 and [16, Lemma 5] (also see [62, Lemma 4.3 and line 12 in p. 100]), there exists a sufficiently large natural number $\tilde{N}_0$ such that, if $N>\tilde{N}_0$, then we have $I_3(f)=I_4(f)=I_5(f)=0$ and only the central elements contribute to $I_6(f)$. When N moves between 1 and $\tilde{N}_0$, there are finitely many $M(\mathbb{Q})$-conjugacy classes ($M\in\mathcal{L}$) which contribute to $I_3(f)$, $I_4(f)$ and $I_6(f)$, and finitely many (G,S)-equivalence classes which contribute to $I_5(f)$ (cf. [62, Proof of Theorem 4.11]). Thus, we get the estimation for $I_3(f)$ by Lemma 4.13 and the estimation for $I_4(f)+I_5(f)$ by Lemmas 4.5, 4.8, 4.9. By these facts and Lemma 4.4, the remaining work is to find the growth of $I_6(f)$ with respect to N. For each proper Levi subgroup M in $\mathcal{L}$ and each element $z\in Z_G(\mathbb{Q})$, we have

$$|J_M^G(z,h_P)|=(\text{constant})\times\int_{\prod_{v|N}N_P(\mathbb{Q}_v)}f_{z'_\alpha K(N)}(n)\,dn\leqslant(\text{constant})\times N^{-3},\tag{5.2}$$

since $K(N)$ is a normal subgroup in $\prod_{v|N}\mathbf{K}_v$. Hence, this proposition is proved. □

We set

$$p_{S'}=\prod_{p\in S'}p,\qquad H^{\mathrm{ur}}(G(\mathbb{Q}_{S'}))^\kappa=\bigoplus_{p\in S'}H^{\mathrm{ur}}(G(\mathbb{Q}_p))^\kappa.$$

Proposition 5.2. (Level aspect) *There exist positive constants* a, b, *and* N_0 *such that*

$$I_2(f)\times \mathrm{vol}(K(N))^{-1}\times|\nu(\alpha)|_{S'}^{-(k_1+k_2)/2}=O(p_{S'}^{\kappa}(k_1-k_2+1)(k_1+k_2-3)\varphi(N)N^8)$$
$$I_6(f)\times \mathrm{vol}(K(N))^{-1}\times|\nu(\alpha)|_{S'}^{-(k_1+k_2)/2}=O(p_{S'}^{a\kappa+b}(k_1+k_2-3)\varphi(N)N^7)$$
$$I_3(f)=I_4(f)=I_5(f)=I_7(f)=0,$$

for any (k_1,k_2), $N>0$, $\kappa\geqslant 1$, S', *and* $f_{S',\alpha}$, *which satisfy the conditions* $k_1\geqslant k_2\geqslant 3$, $f_{S',\alpha}\in H^{\mathrm{ur}}(G(\mathbb{Q}_{S'}))^\kappa$, N *is prime to* $\prod_{p\in S'}p$, *and* $N\geqslant N_0\prod_{p\in S'}p^{10\kappa}$.

Proof. According to [63, §8], a faithful algebraic representation $\Xi:PGSp_4\to GL_m$ is needed. Here we take the adjoint representation $\mathrm{Ad}:PGSp_4\to GL(\mathrm{Lie}(Sp_4))$, i.e., $\Xi=\mathrm{Ad}:PGSp_4\to GL_{10}$. By [63, Lemma 8.4], there exists a natural number N_0 such that, if $N\geqslant N_0\prod_{p\in S'}p^{10\kappa}$, then we have $I_3(f)=I_4(f)=I_5(f)=0$, and the contributions of the non-central elements to $I_6(f)$ are zero. Hence, the estimation for $I_6(f)$ can be proved by [63, Lemma 2.14] and (5.2). The estimation for $I_2(f)$ follows from Lemmas 4.11 and 4.12. □

Proposition 5.3. *(Weight aspect) Fix a level $N > 0$. There exist positive constants a' and b' such that*

$$I_2(f) \times \mathrm{vol}(K(N))^{-1} \times |\nu(\alpha)|_{S'}^{-(k_1+k_2)/2} = O(p_{S'}^{\kappa}(k_1-k_2+1)(k_1+k_2-3))$$
$$I_3(f) \times \mathrm{vol}(K(N))^{-1} \times |\nu(\alpha)|_{S'}^{-(k_1+k_2)/2} = O(p_{S'}^{\kappa}(k_1-1)(k_2-2)),$$
$$\{I_4(f)+I_6(f)\} \times \mathrm{vol}(K(N))^{-1} \times |\nu(\alpha)|_{S'}^{-(k_1+k_2)/2} = O(p_{S'}^{a'\kappa+b'}(k_1+k_2-3)),$$
$$I_5(f) \times \mathrm{vol}(K(N))^{-1} \times |\nu(\alpha)|_{S'}^{-(k_1+k_2)/2} = O(p_{S'}^{a'\kappa+b'}(k_1+k_2-3))$$

for any (k_1, k_2), $\kappa \geqslant 1$, S', and $f_{S',\alpha}$, which satisfy the conditions $k_1 \geqslant k_2 \geqslant 3$, $f_{S',\alpha} \in H^{\mathrm{ur}}(G(\mathbb{Q}_{S'}))^{\kappa}$, and N is prime to $\prod_{p\in S'} p$.

Proof. For each M in $\mathcal{L}$, let Y_M denote the set of $M(\mathbb{A})$-conjugacy classes of semisimple $\mathbb{R}$-elliptic elements of $M(\mathbb{Q})$ whose contributions to $I_{\mathrm{geom}}(f)$ are non-zero. By [63, Proposition 8.7], one finds $|Y_M| = O(p_{S'}^{a_1\kappa+b_1})$ where a_1 and b_1 are certain positive numbers. Furthermore, there exist positive numbers a_2 and b_2 such that $a^M(S,\gamma)\, J_M(\gamma, h_P) = O(p_{S'}^{a_2\kappa+b_2})$ holds for each semisimple $\mathbb{R}$-elliptic element γ in Y_M. This fact is due to [63, Proof of Theorem 9.19]. Hence, the estimation for $I_4(f)$ and $I_6(f)$ follows from these results of [63] and Lemmas 4.4 and 4.5.

Here $I_3(f)$ is the total contribution of $z\delta_1$ ($z \in Z_G(\mathbb{Q})$) and the center z must satisfy $\nu(\alpha) = z^2$. Hence, only the $G(\mathbb{Q})$-conjugacy class of $z_\alpha\delta_1$ can contribute to $I_3(f)$. Therefore, the estimation for $I_3(f)$ is deduced from Lemmas 4.13 and 4.14. By the same argument, only the $G(\mathbb{Q})$-conjugacy class of $z_\alpha u_{\mathrm{min}}$ contributes to $I_2(f)$. Hence, the estimation for $I_2(f)$ follows from Lemmas 4.11 and 4.12.

Next we shall consider the term $I_5(f)$. By Lemma 4.15, it is enough to treat the (G, S)-conjugacy class γ such that G_{γ_s} is isomorphic to G_1, G_2, or G_3. By Lemmas 4.17, 4.18, 4.19, each the contribution is bounded by the product of a constant, $\prod_{p\in S' \text{ or } p|N} p^{\epsilon}$ and a semisimple $\mathbb{R}$-elliptic orbital integral of h_P for a proper standard parabolic subgroup P. Therefore, we can reduce the estimation for $I_5(f)$ to the semisimple case as above. Hence, the proof is completed. □

Finally we treat $I_1(f)$. Suppose $\gamma = z_\gamma \cdot I_4 \in Z_G(\mathbb{Q})$. By (4.12) it satisfies

$$\gamma \in \mathrm{Supp}(h) = (z'_\alpha K(N)) \times G(\mathbb{Z}_{S'})\alpha G(\mathbb{Z}_{S'}) \times \prod_{p\in S\setminus S'\cup\{v|N\infty\}} K_p.$$

Further γ can happen exactly when $\nu(\alpha) \in (\mathbb{Q}_{S'}^{\times})^2$. In this case, it follows that

$$\gamma = \begin{cases} z_\alpha \cdot E_4 & \text{if } N \geqslant 3 \\ \pm z_\alpha \cdot E_4 & \text{otherwise,} \end{cases}$$

since S is sufficiently large (see (5.1) for $z'_\alpha \in Z_G^{+}(\mathbb{Q})$). Furthermore it follows from (4.12) again that

$$\nu(\gamma) = |z_\gamma|_{\mathbb{R}}^2 = |\nu(\alpha)|_{S'}^{-1}.$$

We define

$$\varepsilon(\alpha) = \begin{cases} 2 & \text{if } N = 1, 2 \text{ and } \nu(\alpha) \in (\mathbb{Q}_{S'}^{\times})^2 \\ 1 & \text{otherwise} \end{cases}$$

which is nothing but $\varepsilon_{S'}(\alpha)$ in (3.13) for $U = K(N)$. By Plancherel formula and the limit formula we have

$$\begin{aligned}(-1)^{q(G(\mathbb{R}))} I_1(f) &= \overline{\mu}(G(\mathbb{Q})A_{G,\infty}\backslash G(\mathbb{A})) \dim \xi \sum_{\gamma} \gamma^{-(k_1+k_2)} \cdot f_{S',\alpha}(\gamma) \\ &= \overline{\mu}(G(\mathbb{Q})A_{G,\infty}\backslash G(\mathbb{A})) \dim \xi \sum_{\gamma} |\gamma|_{\mathbb{R}}^{-(k_1+k_2)} \cdot f_{S',\alpha}(|\gamma|_{\mathbb{R}}) \quad \text{(by (2.11))} \\ &= \overline{\mu}(G(\mathbb{Q})A_{G,\infty}\backslash G(\mathbb{A})) \cdot \dim \xi \sum_{\gamma} \widehat{\mu}^{\mathrm{pl}}_{S'}(\widehat{f}^{\gamma}_{S',\alpha}) \cdot |\nu(\alpha)|_{S'}^{(k_1+k_2)/2} \\ &= \varepsilon(\alpha) \cdot \overline{\mu}(G(\mathbb{Q})A_{G,\infty}\backslash G(\mathbb{A})) \cdot \dim \xi \cdot \widehat{\mu}^{\mathrm{pl}}_{S'}(\widehat{f}^{z'_\alpha}_{S',\alpha}) \cdot |\nu(z'_\alpha)|_{S'}^{(k_1+k_2)/2} \qquad (5.3)\end{aligned}$$

where $f^{\gamma}_{S',\alpha}$ (respectively $f^{z'_\alpha}_{S',\alpha}$) is the translation by $|\gamma|_{\mathbb{R}}$ (respectively $|z'_\alpha|_{\mathbb{R}}$) of $f_{S',\alpha}$ and we used

$$\begin{aligned}\widehat{\mu}^{\mathrm{pl}}_{S'}(\widehat{f}^{\gamma}_{S',\alpha}) &= \int_{\widehat{G(\mathbb{Q}_{S'})}} \widehat{f}^{\gamma}_{S',\alpha}(\pi)\widehat{\mu}^{\mathrm{pl}}_{S'}(\pi) = \int_{\widehat{G(\mathbb{Q}_{S'})}} \mathrm{tr}(\pi(f^{\gamma}_{S',\alpha}))\widehat{\mu}^{\mathrm{pl}}_{S'}(\pi) \\ &= f^{\gamma}_{S',\alpha}(1) = f_{S',\alpha}(|\gamma|_{\mathbb{R}}) = f_{S',\alpha}(|z'_\alpha|_{\mathbb{R}}) = f^{z'_\alpha}_{S',\alpha}(1) = \widehat{\mu}^{\mathrm{pl}}_{S'}(\widehat{f}^{z'_\alpha}_{S',\alpha}).\end{aligned}$$

Hence we have

$$\frac{|\nu(\alpha)|_{S'}^{-(k_1+k_2)/2} I_1(f)}{\varepsilon(\alpha) \cdot \overline{\mu}(G(\mathbb{Q})A_{G,\infty}\backslash G(\mathbb{A})) \cdot \dim \xi} = (-1)^{q(G(\mathbb{R}))}\widehat{\mu}^{\mathrm{pl}}_{S'}(\widehat{f}^{z'_\alpha}_{S',\alpha}) = -\widehat{\mu}^{\mathrm{pl}}_{S'}(\widehat{f}^{z'_\alpha}_{S',\alpha}). \qquad (5.4)$$

Note that if $\nu(\alpha) \notin (\mathbb{Q}^{\times}_{S'})^2$, then both sides of (5.4) are zero giving the trivial identity.

