J. Eur. Math. Soc. 24, 3629–3678 (2022) DOI 10.4171/JEMS/1172

Chen Wan

On a multiplicity formula for spherical varieties

Received August 9, 2019

Abstract. In this paper, we propose a conjectural multiplicity formula for general spherical varieties. For all the cases where a multiplicity formula has been proved, including Whittaker models, Gan–Gross–Prasad models, Ginzburg–Rallis models, Galois models and Shalika models, we show that the multiplicity formulas in our conjecture are the same as the multiplicity formulas that have been proved. We also prove the conjectural multiplicity formula in two new cases.

Keywords. Multiplicity of spherical varieties, representation of reductive group over local field

1. Introduction

Let F be a local field of characteristic 0, G be a connected reductive group defined over F, H be a connected closed subgroup of G, and χ be a unitary character of $H(F)$. Assume that H is a spherical subgroup of G (i.e. H admits an open orbit in the flag variety of G). For every irreducible smooth representation π of $G(F)$, we define the multiplicity

$$m(\pi,\chi) := \dim(\mathrm{Hom}_{H(F)}(\pi,\chi)).$$

One of the fundamental problems in the *Relative Langlands Program* is to study the multiplicity $m(\pi,\chi)$. In general, one expects $m(\pi,\chi)$ to be finite and to detect some functorial structures of π. We refer the reader to [19] for a detailed discussion of this kind of problems.

In his pioneering works [21] and [22], Waldspurger developed a new method to study the multiplicities. His idea is to prove a local trace formula $I_{\mathrm{geom}}(f) = I(f) = I_{\mathrm{spec}}(f)$ for the model (G,H), which would imply a multiplicity formula $m(\pi,\chi) = m_{\mathrm{geom}}(\pi,\chi)$. Here $m_{\mathrm{geom}}(\pi,\chi)$ is defined via the Harish-Chandra character θ_π of π and is called the geometric multiplicity. In [21] and [22], Waldspurger applied this method to orthogonal Gan–Gross–Prasad models over the p-adic field. By proving the trace formula and the multiplicity formula, he showed that for orthogonal Gan–Gross–Prasad models, the sum of the multiplicities is always equal to 1 for all tempered local Vogan L-packets. Later

Chen Wan: Department of Mathematics & Computer Science, Rutgers University–Newark, Newark, NJ 07102, USA; chen.wan@rutgers.edu

Mathematics Subject Classification (2020): Primary 22E50

his idea was adapted by Beuzart-Plessis [2], [4] to unitary Gan–Gross–Prasad models, and by the author [23], [24] to Ginzburg–Rallis models. Subsequently, in [3], Beuzart-Plessis applied this method to Galois models; in a joint work with Beuzart-Plessis [5], we applied this method to Shalika models; and in a joint work with Zhang [25], we applied the method to unitary Ginzburg–Rallis models.

For all the cases above, the most crucial step is to prove the local trace formula $I_{\mathrm{geom}}(f) = I(f) = I_{\mathrm{spec}}(f)$. However, the proofs of these trace formulas, especially the geometric side (i.e. $I(f) = I_{\mathrm{geom}}(f)$), have each time been done in some ad hoc way pertaining to the particular features of the case at hand. It makes now little doubt that the local trace formulas and the multiplicity formulas should exist in some generality. However, until this moment, it is not clear (even conjecturally) what would the formulas look like for general spherical varieties. The reason is that although we can easily give a uniform definition of the multiplicity $m(\pi, \chi)$, of the distribution $I(f)$ and of the spectral expansion $I_{\mathrm{spec}}(f)$ for all spherical varieties, the geometric multiplicity $m_{\mathrm{geom}}(\pi, \chi)$ and the geometric expansion $I_{\mathrm{geom}}(f)$ are more mysterious. There are no uniform definitions of these objects for general spherical varieties.

Remark 1.1. The definitions of $m_{\mathrm{geom}}(\pi, \chi)$ and $I_{\mathrm{geom}}(f)$ are very similar to each other. So one only needs to define $m_{\mathrm{geom}}(\pi, \chi)$ for general spherical varieties, which will lead to the definition of $I_{\mathrm{geom}}(f)$.

In this paper, we propose a uniform definition of $m_{\mathrm{geom}}(\pi, \chi)$ (and hence $I_{\mathrm{geom}}(f)$) for general spherical varieties. To justify our definitions, we show that for all the cases where the multiplicity formulas have been proved, including Whittaker models, Gan–Gross–Prasad models, Ginzburg–Rallis models, Galois models, and Shalika models, our definitions are the same as the ones in the known multiplicity formulas. We will also prove the conjectural multiplicity formula for two new cases. We hope our definitions will give a better understanding of the multiplicity formula and the local trace formula, and shed some light on a potential proof of these formulas for general spherical varieties.

1.1. Main results

Let $F, G, H, \chi, m(\pi, \chi)$ be as above. Our goal is to define the geometric multiplicity $m_{\mathrm{geom}}(\pi, \chi)$ for general spherical varieties. Before explaining our definition, let us first consider the baby case when G is a finite group. In this case, let $\theta_\pi(g) = \mathrm{tr}(\pi(g))$ be the character of π. By the representation theory of finite groups, we know that $m(\pi, \chi) = m_{\mathrm{geom}}(\pi, \chi)$ where

$$m_{\mathrm{geom}}(\pi, \chi) := \frac{1}{|H|} \sum_{h \in H} \theta_\pi(h)\chi^{-1}(h) = \sum_x \frac{1}{|Z_H(x)|} \theta_\pi(h)\chi^{-1}(h). \tag{1.1}$$

Here the second summation is over a set of representatives of conjugacy classes of H, and $Z_H(x)$ is the centralizer of x in H.

Guided by the finite group case and all the known cases, it is natural to expect that for a general spherical pair (G, H), $m_{\mathrm{geom}}(\pi, \chi)$ should be an integral over certain semisimple conjugacy classes of $H(F)$ of the Harish-Chandra character θ_π. However, compared with the finite group case, there are three difficulties in the definition of $m_{\mathrm{geom}}(\pi,\chi)$ for spherical varieties over a local field.

First, unlike the finite group case, the Harish-Chandra character θ_π is only defined on the set of regular semisimple elements of $G(F)$. On the other hand, many semisimple conjugacy classes of $H(F)$ are not regular in $G(F)$, which means that θ_π is not defined in those conjugacy classes. In order to solve this issue, we need to use the germ expansions of θ_π. Roughly speaking, near every semisimple element (not necessarily regular) of $G(F)$, θ_π can be written as a linear combination of the Fourier transforms of nilpotent orbital integrals. The coefficients associated to regular nilpotent orbits in this linear combination are called the regular germs of θ_π (see Section 2.4 for details). In order to define θ_π at nonregular semisimple conjugacy classes, we need to use the regular germs of θ_π. This creates the first difficulty: in general when $F\neq\mathbb{C}$, we may have more than one F-rational regular nilpotent orbit. Hence for each spherical pair (G,H), we need to define a subset of regular nilpotent orbits whose regular germs will contribute to the geometric multiplicity. This will be done in Section 6 by using the conjugacy classes in the tangent space of G/H.

Secondly, we need to define the support (i.e. a subset of semisimple conjugacy classes of $H(F)$) of the geometric multiplicity. In the finite group case, the support contains all conjugacy classes of H. But this will not be the case for spherical varieties over local fields. As we will see in Section 4, the geometric multiplicity is only supported on those "elliptic conjugacy classes" $t\in H(F)$ satisfying the following two conditions:

- The centralizers of t in G and H, denoted by (G_t,H_t), form a *minimal spherical variety* (we refer the reader to Section 2.6 for the definition).
- The group G_t is quasi-split over F.

The quasi-split condition provides the existence of regular germs since the existence of regular nilpotent orbits is equivalent to the group being quasi-split. On the other hand, the minimal spherical variety condition on the centralizer (G_t,H_t) ensures that the "homogeneous degree" of the spherical variety $X=G/H$ near t (which is equal to the dimension of H_t minus the dimension of the center) is equal to the homogeneous degree of the regular germs of the Harish-Chandra characters near t (which is equal to the dimension of the maximal unipotent subgroup of G_t). We refer the reader to Section 4 for details.

Thirdly, in the finite group case, we normalize the character θ_π by the number $\frac{1}{|Z_H(x)|}$. For general spherical varieties, we would need an extra number $d(G,H,F)$ which characterizes how the $G(\bar{F})$-conjugacy classes (i.e. stable conjugacy classes) in the tangent space of G/H decompose into $H(F)$-conjugacy classes. We refer the reader to Section 5 for details.

After we have solved the three difficulties above, we are able to write down the definition of $m_{\mathrm{geom}}(\pi,\chi)$ (and hence $I_{\mathrm{geom}}(f)$) for all spherical varieties in Section 7. We will state the conjectural multiplicity formula in Conjecture 7.4. In Section 8, we will show that for all the known cases, our definitions of $m_{\mathrm{geom}}(\pi,\chi)$ are the same as the ones in the known multiplicity formulas.

Theorem 1.2. *Assume that F is p-adic. For Whittaker models, Gan–Gross–Prasad models, Ginzburg–Rallis models, Galois models, or Shalika models, the geometric multiplicities defined in Definition 7.1 are the same as the ones in the multiplicity formulas that have been proved. In particular, Conjecture 7.4 holds for all these models.*

Our proof of Theorem 1.2 uses some Lie algebra version of the local trace formula for Gan–Gross–Prasad models and Ginzburg–Rallis models, as well as a relation between the Shalika germs and Kostant sections proved by Kottwitz (see Lemma 6.10). In general if one can extend Lemma 6.10 to the Archimedean case, then we can also prove Theorem 1.2 when $F = \mathbb{R}$ (the case when $F = \mathbb{C}$ is trivial).

Remark 1.3. Unlike the finite group case, we do not expect the multiplicity formula $m(\pi, \chi) = m_{\mathrm{geom}}(\pi, \chi)$ to hold for all irreducible smooth representations of $G(F)$. For example, in the case when $(G, H) = (\mathrm{GL}_2, \mathrm{GL}_1)$, the geometric multiplicity is just the regular germ of θ_π at the identity element and one can show that the multiplicity formula holds for all generic representations of $G(F) = \mathrm{GL}_2(F)$. However, it is easy to see that the multiplicity formula fails for nongeneric (i.e. finite-dimensional) representations of $\mathrm{GL}_2(F)$.

In general, the multiplicity formula should always hold for all supercuspidal representations. When the spherical pair is tempered, it should hold for all discrete series representations and for almost all tempered representations. When the spherical pair is strongly tempered, it should hold for all tempered representations. We refer the reader to Definition 7.3 for the definitions of tempered and strongly tempered spherical varieties.

Moreover, as observed by Prasad [17], if we want to make the multiplicity formula hold for all irreducible smooth representations of $G(F)$, we need to replace the multiplicity $m(\pi, \chi)$ by the Euler–Poincaré pairing $\mathrm{EP}(\pi, \chi)$. We refer the reader to Section 7 for details.

Finally, all of our discussion so far also makes sense when χ is a finite-dimensional representation of $H(F)$. In particular, we can define the geometric multiplicity and formulate the conjectural multiplicity formula when χ is a finite-dimensional representation of $H(F)$. When F is p-adic, this is not interesting since characters are the only irreducible finite-dimensional representations of $H(F)$. The case we are interested in is when $F = \mathbb{R}$ and $H(\mathbb{R}) = K$ is a maximal connected compact subgroup of $G(\mathbb{R})$. In this case, our definition of the geometric multiplicity $m_{\mathrm{geom}}(\pi, \chi)$ gives a conjectural multiplicity formula $m(\pi, \chi) = m_{\mathrm{geom}}(\pi, \chi)$ for K-types of all irreducible smooth representations of $G(\mathbb{R})$ (note that since $H(\mathbb{R})$ is compact, we have $m(\pi, \chi) = \mathrm{EP}(\pi, \chi)$ for all π). We refer the reader to Sections 7.3 for more details. In Sections 8 and 9, we will prove this conjectural multiplicity formula for K-types for $\mathrm{GL}_n(\mathbb{R})$ and for all complex reductive groups.

Theorem 1.4. *The conjectural multiplicity formula for K-types* (*i.e. Conjecture 7.12*) *holds when*

(1) $G(F) = \mathrm{GL}_n(\mathbb{R})$,

(2) $G = \mathrm{Res}_{\mathbb{C}/\mathbb{R}} H$ *is a complex reductive group.*

In particular, Conjecture 7.4 holds for these two cases.

The key ingredient in our proof of Theorem 1.4 is to show that both the multiplicities and the geometric multiplicities behave nicely under parabolic induction. For the multiplicities, this follows from the Iwasawa decomposition and the reciprocity law. For the geometric multiplicities, it follows from some formulas for the Harish-Chandra characters of induced representations (Proposition 2.7). After we have proved these facts, we can use induction to finish the proof of Theorem 1.4. The upshot is that when $G = \mathrm{GL}_n$ ($n > 2$) or when $G = \mathrm{Res}_{\mathbb{C}/\mathbb{R}} H$ is a nonabelian complex reductive group, the Grothendieck group of finite length smooth representations of $G(\mathbb{R})$ is generated by the induced representations.

1.2. Organization of the paper

The paper is organized as follows: In Section 2, we introduce basic notation and conventions used in this paper. In Section 3, we will use some lower rank examples to explain and motivate our definition of the geometric multiplicity. In Section 4, we will define a subset of conjugacy classes of $H(F)$, which is the support of the geometric multiplicity. In Section 5, we introduce a constant $d(G, H, F)$ associated to minimal spherical varieties. It characterizes how the $G(\bar{F})$-conjugacy classes in the tangent space of G/H decompose into $H(F)$-conjugacy classes. In Section 6, we define a subset of regular nilpotent orbits associated to minimal spherical varieties. The regular germs of these nilpotent orbits will contribute to the geometric multiplicity. Then in Section 7, combining the work of Sections 4–6, we will define the geometric multiplicity $m_{\mathrm{geom}}(\pi, \chi)$ and the geometric expansion of the trace formula $I_{\mathrm{geom}}(f)$ for general spherical varieties. In Section 8, we will show that for all the known cases, our definitions are the same as the ones in the known multiplicity formulas. Finally, in Sections 9 and 10, we will prove the conjectural multiplicity formula for K-types for $\mathrm{GL}_n(\mathbb{R})$ and for all complex reductive groups.

2. Preliminaries

2.1. Notation

Let F be a local field of characteristic 0, and $\psi : F \to \mathbb{C}^\times$ be a nontrivial additive character. Let G be a connected reductive group defined over F, $\mathfrak{g}$ be the Lie algebra of G, Z_G be the center of G, and $A_G(F)$ be the maximal split torus of $Z_G(F)$. We use G_{ss}, G_{reg} (resp. $\mathfrak{g}_{\mathrm{ss}}$, $\mathfrak{g}_{\mathrm{reg}}$) to denote the sets of semisimple and of regular semisimple elements of G (resp. $\mathfrak{g}$). For $x \in G_{\mathrm{ss}}$ (resp. $X \in \mathfrak{g}_{\mathrm{ss}}$), let $Z_G(x)$ (resp. $Z_G(X) = G_X$) be the centralizer of x (resp. X) in G and let G_x be the neutral component of $Z_G(x)$. Similarly, for any abelian subgroup T of G, let $Z_G(T)$ be the centralizer of T in G and let G_T be the neutral component of $Z_G(T)$. We say $x \in G_{\mathrm{reg}}(F)$ is *elliptic* if $G_x(F)$ is a maximal elliptic torus of $G(F)$ (i.e. $G_x(F)/Z_G(F)$ is compact). We use $G_{\mathrm{ell}}(F)$ to denote the set of regular semisimple elliptic elements of $G(F)$. Finally, for $x \in G_{\mathrm{ss}}(F)$ (resp. $X \in \mathfrak{g}_{\mathrm{ss}}(F)$), let $D^G(x) = |\det(1-\mathrm{Ad}(x))_{|\mathfrak{g}/\mathfrak{g}_x}|_F$ (resp. $D^G(X) = |\det(\mathrm{Ad}(X))_{|\mathfrak{g}/\mathfrak{g}_X}|_F$) be the Weyl determinant where $|\ |_F$ is the normalized absolute value on F.

We say a subset $\Omega \subset G(F)$ (resp. $\omega \subset \mathfrak{g}(F)$) is *$G$-invariant* if it is invariant under $G(F)$-conjugation. For any subset $\Omega \subset G(F)$ (resp. $\omega \subset \mathfrak{g}(F)$), we define the G-invariant subset

$$\Omega^G := \{g^{-1}\gamma g \mid g \in G(F), \gamma \in \Omega\}, \quad \omega^G := \{g^{-1}\gamma g \mid g \in G(F), \gamma \in \omega\}.$$

We say a G-invariant subset Ω of $G(F)$ (resp. ω of $\mathfrak{g}(F)$) is *compact modulo conjugation* if there exists a compact subset Γ of $G(F)$ (resp. $\mathfrak{g}(F)$) such that $\Omega \subset \Gamma^G$ (resp. $\omega \subset \Gamma^G$). A *G-domain* on $G(F)$ (resp. $\mathfrak{g}(F)$) is an open subset of $G(F)$ (resp. $\mathfrak{g}(F)$) invariant under $G(F)$-conjugation.

Finally, we fix a minimal Levi subgroup (resp. parabolic subgroup) $M_0(F)$ (resp. $P_0(F) = M_0(F)N_0(F)$) of $G(F)$. We say a parabolic subgroup of $G(F)$ is *standard* if it contains $P_0(F)$. We say a Levi subgroup of $G(F)$ is standard if it is a Levi subgroup of a standard parabolic subgroup and it contains $M_0(F)$. For two Levi subgroups $L_1(F)$ and $L_2(F)$ of $G(F)$, we say that $L_1(F)$ *contains* $L_2(F)$ *up to conjugation* if there exists $g \in G(F)$ such that $L_2(F) \subset gL_1(F)g^{-1}$.

2.2. *Useful function spaces*

Let $C_c^\infty(G(F))$ be the space of smooth compactly supported functions on $G(F)$. We use $\mathcal{C}(G(F))$ to denote the Harish-Chandra–Schwartz space of $G(F)$ (see [4, Section 1.5] for details). On the Lie algebra level, let $C_c^\infty(\mathfrak{g}(F))$ (resp. $\mathcal{S}(\mathfrak{g}(F))$) be the space of smooth compactly supported functions (resp. Schwartz functions) on $\mathfrak{g}(F)$. When F is p-adic, we have $C_c^\infty(\mathfrak{g}(F)) = \mathcal{S}(\mathfrak{g}(F))$.

Let $C_{c,\mathrm{scusp}}^\infty(G(F)) \subset C_c^\infty(G(F))$ be the subspace of strongly cuspidal functions in $C_c^\infty(G(F))$. Similarly we can define the spaces $\mathcal{C}_{\mathrm{scusp}}(G(F))$, $C_{c,\mathrm{scusp}}^\infty(\mathfrak{g}(F))$, and $\mathcal{S}_{\mathrm{scusp}}(\mathfrak{g}(F))$. We refer the reader to [4, Section 5] for the definition and basic properties of strongly cuspidal functions. We say a function $f \in \mathcal{C}(G(F))$ is a *cusp form* if all right translations of f are also strongly cuspidal. We use ${}^\circ\mathcal{C}(G(F))$ to denote the space of cusp forms on $G(F)$.

Finally, we can define the above function spaces with central character. For a unitary character χ of $Z_G(F)$, let $C_c^\infty(G(F),\chi)$ be the Mellin transform of the space $C_c^\infty(G(F))$ with respect to χ. Similarly, we can also define the spaces $\mathcal{C}(G(F),\chi)$, $C_{c,\mathrm{scusp}}^\infty(G(F),\chi)$, $\mathcal{C}_{\mathrm{scusp}}(G(F),\chi)$, and ${}^\circ\mathcal{C}(G(F),\chi)$.

2.3. *Representations*

When F is p-adic, we say a representation π of $G(F)$ is *smooth* if for every $v \in \pi$, the function

$$f : G(F) \to \pi, \quad f(g) = \pi(g)v,$$

is locally constant. When F is Archimedean, we say a representation π of $G(F)$ is *irreducible smooth* (resp. *finite length smooth*) if it is an irreducible (resp. finite length) Casselman–Wallach representation of $G(F)$. We say a finite length smooth representation π of $G(F)$ is an *induced representation* if there exists a proper parabolic subgroup $P = MN$ of G and a finite length smooth representation τ of $M(F)$ such that $\pi = I_P^G(\tau)$. Here $I_P^G(\cdot)$ is the normalized parabolic induction.

We use $\mathcal{R}(G)$ to denote the Grothendieck group of finite length smooth representations of $G(F)$, and we write $\mathcal{R}(G)_{\mathrm{ind}}$ for the subspace of $\mathcal{R}(G)$ generated by the induced representations. The following proposition will be used in the proof of Theorem 1.4.

Proposition 2.1. *Assume that $F = \mathbb{R}$. If $G = \mathrm{GL}_n$ with $n > 2$ or $G = \mathrm{Res}_{\mathbb{C}/\mathbb{R}} H$ where H is a connected reductive group defined over $\mathbb{R}$ that is not abelian, then $\mathcal{R}(G) = \mathcal{R}(G)_{\mathrm{ind}}$. In other words, $\mathcal{R}(G)$ is generated by the induced representations.*

Proof. This follows from the fact that $G_{\mathrm{ell}}(\mathbb{R}) = \emptyset$ when $G = \mathrm{GL}_n$ $(n > 2)$ or when $G = \mathrm{Res}_{\mathbb{C}/\mathbb{R}} H$ where H is a connected reductive group defined over $\mathbb{R}$ that is not abelian. More specifically, since $G_{\mathrm{ell}}(\mathbb{R}) = \emptyset$, $G(\mathbb{R})$ does not have any elliptic representations. This implies that all tempered representations of $G(\mathbb{R})$ are generated by induced representations. Together with the Langlands classification, we get $\mathcal{R}(G) = \mathcal{R}(G)_{\mathrm{ind}}$. ■

2.4. *Quasi-characters and germ expansions*

We fix a nondegenerate, symmetric, G-invariant bilinear form $\langle\ ,\ \rangle$ (i.e. the Killing form) on $\mathfrak{g}$. For any complex valued Schwartz function f on $\mathfrak{g}(F)$, we define its Fourier transform $\hat{f}$ (which is also a Schwartz function on $\mathfrak{g}(F)$) to be

$$\hat{f}(X) = \int_{\mathfrak{g}(F)} f(Y)\psi(\langle X, Y\rangle)\, dY$$

where dY is the selfdual Haar measure on $\mathfrak{g}(F)$ such that $\hat{\hat{f}}(X) = f(-X)$.

Let $\mathrm{Nil}(\mathfrak{g}(F))$ be the set of nilpotent orbits of $\mathfrak{g}(F)$ and $\mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}(F))$ be the set of regular nilpotent orbits of $\mathfrak{g}(F)$. In particular, $\mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}(F))$ is empty unless $G(F)$ is quasi-split. For every $\mathcal{O} \in \mathrm{Nil}(\mathfrak{g}(F))$ and $f \in \mathcal{S}(\mathfrak{g}(F))$, we use $J_{\mathcal{O}}(f)$ to denote the nilpotent orbital integral of f associated to $\mathcal{O}$. Harish-Chandra proved that there exists a unique smooth function $Y \mapsto \hat{j}(\mathcal{O}, Y)$ on $\mathfrak{g}_{\mathrm{reg}}(F)$, which is invariant under $G(F)$-conjugation, and locally integrable on $\mathfrak{g}(F)$, such that for every $f \in \mathcal{S}(\mathfrak{g}(F))$, we have

$$J_{\mathcal{O}}(\hat{f}) = \int_{\mathfrak{g}(F)} f(Y)\hat{j}(\mathcal{O}, Y)\, dY.$$

On the other hand, for $X \in \mathfrak{g}_{\mathrm{reg}}(F)$ and $f \in \mathcal{S}(\mathfrak{g}(F))$, let $J_G(X, f)$ be the orbital integral. Harish-Chandra proved that there exists a unique smooth function $Y \mapsto \hat{j}(X, Y)$ on $\mathfrak{g}_{\mathrm{reg}}(F)$, which is invariant under $G(F)$-conjugation, and locally integrable on $\mathfrak{g}(F)$, such that for every $f \in \mathcal{S}(\mathfrak{g}(F))$, we have

$$J_G(X, \hat{f}) = \int_{\mathfrak{g}(F)} f(Y)\hat{j}(X, Y)\, dY.$$

Definition 2.2. Assume that F is p-adic. Let θ be a smooth function on $G_{\mathrm{reg}}(F)$ that is invariant under $G(F)$-conjugation. We say θ is a *quasi-character* if for every $x \in G_{\mathrm{ss}}(F)$, there is a good neighborhood ω_x of 0 in $\mathfrak{g}_x(F)$, and for every $\mathcal{O} \in \mathrm{Nil}(\mathfrak{g}_x(F))$, there exists $c_{\theta,\mathcal{O}}(x) \in \mathbb{C}$ such that

$$\theta(x \exp(X)) = \sum_{\mathcal{O} \in \mathrm{Nil}(\mathfrak{g}_x(F))} c_{\theta,\mathcal{O}}(x) \hat{j}(\mathcal{O}, X)$$

for every $X \in \omega_{x,\mathrm{reg}}$. The coefficients $\{c_{\theta,\mathcal{O}}(x) \mid \mathcal{O} \in \mathrm{Nil}(\mathfrak{g}_x(F))\}$ (resp. $\{c_{\theta,\mathcal{O}}(x) \mid \mathcal{O} \in \mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}_x(F))\}$) are called the *germs* (resp. *regular germs*) of θ at x.

