**Zeitschrift:** Commentarii Mathematici Helvetici

**Herausgeber:** Schweizerische Mathematische Gesellschaft

**Band:** 78 (2003)

**Artikel:** Generalized Whittaker vectors for holomorphic and quaternionic representations

**Autor:** Wallach, Nolan R.

**DOI:** https://doi.org/10.5169/seals-58757

Comment. Math. Helv. 78 (2003) 266–307
0010-2571/03/020266-42

**Commentarii Mathematici Helvetici**

# Generalized Whittaker vectors for holomorphic and quaternionic representations

Nolan R. Wallach

**Abstract.** The conjugacy class of parabolic subgroups with Heisenberg unipotent radical in a simple Lie groups over $\mathbb{C}$ not of type $C_n$ contains an element defined over $\mathbb{R}$ for each quaternionic real form. In this paper we study the Whittaker models for quaternionic discrete series of these real forms and prove results analogous and by analogous methods to the case of simple Lie groups over $\mathbb{R}$ that are the automorphism groups of tube type Hermitian symmetric domain and (so-called Bessel models) for holomorphic representations. In particular we calculate the decomposition of the space of Whittaker vectors under the action of the stabilizer of the corresponding character in a Levi factor of the Heisenberg parabolic subgroup.

**Mathematics Subject Classification (2000).** 22E46, 22E47.

**Keywords.** Representations, Lie groups, Whittaker vectors, quaternionic discrete series.

## 1. Introduction

The Fourier coefficients of a classical automorphic function at a cusp can be interpreted as defining certain classes of continuous functionals on the space of smooth vectors of the corresponding automorphic representation. The constant term (zero Fourier transform) yields a conical vector. The other Fourier coefficients yield Whittaker vectors. In the classical theory the underlying group (at the infinite place) is $SL(2,\mathbb{R})$ and the representations occurring are principal series for Maas forms and discrete series for the holomorphic forms. Recently, B. Gross has extended the circle of ideas involving classical holomorphic automorphic forms to the forms on arithmetic quotients of split $G_2$ corresponding to quaternionic discrete series. Here the "Fourier coefficients" are replaced by generalized Whittaker vectors parametrized by binary forms over $\mathbb{R}$ of degree 3. His theory suggested that the classical holomorphy condition (the only non-zero Fourier coefficients are the positive ones) be replaced with a quaternionic condition which translates to positive discriminant. That this is true is a consequence of the main theorem of this paper.

Research partially supported by an NSF Summer research grant.

Before we give a description of the main results of the paper we will explain what we mean by a generalized Whittaker vector. Let $G$ be a real reductive group and let $P$ be a proper parabolic subgroup with unipotent radical $N$. We fix a Levi subgroup of $P$ (that is a Lie subgroup, $M$, of $P$ such that the map $M \times N \to P$, $m, n \mapsto mn$ is bijective). Let $\psi : N \to S^1$ be a unitary character. The abelianization of $N$, $V = N/[N, N]$, is isomorphic as a Lie group with an $n$-dimensional vector space over $\mathbb{R}$. The set of unitary characters can be identified with the real dual space of $V$, $V^*$ as follows: if $\lambda \in V^*$ then the corresponding unitary character is $n \mapsto \exp(i\lambda(n[N, N]))$. We will say that the character corresponding to $\lambda$ is generic if the orbit of $\lambda$ under the action of $M$ on $N$ by conjugation is open in $V^*$. If $(\pi, H)$ is a unitary representation of $G$ then we will use the notation $H^\infty$ for the smooth Fréchet representation of $G$ on the $C^\infty$ vectors in $H$. We will denote by $Wh_\psi^\infty(H)$ the space of all continuous linear functionals, $\nu$, on $H^\infty$ such that $\nu(\pi(n)h) = \psi(n)^{-1}\nu(h)$ for all $h \in H^\infty$ and $n \in N$. In the case of $G = SL(2, \mathbb{R})$ with $P$ the standard parabolic subgroup consisting of upper triangular matrices then

$$N = \left\{ \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} |x \in \mathbb{R} \right\}, M = \left\{ \begin{bmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{bmatrix} |x \in \mathbb{R} \right\}$$

and it is clear that the condition that a unitary character of $N$ be generic is just that it be non-trivial. It follows from a result of C.Moore that if $(\pi, H)$ is an appropriate choice holomorphic discrete series representation for $G$ then all of the generic $\psi$ such that $Wh_\psi^\infty(H) \neq 0$ are of the form $\psi\left(\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}\right) = \exp(itx)$ with all $t > 0$ (an independent proof will be found in section 2 furthermore the choice will be made). The analogous situation for $G_2$ is that one of the two parabolic subgroups of dimension 9 has a 5 dimensional Heisenberg group for its unipotent radical. We will fix this parabolic subgroup (that this is a correct choice for quaternionic discrete series is clear from the paper of [Gr-W] also there will be evidence later in this introduction) the identity component of $M$ has commutator group locally $SL(2, \mathbb{R})$ and its action on the unitary characters on $N$ is via the 4 dimensional, irreducible representation (the binary forms of degree 3). The condition that a unitary character be generic is that the discriminant of the corresponding form be non-zero. We prove (following Gross's suggestion) that the condition that there exist a non-zero Whittaker vector for a quaternionic discrete series transforming by $\psi$ is that the discriminant be positive. The structure of the space of such vectors is also described (also conjectured by Gross). We will describe this aspect in full generality later in this introduction.

We will now describe the main results of this paper. Let $G$ be a connected simple Lie group over $\mathbb{R}$ let $K$ be a maximal subgroup of $G$ subject to the condition that if $Z$ is the center of $G$ then $K/Z$ is compact. We say that $G$ is of Hermitian type if $K$ contains a one dimensional center (denoted $T$) it is said to be of quaternionic type if $K$ contains a normal subgroup, $K_1$, isomorphic with $SU(2)$ such that the isotropy representation of $SU(2)$ on the tangent space to the identity coset in $G/K$ is equivalent to a multiple of the action of the unit quaternions on the quaternions under left multiplication. If $G$ is of Hermitian type then a unitary irreducible representation is said to be holomorphic or antiholomorphic if it is admissible when restricted to $T$ (this means that as a representation of $T$ it splits into a direct sum of one dimensional representations with finite multiplicity). If $G$ is quaternionic then a unitary irreducible representation of $G$ is said to be quaternionic if it is admissible when restricted to $K_1$. We will first describe the results in the Hermitian case. We note that if $G$ is simply connected then $T$ is isomorphic with $\mathbb{R}$ under addition. Thus the group of unitary characters of $T$ is also isomorphic with $\mathbb{R}$ under addition. If $(\pi, H)$ is an irreducible unitary representation of $G$ that is admissible with respect to $T$ then the set of unitary characters that occur in the decomposition of $H$ when restricted to $T$ all either positive or negative. We say that $G$ is of tube type if there exists a connected subgroup $G_1$ of $G$ that is locally isomorphic with $SL(2,\mathbb{R})$ and $T \subset G_1$ (we may and do assume that $T$ is locally $SO(2)$ in $SL(2,\mathbb{R})$). We will assume that $G$ is tube type. Fix $P_1$ in $G_1$ corresponding to the upper triangular parabolic subgroup. with fix $h \in Lie(G_1)$ such that $h$ corresponds to the diagonal matrix $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ in $Lie(SL(2,\mathbb{R}))$. Then using the element $h$ we can construct a parabolic subgroup $P$ of $G$. Let $y \in Lie(G_1)$ be such that $y$ corresponds to the matrix

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Then we may look upon $y$ as defining a unitary character of $N$ via $\psi_o(\exp(x)) = \exp(iB(y,x))$ (here $N$ is abelian and we can therefore identify it with its Lie algebra). One can show that the orbit of $\psi_o$ is an open convex cone, $C$, and that if $\psi$ is such that $Wh_\psi^\infty(H) \neq 0$ for some holomorphic representation (after we have made a choice consistent with that for $G_1$), $(\pi, H)$, then $\psi$ is in the closure of the orbit of $\psi_0$. If $H$ is generic in a suitable sense (e.g. it is a holomorphic discrete series representation) then $Wh_\psi^\infty(H) \neq 0$ for all $\psi$ in the closure of the orbit and the action of the stabilizer, $M_\psi$, of $\psi$ on the space $Wh_\psi^\infty(H) \neq 0$ is equivalent with its restriction to a specific finite representation of $M$ to $M_\psi$ that is explicitly computable from the minimal $K$-type (it is in fact the action of $M$ on $Wh_1^\infty(H)$). The point of this discussion is to point out the critical role of one specific subgroup of $G$ that is locally isomorphic with $SL(2,\mathbb{R})$. The key is that groups locally isomorphic with $SL(2,\mathbb{R})$ are the smallest non-compact simple Lie groups that are of Hermitian type.

We will now look at the quaternionic case. Here the smallest such group that is non-compact is locally isomorphic with $SU(2,1)$. We now assume that $G$ is quaternionic. We note that if $G$ is not locally $Sp(n,1)$ then there exists a Lie subgroup of $G$, $S$ that is locally isomorphic with $SU(2,1)$ containing $K_1$ and invariant under the Cartan involution of $G$ corresponding to $K$, $\theta$. We fix a minimal parabolic subgroup, $P_1$, of $S$. Let $P_1 = M_1N_1$ be a Levi decomposition and let $A_1$ be the identity component of the center of $M_1$. Fix $h \in Lie(A_1)$ such that the eigenvalues of $ad(h)$ on $Lie(N_1)$ are strictly positive. Then using $h$ we can use the standard method to construct a parabolic subgroup, $P = MN$ with $M = \{g \in G | ad(g)h = h\}$. This real parabolic subgroup is up to conjugacy the one that was studied in [Gr-W]. The generic unitary characters of $N_1$ form a single orbit of $M_1$ (indeed they are just the non-trivial characters). We will say that a unitary character $\psi$ of $N$ is admissible if the restriction of $\psi$ to $mN_1m^{-1}$ is non-trivial for all $m \in M$. The admissible unitary characters of $N$ form a single orbit of $M$ (which in the special case of $G_2$ is the orbit of positive discriminant and will be described in general in section 7 of this paper). The main theorem is that if $(\pi, H)$ is a quaternionic discrete series of $G$ and if $\psi$ is a generic unitary character of $N$ then $Wh_\psi^\infty(H) \neq 0$ if and only if $\psi$ is admissible and if so the representation of $M_\psi$ on $Wh_\psi^\infty(H)$ is determined as the restriction to $M_\psi$ of a finite dimensional representation of $M$ constructed from the minimal $K$-type of $H$ (in particular it depends only on $H$). The first example that indicates the full subtlety of this result is $G$ locally isomorphic with $SO(4,4)$, that is, the smallest real rank 4 quaternionic group. The result rests on a subrepresentation theorem for quaternionic discrete series in (degenerate) principal series induced from finite dimensional representations of $P$ and an analogue of the multiplicity 1 theorem for generalized Whittaker model for these induced representations. The subrepresentation theorem and the result on Whittaker models (Theorem 12) are of independent interest. The latter is a special case of a general theorem on such models (based on a strengthening of Bruhat theory due to Kolk and Varadarajan [Ko-V]) which will appear in a later paper. However, full details for the case at hand are in this paper.

The literature in the holomorphic case is very rich and we are certain to have missed appropriate references to important contributions. (One such piece of work is [Gr-D] which also contains a substantial bibliography.) We have given an exposition of this case to emphasize the similarities and differences with the quaternionic case. Also, the material of section 10 of this paper is perhaps new (even in the case of the holomorphic discrete series since it involves the full spectrum from conical to Whittaker).

We would like to take this occasion to heartily thank the first reviewer of this paper for his list of "clarifications requested" and his second set of questions which turned out to involve gaps and one error in the original manuscript. The reviewer suggested that we look at the main result of [Ko-V] as a method of fixing the error. This important extension of Bruhat theory did indeed help significantly. Fortunately, no statement of a main result was changed in fixing these lapses. As indicated above, the material of this paper stems from some questions and conjectures of B. Gross. We would like to thank him for the initial question, for his amazing intuition that consistently helped us stay on the correct path and for his encouragement during this project.

## 2. The case of $G$ locally $SL(2,\mathbb{R})$

In this section we will collect some results about Whittaker models for $G$ a covering group of $SL(2,\mathbb{R})$ with covering homomorphism $\nu : G \to SL(2,\mathbb{R})$. Some of them are not so readily accessible so we will give a relatively self contained exposition. We take the usual basis

$$X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, Y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

for $Lie(G)$ which we identify with $Lie(SL(2,\mathbb{R}))$ via the differential of $\nu$. Let $P$ be the usual parabolic subgroup consisting of the normalizer in $G$ of span $H, X$. Then $P = MN$ with $M = Z\exp(\mathbb{R}H)$ and $N = \exp(\mathbb{R}X)$ and $Z$ is the center of $G$. Let $\sigma : M \to \mathbb{C}^\times$ be a group homomorphism. We extend $\sigma$ to $P$ by setting $\sigma(n) = 1$, $n \in N$. Let $I(\sigma)$ denote the representation obtained by inducing (unnormalized) $\sigma$ from $P$ to $G$. Let $I^\infty(\sigma)$ be the space of all $C^\infty$ vectors in $I(\sigma)$ with the usual topology. This space can be described as the space of all $C^\infty$ mappings, $f$, from $G$ to $\mathbb{C}$ satisfying

$$f(pg) = \sigma(p)f(g)$$

with the $C^\infty$ topology. $G$ acts on $I^\infty(\sigma)$ by the right regular representation $\pi_\sigma(g)f(x) = f(xg)$. Let $K = SO(2)$. Then $G/P = K/Z = SO(2)/\{\pm 1\}$. We set $I(\sigma)_K$ equal to the space of $K$ finite vectors in $I^\infty(\sigma)$.

If $\chi$ is a unitary character of $N$ then there exists $r_\chi \in \mathbb{R}$ such that $\chi(\exp xX) = e^{ir_\chi x}$ for $x \in \mathbb{R}$. We denote by $Wh^\infty_\chi(I(\sigma))$ the space of all continuous linear functionals $\lambda : I^\infty(\sigma) \to \mathbb{C}$ such that $\lambda(\pi(n)f) = \chi(n)^{-1}\lambda(f)$ for $n \in N$ and $f \in I^\infty(\sigma)$. The following result is well known (see [J]).

**Theorem 1.** *If* $r_\chi \neq 0$ *then* $\dim Wh^\infty_\chi(I(\sigma)) = 1$.

We will now discuss the implication of this result to Whittaker models for the discrete series of $G$. We set $\mathcal{H} = \{z \in \mathbb{C} | \operatorname{Im} z > 0\}$. Let $n$ be in $\mathbb{R}$ and let $H^n_+$ denote the space of all holomorphic $f : \mathcal{H} \to \mathbb{C}$ such that

$$\int_{-\infty}^{\infty}\int_0^{\infty} |f(x+iy)|^2 y^{n-2}\,dydx < \infty.$$

We set

$$\langle f, g\rangle = \int_{-\infty}^{\infty}\int_0^{\infty} f(x+iy)\overline{g(x+iy)}y^{n-2}\,dydx$$

for $f, g$ in $H^n_+$. The group $G$ acts on the Hilbert space $H^n_+$ as follows.

$$D^+_n(g)f(z) = (-cz+a)^{-n} f\Big(\frac{dz-b}{-cz+a}\Big)$$

if

$$\nu(g) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Here we use the branch of $(-cz+a)^{-n}$ that agrees with the usual rational function for $n \in \mathbb{Z}$. Then one can check that this action defines a non-zero unitary representation of $G$ if $n \geq 1$. Setting

$$f_{k,n}(z) = \frac{(z-i)^k}{(z+i)^{k+n}}$$

and

$$k(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

then

$$D_n^+(k(\theta))f_{k,n} = e^{-(n+2k)i\theta} f_{k,n}.$$

The functions $f_{k,n}$ form a basis of the space of $K$-finite vectors in $H_+^n$, $(H_+^n)_K$. We set $H_-^n$ equal to the space of all $\bar{f}$ with $f \in H_+^n$. With $g \in G$ (given as above) acting by

$$D_n^-(g)f(z) = (-c\bar{z}+a)^{-n} f\Big(\frac{dz-b}{-cz+a}\Big).$$

Thus $(D_n^+, H_+^n)$ and $(D_n^-, H_-^n)$ are dual representations under the pairing

$$\langle f|g\rangle = \int_{-\infty}^{\infty}\int_0^{\infty} f(x+iy)g(x+iy)y^{n-2}dydx$$

with $f \in H_+^n$ and $g \in H_-^n$.

As above, if $(\pi, H)$ is a (strongly continuous) representation of $G$ with, $H^\infty$, the space of $C^\infty$ of $\pi$ and if $\chi$ is a unitary character of $N$ then we will use $Wh_\chi^\infty(H)$ for the space of all $\lambda : H^\infty \to \mathbb{C}$ that are continuous in the standard topology of $H^\infty$ and

$$\lambda(\pi(n)v) = \chi(n)^{-1}\lambda(v)$$

for $v \in H^\infty$ and $n \in N$. The following result is a consequence of a theorem of C. Moore. We will now outline a more direct proof.