Summing up we have obtained the following results which follow from Propositions 5.2, 5.3, and (5.4).

Theorem 5.4. *(Level aspect) Keep the notation as in Proposition 5.2. Then*

$$\frac{|\nu(\alpha)|_{S'}^{-(k_1+k_2)/2} I_{\mathrm{geom}}(f)}{\varepsilon(\alpha) \cdot \overline{\mu}(G(\mathbb{Q})A_{G,\infty}\backslash G(\mathbb{A})) \cdot \dim \xi} = -\widehat{\mu}^{\mathrm{pl}}_{S'}(\widehat{f}^{z'_\alpha}_{S',\alpha}) + A + O(p_{S'}^{a\kappa+b}\varphi(N)N^{-3}),$$

where $A = O(p^{\kappa}_{S'}\varphi(N)N^{-2})$, $(N, p_{S'}) = 1$, *and* $N \geqslant N_0 p_{S'}^{10\kappa}$.

Theorem 5.5. *(Weight aspect) Keep the notation as in Proposition 5.3. Then*

$$\begin{aligned}\frac{|\nu(\alpha)|_{S'}^{-(k_1+k_2)/2} I_{\mathrm{geom}}(f)}{\varepsilon(\alpha) \cdot \overline{\mu}(G(\mathbb{Q})A_{G,\infty}\backslash G(\mathbb{A})) \cdot \dim \xi} &= -\widehat{\mu}^{\mathrm{pl}}_{S'}(\widehat{f}^{z'_\alpha}_{S',\alpha}) + B_1 + B_2 \\ &\quad + O\left(\frac{p_{S'}^{a'\kappa+b'}}{(k_1-k_2+1)(k_1-1)(k_2-2)}\right),\end{aligned}$$

where $B_1 = O(\frac{p^{\kappa}_{S'}}{(k_1-1)(k_2-2)})$ *and* $B_2 = O(\frac{p^{\kappa}_{S'}}{(k_1-k_2+1)(k_1+k_2-3)})$, *and* $(N, p_{S'}) = 1$.

Remark 5.6. Shin's condition in the weight aspect in [62] becomes:

$$k_1 - k_2 \longrightarrow \infty, \ \ k_2 \longrightarrow \infty.$$

Theorem 5.5 can also treat the case where $k_1 - k_2$ is constant while k_2 tends to infinity. This is a new direction of the weight aspect which has not been studied.

Remark 5.7. In Theorem 5.4, if we restrict automorphic forms to those with a fixed central character $\chi \in (\widehat{\mathbb{Z}/N\mathbb{Z}})^\times$, then we have a better result without $\varphi(N)$ on the right hand side. Accordingly the dimension of Siegel cusp forms with a fixed central character is smaller by a factor of $\varphi(N)$. (See Proposition 2.2.)

6. Proof of the main theorem

In this section we give a proof of Theorem 1.1 in the introduction.

Fix a prime $p \nmid N$. Let $U = K(N)$ and $S' = \{p\}$. For any $\kappa \in \mathbb{Z}_{\geqslant 0}$, let f_p be the characteristic function of $K_p \mathrm{diag}(p^{-a_1}, p^{-a_2}, p^{a_1-\kappa}, p^{a_2-\kappa})K_p$, $0 \leqslant a_2 \leqslant a_2 \leqslant \kappa$.

If $k_1 \geqslant k_2 \geqslant 4$, it is immediate by Proposition 4.3, Theorems 5.4 and 5.5.

If $k_2 = 3$, we need to estimate the traces of Hecke operators on the residual spectrum and a part of cuspidal space related to non-tempered representations ω_{l_1}.

If $k_1 > k_2 = 3$, by Propositions 3.1 and 3.7,

$$\sum_{*\in\{\mathrm{cusp},\mathrm{res}\}} \widehat{\mu}_{K(N),\xi_{\underline{k}},\omega_{l_1},*}(\widehat{f_p}) = (\dim \xi_{\underline{k}})^{-1}(O(p^{\kappa/2}N^{-4+\varepsilon}) + O(p^{\kappa/2}k_1 N^{-6})).$$

If $k_1 = k_2 = 3$, by (3.6), Propositions 3.1 and 3.7,

$$\begin{aligned} &\widehat{\mu}_{K(N),\xi_{\underline{k}},\mathbf{1}}(\widehat{f_p}) + \sum_{*\in\{\mathrm{cusp},\mathrm{res}\}} \widehat{\mu}_{K(N),\xi_{\underline{k}},\omega_{l_1},*}(\widehat{f_p}) \\ &= (\dim \xi_{\underline{k}})^{-1}(O(p^{\kappa/2}N^{-4+\varepsilon}) + O(p^{\kappa/2}k_1 N^{-6}) + O(p^{3\kappa/2}N^{-10})). \end{aligned}$$

Notice that the above error terms are subsumed in the error terms in Theorems 5.4 and 5.5. Therefore, Theorem 1.1 follows from Proposition 4.3, Theorems 5.4 and 5.5. This completes a proof of the main theorem.

7. Applications

7.1. The classical formulation

In this subsection we reformulate Theorem 1.1 in terms of classical Siegel modular forms. As we see later, this will be used in the study of Hecke fields and low-lying zeros.

Let $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \longrightarrow \mathbb{C}^\times$ be a Dirichlet character. Fix a square root $\chi(p)^{1/2}$ for each fixed $p \nmid N$. We write $\chi(p)^{i/2} = (\chi(p)^{1/2})^i$. Put $V_{\underline{k},N} = S_{\underline{k}}(\Gamma(N))$ or $S_{\underline{k}}(\Gamma(N),\chi)$ for $k_1 \geqslant k_2 \geqslant 3$. Let us recall the Hecke operators T_m or $T(p^i)$ on $V_{\underline{k},N}$ for $m \in \Delta_n(N)$ and $p \nmid N$. We normalize them as $T'_m = T_m/\nu(m)^{(k_1+k_2-3)/2}$ on $S_{\underline{k}}(\Gamma(N))$ and $T'(p^i) = T(p^i)/(p^{(k_1+k_2-3)/2}\chi(p)^{1/2})^i$ on $S_{\underline{k}}(\Gamma(N),\chi)$. Put $d_{\underline{k},N} = \dim_{\mathbb{C}} V_{\underline{k},N}$. Clearly if $m = p^\kappa E_4$, then

$$\frac{1}{d_{\underline{k},N}} \mathrm{tr}(p^{3\kappa} T'_{p^\kappa E_4} | V_{\underline{k},N}) = 1.$$

Then by Theorem 1.1, we have the following theorem:

Theorem 7.1. *There exist absolute constants a_1, b_1, a_1', b_1' such that for a prime $p \nmid N$ and $m = \mathrm{diag}(p^{u_1}, p^{u_2}, p^{-u_1+\kappa}, p^{-u_2+\kappa})$, $u_1, u_2, \kappa \in \mathbb{Z}$ satisfying $0 \leqslant u_2 \leqslant u_1 \leqslant \kappa$ and $m \notin Z_G(\mathbb{Q})$,*

(1) *(Level aspect)*

$$\frac{1}{d_{\underline{k},N}}\mathrm{tr}(T_m'|V_{\underline{k},N}) = A + O(p^{a_1\kappa+b_1}N^{-3}), \quad A = O(p^{-\kappa/2}N^{-2}), N \gg p^{10\kappa}.$$

(2) *(Weight aspect)*

$$\frac{1}{d_{\underline{k},N}}\mathrm{tr}(T_m'|V_{\underline{k},N}) = B_1 + B_2 + O\left(\frac{p^{a_1'\kappa+b_1'}}{(k_1-k_2+1)(k_1-1)(k_2-2)}\right) \quad (k_1+k_2 \to \infty),$$

$$B_1 = O\left(\frac{p^{-\kappa/2}}{(k_1-1)(k_2-2)}\right), \quad B_2 = O\left(\frac{p^{-\kappa/2}}{(k_1-k_2+1)(k_1+k_2-3)}\right)$$

Proof. Since $\nu(m) = p^\kappa$, by (2.23) the classical Hecke operator T_m' is interpreted as the action of

$$f = f_{K(N)}(p^{-(3/2)\kappa}[K_p m^{-1} K_p])\left(\bigotimes_{\ell \in S\setminus\{v|pN\infty\}} \mathrm{char}_{K_\ell}\right)$$

on the spectral side. Then the LHS of the main theorem is exactly $\frac{1}{d_{\underline{k},N}}\mathrm{tr}(T_m'|V_{\underline{k},N})$.

Notice that

$$\dim S_{\underline{k}}(\Gamma(N)) \sim \frac{(\dim\xi_{\underline{k}})\varphi(N)}{\mathrm{vol}(K(N))}, \quad \dim S_{\underline{k}}(\Gamma(N),\chi) \sim \frac{\dim\xi_{\underline{k}}}{\mathrm{vol}(K(N))}.$$

Then multiplying the RHS of the equations in Theorems 5.4 and 5.5 by $p^{-(3/2)\kappa}$, Remark 5.7 implies the result. □

Remark 7.2. The weight aspect depends on how we increase the weight. For instance, if $k_1 = k_2$ goes to infinity, then B_2 becomes the second main term and B_1 is subsumed into the error term. On the other hand, if k_2 is fixed and k_1 goes to infinity, then B_1 becomes the second main term and B_2 is subsumed into the error term. This aspect would be a new case which has not been studied before.

7.2. The vertical Sato–Tate theorem; Proof of Theorem 1.3

Let K be the maximal open compact subgroup of $G(\mathbb{Q}_p) = GSp_4(\mathbb{Q}_p)$. Let us first recall the Plancherel measure $\widehat{\mu}_p^{\mathrm{pl}}$ for the unitary dual of $G(\mathbb{Q}_p)$. For our purpose it suffices to consider its restriction to the unramified tempered classes $\widehat{G(\mathbb{Q}_p)}_\chi^{\mathrm{ur,temp}}$ with a fixed unitary central character $\chi : \mathbb{Q}_p^\times \longrightarrow \mathbb{C}^1$. We denote it by $\widehat{\mu}_{p,\chi}^{\mathrm{pl,temp}}$. Then by [63, Lemma 3.2], we have a natural bijection

$$\widehat{G(\mathbb{Q}_p)}_\chi^{\mathrm{ur,temp}} \simeq [0,\pi]^2/\mathfrak{S}_2 \tag{7.1}$$

which is in fact a topological isomorphism. By [63, Proposition 3.3], for a usual parameter (θ_1, θ_2) of $[0, \pi]^2/\mathfrak{S}_2$, we have

$$\widehat{\mu}^{\mathrm{pl,temp}}_{p,\chi}(\theta_1, \theta_2) = \frac{(p+1)^4}{p^4\pi^2} \cdot \left| \frac{(1-e^{2\sqrt{-1}\theta_1})(1-e^{2\sqrt{-1}\theta_2})(1-e^{\sqrt{-1}(\theta_1+\theta_2)})(1-e^{\sqrt{-1}(\theta_1-\theta_2)})}{(1-p^{-1}e^{2\sqrt{-1}\theta_1})(1-p^{-1}e^{2\sqrt{-1}\theta_2})(1-p^{-1}e^{\sqrt{-1}(\theta_1+\theta_2)})(1-p^{-1}e^{\sqrt{-1}(\theta_1-\theta_2)})} \right|^2 d\theta_1\, d\theta_2.$$

Note that $\lim_{p\to\infty} \widehat{\mu}^{\mathrm{pl,temp}}_{p,1} = \mu^{\mathrm{ST}}_{\infty}$. By transforming (θ_1, θ_2) into $(x, y) = (2\cos\theta_1, 2\cos\theta_2)$, one has the measure μ_p on $\Omega = [-2, 2]^2/\mathfrak{S}_2$ in the introduction.

By Stone–Weierstrass theorem, the natural map $H^{\mathrm{ur}}(G(\mathbb{Q}_p)) \hookrightarrow C^{\infty}_c(G(\mathbb{Q}_p)) \longrightarrow C^0(\Omega, \mathbb{R})$ has the dense image where the second map is given by the restriction of the correspondence $f \mapsto \widehat{f}$ to Ω via (7.1).

Now apply Theorem 7.1 with $m = p$. Then Theorem 1.3 follows from Theorem 1.1.

