COMPOSITIO MATHEMATICA

Lower bounds for Maass forms on semisimple groups

Farrell Brumley and Simon Marshall

Lower bounds for Maass forms on semisimple groups

Farrell Brumley and Simon Marshall

Abstract

Let G be an anisotropic semisimple group over a totally real number field F. Suppose that G is compact at all but one infinite place v_0. In addition, suppose that G_{v_0} is $\mathbb{R}$-almost simple, not split, and has a Cartan involution defined over F. If Y is a congruence arithmetic manifold of non-positive curvature associated with G, we prove that there exists a sequence of Laplace eigenfunctions on Y whose sup norms grow like a power of the eigenvalue.

1. Introduction

Let Y be a closed Riemannian manifold of dimension n and with Laplace operator Δ. Let $\{\psi_i\}$ be an orthonormal basis of Laplace eigenfunctions for $L^2(Y)$, which satisfy $\|\psi_i\|_2 = 1$ and $(\Delta + \lambda_i^2)\psi_i = 0$. We assume that $\{\psi_i\}$ are ordered by eigenvalue, so that $0 = \lambda_1 \leqslant \lambda_2 \leqslant \cdots$. It is an important question in harmonic analysis to determine the asymptotic size of ψ_i, that is, the growth rate of $\|\psi_i\|_\infty$ in terms of λ_i. The basic upper bound for $\|\psi_i\|_\infty$, proved by Avacumović [Ava56] and Levitan [Lev52], is given by

$$\|\psi_i\|_\infty \ll \lambda_i^{(n-1)/2}. \tag{1}$$

This bound is sharp on the round n-sphere. Indeed, the zonal spherical harmonics have peaks of maximal size at the poles of the axis of rotation. More generally, Sogge and Zelditch [SZ02] have shown that the compact Riemannian manifolds saturating (1) necessarily have points which are fixed by an appropriately large number of geodesic returns, in the sense that a positive measure subset of geodesics passing through such a point are loops.

On the other hand, if Y is negatively curved then its geodesic flow is highly chaotic, and one expects this to be reflected in the asymptotics of the eigenfunctions. For example, the quantum ergodicity theorem of Schnirelman [Sch74], Colin de Verdière [Col85] and Zelditch [Zel87] states that the L^2-mass of a density-one sequence of Laplacian eigenfunctions on a negatively curved manifold equidistributes to the uniform measure. It is likewise reasonable to expect that if Y has negative curvature then the strong pointwise delocalization bound

$$\|\psi_i\|_\infty \ll_\epsilon \lambda_i^\epsilon \tag{2}$$

holds with density one. This is akin to the Ramanujan conjecture in the theory of automorphic forms [Sar95]: a generic sequence of eigenfunctions is tempered. Any sequence violating (2) will be called *exceptional*.

F. Brumley and S. Marshall

Unlike the weak-* setting of the quantum unique ergodicity conjecture [RS94], which posits that in negative curvature the L^2-mass of *any* sequence of eigenfunctions equidistributes to the uniform measure, there do in fact exist compact manifolds Y of negative curvature violating (2). The first examples were given in [RS94], with Y being an arithmetic hyperbolic 3-manifold and the exceptional sequences, of density zero, lying in the image of the theta correspondence from Sp_2.

For surfaces the situation is quite different: Iwaniec and Sarnak [IS95] conjecture that hyperbolic surfaces do not support exceptional sequences. This is a very difficult problem which would imply the classical Lindelöf conjecture on the Riemann zeta function in the case of the modular surface.

This leads to the question of when exactly one should expect to find (zero-density) violations of (2). In this paper we give sufficient conditions for a certain class of negatively curved manifolds to support exceptional sequences. Although the question is of interest in this general setting, our techniques are limited to arithmetic locally symmetric spaces. Put succinctly, we show that an arithmetic manifold supports exceptional sequences whenever it has a point with strong Hecke return properties.

1.1 Statement of results

Our main theorem is modelled on a result of Milićević [Mil11], which, building on [RS94], provides a structural framework for the class of arithmetic hyperbolic 3-manifolds supporting exceptional sequences.

First recall that an arithmetic hyperbolic 3-manifold arises from the following general construction. Let E be a number field having exactly one complex embedding, up to equivalence, and let F be its maximal totally real subfield. For a division quaternion algebra B over E, ramified at all real places of E, denote by G the restriction of scalars of B^1 from E to F. Then any arithmetic hyperbolic 3-manifold is commensurable with a congruence manifold associated with G.

Following [Mil11], an arithmetic hyperbolic 3-manifold as above is said to be of Maclachlan–Reid type if E is quadratic over F and there exists a quaternion division algebra A over F satisfying $B = A \otimes_F E$. The main result of [Mil11] is that Maclachlan–Reid type manifolds support exceptional sequences (in fact, satisfying the same lower bounds as the examples of Rudnick and Sarnak).

Notice that when $B = A \otimes_F E$, the following properties hold. Let v_0 be the unique archimedean place of F which ramifies in E. By [MR03, Theorem 9.5.5] we may assume that A is ramified at v_0. Then:

(i) $G_{v_0} = \mathrm{SL}_2(\mathbb{C})$ is non-compact and non-split (as an $\mathbb{R}$-group);

(ii) $G_v = \mathbf{H}^1$ (the norm-one Hamiltonian quaternions) is compact for all real $v \neq v_0$;

(iii) the global involution $\theta : g \mapsto \sigma(g)$ of G, where σ is the unique non-trivial element in the Galois group of E over F, induces a Cartan involution on G_{v_0}. Indeed, $G^\theta = A^1$, so that $G^\theta(F_{v_0}) = \mathbf{H}^1$ is the maximal compact SU(2) inside $G_{v_0} = \mathrm{SL}_2(\mathbb{C})$.

Our main result is an extension of this to a wide range of compact congruence manifolds.

THEOREM 1.1. *Let F be a totally real number field, and let v_0 be a real place of F. Let G/F be a connected anisotropic semisimple F-group. We make the following additional assumptions on G:*

(i) G_{v_0} *is non-compact, non-split and $\mathbb{R}$-almost simple;*

960

(ii) G_v is compact for all real $v \neq v_0$;

(iii) there is an involution θ of G defined over F that induces a Cartan involution of G_{v_0}.

Let Y be a congruence manifold associated with G as in §2.10. Then there exist $\delta > 0$, depending only on G, and a sequence of linearly independent Laplacian eigenfunctions ψ_i on Y that satisfy

$$\|\psi_i\|_2 = 1, \quad (\Delta + \lambda_i^2)\psi_i = 0, \quad \text{and} \quad \|\psi_i\|_\infty \gg \lambda_i^\delta.$$

We note that the exponent δ in Theorem 1.1, and Theorem 1.2 below, is currently logically ineffective. This comes from using a result of Cluckers, Gordon and Halupczok [ST16, Theorem 14.1] to bound orbital integrals. (Note that Shin and Templier present an alternative, effective proof of [ST16, Theorem 14.1] in Theorem 7.3 of the same paper, but this has an error as described on S.-W. Shin's webpage.) The bound of Cluckers, Gordon and Halupczok is ineffective due to their use of model theory. Julia Gordon informs us that, if the residue characteristic is larger than a logically ineffective bound, it may be possible to transfer the constants in their result from the case of positive characteristic and explicate them in that case. However, this is presently unknown. Moreover, even if this were possible, the implied constant in Theorems 1.1 and 1.2 would still be logically ineffective because of the condition that the residue characteristic be sufficiently large.

An additional source of ineffectiveness of δ comes from using theorems of Finis and Lapid [FL18] when the congruence manifold varies as in Theorem 1.2, although this ineffectiveness is in the weaker sense that the constants are not numerically explicit. Finis and Lapid inform us that their results could probably be made effective with extra work; in fact, as the referee has pointed out, the particular shape of the congruence subgroups, which depend on θ, in Theorem 1.2 should allow for the exponents arising from their work to be more readily explicated.

1.2 Remarks on the main theorem

A well-known theorem of Borel [Bor63] addresses the question of whether one can find many groups satisfying the rationality hypothesis (iii). One consequence of his theorem is that for any connected, simply connected, semisimple algebraic $\mathbb{R}$-group G satisfying condition (i), Theorem 1.1 produces a manifold Y of the form $\Gamma \backslash G/K$ with an exceptional sequence of eigenfunctions. See §1.4 for more details and a concrete example.

Theorem 1.1 goes some distance towards answering the basic question of determining the precise conditions under which one should expect a Lindelöf type bound on a compact congruence negatively curved manifold. The three numbered conditions on the group G are a particularly convenient way of asking that a large enough compact subgroup of G_∞ admits a rational structure, which is a key ingredient in our proof. Although the condition that G_{v_0} is not split should be necessary, we expect that the other conditions can be relaxed somewhat. For example, throughout most of the paper, the condition that G_{v_0} is $\mathbb{R}$-almost simple could be weakened to G being F-almost simple. The stronger form of this condition is only used in Lemma 4.2, to simplify the application of a theorem of Blomer and Pohl [BP16, Theorem 2] and Matz and Templier [MT15, Proposition 7.2].

Besides the results of Rudnick and Sarnak and of Milićević that we have already mentioned, both in the context of arithmetic hyperbolic 3-manifolds, there are other results in the literature which provide examples of arithmetic manifolds supporting exceptional sequences. For instance, the techniques of Rudnick and Sarnak were generalized to n-dimensional hyperbolic manifolds for $n \geqslant 5$ by Donnelly [Don07]. Later, Lapid and Offen in [LO07] discovered a series of arithmetic quotients of $\mathrm{SL}(n, \mathbb{C})/\mathrm{SU}(n)$ admitting large eigenforms through the link with automorphic

L-functions (conditionally on standard conjectures on the size of automorphic L-functions at the edge of the critical strip). To our knowledge [LO07] is the only example of power growth lower bounds for sup norms on non-compact quotients, at least if one restricts to a compact set so that the large peaks of eigenfunctions in the cusp do not play a role. Note that Theorem 1.1 includes the examples of Rudnick and Sarnak, of Donnelly, and of Milićević, although without explicit exponents. It is unable to reproduce the examples of Lapid and Offen due to the compactness requirement, but (as was indicated above) it can produce compact quotients of $\mathrm{SL}(n, \mathbb{C})/\mathrm{SU}(n)$ with an exceptional sequence of eigenfunctions. In fact, non-compact quotients (of suitable level structure) should also be amenable to our techniques, via an application of simple trace formulae, but we have not pursued this here.

Finally, while our approach was largely inspired by that of Milićević, we have made an effort to emphasize (in Appendix A) the common features it shares with the techniques of Rudnick and Sarnak and of Lapid and Offen. A synthesis of the subject, as well as a general conjecture restricting the possible limiting exponents for exceptional sequences, can be found in Sarnak's influential letter [Sar04].

1.3 A hybrid result in the level and eigenvalue aspects

We in fact prove a stronger result than that described in Theorem 1.1, establishing a lower bound in the level and eigenvalue aspects simultaneously. We present this separately, as it requires more care to state; indeed, any notion of non-trivial lower bound must overcome the lower bound one may prove when the eigenspaces have large dimension. More precisely, if M is a compact Riemannian manifold and V is the space of $\psi \in L^2(M)$ with a given Laplace eigenvalue, one may show that there is $\psi \in V$ satisfying $\|\psi\|_\infty \geqslant \sqrt{\dim V}\|\psi\|_2$.

If we consider a tower of congruence covers Y_N of Y, then the Laplace eigenspaces will have growing dimension because of multiplicities in the corresponding representations at places dividing N. Computationally, one observes that this (and its stronger form involving Arthur packets) is the only source of dimension growth. Although we believe that the dimensions of the joint eigenspaces we consider should be small (partly as a result of our choice of 'large' congruence subgroup), we do not know how to prove this in general. As a result, we shall be satisfied if we can beat the bound $\sqrt{\dim V}$, where V is now a space of Hecke–Maass forms with the same Laplace and Hecke eigenvalues. This motivates the following definitions.

Let G be as in Theorem 1.1. Let H be the identity component of the group of fixed points of θ. We let D be a positive integer such that G and H are unramified at places away from D and ∞; see §2.1 for a precise definition. Let K and K_H be compact open subgroups of $G(\mathbb{A}_f)$ and $H(\mathbb{A}_f)$ that are hyperspecial away from D, as in §§2.4 and 2.10. If N is a positive integer prime to D, we let $K(N)$ be the corresponding principal congruence subgroup of K, and define $Y_N = G(F)\backslash G(\mathbb{A})/K(N)K_H K_\infty$. We give each Y_N the probability volume measure.

Let $A \subset G_\infty$ be a maximal $\mathbb{R}$-split torus with real Lie algebra $\mathfrak{a}$ and Weyl group W. We let $\mathfrak{a}_\mathbb{C} = \mathfrak{a} \otimes \mathbb{C}$. Let G_∞^0 be the connected component of G_∞ in the real topology. Any unramified irreducible unitary representation of G_∞^0 gives rise to an element $\xi \in \mathfrak{a}_\mathbb{C}^*/W$ via the Harish-Chandra isomorphism, which we have normalized so that the tempered spectrum corresponds to $\mathfrak{a}^*/W$. We let $\|\cdot\|$ be the norm on $\mathfrak{a}$ and $\mathfrak{a}^*$ coming from the Killing form and extend it naturally to their complexifications. If $\mu, \lambda \in \mathfrak{a}_\mathbb{C}^*/W$ we will sometimes abuse notation and write $\|\mu - \lambda\|$ to mean the minimum of this norm over representatives for the W-orbits.

By a Hecke–Maass form we mean a joint eigenfunction $\psi \in L^2(Y_N)$ for the Hecke algebra (away from N and D) and the ring of invariant differential operators $\mathcal{D}$ on Y_N. We may view the associated eigenvalues as elements in the unramified unitary dual of G_v at finite places v

962

(via the Satake isomorphism), while at infinity they determine an element $\xi \in \mathfrak{a}_{\mathbb{C}}^*/W$. We define a spectral datum c for (G, N) to be a choice of element $\xi(c) \in \mathfrak{a}_{\mathbb{C}}^*/W$ and an element $\pi_v(c)$ in the unramified unitary dual of G_v for all $v \nmid ND\infty$. Given a spectral datum c for (G, N), we define $V(N, c)$ to be the space of Hecke–Maass forms on Y_N whose $\mathcal{D}$-eigenvalues are given by $\xi(c)$ (the spectral parameter) and whose Hecke eigenvalues at $v \nmid ND\infty$ are given by $\pi_v(c)$.

THEOREM 1.2. *With the notation and hypotheses of Theorem 1.1, there exist $\delta > 0$, depending only on G and θ, and $Q > 1$ with the following property. For any positive integer N with $(N, D) = 1$ and spectral parameter $\xi \in \mathfrak{a}^*$ such that $N(1 + \|\xi\|)$ is sufficiently large, there exist a spectral datum c for (G, N) with $\|\xi(c) - \xi\| \leqslant Q$ and a Hecke–Maass form $\psi \in V(N, c)$ such that*

$$\|\psi\|_\infty \gg N^\delta (1 + \|\xi\|)^\delta \sqrt{\dim V(N, c)} \|\psi\|_2.$$

Note that a Hecke–Maass form as in Theorem 1.1 has Laplacian eigenvalue of size roughly $(1 + \|\xi\|)^2$. Theorem 1.2 therefore implies Theorem 1.1. Moreover, taking ξ at distance at least Q from the root hyperplanes ensures that the eigenfunction ψ produced by the theorem is tempered at infinity. As in Theorem 1.1, the exponent δ is ineffective.

The only previous results giving lower bounds in the level aspect are for GL_2 over a number field, due to Saha [Sah15] and Templier [Tem14]. They use the fact that local Whittaker functions of highly ramified p-adic representations are large high in the cusp, and in particular rely on the non-compactness of the manifold.

1.4 Borel's theorem and a concrete example

A classical theorem of Borel on the existence of rationally defined Cartan involutions on real semisimple Lie algebras can be used to provide examples of groups satisfying the hypotheses of Theorem 1.1.[1] We state this as the following result and provide details for how to extract this statement from the work of Borel in §2.9.

PROPOSITION 1.3. *Let $G'/\mathbb{R}$ be connected, simply connected and $\mathbb{R}$-almost simple. Let F be a totally real number field, and let v_0 be a real place of F. There is a connected semisimple group G/F with $G_{v_0} \simeq G'$ that satisfies conditions (ii) and (iii) of Theorem 1.1.*

We now give a concrete example of a family of manifolds to which our theorem can be applied, and which to our knowledge does not already appear in the literature.

Let F be a totally real number field, and let E be a CM extension of F. Let the rings of integers of these fields be $\mathcal{O}_F$ and $\mathcal{O}_E$, respectively. Let v_0 be a distinguished real place of F, and let w_0 be the place of E over v_0. Let V be a vector space of dimension $n + 1$ over E with a Hermitian form $\langle \cdot, \cdot \rangle$ with respect to E/F. Assume that $\langle \cdot, \cdot \rangle$ has signature $(n, 1)$ at w_0 and is definite at all other infinite places of E. Let G be the F-algebraic group $\mathrm{SU}(V, \langle \cdot, \cdot \rangle)$, so that $G_{v_0} \simeq \mathrm{SU}(n, 1)$.

Let $L \subset V$ be an $\mathcal{O}_E$ lattice on which the form $\langle \cdot, \cdot \rangle$ is integral. Let L^* be the dual lattice $L^* = \{x \in V : \langle x, y \rangle \in \mathcal{O}_E \text{ for all } y \in L\}$. Let Γ be the group of isometries of V that have determinant 1, preserve L, and act trivially on L^*/L. If $F \neq \mathbb{Q}$, completion at w_0 allows us to consider Γ as a discrete, cocompact subgroup of $\mathrm{SU}(n, 1)$, which will be torsion-free if L is chosen sufficiently small.

---

[1] It is an interesting question whether condition (iii) on the existence of a rational Cartan involution is automatic or not. We believe that it is not when G is almost simple of type A_n, D_n, or E_6, but are unsure otherwise.

One may associate a complex hyperbolic manifold with Γ in the following way. Let D denote the space of lines in V_{w_0} on which the Hermitian form is negative definite. $\mathrm{SU}(n,1)$ acts on D, and D carries a natural $\mathrm{SU}(n,1)$-invariant metric under which it becomes a model for complex hyperbolic n-space. The quotient $Y = \Gamma \backslash D$ is then a compact complex hyperbolic n-manifold, and is an example of a congruence manifold associated with G as in Theorem 1.1.

If $n \geqslant 2$, G satisfies conditions (i) and (ii) of Theorem 1.1. We now show that (iii) is satisfied. Let $W \subset V$ be a codimension-1 subspace defined over E such that the Hermitian form is positive definite on W_{w_0}. Let θ be the isometry of reflection in W. Then $g \mapsto \theta g \theta^{-1}$ gives an F-involution of G that is a Cartan involution on G_{v_0}, as required. Theorem 1.1 then implies that there is a sequence of Laplace eigenfunctions $\{\psi_i\}$ on Y satisfying $\|\psi_i\|_\infty \gg \lambda_i^\delta \|\psi_i\|_2$.

1.5 The method of proof

The proof of power growth for arithmetic hyperbolic 3-manifolds of Maclachlan–Reid type by Milićević [Mil11] compares an amplified trace and pre-trace formula. Our proof works by extending this method to general groups. The bulk of the work lies in proving asymptotics for the trace formula.

More precisely, our proof of Theorem 1.2 proceeds by comparing a trace formula on G with a relative trace formula for G with respect to H. If we choose a test function $k \in C_c^\infty(G(\mathbb{A}))$, the main geometric terms of these trace formulae are $k(1)$ and $\Pi_H k(1)$ respectively, where $\Pi_H : L^1(G(\mathbb{A})) \to L^1(X(\mathbb{A}))$ is given by integration over H. In fact, we need to choose test functions of the form $k = \omega * \omega^*$ so that their action on $L^2(G(F)\backslash G(\mathbb{A}))$ is positive semidefinite. After controlling the other geometric terms, we wish to find ω that makes $\Pi_H k(1)$ large relative to $k(1)$. At finite places, we shall take for ω_f an appropriately large sum of L^2-normalized basic Hecke operators $\tau(\nu, v)$, supported on $K_v \nu(\varpi_v) K_v$, where ν is a cocharacter of G. The condition that $\Pi_H k(1)$ be large then boils down to

$$(\Pi_H \tau(\nu, v))(1) = \frac{\mathrm{vol}(H_v \cap K_v \nu(\varpi_v) K_v)}{\mathrm{vol}(K_v \nu(\varpi_v) K_v)^{1/2}}$$

being large, for enough places v. Note that this corresponds to our informal description in terms of Hecke returns at the beginning of the introduction: if the projection of the H-period onto the given locally symmetric Y is simply a point p (rather than a finite collection of such), the right-hand side above is roughly the multiplicity with which p appears in its image by the Hecke correspondence $\tau(\nu, v)$. In any case, we bound this quantity from below in Lemma 4.6, in terms of the H-relative size of ν. The latter is a certain cocharacter inequality, which we show in Lemma 2.2 is verified under the conditions of our main theorem.

While we were writing this paper, Erez Lapid pointed out to us that there was another approach to proving Theorems 1.1 and 1.2 based on a theorem of Sakellaridis on the unramified C^∞ spectrum of symmetric varieties. We have included a discussion of this in Appendix A. We have also included an explanation of why the condition of G being non-split at v_0 is natural, and motivated our choice of test functions in the trace formula, based on a related conjecture of Sakellaridis and Venkatesh on the L^2 spectrum.

2. Notation and proof of Proposition 1.3

2.1 Algebraic groups

Let $\mathcal{O}$ denote the ring of integers of F. Let $\mathbb{A}$ and $\mathbb{A}_f$ be the adeles and finite adeles of F.

Let G and θ be as in Theorems 1.1 and 1.2. Let H again denote the identity component of the group of fixed points of θ. We let $T \subset G$ and $T_H \subset H$ be maximal tori defined over F with $T_H \subset T$. We fix an F-embedding $\rho : G \to \mathrm{SL}_d$. Let Z be the centre of G.

Consider SL_d as a group scheme over $\mathcal{O}$ in the standard way, and take the schematic closures of T_H, T, H, and G in $\mathrm{SL}_d/\mathcal{O}$. These closures are group schemes over $\mathcal{O}$, and all the inclusions between them over F extend to closed embeddings over $\mathcal{O}$. By [Tit79, 3.9], we may choose $D \in \mathbb{N}$ such that, over $\mathcal{O}[1/D]$, all fibers of these group schemes are smooth, connected and reductive.

Let $X^*(T)$ and $X_*(T)$ denote the group of characters and cocharacters of $T \times_F \overline{F}$. Let Δ be the set of roots of T in G, and let Δ^+ be a choice of positive roots. Let W be the Weyl group of (G, T) over $\overline{F}$. We define

$$X_*^+(T) = \{\mu \in X_*(T) : \langle \mu, \alpha \rangle \geqslant 0, \alpha \in \Delta^+\}.$$

Similarly, we may define Δ_H, Δ_H^+, W_H and $X_*^+(T_H)$. Letting ρ and ρ_H denote, as usual, the half-sum of positive roots for G and H, we introduce the functions

$$\|\mu\|^* = \max_{w \in W} \langle w\mu, \rho \rangle,$$
$$\|\mu\|_H^* = \max_{w \in W_H} \langle w\mu, \rho_H \rangle$$

on $X_*(T)$ and $X_*(T_H)$, respectively. Our assumption that G is semisimple implies that $\|\mu\|^*$ is a norm; the condition that $\|\mu\|^* = \|-\mu\|^*$ follows from the fact that ρ and $-\rho$ lie in the same Weyl orbit. While $\|\mu\|_H^*$ also has the property that $\|\mu\|_H^* = \|-\mu\|_H^*$, it is trivial on cocharacters arising from the centre of H.