**Theorem 2.** *If* $r_\chi > 0$ *then* $\dim Wh_\chi^\infty(H_+^n) = 1$ *and* $Wh_\chi^\infty(H_-^n) = 0$. *If* $r_\chi < 0$ *then* $\dim Wh_\chi^\infty(H_+^n) = 1$ *and* $Wh_\chi^\infty(H_-^n) = 0$.

*Proof.* The argument is based on a realization of the representations $H_\pm^n$ that is joint work with B. Kostant. We define the following operators on $C^\infty(\mathbb{R}^{>0})$ $(\mathbb{R}^{>0} = (0,\infty))$depending on a parameter $\alpha \in \mathbb{R}$.

$$hf(t) = 2t\frac{df}{dt}(t) + f(t), xf(t) = \frac{it}{2}f(t),$$
$$Jf(t) = 2i\Big(t\frac{d^2f}{dt^2}(t) + \frac{df}{dt}(t)\Big) + i\Big(-\frac{t}{2} + \frac{\alpha^2}{2t}\Big)f(t).$$

We note that all of these operators are formally skew-adjoint on $L^2(\mathbb{R}^{>0}, dt)$. Set $y = J + x$. Then one can check that

$$[x,y] = h, [h,x] = 2x, [h,y] = -2y.$$

We note that if $g_\alpha(t) = e^{-t/2}t^{\alpha/2}$ then

$$Jg_\alpha = -i(\alpha+1)g_\alpha.$$

We set $e = \frac{h-i(x+y)}{2}$, $f = \frac{h+i(x+y)}{2}$. Then

$$[J,e] = 2ie, [J,f] = -2if, [e,f] = -iJ.$$

We also observe that

$$eg_\alpha = 0.$$

Set $g_{n,\alpha} = f^n g_\alpha$ then

$$Jg_{n,\alpha} = -i(\alpha+1+2n)g_{n,\alpha}.$$

We set $\gamma_\alpha(X) = x$, $\gamma_\alpha(Y) = y$ and $\gamma_\alpha(H) = h$. Let $V_\alpha$ be the span over $\mathbb{C}$ of the functions $g_{n,\alpha}$ for $n \geq 0$. Then $(\gamma_\alpha, V_\alpha)$ is a representation of $Lie(SL(2,\mathbb{R}))$. We note that $J = \gamma_\alpha(Y-X)$ (note that $\exp(\theta(Y-X)) = k(\theta)^{-1}$). This implies that $(\gamma_\alpha, V_\alpha)$ for $\alpha \geq 0$, $\alpha \in \mathbb{Z}$ is equivalent with the underlying $(\mathfrak{g}, K)$ module of $H_-^{\alpha+1}$. Since an irreducible $(\mathfrak{g}, K)$ module can have at most one invariant non-zero Hermitian form up to multiple this implies that there exists a unitary operator $T : H_-^{\alpha+1} \to L^2(\mathbb{R}^{>0}, dt)$ that maps the space of $K$-finite vectors (the span of the $f_{k,\alpha+1}$) onto $V_\alpha$. We note that this implies that we can define a unique unitary representation of $G$ on $L^2(\mathbb{R}^{>0}, dt)$ with underlying $(\mathfrak{g}, K)$ module $V_\alpha$. We will use the same notation $\gamma_\alpha$ for this representation. Then $T$ defines a unitary intertwining operator between $D_{\alpha+1}^-$ and $\gamma_\alpha$. We also note that

$$\gamma_\alpha\left(\begin{bmatrix} 1 & u \\ 0 & 1 \end{bmatrix}\right) f(t) = e^{iut/2} f(t).$$

We finally come to the key point. If $r \in \mathbb{R}^{>0}$ then let $\delta_r(f) = f(r)$. Then $\delta_r$ defines a continuous functional on the space of $C^\infty$-vectors for $\gamma_\alpha$. We also note that

$$\delta_r\left(\gamma_\alpha\left(\begin{bmatrix} 1 & u \\ 0 & 1 \end{bmatrix}\right) f\right) = e^{iru/2}\delta_r(f).$$

This implies that if $r > 0$ and if

$$\chi\left(\begin{bmatrix} 1 & u \\ 0 & 1 \end{bmatrix}\right) = e^{-iru/2}$$

then $\delta_r \in Wh_\chi^\infty(H_-^{\alpha+1})$. If we apply complex conjugation we have proved that if $r_\chi > 0$ then $Wh_\chi^\infty(H_+^{\alpha+1}) \neq 0$ and if $r_\chi < 0$ then $Wh_\chi^\infty(H_-^{\alpha+1}) \neq 0$.

As is well known, if $\alpha \geq 0, \alpha \in \mathbb{N}$ then the representation

$$(H_+^{\alpha+1})^\infty \bigoplus (H_-^{\alpha+1})^\infty$$

is a quotient of $I^\infty(\sigma)$ for $\sigma$ defined as follows:

$$\sigma\left(\begin{bmatrix} h & 0 \\ 0 & h^{-1} \end{bmatrix}\right) = h^{-(\alpha+1)}$$

for $h \in \mathbb{R}^{\times}$. Fix $\alpha$ and the corresponding $\sigma$ then for each character $\chi$ on $N$ we have

$$1 = \dim Wh_{\chi}^{\infty}(I(\sigma)) \geq \dim Wh_{\chi}^{\infty}(H_{+}^{\alpha+1}) + \dim Wh_{\chi}^{\infty}(H_{-}^{\alpha+1}).$$

The theorem now follows. □

**Remark 1.** The theory of Verma modules implies that the Harish-Chandra modules $(\gamma_\alpha, V_\alpha)$ are irreducible for $\alpha > -1$. Thus these modules give an implementation of the analytic continuation of the discrete series. We refer the reader to [K2] for an account of the implications of this model. Also see section 10 in this paper.

## 3. Some properties of holomorphic representations

We will freely use the notation of the previous section. Let $\mathfrak{g}_C$ be a simple Lie algebra over $\mathbb{C}$. Then then a real form, $\mathfrak{g}$, of $\mathfrak{g}_C$ is said to be of *Hermitian type* if each Cartan involution of $\mathfrak{g}$, $\theta$, has the property that the fixed point set $\mathfrak{k}$ has a one dimensional center, $\mathfrak{c}$. Fix $\mathfrak{g}$ of Hermitian type and a Cartan involution $\theta$ of $\mathfrak{g}$. Then we say that $\mathfrak{g}$ is of *tube type* if there exists a homomorphism $\phi : \mathfrak{sl}(2,\mathbb{R}) \to \mathfrak{g}$ such that $\phi(-v^T) = \theta(\phi(v))$ for $v \in \mathfrak{sl}(2,\mathbb{R})$ (the superscript $T$ corresponds to transpose) and

$$\phi(\mathbb{R}(X - Y)) = \mathfrak{c}.$$

Throughout the rest of this section we will assume that $\mathfrak{g}$ is of tube type and we fix $\phi$ as in our definition. We set $\mathfrak{p}_C = \{x \in \mathfrak{g}_C | \theta x = -x\}$. Set $h_o = i\phi(X - Y)$. Then $ad(h_o)$ has two eigenvalues on $\mathfrak{p}_C$, $\pm 2$. Let $\mathfrak{p}^+$ denote the eigenspace with eigenvalue 2 and $\mathfrak{p}^-$ with eigenvalue $-2$. Then $\mathfrak{p}_C = \mathfrak{p}^+ \bigoplus \mathfrak{p}^-$. We note that both $\mathfrak{p}^+$ and $\mathfrak{p}^-$ are commutative Lie subalgebras of $\mathfrak{g}$. We set $\mathfrak{q} = \mathfrak{k}_C \bigoplus \mathfrak{p}^+$. Then $\mathfrak{q}$ is a parabolic subalgebra of $\mathfrak{g}_C$. If $W$ is a $\mathfrak{t}_C$ module then we extend the action to $\mathfrak{q}$ by letting $\mathfrak{p}^+$ act by 0. Let $U(\mathfrak{h})$ denote the universal enveloping algebra of a Lie algebra $\mathfrak{h}$. We will identify $U(\mathfrak{q})$ with its image in $U(\mathfrak{g}_C)$ under the natural inclusion. We define a $\mathfrak{g}_C$ module

$$N(W) = U(\mathfrak{g}_C) \bigotimes_{U(\mathfrak{q})} W.$$

Assume that $G$ is a connected Lie group with Lie algebra $\mathfrak{g}$ and that $K$ is the connected subgroup of $G$ with Lie algebra $\mathfrak{k}$. If $W$ is the differential of a finite dimensional $K$ module then $N(W)$ is a $(\mathfrak{g}, K)$-module under the action

$$k \cdot (g \bigotimes w) = \mathrm{Ad}(k)g \bigotimes kw$$

for $g \in U(\mathfrak{g}_C)$ and $w \in W$. If $W$ is irreducible then the module $N(W)$ is called a holomorphic $(\mathfrak{g},K)$ module.

We will use the notation $H = \phi(H)$. Then $ad(H)$ has three eigenvalues on $\mathfrak{g}$, $0, 2, -2$. We set $\mathfrak{n}$ equal to the eigenspace for eigenvalue 2 and $\mathfrak{m}$ the eigenspace for eigenvalue 0. Then $\mathfrak{m}$ is the Lie algebra of $M = \{g \in G | \mathrm{Ad}(g)H = H\}$ and $\mathfrak{n}$ is an abelian Lie algebra. Let $N = \exp \mathfrak{n}$ (we apologize for the two uses of the letter $N$ there should be no confusion since the group and the module will be clearly separated by context). Then $Q_1 = MN$ is a parabolic subgroup of $G$. Set $\mathfrak{q}_1 = Lie(M)_C \bigoplus Lie(N)_C$. Then one can show (see [Wa2] and [E-W]) that $\mathfrak{q}$ and $\mathfrak{q}_1$ are opposite parabolic subalgebras (that is $\mathfrak{q} \cap \mathfrak{q}_1$ is a Levi factor of both $\mathfrak{q}$ and $\mathfrak{q}_1$). This implies that there exists an element $n \in N_C$ and $p \in \exp(\mathfrak{p}^+)$ such that $\mathrm{Ad}(p)Lie(K)_C = \mathrm{Ad}(n)Lie(M)_C$. We observe that if $x \in \mathfrak{p}^+$ and $v \in N(W)$ then there exists $k$ depending on $x$ and $v$ such that $x^k v = 0$. We therefore see that $N(W)$ has a $K_C \exp(\mathfrak{p}^+)$-module structure. We can thus define a $Lie(M)_C$ module structure on $W$ given by $1 \bigotimes uw = \mathrm{Ad}(n)(u)(1 \bigotimes w)$ for $w \in W$ and $u \in Lie(M)_C$. We will use the notation $\widetilde{W}$ for any element in the isomorphism class of this module. Let $G_C$ be a connected simply connected Lie group with Lie algebra $\mathfrak{g}_C$ and let $G$ be the connected subgroup of $G_C$ with Lie algebra $\mathfrak{g}$. Suppose that $W$ is a finite dimensional $K$-module which we extend to a holomorphic $K_C$ module. Then $n^{-1}Mn$ is contained in $K_C$ and thus we have an isomorphism class of $M$-modules whose differentials give the class of the $\widetilde{W}$. We will also use the notation $\widetilde{W}$ for any element of this isomorphism class of $M$-modules. The fact that the two parabolic subalgebras are opposite now has the following consequence.

**Theorem 3.** *Assume that $W$ is a finite dimensional $\mathfrak{k}_C$ module then as $(q_1, K \cap M)$-module $N(W)$ is isomorphic with $U(\mathfrak{n}_C) \bigotimes \widetilde{W}$.*

If $V$ is a $\mathfrak{g}$-module and if $\psi : \mathfrak{n} \to \mathbb{C}$ is an $\mathbb{R}$-linear map (hence a Lie algebra homomorphism since $\mathfrak{n}$ is commutative) then we denote by $Wh_\psi(V)$ ($Wh$ for Whittaker vector, as usual) the space of all elements $\lambda \in V^*$ such that $\lambda(xv) = -\psi(x)\lambda(v)$ for $v \in V$ and $x \in \mathfrak{n}$. If $U$ is a finite dimensional $(\mathfrak{m}, K \cap M)$ module then there exists a unique representation of $M$ on $U$ whose differential yields that module structure. We will use the same notation for that $M$-module. If $\psi \in \mathfrak{n}^*$ then we set $M_\psi = \{m \in M | \psi \circ \mathrm{Ad}(m) = \psi\}$. With all of this notation and these conventions in place the above theorem has the following immediate consequence.

**Corollary 1.** *Let $W$ be a finite dimensional $K$-module and let $\psi$ be an element of $\mathfrak{n}_C^*$ then $Wh_\psi(N(W))$ is isomorphic as an $M_\psi$-module with the module contragredient to $\widetilde{W}$.*

This result says that there are always formal (e.g. algebraic) Whittaker vectors. We have seen in the previous section that under some regularity conditions such vectors do not necessarily exist. We will now derive a theorem that generalizes Theorem 2. We first recall the structure of the holomorphic discrete series. We won't be as explicit as we were in the case of $SL(2,\mathbb{R})$. We fix a maximal abelian subalgebra, $\mathfrak{t}$, of $\mathfrak{k}$ then $\mathfrak{t}_C$ is a Cartan subalgebra of $\mathfrak{g}_C$ and of $\mathfrak{k}_C$. Let $\Phi$ denote the root system of $\mathfrak{g}_C$ with respect to $\mathfrak{t}_C$ and let $\Phi_k$ be the root system of $\mathfrak{k}_C$ with respect to $\mathfrak{t}_C$. Let $\Phi_k^+$ be a fixed system of positive roots for $\mathfrak{k}_C$ and let $\Phi^+$ be the system of positive roots in $\Phi$ given by $\Phi_k^+ \cup \{\alpha \in \Phi | \alpha(h_o) = 2\}$. Let $\rho$ denote the half sum of the positive roots of $\Phi^+$ and let $\alpha_o$ denote the unique simple root in $\Phi^+$ that is not in $\Phi_k^+$. Let $Z$ denote the center of $G$. Recall that a unitary representation, $(\pi, H)$, of $G$ is said to be square integrable modulo the center if

$$\int_{G/Z} |\langle \pi(g)v, v\rangle|^2 dg < \infty$$

for all $v \in H$. The basic result on the holomorphic discrete series is the following theorem of Harish-Chandra [HC].

**Theorem 4.** *Let $W$ be an irreducible, unitary $K$-module with highest weight $\lambda$ relative to $\Phi_k^+$. Then a necessary and sufficient condition that $N(W)$ be the underlying $(\mathfrak{g}, K)$ module of a representation square integrable modulo the center of $G$ is that $(\lambda + \rho, \alpha_o) < 0$.*

If $W$ satisfies the condition of the above theorem then we will use the notation $(\pi_{\lambda+\rho}, H^{\lambda+\rho})$ for a choice of an irreducible square integrable representation of $G$ with underlying $(\mathfrak{g}, K)$-module $N(W)$.

Let $N$ denote the connected subgroup of $G$ with Lie algebra $\mathfrak{n}$ (that is $N = \exp \mathfrak{n}$) as above. If $\psi$ is a unitary character of $N$ we will also use the notation $\psi$ for its differential. The non-degeneracy of $B$ implies that there exists $y_\psi \in \theta\mathfrak{n}$ such that $\psi(x) = iB(y_\psi, x)$ for $x \in \mathfrak{n}$.

If $(\pi, V)$ is a (strongly continuous) representation of $G$ on a Hilbert space the $H^\infty$ will denote the space of all $C^\infty$ vectors with the usual topology. If $\psi$ is a unitary character of $N$ then we will use the notation $Wh_\psi^\infty(V)$ to denote the space of all continuous functionals $\lambda$ on $H^\infty$ such that $\lambda(\pi(n)v) = \psi(n)^{-1}\lambda(v)$ for all $n \in N$ and $v \in H^\infty$. Then we note that ($N(W)$ the space of $K$-finite vectors of $H^{\lambda+\rho}$).

$$Wh_\psi^\infty(H^{\lambda+\rho})_{|N(W)} \subset Wh_\psi(N(W))$$

as an $M_\psi$-submodule.

Let $S_o$ denote the simply connected covering group $SL(2, \mathbb{R})$. Let $\phi : S_o \to G$ be the homomorphism whose differential is what we have been calling $\phi$. By going to a covering group of $G$ we may assume that $\phi$ is injective. Set $S = \phi(S_o)$. We note that the representations $(\pi_{\lambda+\rho}, H^{\lambda+\rho})$ are unitary and when restricted to $C$, the center of $K$ are admissible. Thus since $S$ contains $C$ it and the eigenvalues of $-i\phi(X - Y)$ are strictly negative it follows that as a representation of $S$.

$$H^{\lambda+\rho} \cong \bigoplus_n m_n(\lambda) H_+^n$$

a unitary countable direct sum with finite multiplicities $m_n(\lambda)$.