7.3. Hecke fields; Proof of Corollary 1.4 and 1.5

We first prove Corollary 1.4. Fix a weight $\underline{k} = (k_1, k_2)$ with $k_1 \geqslant k_2 \geqslant 3$ and a prime p. Suppose $[\mathbb{Q}_F : \mathbb{Q}]$ is bounded for any $F \in HE_{\underline{k}}(\Gamma(N), \chi)^{\mathrm{tm}}$, $p \nmid N$. Let $\lambda_F(p) = a_{F,p} + b_{F,p}$ be the Hecke eigenvalue of $T(p)$. By the Ramanujan bound (proved by Laumon–Weissauer [39, 40, 74]) we have

$$|\sigma(\lambda_F(p))| \leqslant 4p^{(k_1+k_2-3)/2}$$

for any $\sigma \in \mathrm{Aut}(\mathbb{C})$. It follows from this with the integrality of $\lambda_F(p)$ (cf. [68, Lemma 2.1]) that there are only finitely many possibilities of $\lambda_F(p)$ when F varies since the weight (k_1, k_2) and p are fixed. Here we also used the assumption that $[\mathbb{Q}_F : \mathbb{Q}]$ is bounded. Let $\lambda_1, \ldots, \lambda_r$ be such possible algebraic numbers for $\lambda_F(p)$. Then the set

$$\{(a_{F,p}, b_{F,p}) \in \Omega \mid F \in HE_{\underline{k}}(\Gamma(N), \chi)^{\mathrm{tm}}\}$$

lies in the union $\bigcup_{i=1}^{r} \{(x, y) \in \Omega \mid x + y = \lambda_i p^{-(k_1+k_2-3)/2}\}$ of hypersurfaces which has the Lebesgue measure zero in Ω. This contradicts to Theorem 1.3. Hence Corollary 1.4 is proved.

To prove Corollary 1.5, we need to estimate the dimension of the endoscopic lifts to $S_{\underline{k}}(\Gamma(N))$. It matches the contribution of the second main term of the geometric side. (See the second term of the right hand side in Theorem 1.1.) Let $S_{\underline{k}}(\Gamma(N))^{\mathrm{en}}$ be the subspace of $S_{\underline{k}}(\Gamma(N))^{\mathrm{tm}}$ generated by Hecke eigenform F such that π_F is endoscopic. By Theorem 3.2, we see that

$$\frac{\dim S_{\underline{k}}(\Gamma(N))^{\mathrm{en}}}{\dim S_{\underline{k}}(\Gamma(N))} = O(((k_1-1)(k_1-2))^{-1}N^{-2+\epsilon}), \quad \text{as } k_1 + k_2 + N \to \infty, \ (N, 11!) = 1.$$

Corollary 1.5 now follows from this with Theorem 1.3.

8. Properties of L-functions of Siegel cusp forms on GSp_4

Put $S_{\underline{k}}(N) = S_{\underline{k}}(\Gamma(N), 1)$ and $HE_{\underline{k}}(N) = HE_{\underline{k}}(\Gamma(N), 1)$ as in the introduction. Given a Siegel cusp form $F \in HE_{\underline{k}}(N)$, let π_F be the associated cuspidal representation of $GSp_4(\mathbb{A})$. Throughout this section, we assume that the central character of π_F is trivial and the level of F satisfies $(N, 11!) = 1$ due to [20] to control the conductor under the functorial lift from the endoscopic subgroup of GSp_4. The result of [20] can be also applied to other groups to study asymptotic behaviors of such lifts. See [44] for $U(3)$.

8.1. Degree 4 spinor L-functions

Let us first assume that π_F is a CAP representation. Since $k_2 \geqslant 3$, by the classification of CAP representations, we must have $k := k_1 = k_2 \geqslant 3$ and it is associated to Siegel parabolic subgroup. As seen in § 3.5, if $(N, 11!) = 1$ there exists a newform f with trivial central character in $S_{2k-2}(\Gamma^1(N))$ so that π_F is defined by π_f and a subset S of $S(\pi_f)$. Let us use the notations for Waldspurger's A-packets in § 3.5 describing π_F. By [60, Remark 3.2] with the local Langlands correspondence for GSp_4 [23] we define the spinor L-function for such π_F by

$$L(s, \pi_F, \mathrm{Spin}) := L(s, \pi_f)L(s, \pi_S)$$

whose local factor coincides with the one defined by the local L-parameter. (See [60, Proposition 1.1] for $L(s, \pi_S)$.) We define the conductor $q(F)$ of $L(s, \pi_F, \mathrm{Spin})$ to be $q(\pi_f)q(\pi_S)$, where $q(\pi_f), q(\pi_S)$ are conductors of $L(s, \pi_f), L(s, \pi_S)$, respectively. Since $q(\pi_S) = \prod_{v \in S} p_v$, we have $q(F)|N^3$.

Next we assume that π_F is endoscopic. As seen in § 3.4 it can be obtained by a theta lift from $H(\mathbb{A})$ where $H = GSO(4)$ or $H = GSO(2, 2)$. We may put $\pi_F = \theta(\tau) = \otimes'_v \theta_v(\tau_v)$ for some cuspidal representation of $\tau = \otimes'_v \tau_v$ on $H(\mathbb{A})$. Since π_F is non-generic, only the case $H = GSO(4)$ happens. Let (π_1, π_2) be a pair of two cuspidal automorphic representations of $GL_2(\mathbb{A})$ with the same central character obtained from τ via Jacquet–Langlands correspondence. As seen before, by [20, Théorème 3.2.3], it turns out that π_i has a fixed vector under the action of $K^1(N)$ under the assumption $(N, 11!) = 1$. We define the spinor L-function

$$L(s, \pi_F, \mathrm{Spin}) := L(s, \tau).$$

The conductor $q(F)$ of $L(s, \pi_F, \mathrm{Spin})$ is defined to be the product of the local conductors. According to the construction of the theta lifting given in § 3.4 we can describe the local L-parameters as follows. Let us follow the notation there. We see that except for the case of (1)-(a), $L(s, \pi_F, \mathrm{Spin})_p = L(s, \pi_{1,p})L(s, \pi_{2,p})$, and $q(F)|N^4$.

In the case of (1)-(a) (when $\pi_{1,p} \simeq \pi_{2,p}$), $\pi_{F,p} \simeq I_{Q(Z)}(1, \pi^{JL}_{1,p})$. The local L-parameter $\rho_p : W_{\mathbb{Q}_p} \longrightarrow GSp_4(\mathbb{C})$ is given in [24, § 13] where $W_{\mathbb{Q}_p}$ stands for the Weil group at p. Indeed it is given by $\rho_p = \rho_p(\pi_{1,p}) \oplus \rho_p(\pi_{1,p})$ where $\rho_p(\pi_{1,p})$ the local Langlands parameter of $\pi_{1,p}$. Therefore, $L(s, \pi_F, \mathrm{Spin})_p := L(s, \rho_p) = L(s, \pi_{1,p})L(s, \pi_{1,p})$. In this case the (local) conductor of ρ_p is less than $\mathrm{ord}_p(N^2)$; in the case of (1)-(b), $\pi_{F,p} \simeq \theta(\pi^{JL}_{1,p} \boxtimes \pi^{JL}_{2,p})$. According to the definition given in [24, § 13], the L-parameter is given by $\rho_p = \rho_p(\pi_{1,p}) \oplus \rho_p(\pi_{2,p})$. Therefore, $L(s, \pi_F, \mathrm{Spin})_p = L(s, \pi_{1,p})L(s, \pi_{2,p})$. In this case the (local) conductor of ρ_p is less than $\mathrm{ord}_p(N^4)$; in the case of (2)-(a), $\pi_{F,p} \simeq I_{P(Y)}(\pi_{1,p} \otimes$

$\chi^{-1}, \chi)$ where $\pi_{1,p}$ is a discrete series and $\pi_{2,p} = \pi(\chi, \chi')$. As seen before, [24, § 13] shows that the L-parameter is given by $\rho_p = \rho_p(\pi_{1,p}) \oplus \rho_p(\chi') \oplus \rho_p(\chi)$ where $\rho_p(\chi)$ is the local Langlands parameter for the character χ given by the local class field theory. Therefore, $L(s, \pi_F, \mathrm{Spin})_p = L(s, \pi_{1,p})L(s, \pi_{2,p})$. In this case the (local) conductor of ρ_p is less than $\mathrm{ord}_p(N^4)$; in the case of (2)-(b), similarly we have $L(s, \pi_F, \mathrm{Spin})_p = L(s, \pi_{1,p})L(s, \pi_{2,p})$ and the (local) conductor of ρ_p is less than $\mathrm{ord}_p(N^4)$.

Finally we assume that π_F is neither CAP nor endoscopic. Since F has a cohomological weight, by [74], π_F is weakly equivalent to a unique globally generic cuspidal representation π of $GSp_4(\mathbb{A})$ so that $\{\pi_{F,\infty}, \pi_\infty\}$ makes up an L-packet of $\Pi(GSp_4(\mathbb{R}))$. Note that the uniqueness follows from [31] once it does exist. Since F is non-endoscopic, the ℓ-adic Galois representation associated to F is irreducible by Chebotarev density theorem and [12], and then we have $\rho_{F,\ell} \sim \rho_{\pi,\ell}$ where $\rho_{F,\ell}$ and $\rho_{\pi,\ell}$ are the corresponding ℓ-adic Galois representations of $G_{\mathbb{Q}}$ attached to π_F, π respectively. By the construction of $\pi = \otimes'_p \pi_p$ in the proof of Theorem 4 in [74], π has a $K(N, M)$-vector for some integer $M > 0$, where $K(N, M) = K(N) \cap \prod_{v|M} I_v$ and I_v is the Iwahori subgroup at v. On the other hand $\rho_{F,\ell}$ is unramified outside N and so is $\rho_{\pi,\ell}$. Therefore, one can conclude $M = 1$ and as a result π has a $K(N)$-fixed vector. Then we define the spinor L-function of π_F by

$$L(s, \pi_F, \mathrm{Spin}) := L(s, \pi, \mathrm{Spin})$$

where the RHS is defined by using Novodvorsky L-function (see [67]). We define the conductor $q(F)$ of $L(s, \pi_F, \mathrm{Spin})$ to be the conductor $q(\pi)$ of $L(s, \pi, \mathrm{Spin})$. For each prime $p \nmid N$, we may write

$$L(s, \pi_F, \mathrm{Spin})_p^{-1} = (1 - \alpha_{0p} p^{-s})(1 - \alpha_{0p}\alpha_{1p} p^{-s})(1 - \alpha_{0p}\alpha_{2p} p^{-s})(1 - \alpha_{0p}\alpha_{1p}\alpha_{2p} p^{-s}),$$
$$\widetilde{\lambda}_F(p) = \alpha_{0p} + \alpha_{0p}\alpha_{1p} + \alpha_{0p}\alpha_{2p} + \alpha_{0p}\alpha_{1p}\alpha_{2p} = \lambda_F(p) p^{-(k_1+k_2-3)/2}.$$

We note that $\widetilde{\lambda}_F(p^2) = \lambda_F(p^2) p^{-(k_1+k_2+3)} + p^{-1}$.

Since the central character is trivial, one has a relation $\alpha_{0p}^2 \alpha_{1p}\alpha_{2p} = 1$. Let $\Gamma_{\mathbb{R}}(s) = \pi^{-s/2}\Gamma(\frac{s}{2})$ and $\Gamma_{\mathbb{C}}(s) = 2(2\pi)^{-s}\Gamma(s)$. Then we have

Lemma 8.1. *Let* $\Lambda(s, \pi_F, \mathrm{Spin}) = q(F)^{s/2}\Gamma_{\mathbb{C}}(s + \frac{k_1+k_2-3}{2})\Gamma_{\mathbb{C}}(s + \frac{k_1-k_2+1}{2})L(s, \pi_F, \mathrm{Spin})$. *Then*

$$\Lambda(s, \pi_F, \mathrm{Spin}) = \epsilon(\pi_F)\Lambda(1 - s, \pi_F, \mathrm{Spin}),$$

where $\epsilon(\pi_F) \in \{\pm 1\}$ *and* $N \leqslant q(F) \leqslant N^4$.

