Acta Math., 167 (1991), 229–285

Poincaré series for $SO(n, 1)$

by

J. COGDELL([1])

Oklahoma State University
Stillwater, OK, U.S.A.

J.-S. LI

University of Maryland
College Park, MD, U.S.A.

I. PIATETSKI-SHAPIRO

Yale University
New Haven, CT, U.S.A.
&
Tel Aviv University
Tel Aviv, Israel

P. SARNAK

Stanford University
Stanford, CA, U.S.A.

§ 1. Introduction

The theory of holomorphic Poincaré series has been developed and studied quite generally [29]. It allows one to construct explicitly in terms of infinite series, holomorphic cusp forms. "Nonholomorphic" Poincaré series were introduced by Selberg [27] for $SL_2(\mathbf{R})$. If K is a maximal compact subgroup of $SL_2(\mathbf{R})$ and Γ a nonuniform lattice in $SL_2(\mathbf{R})$ then by expanding these series once spectrally in $L^2(\Gamma\backslash SL_2(\mathbf{R})/K)$ and once directly in a Fourier series in a cusp, one obtains a relation between the L^2 spectrum and sums of Kloosterman sums. In this way using bounds on Kloosterman sums due to Weil [32], Selberg established the well known estimate

$$\lambda_1 \geq \frac{3}{16} \tag{1.1}$$

for the second smallest eigenvalue of the Laplacian for any congruence subgroup of $SL_2(\mathbf{Z})$. The bound (1.1) above goes part of the way towards the "Ramanujan Conjec-

([1]) The work of the 1st author was supported in part by the NSF under the Grant No. DMS-88-03085, the 3rd author by the NSF Grant No. DMS-88-07336 and the 4th author by the NSF Grant No. DMS-84-51759.

ture'' $\lambda_1 \geqslant 1/4$, which remains a basic unsolved problem. See Iwaniec [12] for some recent progress.

The above relation and set up has a number of striking applications [11, 13] which we do not enter into here. Suffice it to say that it is desirable to develop such a theory for more general groups. For GL_n, $n \geqslant 3$, computations have been done by Bump–Friedberg–Goldfeld [2] and Stevens [30]. However it appears that the direct estimation of the resulting exponential sums does not produce results better than what follows from quantitative versions of property T, see Jacquet–Shalika [14] (and these have nothing to do with arithmetic!).

In this paper we develop a theory of Poincaré series for $SO(n, 1)$. The main applications of the theory are to the meromorphic continuation to $\mathbf{C}$ of the 'Kloosterman–Selberg' zeta function for general $\Gamma \leqslant G$ and the analogue of (1.1) for congruence subgroups of isotropic unit groups of rational quadratic forms. These results were announced in [18]. Some related results and in particular Theorem 4.10 below have also been announced by Elstrodt, Grunewald and Mennicke [5].

A general theory of Poincaré series on real reductive groups of rank 1 has been developed by Miatello and Wallach [20]. Their theory is slightly different from the one developed here. In particular, their Poincaré series are not L^2 while it is essential for our purposes that our Poincaré series are L^2. Their proof of analytic continuation involves a detailed analysis of the special functions involved. Our proof of a similar statement on the continuation of the Kloosterman–Selberg zeta function involves a shift equation which was originally suggested by Selberg. In order to apply this type of argument to our situation we use ideas suggested by J. Bernstein.

We now turn to a more precise description of our results as well as introduce the notation to be used in this paper. Let G be the real special orthogonal group of a form of signature $(r+1, 1)$ with $r \geqslant 2$, i.e. $G = SO(r+1, 1)$. We may realize G as

$$G = \left\{ g \in SL_{r+2} \,\middle|\, {}^t g \begin{pmatrix} & & 1 \\ & \mathbf{1}_r & \\ 1 & & \end{pmatrix} g = \begin{pmatrix} & & 1 \\ & \mathbf{1}_r & \\ 1 & & \end{pmatrix} \right\}, \tag{1.2}$$

where $\mathbf{1}_r$ is the $r \times r$ identity matrix. Denote by A, H, U the following subgroups

$$A = \left\{ \begin{pmatrix} a & & \\ & \mathbf{1}_r & \\ & & a^{-1} \end{pmatrix} \,\middle|\, a \in \mathbf{R}^* \right\}. \tag{1.3}$$

$H = SO(r)$ which we embedded in G via

$$h\to\begin{pmatrix}1&&\\&h&\\&&1\end{pmatrix} \tag{1.4}$$

$$U=\left\{\begin{pmatrix}1&-{}^t u&-1/2(u,u)\\&\mathbf{1}_r&u\\&&1\end{pmatrix}\middle|\, u\in\mathbf{R}^r\right\}. \tag{1.5}$$

Here u is a column vector and (,) is the usual inner product on $\mathbf{R}^r$. With this notation

$$P=UAH \tag{1.6}$$

is a parabolic subgroup of G with unipotent radical $U\cong\mathbf{R}^r$ and Levi component

$$M=AH. \tag{1.7}$$

Let $\Gamma\leqslant G$ be a discrete subgroup of finite co-volume but not co-compact. We may take one of its cusps to correspond to U i.e.

$$\Gamma^\infty=\Gamma\cap U \tag{1.8}$$

is a full rank lattice in U. For notational simplicity assume that in fact $\Gamma\cap P=\Gamma^\infty$ and that this is the only cuspidal subgroup of Γ (in general there may be a finite number of nonconjugate such subgroups). Fix nontrivial unitary characters ψ and η of U, trivial on Γ^∞. The Bruhat decomposition of G asserts that

$$G=P\cup PwU \tag{1.9}$$

where

$$w=\begin{pmatrix}&&+1\\&\varepsilon_r&\\1&&\end{pmatrix}\quad\text{and}\quad \varepsilon_r=\begin{pmatrix}-1&\\&\mathbf{1}_{r-1}\end{pmatrix}.$$

Hence we may write any $\gamma\in\Gamma$, $\gamma\notin P$ uniquely as

$$\gamma=u(\gamma)\,a(\gamma)\,h(\gamma)\,wv(\gamma)$$

or in terms of M as

$$\gamma=u(\gamma)\,m(\gamma)\,wv(\gamma) \tag{1.10}$$

with the obvious notations concerning membership. Let

$$M(\Gamma)=\{m(\gamma)|\gamma\in\Gamma,\gamma\notin\Gamma^{\infty}\}. \tag{1.11}$$

One may choose a set of representatives for $\Gamma^{\infty}\backslash\Gamma/\Gamma^{\infty}$, $\gamma\notin\Gamma^{\infty}$, of the form

$$umwv \tag{1.12}$$

where for each $m\in M(\Gamma)$, u and v run over a finite set in $U\times U$. The 'Kloosterman' sum $\mathrm{Kl}(\psi,\eta,m)$ for $m\in M(\Gamma)$ is defined by

$$\mathrm{Kl}(\psi,\eta,m)=\sum\psi(u)\,\eta(v) \tag{1.13}$$

where the summation ranges over the finite set of u's and v's above. The sum is clearly independent of the choice of these representatives. Let τ be an irreducible unitary representation of H, we define the (matrix valued) Kloosterman–Selberg zeta function by

$$Z(\psi,\eta,\tau,s)=\sum_{m\in M(\Gamma)}\mathrm{Kl}(\psi,\eta,m)\,\tau(h(m))\,|a(m)|^{s+r/2}. \tag{1.14}$$

We will see that the series above converges absolutely for $\mathrm{Re}(s)>r/2$ and hence $Z(\psi,\eta,\tau,s)$ is holomorphic in this region. Our first result is the meromorphic continuation of this function to the complex plane. To describe the location of poles we need to decompose the right regular representation of G on $L^2(\Gamma\backslash G)$. $L^2(\Gamma\backslash G)$ breaks up into the orthogonal sum of invariant spaces

$$L^2(\Gamma\backslash G)=L^2_{\mathrm{disc}}(\Gamma\backslash G)\oplus L^2_{\mathrm{cont}}(\Gamma\backslash G). \tag{1.15}$$

Here $L^2_{\mathrm{cont}}(\Gamma\backslash G)$ is spanned by unitary Eisenstein series while its orthogonal complement $L^2_{\mathrm{disc}}(\Gamma\backslash G)$ decomposes into a countable direct sum of irreducible representations of G. The representations of most relevance here are the principal and complementary series. The principal series of the form

$$\mathrm{Ind}_P^G(\sigma\otimes|\ |^{\varrho}\,\mathrm{sgn}^{\varepsilon}) \tag{1.16}$$

where $\varepsilon=0$ or 1, $\sigma\in\hat{H}$, $\mathrm{Re}(\varrho)=0$, and $|\ |^{\varrho}\,\mathrm{sgn}^{\varepsilon}$ is a character of A. The complementary series are of the same form except that $\sigma=$identity and $-r/2<\varrho\leq r/2$. Let R denote the standard representation of $H=SO(r)$ in $\mathbf{R}^r$ and for any $\sigma\in\hat{H}$ let $\sigma^{\iota}(h)=\sigma(\varepsilon_r h\varepsilon_r)$.

THEOREM 2.16. *$Z(\psi,\eta,\tau,s)$ has a meromorphic continuation to* $\mathbf{C}$ *with the set of poles contained in*

(1) $-k$, $k\geq 0$, $k\in\mathbf{Z}$.

(2) $\varrho-k$, $k\geq 0$ *an integer and where* $\pi\subset\mathrm{Ind}_P^G(\sigma\otimes|\ |^{\varrho}\operatorname{sgn}^{\varepsilon})$ *is a constituent of* $L^2_{\mathrm{disc}}(\Gamma\backslash G)$ and *also* τ *is a constituent of* $\sigma^{\imath}\otimes R^k$.

(3) $\varrho-k$, $k\geq 0$ *an integer and* ϱ *is a pole of* $E(\sigma,g,s)$ (*the Eisenstein series, see* § 2.6) *and also* τ *is a constituent of* $\sigma^{\imath}\otimes R^k$.

The set of possible poles above is in fact a discrete set. Moreover the question of whether a given number in this set is a pole of Z is equivalent to the vanishing of a certain Fourier coefficient, see § 2.

The proof of Theorem 2.16 is based on the following Poincaré series. Let b_1, b_2 be Schwartz functions on U and M respectively. Define the function f on G by

$$\begin{cases} f(u_1wu_2m)=\psi(u_1)\,b_1(u_2)\,b_2(m)\,\tau(h) \\ f(g)=0 \quad \text{if } g \text{ is not of the above form.}\end{cases} \tag{1.17}$$

The Poincaré series $P_f(g)$ on $\Gamma\backslash G$ is defined by

$$P_f(g)=\sum_{\Gamma^{\infty}\backslash\Gamma} f(\gamma g). \tag{1.18}$$

The spectral decomposition and Fourier development of P_f as well as the proof of Theorem 2.16 are carried out in Section 2. The zeta function is obtained as a Mellin transform of P_f along a suitable subgroup.

The most interesting case of the zeta functions $Z(\psi,\eta,\tau,s)$ is that of τ being the identity representation. In this case there may well be poles of Z in $\mathrm{Re}(s)>r/2-1$. These poles correspond to complementary series occurrences of the right regular representation in $L^2_{\mathrm{disc}}(\Gamma\backslash G)$. The duality theorem [6], which we will often use, asserts that the multiplicity of the constituent $\mathrm{Ind}_P^G(1\otimes|\ |^{\varrho})$ in $L^2_{\mathrm{disc}}(\Gamma\backslash G)$ is precisely the multiplicity of the eigenvalue $\lambda=(r/2-\varrho)(r/2+\varrho)$ of the Laplacian Δ on the hyperbolic space $\Gamma\backslash H^{r+1}\cong\Gamma\backslash G/K$, $K=SO(r+1)$. In this way the poles of Z above correspond to eigenvalues $0<\lambda<r-1$.

More generally we call poles of $Z(\psi,\eta,1,s)$ in $0<\mathrm{Re}(s)<r/2$, or equivalently eigenvalues λ of Δ in $(0,(r/2)^2)$, exceptional spectrum. In Section 6 it is shown that exceptional spectrum and even poles arbitrarily close to $r/2$ may occur for the general Γ. For Γ a congruence group the situation is much better.

For the rest of the paper we consider only congruence subgroups of unit groups of rational quadratic forms. To begin with we stick to the orthogonal group G in (1.2) except now we consider this group over other fields. The modifications needed to deal with the general quadratic form of signature $(r+1, 1)$ is straightforward.

Let

$$\Gamma = G \cap SL_{r+2}(\mathbf{Z}) \tag{1.19}$$

be the group of integral automorphs of

$$\begin{bmatrix} & & 1 \\ & \mathbf{1}_r & \\ 1 & & \end{bmatrix}.$$

It is a nonuniform lattice in G [3]. For $D \in \mathbf{Z}$ let $\Gamma(D)$ be the congruence subgroup defined by

$$\Gamma(D) = \{\gamma \in \Gamma \mid \gamma \equiv \mathbf{1} \pmod{D}\}. \tag{1.20}$$

Let $Y_D = \Gamma(D)\backslash G/K$ be the corresponding hyperbolic manifold. Our main result concerning the exceptional spectrum is the following

THEOREM 4.10. *Let $\lambda_1(Y_D)$ be the smallest nonzero eigenvalue of the Laplacian on functions on Y_D then for $r \geq 2$,*

$$\lambda_1(Y_D) \geq \frac{1}{2}\left(r - \frac{1}{2}\right).$$

The method of proof is to show that in these cases, $Z(\psi, \eta, \mathbf{1}, s)$ has no poles in $\mathrm{Re}(s) > (r-1)/2$. To pick up the arithmetic structure of $\Gamma(D)$, especially the local part, we find it both convenient and indispensable to work adelically. In particular we introduce adelic Poincaré series.

Let $G(\mathbf{A})$, $G(\mathbf{Q})$, $P(\mathbf{A})$, $P(\mathbf{Q})$, $U(\mathbf{A})$, $U(\mathbf{Q})$ ... denote the adelic, respectively rational points of the corresponding subgroups. Let $\psi = \Pi_p \psi_p$ be a standard character for $\mathbf{Q}\backslash\mathbf{A}$. For $\xi \in (1/D)\mathbf{Z}^r$ we define a character ψ^ξ of $U(\mathbf{Q})\backslash U(\mathbf{A})$ by

$$\psi^\xi(u) = \psi((u, \xi)), \quad u \in U(\mathbf{A}). \tag{1.21}$$

Clearly we may decompose ψ^ξ as $\psi^\xi = \Pi_p \psi_p^\xi$. To introduce the Poincaré series on $G(\mathbf{Q})\backslash G(\mathbf{A})$ we define the function $f^\xi(g)$, $g \in G(\mathbf{A})$ as follows:

(1) For $p = \infty$ we set

$$f_\infty^\xi(u\,y\,k) = \psi_\infty^\xi(u)\, y^{s+r/2} e^{-2\pi|\xi| y} \tag{1.22}$$

where $u \in U(\mathbf{R})$, $k \in K_\infty$ (the maximal compact)

$$y = \begin{pmatrix} y & & \\ & \mathbf{1}_r & \\ & & y^{-1} \end{pmatrix}$$

and uyk is the Iwasawa factorization of an element in $G(\mathbf{R})$.

(2) For $p<\infty$ let

$$K_p = \{g \in G(\mathbf{Z}_p) | \, g \equiv \mathbf{1} \mod p^l \text{ where } p^l \| D\}$$

then set

$$f_p^\xi(g) = \begin{cases} \psi_p^\xi(u) & \text{if} \quad g = uk, \ u \in U(\mathbf{Q}_p), \ k \in K_p \\ 0 & \text{if} \quad g \text{ is not of this form.} \end{cases} \tag{1.23}$$

Define

$$f^\xi(g) = \prod_p f_p^\xi(g_p) \quad \text{for} \quad g = (g_\infty, g_2, \ldots) \in G(\mathbf{A}). \tag{1.24}$$

Clearly f^ξ satisfies

$$f^\xi(ug) = \psi^\xi(u) f^\xi(g). \tag{1.25}$$

Finally the adelic Poincaré series is defined by

$$P_\xi(g,s) = \sum_{\gamma \in U(\mathbf{Q}) \backslash G(\mathbf{Q})} f^\xi(\gamma g). \tag{1.26}$$

The convergence and meromorphic continuation of this series is investigated in Sections 3 and 4. We note that these series and the "real" ones (1.18) are defined differently. We have found the above definition (1.26) most convenient for purposes of estimation and corresponding proof of holomorphicity of the zeta function (which is closely tied to P_ξ) in the region $\mathrm{Re}(s) > r/2 - 1/2$. The series in (1.18) or rather their Mellin transform over a certain subgroup, is most suitable in developing the meromorphic continuation of $Z(s)$ to *all* of $\mathbf{C}$.

In Section 3 we spectrally analyze $P_\xi(g,s)$. The analysis is similar to Section 2. The heart of the paper is Section 4 which consists of the computation of the Fourier coefficients of $P_\xi(g,s)$. These give rise to local Kloosterman integrals, see (4.4). These of course are special cases of the Kloosterman sums introduced earlier but the arithmetic is properly captured in this form. These exponential sums over finite fields are varied in type. Of theese one family reduces to the classical Kloosterman sum, see Lemma 4.6. To these we apply Weil's bound [32]. For the rest which are multidimensional sums we apply the "trivial bound". By a delicate analysis which requires a careful examination of those γ's for which $f^{\xi}(\gamma wu)\neq 0$, we show in Section 5 that $Z(s)$ is holomorphic in $\mathrm{Re}(s)>(r-1)/2$. This leads to the proof of Theorem 4.10.

The question as to where (if any) are the exceptional poles in $0<s\leq(r-1)/2$ for the manifolds Y_D, remains open. In this context the analogue of the Ramanujan Conjecture would assert $\lambda_1\geq(r/2)^2$ (i.e. no exceptional spectrum at all). This statement however is false. In Section 6 we give an explicit construction via theta-liftings from $SL(2)$ (or $\widetilde{SL(2)}$ depending on the parity of r), of exceptional spectrum for Y_D for D large. These exceptional poles occur at $s=r/2-1$, $r/2-3\ldots$. These "counter examples" to the Ramanujan conjecture are similar to those constructed by Howe and Piatetski-Shapiro [10].

Finally the case of $r=2$ has been treated in part by Sarnak [26] while adelic Poincaré series for $GL(2)$ were introduced by Piatetski-Shapiro [22].

§ 2. Analytic continuation

The goal of this section is to prove the meromorphic continuation of the Kloosterman–Selberg zeta function. The notation used is the same as in Section 1 (1.2).