**Lemma 1.** *Let $\psi$ be a nontrivial unitary character of $N$. Let $(\pi, V)$ be a representation of $G$ on a Banach space and let $\lambda$ be a continuous linear functional on*

*$V^\infty$ such that*

$$\lambda(\pi(n)v) = \psi(n)^{-1}\lambda(v)$$

*for $n \in N$ and $v \in V^\infty$. Then for each $k > 0$ there exists a continuous seminorm $\mu_k$ on $V^\infty$ such that*

$$\lambda(\pi(\exp(tH)v) \leq e^{-kt}\mu_k(v)$$

*for all $t > 0$ and $v \in V^\infty$.*

*Proof.* We observe that there exists a continuous seminorm $\kappa$ on $V^\infty$ such that

$$|\lambda(\pi(g)v)| \leq \|\nu(g)\|^r\kappa(v)$$

for all $g \in G$ and $v \in V^\infty$ (here $\nu$ is a covering homomorphism onto a linear Lie group with Lie algebra isomorphic with $Lie(G)$ and $\|\ldots\|$is a norm on $\nu(G)$ see [Wa3]). We will write $\|g\| = \|\nu(g)\|$Since $\|\ldots\|$ is a norm it follows that there exists $m \in \mathbb{N}$ and $C > 0$ such that

$$\|\exp(tH)\| \leq Ce^{mt}$$

for $t > 0$. Since $\psi$ is nontrivial there exists $X \in \mathfrak{n}$ such that $d\psi(X) = -i$. Thus

$$i\lambda(v) = \lambda(\pi(X)v)$$

for all $v \in V^\infty$. If $t \in \mathbb{R}$ then we have

$$i\lambda(\pi(\exp(tH)v) = \lambda(\pi(X)\pi(\exp(tH)v) =$$
$$\lambda(\pi(\exp(tH)\pi(\mathrm{Ad}(\exp(-tH))X)v) = e^{-2t}\lambda(\pi(\exp(tH)\pi(X)v).$$

Hence

$$|\lambda(\pi(\exp(tH)v)| \leq Ce^{(m-2)t}\kappa(\pi(X)v)$$

for all $v \in V^\infty$ and $t \geq 0$. Set $\xi(v) = C\kappa(\pi(X)v)$. Repeating this argument we find

$$|\lambda(\pi(\exp(tH)v)| \leq e^{(m-4)t}\xi(\pi(X)v).$$

The obvious iteration of this idea implies the result. □

If $x \in \mathfrak{n}$ then we will say that $x$ is *positive* if $x \in \mathrm{Ad}(M)X$. We will write $x > 0$ if $x$ is positive. We will say that a unitary character, $\psi$ of $N$ is positive if $-id\psi(x) > 0$ whenever $x > 0$. The main result of this section is

**Theorem 5.** *Let $\lambda+\rho$ be a Harish-Chandra parameter for a Holomorphic discrete series. A necessary and sufficient condition that*

$$Wh_\psi^\infty(H^{\lambda+\rho}) \neq 0$$

*is that $\psi$ be either positive or the trivial character if this is so then*

$$Wh_\psi^\infty(H^{\lambda+\rho})_{|N(W)} = Wh_\psi(N(W))$$

*as an $M_\psi$-module.*

**Remark 2.** In light of Corollary 1 this theorem gives a complete description of continuous generalized Whittaker models for holomorphic discrete series.

*Proof.* We first observe that as a representation of $S$ the representation $\pi_{\lambda+\rho} \circ \phi$ splits into a direct sum with finite multiplicities of the representations $H_+^k$ (see the observations above). If there exists $x \in \mathrm{Ad}(M)X$ such that $-i\psi(x) < 0$ then replacing $\phi$ by $\phi \circ \mathrm{Ad}(m)^{-1}$ the restriction of any element of $Wh_\psi(N(W))$ cannot extend to a continuous linear functional on any of the spaces $(H_+^k)^\infty$. But this implies that $Wh_\psi^\infty(H^{\lambda+\rho}) = 0$. If there exists $x \in \mathrm{Ad}(M)X$ such that $\psi(x) = 0$ but $\psi \neq 0$ then replacing $\psi$ by $\psi \circ \mathrm{Ad}(m)^{-1}$ (if necessary) we may assume $\psi(X) = 0$. Thus if $V$ is an irreducible $S$ direct summand of $H^{\lambda+\rho}$ then $V$ is equivalent with $H_+^k$ for some $k$. The space of $K \cap S$-finite vectors in $V$ consists of $K$-finite vectors, which we will denote $V_1$. If $\lambda \in Wh_\psi^\infty(H^{\lambda+\rho})$ then $\lambda_{|V_1}$ factors through $V_1/XV_1$ which is a one dimensional space. The element $H$ thus acts a scalar $\mu$. This implies that if $v \in V_1$ then $\lambda(\pi(\exp(tH))v) = e^{-\mu t}\lambda(v)$. Now the previous lemma implies that $\lambda_{|V_1} = 0$. Since the sum of the spaces $V_1$ (as described) is dense in $(H^{\lambda+\rho})^\infty$. Hence $Wh_\psi^\infty(H^{\lambda+\rho}) = 0$ in this case also.

To complete the proof we must (only) show that if $\psi$ is positive or if $\psi = 0$ then $Wh_\psi^\infty(H^{\lambda+\rho}) \neq 0$. If $\psi = 0$ this is an immediate consequence of the fact that $N(W)/\mathfrak{n}N(W) \neq 0$. We may therefore assume that $\psi$ is positive.

To complete the proof of the theorem we will need several structural results that we shall see have analogues in the quaternionic case. We will therefore interrupt the present proof to present these ideas.

We denote by $P_\mathfrak{p}$ the projection of $\mathfrak{g}$ onto $\mathfrak{p}$ given by $P_\mathfrak{p}(x) = \frac{x-\theta x}{2}$ The key result is

**Lemma 2.** *Set* $\mathfrak{k}_1 = \mathfrak{k} \cap \mathfrak{m}$ *and* $\mathfrak{p}_1 = P_\mathfrak{p}\mathfrak{n}$. *Then we have*

$$[\mathfrak{p}_1, \mathfrak{p}_1] \subset \mathfrak{k}_1, [\mathfrak{k}_1, \mathfrak{p}_1] \subset \mathfrak{p}_1.$$

*In other words, if we set* $\mathfrak{g}_1 = \mathfrak{k}_1 \bigoplus \mathfrak{p}_1$, *then it is a reductive Lie algebra over* $\mathbb{R}$ *with given Cartan decomposition.*

The proof is a direct calculation using $[\mathfrak{n}, \mathfrak{n}] = 0$ and $[H, [\mathfrak{n}, \theta\mathfrak{n}]] = 0$.

We now consider the decomposition $\mathfrak{g} = \theta\mathfrak{n} \bigoplus \mathfrak{m} \bigoplus \mathfrak{n}$. Let $P_\mathfrak{n}$ denote the projection onto $\mathfrak{n}$ corresponding to this decomposition. We note that all of the projections that we have defined intertwine the adjoint action of $M \cap K$ on $\mathfrak{g}$. This implies that if $\mathfrak{a}$ is a maximal abelian subspace of $\mathfrak{p}_1$ then since $\mathrm{Ad}(M \cap K)\mathfrak{a} = \mathfrak{p}_1$ it follows that $\mathrm{Ad}(M \cap K)P_\mathfrak{n}\mathfrak{a} = \mathfrak{n}$. Fix a maximal torus, $T$, of $K$ and let $\Phi^+$ be a system of positive roots compatible with the parabolic subalgebra $\mathfrak{q} = \mathfrak{k} \bigoplus \mathfrak{p}^+$. Let $\gamma_1 \dots, \gamma_l$ be the system of strongly orthogonal roots constructed as in [He, pp. 385–386]. Let $\mathfrak{u}_i$ be the Lie subalgebra of $\mathfrak{g}_\mathbb{C}$ isomorphic with $\mathfrak{sl}(2.\mathbb{C})$ that has roots $\pm\gamma_i$. We note that the complexification of our original algebra with basis $X, Y, H$ is the diagonal subalgebra of the direct sum of the $\mathfrak{u}_i$. We may choose

$0 \neq X_i \in \mathfrak{u}_i \cap \mathfrak{n}$ such that $X_i - \theta X_i \in \mathfrak{p}_1$. Then $\sum_i \mathbb{R}(X_i - \theta X_i)$ is maximal abelian in $\mathfrak{p}$ hence in $\mathfrak{p}_1$. This implies that

$$\mathrm{Ad}(M \cap K)\left(\sum_i \mathbb{R}X_i\right) = \mathfrak{n}.$$

We also note that we can choose $X_i$ such that $\sum X_i = X$. Furthermore, $\mathrm{Ad}(M \cap K)X = X$.

The Cartan decomposition implies that $M = (K \cap M)\exp(\mathfrak{a})(K \cap M)$ we assert that this implies

**Lemma 3.** *$\psi$ is positive or trivial if and only if $\psi(x) = iB(y,x)$ with $y = -\mathrm{Ad}(k)\left(\sum_i y_i \theta X_i\right)$ with $y_i \geq 0$ and $k \in M \cap K$.*

*Proof.* By the above we see that since $X = \sum X_i$ we have

$$\mathrm{Ad}(M)X = \{\mathrm{Ad}(k)\sum x_i X_i | x_i > 0, k \in K \cap M\}.$$

If $Y \in \theta\mathfrak{n}$ then $Y = -\mathrm{Ad}(k)\sum y_i \theta X_i$ with $y_i \in \mathbb{R}$ and $k \in M \cap K$. Thus the condition of positivity comes down to

$$B\left(\sum y_i \theta X_i, \mathrm{Ad}(k)\sum x_i X_i\right) < 0$$

for all $k \in K \cap M$ and all $x_i > 0$. In particular if we take $k = 1$ this implies that $y_i \geq 0$. If all of the $y_i > 0$ then $y = -\mathrm{Ad}(m_1)\left(\sum \theta X_i\right)$ with $m_1 \in M$. Thus

$$\begin{aligned} B(y, \mathrm{Ad}(m)X) = -B(\theta X, \mathrm{Ad}(m_1)^{-1}\mathrm{Ad}(m)X) = \\ - B(\theta X, \mathrm{Ad}(k)\mathrm{Ad}(a)X) \end{aligned}$$

since $m_1^{-1}m = kak_1$ with $k, k_1 \in K \cap M$ and $a \in A$. Since $\mathrm{Ad}(K \cap M)X = X$ we have

$$B(y, \mathrm{Ad}(m)X) = -B(\theta X, \mathrm{Ad}(a)X) = -\sum a^{\gamma_i} B(\theta X_i, X_i) > 0.$$

□

We will now complete the proof of Theorem 5. We note that $M$ acts transitively on the interior of the set of positive elements. This implies that if $Wh_{\psi}^{\infty}(H^{\lambda+\rho}) \neq 0$ for one positive element in the interior then it is nonzero for all of them. If $v \in N(W)$ and $v \neq 0$ then the function

$$f_v(n) = \langle \pi(n)v, v \rangle$$

is nonzero and in $L^2(N) \cap L^1(N)$ (cf. [Wa3, Theorem 4.5.4, p. 126]). This implies that there exists an open subset of unitary characters $\chi$ such that

$$\int_N \chi(n)^{-1} f_v(n) dn \neq 0.$$

Now the functional

$$\lambda(w) = \int_N \chi(n)^{-1} \langle \pi(n)w, v\rangle \, dn$$

for $w \in (H^{\lambda+\rho})^\infty$ defines an element of $Wh_\chi^\infty(H^{\lambda+\rho})$. Thus there is an open subset of $\chi$ such that $Wh_\psi^\infty(H^{\lambda+\rho}) \neq 0$. Now the above vanishing results imply that if $Wh_\psi^\infty(H^{\lambda+\rho}) = 0$ for all $\psi$ positive in the interior then we have a contradiction. We now look at the boundary of the set. Suppose that $\psi(X) = iB(Y, X)$. Then up to conjugacy by an element of $M$ we may assume that $Y = -\sum y_i \theta X_i$ with $y_i > 0$ for $i \leq r$ and $y_i = 0$ for $i > r$. Set $h_r = H_{r+1} + \cdots + H_l$ ($H_i = -[X_i, \theta X_i]$). Let $\mathfrak{n}_r$ be the sum of the eigenspaces for (strictly) positive eigenvalues of $ad(h_r)$. Let $\widetilde{\mathfrak{g}_r}$ be the centralizer of $h_r$ in $\mathfrak{g}$. We denote by $\mathfrak{g}_r$ the ideal in $\widetilde{\mathfrak{g}_r}$ generated by $\mathfrak{n} \cap \widetilde{\mathfrak{g}_r}$ and $\theta(\mathfrak{n} \cap \widetilde{\mathfrak{g}_r})$. Then $\mathfrak{g}_r$ is a Lie algebra of the type we have been studying and $(H^{\lambda+\rho})^\infty / \overline{\theta\mathfrak{n}_r (H^{\lambda+\rho})^\infty}$ is the space of smooth vectors for a holomorphic discrete series for the connected real Lie subgroup $G_r$ of $G$ with Lie algebra $\mathfrak{g}_r$. In this context the restriction of $\psi$ to $N \cap G_r$ is in the interior of the set of positive characters. The preceding argument now implies the full result.

In section 10 we will show how these results and the techniques of section 9 can be used to prove the theorem announced in the introduction.

## 4. The case of $SU(2,1)$

In this section we will prove several results analogous to those of section 2 (the case of $SL(2,\mathbb{R})$) for the quaternionic (hence in this case) generic discrete series for $G = SU(2,1)$. To fix notation we take $(z, w) = z_1\overline{w}_1 + z_2\overline{w}_2 - z_3\overline{w}_3$. Then $G$ is the group of all elements of $SL(3,\mathbb{C})$ such that $(gz, gw) = (z, w)$ for all $z, w \in \mathbb{C}^3$. We will write $\mathfrak{g} = Lie(G)$. We take $K$ to be the subgroup of all matrices in $G$ of the form

$$\begin{bmatrix} u & 0 \\ 0 & \det(u)^{-1} \end{bmatrix}$$

with $u \in U(2)$. We take $H$ to be the matrix

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

and $\mathfrak{a} = \mathbb{R}H$, $A = e^{\mathbb{R}H}$ in $G$. We denote by $M$ the subgroup of $K$ consisting of matrices of the form

$$\begin{bmatrix} u & 0 & 0 \\ 0 & u^{-2} & 0 \\ 0 & 0 & u \end{bmatrix}$$

with $u \in \mathbb{C}$ and $|u| = 1$. The element $ad(H)$ has eigenvalues $0, \pm 1, \pm 2$ on $\mathfrak{g}$. We set $\mathfrak{n}$ equal to the direct sum of the eigenspaces for $adH$ corresponding to the eigenvalues 1 and 2. We put $N = e^{\mathfrak{n}}$. Then $KAN$ is an Iwasawa decomposition of $G$ and $P = MAN$ is a proper parabolic subgroup of $G$ given in terms of its standard Langlands decomposition corresponding to our choice of $K$. Let

$$I_{2,1} = \begin{bmatrix} I & 0 \\ 0 & -1 \end{bmatrix}.$$

Then the Cartan involution of $G$ corresponding to $K$ is $\theta(g) = I_{2,1} g I_{2,1}$.

We set

$$X = \frac{1}{2} \begin{bmatrix} i & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & -i \end{bmatrix}$$

and

$$Y = \frac{1}{2} \begin{bmatrix} -i & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & i \end{bmatrix}.$$

Then $\mathbb{R}X = [\mathfrak{n}, \mathfrak{n}]$ and we have

$$[X, Y] = H, [H, X] = 2X, [H, Y] = -2Y.$$

Also $\mathbb{R}X + \mathbb{R}H + \mathbb{R}Y$ form the Lie algebra of the subgroup $H_o \cong SU(1,1)$ imbedded in $G_o$ as

$$\begin{bmatrix} a & b \\ \overline{b} & \overline{a} \end{bmatrix} \longmapsto \begin{bmatrix} a & 0 & b \\ 0 & 1 & 0 \\ \overline{b} & 0 & \overline{a} \end{bmatrix}.$$

We note that $\mathbb{C}X + \mathbb{C}Y + \mathbb{C}H$ is just the Lie algebra of $SL(2, \mathbb{C})$ imbedded in $\mathfrak{g}_C = \mathfrak{sl}(3, \mathbb{C})$ as follows

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix} \longmapsto \begin{bmatrix} a & 0 & b \\ 0 & 0 & 0 \\ c & 0 & -a \end{bmatrix}.$$

Let

$$x = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; h = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Let $\mathfrak{b}$ denote the Borel subalgebra of $\mathfrak{g}_C$ consisting of trace 0 and of the form

$$\begin{bmatrix} * & * & * \\ 0 & * & 0 \\ 0 & * & * \end{bmatrix}.$$

Let $\mathfrak{h}$ denote the diagonal Cartan subalgebra of $\mathfrak{g}_C$. Then $\mathfrak{b}$ is $\theta$-stable and $\mathfrak{b} \cap \mathfrak{k}_C = \mathfrak{h} + \mathbb{C}x$. Thus the simple positive roots corresponding to this choice of Cartan subalgebra and Borel subalgebra are both non-compact. We write $\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{u}$ with $\mathfrak{u}$ the nilradical of $\mathfrak{b}$. We note that, in the usual $\varepsilon$ notation ([Bour]) the simple roots are $\alpha_1 = \varepsilon_1 - \varepsilon_3$ and $\alpha_2 = \varepsilon_3 - \varepsilon_2$ and the compact positive root is $\beta = \varepsilon_1 - \varepsilon_3 = \alpha_1 + \alpha_2$. The coroot associated with $\beta$ is $h$. We also write $\Lambda_1, \Lambda_2$ dominant integral forms defined by $\Lambda_i(\check{\alpha}_j) = \delta_{ij}$ (the upper check indicates coroot).