Proof. Since the functional equation is well known, we prove only the bound on the conductor $q(F)$. When π_F is either CAP or endoscopic, it was proved above. Otherwise, let $\pi' = \otimes'_p \pi'_p$ of $GL_4(\mathbb{A})$ be the strong transfer of π_F. The conductor $q(F)$ can be written in terms of the local conductor of π'. By [65, Proposition 1] (see also the last few lines of its proof) which is still true for not only supercuspidal representations but also square integrable representations, the depth is preserved under the above transfer. For a prime p such that $\pi_{F,p}$ is square integrable, we have

$$\mathrm{ord}_p(q(F)) = c(\pi_{F,p}) = c(\pi'_p) = 4(\mathrm{depth}(\pi'_p) + 1) = 4(\mathrm{depth}(\pi_{F,p}) + 1)$$

where we applied [38, Proposition 2.2] (respectively Main Theorem of [65]) to get the third (respectively the second) equality. By definition of depth, $\pi_F^{K(N)} \neq 0$ implies $\mathrm{depth}(\pi'_p) \leqslant \mathrm{ord}_p(N) - 1$. This gives $\mathrm{ord}_p(q(F)) \leqslant 4 \cdot \mathrm{ord}_p(N)$ provided if $\pi_{F,p}$ is square integrable or unramified. In the remaining cases we can directly compute the conductor by using the [51, Table A.9] under the condition $\pi_F^{K(N)} \neq 0$. □

Let

$$-\frac{L'}{L}(s, \pi_F, \mathrm{Spin}) = \sum_{n=1}^{\infty} \Lambda(n) a_F(n) n^{-s},$$

where $\Lambda(n)$ is the von Mangoldt function, and if $p \nmid N$,

$$a_F(p^d) = \alpha_{0p}^d + (\alpha_{0p}\alpha_{1p})^d + (\alpha_{0p}\alpha_{2p})^d + (\alpha_{0p}\alpha_{1p}\alpha_{2p})^d.$$

For each $m \in T(\mathbb{Q})$ we normalize the Hecke operator T_m so that $T'_m := T_m \nu(m)^{-(k_1+k_2-3)/2}$ and accordingly $T'(p^n) = T(p^n) p^{-n(k_1+k_2-3)/2}$. For a Hecke eigen form F, we denote by $\lambda'_{F,m}$ (respectively $\lambda'_F(p^n)$) the Hecke eigenvalue of F for T'_m (respectively $T'_F(p^n)$). By using the relations (2.7) it is easy to see that

$$\begin{aligned} a_F(p) &= \lambda'_F(p), \\ a_F(p^2) &= \lambda'_{F,t_1^2} - (p-1)\lambda'_{F,t_2} - \left(1 - \frac{1}{p}\right)\left(1 + \frac{1}{p^2}\right), \end{aligned}$$

where $t_1 = \mathrm{diag}(1, 1, p, p)$ and $t_2 = \mathrm{diag}(1, p, p^2, p)$. To apply Theorem 7.1 we have to express these values in terms of linear combinations of the eigenvalues for the Hecke operators which take the shape of T'_m as in Theorem 7.1. Then we find the corresponding operators

$$\begin{aligned} T'(p) &= T'_{t_1}, \\ T'(p^2) &= T'_{t_1^2} - 2pT'_{t_2} - 2(p^{-1} + p^{-2})(p^3 T'_{pE_4}), \end{aligned}$$

respectively. Note that $p^3 T'_{pE_4}$ acts on $S_{\underline{k}}(\Gamma(N))$ as the identity map. Therefore, we have

Proposition 8.2. *Assume* $(N, 11!) = 1$. *Let* $p \nmid N$. *Put* $\underline{k} = (k_1, k_2)$, $k_1 \geqslant k_2 \geqslant 3$ *and* $d_{\underline{k},N} := \dim S_{\underline{k}}(N)$. *There exist absolute constants* $a''_1, a''_2, b''_1, b''_2, c''_1, c''_2, v_1, v'_1, w_1, w'_1$ *such that*

(1) (a) *(level aspect) Fix* k_1, k_2. *Then for* $N \gg p^{10}$,

$$\frac{1}{d_{\underline{k},N}} \sum_{F \in HE_{\underline{k}}(N)} a_F(p) = O(p^{-1/2} N^{-2}) + O(p^{v_1} N^{-3});$$

(b) *(weight aspect) Fix* N. *Then as* $k_1 + k_2 \to \infty$,

$$\frac{1}{d_{\underline{k},N}} \sum_{F \in HE_{\underline{k}}(N)} a_F(p) = B_1 + B_2 + O\left(\frac{p^{v'_1}}{(k_1 - k_2 + 1)(k_1 - 1)(k_2 - 2)}\right),$$

$$B_1 = O\left(\frac{p^{-1/2}}{(k_1 - 1)(k_2 - 2)}\right), \quad B_2 = O\left(\frac{p^{-1/2}}{(k_1 - k_2 + 1)(k_1 + k_2 - 3)}\right).$$

(2) (a) *(level aspect) Fix k_1, k_2. Then for $N \gg p^{10}$,*

$$\frac{1}{d_{\underline{k},N}} \sum_{F \in HE_{\underline{k}}(N)} a_F(p^2) = -\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p^2}\right) + O\left((a_1'' p^{-1/2} + a_2'' p^{1/2}) N^{-2}\right) + O(p^{w_1} N^{-3}).$$

(b) *(weight aspect) Fix N. Then as $k_1 + k_2 \to \infty$,*

$$\frac{1}{d_{\underline{k},N}} \sum_{F \in HE_{\underline{k}}(N)} a_F(p^2) = -\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p^2}\right) + B_1 + B_2 + O\left(\frac{p^{w_1'}}{(k_1-k_2+1)(k_1-1)(k_2-2)}\right),$$

$$B_1 = O\left(\frac{p^{-1/2} b_1'' + p^{1/2} b_2''}{(k_1-1)(k_2-2)}\right), \quad B_2 = O\left(\frac{p^{-1/2} c_1'' + p^{1/2} c_2''}{(k_1-k_2+1)(k_1+k_2-3)}\right).$$

Proof. The claim follows from Theorem 7.1. □

8.2. Degree 5 standard L-functions

Let us first recall the standard L-function for $\pi_F = \otimes'_p \pi_{F,p}$. Let $\tau = \pi_F|_{Sp_4}$ be the restriction of π_F to $Sp_4(\mathbb{A})$, and Π be the transfer of τ corresponding to $\omega_2|_{Sp_4(\mathbb{C})}$, where $\omega_2 : GSp_4(\mathbb{C}) \longrightarrow GL_5(\mathbb{C})$ is the homomorphism attached to the second fundamental weight. Note that if $\iota : GSp_4(\mathbb{C}) \hookrightarrow GL_4(\mathbb{C})$, $\wedge^2 \circ \iota = \omega_2 \oplus 1$. Therefore, $\wedge^2 \pi = \Pi \boxplus 1$, and $L(s, \pi_F, \mathrm{St}) = L(s, \tau, \mathrm{St}) = L(s, \Pi)$. We define the conductor $q(F, \mathrm{St})$ to be the conductor of $L(s, \tau, \mathrm{St})$. For any unramified prime $p \nmid N$,

$$L(s, \pi_F, \mathrm{St})_p^{-1} = (1-p^{-s})(1-\alpha_{1p} p^{-s})(1-\alpha_{2p} p^{-s})(1-\alpha_{1p}^{-1} p^{-s})(1-\alpha_{2p}^{-1} p^{-s}).$$

Lemma 8.3. *Let $\Lambda(s, \pi_F, \mathrm{St}) = q(F, \mathrm{St})^{s/2} \Gamma_{\mathbb{R}}(s) \Gamma_{\mathbb{C}}(s+k_1-1) \Gamma_{\mathbb{C}}(s+k_2-2) L(s, \pi_F, \mathrm{St})$. Then*

$$\Lambda(s, \pi_F, \mathrm{St}) = \epsilon(\pi_F, \mathrm{St}) \Lambda(1-s, \pi_F, \mathrm{St}),$$

where $\epsilon(\pi_F, \mathrm{St}) \in \{\pm 1\}$, and $N \leqslant q(F, \mathrm{St}) \leqslant N^{28}$.

(Henniart noted in a private communication that we would have $q(F, \mathrm{St}) \ll (N^4)^{3/2} = N^6$.)

Proof. We bound the conductor $q(F, \mathrm{St})$ since others are well known. By Lemma 8.1 we know that $q(F) \leqslant N^4$. Let π' be the strong transfer of π_F to $GL_4(\mathbb{A})$. Then the global conductor $q(\pi')$ coincides with $q(F)$. Then the conductor is roughly estimated by the main theorem of [11] as follows:

$$q(F, \mathrm{St}) \leqslant q(\pi' \otimes \pi') \leqslant N^{4(2\cdot 4-1)} = N^{28}.$$

This gives us the claim. □

Let

$$L(s,\pi_F,\mathrm{St})=\sum_{n=1}^{\infty}\mu_F(n)n^{-s}.$$

Then if $p\nmid N$,

$$\mu_F(p)=1+\alpha_{1p}+\alpha_{2p}+\alpha_{1p}^{-1}+\alpha_{2p}^{-1},\quad \lambda'_F(p)^2-\lambda'_F(p^2)-p^{-1}=\mu_F(p)+1.$$

Let

$$-\frac{L'}{L}(s,\pi_F,\mathrm{St})=\sum_{n=1}^{\infty}\Lambda(n)b_F(n)n^{-s},$$

where if $p\nmid N$, $b_F(p^d)=1+\alpha_{1p}^d+\alpha_{2p}^d+\alpha_{1p}^{-d}+\alpha_{2p}^{-d}$. Note that $a_F(p^2)=\lambda'_F(p)^2-b_F(p)-1=\lambda'_F(p^2)+p^{-1}$, and

$$b_F(p)=\mu_F(p)=p^{-1}\lambda'_{F,t_2}+p^{-2}. \tag{8.1}$$

By using the relations (2.7), we see that

$$\begin{aligned}(\lambda'_{F,t_2})^2 &= \lambda'_{F,\mathrm{diag}(1,p^2,p^4,p^2)}+(p+1)\lambda'_{F,\mathrm{diag}(p,p,p^3,p^3)}\\ &\quad+(p^2-1)\lambda'_{F,\mathrm{diag}(p,p^2,p^3,p^2)}+p^{-6}(1+p+p^3+p^4).\end{aligned}$$

Therefore,

$$\begin{aligned}b_F(p^2) &= b_F(p)^2-2a_F(p^2)-2b_F(p)-2\\ &= (p\lambda'_{F,t_2}+p^{-2})^2-2(\lambda'_F(p^2)+p^{-1})-2a_F(p^2)-2\\ &= p^2(\lambda'_{F,t_2})^2+2p^{-1}\lambda'_{F,t_2}-2\lambda'_{F,t_2}-2a_F(p^2)+p^{-4}-2p^{-1}-2\\ &= p^2\lambda'_{F,\mathrm{diag}(1,p^2,p^4,p^2)}+p^2(p+1)\lambda'_{F,\mathrm{diag}(p,p,p^3,p^3)}+p^2(p^2-1)\lambda'_{F,\mathrm{diag}(p,p^2,p^3,p^2)}\\ &\quad+(2p^{-1}-2)\lambda'_{F,t_2}-2\left(\lambda'_{F,t_1^2}-(p-1)\lambda'_{F,t_2}-\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p^2}\right)\right)\\ &\quad-1-p^{-1}+p^{-3}+2p^{-4},\\ &= p^2\lambda'_{F,\mathrm{diag}(1,p^2,p^4,p^2)}+p^2(p+1)\lambda'_{F,\mathrm{diag}(p,p,p^3,p^3)}+p^2(p^2-1)\lambda'_{F,\mathrm{diag}(p,p^2,p^3,p^2)}\\ &\quad+2(p+p^{-1}-2)\lambda'_{F,t_2}-2\lambda'_{F,t_1^2}+1-3p^{-1}+2p^{-2}-p^{-3}+2p^{-4},\end{aligned} \tag{8.2}$$

where $t_1=\mathrm{diag}(1,1,p,p)$ and $t_2=\mathrm{diag}(1,p,p^2,p)$. To obtain an estimation for the average of $b_F(p^2)$, according to (8.2), we apply Theorem 7.1 to $\lambda'_{F,\mathrm{diag}(1,p^2,p^4,p^2)}$, $\lambda'_{F,\mathrm{diag}(p,p,p^3,p^3)}$, $\lambda'_{F,\mathrm{diag}(p,p^2,p^3,p^2)}$, λ_{F,t_2}, and λ_{F,t_1^2}.

Remark 8.4. We can see easily that $L(s,\Pi,\wedge^2)=L(s,\pi,\mathrm{Sym}^2)$. Under the Langlands functoriality conjecture, we expect $\mathrm{Sym}^2(\pi)$ to be an automorphic representation of GL_{10}. Since $L(s,\pi,\wedge^2)$ has a pole at $s=1$, $L(s,\pi,\mathrm{Sym}^2)$ has no pole at $s=1$. Let $L(s,\pi,\mathrm{Sym}^2)=\sum_{n=1}^{\infty}\lambda_{\mathrm{Sym}^2\pi}(n)n^{-s}$. Then $\lambda_{\mathrm{Sym}^2\pi}(p)=\widetilde{\lambda}_F(p^2)=\lambda'_F(p^2)+p^{-1}$.