We refer the reader to [21, Section 3] for the definition of good neighborhoods. Similarly, we can define quasi-characters on a Lie algebra.

Definition 2.3. Assume that F is p-adic. Let θ be a smooth function on $\mathfrak{g}_{\mathrm{reg}}(F)$ that is invariant under $G(F)$-conjugation. We say it is a *quasi-character* on $\mathfrak{g}(F)$ if for every $X \in \mathfrak{g}_{\mathrm{ss}}(F)$, there exists an open G_X-invariant neighborhood $\omega_X \subset \mathfrak{g}_X(F)$ of 0, and for every $\mathcal{O} \in \mathrm{Nil}(\mathfrak{g}_X(F))$, there exists $c_{\theta,\mathcal{O}}(X) \in \mathbb{C}$ such that

$$\theta(X+Y) = \sum_{\mathcal{O} \in \mathrm{Nil}(\mathfrak{g}_X(F))} c_{\theta,\mathcal{O}}(X) \hat{j}(\mathcal{O}, Y)$$

for every $Y \in \omega_{X,\mathrm{reg}}$.

When F is Archimedean, we refer the reader to [4, Sections 4.2–4.4] for the definition of quasi-characters. In this case, the germ expansions become

$$\begin{aligned} &D^G(x \exp(X))^{1/2} \theta(x \exp(X)) \\ &\qquad = D^G(x \exp(X))^{1/2} \sum_{\mathcal{O} \in \mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}_x(F))} c_{\theta,\mathcal{O}}(x) \hat{j}(\mathcal{O}, X) + O(|X|), \\ &D^G(X+Y)^{1/2} \theta(X+Y) \\ &\qquad = D^G(X+Y)^{1/2} \sum_{\mathcal{O} \in \mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}_X(F))} c_{\theta,\mathcal{O}}(X) \hat{j}(\mathcal{O}, Y) + O(|Y|). \end{aligned}$$

The most important examples of quasi-characters on $G(F)$ are the Harish-Chandra characters of finite length smooth representations of $G(F)$. Examples of quasi-characters on $\mathfrak{g}(F)$ are the functions $\hat{j}(X, \cdot)$ ($X \in \mathfrak{g}_{\mathrm{reg}}(F)$) and $\hat{j}(\mathcal{O}, \cdot)$ ($\mathcal{O} \in \mathrm{Nil}(\mathfrak{g}(F))$) defined above.

Definition 2.4. For $X \in \mathfrak{g}_{\mathrm{reg}}(F)$, we use $\Gamma_{\mathcal{O}}(X)$ ($\mathcal{O} \in \mathrm{Nil}(\mathfrak{g}(F))$ in the p-adic case and $\mathcal{O} \in \mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}(F))$ in the Archimedean case to denote the germ of the quasi-character $\hat{j}(X, \cdot)$ at $0 \in \mathfrak{g}(F)$.

The germs $\Gamma_{\mathcal{O}}(X)$ are called *Shalika germs* and we have the germ expansions

$$\hat{j}(X, Y) = \sum_{\mathcal{O} \in \mathrm{Nil}(\mathfrak{g}(F))} \Gamma_{\mathcal{O}}(X) \hat{j}(\mathcal{O}, Y), \quad F \ p\text{-adic},$$

and

$$D^G(X+Y)^{1/2}\hat{j}(X,Y) = D^G(X+Y)^{1/2}\sum_{\mathcal{O}\in \mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}(F))}\Gamma_{\mathcal{O}}(X)\hat{j}(\mathcal{O},Y)+O(|Y|), \quad F \text{ Archimedean},$$

for $Y \in \mathfrak{g}_{\mathrm{reg}}(F)$ close to 0.

Finally, for $f \in \mathcal{C}_{\mathrm{scusp}}(G(F))$ (resp. $f \in \mathcal{S}_{\mathrm{scusp}}(\mathfrak{g}(F))$), let θ_f be the quasi-character on $G(F)$ (resp. $\mathfrak{g}(F)$) defined via the weighted orbital integrals of f. For $f \in \mathcal{S}_{\mathrm{scusp}}(\mathfrak{g}(F))$, let $\hat{\theta}_f = \theta_{\hat{f}}$ be the Fourier transform of θ_f. We refer the reader to [4, Sections 5.2 and 5.6] for details.

2.5. *Regular germs under parabolic induction*

Let π be a finite length smooth representation of $G(F)$ and let θ_π be its Harish-Chandra character.

Definition 2.5. For $x \in G_{\mathrm{ss}}(F)$, define

$$c_\pi(x) = \begin{cases} \frac{1}{|\mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}_x(F))|}\sum_{\mathcal{O}\in \mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}_x(F))} c_{\theta_\pi,\mathcal{O}}(x) & \text{if } \mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}_x(F)) \neq \emptyset; \\ 0 & \text{if } \mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}_x(F)) = \emptyset. \end{cases}$$

Remark 2.6. (1) The set $\mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}_x(F))$ is nonempty if and only if $G_x(F)$ is quasi-split.

(2) For $x \in G_{\mathrm{reg}}(F)$, $c_\pi(x)$ is just $\theta_\pi(x)$.

(3) If $\mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}_x(F))$ only contains a unique element $\mathcal{O}_x$, then $c_\pi(x) = c_{\theta_\pi,\mathcal{O}_x}(x)$.

Let $P = MN$ be a parabolic subgroup of G, τ be a finite length smooth representation of $M(F)$, and $\pi = I_P^G(\tau)$ be the normalized parabolic induction. For all $x \in G_{\mathrm{ss}}(F)$, let $\mathcal{X}_M(x)$ be a set of representatives for the $M(F)$-conjugacy classes of elements in $M(F)$ that are $G(F)$-conjugate to x. The following result was proved in [4, Proposition 4.7.1]; it gives the behavior of $c_\pi(x)$ under parabolic induction.

Proposition 2.7. *For all $x \in G_{\mathrm{ss}}(F)$,*

$$D^G(x)^{1/2}c_\pi(x) = |Z_G(x)(F) : G_x(F)| \sum_{y\in \mathcal{X}_M(x)} |Z_M(y)(F) : M_y(F)|^{-1} D^M(y)^{1/2}c_\tau(y).$$

In particular, $c_\pi(x) = 0$ if the set $\mathcal{X}_M(x)$ is empty.

Remark 2.8. When $G = \mathrm{GL}_n$ or when $x \in G_{\mathrm{reg}}(F)$, the numbers $|Z_G(x)(F) : G_x(F)|$ and $|Z_M(y)(F) : M_y(F)|$ are always equal to 1. Hence the equality above becomes

$$D^G(x)^{1/2}c_\pi(x) = \sum_{y\in \mathcal{X}_M(x)} D^M(y)^{1/2}c_\tau(y).$$

2.6. Spherical subgroups

Let $H \subset G$ be a connected closed subgroup also defined over F. We say that H is a *spherical subgroup* if there exists a Borel subgroup B of G (not necessarily defined over F since $G(F)$ may not be quasi-split) such that BH is Zariski open in G. Such a Borel subgroup is unique up to $H(\bar{F})$-conjugation. If this is the case, then we say (G, H) is a *spherical pair* and $X = G/H$ is a *spherical variety* of G.

From now on, we assume that H is a spherical subgroup. We say the spherical pair (G, H) is *minimal* if the stabilizers of all open Borel orbits are finite modulo the center. In other words, $B \cap H/Z_G \cap H$ is finite for all Borel subgroups $B \subset G$ with BH open in G. Examples of minimal spherical varieties are Whittaker models, Gan–Gross–Prasad models, Ginzburg–Rallis models, and all split symmetric spaces. The following lemma follows from the definition of minimal spherical pairs.

Lemma 2.9. *Assume that (G, H) is a spherical pair. Let $B \subset G$ be a Borel subgroup. Then $\dim(H) - \dim(Z_G \cap H) \geq \dim(G) - \dim(B)$. Moreover, equality holds if and only if (G, H) is minimal. In other words, (G, H) is minimal if and only if the dimension of $H/(H \cap Z_G)$ is equal to the dimension of the maximal unipotent subgroup of G.*

Definition 2.10. Let $P = MN$ be a proper parabolic subgroup of G. For a character $\xi : N(F) \to \mathbb{C}^\times$ of $N(F)$, we use M_ξ to denote the neutral component of the stabilizer of ξ in M (under the adjoint action). For $m \in M(F)$, let ${}^m\xi$ be the character of $N(F)$ defined by ${}^m\xi(n) = \xi(m^{-1}nm)$. We say ξ is a *generic character* if $\dim(M_\xi)$ is minimal, i.e. $\dim(M_\xi) \leq \dim(M_{\xi'})$ for any characters $\xi' : N(F) \to \mathbb{C}^\times$ of $N(F)$.

It is easy to see that if ξ is a generic character, so is ${}^m\xi$ for all $m \in M(F)$. Moreover, there are finitely many generic characters of $N(F)$ up to $M(F)$-conjugation (which are in bijection with the open $M(F)$-orbits in $\mathfrak{n}(F)/[\mathfrak{n}(F), \mathfrak{n}(F)]$ under the adjoint action).

In this paper, we restrict ourselves to the same setting as in [19]. In other words, we consider two types of spherical varieties.

- The reductive case, i.e. H is reductive.
- The Whittaker induction of the reductive case: there exists a parabolic subgroup $P = MN$ of G, and a generic character $\xi : N(F) \to \mathbb{C}^\times$ such that $H = H_0 \ltimes N$ where $H_0 = M_\xi \subset M$ is the neutral component of the stabilizer of ξ in M and H_0 is a reductive spherical subgroup of M.

  In this case, we let $G_0 = M$ and we say that (G, H) is the *Whittaker induction* of (G_0, H_0, ξ). If H is already reductive, we just let $(G_0, H_0, \xi) = (G, H, 1)$. It is easy to see that (G, H) is minimal if and only if (G_0, H_0) is.

Remark 2.11. In general the stabilizer of a generic character is not necessarily reductive (e.g. the parabolic subgroup of GL_3 whose Levi subgroup is $GL_2 \times GL_1$) and also not necessarily a spherical subgroup of M (e.g. the parabolic subgroup of GL_9 whose Levi subgroup is $GL_3 \times GL_3 \times GL_3$).

We use W_G to denote the Weyl group of $G(\bar{F})$. When H is reductive, we use W_X to denote the little Weyl group of the spherical variety $X = G/H$ (defined in [11, pp. 12–13]). The little Weyl group W_X can be identified with a subgroup of W_G. Finally, let $Z_{G,H} = Z_G \cap H$ and $A_{G,H}(F)$ be the maximal split torus of $Z_{G,H}(F)$.

3. Some lower rank examples

In this section, we will give some lower rank examples to motivate and explain the definition of the geometric multiplicity in the next four sections.

Assume that F is a p-adic field. Let $E = F(\sqrt{\delta})$ be a quadratic extension of F, $x \mapsto \bar{x}$ be the conjugation map on E and $N_{E/F}$ (resp. $\mathrm{tr}_{E/F}$) be the norm map (resp. trace map). Let $U_2(F) \subset \mathrm{GL}_2(E)$ be the quasi-split unitary group of two variables defined by

$$U_2(F) = \{g \in \mathrm{GL}_2(E) \mid \bar{g}^t w_2 g = w_2\}, \quad w_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The Lie algebra of $U_2(F)$ has two regular nilpotent orbits $\mathcal{O}_+$ and $\mathcal{O}_-$ with

$$\begin{pmatrix} 0 & i \\ 0 & 0 \end{pmatrix} \in \mathcal{O}_+, \quad \begin{pmatrix} 0 & i\alpha \\ 0 & 0 \end{pmatrix} \in \mathcal{O}_-.$$

Here $i = \sqrt{\delta}$ and $\alpha \in F^\times - \mathrm{Im}(N_{E/F})$. We are going to discuss the multiplicity formulas for five spherical pairs related to the group $U_2(F)$.

Case 1: Let $G(F) = U_2(F)$ and

$$H(F) = \left\{ \begin{pmatrix} 1 & ix \\ 0 & 1 \end{pmatrix} \;\middle|\; x \in F \right\}$$

be a maximal unipotent subgroup of $G(F)$. Up to conjugation, there are two generic characters on the unipotent group $H(F)$ given by

$$\xi_+\left(\begin{pmatrix} 1 & ix \\ 0 & 1 \end{pmatrix}\right) = \psi(x), \quad \xi_-\left(\begin{pmatrix} 1 & ix \\ 0 & 1 \end{pmatrix}\right) = \psi(\alpha x)$$

where ψ is a fixed additive character of F. This gives us two spherical pairs (G, H, ξ_+) and (G, H, ξ_-). The model (G, H, ξ_+) (resp. (G, H, ξ_-)) is the Whittaker induction of $(G_0, H_0, \xi_+) = (T, 1, \xi_+)$ (resp. $(G_0, H_0, \xi_-) = (T, 1, \xi_-)$) where T is the diagonal torus of G. They are the Whittaker models of $U_2(F)$. For an irreducible smooth representation π of $G(F)$, we use $m_+(\pi)$ (resp. $m_-(\pi)$) to denote the multiplicity with respect to the pair (G, H, ξ_+) (resp. (G, H, ξ_-)), i.e.

$$m_+(\pi) = \dim(\mathrm{Hom}_{H(F)}(\pi, \xi_+)), \quad m_-(\pi) = \dim(\mathrm{Hom}_{H(F)}(\pi, \xi_-)).$$

The multiplicity formula in this case was proved by Mœglin and Waldspurger [16, Corollary I.17]:

$$m_+(\pi) = m_{+,\mathrm{geom}}(\pi) := c_{\theta_\pi, \mathcal{O}_+}, \quad m_-(\pi) = m_{-,\mathrm{geom}}(\pi) := c_{\theta_\pi, \mathcal{O}_-}$$

for all irreducible smooth representations of $G(F)$.

Case 2: Let $G(F) = U_2(F) \times U_2(F)$ and

$$H(F) = \{(h,h) \mid h \in U_2(F)\} \simeq U_2(F).$$

Given an irreducible smooth representation $\pi = \pi_1 \otimes \pi_2$ of $G(F)$, let

$$m(\pi) = \dim(\mathrm{Hom}_{H(F)}(\pi, 1))$$

be the multiplicity for the model (G, H). Assume that the central character of π is trivial on $Z_H(F)$ (otherwise the multiplicity is trivially zero). The multiplicity formula in this case was proved by Clozel [6, Theorem 3]:

$$m(\pi) = m_{\mathrm{geom}}(\pi) := \sum_T |W(H,T)|^{-1} \int_{T(F)} D^H(t)\theta_\pi(t)\,dt$$

for all discrete series representations of $G(F)$. Here T runs over a set of representatives of maximal elliptic tori of $H(F)$, and $W(H,T)$ is the Weyl group. It is easy to see that this formula will fail for some non-discrete-series representations.

Remark 3.1. Note that for all the other cases in this section, H is abelian, and that is why the Weyl group and the Wyel determinant do not show up in the geometric multiplicities (because both are trivial in the abelian case).

Case 3 (a): Let $G(F) = U_2(F)$ and $H(F) \simeq U_1(F) \times U_1(F)$ be a maximal elliptic torus of $G(F)$. The model (G, H) is a special case of the unitary Gan–Gross–Prasad models. For an irreducible smooth representation π of $G(F)$ and a character χ of $H(F)$, we use $m(\pi, \chi)$ to denote the multiplicity with respect to the pair (G, H). Assume that the central character of π is equal to the restriction of χ to $Z_G(F) \subset H(F)$ (otherwise the multiplicity is trivially zero). The multiplicity formula in this case was proved by Beuzart-Plessis [4, Theorem 11.2.2]:

$$m(\pi,\chi) = m_{\mathrm{geom}}(\pi,\chi) := \frac{c_{\theta_\pi,\mathcal{O}_+} + c_{\theta_\pi,\mathcal{O}_-}}{2} + \int_{H(F)} \theta_\pi(t)\,dt$$

for all tempered representations of $G(F)$. Moreover, one can easily show that this formula actually holds for all irreducible smooth representations of $G(F)$.

Case 3 (b): If we replace the quasi-split unitary group $U_2(F)$ in Case 3 (a) by the non-quasi-split unitary group $U_2'(F)$, then the Lie algebra of $G(F)$ will no longer have regular nilpotent orbit (because the group is not quasi-split). The multiplicity formula in this case (also proved by Beuzart-Plessis [4, Theorem 11.2.2]) is

$$m(\pi,\chi) = m_{\mathrm{geom}}(\pi,\chi) := \int_{H(F)} \theta_\pi(t)\,dt$$

for all irreducible smooth representations of $G(F)$.

Case 4: Let $G(F) = U_2(F)$ and $H(F) \simeq \mathrm{GL}_1(E)$ be a maximal quasi-split torus of $G(F)$. For an irreducible smooth representation π of $G(F)$ and a character χ of $H(F)$, we use $m(\pi, \chi)$ to denote the multiplicity with respect to the pair (G, H). Assume that the central character of π is equal to the restriction of χ to $Z_G(F) \subset H(F)$ (otherwise the multiplicity is trivially zero). The multiplicity formula in this case can be proved by an argument that is similar to (but much easier than) the argument for the unitary Ginzburg–Rallis model case in [25]:

$$m(\pi, \chi) = m_{\mathrm{geom}}(\pi, \chi) := c_{\theta_\pi, \mathcal{O}_+} + c_{\theta_\pi, \mathcal{O}_-} = 2 \cdot \frac{c_{\theta_\pi, \mathcal{O}_+} + c_{\theta_\pi, \mathcal{O}_-}}{2}$$

for all irreducible infinite-dimensional representations of $G(F)$. On the other hand, it is easy to see that this formula will fail for some finite-dimensional representations of $G(F)$.

For the rest of this section, by using the examples above, we will explain the obstacles in the definition of the geometric multiplicity.

The first difficulty is the support of the geometric multiplicity. In Cases 1 and 3 (a), the geometric multiplicity is supported on all semisimple conjugacy classes of $H(F)$; in Case 2, it is only supported on the regular elliptic conjugacy classes of $H(F)$; in Case 3 (b), it is supported on all semisimple conjugacy classes of $H(F)$ except the center $H(F) \cap Z_G(F)$; in Case 4, it is only supported on the center $H(F) \cap Z_G(F)$.

In general, the geometric multiplicity will be supported on certain "elliptic" conjugacy classes x of $H(F)$ whose centralizers in G and H satisfying the following two conditions:

(1) The group G_x is quasi-split over F.

(2) The pair (G_x, H_x) is a minimal spherical pair.

For Cases 1 and 3 (a), it is easy to see that all the conjugacy classes satisfy these conditions. For Case 2, $H(F)$ has three types of conjugacy classes: the center, regular non-elliptic conjugacy classes, and regular elliptic conjugacy classes. For $x \in Z_H(F)$, we have $(G_x, H_x) = (G, H)$, which is not a minimal spherical pair. That is why the geometric multiplicity is not supported on the central elements. For a regular non-elliptic conjugacy class $x \in H(F)$, the centralizer (G_x, H_x) is a minimal spherical pair, but x is not "elliptic" and hence the geometric multiplicity is not supported on it. As a result, for Case 2, the geometric multiplicity is only supported on the regular elliptic conjugacy classes. For Case 3 (b), all conjugacy classes of $H(F)$ satisfy the "elliptic" condition and the minimal spherical pair condition. But when x belongs to the center $H(F) \cap Z_G(F)$, $G_x(F) = G(F)$ is not quasi-split. Hence the geometric multiplicity is supported on all semisimple conjugacy classes of $H(F)$ except the center $H(F) \cap Z_G(F)$. Finally, for Case 4, all conjugacy classes of $H(F)$ satisfy the quasi-split condition and the minimal spherical pair condition. But if the conjugacy class does not belong to the center $H(F) \cap Z_G(F)$, it violates the "elliptic" condition and that is why the geometric multiplicity is only supported on the center $H(F) \cap Z_G(F)$.

We refer the reader to Section 4 for a detailed definition of the support of the geometric multiplicity.

The second difficulty is to determine which regular germs will contribute to the geometric multiplicity. In Cases 3 (a) and 4, the germs associated to both regular nilpotent orbits of $\mathfrak{g}(F)$ contribute to the geometric multiplicity. On the other hand, in Case 1, only one of regular germs contributes to the geometric multiplicity. This will be discussed in Section 6. We will use the conjugacy classes in the tangent space of the spherical variety $X = G/H$ and the Kostant sections associated to regular nilpotent orbits to determine which regular germs will contribute to the geometric multiplicity. Roughly speaking, the regular germ associated to a regular nilpotent orbit will contribute to the geometric multiplicity if and only if certain conjugacy classes in the Kostant section associated to this nilpotent orbit are contained in the tangent space of the spherical variety $X = G/H$. We refer the reader to Section 6 for more details.

The third obstacle is an extra factor of the regular germs. For Cases 1–3, we just have the regular germs (or the average of the regular germs in Case 3 (a)); while for Case 4, we have the average of the regular germs times 2. So this extra factor is equal to 1 for Cases 1–3, and to 2 for Case 4. This extra factor is related to the number of open Borel orbits in $G(F)/H(F)$, the Weyl group of $G(\bar{F})$ and the little Weyl group of the spherical variety $X = G/H$. Another way to explain this factor is that it characterizes how the stable conjugacy classes in the tangent space of X decompose into rational $H(F)$-conjugacy classes. We refer the reader to Section 5 for more details.

Lastly, as shown in the examples above, the multiplicity formula may not work for all the smooth irreducible representations, sometimes it only works for discrete series or tempered representations. This is related to certain analytic behaviors of spherical varieties and we refer the reader to Section 7.1 for more details.

4. The support of geometric multiplicity

In this section, let (G, H) be a spherical pair which is the Whittaker induction of the reductive spherical pair (G_0, H_0, ξ). Recall that when H is reductive, we let $(G_0, H_0, \xi) = (G, H, 1)$. We are going to define a subset of semisimple conjugacy classes of $H_0(F)$, which will be the support of the geometric multiplicity. We will also define a measure on this subset.

Definition 4.1 (the support of geometric multiplicity). Let $\mathcal{S}(G, H)$ be the set of $H_0(F)$-conjugacy classes $x \in H_0(F)$ satisfying the following three conditions:

(1) (elliptic condition) The quotient $(A_{G_x}(F) \cap H(F))/A_{G,H}(F)$ is compact.

(2) The pair (G_x, H_x) is a minimal spherical pair.

(3) The group $G_x(F)$ is quasi-split.

The set $\mathcal{S}(G, H)$ is the *support of* the geometric multiplicity.

Remark 4.2. For a semisimple conjugacy class $x \in H_0(F)$, the distribution of the local trace formula associated to (G, H) has homogeneous degree $\dim(H_x) - \dim(Z_{G_x} \cap H_x)$ near x. Meanwhile, the germ expansion in Section 2.4 tells us that near x_0, every Harish-Chandra character is a combination of distributions with homogeneous degrees less than or equal to the dimension of the maximal unipotent subgroup of G_x, while equality only occurs when $G_x(F)$ is quasi-split and the distributions are associated to the regular nilpotent orbits. As a result, near x, the homogeneous degree of the distribution of the local trace formula associated to (G, H) will always be greater than or equal to the homogeneous degrees of the distributions in the germ expansions of the Harish-Chandra characters, and equality only occurs when the pair (G_x, H_x) is minimal, $G_x(F)$ is quasi-split, and the distributions are associated to the regular nilpotent orbits (see Lemma 2.9). That is why we have the second and third conditions in the definition. That is also why only the regular germs of Harish-Chandra characters will contribute to the geometric multiplicity.