DEFINITION 2.1. We say that G is H-large if there exists a non-zero $\mu \in X_*(T_H)$ such that

$$2\|\mu\|_H^* \geqslant \|\mu\|^*. \tag{3}$$

Note that this definition is exactly the opposite of the definition of H being small in G given in [Mar14, Definition 8.1]. By [Mar14, Lemma 8.2], the definition of H-large is independent of the choice of tori $T_H \subset T$.

LEMMA 2.2. *A group G satisfying the hypotheses of Theorem 1.1 is H-large.*

Proof. This follows from [Mar14, Theorem 8.3]. Indeed, [Bor91, p. 277] implies that the compact subgroup H_{v_0} is connected in the real topology; by passing to the Lie algebra, this implies that H_{v_0} is a maximal compact connected subgroup of G_{v_0}. We then deduce from [Mar14, Theorem 8.3] that H_{v_0} is not small in G_{v_0}, and invariance of smallness under field extension then implies that G is H-large. $\square$

2.2 Local fields

If v is a place of F, we denote the completion by F_v. If v is finite, we denote the ring of integers, uniformizer and cardinality of the residue field by $\mathcal{O}_v$, ϖ_v and q_v, respectively. If $v \nmid D\infty$, we have the following consequences of our assumptions on D and ρ above.

(i) We have $G(\mathcal{O}_v) = \rho^{-1}(\rho(G_v) \cap \mathrm{SL}_d(\mathcal{O}_v))$ and $H(\mathcal{O}_v) = \rho^{-1}(\rho(H_v) \cap \mathrm{SL}_d(\mathcal{O}_v))$, so that $G(\mathcal{O}_v) \cap H_v = H(\mathcal{O}_v)$.

(ii) $G(\mathcal{O}_v)$ and $H(\mathcal{O}_v)$ are hyperspecial maximal compact subgroups of G_v and H_v, respectively.

(iii) If T (and hence T_H) split at v, the subgroups $G(\mathcal{O}_v)$ and $H(\mathcal{O}_v)$ correspond to points in the Bruhat–Tits buildings of G_v and H_v that lie in the apartments of T and T_H, respectively.

Property (i) is a consequence of the inclusions $H \subset G \subset \mathrm{SL}_d$ extending to closed embeddings over $\mathcal{O}_v$. Property (ii) is [Tit79, 3.8] or [BT84, 5.1.40]. Property (iii) is [BT84, 5.1.33], once we use [BT84, 5.1.40] to show that the $\mathcal{O}_v$ group scheme $T/\mathcal{O}_v$ is the standard one associated with T_v. We let $\mathcal{P}$ be the set of finite places of F that do not divide D and at which T splits. If $v \in \mathcal{P}$, our assumptions imply that G_v has a Cartan decomposition

$$G_v = \coprod_{\mu \in X_*^+(T)} G(\mathcal{O}_v) \mu(\varpi_v) G(\mathcal{O}_v)$$

with respect to $G(\mathcal{O}_v)$ and T.

2.3 Metrics
For any place v of F and $g \in G(F_v)$ let $\|g\|_v$ denote the maximum of the v-adic norms of the matrix entries of $\rho(g)$. For $g \in G(\mathbb{A}_f)$, let $\|g\|_f = \prod_{v \nmid \infty} \|g_v\|_v$. Fix a left-invariant Riemannian metric on $G(F_{v_0})$. Let $d(\cdot, \cdot)$ be the associated distance function. We define $d(x, y) = \infty$ when x and y are in different connected components of $G(F_{v_0})$ with the topology of a real manifold.

2.4 Compact subgroups
We choose compact subgroups K_v of G_v for all v such that:
- $K_{v_0} = H_{v_0}$;
- $K_v = G_v$ for all other real places;
- $\rho(K_v) \subset \mathrm{SL}_d(\mathcal{O}_v)$ for all finite v;
- $K_v = G(\mathcal{O}_v)$ for finite places $v \nmid D$; and
- $K = \prod_{v \nmid \infty} K_v$ is open in $G(\mathbb{A}_f)$.

By the remarks in the proof of Lemma 2.2, K_{v_0} is a maximal compact connected subgroup of G_{v_0}. We shall suppose that K_v for $v | D$ is sufficiently small to ensure that the finite group $Z(F) \cap K$ is reduced to $\{e\}$.

2.5 Measure normalizations
For any place v of F, let $\mu_{G,v}^{\mathrm{can}}$ be the canonical measure on $G(F_v)$ as defined by Gross in [Gro97, §11]; we recall this construction in §6.1. Then for all finite places $v \nmid D$ we have $\mu_{G,v}^{\mathrm{can}}(K_v) = 1$. We may then form the product measure $\mu_G^{\mathrm{can}} = \prod_v \mu_{G,v}^{\mathrm{can}}$ on $G(\mathbb{A})$. All convolutions (local and global) on G will be defined with respect to these measures. If $f \in C_c^\infty(G(\mathbb{A}))$, we define the operator $\pi(f)$ on $L^2(G(F) \backslash G(\mathbb{A}))$ by

$$[\pi(f)\phi](x) = \int_{G(\mathbb{A})} \phi(xg) f(g) \, d\mu_G^{\mathrm{can}}(g).$$

If $f \in C_c^\infty(G(\mathbb{A}))$, we define f^* by $f^*(g) = \overline{f}(g^{-1})$, so that $\pi(f)$ and $\pi(f^*)$ are adjoints.

The choice of canonical measure for G is imposed by the use of the Arthur–Selberg trace formula in §6; indeed, we want a uniform way of normalizing measures on the collection of connected reductive groups appearing as centralizers. We can afford to be more casual with measure normalizations for H, in light of our treatment of the geometric side of the relative trace formula in §5. For finite places v we choose Haar measures dh_v on $H(F_v)$ so that $H(F_v) \cap K_v$ is assigned measure 1. Because H_v is compact for archimedean v, at these places we choose Haar measures dh_v so that H_v has volume 1. We set $dh = \prod_v dh_v$.

966

2.6 Hecke algebras

If S is any finite set of finite places prime to D, let $\mathcal{H}_S$ be the convolution algebra of functions on $G(F_S)$ that are compactly supported and bi-invariant under K_S. Similarly, $\mathcal{H}_f$ is the convolution algebra of functions on $G(\mathbb{A}_f^D)$ that are compactly supported and bi-invariant under K^D. We identify $\mathcal{H}_S$ and $\mathcal{H}_f$ with subalgebras of $C_c^\infty(G(\mathbb{A}_f))$ in the natural way. If G_∞^0 denotes the connected component of the identity in G_∞ in the real topology, we define $\mathcal{H}_\infty$ to be the subspace of $C_c^\infty(G_\infty^0)$ consisting of functions that are bi-invariant under K_∞, and define $\mathcal{H} = \mathcal{H}_\infty \otimes \mathcal{H}_f$.

To control for the degree of Hecke operators in our later estimates, it will be convenient to work with the truncated Hecke algebras defined in [ST16]. Recall from §2.1 the Weyl-invariant norm $\|\cdot\|^*$ on $X_*(T)$. Let v be a finite place not dividing D. Because G_v is unramified, there is a maximal F_v-split torus A_v in G_v such that K_v corresponds to a point in the apartment of A_v. We may conjugate A_v inside T over $\overline{F_v}$, and obtain a norm $\|\cdot\|_v$ on $X_*(A_v)$. We then define

$$\mathcal{H}_v^{\leqslant \kappa} = \operatorname{span}_{\mathbb{C}} \{ \mathbf{1}_{K_v \mu(\varpi_v) K_v} : \mu \in X_*(A_v),\ \|\mu\|_v \leqslant \kappa \};$$

note that here and elsewhere, $\mathbf{1}_A$ denotes the characteristic function of the set A. If S is any finite set of finite places not dividing D, we set $\mathcal{H}_S^{\leqslant \kappa} = \bigotimes_{v \in S} \mathcal{H}_v^{\leqslant \kappa}$.

We note that $\mathcal{H}_v^{\leqslant \kappa}$ is independent of the choices made in its definition. First, the norm $\|\cdot\|_v$ is independent of the conjugation of A_v inside T. To see this, we note that $\|\lambda\|^*$ can also be defined by taking $\mathbb{G}_m \xrightarrow{\lambda} T \xrightarrow{\mathrm{Ad}} \mathrm{GL}(\mathfrak{g})$, decomposing the representation of $\mathbb{G}_m$ on $\mathfrak{g}$ into characters in $X^*(\mathbb{G}_m) \simeq \mathbb{Z}$, and taking half the sum of the positive terms. This lets us extend $\|\cdot\|^*$ to all of $X_*(G)$ in a conjugation-invariant way, and $\|\cdot\|_v$ is the restriction of this to $X_*(A_v)$. Secondly, if A_v and A_v' are two tori whose apartments contain the point corresponding to K_v, then by [Tit79, §2.2.1] they are conjugate by an element of K_v, and this conjugation respects the norms on $X_*(A_v)$ and $X_*(A_v')$. This implies that $\mathcal{H}_v^{\leqslant \kappa}$ is independent of A_v.

2.7 A lemma on double cosets

We shall need bounds on the volume of the support of Hecke operators on G_v and H_v at places where these groups are split. Such bounds are proved by Macdonald [Mac71] for split simply connected semisimple groups, but we shall need them in the general split reductive case, and derive them now from a result of Gross [Gro96].

In this subsection we let F be a p-adic field with integer ring $\mathcal{O}$, uniformizer ϖ, and residue field k, and let $q = \#k$. We let G be a split connected reductive group over F. Let $T \subset G$ be a maximal split torus, and $K \subset G$ a hyperspecial maximal compact subgroup corresponding to a point in the apartment of T. Let $\|\cdot\|^*$ be the function on $X_*(T)$ as in §2.1.

LEMMA 2.3. *We have $\#K\lambda(\varpi)K/K \sim q^{2\|\lambda\|^*}$ for all $\lambda \in X_*(T)$, where the implied constants depend only on $\dim G$.*

Proof. Let P_λ be the parabolic subgroup associated with λ, so that

$$\mathrm{Lie}(P_\lambda) = \mathrm{Lie}(T) + \bigoplus_{\langle \alpha, \lambda \rangle \geqslant 0} \mathrm{Lie}(G)_\alpha.$$

By Proposition 7.4 of [Gro96], we have $\#K\lambda(\varpi)K/K = q^{2\|\lambda\|^* - \dim(G/P_\lambda)} \#(G/P_\lambda)(k)$. Now

$$\#(G/P_\lambda)(k) = \#G(k)/\#P_\lambda(k) \quad \text{and} \quad \#P_\lambda(k) = \#M_\lambda(k) q^{\dim U_\lambda},$$

967

where $P_\lambda = M_\lambda U_\lambda$ is the standard Levi decomposition, and the first equality follows from [Bor91, Corollary 16.5]. It follows that

$$q^{-\dim(G/P_\lambda)} \#(G/P_\lambda)(k) = q^{\dim M_\lambda - \dim G} \#G(k)/\#M_\lambda(k).$$

The formula of Steinberg [Ste68] for the number of points of a reductive group over a finite field implies that $\#G(k)/q^{\dim G} = f(q^{-1})$ and $\#M_\lambda(k)/q^{\dim M_\lambda} = f_\lambda(q^{-1})$, where f and f_λ are polynomials with constant term 1 that depend only on the root data of G and M_λ, respectively. This completes the proof. □

2.8 Lie algebras

Let $\mathfrak{g}$ be the real Lie algebra of $G(F_\infty)$, and let $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ be the Cartan decomposition associated with K_∞. Let $\mathfrak{a} \subset \mathfrak{p}$ be a maximal abelian subalgebra. We let $\Delta_\mathbb{R}$ be the roots of $\mathfrak{a}$ in $\mathfrak{g}$, and let $\Delta_\mathbb{R}^+$ be a choice of positive roots. We let $W_\mathbb{R}$ be the Weyl group of $\Delta_\mathbb{R}$. For $\alpha \in \Delta_\mathbb{R}$, we let $m(\alpha)$ denote the dimension of the corresponding root space. We denote the Killing form on $\mathfrak{g}$ and $\mathfrak{g}^*$ by $\langle \cdot, \cdot \rangle$, and let $\| \cdot \|$ be the associated norm on $\mathfrak{a}$ and $\mathfrak{a}^*$.

Let $c(\xi)$ be Harish-Chandra's c-function; see [Hel08, ch. II, §3.3]. Then it is known [DKV79, (3.44)] that the upper bound

$$|c(\xi)|^{-2} \ll \beta(\xi) \tag{4}$$

holds for all $\xi \in \mathfrak{a}^*$, where we have put

$$\beta(\xi) = \prod_{\alpha \in \Delta_\mathbb{R}^+} (1 + |\langle \alpha, \xi \rangle|)^{m(\alpha)}.$$

2.9 Proof of Proposition 1.3

We now recall Borel's theorem on the existence of rationally defined Cartan involutions on real semisimple Lie algebras. This is the central ingredient in the proof of Proposition 1.3.

PROPOSITION 2.4 (Borel). *Let F be a totally real number field, and let v_0 be a real place of F. Let $\mathfrak{g}'$ be a real semisimple Lie algebra. There exists $\mathfrak{g}/F$ with an involution θ defined over F such that $\mathfrak{g}_{v_0} \simeq \mathfrak{g}'$, $\mathfrak{g}_v$ is of compact type for all real $v \neq v_0$, and θ induces a Cartan involution of $\mathfrak{g}_{v_0}$.*

Proof. The statement can be extracted from the proof of [Bor63, Proposition 3.8]. For completeness we provide details.

By [Bor63, Proposition 3.7], there exist $\mathfrak{g}_0/\mathbb{Q}$ and a $\mathbb{Q}$-involution θ of $\mathfrak{g}_0$ such that $\mathfrak{g}_0 \otimes \mathbb{R} \simeq \mathfrak{g}'$, and θ induces a Cartan involution of $\mathfrak{g}_0 \otimes \mathbb{R}$. Let $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ be the rational Cartan decomposition associated with θ. Choose $\alpha \in F$ such that $F = \mathbb{Q}(\alpha)$ and $\alpha_{v_0} > 0$, and $\alpha_v < 0$ for all other real v. Define $\mathfrak{g} = \mathfrak{k}_0 \otimes F + \sqrt{\alpha}\, \mathfrak{p}_0 \otimes F \subset \mathfrak{g}_0 \otimes \overline{\mathbb{Q}}$. We extend θ to an involution of $\mathfrak{g}$ defined over F.

We have $\mathfrak{g}_{v_0} \simeq \mathfrak{g}_0 \otimes \mathbb{R} \simeq \mathfrak{g}'$, and it is clear that θ induces a Cartan involution of $\mathfrak{g}_{v_0}$. If $v \neq v_0$ is real, we have $\mathfrak{g}_v = \mathfrak{k}_0 \otimes \mathbb{R} + i\mathfrak{p}_0 \otimes \mathbb{R} \subset \mathfrak{g}_0 \otimes \mathbb{C}$ so that $\mathfrak{g}_v$ is of compact type. □

The corresponding statement for simply connected semisimple groups, Proposition 1.3, can be deduced from the Lie algebra version, as follows.

Let G' be as in Proposition 1.3 and let $\mathfrak{g}'/\mathbb{R}$ be its Lie algebra. Let F be a totally real number field and v_0 a real place of F. Applying Proposition 2.4 to $\mathfrak{g}'$ with this choice of F and v_0, we obtain a corresponding $\mathfrak{g}/F$ and involution θ. Let $\mathrm{Int}(\mathfrak{g})$ be the identity component of $\mathrm{Aut}(\mathfrak{g})$, and let G be the simply connected cover of $\mathrm{Int}(\mathfrak{g})$. Then G has Lie algebra $\mathfrak{g}$, and θ induces

968

an involution of G which we also denote by θ (as the derivative at the identity of the former is equal to the latter). Because G' and G_{v_0} are connected, simply connected semisimple groups with isomorphic Lie algebras, they are isomorphic. If $v \neq v_0$ is real, then $\mathfrak{g}_v$ is of compact type and $G(F_v)$ is compact.

By Proposition 2.4, θ induces a Cartan involution of $\mathfrak{g}_{v_0}$. Because G is simply connected, G_{v_0} is connected in the real topology by the Cartan connectedness theorem [OV94, ch. 4, § 2.2, Theorem 2.2], and it follows that θ also induces a Cartan involution of G_{v_0} (in the sense that the fixed point set $G_{v_0}^\theta$ is a maximal compact subgroup of G_{v_0}). This completes the proof of Proposition 1.3.

2.10 Congruence subgroups and adelic quotients

Recall from § 2.1 the integer D, and from § 2.4 the compact open subgroups K_v for finite places v and their product $K = \prod_{v \nmid \infty} K_v$ inside $G(\mathbb{A}_f)$. We now define congruence subgroups of K_v and K to be used throughout the paper.

Fix an integer $N \geqslant 1$ prime to D. For $v \nmid D\infty$ we write $K_v(N)$ for the level N principal congruence subgroup, given by $K_v(N) = \rho^{-1}(\rho(G_v) \cap \mathrm{SL}_d(\mathcal{O}_v)(N))$. Here $\mathrm{SL}_d(\mathcal{O}_v)(N)$ denotes the principal congruence subgroup of $\mathrm{SL}_d(\mathcal{O}_v)$. Let S be a finite set of finite places prime to D. Then we put

$$K(N) = \prod_{v \nmid D\infty} K_v(N) \prod_{v | D} K_v \quad \text{and} \quad K^S(N) = \prod_{\substack{v \nmid D\infty \\ v \notin S}} K_v(N) \prod_{v | D} K_v.$$

For every finite v we define $K_{H,v} = H_v \cap K_v$. Letting

$$K_H = \prod_{v \nmid \infty} K_{H,v} \quad \text{and} \quad K_H^S = \prod_{v \notin S \cup \infty} K_{H,v},$$

then the congruence manifold of interest to us is

$$Y_N = G(F) \backslash G(\mathbb{A}) / K(N) K_H K_\infty.$$

Let Vol_N be the volume assigned by $\prod_{v \nmid \infty} \mu_{G,v}^{\mathrm{can}}$ to the compact open subgroup $K(N) K_H$.

Finally, we write

$$[G] = G(F) \backslash G(\mathbb{A}) \quad \text{and} \quad [H] = H(F) \backslash H(\mathbb{A})$$

for the automorphic spaces associated with G and H. We shall identify $L^2(Y_N)$ with the functions in $L^2([G])$ fixed under $K(N) K_H K_\infty$.

2.11 Hecke–Maass forms

Let $\mathcal{D}$ be the algebra of differential operators on G_∞^0 / K_∞ that are invariant under the left action of G_∞^0. Note that if we define $K_{v_0}^+$ to be the group of fixed points of θ on G_{v_0}, and let K_∞^+ be the group obtained by replacing K_{v_0} with $K_{v_0}^+$ in the definition of K_∞, then K_∞^+ is a maximal compact subgroup of G_∞, we have $G_\infty / K_\infty^+ \simeq G_\infty^0 / K_\infty$, and the elements of $\mathcal{D}$ are also invariant under the larger group G_∞. It follows that $\mathcal{D}$ descends to an algebra of operators on Y_N in a natural way.

We define a Hecke–Maass form to be a function on Y_N that is an eigenfunction of the ring $\mathcal{D}$ on Y_N and the Hecke algebra $\mathcal{H}_f$ (and hence of $\mathcal{H}$). If ψ is a Hecke–Maass form and $\omega \in \mathcal{H}$, then, using the notation in § 2.5, we define $\widehat{\omega}(\psi)$ by the equation $\pi(\omega)\psi = \widehat{\omega}(\psi)\psi$. We define the spectral parameter of ψ to be the unique $\xi \in \mathfrak{a}_{\mathbb{C}}^* / W_{\mathbb{R}}$ such that ψ has the same eigenvalues under the action of $\mathcal{D}$ as the associated spherical function φ_ξ. The Laplace eigenvalue of ψ is given by $(\Delta + C_1(G) + \langle \xi, \xi \rangle)\psi = 0$ for some $C_1(G) \in \mathbb{R}$.

3. A comparison of trace formulae

In the next two sections, we state without proof a trace formula and relative trace formula, and deduce Theorems 1.1 and 1.2 from them. The trace formulae will be proved in §§ 5 and 6.

3.1 The trace formula

Let N be a positive integer prime to D. Let D' be a sufficiently large multiple of D, depending only on G, to be chosen in § 6.4. Let T be a finite set of finite places prime to D' and N. Let $\xi \in \mathfrak{a}^*$. We will use the trace formula to count Hecke–Maass forms on Y_N with spectral parameter near ξ, weighted by a Hecke operator at T. To do this, we shall consider test functions of the form $\phi = \mathbf{1}_N^T \otimes k_T \otimes k_\xi$, for $k_T \in \mathcal{H}_T$ and $k_\xi \in \mathcal{H}_\infty$, where we have put

$$\mathbf{1}_N^T = \mathbf{1}_{K^T(N)K_H^T}. \tag{5}$$

We introduce the following condition on k_ξ.

(P_∞): k_ξ is supported in $\{g \in G(F_\infty) : d(g, H_{v_0}) < 1\}$ and satisfies

$$k_\xi(g) \ll \beta(\xi)(1 + \|\xi\| d(g, H_{v_0}))^{-1/2}.$$

If U is any finite set of finite places, we define $q_U = \prod_{v \in U} q_v$. Our asymptotic trace formula is given in the following theorem.

THEOREM 3.1. *There are constants $A, B, \delta, \eta > 0$ such that the following statement holds. Let N, T and ξ be as above, and let $\kappa \geqslant 0$. For any $k_T \in \mathcal{H}_T^{\leqslant \kappa}$ and any $k_\xi \in \mathcal{H}_\infty$ satisfying (P_∞) we have*

$$\mathrm{Vol}_N \sum_{i \geqslant 0} \widehat{k_T \otimes k_\xi}(\psi_i) = \mu_G^{\mathrm{can}}([G])k_T(1)k_\xi(1) + O(N^{-\delta} q_T^{A\kappa + B} \beta(\xi)(1 + \|\xi\|)^{-\eta} \|k_T\|_\infty),$$

where $\{\psi_i\}$ runs over an orthonormal basis of Hecke–Maass forms for $L^2(Y_N)$. The implied constant depends only on G, θ and K.

We note that the exponents A, B and δ are ineffective, and this is the only source of ineffectiveness of the exponent in Theorems 1.1 and 1.2. As discussed in § 1.1, this comes from using theorems of Finis and Lapid and of Cluckers, Gordon and Halupczok to bound orbital integrals.

3.2 The relative trace formula

For $\psi \in C^\infty(G(F)\backslash G(\mathbb{A}))$ we consider the H-automorphic period

$$\mathcal{P}_H(\psi) = \int_{H(F)\backslash H(\mathbb{A})} \psi(h)\, dh.$$

Our relative trace formula will give the average size of these periods over an orthonormal basis $\{\psi_i\}$ of Hecke–Maass forms for $L^2(Y_N)$.

Let N be a positive integer prime to D, S a finite set of finite places prime to D and N, and $\xi \in \mathfrak{a}^*$. We now consider test functions of the form $\phi = \mathbf{1}_N^S \otimes k_S \otimes k_\xi$, for $k_S \in \mathcal{H}_S$ and $k_\xi \in \mathcal{H}_\infty$. As well as condition (P_∞), we shall use the following condition on k_S.

(P_S): There is $R > 1$ such that $\|g\|_S \leqslant R$ for all $g \in \mathrm{supp}(k_S)$.

The main term of our formula is expressed in terms of the averaging map

$$\Pi_H : L^1(G(F_S)) \longrightarrow L^1(G(F_S)/H(F_S))$$

given by

$$\Pi_H k_S(g) = \int_{H(F_S)} k_S(gh)\, dh.$$

With this notation in hand, we may state our asymptotic relative trace formula.