Recall that $S_{\underline{k}}(N)=S_{\underline{k}}(\Gamma(N),1)$ and $HE_{\underline{k}}(N)=HE_{\underline{k}}(\Gamma(N),1)$. Then by Theorem 7.1 we have the following:

Proposition 8.5. *Let $p \nmid N$. Put $\underline{k} = (k_1, k_2)$ and $d_{\underline{k},N} := \dim S_{\underline{k}}(N)$. There exist absolute constants v_1, v_1', w_1, w_1' such that*

(1) (a) *(level aspect) Fix k_1, k_2. Then for $N \gg p^{30}$,*

$$\frac{1}{d_{\underline{k},N}} \sum_{F \in HE_{\underline{k}}(N)} b_F(p) = p^{-2} + O(p^{-3/2}N^{-2}) + O(p^{v_1}N^{-3}).$$

(b) *(weight aspect) Fix N. Then as $k_1 + k_2 \to \infty$,*

$$\frac{1}{d_{\underline{k},N}} \sum_{F \in HE_{\underline{k}}(N)} b_F(p) = p^{-2} + B_1 + B_2 + O\left(\frac{p^{v_1'}}{(k_1 - k_2 + 1)(k_1 - 1)(k_2 - 2)}\right),$$

$$B_1 = O\left(\frac{p^{-3/2}}{(k_1 - 1)(k_2 - 2)}\right), \quad B_2 = O\left(\frac{p^{-3/2}}{(k_1 - k_2 + 1)(k_1 + k_2 - 3)}\right).$$

(2) (a) *(level aspect) Fix k_1, k_2. Then for $N \gg p^{10}$,*

$$\frac{1}{d_{\underline{k},N}} \sum_{F \in HE_{\underline{k}}(N)} b_F(p^2) = 1 - 3p^{-1} + 2p^{-2} - p^{-3} + 2p^{-4} + O(p^2 f_A(p^{-1})N^{-2}) + O(p^{w_1}N^{-3}).$$

(b) *(weight aspect) Fix N. Then as $k_1 + k_2 \to \infty$,*

$$\frac{1}{d_{\underline{k},N}} \sum_{F \in HE_{\underline{k}}(N)} b_F(p^2) = 1 - 3p^{-1} + 2p^{-2} - p^{-3} + 2p^{-4} + B_1 + B_2 + O\left(\frac{p^{w_1'}}{(k_1 - k_2 + 1)(k_1 - 1)(k_2 - 2)}\right),$$

$$B_1 = O\left(\frac{p^2 f_{B_1}(p^{-1})}{(k_1 - 1)(k_2 - 2)}\right), \quad B_2 = O\left(\frac{p^2 f_{B_1}(p^{-1})}{(k_1 - k_2 + 1)(k_1 + k_2 - 3)}\right),$$

where $f_A(X), f_{B_1}(X), f_{B_2}(X)$ are the polynomials over $\mathbb{Q}$ of the degree 4 in X whose coefficients are independent of p, k_1, k_2, and N.

9. One-level density

We follow the exposition in [14]. Katz and Sarnak [32] proposed a conjecture on low-lying zeros of L-functions in natural families $\mathfrak{F}$, which says that the distributions of the low-lying zeros of L-functions in a family $\mathfrak{F}$ is predicted by a symmetry group $G(\mathfrak{F})$ attached to $\mathfrak{F}$. See [58] for a refined formulation: For a given entire L-function $L(s, \pi)$, we denote the non-trivial zeros of $L(s, \pi)$ by $\frac{1}{2} + \gamma_j\sqrt{-1}$. Since we do not assume GRH for $L(s, \pi)$, γ_j can be a complex number. Let $\phi(x)$ be a Schwartz function which is even and whose Fourier transform

$$\hat{\phi}(y) = \int_{-\infty}^{\infty} \phi(x) e^{-2\pi xy\sqrt{-1}}\, dx$$

has a compact support. We define

$$D(\pi,\phi)=\sum_{\gamma_j}\phi\left(\frac{\gamma_j}{2\pi}\log c_\pi\right)$$

where c_π is the analytic conductor of $L(s,\pi)$.

Let $\mathfrak{F}(X)$ be the set of L-functions in $\mathfrak{F}$ such that $X<c_\pi<2X$. The one-level density conjecture says that, for a Schwartz $\phi(x)$ which is even and whose Fourier transform $\hat\phi(y)$ is compactly supported,

$$\lim_{X\to\infty}\frac{1}{\#\mathfrak{F}(X)}\sum_{\pi\in\mathfrak{F}(X)}D(\pi,\phi)=\int_{-\infty}^{\infty}\phi(x)W(G(\mathfrak{F}))\,dx,$$

where $W(G(\mathfrak{F}))$ is the one-level density function. There are five possible symmetry type of families of L-functions: U, SO(even), SO(odd), O, and Sp. The corresponding density functions $W(G)$ are as in Theorem 1.6 [32].

By Plancherel's formula (and because ϕ is even),

$$\int_{-\infty}^{\infty}\phi(x)W(G)(x)\,dx=\int_{-\infty}^{\infty}\hat\phi(x)\widehat{W}(G)(x)\,dx.$$

It is useful to record that

$$\widehat{W}(\mathrm{U})(x)=\delta_0(x),\quad \widehat{W}(\mathrm{SO(even)})(x)=\delta_0(x)+\tfrac{1}{2}\chi_{[-1,1]}(x),\quad \widehat{W}(\mathrm{O})(x)=\delta_0(x)+\tfrac{1}{2}$$
$$\widehat{W}(\mathrm{SO(odd)})(x)=\delta_0(x)-\chi_{[-1,1]}(x)+1,\quad \widehat{W}(\mathrm{Sp})(x)=\delta_0(x)-\tfrac{1}{2}\chi_{[-1,1]}(x).$$

We study one-level density of the family $HE_{\underline{k}}(N)$. Here we can assume that $F\in HE_{\underline{k}}(N)$ is not a CAP form. If F is a CAP form, $|a_F(p)|\leqslant 4p^{1/2}$ and $|a_F(p^2)|\leqslant 4p$ in (9.2) and (9.3). Hence if the support of ϕ is smaller than $(-1,1)$, then the sum over p is $O(N)$. But the dimension of the space of CAP forms is $O(N^{7+\epsilon})$. Hence it is negligible.

9.1. Degree 4 spinor L-functions

We denote the non-trivial zeros of $L(s,\pi_F,\mathrm{Spin})$ by $\sigma_F=\frac{1}{2}+\sqrt{-1}\gamma_F$. We do not assume GRH, and hence γ_F can be a complex number. Let ϕ be a Schwartz function which is even and whose Fourier transform has a compact support. Define

$$D(\pi_F,\phi,\mathrm{Spin})=\sum_{\gamma_F}\phi\left(\frac{\gamma_F}{2\pi}\log c_{\underline{k},N}\right),$$

where $\log c_{\underline{k},N}=\dfrac{1}{d_{\underline{k},N}}\displaystyle\sum_{F\in HE_{\underline{k}}(N)}\log c(F,\mathrm{Spin})$ for $\underline{k}=(k_1,k_2)$, and $c(F,\mathrm{Spin})=(k_1+k_2)^2(k_1-k_2+1)^2q(F)$ is the analytic conductor (cf. [14]).

Proposition 9.1. *Assume* $(N,11!)=1$. *Let* ϕ *be a Schwartz function which is even and whose Fourier transform has a support sufficiently smaller than* $(-1,1)$.

$$\lim_{k_1+k_2+N\to\infty}\frac{1}{d_{\underline{k},N}}\sum_{F\in HE_{\underline{k}}(N)}D(\pi_F,\phi,\mathrm{Spin})=\hat\phi(0)+\frac{1}{2}\phi(0)=\int_{\mathbb{R}}\phi(x)W(G)(x)\,dx,$$

where $G=$ SO(even), SO(odd), or O type. More precisely, let v_1, w_1, v_1', w_1' be as in Proposition 8.2.

(1) *(level aspect) Fix k_1, k_2. Then for ϕ whose Fourier transform $\hat{\phi}$ has support in $(-u,u)$, where $u=\min\{\frac{3}{4v_1+2}, \frac{3}{4w_1}, \frac{1}{40}\}$, as $N\to\infty$,*

$$\frac{1}{d_{\underline{k},N}}\sum_{F\in HE_{\underline{k}}(N)} D(\pi_F,\phi,\mathrm{Spin}) = \hat{\phi}(0)+\frac{1}{2}\phi(0)+O\left(\frac{1}{\log\log N}\right).$$

(2) *(weight aspect) Fix N. Then for ϕ whose Fourier transform $\hat{\phi}$ has support in $(-u,u)$, where $u=\min\{\frac{1}{2v_1'+1}, \frac{1}{2w_1'}\}$, as $k_1+k_2\to\infty$,*

$$\frac{1}{d_{\underline{k},N}}\sum_{F\in HE_{\underline{k}}(N)} D(\pi_F,\phi,\mathrm{Spin}) = \hat{\phi}(0)+\frac{1}{2}\phi(0)+O\left(\frac{1}{\log((k_1-k_2+2)k_1k_2)}\right).$$

Proof. For $G(s)=\phi\big((s-\frac{1}{2})\frac{\log c_{k,N}}{2\pi\sqrt{-1}}\big)$, by Cauchy's theorem,

$$D(\pi_F,\phi,\mathrm{Spin})=\sum_{\gamma_F} G(\sigma_F)=\frac{1}{2\pi\sqrt{-1}}\int_{(2)} 2G(s)\frac{\Lambda'(s,\pi_F,\mathrm{Spin})}{\Lambda(s,\pi_F,\mathrm{Spin})}\,ds.$$

We have

$$\begin{aligned}\frac{\Lambda'(s,\pi_F,\mathrm{Spin})}{\Lambda(s,\pi_F,\mathrm{Spin})} &= \frac{1}{2}\log q(F)+\psi\left(s+\frac{k_1+k_2-3}{2}\right)+\psi\left(s+\frac{k_1-k_2+1}{2}\right)\\ &\quad -\sum_{n=1}^{\infty}\frac{\Lambda(n)a_K(n)}{n^s}\end{aligned}$$

where $\psi(s)=\frac{\Gamma_{\mathbb{C}}'(s)}{\Gamma_{\mathbb{C}}(s)}$.

The contribution coming from the logarithmic derivative of $L(s,\pi_F,\mathrm{Spin})$ is

$$\begin{aligned}&\frac{1}{2\pi\sqrt{-1}}\int_{(2)} 2G(s)\left(-\sum_{n=1}^{\infty}\frac{\Lambda(n)a_F(n)}{n^s}\right)ds\\ &=-\frac{2}{\log c_{\underline{k},N}}\sum_{n=1}^{\infty}\Lambda(n)a_F(n)\frac{1}{2\pi\sqrt{-1}}\int_{(2)}\phi\left(\left(s-\frac{1}{2}\right)2\pi\sqrt{-1}\right)n^{-s}\,ds\\ &=-\frac{2}{\log c_{\underline{k},N}}\sum_{n=1}^{\infty}\frac{\Lambda(n)a_F(n)}{\sqrt{n}}\int_{-\infty}^{\infty}\phi(y)e^{-y(2\pi\log n/\log c_{\underline{k},N})\sqrt{-1}}\,dy\\ &=-\frac{2}{\log c_{\underline{k},N}}\sum_{n=1}^{\infty}\frac{\Lambda(n)a_F(n)}{\sqrt{n}}\widehat{\phi}\left(\frac{\log n}{\log c_{\underline{k},N}}\right). \qquad (9.1)\end{aligned}$$

The contribution of the constant term $A = \frac{1}{2}\log q(F)$ is

$$\frac{1}{2\pi i}\int_{(2)} 2G(s)A\,ds = \frac{\log q(F)}{2\pi}\int_{-\infty}^{\infty}\phi\left(\frac{\log c_{\underline{k},N}}{2\pi}y\right)dy = \frac{\log q(F)}{2\log c_{\underline{k},N}}\int_{-\infty}^{\infty}\phi(y)\,dy = \frac{\log q(F)}{\log c_{\underline{k},N}}\widehat{\phi}(0).$$

For the Gamma factors' contribution, we use, for $a, t \in \mathbb{R}$, $a > 0$, (cf. [14])

$$\frac{\Gamma'}{\Gamma}(a+t\sqrt{-1}) + \frac{\Gamma'}{\Gamma}(a-t\sqrt{-1}) = 2\frac{\Gamma'}{\Gamma}(a) + O(t^2a^{-2}).$$