2.1. Let π be a Banach representation of G, realized on the Banach space H. Let H_∞ be the space of smooth vectors. Denote by $U(\mathfrak{g})$ the universal enveloping algebra of $\mathfrak{g}$, the Lie algebra of G. We recall that a vector $\phi\in H$ is in H_∞ if and only if $D\cdot\phi\in H$ for all $D\in U(\mathfrak{g})$. The space H_∞ is endowed with its usual Fréchet topology.

Let λ_η be a continuous linear functional on H_∞, such that

$$\lambda_\eta(\pi(u)\phi)=\eta(u)\lambda_\eta(\phi) \tag{2.1}$$

for all $u\in U$, $\phi\in H_\infty$. Note that the character η may be represented as

$$\eta(u)=e((\alpha,u)) \tag{2.2}$$

for suitable $0\neq\alpha\in\mathbf{R}^r$. Set

$$W_\phi(g)=\lambda_\eta(\pi(g)\phi) \tag{2.3}$$

for $g\in G$ and $\phi\in H_\infty$.

LEMMA 2.1. *There is a finite collection $\{D_\alpha\}\subseteq U(\mathfrak{g})$ and an integer $k>0$ such that*

$$\left|W_\phi\begin{pmatrix}a & & \\ & h & \\ & & a^{-1}\end{pmatrix}\right|\leq(a^k+a^{-k})\Big(\sum_\alpha\|D_\alpha\cdot\phi\|\Big).$$

Proof. For an integer $\nu \geq 0$ let $U(\mathfrak{g})_\nu$ be the subspace of $U(\mathfrak{g})$ spanned by elements which are products of no more than ν elements of $\mathfrak{g}$. The fact that λ_η is continuous implies that there is some ν and a basis $\{D_i\}$ of the finite dimensional space $U(\mathfrak{g})_\nu$ such that

$$|W_\phi(1)| \leq \sum_i \|D_i \cdot \phi\|. \tag{2.4}$$

Let

$$g = \begin{pmatrix} a & & \\ & h & \\ & & a^{-1} \end{pmatrix}.$$

We know from Wallach [31, Lemma 2.2] that there exists $c>0$, $r \geq 0$ so that

$$|\pi(g)| \leq c(a^r + a^{-r}) \tag{2.5}$$

where $|\pi(g)|$ denotes the operator norm of $\pi(g)$. We have $W_\phi(g) = W_{\pi(g)\phi}(1)$ so

$$\begin{aligned} \left| W_\phi \begin{pmatrix} a & & \\ & h & \\ & & a^{-1} \end{pmatrix} \right| &\leq \sum_i \|D_i \cdot (\pi(g) \cdot \phi)\| \\ &= \sum_i \|\pi(g) \cdot \mathrm{Ad}(g)^{-1}(D_i) \cdot \phi\| \\ &\leq c(a^r + a^{-r}) \cdot \sum_i \|\mathrm{Ad}(g)^{-1}(D_i) \cdot \phi\| \end{aligned} \tag{2.6}$$

where Ad denotes the adjoint representation.

The subspace $U(\mathfrak{g})_\nu$ is stable under the adjoint action. Thus

$$\mathrm{Ad}(g)^{-1}(D_i) = \sum_j a_{ij}(g) D_j.$$

An explicit look at the adjoint action shows that the matrix coefficents $a_{ij}(g)$ satisfy the estimation

$$|a_{ij}(g)| \leq c_1(a^\nu + a^{-\nu})$$

with c_1 a constant. Combined with (2.6) we see that our lemma is valid for $k = r + \nu$ and $\{D_\alpha\}$ a constant multiple of the collection $\{D_i\}$. Q.E.D.

LEMMA 2.2. *For any integer $N>0$ there exists elements $D_1, \ldots, D_m \in U(\mathfrak{g})$ such that* *for $a \geqslant 1$ and $\phi \in H_\infty$*

$$\left| W_\phi \begin{pmatrix} a & & \\ & h & \\ & & a^{-1} \end{pmatrix} \right| \ll a^{-N} \left(\sum_{j=1}^{m} \|D_j \cdot \phi\| \right).$$

Proof. Let $e_1, \ldots, e_r$ be a basis of $\mathbf{R}^r$. The function

$$f(h) = \max_{1 \leqslant j \leqslant r} |(\alpha, he_j)|$$

is continuous and nowhere vanishing on H. Hence, H being compact, there is a positive constant c_0 such that

$$f(h) \geqslant c_0$$

for all $h \in H$.

Set

$$X_j = \begin{pmatrix} 0 & -{}^{t}e_j & 0 \\ 0 & 0 & e_j \\ 0 & 0 & 0 \end{pmatrix}, \quad j = 1, \ldots, r.$$

Given $h \in H$ there is at least one index j such that

$$|(\alpha, he_j)| \geqslant c_0. \tag{2.7}$$

Now

$$\begin{aligned} W_{X_j \cdot \phi} \begin{pmatrix} a & & \\ & h & \\ & & a^{-1} \end{pmatrix} &= \frac{d}{dt} W_\phi \left(\begin{pmatrix} a & & \\ & h & \\ & & a^{-1} \end{pmatrix} e^{tX_j} \right) \Bigg|_{t=0} \\ &= \frac{d}{dt} e(ta(\alpha, he_j))\, W_\phi \begin{pmatrix} a & & \\ & h & \\ & & a^{-1} \end{pmatrix} \Bigg|_{t=0} \\ &= 2\pi i a(\alpha, he_j)\, W_\phi \begin{pmatrix} a & & \\ & h & \\ & & a^{-1} \end{pmatrix}. \end{aligned}$$

Therefore

$$W_\phi \begin{pmatrix} a & & \\ & h & \\ & & a^{-1} \end{pmatrix} = \frac{1}{2\pi i a(\alpha, he_j)} W_{X_j \cdot \phi} \begin{pmatrix} a & & \\ & h & \\ & & a^{-1} \end{pmatrix}.$$

Repeating this k times gives

$$W_\phi\begin{pmatrix} a & & \\ & h & \\ & & a^{-1}\end{pmatrix} = \frac{1}{a^k(2\pi i(\alpha, he_j))^k} W_{X_j^k\cdot\phi}\begin{pmatrix} a & & \\ & h & \\ & & a^{-1}\end{pmatrix}.$$

Applying Lemma 2.1 and (2.7) we obtain

$$\left| W_\phi\begin{pmatrix} a & & \\ & h & \\ & & a^{-1}\end{pmatrix}\right| \leqslant 2(2\pi c_0)^{-k}\cdot a^{k_0-k}\left(\sum_\alpha \|D_\alpha\cdot X_j^k\cdot\phi\|\right)$$

where k_0 is a fixed integer. Now we obtain our lemma upon taking $k=k_0+N$ and the collection $D_1,\dots,D_m$ to be $2(2\pi c_0)^{-k}\cdot D_\alpha\cdot X_j^k$ for all α and $1\leqslant j\leqslant r$. Q.E.D.

For $\phi\in H_\infty$ we let

$$I_\phi(\tau,s) = \int_H\int_0^{+\infty} W_\phi(m)\overline{\chi_\tau(h_m)}\,|a_m|^{s-r/2}d^*a\,dh \tag{2.8}$$

where χ_τ is the character of the finite dimensional representation τ. More generally for ξ_τ any matrix coefficient of τ we set

$$I_\phi(\xi_\tau,s) = \int_H\int_0^{\infty} W_\phi(m)\overline{\xi_\tau(h_m)}\,|a_m|^{s-r/2}d^*a\,dh. \tag{2.9}$$

Proposition 2.3. *For $\phi\in H_\infty$, $I_\phi(\tau,s)$ and $I_\phi(\xi_\tau,s)$ converge absolutely for* Re(s) *large. The maps $\phi\to I_\phi(\tau,s)$ and $\phi\to I_\phi(\xi_\tau,s)$ are continuous in the topology of H_∞.*

Proof. The convergence follows from the estimation of Lemma 2.2. The continuity follows from the dependence of these estimates on the derivatives of ϕ. Q.E.D.

2.2. *A shift equation.* Let $e_1,\dots,e_r$ denote the standard basis of $\mathbf{R}^r$ and set

$$\begin{cases} X_j = \begin{pmatrix} 0 & -{}^te_j & 0\\ 0 & 0 & e_j\\ 0 & 0 & 0\end{pmatrix}\\ T = \begin{pmatrix} 1 & & \\ & 0 & \\ & & -1\end{pmatrix}\\ \check{X}_j = {}^tX_j\end{cases} \tag{2.10}$$

Let $\mathfrak{g}, \mathfrak{m}, \mathfrak{a}, \mathfrak{u}, \mathfrak{h}$ be the Lie algebras of G, M, etc. Let $\mathfrak{u}^-$ be the span of the $\check{X}_j$. We use the subscript $\mathbf{C}$ to denote complexifications. Let $Z(\mathfrak{g}_\mathbf{C})$ be the center of $U(\mathfrak{g}_\mathbf{C})$.

LEMMA 2.4. *Let $D \in Z(\mathfrak{g}_\mathbf{C})$, then D can bes expressed as*

$$D = D_\mathfrak{m} + \sum_{j=1}^{r} X_j D_j$$

where $D_\mathfrak{m} \in Z(\mathfrak{m}_\mathbf{C})$ and $D_j \in U(\mathfrak{g}_\mathbf{C})$.

Proof. We have

$$\mathfrak{g}_\mathbf{C} = u_\mathbf{C}^- + \mathfrak{m}_\mathbf{C} + \mathfrak{u}_\mathbf{C}$$

and, by the Poincaré–Birkhoff–Witt theorem, a direct sum decomposition

$$U(\mathfrak{g}_\mathbf{C}) = U(\mathfrak{m}_\mathbf{C}) \oplus (\mathfrak{u}_\mathbf{C} \cdot U(\mathfrak{g}_\mathbf{C}) + U(\mathfrak{g}_\mathbf{C}) \cdot \mathfrak{u}_\mathbf{C}^-). \tag{2.11}$$

It is standard that if $D \in Z(\mathfrak{g}_\mathbf{C})$ then its projection onto the first factor in (2.11) lies in $Z(\mathfrak{m}_\mathbf{C})$. Hence

$$Z(\mathfrak{g}_\mathbf{C}) \subseteq Z(\mathfrak{m}_\mathbf{C}) \oplus (\mathfrak{u}_\mathbf{C}\, U(\mathfrak{g}_\mathbf{C}) + U(\mathfrak{g}_\mathbf{C})\, \mathfrak{u}_\mathbf{C}^-). \tag{2.12}$$

Given $D \in Z(\mathfrak{g}_\mathbf{C})$ we may write

$$D = D_0 + f$$

where $D_0 \in Z(\mathfrak{m}_\mathbf{C}), f \in \mathfrak{u}_\mathbf{C}\, U(\mathfrak{g}_\mathbf{C}) + U(\mathfrak{g}_\mathbf{C})\, \mathfrak{u}_\mathbf{C}^-$. Clearly f commutes with T. Now by the PBW theorem f is a linear combination of monomials of the form

$$X^\alpha D_1 \check{X}^\beta$$

where $\alpha, \beta \in \mathbf{Z}_+^r$,

$$X^\alpha = \prod X_j^{\alpha_j}, \quad \check{X}^\beta = \prod \check{X}_j^{\beta_j}$$

and $D_1 \in U(\mathfrak{m}_\mathbf{C})$. We have clearly

$$[T, X^\alpha D_1 \check{X}^\beta] = (|\alpha| - |\beta|)\, X^\alpha D_1 \check{X}^\beta$$

where $|\alpha| = \alpha_1 + \ldots + \alpha_r$, $|\beta| = \beta_1 + \ldots + \beta_r$. Since f commutes with any power of T we may assume that each monomial as above occurring in f has $|\alpha| = |\beta|$. Then since (2.11) is a direct sum we must have

$$|\alpha| = |\beta| > 0.$$

This proves the lemma. Q.E.D.

A representation of G is called quasi-simple if $Z(\mathfrak{g}_{\mathbf{C}})$ acts by scalars. Let π be quasi-simple. Let λ_π be the infinitesimal character of π determined by relation

$$\pi(D)\cdot\phi = \lambda_\pi(D)\,\phi \tag{2.13}$$

for all $\phi \in H_\infty$, $D \in Z(\mathfrak{g}_{\mathbf{C}})$.

If τ is an irreducible representation of H as above and $s \in \mathbf{C}$, then as is well known, the induced representation

$$\pi(\tau, s) = \mathrm{Ind}_P^G(\tau \otimes |\cdot|^{-s}) \tag{2.14}$$

is quasi-simple. Let $\lambda_{\tau,s}$ be its infinitesimal character.

Proposition 2.5. *Let $D \in Z(\mathfrak{g}_{\mathbf{C}})$ and write*

$$D = D_m + \sum_{i=1}^{r} X_i D_i$$

as in Lemma 2.4. Let $r_i(h)$ be the matrix coefficient of the standard representation of H given by

$$r_i(h) = (\alpha, he_i)$$

(see (2.2)). Then for Re(s) *large*

$$I_\phi(\xi_\tau, s) = \frac{2\pi i}{\lambda_\pi(D) - \lambda_{\tau,s}(D)} \sum_{j=1}^{r} I_{D_j\phi}(r_j\xi_\tau, s+1). \tag{2.15}$$

Proof. From the proof of Lemma 2.2 we know

$$W_{\pi(X_j)\phi}\begin{pmatrix} a & & \\ & h & \\ & & a^{-1} \end{pmatrix} = 2\pi i a r_j(h)\, W_\phi\begin{pmatrix} a & & \\ & h & \\ & & a^{-1} \end{pmatrix}.$$

Hence

$$I_{\pi(X_j)\phi}(\xi_\tau, s) = 2\pi i\, I_\phi(r_j\xi_\tau, s+1). \tag{2.16}$$

Now on the one hand we have

$$I_{\pi(D)\phi}(\xi_\tau, s) = \lambda_\pi(D) I_\phi(\xi_\tau, s). \tag{2.17}$$

On the other

$$\begin{aligned} I_{\pi(D)\phi}(\xi_\tau, s) &= I_{D_{\mathfrak{m}}\cdot\phi}(\xi_\tau, s) + \sum_{j=1}^{r} I_{X_j D_j\cdot\phi}(\xi_\tau, s) \\ &= I_{D_{\mathfrak{m}}\cdot\phi}(\xi_\tau, s) + 2\pi i \sum_{j=1}^{r} I_{D_j\cdot\phi}(r_j \xi_\tau, s+1). \end{aligned} \tag{2.18}$$

It therefore remains to show that

$$I_{D_{\mathfrak{m}}\cdot\phi}(\xi_\tau, s) = \lambda_{\tau,s}(D) I_\phi(\xi_\tau, s). \tag{2.19}$$

Let $M^+ = HA^+$ be the connected component of M. Let γ be the representation of M^+ given by

$$\gamma = \check{\tau} \otimes |\cdot|^{s-r/2} \tag{2.20}$$

where $\check{\tau}$ is the contragredient representation of τ. Let $V_{\check{\tau}}$ be the space on which $\check{\tau}$ acts. Then γ can be realized on $V_{\check{\tau}}$. We define a bilinear form $\langle\ ,\ \rangle$ on $H_\infty \times V_{\check{\tau}}$ by

$$\langle \phi, \xi \rangle = \int_H \int_0^\infty W_\phi \begin{pmatrix} a & & \\ & h & \\ & & a^{-1} \end{pmatrix} \xi(h) |a|^{s-r/2} d^*a\, dh.$$

(Here we are identifying $V_{\check{\tau}}$ with a space of matrix coefficients of $\check{\tau}$.) Then

$$I_\phi(\xi_\tau, s) = \langle \phi, \xi_\tau \rangle. \tag{2.21}$$

From invariance of Haar measure on M^+ we see the pairing $\langle\ ,\ \rangle$ is invariant for the action of M^+. On the Lie algebra level this says

$$\langle X\phi, \xi \rangle + \langle \phi, X\xi \rangle = 0 \tag{2.22}$$

for all $X \in \mathfrak{m}$. Define an involution on $\mathfrak{m}$ by

$$X \to \check{X} = -X.$$

This extends to a unique complex linear involution on $U(\mathfrak{m}_{\mathbb{C}})$, which we again denote by

$D_0 \to \check{D}_0$. Then (2.22) implies

$$\langle D_0\phi, \xi\rangle = \langle \phi, \check{D}_0\xi\rangle. \tag{2.23}$$

Let λ_γ be the infinitesimal character of the irreducible representation γ. Then if $D_0 \in Z(\mathfrak{m}_C)$ we get from (2.23)

$$\langle D_0\phi, \xi\rangle = \lambda_\gamma(\check{D}_0)\langle \phi, \xi\rangle. \tag{2.24}$$

In particular

$$\begin{aligned} I_{D_{\mathrm{m}}\cdot\phi}(\xi_\tau, s) &= \langle D_{\mathrm{m}}\cdot\phi, \bar{\xi}_\tau\rangle \\ &= \lambda_\gamma(\check{D}_{\mathrm{m}})I_\phi(\xi_\tau, s). \end{aligned}$$

Thus (2.19) will follow once we show that

$$\lambda_\gamma(\check{D}_{\mathrm{m}}) = \lambda_{\tau,s}(D). \tag{2.25}$$

But this follows from a discussion of the Harish-Chandra homomorphism contained in the next lemma. Q.E.D.

LEMMA 2.6. *The relation* (2.25) *is valid for all* $D \in Z(\mathfrak{g}_C)$.

Proof. Let $\mathfrak{t}$ be a Cartan subalgebra of $\mathfrak{h}$. Then $\mathfrak{l} = \mathfrak{a} + \mathfrak{t}$ is a Cartan subalgebra of $\mathfrak{g}$. Let W be the Weyl group of the pair $(\mathfrak{l}_C, \mathfrak{g}_C)$. The Harish-Chandra homomorphism gives us an isomorphism

$$\psi: Z(\mathfrak{g}_C) \xrightarrow{\sim} U(\mathfrak{l}_C)^W. \tag{2.26}$$

In particular a character of $Z(\mathfrak{g}_C)$ is given by (the W-conjugacy class of) an element of $\mathfrak{l}_C^*$. By convention the element $\mathfrak{l}_C^*$ which corresponds to the infinitesimal character of a representation π will itself be called the infinitesimal character of π.