PROPOSITION 3.2. *There is $A > 0$ such that the following statement holds. Let N, S and ξ be as above. Let $k_S \in \mathcal{H}_S$ satisfy (P_S) and $k_\xi \in \mathcal{H}_\infty$ satisfy (P_∞). Then*

$$\mathrm{Vol}_N \sum_{i \geqslant 0} \widehat{k_S \otimes k_\xi}(\psi_i)|\mathcal{P}_H(\psi_i)|^2$$
$$= \mathrm{vol}([H])\Pi_H k_S(1)k_\xi(1) + O\big(\#(\operatorname{supp} k_S/K_S)\beta(\xi)(1 + \|\xi\|)^{-1/4}N^{-1/4}R^A\|k_S\|_\infty\big),$$

where $\{\psi_i\}$ runs over an orthonormal basis of Hecke–Maass forms for $L^2(Y_N)$.

3.3 Proof of Theorems 1.1 and 1.2
Recall from § 2.2 that $\mathcal{P}$ denotes the set of finite places of F that do not divide D and at which T splits. For $v \in \mathcal{P}$ and $\mu \in X_*(T)$, we define $\tau(v, \mu) \in \mathcal{H}_v$ to be the function supported on $K_v \mu(\varpi_v) K_v$ and taking the value $q_v^{-\|\mu\|^*}$ there. Let P be a positive integer and put

$$S = \{v \in \mathcal{P} : P/2 \leqslant q_v < P,\ v \nmid N,\ v \nmid D'\}. \tag{6}$$

Choose any non-zero $\nu \in X_*(T)$, and let

$$\omega_S = \sum_{v \in S} \tau(v, \nu),$$

where, as in § 2.6, we are identifying $\mathcal{H}_v$ with a subalgebra of $\mathcal{H}_S$ in the natural way.

In § 4 we will deduce the following weighted bounds from our two trace formulae.

PROPOSITION 3.3. *For any ν and P, we have*

$$\mathrm{Vol}_N \sum_{\|\Re\xi_i - \xi\| \leqslant 1} |\widehat{\omega_S}(\psi_i)|^2 \ll P\beta(\xi)(1 + P^A N^{-\delta}(1 + \|\xi\|)^{-\delta}). \tag{7}$$

In addition, if $\nu \in X_*(T_H)$ satisfies the H-large inequality (3), then there is $Q > 0$ depending on ν such that

$$\mathrm{Vol}_N \sum_{\|\Re\xi_i - \xi\| \leqslant Q} |\widehat{\omega_S}(\psi_i)|^2 |\mathcal{P}_H(\psi_i)|^2 \gg_\epsilon P^{2-\epsilon}\beta(\xi)(1 + O(P^A N^{-\delta}(1 + \|\xi\|)^{-\delta})). \tag{8}$$

The exponent A and the implied constants depend on the choice of ν.

Theorems 1.1 and 1.2 follow easily from Proposition 3.3, as we now explain. Choose $0 \neq \nu \in X_*(T_H)$ satisfying (3), and choose P to be a small power of $(1 + \|\xi\|)N$. We then find $\delta > 0$ such that for $(1 + \|\xi\|)N$ sufficiently large,

$$\mathrm{Vol}_N \sum_{\|\Re\xi_i - \xi\| \leqslant Q} |\widehat{\omega_S}(\psi_i)|^2 \ll \beta(\xi)(1 + \|\xi\|)^\delta N^\delta \tag{9}$$

971

and

$$\text{Vol}_N \sum_{\|\Re\xi_i-\xi\|\leqslant Q} |\widehat{\omega_S}(\psi_i)|^2 |\mathcal{P}_H(\psi_i)|^2 \gg_\epsilon \beta(\xi)(1+\|\xi\|)^{2\delta-\epsilon} N^{2\delta-\epsilon}. \tag{10}$$

Comparing (9) and (10) we find that there exist $\delta > 0$ and $Q > 1$ such that, for every $\xi \in \mathfrak{a}^*$ and every N with $(1+\|\xi\|)N$ large, there is a Hecke–Maass form ψ_i on Y_N with spectral parameter $\|\xi_i - \xi\| \leqslant Q$ and satisfying $|\mathcal{P}_H(\psi_i)| \gg (1+\|\xi_i\|)^\delta N^\delta$. This implies the same lower bound on $\|\psi_i\|_\infty$, and in particular proves Theorem 1.1.

It remains to refine this power growth to obtain the stated lower bounds of Theorem 1.2. For this, we will make a special choice of orthonormal basis of Hecke–Maass forms for $L^2(Y_N)$. First recall that we have a Hilbert direct sum decomposition

$$L^2(Y_N) = \bigoplus_\pi m(\pi)\pi^{K(N)K_H K_\infty},$$

the sum ranging over irreducible representations of $G_\infty^0 \times G(\mathbb{A}_f)$ in $L^2(G(F)\backslash G(\mathbb{A}))$ satisfying $\pi^{K(N)K_H K_\infty} \neq 0$, each occurring with multiplicity $m(\pi)$. Each π gives rise to a spectral datum c_π (as defined in the paragraph preceding Theorem 1.2), and this assignment $\pi \mapsto c_\pi$ is finite-to-one. Given a spectral datum c for (G, N), the space $V(N, c)$ is the direct sum of $m(\pi)\pi^{K(N)K_H K_\infty}$ over all π in the fiber over c.

Let $\{c_i\}$ be an enumeration of the spectral data, and for each c_i write $V_i = V(N, c_i)$ and $\xi_i = \xi(c_i)$. The automorphic period $\mathcal{P}_H$ defines a linear functional on each finite-dimensional vector space V_i, and its kernel is of codimension at most 1. If the kernel is of codimension 1 we let $\phi_i \in V_i$ be a unit normal to it, and otherwise choose $\phi_i \in V_i$ to be an arbitrary unit vector. We may complete this set of vectors to an orthonormal basis of Hecke–Maass forms for $L^2(Y_N)$, to which we apply (9) and (10). Since $\mathcal{H}$ acts as a scalar on V_i we may write the left-hand side of (9) as

$$\text{Vol}_N \sum_{\substack{c_i \\ \|\Re\xi_i-\xi\|\leqslant Q}} |\widehat{\omega_S}(\phi_i)|^2 \dim V_i,$$

and the left-hand side of (10) as

$$\text{Vol}_N \sum_{\substack{c_i \\ \|\Re\xi_i-\xi\|\leqslant Q}} |\widehat{\omega_S}(\phi_i)|^2 |\mathcal{P}_H(\phi_i)|^2.$$

We obtain Theorem 1.2 by comparing the right-hand sides of (9) and (10) as before. $\quad\square$

4. Bounds for average periods

We now deduce the average period bounds of Proposition 3.3 from Theorem 3.1 and Proposition 3.2. For this, we must choose test functions $k_\xi \in \mathcal{H}_\infty$ and $k_S \in \mathcal{H}_S$ to insert into Theorem 3.1 and Proposition 3.2 and explicate the error terms in those results.

4.1 The archimedean test function

Let G_∞^0 denote the connected component of the identity in G_∞ in the real topology. If $\mu \in \mathfrak{a}_\mathbb{C}^*$, we define φ_μ to be the corresponding spherical function on G_∞^0. If $k_\infty \in C_c^\infty(G_\infty^0)$, we define its Harish-Chandra transform by

$$\widehat{k}_\infty(\mu) = \int_{G_\infty^0} k_\infty(g)\varphi_{-\mu}(g)\, d\mu_{G,\infty}^{\text{can}}(g).$$

972

Note that this is related to our earlier notation as follows: if ψ is a Maass form with spectral parameter μ and k_∞ is bi-K_∞-invariant, then $\widehat{k}_\infty(\psi) = \widehat{k}_\infty(-\mu)$.

We shall choose k_ξ so that its Harish-Chandra transform concentrates around $-\xi \in \mathfrak{a}^*$. For this we first take a function $h_0 \in C^\infty(\mathfrak{a}^*)$ of Paley–Wiener type that is real, non-negative and satisfies $h_0(0) = 1$. We implicitly extend h_0 to a function on $\mathfrak{a}^*_\mathbb{C}$. Let

$$h_\xi^0(\nu) = \sum_{w \in W_\mathbb{R}} h_0(w\nu - \xi),$$

and let k_ξ^0 be the bi-K_∞-invariant function on G_∞^0 satisfying $\widehat{k}_\xi^0(-\mu) = h_\xi^0(\mu)$. We define $k_\xi = k_\xi^0 * k_\xi^0$ and $h_\xi = (h_\xi^0)^2$. We have $\widehat{k}_\xi(-\mu) = h_\xi(\mu)$, so that if ψ has spectral parameter μ as before then $\widehat{k}_\xi(\psi) = h_\xi(\mu)$. We define a unitary spectral parameter to be a spectral parameter of a spherical unitary representation of G_∞^0. We have $\|\Im\mu\| \leqslant \|\rho\|$ for any unitary spectral parameter μ [Hel00, ch. IV, Theorem 8.1]. If $\mu \in \mathfrak{a}^*_\mathbb{C}$ is unitary, Lemma 4.1 implies that $h_\xi(\mu) = h_\xi^0(\mu)h_\xi^0(\overline{\mu}) = |h_\xi^0(\mu)|^2 \geqslant 0$.

LEMMA 4.1. *If $\mu \in \mathfrak{a}^*_\mathbb{C}$ is a unitary spectral parameter, then $W_\mathbb{R}\mu = W_\mathbb{R}\overline{\mu}$.*

Proof. It suffices to show that $\varphi_\mu = \varphi_{\overline{\mu}}$. Realizing φ_μ as a matrix coefficient of a unitary representation gives $\varphi_\mu(g) = \overline{\varphi}_\mu(g^{-1})$ for all g. The usual formula for φ_μ as a K_∞-integral gives $\overline{\varphi}_\mu(g^{-1}) = \varphi_{-\overline{\mu}}(g^{-1})$. The identity $\varphi_{-\overline{\mu}}(g^{-1}) = \varphi_{\overline{\mu}}(g)$, obtained for example by setting $h = 1$ in [Hel00, ch. IV, Lemma 4.4] and comparing with Theorem 4.3 there, completes the proof. $\square$

LEMMA 4.2. *The function k_ξ satisfies property (P_∞).*

Proof. Note that k_ξ is of compact support independent of ξ from the Paley–Wiener theorem of [Gan71]; we may thus take h_0 so that the support of k_ξ lies in $\{g \in G(F_\infty) : d(g, H_{v_0}) < 1\}$ for all ξ.

We have the inversion formula

$$k_\xi(g) = \frac{1}{|W_\mathbb{R}|} \int_{\mathfrak{a}^*} \widehat{k}_\xi(\mu)\varphi_\mu(g)|c(\mu)|^{-2} \, d\mu. \tag{11}$$

We can now quote either [BP16, Theorem 2] or [MT15, Proposition 7.2] and apply our conditions on G_∞, to find

$$\varphi_\mu(x) \ll_C (1 + \|\mu\|d(x, K_\infty))^{-1/2}, \tag{12}$$

for $\mu \in \mathfrak{a}^*$ and x in a compact set $C \subset G_\infty$. In the inversion formula (11), we apply (4) to bound the Plancherel density function, (12) to bound the spherical function, and recall the concentration of $\widehat{k}_\xi$ around $-\xi$, to obtain

$$k_\xi(x) \ll \beta(\xi)(1 + \|\xi\|d(x, K_\infty))^{-1/2},$$

whence the claim. $\square$

The following lemma will be helpful when truncating spectral sums. We let $B(\mu) \subset \mathfrak{a}^*$ denote the unit ball around $\mu \in \mathfrak{a}^*$.

LEMMA 4.3. *We may choose the function h_0 so that the functions h_μ, $\mu \in \mathfrak{a}^*$, have the following property: if λ is a unitary spectral parameter with $\Re\lambda \in B(\mu)$, then $h_\mu(\lambda) \geqslant c > 0$ for some c depending only on G.*

973

Proof. Let $0 < \delta < 1$. Let $b \in C_0^\infty(\mathfrak{a})$ be non-negative, supported in the δ-ball around 0, and satisfy $\int b = 1$. We further assume that $b = b_0 * b_0$ for some even real-valued b_0, which ensures that $\widehat{b}(\nu)$ is non-negative for $\nu \in \mathfrak{a}^*$. We will show that $h_0 = \widehat{b}$ suffices in our construction of h_μ.

We let $C > 0$ be a constant depending only on G that may vary from line to line. We start by showing that $\Re\widehat{b}(\nu) \geqslant -C\delta$ for all $\nu \in \mathfrak{a}_{\mathbb{C}}^*$ with $\|\Im\nu\| \leqslant \|\rho\|$. We have

$$
\begin{aligned}
\widehat{b}(\nu) &= \int_{\mathfrak{a}} b(H) e^{-i\nu(H)} \, dH \\
&= \int_{\mathfrak{a}} b(H) [e^{-i\Re\nu(H)} + (e^{-i\nu(H)} - e^{-i\Re\nu(H)})] \, dH \\
&= \widehat{b}(\Re\nu) + \int_{\mathfrak{a}} b(H) e^{-i\Re\nu(H)} (e^{\Im\nu(H)} - 1) \, dH.
\end{aligned}
$$

As $\|\Im\nu\| \leqslant \|\rho\|$, we have $|e^{\Im\nu(H)} - 1| \leqslant C\delta$ for all $H \in \mathrm{supp}(b)$, so that

$$
\Re\widehat{b}(\nu) \geqslant \widehat{b}(\Re\nu) - C\delta \geqslant -C\delta
$$

as required.

We now take $h_0 = \widehat{b}$, and construct h_μ^0 and h_μ as before. If we choose δ small, we will have $\Re h_0(\nu) \geqslant 1/2$ for all ν with $\|\Re\nu\| \leqslant 1$, $\|\Im\nu\| \leqslant \|\rho\|$. Therefore if $\lambda \in \mathfrak{a}_{\mathbb{C}}^*$ is unitary with $\Re\lambda \in B(\mu)$, the bound $\|\Im\lambda\| \leqslant \|\rho\|$ implies that

$$
\begin{aligned}
\Re h_\mu^0(\lambda) &= \sum_{w \in W_{\mathbb{R}}} \Re h_0(w\lambda - \mu) \\
&\geqslant \Re h_0(\lambda - \mu) - C\delta \geqslant 1/4.
\end{aligned}
$$

This gives $h_\mu(\lambda) = |h_\mu^0(\lambda)|^2 \geqslant 1/16$ as required. $\qquad\square$

4.2 The S-adic test function

We choose the test function at places in S to be $k_S = \omega_S \omega_S^* \in \mathcal{H}_S$, where ω_S is as in § 3.3.

LEMMA 4.4. *The function k_S satisfies the following properties.*

(a) *There is $B > 0$ such that $\|g\|_S \ll P^B$ for all $g \in \mathrm{supp}(k_S)$.*
(b) *We have $\|k_S\|_\infty \ll P$.*
(c) *There is $C > 0$ such that $\#(\mathrm{supp}\, k_S/K_S) \ll P^C$.*

The exponents B and C depend on the underlying choice of ν in the definition of k_S. All implied constants depend on G and ν.

Proof. (a) Because supports add under convolution, it suffices to show this for $g \in \mathrm{supp}(\tau(v, \nu))$ (or $\mathrm{supp}(\tau(v, -\nu))$, which is similar and we omit) for some $v \in S$. Our assumption that $\rho(K_w) \subset \mathrm{SL}_d(\mathcal{O}_w)$ for all finite w means that we only need to estimate $\|\rho(g)\|_v$. Consider ρ as a representation of G, and let $\Omega \subset X^*(T)$ be the multiset of weights of this representation. It follows that $\rho(\nu(\varpi_v))$ is semisimple with eigenvalues $\{\varpi_v^{\langle\omega,\nu\rangle} : \omega \in \Omega\}$. By [ST16, Lemma 2.17], there is $x \in \mathrm{GL}_d(\mathcal{O}_v)$ such that $x\rho(T_v)x^{-1}$ is diagonal, and so if we define $A = \max\{-\langle\omega,\nu\rangle : \omega \in \Omega\}$ then we have $\|\rho(\nu(\varpi_v))\|_v = \|x\rho(\nu(\varpi_v))x^{-1}\|_v = q_v^A < P^A$. Because $\rho(G(\mathcal{O}_v)) \subset \mathrm{SL}_d(\mathcal{O}_v)$, the same holds for $g \in \mathrm{supp}(\tau(v, \nu))$.

(b) We note that

$$
k_S = \sum_{v \in S} \tau(v, \nu)\tau(v, \nu)^* + \sum_{\substack{v, w \in S \\ v \neq w}} \tau(v, \nu)\tau(w, \nu)^*. \tag{13}
$$

The bound is clear for the second sum, because the terms satisfy $\|\tau(v,\nu)\tau(w,\nu)^*\|_\infty \leqslant 1$ and the supports of the terms are disjoint. For the first sum, we have

$$\|\tau(v,\nu)\tau(v,\nu)^*\|_\infty \leqslant \|\tau(v,\nu)\|_2^2,$$

and $\|\tau(v,\nu)\|_2 \ll 1$ follows from Lemma 2.3.

(c) We may write the first sum in (13) as a linear combination of $\tau(v,\mu)$ with μ lying in a finite set depending on ν. The bound now follows from Lemma 2.3. The second sum may be treated similarly. $\qquad\square$

4.3 Proof of Proposition 3.3

We are now in a position to prove Proposition 3.3. We first note that $k_\xi(1) \asymp \beta(\xi)$, where the upper and lower bounds follow from property (P_∞) and [DKV79, (3.44a)], respectively. Moreover, $\beta(\xi)$ is bounded above (and below) by a power of $(1 + \|\xi\|)$.

We now apply Proposition 3.2 with test functions $k_S k_\xi$. Moreover, for every pair $v, w \in S$ we put $T = \{v, w\}$ and apply Theorem 3.1 with test function $k_T k_\xi$ where $k_T = \tau(v,\nu)\tau(w,\nu)^*$; we then sum over such pairs v, w. As a result, we deduce the existence of constants $A > 2$ (depending on ν) and $\delta > 0$ such that

$$\mathrm{Vol}_N \sum_{i \geqslant 0} |\widehat{\omega_S}(\psi_i)|^2 h_\xi(\xi_i) \asymp \omega_S \omega_S^*(1)\beta(\xi) + O(P^A N^{-\delta}\beta(\xi)(1 + \|\xi\|)^{-\delta}) \tag{14}$$

and

$$\mathrm{Vol}_N \sum_{i \geqslant 0} |\widehat{\omega_S}(\psi_i)|^2 |\mathcal{P}_H(\psi_i)|^2 h_\xi(\xi_i) \asymp \Pi_H \omega_S \omega_S^*(1)\beta(\xi) + O(P^A N^{-\delta}\beta(\xi)(1 + \|\xi\|)^{-\delta}). \tag{15}$$

The error term in (14) was obtained by observing that $\tau(v,\nu)\tau(w,\nu)^* \in \mathcal{H}_T^{\leqslant\kappa}$, for some κ depending only on ν, and $q_T = q_v q_w < P^2$, so that $q_T^{A\kappa+B}$ is bounded by a power of P; we must also insert the L^∞ norm estimate for k_T coming from the proof of Lemma 4.4. The error term in (15) was obtained by taking B as in Lemma 4.4 and setting $R = P^B$ in condition (P_S), and inserting the L^∞ norm estimate of Lemma 4.4. It remains then to explicate the size of $\omega_S \omega_S^*(1)$ and $\Pi_H \omega_S \omega_S^*(1)$ (the latter under our extra condition on ν), and truncate the spectral sums. We continue to use the convention that the values of the exponents $A, \delta > 0$ can vary from line to line.

We begin with (14). Lemma 4.4 (b) gives $\omega_S \omega_S^*(1) \ll P$. By positivity we may truncate the spectral sum to obtain the upper bound

$$\mathrm{Vol}_N \sum_{\|\Re\xi_i - \xi\| \leqslant 1} |\widehat{\omega_S}(\psi_i)|^2 h_\xi(\xi_i) \ll P\beta(\xi)(1 + P^A N^{-\delta}(1 + \|\xi\|)^{-\delta}).$$

Combined with Lemma 4.3, this implies (7).

We next examine (15). We begin by observing that if $\mu \in \mathfrak{a}^*$, then we may again truncate the spectral sum and apply Lemma 4.3 to obtain

$$\mathrm{Vol}_N \sum_{\|\Re\xi_i - \mu\| \leqslant 1} |\widehat{\omega_S}(\psi_i)|^2 |\mathcal{P}_H(\psi_i)|^2 \ll \Pi_H \omega_S \omega_S^*(1)\beta(\mu) + P^A \beta(\mu)(1 + \|\mu\|)^{-\delta} N^{-\delta}. \tag{16}$$

We next use this to truncate the spectral sum about ξ.

975

LEMMA 4.5. *For any $Q > 1$ the left-hand side of (15) can be written as*

$$\mathrm{Vol}_N \sum_{\|\Re\xi_i - \xi\| \leqslant Q} |\widehat{\omega_S}(\psi_i)|^2 |\mathcal{P}_H(\psi_i)|^2 h_\xi(\xi_i)$$
$$+ O_M(\Pi_H \omega_S \omega_S^*(1)\beta(\xi)Q^{-M}) + O(P^A N^{-\delta}\beta(\xi)(1 + \|\xi\|)^{-\delta}).$$

Proof. We must show that the sum

$$\mathfrak{S} = \mathrm{Vol}_N \sum_{\|\Re\xi_i - \xi\| > Q} |\widehat{\omega_S}(\psi_i)|^2 |\mathcal{P}_H(\psi_i)|^2 h_\xi(\xi_i)$$

satisfies

$$\mathfrak{S} = O_M(\Pi_H \omega_S \omega_S^*(1)\beta(\xi)Q^{-M}) + O(P^A N^{-\delta}\beta(\xi)(1 + \|\xi\|)^{-\delta}).$$

Break the region in the positive chamber $\mathfrak{a}_+^*$ defined by $\|\mu - \xi\| > Q$ into an overlapping union of unit balls $B(\mu_n)$ centred at points $\mu_n \in \mathfrak{a}^*$. Because unitary spectral parameters μ satisfy $\|\Im\mu\| \leqslant \|\rho\|$, they also satisfy the rapid decay estimate $h_\xi(\mu) \ll_M \|\Re\mu - \xi\|^{-M}$. Applying this on each ball gives

$$\sum_{\Re\xi_i \in B(\mu_n)} |\widehat{\omega_S}(\psi_i)|^2 |\mathcal{P}_H(\psi_i)|^2 h_\xi(\xi_i) \ll_M \|\mu_n - \xi\|^{-M} \sum_{\Re\xi_i \in B(\mu_n)} |\widehat{\omega_S}(\psi_i)|^2 |\mathcal{P}_H(\psi_i)|^2.$$

Applying (16) and summing over n, we obtain

$$\mathfrak{S} \ll_M \Pi_H \omega_S \omega_S^*(1) \sum_n \|\mu_n - \xi\|^{-M}\beta(\mu_n) + P^A N^{-\delta} \sum_n \|\mu_n - \xi\|^{-M}(1 + \|\mu_n\|)^{-\delta}\beta(\mu_n).$$

From $\beta(\mu_n) \ll \|\mu_n - \xi\|^k \beta(\xi)$, where k is the number of roots of G counted with multiplicity, we may simplify this to

$$\mathfrak{S} \ll_M \Pi_H \omega_S \omega_S^*(1)\beta(\xi) \sum_n \|\mu_n - \xi\|^{-M} + \beta(\xi)P^A N^{-\delta} \sum_n \|\mu_n - \xi\|^{-M}(1 + \|\mu_n\|)^{-\delta}.$$

The first sum is $\ll_M Q^{-M}$. We bound the second sum by breaking it into $\|\mu_n\| \leqslant \|\xi\|/2$ and the complement. The first sum is $\ll (1 + \|\xi\|)^{-M}$, and the second is $\ll Q^{-M}(1 + \|\xi\|)^{-\delta}$. Both of these are dominated by $(1 + \|\xi\|)^{-\delta}$, which finally shows that

$$\mathfrak{S} \ll_M \Pi_H \omega_S \omega_S^*(1)\beta(\xi)Q^{-M} + P^A N^{-\delta}\beta(\xi)(1 + \|\xi\|)^{-\delta},$$

as desired. □

We combine Lemma 4.5 (taking any $M > 0$ and large enough Q) and (15) to obtain

$$\mathrm{Vol}_N \sum_{\|\Re\xi_i - \xi\| \leqslant Q} |\widehat{\omega_S}(\psi_i)|^2 |\mathcal{P}_H(\psi_i)|^2 h_\xi(\xi_i) \asymp \Pi_H \omega_S \omega_S^*(1)\beta(\xi) + O(P^A N^{-\delta}\beta(\xi)(1 + \|\xi\|)^{-\delta}).$$

We now make use of the critical assumptions that G is H-large, and $\nu \in X_*(T_H)$ satisfies (3), to bound the right-hand side from below.