In order to define a measure on $\mathcal{S}(G, H)$, we will give an equivalent definition of $\mathcal{S}(G, H)$. More precisely, we will define $\mathcal{S}(G, H)$ as a union of translations of subtori of $H_0(F)$. Then we can define a measure on $\mathcal{S}(G, H)$ by using the Haar measures on the subtori.

Definition 4.3. Let $\mathcal{T}(G, H)$ be the set of all those closed (not necessarily connected) abelian subgroups $T(F)$ of $H_0(F)$ (up to $H_0(F)$-conjugation) that satisfy the following four conditions.

(1) Every element of $T(F)$ is semisimple and (G_T, H_T) is a minimal spherical variety with $G_T(F)$ quasi-split.

(2) $T(F) = Z_{Z_G(T)}(F) \cap H(F)$ where $Z_{Z_G(T)}(F)$ is the center of $Z_G(T)(F)$. In particular, we have $Z_{G,H}(F) \subset T(F)$ and $A_{G,H}(F) \subset T^\circ(F)$. Here $T^\circ(F)$ is the neutral component of $T(F)$, which is a subtorus of $H_0(F)$.

(3) The quotient $T(F)/Z_{G,H}(F)$ (or equivalently $T^\circ(F)/A_{G,H}(F)$) is compact. This is equivalent to $H(F) \cap A_{G_T}(F)/A_{G,H}(F)$ being finite.

(4) There exists $t \in T(F)$ such that $(G_t, H_t) = (G_T, H_T)$.

Let $\mathcal{T}(G, H)^\circ = \{T(F) \in \mathcal{T}(G, H) \mid T(F) = T^\circ(F) Z_{G,H}(F)\}$.

Remark 4.4. Condition (1) in Definition 4.3 is an analogue of conditions (2) and (3) in Definition 4.1, while condition (3) in Definition 4.3 is an analogue of condition (1) in Definition 4.1. Condition (4) ensures that $T(F)$ contains no elements of the support $\mathcal{S}(G, H)$ while (2) ensures that $T(F)$ is large enough to contain all elements of $\mathcal{S}(G, H)$.

For $T(F) \in \mathcal{T}(G, H)$, there exists a nonempty (this follows from Definition 4.3 (4)) subset $C(T, H)$ of the component group $T(F)/T^\circ(F)$ satisfying the following two conditions:

- For $\gamma \in C(T, H)$, $(G_t, H_t) = (G_T, H_T)$ for almost all $t \in \gamma T^\circ(F)$.
- For $\gamma \in T(F)/T^\circ(F) - C(T, H)$, $(G_t, H_t) \neq (G_T, H_T)$ for all $t \in \gamma T^\circ(F)$.

Definition 4.5. For $T(F) \in \mathcal{T}(G, H)$, let $T_H(F) = \bigcup_{\gamma \in C(T,H)} \gamma T^\circ(F) \subset T(F) \subset H_0(F)$. Let $T_H(F)'$ be the Zariski open subset of $T_H(F)$ consisting of those elements $t \in T_H(F)$ such that $(G_t, H_t) = (G_T, H_T)$.

Remark 4.6. For $T(F) \in \mathcal{T}(G, H)^\circ$, $(G_t, H_t) = (G_T, H_T)$ for almost all $t \in T(F)$, which implies that $T_H(F) = T(F)$.

Lemma 4.7. *The support* $\mathcal{S}(G, H)$ *of the geometric multiplicity is equal to* $\bigcup_{T(F)\in\mathcal{T}(G,H)} T_H(F)'$.

Proof. From the definition it is clear that $T_H(F)'$ is contained in $\mathcal{S}(G, H)$ for all $T(F) \in \mathcal{T}(G, H)$. For the other direction, given $t \in \mathcal{S}(G, H)$, let $T(F) = Z_{Z_G(t)}(F) \cap H(F)$. Then it is easy to see that $T(F) \in \mathcal{T}(G, H)$ and $t \in T_H(F)'$. This proves the lemma. ■

Remark 4.8. The lemma above gives us a natural measure on $\mathcal{S}(G, H)$. More specifically, since $T_H(F)$ is a finite union of translations of subtori $T^\circ(F)$, the Haar measure on $T^\circ(F)$ induces a measure on $T_H(F)$ such that $T_H(F) - T_H(F)'$ has measure zero (because $T_H(F)'$ is a Zariski open subset of $T_H(F)$). This gives us a measure on $T_H(F)'$ and hence a measure on $\mathcal{S}(G, H)$.

For example, for the model $(G, H) = (U_2 \times U_2, U_2)$ in the previous section, the geometric multiplicity is supported on the elliptic regular semisimple conjugacy classes of $H(F)$. The set $\mathcal{T}(G, H)$ is equal to $\mathcal{T}_{\mathrm{ell}}(H)$, a set of representatives of maximal elliptic tori of $H(F)$. For $T(F) \in \mathcal{T}(G, H)$, we have $T_H(F) = T(F)$, and $T_H(F)' = T_{\mathrm{reg}}(F)$ is the set of regular semisimple elements in $T(F)$ (which is a Zariski open subset). The measure on $T_H(F)' = T_{\mathrm{reg}}(F)$ is induced from the Haar measure on the torus $T(F) = T^\circ(F) = T_H(F)$.

Remark 4.9. For $t \in H_{0,\mathrm{ss}}(F)$, (G_t, H_t) is the Whittaker induction of $(G_{0,t}, H_{0,t}, \xi)$. Hence $\mathcal{S}(G, H) = \mathcal{S}(G_0, H_0)$, $\mathcal{T}(G, H) = \mathcal{T}(G_0, H_0)$ and $T_H(F) = T_{H_0}(F)$ for all $T(F) \in \mathcal{T}(G, H) = \mathcal{T}(G_0, H_0)$. In other words, the geometric multiplicity of (G, H) has the same support as the geometric multiplicity of (G_0, H_0).

Remark 4.10. When the spherical variety $X = G/H$ has no Type N spherical root (we refer the reader to [19, Section 3.1] for the definitions of spherical roots and Type N spherical roots), we expect that (although we cannot prove it at this moment) $T(F) = T^\circ(F)Z_{G,H}(F)$ for all $T(F) \in \mathcal{T}(G, H)$ (i.e. $\mathcal{T}(G, H) = \mathcal{T}(G, H)^\circ$). In other words, the geometric multiplicity is essentially supported on some subtori of $H_0(F)$. On the other hand, when $X = G/H$ has a Type N spherical root, the geometric multiplicity may be supported on some translations of subtori of $H_0(F)$.

For example, as we will see in Section 9, the geometric multiplicity of the model $(\mathrm{GL}_n(\mathbb{R}), \mathrm{SO}_n(\mathbb{R}))$ (which has a Type N spherical root when $n > 2$) is supported on the set (not necessarily connected when $n > 2$)

$$\{\mathrm{diag}(I_{n_1}, -I_{2n_2}, t) \mid t \in T(\mathbb{R})\}$$

where (n_1, n_2) runs over the set

$$I(n_1, n_2) := \{(n_1, n_2) \in \mathbb{Z}_{\geq 0} \mid n - n_1 - 2n_2 \text{ is a nonnegative even number}\}$$

and $T(\mathbb{R})$ is a maximal elliptic torus of $\mathrm{SO}_{n-n_1-2n_2}(\mathbb{R})$. In particular, when $n > 2$, the support of the geometric multiplicity contains some translations of subtori of $\mathrm{SO}_n(\mathbb{R})$. The multiplicity formula for this case will be proved in Section 9.

5. The constant $d(G, H, F)$ for minimal spherical varieties

In this section, we assume that (G, H) is a minimal spherical pair with H reductive. Moreover, we assume that G is quasi-split over F. Then we can find a Borel subgroup $B = TN \subset G$ defined over F such that BH is open in G and $B \cap H$ is finite modulo the center. The goal of this section is to define a constant positive integer $d(G, H, F)$ associated to the spherical pair. This constant is the extra factor for the regular germs in the formula for geometric multiplicity. We will first define this constant using the number of open Borel orbits and the Weyl groups. Then we will show that this number also characterizes how the stable conjugacy classes in the tangent space of the spherical variety $X = G/H$ decompose into $H(F)$-conjugacy classes. We will also define another constant $c(G, H, F)$ which is an analogue of the stabilizer $|Z_H(x)|$ for the finite group case (1.1).

We use $\mathfrak{g}, \mathfrak{z} = \mathfrak{z}_{\mathfrak{g}}, \mathfrak{h}, \mathfrak{b}, \mathfrak{t}, \mathfrak{n}$ to denote the Lie algebras of G, Z_G, H, B, T, N. By our choice of H and B, we have

$$\mathfrak{h} \cap \mathfrak{b} = \mathfrak{h} \cap \mathfrak{z}, \quad \mathfrak{g} = \mathfrak{h} + \mathfrak{b}.$$

Let $\mathfrak{h}' = \{X \in \mathfrak{h} \mid \langle X, Y\rangle = 0 \text{ for all } Y \in \mathfrak{z} \cap \mathfrak{h}\}$ and $\mathfrak{h}^{\perp} = \{X \in \mathfrak{g} \mid \langle X, Y\rangle = 0 \text{ for all } Y \in \mathfrak{h}'\}$. The space $\mathfrak{h}^{\perp}$ can be viewed as the tangent space of the spherical variety G/H at the identity component $1 \cdot H$. We have

$$\mathfrak{h} = \mathfrak{h}' \oplus (\mathfrak{z} \cap \mathfrak{h}), \quad \mathfrak{g} = \mathfrak{h}' \oplus \mathfrak{b}, \quad \mathfrak{g} = \mathfrak{h}^{\perp} \oplus \mathfrak{n}.$$

Let $\mathfrak{t}_H$ be the image of $\mathfrak{t}$ under the projection map $\mathfrak{g} = \mathfrak{h}^{\perp} \oplus \mathfrak{n} \to \mathfrak{h}^{\perp}$. Then $\dim(\mathfrak{t}_H) = \dim(\mathfrak{t}) - \dim(\mathfrak{t} \cap \mathfrak{n}) = \dim(\mathfrak{t})$ and we have $\mathfrak{t}_H = \mathfrak{b} \cap \mathfrak{h}^{\perp}$. In particular, $\mathfrak{t}_H$ is independent of the choice of T.

Lemma 5.1. *If $\mathfrak{t}_{\mathrm{reg}} \cap \mathfrak{t}_H \neq \emptyset$, then $H \cap B \subset T$. In particular, $H \cap B$ is abelian.*

Proof. Fix $t \in \mathfrak{t}_{\mathrm{reg}} \cap \mathfrak{t}_H$. Let $\gamma \in H \cap B$. In order to show that $\gamma \in T$, it is enough to show that γ commutes with t. Since $\gamma \in B$, we know that $\gamma t \gamma^{-1} = t + n$ for some $n \in \mathfrak{n}$. Since $\gamma \in H$ and $t \in \mathfrak{h}^{\perp}$, we know that $t + n = \gamma t \gamma^{-1} \in \mathfrak{h}^{\perp}$. This implies that $n = 0$. Hence γ commutes with t. This proves the lemma. ∎

Definition 5.2. Let $c(G, H, F)$ be the number of connected components of $B(F) \cap H(F)$.

Lemma 5.3. *The number $c(G, H, F)$ is independent of the choice of B.*

Proof. Let $B = TN$ and $B' = T'N'$ be two Borel subgroups of G defined over F with BH and $B'H$ being Zariski open in G. In order to prove the lemma, it is enough to show that the group $B(F) \cap H(F)$ is isomorphic to the group $B'(F) \cap H(F)$.

By Lemma 5.1, up to conjugating T (resp. T') by an element of $N(F)$ (resp. $N'(F)$), we may assume that $B \cap H \subset T$ (resp. $B' \cap H \subset T'$). Since BH and $B'H$ are Zariski open in G, there exists $h \in H(\bar{F})$ such that $B = h^{-1} B' h$. Then the morphism

$$B' \cap H \ni t \mapsto h^{-1} t h \in B \cap H$$

is an isomorphism. So it is enough to show that for all $t \in B'(F) \cap H(F)$, we have $h^{-1}th \in B(F) \cap H(F)$.

For $\sigma \in \mathrm{Gal}(\bar{F}/F)$, since both B and B' are defined over F, we have $h^{-1}B'h = B = \sigma(h)^{-1}B'\sigma(h)$. This implies that $B' = h\sigma(h)^{-1}B'\sigma(h)h^{-1}$. Hence $h\sigma(h)^{-1} \in B' \cap H' \subset T'$. Together with the fact that $B'(F) \cap H(F) \subset T'(F)$, we have

$$\sigma(h^{-1}th) = \sigma(h)^{-1}t\sigma(h) = h^{-1}(h\sigma(h)^{-1}t\sigma(h)h^{-1})h = h^{-1}th$$

for all $t \in B'(F) \cap H(F)$. This implies that $h^{-1}th \in B(F) \cap H(F)$. ■

The lemma above shows that the constant $c(G, H, F)$ is well defined, i.e. it only depends on the groups G, H and the field F. Now we define the constant $d(G, H, F)$. We start with a lemma about the open Borel orbits.

Lemma 5.4. *There is a bijection between open orbits in $B(F)\backslash G(F)/H(F)$ and* $\ker(H^1(F, H \cap B) \to H^1(F, H))$. *We use $d'(G, H, F)$ to denote the number of open orbits in $B(F)\backslash G(F)/H(F)$.*

Proof. Let $X = BH$, which is an open subvariety of G. Then open orbits in $B(F)\backslash G(F)/H(F)$ are just the orbits in $B(F)\backslash X(F)/H(F)$. Let $B(F)\backslash X(F)/H(F) = \bigcup_{i=1}^{l} B(F)\gamma_i H(F)$. For each i, there exist $b_i \in B(\bar{F})$ and $h_i \in H(\bar{F})$ such that $\gamma_i = b_i h_i$. Then it is easy to see that the map

$$\mathrm{Gal}(\bar{F}/F) \ni \sigma \mapsto b_i^{-1}\sigma(b_i) = h_i\sigma(h_i)^{-1} \in H \cap B$$

is a cocycle whose image in $H^1(F, H \cap B)$ only depends on the orbit $B(F)\gamma_i H(F)$. Also by definition, this cocycle becomes a coboundary in H. This gives a well defined map from $B(F)\backslash X(F)/H(F)$ to $\ker(H^1(F, H \cap B) \to H^1(F, H))$. One can easily check that this map is a bijection. ■

Definition 5.5. We define

$$d(G, H, F) = d'(G, H, F) \times \frac{|W_G|}{|W_X|}.$$

Recall that W_X is the little Weyl group of the spherical variety $X = G/H$ and W_G is the Weyl group of $G(\bar{F})$.

Remark 5.6. Since (G, H) is a minimal spherical pair, it is wavefront if and only if $W_G = W_X$. If this is the case, we have

$$d(G, H, F) = d'(G, H, F) = |\ker(H^1(F, H \cap B) \to H^1(F, H))|.$$

We refer the reader to [19, Section 2.1] for the definition of wavefront spherical varieties.

Remark 5.7. For all the models considered in Section 3, the constant $d(G, H, F)$ is equal to $d'(G, H, F)$ since all the models there are symmetric pairs (in particular, wavefront).

For all the models (G, H) in Cases 1–3 of Section 3 and for all $t \in \mathcal{S}(G, H)$, one can easily see that the spherical pair (G_t, H_t) has only one open Borel orbit. That is why the constant $d(G, H, F)$ is equal to 1 for all these cases. For Case 4, the spherical pair (G, H) has two open Borel orbits corresponding to $F^\times/\mathrm{Im}(N_{E/F})$ and hence $d(G, H, F) = 2$ for this case. That is why in the geometric multiplicity for Case 4, we have the average of the regular germs times 2.

Remark 5.8. Here is an example of a non-wavefront spherical pair. Consider the pair $(G, H) = (\mathrm{GL}_3, \mathrm{SL}_2)$. It is easy to see that there is only one open orbit in $B(F)\backslash G(F)/H(F)$, i.e. $d'(G, H, F) = 1$. On the other hand, the Weyl group W_G is equal to S_3 while the little Weyl group W_X is equal to S_2 (see [12, Table 3]). As a result, we have

$$d(G, H, F) = d'(G, H, F) \times \frac{|W_G|}{|W_X|} = 1 \times 3 = 3.$$

The rest of this subsection is to study the relation between the number $d(G, H, F)$ and the slice representation (i.e. the conjugation action of $H(F)$ on the tangent space $\mathfrak{h}^\perp(F)$). We are going to show that almost all quasi-split regular semisimple $G(\bar{F})$-conjugacy classes (i.e. stable conjugacy classes) in $\mathfrak{h}^\perp(F)$ break into $d(G, H, F)$-many $H(F)$-conjugacy classes.

Lemma 5.9. *There exists a W_G-invariant Zariski open subset $\mathfrak{t}^0$ of $\mathfrak{t}_{\mathrm{reg}}$ such that for all $t \in \mathfrak{t}^0(\bar{F})$, the $G(\bar{F})$-conjugacy class of t in $\mathfrak{h}^\perp(\bar{F})$ breaks into $\frac{|W_G|}{|W_X|}$-many $H(\bar{F})$-conjugacy classes.*

Proof. By quotienting H and G by the center $Z_{G,H} = H \cap Z_G$, we may assume that $H \cap Z_G = \{1\}$. Then $B \cap H$ is finite. We denote by $\mathcal{X}(T)$ the group of rational characters of T, and define $\mathfrak{a} = \mathrm{Hom}(\mathcal{X}(T), \mathbb{R})$. Let $\mathcal{X}(X)$ be the group of T-eigencharacters on $\bar{F}(X)^{(B)}$ where $\bar{F}(X)^{(B)}$ is the multiplicative group of nonzero B-eigenfunctions on $\bar{F}(X)$ and $\bar{F}(X)$ is the field of rational functions on $X(\bar{F})$. Finally, let $\mathfrak{a}_X = \mathrm{Hom}(\mathcal{X}(X), \mathbb{R})$. Since $H \cap B$ is finite, we have $\mathfrak{a} = \mathfrak{a}_X$. Let $\mathfrak{a}^* = \mathfrak{a}_X^*$ be the dual of $\mathfrak{a} = \mathfrak{a}_X$, and let $T^*X = \mathfrak{h}^\perp \times_H G$ be the cotangent bundle of X. By [11, Satz 7.1 and Korollar 7.2], we have $\mathfrak{h}^\perp /\!/ H = T^*X /\!/ G = \mathfrak{a}_X^* /\!/ W_X = \mathfrak{a}^* /\!/ W_X$. Meanwhile, $\mathfrak{g} /\!/ G = \mathfrak{a}^* /\!/ W_G$. This proves the lemma. ∎

Remark 5.10. When (G, H) is a symmetric pair (which is wavefront), we have $W_G = W_X$. By the work of Kostant–Rallis [14, Theorem 1], we can even take $\mathfrak{t}^0$ to be $\mathfrak{t}_{\mathrm{reg}}$. Examples of non-wavefront minimal spherical pairs are $(\mathrm{SO}_{2n+1}, \mathrm{GL}_n)$ and $(\mathrm{GL}_{2n+1}, \mathrm{Sp}_{2n})$.

Definition 5.11. We define $\mathfrak{h}^{\perp,0}$ to be the set of elements in $\mathfrak{h}^\perp$ that are G-conjugate to an element in $\mathfrak{t}^0$.

Since $\mathfrak{t}^0$ is Zariski open in $\mathfrak{t}_{\mathrm{reg}}$, we know that $\mathfrak{h}^{\perp,0}$ is a Zariski open subset of $\mathfrak{h}^\perp$. By Lemma 5.9, each $G(\bar{F})$-conjugacy class in $\mathfrak{h}^{\perp,0}(\bar{F})$ breaks into $\frac{|W_G|}{|W_X|}$-many $H(\bar{F})$-conjugacy classes.

Lemma 5.12. *For every $t \in \mathfrak{t}_H(F)$ regular semisimple, the $H(\bar{F})$-conjugacy class of t in $\mathfrak{h}^{\perp}(F)$ breaks into $d'(G,H,F)$-many $H(F)$-conjugacy classes.*

Proof. By conjugating T we may assume that $t \in \mathfrak{t}_{\mathrm{reg}}(F)$. By Lemma 5.1, we know that $H \cap B \subset T$. Let $t' \in \mathfrak{h}^{\perp}(F)$ be $H(\bar{F})$-conjugate to t. Then there exists $h \in H(\bar{F})$ such that $ht'h^{-1} = t$. For all $\sigma \in \mathrm{Gal}(\bar{F}/F)$, we have

$$\sigma(h)t'\sigma(h)^{-1} = ht'h^{-1} = t.$$

In particular, $\sigma(h)h^{-1}$ commutes with t. This implies that $\sigma(h)h^{-1} \in H \cap T = H \cap B$. Then it is easy to see that the map

$$\mathrm{Gal}(\bar{F}/F) \ni \sigma \mapsto \sigma(h)h^{-1} \in H \cap B$$

is a cocycle whose image in $H^1(F, H \cap B)$ only depends on the $H(F)$-conjugacy class of t'. Also it is easy to see that this cocycle becomes a coboundary in H. This gives a well defined map from the set of $H(F)$-conjugacy classes in the $H(\bar{F})$-conjugacy class of t in $\mathfrak{h}^{\perp}(F)$ to $\ker(H^1(F, T_0) \to H^1(F, H))$. One can easily check that this map is a bijection. ∎

Combining the lemmas above, we have proved the following proposition.

Proposition 5.13. *For every $t \in \mathfrak{h}^{\perp,0}(F)$, if $G_t(F)$ is a maximal quasi-split torus of $G(F)$ (i.e. the conjugacy class of t is "quasi-split"), then the $G(\bar{F})$-conjugacy class of t (i.e. the stable conjugacy class of t) in $\mathfrak{h}^{\perp}(F)$ breaks into $d(G,H,F) = d'(G,H,F) \times \frac{|W_G|}{|W_X|}$-many $H(F)$-conjugacy classes.*

Remark 5.14. If $H \cap B \subset Z_G$, then by the same argument as above, we can even show that every $G(\bar{F})$-conjugacy class (not necessarily quasi-split) in $\mathfrak{h}^{\perp,0}(F)$ breaks into $d(G,H,F)$-many $H(F)$-conjugacy classes.

Remark 5.15. In general, if (G,H) is the Whittaker induction of (G_0, H_0, ξ) with (G_0, H_0) minimal, we can also define an analogue of the space $\mathfrak{h}^{\perp}(F)$ by adding the information of ξ (see Section 6.4). We will denote this space by $\Xi + \mathfrak{h}_0^{\perp}(F) + \mathfrak{n}(F)$; we are still interested in how the stable conjugacy classes in $\Xi + \mathfrak{h}_0^{\perp}(F) + \mathfrak{n}(F)$ decompose into $H(F)$-conjugacy classes.

For most known cases, the stable conjugacy classes in $\Xi + \mathfrak{h}_0^{\perp}(F) + \mathfrak{n}(F)$ coincide with the $H(F)$-conjugacy classes, i.e. $d(G_0, H_0, F) = 1$. In other words, two regular semisimple elements in $\Xi + \mathfrak{h}_0^{\perp}(F) + \mathfrak{n}(F)$ are $G(\bar{F})$-conjugate to each other if and only if they are $H(F)$-conjugate to each other. For Whittaker models, this follows from the theory of Kostant sections ([13, Proposition 19], see also the summary in [15, Section 2.4]). For Gan–Gross–Prasad models, it was proved in [21, Section 9] (the orthogonal case) and in [4, Section 10] (the unitary case). For Ginzburg–Rallis models, it was proved in [23, Section 8]. This property is crucial in the proofs of the local trace formula for those cases.

The only exception among the known cases is the Ginzburg–Rallis model for the unitary group (see Section 8.3). In that case, the number $d(G_0, H_0, F)$ is equal to 2, which means that every $G(\bar{F})$-conjugacy class in $\Xi + \mathfrak{h}_0^{\perp}(F) + \mathfrak{n}(F)$ breaks into two $H(F)$-conjugacy classes. However, although we have proved the multiplicity formula for this model in [25], it was not proved by the trace formula method. Instead, we first considered the Ginzburg–Rallis model for the unitary similitude group (where $d(G_0, H_0, F) = 1$). We proved the trace formula and the multiplicity formula for the unitary similitude group case. Hence we deduced the multiplicity formula for the unitary group case.