Associated with this choice of Borel subalgebra is the generic discrete series. Which we will now study. Let $T \subset G$ be the maximal torus of $K$ with complexified Lie algebra $\mathfrak{h}$. That is, the diagonal elements of $G$. If $\chi$ is a unitary character of $T$ with $\Lambda = d\chi \in \mathfrak{h}^*$ then we write $\chi = \chi_\Lambda$ and we will also use the notation $t^\Lambda = \chi(t)$. If $\Lambda$ is regular and dominant with respect to the choice of $\mathfrak{b}$ then there is an associated an equivalence class of square integrable representations of $G$. We choose a representative $(\pi_\chi, H^\chi) = (\pi_\Lambda, H^\Lambda)$ of the class.

At the other extreme there is the principal series of representations of $G$. Let $\xi$ be a unitary character of $M$ and let $\nu$ be a (quasi-)character of $A$ which we will write as $a \longmapsto a^\nu$ with $\nu \in \mathfrak{a}_C^*$. We denote by $(\pi_{\xi,\nu}, H^{\xi,\nu})$ the representation of $G$ induced from the quasi-character $man \longmapsto \xi(m)a^\nu$. The space of $H^{\xi,\nu}$ is a subspace of the functions $f : G \to \mathbb{C}$ such that $f(mang) = \xi(m)a^{\nu+\rho}f(g)$ for all $m \in N$, $a \in A$ and $n \in N$ and the action by $G$ is by right translation. Here, as usual, $\rho(h) = \frac{1}{2}\mathrm{tr}(adh_{|n})$ for $h \in \mathfrak{a}$. Equivalently, $\rho(H) = 2$. We note that $Lie(P)_C$ is a Borel subalgebra of $\mathfrak{g}_C$ and $\mathfrak{h}_1 = (\mathfrak{a} \oplus Lie(\mathfrak{m}))_C$ is a Cartan subalgebra of $\mathfrak{g}_C$.

Let $\sigma$ be an inner automorphism of $\mathfrak{g}_C$ such that $\sigma(\mathfrak{b}) = Lie(P)_C$ and $\sigma(\mathfrak{h}) = \mathfrak{h}_1$. It is unique up to multiplication by an automorphism of the form $e^{ad(h)}$ with $h \in \mathfrak{h}$. We choose it as $\sigma = \mathrm{Ad}(g)$ with

$$g = \begin{bmatrix} \frac{i}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & -1 \\ \frac{i}{\sqrt{2}} & \frac{-i}{\sqrt{2}} & 0 \end{bmatrix}.$$

Then $\Lambda \circ \sigma$ defines a Lie algebra homomorphism of $Lie(P)_C$ to $\mathbb{C}$. This homomorphism exponentiates to a quasi-character of $P$. We will write $ma \longmapsto \xi_\Lambda(m)a^{\nu_\Lambda}$ for the corresponding character of $MA$. We observe that if $\mu$ is a dominant integral weight and if $\Phi^\mu$ and $\Phi_{-\mu}$ are the corresponding Zuckerman translation functors (cf. [Wa3], [K-V]) than (sub $K$ denotes $K$-finite vectors) $\Phi^\mu((H^\Lambda)_K) \cong (H^{\Lambda+\mu})_K$, $\Phi_{-\mu}((H^{\Lambda+\mu})_K) \cong (H^\Lambda)_K$ and if $\sigma$ is an element of the Weyl group then $\Phi^{s\mu}((H^{\xi_{s\Lambda},\nu_{s\Lambda}})_K) \cong (H^{\xi_{s\Lambda+s\mu},\nu_{s\Lambda+s\mu}})_K$ and $\Phi_{-s\mu}((H^{\xi_{s\Lambda+s\mu},\nu_{s\Lambda+s\mu}})_K) \cong (H^{\xi_{s\Lambda},\nu_{s\Lambda}})_K$. These assertions will be abbreviated to the statement that the families $\Lambda \to H_K^\Lambda$ and $\Lambda \to H^{\xi_\Lambda,\nu_\Lambda}$ are coherent.

If $(\pi, H)$ is a representation of $G$ we will use the notation $H_\infty$ for the space if $C^\infty$ vectors in $H$ relative to the action $\pi$ and $H_K$ for the space of $K$-finite vectors.

**Theorem 6.** *Let $\Lambda$ be a unitary character of $T$ that is dominant and regular relative to $\mathfrak{b}$ which we write as $m_1\Lambda_1 + m_2\Lambda_2$. There exist continuous injective $G$-intertwining operator $R_\Lambda^j$ from $(H^\Lambda)_\infty$ into $(H^{\xi_{\mu_j},\nu_{\mu_j}})_\infty$ for $j = 1, 2, 3$ and $\mu_1 = \Lambda$, $\mu_2 = (m_1 + m_2)\Lambda_1 - m_2\Lambda_2$ and $\mu_3 = -m_1\Lambda_1 + (m_1 + m_2)\Lambda_2$. We write $R_\Lambda$ for $R_\Lambda^1$.*

*Proof.* In [Wa1, p. 183] on can find this result stated for $K$-finite vectors. The $K$-finite case implies the result for $C^\infty$ in light of [Ca], [Wa4]. Since there are no proofs in [Wall1] we will give a sketch of a proof based on [J-W] and [K-W]. [K-W] gives 2 Szegő quotient maps for this discrete series. The conjugate dual parameters are exactly the ones corresponding to $\mu_2$ and $\mu_3$ in the statement (see [K-W, Theorem 10.8, p. 195]). We now consider $R_\Lambda$. We first look at the case when $\xi_\Lambda = 1$, that is, the corresponding principal series representation is spherical. Then we must have $\Lambda = m\beta$ with $m > 0$ and $m \in \mathbb{Z}$. Hence $\nu = m\rho$. Then in this case the result is now an immediate consequence of Theorem 5.1 (3) and Theorem 8.2 in [J-W]. To complete the proof we observe that both of the families $\Lambda \to H_K^\Lambda$ and $\Lambda \to H^{\xi_\Lambda,\nu_\Lambda}$ are coherent (see the discussion preceding the statement of the theorem). □

We will now examine in more detail the case (that appears in the proof above) when we have an embedding into a spherical principal series representation (that is $\xi \equiv 1$). As in [Wa1] this corresponds to the case $\langle\Lambda,\alpha_1\rangle = \langle\Lambda,\alpha_2\rangle$. This implies (in light of the integrality and dominance assumption) that $\Lambda = m\beta$ with $m > 0$. In this case we have $\xi = 1$ and $\nu = m\rho$.

We note that the conjugate dual principal series to $\pi_{1,m\rho}$ is $\pi_{1,-m\rho}$. Thus $\pi_\Lambda$ appears as a quotient of $\pi_{1,-m\rho}$. Let $\psi : N \to S^1$ be a unitary character. Then we have the following spaces:

1. $Wh_\psi^K(H^{\xi,\nu}) = \{\lambda \in (H_K^{\xi,\nu})^* | X \cdot \lambda = d\psi(X)\lambda\}$.
2. $Wh_\psi^\infty(H^{\xi,\nu}) = \{\lambda \in (H_\infty^{\xi,\nu})' | \lambda \circ \pi_{\xi,\nu}(n)^{-1} = \psi(n)\lambda\}$.

Here the "prime" indicates the continuous dual of the Fréchet space $H_\infty^{\xi,\nu}$. If $\psi$ is generic (which means in this special case that $\psi \neq 1$) and if $m > 0$ then since the only generic subquotient of $H_K^{1,\frac{m}{2}\rho}$ is $H_K^{\frac{m}{2}\beta}$ the maps

$$H_K^{m\beta} \to H_K^{1,m\rho}, H_K^{1,-m\rho} \to H_K^{m\beta}$$

induce respectively (via pullback) isomorphisms [Ko1] implies that

$$Wh_\psi^K(H^{1,m\rho}) \cong Wh_\psi^K(H^{m\beta}), Wh_\psi^K(H^{-m\beta}) \cong Wh_\psi^K(H^{1,m\rho}).$$

Also at the level of $C^\infty$ vectors the work of [Ca] and [Wa3, Chapter 11] imply that we have continuous $G$-homomorphisms

$$H_\infty^{m\beta} \to H_\infty^{1,m\rho}, H_\infty^{1,-m\rho} \to H_\infty^{m\beta}$$

which induce

$$Wh_\psi^\infty(H^{1,m\rho}) \cong Wh_\psi^\infty(H^{m\beta}), Wh_\psi^\infty(H^{-m\beta}) \cong Wh_\psi^\infty(H^{1,m\rho}).$$

Let $s_o$ be the non-trivial element of the Weyl group of $G$ acting on $MA$. (Then $s_{o|\mathfrak{a}} = -1$.) We will fix a representative an element $s^* \in K$ such that $\mathrm{Ad}(s^*) = s_o$. We set for $f \in H_\infty^{\xi,\nu}$

$$J_{\xi,\nu}^\psi(f) = \int_N \psi(n) f(s^* n) dn.$$

Here we have made a choice of Haar measure on $N$ and we observe that standard estimates of Harish-Chandra (cf. [Wa3, Theorem 4.5.4, p. 126]) imply that this integral converges absolutely and uniformly in compacta of $\{\nu \in \mathfrak{a}_C^* | \operatorname{Re}\langle \nu, \rho\rangle > 0\}$ and $J_{\xi,\nu}^{\psi} \in Wh_{\psi}^{\infty}(H^{\xi,\nu})$ for $\nu$ in this set.

**Theorem 7.** *Let $\Lambda$ be a unitary character of $T$ that is dominant and regular relative to $\mathfrak{b}$ and let $\psi$ be a generic unitary character of $N$ then*

1. $\dim Wh_{\psi}^{K}(H^{\Lambda}) = \dim\{\lambda \in (H_K^{\Lambda})^* | X \cdot \lambda = d\psi(X)\lambda\} = 2.$
2. $Wh_{\psi}^{\infty}(H^{\Lambda}) = \{\lambda \in (H_{\infty}^{\Lambda})' | \lambda \circ \pi_{\xi,\nu}(n)^{-1} = \psi(n)\lambda\} = \mathbb{C}J_{\xi_\Lambda,\nu_\Lambda}^{\psi} \circ R_{\Lambda}.$
3. *Let $v \in (H^{\Lambda})_{\infty}$ then there exists a constant $C(v) \in \mathbb{C}$ such that*

$$\int_N \psi(n)\,\langle \pi(n)w, v\rangle\, dn = C(v) J_{\xi_\Lambda,\nu_\Lambda}^{\psi}(R_\Lambda w)$$

*for $w \in (H^{\Lambda})_{\infty}$. Furthermore, the integral in the left hand side of this equation is absolutely convergent and the conjugate linear map $v \longmapsto C(v)$ is not identically* $0$.

*Proof.* The first result is due to Kostant [Ko1]. The second can be proved as follows. If $\Lambda = m\beta$ with $m > 0$ then the result follows from the multiplicity 1 theorem for Whittaker models of spherical principal series. For the general case follows by using the observation of coherence of both the principal series and the discrete series (as in the proof of Theorem 6) and the methods of the proof of Theorem 7.2 in [Wa5]. The last assertion follows from Lemma 15.3.7, Lemma 15.7.5 and the proof of Theorem 15.7.1 (in 15.7.6) of [Wa3]. □

## 5. Some properties of quaternionic real forms

Let $G_C$ be a connected simply connected simple Lie group over $\mathbb{C}$. Let $G$ be a quaternionic real form of $G_C$ and let $K \subset G$ be a maximal compact subgroup. We recall that (up to conjugacy) there is only one quaternionic real form and one is described as follows. Fix $T_C$ a Cartan subgroup of $G_C$ with Lie algebra $\mathfrak{h}$. Let $\mathfrak{g}_C = Lie(G)$ identified with the complexification of $\mathfrak{g} = Lie(G)$. Let $\Phi$ be the root system of the $h$ acting on $\mathfrak{g}_C$. Fix $\Phi^+$ a system of positive roots and $\mathfrak{g}_u$ a compact real form of $\mathfrak{g}$ such that $\mathfrak{h} \cap \mathfrak{g}_u$ is a maximal abelian subalgebra. Let $\beta$ be the largest root of $\Phi^+$ and let $\check{\beta} \in \mathfrak{h}$ be the corresponding coroot. Set $\theta = e^{\pi i ad(\check{\beta})}$. Set $\mathfrak{g}_u^{\pm} = \{X \in \mathfrak{g}_u | \theta X = \pm X\}$. Then $G$ is the connected subgroup of $G_C$ such that $Lie(G) = \mathfrak{g}_u^{+} + i\mathfrak{g}_u^{-}$. Let $\Delta$ be the set of simple roots of $\Phi^+$ and $\Delta_{nc} = \{\alpha \in \Delta | \alpha(\check{\beta}) > 0\}$. We write $\beta = \sum_{\alpha\in\Delta} m_\alpha a$. The following result is a combination of Table 2.5 and Corollary 4.6 in [Gr-W].

**Lemma 4.** *If $G_c$ is of type $A_n$ then $\Delta_{nc}$ consists of two elements. Otherwise $\Delta_{nc} = \{\alpha_o\}$ (i.e. consists of one element) and in all cases $\sum_{\alpha\in\Delta_{nc}} m_\alpha = 2$.*

The next result is a critical step in our analysis of principal series embeddings and generalized Whittaker models for quaternionic discrete series. We will denote by $\langle\ ,\ \rangle$ the Weyl group invariant form induced on $\mathfrak{h}$ by a multiple of the Killing form such that $\langle\beta,\beta\rangle = 2$. If $\alpha \in \Phi$ then we use the notation $(\mathfrak{g}_C)_\alpha$ for the corresponding root space.

**Proposition 1.** *Assume that $G_C$ is not of type $A_n$. If $\langle\beta,\beta\rangle = \langle\alpha_o,\alpha_o\rangle = 2$ then $\Phi_o = \{\pm\alpha_o, \pm(\beta-\alpha_o), \pm\beta\}$ is a root subsystem of $\Phi$ of type $A_2$ and if $\mathfrak{u}_C = \mathbb{C}\check{\alpha}_o + \mathbb{C}\check{\beta} + \sum_{\alpha\in\Phi_o}(\mathfrak{g}_C)_\alpha$ then $\mathfrak{u}_C \cap \mathfrak{g}$ is a real form of $\mathfrak{u}_C$ isomorphic with $\mathfrak{su}(2,1)$. If $G_C$ is of type $A_n$ then there are 2 non-compact simple roots $\alpha_1, \alpha_2$ and we have $\Phi_i = \{\pm\alpha_i, \pm(\beta-\alpha_i), \pm\beta\}$ is a root system of type $A_2$ for $i = 1, 2$ and for either $i$ $\mathfrak{u}_C = \mathbb{C}\check{\alpha}_o + \mathbb{C}\check{\beta} + \sum_{\alpha\in\Phi_i}(\mathfrak{g}_C)_\alpha$ then $\mathfrak{u}_C \cap \mathfrak{g}$ is a real form of $\mathfrak{u}_C$ isomorphic with $\mathfrak{su}(2,1)$.*

*Proof.* We first consider the case when $G_C$ is not of type $A_n$. We note that with our normalization $\langle\beta,\alpha_o\rangle = 1$. So if we do the algebra we find that $\langle\beta-\alpha_o,\beta-\alpha_o\rangle = 2$ and $\langle\alpha_o,\beta-\alpha_o\rangle = -1$. This implies that $\beta - 2\alpha_o$ is not a root. Hence $\Phi_o$ is indeed a root subsystem of type $A_2$. The real form $\mathfrak{g}\cap\mathfrak{u}_C$ is clearly the quaternionic real form of $A_2$ (i.e. is constructed as above using $\mathfrak{u}_C \cap \mathfrak{h}$, $\Phi_o^+ = \{\alpha_o, (\beta-\alpha_o), \beta\}$) which is isomorphic with $\mathfrak{su}(2,1)$. The case of $A_n$ is even easier. $\square$

**Note.** If we look at Table 2.5 in [Gr-W] then we see that if $G_C$ is not of type $C_n$ then the hypothesis of the Proposition are satisfied. Hence there is a "canonical" $\mathfrak{su}(2,1)$ contained in $\mathfrak{g}$. The exceptional cases corresponds to the real Lie groups $Sp(n-1,1)$, $n \geq 2$.

For the rest of this section we will assume that $G_C$ is not of type $C_n$ and that if $G_C$ is of type $A_n$ then we have chosen $\alpha_o$ to be one of the two non-compact simple roots. Thus we have made a choice of a $\mathfrak{u} \cong \mathfrak{su}(2,1)$ in $\mathfrak{g}$ for each case. As in [Gr-W] we set $2d = |\{\alpha \in \Phi^+ | \beta - \alpha \in \Phi\}|$.