LEMMA 4.6. *If $\nu \in X_*(T_H)$ satisfies (3) then $\Pi_H \omega_S \omega_S^*(1) \gg_\epsilon P^{2-\epsilon}$.*

976

Proof. Note that for any $v \in \mathcal{P}$ and $\nu \in X_*(T_H)$ we have

$$q_v^{\|\nu\|^*} \int_{H_v} \tau(v, \nu)(x)\, dx = \mathrm{vol}(H_v \cap K_v \nu(\varpi_v) K_v) \geqslant \mathrm{vol}(K_{H,v} \nu(\varpi_v) K_{H,v}) \gg q_v^{2\|\nu\|_H^*},$$

where we have used our assumption that T, and hence T_H, is split at v, and applied the lower bound from Lemma 2.3. If $\nu \in X_*(T_H)$ satisfies (3), then so does $-\nu$, and applying the above bound with $\pm\nu$, we get $\int_{H(F_S)} \tau(v, \nu) \tau(w, \nu)^* \gg 1$ if $v \neq w$. Summing over $v, w \in S$ yields the claim. $\qquad\square$

We deduce from the above lemma that

$$\mathrm{Vol}_N \sum_{\|\Re\xi_i - \xi\| \leqslant Q} |\widehat{\omega_S}(\psi_i)|^2 |\mathcal{P}_H(\psi_i)|^2 h_\xi(\xi_i) \gg_\epsilon P^{2-\epsilon} \beta(\xi)(1 + O(P^A N^{-\delta}(1 + \|\xi\|)^{-\delta})).$$

The bound $h_\xi(\xi_i) \ll 1$ completes the proof of (8).

5. The amplified relative trace formula

We now prove Proposition 3.2. Recall the notation introduced in § 3.2. For $\phi = \mathbf{1}_N^S \otimes k_S \otimes k_\xi$ we let

$$K(x, y) = \sum_{\gamma \in G(F)} \phi(x^{-1}\gamma y) \quad \text{and} \quad K_H(x, y) = \sum_{\gamma \in H(F)} \phi(x^{-1}\gamma y).$$

Integrating the spectral expansion

$$K(x, y) = \mathrm{Vol}_N \sum_{i \geqslant 0} \widehat{k_S \otimes k_\xi}(\psi_i) \psi_i(x) \overline{\psi_i(y)}$$

over $[H] \times [H]$, we obtain

$$\int_{[H] \times [H]} K(x, y)\, dx\, dy = \mathrm{Vol}_N \sum_{i \geqslant 0} \widehat{k_S \otimes k_\xi}(\psi_i) |\mathcal{P}_H(\psi_i)|^2.$$

On the other hand, by unfolding we have

$$\int_{[H] \times [H]} K_H(x, y)\, dx\, dy = \mathrm{vol}([H]) \int_{H(\mathbb{A})} \phi(x)\, dx = \mathrm{vol}([H]) \Pi_H k_S(1) k_\xi(1).$$

We have used the fact that $K^S(N) K_H^S \cap H(\mathbb{A}_f^S) = K_H^S$ and the volume of this is $\mathrm{vol}\, K_H^S = 1$.

For the remaining terms, first observe that if $x, y \in H(\mathbb{A})$ we have

$$\#\{G(F) \cap \mathrm{supp}(\phi(x^{-1} \cdot y))\} = \#\{x^{-1}G(F)y \cap \mathrm{supp}(\phi)\} \ll \#(\mathrm{supp}\, k_S / K_S), \qquad (17)$$

uniformly in x, y and N. Note that the left-hand side of this bound only depends on the image of x and y in $[H]$, so we may assume that they lie in a fixed compact set. To prove (17), let $g, g' \in x^{-1}G(F)y \cap \mathrm{supp}(\phi)$ with $g = x^{-1}\gamma y$, $g' = x^{-1}\gamma' y$, and suppose that g, g' lie in the same coset in $\mathrm{supp}\, k_S / K_S$. Then $g_\infty^{-1} g'_\infty$ lies in a fixed compact set and $g_f^{-1} g'_f \in K$, so that $g^{-1}g'$ also lies in a fixed compact set. We have $g^{-1}g' = y^{-1}\gamma^{-1}\gamma' y$, so that $\gamma^{-1}\gamma'$ lies in a fixed compact set, and so if g is fixed there are only finitely many possibilities for g'. Therefore the map $x^{-1}G(F)y \cap \mathrm{supp}(\phi) \to \mathrm{supp}\, k_S / K_S$ has $O(1)$ fibres, which proves (17). Using this, we simply estimate

$$\int_{[H] \times [H]} \sum_{\gamma \in G(F) - H(F)} \phi(x^{-1}\gamma y)\, dx\, dy$$

with the pointwise bounds of Corollary 5.2.

5.1 Bounding the off-diagonal contributions

In this section we establish Corollary 5.2, which was used in the proof of Proposition 3.2 above. It is based on the following Diophantine lemma which shows, roughly speaking, that any $\gamma \in G(F) - H(F)$ cannot be too close to H_v for various v. We shall use the notation introduced in § 2.3.

LEMMA 5.1. *If $C > 0$ is given, there are $A, C_1, C_2 > 0$ such that the following properties hold for any $\gamma \in G(F) - H(F)$:*

(i) $d(\gamma, H_{v_0}) \geqslant C_1 \|\gamma\|_f^{-A}$;

(ii) *if $(N, D) = 1$ is such that $N > C_2 \|\gamma\|_f^A$, and $d(\gamma, H_{v_0}) < C$, then there is a place $v | N$ such that $\gamma_v \notin K(N)_v K_{H,v}$.*

Proof. We consider H and G as subvarieties of F^{d^2} via the embedding ρ. Let $p_1, \ldots, p_k \in \mathcal{O}[1/D, x_1, \ldots, x_{d^2}]$ be a set of defining polynomials for H that are integral over $\mathcal{O}[1/D]$. Now if $\gamma \in G(F) - H(F)$ then $p_i(\gamma) \neq 0$ for some i.

There are $A, C_3 > 0$ such that for all $\gamma \in G(F)$ we have $\prod_{v \nmid \infty} |p_i(\gamma)|_v \leqslant C_3 \|\gamma\|_f^A$. For $\gamma \in G(F) - H(F)$ the product formula applied to $p_i(\gamma) \in F^\times$ implies that the archimedean norms satisfy $\prod_{v | \infty} |p_i(\gamma)|_v \geqslant \|\gamma\|_f^{-A}/C_3$. Because G_v is compact for all $v | \infty$ other than v_0, $|p_i(\gamma)|_v$ is bounded for all such v. But then we have $|p_i(\gamma)|_{v_0} \geqslant C_1 \|\gamma\|_f^{-A}$ for some $C_1 > 0$, and so $d(\gamma, H_{v_0})$ satisfies the same bound. This establishes (i).

As above, there are $A_1, C_3 > 0$ such that $\prod_{v \nmid \infty N} |p_i(\gamma)|_v \leqslant C_3 \|\gamma\|_f^{A_1}$ for all $\gamma \in G(F)$. Moreover, we have $|p_i(\gamma)|_v \ll 1$ for $v | \infty$ by our assumption $d(\gamma, H_{v_0}) < C$. Now p_i descends to a map $K_v/K(N)_v \to \mathcal{O}_v/N\mathcal{O}_v$ which is trivial on $K_{H,v}$. Thus if $\gamma_v \in K(N)_v K_{H,v}$ then $|p_i(\gamma)|_v \leqslant |N|_v$. If we then suppose that $\gamma_v \in K(N)_v K_{H,v}$ for all $v | N$ it would follow that $\prod_{v | N} |p_i(\gamma)|_v \leqslant N^{-|F:\mathbb{Q}|}$. But if $N > C_2 \|\gamma\|_f^A$ for suitable $A, C_2 > 0$, we obtain a contradiction by again applying the product formula to $p_i(\gamma) \in F^\times$. This establishes (ii). $\square$

COROLLARY 5.2. *There is $A > 0$ such that the following statement holds. Let N, S and ξ be as in § 3.2. Let $k_S \in \mathcal{H}_S$ satisfy (P_S) and $k_\xi \in \mathcal{H}_\infty$ satisfy (P_∞). Put $\phi = \mathbf{1}_N^S \otimes k_S \otimes k_\xi$. Then, for all $\gamma \in G(F) - H(F)$ and all $x, y \in H(\mathbb{A})$, we have*

$$\phi(x^{-1}\gamma y) \ll \beta(\xi)(1 + \|\xi\|)^{-1/4} N^{-1/4} R^A \|k_S\|_\infty.$$

Proof. Let $\Omega_H \subset H(\mathbb{A})$ be a compact set containing a fundamental domain for $[H]$. We assume that $\Omega_H = \Omega_{H, D\infty} \times \prod_{v \nmid D\infty} K_{H,v}$ after possibly enlarging D. Because $G(F) - H(F)$ is bi-invariant under $H(F)$, we may assume that $x, y \in \Omega_H$.

We may also clearly assume that $\phi(x^{-1}\gamma y) \neq 0$. It then follows from Property (P_S) that $\|x^{-1}\gamma y\|_S \leqslant R$; in fact we have $\|x^{-1}\gamma y\|_f \leqslant R$, using the condition $x^{-1}\gamma y \in K^S(N)K_H^S \subset K^S$. When combined with $x, y \in \Omega_H$, this gives $\|\gamma\|_f \ll R$. We may now apply part (i) of Lemma 5.1, to find that $d(\gamma, H_{v_0}) \gg R^{-A}$. All together, since $x, y \in H_{v_0}$, we deduce that $d(x^{-1}\gamma y, H_{v_0}) \gg R^{-A}$. Similarly, from $k_\xi(x^{-1}\gamma y) \neq 0$ and Property (P_∞) it follows that $d(x^{-1}\gamma y, H_{v_0}) < 1$, and hence $d(\gamma, H_{v_0}) \ll 1$.

Suppose that $1 + \|\xi\| \geqslant N$. We have $(k_S \otimes k_\xi)(x^{-1}\gamma y) \leqslant \|k_S\|_\infty k_\xi(x^{-1}\gamma y)$. We then combine (P_∞) with $d(x^{-1}\gamma y, H_{v_0}) \gg R^{-A}$ to get

$$k_\xi(x^{-1}\gamma y) \ll \beta(\xi)(1 + \|\xi\| d(x^{-1}\gamma y, H_{v_0}))^{-1/2}$$
$$\ll \beta(\xi)(1 + \|\xi\| R^{-A})^{-1/2}$$

$$\ll \beta(\xi)(1 + \|\xi\|)^{-1/2} R^{A/2}$$
$$\ll \beta(\xi)(1 + \|\xi\|)^{-1/4} N^{-1/4} R^{A/2},$$

which completes the proof in this case.

Now suppose that $1 + \|\xi\| < N$. Because $\|\gamma\|_f \ll R$, part (ii) of Lemma 5.1 implies that there are $A, C > 0$ such that if $N > CR^A$, then there is a place $v|N$ for which $\gamma_v \notin K(N)_v K_{H,v}$. Because $x, y \in \Omega_H$, we have $x, y \in K_{H,v}$, and so $x^{-1}\gamma y \notin K(N)_v K_{H,v}$. It follows that if $N > CR^A$, then $\phi(x^{-1}\gamma y) = 0$. We may rephrase this as saying that

$$\phi(x^{-1}\gamma y) \leqslant \|k_S k_\xi\|_\infty N^{-1} C R^A \ll \|k_S\|_\infty \beta(\xi) N^{-1} R^A,$$

and the bound $N^{-1} \leqslant N^{-1/2}(1 + \|\xi\|)^{-1/2}$ completes the proof. $\square$

6. The amplified trace formula

We now prove Theorem 3.1, our trace formula asymptotic with uniform error term. Throughout this section, we can and will relax our condition that G_{v_0} is $\mathbb{R}$-almost simple to the condition that G is F-almost simple.

Our proof relies crucially on recent work of Shin and Templier and of Cluckers, Gordon and Halupczok [ST16] on bounding centralizer volumes and p-adic orbital integrals, as well as work of Finis and Lapid [FL18] bounding intersection volumes of conjugacy classes with congruence subgroups. We must supply our own bounds on archimedean orbital integrals; these are proven in § 7.

6.1 Canonical and Tamagawa measures

If G is a general connected reductive group over F, Gross [Gro97, (1.5)] attaches to G an Artin–Tate motive

$$M_G = \bigoplus_{d \geqslant 1} M_{G,d}(1 - d)$$

with coefficients in $\mathbb{Q}$. Here $M_{G,d}$ is an Artin motive and $(1 - d)$ denotes the Tate twist. We let $\epsilon(M_G)$ be the ϵ-factor of this motive, which is given by

$$\epsilon(M_G) = |d_F|^{\dim G/2} \prod_{d \geqslant 1} \mathcal{N}_{F/\mathbb{Q}}(\mathfrak{f}(M_{G,d}))^{d-1/2},$$

where $\mathfrak{f}(M_{G,d})$ denotes the conductor of the Artin motive $M_{G,d}$ (see [Gro97, (9.8)]). We let $L(M_{G_v}^\vee(1))$ denote the L-function of the local motive $M_{G_v}^\vee(1)$, and $L(M_G^\vee(1))$ and $\Lambda(M_G^\vee(1))$ denote the finite and completed L-functions of $M_G^\vee(1)$. Then $L(M_{G_v}^\vee(1))$ is a positive real number, and $L(M_G^\vee(1))$ and $\Lambda(M_G^\vee(1))$ are finite if $Z(G)$ does not contain an F-split torus (see [Gro97, Proposition 9.4]). From now on we shall assume that G satisfies this condition.

In [Gro97, § 11] Gross defines a canonical measure $|\omega_{G_v}|$ on G_v at any place of F. We define $\mu_{G,v}^{\mathrm{can}} = L(M_{G_v}^\vee(1)) \cdot |\omega_{G_v}|$ as in [Gro97]. When v is finite and G is unramified at v, $\mu_{G,v}^{\mathrm{can}}$ assigns volume 1 to a hyperspecial subgroup of $G(F_v)$, and so we can define the measure $\mu_G^{\mathrm{can}} = \prod_v \mu_{G,v}^{\mathrm{can}}$ on $G(\mathbb{A})$.

Now let ω be a non-zero differential form of top degree on G defined over F. For each v, one may associate with ω a Haar measure $|\omega|_v$ on $G(F_v)$. For almost all v, $L(M_{G_v}^\vee(1)) \cdot |\omega|_v$ assigns volume 1 to a hyperspecial subgroup of $G(F_v)$. Let μ_G^{Tam} be the Tamagawa measure on $G(\mathbb{A})$, which is defined by

$$\mu_G^{\mathrm{Tam}} = \Lambda(M_G^\vee(1))^{-1} |d_F|^{-\dim G/2} \bigotimes_v L(M_{G_v}^\vee(1)) |\omega|_v$$

979

(see [Gro97, (10.2)]) and satisfies [Kot88, p. 629]

$$\mu_G^{\mathrm{Tam}}(G(F)\backslash G(\mathbb{A})) = |\pi_0(Z(\widehat{G})^{\mathrm{Gal}(\overline{F}/F)})||\ker^1(F, Z(\widehat{G}))|^{-1}.$$

The comparison between μ_G^{can} and μ_G^{Tam} is given by [Gro97, Theorem 11.5]

$$\frac{\mu_G^{\mathrm{can}}}{\mu_G^{\mathrm{Tam}}} = \epsilon(M_G)\Lambda(M_G^\vee(1)). \tag{18}$$

6.2 The trace formula
We now assume that G is anisotropic. The trace formula is a distributional identity

$$I_{\mathrm{spec}}(\phi, \mu_G^{\mathrm{can}}) = I_{\mathrm{geom}}(\phi, \mu_G^{\mathrm{can}}),$$

for $\phi \in C_c^\infty(G(\mathbb{A}))$. More precisely,

$$I_{\mathrm{spec}}(\phi, \mu_G^{\mathrm{can}}) = \sum_\pi m(\pi)\mathrm{tr}(\pi(\phi)),$$

where π runs over all irreducible subrepresentations of $L^2(G(F)\backslash G(\mathbb{A}))$ occurring with multiplicity $m(\pi)$, and

$$I_{\mathrm{geom}}(\phi, \mu_G^{\mathrm{can}}) = \sum_{\{\gamma\}} \frac{\mu_{I_\gamma}^{\mathrm{can}}(I_\gamma(F)\backslash I_\gamma(\mathbb{A}))}{|G_\gamma(F) : I_\gamma(F)|} O_\gamma(\phi),$$

where $\{\gamma\}$ runs over all $G(F)$-conjugacy classes, G_γ is the centralizer of γ in G, I_γ is the connected component of G_γ, and

$$O_\gamma(\phi) = \int_{I_\gamma(\mathbb{A})\backslash G(\mathbb{A})} \phi(x^{-1}\gamma x)\, d\mu_\gamma(x).$$

(See, for example, [Kot86, §9.1] for this formulation of the geometric side.) The measure μ_γ above denotes the quotient measure $d\mu_G^{\mathrm{can}}/d\mu_{I_\gamma}^{\mathrm{can}}$.

We shall bound the terms in $I_{\mathrm{geom}}(\phi, \mu_G^{\mathrm{can}})$ using the Weyl discriminant. For any v and $\gamma \in G_v$, this is defined by

$$D_v(\gamma) = |\det(1 - \mathrm{Ad}(\gamma)_{|\mathfrak{g}_v/\mathfrak{g}_{v,\gamma}})|_v,$$

where $\mathfrak{g}_{v,\gamma}$ denotes the centralizer of γ in $\mathfrak{g}_v$. If S is any set of places and $\gamma \in G(F)$, we define $D_S(\gamma) = \prod_{v\in S} D_v(\gamma)$ and $D^S(\gamma) = \prod_{v\notin S} D_v(\gamma)$.

6.3 Bounding volumes
We again let G denote a group satisfying the conditions of Theorem 1.1. Throughout the rest of this section, A, B and C will denote sufficiently large positive constants that may vary from line to line, and will never depend on a choice of place of F.

In preparation for the following result, we introduce some additional notation. Given $\kappa \geqslant 0$ and a finite set of finite places T away from D, we write $U_T^{\leqslant\kappa}$ for the open compact subset $\mathrm{supp}\,\mathcal{H}_T^{\leqslant\kappa}$. Furthermore, we denote by $\mathcal{C}_T^{\leqslant\kappa}$ the set of $G(F)$-conjugacy classes of elements in $G(F) - Z(F)$ whose $G(\mathbb{A})$-conjugacy classes have non-empty intersection with $K^T \cdot U_T^{\leqslant\kappa} \cdot U_\infty$, where $U_\infty = \{g \in G(F_\infty) : d(g, H_{v_0}) < 1\}$.

980

PROPOSITION 6.1. *There exist $A, B > 0$ such that for any $\kappa \geqslant 0$, any finite set of finite places T away from D, and any $\{\gamma\} \in \mathcal{C}_T^{\leqslant \kappa}$, we have*

$$\mu_{I_\gamma}^{\mathrm{can}}(I_\gamma(F)\backslash I_\gamma(\mathbb{A})) \ll q_T^{A\kappa + B}.$$

The implied constant depends only on G.

Proof. Let S_D denote the set of places dividing D. Put $S_\gamma = \{v \notin S_D \cup \infty \cup T : D_v(\gamma) \neq 1\}$. We begin by noting that for any $\{\gamma\} \in \mathcal{C}_T^{\leqslant \kappa}$ we have

$$\mu_{I_\gamma}^{\mathrm{can}}(I_\gamma(F)\backslash I_\gamma(\mathbb{A})) \ll q_{S_\gamma}^B q_T^B, \tag{19}$$

where the implied constant depends only on G. Indeed, from the proof of [ST16, Corollary 6.16] we have

$$\epsilon(M_{I_\gamma}) L(M_{I_\gamma}^\vee(1)) \ll \prod_{v \in \mathrm{Ram}(I_\gamma)} q_v^B \ll q_{S_\gamma}^B q_T^B,$$

where $\mathrm{Ram}(I_\gamma)$ is the set of finite places where I_γ is ramified. Note that the last bound follows from the inclusion $\mathrm{Ram}(I_\gamma) \subset S_D \cup S_\gamma \cup T$, which follows from [Kot86, Proposition 7.1]. Moreover, from the definition of the local archimedean factors in [Gro97, (7.1) and (7.2)], combined with [ST16, Proposition 6.3], we find that $L_\infty(M_{I_\gamma}^\vee(1)) \ll 1$, the implied constant depending only on G. Finally, Corollary 8.12 and Lemma 8.13 of [ST16] imply $\mu_{I_\gamma}^{\mathrm{Tam}}(I_\gamma(F)\backslash I_\gamma(\mathbb{A})) \ll 1$. By combining these estimates with (18) we obtain (19).

Now for $\{\gamma\} \in \mathcal{C}_T^{\leqslant \kappa}$ we have

$$D_v(\gamma) \leqslant \begin{cases} q_v^{A\kappa + B} & \text{for } v \in T, \\ C & \text{for } v \mid \infty, \\ 1 & \text{for } v \notin T \cup \infty. \end{cases}$$

From this and the product formula we deduce that

$$1 = \prod_{v \in T} D_v(\gamma) \prod_{v \in S_D \cup \infty} D_v(\gamma) \prod_{v \in S_\gamma} D_v(\gamma) \ll q_T^{A\kappa + B} q_{S_\gamma}^{-1}, \tag{20}$$

since $D_v(\gamma) \leqslant q_v^{-1}$ for every $v \in S_\gamma$. Inserting this into (19) gives the proposition. □

6.4 Bounding adelic orbital integrals

We shall prove Theorem 3.1 by inserting test functions of the form $\phi = \mathbf{1}_N^T \otimes k_T \otimes k_\xi$ into the trace formula, where $k_T \in \mathcal{H}_T$, $k_\xi \in \mathcal{H}_\infty$, and $\mathbf{1}_N^T$ was defined in (5). We now estimate the orbital integrals of these functions.

PROPOSITION 6.2. *There are constants $A, B, \delta, \eta > 0$ such that the following statement holds. Let T be a finite set of finite places away from N and D'. Let $\xi \in \mathfrak{a}^*$ and $\kappa \geqslant 0$. Then, for $k_T \in \mathcal{H}_T^{\leqslant \kappa}$, any $k_\xi \in \mathcal{H}_\infty$ satisfying (P_∞), and any $\gamma \in G(F) - Z(F)$, we have*

$$O_\gamma(\mathbf{1}_N^T \otimes k_T \otimes k_\xi) \ll N^{-\delta}\beta(\xi)(1 + \|\xi\|)^{-\eta} q_T^{A\kappa + B}\|k_T\|_\infty,$$

where the implied constant depends only on G, θ and K.

