Vol. 71, No. 1 DUKE MATHEMATICAL JOURNAL © July 1993

DENSITY OF INTEGER POINTS ON AFFINE HOMOGENEOUS VARIETIES

W. DUKE, Z. RUDNICK, AND P. SARNAK

Section 1. Let V be an affine variety defined over $\mathbf{Z}$ by integral polynomials $f_j \in \mathbf{Z}[x_1, \dots, x_n]$:

$$V = \{x \in \mathbf{C}^n : f_j(x) = 0, j = 1, \dots, v\} \tag{1.1}$$

A basic problem of diophantine analysis is to investigate the asymptotics as $T \to \infty$ of

$$N(T, V) = \{m \in V(\mathbf{Z}) : \|m\| \leqslant T\} \tag{1.2}$$

where we denote by $V(A)$, for any ring A, the set of A-points of V. Hence $\|\cdot\|$ is some Euclidean norm on $\mathbf{R}^n$.

The only general method available for such problems is the Hardy-Littlewood circle method, which however has certain limitations, requiring roughly that the codimension of V in the ambient space $\mathbf{A}^n$, as well as the degree of the equations (1.1), be small relative to n. Furthermore, there are restrictions on the size of the singular sets of the related varieties:

$$V_\mu = \{x \in \mathbf{C}^n : f_j(x) = \mu_j, j = 1, \dots, v\}, \qquad \mu = (\mu_j) \in \mathbf{C}^n.$$

We refer to [Bi] and [Sch] for a discussion of the restriction. Regardless of these restrictions, one hopes that for many more cases $N(T, V)$ can be given in the form predicted by the Hardy-Littlewood method, that is, as a product of local densities:

$$N(T, V) \sim \prod_{p < \infty} \mu_p(V) \cdot \mu_\infty(T, V), \tag{$*$}$$

where the "singular series" $\prod_{p<\infty} \mu_p(V)$ is given by p-adic densities:

$$\mu_p(V) = \lim_{k \to \infty} \frac{\# V(\mathbf{Z}/p^k\mathbf{Z})}{p^{k \dim V}}$$

and $\mu_\infty(T, V)$ is a real density—the "singular integral." Following Schmidt [Sch], we say that V is a Hardy-Littlewood system if the above asymptotics ($*$) is valid.

Received 25 September 1992. Revision received 6 November 1992.
Supported by NSF grants No. DMS-8902992, DMS-9102082, and the Sloan Foundation.

In complete generality, the above problem (1.2) is hopeless, and so one seeks to solve it for a special but rich family of varieties. In this paper we consider varieties defined via actions of linear algebraic groups and, in particular, such varieties defined by invariant theory. In a recent paper [FMT] Franke-Manin-Tschinkel consider the problem of counting rational points of height $\leqslant T$ on flag varieties $V = P\backslash G$ where G is reductive (over $\mathbf{Q}$) and P is a parabolic subgroup. In their case, the corresponding Eisenstein series is the key tool in determining the asymptotics. In the theory developed below, the full harmonic analysis of $L^2(G(\mathbf{Z})\backslash G(\mathbf{R}))$ comes into play.

We now formulate the main result. Let G be a linear algebraic semisimple group defined over $\mathbf{Q}$. Let $\rho\colon G \to GL(W)$ be a rational representation of G defined over $\mathbf{Q}$, W being a $\mathbf{Q}$ vector space. Let $w_0 \in W_Q$ be a vector whose orbit $V = w_0\rho(G)$ is Zariski closed. Then the stabilizer $H \subset G$ of w_0 is reductive and V is isomorphic to $H\backslash G$. It is this family of varieties that we investigate in connection with the basic problem. The reason that the problem is at all approachable is that a fundamental theorem of Borel–Harish-Chandra [B-HC] asserts that $V(\mathbf{R})$ breaks up into finitely many $G(\mathbf{R})$-orbits and, more surprisingly, $V(\mathbf{Z})$ into finitely many $G(\mathbf{Z})$-orbits. Thus the points of $V(\mathbf{Z})$ are parametrized by cosets of $G(\mathbf{Z})$.

For our purpose of studying (1.2), it suffices to fixate on one $G(\mathbf{Z})$ orbit $\mathcal{O}$, say $\mathcal{O} = w_0 G(\mathbf{Z})$ with $w_0 \in V(\mathbf{Z})$. The stabilizer of w_0 in $G(\mathbf{Z})$ is $H(\mathbf{Z}) = H \cap G(\mathbf{Z})$, where H is the stabilizer of w_0. The counting problem in question becomes

$$N(T, \mathcal{O}) = |\{\gamma \in H(\mathbf{Z})\backslash G(\mathbf{Z})\colon \|w_0\gamma\| \leqslant T\}|. \tag{1.3}$$

To state the main theorem we will need to make a further restriction, one which is satisfied by many interesting examples. We say $V(\mathbf{R}) = H(\mathbf{R})\backslash G(\mathbf{R})$ is *symmetric* if $H(\mathbf{R})$ is the fixed point set of some involution τ of $G(\mathbf{R})$. Note that τ need not be a Cartan involution, and so we are not assuming that $V(\mathbf{R})$ is a Riemannian symmetric space. While $\mathrm{Vol}(G(\mathbf{Z})\backslash G(\mathbf{R})) < \infty$, $\mathrm{Vol}(H(\mathbf{Z})\backslash H(\mathbf{R}))$ need not be finite as H need only be reductive. For this paper we will assume that

$$\mathrm{Vol}(H(\mathbf{Z})\backslash H(\mathbf{R})) < \infty. \tag{1.4}$$

This assumption can be removed by a refinement of these methods. Indeed, the basic asymptotics below change by factors of $\log T$ if (1.4) fails.

We will usually assume that G is a $\mathbf{Q}$-simple, connected $\mathbf{Q}$-group, with $G(\mathbf{R})$ noncompact (see however Example 1.6). This is to guarantee that nontrivial spherical constituents of $L^2(\Gamma\backslash G)$ have matrix coefficients which decay at infinity (see Theorem 2.6).

We normalize dg on $G(\mathbf{R})$ and dh on $H(\mathbf{R})$ so that

$$\mathrm{Vol}(G(\mathbf{Z})\backslash G(\mathbf{R})) = \mathrm{Vol}(H(\mathbf{Z})\backslash H(\mathbf{R})) = 1. \tag{1.5}$$

Note that such a normalization builds in the arithmetical constants involved in volumes of fundamental domains.

With this normalization we get a unique $G(\mathbf{R})$ invariant measure $d\dot{g}$ on $V(\mathbf{R}) = H(\mathbf{R})\backslash G(\mathbf{R})$ satisfying

$$dg = dh\, d\dot{g}. \tag{1.6}$$

Let $\mu(T)$ be defined by

$$\mu(T) = \int_{\|w_0\dot{g}\| \leqslant T} d\dot{g}. \tag{1.7}$$

Our main result is the following theorem.

THEOREM 1.2. *Assuming $V(\mathbf{R})$ is (affine) symmetric, we have*

$$N(T, \mathcal{O}) \sim \mu(T) \qquad \text{as } T \to \infty. \tag{1.8}$$

Remarks

1.3. We will prove Theorem 1 in the case that $H \cap \Gamma\backslash H$ is compact and G is classical. The assumption that G is classical is imposed to avoid technical issues in Lemma 2.7. Our proof of the general case ($H \cap \Gamma\backslash H$ noncompact) is very involved as it requires a regularization of the period integrals of Eisenstein series. Since in the meantime Eskin and McMullen have given a technically much simpler proof of the theorem, we see no reason at this time to present our original proof of Theorem 1 in the case that $H \cap \Gamma\backslash H$ is noncompact. Because of its possible use in other contexts, we nevertheless state without proof the general regularization of periods of Eisenstein series (see Theorem 1.11).

1.3′. The assumption that $V(\mathbf{R})$ be symmetric cannot be dropped since Eskin has recently found a nonsymmetric example where (1.2) fails. See example (1.8) for a nonsymmetric example where (1.2) holds.

1.4. Combining the asymptotics of $N(T, \mathcal{O}_j)$ over the finite number of orbits gives the asymptotics for $N(T, V)$. The constants c coming up in Proposition 1.1 via the normalization (1.5) are of considerable interest since the corresponding weighted sum over the orbits gives one side of a "mass formula" à la Siegel. Indeed, the other side is supplied if V is a Hardy-Littlewood system. Theorem 1.2 may be used to give a new proof of Siegel's mass formula for rational quadratic forms or, what is the same, that the Tamagawa number of an orthogonal group is 2 [ERS]. We also find that many homogeneous affine V's are Hardy-Littlewood systems or are not far from being such.

We give some concrete examples of Theorem 1.2.

Example 1.5. The well-studied case of quadratic forms is included in the above. Let

$$F(x_1, \ldots, x_n) = \sum_{i,j} f_{ij} x_i x_j$$

be a nondegenerate integral quadratic form, whose signature over $\mathbf{R}$ is (p, q), $p + q = n$, $pq \neq 0$. Let

$$W_k = \{x \mid F(x) = k\}, \qquad k \neq 0.$$

W_k is a hyperboloid and is a symmetric space of the form $SO(p-1, q)\backslash SO(p, q)$. As long as $\operatorname{vol}(H(\mathbf{Z})\backslash H(\mathbf{R})) < \infty$, which certainly is the case if $n \geqslant 4$, then by Theorem 1.2

$$N(T, W_k) \sim C_{k,F} T^{n-2}. \tag{1.9}$$

$C_{k,F}$ is explicitly computable and is 0 if (and only if) there are no integral points on V_k. The case of $n = 3$ is instructive. Let

$$F(x_1, x_2, x_3) = -x_1^2 - x_2^2 + dx_3^2, \qquad d > 0.$$

Let

$$V = V_1 = \{x \mid F(x) = -1\}. \tag{1.10}$$

$V(\mathbf{R})$ is a one-sheeted hyperboloid, and the group $G(\mathbf{R}) \cong SO_0(1, 2)$ acts transitively on $V(\mathbf{R})$ with $w_0 = (1, 0, 0)$ and stabilizer $H(\mathbf{R}) \cong SO(1, 1)$. H is the orthogonal group of the form

$$Q(x, y) = -x^2 + dy^2. \tag{1.11}$$

If for $\| \ \|$ on $\mathbf{R}^3$ we choose

$$\|x\|^2 = x_1^2 + x_2^2 + dx_3^2,$$

then

$$N(T, V) = \sum_{|n| \leqslant T} r_2(1 + dn^2)$$

where

$$r_2(n) = |\{(x, y) \in \mathbf{Z}^2 \mid x^2 + y^2 = n\}|.$$

In this case the asymptotics of $N(T, V)$ may also be deduced in an elementary fashion

$$N(T, V) \sim \begin{cases} c_d T, & \text{if } d \text{ is not a square} \\ c_d T \log T, & \text{if } d \text{ is a square} \end{cases}$$

where $c_d > 0$ [Sco]. The last dichotomy corresponds to exactly the cases, $\mathrm{vol}(H(\mathbf{Z})\backslash H(\mathbf{R}))$ being finite or not. Theorem 1.2 gives the first case and the modification mentioned after (1.4) gives the latter.

Of course, for this example (1.5) the Hardy-Littlewood method works if $n \geqslant 5$ while the theory of θ-functions can also be used to analyze this case.

Example 1.6. Let $V_{n,k} = \{x \in Mat_n | \det x = k\}$, $k \neq 0$. The group $G^* = SL_n \times SL_n$ acts on $V_{n,k}$ by $x(g_1, g_2) = g_1^{-1} x g_2$. Over $\mathbf{C}$, $V_{n,k} \cong \Delta\backslash G^*$ where $\Delta = \{(g, g) | g \in SL_n\} \cong SL_n$ is the diagonal subgroup. It is the fixed point set of the involution $(g_1, g_2)^\sigma = (g_2, g_1)$. Thus $V(\mathbf{R})$ is symmetric. The Euclidean norm on $V_{n,k}$, $\|x\|^2 = \mathrm{tr}({}^t xx)$ is invariant under $K^* = SO(n) \times SO(n)$, which is a maximal compact subgroup. Theorem 1.2 gives

$$N(T, V_{n,k}) \sim c_{n,k} T^{n^2-n} \tag{1.12}$$

where, if $k = p_1^{a_1} \cdots p_r^{a_r}$, then

$$c_{n,k} = \frac{\pi^{n^2/2} k^{-(n-1)}}{\Gamma\left(\frac{n^2-n+2}{2}\right)\Gamma\left(\frac{n}{2}\right)\zeta(2)\cdots\zeta(n)} \prod_{j=1}^{r} \frac{(p_j^{a_j+1}-1)\cdots(p_j^{a_j+n-1}-1)}{(p_j-1)\cdots(p_j^{n-1}-1)}$$

$$= \frac{\pi^{n^2/2}}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n^2-n+2}{2}\right)} \zeta(2)^{-1}\cdots\zeta(n)^{-1} \sum_{d_1\cdots d_n=k} d_2^{-1} d_3^{-2} \cdots d_n^{-(n-1)}.$$

Note that $N(T, V_{n,1})$ just counts

$$\sum_{\substack{\gamma \in SL_n(\mathbf{Z}) \\ \|\gamma\| \leqslant T}} 1. \tag{1.14}$$

The latter, when $n = 2$, is a quadratic equation and so falls into the previous example. In this case it also corresponds to a non-Euclidean lattice point problem, and the result and indeed our method, via non-Euclidean harmonic analysis, goes back to Delsarte [D]. For lattice points in a non-Euclidean ball in hyperbolic spaces see [LP], and see [Ba] for non-Euclidean balls in more general symmetric spaces when the lattice Γ is cocompact.