Let

$$\mathfrak{m}_C = \mathfrak{n}_C \oplus \mathfrak{l}_C \oplus \mathfrak{n}_C^-$$

be a triangular decomposition of $\mathfrak{m}_C$, in such a way that $\mathfrak{n}_C + \mathfrak{u}_C$ and $\mathfrak{n}_C^- + \mathfrak{u}_C^-$ are maximal unipotent subalgebras of $\mathfrak{g}_C$. Let $\delta_{\mathrm{n}}, \delta_{\mathrm{u}}$ be the half sum of positive roots of $\mathfrak{l}_C$ in $\mathfrak{n}_C$ and $\mathfrak{u}_C$ respectively. We note that δ_{n} vanishes on $\mathfrak{a}_C$ while δ_{u} vanishes on $\mathfrak{t}_C$. Let Λ_τ be the highest weight of τ with respect to $\mathfrak{n}_C$. From Knapp [15, p. 225] we know that the representation (2.14) has infinitesimal character given by

(2.27) $$\Lambda_\tau+\delta_{\mathrm{n}}-s$$

where s is identified with the functional on $\mathfrak{a}_{\mathrm{C}}$ which takes the basis element $T\in\mathfrak{a}_{\mathrm{C}}$ to s.

Let us identify $U(\mathfrak{l}_{\mathrm{C}})$ with polynomial functions on $\mathfrak{l}_{\mathrm{C}}^*$:

$$U(\mathfrak{l}_{\mathrm{C}})\simeq P(\mathfrak{l}_{\mathrm{C}}^*).$$

Define an automorphism

$$\beta\colon P(\mathfrak{l}_{\mathrm{C}}^*)\to P(\mathfrak{l}_{\mathrm{C}}^*)$$

by the recipe

$$\beta(f)(\nu)=f(\nu+\delta_{\mathrm{u}})$$

for $\nu\in\mathfrak{l}_{\mathrm{C}}^*$. If

$$\psi_{\mathrm{m}}\colon Z(\mathfrak{m}_{\mathrm{C}})\to U(\mathfrak{l}_{\mathrm{C}})$$

denotes the Harish-Chandra homomorphism for $\mathfrak{m}_{\mathrm{C}}$ it is easy to see that the map (2.26) is given by

(2.28) $$\psi(D)=\beta\circ\psi_{\mathrm{m}}(D_{\mathrm{m}})$$

where D_{m} is related to D as in Proposition 2.5. To conclude the proof we need to show, by (2.27)

(2.29) $$\psi(D)(\Lambda_\tau+\delta_{\mathrm{n}}-s)\ \lambda_\gamma(\check{D}_{\mathrm{m}}).$$

But the map

$$D_{\mathrm{m}}\to\lambda_\gamma(\check{D}_{\mathrm{m}})$$

corresponds to the infinitesimal character of the contragredient representation of γ, which is

$$\check{\gamma}=\tau\otimes|\cdot|^{-(s-r/2)}.$$

The representation $\check{\gamma}$ has infinitesimal character

$$\Lambda_\tau+\delta_{\mathrm{n}}-(s-r/2)=\Lambda_\tau+\delta_{\mathrm{n}}-s+\delta_{\mathrm{u}}.$$

Thus (2.29) amounts to

$$\psi(D)(\Lambda_\tau+\delta_{\mathrm{n}}-s)=\psi_{\mathrm{m}}(D_{\mathrm{m}})(\Lambda_\tau+\delta_{\mathrm{n}}-s+\delta_{\mathrm{u}})$$

which is obviously true in view of (2.28). This proves the lemma. Q.E.D.

Remark. If $\mathfrak{g}$ is of type B, i.e., if r is odd, then $-s$ in (2.14) can be replaced by s. For in such case there is an element in the Weyl group which acts as -1 on $\mathfrak{a}_{\mathbf{C}}$ and as identity on $\mathfrak{t}_{\mathbf{C}}$.

We shall adopt the notations and terminologies introduced in the proof of Lemma 2.6.

Let π be a quasi-simple representation. Let $\chi_\pi\in\mathfrak{t}_{\mathbf{C}}^*$ be the infinitesimal character of π. Given $u, v\in\mathfrak{t}_{\mathbf{C}}^*$ we write

$$u\overset{W}{\sim}v$$

if there is an element of W taking u to v.

PROPOSITION 2.7. *Let notations and assumptions be as above. Then*

(i) *$I_\phi(\xi_\tau, s)$ has a meromorphic continuation to all of* $\mathbf{C}$.

(ii) *Let R be the standard representation of $H=SO(r)$ on $\mathbf{R}^r$. Then $I_\phi(\xi_\tau, s)$ has a possible pole at s only if the following conditions are satisfied: There is a highest weight Λ occurring in $\tau\otimes R^k$, k an integer $\geqslant 0$, such that*

$$\Lambda+\delta_{\mathrm{n}}-(s+k)\overset{W}{\sim}\chi_\pi. \tag{2.30}$$

Proof. By Lemmas 2.1, 2.2, there is a real number s_0 depending on π, such that $I_\phi(\xi_\tau, s)$ is holomorphic in the half plane $\mathrm{Re}(s)>s_0$ for any ϕ and ξ_τ. Then from equation (2.15) we see that $I_\phi(\xi_\tau, s)$ is holomorphic in the region $\mathrm{Re}(s)>s_0-1$ except at those points s with $s_0-1<\mathrm{Re}(s)\leqslant s_0$, such that

$$\lambda_\pi(D)=\lambda_{\tau,s}(D)$$

for all $D\in Z(\mathfrak{g}_{\mathbf{C}})$. In view of the Harish-Chandra homomorphism (2.26) the above implies

$$\Lambda_\tau+\delta_{\mathrm{n}}-s\overset{W}{\sim}\chi_\pi.$$

where $\Lambda_\tau+\delta_n-s$ corresponds to the infinitesimal character $\lambda_{\tau,s}$, as we have seen in the proof of Lemma 2.6.

Now $r_j\cdot\xi_\tau$ in (2.15) is a matrix coefficient of $\tau\otimes R$. By decomposing $\tau\otimes R$ we may write

$$r_j \xi_\tau = \sum_{\tau'} \xi_{\tau'}$$

where $\xi_{\tau'}$ is a matrix coefficient of the irreducible representation τ' occurring in $\tau \otimes R$. Then

$$I_\phi(r_j \xi_\tau, s+1) = \sum_{\tau'} I_\phi(\xi_{\tau'}, s+1),$$

and we can repeat the above argument with each $I_\phi(\xi_{\tau'}, s+1)$. This proves the proportion. Q.E.D.

Remark 2.8. Let π be irreducible. It is well known that π is then a subquotient (or even a subrepresentation) of an induced representation

$$\mathrm{Ind}_P^G(\sigma \otimes |\cdot|^\varrho \cdot \mathrm{sgn})$$

where σ is an irreducible representation of H, $\varrho \in \mathbf{C}$ and $\varepsilon=0$ or 1. The infinitesimal character of π is then

$$\chi_\pi = \Lambda_\sigma + \delta_{\mathrm{n}} + \varrho.$$

If

$$\varrho \not\equiv \frac{r}{2} \mod \mathbf{Z}$$

then (2.30) implies

(2.31) $$s+k = \pm \varrho.$$

If

$$\varrho \equiv \frac{r}{2} \mod \mathbf{Z}$$

then (2.30) implies

(2.32) $$s \equiv \frac{r}{2} \mod \mathbf{Z}.$$

These relations can be easily obtained by looking at the explicit action of the Weyl group.

2.3. We now apply the above to the case of interest. The group G acts on $L^2(\Gamma\backslash G)$ in the usual way and defines a unitary representation there. Let $S(\Gamma\backslash G)$ be the space of smooth functions in $L^2(\Gamma\backslash G)$ with respect to the action of G. For $\phi \in S(\Gamma\backslash G)$ we define its Fourier coefficient $W_\phi(g)$ by

$$W_\phi(g) = \int_{\Gamma\cap U\backslash U} \phi(ug)\,\eta^{-1}(u)\,du \tag{2.33}$$

where η is again a non-trivial character of U, but now assumed to be trivial on $\Gamma\cap U$.

LEMMA 2.9. *The functional*

$$\phi \to W_\phi(1)$$

is continuous with respect to the topology of $S(\Gamma\backslash G)$.

Proof. This is an application of Sobolev's lemma. In fact, if $l > \frac{1}{2}\dim \mathfrak{g}$ then for a suitable basis $\{D_j\}$ of $U(\mathfrak{g})_l$ we have

$$|W_\phi(1)| \leq \sum_j \|D_j\phi\| \tag{2.34}$$

for all $\phi \in S(\Gamma\backslash G)$. Q.E.D.

2.4. *Spectral decomposition.* In order to meromorphically continue $I_\phi(s)$ we need to expand ϕ spectrally. Consider the spectral decomposition of $L^2(\Gamma\backslash G)$. We have

$$\begin{aligned} L^2(\Gamma\backslash G) &= L_0^2(\Gamma\backslash G)\oplus L^2_{\mathrm{res}}(\Gamma\backslash G)\oplus L^2_{\mathrm{cont}}(\Gamma\backslash G) \\ &= L^2_{\mathrm{disc}}(\Gamma\backslash G)\oplus L^2_{\mathrm{cont}}(\Gamma\backslash G) \end{aligned} \tag{2.35}$$

$L_0(\Gamma\backslash G)$ decomposes as a discrete direct sum of irreducibles each occurring with finite multiplicity.

$L^2_{\mathrm{res}}(\Gamma\backslash G)$ decomposes first as

$$L^2_{\mathrm{res}}(\Gamma\backslash G) = \bigoplus_{\tau\in\hat{H}} L^2_{\mathrm{res},\tau}(\Gamma\backslash G) \tag{2.36}$$

where $\hat{H}$ is the unitary dual of H and $L^2_{\mathrm{res},\tau}(\Gamma\backslash G)$ is spanned by the residues of Eisenstein series of the form $\mathrm{Ind}_P^G(\tau\otimes|\ |^{\varrho}\,\mathrm{sgn}^{\varepsilon})$, $\varepsilon=0$ or 1.

For each τ, the space $L^2_{\mathrm{res},\tau}(\Gamma\backslash G)$ decomposes as a direct sum of a finite number of irreducible representations (of the complementary series). $L^2_{\mathrm{cont}}(\Gamma\backslash G)$ is a direct integral of the form

$$L^2_{\mathrm{cont}}(\Gamma\backslash G)=\sum_{\varepsilon=0}^{1}\sum_{\tau\in\hat{H}}\int_{-\infty}^{\infty}\mathrm{Ind}_P^G(\tau\otimes\mathrm{sgn}^{\varepsilon}|\;|^{ir})\frac{dr}{4\pi}. \tag{2.37}$$

Note that $\tau\otimes\mathrm{sgn}^{\varepsilon}|\;|^{\varrho}$ is naturally a representation of $H\times A$ and thus of P via its quotient M. (For notational simplicity only we have assumed Γ has only one cusp.)

As $\phi\in L^2(\Gamma\backslash G)$ we have the corresponding decomposition of ϕ. We may write

$$\phi=\sum_{\pi\subset L^2_{\mathrm{disc}}(\Gamma\backslash G)}F_\pi(\phi)+\sum_{\tau,\varepsilon}\int F_{\pi(\tau,ir,\varepsilon)}(\phi)\frac{dr}{4\pi} \tag{2.38}$$

with $F_\pi(\phi)$ the projection of ϕ into the corresponding π component of $L^2(\Gamma\backslash G)$. If in addition $\phi\in S(\Gamma\backslash G)$, then one can use the Dixmier–Malliavin theorem [4] to conclude that each $F_\pi(\phi)$ is a smooth vector in π and furthermore that the above decomposition converges in the $S(\Gamma\backslash G)$ topology.

From Proposition 2.3 we have

PROPOSITION 2.10. *For $\phi\in S(\Gamma\backslash G)$ and* Re(s) *large we have the absolutely convergent representations*

$$I_\phi(\xi_\tau,s)=\sum_{\pi}I_{F_\pi(\phi)}(\xi_\tau,s)+\sum_{\sigma,\varepsilon}\int_{-\infty}^{\infty}I_{F_{\pi(\sigma,ir,\varepsilon)}(\phi)}(\xi_\tau,s)\frac{dr}{4\pi}$$

and similarly for $I_\phi(\tau,s)$.

In the next subsection we show that the analytic continuation of $Z(s)$ can be achieved through the continuation of $I_\phi(\xi_\tau,s)$. The representation in Proposition 2.10 will be used to do this. In fact later we show that for s in a compact set all but a finite number of the terms in the above series are analytic in s. For the finitely many terms we have studied the meromorphicity of the $I_{F_\pi(\phi)}(\xi_\tau,s)$ individually.

2.5. *Poincaré series and the zeta function.* Let $\nu\in\mathcal{S}(\mathbf{R}^r)$ and $\mu\in C_0^\infty(\mathbf{R}^*)$. Define $f\in S(U\backslash G,\psi)$ as follows: First recall the Bruhat decomposition

$$G=P\cup UwP.$$

Set

$$f(g)=\begin{cases}\psi(u_1)\nu(u_2)\chi_\tau(h_m)\mu(a_m) & \text{if } g=u_1wu_2m\\ 0 \text{ for } g\in P.\end{cases} \tag{2.39}$$

where τ is a fixed irreducible representation of H. Define

$$P_f(g) = \sum_{\gamma \in \Gamma \cap U \backslash \Gamma} f(\gamma g). \tag{2.40}$$

Clearly $P_f \in S(\Gamma \backslash G)$. Hence we may compute its Whittaker–Mellin transform $I_{P_f}(\tau, s)$. We begin with the computation of its Whittaker function.

LEMMA 2.11. *If $m \in M$ then*

$$W_{P_f}(m) = \sum_{m' \in M(\Gamma)} \mathrm{Kl}(\psi, \eta, m')\, \hat{\nu}(a_{m'} \tilde{h}_{m'}^{-1} \alpha)\, \chi_\tau(\tilde{h}_{m'} h_m)\, \mu\left(\frac{a_m}{a_{m'}}\right) \cdot |a_{m'}|^r$$

where $\hat{\nu}$ is the Fourier transform on $\mathbf{R}^r$ with respect to η^{-1}, $\tilde{h}_{m'} = \varepsilon_r h_{m'} \varepsilon_r$, α is defined in (2.5) *and* Kl *is the Kloosterman sum of Section* 1.

Proof.

$$\begin{aligned} W_{P_f}(m) &= \int_{\Gamma^\infty \backslash U} P_f(um)\, \eta^{-1}(u)\, du \\ &= \int_{\Gamma^\infty \backslash U} \sum_{\Gamma^\infty \backslash \Gamma} f(\gamma um)\, \eta^{-1}(u)\, du. \end{aligned}$$

Now for $\gamma \in \Gamma$, $\gamma \notin \Gamma^\infty$ we may write

$$\gamma = u_1(\gamma)\, m(\gamma)\, w u_2(\gamma).$$

We are assuming $\Gamma \cap P = \Gamma \cap U = \Gamma^\infty$, so by the Bruhat decomposition

$$\Gamma^\infty \backslash \Gamma = 1 + \bigcup_{m' \in M(\Gamma)} \Gamma^\infty \backslash U m' w U$$

therefore

$$W_{P_f}(m) = \int_{\Gamma^\infty \backslash U} f(um)\, \eta^{-1}(u)\, du + \sum_{m' \in M(\Gamma)} \int_U \sum_{\Gamma^\infty \backslash U m' w U / \Gamma^\infty} f(u_1(\gamma)\, m' w u_2(\gamma)\, um)\, \eta^{-1}(u)\, du.$$

The first integral is 0 since f vanishes on the small Bruhat cell. The second gives

$$\sum_{m' \in M(\Gamma)} \mathrm{Kl}(\psi, \eta, m')\, \hat{\nu}(a_{m'} \tilde{h}_{m'}^{-1} \alpha)\, |a_{m'}|^r \mu\left(\frac{a_m}{a_{m'}}\right) \chi_\tau(\tilde{h}_{m'} h_m). \qquad \text{Q.E.D.}$$

Since Γ is discrete, there is a positive constant C such that for $m \in M(\Gamma)$ we have

$|a_m^{-1}|>C$. Since H is compact we see that

$$\{\xi = a_m \bar{h}_m^{-1} \alpha \mid m \in M(\Gamma)\}$$

is contained in a closed ball B in $\mathbf{R}^r$. Hence we can find $\nu \in \mathcal{S}(\mathbf{R}^r)$ such that $\hat{\nu}(\xi)=1$ for all $\xi \in B$. Fix such a choice of ν. Then the series for W in the previous lemma takes the form

$$W_{P_f}(m) = \sum_{m' \in M(\Gamma)} \mathrm{Kl}(\psi, \eta, m') \mu\left(\frac{a_m}{a_{m'}}\right) |a_{m'}|^r \chi_\tau(\bar{h}_{m'} h_m).$$

We can now proceed to compute the Mellin transform. To simplify the computations assume that $\mu(a)=\mu(|a|)$.

PROPOSITION 2.12. *For* $\mathrm{Re}(s)>r/2+k$ *(or simply large),* $\mu \in C_0^\infty(\mathbf{R}^*)$ *and even,* ν *as above*

$$I_{P_f}(\tau, s) = \frac{\tilde{\mu}(s-r/2)}{\dim \tau} Z(\psi, \eta, \tau^\iota, s)$$

where $\tilde{\mu}$ *denotes the Mellin transform and* $\tau^\iota(h)=\tau(\varepsilon_r h \varepsilon_r)=\tau(\bar{h})$.

Proof. For $\mathrm{Re}(s)$ large

$$\begin{aligned} I_{P_f}(\tau, s) &= \int_{M^+} W_{P_f}(m) \overline{\chi_\tau(h_m)}\, |a_m|^{s-r/2}\, dm \\ &= \sum_{m' \in M(\Gamma)} \mathrm{Kl}(\psi, \eta, m') |a_{m'}|^r \cdot \int_0^\infty \mu\left(\frac{a_m}{a_{m'}}\right) |a_m|^{s-r/2}\, d^* a_m \\ &\qquad \int_H \chi_\tau(\bar{h}_{m'} h_m) \overline{\chi_\tau(h_m)}\, dh_m. \end{aligned}$$

Now

$$\int_0^\infty \mu\left(\frac{a_m}{a_{m'}}\right) |a_m|^{s-r/2}\, d^* a_m = |a_{m'}|^{s-r/2} \tilde{\mu}(s-r/2)$$

while

$$\int_H \chi_\tau(\bar{h}_{m'} h_m) \overline{\chi_\tau(h_m)}\, dh_m = \frac{1}{\dim(\tau)} \chi_\tau(\bar{h}_{m'}).$$

The proposition follows. Q.E.D.

Now $\mu \in C_0^\infty(\mathbf{R}^*)$ is arbitrary and so can be chosen so that $\tilde{\mu}$ does not vanish at any given s. Hence the meromorphic continuation of Z and location of poles is dictated by $I_{P_f}(\tau, s)$. A similar analysis to the above can be carried out for the matrix Kloosterman zeta function rather than its trace (i.e. for χ_τ replaced by matrix coefficients).