981

Proof. We may write k_T as a linear combination of at most $q_T^{A\kappa+B}$ terms of the form $\bigotimes_{v \in T} 1_{K_v \mu_v(\varpi_v) K_v} \in \mathcal{H}_T^{\leqslant \kappa}$, whose sup norms are all bounded by $\|k_T\|_\infty$. We may therefore assume that k_T is a multiple of $\bigotimes_{v \in T} 1_{K_v \mu_v(\varpi_v) K_v}$. This assumption implies that the orbital integral factorizes as $\prod_v O_\gamma(\phi_v)$, where for any $\gamma_v \in G(F_v)$ we have

$$O_{\gamma_v}(\phi_v) = \int_{I_{\gamma_v}(F_v) \backslash G(F_v)} \phi_v(x_v^{-1} \gamma_v x_v) \, d\mu_{\gamma,v}(x_v)$$

and $\mu_{\gamma,v} = \mu_{G,v}^{\mathrm{can}}/\mu_{I_{\gamma_v},v}^{\mathrm{can}}$. It therefore suffices to work place by place.

In [ST16, Theorem 14.1] it is shown that, if we choose D' to be a sufficiently large (ineffective) multiple of D depending only on G, then

$$|O_\gamma(k_T)| \leqslant q_T^{A\kappa+B} D_T(\gamma)^{-1/2} \|k_T\|_\infty. \tag{21}$$

We may prove the following bound for the integral at infinity using the results of § 7.

LEMMA 6.3. *For any $0 < \eta < 1/2$, we have the bound*

$$O_\gamma(k_\xi) \ll \beta(\xi)(1 + \|\xi\|)^{-\eta} D_\infty(\gamma)^{-3/4}.$$

Proof. Let G_{cpt} and G_{cpt,v_0} be the groups associated with G_∞ and G_{v_0} by Definition 7.2. We begin by showing that, as a consequence of Proposition 7.3, the following statement holds. Let $0 < \eta < 1/2$, and let $k_\xi \in \mathcal{H}_\infty$ satisfy property (P_∞). Then

$$O_\gamma(k_\xi) \ll \beta(\xi)(1 + \|\xi\|)^{-\eta} D_\infty(\gamma)^{-3/4} \tag{22}$$

for every semisimple $\gamma \in G_\infty - G_{\mathrm{cpt}}$.

To see how (22) follows from Proposition 7.3, first note that for any non-negative $f \in C_c^\infty(G_\infty^0)$ such that $f = 1$ on the support of k_ξ, we have, using property (P_∞),

$$\begin{aligned} k_\xi(g) &\ll \beta(\xi)(1 + \|\xi\| d(g, H_{v_0}))^{-1/2} f(g) \\ &\leqslant \beta(\xi)(1 + \|\xi\| d(g, H_{v_0}))^{-\eta} f(g) \\ &\ll \beta(\xi)(1 + \|\xi\|)^{-\eta} d(g, H_{v_0})^{-\eta} f(g). \end{aligned}$$

Thus

$$O_\gamma(k_\xi) \ll \beta(\xi)(1 + \|\xi\|)^{-\eta} O_\gamma(f d(\cdot, H_{v_0})^{-\eta}),$$

to which we may apply (27). Indeed, since $Z(G)$ is finite, the function $\|X(g)\|_0$ used there satisfies $\|X(g)\|_0 \ll d(g, H_{v_0})$.

The proof of Lemma 6.3 then follows from (22) once we have verified that a non-central element $\gamma \in G(F)$ cannot lie in G_{cpt}. Because $G_\infty = \prod_{v | \infty} G_v$, we have $G_{\mathrm{cpt}} = G_{\mathrm{cpt},v_0} \times \prod_{v \neq v_0} G_v$, and so it suffices to verify that $\gamma_{v_0} \notin G_{\mathrm{cpt},v_0}$.

In the case at hand, G_{cpt,v_0} is a normal subgroup of G_{v_0} which, as G is semisimple, is compact. If we let H^+ be the fixed point set of θ in G (we write H^+ to distinguish it from its identity component H), then $H_{v_0}^+$ is a maximal compact subgroup of G_{v_0}, and so we have $G_{\mathrm{cpt},v_0} \subset g H_{v_0}^+ g^{-1}$ for all $g \in G_{v_0}$. Thus, if $\gamma_{v_0} \in G_{\mathrm{cpt},v_0}$, we have $\gamma \in g H^+(F) g^{-1}$ for all $g \in G(F)$. The group

$$\bigcap_{g \in G(F)} g H^+ g^{-1}$$

is a proper F-subgroup of G; it is normal, since its normalizer contains the Zariski-dense set $G(F)$. Our assumption that G is F-almost simple implies that it must be contained in $Z(G)$. We therefore have $\gamma \in Z(F)$, a contradiction. $\qquad\square$

982

It remains then to address the size of the orbital integral at finite places away from T. This is provided by the following result. We retain the set-up from the statement of Proposition 6.2; in particular, we recall the notation $\mathbf{1}_N^T$ from (5).

LEMMA 6.4. *There are constants $\delta, A, B, C > 0$ such that for $\{\gamma\} \in \mathcal{C}_T^{\leqslant \kappa}$ we have*

$$O_\gamma(\mathbf{1}_N^T) \ll N^{-\delta} q_T^{A\kappa + B} D^{T\infty}(\gamma)^{-C}. \tag{23}$$

Proof. Recall the sets S_D and S_γ from the proof of Proposition 6.1. Let S_N denote the set of places dividing N.

- If $v \notin S_D \cup S_N \cup S_\gamma \cup T \cup \infty$, then $K_v(N)K_{H_v} = K_v$ and we have $O_\gamma(\mathbf{1}_{K_v}) = 1$; see, for example, [Kot86, Corollary 7.3].
- If $v \in S_D$, then $K_v(N)K_{H_v} = K_v K_{H_v}$ and a general bound of Kottwitz [ST16, Theorem 13.1] establishes that $O_\gamma(\mathbf{1}_{K_v K_{H_v}}) \ll_v D_v(\gamma)^{-1/2}$.
- If $v \in S_N \cup S_\gamma$ we argue as follows.

We begin by estimating the orbital integrals at places $v \in S_N$, for which we will use as a critical input the work of Finis and Lapid [FL18]. As the setting of [FL18] is that of $\mathbb{Q}$-groups, we shall need to restrict scalars from F to $\mathbb{Q}$ to properly invoke their results. We thus let p denote the rational prime over which v lies, and we note that all places lying over p belong to S_N. We set $K_p = \prod_{v|p} K_v$, and $K_p(N)K_{H,p} = \prod_{v|p} K_v(N)K_{H,v}$. Factorize $N = \prod_{p|N} p^{n_p}$ and put $N_p = p^{n_p}$.

Let $\mu_{G,p}^{\mathrm{can}}$ (respectively, $\mu_{I_\gamma,p}^{\mathrm{can}}$) be the product measure on $G_p = \prod_{v|p} G_v$ (respectively, $I_{\gamma,p} = \prod_{v|p} I_{\gamma,v}$), and let $\mu_{\gamma,p}$ be the quotient measure. Let $\mathrm{Ad} : I_{\gamma,p} \backslash G_p \to G_p$ be the map $g \mapsto g^{-1}\gamma g$, which is a proper map as γ is semisimple. By definition,

$$O_\gamma(\mathbf{1}_{K_p(N)K_{H_p}}) = \mu_{\gamma,p}(\mathrm{Ad}^{-1}(K_p(N)K_{H_p})).$$

As $\mathrm{Ad}^{-1}(K_p)$ is compact and right-invariant under K_p, we may break it up into a finite number of orbits $\coprod \overline{g}_i K_p$, where $g_i \in G_p$ and $\overline{g}_i$ denotes its image in $I_{\gamma,p} \backslash G_p$. This gives

$$O_\gamma(\mathbf{1}_{K_p(N)K_{H_p}}) = \sum_i \mu_{\gamma,p}(\overline{g}_i K_p \cap \mathrm{Ad}^{-1}(K_p(N)K_{H_p})).$$

Let $g \in G_p$ be such that $x := g^{-1}\gamma g \in K_p$. The fibers of the reduction map $gK_p \to \overline{g}K_p$ are the left orbits of $I_{\gamma,p} \cap gK_p g^{-1}$, so for any $K_p' \subset K_p$ we have

$$\begin{aligned}
\mu_{\gamma,p}(\overline{g}K_p \cap \mathrm{Ad}^{-1}(K_p')) &= \frac{\mu_{G,p}^{\mathrm{can}}(h \in gK_p : h^{-1}\gamma h \in K_p')}{\mu_{I_\gamma,p}^{\mathrm{can}}(I_{\gamma,p} \cap gK_p g^{-1})} \\
&= \frac{\mu_{G,p}^{\mathrm{can}}(k \in K_p : k^{-1}xk \in K_p')}{\mu_{I_\gamma,p}^{\mathrm{can}}(I_{\gamma,p} \cap gK_p g^{-1})}.
\end{aligned}$$

Taking the quotient of these with K_p' equal to K_p and $K_p(N)K_{H_p}$, and using the fact that $\mu_{G,p}^{\mathrm{can}}(K_p) = 1$ for all $v \notin S_D \cup \infty$, gives

$$\frac{\mu_{\gamma,p}(\overline{g}K_p \cap \mathrm{Ad}^{-1}(K_p(N)K_{H_p}))}{\mu_{\gamma,p}(\overline{g}K_p)} = \mu_{G,p}^{\mathrm{can}}(k \in K_p : k^{-1}xk \in K_p(N)K_{H_p}).$$

One can deduce from [FL18, Propositions 5.10 and 5.11] (see §6.5 below) that there are constants $\epsilon, \delta > 0$ (independent of p) such that if $x \in K_p$ is conjugate to an element of $G(F) - Z(F)$ and satisfies $D_p(x) > N_p^{-\epsilon}$ then

$$\mu_{G,p}^{\mathrm{can}}(k \in K_p : k^{-1}xk \in K_p(N)K_{H,p}) \ll N_p^{-\delta}. \tag{24}$$

(In particular, the implied constant is independent of p.) On the other hand, $\sum_i \mu_{\gamma,p}(\overline{g}_i K_p) = \mu_{\gamma,p}(\mathrm{Ad}^{-1}(K_p)) = O_\gamma(\mathbf{1}_{K_p})$. Noting that $D_p(\gamma) = D_p(g_i^{-1}\gamma g_i)$ for all i, we deduce that for $D_p(\gamma) > N_p^{-\epsilon}$ we have

$$O_\gamma(\mathbf{1}_{K_p(N)K_{H_p}}) \ll N_p^{-\delta} \sum_i \mu_{\gamma,p}(\overline{g}_i K_p) = N_p^{-\delta} O_\gamma(\mathbf{1}_{K_p}). \tag{25}$$

In the remaining range $D_p(\gamma) \leqslant N_p^{-\epsilon}$, we have

$$O_\gamma(\mathbf{1}_{K_p(N)K_{H_p}}) \leqslant O_\gamma(\mathbf{1}_{K_p}) \leqslant N_p^{-\delta} D_p(\gamma)^{-\delta/\epsilon} O_\gamma(\mathbf{1}_{K_p}).$$

Since $D_v(\gamma) \leqslant 1$ for all $v \notin T \cup \infty$, we may combine these as

$$\prod_{v \in S_N} O_\gamma(\mathbf{1}_{K_v(N)K_{H_v}}) \ll N^{-\delta} D_{S_N}(\gamma)^{-C} \prod_{v \in S_N} O_\gamma(\mathbf{1}_{K_v}).$$

Note that we may shrink δ to absorb the implied constant in (25) for p sufficiently large, so that the implied constant above only depends on G and not the number of factors of N.

We now return to the product of orbital integrals over all $v \in S_N \cup S_\gamma$. Recalling that $O_\gamma(\mathbf{1}_{K_v}) = 1$ for $v \notin S_\gamma$, we have just shown that

$$\prod_{v \in S_N} O_\gamma(\mathbf{1}_{K_v(N)K_{H_v}}) \prod_{\substack{v \in S_\gamma \\ v \notin S_N}} O_\gamma(\mathbf{1}_{K_v}) \ll N^{-\delta} D_{S_N}(\gamma)^{-C} \prod_{v \in S_N} O_\gamma(\mathbf{1}_{K_v}) \prod_{\substack{v \in S_\gamma \\ v \notin S_N}} O_\gamma(\mathbf{1}_{K_v})$$

$$= N^{-\delta} D_{S_N}(\gamma)^{-C} \prod_{v \in S_\gamma} O_\gamma(\mathbf{1}_{K_v}).$$

For $v \in S_\gamma$, we proceed as follows. If $v \nmid D'$ we again apply [ST16, Theorem 14.1] to get $O_\gamma(\mathbf{1}_{K_v}) \leqslant q_v^B D_v(\gamma)^{-1/2}$, and if $v \mid D'$ we apply [ST16, Theorem 13.1] to get $O_\gamma(\mathbf{1}_{K_v}) \ll_v D_v(\gamma)^{-1/2}$. Combining these gives

$$\prod_{v \in S_\gamma} O_\gamma(\mathbf{1}_{K_v}) \leqslant q_{S_\gamma}^B D_{S_\gamma}(\gamma)^{-C}.$$

Since $\{\gamma\} \in \mathcal{C}_T^{\leqslant \kappa}$ we may invoke (20) to obtain

$$\prod_{v \in S_\gamma} O_\gamma(\mathbf{1}_{K_v}) \leqslant q_{S_\gamma}^B D_{S_\gamma}(\gamma)^{-C} \ll q_T^{A\kappa+B} D_{S_\gamma}(\gamma)^{-C}.$$

Putting these estimates together completes the proof of Lemma 6.4. $\qquad\square$

Remark 6.5. As the authors point out in [ST16, Remark 7.4], the bound [ST16, Theorem 14.1] is uniform in the place $v \nmid D'$ whereas the bound [ST16, Theorem 13.1] of Kottwitz applies to $v \mid D'$ but is not uniform in v. As we allow the implied constant in Proposition 6.2 to depend on the group, this non-uniformity is not an issue.

Taken together (and using the product rule for the product of Weyl discriminants), the estimates in (21), Lemma 6.3 and Lemma 6.4 imply Proposition 6.2. $\qquad\square$

6.5 The work of Finis and Lapid

We now explain how to extract from [FL18, Propositions 5.10 and 5.11] the bound we stated in (24). Recall the notation $G_p = \prod_{v|p} G_v$, $K_p = \prod_{v|p} K_v$, and so on, from the proof of Lemma 6.4. In what follows, we simplify notation by writing K'_p for $K_p(N)K_{H,p}$. We recall that $N = \prod_{p|N} p^{n_p}$ and that x is taken to lie in K_p.

We first remark that we may assume that x lies in K'_p, for if there is no such representative then the left-hand side of (24) is zero. We then have

$$\mu_{G,p}^{\mathrm{can}}(k \in K_p : k^{-1}xk \in K'_p) = \mu_{G,p}^{\mathrm{can}}(k \in K_p : [k,x] \in K'_p),$$

where $[k,x] := kxk^{-1}x^{-1}$ is the commutator. Finis and Lapid's bounds are stated in terms of the adjoint group G^{ad}. They choose an embedding $\rho_{\mathbb{Q}}^{\mathrm{ad}} : \mathrm{Res}_{F/\mathbb{Q}} G^{\mathrm{ad}} \to \mathrm{GL}(N^{\mathrm{ad}})$, and for every p define $K_p^{\mathrm{ad}} = (\rho_{\mathbb{Q}}^{\mathrm{ad}})^{-1}(\mathrm{GL}(N^{\mathrm{ad}}, \mathbb{Z}_p))$. In [FL18, Definition 5.1], they define the function

$$\phi_{K'_p}(x) = \mathrm{vol}(k \in K_p^{\mathrm{ad}} : [k,x] \in K'_p)$$

for $x \in K_p^{\mathrm{ad}}$, where vol is the probability Haar measure on K_p^{ad}.[2] We first relate the two quantities.

LEMMA 6.6. *Let* $\pi : G \to G^{\mathrm{ad}}$ *be the natural projection map. Then for any* $x \in K_p$,

$$\mu_{G,p}^{\mathrm{can}}(k \in K_p : [k,x] \in K'_p) \ll \phi_{K'_p}(\pi(x)),$$

the implied constant depending only on G.

Proof. We first claim that $\pi(K_p) \subset K_p^{\mathrm{ad}}$ for almost all p. To see this, we choose a $\mathbb{Q}$-embedding $\rho_{\mathbb{Q}} : \mathrm{Res}_{F/\mathbb{Q}} G \to \mathrm{GL}(N')$ with $K_p = \rho_{\mathbb{Q}}^{-1}(\mathrm{GL}(N', \mathbb{Z}_p))$ for almost all p. The embedding

$$\Delta : \mathrm{Res}_{F/\mathbb{Q}} G \xrightarrow{\mathrm{id} \times \pi} \mathrm{Res}_{F/\mathbb{Q}} G \times \mathrm{Res}_{F/\mathbb{Q}} G^{\mathrm{ad}} \xrightarrow{\rho_{\mathbb{Q}} \times \rho_{\mathbb{Q}}^{\mathrm{ad}}} \mathrm{GL}(N' + N^{\mathrm{ad}})$$

is such that $\Delta^{-1}(\mathrm{GL}(N' + N^{\mathrm{ad}}, \mathbb{Z}_p)) \subset K_p$ is hyperspecial for almost all p, whence the claim. Increasing D if necessary, we may then assume that $\pi(K_p) \subset K_p^{\mathrm{ad}}$ for all $p \nmid D$.

For convenience, let us write

$$U = \{k \in K_p : [k,x] \in K'_p\} \quad \text{and} \quad V = \{k \in K_p^{\mathrm{ad}} : [k,\pi(x)] \in K'_p\}.$$

We want to show that $\mu_{G,p}^{\mathrm{can}}(U) \ll \mathrm{vol}(V)$. From the inclusion $\pi(K_p) \subset K_p^{\mathrm{ad}}$ it follows that $\pi(U)$ is contained in both $\pi(K_p)$ and V, yielding

$$\mu_{G,p}^{\mathrm{can}}(U) \leqslant \pi_* \mu_{G,p}^{\mathrm{can}}(\pi(U)) \leqslant \pi_* \mu_{G,p}^{\mathrm{can}}(V \cap \pi(K_p)).$$

Now, on $\pi(K_p)$ the push-forward measure $\pi_* \mu_{G,p}^{\mathrm{can}}$ is just $[K_p^{\mathrm{ad}} : \pi(K_p)]\mathrm{vol}$. Indeed, since K_p is maximal compact we have $\pi^{-1}(K_p^{\mathrm{ad}}) = K_p$ so that $\pi_* \mu_{G,p}^{\mathrm{can}}(\pi(K_p)) = \mu_{G,p}^{\mathrm{can}}(K_p) = 1$, while $\mathrm{vol}(\pi(K_p)) = [K_p^{\mathrm{ad}} : \pi(K_p)]^{-1}$. We deduce that

$$\pi_* \mu_{G,p}^{\mathrm{can}}(V \cap \pi(K_p)) = [K_p^{\mathrm{ad}} : \pi(K_p)]\mathrm{vol}(V \cap \pi(K_p)) \leqslant [K_p^{\mathrm{ad}} : \pi(K_p)]\mathrm{vol}(V).$$

Finally, $\pi(K_p) = K_p^{\mathrm{ad}} \cap \pi(G_p)$ so that $[K_p^{\mathrm{ad}} : \pi(K_p)] \leqslant [G_p^{\mathrm{ad}} : \pi(G_p)]$, which is bounded in terms of G. $\qquad\square$

---

[2] Note that Finis and Lapid consider the commutator as a map $G^{\mathrm{ad}} \times G^{\mathrm{ad}} \to G$ in this definition.

We now fix a $\mathbb{Z}$-lattice Λ in $\mathrm{Res}_{F/\mathbb{Q}}(\mathfrak{g})$ such that $\Lambda \otimes \widehat{\mathbb{Z}}$ is K^{ad}-stable. For $y \in G_p^{\mathrm{ad}}$ define $\lambda_p(y)$ as in [FL18, Definition 5.2], namely,

$$\lambda_p(y) = \max\{n \in \mathbb{Z} \cup \{\infty\} : (\mathrm{Ad}(y) - 1)\mathrm{Pr}_{\mathfrak{h}}(\Lambda \otimes \mathbb{Z}_p) \subset p^n(\Lambda \otimes \mathbb{Z}_p) \text{ for some } \mathfrak{h} \neq 0\},$$

where $\mathfrak{h}$ ranges over the non-trivial $\mathbb{Q}_p$-ideals of $\mathfrak{g}_p = \prod_{v|p} \mathfrak{g}_v$ and $\mathrm{Pr}_{\mathfrak{h}}$ denotes the corresponding projection $\mathfrak{g}_p \to \mathfrak{h} \subset \mathfrak{g}_p$.

We first show how $\lambda_p(y)$ controls the size of $\phi_{K_p'}(y)$, using a combination of Propositions 5.10 and 5.11[3] of [FL18]. The argument is already present in [FL18, § 5.2] in the deduction of the global result [FL18, Theorem 5.3] from these two local results.

LEMMA 6.7. *For every $\epsilon > 0$ small enough there is $\delta > 0$ such that whenever $y \in K_p^{\mathrm{ad}}$ satisfies $\lambda_p(y) < \epsilon n_p$ we have $\phi_{K_p'}(y) \ll_G p^{-\delta n_p}$.*

Proof. From [FL18, Proposition 5.11] it follows that there are positive constants $a, b > 0$ and $c \geqslant 0$ (depending only on G) such that $\phi_{K_p'}(y) \leqslant p^{a(c+\lambda_p(x)-bn_p)}$. The constant c renders this bound useless for small n_p. Taking ϵ small enough to satisfy $0 < \epsilon < (c+1)^{-1}b$, we shall apply this bound only in the range $n_p > \epsilon^{-1}$; we obtain $\phi_{K_p'}(y) \leqslant p^{-a(b-\epsilon(c+1))n_p}$. In the remaining range $n_p \leqslant \epsilon^{-1}$ we see that $\lambda_p(y) < 1$ so that in fact $\lambda_p(y) = 0$. In this case [FL18, Proposition 5.10] ensures that $\phi_{K_p'}(y) \ll_G p^{-1} \leqslant p^{-\epsilon n_p}$. Taking $\delta = \min\{\epsilon, a(b - \epsilon(c+1))\}$, we establish the claim. $\square$

It remains to understand the relation between $D_p(y)$ and $\lambda_p(y)$, for $y \in K_p^{\mathrm{ad}}$. For this, the following lemma will be helpful.

LEMMA 6.8. *If $\gamma \in G(F) - Z(F)$, $\mathrm{Ad}(\gamma)$ acts non-trivially on every non-trivial $\mathbb{Q}_p$-ideal of $\mathfrak{g}_p$.*

Proof. Suppose that $\mathfrak{h} \subset \mathfrak{g}_p$ is a non-trivial $\mathbb{Q}_p$-ideal on which $\mathrm{Ad}(\gamma)$ acts trivially. Let $v|p$ be a place for which $\mathrm{Pr}_{\mathfrak{g}_v}\mathfrak{h} \neq 0$, and let $\mathfrak{h}_v$ be the F_v-linear span of $\mathrm{Pr}_{\mathfrak{g}_v}\mathfrak{h}$. Then $\mathfrak{h}_v$ is a non-trivial ideal of $\mathfrak{g}_v$ on which $\mathrm{Ad}(\gamma)$ acts trivially, or equivalently such that $\mathfrak{h}_v \subset \mathfrak{g}_{v,\gamma}$. Since G is F-almost simple we have $\bigcap_{g \in G(F)} gG_\gamma g^{-1} \subset Z(G)$ as in Lemma 6.3. This implies that $\bigcap_{g \in G(F)} \mathrm{Ad}(g)\mathfrak{g}_\gamma = 0$, and hence $\bigcap_{g \in G(F)} \mathrm{Ad}(g)\mathfrak{g}_{v,\gamma} = 0$. However, this is a contradiction as $\mathrm{Ad}(g)\mathfrak{h}_v = \mathfrak{h}_v$ for all $g \in G_v$. $\square$

Finally, we complete the proof of (24). Note that we have assumed that the $x \in K_p$ to which we apply (24) are G_p-conjugate to an element in $G(F) - Z(F)$. We may therefore apply the conclusion of Lemma 6.8 to x, and thus to $y = \pi(x) \in K_p^{\mathrm{ad}}$.