Example 1.7. Let $S_{n,k}$ denote the space of symmetric $n \times n$ matrices of determinant $k \neq 0$. SL_n acts on $S_{n,k}$ by $A \cdot g = {}^t g A g$. Then $S_{n,k}(\mathbf{R})$ is a union of symmetric spaces $SO(p, q)\backslash SL_n(\mathbf{R})$ with $p + q = n$, q even or odd depending on $\mathrm{sign}(k)$. Theorem 1.2 yields

$$N(T, S_{n,k}) \sim d_{n,k} T^{n(n-1)/2}. \tag{1.14}$$

This example generalizes easily to the variety of nondegenerate skew symmetric matrices of order $2n$ and of fixed Pfaffian.

Example 1.8. Let W_n denote the vector space of binary forms of degree $n \geqslant 3$ ($n = 2$ is again quadratic):

$$W_n = \{f(x, y) = a_0 x^n + a_1 x^{n-1} y + \cdots + a_n y^n\}. \tag{1.15}$$

$SL_2(\mathbf{C})$ acts on $W_n(\mathbf{C})$ by linear substitutions and the stabilizer of a generic form is finite. The corresponding orbits are closed and so are described by a finite number of polynomials in the coefficients $a_0, a_1, \ldots, a_n$, these being generators for the ring of invariants $\mathbf{C}[W_n]^{SL_2}$. On $W_n(\mathbf{R})$ define the Euclidean norm

$$\|(a_0, a_1, \ldots, a_n)\|^2 = \sum_{i=0}^{n} \binom{n}{i}^{-1} a_i^2. \tag{1.16}$$

For an $SL_2(\mathbf{C})$ orbit $\mathcal{O} \subset W_n$, let as usual

$$N(T, \mathcal{O}) = |\{f \in \mathcal{O}_{\mathbf{Z}}: \|f\| \leqslant T\}|. \tag{1.16}$$

This example falls into the $H\backslash G$ setup with H finite; however $H(\mathbf{R})\backslash G(\mathbf{R})$ is not (affine) symmetric so that Theorem 1.2 does not apply. We will prove the following statement.

THEOREM 1.9.

$$N(T, \mathcal{O}) \sim C_{\mathcal{O}} T^{2/n} \qquad as\ T \to \infty.$$

For $n = 3$ the discriminant $D(a_0, a_1, a_2, a_3) = a_0^2 a_3^2 + 18 a_0 a_1 a_2 a_3 - 4 a_0 a_2^3 - 4 a_1^3 a_3 - 27 a_1^2 a_2^2$ generates the ring of SL_2 invariants. Thus the varieties we obtain are

$$V_k = \{(a_0, a_1, a_2, a_3) | D(a) = k\} \tag{1.17}$$

and

$$N(T, V_k) \sim C_k T^{2/3}. \tag{1.18}$$

The constants $C_{\mathcal{O}}$ above, besides involving the usual arithmetical constants, also involve the numbers

$$C_f = \int_{-\infty}^{\infty} \frac{dx}{|f(1, x)|^{2/n}}; \tag{1.19}$$

see Section 4. Interestingly, C_f is invariant under $SL_2(\mathbf{R})$ but not $SL_2(\mathbf{C})$.

In these examples of varieties defined by level sets of a homogeneous form F of degree d in n variables, one expects heuristically cT^{n-d} for the asymptotics. The reasoning being that F assumes values in $[-c'T^d, cT^d]$ as $m \in \mathbf{Z}^n$ ranges in a ball of radius T. Furthermore, there are order T^n such points and one expects each value is assumed (roughly) equally often. This heuristic (up to factors of $\log T$) is accurate for Examples 1.5, 1.6, and 1.7, but (1.18) has an unexpectedly large number of points.

The method of proof of Theorem 1.2 in principal also allows us to obtain a remainder term. We have pursued this for the varieties $V_{n,k}$ of Example 1.6.

THEOREM 1.10.

$$N(T, V_{n,k}) = \mu(T) + O(T^{n^2-n-1/(n+1)+\eta}) \qquad \text{for all } \eta > 0.$$

For $n = 2$, which is the classical case of the upper half-plane, our remainder term of $O(T^{5/3})$ falls short of the best known remainder of $O(T^{4/3})$ due to Selberg [LP].

We now outline the contents of the rest of the paper. In Section 2 we prove Theorem 1.2. The problem is reduced to estimating the number of $\gamma \in G(\mathbf{Z})$ which lie in a certain family of regions R_T in $G(\mathbf{R})$. To do this one applies the spectral theory of functions on $G(\mathbf{Z})\backslash G(\mathbf{R})$. For simplicity we assume that the Euclidean norm on W satisfies

$$\|wk\| = \|w\| \qquad \text{for } k \in K \tag{1.20}$$

where $K \subset G(\mathbf{R})$ is a maximal compact subgroup[1]. This assumption can be removed, and we will indicate how to do so at the end of Section 2. An important issue is the study of the H-periods of $G(\mathbf{Z})\backslash G(\mathbf{R})/K$ eigenfunctions of the ring of invariant differential operators $\mathscr{D}(S)$ on $S = G(\mathbf{R})/K$. Precisely, if ϕ is such a function, let

$$\phi^H(g) = \int_{H(\mathbf{Z})\backslash H(\mathbf{R})} \phi(hg)\, dh \tag{1.21}$$

If this integral converges, as will be the case if ϕ is a cusp form, then ϕ^H is a left $H(\mathbf{R})$ and right K-invariant eigenfunction whose asymptotics at infinity may be studied. The corresponding integral, if ϕ is an Eisenstein series and $H \cap \Gamma\backslash H$ noncompact, often diverges and requires an elaborate regularization. We state the main result that will be needed here.

Let $P \subset G$ be a $\mathbf{Q}$-parabolic subgroup with Langlands decomposition $P = NAM$. Let $v \in C_{00}^{\infty}(A)$ and ϕ a cusp form on $M \cap \Gamma\backslash M$, $\Gamma = G(\mathbf{Z})$. Let

$$E_v(g) = \sum_{\gamma \in P \cap \Gamma\backslash\Gamma} \phi(m(\gamma g))v(a(\gamma g)). \tag{1.22}$$

---

[1] In many examples, such a norm is unique (up to scalar multiples) and is the "natural" norm for the problem.

Fourier inversion gives

$$E_v(g) = (2\pi)^{-\dim A} \int_{\mathrm{Re}(\lambda)=\lambda_0} \hat{v}(\lambda) E(\lambda, \phi, g) |d\lambda| \tag{1.23}$$

where

$$E(\lambda, \phi, g) = \sum_{\gamma \in \Gamma \cap P \backslash \Gamma} \phi(m(\gamma g)) e^{(\lambda+\rho)H(\gamma g)} \tag{1.24}$$

for $\lambda \in a_{\mathbf{C}}^*$ in the region of absolute convergence of the Eisenstein series (1.24).

Denote by $C_\lambda(H(\mathbf{R})\backslash G(\mathbf{R})/K)$ the space of eigenfunctions of the center $\mathscr{Z}(g)$ of the universal enveloping algebra of G, with infinitesimal character $\lambda \in (a_1)_{\mathbf{C}}^*/W$ (where $a_1 = Lie(A_1)$, $N_1 A_1 K$ an Iwasawa decomposition of $G(\mathbf{R})$), which are left $H(\mathbf{R})$ and right K-invariant. Let Ω be the convex hull of $\{w\rho | w \in W(G(\mathbf{R}), A_1)\}$ where ρ is half the sum of the positive roots. The general regularization reads as follows:

THEOREM 1.11. *There are measures $d\mu_j$, $j = 1, \ldots, v$ on $a_{\mathbf{C}}^*$ and meromorphic functions $E_j^H(g, \lambda)$, $\lambda \in a_{\mathbf{C}}^*$ such that*

$$E_v^H(g) = \langle E_v, 1 \rangle + \sum_{j=1}^{v} \int E_j^H(g, \lambda) \hat{v}(\lambda)\, d\mu_j(\lambda)$$

where $E_j^H(g, \lambda) \in C_{B_j(\lambda)}(H(\mathbf{R})\backslash G(\mathbf{R})/K)$ for $\lambda \in \mathrm{supp}(\mu_j)$ and $\mathrm{Re}(B_j(\lambda)) \in \Omega^0$ for $\lambda \in$ support μ_j.

The measures $d\mu_j$ correspond to contour integrals of varying dimensions and may be described explicitly. They arise from regularizing (1.21). As explained following Theorem 1.1, we will not prove this regularization in general. Of course, when $H \cap \Gamma\backslash H$ is compact, Theorem 1.11 is obvious by shifting contours.

In Section 3 we prove Theorem 1.10 and in Section 4, Theorem 1.9. In Appendix 1 we estimate certain integrals, and in Appendix 2 we prove a special case of Theorem 1.11.

Acknowledgement. We would like to thank E. van den Ban, M. Burger, A. Eskin, R. Howe, H. Schlichtkrull, and J. Silverman for their comments.

Section 2. In this section we prove Theorem 1.2. We assume $H(\mathbf{R})$ is symmetric and (1.5) holds. $\mathcal{O}$ is the $G(\mathbf{Z})$ orbit $w_0 G(\mathbf{Z})$ in W.

Define the function F_T on $G(\mathbf{R})$ by

$$F_T(g) = \sum_{\gamma \in H(\mathbf{Z})\backslash G(\mathbf{Z})} \chi_T(w_0 \gamma g) \tag{2.1}$$

where $\chi_T(w)$ is the characteristic function of $B_T = \{\|w\| \leqslant T\}$ in W. Clearly, we have

$$F_T(\gamma g) = F_T(g) \tag{2.2}$$

for $\gamma \in G(\mathbf{Z})$; that is, F_T lives on $G(\mathbf{Z})\backslash G(\mathbf{R})$. Note that

$$N(T, \mathcal{O}) = F_T(e). \tag{2.3}$$

For $\psi \in L^\infty(G(\mathbf{Z})\backslash G(\mathbf{R}))$ we have

$$\begin{aligned}\langle F_T, \psi\rangle &= \int_{G(\mathbf{Z})\backslash G(\mathbf{R})} F_T(g)\overline{\psi(g)}\,dg \\ &= \int_{G(\mathbf{Z})\backslash G(\mathbf{R})} \Bigg(\sum_{\gamma \in H(\mathbf{Z})\backslash G(\mathbf{Z})} \chi_T(w_0 \gamma g)\Bigg)\overline{\psi(g)}\,dg \\ &= \int_{H(\mathbf{Z})\backslash G(\mathbf{R})} \chi_T(w_0 g)\overline{\psi(g)}\,dg \\ &= \int_{H(\mathbf{R})\backslash G(\mathbf{R})} \chi_T(w_0 \dot{g}) \int_{H(\mathbf{Z})\backslash H(\mathbf{R})} \overline{\psi(h\dot{g})}\,dh\,d\dot{g}.\end{aligned}$$

Thus

$$\langle F_T, \psi\rangle = \int_{H(\mathbf{R})\backslash G(\mathbf{R})} \chi_T(w_0 \dot{g})\overline{\psi^H(\dot{g})}\,d\dot{g} \tag{2.4}$$

where

$$\psi^H(g) = \int_{H(\mathbf{Z})\backslash H(\mathbf{R})} \psi(hg)\,dh \tag{2.5}$$

is the H-period of ψ, which is plainly left $H(\mathbf{R})$-invariant.

Applying (2.4) to $\psi \equiv 1$ and using $F_T \geqslant 0$, we have

$$\|F_T\|_{L^1} = \mu(T). \tag{2.6}$$

In general, this is all we can say about F_T. That is, in general, it need not be in L^2 (see Appendix 2 for an example). It is this that makes the spectral analysis of F_T delicate. Set

$$\tilde{F}_T = \frac{1}{\mu(T)} F_T. \tag{2.7}$$

We may view $\tilde{F}_T$ as a probability measure on $G(\mathbf{R})\backslash G(\mathbf{R})$, and in view of (2.3) Theorem 1.2 will follow from the following theorem.

THEOREM 2.1. *$\tilde{F}_T(g) \to 1$ for each fixed g.*

We can rewrite F_T in the form

$$F_T(g) = \sum_{\substack{m \in \mathcal{C} \\ mg \in B_T}} 1 = \sum_{\substack{m \in \mathcal{C} \\ m \in B_T g^{-1}}} 1. \tag{2.8}$$

LEMMA 2.2. *Let $L \subset G(\mathbf{Z})\backslash G(\mathbf{R})$ be compact. Then there is $\kappa = \kappa(L) > 1$ such that for $g_1, g_2 \in L$ and $T \geqslant 1$*

$$F_{\kappa^{-1}T}(g_2) \leqslant F_T(g_1) \leqslant F_{\kappa T}(g_2).$$

Moreover, as $L \to \{e\}$, $\kappa(L) \to 1$.

Proof. Since G acts linearly on W, it is clear that for L compact there is $\kappa(L)$ such that

$$B_{\kappa^{-1}T} g_2^{-1} \subset B_T g_1^{-1} \subset B_{\kappa T} g_2^{-1}$$

and that $\kappa \to 1$ as $L \to \{e\}$. The lemma follows immediately.