Thus our problem is reduced to meromorphically continuing $I_{P_f}(\xi_\tau, s)$ (which contains the continuation of $I_{P_f}(\tau, s)$ in the obvious way). In fact the explicit form of P_f will no longer play a role. We are reduced to the more general problem: For $\phi \in S(\Gamma \backslash G)$ to analytically continue $I_\phi(\xi_\tau, s)$ to $\mathbf{C}$ and determine the location of the poles.

2.6. *Analytic continuation.* Recall that from Proposition 2.10 we have for $\mathrm{Re}(s)$ large the representation

$$I_\phi(\xi_\tau, s) = \sum_{\pi \in L^2_{\mathrm{disc}}(\Gamma \backslash G)} I_{F_\pi(\phi)}(\xi_\tau, s) + \sum_{\sigma, \varepsilon} \int_{-\infty}^{+\infty} I_{F_{\pi(\sigma, ir, \varepsilon)}(\phi)}(\xi_\tau, s) \frac{dr}{4\pi}. \tag{2.41}$$

The first step is to use the shift equation of Proposition 2.5 to prove

PROPOSITION 2.13. *Let $\mathcal{O}$ be an open set with compact closure in $\mathbf{C}$. Then there is a finite set F_1 of π's and F_2 of (σ, ε)'s such that*

$$I_\phi(\xi_\tau, s) - \sum_{\pi \in F_1} I_{F_\pi(\phi)}(\xi_\tau, s) - \sum_{(\sigma, \varepsilon) \in F_2} \int_{-\infty}^{\infty} I_{F_{\pi(\sigma, ir, \varepsilon)}(\phi)}(\xi_\tau, s) \frac{dr}{4\pi}$$

has an analytic continuation to $\mathcal{O}$.

With this proposition the meromorphic continuation of $I_\phi(\xi_\tau, s)$ and location of poles, is reduced to studying the finite sum of integrals above which are easily handled by Proposition 2.7.

For a moment, let us return to the setting of Lemma 2.6 and Proposition 2.7.

Let $\mathfrak{l}_{\mathbf{C}}^*/W$ denote the set of orbits of W in $\mathfrak{l}_{\mathbf{C}}^*$. This is an affine variety, in fact even an affine space. Recall that to each quasisimple representation π one can associate its infinitesimal character $\chi_\pi \in \mathfrak{l}_{\mathbf{C}}^*/W$.

LEMMA 2.14. *Let K be a fixed compact subset of $\mathfrak{l}_{\mathbf{C}}^*/W$. Then*

(a) *$\{\chi_\pi | \pi \subset L^2_{\mathrm{disc}}(\Gamma \backslash G)\} \cap K$ is finite.*

(b) *If $\pi(\sigma, ir, \varepsilon) = \mathrm{Ind}_P^G(\sigma \otimes |\ |^{ir} \mathrm{sgn}^\varepsilon)$, then at most a finite number of the curves (in r) $\chi_\pi(\sigma, ir, \varepsilon)$ with $(\sigma, \varepsilon) \in \hat{H} \times \{0, 1\}$ meet K and those curves that do meet K, do so with bounded r.*

Proof. (a) The spectrum of G acting on $L^2_{\mathrm{disc}}(\Gamma\backslash G)$ is discrete, and this implies that the set

$$\{\chi_\pi \mid \pi \subset L^2_{\mathrm{disc}}(\Gamma\backslash G)\}$$

is discrete in $\mathfrak{l}^*_{\mathbf{C}}/W$. The statement (b) follows more easily. For as we know from Knapp [15, p. 225], the infinitesimal character of $\pi(\sigma, ir, \varepsilon)$ is represented by

$$\Lambda_\sigma + \delta_{\mathrm{n}} + ir \in \mathfrak{l}^*_{\mathbf{C}}$$

(see (2.27)). The fact that $\Lambda_\sigma + \delta_{\mathrm{n}} + ir$ represents an element of the compact set K implies that $\Lambda_\sigma + \delta_{\mathrm{n}} + ir$ is bounded in $\mathfrak{l}^*_{\mathbf{C}}$. In particular

$$-R \leq r \leq R. \tag{2.42}$$

But Λ_σ, being the highest weight of σ, lies in the "weight lattice" inside $\mathfrak{t}^* \subseteq \mathfrak{l}^*$. Hence there can be only finitely many of them within a bounded set. This proves (b). Q.E.D.

The fact that $\mathfrak{l}^*_{\mathbf{C}}/W$ is an affine space corresponds to the fact that $Z(\mathfrak{g}_{\mathbf{C}}) \simeq U(\mathfrak{l}_{\mathbf{C}})^W$ is a polynomial algebra. Let

$$m = \dim_{\mathbf{C}}(\mathfrak{l}_{\mathbf{C}}) = \left[\frac{r}{2}\right] + 1.$$

Choose a set of generators $D_1, \ldots, D_m$ for the algebra $Z(\mathfrak{g}_{\mathbf{C}})$. Given $\chi \in \mathfrak{l}^*_{\mathbf{C}}/W$ let λ be the character of $Z(\mathfrak{g}_{\mathbf{C}})$ corresponding to χ. The map

$$\chi \to \begin{pmatrix} \lambda(D_1) \\ \vdots \\ \lambda(D_m) \end{pmatrix}$$

then gives us an isomorphism

$$\mathfrak{l}^*_{\mathbf{C}}/W \xrightarrow{\sim} \mathbf{C}^m. \tag{2.43}$$

From now on for a quasi-simple representation of G we shall write

$$\chi_\pi = \begin{pmatrix} \lambda_\pi(D_1) \\ \vdots \\ \lambda_\pi(D_m) \end{pmatrix} \in \mathbf{C}^m. \tag{2.44}$$

We can now prove Proposition 2.13. Let $\lambda_{\tau,s}$ be as in (2.15) and let $\chi_{\tau,s}$ be the corresponding element given by (2.44). Set

$$K = \text{closure of } \{\chi_{\tau,s} \mid s \in \mathcal{O}\}.$$

By Lemma 2.14, this meets $\{\chi_\pi \mid \pi \in L^2_{\text{disc}}(\Gamma\backslash G)\}$ in a finite set F_1 of π's. We claim

$$f(s) = \sum_{\substack{\pi \in L^2_{\text{disc}}(\Gamma\backslash G) \\ \pi \notin F_1}} I_{F_\pi(\phi)}(\xi_{\tau,s})$$

is holomorphic in $\mathcal{O}$ for any $\phi \in S(\Gamma\backslash G)$. Firstly this series converges absolutely for $\text{Re}(s) > \sigma_0$. Now for each $\pi \notin F_1$ there is $j = j(\pi) \in \{1, \dots, m\}$ such that

$$|\lambda_{\tau,s}(D_j) - \lambda_\pi(D_j)| \geq \varepsilon_0 > 0 \tag{2.45}$$

where $\varepsilon_0 > 0$ is fixed.

Hence applying (2.15) we find the representation for f

$$f(s) = \sum_{\substack{\pi \in L^2_{\text{disc}}(\Gamma\backslash G) \\ \pi \notin F_1}} \frac{2\pi i}{\lambda_\pi(D_{j(\pi)}) - \lambda_{\tau,s}(D_{j(\pi)})} \cdot \sum_{i=1}^{r} I_{\pi(Y_i)\phi}(r_i \xi_\tau, s+1).$$

In view of (2.45) this series converges absolutely and renders f holomorphic in $\{\text{Re}(s) > \sigma_0 - 1\} \cap \mathcal{O}$. Repeating this gives the holomorphicity of f in $\mathcal{O}$. The argument for the continuous part is similar. This completes the proof of Proposition 2.13. Q.E.D.

We can now state the main theorem of this section:

THEOREM 2.15. *Let $\phi \in S(\Gamma\backslash G)$ then $I_\phi(\xi_\tau, s)$ has a meromorphic continuation to* $\mathbf{C}$ *with the set of possible poles contained in*

(1) *$\varrho - k$, $k \geq 0$, $k \in \mathbf{Z}$ and where $\Pi \subseteq \text{Ind}_P^G(\sigma \otimes |\ |^\varrho \operatorname{sgn}^\varepsilon)$ is a constituent of $L^2_{\text{disc}}(\Gamma\backslash G)$ and also τ is constituent of $\sigma \otimes R^k$.*

(2) *$\varrho - k$, $k \geq 0$ an integer and ϱ a pole of $c(\sigma, s)$, $\sigma \in \hat{H}$, where $c(\sigma, s)$ is the constant term of the Eisenstein series, and here too τ must be a constituent of $\sigma \otimes R^k$.*

Proof. In view of Proposition 2.13 the theorem is reduced to studying the meromorphic continuation and poles of a finite set of $I_{F_\pi(\phi)}(\xi_\tau, s)$'s and a finite set of integrals. Hence they may be dealt with term by term. Proposition 2.7 together with Remark 2.8 immediately gives (1) above which come from the sum over F_1. To meromorphically continue the integral consider one such

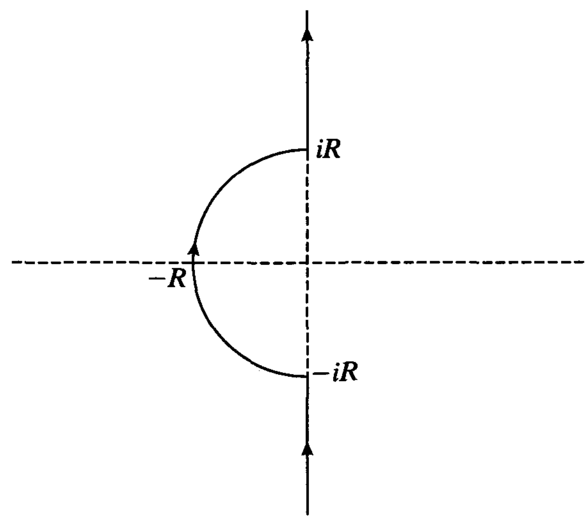

Fig. 1

$$\int_{-\infty}^{\infty} I_{F_{\pi(\sigma, ir, \xi)}(\phi)}(\xi_\tau, s)\, dr, \tag{2.46}$$

or, more generally, an integral of the form

$$I_R(s) = \int_{C_R} I_{F_{\pi(\sigma, \varrho, \xi)}(\phi)}(\xi_\tau, s) \frac{d\varrho}{i}$$

where C_R is the contour in the ϱ-plane given by

$$C_R = \{\varrho = ir : |r| \geqslant R\} \cup \left\{\varrho = Re^{i\theta} : \frac{\pi}{2} < \theta < \frac{3\pi}{2}\right\}$$

oriented so that $\mathrm{Im}(\varrho)$ increases, see Figure 1.

We assume C_R avoids all poles of the Eisenstein series associated to elements of $\mathrm{Ind}_P^G(\sigma \otimes |\ |^{\varrho} \mathrm{sgn}^{\varepsilon})$, or equivalently, the c-function $c(\sigma, \varrho)$. (We can always make small symmetric perturbations of C_R to guarantee this without effecting the validity of what follows.) We have the convergence of $I_0(s)$ for $\mathrm{Re}(s) \gg 0$ from the spectral theory, and since C_R differs from C_0 only on a compact set which avoids the poles in ϱ of the integrand, we have the convergence of $I_R(s)$ for $\mathrm{Re}(s) \gg 0$ as well.

As a function of s, it follows from Proposition 2.7 and Remark 2.8 that the integrand of $I_R(s)$ is a holomorphic function of s on the complement of the curves $C_R - k$

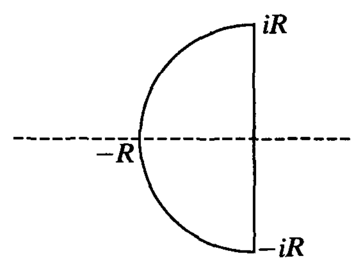

Fig. 2

for $k=0, 1, 2, \ldots$. We claim that in fact $I_R(s)$ converges absolutely, uniformly on compact subsets, on the compliment of the C_R-k to a holomorphic function of s. For if s does not lie on a curve C_R-k for some k, then there is an element $D \in U(\mathfrak{g})$ such that the function $\lambda_\pi(D)-\lambda_{\tau,s}(D)$ does not vanish at s for $\pi=\pi(\sigma, \varrho, \varepsilon)$ with $\varrho \in C_R$. Since $\lambda_\pi(D)-\lambda_{\tau,s}(D)$ is polynomial in both ϱ and s, there is a relatively compact neighborhood U of s such that $\lambda_\pi(D)-\lambda_{\tau,s}(D)$ is bounded away from 0 for $(\varrho, s) \in C_R \times U$. Applying the shift equation (2.15) for D we find

$$\begin{aligned}\int_{C_R} |I_{F_{\pi(\sigma,\varrho,\varepsilon)}(\phi)}(\xi_{\tau,s})|\,|d\varrho| &\leqslant 2\pi \sum_j \int_{C_R} \frac{1}{|\lambda_\pi(D)-\lambda_{\tau,s}(D)|} |I_{\pi(D_j)F_\pi(\phi)}(\xi_\tau r_j, s+1)|\,|d\varrho| \\ &\leqslant 2\pi M \sum_j \int_{C_R} |I_{\pi(D_j)F_\pi(\phi)}(\xi_\tau r_j, s+1)|\,|d\varrho|\end{aligned}$$

where $\pi=\pi(\sigma, \varrho, \varepsilon)$ and the D_j are as in Proposition 2.5. Hence $I_R(s)$ converges absolutely for $s \in U$ iff the $I_R(s)$ associated to $D_j(\phi)$ converge absolutely for $s \in U+1$. Applying this repeatedly we may move s into the half-plane of absolute convergence for the $I_R(s)$. This proves the claim. Note in particular that $I_R(s)$ converges absolutely and uniformly on compact subsets, to the right of C_R.

Now let D_R be the region of the ϱ-plane bounded by the curves $\{\varrho=ir: |r| \leqslant R\}$ and $\{\varrho=Re^{i\theta}: \pi/2 \leqslant \theta \leqslant 3\pi/2\}$, see Figure 2.

Let P_R denote the set of poles in ϱ of the integrand $I_{F_{\pi(\sigma,\varrho,\varepsilon)}(\phi)}(\xi_\tau, s)$ in D_R. These are the same as the poles of the c-function $c(\sigma, \varrho)$ in D_R, and hence finite in number. By the Cauchy residue theorem, for $\mathrm{Re}(s) >> 0$ we have

$$I_0(s) = I_R(s) + 2\pi i \sum_{\varrho_j \in P_R} \operatorname*{Res}_{\varrho=\varrho_j} I_{F_{\pi(\sigma,\varrho,\varepsilon)}(\phi)}(\xi_\tau, s). \tag{2.47}$$

As the integral $I_{F_{\pi(\sigma,\varrho,\varepsilon)}(\phi)}(\xi_\tau, s)$ is absolutely convergent, we may move the residue computation inside the integral. Let $\pi'(\varrho_j')$ denote the irreducible quotient of $\pi(\sigma,\varrho,\varepsilon)$ at $\varrho=\varrho_j$. Then we have

$$\operatorname*{Res}_{\varrho=\varrho_j} I_{F_{\pi(\sigma,\varrho,\varepsilon)}(\phi)}(\xi_\tau, s) = I_{F_{\pi'(\varrho_j)}(\phi)}(\xi_\tau, s).$$

Each $I_{F_{\pi'(\varrho_j)}(\phi)}(\xi_\tau, s)$ has a meromorphic continuation to all of $\mathbf{C}$ with poles at $s=\varrho_j-k$, $k=0,1,2,\ldots$, if τ is a constituent of $\sigma\otimes R^k$ by Proposition 2.7 and Remark 2.8. Hence the expression on the right hand side of (2.47) gives a meromorphic continuation of our integral to the region of the plane to the right of C_R with the stated poles. By taking R sufficiently large, this extends (2.46) to any relatively compact region of $\mathbf{C}$. This affords the meromorphic continuation of (2.46) and completes the proof of the theorem. Q.E.D.

Combining Theorem 2.15 with Proposition 2.12 gives the proof of Theorem 2.16 stated in the introduction.

The method of proof here generalizes to give the analytic continuation of a Kloosterman–Selberg zeta function for an arbitrary real reductive group G. We will return to this topic in a future paper.

§ 3. Adelic Poincaré series

The notation used here is the same as in Section 1. $D>0$ is an integer. For $p\neq\infty$

$$K_p = \{g\in G(\mathbf{Z}_p) \mid g\equiv 1 \,(\mathrm{mod}\, p^l)\}$$

where p^l is the highest power of p dividing D. K_p is a compact subgroup of $G(\mathbf{Z}_p)$. For $p\nmid D$ clearly $K_p=G(\mathbf{Z}_p)$. For $p=\infty$, we let K_∞ be the maximal compact subgroup of $G(\mathbf{R})$ defined by

$$K_\infty = SO_{r+1,1}\cap O_{r+2} = \{g\in G(\mathbf{R}) \mid {}^t g g = I_{r+2}\}.$$

Set

$$K_D = \prod_p K_p.$$

If G' is an arbitrary $\mathbf{Q}$-subgroup of G we let

$$G'(D) = \{g\in G'(\mathbf{Z}) \mid g\equiv 1 \bmod D\}.$$

In this notion $\Gamma(D)$ of Section 1 is denoted $G(D)$. Let

$$A^{+}(\mathbf{R}) = \left\{ \begin{pmatrix} y & & \\ & \mathbf{1}_r & \\ & & y^{-1} \end{pmatrix} \middle| \, y > 0 \right\}$$

be the connected component of $A(\mathbf{R})$. Let us make a convention that will be convenient: the element

$$\begin{pmatrix} y & & \\ & \mathbf{1}_r & \\ & & y^{-1} \end{pmatrix}$$

of $A^{+}(\mathbf{R})$ will always be denoted y. The Iwasawa decomposition for $G(\mathbf{R})$ is

$$G(\mathbf{R}) = U(\mathbf{R}) A(\mathbf{R})^{+} K_{\infty}.$$

LEMMA 3.1. *The Iwasawa decomposition of a general element $g \in G(\mathbf{R})$ is given as follows. Put g in block matrix form*

$$g = \begin{array}{c} \\ 1 \\ r \\ 1 \end{array} \!\!\begin{array}{c} \begin{array}{ccc} 1 & r & 1 \end{array} \\ \begin{pmatrix} * & * & * \\ a & b & c \\ d & e & f \end{pmatrix} \end{array} = (g_{ij}).$$

Then

$$g = \begin{pmatrix} 1 & -{}^{t}u & -\tfrac{1}{2}(u, u) \\ & \mathbf{1}_r & u \\ & & 1 \end{pmatrix} \begin{pmatrix} y & & \\ & \mathbf{1}_r & \\ & & y^{-1} \end{pmatrix} k$$

($k \in K_{\infty}$) where

$$y = (d^2 + (e, e) + f^2)^{-1/2} = (g_{r+2,1}^2 + \ldots + g_{r+2,r+2}^2)^{-1/2}$$

$$u = y^2 (d \cdot a + b \cdot {}^{t}e + f \cdot c)$$

Proof. $g = uyk$ implies

$$g\,{}^{t}g = uyk\,{}^{t}k\,{}^{t}y\,{}^{t}u = uy\,{}^{t}(uy)$$

the formulas follow immediately from this equation. Q.E.D.