LEMMA 6.9. *Let $y \in K_p^{\mathrm{ad}}$ be semisimple. Assume that $\mathrm{Ad}(y)$ is non-trivial on every non-trivial $\mathbb{Q}_p$-ideal of $\mathfrak{g}_p$ and $D_p(y) > p^{-\epsilon n_p}$. Then $\lambda_p(y) < \epsilon n_p$.*

Proof. Let $l = \lceil \epsilon n_p \rceil$. Now if $\lambda_p(y) \geqslant \epsilon n_p$ then, by definition, there is some non-trivial $\mathbb{Q}_p$-ideal $\mathfrak{h}$ of $\mathfrak{g}_p$ such that

$$(\mathrm{Ad}(y) - 1)\mathrm{Pr}_{\mathfrak{h}}(\Lambda \otimes \mathbb{Z}_p) \subset p^l(\Lambda \otimes \mathbb{Z}_p).$$

---

[3] Note that [FL18, Propositions 5.10 and 5.11] assume that G is simply connected. This assumption can be dropped for those subgroups not containing the intersection of K_p with G_p^+. Here, G_p^+ denotes the image in G_p of the $\mathbb{Q}_p$-points of the simply connected cover of G_p. That our subgroups $K_p' = K_p(N)K_{H,p}$ satisfy this condition (for p large enough with respect to G and F) can be seen from comparing indices. The subgroups $K_p(N)K_{H,p}$ have indices growing like a power of p, whereas those containing $K_p \cap G_p^+$ are of index bounded in terms of G and F.

As $\mathrm{Ad}(y)$ preserves $\mathfrak{h}$, this implies that

$$(\mathrm{Ad}(y) - 1)\mathrm{Pr}_{\mathfrak{h}}(\Lambda \otimes \mathbb{Z}_p) \subset p^l \mathrm{Pr}_{\mathfrak{h}}(\Lambda \otimes \mathbb{Z}_p).$$

This implies that all eigenvalues of $\mathrm{Ad}(y) - 1$, considered as an endomorphism of the $\mathbb{Q}_p$ vector space $\mathfrak{h}$, must have p-adic valuation at least l. Moreover, as y is semisimple and $\mathrm{Ad}(y)$ is non-trivial on $\mathfrak{h}$, one of these eigenvalues must be non-zero. As $D_p(y) = |\det(1 - \mathrm{Ad}(y))_{\mathfrak{g}_p/\mathfrak{g}_{p,y}}|_p$, we therefore have $D_p(y) \leqslant p^{-\epsilon n_p}$, a contradiction. $\qquad\square$

6.6 Proof of Theorem 3.1

We recall the test function ϕ from the statement of Proposition 6.2, and the set $\mathcal{C}_T^{\leqslant\kappa}$ from §6.3. Then

$$I_{\mathrm{spec}}(\phi, \mu_G^{\mathrm{can}}) = \mathrm{Vol}_N \sum_{i \geqslant 0} \widehat{k_T \otimes k_\xi}(\psi_i),$$

the sum ranging over an orthonormal basis of Hecke–Maass forms for Y_N, and

$$I_{\mathrm{geom}}(\phi, \mu_G^{\mathrm{can}}) = \mu_G^{\mathrm{can}}([G])k_T(1)k_\xi(1) + \sum_{\{\gamma\} \in \mathcal{C}_T^{\leqslant\kappa}} \frac{\mu_{I_\gamma}^{\mathrm{can}}(I_\gamma(F) \backslash I_\gamma(\mathbb{A}))}{|G_\gamma(F) : I_\gamma(F)|} O_\gamma(\phi).$$

Here we have used the hypothesis on K from §2.4 that $Z(F) \cap K = \{e\}$. Now by [ST16, Corollary 8.10] we have $|\mathcal{C}_T^{\leqslant\kappa}| \ll q_T^{A\kappa+B}$. From this and Propositions 6.1 and 6.2 we find

$$I_{\mathrm{geom}}(\phi, \mu_G^{\mathrm{can}}) = \mu_G^{\mathrm{can}}([G])k_T(1)k_\xi(1) + O(q_T^{A\kappa+B}N^{-\delta}\beta(\xi)(1 + \|\xi\|)^{-\eta}\|k_T\|_\infty),$$

as desired. This completes the proof of Theorem 3.1.

7. Bounds for real orbital integrals

The aim of this section is to prove Proposition 7.3, which establishes the uniform bounds on real orbital integrals that were used in the proof of the global bounds of Proposition 6.2.

7.1 Notation

We adopt the following notation in this section.

- G is a connected reductive group over $\mathbb{R}$ with real Lie algebra $\mathfrak{g}$.
- θ is a Cartan involution of G.
- K is the fixed point set of θ, so that K is a maximal compact subgroup of G.
- $\mathfrak{g} = \mathfrak{p} + \mathfrak{k}$ is the Cartan decomposition associated with θ.
- A and A_G are maximal split tori in G and $Z(G)$, respectively. We assume that θ acts by -1 on A and A_G; see, for example, [Bor91, §24.C].
- A^0 is the connected component of A in the real topology.
- $\mathfrak{a}$ and $\mathfrak{a}_G$ are the Lie algebras of A and A_G.
- W is the Weyl group of $\mathfrak{a}$. We recall [Bor91, §24.C] that this is equal to $N_G(A)/Z_G(A)$, $N_K(A)/Z_K(A)$, and the group generated by reflections in the roots of $\mathfrak{a}$ in $\mathfrak{g}$.
- $\mathfrak{a}^+$ is a choice of open Weyl chamber in $\mathfrak{a}$.
- $\langle\cdot,\cdot\rangle$ is an Ad-invariant bilinear form on $\mathfrak{g}$ that is positive on $\mathfrak{p}$ and negative on $\mathfrak{k}$.
- $\langle\cdot,\cdot\rangle_+$ is the positive definite inner product on $\mathfrak{g}$ given by $\langle u, v\rangle_+ = -\langle\theta u, v\rangle$. We let $\|\cdot\|$ be the associated norm.

987

For a semisimple element $\gamma \in G$ let G_γ be its centralizer, and let I_γ be the neutral component of G_γ. The Lie algebra of G_γ is denoted by $\mathfrak{g}_\gamma$. By [Kna02, Theorem 7.39], G has a Cartan decomposition $G = KA^0K$, and any $g \in G$ may be written as $g = k_1 e^H k_2$ for a unique $H \in \overline{\mathfrak{a}}^+$. We use this to define a map $X : G \to \overline{\mathfrak{a}}^+$ by $g \in Ke^{X(g)}K$. We let D^G (or D if there is no confusion) denote the Weyl discriminant. A Levi subgroup or parabolic of G will be called semi-standard if it contains A. We will always choose the Levi of a semi-standard parabolic to be semi-standard.

7.2 Orbital integrals of continuous functions

We shall derive Proposition 7.3 from the following result, which bounds $O_\gamma(f)$ for semisimple $\gamma \in G$ and non-singular test functions $f \in C(G)$. It is the archimedean analogue of a result of Kottwitz [ST16, Theorem 13.1] in the p-adic case.

PROPOSITION 7.1. *If $f \in C^\infty(G)$ is bounded and compactly supported modulo centre, then we have $O_\gamma(f) \ll_f D(\gamma)^{-1/2}$ for every semisimple $\gamma \in G$.*

Proof. The stated inequality is proved in [Har57, Theorem 2] for γ regular semisimple. The extension of this bound to all semisimple γ follows immediately from Harish-Chandra's derivative formula [Har66, Lemma 23], which expresses $O_\gamma(f)$ as the derivative of (a quantity closely related to) $D(\sigma)^{1/2}O_\sigma(f)$, for regular semisimple σ in a neighbourhood of γ. The idea is that the latter is bounded independently of γ, and examining the constant that appears there gives the proposition.

We now explain this idea in more detail. We recall that a maximal torus T in a connected reductive group H is called fundamental if the compact part of T has maximum possible dimension. For any semisimple $\gamma \in G$, we may choose a fundamental maximal torus T in I_γ. Note that $\gamma \in I_\gamma$, which implies that $\gamma \in T$. As there are only finitely many possibilities for T up to conjugacy, we may assume that T is fixed. Once T is fixed there are finitely many possibilities for I_γ, so we also assume I_γ is fixed. Note that we need to choose T fundamental in order for (26) below to hold with a non-zero constant c.

We let Δ and Δ_γ be the roots of T in G and I_γ. We choose a system of positive roots Δ^+, and let $\Delta_\gamma^+ = \Delta_\gamma \cap \Delta^+$. We let ρ be a branch of the square root of the modular character, that is, a smooth function on T defined near γ so that $\rho^2 = \prod_{\alpha \in \Delta^+} \alpha$. Let G_{reg} be the set of strongly regular elements in G, which we recall is the set of σ such that G_σ is a torus, and define $T_{\mathrm{reg}} = T \cap G_{\mathrm{reg}}$. Let Ω be a connected component of T_{reg} such that $\gamma \in \overline{\Omega}$, and let $V \subset \Omega$ be the intersection of Ω with an open neighbourhood of γ in T. For $\sigma \in V$, we define

$$F_f(\sigma) = \rho(\sigma) \prod_{\alpha \in \Delta^+} (1 - \alpha(\sigma)^{-1}) \int_{T \backslash G} f(x^{-1} \sigma x)\, dx.$$

For any $\alpha \in \Delta_\gamma$ let $H_\alpha \in \mathrm{Lie}(T)$ be as defined in [Har64, §4]. Define D_I to be the translation-invariant differential operator on T corresponding to $\prod_{\alpha \in \Delta_\gamma^+} H_\alpha$. By [Har66, Lemma 23] (adapted to the real algebraic case), we may show that

$$\lim_{\sigma \to \gamma} D_I F_f(\sigma) = c\rho(\gamma) \prod_{\alpha \in \Delta^+ - \Delta_\gamma^+} (1 - \alpha(\gamma)^{-1}) \int_{I_\gamma \backslash G} f(x^{-1} \gamma x)\, dx, \qquad (26)$$

where $c \neq 0$ depends only on the isomorphism class of I_γ.

988

On the other hand, it is known [Har57, Theorem 2] that $F_f \in C^\infty(V)$, and if X is any translation-invariant differential operator on T, we have

$$\sup_{\sigma \in V} |XF_f(\sigma)| \ll_{X,f} 1.$$

Combined with (26), this gives the proposition. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

7.3 Orbital integrals of singular functions

We now arrive at the main result of this section. We define $\|\cdot\|_0$ to be the seminorm on $\mathfrak{a}$ obtained from the Killing form, which descends to a norm on $\mathfrak{a}/\mathfrak{a}_G$.

DEFINITION 7.2. We define G_{cpt} to be the elements of G such that $\mathrm{Ad}(g)$ is trivial on all the $\mathbb{R}$-simple factors of $\mathfrak{g}$ of non-compact type.

It is clear that G_{cpt} is a normal subgroup of G, and because its image under Ad is compact, G_{cpt} is compact if and only if A_G is trivial.

PROPOSITION 7.3. *Suppose that G is not a torus. Let $0 < \eta < 1/2$ and let $f \in C(G)$ be bounded and compactly supported modulo centre. Then $O_\gamma(f\|X(\cdot)\|_0^{-\eta})$ converges absolutely, and there is a constant $c(\eta, f) > 0$ such that*

$$|O_\gamma(f\|X(\cdot)\|_0^{-\eta})| < c(\eta, f)D(\gamma)^{-3/4} \tag{27}$$

for every semisimple $\gamma \in G - G_{\mathrm{cpt}}$.

Remark 7.4. It may be seen that the bound (27) cannot hold for all semisimple $\gamma \in G$. For instance, if γ lies in a compact normal subgroup of G (and hence in G_{cpt}) then the conjugacy class of γ will be contained in K, and the function $f\|X(\cdot)\|_0^{-\eta}$ will be singular everywhere on the conjugacy class.

We sketch the proof of Proposition 7.3, which will occupy the remaining subsections. In § 7.4 we begin by reducing to the case where $Z(G)$ is anisotropic. This is simple; if A_G is the maximal split torus in $Z(G)$, we simply push the orbital integrals forward to G/A_G. We then divide the proof into two cases, based on the following definition.

DEFINITION 7.5. We say that $\gamma \in G$ is elliptic if it is semisimple and $Z(G_\gamma)$ is anisotropic.

If our semisimple γ is not elliptic, we may choose a non-trivial split torus $S \subset Z(G_\gamma)$ and define M to be the centralizer of S in G. Because we have assumed that $Z(G)$ is anisotropic, M is a proper Levi subgroup satisfying $G_\gamma \subset M$, and we may apply parabolic descent as well as Proposition 7.1. This is executed in § 7.5.

If γ is elliptic, we may assume without loss of generality that $\gamma \in K$. For $\epsilon > 0$, define $K(\epsilon) = \{g \in G : \|X(g)\| < \epsilon\}$. Proving Proposition 7.3 in this case is roughly equivalent to controlling the volume of the conjugacy class of γ that lies inside $K(\epsilon)$, uniformly in ϵ and $D(\gamma)$. This turns out to be equivalent to bounding the set of points in G/K that are moved a distance at most ϵ by the rotation γ. Note that this set will be non-compact if I_γ is non-compact. We do this in § 7.6 using a geometric argument and Proposition 7.1.

7.4 Reduction to the case of $Z(G)$ anisotropic

Let $\overline{G} = G/A_G$. We wish to show that if Proposition 7.3 holds for $\overline{G}$ then it holds for G. We have an exact sequence $1 \to A_G \to G \to \overline{G} \to 1$ of algebraic groups over $\mathbb{R}$, and by Hilbert 90 this gives

an exact sequence on points. If $\gamma \in G$, we denote its image in $\overline{G}$ by $\overline{\gamma}$. We denote the connected centralizer of $\overline{\gamma}$ by $\overline{I}_{\overline{\gamma}}$. There is an exact sequence $1 \to A_G \to I_\gamma \to \overline{I}_{\overline{\gamma}} \to 1$, and so $G \to \overline{G}$ induces a bijection $I_\gamma \backslash G \simeq \overline{I}_{\overline{\gamma}} \backslash \overline{G}$. Moreover, under this bijection we have $\mu_G^{\mathrm{can}}/\mu_{I_\gamma}^{\mathrm{can}} = C \mu_{\overline{G}}^{\mathrm{can}}/\mu_{\overline{I}_{\overline{\gamma}}}^{\mathrm{can}}$ for some $C > 0$, and the function $\|X(\cdot)\|_0$ pushes forward to $\|X(\cdot)\|$. It follows that if $f \in C(G)$ is invariant under A_G, and its reduction $\overline{f}$ lies in $C_c(\overline{G})$, we have $O_\gamma^G(f\|X(\cdot)\|_0^{-\eta}) = C \cdot O_{\overline{\gamma}}^{\overline{G}}(\overline{f}\|X(\cdot)\|^{-\eta})$.

Let $f \in C(G)$ be bounded and compactly supported modulo centre. We may assume without loss of generality that $f \geqslant 0$, by replacing f by its absolute value. Since f is compactly supported modulo centre, we may choose $h \in C(G)$ that is invariant under A_G, satisfies $h \geqslant f$, and has reduction $\overline{h}$ lying in $C_c(\overline{G})$. We have

$$O_\gamma^G(f\|X(\cdot)\|_0^{-\eta}) \leqslant O_\gamma^G(h\|X(\cdot)\|_0^{-\eta}) = C \cdot O_{\overline{\gamma}}^{\overline{G}}(\overline{h}\|X(\cdot)\|^{-\eta}).$$

Because $D(\gamma) = D(\overline{\gamma})$, if Proposition 7.3 holds for $\overline{G}$ then it holds for G.

We may henceforth assume that $Z(G)$ is anisotropic. We may then take the function f of Proposition 7.3 to lie in $C_c(G)$, and work with $\|X(\cdot)\|$ instead of $\|X(\cdot)\|_0$. As in §6, we write $\mu_\gamma = \mu_G^{\mathrm{can}}/\mu_{I_\gamma}^{\mathrm{can}}$.

7.5 The case when γ is not elliptic

We handle the case when γ is not elliptic by the process of parabolic descent, which we now recall. Let $P = MN$ be a semi-standard parabolic subgroup of G. Choose Haar measures on N and K so that dk gives K measure 1, and for which $d\mu_G^{\mathrm{can}} = d\mu_M^{\mathrm{can}} \, dn \, dk$ in Langlands MNK co-ordinates. For functions in $C_c(G)$, the parabolic descent along P is defined by

$$f \in C_c(G) \mapsto f^P \in C_c(M),$$

where

$$f^P(m) = \delta_P^{1/2}(m) \int_N \int_K f(k^{-1}mnk) \, dn \, dk.$$

If $\gamma \in M$, we define $D_M^G(\gamma)$ by choosing a maximal torus $\gamma \in T \subset M$, letting Δ and Δ_M be the roots of T in G and M, and setting $D_M^G(\gamma) = \prod_{\alpha \in \Delta - \Delta_M} |\alpha(\gamma) - 1|$. It may be seen that this is independent of the choice of T. We say that $\gamma \in M$ is (G, M)-regular if $D_M^G(\gamma) \neq 0$. We recall the descent relation between the orbital integrals of f and f^P.

LEMMA 7.6. If $\gamma \in M$ is (G, M)-regular and $f \in C_c(G)$, we have

$$D_M^G(\gamma)^{1/2} O_\gamma(f) = O_\gamma^M(f^P).$$

Proof. Because γ is (G, M)-regular, we have $I_\gamma \subset M$. Let $\mu_\gamma^M = \mu_M^{\mathrm{can}}/\mu_{I_\gamma}^{\mathrm{can}}$. We may parametrize $I_\gamma \backslash G$ in Langlands coordinates as $(I_\gamma \backslash M)NK$, which allows us to write

$$\begin{aligned}
O_\gamma(f) &= \int_{I_\gamma \backslash G} f(x^{-1}\gamma x) \, d\mu_\gamma(x) \\
&= \int_{I_\gamma \backslash M} \int_N \int_K f(k^{-1}n^{-1}m^{-1}\gamma mnk) \, d\mu_\gamma^M(m) \, dn \, dk \\
&= D_M^G(\gamma)^{-1/2} \delta_P(\gamma)^{1/2} \int_{I_\gamma \backslash M} \int_N \int_K f(k^{-1}m^{-1}\gamma mnk) \, d\mu_\gamma^M(m) \, dn \, dk \\
&= D_M^G(\gamma)^{-1/2} O_\gamma^M(f^P),
\end{aligned}$$

as desired. $\qquad \square$

990

We shall need a version of Lemma 7.6 that can be applied to the singular functions $f\|X(\cdot)\|^{-\eta}$. This is Lemma 7.10 below, which we prove by adapting methods from [MT15], where similar statements were established for $A_G\backslash\mathrm{GL}_n(\mathbb{R})$. If $A \in \mathrm{GL}(\mathfrak{g})$, define $^{\mathrm{t}}A$ to be the transpose with respect to $\langle\cdot,\cdot\rangle_+$. As in [MT15, §4.2], we define the functions $\mathcal{L},\mathcal{N} : \mathrm{GL}(\mathfrak{g}) \to \mathbb{R}$ by $\mathcal{L}(A) = \log(\mathrm{tr}(A\,{}^{\mathrm{t}}A)/\dim\mathfrak{g})$ and $\mathcal{N}(A) = \mathrm{tr}(A\,{}^{\mathrm{t}}A)/\dim\mathfrak{g}$.

LEMMA 7.7. *We have $0 \leqslant \mathcal{L}(\mathrm{Ad}(g)) \leqslant 2\|X(g)\|$ for all $g \in G$.*

Proof. We follow [MT15, Lemma 4.2]. It may be seen that $^{\mathrm{t}}\mathrm{Ad}(g) = \mathrm{Ad}(\theta(g))^{-1}$ for $g \in G$. We write $g \in G$ as $k_1 e^{X(g)} k_2$, so that $\mathrm{Ad}(g)\,{}^{\mathrm{t}}\mathrm{Ad}(g) = \mathrm{Ad}(g\theta(g)^{-1}) = \mathrm{Ad}(k_1 e^{2X(g)} k_1^{-1})$. Taking traces gives

$$\mathcal{N}(\mathrm{Ad}(g)) = \mathrm{tr}(\mathrm{Ad}(e^{2X(g)}))/\dim\mathfrak{g} \leqslant e^{2\|X(g)\|},$$

and taking logs gives the upper bound. One sees that $\mathrm{tr}(\mathrm{Ad}(e^{2X(g)})) \geqslant \dim\mathfrak{g}$ by applying the arithmetic mean–geometric mean inequality to opposite pairs of root spaces of $\mathfrak{g}$, which gives the lower bound. $\qquad\square$

We let $P = MN$ be as above, and let $\mathfrak{n}$ be the Lie algebra of N.

LEMMA 7.8. *Let $B_M \subset M$ and $B_{\mathfrak{n}} \subset \mathfrak{n}$ be compact. There is $C > 0$ depending on B_M and $B_{\mathfrak{n}}$ such that for all $m \in B_M$, $V \in B_{\mathfrak{n}}$, we have $\mathcal{N}(\mathrm{Ad}(me^V)) \geqslant 1 + C\|V\|^2$.*

Proof. Choose a basis for $\mathfrak{g}$ subordinate to the root space decomposition $\mathfrak{g} = Z_{\mathfrak{g}}(\mathfrak{a}) \oplus \bigoplus_{\alpha\in\Delta}\mathfrak{g}_\alpha$ that is orthonormal with respect to $\langle\cdot,\cdot\rangle_+$, where Δ are the roots of $\mathfrak{a}$ in $\mathfrak{g}$. It may be seen that the non-zero entries of $\mathrm{Ad}(m)$ and $\mathrm{Ad}(m)(\mathrm{Ad}(e^V) - 1)$ with respect to this basis are disjoint, so we have

$$\mathcal{N}(\mathrm{Ad}(me^V)) = \mathcal{N}(\mathrm{Ad}(m)) + \mathcal{N}(\mathrm{Ad}(m)(\mathrm{Ad}(e^V) - 1)) \geqslant 1 + \mathcal{N}(\mathrm{Ad}(m)(\mathrm{Ad}(e^V) - 1)).$$

It follows that if $\mathcal{N}(\mathrm{Ad}(me^V)) = 1$ then $\mathcal{N}(\mathrm{Ad}(m)(\mathrm{Ad}(e^V)-1)) = 0$, so $V = 0$. By a compactness argument, we may therefore assume that $\|V\|$ is less than an arbitrarily small constant. We deal with small V by applying the Taylor expansion of e^V, which gives

$$\mathrm{Ad}(m)(\mathrm{Ad}(e^V) - 1) = \mathrm{Ad}(m)\mathrm{ad}(V) + O_{B_M}(\|V\|^2),$$
$$\mathrm{Ad}(m)(\mathrm{Ad}(e^V) - 1)\,{}^{\mathrm{t}}[\mathrm{Ad}(m)(\mathrm{Ad}(e^V) - 1)] = \mathrm{Ad}(m)\mathrm{ad}(V)\,{}^{\mathrm{t}}[\mathrm{Ad}(m)\mathrm{ad}(V)] + O_{B_M}(\|V\|^3),$$
$$\mathcal{N}(\mathrm{Ad}(m)(\mathrm{Ad}(e^V) - 1)) = \mathcal{N}(\mathrm{Ad}(m)\mathrm{ad}(V)) + O_{B_M}(\|V\|^3).$$

By compactness, we have $\mathcal{N}(\mathrm{Ad}(m)\mathrm{ad}(V)) \geqslant C > 0$ for all $\|V\| = 1$ and $m \in B_M$. As $\mathcal{N}$ is quadratic, we have $\mathcal{N}(\mathrm{Ad}(m)\mathrm{ad}(V)) \geqslant C\|V\|^2$ for all V and $m \in B_M$, which completes the proof. $\qquad\square$

We next show that the parabolic descent integral of $f\|X(\cdot)\|^{-\eta}$ along P converges and defines a bounded function on M.