LEMMA 2.3. *In order to prove Theorem 2.1, it suffices to prove that $\tilde{F}_T \to 1$ in the w-star topology, that is*

$$\langle \tilde{F}_T, \psi \rangle \to \langle 1, \psi \rangle = \int_{G(\mathbf{Z})\backslash G(\mathbf{R})} \overline{\psi(g)}\, dg \tag{2.9}$$

for any fixed $\psi \in C_0(G(\mathbf{Z})\backslash G(\mathbf{R}))$.

Proof. Fix $g_0 \in G(\mathbf{R})$. Let $\psi_\varepsilon \geqslant 0$ be an approximation to the identity near g_0. Precisely, $\psi_\varepsilon \in C_{00}(G(\mathbf{Z})\backslash G(\mathbf{R}))$, $\int_{G(\mathbf{Z})\backslash G(\mathbf{R})} \psi_\varepsilon(g)\, dg = 1$ and $\psi_\varepsilon = 0$ outside a compact neighborhood U_ε of g_0 where $U_\varepsilon \to \{g_0\}$ as $\varepsilon \to 0$. Applying Lemma 2.2, we have

$$\tilde{F}_T(g_0) \leqslant \frac{\mu(\kappa_\varepsilon T)}{\mu(T)} \int_{G(\mathbf{Z})\backslash G(\mathbf{R})} \psi_\varepsilon(g) \tilde{F}_{\kappa_\varepsilon T}(g)\, dg. \tag{2.10}$$

Now by Appendix 1, we have

$$b(\kappa_\varepsilon) \leqslant \liminf \frac{\mu(\kappa_\varepsilon T)}{\mu(T)} \leqslant \limsup \frac{\mu(\kappa_\varepsilon T)}{\mu(T)} \leqslant a(\kappa_\varepsilon)$$

with $a(\kappa)$, $b(\kappa) \to 1$ as $\kappa \to 1$, while by assumption

$$\langle \tilde{F}_{\kappa_\varepsilon T}, \psi_\varepsilon \rangle \to \langle 1, \psi_\varepsilon \rangle = 1.$$

Thus letting $T \to \infty$ in (2.10) yields

$$\limsup_{T\to\infty} \tilde{F}_T(g_0) \leqslant 1 .$$

In a similar way, one deduces that

$$\liminf_{T\to\infty} \tilde{F}_T(g_0) \geqslant 1 ,$$

completing the proof of Lemma 2.3.

One further remark: Since $\|\tilde{F}_T\| = 1$, it clearly suffices to check (2.9) for a dense set of functions in $C_0(G(\mathbf{Z})\backslash G(\mathbf{R}))$. We will do so for certain eigenfunctions of the center of the universal enveloping algebra. At this point we assume that (1.20) holds, and so $\tilde{F}_T(g)$ satisfies

$$\tilde{F}_T(gk) = \tilde{F}_T(g), \qquad \text{for } k \in K . \tag{2.11}$$

This assumption can easily be relaxed to K-finite functions and hence to deal with the general $\| \ \|$ on W.[2]

Thus $\tilde{F}_T \in L^1(G(\mathbf{Z})\backslash G(\mathbf{R})/K)$, and we may stick to ψ's on the same space. The ψ's we consider are eigenfunctions of the center of the universal enveloping algebra $\mathscr{Z}(g)$. The infinitesimal character of such an eigenfunction corresponds to a point $\wedge \in a_{\mathbf{C}}^*/W$ where $NAK = G(\mathbf{R})$ is an Iwasawa decomposition, $a = Lie(A)$. Since $G(\mathbf{Z})\backslash G(\mathbf{R})$ need not be compact, we need to include certain Eisenstein wave packets in order to get a dense set. Precisely, for each $\mathbf{Q}$-parabolic subgroup P of G, let

$$P = N_P A_P M_P \tag{2.12}$$

be a Langlands decomposition of P. For $v \in C_{00}^\infty(A_P)$ and ϕ a cusp form [HC] on $M_P \cap \Gamma\backslash M_P/M_P \cap K$, $\Gamma = G(\mathbf{Z})$, let

$${}_PE_v(\phi, g) = \sum_{\gamma \in \Gamma \cap P\backslash\Gamma} \phi(m_P(\gamma g)) v(a_P(\gamma g)) . \tag{2.13}$$

This is a finite series and the resulting function ${}_PE_v \in C_0(G(\mathbf{Z})\backslash G(\mathbf{R})/K)$. According to standard convention, $P = G$ is also allowed, in which case the above function is just a cusp form on $G(\mathbf{Z})\backslash G(\mathbf{R})/K$.

LEMMA 2.4. *The functions ${}_PE_v(\phi, g)$, as P runs over parabolics, $v \in C_{00}^\infty(A)$, and ϕ runs over cusp forms on $M_P \cap \Gamma\backslash M_P/K \cap M_P$, are dense in $C_0(G(\mathbf{Z})\backslash G(\mathbf{R})/K)$.*

Proof. Let ν be a measure of finite (total) variation on $G(\mathbf{Z})\backslash G(\mathbf{R})/K$. We must show that, if $\langle \nu, {}_PE_v(\phi, \cdot)\rangle = 0$ for all P, v, ϕ, then $\nu = 0$.

---

[2] Theorem 2.6 below holds for K-finite functions, while the rest of the argument can be carried out along the lines of §4.

Let $k_\varepsilon(z, w)$ be a point pair invariant on $S \times S$ [Se] of compact support which as $\varepsilon \downarrow 0$ is an approximation to the identity. Let

$$K_\varepsilon(z, w) = \sum_{\gamma \in G(\mathbf{Z})} k_\varepsilon(\gamma z, w) \tag{2.14}$$

and

$$f_\varepsilon(z) = \int_{G(\mathbf{Z})\backslash G(\mathbf{R})} K_\varepsilon(z, w)\, dv(w). \tag{2.15}$$

Then $f_\varepsilon(z) \in C^\infty(G(\mathbf{Z})\backslash G(\mathbf{R})/K)$ and is of moderate growth. Moreover,

$$\langle f_\varepsilon, \psi \rangle \to \langle v, \psi \rangle \qquad \text{as } \varepsilon \to 0 \tag{2.16}$$

for all $\psi \in C_{00}(G(\mathbf{Z})\backslash G(\mathbf{R})/K)$. It is easy to see that, for each ε, f_ε satisfies the same hypothesis as v, i.e.

$$\langle f_\varepsilon, {}_PE_v(\phi, \cdot) \rangle = 0 \qquad \text{for all } P, v, \phi. \tag{2.17}$$

The last implies $f_\varepsilon \equiv 0$ by a theorem of Langlands (see [HC, Theorem 4]). Hence in view of (2.16), $v = 0$ as needed.

This lemma allows us to deal with cuspidal Eisenstein series only. This is an important technical point since we avoid the difficulties associated with residual Eisenstein series. Indeed, those are needed for L^2-decompositions, which is not appropriate in our problem since $\tilde{F}_T$ is not necessarily in L^2.

Our analysis shows that it suffices to prove the following proposition.

Proposition 2.5.

$$\langle \tilde{F}_T, {}_PE_v(\phi, \cdot) \rangle \to \langle 1, {}_PE_v(\phi, \cdot) \rangle$$

for all such E's.

We begin by proving this when $P = G$, i.e. ϕ is a cusp form on $G(\mathbf{Z})\backslash G(\mathbf{R})/K$. If $G(\mathbf{Z})\backslash G(\mathbf{R})$ is compact, this case would be the whole story. ("Cusp form" then means an eigenfunction.) The function ϕ (as well as the general $E_v(\phi, \cdot)$) is rapidly decreasing in the cusps of $G(\mathbf{Z})\backslash G(\mathbf{R})$ [HC], so that calculation (2.4) applies and yields

$$\langle \tilde{F}_T, \bar{\phi} \rangle = \frac{1}{\mu(T)} \int_{H(\mathbf{R})\backslash G(\mathbf{R})} \chi_T(w_0 \dot{g}) \phi^H(\dot{g})\, d\dot{g}. \tag{2.18}$$

Also from (2.5) (which converges) $\phi^H(g)$ is an eigenfunction of $\mathscr{Z}(g)$ and is both left $H(\mathbf{R})$ and right K-invariant. Denote by $C_\Lambda(H(\mathbf{R})\backslash G(\mathbf{R})/K)$ the linear space of such eigenfunction whose infinitesimal character is $\wedge \in \mathfrak{a}_{\mathbf{C}}^*/W$. Now we use heavily the fact that $H(\mathbf{R})$ is symmetric. This allows one to conclude [FJ] that $C_\wedge(H(\mathbf{R})\backslash G(\mathbf{R})/K)$ is finite-dimensional. Let $\Omega \subset \mathfrak{a}_{\mathbf{R}}^*$ be the convex hull of $\{w\rho \mid w \in W\}$ and $\rho \in \mathfrak{a}_{\mathbf{R}}^*$ is half the sum of the positive roots. Let Ω^0 be its interior. We will need the following theorem due to Rudnick and Schlichtkrull [RS].

THEOREM 2.6. *If* $\operatorname{Re}(\wedge) \in \Omega^0$ *and* $\psi \in C_\wedge(H(\mathbf{R})\backslash G(\mathbf{R})/K)$, *then* $\psi(\dot{g}) \to 0$ *as* $\dot{g} \to \infty$ *in* $H(\mathbf{R})\backslash G(\mathbf{R})$.

Returning to (2.18), if ϕ is a (nonconstant) cusp form on $G(\mathbf{Z})\backslash G(\mathbf{R})/K$ with eigenvalue $\wedge$, then since ϕ appears in the L^2 spectrum, the Howe-Moore theorem [HM], [BW] implies that $\operatorname{Re}(\wedge) \in \Omega^0$ and hence that

$$\phi^H(\dot{g}) \to 0, \qquad \text{as } \dot{g} \to \infty \text{ in } H(\mathbf{R})\backslash G(\mathbf{R}). \tag{2.19}$$

This is where we use the assumption that G is a $\mathbf{Q}$-simple connected $\mathbf{Q}$-group; this ensures Howe-Moore for nontrivial spherical constituents of $L^2(\Gamma\backslash G)$. In the setting of Example 1.6, where $\Gamma\backslash G = \Gamma_1\backslash G_1 \times \Gamma_1\backslash G_1$, and $H = \Delta \cong G_1$ is the diagonal subgroup, we further need to note that, although there are representations in $L^2(\Gamma\backslash G)$ whose matrix coefficients do not decay at infinity, the only irreducible nontrivial representations with nonzero H-periods are of the form $\Pi = \pi_1 \otimes \tilde{\pi}_1$, with π_1 an irreducible unitary representation of G_1, and so except for the constants, all eigenfunctions ϕ for which $\phi^H \neq 0$ have their Langlands parameter in Ω^0 and so satisfy (2.19).

From (2.18) it follows that $\langle \tilde{F}_T, \bar{\phi} \rangle \to 0 = \langle 1, \bar{\phi} \rangle$ as $T \to \infty$ as needed. We note that, if $\wedge \in i\mathfrak{a}_{\mathbf{R}}^*/W$, that is ϕ belongs to the tempered spectrum (which is what is expected for most cusp forms by the general Ramanujan conjectures; see [BLS], [Sa]), then the result of Rudnick-Schlichtkrull mentioned earlier ensures that, for $\varepsilon > 0$,

$$\phi^H \in L^{2+\varepsilon}(H(\mathbf{R})\backslash G(\mathbf{R})).$$

Thus

$$\langle F_T, \phi \rangle = O_\varepsilon(\mu(T)^{1/2+\varepsilon}).$$

That is, these frequencies contribute at most the square-root of the leading term!

We now deal with the more general ${}_PE_v(\phi, g)$. One cannot directly deal with the Eisenstein series (1.24) since (2.5) may not converge. Instead, we apply (2.4) to ${}_PE_v(\phi, g)$ and use the regularization Theorem 1.11. This gives

$$\begin{aligned} &\langle \tilde{F}_T, E_v(\phi, \cdot) \rangle = \langle E_v(\phi, \cdot), 1 \rangle \\ &\qquad + \sum_{j=1}^{v} \int \left(\frac{1}{\mu(T)} \int_{H(\mathbf{R})\backslash G(\mathbf{R})} \chi_T(w_0 \dot{g}) E_j^H(\dot{g}, \wedge)\, d\dot{g} \right) \hat{\vartheta}(\wedge)\, d\mu_j(\wedge). \end{aligned} \tag{2.20}$$

The integrals are all absolutely convergent, a fact which follows from $\hat{v}(\wedge)$ being rapidly decreasing in $\mathrm{Im}(\wedge)$, while $E_j^H(g, \wedge)$ is polynomially bounded in $\mathrm{Im}(\wedge)$.