We review the definition of the Poincaré series from Section 1. Let $\xi \in (1/D)\mathbf{Z}^r \cdot \psi^\xi$ is a character of $U(\mathbf{Q})\backslash U(\mathbf{A})$ by

$$\psi^\xi(u) = \psi((u, \xi))$$

ψ_p^ξ is a character of $U(\mathbf{Q}_p)$ defined analogously so that $\psi^\xi = \Pi_p \psi_p^\xi$. The functions f_p^ξ on $G(\mathbf{Q}_p)$ are defined by

(i) $p=\infty$, then

$$f_\infty^\xi(uyk) = \psi_\infty^\xi(u)\, y^{s+r/2} e^{-2\pi|\xi| y} \quad (u \in U(\mathbf{R}),\ y \in A(\mathbf{R})^+,\ k \in K_\infty),$$

s a complex parameter.

(ii) $p<\infty$,

$$f_p^\xi(uk) = \psi_p^\xi(u) \quad (u \in U(\mathbf{Q}_p),\ k \in K_p)$$

and set f_p^ξ to be 0 outside the open set $U(\mathbf{Q}_p)K_p$.

For $g=(g_\infty, g_2, g_3 \dots) \in G(\mathbf{A})$ put

$$f^\xi(g) = \prod_p f_p^\xi(g_p).$$

Clearly

$$f^\xi(ug) = \psi^\xi(u) f^\xi(g) \quad (u \in U(\mathbf{A}))$$

and hence in particular is left invariant by $U(\mathbf{Q})$.

Definition 3.2. The Poincaré series $P_\xi(g, s)$ is defined to be

$$P_\xi(g, s) = \sum_{\gamma \in U(\mathbf{Q})\backslash G(\mathbf{Q})} f^\xi(\gamma g).$$

LEMMA 3.3. *The above series converges absolutely for* $\mathrm{Re}(s)>r/2$, *uniformly for g and s in compact subsets.*

Proof. It is both efficient and enlightening to prove this by comparing the series with a particular Eisenstein series. If $g_p \in G(\mathbf{Q}_p)$ we may write

$$g_p = \begin{pmatrix} a & * & * \\ 0 & h & * \\ 0 & 0 & a^{-1} \end{pmatrix} \cdot k$$

with $k \in G(\mathbf{Z}_p)$ if p is finite and $k \in K_\infty$ if $p=\infty$. Define $\phi_p(g_p)$ to be $|a|_p$ in this factorization. For $g \in G(\mathbf{A})$ we set

$$\phi(g) = \prod_p \phi_p(g_p).$$

The Eisenstein series is defined by

$$E(g, s) = \sum_{\gamma \in P(\mathbf{Q})\backslash G(\mathbf{Q})} \phi(\gamma g)^{s+r/2}. \tag{3.0}$$

From the general theory of Eisenstein series [16] we know that the statement of Lemma 3.3 is true for $E(g, s)$.

We have

$$G(\mathbf{A}) = G(\mathbf{Q})\, G(\mathbf{R}) \prod_{p<\infty} G(Z_p). \tag{3.1}$$

Thus in Definition 3.2 it suffices to take $g=g_\infty \cdot k$ with $g_\infty \in G(\mathbf{R})$ and $k \in \Pi_{p<\infty} G(\mathbf{Z}_p)$. Since f^ξ is also right invariant by K_∞, we may even take g to be

$$z = x \cdot y = \begin{pmatrix} 1 & -{}^t x & -\frac{1}{2}(x,x) \\ & \mathbf{1}_r & x \\ & & 1 \end{pmatrix} \begin{pmatrix} y & & \\ & \mathbf{1}_r & \\ & & y^{-1} \end{pmatrix} \tag{3.1'}$$

$(x \in \mathbf{R}^r, y>0)$ which is naturally a variable on the hyperbolic $(r+1)$-space $H^{r+1} \cong G/K_\infty$.

Let

$$w = \begin{pmatrix} & & 1 \\ & \varepsilon_r & \\ 1 & & \end{pmatrix} \quad \text{with} \quad \varepsilon_r = \begin{pmatrix} -1 & 0 \\ 0 & \mathbf{1}_{r-1} \end{pmatrix}.$$

The Bruhat decomposition for $G(\mathbf{Q})$ reads

$$G(\mathbf{Q}) = P(\mathbf{Q}) \cup P(\mathbf{Q})\, wU(\mathbf{Q}).$$

Hence we may write $P_\xi(g, s)$ as

$$P_\xi(g, s) = \sum_{a,h} f^\xi(ah \cdot g) + \sum_{a,h,u} f^\xi(ahwug) \tag{3.2}$$

$(a \in A(\mathbf{Q}),\ h \in H(\mathbf{Q}),\ u \in U(\mathbf{Q}))$. From the definition of f^ξ we see that $f^\xi(ah \cdot g) \neq 0$ only if

$ahk_p \in K_p$ for all $p<\infty$, and in such a case we have

$$|f^{\xi}(ah \cdot g)| = y^{\sigma+r/2} e^{-2\pi|\xi| y} \quad (\sigma = \mathrm{Re}(s)).$$

It follows that the right hand side of (3.2) is dominated by

$$c_1 y^{\sigma+r/2} e^{-cy} + \sum_{a,h,u} |f^{\xi}(ahwu \cdot z)| \quad (c = 2\pi|\xi|) \tag{3.3}$$

with a positive constant c_1. Since

$$f^{\xi}(ahwug) = f^{\xi}_{\infty}(ahwuz) \prod_{p<\infty} f^{\xi}_{p}(ahwuk_p)$$

we have $f^{\xi}(ahwug) \neq 0$ iff $f^{\xi}_{p}(ahwuk_p) \neq 0$ for all $p \neq \infty$ and in such a case

$$|f^{\xi}(ahwug)| = |f^{\xi}_{\infty}(ahwuz)| \leqslant |a|_{\infty}^{\sigma+r/2} \phi_{\infty}(wuz)^{\sigma+r/2}. \tag{3.4}$$

Now $f^{\xi}_{p}(ahwuk_p) \neq 0$ means that there is a $v \in U(\mathbf{Q}_p)$ such that $vahwuk_p \in K_p$, or

$$wu = (vah)^{-1} k'_p \quad (k'_p \in G(\mathbf{Z}_p))$$

this implies

$$\phi_p(wu) = |a|_p^{-1}.$$

The product formula gives

$$|a|_{\infty} = \prod_{p<\infty} |a|_p^{-1} = \prod_{p<\infty} \phi_p(wu).$$

Therefore (3.4) can be written

$$|f^{\xi}(ahwug)| \leqslant \phi(wuz)^{\sigma+r/2}. \tag{3.5}$$

We now need the following crucial lemma.

LEMMA 3.4. (a) *For $p \neq \infty$, given $g_p \in G(\mathbf{Q}_p)$, the condition*

$$f^{\xi}_{p}(ahg_p) \neq 0 \quad (a \in A(\mathbf{Q}_p), h \in H(\mathbf{Q}_p))$$

determines a (respectively h) up to a multiplication on the left by $A(\mathbf{Q}_p) \cap K_p$ (respectively $H(\mathbf{Q}_p) \cap K_p$).

(b) *Given* $g \in G(\mathbf{A})$ *the condition* $f^{\xi}(ahg) \neq 0$, ($a \in A(\mathbf{Q})$, $h \in H(\mathbf{Q})$) *determines* ah *up to multiplication on the left by* $M(D)$.

Proof. Clearly (b) follows from (a). To prove (a) note that $f_p^{\xi}(ahg_p) \neq 0$ means that there is a $v \in U(\mathbf{Q}_p)$ such that $vahg_p \in K_p$. If also $f_p^{\xi}(a'h'g_p) \neq 0$ then $v'a'h'g_p \in K_p$ for some $v' \in U(\mathbf{Q}_p)$. Thus $v'a'h'(ah)^{-1}v^{-1} = v'a'h'g_p(vahg_p)^{-1} \in K_p$. Since $M = AH$ normalizes U, the condition $v'a'h'(ah)^{-1}v^{-1} \in K_p$ implies $a'h'(ah)^{-1} \in K_p$. The lemma follows. Q.E.D.

Combining (3.3), (3.5) and Lemma 3.4 we find that the right hand side of (3.2) is bounded by a constant multiple of

$$y^{\sigma+r/2}e^{-cy} + \sum_{u \in U(\mathbf{Q})} \phi(wuz)^{\sigma+r/2}. \tag{3.6}$$

On the other hand, the Bruhat decomposition also enables us to write

$$E(z, \sigma) = y^{\sigma+r/2} + \sum_{u \in U(\mathbf{Q})} \phi(wuz)^{\sigma+r/2}.$$

Comparing this with (3.6), we see that Lemma 3.3 follows from familiar facts about Eisenstein series [16]. Q.E.D.

We turn now to the spectral analysis and analytic continuation of $P_{\xi}(g, s)$. For simplicity we will treat only the case $D=1$ in detail. $D>1$ can be handled similarly, but the book keeping involved is more tedious. So we set $\Gamma = \Gamma(1)$. We have seen that for $\mathrm{Re}(s) > r/2$, $P_{\xi}(g, s)$ is left $G(\mathbf{Q})$ and right $K = \Pi_p K_p$ invariant. Hence as in (3.1) it may be considered as a function on $G(\mathbf{Q})\backslash G(\mathbf{A})/K \cong \Gamma\backslash H^{r+1}$. When viewed as such we write it as $P_{\xi}(z, s)$, $z \in H^{r+1}$ as in (3.1′). Our first calculation is that of the spectral components of $P_{\xi}(z, s)$. We see from the proof of convergence that $P_{\xi}(z, s)$ is of moderate growth (since the Eisenstein series are). The spectral calculation will allow us to conclude much more.

The spectrum of Δ on $\Gamma\backslash H^{r+1}$ consists of a discrete set of $L^2(\Gamma\backslash H^{r+1})$ eigenfunctions as well as the unitary Eisenstein series [3] and [16]. Let $u_0(z)$, $u_1(z) \ldots$ be an orthonormal basis spanning the discrete spectrum. We write $\lambda_0 < \lambda_1 \leq \lambda_2 \ldots$ with

$$\begin{cases} \Delta u_j + \lambda_j u_j = 0 \\ \lambda_j = \left(\dfrac{r}{2} - \nu_j\right)\left(\dfrac{r}{2} + \nu_j\right). \end{cases} \tag{3.7}$$

Note that $\nu_0=r/2$ corresponds to the constant eigenfunction $u_0(z)$. Since all but a finite number of the λ_j's are less than $(r/2)^2$ it follows that all but a finite number of the ν_j's are purely imaginary. Also the Weyl asymptotics for the numbers λ_j [3] ensures that

$$|\{j\colon |\nu_j|\leq T\}| = O(T^{r+1}). \tag{3.8}$$

The rest of the spectrum is spanned by Eisenstein series. Since $D=1, \Gamma\backslash H^{r+1}$ has only one cusp. We have for this cusp the Eisenstein series $E(z, s)$ [3], [16]. The Eisenstein series are holomorphic on $\mathrm{Re}(s)=0$ and the functions $E(z, it)$, $t\in\mathbf{R}$ furnish the rest of the L^2 spectrum [3], [16].

To compute the inner products $\langle P_\xi(\cdot, s), u_j\rangle$ and $\langle P_\xi(\cdot, s), E(\cdot, it)\rangle$ we need the Fourier developments of these eigenfunctions. Let $L=\Gamma^\infty=\Gamma\cap U(\mathbf{R})$ as usual. Clearly $L\cong\mathbf{Z}^r$. The functions $u_j(z)$ and $E(z, it)$, being L periodic, may be developed in a Fourier series. A simple calculation (separating variables) shows using (3.7), (see [3]), that

$$u_j(z) = \sum_{\eta\in L^*}{}' \varrho_j(\eta)\, y^{r/2} K_{\nu_j}(2\pi|\eta|y)\, e((\eta, x)) + \text{possible zero coefficient term} \tag{3.9}$$

and

$$E(z, w) = y^{r/2+w} + c_\infty(w)\, y^{r/2-w} + y^{r/2} \sum_{\eta\in L^*}{}' c(w, \eta)\, K_w(2\pi|\eta|\, y)\, e((\eta, x)) \tag{3.10}$$

where $e(z)=e^{2\pi iz}$ and $K_\nu(y)$ is the Bessel function

$$K_\nu(y) = \int_0^\infty e^{-z\cosh t}\cosh \nu t\, dt.$$

PROPOSITION 3.5. *For* $\mathrm{Re}(s)$ *large the inner products* $\langle P, u_j\rangle$ *and* $\langle P, E\rangle$ *converge and*

$$\langle P_\xi(\cdot, s), u_j\rangle = \frac{\overline{\varrho_j(\xi)}\,\Gamma(s-\nu_j)\,\Gamma(s+\nu_j)}{(4\pi|\xi|)^s\,\Gamma(s+1/2)} \tag{3.11}$$

$$\langle P_\xi(\cdot, s), E(\cdot, it)\rangle = \frac{\overline{c(it, \xi)}\,\Gamma(s-it)\,\Gamma(s+it)}{(4\pi|\xi|)^s\,\Gamma(s+1/2)}. \tag{3.12}$$

Proof. The functions $u_j(z)$ and $E(z, it)$ defined on H^{r+1}, have unique extensions to $G(\mathbf{A})$ so as to be $G(\mathbf{Q})$ automorphic and right K invariant. We denote these extensions by $u_j(g)$ and $E(g, it)$. Consider first $\langle P, u_j\rangle$.

$$\langle P_\xi(\cdot, s), u_j\rangle = \int_{\Gamma\backslash H^{r+1}} P_\xi(z, s)\overline{u_j(z)}\,dz \tag{3.13}$$

$$= \int_{G(\mathbf{Q})\backslash G(\mathbf{A})} P_\xi(g, s)\overline{u_j(g)}\,dg$$

$$= \int_{G(\mathbf{Q})\backslash G(\mathbf{A})} \sum_{\gamma\in U(\mathbf{Q})\backslash G(\mathbf{Q})} f^\xi(\gamma g)\overline{u_j(g)}\,dg$$

$$= \int_{U(\mathbf{Q})\backslash G(\mathbf{A})} f^\xi(g)\overline{u_j(g)}\,dg$$

$$= \int_{U(\mathbf{A})\backslash G(\mathbf{A})}\int_{U(\mathbf{Q})\backslash U(\mathbf{A})} f^\xi(g)\,\psi^\xi(u)\,\overline{u_j(ug)}\,du\,dg$$

$$= \int_{U(\mathbf{A})\backslash G(\mathbf{A})} f^\xi(g)\overline{u_{j,\xi}(g)}\,dg \tag{3.14}$$

where

$$u_{j,\xi}(g) = \int_{U(\mathbf{Q})\backslash U(\mathbf{A})} u_j(ug)\,\psi(-(u,\xi))\,du$$

by weak approximation for $U(\mathbf{A})$ the last

$$= \int_{L\backslash U(\mathbf{R})} u_j(u_\infty z)\,\psi(-(u,\xi)_\infty))\,du$$

$$= \varrho_j(\xi)\,y^{r/2}K_{\nu_j}(2\pi|\xi|y), \quad \text{using 3.9.}$$

Hence returning to (3.14) we have

$$\langle P_\xi(\cdot, s), u_j\rangle = \overline{\varrho_j(\xi)}\int_0^\infty f^\xi(y)\,\overline{y^{r/2}K_{\nu_j}(2\pi|\xi|y)}\,\frac{dy}{y^{r+1}}$$

$$= \overline{\varrho_j(\xi)}\int_0^\infty y^s e^{-2\pi|\xi|y}\,\overline{K_{\nu_j}(2\pi|\xi|y)}\,\frac{dy}{y}.$$

The last integral may be evaluated, see Gradshteyn–Ryzhik [7, p. 712] giving

$$= \overline{\varrho_j(\xi)}\,\frac{\Gamma(s-\nu_j)\,\Gamma(s+\nu_j)}{(4\pi|\xi|)^s\,\Gamma(s+1/2)}.$$

This verifies (3.11) except for a word of explanation concerning the interchange of integrals above. In fact one can easily check the absolute convergence of (3.13) above.

Carrying out the computations above we find that it is dominated by

$$\int_0^\infty e^{-cy} y^\sigma \max_x |u_j(x,y)| \frac{dy}{y^{r+1}}$$

which is finite for σ large. The calculation with the Eisenstein series is identical. Q.E.D.

LEMMA 3.6. *There is N_0 depending on r only such that*

$$|\varrho_j(\xi)| \ll e^{\pi|\nu_j|/2} |\nu_j|^{N_0}$$

$$|c(it,\xi)| \ll e^{\pi|t|/2} |t|^{N_0}.$$

Proof. From (3.9) we have

$$|\varrho_j(\xi) K_{\nu_j}(2\pi|\xi| y)| \ll \int_{L\setminus U(\mathbf{R})} |u_j(x,y)|\, dx.$$

Hence for $|\nu_j|$ large;

$$|\varrho_j(\xi)|^2 \int_{|\nu_j|-2|\nu_j|^{1/2}}^{|\nu_j|-|\nu_j|^{1/2}} |K_{\nu_j}(y)|^2 \frac{dy}{y^{r+1}} \ll \int_{L\setminus U(\mathbf{R})} \int_{A_0}^\infty |u_j|^2 \frac{dx\,dy}{y^{r+1}} \ll 1. \tag{3.15}$$

Now for y in this range we have

$$K_{\nu_j}(y) \sim \frac{1}{3} \pi e^{-\pi|\nu_j|/2} \left(\frac{2(i\nu_j - y)}{y}\right)^{1/2} \cdot [J_{1/3}(y) + J_{-1/3}(y)]$$

see Magnus–Oberhettinger–Soni [19, p. 142]. From this and (3.15) we conclude

$$|\varrho_j(\xi)| \ll |\nu_j|^{N_0} e^{\pi|\nu_j|/2}.$$

The estimation of $c(it,\xi)$ is similar. Q.E.D.