Hence if one wants to prove the multiplicity formula and the local trace formula for general spherical varieties, one of the important steps is to develop a method to deal with the cases when $d(G_0, H_0, F) \neq 1$. Roughly speaking, we need to “stabilize” the trace formula.

6. Nilpotent orbits associated to minimal spherical varieties

The goal of this section is to solve the last obstacle in the definition of geometric multiplicity. We will determine the regular germs that contribute to the geometric multiplicity. Let (G, H) be a minimal spherical pair with $G(F)$ quasi-split. The goal is to define a subset $\mathcal{N}(G, H, \xi)$ (note that $\xi = 1$ when H is reductive) of $\mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}(F))$.

In Section 6.1, we define a notion of *null* conjugacy classes which plays a key role in our definition of $\mathcal{N}(G, H, \xi)$. Then in Section 6.2, we discuss the conjugacy classes associated to regular nilpotent orbits (i.e. Kostant sections). Finally, we define $\mathcal{N}(G, H, \xi)$ in Section 6.3 for the reductive case and in Section 6.4 for the nonreductive case.

6.1. Null conjugacy classes

Definition 6.1. Let $\mathcal{L}(G, H)$ be the set of standard Levi subgroups $L(F)$ of $G(F)$ satisfying the following condition:

- There exists $T(F) \in \mathcal{T}(G, H)^{\circ}$ with $T(F) \neq Z_{G,H}(F)$ such that $L(F)$ is conjugate to the Levi subgroup $Z_G(A_T)(F)$ where $A_T(F)$ is a maximal split torus of $G_T(F)$.

Here the set $\mathcal{T}(G, H)^{\circ}$ is defined in Section 4.

Definition 6.2. For $t \in G_{\mathrm{reg}}(F)$, let $T(F) = G_t(F)$, $A_T(F)$ be the maximal split subtorus of $T(F)$, and $L(t)(F) = Z_G(A_T)(F)$, which is a Levi subgroup of $G(F)$. In particular, t is elliptic regular if and only if $L(t) = G$. Similarly we can define $L(X)(F)$ for $X \in \mathfrak{g}_{\mathrm{reg}}(F)$.

Definition 6.3. We say $X \in \mathfrak{g}_{\mathrm{reg}}(F)$ is *null with respect to* H if $L(X)$ does not contain any element of $\mathcal{L}(G, H)$ up to conjugation. Clearly, this definition only depends on the $G(\bar{F})$-conjugacy class (i.e. the stable conjugacy class) of X. As a result, we say a regular semisimple conjugacy class (resp. stable conjugacy class) of $\mathfrak{g}(F)$ is null with respect to H if every element in it is null with respect to H.

Remark 6.4. If $\mathcal{T}(G,H)^\circ = \{Z_{G,H}(F)\}$ or $\emptyset$ (e.g. the Whittaker models), the set $\mathcal{L}(G,H)$ is empty, which implies that every regular semisimple element in $\mathfrak{g}(F)$ is null with respect to H.

Remark 6.5. Another way to understand the notion of null is via the quasi-character $\theta = \hat{j}(X,\cdot)$ ($X \in \mathfrak{g}_{reg}(F)$) on $\mathfrak{g}(F)$, defined in Section 2.4. By the definition of null and [4, Proposition 4.7.1], if X is null with respect to H, then the regular germ of θ at $\mathfrak{t}(F)$ is zero for all $T(F) \in \mathcal{T}(G,H)^\circ$ with $T(F) \neq Z_{G,H}(F)$. Here $\mathfrak{t}(F)$ is the Lie algebra of $T^\circ(F)$.

In Section 8, we are going to use this property of null (together with some local trace formulas on the Lie algebra) to show that our definitions of the geometric multiplicities are the same as the ones that have already been proved for Gan–Gross–Prasad models and Ginzburg–Rallis models.

6.2. Conjugacy classes associated to regular nilpotent orbits

Fix a regular nilpotent orbit $\mathcal{O}$ of $\mathfrak{g}(F)$. For $\Xi \in \mathcal{O}$, by the theory of $\mathfrak{sl}_2$-triples, there exists a homomorphism

$$\varphi : F^\times \to G(F)$$

such that for all $s \in F^\times$, we have $\varphi(s)\Xi\varphi(s)^{-1} = s^{-2}\Xi$.

Since $\mathcal{O}$ is regular, φ is unique up to the center (i.e. two different choices of φ differ by an element of $\mathrm{Hom}(F^\times, Z_G(F))$). Let $N(F)$ and $\bar{N}(F)$ be the unipotent subgroups of $G(F)$ whose Lie algebras are given by

$$\mathfrak{n}(F) = \left\{X \in \mathfrak{g}(F) \;\middle|\; \lim_{s\to 0} \varphi(s)X\varphi(s)^{-1} = 0\right\},$$

$$\bar{\mathfrak{n}}(F) = \left\{X \in \mathfrak{g}(F) \;\middle|\; \lim_{s\to 0} \varphi(s)^{-1}X\varphi(s) = 0\right\}.$$

In particular, $\Xi \in \bar{\mathfrak{n}}(F)$. Finally, let $T(F)$ be the centralizer of $\mathrm{Im}(\varphi)$ in $G(F)$. Since $\mathcal{O}$ is regular, we know that $N(F)$ and $\bar{N}(F)$ are maximal unipotent subgroups of $G(F)$, $T(F)$ is a maximal torus of $G(F)$, $B(F) = T(F)N(F)$ and $\bar{B}(F) = T(F)\bar{N}(F)$ are Borel subgroups of $G(F)$, and $B(F)$ and $\bar{B}(F)$ are opposite to each other.

Remark 6.6. Let us consider an easy example when $G = \mathrm{SL}_2$ and $\Xi = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. In this case, we can define the map φ to be $\varphi(a) = \mathrm{diag}(a^{-1}, a)$. Then T is the diagonal torus of SL_2, $\bar{N}$ is the upper triangular unipotent subgroup of SL_2, and N is the lower triangular unipotent subgroup of SL_2.

Remark 6.7. The map

$$\xi : N(F) \to \mathbb{C}^\times, \quad \xi(\exp(X)) = \psi(\langle \Xi, X\rangle),\ X \in \mathfrak{n}(F),$$

is a generic character of $N(F)$.

Definition 6.8. For $X \in \mathfrak{g}_{reg}(F)$, we say that X is *associated to* $\mathcal{O}$ if X is $G(F)$-conjugate to an element in $\Xi + \mathfrak{b}(F)$. We say a regular semisimple conjugacy class of $\mathfrak{g}(F)$ is associated to $\mathcal{O}$ if all its elements are associated to $\mathcal{O}$. It is easy to see that this definition does not depend on the choice of Ξ. Moreover, $\Xi + \mathfrak{b}(F)$ is called the *Kostant section* associated to $\mathcal{O}$.

Remark 6.9. By the theory of Kostant sections ([13, Proposition 19], see also the summary in [15, Section 2.4]), for every stable regular semisimple conjugacy class of $\mathfrak{g}(F)$, there is a unique conjugacy class inside it that is associated to $\mathcal{O}$. Later in Section 8.1, we will show that for any two different regular nilpotent orbits $\mathcal{O}_1, \mathcal{O}_2 \in \mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}(F))$, there exists a regular semisimple conjugacy class of $\mathfrak{g}(F)$ that is associated to $\mathcal{O}_1$, but not to $\mathcal{O}_2$.

Lemma 6.10. *When F is p-adic, for all regular semisimple conjugacy classes $\{gXg^{-1} \mid g \in G(F)\}$ of $\mathfrak{g}(F)$, $\Gamma_{\mathcal{O}}(X) = 1$ if and only if X is associated to $\mathcal{O}$. Here $\Gamma_{\mathcal{O}}(X)$ is the Shalika germ defined in Section 2.4.*

Proof. This was proved by Kottwitz [15, Theorem 5.1 and Corollary 5.2]. See [7, Proposition 4.2] for a different proof. ∎

Remark 6.11. In general we expect that the above lemma also holds when $F = \mathbb{R}$ (the case of $F = \mathbb{C}$ is trivial).

6.3. *The reductive case*

We first consider the case when H is reductive. In the previous section, we have defined the subspace $\mathfrak{h}^{\perp}(F)$ of $\mathfrak{g}(F)$.

Definition 6.12. Let $\mathcal{N}(G, H, 1)$ be the subset of $\mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}(F))$ consisting of all elements $\mathcal{O} \in \mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}(F))$ satisfying the following condition:

- For almost all regular semisimple conjugacy classes of $\mathfrak{g}(F)$, if the conjugacy class is null with respect to H and is associated to $\mathcal{O}$, then this class intersects $\mathfrak{h}^{\perp}(F)$.

We refer the reader to Definition 6.3 for the definition of null.

6.4. *The nonreductive case*

Now we consider the nonreductive case. Let (G, H) be the parabolic induction of (G_0, H_0, ξ). In other words, there exists a parabolic subgroup of $P = MN$ of G, and a generic character $\xi : N(F) \to \mathbb{C}^{\times}$ of $N(F)$ such that

- $G_0 = M$ and $H = H_0 \ltimes N$ where $H_0 \subset G_0 = M$ is the neutral component of the stabilizer of the character ξ.

Let $\bar{P} = M\bar{N}$ be the opposite parabolic subgroup and let $\Xi \in \bar{\mathfrak{n}}(F)$ be the unique element such that

$$\xi(\exp(X)) = \psi(\langle \Xi, X \rangle), \quad \forall X \in \mathfrak{n}(F).$$

Since (G, H) is minimal, so is (G_0, H_0). By the discussion of the reductive case, we can define the subspace $\mathfrak{h}_0^\perp(F)$ of $\mathfrak{g}_0(F) = \mathfrak{m}(F)$ associated to the minimal spherical pair (G_0, H_0).

Definition 6.13. With the notation above, let $\mathcal{N}(G, H, \xi)$ be the subset of $\mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}(F))$ consisting of elements $\mathcal{O} \in \mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}(F))$ satisfying the following condition:

- For almost all regular semisimple conjugacy classes of $\mathfrak{g}(F)$, if the conjugacy class is null with respect to H and is associated to $\mathcal{O}$, then this conjugacy class intersects $\Xi + \mathfrak{h}_0^\perp(F) + \mathfrak{n}(F)$.

Remark 6.14. This definition depends on the generic character ξ.

Conjecture 6.15. *The set $\mathcal{N}(G, H, \xi)$ is nonempty.*

To end this section, we point out that the notion of null is crucial in our definition of the set $\mathcal{N}(G, H, \xi)$. The reason is that in most cases, the tangent space $\mathfrak{h}^\perp(F)$ (or $\Xi + \mathfrak{h}_0^\perp(F) + \mathfrak{n}(F)$ in the nonreductive case) does not contain all regular semisimple stable conjugacy classes of $\mathfrak{g}(F)$, but we do expect it contains almost all those regular semisimple stable conjugacy classes that are null with respect to H. Here are some examples.

For the model

$$(G(F), H(F)) = (\mathrm{GL}_{2n}(\mathbb{R}), \mathrm{SO}_{2n}(\mathbb{R})),$$

the set $\mathcal{T}(G, H)^\circ$ consists of subgroups of the form $\pm I_{2n-2m} \times (\mathbb{C}^1)^m$ with $0 \le m \le n$ (see Lemma 9.2). Here $\mathbb{C}^1$ is the norm 1 elements in $\mathbb{C}^\times$ identified with a torus of $\mathrm{GL}_2(\mathbb{R})$ via $e^{i\theta} \mapsto \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. As a result, the set $\mathcal{L}(G, H)$ consists of all standard Levi subgroups of $\mathrm{GL}_{2n}(\mathbb{R})$ of the form $(\mathrm{GL}_2(\mathbb{R}))^m \times (\mathrm{GL}_1(\mathbb{R}))^{2n-2m}$ for $1 \le m \le n$. This implies that a regular semisimple conjugacy class in $\mathfrak{g}(\mathbb{R}) = \mathfrak{gl}_{2n}(\mathbb{R})$ is null with respect to H if and only if all its eigenvalues are real numbers. On the other hand, by basic linear algebra, all eigenvalues of a symmetric real matrix are real numbers. This implies that $\mathfrak{h}^\perp(\mathbb{R})$ only contains those conjugacy classes that are null with respect to H. A similar discussion also holds for the model $(G(F), H(F)) = (\mathrm{GL}_{2n+1}(\mathbb{R}), \mathrm{SO}_{2n+1}(\mathbb{R}))$.

For the model

$$(G, H) = (\mathrm{GL}_3, \mathrm{SL}_2),$$

the set $\mathcal{T}(G, H)^\circ$ consists of all the maximal elliptic tori of $\mathrm{SL}_2(F)$ and the trivial torus. Hence the set $\mathcal{L}(G, H)$ contains all standard Levi subgroups of GL_3 of the form $\mathrm{GL}_2 \times \mathrm{GL}_1$. As a result, a regular semisimple conjugacy class in $\mathfrak{g}(F) = \mathfrak{gl}_3(F)$ is null with respect to H if and only if all of its eigenvalues belong to F (i.e. its centralizer in $G(F)$ is a split torus). On the other hand, it is easy to see that a regular semisimple conjugacy class appears in $\mathfrak{h}^\perp(F)$ if and only if at least one of its eigenvalues belongs to F (i.e. it is not elliptic). In particular, $\mathfrak{h}^\perp(F)$ does not contain all regular semisimple conjugacy classes of $\mathfrak{g}(F)$, but it does contain all those regular semisimple conjugacy classes that are null with respect to H.

7. The conjectural multiplicity formula and trace formula

7.1. The multiplicity formula

Let (G, H) be a spherical variety that is the parabolic induction of the reductive pair (G_0, H_0, ξ) (as in the previous sections, if (G, H) is reductive, we just let $(G_0, H_0, \xi) = (G, H, 1)$). Let $\omega : H_0(F) \to \mathbb{C}^\times$ be a unitary character. Then $\omega \otimes \xi$ is a character on $H(F) = H_0(F) \ltimes N(F)$. For any irreducible smooth representation π of $G(F)$, we define

$$m(\pi, \omega \otimes \xi) := \dim(\mathrm{Hom}_{H(F)}(\pi, \omega \otimes \xi)).$$

Recall that $Z_{G,H}(F) = Z_G(F) \cap H(F)$ and $A_{G,H}(F)$ is the maximal split torus of $Z_{G,H}(F)$. Let η be the restriction of the character ω to $A_{G,H}(F)$. Then we know that $m(\pi, \omega \otimes \xi) = 0$ unless the central character of π is equal to η on $A_{G,H}(F)$. We fix a central character $\chi : Z_G(F) \to \mathbb{C}^\times$ with $\chi|_{A_{G,H}(F)} = \eta$. Let $\mathrm{Irr}(G, \chi)$ be the set of all those irreducible smooth representations of $G(F)$ whose central character is χ. We use $\Pi_{\mathrm{temp}}(G, \chi)$ (resp. $\Pi_{\mathrm{disc}}(G, \chi)$, $\Pi_{\mathrm{cusp}}(G, \chi)$) to denote the set of tempered representations (resp. discrete series, supercuspidal representations) in $\mathrm{Irr}(G, \chi)$.

For $T(F) \in \mathcal{T}(G, H)$, we have defined $T_H(F) = \bigcup_{\gamma \in C(T,H)} \gamma T^\circ(F)$ in Section 4. Let dt be the Haar measure on $T^\circ(F)/A_{G,H}(F)$ such that the total volume is 1 (note that $T^\circ(F)/A_{G,H}(F)$ is compact). This induces a measure dt on $T_H(F)/A_{G,H}(F) = \bigcup_{\gamma \in C(T,H)} \gamma \cdot T^\circ(F)/A_{G,H}(F)$.

Now we are ready to define the geometric multiplicity.

Definition 7.1. Let θ be a quasi-character on $G(F)$ with central character χ (i.e. $\theta(zg) = \chi(z)\theta(g)$ for $z \in Z_G(F)$ and $g \in G_{\mathrm{reg}}(F)$). Define

$$\begin{aligned} m_{\mathrm{geom}}(\theta) = &\sum_{T(F) \in \mathcal{T}(G,H)} |W(H_0, T)|^{-1} \\ &\times \int_{T_H(F)/A_{G,H}(F)} \omega^{-1}(t) D^H(t) \frac{d(G_{0,T}, H_{0,T}, F)}{|Z_{H_0}(T)(F) : H_{0,T}(F)| \times c(G_{0,T}, H_{0,T}, F)} \\ &\times \frac{1}{|\mathcal{N}(G_T, H_T, \xi)|} \sum_{\mathcal{O} \in \mathcal{N}(G_T, H_T, \xi)} c_{\theta,\mathcal{O}}(t)\, dt. \end{aligned}$$

Here dt is the Haar measure on $T_H(F)/A_{G,H}(F)$ defined above, the numbers $d(G_{0,T}, H_{0,T}, F)$ and $c(G_{0,T}, H_{0,T}, F)$ are defined in Section 5, and $W(H_0, T) = N_{H_0}(T)(F)/Z_{H_0}(T)(F)$ where $N_{H_0}(T)(F)$ is the normalizer of $T(F)$ in $H_0(F)$.

For $\pi \in \mathrm{Irr}(G, \chi)$, we define the geometric multiplicity

$$m_{\mathrm{geom}}(\pi, \omega \otimes \xi) = m_{\mathrm{geom}}(\theta_\pi).$$

The number

$$\frac{1}{|Z_{H_0}(T)(F) : H_{0,T}(F)| \times c(G_{0,T}, H_{0,T}, F)}$$

is an analogue of $\frac{1}{|Z_H(x)|}$ for the finite group case in (1.1).

Remark 7.2. In general, the integral defining $m_{\mathrm{geom}}(\pi, \omega \otimes \xi)$ may not be absolutely convergent, and one would need to regularize it.

Among all the known cases (i.e. Whittaker models, Gan–Gross–Prasad models, Ginzburg–Rallis models, Galois models, and Shalika models), the integral defining $m_{\mathrm{geom}}(\pi, \omega \otimes \xi)$ is convergent for Whittaker models (this is trivial), orthogonal Gan–Gross–Prasad models [21, Proposition 7.3], Ginzburg–Rallis models [23, Proposition 5.2], Galois models [3, Section 4.1], and Shalika models [5, Lemma 3.2]. For unitary Gan–Gross–Prasad models, the integral is not convergent and one needs to regularize it ([2, Section 5] and [4, Section 11.1]).

Definition 7.3. When H is reductive, we say (G, H) is *tempered* (resp. *strongly tempered*) if all matrix coefficients of discrete series (resp. tempered) representations of $G(F)$ are integrable on $H(F)/A_{G,H}(F)$. In general, if (G, H) is the Whittaker induction of (G_0, H_0, ξ), we say (G, H) is tempered (resp. strongly tempered) if (G_0, H_0) is tempered (resp. strongly tempered).

Conjecture 7.4. (1) $m(\pi) = m_{\mathrm{geom}}(\pi)$ *for all* $\pi \in \Pi_{\mathrm{cusp}}(G, \chi)$.

(2) *If* (G, H) *is tempered, then* $m(\pi) = m_{\mathrm{geom}}(\pi)$ *for all* $\pi \in \Pi_{\mathrm{disc}}(G, \chi)$. *Moreover, let* $d\pi$ *be the natural measure on the set* $\Pi_{\mathrm{temp}}(G, \chi)$ *as defined in* [4, *Section* 2.6]. *Then* $m(\pi) = m_{\mathrm{geom}}(\pi)$ *for almost all* $\pi \in \Pi_{\mathrm{temp}}(G, \chi)$ (*under the measure* $d\pi$).

(3) *If* (G, H) *is strongly tempered, then* $m(\pi) = m_{\mathrm{geom}}(\pi)$ *for all* $\pi \in \Pi_{\mathrm{temp}}(G, \chi)$.

Remark 7.5. In the last case of the conjecture, we expect that the multiplicity formula holds not only for all the tempered representations, but also for all representations in the generic L-packets. Note that we say an L-packet is *generic* if it contains a generic representation.

As we said in the introduction, in general, if we want the multiplicity formula to hold for all irreducible smooth representations (or even all finite length smooth representations) of $G(F)$, we need to replace the multiplicity by the Euler–Poincaré pairing. One reason is that both the Harish-Chandra character and the Euler–Poincaré pairing behave nicely under short exact sequences, while the multiplicity does not. This was first observed by Prasad [17]. To be specific, for two smooth (not necessarily finite length) representations π and π' of $G(F)$, we define the Euler–Poincaré pairing

$$\mathrm{EP}_G[\pi, \pi'] = \sum_i (-1)^i \dim(\mathrm{Ext}^i_G[\pi, \pi']).$$

Then for a finite length smooth representation π of $G(F)$, we define (here for simplicity we assume that the split center $A_{G,H}(F)$ is trivial)

$$\mathrm{EP}(\pi, \omega \otimes \xi) = \mathrm{EP}_G(\pi, \mathrm{Ind}_H^G(\omega \otimes \xi)).$$

Conjecture 7.6. *Given a finite length smooth representation* π *of* $G(F)$, *the following hold:*

(1) *The Euler–Poincaré pairing* $\mathrm{EP}(\pi,\omega\otimes\xi)$ *is well defined. In other words,* $\mathrm{Ext}^i_G(\pi,\mathrm{Ind}_H^G(\omega\otimes\xi))$ *is finite-dimensional for all* $i\geq 0$.

(2) $\mathrm{EP}(\pi,\omega\otimes\xi)=m_{\mathrm{geom}}(\pi,\omega\otimes\xi)$.

When F is p-adic, the first part of the conjecture was proved by Aizenbud and Sayag [1].

Remark 7.7. When π is supercuspidal, we have $\mathrm{Ext}^i_G(\pi,\mathrm{Ind}_H^G(\omega\otimes\xi))=0$ for $i>0$, which implies that $\mathrm{EP}(\pi,\omega\otimes\xi)=m(\pi,\omega\otimes\xi)$. That is why the multiplicity formula $m(\pi,\omega\otimes\xi)=m_{\mathrm{geom}}(\pi,\omega\otimes\xi)$ should always hold in the supercuspidal case.

Remark 7.8. For the examples in Section 3, the model in Case 2 is tempered but not strongly tempered; that is why the multiplicity formula only holds for discrete series representations. The models in the remaining cases are strongly tempered, so the multiplicity formula holds for all the representations in the generic L-packets (for $U_2(F)$, a representation belongs to a generic L-packet if and only if it is infinite-dimensional). For the Whittaker model in Case 1, the Euler–Poincaré pairing is equal to the multiplicity [17, Proposition 2.8] and hence the multiplicity formula holds for all irreducible smooth representations. For Case 3, the Euler–Poincaré pairing is equal to the multiplicity because the group $H(F)$ is compact. So the multiplicity formula in this case also holds for all irreducible smooth representations.

In Section 8, we will show that Conjecture 7.4 holds for Whittaker models, Gan–Gross–Prasad models, Ginzburg–Rallis models, Galois models and Shalika models. For each of these cases, there is a multiplicity formula that has already been proved. Hence in order to prove Conjecture 7.4, we just need to show that our definition of the geometric multiplicity is the same as the one in the known multiplicity formula. On the other hand, Conjecture 7.6 is more difficult. The only known cases are the group case $(G,H)=(H\times H,H)$, the Whittaker models, and the Gan–Gross–Prasad models for general linear groups (see [17, Propositions 2.1, 2.8 and Theorem 4.2]).

7.2. The trace formula

We use the same notation as in the previous subsection. We first need to define the space of test functions. When (G,H) is tempered, we require $f\in\mathcal{C}_{\mathrm{scusp}}(G(F),\chi)$. When (G,H) is not tempered, we require $f\in{}^{\circ}\mathcal{C}(G(F),\chi)\cap C_c^\infty(G(F),\chi)$. For such a test function f, we define the distribution $I(f)$ of the trace formula to be

$$I(f)=\int_{H(F)\backslash G(F)}\int_{H(F)/A_{G,H}(F)} f(g^{-1}hg)\omega\otimes\xi(h)^{-1}\,dh\,dg.$$

In general the double integral above is not absolutely convergent (although each individual integral is usually convergent) and one needs to introduce some truncation functions on $H(F)\backslash G(F)$.