Now for $\wedge$ fixed in the support of $d\mu_j$,

$$\frac{1}{\mu(T)} \int_{H\backslash G} \chi_T(w_0 \dot{g}) E_j^H(\dot{g}, \wedge)\, d\dot{g} \to 0$$

as $T \to \infty$. The reason is that, by Theorem 1.11, $E_j^H(g, \wedge) \in C_{B_j(\wedge)}(H(\mathbf{R})\backslash G(\mathbf{R})/K)$ with $\mathrm{Re}\, B_j(\wedge) \in \Omega^0$, and so Theorem 2.6 applies. Thus the result that

$$\langle \tilde{F}_T, {}_PE_v(\phi, \cdot) \rangle \to \langle 1, {}_PE_v(\phi, \cdot) \rangle$$

follows from the convergence theorem and the following lemma [Ru]:

Lemma 2.7. *Assume G is classical. Then for C a large compact set in $G(\mathbf{R})$, there is a constant c such that, for* $\mathrm{Re}(\wedge) \in \Omega$ *and an $\phi \in C_\wedge(H(\mathbf{R})\backslash G(\mathbf{R})/K)$,*

$$|\phi(g)| \leqslant c \sup_{g_1 \in C} |\phi(g_1)|.$$

Section 3. In this section we prove Theorem 1.10. We do so for $V_{n,1} = \{(x_{ij}) | \det x_{ij} = 1\}$, $G = SL_n$. The more general case of $V_{n,k}$ is dealt with in a similar way by considering each of the $\Gamma = SL_n(\mathbf{Z})$ orbits on matrices of determinant k, separately. For $V_{n,1}$ we prove the following more general theorem.

Theorem 3.1. *Let $n \geqslant 3$ and $\Gamma \subset SL_n(\mathbf{R})$ be any lattice. Set*

$$N(T, \Gamma) = \sum_{\|\gamma\| \leqslant T} 1$$

where $\|g\|^2 = \mathrm{tr}({}^tgg)$. Then

$$N(T, \Gamma) = \mu(T) + O_\eta(T^{n^2-n-(1/(n+1))+\eta}), \qquad \textit{for all } \eta > 0.$$

Here $\mu(T)$ as in (1.7) is defined to be

$$\mu(T) = \int_{\substack{SL_n(\mathbf{R}) \\ \|g\| \leqslant T}} dg.$$

Its asymptotics are described in Appendix 1.

The above theorem is not valid for $n = 2$ since $\Gamma\backslash SL_2(\mathbf{R})$ may have small eigenvalues [Ra]. However, if $\Gamma = SL_2(\mathbf{Z})$ or a congruence subgroup thereof, then the result is true with the remainder term of the form $O(T^{4/3})$, a result due to Selberg [LP].

We turn to the proof of Theorem 3.1. Here $G = SL_n(\mathbf{R})$, $K = SO_n(\mathbf{R})$, and G/K is the Riemannian symmetric space of positive definite matrices of determinant 1. What we exploit here over and above the techniques of the last section is the multiplicity one of zonal spherical functions. That is, for an $\wedge \in a_{\mathbf{C}}^*$, $C_\wedge(K\backslash G/K)$ is exactly one-dimensional [Se]. In fact, its unique member $\phi_\wedge(g)$ which is 1 at $g = e$ is given by

$$\phi_\wedge(g) = \int_K e^{(\wedge+\rho)H(kg)}\,dk \tag{3.1}$$

where ρ is as usual and elements of $a = \{\text{trace zero diagonal matrices}\}$ are denoted by H.

Let k_T be defined on $G(\mathbf{R})$ by

$$k_T(g) = \begin{cases} 1, & \text{if } \|g\| \leqslant T \\ 0, & \text{otherwise}. \end{cases} \tag{3.2}$$

Note that $k_T(k_1 g k_2) = k_T(g)$, $k_1, k_2 \in K$. Set

$$K_T(g, h) = \sum_{\gamma \in \Gamma} k_T(g^{-1}\gamma h). \tag{3.3}$$

So

$$N(T, \Gamma) = K_T(I, I) \tag{3.4}$$

and

$$K_T(\gamma_1 g k_2, \gamma_2 h k_2) = K_T(g, h) \tag{3.5}$$

for $\gamma_1, \gamma_2 \in \Gamma$, $k_1, k_2 \in K$.

We examine the precise behavior of $K_T(x, y)$ near (I, I) in the space $G/K \times G/K$ of positive definite matrices of determinant equal to 1.

Define $\| \ \|_\infty$ on $G(\mathbf{R})$ by

$$\|b\|_\infty^2 = \lambda_{\max}({}^t bb). \tag{3.6}$$

Let

$$B_\varepsilon = \{x = {}^t gg \colon \|g\|_\infty < 1 + \varepsilon,\ \|g^{-1}\|_\infty < 1 + \varepsilon\}. \tag{3.7}$$

LEMMA 3.2. *For $x, y \in B_\varepsilon$,*

$$K_{T(1+\varepsilon)^{-2}}(x, y) \leqslant K_T(I, I) \leqslant K_{T(1+\varepsilon)^2}(x, y).$$

This will follow from the next lemma.

LEMMA 3.3. *For $b, c \in G(\mathbf{R})$*

$$\|bc\| \leqslant \|b\|_\infty \|c\| \quad \textit{and} \quad \|bc\| \leqslant \|b\|\,\|c\|_\infty .$$

Proof. Let $0 \leqslant \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ be the eigenvalues of ${}^t cc$ and $0 \leqslant \lambda_1' \leqslant \lambda_2' \leqslant \cdots \leqslant \lambda_n'$, those of ${}^t(bc)bc$. We first show that

$$\lambda_i' \leqslant \|b\|_\infty^2 \lambda_i \qquad i = 1, \ldots, n. \tag{3.8}$$

Write ${}^t bb = {}^t k\, dk$ where $k \in O_n(\mathbf{R})$ and $d \geqslant 0$ is diagonal. Define

$$\tau = {}^t k(\|b\|_\infty^2 I - d)k \geqslant 0$$

by (3.6). But

$$\|b\|_\infty^2\, {}^t cc = {}^t c({}^t bb + \tau)c = {}^t(bc)bc + {}^t c\tau c.$$

Since ${}^t c\tau c \geqslant 0$, we deduce (3.8).

Now we can write the singular value decomposition of $c = k_1 a k_2$ where a is diagonal. Of course $\|c\| = \|a\|$. Further, $\|a\|$ defines a norm on $\mathbf{R}^n$ which depends only on the absolute values of the components of a. Thus

$$\begin{aligned} \|bc\| &= \|(\lambda_1')^{1/2}, \ldots, (\lambda_n')^{1/2}\| \\ &\leqslant \|b\|_\infty \|(\lambda_1^{1/2}, \ldots, \lambda_n^{1/2})\| \\ &= \|b\|_\infty \|c\| \qquad \text{by (3.8)}. \end{aligned}$$

The second inequality follows from the first together with the obvious fact that $\|{}^t b\| = \|b\|$.

Note 3.4. The above proof, as well as what follows, would apply just as well to any norm $\|\ \|$ on $G(\mathbf{R})$ satisfying

$$\|k_1 g k_2\| = \|g\|, \qquad k_1, k_2 \in K.$$

We turn to the proof of Theorem 3.1. Choose $\psi \in C_0^\infty(\Gamma\backslash G(\mathbf{R})/K)$ supported in B_ε and such that $\psi \geqslant 0$, $\int_{\Gamma\backslash G(\mathbf{R})/K} \psi(x)\,dx = 1$. According to Lemma 3.2,

$$H(T(1+\varepsilon)^{-2}) \leqslant N(T, \Gamma) \leqslant H(T(1+\varepsilon)^2) \tag{3.9}$$

where

$$H(T) = \iint\limits_{\Gamma\backslash G(\mathbf{R})/K \times \Gamma\backslash G(\mathbf{R})/K} \psi(x)\overline{\psi(y)} K_T(x, y)\,dx\,dy. \tag{3.10}$$

On the other hand, $H(T)$ can be expanded in the spectrum of $L^2(\Gamma\backslash G(\mathbf{R})/K)$. In this case it would be in terms of cusp forms, unitary Eisenstein series as well as residual Eisenstein series [L]. However, (3.10) is a purely L^2-statement, and so one can simply appeal to the abstract spectral theorem for the selfadjoint ring $\mathscr{D}(G(\mathbf{R})/K)$ to get

$$H(T) = \int_\sigma h_T(\wedge)|\hat{\psi}(\wedge)|^2\, d\mu(\wedge) \tag{3.11}$$

where $\sigma \subset a_{\mathbf{C}}^*$ denotes the spectrum of $L^2(\Gamma\backslash G(\mathbf{R})/K)$, $h_T(\wedge)$ is the Selberg transform

$$h_T(\wedge) = \int_{G(\mathbf{R})} \phi_\wedge(g)k_T(g)\, dg, \tag{3.12}$$

and $\hat{\psi}(\wedge)$ is the spectral transform of $\psi(g)$. In particular, we have Parseval's formula

$$\int_{\Gamma\backslash G} |\psi(g)|^2\, dg = \int_\sigma |\hat{\psi}(\wedge)|^2\, d\mu(\wedge). \tag{3.13}$$

Note that (3.11) uses the multiplicity-one theorem for zonal spherical functions. In (3.11) we separate out the main term which comes from the constant function—that is, from $\wedge = \rho$ in σ.

$$H(T) = h_T(\rho) + \mathrm{Rem}(T) \tag{3.14}$$

where

$$\mathrm{Rem}(T) = \int_{\substack{\wedge \neq \rho \\ \wedge \in \sigma}} h_T(\wedge)|\hat{\psi}(\wedge)|^2\, d\mu(\wedge). \tag{3.15}$$

Of course,

$$h_T(\rho) = \int_{G(\mathbf{R})} k_T(g)\, dg = \mu(T). \tag{3.16}$$

We know that $\mu(T) = O(T^{n^2-n})$; hence it follows from (3.9) and (3.14) that

$$N(T, \Gamma) = \mu(T) + O(\varepsilon T^{n^2-n} + \mathrm{Rem}(T)). \tag{3.17}$$

To estimate $\mathrm{Rem}(T)$ we need to know a little about the spectrum σ, which we note is contained in the unitary spherical dual of $SL_n(\mathbf{R})$. From the classification due to Vogan [V] (see also [Sca]), one can show that, for $\wedge \in a_{\mathbf{C}}^*/W$, $\wedge \neq \rho$, and $\phi_\wedge(g)$

the corresponding spherical function, we have uniformly

$$\|\phi_\Lambda\|_{L^p} \leqslant C_p < \infty \tag{3.18}$$

for $p > 2(n-1)$.

Thus for $\Lambda \in \sigma$, $\Lambda \neq \rho$, we have

$$|h_T(\Lambda)| \leqslant \left(\int_G |k_T(g)|^q \, dy\right)^{1/q} \cdot C_p^{1/p} \leqslant (\mu(T))^{1/q} C_p^{1/p} \tag{3.19}$$

where $1/p + 1/q = 1$ and $p > 2(n-1)$ is fixed. Thus from (3.15)

$$\operatorname{Rem}(T) \ll (\mu(T))^{1/q} \int_{\sigma - \{\rho\}} |\hat{\psi}(\Lambda)|^2 \, d\Lambda. \tag{3.20}$$

From (3.13)

$$\begin{aligned}\int_{\sigma - \{\rho\}} |\hat{\psi}(\Lambda)|^2 \, d\Lambda &\leqslant \int_{\Gamma\backslash G/K} |\psi_\varepsilon(g)|^2 \, dg \\ &= O(\varepsilon^{1-(n+1)n/2}).\end{aligned}$$

Since ψ_ε is an approximation to the identity on an $(n(n+1)/2 - 1)$-dimensional space, it can be chosen to be of L^2-norm as above. This gives

$$\operatorname{Rem}(T) \ll_\eta T^{n^2 - (3/2)n + \eta} \varepsilon^{1 - n(n+1)/2}$$

for any $\eta > 0$ (using $p > 2(n-1)$). Hence returning to (3.17),

$$N(T, \Gamma) = \mu(T) + O_\eta(\varepsilon T^{n^2 - n} + T^{n^2 - (3/2)n + \eta} \varepsilon^{1 - n(n+1)/2}).$$

To minimize the remainder choose

$$\varepsilon = T^{-1/(n+1)},$$

and we obtain Theorem 3.1.

As mentioned before, the above arguments are exactly adapted to deal with other K-bi-invariant norms. For example,

$$\|g\|^2 = \operatorname{tr}(({}^t g g)^2)$$

arises naturally from the action of $G = SL_n$ on positive definite symmetric matrices $sym^+(n)$. We obtain in this way the following theorem.

THEOREM 3.5. *Let*

$$N(T) = \sum_{\substack{B \in sym^+(n, \mathbf{Z}) \\ \det B = k \\ \|B\| \leqslant T}} 1$$

where $\|B\|^2 = \operatorname{tr}({}^tBB)$; *then*

$$N(T) = \mu_1(T) + O_\eta(T^{(n^2-n)/2-1/2(n+1)+\eta}) \qquad \textit{for } \eta > 0.$$

Here

$$\mu_1(T) \sim c_{n,k} T^{(n^2-n)/2}$$

for a suitable nonzero constant $c_{n,k}$ (assuming of course that $k \geqslant 1$).

Section 4. Our aim in this section is to prove Theorem 1.9 and some related results. This case involves $G = SL_2$ and H finite. We hope the method below will form a basis for the general case when $H(\mathbf{Z})$ is finite.