LEMMA 3.7. *For z lying in a compact subset of H^{r+1} we have the uniform estimates*

$$|u_j(z)| \ll (|\nu_j|+1)^{2r}$$

$$|E(z,it)| \ll (|t|+1)^{2r}$$

where the implied constants depend on the compact set.

Proof. This follows easily from the equations $\Delta u_j+\lambda_j u_j=0$ and $\Delta E+s(r-s)E=0$, see Cohen–Sarnak [3]. Q.E.D.

We can prove the basic proposition concerning the spectral expansion of $P_\xi(g, s)$.

PROPOSITION 3.8. *The function $P_\xi(g,s)$ has a meromorphic continuation to* $\mathrm{Re}(s)>0$ *with series representation*

$$P_\xi(g,s)=\sum_{j=0}^{\infty}\frac{\overline{\varrho_j(\xi)}\,\Gamma(s-\nu_j)\,\Gamma(s+\nu_j)}{(4\pi|\xi|)^s\,\Gamma(s+1/2)}u_j(g)+\frac{1}{4\pi}\int_{-\infty}^{\infty}\frac{\overline{c(it,\xi)}\,\Gamma(s-it)\,\Gamma(s+it)}{(4\pi|\xi|)^s\,\Gamma(s+1/2)}E(g,it)\,dt.$$

The above series and integral converge absolutely and uniformly on compact subsets of $G(\mathbf{A})$ *and of* $\mathrm{Re}(s)>0$. *In particular the possible poles (which are simple) of $P_\xi(g,s)$ in* $\mathrm{Re}(s)>0$ *are at $s=\nu_j$, $\nu_j>0$, that is the exceptional spectrum. The corresponding residue is*

$$\frac{\overline{\varrho_j(\xi)}\,\Gamma(2\nu_j)}{(4\pi|\xi|)^{\nu_j}\,\Gamma(\frac{1}{2}+\nu_j)}u_j(g).$$

Proof. The convergence of the right hand side above follows from the estimates in Lemmas 3.6 and 3.7, Weyl's law (3.8) and Stirlings series

$$|\Gamma(c+it)|\sim e^{-\pi/2|t|}|t|^{\sigma-1/2}\sqrt{2\pi}\quad\text{as}\quad |t|\to\infty.$$

Now for $\mathrm{Re}(s)>r/2$ the right hand side $H(z,s)$ represents a smooth L^2 function whose components alone $u_j(z)$ and $E(z, it)$ are identical with $P_\xi(z,s)$. It follows that

$$\int_{\Gamma\backslash H^{r+1}}(H(z,s)-P_\xi(z,s))\,\phi(z)\,dz=0$$

for any $\phi\in L^2(\Gamma\backslash H^{r+1})$ of compact support. Clearly then $H(z,s)\equiv P_\xi(z,s)$ for all s and z, proving the proposition. Q.E.D.

Note that $\nu_0=r/2$ is not a pole of $P_\xi(g,s)$ since $\varrho_0(\xi)=0$.

COROLLARY 3.9. *For $y\in A(\mathbf{R})^+$ the function*

$$F(\xi,s)=\int_{U(\mathbf{Q})\backslash U(\mathbf{A})}P_\xi(uy,s)\,\psi^\xi(-u)\,du$$

is meromorphic in $\mathrm{Re}(s)>0$ *and has simple poles at the exceptional spectrum ν_j with* *residues*

$$\left(\sum_{\nu_j'=\nu_j} |\varrho_{j'}(\xi)|^2\right) \cdot \frac{\Gamma(2\nu_j)\, y^{r/2} K_{\nu_j}(2\pi|\xi|y)}{(4\pi|\xi|)^{\nu_j}\Gamma(\frac{1}{2}+\nu_j)}.$$

The important point to note here is that if ν_j is in the exceptional spectrum then for some ξ, $F(\xi, s)$ does have a pole at ν_j.

If $D>1$ the situation is essentially the same. For now we will have

$$G(\mathbf{Q})\backslash G(\mathbf{A})/K = \bigcup_{i=1}^{h} \Gamma_i\backslash H^{r+1}$$

and each Γ_i is a congruence subgroup of $\Gamma(1)$ and may have more than one cusp. Taking $\tilde{\Gamma}=\bigcap_{i=1}^{h}\Gamma_i$ one has that $P_\xi(g, s)$ lives on $\tilde{\Gamma}\backslash\mathbf{H}^{r+1}$ and the spectral expansion and location of poles is similar.

PROPOSITION 3.10. *For any level $D\geq 1$ and for $y\in A(\mathbf{R})^+$ the function*

$$F(\xi, s) = \int_{U_{\mathbf{Q}}\backslash U_A} P_\xi(uy, s)\, \psi^\xi(-u)\, du$$

is meromorphic in $\mathrm{Re}(s)>0$ *and has possible simple poles at the exceptional spectrum ν_j. If ν_j is in the exceptional spectrum, then for some ξ, $F(\xi, s)$ does have a pole at ν_j.*

§ 4. Fourier coefficients and local Kloosterman integrals

In this section we compute the Fourier coefficient $F(\xi, s)$ of Proposition 3.10, in terms of Kloosterman integrals. We have

$$F(\xi, s) = \int_{U(\mathbf{Q})\backslash U(\mathbf{A})} P_\xi(uy, s)\, \psi^\xi(-u)\, du.$$

Unfolding $P_\xi(uy, s)$ using the Bruhat decomposition gives

$$F(\xi, s) = \sum_{\gamma\in M(\mathbf{Q})} \int_{U(\mathbf{Q})\backslash U(\mathbf{A})} f^\xi(\gamma uy)\, \psi^\xi(-u)\, du + \sum_{\gamma\in M(\mathbf{Q})} \int_{U(\mathbf{A})} f^\xi(\gamma wuy)\, \psi^\xi(-u)\, du. \tag{4.1}$$

Since $f^\xi(\gamma uy)\neq 0$ only when $\gamma\in M(D)$, the first sum is simply

$$y^{s+r/2}e^{-cy}\sum_{\gamma\in M(D)}\int_{U(\mathbf{Q})\backslash U(\mathbf{A})}\psi^{\xi}(\gamma(u)-u)\,du$$

(where $\gamma(u)=\mathrm{Ad}(\gamma)(u)$). This last expression is clearly entire in s and plays no further role in our discussion.

We denote the second term in (4.1) by $\tilde{Z}(y,s)$ (we drop the notational dependence on ξ which is fixed). As in the last section we may write $\gamma\in M(\mathbf{Q})$ as

$$\gamma=ah' \quad (a\in A(\mathbf{Q}),\ h'\in H(\mathbf{Q})).$$

Put

$$h=h'\varepsilon_r \quad \left(\text{recall that } w=\begin{pmatrix} & & 1\\ & \varepsilon_r & \\ 1 & & \end{pmatrix},\ \varepsilon_r=\begin{pmatrix}-1 & \\ & \mathbf{1}_{r-1}\end{pmatrix}\right)$$

so that $h\in O_r(\mathbf{Q})$.

Using Lemma 3.1 we find that

$$\begin{aligned}I_\gamma(y,s)&\overset{\Delta}{=}\int_{U(\mathbf{R})}f^{\xi}_{\infty}(\gamma wuy)\,\psi^{\xi}_{\infty}(-u)\,du\\ &=\int_{\mathbf{R}^r}\left(\frac{|a|y}{y^2+\frac{1}{2}(u,u)}\right)^{s+r/2}\exp\left(-\frac{c|a|y}{y^2+\frac{1}{2}(u,u)}\right)e\left(-(u,\xi)-\frac{a(hu,\xi)}{y^2+\frac{1}{2}(u,u)}\right)du.\end{aligned} \tag{4.3}$$

For $p<\infty$ we let

$$\mathrm{Kl}_p(\gamma)=\int_{U(\mathbf{Q}_p)}f^{\xi}_{p}(\gamma wu)\,\psi^{\xi}_{p}(-u)\,du \tag{4.4}$$

and put

$$\mathrm{Kl}(\gamma)=\prod_{p\neq\infty}\mathrm{Kl}_p(\gamma)$$

then

$$\tilde{Z}(y,s)=\sum_{\gamma\in M(\mathbf{Q})}I_\gamma(y,s)\,\mathrm{Kl}(\gamma). \tag{4.5}$$

We first compute the "local Kloosterman integrals" (4.4). These turn out to be generalized Kloosterman sums. The integrand (4.4) is zero unless there is $v\in U(\mathbf{Q}_p)$

such that $v\gamma wu \in K_p$. Direct calculation gives

$$(4.6)\qquad v\gamma wu = \begin{pmatrix} -\frac{1}{2a}(v,v) & -{}^t vh + \frac{(v,v)}{2a}{}^t u & a + \frac{1}{4a}(u,u)(v,v) - (v,hu) \\ \frac{1}{a}v & h - \frac{1}{a}v{}^t u & hu - \frac{(u,u)}{2a}v \\ \frac{1}{a} & -\frac{1}{a}{}^t u & -\frac{1}{2a}(u,u) \end{pmatrix}.$$

In what follows we use the notation that for any matrix over $\mathbf{Q}_p$, we let $|\ |_p$ denote the maximum of the absolute values of its entries.

We now look at the conditions on γ and u which are forced by the requirement that $v\gamma wu \in K_p$. A glance at the last row of the matrix (4.6) shows that

$$(4.7)\qquad |a|_p \geq 1,\quad |u|_p \leq |a|_p,\quad |(u,u)|_p \leq |2a|_p \quad \text{and also} \quad |v|_p \leq |a|_p.$$

From the center entry, one has

$$h = \frac{1}{a}v{}^t u + z$$

with $z \in M_r(\mathbf{Z}_p)$. Therefore

$$|h|_p \leq \max\left(\frac{|u|_p |v|_p}{|a|_p}, |z|_p\right).$$

By the above,

$$\frac{|u|_p |v|_p}{|a|_p} \leq \frac{|a|_p^2}{|a|_p} = |a|_p$$

and by assumption $|z|_p \leq 1$. Hence $|z|_p \leq \max(|a|_p, 1)$. So if $|a|_p > 1$ we have $|h|_p \leq |a|_p$. Since ${}^t hh = 1$, we always have $|h|_p \geq 1$. Hence we have three cases: (I) $|a|_p = 1$; (II) $|a|_p > |h|_p \geq 1$; (III) $|a|_p = |h|_p > 1$.

Let us first assume $p \nmid D$, $p \neq 2$ and examine the various cases separately:

Case I. $|a|_p = 1$.

Looking at (4.6) we see that $\gamma \in M(\mathbf{Z}_p)$, $u \in U(\mathbf{Z}_p)$, $v \in U(\mathbf{Z}_p)$. In this case we have

$$\mathrm{Kl}_p(\gamma) = 1.$$

Note that for a given $\gamma \in M(\mathbf{Q})$ this is the case for almost all p.

Case II. $|a|_p > |h|_p$.

(Since ${}^t hh = I_r \Rightarrow |h|_p \geq 1$, the condition $|a|_p > |h|_p$ implicitly implies $|a|_p > 1$.) In this case we claim that either $|u|_p < |a|_p$ or $|v|_p < |a|_p$. For otherwise we would have

$$|u|_p = |v|_p = |a|_p \quad \text{(see 4.7)}$$

and this implies

$$\left|\frac{1}{a} v^t u\right|_p = \frac{|v|_p |u|_p}{|a|_p} = |a|_p > 1.$$

But $h - (1/a)v^t u$ (which is the central block in the matrix (4.6)) is integral. Hence

$$|h|_p = \left|\frac{1}{a} v^t u\right|_p = |a|_p$$

contrary to our assumption.

Let's assume that $|u|_p < |a|_p$. (If we start with the assumption $|v|_p < |a|_p$ it will any way turn out that $|u|_p < |a|_p$ holds.) Looking at the last row in (4.6) we see that $(u, u)/2a$ must be a unit in $\mathbf{Z}_p$. Since $hu - (u, u)v/2a$ (which is located in the 3rd column of 4.6) must be integral, we may conclude

$$v \equiv \frac{2a}{(u,u)} hu \quad (\operatorname{mod} \mathbf{Z}_p^r). \tag{4.8}$$

Hence

$$h - \frac{1}{a} v^t u = h\left(1 - \frac{2}{(u,u)} u^t u\right) \quad \text{modulo integral part.}$$

The element

$$R(u) = 1 - \frac{2}{(u,u)} u^t u$$

is a reflection in $\mathbf{Q}_p^r$ determined by the vector u. So we must have $hR(u) \in H(\mathbf{Z}_p)$. Summing up we obtain the following conditions

$$\begin{cases} \text{(i)} \left|\dfrac{(u,u)}{2a}\right|_p = 1, \\ \text{(ii)} \ R(u) \in H(\mathbf{Z}_p)\, h \end{cases} \tag{4.9}$$

(since $R(u)^2 = 1$, the condition $hR(u) \in H(\mathbf{Z}_p)$ is equivalent to $h \in H(\mathbf{Z}_p) R(u)$ which in turn is equivalent to $R(u) \in H(\mathbf{Z}_p) h$, etc.). One checks that once (4.8) and (4.9) are satisfied then indeed $v\gamma wu \in K_p$. We have

LEMMA 4.1. *If* $|h|_p<|a|_p$, *then*

$$\mathrm{Kl}_p(\gamma) = \int \psi_p^\xi\left(u+\frac{2a}{(u,u)}hu\right)du$$

where the integral is over the set of u's described by (4.9).

We need an upper bound for $|\mathrm{Kl}_p(\gamma)|$. The "trivial" estimation of this integral is an estimate for the volume of the domain of integration.

LEMMA 4.2. *Suppose that u satisfies* (4.9) *then*

$$|u|_p^2 = |a|_p\,|h|_p. \tag{4.10}$$

Indeed (4.9)(ii) implies that

$$|h|_p = |R(u)|_p = \frac{|u|_p^2}{|(u,u)|_p}.$$

Together with (4.9)(i) this implies (4.10). Q.E.D.

LEMMA 4.3. *Let* $|h|_p=p^k$. *Fix an element* u_0 *satisfying* (4.9). *Then the set of u's satisfying* (4.9) *can be described by*

$$\begin{cases} \text{(i) } |(u,u)|_p = |2a|_p = |a|_p \textit{ since } p \neq 2, \\ \text{(ii) } u = \lambda u_0 + u' \textit{ where } \lambda \textit{ runs through a set of} \\ \quad \textit{representatives of units in } \mathbf{Z}_p \textit{ modulo } p^k, \\ \quad \textit{and } |u'|_p \leq |a|_p^{1/2}\,|h|_p^{-1/2}. \end{cases} \tag{4.11}$$

Proof. If $|h|_p=1$ then (4.11) comes down to $|(u,u)|_p=|a|_p$ and $|u|_p\leq|a|_p^{1/2}$, which is easily seen to be the same as (4.9) in this case.

Now let $|h|_p>1$. If u satisfies (4.9) then

$$R(u_0) \in H(\mathbf{Z}_p)\,R(u) \quad \text{i.e.} \quad R(u_0)\,R(u) \in H(\mathbf{Z}_p).$$

We have

$$\begin{aligned} R(u_0)\,R(u) &= \left(1-\frac{2u_0\,{}^tu_0}{(u_0,u_0)}\right)\left(1-\frac{2u\,{}^tu}{(u,u)}\right) \\ &= 1-\frac{2u_0\,{}^tu_0}{(u_0,u_0)}-\frac{2u\,{}^tu}{(u,u)}+\frac{4(u_0,u)\,u_0\,{}^tu}{(u_0,u_0)\,(u,u)}. \end{aligned}$$

Let us write u_0^j, u^j for the jth entry of u_0 and u respectively. Choose j so that $|u^j|_p=|u|_p$. The condition $R(u_0)R(u)\in H(\mathbf{Z}_p)$ implies in particular

$$\left|-\frac{2u_0u_0^j}{(u_0,u_0)}-\frac{2uu^j}{(u,u)}+\frac{4(u_0,u)\,u_0u^j}{(u_0,u_0)(u,u)}\right|_p\leqslant 1.$$

The last inequality can be written

$$|u-\lambda u_0|_p\leqslant\frac{|(u,u)|_p}{|2u^j|_p}\tag{4.12}$$

where

$$\lambda=\frac{(u,u)}{2u^j}\left(\frac{4(u_0,u)\,u^j}{(u_0,u_0)(u,u)}-\frac{2u_0^j}{(u_0,u_0)}\right).$$

Since $|2u^j|_p=|u^j|_p=|u|_p$, (4.12) can be written

$$|u-\lambda u_0|_p\leqslant\frac{|(u,u)|_p}{|u|_p}=|a|_p^{1/2}\,|h|_p^{-1/2}.\tag{4.13}$$

A moment's thought shows that (4.10), (4.13) together with $|h|_p>1$ imply $|\lambda|_p=1$ i.e. λ is a unit in $\mathbf{Z}_p$. Furthermore if $\lambda'\equiv\lambda\ (\mathrm{mod}\,p^k)$ then

$$\begin{aligned}|\lambda'u_0-\lambda u_0|_p&=|\lambda'-\lambda|_p\,|u_0|_p\\&\leqslant|h|_p^{-1}(|a|_p\,|h|_p)^{+1/2}\\&=|a|_p^{1/2}\,|h|_p^{-1/2}\end{aligned}$$

i.e. $\lambda'u_0=\lambda u_0+u'$ with

$$|u'|_p\leqslant|a|_p^{1/2}\,|h|_p^{-1/2}.$$

We have shown that (4.9)$\Rightarrow$(4.11). The other direction is easier and left to the reader.

Q.E.D.

COROLLARY 4.4. *In Case* II

$$|\mathrm{Kl}_p(\gamma)|\leqslant|h|_p^{1-r/2}\cdot|a|_p^{r/2}.\tag{4.14}$$

Indeed this is essentially the volume bound. We turn to

Case III. $|h|_p=|a|_p>1$.