For the geometric expansion, let θ_f be the quasi-character on $G(F)$ defined via the weighted orbital integrals of f. We define the geometric expansion of the trace formula to be

$$I_{\mathrm{geom}}(f) = m_{\mathrm{geom}}(\theta_f)$$

where $m_{\mathrm{geom}}(\theta_f)$ was defined in Definition 7.1.

For the spectral expansion, when (G, H) is not tempered, let

$$I_{\mathrm{spec}}(f) = \sum_{\pi \in \Pi_{\mathrm{cusp}}(G,\chi)} m(\pi, \omega \otimes \xi)\, \mathrm{tr}(\pi^\vee(f)) \tag{7.1}$$

where $\pi^\vee$ is the contragredient of π. When (G, H) is tempered, let

$$I_{\mathrm{spec}}(f) = \int_{\mathcal{X}(G,\chi)} D(\pi)\theta_f(\pi^\vee) m(\pi, \omega \otimes \xi)\, d\pi. \tag{7.2}$$

Here $\mathcal{X}(G, \chi)$ is a set of virtual tempered representations of $G(F)$ with central character χ defined in [4, Section 2.7], where the number $D(\pi)$ and the measure $d\pi$ are also defined, and $\theta_f(\pi^\vee)$ is defined in [4, Section 5.4] via weighted characters. Now we are ready to state the conjectural trace formula.

Remark 7.9. When $f \in {}^\circ\mathcal{C}(G(F), \chi) \cap C_c^\infty(G(F), \chi)$, the expression on the right hand side of (7.2) is equal to the one on the right hand side of (7.1).

Conjecture 7.10. (1) *If (G, H) is tempered, then*

$$I_{\mathrm{geom}}(f) = I(f) = I_{\mathrm{spec}}(f) \quad \textit{for all } f \in \mathcal{C}_{\mathrm{scusp}}(G(F), \chi).$$

(2) *(G, H) is not tempered, then*

$$I_{\mathrm{geom}}(f) = I(f) = I_{\mathrm{spec}}(f) \quad \textit{for all } f \in {}^\circ\mathcal{C}(G(F), \chi) \cap C_c^\infty(G(F), \chi).$$

Like the conjectural multiplicity formula, by our discussion in Section 8, we know that Conjecture 7.10 holds for Whittaker models, Gan–Gross–Prasad models, Ginzburg–Rallis models, Galois models and Shalika models.

Remark 7.11. Although the trace formulas are the same for the tempered case and the strongly tempered case, the multiplicity formulas for these two cases behave differently. As discussed in Conjecture 7.4, for the strongly tempered case, the multiplicity formula should hold for all tempered representations; while for the tempered case, it only holds for all discrete series representations and for almost all tempered representations. An easy example of this kind would be the Shalika models (see [5, Remark 3.4]).

7.3. *The case when ω is not a character*

In this subsection, we assume that $F = \mathbb{R}$ and $H(\mathbb{R}) = K$ is a maximal connected compact subgroup of $G(\mathbb{R})$. Let ω be a finite-dimensional representation of $H(\mathbb{R})$. For a finite length smooth representation π of $G(\mathbb{R})$, we can still define the multiplicity $m(\pi, \omega)$ and the Euler–Poincaré pairing $\mathrm{EP}(\pi, \omega)$ as in the previous subsections. Moreover, since $H(\mathbb{R})$ is compact, we have $m(\pi, \omega) = \mathrm{EP}(\pi, \omega)$.

Meanwhile, let $\omega^\vee$ be the dual representation of ω and let

$$\theta_{\omega^\vee}(h) = \mathrm{tr}(\omega^\vee(h)), \quad h \in H(\mathbb{R}),$$

be the character of $\omega^\vee$. Then we can define the geometric multiplicity $m_{\mathrm{geom}}(\pi,\omega)$ as in the character case in Definition 7.1. The only difference is that we replace ω^{-1} by $\theta_{\omega^\vee}$. To be specific, we define

$$m_{\mathrm{geom}}(\pi,\omega) = \sum_{T(F)\in\mathcal{T}(G,H)} |W(H,T)|^{-1} \int_{T_H(F)/A_{G,H}(F)} \theta_{\omega^\vee}(t) D^H(t) \\ \times \frac{d(G_T,H_T,F)}{|Z_H(T)(F):H_T(F)| \times c(G_T,H_T,F)} \sum_{\mathcal{O}\in\mathcal{N}(G_T,H_T,1)} \frac{c_{\theta_\pi,\mathcal{O}}(t)}{|\mathcal{N}(G_T,H_T,1)|}\, dt.$$

Conjecture 7.12. *For all finite length smooth representations π of $G(\mathbb{R})$, we have $m(\pi,\omega) = m_{\mathrm{geom}}(\pi,\omega)$.*

Conjecture 7.12 gives a conjectural multiplicity formula for K-types of all finite length smooth representations of $G(\mathbb{R})$. In Sections 9 and 10, we will prove Conjecture 7.12 when $G(\mathbb{R}) = \mathrm{GL}_n(\mathbb{R})$ and when $G = \mathrm{Res}_{\mathbb{C}/\mathbb{R}} H$ is a complex reductive group. Clearly, it is enough to prove the conjecture when π and ω are irreducible.

8. The known cases

In this section, we assume that F is p-adic. We will show that for each of the known cases, the geometric multiplicity defined in Definition 7.1 is the same as the one in the multiplicity formula that has been proved. This implies that Conjectures 7.4 and 7.10 hold for all these cases. We consider Wittaker models in Section 8.1, Gan–Gross–Prasad models in Section 8.2, Ginzburg–Rallis models in Section 8.3, Galois models in Section 8.4, and Shalika models in Section 8.5.

We point out that none the models above has a Type N root. And for all these models, we have $\mathcal{T}(G,H) = \mathcal{T}(G,H)^\circ$ (i.e. the geometric multiplicity is only supported on tori of $G(F)$). This matches the discussion in Remark 4.10.

8.1. Whittaker models

Let G be a connected reductive group defined over F. Assume that $G(F)$ is quasi-split. Let $B = TN$ be a Borel subgroup of G, $\bar{B} = T\bar{N}$ be the opposite Borel subgroup, and $\xi : N(F) \to \mathbb{C}^\times$ be a generic character. Then there exists a unique element $\Xi \in \bar{\mathfrak{n}}(F)$ such that

$$\xi(\exp(X)) = \psi(\langle X, \Xi\rangle), \quad X \in \mathfrak{n}(F).$$

Without loss of generality, we assume that $G(F)$ has finite center (otherwise, we just need to replace $N(F)$ by $N(F)Z_G^\circ(F)$ where $Z_G^\circ(F)$ is the neutral component of $Z_G(F)$). For any irreducible smooth representation π of $G(F)$, define

$$m(\pi,\xi)=\dim(\mathrm{Hom}_{N(F)}(\pi,\xi)).$$

The model $(G,H,\xi)=(G,N,\xi)$ is called a *Whittaker model* of G and it is the Whittaker induction of the model $(G_0,H_0,\xi)=(T,1,\xi)$.

Let $\mathcal{O}\in \mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}(F))$ be the nilpotent orbit containing Ξ. By the work of Rodier [18, theorem p. 161 and Remark 2 p. 162] for the split case and the work of Moeglin–Waldspurger [16, Corollary I.17] for the general case, we have the multiplicity formula

$$m(\pi,\xi)=c_{\theta_\pi,\mathcal{O}}(1).$$

The goal of this subsection is to show that

$$m_{\mathrm{geom}}(\pi,\xi)=c_{\theta_\pi,\mathcal{O}}(1).$$

First, it is easy to see that the set $\mathcal{T}(G,N)$ only contains the trivial torus. Combining this with the fact that the Whittaker model is the Whittaker induction of the model $(T,1)$, we have

$$m_{\mathrm{geom}}(\pi,\xi)=\frac{1}{|\mathcal{N}(G,N,\xi)|}\sum_{\mathcal{O}'\in\mathcal{N}(G,N,\xi)}c_{\theta_\pi,\mathcal{O}'}(1).$$

Hence it is enough to show that

$$\mathcal{N}(G,N,\xi)=\{\mathcal{O}\}.$$

By the definition of $\mathcal{N}(G,N,\xi)$, we have $\mathcal{O}\in\mathcal{N}(G,N,\xi)$. Let $\mathcal{O}'\in\mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}(F))$ with $\mathcal{O}'\neq\mathcal{O}$. It is enough to show that $\mathcal{O}'\notin\mathcal{N}(G,N,\xi)$. In this case, $\mathcal{T}(G,N)=\{1\}$, which implies that all regular semisimple conjugacy classes of $\mathfrak{g}(F)$ are null with respect to N (Remark 6.4). Combining this with Lemma 6.10, in order to show that $\mathcal{O}'\notin\mathcal{N}(G,N,\xi)$, it is enough to prove the following lemma.

Lemma 8.1. *There exists a regular semisimple element $X\in\mathfrak{g}_{\mathrm{reg}}(F)$ such that*

$$\Gamma_{\mathcal{O}}(X)=1,\quad \Gamma_{\mathcal{O}'}(X)=0.$$

Here $\Gamma_{\mathcal{O}}(\cdot)$ and $\Gamma_{\mathcal{O}'}(\cdot)$ are the Shalika germs defined in Section 2.4.

Proof. By the result of Shelstad [20, p. 276], the regular Shalika germ is either 0 or 1. Hence if the statement of the lemma is false, we have $\Gamma_{\mathcal{O}}(X)=\Gamma_{\mathcal{O}'}(X)$ for all regular semisimple elements in $\mathfrak{g}(F)$. Since the distributions of nilpotent orbital integrals $\{J_{\mathcal{O}}(\cdot)\mid \mathcal{O}\in\mathrm{Nil}(\mathfrak{g}(F))\}$ are linearly independent [10, Lemma 3.8], there exists $f\in C_c^\infty(\mathfrak{g}(F))$ such that $J_{\mathcal{O}}(f)=1$, $J_{\mathcal{O}'}(f)=-1$ and $J_{\mathcal{O}_0}(f)=0$ for all other nilpotent orbits (not necessarily regular). By replacing f by $f\cdot 1_\omega$ where ω is a small G-invariant neighborhood of 0 in $\mathfrak{g}(F)$, we may assume that for all $X\in\mathrm{Supp}(f)\cap\mathfrak{g}_{\mathrm{reg}}(F)$, we have

$$J_G(X,f)=\sum_{\mathcal{O}_0\in\mathrm{Nil}(\mathfrak{g}(F))}\Gamma_{\mathcal{O}_0}(X)J_{\mathcal{O}_0}(f).$$

This implies that

$$J_G(X,f) = \sum_{\mathcal{O}_0 \in \mathrm{Nil}(\mathfrak{g}(F))} \Gamma_{\mathcal{O}_0}(X) J_{\mathcal{O}_0}(f) = \Gamma_{\mathcal{O}}(X) - \Gamma_{\mathcal{O}'}(X) = 0$$

for all $X \in \mathrm{Supp}(f) \cap \mathfrak{g}_{\mathrm{reg}}(F)$. Hence $J_G(X,f) = 0$ for all $X \in \mathfrak{g}_{\mathrm{reg}}(F)$. By [10, Theorem 3.1], we know that $J_{\mathcal{O}}(f) = J_{\mathcal{O}'}(f) = 0$. This is a contradiction. ∎

8.2. Gan–Gross–Prasad models

We only consider the orthogonal groups case; the unitary groups case is similar. We first recall the definition of the model from [21, Section 7]. Let V be a vector space of dimension d, and q be a nondegenerate symmetric bilinear form on V. Let $r \in \mathbb{N}$ with $2r+1 \le d$. Suppose we have an orthogonal decomposition $V = W \oplus D \oplus Z$ where D is a one-dimensional anisotropic subspace and Z is a hyperbolic subspace of dimension $2r$. We fix a basis v_0 of D and a basis $(v_i)_{i=\pm 1,\dots,\pm r}$ of Z with $q(v_i, v_j) = \delta_{i,-j}$. Let A be the maximal split torus of $\mathrm{SO}(Z)$ that preserves the subspace Fv_i. Let $G = \mathrm{SO}(V)$, and let $P = MN$ be the parabolic subgroup of G that preserves the filtration

$$Fv_r \subset Fv_r \oplus Fv_{r-1} \subset \cdots \subset Fv_r \oplus \cdots \oplus Fv_1$$

with $A \subset M$. In particular, $M = AG_0$ with $G_0 = \mathrm{SO}(V_0)$ and $V_0 = W \oplus D$. Let $\xi : N(F) \to \mathbb{C}^\times$ be the generic character defined in [21, Section 7.2]. Its stabilizer in $M(F)$ is $H_0^+(F) = \mathrm{O}(W)$. Let $H_0 = \mathrm{SO}(W)$ be the neutral component of H_0^+ and $H = H_0 \ltimes N$. The model $(G \times H_0, H, \xi)$ is a Gan–Gross–Prasad model for orthogonal groups (the embedding $H \to G \times H_0$ comes from the diagonal embedding $H_0 \to G_0 \times H_0$ and the embedding $N \to G$) defined by Gross and Prasad [9]. It is the Whittaker induction of the model $(G_0 \times H_0, H_0, \xi)$ (which is also a Gan–Gross–Prasad model). Let π (resp. σ) be an irreducible smooth representation of $G(F)$ (resp. $H_0(F)$). Define

$$m(\pi \otimes \sigma, \xi) = \dim(\mathrm{Hom}_{H(F)}(\pi \otimes \sigma, \xi)).$$

The multiplicity formula for this model was proved by Waldspurger [21], [22]. The goal of this subsection is to show that the geometric multiplicity $m_{\mathrm{geom}}(\pi \otimes \sigma, \xi)$ defined in Section 7 is the same as Waldspurger's definition in [21, Section 13.1]. We use $m'_{\mathrm{geom}}(\pi \otimes \sigma, \xi)$ to denote the geometric multiplicity defined by Waldspurger.

Remark 8.2. $(G_0 \times H_0, H_0)$ is a minimal wavefront spherical variety. Moreover, it is easy to see that there is only one open Borel orbit in $G_0(F) \times H_0(F)/H_0(F)$ and it has trivial stabilizer. In particular, $d(G_0 \times H_0, H_0, F) = c(G_0 \times H_0, H_0, F) = 1$.

Proposition 8.3. *The set $\mathcal{T}(G \times H_0, H)$ consists of tori $T(F)$ of $H_0(F)$ (up to conjugation) such that there exists an orthogonal decomposition $W = W' \oplus W''$ of W satisfying the following conditions:*

(1) *The dimension of W' is an even number.*

(2) *$T(F)$ is a maximal elliptic torus of $H_0'(F) = \mathrm{SO}(W')(F)$.*

(3) *If d is odd, the anisotropic rank of $V'' = W'' \oplus D \oplus Z$ is* 1. *If d is even, the anisotropic rank of W'' is* 1. *This is equivalent to saying that* $\mathrm{SO}(V'')(F)$ *and* $\mathrm{SO}(W'')(F)$ *are quasi-split.*

In particular, $\mathcal{T}(G \times H_0, H) = \mathcal{T}(G \times H_0, H)^{\circ}$.

Remark 8.4. The proposition implies that $\mathcal{T}(G \times H_0, H)$ is equal to the set $\underline{\mathcal{T}}$ defined in [21, Section 7.3], i.e. our definition of the support of the geometric multiplicity is the same as Waldspurger's definition for orthogonal Gan–Gross–Prasad models.

Proof of Proposition 8.3. It is easy to see that if a torus satisfies (1)–(3), it belongs to $\mathcal{T}(G, H)$. So we only need to prove the other direction: for given $T(F) \in \mathcal{T}(G, H)$, we need to show that $T(F)$ satisfies (1)–(3). Let W'' be the intersection of the kernel of $t - 1$ for $t \in T(F)$. Then for almost all $t \in T_H(F)$, W'' is the kernel of $t - 1$. In particular, $q|_{W''}$ is nondegenerate and $\dim(W) - \dim(W'')$ is even. Let W' be the orthogonal complement of W'' in W (i.e. $W = W' \oplus W''$), and $V'' = W'' \oplus D \oplus Z$. Then $T(F)$ is an abelian subgroup of $\mathrm{SO}(W')(F)$, $G_T = \mathrm{SO}(W')_T \times \mathrm{SO}(V'')$, $H_{0,T} = \mathrm{SO}(W')_T \times \mathrm{SO}(W'')$ and $H_T = \mathrm{SO}(W')_T \times (\mathrm{SO}(W'') \ltimes N'')$ where $N'' = N \cap \mathrm{SO}(V'')$ is the unipotent radical of the parabolic subgroup $P'' = P \cap \mathrm{SO}(V'')$ of $\mathrm{SO}(V'')$. In particular, $(\mathrm{SO}(V'') \times \mathrm{SO}(W''), \mathrm{SO}(W'') \ltimes N'')$ is the Gan–Gross–Prasad model associated to the decomposition $V'' = W'' \oplus D \oplus Z$. We will show that the decomposition $W = W' \oplus W''$ satisfies conditions (1)–(3).

Condition (1) follows from the fact that $\dim(W) - \dim(W'')$ is even. Since $G_T(F)$ and $H_{0,T}(F)$ are quasi-split, so are $\mathrm{SO}(V'')(F)$ and $\mathrm{SO}(W'')(F)$. This proves (3). It remains to prove (2). The following two statements follow from the definition of minimal spherical varieties:

- If (G_1, H_1) and (G_2, H_2) are spherical pairs, then $(G_1 \times G_2, H_1 \times H_2)$ is minimal if and only if (G_1, H_1) and (G_2, H_2) are minimal.
- For any connected reductive group H_1, the spherical pair $(H_1 \times H_1, H_1)$ is minimal if and only if H_1 is abelian (i.e. it is a torus).

Since $T(F) \in \mathcal{T}(G, H)$, $(G_T \times H_{0,T}, H_T)$ is minimal. By the statements above, we know that $\mathrm{SO}(W')_T$ is abelian, which implies that $\mathrm{SO}(W')_T$ is a maximal torus of $\mathrm{SO}(W')$. By Definition 4.3 (3), we know that $T(F)$ is the intersection of $H(F)$ with the center of $Z_G(T)(F) \times Z_{H_0}(T)(F)$, which implies that $T(F) = \mathrm{SO}(W')_T(F)$ (i.e. $T(F) = T^{\circ}(F)$ is a maximal torus of $\mathrm{SO}(W')(F)$). Finally, by Definition 4.3, $T(F)$ is compact, which implies that it is a maximal elliptic torus of $\mathrm{SO}(W')(F)$. This proves (2) and finishes the proof of the proposition. ■

Given $T(F) \in \mathcal{T}(G \times H_0, H)$ and let $W = W' \oplus W''$ be the decomposition associated to T. Then the model $(G_T \times H_{0,T}, H)$ is the product of the abelian model $(\mathrm{SO}(W')_T, \mathrm{SO}(W')_T) = (T, T)$ and the Gan–Gross–Prasad model associated to the decomposition $V'' = W'' \oplus D \oplus Z$. By Remark 8.2, we know that the constants $d(G_{0,T} \times H_{0,T}, H_{0,T}, F)$ and $c(G_{0,T} \times H_{0,T}, H_{0,T}, F)$ associated to the Gan–Gross–Prasad models are equal to 1. Moreover, since $Z_{H_0}(T) = H_{0,T}$, the constant

$|Z_{H_0}(T)(F) : H_{0,T}(F)|$ in the definition of geometric multiplicity is also equal to 1. Hence in order to prove $m_{\mathrm{geom}}(\pi \otimes \sigma, \xi) = m'_{\mathrm{geom}}(\pi \otimes \sigma, \xi)$, it remains to show that our choice of nilpotent orbits in Section 6 is the same as Waldspurger's in [21, Section 7.3].

Proposition 8.5. *Assume that $G(F)$ and $H_0(F)$ are quasi-split. Let $\mathcal{O}_G$ (resp. $\mathcal{O}_H$) be the regular nilpotent orbit of $\mathfrak{g}(F)$ (resp. $\mathfrak{h}_0(F)$) defined in* [21, Section 7.3]. *Then*

$$\mathcal{N}(G \times H_0, H, \xi) = \{\mathcal{O}_G \times \mathcal{O}_H\}.$$

Proof. Let $\Xi + \mathfrak{h}_0^{\perp}(F) + \mathfrak{n}(F) \subset \mathfrak{g}(F) \oplus \mathfrak{h}_0(F)$ be the space associated to the model $(G \times H_0, H, \xi)$ as in Section 6.4. By Lemma 6.10 together with [21, Sections 11.4–11.6], we know that $\mathcal{O} \notin \mathcal{N}(G \times H_0, H, \xi)$ for any $\mathcal{O} \in \mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}(F) \times \mathfrak{h}_0(F))$ with $\mathcal{O} \neq \mathcal{O}_G \times \mathcal{O}_H$. In fact, for any $\mathcal{O} \in \mathrm{Nil}_{\mathrm{reg}}(\mathfrak{g}(F) \times \mathfrak{h}_0(F))$ with $\mathcal{O} \neq \mathcal{O}_G \times \mathcal{O}_H$, in [21, Sections 11.4–11.6] Waldspurger has constructed an open subset $\mathfrak{t}_G(F)$ (resp. $\mathfrak{t}_H(F)$) of the regular semisimple conjugacy classes of $\mathfrak{g}(F)$ (resp. $\mathfrak{h}_0(F)$) such that for all $X_G \times X_H \in \mathfrak{t}_G(F) \times \mathfrak{t}_H(F)$, the following hold:

- $\Gamma_{\mathcal{O}}(X_G \times X_H) = 1$ and the conjugacy class $X_G \times X_H$ does not intersect $\Xi + \mathfrak{h}_0^{\perp}(F) + \mathfrak{n}(F)$.
- The conjugacy class $X_G \times X_H$ is null with respect to H.

Combining this with Lemma 6.10, we find that $\mathcal{O} \notin \mathcal{N}(G \times H_0, H, \xi)$.

Now it remains to show that

$$\mathcal{O}_G \times \mathcal{O}_H \in \mathcal{N}(G \times H_0, H, \xi). \tag{8.1}$$

The idea is to use the Lie algebra version of the local trace formula proved in [21]. Let f_G (resp. f_H) be a smooth compactly supported strongly cuspidal function on $\mathfrak{g}(F)$ (resp. $\mathfrak{h}_0(F)$). Let θ_{f_G} (resp. θ_{f_H}) be the quasi-character on $\mathfrak{g}(F)$ (resp. $\mathfrak{h}_0(F)$) associated to f_G (resp. f_H), and $\hat{\theta}_{f_G}$ (resp. $\hat{\theta}_{f_H}$) be its Fourier transform. By the local trace formula proved in [21, Sections 7.9 and 11.2], we have

$$I(\theta_{f_H}, \theta_{f_G}) = \sum_{T \in \mathcal{T}} |W(G,T)|^{-1} \int_{\mathfrak{t}(F)^H} D^{G \times H_0}(t)^{1/2} \hat{\theta}_{f_G} \times \hat{\theta}_{f_H}(t)\, dt \tag{8.2}$$

where $I(\theta_{f_H}, \theta_{f_G})$ is the Lie algebra analogue of the geometric multiplicity defined in [21, Section 7.9], $\mathcal{T}$ is a set of representatives of maximal tori of $G(F) \times H_0(F)$, and $W(G,T) = N_G(T)(F)/Z_G(T)(F)$ is the Weyl group. For $T \in \mathcal{T}$, $\mathfrak{t}^H(F)$ is the set of elements in $\mathfrak{t}_{\mathrm{reg}}(F)$ that are conjugate to an element in $\Xi + \mathfrak{h}_0^{\perp}(F) + \mathfrak{n}(F)$ (which is an open subset of $\mathfrak{t}_{\mathrm{reg}}(F)$).

If $\mathcal{O}_G \times \mathcal{O}_H \notin \mathcal{N}(G \times H_0, H, \xi)$, by Lemma 6.10 and the definition of $\mathcal{N}(G \times H_0, H, \xi)$, there exists $T_0 \in \mathcal{T}$ and a small open compact subset ω of $\mathfrak{t}_{0,\mathrm{reg}}(F)$ satisfying the following two conditions.

- For all $X \in \omega$, X is null with respect to H and X is associated to $\mathcal{O}_G \times \mathcal{O}_H$.
- The set $\omega' = \{X \in \omega \mid X \notin \mathfrak{t}_0(F)^H\}$ has nonzero measure.