**Proposition 2.** *As a $\mathfrak{u}$-module $\mathfrak{g}_C = \mathfrak{u}_C \oplus \{x \in \mathfrak{g}_C | [x,\mathfrak{u}] = 0\} \oplus V \oplus V^*$ with $V$ equivalent as a $\mathfrak{u}$ ($\cong \mathfrak{su}(2,1)$)-module with $(d-1)$-copies of the standard 3 dimensional irreducible module.*

*Proof.* We note that $\check{\beta} \in \mathfrak{u}_C$ and that the eigenvalues of $\mathrm{ad}(\check{\beta})$ on $\mathfrak{g}_C$ are $0, \pm1, \pm2$ with the $\pm2$ eigenspaces each 1 dimensional and contained in $\mathfrak{u}_C$. Thus on $\mathfrak{g}_C/\mathfrak{u}_C$, $\mathrm{ad}(\check{\beta})$ has eigenvalues $\pm1$ and 0. In the notation of the previous section $\check{\beta}$ corresponds to the element $h \in \mathfrak{su}(2,1)$. If we use the system of positive roots for $\mathfrak{u}_C \cap \mathfrak{h}$ such that $\alpha_o, \beta-\alpha_o$ are the simple roots and if $\Lambda_1, \Lambda_2$ are the corresponding basic highest weights then if $\Lambda = m_1\Lambda_1 + m_2\Lambda_2$ then we have $\Lambda(\check{\beta}) = m_1 + m_2$. Hence if $\Lambda$ were a highest weight of a non-trivial irreducible representation occurring in $\mathfrak{g}_C/\mathfrak{u}_C$ then we must have $m_1 + m_2 = 1$. Hence $m_1 = 1$ and $m_2 = 0$ or vice-versa. Now if $F$ is the finite dimensional irreducible representation with highest weight $\Lambda_1$ then $F$ is the 3-dimensional standard representation and $F^*$

has highest weight $\Lambda_2$. Since the Killing form of $\mathfrak{g}_C$ is non-degenerate on $\mathfrak{u}_C$ we see that the representation of $\mathfrak{u}_C$ on $\mathfrak{g}_C/\mathfrak{u}_C$ is self-dual. Thus the multiplicity of $F$ must be the same as that of $F^*$. It is also clear that multiplicity must be $\frac{1}{2}$ the dimension of the eigenspace for 1 for $\mathrm{ad}(\check{\beta})$. Since the eigenspace for 1 in $\mathfrak{u}_C$ is 2-dimensional we see that the multiplicity of $F$ in $\mathfrak{g}_C/\mathfrak{u}_C$ is $\frac{2d-2}{2} = d-1$ as asserted. The multiplicity of the trivial one dimensional representation of $\mathfrak{u}_C$ in $\mathfrak{g}_C/\mathfrak{u}_C$ is equal to $\dim\{x \in \mathfrak{g}_C | [x, \mathfrak{u}] = 0\}$ since $\mathfrak{u}_C$ is reductive. $\square$

We note that in particular, the result implies that if $U$ is the connected subgroup of $G$ such that $Lie(U) = \mathfrak{u}$ then $U$ is isomorphic with $SU(2,1)$. Let $U_C$ be the connected subgroup of $G_C$ with $Lie(U_C) \cong \mathfrak{u}_C$. Then $U_C$ is isomorphic with $SL(3,\mathbb{C})$.

In the previous section we began with a maximally split Cartan subalgebra of $Lie(SU(2,1))$ and derived from it a maximal torus of a maximal compact subgroup. Here we will trace our steps in the opposite direction. If we identify $\mathfrak{u}_C \cap (\mathfrak{h} \bigoplus \bigoplus_{\alpha\in\Phi^+} \mathfrak{g}_\alpha)$ with the theta stable Borel subalgebra, $\mathfrak{b}$, of $Lie(SU(2,1))_C$ of the previous section then we may retrace the steps in its construction finding elements $X, Y, H \in \mathfrak{u}$ such that $[X,Y] = H$, $[H,X] = 2X$, $[H,Y] = -2Y$, $\theta H = -H$, $\theta X = -Y$. Further, $adH$ has eigenvalues $-2,-1,0,1,2$ and $\{x \in \mathfrak{u} | [H,x] = 2x\} = \mathbb{R}X$. Indeed, $H$ is the coroot of the largest root relative to a Borel subalgebra of $\mathfrak{u}$ that is the complexification of a minimal parabolic subalgebra of $\mathfrak{u}$. So these assertions follow from the previous proposition.

We set $\mathfrak{m} = \{x \in \mathfrak{g} | ad(H)x = 0\}$ and $\mathfrak{n}_1 = \{x \in \mathfrak{g} | ad(H)x = x\}$, $\mathfrak{n}_2 = \{x \in \mathfrak{g} | ad(H)x = 2x\}$. Put $\mathfrak{n} = \mathfrak{n}_1 + \mathfrak{n}_2$. We set $\mathfrak{p} = \mathfrak{m} + \mathfrak{n}$ and note that $\mathfrak{p}$ is a parabolic subalgebra of $\mathfrak{g}$. We also note that $\mathfrak{p} \cap \mathfrak{u}$ is a minimal parabolic subalgebra of $\mathfrak{u}$. We set $P = \{g \in G | \mathrm{Ad}(g)\mathfrak{p} \subset \mathfrak{p}$ and $N = \exp \mathfrak{n}$ the nilradical of $P$.

**Lemma 5.** *We can write $\mathfrak{n} = \mathfrak{n} \cap \mathfrak{u} \oplus W$ such that relative to the decomposition in the preceding Proposition, $W \subset (V \oplus V^*)$.*

*Proof.* We note that if $z, w \in \mathfrak{n}_1$ then $[z,w] = \omega(z,w)X$ with $\omega$ a symplectic structure on $\mathfrak{n}_1$. Let $W = \{w \in \mathfrak{n}_1 | \omega(w, \mathfrak{u} \cap \mathfrak{n}_1) = 0\}$. Then since $\omega$ is nondegenerate when restricted to $u \cap n_1$ we have $\mathfrak{n}_1 = \mathfrak{u} \cap \mathfrak{n}_1 \oplus W$. Now $\mathfrak{n}_2 = \mathbb{R}X \subset \mathfrak{u}$. Thus $\mathfrak{n} = \mathfrak{u} \cap \mathfrak{n} \oplus W$. The definition of $\omega$ also implies that $[\mathfrak{u} \cap \mathfrak{n}, W] = 0$. We therefore see that the $\mathfrak{u}_C$ cyclic space (under $ad$) of a non-zero vector in $W$ is a sum of irreducible subspaces that have highest weight either $\Lambda_1$ or $\Lambda_2$. Thus $W \subset (V \oplus V^*)$ as asserted. $\square$

We set $M_C = \{g \in G_C | \mathrm{Ad}(g)H = H\}$. We put $M = M_C \cap G$. Then $P = MN$. We will determine the set of open orbits of the action of $M$ on $\mathfrak{n}_1$. Since $\theta(M) = M$, these results will also determine the orbit structure of $M$ acting on $\theta\mathfrak{n}_1$. We will say that $x \in \mathfrak{n}_1$ is *generic* if $\mathrm{Ad}(M)x$ is open in $\mathfrak{n}_1$. We will use the notation $Lie(M)$ and $\mathfrak{m}$ interchangeably.

In [Gr-W] we proved that the ring of polynomial semi-invariants for $\mathrm{Ad}(M_C)$ acting on $(\mathfrak{n}_1)_C$ is $\mathbb{C}[u]$ with $u$ homogeneous of degree 4 and that $\mathrm{Ad}(M_C)$ has a unique Zariski open orbit in $(\mathfrak{n}_1)_C$ given by the subset of $x$ such that $u(x) \neq 0$.

**Lemma 6.** *Suppose that $x \in \mathfrak{n}_1$ and there exists $y \in \theta\mathfrak{n}_1$ such that $[x, y] = 2H$. Then $x$ is generic.*

*Proof.* We note that $\{x, y, 2H\}$ forms a standard basis of a Lie algebra isomorphic with $\mathfrak{sl}(2,\mathbb{R})$. If we apply the representation theory of $\mathfrak{sl}(2,\mathbb{R})$ we have that the eigenspace for $2ad(H)$ with eigenvalue 2 is equal to the image under $ad(x)$ of the 0 eigenspace. The 0 eigenspace is $Lie(M)$ and the eigenspace for eigenvalue 2 is $\mathfrak{n}_1$. Thus $[Lie(M), x] = ad(x)Lie(M) = \mathfrak{n}_1$. Hence $\mathrm{Ad}(M)x$ is open. □

We now observe that if $x \in \mathfrak{u} \cap \mathfrak{n}_1$ is non-zero then there exists $y \in \mathfrak{u} \cap \theta\mathfrak{n}_1$ such that $[x, y] = 2H$. This every non-zero element of $\mathfrak{u} \cap \mathfrak{n}_1$ is generic. If $x \in \mathfrak{n}_1$ then we note that $ad(x)^4 Y$ is a multiple of $X$. If $x$ is a non-zero element of $\mathfrak{u} \cap \mathfrak{n}_1$ then this multiple is non-zero. We define a polynomial of degree 4 on $(\mathfrak{n}_1)_C$ by $ad(x)^4 Y = f(x)X$. Then $f$ is real valued on $\mathfrak{n}_1$ and non-zero. Furthermore, if we changed the choice of $X, Y$ (maintaining the commutation relation $[X, Y] = H$) we would only change $f$ by a positive real number. We note that $f$ is a non-zero multiple of $u$ defined above.

**Proposition 3.** *Let $x \in \mathfrak{n}_1$ then the following are equivalent:*

1) *$x$ is generic.*
2) *There exists $y \in \theta\mathfrak{n}_1$ such that $[x, y] = 2H$.*
3) *$f(x) \neq 0$.*

*Furthermore, under the condition 1) $ad(x)$ is injective on $\theta\mathfrak{n}$. In particular the element $y$ in 2) is uniquely determined by $x$.*

*Proof.* Suppose that $x$ is generic. Then $[Lie(M), x] = \mathfrak{n}_1$. Thus if we complexify we see that $\mathrm{Ad}(M_C)x$ is open in $(\mathfrak{n}_1)_C$. Hence there exists $m \in M_C$ such that $\mathrm{Ad}(m)x$ is in $\mathfrak{u} \cap \mathfrak{n}_1$. Thus the observations above imply that there exists $u \in (\theta\mathfrak{n}_1)_C$ such that $[x, u] = 2H$. Now $u = y + iw$ with $y, w \in \theta\mathfrak{n}_1$. So $[x, u] = [x, y] + i[x, w]$. Since $H, [x, y], [x, w] \in \mathfrak{g}$ we see that $[x, w] = 0$ and $[x, y] = 2H$. This proves that 1) implies 2). That 2) implies 1) is the content of the previous lemma.

We will now prove the equivalence of the third condition. We note that if $m \in M_C$ then $f(\mathrm{Ad}(m)z) = \chi(m)^2 f(z)$ with $\chi(m)$ defined by $\mathrm{Ad}(m)X = \chi(m)X$. If we apply the argument above to $x$ generic then we find that there exists $m \in M_C$ such that $f(\mathrm{Ad}(m)x) \neq 0$. Thus $f(x) \neq 0$. Thus 1) implies 3). To prove that 3) implies 2) it is enough to observe (as we have above) that $\{z \in (\mathfrak{n}_1)_C | f(z) \neq 0\}$ consists of a single $M_C$ orbit.

We will now prove the last assertion. We note that $\theta\mathfrak{n}$ is the direct sum of the $-2$ and the $-4$ eigenspace for $ad(2H)$. Thus the representation theory of $\mathfrak{sl}(2,\mathbb{R})$ implies that $ad(x)$ is injective on $\theta\mathfrak{n}$. If $[x, y] = [x, v] = 2H$. Then $[x, y - v] = 0$.

So if $v \in \theta\mathfrak{n}_1$, $y - v = 0$. $\square$

**Proposition 4.** *Let $x \in \mathfrak{n}_1$ be generic. Let $y$ denote the element of $\theta\mathfrak{n}_1$ as in 2) in the proposition above. Then the Lie algebra generated by $\{x, y, X, Y\}$ is either isomorphic with $\mathfrak{sl}(3,\mathbb{R})$ or $\mathfrak{su}(2,1)$. Furthermore, it is isomorphic with $\mathfrak{sl}(3,\mathbb{R})$ if $f(x) > 0$ and isomorphic with $\mathfrak{su}(2,1)$ if $f(x) < 0$.*

*Proof.* Let $\mathfrak{h}$ be the Lie algebra described in the statement. Then if we follow the proof of the previous corollary we and make the same identifications see that there exists $m \in M_C$ such that $\mathrm{Ad}(m)\mathfrak{h}_C$ is isomorphic with the Lie algebra

$$\mathbb{C}H_{\alpha_o} \bigoplus \mathbb{C}H_\beta \bigoplus \mathfrak{g}_{\alpha_o} \bigoplus \mathfrak{g}_{\beta-\alpha_o} \bigoplus \mathfrak{g}_{-\alpha_o} \bigoplus \mathfrak{g}_{\alpha_o-\beta} \bigoplus \mathfrak{g}_\beta \bigoplus \mathfrak{g}_{-\beta}.$$

This Lie algebra is of type $A_2$. Thus the Lie algebra that we are studying is a real form of $A_2$. Since it contains nilpotent elements it must be either isomorphic with $\mathfrak{sl}(3,\mathbb{R})$ or $\mathfrak{su}(2,1)$.

To prove the last assertion of the proposition we will give another formula for $f(x)$. Let $B$ denote the Killing form of $\mathfrak{g}$. We note that the element $ad(x)^2Y \in Lie(M) \cap \mathfrak{h}$ and $B(ad(x)^2Y, H) = 0$. Since the Lie algebras $\mathfrak{sl}(3,\mathbb{R})$ and $\mathfrak{su}(2,1)$ are respectively split or quasi-split over $\mathbb{R}$ and since $\mathfrak{p} \cap \mathfrak{h}$ is a minimal parabolic subalgebra of $\mathfrak{h}$. We see that the element $ad(x)^2Y$ is semi-simple and has real eigenvalues in the case of $\mathfrak{sl}(3,\mathbb{R})$ and purely imaginary eigenvalues in the case of $\mathfrak{su}(2,1)$. Let $g(z) = B(ad(z)^2Y, ad(z)^2Y) = tr(ad(ad(z)^2Y)^2)$ for $z \in (\mathfrak{n}_1)_C$. We note that $g(z) = B(ad(z)^4Y, Y) = f(z)B(X,Y)$. Since $[X,Y] = H$ this says that $B(H,H) = B([X,Y],H) = B(X,Y)$. Thus $g(z) = B(H,H)f(z)$, a strictly positive multiple of $f$. Since $g(x) \geq 0$ in the case of $\mathfrak{sl}(3,\mathbb{R})$ and $g(x) \leq 0$ in the case of $\mathfrak{su}(2,1)$ the proof is complete. $\square$

Our final goal of this section is to give a description of the open orbits of $\mathrm{Ad}(M)$ in $\mathfrak{n}_1$ to the extent that will be necessary for our analysis of Heisenberg–Whittaker models (however Lemma 7 is not directly related to this goal but it uses some of the ingredients in our analysis and might be useful in other contexts). As in [Gr-W] we will see that there is a uniformity in the answer only for the quaternionic real forms of real rank 4. If $x \in \mathfrak{n}_1$ is generic and if the Lie algebra generated by $\{x, y, X, Y\}$ (as above) is isomorphic with $\mathfrak{su}(2,1)$ then we will say that $x$ is of *complex type.* Note that if $x \in \mathfrak{u} \cap \mathfrak{n}_1$ is non-zero then $x$ is generic and of complex type by the results above.

**Proposition 5.** *Let $x \in \mathfrak{n}_1$ be of complex type. If $\mathrm{Ad}(M)x \cap \mathfrak{u} = \emptyset$ then there exists $m \in M$ such that $B(\mathrm{Ad}(m)y, \mathfrak{u} \cap \mathfrak{n}_1) = \{0\}$.*

*Proof.* Assume that $x \in \mathfrak{n}_1$ is of complex type. Let $\mathfrak{v}$ denote the Lie algebra generated by $\{x, y, X, Y\}$. Set $j = ad(x)^2Y$ then $B(j, H) = 0$ and $\{j, H\}$ span a Cartan subalgebra of $\mathfrak{v}$. We note that $\mathfrak{v}_C$ is isomorphic with $\mathfrak{sl}(3,\mathbb{C})$. Choosing the positive roots of $\mathfrak{v}_C$ to be compatible with $\mathfrak{n} \cap \mathfrak{v}$ the argument given in Proposition 2 implies that $\mathfrak{g}_C$ decomposes into the direct sum of the adjoint representation, the centralizer of $\mathfrak{v}_C$ and the direct sum of $d-1$ copies of the standard 3 dimensional representation direct sum with its dual. This implies that if we multiply $x$ by a positive real number (which doesn't change its $M$-orbit) we may assume that the eigenvalues of $ad(j)$ are given as follows: in $\mathfrak{m}_C$ they are $0, \pm 2i$ in $(\mathfrak{n}_1)_C$ they are $\pm 3i$ each with multiplicity 1 and $\pm i$ each of multiplicity $d-1$. Furthermore, $\mathfrak{n}_1 \cap \mathfrak{v} = \{z \in \mathfrak{n}_1 | ad(j)^2 z = -9z\}$.