Since $|h|_p = |a|_p > 1$ the integrality of $h-(1/a)\,v\,{}^t u$ implies

$$|h|_p = \left|\frac{1}{a} v\,{}^t u\right|_p = \frac{|v|_p\,|u|_p}{|a|_p}.$$

Together with (4.7) we see that

$$|u|_p = |v|_p = |a|_p.$$

Let

$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_r \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ \vdots \\ v_r \end{pmatrix}$$

$$h_i = i\text{th row of } h$$

$$h^j = j\text{th column of } h$$

$$h_{ij} = i,j \text{ entry of } h.$$

With this notation the integrality of $h-(1/a)\,v^t u$ is expressed as

$$h_{ij} - \frac{v_i u_j}{a} \in \mathbf{Z}_p \quad \text{for all } i,j. \tag{4.15}$$

Fix k, l so that $|h_{kl}|_p = |a|_p$. Then also

$$|v_k|_p = |u_l|_p = |a|_p.$$

Set $\lambda = a/v_k$. We have $|\lambda|_p = 1$ and (4.15) implies

$$\begin{cases} \text{(i) } u = \lambda\,{}^t h_k \mod \mathbf{Z}_p^r \\ \text{(ii) } v = \dfrac{1}{\lambda}\dfrac{a}{h_{kl}} h^l \mod \mathbf{Z}_p^r. \end{cases} \tag{4.16}$$

Substituting these into (4.15) we find that

$$h - \frac{1}{h_{kl}} h^l h_k \quad \text{is integral.} \tag{4.17}$$

One verifies that the conditions (4.16)–(4.17) imply that $v\gamma w u \in K_p$. Hence we have

LEMMA 4.5. $\mathrm{Kl}_p(\gamma)$ *is non-zero only when* (4.17) *is satisfied, and in such a case we* *have*

$$\mathrm{Kl}_p(\gamma) = \sum_{\lambda\in(\mathbf{Z}_p/a^{-1}\mathbf{Z}_p)^*} \psi_p^{\xi}\left(\lambda\, {}^t h_k + \frac{1}{\lambda}\frac{a}{h_{kl}} h^l\right). \tag{4.18}$$

For later purposes it is convenient to state the preceeding lemma another way.

LEMMA 4.6. *Suppose that* $|h|_p=|a|_p>1$. *Let* $u_0\in U(\mathbf{Q}_p)$ *be such that* $f_p^{\xi}(\gamma w u_0)\neq 0$ *then for any* $u\in U(\mathbf{Q}_p)$ *we have* $f_p^{\xi}(\gamma w u)\neq 0$ *iff*

$$u=\lambda u_0 \mod \mathbf{Z}_p^r, \quad \lambda\in \mathbf{Z}_p^*. \tag{4.19}$$

There is $v_0\in U(\mathbf{Q}_p)$ *such that* $v_0\gamma w u_0\in K_p$ *the class of* $v_0 \bmod \mathbf{Z}_p^r$ *is determined by* u_0, *we have* $|u_0|_p=|v_0|_p=|a|_p$ *and*

$$\mathrm{Kl}_p(\gamma) = \sum_{\lambda\in(\mathbf{Z}_p/a^{-1}\mathbf{Z}_p)^*} \psi_p^{\xi}(\lambda u_0+\lambda^{-1}v_0). \tag{4.20}$$

Finally we take up the remaining primes; $p|D$ or $p=2$. We can assume that $2|D$. The situation here is similar to Case II above (i.e. we will need only the volume estimate) and we state the results only.

LEMMA 4.7. *Let* p^l $(l>0)$ *be the highest power of* p *dividing* D, *then* $f_p^{\xi}(\gamma w u)$ *is non-zero only when*

$$|a|_p \geq |2|_p p^{2l}, \quad |h|_p \leq \max\left(1, \frac{|a|_p}{p^{2l}}\right) \tag{4.21}$$

and

$$\begin{cases} \text{(i) } \dfrac{(u,u)}{2a} \equiv -1 \pmod{p^l} \\ \text{(ii) } hR(u)\in H(\mathbf{Q}_p)\cap K_p. \end{cases} \tag{4.22}$$

The relation $|h|_p=\max(1, |u|_p^2/|a|_p)$ *holds. When* $p\neq 2$ *we have* $|h|_p=|u|_p^2/|a|_p)$ *as in* (4.10). *If* (4.21) *and* (4.22) *hold, then*

$$\mathrm{Kl}_p(\gamma) = \int \psi_p^{\xi}\left(u+\frac{2a}{(u,u)}hu\right) du \tag{4.23}$$

where the integral is over the set of u's satisfying (4.22).

LEMMA 4.8. *Let* $|h|_p=p^k$. *Fix one* u_0 *satisfying* (4.22). *Then the set of* u*'s defined by* (4.22) *is given by the following conditions*

(i) $(u,u)/2a\equiv -1 \bmod p^l$

(ii) $u=\lambda u_0+u'$, *where* λ *goes through a set of representatives of the units in* $\mathbf{Z}_p$ $\bmod p^{l+k}$ *and* $|u'|\leqslant p^{-l}|h|_p^{-1/2}|a|_p^{1/2}$.

COROLLARY 4.9. *For* $p|D$ *and* γ *satisfying* (4.21) *we have*

$$|\mathrm{Kl}_p(\gamma)| \ll |h|_p^{1-r/2}\,|a|_p^{r/2}$$

where the constant depends on the normalization of Haar measure on $U(\mathbf{Q}_p)$ *and is not important for us.*

With these preliminaries we can now return to $\tilde{Z}(y,s)$.

From (4.7) we have that $|a(\gamma)|_p\geqslant 1$ for all but a finite number of p (and for the exceptional p the numerator is bounded). Hence essentially $a(\gamma)=1/n$ with n running over the integers. In this case $|a|$ tends to 0 and we have from (4.3)

$$I_\gamma(y,s)=\int_{\mathbf{R}^r}\left(\frac{|a|\,y}{y^2+\frac{1}{2}(u,u)}\right)^{s+r/2} e((u,\xi))\,du+O(|a|^{s+r/2+1}).$$

Hence clearly

$$\begin{aligned}\tilde{Z}(y,s)=&\left(\sum_{\gamma\in M(\mathbf{Q})}\mathrm{Kl}(\gamma)\,|a|^{s+r/2}\right)\int_{\mathbf{R}^r}\left(\frac{y}{y^2+\frac{1}{2}(u,u)}\right)^{s+r/2} e((u,\xi))\,du\\ &+O\left(\sum_{\gamma\in M(\mathbf{Q})}|\mathrm{Kl}(\gamma)|\,|a|^{s+r/2+1}\right).\end{aligned}\tag{4.24}$$

Let

$$Z(s)=\sum_{\gamma\in M(\mathbf{Q})}\mathrm{Kl}(\gamma)\,|a|^{s+r/2}.\tag{4.25}$$

This zeta function is of course a special case of the ones introduced in Section 2 and hence is meromorphic in $\mathbf{C}$. In the next section we will prove using the evaluations of the $\mathrm{Kl}_p(\gamma)$ carried out in this section that $Z(s)$ is holomorhic in $\mathrm{Re}(s)>r/2-1/2$. In fact we will show that the series (4.25) converges absolutely in this region. It follows that the O term in (4.24) is holomorphic in $\mathrm{Re}(s)>r/2-3/2$. Hence we learn that $\tilde{Z}(y,s)$ is holomorphic in $\mathrm{Re}(s)>r/2-1/2$ and hence from (4.1) and the remark following it that $F(\xi,s)$ is holomorphic in the same region. Combining this with Proposition 3.10 yields

THEOREM 4.10. *If Y_D is the hyperbolic manifold $\Gamma(D)\backslash H^{r+1}$ and λ_1 the smallest non-zero eigenvalue of the Laplacian on Y_D then*

$$\lambda_1 \geqslant 1/2(r-1/2).$$

§ 5. Estimation

THEOREM 5.1. *The Kloosterman–Selberg zeta function*

$$Z(s)=\sum_{\gamma\in M(\mathbf{Q})} \mathrm{Kl}(\gamma)\,|a_\gamma|^{s+r/2}$$

is absolutely convergent in $\mathrm{Re}(s)>r/2-1/2$.

Proof. Let $\mathbf{A}_f$ denote the finite adeles of $\mathbf{Q}$. Given

$$\gamma_p=\begin{pmatrix} a & & \\ & h & \\ & & a^{-1}\end{pmatrix}\in M(\mathbf{Q}_p)$$

set

$$n(\gamma_p)=|a|_p.$$

For $\gamma=(\gamma_2,\gamma_3,\dots)\in M(\mathbf{A}_f)$ we put

$$n(\gamma)=\prod_{p<\infty} n(\gamma_p)$$

then for $\gamma\in M(\mathbf{Q})\subseteq M(A_f)$ we have

$$n(\gamma)=|a_\gamma|_\infty^{-1}\quad\text{(product formula).}$$

Thus we may write

$$Z(s)=\sum_{\gamma\in M(\mathbf{Q})}\frac{\mathrm{Kl}(\gamma)}{n(\gamma)^{s+r/2}}. \tag{5.1}$$

Our strategy now is as follows. To each $\gamma_p\in M(\mathbf{Q}_p)$ we will find a majorant

$$|\mathrm{Kl}_p(\gamma_p)|\leqslant\Phi_p(\gamma_p) \tag{5.2}$$

such that Φ_p depends only on the coset $(M(\mathbf{Q}_p)\cap K_p)\gamma_p$, i.e. Φ_p is a function defined on $M(\mathbf{Q}_p)\cap K_p\backslash M(\mathbf{Q}_p)$. Set

$$\Phi(\gamma) = \prod \Phi_p(\gamma_p) \quad \text{for} \quad \gamma \in M(\mathbf{A}_f),$$

then Φ is a function on $M_f \backslash M(\mathbf{A}_f)$ where

$$M_f = \prod_{p<\infty} (M(\mathbf{Q}_p) \cap K_p).$$

In particular it restricts to a function in $M(D)\backslash M(\mathbf{Q}) \subseteq M_f \backslash M(\mathbf{A}_f)$. Now

$$\sum_{\gamma \in M(\mathbf{Q})} \left| \frac{\mathrm{Kl}(\gamma)}{n(\gamma)^{s+r/2}} \right| \leq \# M(D) \sum_{\gamma \in M(D)\backslash M(\mathbf{Q})} \frac{\Phi(\gamma)}{n(\gamma)^{\sigma+r/2}}. \tag{5.3}$$

Enlarging the index set on the right to $M_f \backslash M(\mathbf{A}_f)$ we find that the series is dominated by

$$\sum_{\gamma \in M_f \backslash M(\mathbf{A}_f)} \frac{\Phi(\gamma)}{n(\gamma)^{\sigma+r/2}} = \prod_{p<\infty} \zeta(p, \sigma) \tag{5.4}$$

where

$$\zeta(p, \sigma) = \sum_{\gamma \in M(\mathbf{Q}_p) \cap K_p \backslash M(\mathbf{Q}_p)} \frac{\Phi(\gamma_p)}{n(\gamma_p)^{\sigma+r/2}}.$$

Remark. Replacing $M(D)\backslash M(\mathbf{Q})$ by $M_f \backslash M(\mathbf{A}_f)$ enables us to gain the Euler product ($Z(s)$ does not have an Euler product!). This technique was first used in [17] in the analogous situation for the unitary group $U(n, 1)$.

The local calculations of the previous sections show that $\mathrm{Kl}_p(\gamma_p)=0$ if $n(\gamma_p)<1$. Take $\Phi_p(\gamma_p)=0$ in this case. If $n(\gamma_p)=1$ and $\mathrm{Kl}_p(\gamma_p)\neq 0$ then $p \nmid D$ and for $v_p \in M(\mathbf{Q}_p) \cap K_p$ we have

$$\mathrm{Kl}_p(\gamma_p) = \begin{cases} 0 & \text{if } p \mid D \\ 1 & \text{if } p \nmid D. \end{cases}$$

Thus we take $\Phi_p(\gamma_p)$ to be 0 and 1 respectively in these two cases. For each integer l set

$$S(p^l) = \sum_{n(\gamma_p)=p^l} \Phi_p(\gamma_p), \quad \gamma_p \in M(\mathbf{Q}_p) \cap K_p \backslash M(\mathbf{Q}_p) \tag{5.5}$$

then

$$\zeta(p, \sigma) = \sum_{l=0}^{\infty} \frac{S(p^l)}{p^{l(\sigma+r/2)}} = 1 + \sum_{l=1}^{\infty} \frac{S(p^l)}{p^{l(\sigma+r/2)}}. \tag{5.6}$$

Theorem 5.1 will follow from the following lemma.

LEMMA 5.2. *Let* $\sigma>r/2-1/2$ *then*

$$\sum_{l=1}^{\infty}\frac{S(p^l)}{p^{l(\sigma+r/2)}}\ll_{\varepsilon}p^{-1-\varepsilon}\quad(\varepsilon>0).$$

For the rest of this section we will define the function Φ_p satisfying (5.2) and verify Lemma 5.2.

Consider first $\mathrm{Kl}_p(\gamma_p)$ with $p\nmid D$. Write

$$\gamma_p=\begin{pmatrix}a&&\\&h&\\&&a^{-1}\end{pmatrix}$$

and put as before

$$n(\gamma_p)=|a|_p=p^l,\quad |h|_p=p^k\quad(0\leqslant k\leqslant l)$$

else Kl is zero by the analysis in Section 4.

Assume first that $l>1$. If $k<l$ then (4.14) gives

$$|\mathrm{Kl}_p(\gamma_p)|\leqslant p^{k(1-r/2)}p^{lr/2}\triangleq\Phi_p(\gamma_p). \tag{5.7}$$

We may count the number of γ_p's (modulo $M(\mathbf{Z}_p)$ of course) for which $\mathrm{Kl}_p(\gamma_p)\neq0$ using Lemmas 3.4 and 4.3. First observe that 4.10 implies $k+l$ must be even. We may parametrize the relevant γ_p's by means of the vectors u_0 appearing in Lemma 4.3. The r-dimensional vector u_0 must lie in $p^{-((k+l)/2)}\mathbf{Z}_p^r$. (4.11) implies that we may "mod out" u_0 by vectors in $p^{-((l-k)/2)}\mathbf{Z}_p^r$. The quotient space

$$p^{-((l+k)/2)}\mathbf{Z}_p^r\Big/p^{-((l-k)/2)}\mathbf{Z}_p^r$$

is isomorphic to $\mathbf{Z}_p^r/p^k\mathbf{Z}_p^r$. The relation (4.11)(ii) further implies that we may identify multiples of elements from the multiplicative group $(\mathbf{Z}_p/p^k\mathbf{Z}_p)^*$. Finally, there is one more constraint coming from (4.11)(i). Putting all this together, we find that the number of relevant γ_p's is bounded by $p^{k(r-2)}$. Thus for fixed $k<l$ (with $k\equiv l(2)$) the contribution to $S(p^l)$ is bounded by

$$p^{k(r-2)}p^{k(1-r/2)}p^{lr/2}.$$

If $k=l$ then the non-zero $\mathrm{Kl}_p(\gamma_p)$'s are given by (4.20). Trivial estimation gives

$$|\mathrm{Kl}_p(\gamma_p)|\leqslant p^l$$

(this is the same as (5.7) with $k=l$). Similar arguments as before, this time using Lemma 4.6 show that the number of relevant γ_p's is bounded by $p^{l(r-2)}$. Hence we have

$$\begin{aligned} S(p^l) &\leq p^{l(r-2)}p^l + \sum_{\substack{0\leq k<l \\ k\equiv l \bmod 2}} p^{k(r-2)+k(1-r/2)+lr/2} \\ &\leq 2p^{l(r-1)}. \end{aligned}$$

This implies that for $\sigma>r/2-1/2$

$$\sum_{l=2}^{\infty} \frac{S(p^l)}{p^{l(\sigma+r/2)}} << p^{-1-\varepsilon} \tag{5.8}$$

if $\sigma=r/2-1/2+\varepsilon$. Note that so far we have only used the "trivial" estimation of Kl.

Next we consider the case $l=1$. Here the only non-zero $\mathrm{Kl}_p(\gamma_p)$ is given by (4.20). If $(u_0, \xi) \notin \mathbf{Z}_p$ then the sum in question is a nontrivial "classical" Kloosterman sum and Weil's estimation [32] gives

$$|\mathrm{Kl}_p(\gamma_p)| \leq 2\sqrt{p} \overset{\Delta}{=} \Phi_p(\gamma_p). \tag{5.9}$$

The number of such γ_p's is as before bounded by p^{r-2}. Hence the contribution to $S(p)$ is bounded by $p^{1/2}p^{r-2}=p^{r-3/2}$. On the other hand if

$$(u_0, \xi) \in \mathbf{Z}_p \tag{5.10}$$

then we can only use the trivial bound

$$|\mathrm{Kl}_p(\gamma_p)| \leq p.$$

However (5.10) puts on an extra constraint on our parameter u_0 so that the number of non-zero $\mathrm{Kl}_p(\gamma_p)$'s in this case is bounded by p^{r-3}. We conclude that

$$S(p) \leq p^{r-3/2} + pp^{r-3} \leq 2p^{r-3/2}.$$

Therefore for $\sigma>r/2-1/2$

$$\frac{S(p)}{p^{\sigma+r/2}} \leq 2p^{-1-\varepsilon} \tag{5.11}$$

for $\sigma=r/2-1/2+\varepsilon$. The estimates (5.8) and (5.11) clearly give Lemma 5.2 for $p\nmid D$. On the other hand there are only a finite number of p's dividing D. Thus to complete the estimation we need only check that the series

$$\sum_{l=1}^{\infty} \frac{S(p^l)}{p^{l(\sigma+r/2)}}$$

is convergent for $p|D$ and $\sigma>r/2-1/2$. We can do this easily using Lemmas 4.7, 4.8 and Corollary 4.9—the details are left to the reader. Q.E.D.

§ 6. Examples of exceptional spectrum

As we noted in Section 1 the manifolds Y_D (for $r\geq 3$) may have exceptional spectrum. The purpose of this section is to exhibit such examples. The method of doing so is via theta liftings which we review.

Recall that, the metaplectic group $\widetilde{SL_2(\mathbf{A})}$ is a two fold covering of $SL_2(\mathbf{A})$, so that it fits into the following short exact sequence

$$1 \to \{1, \zeta\} \to \widetilde{SL_2(\mathbf{A})} \to SL_2(\mathbf{A}) \to 1. \tag{6.1}$$

Here we have denoted by $\{1, \zeta\}$ the group of two elements. The pair $G(\mathbf{A})$, $SL_2(\mathbf{A})$ is (almost) a reductive dual pair in the sense of Howe [8]. In particular there is an oscillator (or "Weil") representation ω associated to this dual pair.

To be more precise, the choice of ω depends on a nontrivial character of $\mathbf{Q}\backslash\mathbf{A}$. Given the standard character ψ fixed at the beginning of the paper, any other nontrivial character of $\mathbf{Q}\backslash\mathbf{A}$ is of the form ψ^α with $\alpha \in \mathbf{Q}^*$ where

$$\psi^\alpha(x) = \psi(\alpha x) \quad (x \in \mathbf{Q}\backslash\mathbf{A}). \tag{6.2}$$

Let us denote by ω_α the relevant oscillator representation associated to the character ψ^α. We now describe a particular (Schrödinger model) realization of ω_α suitable for our purposes here.