For $\alpha \geqslant \frac{1}{4}$, $\frac{\Gamma'}{\Gamma}(\alpha+\frac{1}{4}) = \log\alpha + O(1)$. Hence the Gamma factors contribute

$$\frac{2\log(k_1+k_2) + 2\log(k_1-k_2+1)}{\log c_{\underline{k},N}}\widehat{\phi}(0) + O\left(\frac{1}{\log^3 c_{\underline{k},N}}\right).$$

It is shown in [15] that the prime powers p^l, $l \geqslant 3$ from (9.1), contribute $O\left(\frac{1}{\log c_{\underline{k},N}}\right)$; If π_F satisfies the Ramanujan conjecture, $|a_F(n)| \leqslant 4$, and it is obvious. Hence

$$\frac{1}{d_{\underline{k},N}}\sum_{F\in HE_{\underline{k}}(N)} D(\pi_F,\phi,\mathrm{Spin}) = \widehat{\phi}(0) - \frac{2}{(\log c_{\underline{k},N})d_{\underline{k},N}}\sum_{F\in HE_{\underline{k}}(N)}\sum_p \frac{a_F(p)\log p}{\sqrt{p}}\widehat{\phi}\left(\frac{\log p}{\log c_{\underline{k},N}}\right) - \frac{2}{(\log c_{\underline{k},N})d_{\underline{k},N}}\sum_{F\in HE_{\underline{k}}(N)}\sum_p \frac{a_F(p^2)\log p}{p}\widehat{\phi}\left(\frac{2\log p}{\log c_{\underline{k},N}}\right) + O\left(\frac{1}{\log c_{\underline{k},N}}\right).$$

Let $\tilde{a}_F(p) = a_F(p^2)+1$. We note, from the prime number theorem,

$$\sum_p \widehat{\phi}\left(\frac{2\log p}{\log c_{k,N}}\right)\frac{2\log p}{p\log c_{\underline{k},N}} = \int_2^{\infty}\widehat{\phi}\left(\frac{2\log t}{\log c_{\underline{k},N}}\right)\frac{2\log t}{t\log c_{k,N}}\,d\pi(t) + O\left(\frac{1}{\log c_{\underline{k},N}}\right) = \frac{1}{2}\phi(0) + O\left(\frac{1}{\log c_{\underline{k},N}}\right).$$

Hence

$$\frac{1}{d_{\underline{k},N}}\sum_{F\in HE_{\underline{k}}(N)} D(\pi_F,\phi,\mathrm{Spin}) = \widehat{\phi}(0) + \frac{1}{2}\phi(0) - \frac{2}{(\log c_{\underline{k},N})d_{\underline{k},N}}\sum_{F\in HE_{\underline{k}}(N)}\sum_p \frac{a_F(p)\log p}{\sqrt{p}}\widehat{\phi}\left(\frac{\log p}{\log c_{\underline{k},N}}\right) \tag{9.2}$$

$$- \frac{2}{(\log c_{\underline{k},N})d_{\underline{k},N}}\sum_{F\in HE_{\underline{k}}(N)}\sum_p \frac{\tilde{a}_F(p)\log p}{p}\widehat{\phi}\left(\frac{2\log p}{\log c_{\underline{k},N}}\right) + O\left(\frac{1}{\log c_{\underline{k},N}}\right). \tag{9.3}$$

Now we exchange the two sums. If $p \nmid N$, use Proposition 8.2. If $p|N$, since $|a_F(p)| \leqslant p^{1/2-1/17}$ and $|a_F(p^2)| \leqslant p^{1-2/17}$ (Note that the Ramanujan bound in [42] is valid also for ramified primes.), by the trivial bound, $\sum_{p|N} \frac{a_F(p)\log p}{\sqrt{p}} \leqslant \omega(N)$ and $\sum_{p|N} \frac{a_F(p^2)\log p}{p} \leqslant \omega(N)$. Hence if the support of $\hat{\phi}$ is $(-u,u)$ for an appropriate $u<1$,

$$\frac{1}{d_{\underline{k},N}} \sum_{F\in HE_{\underline{k}}(N)} D(\pi_F,\phi,\text{Spin}) = \widehat{\phi}(0)+\frac{1}{2}\phi(0)+O\left(\frac{\omega(N)}{\log c_{\underline{k},N}}\right).$$

The well-known bound $\omega(N) \ll \frac{\log N}{\log\log N}$ implies our result. □

Remark 9.2. Since the support of $\hat{\phi}$ is smaller than $(-1,1)$, we cannot distinguish the symmetry type among SO(even), SO(odd), or O type. In order to distinguish them, we need to compute the n-level density (cf. [15]). We show in an upcoming paper that when the root number $\epsilon(\pi_F)=1$, the symmetry type is SO(even); when the root number $\epsilon(\pi_F)=-1$, the symmetry type is SO(odd).

9.2. Degree 5 standard L-functions

As in the degree 4 spin L-function case, denote the non-trivial zeros of $L(s,\pi_F,\text{St})$ by $\sigma_F=\frac{1}{2}+\sqrt{-1}\gamma_F$. We do not assume GRH, and hence γ_F can be a complex number. Let ϕ be a Schwartz function which is even and whose Fourier transform has a compact support. Define

$$D(\pi_F,\phi,\text{St}) = \sum_{\gamma_F} \phi\left(\frac{\gamma_F}{2\pi}\log c_{\underline{k},st,N}\right),$$

where $\log c_{\underline{k},st,N} = \frac{1}{d_{\underline{k},N}}\sum_{F\in HE_{\underline{k}}(N)} \log c(F,\text{St})$, and $c(F,\text{St})=(k_1k_2)^2 q(F,\text{St})$ is the analytic conductor.

As in the degree 4 spinor L-function case, we can show that

$$\begin{aligned}\frac{1}{d_{\underline{k},N}} \sum_{F\in HE_{\underline{k}}(N)} D(\pi_F,\phi,\text{St}) &= \widehat{\phi}(0)-\frac{1}{2}\phi(0)\\ &-\frac{2}{(\log c_{\underline{k},st,N})d_{\underline{k},N}} \sum_{F\in HE_{\underline{k}}(N)}\sum_p \frac{b_F(p)\log p}{\sqrt{p}}\widehat{\phi}\left(\frac{\log p}{\log c_{\underline{k},st,N}}\right) &(9.4)\\ &-\frac{2}{(\log c_{\underline{k},st,N})d_{\underline{k},N}} \sum_{F\in HE_{\underline{k}}(N)}\sum_p \frac{\tilde{b}_F(p)\log p}{p}\widehat{\phi}\left(\frac{2\log p}{\log c_{\underline{k},st,N}}\right)+O\left(\frac{1}{\log c_{\underline{k},st,N}}\right) &(9.5)\end{aligned}$$

where $\tilde{b}_F(p)=b_F(p^2)-1$. (If π_F satisfies the Ramanujan conjecture, then $|b_F(p^l)|\leqslant 5$ and we can show easily that the prime powers p^l, $l\geqslant 3$, contribute to $O(\frac{1}{\log c_{\underline{k},st,N}})$. In the appendix, we show it without any assumptions.)

By interchanging two sums and using Proposition 8.5 as in § 9.1, we see that if the support of $\hat{\phi}$ is $(-u,u)$ for an appropriate $u<1$,

$$\frac{1}{d_{\underline{k},N}} \sum_{F\in HE_{\underline{k}}(N)} D(\pi_F,\phi,\text{St}) = \widehat{\phi}(0)-\frac{1}{2}\phi(0)+O\left(\frac{\omega(N)}{\log c_{\underline{k},st,N}}\right).$$

Hence we have proved

Proposition 9.3. *Let ϕ be a Schwartz function which is even and which its Fourier transform has a support sufficiently smaller than* $(-1, 1)$.

$$\lim_{k_1+k_2+N\to\infty} \frac{1}{d_{\underline{k},N}} \sum_{F\in HE_{\underline{k}}(N)} D(\pi_F, \phi, \mathrm{St}) = \hat{\phi}(0) - \frac{1}{2}\phi(0) = \int_{\mathbb{R}} \phi(x)W(\mathrm{Sp})(x)\,dx.$$

More precisely, let v_1, w_1, v_1', w_1' be as in Proposition 8.5.

(1) *(level aspect) Fix k_1, k_2. Then for ϕ whose Fourier transform $\hat{\phi}$ has support in $(-u, u)$, where $u = \min\{\frac{3}{28v_1+14}, \frac{3}{w_1}, \frac{1}{840}\}$, as $N \to \infty$,*

$$\frac{1}{d_{\underline{k},N}} \sum_{F\in HE_{\underline{k}}(N)} D(\pi_F, \phi, \mathrm{St}) = \hat{\phi}(0) - \frac{1}{2}\phi(0) + O\left(\frac{1}{\log\log N}\right).$$

(2) *(weight aspect) Fix N. Then for ϕ whose Fourier transform $\hat{\phi}$ has support in $(-u, u)$, where $u = \min\{\frac{1}{2v_1'+1}, \frac{1}{2w_1'}, \frac{1}{8}\}$, as $k_1 + k_2 \to \infty$,*

$$\frac{1}{d_{\underline{k},N}} \sum_{F\in HE_{\underline{k}}(N)} D(\pi_F, \phi, \mathrm{St}) = \hat{\phi}(0) - \frac{1}{2}\phi(0) + O\left(\frac{1}{\log((k_1-k_2+2)k_1k_2)}\right).$$

10. Stable vs unstable pseudo-coefficients

In this section we compare Shin's results [62] with ours. This would explain how using a single pseudo-coefficient violates a symmetry, and how the defect corresponds to the non-semisimple contributions on the geometric side and non-holomorphic endoscopic lifts on the spectral side.

Let $(l_1, l_2) = (k_1 - 1, k_2 - 2)$ be the Harish–Chandra parameter and let $D^{\mathrm{large}}_{l_1,l_2}$ be the large discrete series of $G(\mathbb{R}) = GSp_4(\mathbb{R})$ so that $\{D^{\mathrm{hol}}_{l_1,l_2}, D^{\mathrm{large}}_{l_1,l_2}\}$ makes up an L-packet of $\prod(G(\mathbb{R}))$ (see [71, § 2.3] for $D^{\mathrm{large}}_{l_1,l_2}$ and [46] for an interpretation as C^∞ classical forms). Note that irreducible components of $D^{\mathrm{hol}}_{l_1,l_2}|_{Sp_4(\mathbb{R})}$ and $D^{\mathrm{large}}_{l_1,l_2}|_{Sp_4(\mathbb{R})}$ form an L-packet of $\prod(Sp_4(\mathbb{R}))$ which consists of four elements.

Let $\widetilde{S}_{\underline{k}}(N)$ be the set introduced in Remark 1.9.

Suppose $\pi \in \widetilde{S}_{\underline{k}}(N)$ is not a CAP form, and non-endoscopic. Let $\pi = \pi_\infty \otimes \pi_f$. By Laumon [39, 40] and Weissauer [74], if $\pi_\infty \simeq D^{\mathrm{hol}}_{l_1,l_2}$, there exists a cuspidal representation $\pi' = \pi'_\infty \otimes \pi'_f$ such that $\pi'_\infty \simeq D^{\mathrm{large}}_{l_1,l_2}$, $\pi'_{f,p} \simeq \pi_{f,p}$ for any unramified p and π'_f has a non-zero $K_2(N)$-fixed vector. The converse is also true, namely for cuspidal representation π with $\pi_\infty \simeq D^{\mathrm{large}}_{l_1,l_2}$, there exists a cuspidal representation $\pi' = \pi'_\infty \otimes \pi'_f$ such that $\pi'_\infty \simeq D^{\mathrm{hol}}_{l_1,l_2}$, $\pi'_{f,p} \simeq \pi_{f,p}$ for any unramified p and π'_f has a non-zero $K_2(N)$-fixed vector.

Now suppose π is endoscopic and $\pi_\infty \simeq D^{\mathrm{hol}}_{l_1,l_2}$. Then by Roberts [50], there exists a cuspidal representation $\pi' = \pi'_\infty \otimes \pi'_f$ such that $\pi'_\infty \simeq D^{\mathrm{large}}_{l_1,l_2}$ and $\pi'_f \sim \pi_f$. (Here $\sim$ means weak equivalence, and in fact equivalent outside the ramification of π.)