Now choose f_G and f_H such that $\hat{\theta}_{f_G} \times \hat{\theta}_{f_H}$ is the characteristic function on $\omega^{G\times H_0}$. Then the right hand side of (8.2) is equal to

$$\int_{\omega\cap\mathfrak{t}_0(F)^H} D^{G\times H_0}(t)^{1/2}\,dt. \tag{8.3}$$

Since every element in ω is null with respect to H and is associated to $\mathcal{O}_G \times \mathcal{O}_H$, by [4, Propositions 4.1.1, 4.7.1] (here we use the property of null in Remark 6.5) we have

$$\begin{aligned} I(\theta_{f_H},\theta_{f_G}) &= c_{\theta_{f_G}\times\theta_{f_H},\mathcal{O}_G\times\mathcal{O}_H}(0) = \int_{\omega} D^{G\times H_0}(t)^{1/2}\Gamma_{\mathcal{O}_G\times\mathcal{O}_H}(t)\,dt \\ &= \int_{\omega} D^{G\times H_0}(t)^{1/2}\,dt = \int_{(\omega\cap\mathfrak{t}_0(F)^H)\cup\omega'} D^{G\times H_0}(t)^{1/2}\,dt. \end{aligned}$$

This contradicts (8.2) and (8.3) since ω' has nonzero measure. Hence $\mathcal{O}_G \times \mathcal{O}_H \in \mathcal{N}(G\times H_0, H, \xi)$. This finishes the proof of the proposition. ∎

8.3. *Ginzburg–Rallis models*

In this subsection, we consider Ginzburg–Rallis models. We will show that the geometric multiplicities defined in Section 7 are the same as the ones in the multiplicity formulas proved in [23], [24] (general linear groups case) and [25] (unitary groups and unitary similitude groups cases). For simplicity, we only consider the quasi-split unitary group and unitary similitude group cases; the non-quasi-split cases and the general linear groups case follow from a similar and easier argument.

Set $w_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, and $w_n = \begin{pmatrix} 0 & 1 \\ w_{n-1} & 0 \end{pmatrix}$ for $n > 2$. Let E/F be a quadratic extension. We define the unitary groups and unitary similitude groups to be

$$\begin{aligned} \mathrm{U}_n(F) &= \{g \in \mathrm{GL}_n(E) \mid \bar{g}^t w_n g = w_n\}, \\ \mathrm{GU}_n(F) &= \{g \in \mathrm{GL}_n(E) \mid \bar{g}^t w_n g = \lambda w_n, \quad \lambda \in F^\times\}. \end{aligned}$$

We use $\lambda : \mathrm{GU}_n(F) \to F^\times$ to denote the similitude character.

8.3.1. The unitary similitude group case. Let $G(F) = \mathrm{GU}_6(F)$, $P = MN$ be the parabolic subgroup of G with

$$N(F) = \left\{ \left. \begin{pmatrix} I_2 & X & Y \\ 0 & I_2 & -w_2\bar{X}^t w_2 \\ 0 & 0 & I_2 \end{pmatrix} \right| X, Y \in M_2(E),\ w_2 X w_2 \bar{X}^t + w_2 Y w_2 + \bar{Y}^t = 0 \right\},$$

$$M(F) = \left\{ \left. \begin{pmatrix} h & 0 & 0 \\ 0 & g & 0 \\ 0 & 0 & \lambda(g) w_2 (\bar{h}^t)^{-1} w_2 \end{pmatrix} \right| g \in \mathrm{GU}_2(F),\ h \in \mathrm{GL}_2(E) \right\}.$$

Here $M_n = \mathrm{Mat}_{n\times n}$. Let $H(F) = H_0(F) \ltimes N(F)$ with

$$H_0(F) = \left\{ \left. \begin{pmatrix} h & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & \lambda(h) w_2 (\bar{h}^t)^{-1} w_2 \end{pmatrix} \right| h \in \mathrm{GU}_2(F) \right\}.$$

Fix a character χ of $\mathrm{GU}_2(F)$. Define the character $\omega \otimes \xi$ on $H(F)$ by

$$\omega \otimes \xi\left(\begin{pmatrix} h & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & \lambda(h) w_2 (\bar{h}^t)^{-1} w_2 \end{pmatrix}\begin{pmatrix} I_2 & X & Y \\ 0 & I_2 & -w_2 \bar{X}^t w_2 \\ 0 & 0 & I_2 \end{pmatrix}\right) = \chi(h)\psi(\mathrm{tr}_{E/F}(\mathrm{tr}(X))).$$

Let π be an irreducible smooth representation of $G(F)$. Define

$$m(\pi, \omega \otimes \xi) = \dim(\mathrm{Hom}_{H(F)}(\pi, \omega \otimes \xi)).$$

The model (G, H) is the unitary similitude analogue of the Ginzburg–Rallis models defined in [8], and it is the Whittaker induction of the model

$$(G_0, H_0, \xi) = (M, H_0, \xi) = (\mathrm{GU}_2(F) \times \mathrm{GL}_2(E), \mathrm{GU}_2(F), \xi).$$

It is easy to see that both (G, H) and (G_0, H_0) are minimal.

In [25, Section 5.1], we proved the multiplicity formula

$$m(\pi, \omega \otimes \xi) = c_{\theta_\pi, \mathcal{O}_{\mathrm{reg}}}(1) + \sum_{T \in \mathcal{T}_{\mathrm{ell}}(H_0)} |W(H_0, T)|^{-1} \int_{T(F)/A_{H_0}(F)} \omega(t)^{-1} D^H(t) c_{\theta_\pi, \mathcal{O}_t}(t)\, dt$$

where $\mathcal{O}_{\mathrm{reg}}$ is the unique regular nilpotent orbit of $\mathfrak{g}(F)$, $\mathcal{T}_{\mathrm{ell}}(H_0)$ is a set of representatives of maximal elliptic tori of $H_0(F)$, and for $T \in \mathcal{T}_{\mathrm{ell}}(H_0)$ and $t \in T(F)_{\mathrm{reg}}$, $\mathcal{O}_t$ is the unique regular nilpotent orbit in $\mathfrak{g}_t(F)$. The goal of this subsection is to show that

$$m_{\mathrm{geom}}(\pi, \omega \otimes \xi) = c_{\theta_\pi, \mathcal{O}_{\mathrm{reg}}}(1) + \sum_{T \in \mathcal{T}_{\mathrm{ell}}(H_0)} |W(H_0, T)|^{-1} \int_{T(F)/A_{H_0}(F)} \omega(t)^{-1} D^H(t) c_{\theta_\pi, \mathcal{O}_t}(t)\, dt. \tag{8.4}$$

First, we show that $\mathcal{T}(G, H) = \mathcal{T}(G, H)^\circ = \mathcal{T}_{\mathrm{ell}}(H_0) \cup \{1\}$. In fact, there are three types of conjugacy classes in $H_0(F)$: the center, elliptic regular conjugacy classes and nonelliptic regular conjugacy classes. It is easy to see from the definition that the center and the elliptic regular conjugacy classes satisfy all the conditions for the support of the geometric multiplicity. On the other hand, the nonelliptic conjugacy classes violate the "elliptic" condition of the support of the geometric multiplicity. This implies that $\mathcal{T}(G, H) = \mathcal{T}(G, H)^\circ = \mathcal{T}_{\mathrm{ell}}(H_0) \cup \{1\}$.

For $T \in \mathcal{T}_{\mathrm{ell}}(H_0)$, we have $G_T = Z_G(T)$ and $H_{0,T} = Z_{H_0}(T)$, and the model (G_T, H_T, ξ) is just the Whittaker model of G_T. By the result in Section 8.1 for Whittaker models, in order to prove (8.4), we only need to consider the geometric multiplicity at the identity $\{1\}$ and prove the following lemma.

Lemma 8.6. (1) $d(G_0, H_0, F) = c(G_0, H_0, 1) = 1$.

(2) $\mathcal{N}(G, H, \xi) = \{\mathcal{O}_{\mathrm{reg}}\}$.

Proof. It is easy to see that there is only one open Borel orbit in $G_0(F)/H_0(F)$, and the stabilizer of this orbit is the center of $H_0(F)$, which is connected. This implies that $d'(G_0, H_0, F) = c(G_0, H_0, F) = 1$. On the other hand, the model $(G_0(\bar{F}), H_0(\bar{F}))$ is essentially the trilinear GL_2 model $(\mathrm{GL}_2 \times \mathrm{GL}_2 \times \mathrm{GL}_2, \mathrm{GL}_2^{\mathrm{diag}})$, which is wavefront. Hence $d(G_0, H_0, F) = d'(G_0, H_0, F) = 1$. This proves (1).

For (2), since $\mathcal{O}_{\mathrm{reg}}$ is the unique regular nilpotent orbit of $\mathfrak{g}(F)$, it is enough to show that $\mathcal{O}_{\mathrm{reg}} \in \mathcal{N}(G, H, \xi)$. The argument is exactly the same as for Gan–Gross–Prasad models in (8.1); the local trace formula (8.2) for this case was proved in [25, Section 4.3]. This finishes the proof of the lemma and hence the proof of (8.4). ∎

8.3.2. The unitary group case. Let $G(F) = \mathrm{U}_6(F)$, $N \subset G$ be the unipotent subgroup as in the unitary similitude group case, and $P = MN$ be the parabolic subgroup of G with

$$M(F) = \left\{ \begin{pmatrix} h & 0 & 0 \\ 0 & g & 0 \\ 0 & 0 & w_2(\bar{h}^t)^{-1}w_2 \end{pmatrix} \middle| g \in \mathrm{U}_2(F),\ h \in \mathrm{GL}_2(E) \right\} \simeq \mathrm{U}_2(F) \times \mathrm{GL}_2(E).$$

Let $H(F) = H_0(F) \ltimes N(F)$ with

$$H_0(F) = \left\{ \begin{pmatrix} h & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & w_2(\bar{h}^t)^{-1}w_2 \end{pmatrix} \middle| h \in \mathrm{U}_2(F) \right\}.$$

Fix a character χ of $\mathrm{U}_2(F)$. Define the character $\omega \otimes \xi$ on $H(F)$ by

$$\omega \otimes \xi \left(\begin{pmatrix} h & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & w_2(\bar{h}^t)^{-1}w_2 \end{pmatrix} \begin{pmatrix} I_2 & X & Y \\ 0 & I_2 & -w_2\bar{X}^t w_2 \\ 0 & 0 & I_2 \end{pmatrix} \right) = \chi(h)\psi(\mathrm{tr}_{E/F}(\mathrm{tr}(X))).$$

Let π be an irreducible smooth representation of $G(F)$. Define

$$m(\pi, \omega \otimes \xi) = \dim(\mathrm{Hom}_{H(F)}(\pi, \omega \otimes \xi)).$$

The model (G, H) is the unitary analogue of the Ginzburg–Rallis models defined in [8], and it is the Whittaker induction of the model

$$(G_0, H_0, \xi) = (M, H_0, \xi) = (\mathrm{U}_2(F) \times \mathrm{GL}_2(E), \mathrm{U}_2(F), \xi).$$

It is easy to see that both (G, H) and (G_0, H_0) are minimal.

In [25, Proposition 5.4], we proved the multiplicity formula

$$m(\pi, \omega \otimes \xi) = c_{\theta_\pi, \mathcal{O}_{\mathrm{reg},1}}(1) + c_{\theta_\pi, \mathcal{O}_{\mathrm{reg},2}}(1) + \sum_{T \in \mathcal{T}_{\mathrm{ell}}(H_0)} |W(H_0, T)|^{-1} \int_{T(F)} \omega(t)^{-1} D^H(t) c_{\theta_\pi, \mathcal{O}_t}(t)\, dt$$

where $\mathcal{O}_{\mathrm{reg},1}, \mathcal{O}_{\mathrm{reg},2}$ are the regular nilpotent orbits of $\mathfrak{g}(F)$, $\mathcal{T}_{\mathrm{ell}}(H_0)$ is a set of representatives of maximal elliptic tori of $H_0(F)$, and for $T \in \mathcal{T}_{\mathrm{ell}}(H_0)$ and $t \in T(F)_{\mathrm{reg}}$, $\mathcal{O}_t$ is the unique regular nilpotent orbit in $\mathfrak{g}_t(F)$. The goal of this subsection is to show that

$$m_{\mathrm{geom}}(\pi, \omega \otimes \xi) = c_{\theta_\pi, \mathcal{O}_{\mathrm{reg},1}}(1) + c_{\theta_\pi, \mathcal{O}_{\mathrm{reg},2}}(1) + \sum_{T \in \mathcal{T}_{\mathrm{ell}}(H_0)} |W(H_0, T)|^{-1} \int_{T(F)} \omega(t)^{-1} D^H(t) c_{\theta_\pi, \mathcal{O}_t}(t)\, dt. \quad (8.5)$$

By the same argument as in the unitary similitude group case, we only need to prove the following lemma.

Lemma 8.7. (1) $d(G_0, H_0, F) = 2$ *and* $c(G_0, H_0, F) = 1$.

(2) $\mathcal{N}(G, H, \xi) = \{\mathcal{O}_{\mathrm{reg},1}, \mathcal{O}_{\mathrm{reg},2}\}$.

Proof. It is easy to see that there are two open Borel orbits of $G_0(F)/H_0(F)$ (they correspond to $F^\times/\mathrm{Im}(N_{E/F})$ where $N_{E/F} : E^\times \to F^\times$ is the norm map) and the stabilizer of each orbit is the center of $H_0(F)$, which is connected. This implies that $d'(G_0, H_0, F) = 2$ and $c(G_0, H_0, F) = 1$. On the other hand, the model $(G_0(\bar{F}), H_0(\bar{F}))$ is the trilinear GL_2 model which is wavefront. Hence $d(G_0, H_0, F) = d'(G_0, H_0, F) = 2$. This proves (1).

For (2), we cannot use the same argument as in the previous cases. The reason is that in [25], we were not able to prove the local trace formula (8.2) for this model (this is largely due to the fact that $d(G_0, H_0, F) \neq 1$, see Remark 5.15). Instead, we are going to use the result for the unitary similitude group case to prove (2).

Let $\Xi + \mathfrak{h}_0^\perp(F) + \mathfrak{n}(F)$ be the space associated to the model $(G \times H_0, H, \xi)$ as in Section 6.4. Let $\mathfrak{g}'(F)$ be the Lie algebra of $\mathrm{GU}_6(F)$, $\mathcal{O}_{\mathrm{reg}}$ be the unique nilpotent orbit of $\mathfrak{g}'(F)$, and (G', H', ξ) be the model in the unitary similitude group case. Then $\mathcal{O}_{\mathrm{reg}} = \mathcal{O}_{\mathrm{reg},1} \cup \mathcal{O}_{\mathrm{reg},2}$ and $\mathfrak{g}'(F) = \mathfrak{g}(F) \oplus \mathfrak{z}(F)$ where $\mathfrak{z}(F) = \{aI_6 \mid a \in F\}$ is contained in the center of $\mathfrak{g}'(F)$. Moreover, $\Xi + \mathfrak{h}_0^\perp(F) + \mathfrak{n}(F) + \mathfrak{z}(F)$ is the space associated to the model (G', H', ξ).

Since $\mathcal{O} = \mathcal{O}_{\mathrm{reg},1} \cup \mathcal{O}_{\mathrm{reg},2}$, if a regular semisimple element $X \in \mathfrak{g}(F)$ is associated to $\mathcal{O}_{\mathrm{reg},1}$ or $\mathcal{O}_{\mathrm{reg},2}$, then it is associated to $\mathcal{O}$ (as an element in $\mathfrak{g}'(F)$). Moreover, X is null with respect to H if and only if it is null with respect to H'. Hence by Lemma 8.6, for almost all regular semisimple $G(F)'$-conjugacy classes in $\mathfrak{g}(F)$, if the conjugacy class is null with respect to H and if it is associated to $\mathcal{O}_{\mathrm{reg},1}$ or $\mathcal{O}_{\mathrm{reg},2}$, then the class intersects $\Xi + \mathfrak{h}_0^\perp(F) + \mathfrak{n}(F)$. As a result, in order to prove the lemma, it is enough to prove the following statement.

(3) For all regular semisimple elements $X_1, X_2 \in \mathfrak{g}_{\mathrm{reg}}(F)$, if X_1 and X_2 are null with respect to H, then X_1 and X_2 are $G'(F)$-conjugate to each other if and only if they are $G(F)$-conjugate to each other.

Let $T(F) = G'_{X_1}(F)$, and $A_T(F)$ be the maximal split subtorus of $T(F)$. Since X_1 is null with respect to H, $L(F) = Z_{G'}(A_T)(F)$ is contained in a Siegel Levi subgroup of $G'(F)$ and we have $X_1 \in \mathfrak{l}(F)$. In particular, X_1 commutes with $Z_L(F)$. Then (3) follows from the fact that every element $g \in G'(F)$ can be written as $g = g_1 z$ with $g_1 \in G(F)$ and $z \in Z_L(F)$. This finishes the proof of the lemma and hence the proof of (8.5). ∎

8.4. Galois models

Let E/F be a quadratic extension, H be a connected reductive group defined over F, and $G = \mathrm{Res}_{E/F} H$. Let χ be a character of $H(F)$. For an irreducible smooth representation π

of $G(F)$, define

$$m(\pi,\chi)=\dim(\mathrm{Hom}_{H(F)}(\pi,\chi)).$$

In [3, Theorem 3], Beuzart-Plessis proved the multiplicity formula for this model:

$$m(\pi,\chi)=\sum_{T\in\mathcal{T}_{\mathrm{ell}}(H)}|W(H,T)|^{-1}\int_{T(F)/A_H(F)}\chi(t)^{-1}D^H(t)\theta_\pi(t)\,dt$$

where $\mathcal{T}_{\mathrm{ell}}(H)$ is a set of representatives of maximal elliptic tori of $H(F)$. We want to show that

$$m_{\mathrm{geom}}(\pi,\chi)=\sum_{T\in\mathcal{T}_{\mathrm{ell}}(H)}|W(H,T)|^{-1}\int_{T(F)/A_H(F)}\chi(t)^{-1}D^H(t)\theta_\pi(t)\,dt. \tag{8.6}$$

For $T\in\mathcal{T}_{\mathrm{ell}}(H)$, we have $H_T(F)=Z_H(T)(F)=T(F)$ and the model $(G_T(F),H_T(F))$ is equal to the abelian model $(T(E),T(F))$. This implies that $|Z_H(T)(F):H_T(F)|=d(G_T,H_T,F)=c(G_T,H_T,F)=1$ and $\mathcal{N}(G_T,H_T)=\{0\}$ (here 0 is the unique nilpotent orbit of $\mathfrak{g}_T$). Hence in order to prove (8.6), it is enough to show that $\mathcal{T}(G,H)=\mathcal{T}_{\mathrm{ell}}(H)$. It is easy to see from the definition that $\mathcal{T}_{\mathrm{ell}}(H)\subset\mathcal{T}(G,H)$. For the other direction, let $T(F)\in\mathcal{T}(G,H)$. Then $(G_T,H_T)=(\mathrm{Res}_{E/F}H_T,H_T)$. In particular, it is minimal if and only if H_T is abelian (i.e. it is a maximal torus of H). By Definition 4.3 (3), we know that $T(F)=T^\circ(F)=H_T(F)$ is a maximal torus of $H(F)$. By Definition 4.3 (4), $T(F)/A_H(F)$ is compact. This implies that $T\in\mathcal{T}_{\mathrm{ell}}(H)$ and proves (8.6).

8.5. Shalika models

Let $G=\mathrm{GL}_{2n}$, $P=MN$ be a parabolic subgroup of G with

$$M=\left\{\begin{pmatrix}h_1&0\\0&h_2\end{pmatrix}\middle|\,h_i\in\mathrm{GL}_n\right\},$$
$$N=\left\{\begin{pmatrix}I_n&X\\0&I_n\end{pmatrix}\middle|\,X\in\mathrm{Mat}_{n\times n}\right\},$$

and $H=H_0\ltimes N$ with

$$H_0=\left\{\begin{pmatrix}h&0\\0&h\end{pmatrix}\middle|\,h\in\mathrm{GL}_n\right\}.$$

Given a multiplicative character $\chi:F^\times\to\mathbb{C}^\times$, we can define a character $\omega\otimes\xi$ of $H(F)$ by

$$\omega\otimes\xi\left(\begin{pmatrix}h&0\\0&h\end{pmatrix}\begin{pmatrix}I_n&X\\0&I_n\end{pmatrix}\right):=\psi(\mathrm{tr}(X))\chi(\det(h)).$$

For an irreducible smooth representation π of $G(F)$, define

$$m(\pi,\omega\otimes\xi)=\dim(\mathrm{Hom}_{H(F)}(\pi,\omega\otimes\xi)).$$

The pair (G, H) is called a *Shalika model*, it is the Whittaker induction of the model $(G_0, H_0, \xi) = (M, H_0) = (\mathrm{GL}_n \times \mathrm{GL}_n, \mathrm{GL}_n, \xi)$. In a joint work with Beuzart-Plessis [5, Theorem 1.4], we have proved the multiplicity formula

$$m(\pi, \omega \otimes \xi) = \sum_{T \in \mathcal{T}_{\mathrm{ell}}(H_0)} |W(H_0, T)|^{-1} \int_{T(F)/Z_G(F)} \omega(t)^{-1} D^H(t) c_{\theta_\pi, \mathcal{O}_t}(t)\, dt$$

where $\mathcal{T}_{\mathrm{ell}}(H_0)$ is a set of representatives of maximal elliptic tori of $H_0(F)$, and for $T \in \mathcal{T}_{\mathrm{ell}}(H_0)$ and $t \in T(F)_{\mathrm{reg}}$, $\mathcal{O}_t$ is the unique regular nilpotent orbit in $\mathfrak{g}_t(F)$. We want to show that

$$m_{\mathrm{geom}}(\pi, \omega \otimes \xi) = \sum_{T \in \mathcal{T}_{\mathrm{ell}}(H_0)} |W(H_0, T)|^{-1} \int_{T(F)/Z_G(F)} \omega(t)^{-1} D^H(t) c_{\theta_\pi, \mathcal{O}_t}(t)\, dt. \tag{8.7}$$

For $T \in \mathcal{T}_{\mathrm{ell}}(H_0)$, let K/F be the degree n extension such that $T(F) \simeq K^\times$. Then the model (G_T, H_T, ξ) is just the Whittaker model for $\mathrm{GL}_2(K)$. By the result in Section 8.1 for Whittaker models, to prove (8.7) it is enough to show that $\mathcal{T}(G, H) = \mathcal{T}_{\mathrm{ell}}(H_0)$.

Since $\mathcal{T}(G, H) = \mathcal{T}(G_0, H_0)$ (Remark 4.9), we only need to show that $\mathcal{T}(G_0, H_0) = \mathcal{T}_{\mathrm{ell}}(H_0)$. But the model $(G_0, H_0) = (\mathrm{GL}_n \times \mathrm{GL}_n, \mathrm{GL}_n)$ is a special case of the Galois models discussed in the previous subsection (just let $H = \mathrm{GL}_n$ and $E = F \oplus F$). By the result in the previous subsection, $\mathcal{T}(G_0, H_0) = \mathcal{T}_{\mathrm{ell}}(H_0)$. This proves (8.7).

9. The proof of Theorem 1.4 (1)

The goal of this section is to prove the conjectural multiplicity formula for K-types for $\mathrm{GL}_n(\mathbb{R})$, i.e. Theorem 1.4 (1). The proof has two parts. First we can easily prove the formula when $n \leq 2$. The second step is to show that both the multiplicities and the geometric multiplicities are invariant under parabolic induction. Then we can prove the multiplicity formula by using Proposition 2.1.

In Section 9.1, we explicitly write down the geometric multiplicity in this case. Then in Section 9.2, we reduce the proof of the multiplicity formula for $(\mathrm{GL}_n(\mathbb{R}), \mathrm{SO}_n(\mathbb{R}))$ to the proof of the multiplicity formula for $(\mathrm{GL}_n(\mathbb{R}), \mathrm{O}_n(\mathbb{R}))$. The reason is that the models $(\mathrm{GL}_n(\mathbb{R}), \mathrm{O}_n(\mathbb{R}))$ behave nicely under parabolic induction. Finally, in Section 9.3 we prove the multiplicity formula for $(\mathrm{GL}_n(\mathbb{R}), \mathrm{O}_n(\mathbb{R}))$.