LEMMA 7.9. *Let $f \in C_c(G)$, and for $0 < \eta < 1/2$ define $F_\eta = f\|X(\cdot)\|^{-\eta}$. Then:*

  (i) *the integral defining $F_\eta^P(m)$ converges absolutely for any $m \in M$;*
  (ii) *there exists $f_1 \in C_c(M)$ depending only on f and η such that $|F_\eta^P(m)| \leqslant f_1(m)$ for all $m \in M$.*

Proof. We are free to replace f by its absolute value, so that $f \geqslant 0$. It follows that $F_\eta^P(m)$ is defined everywhere as a value in $[0, \infty]$. We may also assume that f is bi-K-invariant. Let $B \subset G$ be a compact set containing $\operatorname{supp}(f)$. There are compact sets $B_M \subset M$ and $B_{\mathfrak{n}} \subset \mathfrak{n}$ depending only on B such that $me^V \in B$ implies $m \in B_M$ and $V \in B_{\mathfrak{n}}$. It follows that $\operatorname{supp}(F_\eta^P) \subset B_M$.

Lemma 7.7 implies that

$$F_\eta^P(m)\delta_P^{-1/2}(m) = \int_N f(mn)\|X(mn)\|^{-\eta}\,dn \ll \int_N f(mn)\mathcal{L}(\operatorname{Ad}(mn))^{-\eta}\,dn.$$

Writing $n = e^V$ and applying Lemma 7.8 gives

$$F_\eta^P(m)\delta_P^{-1/2}(m) \ll \int_{\mathfrak{n}} f(me^V)\|V\|^{-2\eta}\,dV.$$

In view of our assumption $0 < \eta < 1/2$, the right-hand side is bounded by $C(B, \eta)\|f\|_\infty$. This proves (i), and (ii) follows by combining this with $\operatorname{supp}(F_\eta^P) \subset B_M$. $\quad\square$

Combining Lemmas 7.6 and 7.9 gives our next result.

LEMMA 7.10. *Let $f \in C_c(G)$, and for $0 < \eta < 1/2$ define $F_\eta = f\|X(\cdot)\|^{-\eta}$. If $\gamma \in M$ is (G, M)-regular, then the integral $O_\gamma(F_\eta)$ converges absolutely, and we have*

$$D_M^G(\gamma)^{1/2}O_\gamma(F_\eta) = O_\gamma^M(F_\eta^P). \tag{28}$$

We return now to the proof of Proposition 7.3. Let $0 < \eta < 1/2$, and let $f \in C_c(G)$ and $F_\eta = f\|X(\cdot)\|^{-\eta}$. We assume that $f \geqslant 0$. As γ is not elliptic, we may let M be a proper Levi subgroup with $G_\gamma \subset M$. Because $G_\gamma \subset M$, $\operatorname{Ad}(\gamma) - 1$ must be invertible on $\mathfrak{g}/\operatorname{Lie}(M)$, so that γ is (G, M)-regular. Let P be a parabolic with Levi M. By conjugation, we may assume that P and M are semi-standard. We apply Lemma 7.9 to F_η to obtain $f_1 \in C_c(M)$ such that $F_\eta^P(m) \leqslant f_1(m)$ for all $m \in M$. Lemma 7.10 gives

$$O_\gamma^G(F_\eta) = D_M^G(\gamma)^{-1/2}O_\gamma^M(F_\eta^P) \leqslant D_M^G(\gamma)^{-1/2}O_\gamma^M(f_1).$$

From this and Proposition 7.1, applied on M, we obtain

$$O_\gamma^G(F_\eta) < c(\eta, f)D_M^G(\gamma)^{-1/2}D^M(\gamma)^{-1/2} < c'(\eta, f)D^G(\gamma)^{-1/2},$$

as desired.

7.6 The case of γ elliptic

We first observe that, if $\gamma \in G$ is elliptic, then it is conjugate to an element of K. This is because if $Z(G_\gamma)$ is anisotropic then it is compact and therefore conjugate to a subgroup of K.

We shall assume $\gamma \in K - G_{\mathrm{cpt}}$, and note that for such γ we have $D(\gamma) \ll 1$, where the implied constant depends only on G. For $\epsilon > 0$ we put $K(\epsilon) = \{g \in G : \|X(g)\| < \epsilon\}$ and write $1_{K(\epsilon)}$ for the characteristic function of $K(\epsilon)$. Proposition 7.3 will follow from a bound for $O_\gamma(1_{K(\epsilon)})$ that is uniform in ϵ and $D(\gamma)$, as provided by the following lemma.

LEMMA 7.11. *For all $\gamma \in K - G_{\mathrm{cpt}}$ and $0 < \epsilon \leqslant 2D(\gamma)^{1/2}$, we have $O_\gamma(1_{K(\epsilon)}) \ll \epsilon^{d_\gamma}D(\gamma)^{-d_\gamma/2}$, where $d_\gamma \geqslant 1$ is the codimension of $I_\gamma K$ in G.*

992

Proof. We first convert the problem to one on the symmetric space $S = G/K$. Let ds be the metric tensor on S associated with the norm $\|\cdot\|$ on $\mathfrak{p}$. We note that the distance function d_S attached to ds is given by $d_S(g_1, g_2) = \|X(g_1^{-1}g_2)\|$.

We observe that $O_\gamma(1_{K(\epsilon)}) = \mu_\gamma(I_\gamma \backslash G(\epsilon))$, where $G(\epsilon) = \{x \in G : x^{-1}\gamma x \in K(\epsilon)\}$. The set $G(\epsilon)$ is right K-invariant, and $G(\epsilon)/K$ is the set of points $x \in S$ such that $d_S(x, \gamma x) < \epsilon$. Moreover, $G(\epsilon)$ is left I_γ-invariant, and we will see it is roughly a tube around I_γ. Bounding the volume of $I_\gamma \backslash G(\epsilon)$ is therefore roughly equivalent to finding the radius of this tube, in a way which we now make precise.

Because $\theta(\gamma) = \gamma$, $\mathfrak{g}_\gamma$ is θ-stable. We may then write $\mathfrak{g}_\gamma = \mathfrak{p}_\gamma + \mathfrak{k}_\gamma$ where $\mathfrak{k}_\gamma = \mathfrak{k} \cap \mathfrak{g}_\gamma$ and $\mathfrak{p}_\gamma = \mathfrak{p} \cap \mathfrak{g}_\gamma$. Let $\mathfrak{p}_\gamma^\perp$ be the orthocomplement of $\mathfrak{p}_\gamma$ in $\mathfrak{p}$. We note that $\gamma \notin G_{\mathrm{cpt}}$ implies that $\mathfrak{p}_\gamma \neq \mathfrak{p}$. Indeed, if $\mathfrak{p}_\gamma = \mathfrak{p}$ then $\mathrm{Ad}(\gamma)$ fixes $\mathfrak{p}$, and hence $[\mathfrak{p}, \mathfrak{p}]$. However, $\mathfrak{p} + [\mathfrak{p}, \mathfrak{p}]$ is the product of the $\mathbb{R}$-simple factors of $\mathfrak{g}$ of non-compact type (see [Hel01, ch. V, Theorem 1.1] and the subsequent proof). In particular, the codimension d_γ of $I_\gamma K$ in G is at least 1.

Let $x \in G(\epsilon)$. By [Hel01, ch. VI, Theorem 1.4], we may write $x = e^{X_\gamma} e^{X^\gamma} k$ with $X_\gamma \in \mathfrak{p}_\gamma$, $X^\gamma \in \mathfrak{p}_\gamma^\perp$ and $k \in K$. The condition $x^{-1}\gamma x \in K(\epsilon)$ simplifies to $e^{-X^\gamma}\gamma e^{X^\gamma} \in K(\epsilon)$, which (since $\gamma \in K$) is equivalent to $e^{-X^\gamma}e^{\mathrm{Ad}(\gamma)X^\gamma} \in K(\epsilon)$. This implies that $d_S(e^{X^\gamma}, e^{\mathrm{Ad}(\gamma)X^\gamma}) < \epsilon$: the element e^{X^γ} is rotated by γ by a distance at most ϵ. The Cartan–Hadamard theorem [Hel01, ch. I, Theorem 13.1] then implies that $\|\mathrm{Ad}(\gamma)X^\gamma - X^\gamma\| < \epsilon$. Because all eigenvalues of $\mathrm{Ad}(\gamma) - 1$ on $\mathfrak{p}_\gamma^\perp$ have absolute value at most 2, and they occur in complex conjugate pairs, their absolute value is at least $CD(\gamma)^{1/2}$ for $C > 0$ depending only on G. It follows that $\|X^\gamma\| \leqslant C\epsilon D(\gamma)^{-1/2}$. If we let $B_{\mathfrak{p}_\gamma^\perp}(r)$ be the ball of radius r around 0 in $\mathfrak{p}_\gamma^\perp$ with respect to $\|\cdot\|$, it follows that $G(\epsilon) \subset I_\gamma \exp(B_{\mathfrak{p}_\gamma^\perp}(r_0))K$ with $r_0 = C\epsilon D(\gamma)^{-1/2} \leqslant 2C$.

Let $I_\gamma^{\mathrm{c}} \subset I_\gamma$ be a compact set such that $\mu_{I_\gamma}^{\mathrm{can}}(I_\gamma^{\mathrm{c}}) = 1$. Then from the preceding paragraph we deduce that

$$\mu_\gamma(I_\gamma \backslash G(\epsilon)) \leqslant \mu_G(I_\gamma^{\mathrm{c}} \exp(B_{\mathfrak{p}_\gamma^\perp}(r_0))K).$$

Let $B_{\mathfrak{p}}(r)$ denote the ball of radius r around 0 in $\mathfrak{p}$ with respect to $\|\cdot\|$. We have

$$I_\gamma^{\mathrm{c}} \exp(B_{\mathfrak{p}_\gamma^\perp}(r_0))K \subset I_\gamma^{\mathrm{c}} \exp(B_{\mathfrak{p}}(r_0))K = I_\gamma^{\mathrm{c}} K \exp(B_{\mathfrak{p}}(r_0)).$$

Because $I_\gamma^{\mathrm{c}} K$ is compact and contained in $I_\gamma K$, the result follows from the expression for r_0, provided we show that our bounds do not depend on the choice of I_γ^{c}.

We show that, after conjugating γ in G to another element of K, there are only finitely many possibilities for I_γ.[4] Let $T_1, \ldots, T_n$ be representatives for the conjugacy classes of maximal $\mathbb{R}$-tori in G. If T_i^{c} denotes the maximal compact subgroup of $T_i(\mathbb{R})$, we may assume that $T_i^{\mathrm{c}} \subset K$ for all i. As we are free to replace γ by any conjugate, we may assume that $\gamma \in T_i$ for some i, which implies that $\gamma \in T_i^{\mathrm{c}} \subset K$. There are now finitely many possibilities for I_γ. $\qquad\square$

We now return to the proof of Proposition 7.3. Choose a compact set B containing $\mathrm{supp}(f)$, so that

$$f\|X(\cdot)\|^{-\eta} \ll_f 1_B + \sum_{k=1}^\infty 2^{\eta k} 1_{K(2^{-k})}.$$

We deduce that

$$O_\gamma(f\|X(\cdot)\|^{-\eta}) \ll_f O_\gamma(1_B) + \sum_{k=1}^\infty 2^{\eta k} O_\gamma(1_{K(2^{-k})}).$$

---

[4] We have already shown in the proof of Proposition 7.1 that there are finitely many possibilities for I_γ after conjugating γ in G, but we need this stronger statement.

An application of Proposition 7.1 and Lemma 7.11 yields

$$O_\gamma(f\|X(\cdot)\|^{-\eta}) \ll_f D(\gamma)^{-1/2} + \sum_{2^{k+1} \geqslant D(\gamma)^{-1/2}} 2^{\eta k - k} D(\gamma)^{-1/2} + \sum_{2^{k+1} < D(\gamma)^{-1/2}} 2^{\eta k} D(\gamma)^{-1/2}.$$

Our assumption that $0 < \eta < 1/2$ implies that both geometric series are bounded by $D(\gamma)^{-3/4}$, completing the proof of Proposition 7.3.

Appendix A. Distinction principles and the C^∞ spectrum of symmetric varieties

There are several methods that may be used to prove power growth of eigenfunctions on arithmetic manifolds. The original proof of Rudnick and Sarnak uses a distinction principle. This means that, for certain period integrals, if an automorphic form ϕ has a non-zero period then ϕ is exceptional in some sense, which can mean being a transfer from a smaller group, or being non-tempered.

In this appendix we review the method of Rudnick and Sarnak and describe an alternative approach to proving Theorems 1.1 and 1.2 which retains much of the spirit of their approach. It is based on a theorem of Sakellaridis on the unramified C^∞ spectrum of spherical varieties, which we apply in the case of symmetric varieties. In short, we replace the global distinction principle of Rudnick and Sarnak which implies being in the image of a functorial lift by a local distinction principle which implies having non-tempered Satake parameters. What both of these implied properties have in common is that they are rare. Counting arguments can then be used to produce power growth.

A.1 The global distinction argument
We illustrate the distinction argument in a special case, taken from Rudnick and Sarnak's proof [RS94].

Let Q be the quadratic form $Q(x) = x_1^2 + x_2^2 + x_3^2 - 7x_4^2$. If we let $V = \{x \in \mathbb{R}^4 : Q(x) = -1\}$, then V is a two-sheeted hyperboloid and the upper sheet is a model for $\mathbb{H}^3$. If we let Γ be the intersection of $\mathrm{O}(Q, \mathbb{Z})$ with the identity component of $\mathrm{O}(Q, \mathbb{R})$, then $Y = \Gamma \backslash \mathbb{H}^3$ is a compact hyperbolic 3-manifold. The distinction result that Rudnick and Sarnak prove is that if $\psi \in L^2(Y)$ is orthogonal to all theta lifts of cusp forms of weight 1 on $\Gamma_1(28)$, then $\psi((2, 1, 1, 1)) = 0$. The result then follows from the local Weyl law and a counting argument. Indeed, the local Weyl law says that the average size of $|\psi((2, 1, 1, 1))|^2$ must be 1. However, the number of eigenfunctions on X with eigenvalue $\lambda \leqslant R$ is roughly R^3, while the number of theta lifts in this range is roughly R^2. Because the number of non-vanishing eigenfunctions is small, their values must be large to make up the right average.

The generalization of this principle, namely that an automorphic form on $\mathrm{SO}(n, 1)$ that is orthogonal to theta lifts from SL_2 must have vanishing $\mathrm{SO}(n)$ periods, was used by Donnelly [Don07]. It is likely that this could be used to prove Theorem 1.1 on other groups of the form $\mathrm{SO}(m, n)$, $\mathrm{U}(m, n)$, or $\mathrm{Sp}(m, n)$. Another distinction principle that one could apply is due to

Jacquet [Jac05] (and later refined by Feigon, Lapid and Offen in [FLO12, LO07]), which states that a form on $\mathrm{GL}(n, \mathbb{C})$ with a non-vanishing $\mathrm{U}(n)$ period must come from quadratic base change. See [JLR93] for a general discussion of these ideas.

A.2 Symmetric varieties

Let F be a field of characteristic 0. A symmetric variety over F is a variety $X = G/H$ where G is a reductive F-group, θ is an involution of G over F, and H is an open F-subgroup of the fixed point group G^θ. We refer the reader to [HW93, Vus74] for background on these varieties.

If $S \subset G$ is a θ-stable torus, let S^+ and S^- be the neutral component of S^θ and $\{x \in S : \theta(x) = x^{-1}\}$ respectively, so that $S^+ \times S^- \to S$ is an isogeny. A θ-stable torus S is said to be (F, θ)-split if it is F-split and θ acts on it by inversion. A parabolic subgroup P of G is said to be θ-split if P and θP are opposed. In this case, $P \cap \theta(P)$ is the unique θ-stable Levi of both P and θP. All maximal (F, θ)-split tori and all minimal θ-split F-subgroups are conjugate under $G(\overline{F})$. We define the θ-split rank of G to be the common dimension of such tori. We say that G is θ-split if its θ-split rank is equal to its absolute rank, that is, if G contains an (F, θ)-split maximal torus. We say that G is θ-quasi-split if G contains a θ-split Borel defined over F. In this case, $B \cap \theta(B)$ is a θ-stable torus, which may or may not be θ-split.

From now on we shall assume G split over F. The following lemma then implies that G is θ-(quasi-)split if and only if $G \times \overline{F}$ is.

LEMMA A.1. *A maximal (F, θ)-split torus in G is also a maximal θ-split torus in $G \times \overline{F}$, and a minimal θ-split F-parabolic in G is also a minimal θ-split parabolic in $G \times \overline{F}$.*

Proof. Let S be such a torus, and let A be a maximal F-split torus containing S. As in [HW93, Lemma 4.5], let C, M_1 and M_2 be the central, anisotropic and isotropic factors of $Z_G(S)$ over F. Because $Z_G(A) = A$, [HW93] gives $CM_1 \subset A$ so that M_1 is trivial and C is split. We have $S \subset C$, and C is θ-stable, so by the maximality of S we must have $S = C^-$. As $M_2 \subset G^\theta$ by [HW93], θ acts trivially on $Z_G(S)/C^-$ so there is no strictly larger θ-split torus containing S over $\overline{F}$.

Let P be such a parabolic. By [HW93, Lemma 2.4] we may let $A \subset P$ be a θ-stable maximal F-split torus of G. Then [HW93, Proposition 4.7] implies that A^- is a maximal (F, θ)-split torus of G, and $Z_G(A^-) = P \cap \theta(P)$. We have shown that A^- is a maximal θ-split torus in $G \times \overline{F}$, and applying [HW93] again gives that P is minimal θ-split in $G \times \overline{F}$. □

A.3 A result of Sakellaridis on the unramified C^∞ spectrum

In [Sak08, Sak13], Sakellaridis describes the unramified C^∞ spectrum of a spherical variety. We reprise his results below in the symmetric case, and use them to describe an alternative approach to Theorems 1.1 and 1.2 in the case when G_{v_0} is not quasi-split. Throughout, F will be a p-adic field and, as in § A.2, the group G will be split over F.

Let A and B be the θ-stable maximal split torus and Borel of G chosen in Lemma A.4. Let δ denote the modular character of A with respect to B. Let $\check{A}$ be the complex dual torus of A, so that unramified characters of A correspond to elements of $\check{A}$. Let W be the Weyl group of A and $\check{A}$. Recall that the irreducible unramified representations of G are in bijection with $\check{A}/W$, via the map taking π to its Satake parameter. We introduce the torus $A_X = A/A \cap H$ and write $\check{A}_X$ for its dual torus. We then have a map $\iota : \check{A}_X \to \check{A}$ with finite kernel. If χ is an unramified character of A, let $I(\chi)$ be the corresponding unitarily normalized induced representation of G.

The following result of Sakellaridis [Sak08, Theorem 1.2.1] describes the unramified H-distinguished C^∞ spectrum in terms of the image $\iota(\check{A}_X)$ in $\check{A}$, denoted by A_X^*.

995

THEOREM A.2 (Sakellaridis). *There exists a non-zero morphism $C_c^\infty(X) \to I(\chi)$ only if χ lies in a W-translate of $\delta^{-1/2} A_X^*$.*

From this we may deduce the following consequence.

COROLLARY A.3. *If G is not θ-quasi-split, any irreducible unramified representation π of G that occurs as a subrepresentation of $C^\infty(X)$ must be non-tempered.*

To see how Corollary A.3 follows from Theorem A.2, we must convert from Sakellaridis' notation to ours. Sakellaridis defines the tori A and A_X using an open Borel orbit $\mathring{X} \subset X$ and its stabilizer, a parabolic subgroup called the associated parabolic. The following lemma translates his definitions to the symmetric case.

LEMMA A.4. *Let notation be as in §A.2. In particular, G is assumed split. Let $x_0 = H$ denote the identity coset in $X = G/H$. We may choose*

- *a θ-stable maximal F-split torus A,*
- *a Borel subgroup B defined over F, and*
- *a minimal θ-split parabolic F-subgroup P*

such that $A \subset B \subset P$ and

- *the orbit Bx_0 is open,*
- *$P = \mathrm{stab}_G(Bx_0)$ is the parabolic associated with the open orbit, and*
- *the tori A and $A_X = A/A \cap H$ are the same as those associated with B and P in [Sak13, p. 8].*

Proof. Let P be a minimal θ-split parabolic F-subgroup, and let $A \subset P$ be a θ-stable maximal F-split torus of G. We may choose a Borel subgroup $A \subset B \subset P$. By [HW93, Lemma 4.8], Bx_0 is open in X, and P stabilizes this open orbit. It is also known that $P = \mathrm{stab}_G(Bx_0)$; see, for example, [Sak13, p. 8]. Let $L = Z_G(A^-)$ be the θ-stable Levi of P, so that $A \subset B \cap L$. It follows that our choices of A, B, P and L are compatible with those in [Sak13, p. 8] in the special case of a symmetric variety. Sakellaridis defines $A_X = L/L \cap H$ in [Sak13, p. 8]. As A is a maximal torus in L and $[L, L] \subset H$ [Vus74, Proposition 2(ii)], we have an exact sequence $1 \to A \cap H \to A \to L/L \cap H \to 1$ so $A_X \simeq A/A \cap H$, which completes the proof. $\square$

We next show that if G is not θ-quasi-split, then $\delta^{-1/2} A_X^*$ does not intersect the maximal bounded subgroup of $\check{A}$. If there is some unitary $\chi \in \check{A}$ that lies in $\delta^{-1/2} A_X^*$, then we have $\chi \delta^{1/2} \in A_X^*$. As A_X^* is the set of characters of A that arise by pullback from A_X, this implies that $\chi \delta^{1/2}$ is trivial on $A \cap H$, and that δ is unitary on $A \cap H$. However, this contradicts the following lemma.

LEMMA A.5. *δ is unitary on A^+ if and only if G is θ-quasi-split.*

Proof. Let $P = LU$. We have $\delta = \delta_L \delta_U$, where δ_L and δ_U are the modular characters of $B \cap L$ and U, respectively. Because P and $\theta(P)$ are opposed, we have $\theta(\delta_U) = \delta_U^{-1}$, so that $\delta_U^2 = 1$ on A^+; by positivity, $\delta_U = 1$ on A^+. Therefore $\delta|_{A^+}$ is unitary if and only if $\delta_L|_{A^+}$ is. As δ_L is trivial on A^-, this is equivalent to δ_L being unitary, or to L being a torus. This completes the proof. $\square$

We now turn to the proof of Corollary A.3. Assume that G is not θ-quasi-split, and let $\pi \subset C^\infty(X)$ be irreducible, unramified and tempered. Let $\chi \in \check{A}/W$ be the Satake parameter of π. The contragredient $\pi^\vee$ has Satake parameter χ^{-1}, and we have a non-zero map $C_c^\infty(X) \to \pi^\vee$. By [Cas80, Proposition 2.6] we may choose a representative for χ^{-1} in $\check{A}$ such that $\pi^\vee \subset I(\chi^{-1})$. We now have a non-zero map $C_c^\infty(X) \to I(\chi^{-1})$, so that by Theorem A.2, $W\chi^{-1}$ must intersect $\delta^{-1/2} A_X^*$, but this contradicts our assertion above as χ is unitary.

A.4 An alternative approach to Theorems 1.1 and 1.2

We now describe how one might use Corollary A.3 to prove asymptotic lower bounds for periods. The argument is in the same style as that of Rudnick and Sarnak described in § A.1.