We begin with some results concerning lattice points in regions in $\mathscr{H}$—the hyperbolic plane. Let $\Gamma \leqslant SL_2(\mathbf{R})$ be a lattice and let

$$N(\Gamma, R) = |\{\gamma \in \Gamma | d(\gamma i, i) \leqslant R\}| \tag{4.1}$$

where $d(z, w)$ is the non-Euclidean distance. (Of course, one could consider $d(\gamma z, w)$ for $z, w \in \mathscr{H}$ as well.) It is well known (Delsarte [D] in the cocompact case and Selberg [LP] in the finite-volume case), and is also a special case of Theorem 1.2, that

$$N(\Gamma, R) \sim \frac{\pi}{\operatorname{Vol}(\Gamma \backslash \mathscr{H})} e^R \qquad \text{as } R \to \infty. \tag{4.2}$$

The usual argument leading to this may easily be extended to include the case of lattice points in sectors. That is, if (r, θ) are geodesic polar coordinates about i and

$$N(\Gamma, R, I) = |\{\gamma \in \Gamma | d(\gamma i, i) \leqslant R, \theta(\gamma i) \in I\}| \tag{4.3}$$

for $I \subset [0, 2\pi]$, then

$$N(\Gamma, R, I) \sim \frac{l(I) e^R}{2 \operatorname{Vol}(\Gamma \backslash \mathscr{H})}, \tag{4.4}$$

$l(I)$ being the length of I.

We extend (4.4) to cover regions $R_T \subset \mathscr{H}$ of a more complicated nature. Let $p_j(\theta)$, $-n \leqslant j \leqslant n$, be trigonometric polynomials. Let R_T be the region defined by

$$|e^{nr/2}p_n(\theta)|^2 + |e^{(n-2)r/2}p_{n-2}(\theta)|^2 + \cdots + |e^{-nr/2}p_{-n}(\theta)|^2 \leqslant T^2. \tag{4.5}$$

Let

$$N(\Gamma, R_T) = |\{\gamma \in \Gamma \mid \gamma i \in R_T\}|. \tag{4.6}$$

Proposition 4.1. *Assume that*

$$K = \int_0^{2\pi} \frac{d\theta}{|p_n(\theta)|^{2/n}} < \infty \tag{4.7}$$

and that $(p_n(\theta), \ldots, p_{-n}(\theta)) \neq (0, \ldots, 0)$ *for all* $\theta \in [0, 2\pi]$. *Then*

$$N(\Gamma, R_T) \sim \frac{K}{2\,\mathrm{Vol}(\Gamma\backslash\mathscr{H})} T^{2/n} \qquad as\ T \to \infty.$$

Note. If $n \geqslant 3$ and the zeros of p_n are simple, then $K < \infty$; this will be used in applications. Also the assumption about the nonvanishing of $(p_n(\theta), \ldots, p_{-n}(\theta))$ is equivalent to R_T being compact.

We will need the following general upper bound.

Lemma 4.2. *Let R be a connected region in $\mathscr{H}$; then*

$$N(\Gamma, R) \ll \mathrm{Vol}(R) + \mathrm{length}\,(\partial R).$$

Proof. Without loss of generality, we can assume that Γ has no torsion. Let $\varepsilon_0 = \varepsilon_0(\Gamma)$ be such that

$$d(\delta i, \gamma i) = d(i, \delta^{-1}\gamma i) > \varepsilon_0$$

if $\delta \neq \gamma$. For $\gamma \in \Gamma$ such that $\gamma i \in R$ let $B_{\varepsilon_0/2}(\gamma i) = \gamma B_{\varepsilon_0/2}(i)$ be the ball about γi of radius $\varepsilon_0/2$. Now either $B_{\varepsilon_0/2}(\gamma i) \subset R$ in which case let $m(\gamma) = \mathrm{Vol}(B_{\varepsilon_0/2}(\gamma i))$, or $\partial R \cap B_{\varepsilon_0/2}(\gamma i) \neq \emptyset$. In this latter case let C be the connected component of $R \cap B_{\varepsilon_0/2}(\gamma i)$ containing γi (see Figure 4.3).

Since R is connected, C must meet $\partial B_{\varepsilon_0/2}(\gamma i)$. There are two possibilities: either $B_{\varepsilon_0/4}(\gamma i)$ meets ∂C, in which case $|\partial C| \geqslant \varepsilon_0/4$ and $|\partial R \cap B_{\varepsilon_0/2}(\gamma i)| \gg \varepsilon_0^2/4$, or $R \supset B_{\varepsilon_0/4}(\gamma i)$ and $\mathrm{Vol}(R \cap B_{\varepsilon_0/2}(\gamma i)) \geqslant \varepsilon_0^2/4$. In this way we associate to each γ with $\gamma i \in R$ either a subregion $B_\gamma \supset R$ of volume $m(\gamma) \gg 1$ or a subset of ∂R of length $n(\gamma) \gg 1$.

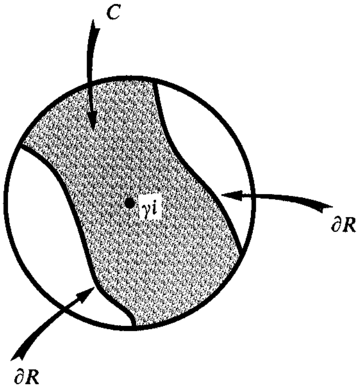

FIGURE 4.3

Moreover, these sets are all disjoint as we vary over $\gamma \in \Gamma$. Hence

$$\sum_{\substack{\gamma \in \Gamma \\ \gamma i \in R}} 1 \ll \sum_{\gamma \in \Gamma} m(\gamma) + n(\gamma) \ll \operatorname{Vol}(R) + \operatorname{length}(\partial R),$$

which proves the lemma.

LEMMA 4.4. *Let f be continuous on $[\alpha, \beta] \subset [0, 2\pi]$ and $f(\theta) \neq 0$ for $\theta \in [\alpha, \beta]$. Let $R_{T,f} = \{(r, \theta) | r \leqslant v \log|T/f(\theta)|, \alpha \leqslant \theta \leqslant \beta\}$ where v is a positive constant. Then*

$$N(\Gamma, R_T, f) \sim \frac{T^v}{2 \operatorname{Vol}(\Gamma \backslash \mathscr{H})} \int_\alpha^\beta \frac{d\theta}{|f(\theta)|^v} \qquad \textit{as } T \to \infty. \tag{4.8}$$

Proof. The lemma follows from (4.3) by an obvious approximation argument.

We now prove Proposition 4.1. Clearly $p_n(\theta) \not\equiv 0$; so let $\theta_1, \ldots, \theta_l$ be its zeros. Fix $\varepsilon > 0$ and define $R_T^{(\varepsilon)}$ and $S_T^{(\varepsilon)}$ by: R_T^ε consists of all (γ, θ) satisfying (4.5) with $\theta \notin I_\varepsilon = \bigcup_{j=1}^l (\theta_j - \varepsilon, \theta_j + \varepsilon)$ and $S_T^{(\varepsilon)} = R_T \backslash R_T^{(\varepsilon)}$. From Lemma 4.4 one deduces that

$$N(\Gamma, R_T^{(\varepsilon)}) \sim \left(\frac{1}{2} \int_{[0, 2\pi] \backslash I_\varepsilon} \frac{d\theta}{|p_n(\theta)|^{2/n}}\right) \frac{T^{2/n}}{\operatorname{Vol}(\Gamma \backslash \mathscr{H})} \tag{4.9}$$

as $T \to \infty$.

We claim that

$$N(\Gamma, S_T^{(\varepsilon)}) \leqslant \alpha(\varepsilon) T^{2/n} \qquad \text{as } T \to \infty \tag{4.10}$$

where $\alpha(\varepsilon) \to 0$ as $\varepsilon \to 0$.

Granting this, we have for each fixed ε

$$\limsup_{T\to\infty} \frac{N(\Gamma, R_T)}{T^{2/n}} \leqslant \left(\frac{1}{2}\int_{[0,2\pi]\setminus I_\varepsilon} \frac{d\theta}{|p_n(\theta)|^{2/n}}\right)\frac{1}{\operatorname{Vol}(\Gamma\backslash\mathscr{H})} + \alpha(\varepsilon)$$

and

$$\liminf_{T\to\infty} \frac{N(\Gamma, R_T)}{T^{2/n}} \geqslant \left(\frac{1}{2}\int_{[0,2\pi]\setminus I_\varepsilon} \frac{d\theta}{|p_n(\theta)|^{2/n}}\right)\frac{1}{\operatorname{Vol}(\Gamma\backslash\mathscr{H})}.$$

Since $K < \infty$ and $\alpha(\varepsilon) \to 0$ as $\varepsilon \to 0$, Proposition 4.1 follows from the last two inequalities.

We turn to (4.10). It suffices to consider each such $(\theta_j - \varepsilon, \theta, +\varepsilon)$ separately. Say $\theta_1 = 0$ is a zero of $p_n(\theta)$ of order m. Now $'R_T^\varepsilon = R_R \cap \{\theta \,|\, |\theta| < \varepsilon\}$ is contained in both

(I) $$r \leqslant \log\left|\frac{T}{p_l(\theta)}\right|^{2/l}, \qquad |\theta| < \varepsilon,$$

(II) $$r \leqslant \log\left|\frac{T}{p_n(\theta)}\right|^{2/n}, \qquad |\theta| < \varepsilon,$$

where $1 \leqslant l < n$ is such that $p_l(0) \neq 0$. (That such l exists follows from the assumption in Proposition 4.1.) Now we use the set II to bound $'R_T^\varepsilon$ in the range $\theta_1(T) < \theta \leqslant \varepsilon$ and I in the range $0 \leqslant \theta \leqslant \theta_1(T)$ where $\theta_1(T)$ is to be determined. Near 0, $p_n(\theta) \approx a\theta^m$ with $a \neq 0$ (and $m < n/2$ since $K < \infty$); hence

$$\operatorname{Vol}(\mathrm{II}) \ll \int_{\theta_1(T)}^{\varepsilon} \left|\frac{T}{p_n(\theta)}\right|^{2/n} d\theta = T^{2/n}\alpha_1(\varepsilon) \tag{4.11}$$

where $\alpha_1(\varepsilon) \downarrow 0$ as $\varepsilon \to 0$. Also

$$|\partial(\mathrm{II})| \ll \int_{\theta_1(T)}^{\varepsilon} \sqrt{\left(\frac{dr}{d\theta}\right)^2 + (\sinh r)^2}\, d\theta$$

where

$$\begin{aligned} r(\theta) &= \log T^{2/n} - \log|p_n(\theta)|^{2/n} \\ &\sim \log T^{2/n} - \log \theta^{2m/n}. \end{aligned}$$

One checks that

$$|\partial(\mathrm{II})| \ll -\log\theta_1(T) + \alpha_2(\varepsilon)T^{2/n} \tag{4.12}$$

with $\alpha_2(\varepsilon) \to 0$ as $\varepsilon \to 0$.

On the other hand,

$$\mathrm{Vol}(\mathrm{I}) \ll \theta_1(T)T^{2/l}$$

$$\text{and} \tag{4.13}$$

$$|\partial(I)| \ll \theta_1(T)T^{2/l}.$$

So choosing $\theta_1(T) = 1/T$ and applying Lemma 4.2, we establish (4.10) and hence Proposition 4.1.

With these results on lattice points we turn to the proof of Theorem 1.9. Let W_n denote the space of binary forms of degree n

$$W_n = \{f(x, y) = a_0x^n + a_1x^{n-1}y + \cdots + a_ix^{n-i}y^i + \cdots + a_ny^n\}. \tag{4.14}$$

For $f \in W_n$, $\mathrm{disc}(f) \neq 0$, denote by $\mathrm{Orb}(f, \mathbf{Z}) = fSL_2(\mathbf{Z})$, $\mathrm{Orb}(f, \mathbf{R}) = fSL_2(\mathbf{R})$ the orbits of f under the action of SL_2. Let $\| \ \|$ be

$$\|f\|^2 = \sum_{i=0}^{n} a_i^2 \binom{n}{i}^{-1};$$

then

$$\|fk\| = \|f\| \qquad \text{for all } k \in SO(2). \tag{4.15}$$

Under the above representation of SL_2 the matrix $\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$ takes the form

$$\begin{bmatrix} \alpha^n & \cdots & \beta^n \\ n\alpha^{n-1}\gamma & \cdots & n\beta^{n-1}\delta \\ \vdots & & \vdots \\ \binom{n}{i}\alpha^{n-i}\gamma^i & \cdots & \\ \vdots & & \vdots \\ \gamma^n & \cdots & \delta^n \end{bmatrix}. \tag{4.16}$$

Hence if $g = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} t & \\ & t^{-1} \end{bmatrix}k$, $k \in K$, $t \geqslant 1$, then

$$f\rho(g) = (t^n p_n(\theta), t^{n-2}p_{n-2}(\theta), \ldots, t^{-n}p_{-n}(\theta)) \tag{4.17}$$

where

$$p_n(\theta) = f(\cos\theta, \sin\theta). \tag{4.18}$$

The condition $\|f\rho(g)\| \leqslant T$ becomes

$$|t^n p_n(\theta)|^2 + |t^{n-2} p_{n-2}(\theta)|^2 + \cdots + |t^{-n} p_{-n}(\theta)|^2 \leqslant T^2 \tag{4.19}$$

If we map $SL_2 \to \mathscr{H}$ by $g \to gi$, then (4.19) becomes

$$|e^{rn/2} p_n(\theta)|^2 + \cdots + |e^{-rn/2} p_{-n}(\theta)|^2 \leqslant T^2$$

where $e^{r/2} = t$ and (r, θ) are polar coordinates about i. That is, we get (4.5). Hence we see that

$$N(T, \mathrm{Orb}(f, \mathbf{Z})) = \frac{1}{|Stab_f(\mathbf{Z})|} N(SL_2(\mathbf{Z}), R_T) \tag{4.20}$$

with R_T as in (4.5). We are assuming $n \geqslant 3$ so that $Stab_f(\mathbf{Z}) = \{\gamma \in SL_2(\mathbf{Z}) | f = f\rho(\gamma)\}$ is finite.