Now let $x_o \in \mathfrak{n}_1 \cap \mathfrak{u}$ be non-zero and let $j_o$ be the element constructed as above for $x_o$. The conjugacy of maximal compact tori in $M$ implies that there exists $m \in M$ such that $\mathrm{Ad}(m)j$ commutes with $j_o$. Replace $j$ by this conjugate. Now $ad(j)(\mathfrak{n}_1 \cap \mathfrak{u}) \subset \mathfrak{n}_1 \cap \mathfrak{u}$, since $\mathfrak{n}_1 \cap \mathfrak{u}$ is the full eigenspace for $ad(j_o)^2$ for eigenvalue $-9$ in $\mathfrak{n}_1$. Thus $ad(j)^2$ has spectrum contained in $\{-1, -9\}$ on this space. If $ad(j)^2$ has one eigenvalue $\mu$ on $\mathfrak{n}_1 \cap \mathfrak{u}$ then there exists $z \in \mathfrak{n}_1 \cap \mathfrak{u}$ such that $z \neq 0$ and $ad(j)^2 z = \mu z$. But $\mathfrak{n}_1 \cap \mathfrak{u}$ is the real span of $z$ and $ad(j_o)z$ thus since $[j, j_o] = 0$ we see that there are two possibilities $ad(j)^2$ is $-1$ on $\mathfrak{n}_1 \cap \mathfrak{u}$ or $ad(j)^2$ is $-9$ on $\mathfrak{n}_1 \cap \mathfrak{u}$. If the latter is true that $\mathfrak{n}_1 \cap \mathfrak{u} = \mathfrak{n}_1 \cap \mathfrak{v}$ and hence $x$ is conjugate to $x_o$. Otherwise, since, $ad(j)^2 y = -9y$ ($\theta\mathfrak{n}_1 \cap \mathfrak{v}$ is the span of $ad(x)Y$ and $ad(j)ad(x)Y$) we see that if $z \in \mathfrak{n}_1 \cap \mathfrak{u}$ then

$$-9B(y,z) = B(ad(j)^2 y, z) = B(y, ad(j)^2 z) = -B(y,z).$$

this completes the proof. □

Let $x \in \mathfrak{n}_1$ and $y \in \theta\mathfrak{n}_1$ be such that $[x, y] = 2H$. Let $\xi : \mathfrak{sl}(2,\mathbb{R}) \to \mathfrak{g}$ be defined by

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \longmapsto x, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \longmapsto y, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \longmapsto 2H.$$

This exponentiates to a group homomorphism (also denoted $\xi$) of $SL(2,\mathbb{R})$ into $G$. Let

$$s = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Define $g_o = \xi(s)$. If we set $h = 2H$. Then the eigenvalues of $ad(h)$ on $\mathfrak{g}$ are 2 with multiplicity $2d$, $-2$ with the same multiplicity 0 with multiplicity $2d + \dim\{u \in \mathfrak{m} | [u, x] = 0\}$, 4 and $-4$ each of multiplicity 1. Thus as a representation of $SL(2,\mathbb{R})$ the composition $\mathrm{Ad} \circ \xi$ splits into $F^4 \bigoplus (2d-1)F^2 \bigoplus (\dim \mathfrak{g} - 6d - 2)F^0$. Here $F^k$ is a representative of the class of the irreducible $k+1$ dimensional representation of $SL(2,\mathbb{R})$. Now $s$ acts by $-1$ on the 0 weight space of $F^2$ and by 1 on the 0 weight space of $F^4$. We therefore have

**Lemma 7.** *Let $\sigma = \mathrm{Ad}(g_o)$ then $\sigma$ is an involutive automorphism of $\mathfrak{m}$ such that $\mathfrak{m}^\sigma = \{z \in \mathfrak{m} | \sigma(z) = z\} = \mathbb{R}ad(y)^2 X \bigoplus \mathfrak{m}^x$ ($\mathfrak{m}^x = \{z \in \mathfrak{m} | [x, z] = 0\}$). Furthermore, $\mathbb{R}ad(y)^2 X = \mathbb{R}ad(x)^2 Y$.*

*Proof.* Since the 0 weight space for $h$ is $\mathfrak{m}$. It is clear, from the above, that $\sigma$ is involutive on $\mathfrak{m}$. Also the 1 eigenspace for $\sigma$ acting on $\mathfrak{m}$ is the direct sum of the zero weight space in $F^4$ and the isotypic space of $F^0$. This latter space is easily seen to be $\mathfrak{m}^x$. Since $X$ (resp. $Y$) is a highest (resp. lowest) weight vector for the component of $\mathfrak{g}$ corresponding to $F^4$, the observations above now complete the proof. □

We will now begin our discussion of the orbits. We will first study the low real rank cases. Suppose that $G$ is of type $A_{n+1}$ with $n \geq 1$. Then $G$ is locally isomorphic with $SU(2,n)$ and the action of $M$ on $\mathfrak{n}_1$ is that of $\mathbb{C}^\times \times SU(1,n-1)$ on $\mathbb{C}^n$ looked upon as a real vector space. Also we have action $(z,u)x = z^2ux$. Let $(\ldots,\ldots)$ denote a $SU(1,n-1)$-invariant Hermitian form of signature $(1,n-1)$. Then one can check that up to positive scalar multiple we have $f(z) = -(z,z)^2$. Thus the generic elements are exactly the non-isotropic vectors. If $n = 1$ it is clear that there is exactly one open orbit (the non-zero vectors). In the case when $n > 1$ there are 2 open orbits $(z,z) > 0$ and $(z,z) < 0$.

The next case we will study is that of BD. Here $G$ is locally isomorphic with the identity component of $SO(4,n)$ with $n \geq 3$. Here $M$ acts on $V = \mathfrak{n}_1$ by the action of $GL(2) \times SO(2,n-2)$ on $\mathbb{R}^2 \bigotimes \mathbb{R}^n$ with the $\mathbb{R}^n$ factor endowed with a non-degenerate symmetric bilinear form $\langle\ ,\ \rangle$ of signature $(2,n-2)$. Let $e_1, e_2$ be the usual basis of $\mathbb{R}^2$. If $v \in V$ then we can write

$$v = e_1 \bigotimes v_1 + e_2 \bigotimes v_2$$

we define $h(v) = \det[\langle v_i, v_j\rangle]$. The generic elements are precisely those vectors $v$ such that $h(v) \neq 0$. Set $F(v)$ equal to the span over $\mathbb{R}$ of $\{v_1, v_2\}$. If $n = 3$, and $h(v) \neq 0$, then there are 2 possible signatures for the form restricted to $F(v)$ : $(2,0)$ and $(1,1)$. If $n \geq 4$ then there are 3 possibilities $(2,0),(1,1),(0,-2)$. Using these observations one can see that in the case of $SO(4,3)$ there are two open orbits and for $SO(4,n)$ with $n \geq 4$ there are 3. A direct calculation shows that up to a positive multiple $f(v) = -h(v)$ (yet another formula for $f$).

We note that at this point we have with relative ease completely dealt with all of the quaternionic real forms of classical type. After we show that the case of $G_2$ is essentially the same as $SO(4,3)$.we will handle the remaining quaternionic real forms (all of real rank 4) by reducing to the case of $SO(4,4)$

We now look at the case of $G_2$. If we consider the imbedding of $G_2$ into $SO(4,3)$ compatible with our choice of Heisenberg parabolic subgroup over $\mathbb{R}$ then each $G_2$ open orbit intersects in an $SO(4,3)$ open orbit. Indeed, the action of the identity component of $M$ for $G_2$ is given through the action of $\mathbb{R}^\times \times SL(2,\mathbb{R})$ on $S^3(\mathbb{R}^2)$ (third symmetric power). This can be made consistent with the "picture" for $SO(4,3)$ by observing that the map $g \mapsto g \bigotimes S^2(g)$ acting on $\mathbb{R}^2 \bigotimes S^2(\mathbb{R}^2)$ factors through $SL(2,\mathbb{R}) \times SO(1,2)$ and then using multiplication in the symmetric algebra to map to the space $S^3(\mathbb{R}^2)$. On checks that under this mapping the orbits of the $M$ for $SO(4,3)$ map onto those for $G_2$. Hence the result for $G_2$ corresponds that for $SO(4,3)$. We are left with the $\mathbb{R}$ rank 4 quaternionic real forms of the exceptional groups.

We will now study the orbit structure on the subset of $\mathfrak{n}_1$ consisting of those $z$ with $f(z) > 0$. To do this analysis we fix a $\theta$-equivariant imbedding of $Lie(SO(4,4))$ into $Lie(G)$ with the property that $Lie(U)$ is a subalgebra of $Lie(SO(4,4))$ constructed as above and $Lie(SO(4,4)$ contains the Lie algebra, $\mathfrak{a}$, of a four dimensional $\mathbb{R}$-split (algebraic) subtorus of $M$. Let $x \in \mathfrak{n}_1$ be such that $f(x) > 0$ then $ad(j)$ has real eigenvalues. The conjugacy of $\mathbb{R}$-split tori implies that there exists $m \in M$ such that $\mathrm{Ad}(m)j \in \mathfrak{a}$. Thus we see that we can conjugate $x$ into the image of $Lie(SO(4,4))$ in the imbedding described as above. As in the study above of the elements of complex type we may replace $x$ by a positive real multiple such that the eigenspace for $(adj)^2$ with eigenvalue 9 in $\mathfrak{v}$ is $\mathfrak{v} \cap \mathfrak{n}_1$. Thus the 2 dimensional eigenspace with eigenvalue 9 for $\mathrm{Ad}(\mathrm{Ad}(m)j)$ in $\mathfrak{n}_1$ is contained in the image of $Lie(SO(4,4))$. Thus $\mathrm{Ad}(m)x$ is contained in this image. The result for $SO(4,4)$ now implies that where $f$ is positive consists of exactly one orbit. In the case when $f(x) < 0$ that is $h(v) > 0$ Proposition 5 implies that if the orbit of $x$ is not equal to $\mathrm{Ad}(M)(\mathfrak{u} \cap \mathfrak{n}_1 - \{0\})$ then there exists $m \in M$ such that $B(\mathrm{Ad}(m)y, \mathfrak{u} \cap \mathfrak{n}_1) = \{0\}$. The same conclusion can be made from calculations in $SO(4,4)$ to see that if $f(x) > 0$ then there exists $m \in M$ such that $B(\mathrm{Ad}(m)y, \mathfrak{u} \cap \mathfrak{n}_1) = \{0\}$. We have proved.

**Proposition 6.** *Assume that $G$ is of real rank 4. Assume that $x$ is generic and $x \notin \mathrm{Ad}(M)(\mathfrak{u} \cap \mathfrak{n}_1 - \{0\})$. Then there exists $m \in M$ such that $B(\mathrm{Ad}(m)y, \mathfrak{u} \cap \mathfrak{n}_1) = \{0\}$.*

## 6. Embeddings of quaternionic discrete series into degenerate principal series

We retain the notation of the previous section. We will also assume that the condition of section 5 is satisfied (i.e. that $G_C$ is not of type $C_n$). The quaternionic discrete series of $G$ is the family of discrete series associated with the system of positive roots $\Phi^+$ (see [Gr-W; section 5]). If $\Lambda \in \mathfrak{h}^*$ is $\Phi^+$-dominant integral then we will write $(\Pi_\Lambda, L^\Lambda)$ for a representative of the corresponding equivalence class of quaternionic discrete series representations. We will first study the case when $\Lambda = \Lambda_k = \rho + (k - d - 1)\beta$ with $k \geq d + 1$. (This weird parametrization is because then $s_\beta \Lambda = \rho - k\beta$. Here as usual, $s_\beta$ is the root reflection about $\beta$.) We will denote this representation as $(\Pi_k, L^k)$.

We recall that $Lie(K) = Lie(K_1) \oplus Lie(K_2)$ with $K_1 \cong SU(2)$. The representations $\Pi_\Lambda$ are admissible as $K_1$-modules. Let $U \subset G$ be the connected subgroup with $Lie(U) = \mathfrak{u}$ as in Proposition 1 in the previous section. The results in that section imply that $U \cong SU(2,1)$ and that $K_1 \subset K$. This implies that as a representation of $U$ the restriction of $\Pi_\Lambda$ splits into a direct sum of irreducible representations with finite multiplicities. We now describe the constituents in this decomposition explicitly. We first need a bit of notation. Let $\mathfrak{q}$ denote the $\theta$-stable parabolic subalgebra of $\mathfrak{g}_C$ given by

$$\{x \in \mathfrak{g}_C | [\check{\beta}, x] = 0\} \bigoplus \{x \in \mathfrak{g}_C | [\check{\beta}, x] = x\} \bigoplus \{x \in \mathfrak{g}_C | [\check{\beta}, x] = 2x\}.$$

We set $V = \{x \in \mathfrak{g}_C | [\check{\beta}, x] = x\}$. Then $V$ is a $K_2$ module under Ad. This notation will only be used in the proof of the following result.

**Theorem 8.** *Assume that $G \neq U$. The representation $(\Pi_{k|U}, L^k)$ splits into the orthogonal direct sum*

$$\bigoplus_{p,q\geq 0} \binom{p+d-2}{d-2}\binom{q+d-2}{d-2} H^{(k-1+p)\Lambda_1+(k-1+q)\Lambda_2}.$$

*Proof.* (The notation in this proof will only be used here. It is consistent with [Gr-W] but not with the previous section.) As a representation of $K_1 \times K_2$ the discrete series $L^k$ splits into the direct sum:

$$\bigoplus_{n\geq 0} S^{2k-2+n}(\mathbb{C}^2) \bigotimes S^n(V)$$

as in [Gr-W] Proposition 5.7. Here $K_1 \cong SU(2)$ with $K_2$ acting trivially and $V$ is a $K_2$ module with $K_1$ acting trivially. Now $U \cap K_2$ is the center of $U \cap K$. We find that as a $U \cap K_2$-module $V = V \cap Lie(U)_C \bigoplus V'$. With $V'$ contained in the orthogonal complement of $Lie(U)$. We write $W = V \cap Lie(U)_C$. Now $U \cap K_2$ is a circle group with character group $\mathbb{Z}\frac{\Lambda_1-\Lambda_2}{2}$. We note that as a $U \cap K_2$-module we have (the subscripts indicate $U \cap K_2$-isotypic components)

$$W = W_{3(\frac{\Lambda_1-\Lambda_2}{2})} \bigoplus W_{-3(\frac{\Lambda_1-\Lambda_2}{2})}, V' = V'_{\frac{\Lambda_1-\Lambda_2}{2}} \bigoplus V'_{-\frac{\Lambda_1-\Lambda_2}{2}}.$$

Here $\dim W_{\pm 3(\frac{\Lambda_1-\Lambda_2}{2})} = 1$ and $\dim V'_{\pm\frac{\Lambda_1-\Lambda_2}{2}} = d-1 > 0$. Thus

$$S^n(V) = \bigoplus_{p+q+k+l=n} S^k(W_{3(\frac{\Lambda_1-\Lambda_2}{2})}) \bigotimes S^l(W_{-3(\frac{\Lambda_1-\Lambda_2}{2})}) \bigotimes S^p(V'_{\frac{\Lambda_1-\Lambda_2}{2}}) \bigotimes S^q(V'_{-\frac{\Lambda_1-\Lambda_2}{2}}).$$

We note that $S^k(W_{\pm 3(\frac{\Lambda_1-\Lambda_2}{2})})$ is a one dimensional $U \cap K_2$ module on which $U \cap K_2$ acts by the character $\pm 3k(\frac{\Lambda_1-\Lambda_2}{2})$. We denote this one dimensional module as $\mathbb{C}_{\pm 3k(\frac{\Lambda_1-\Lambda_2}{2})}$.

The $K$-decomposition described above implies that as a $U \cap K$ module we have

$$H^{k\Lambda_1+l\Lambda_2} = \bigoplus_{n\geq 0} S^{k+l+n}(C^2) \bigotimes \mathbb{C}_{(k-l)(\frac{\Lambda_1-\Lambda_2}{2})} \bigotimes S^n(W).$$

Using this it is an easy matter to see that the decomposition of $\Pi_k$ as a $K \cap U$ module is exactly as described in the assertion of the theorem. Since the $K$-character determines a discrete series representation the result follows. $\square$

We will now use the preceding result to describe an imbedding of $L^k$ into spherical (degenerate) principal series induced from $P$. As usual, if $(\pi, H)$ is an admissible representation of $G$ and if $K'$ is a compact subgroup of $G$ then $H_{K'}$ will denote the $K'$-finite $C^\infty$ vectors of $H$. If $(\sigma, V)$ is a finite dimensional representation of $P$ then we will use the notation $I(\sigma)$ for the representation induced (non-unitarily) from $P$ to $G$. Define $\chi : P \to \mathbb{R}^\times$ by $\mathrm{Ad}(p)X = \chi(p)X$, $p \in P$.