Let W be the two dimensional symplectic vector space over $\mathbf{Q}$ on which SL_2 acts (on the right). Let e_1, e_2 be a symplectic basis for W. Then ω_α will be realized on $\mathcal{S}(W(\mathbf{A})\oplus\mathbf{A}^r)$, the space of Bruhat–Schwartz functions on $W(\mathbf{A})\oplus\mathbf{A}^r$. Note that

$$\mathcal{S}(W(\mathbf{A})\oplus\mathbf{A}^r) \cong \mathcal{S}(W(\mathbf{A}))\otimes\mathcal{S}(\mathbf{A}^r). \tag{6.3}$$

The action of $\widetilde{SL_2(\mathbf{A})}$ is a tensor product of two oscillator representations. It acts on the first factor of (6.3) through the linear action of $SL_2(\mathbf{A})$ on $W(\mathbf{A})$. On the second factor $\mathcal{S}(\mathbf{A}^r)$ its action is just the oscillator representation of $\widetilde{SL_2(\mathbf{A})}$ associated to our r-dimensional quadratic form (,). In particular for $\Phi \in \mathcal{S}(W(\mathbf{A})\oplus\mathbf{A}^r)$

$$\omega_a\left(\begin{pmatrix}1 & n\\ 0 & 1\end{pmatrix}\right)\Phi(w,x)=\psi^{a}(nq(x))\,\Phi\left(w\begin{pmatrix}1 & n\\ 0 & 1\end{pmatrix},x\right) \tag{6.4}$$

($w\in W(\mathbf{A}), x\in\mathbf{A}^r, n\in\mathbf{A}$) where we have put $q(x)=(x,x)$. It is not difficult to describe the action of $U(\mathbf{A})$ in this model. However all we need is the following formula.

$$\omega_a(u)\,\Phi(e_2,x)=\psi^{a}((u,x))\,\Phi(e_2,x). \tag{6.5}$$

Here, as usual, we are using the identification $U(\mathbf{A})\cong\mathbf{A}^r$ in the obvious way.

Remark. (6.5) follows from Howe [9, formula 19] and is easy to verify. For each $\Phi\in\mathcal{S}(W(\mathbf{A})\oplus\mathbf{A}^r)$ we define

$$\theta_\Phi(g,h)=\sum_{\substack{\eta\in W(\mathbf{Q})\\ \xi\in\mathbf{Q}^r}}\omega_a(g,h)\,\Phi(\eta,\xi)\quad (g\in G(\mathbf{A}),\ h\in\widetilde{SL_2(\mathbf{A})}).$$

Then θ_Φ is a slowly increasing automorphic function on $G(\mathbf{A})\times\widetilde{SL_2(\mathbf{A})}$. If ϕ is a cusp form on $\widetilde{SL_2(\mathbf{A})}$, the following integral is absolutely convergent and gives a $G(\mathbf{Q})$ invariant function $f(g)$ on $G(\mathbf{A})$.

$$f(g)=\int_{SL_2(\mathbf{Q})\backslash SL_2(\mathbf{A})}\theta_\Phi(g,h)\,\overline{\phi(h)}\,dh. \tag{6.6}$$

Remark. We assume that the following compatibility condition is satisfied for ϕ (see 6.1)

$$\phi(\zeta h)=\begin{cases}\phi(h) & \text{if } r \text{ is even}\\ -\phi(h) & \text{if } r \text{ is odd.}\end{cases} \tag{6.7}$$

In this way the integrand in (6.6) is in fact a function on $SL_2(\mathbf{Q})\backslash SL_2(\mathbf{A})$.

LEMMA 6.1. *For $r\geq 3$ the function $f(g)$ defined by* (6.6) *is square integrable on* $G(\mathbf{Q})\backslash G(\mathbf{A})$.

Proof (sketch). Proposition 8 of Weil [33] implies that the function $\theta_\Phi(g,h)$ is square integrable in the first variable if $r\geq 3$. Since the cusp from ϕ is rapidly decreasing this property of square integrability is not lost on integrating against ϕ. Actually for essentially this situation in the classical language see Siegel [28], who examines $\theta(g,h)$ in the variable g for $r\geq 3$. Q.E.D.

The next question is whether f is identically zero. Taking Φ of the form $\Phi=\Phi_1\cdot\Phi_2$ where $\Phi_1\in\mathcal{S}(W(\mathbf{A}))$, $\Phi_2\in\mathcal{S}(\mathbf{A}^r)$ we find that

$$\theta_\Phi(1,h)=F_1(h)\,F_2(h)$$

where

$$F_1(h)=\sum_{\eta\in W(\mathbf{Q})}\Phi_1(\eta h)$$

and

$$F_2(h)=\sum_{\xi\in\mathbf{Q}^r}\omega'_a(h)\,\Phi_2(\xi).$$

Here ω'_a is the oscillator representation attached to our r-dimensional quadratic form $(\ ,\ )$. Denote by N the upper triangular unipotent subgroup of SL_2. Then N is precisely the stabilizer of e_2, and we have

$$F_1(h)=\Phi_1(0)+\sum_{\gamma\in N(\mathbf{Q})\backslash SL_2(\mathbf{Q})}\Phi_1(e_2\gamma h).$$

It follows that

$$\begin{aligned}
f(1)&=\int_{SL_2(\mathbf{Q})\backslash SL_2(\mathbf{A})}\Phi_1(0)\,F_2(h)\,\overline{\phi(h)}\,dh\\
&\quad+\int_{N(\mathbf{Q})\backslash SL_2(\mathbf{A})}\Phi_1(e_2h)\,F_2(h)\,\overline{\phi(h)}\,dh\\
&=\Phi_1(0)\int_{SL_2(\mathbf{Q})\backslash SL_2(\mathbf{A})}F_2(h)\,\overline{\phi(h)}\,dh\\
&\quad+\sum_{\xi\in\mathbf{Q}^r}\int_{N(\mathbf{Q})\backslash SL_2(\mathbf{A})}\Phi_1(e_2h)\,\omega'_a(h)\,\phi_2(\xi)\,\overline{\phi(h)}\,dh.
\end{aligned}$$

From (6.4) we see that the integral

$$\int_{N(\mathbf{Q})\backslash SL_2(\mathbf{A})}\Phi_1(e_2h)\,\omega'_a(h)\,\Phi_2(\xi)\,\overline{\phi(h)}\,dh$$

is zero if $\xi=0$ (since ϕ is a cusp form) and otherwise it equals

$$f_\xi(1) \overset{\triangle}{=} \int_{N(\mathbf{A})\backslash SL_2(\mathbf{A})} \Phi_1(e_2 h)\,\Phi_2(\xi)\,\overline{\phi_\xi(h)}\,dh \tag{6.8}$$
$$= \int_{N(\mathbf{A})\backslash SL_2(\mathbf{A})} \omega_\alpha(h)\,\Phi(e_2,\xi)\,\overline{\phi_\xi(h)}\,dh$$

where

$$\phi_\xi(h) = \int_{N(\mathbf{Q})\backslash N(\mathbf{A})} \phi(nh)\,\psi^\alpha(-nq(\xi))\,dn$$
$$= \int_{N(\mathbf{Q})\backslash N(\mathbf{A})} \phi(nh)\,\psi^{\alpha\cdot q(\xi)}(-n)\,dn.$$

Set

$$f_0(1) = \int_{SL_2(\mathbf{Q})\backslash SL_2(\mathbf{A})} \Phi_1(0)\,F_2(h)\,\overline{\phi(h)}\,dh.$$

We have shown that

$$f(1) = f_0(1) + \sum_{\substack{\xi\in\mathbf{Q}^r\\ \xi\neq 0}} f_\xi(1) \tag{6.9}$$

the same formula holds for the general Φ (not necessarily of the form $\Phi_1\cdot\Phi_2$).

Observe that the characters of $U(\mathbf{Q})\backslash U(\mathbf{A})$ can be identified with $\mathbf{Q}^r$ and (6.5) and (6.8) show that (6.9) gives precisely the Fourier expansion of f along U! Note that replacing $f(1)$ by $f(g)$ simply replaces Φ by some other Schwartz function. We have shown

LEMMA 6.2. (a) *f is cuspidal iff ϕ is orthogonal to all theta series attached to the quadratic form $(\ ,\ )$.*

(b) *If ϕ possesses a nonzero Fourier coefficient with respect to the character $\psi^{\alpha\cdot q(\xi)}$ for some $\xi\in\mathbf{Q}^r$, then f is nonzero.*

Remark. (1) Part (a) is a special case of a general criterion due to Rallis [23].

(2) The condition (b) is always satisfied for an appropriate choice of α.

The infinitesimal (Howe duality) correspondence for a general real reductive dual pair is known. In particular there is a formula relating the images of the Casimir elements of the two groups under the oscillator representation. In the present situation let D and D' be the Casimir elements for the groups $G(\mathbf{R})$ and $\widetilde{SL_2(\mathbf{R})}$ respectively. Then

(6.10) $$\omega_\alpha(D') = \frac{r}{4}\omega_\alpha(D) - \frac{1}{8}\left(1-\frac{r^2}{4}\right)$$

(see for example J. D. Adams [1]).

Now take ϕ to be a cusp form on $\widetilde{SL_2(\mathbf{A})}$ which comes from a classical holomorphic cusp form of weight k (with k an integer or half an odd integer according as r is even or odd). Then f will be an eigenfunction on $G(\mathbf{A})$ with eigenvalue (for Δ) λ determined by

$$\frac{1}{2}\frac{k}{2}\left(\frac{k}{2}-1\right) = \frac{r}{4}\left(-\frac{\lambda}{2r}\right) - \frac{1}{8}\left(1-\frac{r^2}{4}\right)$$

i.e.

(6.11) $$\left(\frac{r}{2}\right)^2 - \lambda = (k-1)^2.$$

Remark. The relation between the Casimir elements and the Laplacians is as follows: When expressed in terms of the Iwasawa coordinates $-2D'$ assumes the form

$$-y^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right) - y\frac{\partial^2}{\partial x\,\partial\theta}$$

while when acting on H^{r+1}, $2rD$ takes the form

$$\Delta = -y^2\left(\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial x_1^2}+\dots+\frac{\partial^2}{\partial x_r^2}\right) - (r-1)\,y\frac{\partial}{\partial y}$$

with an appropriate normalization.

So far we know that f is an eigenform for $-2rD$ with eigenvalue λ given by (6.11). However we don't as yet know that f is K_∞-invariant (and hence may be viewed as a function on H^{r+1}). To put it another way the action of Casimir is not enough to determine the representation of $G(\mathbf{R})$ in question.

The additional information needed is provided by Rallis–Schiffmann [24]. The following lemma follows from their Theorem 5.

LEMMA 6.3. *Suppose that $1/2<k\leq r/2$ and that k is of the form $k=r/2-2j$ with j a nonnegative integer. Then the (holomorphic) discrete series representation of $\widetilde{SL_2(\mathbf{R})}$ of lowest weight k corresponds to a spherical representation of $G(\mathbf{R})$.*

Hence if k satisfies the condition of Lemma 6.3 then we may take f to be K_∞-invariant. Such an f corresponds to eigenforms of Δ for some appropriate congruence subgroup. Since $\lambda=r^2/4-s^2$, (6.11) implies $s=k-1$. We have shown

THEOREM 6.4. *For some congruence subgroup there exist exceptional eigenforms whose infinite parameter corresponds to* $s=r/2-1$, $r/2-3, \ldots$.

Remark. With a little more work the levels of these exceptional forms can be determined and one can also arrange these exceptional forms to be cuspidal.

We end this section by giving examples of noncongruence Γ for which $\lambda_1(\Gamma\backslash H^{r+1})$ is arbitrarily small, so that in Theorem 4.10 the congruence assumption is essential. The examples are similar to those of Selberg [27] and Randol [25] for $SL_2(\mathbf{R})$. Millson [21] has shown that there exist D for which $d=\beta_1(Y_D)=\operatorname{rank} H_1(Y_D,\mathbf{Z})$ satisfies $d\geq 1$. For such D it follows that the set of characters (unitary) of $\Gamma(D)$ is $T^d\times F$, where T^d is a real torus of dimension d and F a finite abelian group. Let A^0 denote the connected component of the identity in the above group. For $\chi\in A^0$ consider the Laplacian Δ_χ acting on $L^2(\Gamma(D)\backslash H^{r+1},\chi)$ i.e. f's such that

$$f(\gamma z)=\chi(\gamma)f(z) \quad (\gamma\in\Gamma(D)).$$

The spectrum of Δ_χ in $[0,(r/2)^2)$ is discrete. Let $\lambda_0(\chi)$ denote the smallest eigenvalue of Δ_χ. Thus $\lambda_0(\mathbf{1})=0$ while it is clear that $\lambda_0(\chi)>0$ for $\chi\neq\mathbf{1}$. It is easily seen, say using the minimax principle, that $\lambda_0(\chi)$ is continuous on A^0. Hence for $\varepsilon>0$ we can find $\chi_1\neq\mathbf{1}$, χ_1 of finite order in A^0, such that $0<\lambda_0(\chi)<\varepsilon$. If $\Gamma'=\ker\chi_1$ then Γ' is of finite index in $\Gamma(D)$ and $\lambda_1(\Gamma'\backslash H^{r+1})<\varepsilon$. This gives examples of Γ's with small λ_1. In view of Theorem 4.10, Γ' cannot be a congruence subgroup of $\Gamma(1)$.

References

[1] ADAMS, J. D., Discrete spectrum of the reductive dual pair $(O(p,q), Sp(2m))$. *Invent. Math.*, 74 (1983), 449–475.

[2] BUMP, D., FRIEDBERG, S. & GOLDFELD, D., Poincaré series and Kloostrman sums for $SL(3,\mathbf{Z})$. *Acta Arith.*, 50 (1988), 31–89.

[3] COHEN, P. & SARNAK, P., *Discrete Groups and Automorphic Forms*. Notes, Stanford University, 1980.

[4] DIXMIER, J. & MALLIAVIN, P., Factorisations de fonctions et de vecteurs indéfiniment differentiables. *Bull. Sci. Math.*, 102 (1987), 307–330.

[5] ELSTRODT, J., GRUNEWALD, F. & MENNICKE, J., Poincaré series, Kloosterman sums, and eigenvalues of the Laplacian for congruence groups acting on hyperbolic spaces. *C.R. Acad. Sci. Paris*, 305 (1987), 577–581.

[6] GELFAND, I. M., GRAEV, M. I. & PIATETSKI-SHAPIRO, I. I., *Representation Theory and Automorphic Functions*. W.B. Saunders & Co., Philadelphia, 1969.

[7] GRADSHTEYN, I. S. & RYZHIK, I. M., *Tables of Integrals, Series and Products*. Academic Press, New York, 1980.

[8] HOWE, R., θ-series and invariant theory. *Proc. Sympos. Pure Math.*, 33:1 (1979), 275–285.
[9] — L^2-duality in the stable range. Preprint.
[10] HOWE, R. & PIATETSKI-SHAPIRO, I. I., A counter example to the generalized Ramanujan conjecture for (quasi)-split groups. *Proc. Sympos. Pure Math.*, 33:1 (1979), 315–322.
[11] IWANIEC, H., Non-holomorphic modular forms and their applications, in *Modular Forms*, ed. R. A. Rankin. Ellis Horwood Ltd., W. Sussex, 1984.
[12] — Small eigenvalues of the Laplacian for $\Gamma_0(N)$. *Acta Arith.*, 56 (1990), 65–82.
[13] — Spectral theory of automorphic functions and recent developments in analytic number theory. *Proc. Internat. Congress Math.*, Berkeley, Ca., 1986, pp. 444–456.
[14] JACQUET, H. & SHALIKA, J., On Euler products and the classification of automorphic representations. *Amer. J. Math.*, 103 (1981), 499–558.
[15] KNAPP, A., *Representation Theory of Semisimple Groups*. Princeton Univ. Press, Princeton, 1986.
[16] LANGLANDS, R. P., *On the Functional Equation Satisfied by Eisenstein Series*. Lecture Notes in Math., 544. Springer Verlag, 1976.
[17] LI, J.-S., Kloosterman–Selberg zeta functions on complex hyperbolic spaces. Preprint.
[18] LI, J.-S., PIATETSKI-SHAPIRO, I. I. & SARNAK, P., Poincaré series for $SO(n, 1)$. *Proc. Indian Acad. Sci. Math. Sci.*, 97 (1987), 231–237.
[19] MAGNUS, W., OBERHETTINGER, F. & SONI, R., *Formulas and Theorems for the Special Functions of Mathematical Physics*. Springer Verlag, New York, 1966.
[20] MIATELLO, R. & WALLACH, N., Automorphic forms constructed from Whittaker vectors. Preprint.
[21] MILLSON, J., On the first Betti number of a constant negatively curved manifold. *Ann. of Math.*, 104 (1976), 235–247.
[22] PIATETSKI-SHAPIRO, I. I., Lectures on Poincaré series and Kloosterman sums over a function field. Preprint.
[23] RALLIS, S., On the Howe duality conjecture. *Comput. Math.*, 51 (1984), 333–399.
[24] RALLIS, S. & SCHIFFMANN, G., Discrete spectrum of the Weil representation. *Bull. Amer. Math. Soc.*, 83 (1977), 267–276.
[25] RANDOL, B., Small eigenvalues of the Laplace operator on compact Riemann surfaces. *Bull. Amer. Math. Soc.*, 80 (1974), 996–1000.
[26] SARNAK, P., The arithmetic and geometry of some hyperbolic three manifolds. *Acta Math.*, 151 (1983), 253–295.
[27] SELBERG, A., On the estimation of Fourier coefficients of modular forms. *Number Theory: Proc. Sympos. Pure Math.*, 8 (1965), 1–15.
[28] SIEGEL, C. L., Indefinite quadratische Formen und Funktionen Theorie, I. *Math. Ann.*, 124 (1951), 17–24; II. *Math. Ann.*, 124 (1952), 364–387.
[29] — *Lectures on Advanced Analytic Number Theory*. Tata Institute of Fundamental Research, Bombay, 1961.
[30] STEVENS, G., Poincaré series on $GL(m)$ and Kloosterman sums. *Math. Ann.*, 277 (1987), 25–51.
[31] WALLACH, N., Asymptotic expansions of generalized matrix entries of representations of real reductive groups, in *Lie Group Representations* I. Proceedings, University of Maryland, 1982–1983. Lecture Notes in Math., 1024, pp. 287–369.
[32] WEIL, A., On some exponential sums. *Proc. Nat. Acad. Sci. U.S.A.*, 34 (1948), 204–207.
[33] — Sur la formule de Siegel dans la théorie des groupes classiques. *Acta Math.*, 113 (1965), 1–87.

Received August 23, 1989
Received in revised form June 15, 1990