Now R_T is compact; so by note (4.2) the nonvanishing condition is satisfied. Also we are assuming $\operatorname{disc}(f) = k \neq 0$ so that

$$K_f = \int_0^{2\pi} \frac{d\theta}{|f(\cos\theta, \sin\theta)|^{2/n}} = \int_{-\infty}^{\infty} \frac{dx}{|f(1, x)|^{2/n}} < \infty$$

since $n \geqslant 3$ and the roots of f are simple. Thus Proposition 4.1 applies, and we get

$$N(T, \mathrm{Orb}(f, \mathbf{Z})) \sim \frac{3K_f}{2\pi |Stab_f(\mathbf{Z})|} T^{2/n}. \tag{4.21}$$

This proves Theorem 1.9.

The constant K_f also appears in an asymptotic problem concerning binary forms in the work of Siegel and Mahler [Ma]. We apply (4.21) to the varieties $V \subset W_n$ which are given by specializing (to integers) the values of the invariants. The V_k's in (1.7) are examples of such. For such a V, let $f_1, \ldots, f_h$ be a representative set of the h (h = class number) $SL_2(\mathbf{Z})$-orbits in $V(\mathbf{Z})$. Here h may be zero which happens if $V(\mathbf{Z}) = \varnothing$. From (4.21) we obtain the following result.

THEOREM 4.5.

$$N(T, V) \sim \frac{3}{2\pi} \left(\sum_{j=1}^{h} \frac{K_{f_j}}{|Stab_{f_j}(\mathbf{Z})|} \right) \cdot T^{2/n}.$$

Note that K_f is $SL_2(\mathbf{R})$-invariant; so we may collect together the $SL_2(\mathbf{R})$ classes above. For example, in the case of $n = 3$, there is only one $SL_2(\mathbf{R})$ orbit with a given discriminant. Also if the discriminant

$$D = -27a^2d^2 + 18abcd + b^2c^2 - 4ac^3 - 4bd^3 < 0,$$

then $Stab_f(\mathbf{R}) = 1$ for any f of discriminant D. (The complex roots $f(1, x) = 0$ must be pointwise fixed and hence also the real root.) We conclude that for $k < 0$ and

$$V_k = \{(a, b, c, d) | D(a, b, c, d) = k\},$$

we have

$$N(T, V_k) \sim \frac{3^{3/2} K_1 h}{2\pi(-k)^{1/6}} \cdot T^{2/3} \tag{4.21}$$

where $h = h(-k)$ is the class number and

$$K_1 = \int_{-\infty}^{\infty} \frac{dx}{(1 + x^3)^{2/3}} = \frac{3}{2\pi}\sqrt[3]{16\pi} \int_0^1 \frac{dt}{\sqrt{1 - t^3}}. \tag{4.22}$$

Appendix 1
The volume function $\mu(T)$

In this appendix, we derive some properties of the measure $\mu(T)$ given in (1.7), and show that

$$b(\kappa) \leqslant \liminf \frac{\mu(\kappa T)}{\mu(T)} \leqslant \limsup \frac{\mu(\kappa T)}{\mu(T)} \leqslant a(\kappa) \tag{A1.1}$$

with $a(\kappa)$, $b(\kappa) \to 1$ as $\kappa \to 1$. We also compute $\mu(T)$ explicitly in the case of SL_m.

Structure theory. Let G be semisimple, σ an involution of G with fixed-point group H, θ a Cartan involution of G commuting with σ, and K the corresponding maximal compact subgroup of G. Let $\mathbf{g} = \mathbf{k} \oplus \mathbf{p} = \mathbf{h} \oplus \mathbf{q}$ be the decomposition of the Lie algebra of G into the ± 1-eigenspaces of θ and σ, respectively. Let $\mathbf{a}_q$ be a maximal abelian subspace of $\mathbf{p} \cap \mathbf{q}$, $\Sigma_q = \Sigma(\mathbf{a}_q, \mathbf{g})$ the root system of $\mathbf{a}_q$ in $\mathbf{g}$, Σ_q^+ a system of positive roots, and ρ the corresponding half-sum of the positive roots. Denote by $\mathbf{g}^+$ the fixed points of the involution $\theta\sigma$; it is a reductive subalgebra of $\mathbf{g}$ with $\mathbf{a}_q$ as its Cartan subspace. Let $\Sigma(\mathbf{a}_q, \mathbf{g}^+)$ be the set of restricted roots, and $\Sigma^+(\mathbf{a}_q, \mathbf{g}^+)$ a set of positive roots chosen so that it is contained in Σ_q^+. Let $\mathbf{a}_q^+$ be the positive Weyl chamber determined by the choice of $\Sigma^+(\mathbf{a}_q, \mathbf{g}^+)$. We have a "polar decomposition" $G = KA_q^+H$ and, corresponding to it, an integral formula for Haar measure on G [FJ]:

$$\int_G f(g)\,dg = \int_K \int_{\mathbf{a}_q^+} \int_H f(k \exp(Y)h)\delta(Y)\,dk\,dY\,dh \tag{A1.2}$$

where dh, dk are Haar measures on H and K, dY is Lebesgue measure on $\mathbf{a}_q$, and the Jacobian factor $\delta(Y)$ is defined as follows: For each root $\alpha \in \Sigma_q$, let $\mathbf{g}_\alpha = \mathbf{g}_\alpha^+ \oplus \mathbf{g}_\alpha^-$ be the decomposition of the corresponding root space into the ± 1-eigenspaces of $\theta\sigma$ and let $m_\pm(\alpha)$ be their dimension. Then up to a constant factor, $\delta(Y)$, $Y \in \mathbf{a}_q$, is given by

$$\delta(Y) = \prod_{\alpha \in \Sigma_q^+} \sinh^{m_+(\alpha)} \alpha(Y) \cosh^{m_-(\alpha)} \alpha(Y). \tag{A1.3}$$

The asymptotics of $\mu(T)$. We are given a linear representation of G on a vector space $\mathbf{R}^N$, a vector $v_0 \in \mathbf{R}^N$ with $Stab_G(v_0) = H$, and a K-invariant euclidean norm $\|\cdot\|$ on $\mathbf{R}^N$. For $T > 0$ we let

$$\chi_T(g) = \begin{cases} 1, & \|v_0 g\| \leqslant T \\ 0, & \text{otherwise}. \end{cases} \tag{A1.4}$$

In Lie algebra coordinates, χ_T is a characteristic function on $\mathbf{a}_q$, given by inequalities involving linear forms $\lambda_i \in \mathbf{a}_q^*$. More specifically, in the representation of G on $\mathbf{R}^N$, one can choose an orthonormal basis consisting of eigenvectors $\{v_i\}_{i=1}^N$ for A_q:

$$v_i \exp(Y) = e^{\lambda_i(Y)} v_i, \qquad Y \in \mathbf{a}_q. \tag{A1.5}$$

Thus $\chi_T(\exp Y)$ is the characteristic function of the set

$$S_T = \left\{ Y \in \mathbf{a}_q^+ : \sum_{i=1}^N e^{2\lambda_i(Y)} \leqslant T^2 \right\}. \tag{A1.6}$$

The volume function $\mu(T)$ is written as a (euclidean) integral

$$\mu(T) = \int_{S_T \subset \mathbf{a}_q^+} \delta(Y)\,dY. \tag{A1.7}$$

We may describe S_T in polar coordinates (r, ω) on $\mathbf{a}_q$:

$$S_T = \left\{ (r, \omega) : \sum_i e^{2r\lambda_i(\omega)} \leqslant T^2 \right\}. \tag{A1.8}$$

Let $f_\omega(r) = \sum e^{2\lambda_i(w)r}$. Then $f_\omega(r)$ is increasing for r sufficiently large (independent of ω), and so S_T is star-shaped for T large. For $t \gg 0$, define $r_\omega(t)$ by $f_\omega(r_\omega(t)) = e^t$.

LEMMA. *There is $c_0 > 0$ such that, for all $t \gg 0$, ω, and $\alpha > 0$,*

$$0 \leqslant r_\omega(t+\alpha) - r_\omega(t) \leqslant c_0\alpha. \tag{A1.9}$$

Proof. It suffices to give an upper bound for the derivative of r_ω or, what is the same, a lower bound for the derivative of f_ω. For each ω, let $\lambda_\infty(\omega) = \max_i \lambda_i(\omega)$ and let $\lambda^+ = \min_\omega \lambda_\infty(\omega)$. Then $\lambda^+ > 0$; otherwise, there would be infinite directions on $H\backslash G$ which keep the vector v_0 inside a compact set.

For $r \gg 1$ we have

$$\begin{aligned} f'_\omega(r) &= \sum_{\lambda_i(\omega)>0} 2\lambda_i(\omega)e^{2\lambda_i(\omega)r} - \sum_{\lambda_i(\omega)<0} 2|\lambda_i(\omega)|e^{2\lambda_i(\omega)r} \\ &\geqslant 2\lambda_\infty(\omega)e^{2\lambda_\infty(\omega)r} - c \geqslant 2\lambda^+ e^{2\lambda_\infty(\omega)r} - C \geqslant Ke^{2\lambda_\infty(\omega)r}, \end{aligned}$$

and so we find

$$r'_\omega(t) = \frac{e^t}{f'_\omega} \leqslant \sum_i e^{2\lambda_i(\omega)r_\omega(t)} \cdot \frac{1}{Ke^{2\lambda_\infty(\omega)r_\omega(t)}} \leqslant c_0. \qquad \square$$

The volume $\mu(T)$ can be expressed in polar coordinates as

$$\mu(e^t) = \int_\omega \int_0^{r_\omega(t)} \delta(r,\omega)r^{d-1}\,dr\,d\omega = \int_\omega \mu(e^t,\omega)\,d\omega$$

where $d = \dim \mathbf{a}_q$ and $\mu(e^t,\omega) = \int_0^{r_\omega(t)} \delta(r,\omega)r^{d-1}\,dr$. From (A1.3), it can be seen that

$$\delta(r,\omega)r^{d-1} \leqslant c_1 \int_0^r \delta(s,\omega)s^{d-1}\,ds. \tag{A1.10}$$

Therefore we have by (A1.9)

$$\begin{aligned} \log\frac{\mu(e^{t+\alpha},\omega)}{\mu(e^t,\omega)} &= \log \int_0^{r_\omega(t+\alpha)} \delta(r,\omega)r^{d-1}\,dr - \log \int_0^{r_\omega(t)} \delta(r,\omega)r^{d-1}\,dr \\ &\leqslant \log \int_0^{r_\omega(t)+c_0\alpha} \delta(r,\omega)r^{d-1}\,dr - \log \int_0^{r_\omega(t)} \delta(r,\omega)r^{d-1}\,dr \end{aligned}$$

By the mean value theorem, for some $r < u < r + c_0\alpha$, this equals

$$= c_0\alpha \frac{\delta(u,\omega)u^{d-1}}{\int_0^u \delta(s,\omega)s^{d-1}\,ds} \leqslant c\alpha$$

by (A1.10).

Therefore if $\kappa > 1$, $\mu(\kappa T, \omega) \leqslant \kappa^c \mu(T, \omega)$ for $c > 0$ independent of ω. Integrating over ω, we find

$$\mu(\kappa T) \leqslant \kappa^c \mu(T),$$

which proves (A1.1).

An integral on SL_m. On $GL_m^+(\mathbf{R})$ use the Haar measure

$$dg = e^{2\rho(H)}\, dk\, da\, dn$$

where KAN is as usual, $\int_K dk = 1$, $da = da_1/a_1 \cdots da_m/a_m$, $dn = \prod dn_{ij}$.

$$a = \begin{bmatrix} e^{H_1} & & \\ & \ddots & \\ & & e^{H_m} \end{bmatrix},$$

$$2\rho(H) = (m-1)H_1 + (m-3)H_2 \cdots - (m-1)H_m.$$

Let

$$F(s) = \int_{GL_m^+(\mathbf{R})} e^{-\operatorname{tr}({}^t gg)} (\det g)^s\, dg \tag{A1.13}$$

for $\operatorname{Re}(s)$ large. We first evaluate $F(s)$.

$$\int_N e^{-\operatorname{tr}({}^t n^t aan)}\, dn$$

$$= \int_N e^{-a_1^2 - a_1^2 n_2^2 \cdots - a_1^2 n_{1m}^2 - a_2^2 - a_2^2 n_{23}^2 \cdots - a_2^2 n_{2m}^2 \cdots - a_m^2}\, dn_{ij}$$

$$= e^{-a_1^2 - a_2^2 \cdots - a_m^2} a_m^0 a_{m-1}^{-1} \cdots a_1^{-(m-1)} \pi^{m(m-1)/4}.$$

Hence

$$F(s) = \int_0^\infty \cdots \int_0^\infty \pi^{m(m-1)/4} e^{-(a_1^2 + \cdots + a_m^2)} a_1^{-(m-1)} \cdots a_m^0 a_1^{m-1} \cdots a_m^{-(m-1)}$$

$$\cdot (a_1 a_2 \cdots a_m)^s \frac{da_1}{a_1} \cdots \frac{da_m}{a_m}$$

$$= \frac{\pi^{m(m-1)/4}}{2^m} \prod_{j=1}^m \Gamma\left(\frac{s+1-j}{2}\right).$$

Now let $dg = dg_0(dt/t)$ where $t = \det g$ and dg_0 is the corresponding Haar measure on $SL_m(\mathbf{R})$. Let $H(\lambda)$ be defined by

$$H(\lambda) = \int_{SL_m(\mathbf{R})} e^{-\lambda \operatorname{tr}({}^t g_0 g_0)}\, dg_0 .$$

The asymptotics of $H(\lambda)$ as $\lambda \downarrow 0$ will give us the volume asymptotics.