9.1. The geometric multiplicity

Let $F = \mathbb{R}$, $G = \mathrm{GL}_n$ and $H = \mathrm{SO}_n = \{g \in \mathrm{GL}_n \mid gg^t = I_n, \det(g) = 1\}$. Then $H(\mathbb{R})$ is a maximal connected compact subgroup of $G(\mathbb{R})$. Let π be a finite length smooth representation of $\mathrm{GL}_n(\mathbb{R})$ and ω be a finite-dimensional representation of $\mathrm{SO}_n(\mathbb{R})$. The goal of this section is to prove the multiplicity formula

$$m(\pi, \omega) = m_{\mathrm{geom}}(\pi, \omega)$$

where $m(\pi, \omega) = \dim(\mathrm{Hom}_{H(\mathbb{R})}(\pi, \omega))$ and the geometric multiplicity $m_{\mathrm{geom}}(\pi, \omega)$ was defined in Section 7.3. In this subsection, we will give an explicit expression of $m_{\mathrm{geom}}(\pi, \omega)$; the result is summarized in Proposition 9.7.

Definition 9.1. Let $I(n) = \{(n_1, n_2, k) \in (\mathbb{Z}_{\geq 0})^3 \mid n_1 + 2n_2 + 2k = n\}$. For (n_1, n_2, k) in $I(n)$, if n is even ($\Leftrightarrow n_1$ is even), let $T_{n_1,n_2,k}$ be the abelian subgroup of $\mathrm{SO}_n(\mathbb{R})$ defined by

$$T_{n_1,n_2,k}(\mathbb{R}) = \{\mathrm{diag}(\pm I_{n_1}, \pm I_{2n_2}, t) \mid t \in (\mathbb{C}^1)^k\}$$

where $\mathbb{C}^1$ is the group of norm 1 elements in $\mathbb{C}$ and we identify it with $\mathrm{SO}_2(\mathbb{R})$ via the isomorphism $e^{2\pi i\theta} \mapsto \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. In particular, $t \in (\mathbb{C}^1)^k$ becomes an element of $\mathrm{SO}_{2k}(\mathbb{R}) \subset \mathrm{GL}_{2k}(\mathbb{R})$ and $\mathrm{diag}(\pm I_{n_1}, \pm I_{2n_2}, t)$ are elements of $\mathrm{SO}_n(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{R})$.

Similarly, if n is odd ($\Leftrightarrow n_1$ is odd), we define

$$T_{n_1,n_2,k}(\mathbb{R}) = \{\mathrm{diag}(I_{n_1}, \pm I_{2n_2}, t) \mid t \in (\mathbb{C}^1)^k\} \subset \mathrm{SO}_n(\mathbb{R}).$$

Lemma 9.2. *Assume that n is even. The set $\mathcal{T}(G, H)$ (defined in Definition 4.3) is the union of $T_{n_1,n_2,k}(\mathbb{R})$ where $(n_1, n_2, k) \in I(n)$ with $n_1 \geq 2n_2$.*

Proof. It is easy to see that $T_{n_1,n_2,k}(\mathbb{R}) \in \mathcal{T}(G, H)$. So it is enough to prove the other direction. Let t be a semisimple element of $H(\mathbb{R}) = \mathrm{SO}_n(\mathbb{R})$ such that (G_t, H_t) is a minimal spherical pair. After conjugation, we may assume that $t = \mathrm{diag}(I_{n_1}, -I_{2n_2}, t_0)$ where t_0 is a semisimple element in $\mathrm{SO}_{2k}(\mathbb{R})$ such that $t_0 \pm I_{2k} \in \mathrm{GL}_{2k}(\mathbb{R})$ (i.e. ± 1 are not eigenvalues of t_0). Here $2k = n - n_1 - 2n_2$.

Since ± 1 are not eigenvalues of t_0, the centralizer of t_0 in $\mathrm{GL}_{2k}(\mathbb{R})$ is of the form (note that all eigenvalues of t belong to $\mathbb{C}^1$)

$$\mathrm{GL}_{k_1}(\mathbb{C}) \times \cdots \times \mathrm{GL}_{k_m}(\mathbb{C})$$

with $k = k_1 + \cdots + k_m$. Then

$$G_t(\mathbb{R}) = \mathrm{GL}_{n_1}(\mathbb{R}) \times \mathrm{GL}_{2n_2}(\mathbb{R}) \times \mathrm{GL}_{k_1}(\mathbb{C}) \times \cdots \times \mathrm{GL}_{k_m}(\mathbb{C}),$$
$$H_t(\mathbb{R}) = \mathrm{SO}_{n_1}(\mathbb{R}) \times \mathrm{SO}_{2n_2}(\mathbb{R}) \times \mathrm{U}_{k_1}(\mathbb{R}) \times \cdots \times \mathrm{U}_{k_m(\mathbb{R})}.$$

Since (G_t, H_t) is a minimal spherical pair, we know that $(\mathrm{Res}_{\mathbb{C}/\mathbb{R}}\mathrm{GL}_{k_i}, \mathrm{U}_{k_i})$ is a minimal spherical pair for $1 \leq i \leq m$. This implies that $k_i = 1$ for $1 \leq i \leq m$. In other words, t_0 is a regular semisimple element of $\mathrm{GL}_{2k}(\mathbb{R})$.

Now we are ready to prove the lemma. Let $T(\mathbb{R}) \in \mathcal{T}(G, H)$. By conditions (1) and (4) of Definition 4.3, there exists $t \in T(\mathbb{R})$ such that $(G_T, H_T) = (G_t, H_t)$ is a minimal spherical pair. By the discussion above, up to conjugation, we may assume that $t = \mathrm{diag}(I_{n_1}, -I_{2n_2}, t_0)$ where $t_0 \in \mathrm{SO}_{2k}(\mathbb{R})$ is a regular semisimple element of $\mathrm{GL}_{2k}(\mathbb{R})$ and $(n_1, n_2, k) \in I(n)$. Combining this with condition (2) of Definition 4.3, we have

$$T(\mathbb{R}) = Z_{G_t}(\mathbb{R}) \cap H(\mathbb{R}) = \{\mathrm{diag}(\pm I_{n_1}, \pm I_{2n_2}, t') \mid t' \in T_0(\mathbb{R})\}$$

where $T_0(\mathbb{R})$ is the centralizer of t_0 in $\mathrm{SO}_{2k}(\mathbb{R})$, which is a maximal torus of $\mathrm{SO}_{2k}(\mathbb{R})$. Up to conjugation, we may assume that $n_1 \geq 2n_2$. Then the lemma follows from the fact that every maximal torus of $\mathrm{SO}_{2k}(\mathbb{R})$ is conjugate to $(\mathbb{C}^1)^k$. ∎

Lemma 9.3. *Assume that n is odd. Then the set $\mathcal{T}(G,H)$ is the union of $T_{n_1,n_2,k}(\mathbb{R})$ where $(n_1,n_2,k)\in I(n)$.*

Proof. The proof is similar to that of the previous lemma, so we skip it here. ∎

Corollary 9.4. *The geometric multiplicity $m_{\mathrm{geom}}(\pi,\omega)$ is supported on*

$$\{\mathrm{diag}(I_{n_1},-I_{2n_2},t)\mid t\in(\mathbb{C}^1)^k\}\cup\{\mathrm{diag}(-I_{n_1},I_{2n_2},t)\mid t\in(\mathbb{C}^1)^k\}$$

where $(n_1,n_2,k)\in I(n)$ with $n_1\geq 2n_2$ when n is even; and it is supported on

$$\{\mathrm{diag}(I_{n_1},-I_{2n_2},t)\mid t\in(\mathbb{C}^1)^k\},\ (n_1,n_2,k)\in I(n)$$

when n is odd.

Proof. This is a direct consequence of the previous two lemmas. ∎

Lemma 9.5. (1) *(G,H) is a minimal spherical pair.*

(2) *$d(G,H,\mathbb{R})=1$ and $c(G,H,\mathbb{R})=2^{n-1}$.*

(3) *$\mathcal{N}(G,H,1)=\{\mathcal{O}\}$ where $\mathcal{O}$ is the unique regular nilpotent orbit of $\mathfrak{g}(\mathbb{R})=\mathfrak{gl}_n(\mathbb{R})$.*

Proof. The first part is trivial. For (2), let $B(\mathbb{R})$ be the upper triangular Borel subgroup of $G(\mathbb{R})$. Since (G,H) is a symmetric pair which is wavefront, we know that $d(G,H,\mathbb{R})=d'(G,H,\mathbb{R})$. By the Iwasawa decomposition, we have

$$G(\mathbb{R})=B(\mathbb{R})H(\mathbb{R})\quad\text{and}\quad B(\mathbb{R})\cap H(\mathbb{R})\simeq(\mathbb{Z}/2\mathbb{Z})^{n-1}.$$

This implies that $d(G,H,\mathbb{R})=d'(G,H,\mathbb{R})=1$ and $c(G,H,\mathbb{R})=2^{n-1}$. The last part follows from the arguments at the end of Section 6. ∎

Given $(n_1,n_2,k)\in I(n)$, let $T=T_{n_1,n_2,k}$. Then the model (G_T,H_T) is the product of the models $(\mathrm{GL}_{n_1}(\mathbb{R}),\mathrm{SO}_{n_1}(\mathbb{R}))$, $(\mathrm{GL}_{2n_2}(\mathbb{R}),\mathrm{SO}_{2n_2}(\mathbb{R}))$ and $((\mathbb{C}^1)^k,(\mathbb{C}^1)^k)$. The following lemma is easy to verify.

Lemma 9.6. (1) *The number $|Z_H(T)(\mathbb{R}):H_T(\mathbb{R})|$ is equal to 1 if $n_1n_2=0$, and is equal to 2 if $n_1n_2\neq 0$.*

(2) *If $n_1=n_2=0$ (this only happens when n is even), then*

$$|W(H,T)|=2^{k-1}k!=2^{n-k-n_1-2n_2-1}k!.$$

If $n_1=2n_2\neq 0$ (this only happens when n is even and $n\geq 4$), then

$$|W(H,T)|=2\times 2^k k!=2^{n-k-n_1-2n_2+1}k!.$$

If $n_1\neq 2n_2$, then

$$|W(H,T)|=2^k k!=2^{n-k-n_1-2n_2}k!.$$

Combining Corollary 9.4, Lemma 9.5 and Lemma 9.6, we find that (we set $t_{n_1,n_2} = \mathrm{diag}(I_{n_1}, -I_{2n_2}, t)$ and $t'_{n_1,n_2} = \mathrm{diag}(-I_{n_1}, I_{2n_2}, t)$) $m_{\mathrm{geom}}(\pi,\omega)$ is equal to

$$\sum_{(n_1,n_2,k)\in I(n),\, n_1>2n_2} \frac{1}{2^{n-k-1}k!} \int_{(\mathbb{C}^1)^k} \big(D^{\mathrm{SO}_n}(t_{n_1,n_2}) c_\pi(t_{n_1,n_2}) \theta_{\omega^\vee}(t_{n_1,n_2}) + D^{\mathrm{SO}_n}(t'_{n_1,n_2}) c_\pi(t'_{n_1,n_2}) \theta_{\omega^\vee}(t'_{n_1,n_2})\big)\, dt$$
$$+ \sum_{(n_1,n_2,k)\in I(n),\, n_1=2n_2\neq 0} \frac{1}{2^{n-k}k!} \int_{(\mathbb{C}^1)^k} \big(D^{\mathrm{SO}_n}(t_{n_1,n_2}) c_\pi(t_{n_1,n_2}) \theta_{\omega^\vee}(t_{n_1,n_2}) + D^{\mathrm{SO}_n}(t'_{n_1,n_2}) c_\pi(t'_{n_1,n_2}) \theta_{\omega^\vee}(t'_{n_1,n_2})\big)\, dt$$
$$+ \frac{1}{2^{n-n/2-1}\left(\frac{n}{2}!\right)} \int_{(\mathbb{C}^1)^{\frac{n}{2}}} D^{\mathrm{SO}_n}(t) c_\pi(t) \theta_{\omega^\vee}(t)\, dt$$

when n is even, and is equal to

$$\sum_{(n_1,n_2,k)\in I(n)} \frac{1}{2^{n-k-1}k!} \int_{(\mathbb{C}^1)^k} D^{\mathrm{SO}_n}(t_{n_1,n_2}) c_\pi(t_{n_1,n_2}) \theta_{\omega^\vee}(t_{n_1,n_2})\, dt$$

when n is odd where

- the Haar measure on $\mathbb{C}^1 = \mathrm{SO}_2(\mathbb{R})$ is chosen so that the total volume is 1;
- $c_\pi(t_{n_1,n_2}) = c_\pi(\mathrm{diag}(I_{n_1}, -I_{2n_2}, t))$ is the regular germ of θ_π at $t_{n_1,n_2} = \mathrm{diag}(I_{n_1}, -I_{2n_2}, t)$, and $c_\pi(t'_{n_1,n_2}) = c_\pi(\mathrm{diag}(-I_{n_1}, I_{2n_2}, t))$ is the regular germ of θ_π at $t'_{n_1,n_2} = \mathrm{diag}(-I_{n_1}, I_{2n_2}, t))$ (see Section 2.5 for the definition of regular germs);
- $\omega^\vee$ is the dual representation of ω and $\theta_{\omega^\vee}$ is the character of $\omega^\vee$.

When n is even, we can replace the element $t'_{n_1,n_2} = \mathrm{diag}(-I_{n_1}, I_{2n_2}, t)$ in the expression of $m_{\mathrm{geom}}(\pi,\omega)$ by $\mathrm{diag}(I_{2n_2}, -I_{n_1}, t)$ because they are conjugate to each other in $\mathrm{SO}_n(\mathbb{R})$. Then we have

$$m_{\mathrm{geom}}(\pi,\omega) = \sum_{(n_1,n_2,k)\in I(n)} \frac{1}{2^{n-k-1}k!} \int_{(\mathbb{C}^1)^k} D^{\mathrm{SO}_n}(\mathrm{diag}(I_{n_1}, -I_{2n_2}, t)) \times c_\pi(\mathrm{diag}(I_{n_1}, -I_{2n_2}, t)) \theta_{\omega^\vee}(\mathrm{diag}(I_{n_1}, -I_{2n_2}, t))\, dt.$$

In other words, we get the same expression as in the odd case. To summarize, we have proved the following proposition.

Proposition 9.7. *The geometric multiplicity $m_{\mathrm{geom}}(\pi,\omega)$ is given by the following formula:*

$$m_{\mathrm{geom}}(\pi,\omega) = \sum_{(n_1,n_2,k)\in I(n)} \frac{1}{2^{n-k-1}k!} \int_{(\mathbb{C}^1)^k} D^{\mathrm{SO}_n}(\mathrm{diag}(I_{n_1}, -I_{2n_2}, t)) \times c_\pi(\mathrm{diag}(I_{n_1}, -I_{2n_2}, t)) \theta_{\omega^\vee}(\mathrm{diag}(I_{n_1}, -I_{2n_2}, t))\, dt.$$

9.2. A reduction

Given a finite length smooth representation π of $\mathrm{GL}_n(\mathbb{R})$ and a finite-dimensional representation ω of $\mathrm{SO}_n(\mathbb{R})$, we need to prove the multiplicity formula

$$m(\pi,\omega) = m_{\mathrm{geom}}(\pi,\omega) \tag{9.1}$$

where $m_{\mathrm{geom}}(\pi,\omega)$ was defined in Proposition 9.7.

To prove (9.1), we need a multiplicity formula for the model $(\mathrm{GL}_n(\mathbb{R}), \mathrm{O}_n(\mathbb{R}))$. To be specific, let ω_+ be a finite-dimensional representation of $\mathrm{O}_n(\mathbb{R}) = \{g \in \mathrm{GL}_n(\mathbb{R}) \mid gg^t = I_n\}$, $\omega_+^\vee$ be its dual representation, and $\theta_{\omega_+^\vee} : \mathrm{O}_n(\mathbb{R}) \to \mathbb{C}$ be the character of $\omega_+^\vee$. We use $\mathrm{sgn} : \mathrm{O}_n(\mathbb{R}) \to \{\pm 1\}$ to denote the sign character of $\mathrm{O}_n(\mathbb{R})$. Given a finite length smooth representation π of $\mathrm{GL}_n(\mathbb{R})$, we define the multiplicity

$$m(\pi,\omega_+) = \dim(\mathrm{Hom}_{\mathrm{O}_n(\mathbb{R})}(\pi,\omega_+)),$$

and the geometric multiplicity

$$\begin{aligned} m_{\mathrm{geom}}(\pi,\omega_+) = \sum_{(n_1,n_2,k)\in J(n)} \frac{1}{2^{n-k}k!} \int_{(\mathbb{C}^1)^k} & D^{\mathrm{SO}_n}(\mathrm{diag}(I_{n_1}, -I_{n_2}, t)) \\ & \times c_\pi(\mathrm{diag}(I_{n_1}, -I_{n_2}, t))\theta_{\omega_+^\vee}(\mathrm{diag}(I_{n_1}, -I_{n_2}, t))\, dt \end{aligned} \tag{9.2}$$

where $J(n) = \{(n_1,n_2,k) \in (\mathbb{Z}_{\geq 0})^3 \mid n_1 + n_2 + 2k = n\}$.

Remark 9.8. Here we extend the Weyl determinant $D^{\mathrm{SO}_n}(\cdot)$ from $\mathrm{SO}_n(\mathbb{R})$ to $\mathrm{O}_n(\mathbb{R})$ by the same formula, i.e. for $x \in \mathrm{O}_n(\mathbb{R})_{\mathrm{ss}}$, we define

$$D^{\mathrm{SO}_n}(x) = |\det(1-\mathrm{Ad}(x))_{|\mathfrak{so}_n(\mathbb{R})/\mathfrak{so}_n(\mathbb{R})_x}|$$

where $\mathfrak{so}_n(\mathbb{R})_x$ is the centralizer of x in $\mathfrak{so}_n(\mathbb{R})$.

Remark 9.9. The reason we consider the models $(\mathrm{GL}_n(\mathbb{R}), \mathrm{O}_n(\mathbb{R}))$ is that they behave nicely under parabolic induction. To be specific, the intersection of $\mathrm{O}_n(\mathbb{R})$ with the standard Levi subgroup $\mathrm{GL}_{n'}(\mathbb{R}) \times \mathrm{GL}_{n''}(\mathbb{R})$ $(n = n' + n'')$ of $\mathrm{GL}_n(\mathbb{R})$ is $\mathrm{O}_{n'}(\mathbb{R}) \times \mathrm{O}_{n''}(\mathbb{R})$, while the intersection of $\mathrm{SO}_n(\mathbb{R})$ with $\mathrm{GL}_{n'}(\mathbb{R}) \times \mathrm{GL}_{n''}(\mathbb{R})$ is $\mathrm{S}(\mathrm{O}_{n'}(\mathbb{R}) \times \mathrm{O}_{n''}(\mathbb{R}))$.

Proposition 9.10. *Let ω_+ be a finite-dimensional representation of $\mathrm{O}_n(\mathbb{R})$ and $\omega = \omega_+|_{\mathrm{SO}_n(\mathbb{R})}$, which is a finite-dimensional representation of $\mathrm{SO}_n(\mathbb{R})$. For all finite length smooth representations π of $\mathrm{GL}_n(\mathbb{R})$, we have*

$$m(\pi,\omega) = m(\pi,\omega_+) + m(\pi,\omega_+ \otimes \mathrm{sgn}),$$
$$m_{\mathrm{geom}}(\pi,\omega) = m_{\mathrm{geom}}(\pi,\omega_+) + m_{\mathrm{geom}}(\pi,\omega_+ \otimes \mathrm{sgn}).$$

Proof. The second equation follows from the definitions of $m_{\mathrm{geom}}(\pi,\omega)$ and $m_{\mathrm{geom}}(\pi,\omega_+)$, together with the fact that $\theta_{\omega_+^\vee\otimes \mathrm{sgn}}(h) = \theta_{\omega_+^\vee}(h)\,\mathrm{sgn}(h)$ for all $h \in \mathrm{O}_n(\mathbb{R})$.

For the first equation, we just need to show that the linear map

$$\mathrm{Hom}_{\mathrm{O}_n(\mathbb{R})}(\pi,\omega_+)\oplus\mathrm{Hom}_{\mathrm{O}_n(\mathbb{R})}(\pi,\omega_+\otimes\mathrm{sgn})\to\mathrm{Hom}_{\mathrm{SO}_n(\mathbb{R})}(\pi,\omega),$$
$$l_1\oplus l_2\mapsto l_1+l_2,$$

is an isomorphism. It is clear that this map is injective, so we just need to show that it is surjective. Given $l\in\mathrm{Hom}_{\mathrm{SO}_n(\mathbb{R})}(\pi,\omega)$, we have $l=(l_1+l_2)/2$ where

$$l_1=l+\omega_+(\varepsilon)^{-1}\circ l\circ\pi(\varepsilon),\quad l_2=l-\omega_+(\varepsilon)^{-1}\circ l\circ\pi(\varepsilon),$$
$$\varepsilon=\mathrm{diag}(-1,I_{n-1})\in\mathrm{O}_n(\mathbb{R})-\mathrm{SO}_n(\mathbb{R}).$$

It is enough to show that

$$l_1\in\mathrm{Hom}_{\mathrm{O}_n(\mathbb{R})}(\pi,\omega_+),\quad l_2\in\mathrm{Hom}_{\mathrm{O}_n(\mathbb{R})}(\pi,\omega_+\otimes\mathrm{sgn}).$$

For $v\in\pi$ and $h\in\mathrm{SO}_n(\mathbb{R})$, we have

$$\begin{aligned}l_1(\pi(h)v)&=l(\pi(h)v)+\omega_+(\varepsilon)^{-1}\big(l(\pi(\varepsilon h)v)\big)\\&=\omega(h)l(v)+\omega_+(\varepsilon)^{-1}\big(l(\pi(\varepsilon h\varepsilon^{-1})\pi(\varepsilon)v)\big)\\&=\omega(h)l(v)+\omega_+(\varepsilon)^{-1}\big(\omega(\varepsilon h\varepsilon^{-1})l(\pi(\varepsilon)v)\big)\\&=\omega(h)l(v)+\omega(h)\omega_+(\varepsilon)^{-1}l(\pi(\varepsilon)v)=\omega(h)l_1(v)\end{aligned}$$

and

$$\begin{aligned}l_1(\pi(\varepsilon)v)&=l(\pi(\varepsilon)v)+\omega_+(\varepsilon)^{-1}\big(l(\pi(\varepsilon^2)v)\big)=l(\pi(\varepsilon)v)+\omega_+(\varepsilon)^{-1}\big(\omega(\varepsilon^2)l(v)\big)\\&=l(\pi(\varepsilon)v)+\omega_+(\varepsilon)l(v)=\omega_+(\varepsilon)l_1(v).\end{aligned}$$

This implies $l_1\in\mathrm{Hom}_{\mathrm{O}_n(\mathbb{R})}(\pi,\omega_+)$; and $l_2\in\mathrm{Hom}_{\mathrm{O}_n(\mathbb{R})}(\pi,\omega_+\otimes\mathrm{sgn})$ can be shown similarly. This proves the proposition. ■

The following theorem will be proved in the next subsection. It is the multiplicity formula for $(\mathrm{GL}_n(\mathbb{R}),\mathrm{O}_n(\mathbb{R}))$.

Theorem 9.11. *For all finite length smooth representations π of $\mathrm{GL}_n(\mathbb{R})$ and for all finite-dimensional representations ω_+ of $\mathrm{O}_n(\mathbb{R})$, we have*

$$m(\pi,\omega_+)=m_{\mathrm{geom}}(\pi,\omega_+).\tag{9.3}$$

Now we are ready to prove (9.1). It is enough to consider the case when ω is irreducible. We use ω' to denote the irreducible representation of $\mathrm{SO}_n(\mathbb{R})$ given by $\omega'(h)=\omega(\varepsilon^{-1}h\varepsilon)$ with $\varepsilon=\mathrm{diag}(-1,I_{n-1})$. If $\omega\simeq\omega'$, there exists an irreducible representation ω_+ of $\mathrm{O}_n(\mathbb{R})$ such that $\omega=\omega_+|_{\mathrm{SO}_n(\mathbb{R})}$. Then (9.1) follows from Proposition 9.10 and Theorem 9.11.