However, if π is endoscopic and $\pi_\infty \simeq D^{\mathrm{large}}_{l_1,l_2}$, there does not necessarily exist a cuspidal representation π' such that $\pi'_\infty \simeq D^{\mathrm{large}}_{l_1,l_2}$ and $\pi'_f \sim \pi_f$. (For example, we cannot construct a holomorphic Siegel cusp form from a pair of two elliptic cusp forms of level one, but we can construct a cuspidal representation with the infinity type $D^{\text{large}}_{l_1,l_2}$ of level one.)

Therefore, holomorphic Siegel cusp forms always appear in pairs with cuspidal representations with the infinity type $D^{\text{large}}_{l_1,l_2}$, but there are cuspidal representations with the infinity type $D^{\text{large}}_{l_1,l_2}$, which do not appear in pairs. Let $\widetilde{S}_{\underline{k}}(N)^{\text{en,large}}$ be the subset of $\widetilde{S}_{\underline{k}}(N)$ consisting of Π such that Π is endoscopic and Π_∞ is isomorphic to the large discrete series $D^{\text{large}}_{l_1,l_2}$. Then the same argument in § 3.4 works for the theta lift from $GSO(2,2)$ to GSp_4 and we have

$$\dim \widetilde{S}_{\underline{k}}(N)^{\text{en,large}} = O((k_1-k_2+1)(k_1+k_2-3)N^{8+\epsilon}), \quad \text{as } k_1+k_2+N\to\infty.$$

Therefore,

$$\frac{\dim \widetilde{S}_{\underline{k}}(N)^{\text{en,large}}}{\dim S_{\underline{k}}(N)} = O(((k_1-1)(k_1-2))^{-1}N^{-2+\epsilon}), \quad \text{as } k_1+k_2+N\to\infty,$$

which might be related to the second main term A, B_1 of Theorem 1.1.

For $* \in \{\text{hol}, \text{large}\}$ and each $D^*_{l_1,l_2}$, we choose a pseudo-coefficient $f^*_{\xi_{\underline{k}}} \in C^\infty_c(G(\mathbb{R}))$. Put $f^{\text{tot}}_{\xi_{\underline{k}}} := f^{\text{hol}}_{\xi_{\underline{k}}} + f^{\text{large}}_{\xi_{\underline{k}}}$, where we may call it 'stable' pseudo-coefficient. (This is called Euler–Poincaré function in [62].) Note that if we work on Sp_4, we would consider $f^{\text{tot}}_{\xi_{\underline{k}}} := f^{\text{hol}}_{\xi_{\underline{k}}} + f^{\text{large}}_{\xi_{\underline{k}}} + f^{\text{anti-large}}_{\xi_{\underline{k}}} + f^{\text{anti-hol}}_{\xi_{\underline{k}}}$.

As Shin [62] did, by using $f^{\text{tot}}_{\xi_{\underline{k}}}$ in the Arthur–Selberg trace formula, we can avoid the non-semisimple contributions. However the trace $\text{tr}(f^{\text{tot}}_{\xi_{\underline{k}}})$ collects various automorphic forms both holomorphic cusp forms and non-holomorphic cusp forms. On the other hand in this paper we used a single pseudo-coefficient $f^{\text{hol}}_{\xi_{\underline{k}}}$ to collect only holomorphic cusp forms. Such a pseudo-coefficient might be called 'unstable'. Thereby we had to calculate non-semisimple contributions whose behaviors have not been understood well.

Let $\widehat{\mu}_{\text{Shin}}$ be the measure for $U = K(N)$ introduced in [62]. As in § 3.6, we can define, by using a pseudo-coefficient of $D^{\text{large}}_{l_1,l_2}$ (cf. [71, § 2.3]), the counting measure $\widehat{\mu}^{\text{en}}_{K(N),\xi_{\underline{k}},D^{\text{large}}_{l_1,l_2}}$ on $\widetilde{S}_{\underline{k}}(N)^{\text{en,large}}$. We also define the counting measure $\widehat{\mu}^{P_1}_{K(N),\xi_{\underline{k}},D^{\text{hol}}_{l_1,l_2}}$ for CAP forms associated to P_1.

It is not difficult to estimate non-holomorphic residual spectrum part as in § 3. Then we would have the following: for any $f = f_S$ in Proposition 5.1, the difference

$$\widehat{\mu}_{\text{Shin}}(\widehat{f}) - 2\big(\widehat{\mu}_{K(N),\xi_{\underline{k}},D^{\text{hol}}_{l_1,l_2}}(\widehat{f}) - \widehat{\mu}^{P_1}_{K(N),\xi_{\underline{k}},D^{\text{hol}}_{l_1,l_2}}(\widehat{f})\big) = \widehat{\mu}^{en}_{K(N),\xi_{\underline{k}},D^{\text{large}}_{l_1,l_2}}(\widehat{f}) + (\text{remainder})$$

would be

(1) (level aspect) $A + O(p_{S'}^{a\kappa+b}\varphi(N)N^{-3})$, as $N\to\infty$;

(2) (weight aspect) $B_1 + B_2 + O(\frac{p_{S'}^{a'\kappa+b'}}{(k_1-k_2+1)(k_1-1)(k_2-2)})$, as $k_1+k_2\to\infty$;

where A, B_1, B_2 are the second main terms in Theorem 1.1. Note also that the remainder in RHS comes from CAP representations and residual spectrum, and it would be subsumed into the error term.

Hence we expect that the second terms A, B_1 correspond to $S_{\underline{k}}(\Gamma(N))^{\mathrm{en,large}}$. However, B_2 is still mysterious and it seems interesting to figure out what kind of representations contribute to B_2. We confirm the above speculation elsewhere.

Acknowledgements. We would like to thank J. Arthur, P.-S. Chan, M. Miyauchi, R. Schmidt, S.-W. Shin and N. Templier for helpful discussions. We also thank S. Sugiyama for pointing out an error in the condition of Corollary 1.4 in the previous version. This work started when the authors visited RIMS in Kyoto in February 2015. We also discussed at KIAS in June 2015 and University of Toronto in September 2015. We thank these institutes for their incredible hospitality. We thank the referee for helpful suggestions, especially the correct definition of adelic forms in § 2.3.

Appendix.

In this appendix, we estimate the sum in (10.5) without any assumptions, namely,

$$\sum_p \sum_{l\geqslant 3} \frac{|b_F(p^l)|\log p}{p^{l/2}},$$

converges if F is not a CAP form. In [63], it was proved under the Langlands functoriality for all L-group homomorphism $r : {}^LG \longrightarrow GL_d(\mathbb{C})$ (Hypothesis 10.1). (Note that in [63], the sum in question is (12.21). It follows from (12.13). It is where the functoriality is needed.) In fact, it was also proved in [15, p. 7867] that the functoriality of the exterior square for GL_5 implies the convergence of the sum.

Recall the bound on the Satake parameters [42]: Let π be a cuspidal representation of GL_m, and let $\{\alpha_1(p), \dots, \alpha_m(p)\}$ be Satake parameters. Then $|\alpha_i(p)| \leqslant p^{1/2-1/(m^2+1)}$.

Hence $|b_F(p^l)| \leqslant 5p^{l/2-l/26}$. So

$$\sum_p \sum_{l\geqslant 52} \frac{|b_F(p^l)|\log p}{p^{l/2}} \ll \sum_p \log p \sum_{l\geqslant 52} (p^{-1/26})^l \ll \sum_p \frac{\log p}{p^2} = O(1).$$

Hence it is enough to prove that for each $l \geqslant 3$, the series $\sum_p \frac{|b_F(p^l)|\log p}{p^{l/2}}$ converges.

Recall the classification of spherical generic unitary representations of $GSp_4(\mathbb{Q}_p)$:

(1) $L(\mu_1, \mu_2, \eta)$; μ_1, μ_2, η are unitary characters;
(2) $L(\nu^\beta\mu, \nu^\beta\mu^{-1}, \nu^{-\beta}\eta)$; μ, η are unitary characters and $\mu^2 \neq 1$ and $0 < \beta < \frac{1}{2}$;
(3) $L(\nu^\beta, \mu, \nu^{-\beta/2}\eta)$; $\mu \neq 1, \nu$ are unitary characters and $0 < \beta < 1$;
(4) $L(\nu^{\beta_1}\mu, \nu^{\beta_2}\mu, \nu^{-(\beta_1+\beta_2)/2}\eta)$; μ, η are unitary characters and $\chi^2 = 1$ and $0 < \beta_2 \leqslant \beta_1 < 1$, $\beta_1 + \beta_2 < 1$.

Hence Satake parameters of Π_p are of the form

(1) $S_1 : 1, \alpha_{1p}, \alpha_{2p}, \alpha_{1p}^{-1}, \alpha_{2p}^{-1}$, where $|\alpha_{ip}| = 1$;
(2) $S_2 : 1, p^\beta\alpha_p, p^\beta\alpha_p^{-1}, p^{-\beta}\alpha_p, p^{-\beta}\alpha_p^{-1}$, where $|\alpha_p| = 1$;
(3) $S_3 : 1, p^\beta, p^{-\beta}, \alpha_p, \alpha_p^{-1}$, where $|\alpha_p| = 1$;
(4) $S_4 : 1, p^{\beta_1}\alpha_p, p^{\beta_2}\alpha_p, p^{-\beta_1}\alpha_p^{-1}, p^{-\beta_2}\alpha_p^{-1}$, where $|\alpha_p| = 1$.

Clearly, $\sum_{p\in S_1} \frac{|b_F(p^l)|\log p}{p^{l/2}}$ converges.

For S_2, note that

$$|b_F(p^l)| = |1+(p^{l\beta}+p^{-l\beta})(\alpha_p^l+\alpha_p^{-l})| \leqslant 2p^{l\beta}+3,$$

and $|a_F(p)| = |p^\beta + p^{-\beta} + \alpha_p + \alpha_p^{-1}| \geqslant p^\beta - 3$. Hence

$$|b_F(p^l)| \ll |a_F(p)|^l \ll |a_F(p)|^2 p^{(l-2)/2-(l-2)/17}.$$

Therefore,

$$\sum_{p\in S_2} \frac{|b_F(p^l)|\log p}{p^{l/2}} \ll \sum_p \frac{|a_F(p)|^2}{p^{1+(l-2)/17}}.$$

Since $L(s,\pi\times\pi)$ converges absolutely for $\mathrm{Re}(s)>1$, the above series converges.

For S_3, note that

$$1+p^\beta+p^{-\beta}+|\alpha_p|+|\alpha_p^{-1}| \leqslant |1+p^\beta+p^{-\beta}+\alpha_p+\alpha_p^{-1})|+6.$$

Hence $|b_F(p^l)| \ll |b_F(p)|^l \ll |b_F(p)|^2 p^{(l-2)/2-(l-2)/26}$. So

$$\sum_{p\in S_3} \frac{|b_F(p^l)|\log p}{p^{l/2}} \ll \sum_p \frac{|b_F(p)|^2}{p^{1+(l-2)/26}}.$$

Since $L(s,\Pi\times\Pi)$ converges absolutely for $\mathrm{Re}(s)>1$, the above series converges.

For S_4, note that

$$1+p^{\beta_1}+p^{-\beta_1}+p^{\beta_2}+p^{-\beta_2} \leqslant |1+p^{\beta_1}\alpha_p+p^{\beta_2}\alpha_p+p^{-\beta_1}\alpha_p^{-1}+p^{-\beta_2}\alpha_p^{-1}|+6.$$

Hence $|b_F(p^l)| \ll |b_F(p)|^l$, and it is similar to S_2 case.

The above proof shows also that Hypothesis H in [54] for $L(s,\pi_F,\mathrm{St})$ is satisfied, namely, for each $l\geqslant 2$,

$$\sum_p \frac{|b_F(p^l)|^2(\log p)^2}{p^l}$$

converges.

References

1. T. Arakawa, Vector-valued Siegel's modular forms of degree two and the associated Andrianov L-functions, *Manuscripta Math.* **44**(1–3) (1983), 155–185.
2. J. Arthur, The L^2-Lefschetz numbers of Hecke operators, *Invent. Math.* **97** (1989), 257–290.
3. J. Arthur, *An Introduction to the Trace Formula, Harmonic Analysis, the Trace Formula, and Shimura Varieties*, Clay Mathematics Proceedings, Volume 4, pp. 1–263 (American Mathematical Society, Providence, RI, 2005).
4. J. Arthur, On elliptic tempered characters, *Acta Math.* **171** (1993), 73–138.
5. J. Arthur, On the Fourier transforms of weighted orbital integrals, *J. Reine Angew. Math.* **452** (1994), 163–217.