Now setting $g = t^{1/m} g_0$,

$$\begin{aligned}\int_{GL_m^+(\mathbf{R})} e^{-\operatorname{tr}({}^t g g)} (\det g)^s\, dg &= \int_0^\infty \left(\int_{SL_m(\mathbf{R})} e^{-t^{2/m}\operatorname{tr}({}^t g_0 g_0)}\, dg_0 \right) t^s \frac{dt}{t} \\ &= \int_0^\infty H(t^{2/m}) t^s \frac{dt}{t} = F(s).\end{aligned}$$

Thus if $f(t) = H(t^{2/m})$, then for ξ large

$$f(t) = \frac{1}{2\pi i} \int_{\operatorname{Re}(s)=\xi} \frac{\pi^{m(m-1)/4}}{2^m} \prod_{j=1}^{m} \Gamma\left(\frac{s+1-j}{2}\right) t^{-s}\, ds .$$

Shifting the contour to the left, the first pole occurs when $s = m - 1$. Hence

$$f(t) \sim 2 \frac{\pi^{m(m-1)/4}}{2^m} \prod_{j=1}^{m-1} \Gamma\left(\frac{m-j}{2}\right) t^{-(m-1)}$$

as $t \to 0$. Hence

$$H(t) \sim \frac{\pi^{m(m-1)/4}}{2^{m-1}} \cdot \prod_{j=1}^{m-1} \Gamma\left(\frac{m-j}{2}\right) \cdot t^{-m(m-1)/2} . \tag{A1.14}$$

Setting

$$\psi(x) = \int_{\operatorname{tr}({}^t g_0 g_0) \leqslant x} dg_0 ,$$

we have

$$H(\lambda) = \int_0^\infty e^{-\lambda t}\, d\psi(t).$$

Thus by a standard Tauberian argument [W, p192], (A1.14) implies

$$\psi(x) \sim \frac{\pi^{m(m-1)/4}}{2^{m-1}\Gamma\left(\frac{m^2-m+2}{2}\right)} \prod_{j=1}^{m-1} \Gamma\left(\frac{m-j}{2}\right) x^{m(m-1)/2} \qquad \text{as } x \to \infty . \tag{A1.15}$$

The passage to the asymptotics for $d\tilde{g}_0$ which has $\mathrm{vol}(G(\mathbf{Z})\backslash G(\mathbf{R})) = 1$ is straightforward. First, one deduces the analogue of (A1.15) for the measure (on $GL_m^+(\mathbf{R})$)

$$d\tilde{g} = \frac{\prod_{i,j} dg_{ij}}{|\det g|^m}$$

which differs from dg by a factor of

$$\frac{2^{m-1}\pi^{m^2/2-m(m-1)/4}}{\prod_{j=0}^{m-1}\Gamma\left(\frac{m-j}{2}\right)}.$$

If $d\tilde{g}_0$ is the corresponding measure on $SL_m(\mathbf{R})$, then according to Minkowski (see [Si]),

$$\mathrm{Vol}_{d\tilde{g}_0}(SL_m(\mathbf{Z})\backslash SL_m(\mathbf{R})) = \zeta(2)\zeta(3)\cdots\zeta(m).$$

(1.12) then follows.

APPENDIX 2

Regularizing Eisenstein periods on $SL_2(\mathbf{R})\backslash SL_2(\mathbf{C})$

Let $G = SL_2(\mathbf{C})$, $H = SL_2(\mathbf{R})$, $\Gamma = SL_2(\mathbf{Z}[i])$ the Picard group, and $\Gamma_H = H \cap \Gamma = SL_2(\mathbf{Z})$. Let

$$N = \left\{\begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix} : z \in \mathbf{C}\right\}, \qquad A = \left\{\begin{bmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{bmatrix} : y > 0\right\},$$

$$M = \begin{bmatrix} e^{i\theta} & \\ & e^{-i\theta} \end{bmatrix}, \qquad P = MAN, \qquad K = SU(2).$$

Any $g \in G$ has Iwasawa decomposition $g = n\begin{bmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{bmatrix}k$, $n \in N$, $k \in K$ and $y = y(g) > 0$. In these coordinates, Haar measure on G is given by

$$dg = dn\frac{dy}{y^3}\,dk.$$

Likewise, Haar measure on H is given by $dh = dn\, y^{-2}\, dy\, dk'$.

Remark. This situation gives an example of F_T which is not in L^2:

$$F_T\left(\begin{pmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{pmatrix}\right) \geqslant \sum_{n\in\mathbf{Z}} \chi_T\left(\begin{pmatrix} 1 & in \\ 0 & 1 \end{pmatrix}\begin{pmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{pmatrix}\right)$$

$$= \sum_{n\in\mathbf{Z}} \chi_T\left(\begin{pmatrix} 1 & in/y \\ 0 & 1 \end{pmatrix}\right) \gg y \qquad \text{as } y \to \infty.$$

Since the measure on $\Gamma\backslash G$ is $y^{-3}\, dz\, dy\, dk$, this estimate shows that $F_T \notin L^2$.

Define an Eisenstein series on G by

$$E(g, \lambda) = \sum_{\gamma\in\Gamma\cap P\backslash\Gamma} y(\gamma g)^\lambda$$

which is absolutely convergent for $\operatorname{Re}\lambda > 2$ and has meromorphic continuation with simple pole at $\lambda = 2$, with residue

$$\text{(A2.1)} \qquad \operatorname*{Res}_{\lambda=2} E(g, \lambda) = \frac{\operatorname{Vol}(\Gamma\cap P\backslash N)}{\operatorname{Vol}(\Gamma\backslash G)}.$$

The constant term of $E(g, \lambda)$ along N is given by

$$E^P(g, \lambda) = \int_{\Gamma\cap N\backslash N} E(ng, \lambda)\, dn = y(g)^\lambda + \phi(\lambda) y(g)^{2-\lambda},$$

with

$$\phi(\lambda) = \frac{\xi(\lambda - 1)}{\xi(\lambda)},$$

$\xi(s)$ being the Dedekind zeta function of $\mathbf{Q}(\sqrt{-1})$ (with archimedean factor).

Denote by EDV the space of entire functions $\hat{f}(\lambda)$, rapidly decreasing in vertical strips. For $\hat{f} \in$ EDV, define $f \in C_c^\infty(N\backslash G)$ by

$$f(g) = \frac{1}{2\pi i}\int_{\operatorname{Re}\lambda=\lambda_0} \hat{f}(\lambda) y(g)^\lambda\, d\lambda.$$

Now for $\operatorname{Re}\lambda = \lambda_0 > 2$, let

$$E_f(g) = \sum_{\gamma\in\Gamma\cap P\backslash\Gamma} f(\gamma g) = \frac{1}{2\pi i}\int_{\operatorname{Re}\lambda=\lambda_0} \hat{f}(\lambda) E(g, \lambda)\, d\lambda$$

and define

$$E_f^H(g) = \int_{\Gamma \cap H \backslash H} E_f(hg)\, dh.$$

Mellin inversion shows that

$$\int_{\Gamma \backslash G} E_f(g)\, dg = \mathrm{Vol}(\Gamma \cap P \backslash N) \hat{f}(2). \tag{A2.2}$$

As a special case of Theorem 1.11, we will show the following theorem.

THEOREM. *If $\varepsilon > 0$ is sufficiently small, then*

$$E_f^H(g) = \frac{\mathrm{Vol}(\Gamma_H \backslash H)}{\mathrm{Vol}(\Gamma \backslash G)} \cdot \int_{\Gamma \backslash G} E_f(g)\, dg$$
$$+ \frac{1}{2\pi i} \int_{\mathrm{Re}\, \lambda = 2 - \varepsilon} \hat{f}(\lambda) E^{G,H}(g, \lambda)\, d\lambda - \hat{f}(1) \int_{K \cap H} y(kg)\, dk,$$

where $E^{G,H}(g, \lambda)$ is an H-invariant eigenfunction with central character λ, meromorphic in λ.

We use the standard fundamental domain for $SL(2, \mathbf{Z})$:

$$\mathscr{F} = \left\{ \begin{bmatrix} 1 & x \\ & 1 \end{bmatrix} \begin{bmatrix} y^{1/2} & \\ & y^{-1/2} \end{bmatrix} k \colon -1/2 < x \leqslant 1/2,\, x^2 + y^2 \geqslant 1,\, k \in SO(2) \right\}$$

Let $T > 1$ and decompose $\mathscr{F}$ as $\mathscr{F} = \mathscr{F}_H(T) \cup \mathscr{F}_P(T)$ where

$$\mathscr{F}_H(T) = \{ h \in \mathscr{F} \colon y(h) \leqslant T \}$$

which is compact and

$$\mathscr{F}_P(T) = \{ h \in \mathscr{F} \colon y(h) > T \}.$$

Then

$$E_f^H(g) = \int_{\mathscr{F}_H(T)} E_f(hg)\, dh + \int_{\mathscr{F}_P(T)} E_f(hg)\, dh.$$

In the compact part, we interchange order of integration and write (for $\mathrm{Re}\, \lambda_0 > 2$)

$$\int_{\mathscr{F}_H(T)} E_f(hg)\, dh = \frac{1}{2\pi i} \int_{\mathrm{Re}\, \lambda = \lambda_0} \hat{f}(\lambda) \left\{ \int_{\mathscr{F}_H(T)} E(hg, \lambda)\, dh \right\} d\lambda,$$

which is meromorphic in λ, with a simple pole at $\lambda = 2$ with residue

$$\operatorname*{Res}_{\lambda=2} \int_{\mathscr{F}_H(T)} E_f(hg)\,dh = \frac{\operatorname{Vol}(\Gamma \cap P\backslash N)}{\operatorname{Vol}(\Gamma\backslash G)} \operatorname{Vol}(\mathscr{F}_H(T)).$$

Since for $\operatorname{Re}\lambda > 2$ we have $E(g, \lambda) \sim y(g)^\lambda$ as $y(g) \to \infty$, in the integral $\int_{\mathscr{F}_P(T)} E_f(hg)\,dh$ we *cannot* interchange orders of integration as in the integral over $\mathscr{F}_H(T)$.

Note. For $h \in \mathscr{F}_P(T)$ and g in a Siegel set $\mathscr{S}_P$ relative to P, hg lies in $N\mathscr{S}'_P$.

CONCLUSION. *$E(hg, \lambda) - E^P(hg, \lambda)$ is rapidly decreasing in h as h varies in $\mathscr{F}_P(T)$.*

This is because $E(g', \lambda) - E^P(g', \lambda)$ is rapidly decreasing as $g' \to \infty$ in $N\mathscr{S}_P$.
Now write

$$\text{(A2.3)} \quad E_f(hg) = \frac{1}{2\pi i} \int_{\operatorname{Re}\lambda=\lambda_0} \hat{f}(\lambda)(E - E^P)(hg)\,d\lambda + \frac{1}{2\pi i} \int_{\operatorname{Re}\lambda=\lambda_0} \hat{f}(\lambda)E^P(hg, \lambda)\,d\lambda.$$

Then the integrand is rapidly decreasing for $h \in \mathscr{F}_P(T)$; so the integral over $\mathscr{F}_P(T)$ of the first integrand in (A2.3) is absolutely convergent, and we may write

$$\text{(A2.4)} \quad \int_{\mathscr{F}_P(T)} E_f(hg)\,dh = \frac{1}{2\pi i} \int_{\operatorname{Re}\lambda=\lambda_0} \hat{f}(\lambda)\left\{\int_{\mathscr{F}_P(T)} (E - E^P)(hg, \lambda)\,dh\right\} d\lambda$$
$$+ \int_{\mathscr{F}_P(T)} \left\{\frac{1}{2\pi i} \int_{\operatorname{Re}\lambda=\lambda_0} \hat{f}(\lambda)E^P(hg, \lambda)\,d\lambda\right\} dh.$$

Since $E^P(g, \lambda) = y(g)^\lambda + \phi(\lambda)y(g)^{2-\lambda}$, we have

$$\text{(A2.5)} \quad \int_{\operatorname{Re}\lambda=\lambda_0} \hat{f}(\lambda)E^P(hg, \lambda)\,d\lambda$$
$$= \int_{\operatorname{Re}\lambda=\lambda_0} \hat{f}(\lambda)\{y(h)^\lambda y(\kappa(h)g)^\lambda + \phi(\lambda)y(h)^{2-\lambda}y(\kappa(h)g)^{2-\lambda}\}\,d\lambda.$$

For $\operatorname{Re}\lambda = \lambda_0 > 2$, $y(hg)^{2-\lambda} = y(h)^{2-\lambda} \cdot y(\kappa(h)g)^{2-\lambda}$ is decreasing in $\mathscr{F}_P(T)$ and is integrable over $\mathscr{F}_P(T)$. Note that $\phi(\lambda)$ is bounded in $\operatorname{Re}\lambda > 2$.