If ω is not isomorphic to ω' (this only happens when n is even), then there exists an irreducible representation ω_+ of $\mathrm{O}_n(\mathbb{R})$ such that $\omega\oplus\omega'=\omega_+|_{\mathrm{SO}_n(\mathbb{R})}$. By Proposition 9.10 and Theorem 9.11, we have

$$m(\pi,\omega)+m(\pi,\omega')=m_{\mathrm{geom}}(\pi,\omega)+m_{\mathrm{geom}}(\pi,\omega').$$

Hence in order to prove (9.1), it is enough to show that

$$m(\pi,\omega)=m(\pi,\omega'),\quad m_{\mathrm{geom}}(\pi,\omega)=m_{\mathrm{geom}}(\pi,\omega').$$

The first equality follows from the fact that the linear map

$$\begin{aligned}\mathrm{Hom}_{\mathrm{SO}_n(\mathbb{R})}(\pi,\omega)&\to \mathrm{Hom}_{\mathrm{SO}_n(\mathbb{R})}(\pi,\omega'),\\ l&\mapsto \omega_+(\varepsilon)^{-1}\circ l,\end{aligned}$$

is an isomorphism. The second equality follows from the facts that $\theta_{\omega^\vee}(h)=\theta_{(\omega')^\vee}(\varepsilon^{-1}h\varepsilon)$ for all $h\in \mathrm{SO}_n(\mathbb{R})$ and θ_π is invariant under ε-conjugation. This finishes the proof of (9.1) and hence the proof of Theorem 1.4 (1).

9.3. The proof of Theorem 9.11

In this subsection, we are going to prove Theorem 9.11. To simplify the notation, we will replace ω_+ by ω. We first consider the cases when $n\leq 2$. The case when $n=1$ is trivial. Now let $n=2$. We need to show that for all smooth finite length representations π of $\mathrm{GL}_2(\mathbb{R})$ and for all finite-dimensional representations ω of $\mathrm{O}_2(\mathbb{R})$, we have

$$m(\pi,\omega)=m_{\mathrm{geom}}(\pi,\omega) \tag{9.4}$$

where $m_{\mathrm{geom}}(\pi,\omega)$ is defined to be

$$\frac{c_\pi(I_2)\theta_\omega(I_2)+c_\pi(-I_2)\theta_{\omega^\vee}(-I_2)+2\theta_\pi\left(\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)\right)\theta_{\omega^\vee}\left(\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)\right)}{4} + \frac{1}{2}\int_{\mathrm{SO}_2(\mathbb{R})}\theta_\pi(t)\theta_{\omega^\vee}(t)\,dt.$$

When π is finite-dimensional, by the representation theory of compact groups we have

$$\begin{aligned}m(\pi,\omega)&=\int_{\mathrm{O}_2(\mathbb{R})}\theta_\pi(t)\theta_{\omega^\vee}(t)\,dt\\ &=\frac{\theta_\pi\left(\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)\right)\theta_{\omega^\vee}\left(\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)\right)}{2}+\frac{1}{2}\int_{\mathrm{SO}_2(\mathbb{R})}\theta_\pi(t)\theta_{\omega^\vee}(t)\,dt.\end{aligned}$$

Here the Haar measure on $\mathrm{O}_2(\mathbb{R})$ (resp. $\mathrm{SO}_n(\mathbb{R})$) is chosen so that the total volume is equal to 1. On the other hand, since π is finite-dimensional, we have $c_\pi(I_2)=c_\pi(-I_2)=0$. This proves (9.4).

Then we consider the induced representations. Assume that $\pi=I_B^{\mathrm{GL}_2}(\pi_1\otimes\pi_2)$ where $B=TN$ is the upper triangular Borel subgroup of $\mathrm{GL}_2(\mathbb{R})$ and $\pi_1\otimes\pi_2$ is a finite-dimensional representation of $T(\mathbb{R})=\mathrm{GL}_1(\mathbb{R})\times\mathrm{GL}_1(\mathbb{R})$. By the Iwasawa decomposition $\mathrm{GL}_2(\mathbb{R})=B(\mathbb{R})\mathrm{O}_2(\mathbb{R})$ and the reciprocity law, we have

$$\mathrm{Hom}_{\mathrm{O}_2(\mathbb{R})}(\pi,\omega)=\mathrm{Hom}_{\mathrm{O}_1(\mathbb{R})\times\mathrm{O}_1(\mathbb{R})}(\pi_1\otimes\pi_2,\omega|_{\mathrm{O}_1(\mathbb{R})\times\mathrm{O}_1(\mathbb{R})}).$$

By the representation theory of finite groups (note that $\mathrm{O}_1(\mathbb{R}) = \mathbb{Z}/2\mathbb{Z}$ is a finite group), we have

$$m(\pi,\omega) = \frac{\theta_{\pi_1}(1)\theta_{\pi_2}(1)\theta_{\omega^\vee}(I_2)}{4} + \frac{\theta_{\pi_1}(-1)\theta_{\pi_2}(-1)\theta_{\omega^\vee}(-I_2)}{4} + \frac{\theta_{\pi_1}(1)\theta_{\pi_2}(-1)\theta_{\omega^\vee}\left(\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)\right)}{4} + \frac{\theta_{\pi_1}(-1)\theta_{\pi_2}(1)\theta_{\omega^\vee}\left(\left(\begin{smallmatrix}-1&0\\0&1\end{smallmatrix}\right)\right)}{4}.$$

On the other hand, by Proposition 2.7, we have

$$m_{\mathrm{geom}}(\pi,\omega) = \frac{\theta_{\pi_1}(1)\theta_{\pi_2}(1)\theta_{\omega^\vee}(I_2)}{4} + \frac{\theta_{\pi_1}(-1)\theta_{\pi_2}(-1)\theta_{\omega^\vee}(-I_2)}{4} + \frac{\theta_{\pi_1}(1)\theta_{\pi_2}(-1)\theta_{\omega^\vee}\left(\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)\right)}{4} + \frac{\theta_{\pi_1}(-1)\theta_{\pi_2}(1)\theta_{\omega^\vee}\left(\left(\begin{smallmatrix}-1&0\\0&1\end{smallmatrix}\right)\right)}{4}.$$

This proves (9.4).

Now we prove (9.4) for the general case. It is enough to consider the case when π is irreducible. There are three kinds of irreducible smooth representations of $\mathrm{GL}_2(\mathbb{R})$: finite-dimensional representations, principal series and discrete series representations. The first two cases have already been considered, so it remains to consider the discrete series case. Assume that π is an irreducible discrete series representation. Then there exists a character $\chi_1 \otimes \chi_2$ of $T(\mathbb{R}) = \mathrm{GL}_1(\mathbb{R}) \times \mathrm{GL}_1(\mathbb{R})$ such that π is the unique subrepresentation of $\Pi = I_B^{\mathrm{GL}_2}(\chi_1 \otimes \chi_2)$ and $\pi' = \Pi/\pi$ is a finite-dimensional representation of $\mathrm{GL}_2(\mathbb{R})$. We have

$$m(\Pi,\omega) = m(\pi,\omega) + m(\pi',\omega), \quad m_{\mathrm{geom}}(\Pi,\omega) = m_{\mathrm{geom}}(\pi,\omega) + m_{\mathrm{geom}}(\pi',\omega).$$

By the discussion above, we have $m(\Pi,\omega) = m_{\mathrm{geom}}(\Pi,\omega)$ and $m(\pi',\omega) = m_{\mathrm{geom}}(\pi',\omega)$. Hence (9.4) also holds for discrete series representations. This proves Theorem 9.11 when $n \leq 2$.

Now assume that $n > 2$; we are going to prove Theorem 9.11 for $\mathrm{GL}_n(\mathbb{R})$. By induction, we assume that Theorem 9.11 holds for $\mathrm{GL}_k(\mathbb{R})$ when $k < n$. By Proposition 2.1, in order to prove Theorem 9.11, it is enough to prove the following proposition.

Proposition 9.12. *Theorem 9.11 holds for all induced representations. In other words, if $\pi = I_P^{\mathrm{GL}_n}(\tau)$ is an induced representation with $P = MN$ a proper parabolic subgroup of GL_n and τ a finite length smooth representation of $M(\mathbb{R})$, then $m(\pi,\omega) = m_{\mathrm{geom}}(\pi,\omega)$ for all smooth finite-dimensional representations ω of $\mathrm{O}_n(\mathbb{R})$.*

Proof. Let π be an induced representation of $\mathrm{GL}_n(\mathbb{R})$. Then there exists a maximal upper triangular parabolic subgroup $P = MN$ of $\mathrm{GL}_n(\mathbb{R})$ and a finite length smooth representation τ of $M(\mathbb{R})$ such that $\pi = I_P^{\mathrm{GL}_n}(\tau)$. Since P is maximal, $M(\mathbb{R}) = \mathrm{GL}_{n'}(\mathbb{R}) \times \mathrm{GL}_{n''}(\mathbb{R})$ for some $n', n'' > 0$ with $n = n' + n''$ and $\tau = \tau' \otimes \tau''$ where τ' (resp. τ'') is a finite length smooth representation of $\mathrm{GL}_{n'}(\mathbb{R})$ (resp. $\mathrm{GL}_{n''}(\mathbb{R})$).

By the Iwasawa decomposition $\mathrm{GL}_n(\mathbb{R}) = P(\mathbb{R})\mathrm{O}_n(\mathbb{R})$ and the reciprocity law, we have

$$\mathrm{Hom}_{\mathrm{O}_n(\mathbb{R})}(\pi,\omega) \simeq \mathrm{Hom}_{\mathrm{O}_{n'}(\mathbb{R})\times \mathrm{O}_{n''}(\mathbb{R})}(\tau_1 \otimes \tau_2, \omega|_{\mathrm{O}_{n'}(\mathbb{R})\times \mathrm{O}_{n''}(\mathbb{R})}).$$

Together with the induction hypothesis (applied to the pairs $(\mathrm{GL}_{n'}(\mathbb{R}), \mathrm{O}_{n'}(\mathbb{R}))$ and $(\mathrm{GL}_{n''}(\mathbb{R}), \mathrm{O}_{n''}(\mathbb{R}))$), we have

$$m(\pi,\omega) = \sum_{(n_1',n_2',k')\in J(n'),(n_1'',n_2'',k'')\in J(n'')} \frac{1}{2^{n'-k'}k'!}\frac{1}{2^{n''-k''}k''!}\int_{(\mathbb{C}^1)^{k'}}\int_{(\mathbb{C}^1)^{k''}} D^{\mathrm{SO}_{n'}}(\mathrm{diag}(I_{n_1'},-I_{n_2'},t'))D^{\mathrm{SO}_{n''}}(\mathrm{diag}(I_{n_1''},-I_{n_2''},t''))c_{\pi'}(\mathrm{diag}(I_{n_1'},-I_{n_2'},t')) \times c_{\pi''}(\mathrm{diag}(I_{n_1''},-I_{n_2''},t''))\theta_{\omega^\vee}(\mathrm{diag}(I_{n_1'},-I_{n_2'},t',I_{n_1''},-I_{n_2''},t''))\,dt'\,dt''. \tag{9.5}$$

It remains to show that $m_{\mathrm{geom}}(\pi,\omega)$ is equal to the right hand side of (9.5).

We first recall the definition of $m_{\mathrm{geom}}(\pi,\omega)$ from (9.2):

$$m_{\mathrm{geom}}(\pi,\omega) = \sum_{(n_1,n_2,k)\in J(n)} \frac{1}{2^{n-k}k!}\int_{(\mathbb{C}^1)^k} D^{\mathrm{SO}_n}(\mathrm{diag}(I_{n_1},-I_{n_2},t)) \times c_\pi(\mathrm{diag}(I_{n_1},-I_{n_2},t))\theta_{\omega^\vee}(\mathrm{diag}(I_{n_1},-I_{n_2},t))\,dt. \tag{9.6}$$

For $(n_1,n_2,k)\in J(n) = \{(n_1,n_2,k)\in(\mathbb{Z}_{\geq 0})^3 \mid n_1+n_2+2k=n\}$, let

$$I(n_1,n_2,k) = \{(n_1',n_1'',n_2',n_2'',k',k'')\in\mathbb{Z}^6_{\geq 0} \mid n_1=n_1'+n_1'', n_2=n_2'+n_2'', k=k'+k'',\ (n_1',n_2',k')\in J(n'),\ (n_1'',n_2'',k'')\in J(n'')\}.$$

By Proposition 2.7, for $(n_1,n_2,k)\in J(n)$ and $t=t_1\times\cdots\times t_k\in(\mathbb{C}^1)^k$ with $t_i\neq\pm 1$, $t_i\neq t_j$ and $t_i\neq\overline{t_j}$ for $1\leq i\neq j\leq n$, we have

$$D^{\mathrm{SO}_n}(\mathrm{diag}(I_{n_1},-I_{n_2},t))c_\pi(\mathrm{diag}(I_{n_1},-I_{n_2},t)) = \sum_{(n_1',n_1'',n_2',n_2'',k',k'')\in I(n_1,n_2,k)}\ \sum_{\{i_1,\ldots,i_{k'}\},\{j_1,\ldots,j_{k''}\}} D^{\mathrm{SO}_{n'}}(\mathrm{diag}(I_{n_1'},-I_{n_2'},t')) \times D^{\mathrm{SO}_{n''}}(\mathrm{diag}(I_{n_1''},-I_{n_2''},t''))c_{\pi'}(\mathrm{diag}(I_{n_1'},-I_{n_2'},t'))c_{\pi''}(\mathrm{diag}(I_{n_1''},-I_{n_2''},t'')) \tag{9.7}$$

where

- $i_1<\cdots<i_{k'}$ and $j_1<\cdots<j_{k''}$; $\{i_1,\ldots,i_{k'}\}$ runs over the subsets of $\{1,\ldots,k\}$ containing k' elements and $\{j_1,\ldots,j_{k''}\}=\{1,\ldots,k\}-\{i_1,\ldots,i_{k'}\}$.
- $t'=t_{i_1}\times\cdots\times t_{i_{k'}}$ and $t''=t_{j_1}\times\cdots\times t_{j_{k''}}$.

Combining (9.5)–(9.7), we have $m(\pi,\omega)=m_{\mathrm{geom}}(\pi,\omega)$. This finishes the proof of the proposition and hence the proofs of Theorems 1.4 (1) and 9.11. ∎

10. The proof of Theorem 1.4 (2)

In this section, we are going to prove the conjectural multiplicity formula for K-types of complex reductive groups, i.e. Theorem 1.4 (2). Let H be a connected reductive group defined over $\mathbb{R}$ with $H(\mathbb{R})$ compact and let $G=\mathrm{Res}_{\mathbb{C}/\mathbb{R}}H$. Let π be a finite length smooth representation of $G(\mathbb{R})$ and ω be a finite-dimensional representation of $H(\mathbb{R})$. We have defined the multiplicity

$$m(\pi,\omega)=\dim(\mathrm{Hom}_{H(\mathbb{R})}(\pi,\omega))$$

in previous sections. Moreover, by the discussion in Section 8.4 (note that (G,H) is a special case of the Galois models), we know that the geometric multiplicity in this case is defined by

$$\begin{aligned} m_{\mathrm{geom}}(\pi,\omega) &= |W(H,T)|^{-1}\int_{T(\mathbb{R})} D^H(t)\theta_\pi(t)\theta_{\omega^\vee}(t)\,dt \\ &= |W(G)|^{-1}\int_{T(\mathbb{R})} D^H(t)\theta_\pi(t)\theta_{\omega^\vee}(t)\,dt \end{aligned}$$

where $T(\mathbb{R})$ is a maximal torus of $H(\mathbb{R})$ (which is unique up to $H(\mathbb{R})$-conjugation) and $W(H,T)$ is the Weyl group which is isomorphic to the Weyl group $W(G)$ of $G(\mathbb{R})=H(\mathbb{C})$. The goal of this section is to prove Theorem 1.4 (2). In other words, we need to show that

$$m(\pi,\omega)=m_{\mathrm{geom}}(\pi,\omega). \tag{10.1}$$

When G is abelian, (10.1) is trivial. Hence by induction, we may assume that (10.1) holds for all proper Levi subgroups of G. By Proposition 2.1, it is enough to prove the following proposition.

Proposition 10.1. *Equality* (10.1) *holds for all induced representations. In other words, if $\pi=I_P^G(\tau)$ is an induced representation with $P=MN$ be a proper parabolic subgroup of G and τ be a finite length smooth representation of $M(\mathbb{R})$, then $m(\pi,\omega)=m_{\mathrm{geom}}(\pi,\omega)$ for all finite-dimensional representations ω of $H(\mathbb{R})$.*

Proof. By conjugating M we may assume that $P(\mathbb{R})\cap H(\mathbb{R})=M(\mathbb{R})\cap H(\mathbb{R})$ is a maximal compact subgroup of $M(\mathbb{R})$. Set $H_M=M\cap H$, then $M\simeq \mathrm{Res}_{\mathbb{C}/\mathbb{R}}H_M$. Moreover, we may choose the torus T so that $T\subset H_M$ (i.e. $T(\mathbb{R})$ is also a maximal torus of $H_M(\mathbb{R})$). By the Iwasawa decomposition $G(\mathbb{R})=P(\mathbb{R})H(\mathbb{R})$ and the reciprocity law, we have

$$\mathrm{Hom}_{H(\mathbb{R})}(\pi,\omega)\simeq \mathrm{Hom}_{H_M(\mathbb{R})}(\tau,\omega|_{H_M(\mathbb{R})}).$$

Combining this with our inductional hypothesis (applied to the pair $(M(\mathbb{R}),H_M(\mathbb{R}))$), we have

$$m(\pi,\omega)=|W(M)|^{-1}\int_{T(\mathbb{R})} D^{H_M}(t)\theta_\tau(t)\theta_{\omega^\vee}(t)\,dt \tag{10.2}$$

where $W(M)$ is the Weyl group of $M(\mathbb{R})=H_M(\mathbb{C})$.

For $t\in T(\mathbb{R})\cap G_{\mathrm{reg}}(\mathbb{R})$, we have $D^H(t)=D^G(t)^{1/2}$ and $D^{H_M}(t)=D^M(t)^{1/2}$. Combining this with Proposition 2.7, we have

$$D^H(t)\theta_\pi(t)=\sum_{t_M} D^{H_M}(t_M)\theta_\tau(t_M)$$

where t_M runs over a set of representatives for the $M(\mathbb{R})$-conjugacy classes of elements in $T(\mathbb{R})$ that are $G(\mathbb{R})$-conjugate to t (note that each regular semisimple $G(\mathbb{R})$-conjugacy class decomposes into $\frac{|W(G)|}{|W(M)|}$-many $M(\mathbb{R})$-conjugacy classes). As a result, we have

$$\int_{T(\mathbb{R})} D^H(t)\theta_\pi(t)\theta_{\omega^\vee}(t)\,dt = \frac{|W(G)|}{|W(M)|}\int_{T(\mathbb{R})} D^{H_M}(t)\theta_\tau(t)\theta_{\omega^\vee}(t)\,dt. \tag{10.3}$$

Now the proposition follows from (10.2) and (10.3). This finishes the proof of the proposition and hence the proof of Theorem 1.4 (2). ■

Acknowledgements. The author would like to thank Raphaël Beuzart-Plessis for the helpful comments on the first draft of this paper, and for many helpful discussions which led to the definition of the geometric multiplicity when the spherical variety has a Type N root. The author would also like to thank the anonymous referees for their helpful comments and corrections.

Funding. The author's work is partially supported by the NSF grant DMS-2000192 and DMS-2103720.

References

[1] Aizenbud, A., Sayag, E.: Homological multiplicities in representation theory of p-adic groups. Math. Z. **294**, 451–469 (2020) Zbl 1432.22017 MR 4054816

[2] Beuzart-Plessis, R.: La conjecture locale de Gross-Prasad pour les représentations tempérées des groupes unitaires. Mém. Soc. Math. France (N.S.) vii+191 (2016) Zbl 1357.22008 MR 3676153

[3] Beuzart-Plessis, R.: On distinguished square-integrable representations for Galois pairs and a conjecture of Prasad. Invent. Math. **214**, 437–521 (2018) Zbl 1409.22011 MR 3858402

[4] Beuzart-Plessis, R.: A local trace formula for the Gan–Gross–Prasad conjecture for unitary groups: the archimedean case. Astérisque **418** (2020) Zbl 07269852 MR 4146145

[5] Beuzart-Plessis, R., Wan, C.: A local trace formula for the generalized Shalika model. Duke Math. J. **168**, 1303–1385 (2019) Zbl 1426.22011 MR 3953435

[6] Clozel, L.: Invariant harmonic analysis on the Schwartz space of a reductive p-adic group. In: Harmonic Analysis on Reductive Groups (Brunswick, ME, 1989), Progr. Math. 101, Birkhäuser Boston, Boston, MA, 101–121 (1991) Zbl 0760.22023 MR 1168480

[7] DeBacker, S., Reeder, M.: On some generic very cuspidal representations. Compos. Math. **146**, 1029–1055 (2010) Zbl 1195.22011 MR 2660683

[8] Ginzburg, D., Rallis, S.: The exterior cube L-function for GL(6). Compos. Math. **123**, 243–272 (2000) Zbl 0989.11024 MR 1795291

[9] Gross, B. H., Prasad, D.: On irreducible representations of $\mathrm{SO}_{2n+1} \times \mathrm{SO}_{2m}$. Canad. J. Math. **46**, 930–950 (1994) Zbl 0829.22031 MR 1295124

[10] Harish-Chandra: Admissible Invariant Distributions on Reductive p-adic Groups. Univ. Lecture Ser. 16, Amer. Math. Soc., Providence, RI (1999) Zbl 0928.22017 MR 1702257

[11] Knop, F.: Weylgruppe und Momentabbildung. Invent. Math. **99**, 1–23 (1990) Zbl 0726.20031 MR 1029388

[12] Knop, F., Schalke, B.: The dual group of a spherical variety. Trans. Moscow Math. Soc. **78**, 187–216 (2017) Zbl 1433.14046 MR 3738085

[13] Kostant, B.: Lie group representations on polynomial rings. Amer. J. Math. **85**, 327–404 (1963) Zbl 0124.26802 MR 158024

[14] Kostant, B., Rallis, S.: Orbits and representations associated with symmetric spaces. Amer. J. Math. **93**, 753–809 (1971) Zbl 0224.22013 MR 311837

[15] Kottwitz, R. E.: Transfer factors for Lie algebras. Represent. Theory **3**, 127–138 (1999) Zbl 1044.22011 MR 1703328

[16] Mœglin, C., Waldspurger, J.-L.: Modèles de Whittaker dégénérés pour des groupes p-adiques. Math. Z. **196**, 427–452 (1987) Zbl 0612.22008 MR 913667

[17] Prasad, D.: Ext-analogues of branching laws. In: Proc. Int. Congress of Mathematicians (Rio de Janeiro, 2018), Vol. II. Invited lectures, World Sci., Hackensack, NJ, 1367–1392 (2018) Zbl 1443.11086 MR 3966813

[18] Rodier, F.: Modèle de Whittaker et caractères de représentations. In: Non-commutative harmonic analysis (Actes Colloq., Marseille-Luminy, 1974), 151–171. Lecture Notes in Math., Vol. 466 (1975) Zbl 0339.22014 MR 0393355

[19] Sakellaridis, Y., Venkatesh, A.: Periods and harmonic analysis on spherical varieties. Astérisque viii+360 (2017) Zbl 06847674 MR 3764130

[20] Shelstad, D.: A formula for regular unipotent germs. Astérisque **171-172**, 275–277 (1989) Zbl 0715.22012 MR 1021506

[21] Waldspurger, J.-L.: Une formule intégrale reliée à la conjecture locale de Gross–Prasad. Compos. Math. **146**, 1180–1290 (2010) Zbl 1200.22010 MR 2684300

[22] Waldspurger, J.-L.: Une formule intégrale reliée à la conjecture locale de Gross–Prasad, 2e partie: extension aux représentations tempérées. Astérisque **346**, 171–312 (2012) Zbl 1290.22012 MR 3202558

[23] Wan, C.: A local relative trace formula for the Ginzburg–Rallis model: the geometric side. Mem. Amer. Math. Soc. **261**, v+90 pp. (2019) Zbl 1460.22007 MR 4028458

[24] Wan, C.: Multiplicity one theorem for the Ginzburg–Rallis model: the tempered case. Trans. Amer. Math. Soc. **371**, 7949–7994 (2019) Zbl 1415.22009 MR 3955540

[25] Wan, C., Zhang, L.: The multiplicity problems for the unitary Ginzburg–Rallis models. Israel J. Math., to appear; arXiv:1808.02203 (2018)