**Theorem 9.** *Let $\mathfrak{u}_1$ be a Lie subalgebra of $\mathfrak{g}$ isomorphic with $\mathfrak{su}(2,1)$ and containing $Lie(K_1)$ and $X, Y, H$. Let $k \geq d+1$ then there exists an injective $(\mathfrak{g}, K)$-module homomorphism of $L_K^k$ into $I(\chi^{k-1}|\chi|)_K$.*

*Proof.* In the course of this proof we will use the notation of the previous section. We observe that $\mathfrak{n} = \mathfrak{n} \cap \mathfrak{u} \bigoplus W$ as in Lemma 5 section 5. Since $[\mathfrak{n}, \mathfrak{n}] \subset \mathfrak{n} \cap \mathfrak{u}$ we see that $\mathfrak{n} \cap \mathfrak{u}$ is a normal subalgebra of $\mathfrak{n}$ and the projection of $W$ into $\mathfrak{n}/\mathfrak{n} \cap \mathfrak{u}$ is a bijective linear map onto an abelian Lie algebra. We will now use the previous theorem to analyze the $\mathfrak{m} \cap \mathfrak{u} \bigoplus \mathfrak{n}$-module $L_K^k/(\mathfrak{n} \cap \mathfrak{u})L_K^k$. We will think of this space as a $W_C$ module with $W$ identified with its image in $\mathfrak{n}/\mathfrak{n} \cap \mathfrak{u}$. We note that as an $\mathfrak{m} \cap \mathfrak{u}$-module $W_C = W_{\Lambda_1} \bigoplus W_{\Lambda_2}$ (see Lemma 5 section 5). Also $L_K^k = L_{K_1}^k$. We note that we are identifying characters $\Lambda$ and $\Lambda \circ \sigma$. Thus

$$L_K^k = \bigoplus_{p,q \geq 0} \binom{p+d-2}{d-2} \binom{q+d-2}{d-2} H_{K_1}^{(k-1+p)\Lambda_1 + (k-1+q)\Lambda_2}.$$

Thus if we set $c_{p,q} = \binom{p+d-2}{d-2}\binom{q+d-2}{d-2}$ then we have

$$\begin{aligned} &L_K^k/(\mathfrak{n} \cap \mathfrak{u})L_K^k \\ &= \bigoplus_{p,q \geq 0} c_{p,q} H_{K_1}^{(k-1+p)\Lambda_1+(k-1+q)\Lambda_2} / (\mathfrak{n} \cap \mathfrak{u}) H_{K_1}^{(k-1+p)\Lambda_1+(k-1+q)\Lambda_2} \end{aligned}$$

as a $\mathfrak{m} \cap \mathfrak{u} \bigoplus W$-module. As an $\mathfrak{m} \cap \mathfrak{u}$-module we have (taking into account the $\rho$ shifts)

$$\begin{aligned} &H_{K_1}^{(k-1+p)\Lambda_1+(k-1+q)\Lambda_2} / (\mathfrak{n} \cap \mathfrak{u}) H_{K_1}^{(k-1+p)\Lambda_1+(k-1+q)\Lambda_2} \\ &= \mathbb{C}_{(k+p)\Lambda_1+(k+q)\Lambda_2} \bigoplus \mathbb{C}_{(2k+p+q-1)\Lambda_1-(k+q-2)\Lambda_2} \bigoplus \mathbb{C}_{-(k+p-2)\Lambda_1+(2k+p+q-1)\Lambda_2}. \end{aligned}$$

by Theorem 6. The decomposition above of $W_C$ implies that

$$Z = \bigoplus_{p+q>0} c_{p,q} H_{K_1}^{(k-1+p)\Lambda_1+(k-1+q)\Lambda_2} / (\mathfrak{n} \cap \mathfrak{u}) H_{K_1}^{(k-1+p)\Lambda_1+(k-1+q)\Lambda_2}$$

is a $\mathfrak{m} \cap \mathfrak{u} \bigoplus W$-submodule of $L_K^k/(\mathfrak{n} \cap \mathfrak{u})L_K^k$. By comparing weights we see that $W$ acts trivially on $(L_K^k/(\mathfrak{n} \cap \mathfrak{u})L_K^k)/Z$. Let $\nu$ denote the natural projection of $L_K^k/(\mathfrak{n} \cap \mathfrak{u})L_K^k$ onto $L_K^k/\mathfrak{n}L_K^k$. We therefore have an $(\mathfrak{m} \cap \mathfrak{u}) \bigoplus \mathfrak{n}$-module homomorphism of $L_K^k$ onto $\mathbb{C}_{k\Lambda_1+k\Lambda_2}$ (with $\mathfrak{n} \cdot 1 = 0$). We now compute the $2k$-eigenspace for $H$ in $L_K^k/(\mathfrak{n} \cap \mathfrak{u})L_K^k$. We note that $H$ acts by $2k+p+q$ on $\mathbb{C}_{(k+p)\Lambda_1+(k+q)\Lambda_2}$ and by

$k+p+1$ on $\mathbb{C}_{(2k+p+q-1)\Lambda_1-(k+q-2)\Lambda_2}$ or $\mathbb{C}_{-(k+q-2)\Lambda_1+(2k+p+q-1)\Lambda_2}$. Thus the $2k$ eigenspace is

$$\mathbb{C}_{k(\Lambda_1+\Lambda_2)} \bigoplus \bigoplus_{p\geq 0} c_{k+1,p}(\mathbb{C}_{(3k+p)\Lambda_1-(k+p-2)\Lambda_2} \bigoplus \mathbb{C}_{-(k+p-2)\Lambda_1+(3k+p)\Lambda_2}).$$

We now note that we have as basis of $\mathfrak{m}\cap\mathfrak{u}$ the elements $H$ and $J$ with $\Lambda_1(J)=i$ and $\Lambda_2(J)=-i$. Thus the weights of $J$ on $\mathbb{C}_{(3k+p)\Lambda_1-(k+p-2)\Lambda_2}$ and $\mathbb{C}_{-(k+p-2)\Lambda_1+(3k+p)\Lambda_2}$ are respectively $4k+2p$ and $-(4k+2p)$. We recall that $k\geq d+1\geq 3$. We also note that the eigenvalues of $ad(J)$ on $\mathfrak{g}_C$ are $3i,-3i,i$, $-i,-2i$. Since the $2k$ eigenspace for $H$ on $L_K^k/\mathfrak{n}L_K^k$ is a $J$-module quotient of the $2k$ eigenspace of $L_K^k/(\mathfrak{n}\cap\mathfrak{u})L_K^k$ the above weight theoretic arguments imply that $[\mathfrak{m},\mathfrak{m}]\nu(\mathbb{C}_{k(\Lambda_1+\Lambda_2)})=0$. Thus $Lie(P)$ acts on $\nu(\mathbb{C}_{k(\Lambda_1+\Lambda_2)})$ by $kd\chi$. To complete the proof of the theorem we must show that if $m\in M$ then it acts on $\nu(\mathbb{C}_{k(\Lambda_1+\Lambda_2)})$ by $|\chi(m)|\chi(m)^{k-1}$.

If $m\in M\cap K$ then since $K_1$ is a normal subgroup of $K$ we see that $mK_1m^{-1}\subset K_1$. Also $\mathrm{Ad}(m)H=H$ and $\mathbb{R}\mathrm{Ad}(K_1)H=\{z\in\mathfrak{u}|\theta z=-z\}=\mathfrak{u}_-$. Thus $\mathrm{Ad}(m)$ stabilizes $\mathfrak{u}_-$. Since $\mathfrak{u}=\mathfrak{u}_-+[\mathfrak{u}_-,\mathfrak{u}_-]$ it follows that $\mathrm{Ad}(m)\mathfrak{u}=\mathfrak{u}$. Fix $x\in\mathfrak{u}\cap\mathfrak{n}_1$ with $B(\theta x,x)=-1$. Then if $y=-c\theta x$ we have $B([x,y],H)=-cB(x,[\theta x,H])=-cB(x,\theta x)$. Thus if we take $c=2B(H,H)$ then $[x,y]=2H$. Let $M^o$ denote the identity component of $M$. If $m\in K\cap M$ then $\mathrm{Ad}(m)\mathfrak{u}\cap\mathfrak{n}_1\subset\mathfrak{u}\cap\mathfrak{n}_1$. Let $U$ denote the connected subgroup of $G$ with Lie algebra $\mathfrak{u}$. Then $K\cap M\cap U$ acts transitively on $\{z\in\mathfrak{u}\cap\mathfrak{n}_1|B(z,\theta z)=-1\}$. Thus there exists $u\in K\cap M\cap U$ such that $\mathrm{Ad}(u)\mathrm{Ad}(m)x=x$. Since $K\cap M\cap U$ is connected, this implies that if $(M\cap K)^x=\{m\in K\cap M|\mathrm{Ad}(m)x=x\}$then $(M\cap K)^xM^o=(M\cap K)M^o=M$. We therefore must show that if $m\in(K\cap M)^x$ then it acts on $\nu(\mathbb{C}_{k(\Lambda_1+\Lambda_2)})$ by $\chi(m)^{k-1}$. Fix such an $m$. Then we note that $\mathrm{Ad}(m)y=y$ (since $\mathrm{Ad}(m)H=H$), $\mathrm{Ad}(m)X=\chi(m)X$, and $\mathrm{Ad}(m)Y=\chi(m)Y$. Thus if $\chi(m)=1$ then $\mathrm{Ad}(m)$ acts trivially on $\mathfrak{u}$ so $\beta(m)=1$. In particular, since the subrepresentation $H^{(k-1)(\Lambda_1+\Lambda_2)}$ occurs with multiplicity one in $L^k$ we see that $m$ acts as a scalar on that space. The scalar, can thus be calculated from the action of $K$ on the minimal $K$-type which has highest weight $(k-1)\beta$. We now look at the case when $\chi(m)=-1$ (note that $\chi(m)$ is a unitary character taking real values). This time we have

$$\mathrm{Ad}(m)x=x, \mathrm{Ad}(m)y=y, \mathrm{Ad}(m)X=-X, \mathrm{Ad}(m)Y=-Y.$$

This implies that if we identify $U$ with $SU(2,1)$ then $mum^{-1}=\bar{u}$ (complex conjugation). Hence on $U_C$ (identified with $SL(3,\mathbb{C})$) we have $mum^{-1}=I_{2,1}(u^T)^{-1}I_{2,1}$ (here we use the notation of section 3 and super T is transpose). This implies that $\mathrm{Ad}(m)$ preserves the $\beta$-root space and $\beta(m)=\chi m)$. We observe that since $m\in K$ it preserves the minimal $K$-type. Since it normalizes $\mathfrak{u}$ it preserves the space $H^{(k-1)(\Lambda_1+\Lambda_2)}$, it therefore acts on the space $\nu(\mathbb{C}_{k(\Lambda_1+\Lambda_2)})$ by its value on the minimal $K$-type which is the scalar $\beta(m)^{k-1}$. This completes the proof. $\square$

## 7. Generic characters

In this section we will study generic unitary characters of $N$ with $P = MN$ as in the previous sections. We will use the definition in [Wa5] which we will now recall in our context. Let $\psi : \mathfrak{n}_C \to \mathbb{C}$ be a Lie algebra homomorphism. Then there exists $z_\psi \in (\theta\mathfrak{n}_1)_C$ such that $\psi(u) = B(z_\psi, u)$ for $u \in \mathfrak{n}_C$. We say that $\psi$ is generic $\mathrm{Ad}(M_C)z_\psi$ is open in $(\theta\mathfrak{n}_1)_C$.

If $\psi$ is a unitary character of $N$ then we will also use the notation $\psi$ for its differential. Then $\psi = i\phi$ and so $z_\psi = iz_\phi$. If we interchange $H$ with $-H$ then we note that $\psi$ is generic if and only if $z_\phi$ is generic. The results in section 5 imply that there exists a unique $x_\phi$ in $\mathfrak{n}_1$ such that $[x_\phi, z_\phi] = 2H$. In this section we will determine the generic unitary characters of $N$ that give rise to Whittaker models for the quaternionic discrete series. Before we can do this we will need some simple analytic results.

We first give a simple asymptotic result on the growth of certain generalized matrix entries. Let $\psi : \mathfrak{n} \to i\mathbb{R}$ be a Lie algebra homomorphism. We will identify $\psi$ with the corresponding unitary character of $N$. If $(\pi, V)$ is Hilbert representation of $G$ then we will use the notation

$$Wh_\psi^\infty(V) = \{\lambda \in V_\infty^{'} | \lambda \circ \pi(n)^{-1} = \psi(n)\lambda\}.$$

**Lemma 8.** *Let $(\pi, V)$ be a Hilbert representation of $G$. If $\psi \neq 0$ (i.e. not the trivial unitary character) then for each $k = 1, 2, \ldots$ there exists a continuous seminorm $\nu_k$ on $V_\infty$ such that*

$$|\lambda(\pi(\exp tH)v)| \leq e^{-kt}\nu_k(v)$$

*for $t \geq 0$ and $v \in V_\infty$.*

*Proof.* The argument is essentially the same as that in Lemma 1 section 3. Fix $x \in \mathfrak{n}_1$ such that $\psi(x) \neq 0$. We may assume that $\psi(x) = -i$. Then $\lambda(\pi(x)v) = i\lambda(v)$. Also Lemma 11.5.1 in [Wa3] implies that there exists a continuous seminorm $\mu$ on $V_\infty$ and $c \geq 0$ such that

$$|\lambda(\pi(\exp tH)v)| \leq e^{ct}\mu(v)$$

for all $t \geq 0$ and $v \in V_\infty$. We have

$$\begin{aligned}\lambda(\pi(\exp tH)v) &= -i\lambda(\pi(x)\pi(\exp tH)v) = -i\lambda(\pi(\exp tH)\pi(\mathrm{Ad}(\exp -tH)x)v) \\ &= -ie^{-t}\lambda(\pi(\exp tH)\pi(x)v).\end{aligned}$$

We therefore have

$$|\lambda(\pi(\exp tH)v)| \leq e^{-t}e^{ct}\mu(\pi(x)v.$$

We set $\mu_1(v) = \mu(\pi(x)v$. Then $|\lambda(\pi(\exp tH)v)| \leq e^{-t}e^{ct}\mu_1(v)$ for all $v \in V_\infty, t \geq 0$. We replacing $\mu$ by $\mu_1$ and $c$ by $c-1$ we can repeat this argument. The result follows by the, by now, obvious iteration of this method. ☐

We will now come to the crux of the matter.

**Theorem 10.** *Let $\psi$ be a non-trivial unitary character of $N$. Let $k \geq d-1$ and let $\lambda \in Wh_\psi^\infty(L^k)$. If $\psi(\mathfrak{n} \cap \mathfrak{u}) = 0$ (recall that we are using the same notation for $\psi$ and the differential of $\psi$) then $\lambda = 0$.*

*Proof.* The representation $L^k$ is admissible with respect to $K \cap U$. Thus the decomposition given in Theorem 8 in section 6 applies to the $K$-finite vectors which are contained in the $C^\infty$ vectors. Let $W$ be an irreducible summand of $L^k$ restricted to $U$. Then $W_{K\cap U} \subset L_\infty^k$. Since $\psi(\mathfrak{n} \cap \mathfrak{u}) = 0$ it follows that $\lambda_{|W_{K\cap U}}$ factors through $W_{K\cap U}/((\mathfrak{n} \cap \mathfrak{u})W_{K\cap U})$. This space is finite dimensional and is invariant under the action of $H$. This implies that there exist $c_1, \ldots, c_d \in \mathbb{C}$, non-negative integers $n_1, \ldots, n_d$ and linear functionals $q_{i,j}$, $i = 1, \ldots, d$ and $j = 0, \ldots, n_i$ on $W_{K\cap U}$ such that

$$\lambda(\Pi_k(\exp tH)v) = \sum_{i,j} e^{c_i t} t^j q_{ij}(v)$$

for all $v \in W_{K\cap U}$. We note that if $\phi(t) = \sum_{i,j} e^{c_i t} t^j a_{ij}$ (finite sum) with $a_{ij} \in \mathbb{C}$ and if for each $m \in \mathbb{Z}$, $m > 0$ there exists $C_m$ such that

$$|\phi(t)| \leq C_k e^{-kt}$$

for all $t > 0$ then $a_{ij} = 0$ for all $ij$. Thus Lemma 8 implies that $\lambda_{|W_{K\cap U}} = 0$. This implies that $\lambda_{|L_{K\cap U}^k} = 0$. Since $L_{K\cap U}^k$ is dense in $L_\infty^k$, $\lambda = 0$. □

This leads us to a new definition. Let $\psi$ be a non-degenerate character of $N$. Then we say that $\psi$ is *admissible* if $\psi \circ \mathrm{Ad}(m)_{|\mathfrak{n}\cap\mathfrak{u}} \neq 0$ for all $m \in M$.

The main result of this section is

**Theorem 11.** *Let $\psi$ be a non-degenerate unitary character of $N$. Then the following are equivalent*

1. *There exists $k \geq d+1$ such that $Wh_\psi^\infty(L^k) \neq 0$.*
2. *$Wh_\psi^\infty(L^k) \neq 0$ for all $k \geq d+1$.*
3. *$\psi$ is admissible.*

*Furthermore, the set of admissible characters of $N$ form a single orbit under the action of $M$.*

*Proof.* We note that the preceding theorem implies that if condition 1.is true for $\psi$ then $\psi$ is admissible. Thus we see that 2. implies 1. implies 3. We therefore must only prove that 3. implies 2. This assertion will be proved as follows. We will first show that for each $k$ there exists at least one non-degenerate $\psi$ such that $Wh_\psi^\infty(L^k) \neq 0$. We will then prove the last assertion of the theorem. The combination implies that $Wh_\psi^\infty(L^k) \neq 0$ for all admissible $\psi$. This will complete the proof. We now began the implementation of the plan.