6. J. Arthur, Automorphic representations of $GSp(4)$, in *Contributions to Automorphic Forms, Geometry, and Number Theory*, pp. 65–81 (Johns Hopkins University Press, Baltimore, 2004).
7. M. Assem, Unipotent orbital integrals of spherical functions on p-adic 4×4 symplectic groups, *J. Reine Angew. Math.* **437** (1993), 181–216.
8. A. Borel and H. Jacquet, *Automorphic Forms and Automorphic Representations*, Proceedings of Symposia in Pure Mathematics, Volume XXXIII, (Part 1) pp. 189–207 (American Mathematical Society, Providence, RI, 1977).
9. A. Borel and N. Wallach, *Continuous Cohomology, Discrete Subgroups, and Representations of Reductive Groups*, second edition, Mathematical Surveys and Monographs, Volume 67 (American Mathematical Society, Providence, RI, 2000).
10. M. Bozicević, A limit formula for elliptic orbital integrals, *Duke Math. J.* **113** (2002), 331–353.
11. C. J. Bushnell and G. Henniart, An upper bound on conductors for pairs, *J. Number Theory* **65**(2) (1997), 183–196.
12. F. Calegari and T. Gee, Irreducibility of automorphic Galois representations, *Ann. Inst. Fourier* **63**(5) (2013), 1881–1912.
13. P.-S. Chan and W. T. Gan, The local Langlands conjecture for *GSp*(4) III: Stability and twisted endoscopy, *J. Number Theory* **146** (2015), 69–133.
14. P. J. Cho and H. H. Kim, Low lying zeros of Artin L-functions, *Math. Z.* **279** (2015), 669–688.
15. P. J. Cho and H. H. Kim, n-level densities of Artin L-functions, *Int. Math. Res. Not. IMRN* (17) (2015), 7861–7883.
16. L. Clozel, On limit multiplicities of discrete series representations in spaces of automorphic forms, *Invent. Math.* **83** (1986), 265–284.
17. L. Clozel and P. Delorme, Le theoreme de Paley–Wiener invariant pour les groupes de Lie reductifs II, *Ann. Sci. Éc. Norm. Supér. (4)* **23** (1990), 193–228.
18. M. Dickson, Local spectral equidistribution for degree two Siegel modular forms in level and weight aspects, *Int. J. Number Theory* **11** (2015), 341–396.
19. S. A. Evdokimov, Euler products for congruence subgroups of the Siegel group of genus 2, *Math. USSR Sb.* **28** (1976), 431–458.
20. A. Ferrari, Théorème de l'Indice et Formule des Traces, *Manuscripta Math.* **124** (2007), 363–390.
21. W. Fulton and J. Harris, *Representation Theory, A First Course* (Springer-Verlag, New York, 1991).
22. W. T. Gan, The Saito–Kurokawa space of $PGSp_4$ and its transfer to inner forms, in *Eisenstein Series and Applications*, Progress in Mathematics, Volume 258, pp. 87–123 (Birkhäuser Boston, 2008).
23. W. T. Gan and S. Takeda, The local Langlands conjecture for $GSp(4)$, *Ann. of Math. (2)* **173**(3) (2011), 1841–1882.
24. W. T. Gan and S. Takeda, Theta correspondences for GSp(4), *Represent. Theory* **15** (2011), 670–718.
25. G. van der Geer, *Siegel Modular Forms and their Applications*, The 1-2-3 of modular forms, Universitext, pp. 181–245 (Springer, Berlin, 2008).
26. H. Hecht, The characters of some representations of Harish–Chandra, *Math. Ann.* **219** (1976), 213–226.
27. K. Hiraga, On the multiplicities of the discrete series of semisimple Lie groups, *Duke Math. J.* **85** (1996), 167–181.

28. T. Hirai, Explicit form of the characters of discrete series representations of semisimple Lie groups, in *Harmonic Analysis on Homogeneous Spaces*, Proceedings of Symposia in Pure Mathematics, Volume 26, pp. 281–288 (American Mathematical Society, Providence, RI, 1972).
29. W. Hoffmann and S. Wakatsuki, On the geometric side of the Arthur trace formula for the symplectic group of rank 2, *Mem. Amer. Math. Soc.* (to appear).
30. T. Ibukiyama, Dimension formulas of Siegel modular forms of weight 3 and supersingular abelian varieties, in *Proceedings of the 4th Spring Conference on Modular Forms and Related Topics, Siegel Modular Forms and Abelian Varieties* (ed. T. Ibukiyama), pp. 39–60. (2007).
31. D. Jiang and D. Soudry, The multiplicity-one theorem for generic automorphic forms on $GSp(4)$, *Pacific J. Math.* **229** (2007), 381–388.
32. N. Katz and P. Sarnak, *Random Matrices, Frobenius Eigenvalues and Monodromy*, Amer. Math. Soc. Colloq. Publ., Volume 45, pp. 1–427 (American Mathematical Society, Providence, RI, 1999).
33. H. H. Kim, Residual spectrum of odd orthogonal groups, *Int. Math. Res. Not. IMRN* (17) (2001), 873–906.
34. H. H. Kim and T. Yamauchi, A conditional construction of Artin representations for real analytic Siegel cusp forms of weight $(2, 1)$, in *Advances in the Theory of Automorphic Forms and their L-Functions*, Contemporary Mathematics, Volume 664, pp. 225–260 (American Mathematical Society, Providence, RI, 2016).
35. A. W. Knapp, Representation theory of semisimple groups, An overview based on examples, in *Reprint of the 1986 original, Princeton Landmarks in Mathematics* (Princeton University Press, Princeton, NJ, 2001).
36. E. Kowalski, A. Saha and J. Tsimerman, Local spectral equidistribution for Siegel modular forms and applications, *Compos. Math.* **148** (2012), 335–384.
37. R. P. Langlands, The dimension of spaces of automorphic forms, *Amer. J. Math.* **85** (1963), 99–125.
38. J. Lansky and A. Raghuram, On the correspondence of reps between $GL(n)$ and division algebras, *Proc. Amer. Math. Soc.* **131** (2003), 1641–1648.
39. G. Laumon, Sur la cohomologie a supports compacts des varietes de Shimura pour $GSp(4)_{\mathbb{Q}}$, *Compos. Math.* **105**(3) (1997), 267–359.
40. G. Laumon, Fonctions zetas des varietes de Siegel de dimension trois, Formes automorphes. II. Le cas du groupe GSp(4), *Astérisque* **302** (2005), 1–66.
41. X. Li, Upper bounds on L-functions at the edge of the critical strip, *Int. Math. Res. Not. IMRN* (4) (2010), 727–755.
42. W. Luo, Z. Rudnick and P. Sarnak, *On the Generalized Ramanujan Conjecture for* $GL(n)$, Proceedings of Symposia in Pure Mathematics, Volume 66, pp. 301–310 (American Mathematical Society, Providence, RI, 1999). Part 2.
43. S. Martens, The characters of the holomorphic discrete series, *Proc. Natl. Acad. Sci. USA* **72** (1975), 3275–3276.
44. S. Marshall, Endoscopy and cohomology growth on $U(3)$, *Compos. Math.* **150** (2014), 903–910.
45. M. Miyauchi and T. Yamauchi, An explicit computation of p-stabilized vectors, *J. Théor. Nombres Bordeaux* **26** (2014), 531–558.
46. T. Oda, Cohomology of Siegel modular varieties of genus 2 and corresponding automorphic forms, in *Geometry and Analysis of Automorphic Forms of Several Variables*, pp. 211–253 (World Sci. Publ., Hackensack, NJ, 2012).
47. T. Okazaki and T. Yamauchi, On some Siegel threefold related to the tangent cone of the Fermat quartic surface, *Adv. Theor. Math. Phys.* **21**(3) (2017), 585–630.

48. I. I. Piatetski-Shapiro, On the Saito–Kurokawa lifting, *Invent. Math.* **71** (1983), 309–338.
49. A. Pitale, A. Saha and R. Schmidt, Lowest weight modules of $Sp_4(\mathbb{R})$ and nearly holomorphic Siegel modular forms, preprint, arXiv:1501.00524 (2015).
50. B. Roberts, Global L-packets for $GSp(2)$ and theta lifts, *Doc. Math.* **6** (2001), 247–314.
51. B. Roberts and R. Schmidt, Local Newforms for $GSp(4)$, *Lecture Notes in Math.* **1918** (2007).
52. B. Roberts and R. Schmidt, Some results on Bessel functionals for $GSp(4)$, *Doc. Math.* **21** (2016), 467–553.
53. W. Rossmann, *Nilpotent Orbital Integrals in a Real Semisimple Lie Algebra and Representations of Weyl Groups*, Progress in Mathematics, Volume 92, pp. 263–287 (Birkhäuser, Boston, 1990).
54. Z. Rudnick and P. Sarnak, Zeros of principal L-functions and random matrix theory, *Duke Math. J.* **81** (1996), 269–322.
55. R. Salvati Manni and J. Top, Cusp forms of weight 2 for the group $\Gamma(4, 8)$, *Amer. J. Math.* **115** (1993), 455–486.
56. A. Saha, On ratios of Petersson norms for Yoshida lifts, *Forum Math.* **27** (2015), 2361–2412.
57. P. Sarnak, Statistical properties of eigenvalues of the Hecke operators, in *Analytic Number Theory and Diophantine Problems (Stillwater, OK), 1984*, Progress in Mathematics, Volume 70, pp. 321–331 (Birkhäuser Boston, Boston, MA, 1987).
58. P. Sarnak, S. W. Shin and N. Templier, Families of L-functions and their symmetry, in *Families of Automorphic Forms and the Trace Formula. Proceedings of the Simons Symposium, Puerto Rio*, pp. 531–578 (Springer, 2016).
59. F. Sauvageot, Principe de densité pour les groupes réductifs à Solène, pour son premier sourire, *Compos. Math.* **108** (1997), 151–184.
60. R. Schmidt, The Saito–Kurokawa lifting and functoriality, *Amer. J. Math.* **127**(1) (2005), 209–240.
61. J.-P. Serre, Répartition asymptotique des valeurs propres de l'opérateur de Hecke T_p, *J. Amer. Math. Soc.* **10**(1) (1997), 75–102.
62. S. W. Shin, Automorphic Plancherel density theorem, *Israel J. Math.* **192**(1) (2012), 83–120.
63. S. W. Shin and N. Templier, Sato-Tate theorem for families and low-lying zeros of automorphic L-functions, *Invent. Math.* **203**(1) (2016), 1–177.
64. S. W. Shin and N. Templier, On fields of rationality for automorphic representations, *Compos. Math.* **150**(12) (2014), 2003–2053.
65. C. M. Sorensen, Galois representations attached to Hilbert–Siegel modular forms, *Doc. Math.* **15** (2010), 623–670.
66. D. Soudry, The CAP representations of $GSp(4, \mathbb{A})$, *J. Reine Angew. Math.* **383** (1988), 87–108.
67. R. Takloo-Bighash, Some results on L-functions for the similitude symplectic group of order four $GSp(4)$, in *Siavash Shahshahani's Sixtieth* (ed. K. Lajevardi, P. Safari and Y. Tabesh), pp. 105–122. (2002).
68. R. Taylor, On congruences of modular forms, Thesis (1988).
69. R. Tsushima, An explicit dimension formula for the spaces of generalized automorphic forms with respect to $Sp(2, \mathbb{Z})$, *Proc. Japan Acad. Ser. A Math. Sci.* **59**(4) (1983), 139–142.
70. S. Wakatsuki, Dimension formulas for spaces of vector-valued Siegel cusp forms of degree two, *J. Number Theory* **132** (2012), 200–253.
71. S. Wakatsuki, Multiplicity formulas for discrete series representations in $L^2(\Gamma\backslash \mathrm{Sp}(2, \mathbb{R}))$, *J. Number Theory* **133** (2013), 3394–3425.

72. N. R. Wallach, On the constant term of a square integrable automorphic form, *Operator Algebras and Group Representations, Vol. II (Neptun, 1980)*, Monogr. Stud. Math., Volume 18, pp. 227–237 (Pitman, Boston, MA, 1984).
73. S. Wang, An effective version of the Grunwald–Wang theorem, Thesis at Caltech (2002).
74. R. Weissauer, On the cohomology of Siegel modular threefolds, in *Arithmetic of Complex Manifolds*, Lecture Notes in Mathematics, Volume 1399, pp. 155–171 (Springer, Berlin, 1989).
75. R. Weissauer, Modular forms of genus 2 and weight 1, *Math. Z.* **210** (1992), 91–96.