We separate out the two terms in (A2.5); the second one equals

$$\text{(A2.6)} \quad \frac{1}{2\pi i} \int_{\operatorname{Re}\lambda=\lambda_0} \hat{f}(\lambda)\left(\phi(\lambda) \int_{\mathscr{F}_P(T)} y(hg)^{2-\lambda}\,dh\right) d\lambda.$$

Now we deal with the contribution of the first term of the integrand in (A2.5):

$$\int_{\mathcal{F}_P(T)} \left\{ \frac{1}{2\pi i} \int_{\operatorname{Re} \lambda = \lambda_0} \hat{f}(\lambda) y(hg)^\lambda \, d\lambda \right\} dh. \tag{A2.7}$$

We first shift the contour in this integral from $\operatorname{Re} \lambda = \lambda_0 > 2$ to $\operatorname{Re} \lambda = \lambda_1$, with $\lambda_1 < 1$. This can be done since $\hat{f}(\lambda) \in \mathrm{EDV}$. Having done this, the integral in (A2.7) is absolutely convergent and after interchanging order of integration equals

$$\begin{aligned} &\frac{1}{2\pi i} \int_{\operatorname{Re} \lambda = \lambda_1 < 1} \hat{f}(\lambda) \int_{\mathcal{F}_P(T)} y(hg)^\lambda \, dh \, d\lambda \\ &= \frac{1}{2\pi i} \int_{\operatorname{Re} \lambda = \lambda_1 < 1} \hat{f}(\lambda) \int_{|x| < 1/2} \int_T^\infty \int_{K \cap H} y^\lambda y(kg)^\lambda \, dx \frac{dy}{y^2} \, dk \, d\lambda \\ &= \frac{1}{2\pi i} \int_{\operatorname{Re} \lambda = \lambda_1} \hat{f}(\lambda) \frac{T^{\lambda - 1}}{1 - \lambda} \int_{K \cap H} y(kg)^\lambda \, dk \, d\lambda. \end{aligned} \tag{A2.8}$$

We now shift the contour in (A2.8) back to $\operatorname{Re} \lambda = \lambda_0 > 2$ and pick up a residue at $\lambda = 1$ to get

$$\frac{1}{2\pi i} \int_{\operatorname{Re} \lambda = \lambda_0} \hat{f}(\lambda) \frac{T^{\lambda - 1}}{\lambda - 1} \int_{K \cap H} y(kg)^\lambda \, dk \, d\lambda - \hat{f}(1) \int_{K \cap H} y(kg) \, dk. \tag{A2.9}$$

Combining (A2.3), (A2.4), (A2.6), and (A2.9), we find

$$E_f^H(g) = \frac{1}{2\pi i} \int_{\operatorname{Re} \lambda = \lambda_0 > 2} \hat{f}(\lambda) E^{G,H}(g, \lambda) \, d\lambda - \hat{f}(1) \int_{K \cap H} y(kg) \, dk \tag{A2.10}$$

where

$$\begin{aligned} E^{G,H}(g, \lambda) = &\int_{\mathcal{F}_H(T)} E(hg, \lambda) \, dh + \int_{\mathcal{F}_P(T)} (E - E^P)(hg, \lambda) \, dh \\ &+ \phi(\lambda) \int_{\mathcal{F}_P(T)} y(hg)^{2-\lambda} \, dh + \frac{T^{\lambda - 1}}{1 - \lambda} \int_{K \cap H} y(kg)^\lambda \, dk. \end{aligned} \tag{A2.11}$$

From this formula, we see that $E^{G,H}(g, \lambda)$ is an eigenfunction with infinitesimal character λ, is meromorphic in λ, and holomorphic for $\operatorname{Re} \lambda > 2 - \varepsilon$ for some $\varepsilon > 0$, except for a simple pole at $\lambda = 2$.

CLAIM. *$E^{G,H}(g, \lambda)$ is H-invariant.*

Proof. Indeed, from (A2.10) we have

$$\frac{1}{2\pi i}\int_{\mathrm{Re}(\lambda)=\lambda_0>2} \hat{f}(\lambda)E^{G,H}(g,\lambda)\,d\lambda = E_f^H(g) + \hat{f}(1)\int_{K\cap H} y(kg)\,dk.$$

$E_f^H(g)$ is H-invariant and one easily checks the following statement.

LEMMA. *$\int_{K\cap H} y(kg)\,dk$ is H-invariant.*

Thus $\int_{\mathrm{Re}\,\lambda=\lambda_0}\hat{f}(\lambda)E^{G,H}(g,\lambda)\,d\lambda$ is H-invariant for all $\hat{f}\in$ EDV. Since $E^{G,H}(g,\lambda)$ is holomorphic in $\mathrm{Re}\,\lambda > 2$, of moderate growth in vertical strips, this forces $E^{G,H}(g,\lambda)$ to be H-invariant.

We now shift the contour of integration in (A2.10) from $\mathrm{Re}\,\lambda = \lambda_0 > 2$ to the *left* of $\mathrm{Re}\,\lambda = 2$. (Just a slight shift will suffice for our purposes.) To do so, we need to know that $E^{G,H}(g,\lambda)$ is at most of polynomial growth in $\mathrm{Re}\,\lambda > 2-\varepsilon$. This follows from (A2.11) modulo knowing this for $\phi(\lambda)$ and $E(g,\lambda)$. From (A2.11) we see $E^{G,H}(g,\lambda)$ is holomorphic in $\mathrm{Re}\,\lambda > 2-\varepsilon$ except for a pole at $\lambda = 2$, since the same holds for $\phi(\lambda)$ and $E(g,\lambda)$. Also from (A2.11) we see

$$\text{(A2.12)}\quad \operatorname*{Res}_{\lambda=2} E^{G,H}(g,\lambda)$$

$$= \frac{\mathrm{Vol}(\Gamma\cap P\backslash N)}{\mathrm{Vol}(\Gamma\backslash G)}(\mathrm{Vol}\,\mathscr{F}_H(T) + \mathrm{Vol}\,\mathscr{F}_P(T)) - \operatorname*{Res}_{\lambda=2}\phi(\lambda) + \operatorname*{Res}_{\lambda=2}\phi(\lambda)$$

$$= \frac{\mathrm{Vol}(\Gamma\cap P\backslash N)}{\mathrm{Vol}(\Gamma\backslash G)}\mathrm{Vol}(\Gamma_H\backslash H).$$

Therefore

$$E_f^H(g) = \frac{\mathrm{Vol}(\Gamma\cap P\backslash N)}{\mathrm{Vol}(\Gamma\backslash G)}\mathrm{Vol}(\Gamma_H\backslash H)\hat{f}(2)$$

$$+\frac{1}{2\pi i}\int_{\mathrm{Re}\,\lambda=2-\varepsilon}\hat{f}(\lambda)E^{G,H}(g,\lambda)\,d\lambda - \hat{f}(1)\int_{K\cap H} y(kg)\,dk.$$

Upon using (A2.2), this becomes

$$\text{(A2.13)}\quad E_f^H(g) = \frac{\mathrm{Vol}(\Gamma_H\backslash H)}{\mathrm{Vol}(\Gamma\backslash G)}\int_{\Gamma\backslash G} E_f(g)\,dg$$

$$+\frac{1}{2\pi i}\int_{\mathrm{Re}\,\lambda=2-\varepsilon}\hat{f}(\lambda)E^{G,H}(g,\lambda)\,d\lambda - \hat{f}(1)\int_{K\cap H} y(kg)\,dk.$$

Remark. (A2.13) shows that

$$E_f^H(g) = \frac{\operatorname{Vol}(\Gamma_H\backslash H)}{\operatorname{Vol}(\Gamma\backslash G)} \int_{\Gamma\backslash G} E_f(g)\, dg + \text{terms decaying on } H\backslash G/K$$

since in (A2.13), both $E^{G,H}(g, \lambda)$ and $\int_{K\cap H} y(kg)\, dk$ are eigenfunctions with infinitesimal character having real part in the "convex hull" $0 < \operatorname{Re} \lambda < 2$ and so decay on $H\backslash G/K$ [RS].

REFERENCES

[B] E. P. VAN DEN BAN, *Asymptotic behaviour of matrix coefficients related to reductive symmetric spaces*, Nederl. Akad. Wet. Proc. Ser. A Math. Sci. **90** (1987), 225–249.

[Ba] H. J. BARTELS, *Nichteuklidische Gitterpunktprobleme und Gleichverteilung in linearen algebraischen Gruppen*, Comm. Math. Helv. **57** (1982), 158–172.

[Bi] B. J. BIRCH, *Forms in many variables*, Proc. Roy. Soc. Edinburgh Sect. A **265** (1961/62), 245–263.

[B-HC] A. BOREL AND HARISH-CHANDRA, *Arithmetic subgroups of algebraic groups*, Ann. of Math. **75** (1962), 485–535.

[BW] A. BOREL AND N. WALLACH, *Continuous Cohomology, Discrete Subgroups, and Representations of Reductive Groups*, Ann. of Math. Stud. **94**, Princeton Univ. Press, Princeton, 1980.

[BLS] M. BURGER, J.-S. LI, AND P. SARNAK, *Ramanujan duals and automorphic spectrum I*, Bull. of Amer. Math. Soc. **26** (1992), 253–257.

[D] J. DELSARTE, *Sur le gitter Fuchsien*, C. R. Acad. Sci. Paris **214** (1942), 147–149.

[ERS] A. ESKIN Z. RUDNICK, AND P. SARNAK, *A proof of Siegel's weight formula*, Internat. Math. Res. Notices **1991**, 65–69.

[FJ] M. FLENSTED-JENSEN, *Analysis on Non-Riemannian Symmetric Spaces*, Reg. Conf. Ser. Math. **61**, Amer. Math. Soc., Providence, 1986.

[FMT] J. FRANKE, YU. MANIN, AND YU. TSCHINKEL, *Rational points of bounded height on Fano varieties*, Invent. Math. **95** (1989), 421–435.

[HC] HARISH-CHANDRA, *Automorphic Forms on Semisimple Lie Groups*, Lecture Notes in Math. **62**, Springer-Verlag, Berlin, 1968.

[HM] R. HOWE AND C. C. MOORE, *Asymptotic properties of unitary representations*, J. Funct. Anal. **32** (1979), 72–96.

[L] R. P. LANGLANDS, *On the Functional Equations Satisfied by Eisenstein Series*, Lecture Notes in Math. **544**, Springer-Verlag, Berlin, 1976.

[LP] P. LAX AND R. PHILLIPS, *The asymptotic distribution of lattice points in Euclidean and non-Euclidean spaces*, J. Funct. Anal. **46** (1982), 280–350.

[Ma] K. MAHLER, *Zur Approximation algebraischer Zahlen III*, Acta Math. **62** (1933), 91–166.

[R] B. RANDOL, *Small eigenvalues of the Laplace operator on compact Riemann surfaces*, Bull. Amer. Math. Soc. **80** (1974), 996–1000.

[Ru] Z. RUDNICK, *Uniform decay rates for eigenfunctions on affine symmetric spaces*, preprint, 1992.

[RS] Z. RUDNICK AND H. SCHLICHTKRULL, *Decay of eigenfunctions on semisimple symmetric spaces*, Duke Math. J. **64** (1991), 445–450.

[Sa] P. SARNAK, "Diophantine problems and linear groups" in *Proceedings of the International Congress of Mathematicians, Kyoto, 1990, Volume 1*, Springer-Verlag, Tokyo, 1991, 459–471.

[Sca] R. SCARAMUZZI, *A notion of rank for unitary representations of general linear groups*, Trans. Amer. Math. Soc. **319** (1990), 349–379.

[Sch] W. SCHMIDT, *The density of integer points on homogeneous varieties*, Acta Math. **154** (1985), 243–296.

[Sco] E. J. SCOURFIELD, *The divisors of a quadratic polynomial*, Proc. Glasglow Math. Assoc. **5** (1961), 8–20.

[Se] A. SELBERG, *Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series*, J. Indian Math. Soc. **20** (1956), 47–87.

[Si] C. L. SIEGEL, *Lectures on the Geometry of Numbers*, Springer-Verlag, Berlin, 1989.

[V] D. VOGAN, *The unitary dual of $GL(n)$ over an archimedian field*, Invent. Math. **83** (1986), 449–505.

[W] D. V. WIDDER, *The Laplace Transform*, Princeton Math. Ser. **6**, Princeton Univ. Press, Princeton, 1946.

DUKE: DEPARTMENT OF MATHEMATICS, RUTGERS UNIVERSITY, NEW BRUNSWICK, NEW JERSEY 08903, USA

RUDNICK: DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, STANFORD, CALIFORNIA 94305-2125, USA; CURRENT: DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON, NEW JERSEY 08544, USA

SARNAK: DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, STANFORD, CALIFORNIA 94305-2125, USA; CURRENT: DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON, NEW JERSEY 08544, USA