We return to the global situation of Theorems 1.1 and 1.2, and freely use the notation of those statements. We note that the compact subgroup K_∞ is given by $K_{v_0} = H_{v_0}$ and $K_v = G_v$ at all other infinite places. We shall let $N = 1$ for simplicity, and write $Y = G(F)\backslash G(\mathbb{A})/K_\infty K$. The image of $H(\mathbb{A})$ in Y is a finite number of points, and we shall assume for simplicity that it is a single point p. Let $\mathcal{P}_H : C^\infty(G(F)\backslash G(\mathbb{A})) \to \mathbb{C}$ be the period map $f \mapsto \int_{H(F)\backslash H(\mathbb{A})} f \, dh$. When restricted to $C^\infty(Y)$, $\mathcal{P}_H$ is just evaluation at p.

Note that a general connected real reductive group G' with Cartan involution θ is (quasi-)split over $\mathbb{R}$ if and only if $G' \times \mathbb{C}$ is θ-(quasi-)split. If we assume that G_{v_0} is not quasi-split over $\mathbb{R}$ then this, together with the invariance of θ-(quasi-)splitness under extension of algebraically closed fields, and Lemma A.1, implies that G_v is not θ-quasi-split at any finite v at which G splits. Corollary A.3 then gives that any unramified representation π_v occurring in $C^\infty(G_v/H_v)$ must be non-tempered.

Let π be a cuspidal automorphic representation of G, and let $\psi \in \pi$ be invariant under $K_\infty K$. If $\mathcal{P}_H(\psi) \neq 0$, this implies that each factor π_v admits a non-zero smooth linear functional invariant under H_v. This is equivalent to the existence of an embedding $\pi_v \to C^\infty(G_v/H_v)$, so that if v is finite, G splits at v, and all data are unramified, then π_v is non-tempered. The strategy would then be to use the trace formula to show that the number of such ψ is a power smaller than the total number of ψ. Combining this with the local Weyl law would then produce asymptotic growth.

Note that in the case when G_{v_0} is quasi-split but not split over $\mathbb{R}$, Theorem A.2 only implies that the Satake parameters of the distinguished π lie in a fixed lower-dimensional subset of $\check{A}$. It may also be possible to use this to prove a power saving for the number of such π.

A.5 Higher-dimensional periods

One advantage of the method described in § A.4 is that it applies equally well to periods along positive-dimensional submanifolds of Y arising from rational symmetric subgroups. Note that this would produce a result of the form 'there are certain eigenfunctions whose periods are larger than the average by a power of the eigenvalue', while determining the size of the average period is a separate problem. It should be pointed out that the average size of a positive-dimensional period should be a negative power of the eigenvalue, so even if one could improve over this one would not necessarily obtain power growth of sup norms as a result.

In comparison, the relative trace formula approach we use is more difficult in the positive-dimensional case, because the analysis of the error terms becomes much more complicated. In the case of a point, one needs to bound the value of a spherical function φ_λ away from its centre of symmetry. In the positive-dimensional case, one needs to bound the value of an oscillatory integral whose kernel is constructed from φ_λ, and which is taken over two copies of the submanifold in question. Moreover, the bound obtained must be uniform as the submanifolds move.

997

Appendix B. Hecke operators and the L^2 spectrum of symmetric varieties

We have sketched two proofs of Theorems 1.1 and 1.2, the one described in § 1.5, which we carry out to completion in this paper, and the one outlined in § A.4. In the latter, no Hecke operators appear; in the former, no distinction principle is used. What is their connection, if any?

In this section we describe how a conjecture of Sakellaridis and Venkatesh on the L^2 spectrum of symmetric varieties relates to the method we have used in this paper, in particular to our choices of test functions in the trace formula. As their conjecture is expressed in terms of X-distinguished Arthur parameters, it acts in a sense as a bridge between the two methods.

B.1 Plancherel measures

Recall from § 1.5 that a comparison of trace formulae reduces the problem of finding exceptional sequences of Maass forms to that of finding ω in the global Hecke algebra with the property that $\Pi_H(\omega * \omega^*)(1)$ is large relative to $(\omega * \omega^*)(1)$. Indeed, these are just the identity distributions on the geometric side of the H-relative trace formula and the Arthur–Selberg trace formula. The latter question may in turn be reduced to a local problem, namely that of finding an L^2 normalized $\tau \in \mathcal{H}_v$ satisfying

$$\tau(1) = 0 \quad \text{and} \quad \Pi_H \tau(1) \gg 1, \tag{B.1}$$

for v in a set of places having positive density. In § 1.5 we explained how these two conditions can be interpreted geometrically, in terms of Hecke returns to a fixed point. We would now like to interpret these same conditions spectrally.

Until § B.3 we let F be a p-adic field. We let G be a split reductive group over F, and let K be a hyperspecial maximal compact subgroup. We assume that the Haar measure on G gives K measure 1. We let $\mathcal{H}$ be the spherical Hecke algebra with respect to K. We let $\widehat{G}$ be the unitary dual of G, and let $\widehat{G}^{\mathrm{sph}}$ be the spherical unitary dual with respect to K.

Then the first condition in (B.1) can be expressed, via the Plancherel inversion formula, as

$$\tau(1) = \int_{\widehat{G}^{\mathrm{sph}}} \widehat{\tau}(\nu) \, d\mu_G^{\mathrm{sph}}(\nu) = 0. \tag{B.2}$$

In other words, τ must be such that its Satake transform $\widehat{\tau}$ is oscillatory along the support of μ_G^{sph}, the tempered spectrum $\widehat{G}^{\mathrm{sph,temp}}$.

A similar Plancherel inversion formula can be used to express the second condition in (B.1). Let $X = G/H$ be a symmetric variety. We recall the existence of a Plancherel measure μ_X associated with the separable Hilbert space $L^2(X)$, viewed as a G-representation; roughly speaking, this is a measure on $\widehat{G}$ satisfying

$$L^2(X) = \int_{\widehat{G}} M(\pi) \otimes \pi \, d\mu_X(\pi),$$

where $M(\pi)$ is some multiplicity space. Notice that μ_X is only defined up to absolutely continuous equivalence (we shall only be concerned with its support).

We let $\Pi_H : L^1(G) \to L^1(X)$ be given by integration over H. If we let $v^0 = \Pi_H(\mathbf{1}_K)$, there is a second measure, the spherical Plancherel measure μ_X^{sph}, which satisfies

$$\langle \omega \cdot v^0, v^0 \rangle_{L^2(X)} = \int_{\widehat{G}^{\mathrm{sph}}} \widehat{\omega}(\nu) \, d\mu_X^{\mathrm{sph}}(\nu)$$

998

for all $\omega \in \mathcal{H}$. In particular, the support of μ_X^{sph} is contained in the support of μ_X. Note that $\omega \cdot v^0$ denotes the action of ω on v^0, given by

$$\omega \cdot v^0 = \int_G \omega(g)(g \cdot v^0)\, dg.$$

By a simple unfolding argument, one may show that for any $\omega \in \mathcal{H}$ we have

$$\Pi_H \omega(1) = \int_{\widehat{G}^{\mathrm{sph}}} \widehat{\omega}(\nu)\, d\mu_X^{\mathrm{sph}}(\nu),$$

so that μ_X^{sph} determines the period of ω along H. Using this, we may now rephrase the second condition in (B.1) as

$$\Pi_H \tau(1) = \int_{\widehat{G}^{\mathrm{sph}}} \widehat{\tau}(\nu)\, d\mu_X^{\mathrm{sph}}(\nu) \gg 1. \tag{B.3}$$

In other words, τ must be such that its Satake transform $\widehat{\tau}$ does not oscillate too much on the support of μ_X^{sph}.

Having expressed the two conditions in (B.1) spectrally, we see that the existence of appropriate test functions τ can be read off from the support of μ_X^{sph}, in particular relative to the tempered spectrum. We say that μ_X^{sph} is *tempered* if the support of μ_X^{sph} is contained in $\widehat{G}^{\mathrm{sph,temp}}$ and *strongly tempered* if $\mu_X^{\mathrm{sph}} \leqslant C \mu_G^{\mathrm{sph}}$ for some $C > 0$, where μ_G^{sph} is the spherical Plancherel measure on G. With this terminology, we summarize our discussion as follows:

(ST): if μ_X^{sph} is strongly tempered the oscillation of $\widehat{\tau}$ should prevent one from simultaneously achieving both (B.2) and (B.3);

(T): if μ_X^{sph} is tempered but not strongly tempered, the existence of τ satisfying both (B.2) and (B.3) depends on how singular μ_X^{sph} is relative to μ_G^{sph};

(NT): if μ_X^{sph} is non-tempered, the exponential growth of $\widehat{\tau}(\nu)$ away from $\widehat{G}^{\mathrm{sph,temp}}$ should allow one (barring unforseen cancellation) to ensure both conditions (B.2) and (B.3).

In the next subsection we shall see how recent conjectures of Sakelleridis and Venkatesh relate the tempered properties of μ_X to the weak containment properties of H-distinguished representations.

B.2 The conjecture of Sakellaridis and Venkatesh

Recall the dual torus $\check{A}_X$ and the map $\check{A}_X \to \check{A}$, which were of critical use in describing the local distinction argument of Appendix A. Let $\check{G}$ denote the Langlands dual group of G, containing $\check{A}$ as a maximal torus. In [SV17, §2.2], Sakellaridis and Venkatesh define a dual group $\check{G}_X$ associated with X and a homomorphism $\iota : \check{G}_X \times \mathrm{SL}(2, \mathbb{C}) \to \check{G}$ whose restriction to $\check{G}_X$ has finite kernel. (Note that this requires imposing certain conditions on X, which we shall ignore as this section is purely expository.) The torus $\check{A}_X$ sits in the complex reductive algebraic group $\check{G}_X$ as a maximal torus, and the map $\iota : \check{G}_X \to \check{G}$ restricts to the natural map $\check{A}_X \to \check{A}$ above.
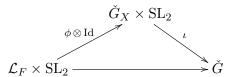
Again under technical assumptions that we shall ignore, $\iota(\check{G}_X)$ is equal to a group constructed by Gaitsgory and Nadler; see [SV17, §3], in particular §3.2. Moreover, in the case of symmetric varieties, the group of Gaitsgory and Nadler is equal to the group $\check{H}$ constructed by Nadler in [Nad05].[5] We continue to assume that X is symmetric and now recall the following facts about

---

[5] Note that Gaitsgory and Nadler consider spherical varieties over $\mathbb{C}$ rather than a p-adic field, but we may ignore this distinction as the dual groups are only defined using root data that are independent of the field. Likewise, Nadler works with real reductive groups, but these are equivalent to complex reductive groups with involution. These equivalences respect rank and quasi-splitness in the natural way.

$\check{G}_X$ and ι.

- The rank of $\check{G}_X$ is equal to the θ-split rank of G.
- $\iota(\check{G}_X) = \check{G}$ if and only if G is θ-split.
- ι is trivial on the $\mathrm{SL}(2, \mathbb{C})$ factor if and only if G is θ-quasi-split.

The first claim is stated in [Nad05, §1.1 and Proposition 10.6.1]. The fact that $\iota(\check{G}_X) = \check{G}$ when G is θ-split is also stated in §1.1 there, and the reverse implication follows by considering ranks. The third claim follows from the condition that ι be a distinguished morphism, as defined in the comment before [SV17, §2.2, Theorem 2.2.3]. Indeed, the group L in that comment is the Levi of a minimal θ-split parabolic, and ρ_L is the half sum of its positive roots, so that ι is trivial on $\mathrm{SL}(2, \mathbb{C})$ if and only if ρ_L is trivial, that is, L is a torus. Sakellaridis and Venkatesh conjecture [SV17, Conjecture 16.2.2] that the support of μ_X may be described in terms of the tempered dual of $\check{G}_X$ and the map ι. They define an X-distinguished Arthur parameter to be a commutative diagram

$$
\begin{array}{ccc}
 & \check{G}_X \times \mathrm{SL}_2 & \\
\phi \otimes \mathrm{Id} \nearrow & & \searrow \iota \\
\mathcal{L}_F \times \mathrm{SL}_2 & \longrightarrow & \check{G}
\end{array}
$$

where $\mathcal{L}_F$ is the local Langlands group of F, and ϕ is a tempered Langlands parameter for $\check{G}_X$. This naturally gives rise to an Arthur parameter for $\check{G}$. We shall say that an Arthur parameter for $\check{G}$ is X-distinguished if it arises from such a diagram, and likewise for an X-distinguished Arthur packet.

CONJECTURE 1 (Sakellaridis and Venkatesh). The support of μ_X is contained in the Fell closure of the union of the X-distinguished Arthur packets for $\check{G}$.

Note that the unramified members of a given Arthur packet should be contained in the associated L-packet, and we assume this from now on without further comment. Thus the above conjecture implies, in particular, that the support of the spherical measure μ_X^{sph} is contained in the Fell closure of the union of the L-packets associated with X-distinguished Arthur parameters. The latter statement has in fact been proved in [Sak13] under certain combinatorial assumptions. Let us now discuss what Conjecture 1 implies for μ_X^{sph} under the assumptions that G is θ-split, θ-quasi-split, or neither.

(ST): If G is θ-split, then $\iota(\check{G}_X) = \check{G}$ and ι is trivial on $\mathrm{SL}(2, \mathbb{C})$. Conjecture 1 then implies that μ_X is supported on the tempered dual of G. In fact, it may be shown in this case that μ_X^{sph} is strongly tempered.

(T): If G is θ-quasi-split but not θ-split, ι is still trivial on the $\mathrm{SL}(2, \mathbb{C})$ factor. This implies that μ_X is still tempered. However, because $\mathrm{rank}(\check{G}_X) < \mathrm{rank}(\check{G})$, if we identify the tempered spherical dual of G with a quotient of a compact torus by the Weyl group, the support of μ_X^{sph} will be contained in a union of lower-dimensional tori. In particular, μ_X^{sph} will not be strongly tempered.

(NT): If G is not θ-quasi-split, then an X-distinguished Arthur parameter has non-trivial $\mathrm{SL}(2, \mathbb{C})$ factor. Its underlying Langlands parameter is therefore non-tempered, and so then are all members of the associated L-packet. It follows that if ψ is X-distinguished with packet Π_ψ, and $\pi \in \Pi_\psi$ is spherical, then π must be non-tempered. From Conjecture 1 we deduce that the same is true for any π in the support of μ_X^{sph}.

1000

B.3 Existence of test functions

Finally, we return to the global situation of Theorems 1.1 and 1.2. We freely use the notation of those statements, with the exception that we drop the assumption that G_{v_0} is not split. Let v be a finite place at which G is split and all data are unramified. We let $\mu_{G,v}^{\mathrm{sph}}$ and $\mu_{X,v}^{\mathrm{sph}}$ denote the spherical Plancherel measures of G_v and $X_v = G_v/H_v$. As before, G_{v_0} is (quasi-)split over $\mathbb{R}$ if and only if G_v is θ-(quasi-)split.

Comparing the conditions from §§ B.1 and B.2, we obtain the following consequences of Conjecture 1 for the existence of τ satisfying both conditions in (B.1).

(ST): If G_{v_0} is split, then $\mu_{X,v}^{\mathrm{sph}}$ is strongly tempered, and such τ should not exist.

(T): If G_{v_0} is quasi-split but not split, the existence of such τ depends on how singular $\mu_{X,v}^{\mathrm{sph}}$ is.

(NT): If G_{v_0} is not quasi-split, $\mu_{X,v}^{\mathrm{sph}}$ has non-tempered support, and such τ should exist.

Recalling our remarks at the end of § A.4, it seems that in the case (T) above the measure $\mu_{X,v}^{\mathrm{sph}}$ is still singular enough to allow (B.1) to be satisfied. This of course aligns with the conditions of our theorems, which only require G_{v_0} to be non-split.

References

Ava56   G. V. Avacumović, *Über die Eigenfunktionen auf geschlossenen Riemannschen Mannigfaltig-keiten*, Math. Z. **65** (1956), 327–344.

BP16   V. Blomer and A. Pohl, *The sup-norm problem on the Siegel modular space of rank two*, Amer. J. Math. **138** (2016), 999–1027.

Bor63   A. Borel, *Compact Clifford–Klein forms of symmetric spaces*, Topology **2** (1963), 111–122.

Bor91   A. Borel, *Linear algebraic groups*, Graduate Texts in Mathematics, vol. 126, second edition (Springer, New York, 1991).

BT84   F. Bruhat and J. Tits, *Groupes réductifs sur un corps local: II. Schémas en groupes. Existence d'une donnée radicielle valué*, Publ. Math. Inst. Hautes Études Sci. **60** (1984), 5–184.

Cas80   W. Casselman, *The unramified principal series of p-adic groups I*, Compos. Math. **40** (1980), 387–406.

Col85   Y. Colin de Verdière, *Ergodicité et fonctions propres du Laplacien*, Comm. Math. Phys. **102** (1985), 497–502.

Don07   H. Donnelly, *Exceptional sequences of eigenfunctions for hyperbolic manifolds*, Proc. Amer. Math. Soc. **135** (2007), 1551–1555.

DKV79   J. J. Duistermaat, J. A. C. Kolk and V. S. Varadarajan, *Spectra of compact locally symmetric manifolds of negative curvature*, Invent. Math. **52** (1979), 27–93.

FLO12   B. Feigon, E. Lapid and O. Offen, *On representations distinguished by unitary groups*, Publ. Math. Inst. Hautes Études Sci. **115** (2012), 185–323.

FL18   T. Finis and E. Lapid, *An approximation principle for congruence subgroups*, J. Eur. Math. Soc. (JEMS) **20** (2018), 1075–1138.

Gan71   R. Gangolli, *On the Plancherel formula and the Paley–Wiener theorem for spherical functions on semisimple Lie groups*, Ann. of Math. (2) **93** (1971), 105–165.

Gro96   B. Gross, *On the Satake isomorphism*, in *Galois representations in arithmetic algebraic geometry (Durham, 1996)*, London Mathematical Society Lecture Note Series, vol. 254 (Cambridge University Press, Cambridge), 223–238.

Gro97   B. Gross, *On the motive of a reductive group*, Invent. Math. **130** (1997), 287–313.

Har57   Harish-Chandra, *A formula for semisimple Lie groups*, Amer. J. Math. **79** (1957), 733–760.

Har64   Harish-Chandra, *Some results on an invariant integral on a semi-simple Lie algebra*, Ann. of Math. (2) **80** (1964), 551–593.

Har66   Harish-Chandra, *Discrete series for semisimple Lie groups II*, Acta Math. **116** (1966), 1–111.

Hel00   S. Helgason, *Groups and geometric analysis*, Mathematical Surveys and Monographs, vol. 83. (American Mathematical Society, Providence, RI, 2000).

Hel01   S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Graduate Studies in Mathematics, vol. 34 (American Mathematical Society, Providence, RI, 2001).

Hel08   S. Helgason, *Geometric analysis on symmetric spaces*, Mathematical Surveys and Monographs, vol. 39 (American Mathematical Society, 2008).

HW93   A. G. Helminck and S. P. Wang, *On rationality properties of involutions of reductive groups*, Adv. Math. **99** (1993), 26–97.

IS95   H. Iwaniec and P. Sarnak, L^∞ *norms of eigenfunctions of arithmetic surfaces*, Ann. of Math. (2) **141** (1995), 301–320.

Jac05   H. Jacquet, *Kloosterman identities over a quadratic extension II*, Ann. Sci. Éc. Norm. Supér. (4) **38** (2005), 609–669.

JLR93   H. Jacquet, K. Lai and S. Rallis, *A trace formula for symmetric spaces*, Duke Math. J. **70** (1993), 305–372.

Kna02   A. Knapp, *Lie groups beyond an introduction*, Progress in Mathematics, vol. 140 (Birkhäuser, Boston, 2002).

Kot86   R. Kottwitz, *Stable trace formula: elliptic singular terms*, Math. Ann. **275** (1986), 365–400.

Kot88   R. Kottwitz, *Tamagawa numbers*, Ann. of Math. (2) **127** (1988), 629–646.

LO07   E. Lapid and O. Offen, *Compact unitary periods*, Compos. Math. **143** (2007), 323–338.

Lev52   B. M. Levitan, *On the asymptoptic behavior of the spectral function of a self-adjoint differential equation of second order*, Isv. Akad. Nauk SSSR Ser. Mat. **16** (1952), 325–352.

Mac71   I. G. Macdonald, *Spherical functions on a group of p-adic type*, Publications of the Ramanujan Institute no. 2 (Ramanujan Institute, Centre for Advanced Study in Mathematics, University of Madras, Madras, 1971).

MR03   C. Maclachlan and A. W. Reid, *The arithmetic of hyperbolic 3-manifolds* (Springer, New York, 2003).

Mar14   S. Marshall, *Upper bounds for Maass forms on semisimple groups*, Preprint (2014), arXiv:1405.7033.

MT15   J. Matz and N. Templier, *Sato–Tate equidistribution for families of Hecke–Maass forms on* $\mathrm{SL}(n, \mathbb{R})/\mathrm{SO}(n)$, Preprint (2015), arXiv:1505.07285.

Mil11   D. Milicevic, *Large values of eigenfunctions on arithmetic hyperbolic 3-manifolds*, Geom. Funct. Anal. **21** (2011), 1375–1418.

Nad05   D. Nadler, *Perverse sheaves on real loop Grassmannians*, Invent. Math. **159** (2005), 1–73.

OV94   A. Onishchik and E. Vinberg, *Lie groups and Lie algebras III*, Encyclopaedia of Mathematical Sciences, vol. 41 (Springer, Berlin, 1994).

RS94   Z. Rudnick and P. Sarnak, *The behaviour of eigenstates of arithmetic hyperbolic manifolds*, Comm. Math. Phys. **161** (1994), 195–213.

Sah15   A. Saha, *Large values of newforms on GL(2) with highly ramified central character*, Int. Math. Res. Not. IMRN **2016** (2016), 4103–4131.

Sak08   Y. Sakellaridis, *On the unramified spectrum of spherical varieties over p-adic fields*, Compos. Math. **144** (2008), 978–1016.

Sak13   Y. Sakellaridis, *Spherical functions on spherical varieties*, Amer. J. Math. **135** (2013), 1291–1381.

SV17   Y. Sakellaridis and A. Venkatesh, *Periods and harmonic analysis on spherical varieties*, Astérisque **396** (2017).

Sar95   P. Sarnak, *Arithmetic quantum chaos*, Israel Mathematical Conference Proceedings, Schur Lectures (Bar-Ilan University, Ramat-Gan, Israel, 1995).

Sar04   P. Sarnak, Letter to Morawetz, available at http://publications.ias.edu/sarnak/paper/480.

Sch74   A. I. Schnirelman, *Ergodic properties of eigenfunctions*, Uspekhi Mat. Nauk **29** (1974), 181–182.

ST16   S.-W. Shin and N. Templier, *Sato–Tate theorem for families and low-lying zeros of automorphic L-functions*, Invent. Math. **203** (2016), 1–177.

SZ02   C. Sogge and S. Zelditch, *Riemannian manifolds with maximal eigenfunction growth*, Duke Math. J. **114** (2002), 387–437.

Ste68   R. Steinberg, *Endomorphisms of linear algebraic groups*, Mem. Amer. Math. Soc. **80** (1968).

Tem14   N. Templier, *Large values of modular forms*, Camb. J. Math. **2** (2014), 91–116.

Tit79   J Tits, *Reductive groups over local fields*, in *Automorphic forms, representations and L-functions*, Proceedings of Symposia in Pure Mathematics, vol. 33 (American Mathematical Society, Providence, RI, 1979), 29–70; Part 1.

Vus74   T. Vust, *Opération de groupes réductifs dans un type de cônes presque homogènes*, Bull. Soc. Math. France **102** (1974), 317–333.

Zel87   S. Zelditch, *Uniform distribution of eigenfunctions on compact hyperbolic surfaces*, Duke Math. J. **55** (1987), 919–941.

Farrell Brumley   brumley@math.univ-paris13.fr

Université Sorbonne Paris Nord, Laboratoire de Géométrie, Analyse et Applications, LAGA, CNRS, UMR 7539, F-93430, Villetaneuse, France

Simon Marshall   marshall@math.wisc.edu

Department of Mathematics, University of Wisconsin – Madison, 480 Lincoln Drive, Madison WI 53706, USA