To prove that there exists some non-degenerate $\psi$ such that $Wh_\psi^\infty(L^k) \neq 0$. We observe that if $\Lambda = (k-1)(\Lambda_1 + \Lambda_2)$ then we have a unitary imbedding of $H^\Lambda_{K\cap U}$ into $L^k_K$ (see Theorem 4). We will identify this space with its image. We will apply Theorem 7 in section 4 which implies that there exists a unitary character $\chi$ of $N \cap U$ and $v, w \in H^\Lambda_{K\cap U}$ such that

$$\int_{N\cap U} \chi(n)^{-1} \langle \pi_\Lambda(n)v, w\rangle \, dn \neq 0.$$

Now, $N \cap U$ is a normal subgroup of $N$. Thus if we set

$$\phi(x) = \int_{N\cap U} \chi(n)^{-1} \langle \Pi_k(xn)v, w\rangle \, dn$$

for $x \in N$. We note that the standard estimates of Harish-Chandra (cf. [Wa3, 7.2.1]) imply that the function $|\phi(x)|$ is in $L^2(N/N\cap U) \cap L^1(N/N\cap U)$. This implies that

$$\Omega(\psi) = \int_N \psi(n) \langle \Pi_k(n)v, w\rangle \, dn.$$

defines a continuous function on the set of unitary characters on $N$ that extend $\chi$. Since the set of non-degenerate characters that extend $\chi$ form an open dense subset we see that there must exist $\psi$ such that $\Omega(\psi) \neq 0$. Hence if we define

$$\lambda(u) = \int_N \psi(n) \langle \Pi_k(n)u, w\rangle \, dn$$

for $u \in L^k_\infty$ then $\lambda \neq 0$ and $\lambda \in Wh_\psi^\infty(L^k)$. We now begin the second step of the proof.

There are several special cases that we must consider before we do the "general case". If $\mathfrak{g}_C$ is of type $A_{n+1}$ then we are considering the real form $SU(2,n)$ with $n \geq 1$. If $n = 1$ then we have seen in the discussion of generic orbits in section 2 that there is only one open orbit this concludes the proof in this case. If $n \geq 2$ then we have seen that there are two open orbits

$$\mathcal{O}_\pm = \{\psi | \pm \langle x_\phi, x_\phi\rangle > 0\}$$

and both are complex. We note that $\mathfrak{u} \cap \mathfrak{n}_1 \subset \{x | \langle x, x\rangle > 0\}$. Proposition 5 section 5 implies that all of the admissible characters are contained in $\mathcal{O}_+$. The theorem now follows in this case.

We next look at the when is $G$ locally isomorphic with (split) $G_2$ or $SO(4,3)$. We have seen that there are exactly two orbits for the action of $M$ on $\mathfrak{n}_1$ in either of these cases. We will now show that in this case admissible implies complex. We note that it is enough to check this for the case of $G_2$. Here we can identify $\mathfrak{n}_1$ with the binary forms of homogenous of degree 3. With the identity component of $[M, M]$ acting through the classical action of $SL(2,\mathbb{R})$ on the binary forms. The invariant $f$ is up to a positive scalar multiple the negative of the discriminant. That is, if

$$a(x,y) = a_0x^3 + a_1x^2y + a_2xy^2 + a_3y^3$$

then the discriminant is

$$a_1^2a_2^2 - 4a_0a_2^3 - 4a_1^3a_3 - 27a_0^2a_3^2 + 18a_0a_1a_2a_3.$$

The space $\mathfrak{u} \cap \mathfrak{n}_1$ is spanned by $p(x,y) = x^3 - xy^2$ and $q(x,y) = -x^2y + y^3$. The $[M,M]$-invariant pairing is given by

$$Q(a,b) = 2a_0b_3 - 3a_3b_0 - a_1b_2 + a_2b_1$$

if

$$a(x,y) = a_0x^3 + a_1x^2y + a_2xy^2 + a_3y^3$$

and

$$b(x,y) = b_0x^3 + b_1x^2y + b_2xy^2 + b_3y^3.$$

The form $h(x,y)$ is admissible if and only if $Q(h \circ g, p)$ or $Q(h \circ g, q)$ is non-zero for each $g \in GL(2,\mathbb{R})$. Consider the form $h(x,y) = x^3 + y^3$ Then if we set

$$g = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

we calculate that both $Q(h \circ g, p)$ and $Q(h \circ g, q)$ are 0.

We are left with the quaternionic real forms of real rank 4. For these Proposition 6 implies that the only possible admissible characters $\psi$ such that $x_\phi$ can be conjugated into $\mathfrak{u} \cap \mathfrak{n}_1$. These form a single orbit. This completes the proof. $\square$

## 8. The Heisenberg–Whittaker vectors for certain degenerate principal series representations

In this section we will give a complete analysis of Heisenberg–Whittaker models for the degenerate principal series associated with the Heisenberg parabolic subgroup. In the next section we will apply these results to the quaternionic discrete series Before we begin we discuss the nature of the stabilizer of a generic character of $N$ in $M$.

If $\psi$ is a unitary character of $N$ then set $M_\psi = \{m \in M | \psi \circ \mathrm{Ad}(m) = \psi\}$. We note that $M_\psi = M_x = \{m \in M | \mathrm{Ad}(m)x = x\}$ with $x = x_\phi$ if $\psi = i\phi$.

We will be using a theorem of [Ko-V] to analyze the space $Wh_\psi^\infty(I_{P,\sigma,\nu})$ for $\psi$ a generic character of $N$. We first observe that the Bruhat Lemma implies that

$$G = \bigcup_{s \in W_M \backslash W / W_M} PwP$$

where $W$ is the (small) Weyl group of $G$ and $W_M$ is the (small) Weyl group of $M$. There is exactly one double coset $Pw_oP$ that is open and all the others have strictly lower dimension. Fix a minimal parabolic subgroup $P_o = M_oN_o$ with $M_o \subset M$ of $G$ contained in $P$ and set $P_o^* = P_o \cap M$. Let $M_o = {}^oM_oA_o$ with $A_o$ the identity component of the $\mathbb{R}$-split torus in $M_o$. We will look upon $W$ as $N_K(A_o)/C_K(A_o)$ (as usual) where $N_K$ denotes normalizer in $K$ and $C_K$ denotes centralizer in $K$ (note $C_K(A_o) = M_o$. Let $w_G$ be the maximal element of $W$ with respect to the choice of $P_o$ and let $w_M$ be the longest element of $W_M$ relative to $P_o^*$. Since $M$ is the centralizer of the coroot of the highest root relative to $P_o$ we see that we can choose a representative $k \in N_K(A_o)$ for $w_G$ such that $kMk^{-1} = M$. We may take $w_o$ to be equal to $w_G w_M$ and it is easy to check that we can choose a representative of $w_o$, $k_o \in N_K(A_o)$, such that $k_o N k_o^{-1} = \theta(N)$. Hence we have

$$Pw_oP = Nk_oP.$$

We note that $H \in Lie(A_o)$. Let $\Phi$ be the set of restricted roots of $A_o$ on $Lie(G)$. If $\alpha$ is a restricted root then we will use the notation $\mathfrak{g}^\alpha$ for the corresponding $A_o$ weight space. Let $\Phi^+$ be the system of positive roots corresponding to $P_o$ and let $\Phi_M^+$ be the positive roots that are roots of $P_o^*$. Now $[\mathfrak{n}, \mathfrak{n}]$ is a one dimensional space that is $A_o$ invariant. Let $\delta$ denote the corresponding restricted root (the coroot of $\delta$). We note that

$$\Phi^+ - \Phi_M^+ = \{\alpha \in \Phi | \langle \delta, \alpha \rangle > 0\}.$$

Let $\Sigma = \Phi^+ - \Phi_M^+ - \{\delta\}$. We can write $\Sigma$ as a disjoint union of subsets $\Sigma_j$ each consisting of one element or it consists of two such that if we set $\mathfrak{n}^j = \sum_{\alpha \in \Sigma_j} \mathfrak{g}^\alpha$ then the form $\omega$ defined by $[u, v] = \omega(u, v)X$ on $\mathfrak{n}_1$ is non-degenerate on each of the spaces $\mathfrak{n}^j$ Furthermore, $[\mathfrak{n}^j, \mathfrak{n}^k] = 0$ if $j \neq k$.

**Lemma 9.** *Let $\psi$ be regular character of $N$. If $w \in W$ is such that $w\Phi_M^+ \subset \Phi^+$ and if $\psi$ is trivial on $kN_ok^{-1} \cap N$ with $k$ a representative of $w$ then $kN_ok^{-1} \cap N = \{1\}$.*

*Proof.* We first make some observations. We assume that $w$ satisfies $w\Phi_M^+ \subset \Phi^+$. We fix an linear order $>$ on $Lie(A_o)^*$ that produces $\Phi^+$.

1. If $w\delta = -\delta$ then $kN_ok^{-1} \cap N = \{1\}$.

If $\alpha \in \Sigma$ then $(\delta, \alpha) = (w^{-1}\delta, w^{-1}\alpha) = -(\delta, w^{-1}\alpha)$. Thus $(\delta, w^{-1}\alpha) < 0$. Hence $w^{-1}\Sigma = -\Sigma$.

2. If $-\Sigma_i \subset w\Phi^+$ for some $i$ then $kNk^{-1} \cap N_o = \{1\}$.

Indeed, if $\Sigma_j = \{\alpha\}$ and if $-\alpha \in w\Phi^+$then $-\alpha = w\gamma$ with $\gamma \in \Sigma \cup \{\delta\}$ (since $w\Phi_M^+ \subset \Phi^+$). Now $2\alpha = \delta$, $[\mathfrak{g}^{-\alpha}, \mathfrak{g}^{-\alpha}] = \mathfrak{g}^{-\delta} \neq 0$. Hence, $[\mathfrak{g}^\gamma, \mathfrak{g}^\gamma] \neq 0$. But then the bracket must be $\mathfrak{g}^\delta$. This implies that $w\delta = -\delta$. Now apply 1. above. If $\Sigma_j = \{\alpha, \lambda\}$ and $w\Phi^+ \supset -\Sigma_j$ then, as before, there must $\gamma, \mu \in \Sigma \cup \{\delta\}$ such that $w\gamma = -\alpha$, $w\mu = -\lambda$. But then (as in the previous case) we must have $[\mathfrak{g}^{-\alpha}, \mathfrak{g}^{-\lambda}] = \mathfrak{g}^{-\delta} \neq 0$ and hence $[\mathfrak{g}^\gamma, \mathfrak{g}^\mu] \neq 0$ so as before $[\mathfrak{g}^\gamma, \mathfrak{g}^\mu] = \mathfrak{g}^\delta$ and hence $w\delta = -\delta$ as in the previous case.

Write $\psi(x) = iB(y, x)$ for all $x \in \mathfrak{n}_1$ with $y \in \theta\mathfrak{n}_1$. Now $y = \sum y_i$ with $y_i \in \theta\mathfrak{g}^{\alpha_i}$ and $\alpha_i \in \Sigma$, giving an enumeration of $\Sigma$. Suppose that $\psi$ is trivial on $kNk^{-1} \cap N$ and that $kNk^{-1} \cap N \neq \{1\}$. We will derive a contradiction. Since $\psi$ is generic we have $ad(y)^4X = cY$ with $c \neq 0$. Hence

$$cY = ad(y)^4X = \sum ad(y_{i_1})ad(y_{i_2})ad(y_{i_3})ad(y_{i_4})X$$

the sum over all indices $1 \leq i_1, i_2, i_3, i_4 \leq |\Sigma|$. There must therefore be a choice of 4 indices (possibly not distinct) $j_1, j_2, j_3, j_4$ such that

$$ad(y_{j_1})ad(y_{j_2})ad(y_{j_3})ad(y_{j_4})X = aY$$

with $a \neq 0$. In particular this implies that $\delta - \alpha_{j_1} - \alpha_{j_2} - \alpha_{j_3} - \alpha_{j_4} = -\delta$. Thus $2\delta = \alpha_{j_1} + \alpha_{j_2} + \alpha_{j_3} + \alpha_{j_4}$. We assert that we may assume that $w^{-1}\alpha_{i_j} < 0$. Indeed, if $w^{-1}\alpha_{i_j} > 0$ then $\mathfrak{g}^{\alpha_{i_j}} \subset \mathrm{Ad}(k)\mathfrak{n}_o \cap \mathfrak{n}$ and so $d\psi(\mathrm{Ad}(k)\mathfrak{n}_o \cap \mathfrak{n}) \neq 0$. This would imply that $w^{-1}\delta < 0$. On the other hand 2. implies that if $\alpha \in \Sigma$ and if $w^{-1}\alpha < 0$ then $w^{-1}(\delta - \alpha) > 0$. However since

$$(\delta - \alpha_{j_1}) + (\delta - \alpha_{j_2}) + (\delta - \alpha_{j_3}) + (\delta - \alpha_{j_4}) = 2\delta.$$

we have the (desired) contradiction $w^{-1}\delta > 0$. □

**Lemma 10.** *Let $\psi$ be a regular character of $N$ and let $m \in M$ and $k \in K$ such that $k$ is a representative of $w \in W$ such that $w\Phi_M^+ \subset \Phi^+$ and $kNk^{-1} \cap N \neq \{1\}$ Then*

$$\psi_{|mkN(mk)^{-1} \cap N} \neq 1.$$

*Proof.* The conclusion is equivalent with the assertion that

$$\psi \circ \mathrm{Ad}(m)_{|kNk^{-1} \cap N} \neq 1.$$

But this follows from the previous lemma since the regularity of $\psi$ implies that of $\psi \circ \mathrm{Ad}(m)$. □

We are almost ready to apply a result of [Ko-V]. To do so we will need a bit more notation.

We set $A = \exp(\mathbb{R}H)$. Let $\mathcal{X}(M)$ denote the set of all continuous homomorphisms from $M$ to $\mathbb{R}^{>0} = \{x \in \mathbb{R} | x > 0\}$. Set $M^o = \cap_{\chi \in \mathcal{X}(M)} Ker(\chi)$. If $(\sigma, H_\sigma)$ is a finite dimensional representation of $M^o$ and if $\nu \in \mathbb{C}$ then we will denote by $\sigma_\nu$ the representation of $P$ on $H_\sigma$ given by $\sigma_\nu(m\exp(tH)n) = e^{(\nu+d+1)t}\sigma(m)$ for $m \in M^o, t \in \mathbb{R}$ and $n \in N$. We set $I^\infty_{P,\sigma,\nu}$ equal to the space of $C^\infty$ vectors in the induced representation $I(\sigma_\nu)$ (see section 6). Let $\pi_{\sigma,\nu}$ be the action of $M$ on the representation space. Set $s_o = w_o^{-1}$ (that is $w_M w_G$). We denote by $U_{\sigma,\nu}$ the space of all $f \in I^\infty_{P,\sigma,\nu}$ supported in $Ns_oP$ and having compact support module $P$. With this notation in hand, we have

**Theorem 12.** *Let $\psi$ be an regular unitary character of $N$ and let $\lambda \in Wh^\infty_\psi(I^\infty_{P,\sigma,\nu})$. Then if $\lambda_{|U_{\sigma,\nu}} = 0$ then $\lambda = 0$.*

*Proof.* We apply Theorem 3.15, p. 82 of [Ko-V]. This theorem is a generalization of Bruhat theory which (fortunately applies to the case at hand). The specific case of the theorem we will use is their case iii). To refer to their notation we consider the Lie group $H = P \times P$ acting on $G$ via $(p_1, p_2)g = p_1 g p_2^{-1}$. (Sorry about all of the $H$'s). We also have $H' = N \times P$ acting as a subgroup of $H$. We consider the representation $\mu$ of $H^{'}$ on $H_\sigma$ given by $\mu(n,p)=\psi(n)\sigma_\nu(p)$. This is a finite dimensional representation of $H^{'}$. $H$ acts on $G$ with a finite number of orbits (see the discussion above). On an orbit $PwP$ with $w\Phi_M^+\subset\Phi^+$ of $H$ the orbits of $H^{'}$ are $NmkP$ with $m\in M$ and $k$ a representative of $w$. Lemma 10 implies that if $PwP$ is not the unique open orbit (which is also an orbit of $H^{'}$) then we have $\psi_{|mkN(mk)^{-1}\cap N}\neq 1$. Let $y=mk$. The stability group of $y$ in $H^{'}$ is the group of all pairs $(n,p)$ with $nmkp^{-1}=mk$. That is $m^{-1}nm=kpk^{-1}$. Since $\mathrm{Ad}(k)$ permutes the root spaces and normalizes $Lie(A_o)$ and since $w\Phi_M^+\subset\Phi^+$ we see that the stability group of $y$ is $H'$ consists of the set of pairs $(n,y^{-1}ny)$ with $n\in\mathrm{Ad}(m)(N\cap kN_ok^{-1})$. Thus if $NyP$ is not the open orbit Lemma 10 implies that equation (3.27) of [Ko-V] is satisfied. Thus (taking appropriate inverses) and using the standard Bruhat theoretic arguments (cf. [Wa5], [Ko-V, p. 88 second paragraph]) the Theorem follows. □

**Corollary 2.** *Let $M_\psi$ act on $Wh_\psi^\infty(I_{P,\sigma,\nu}^\infty)$ by $m\cdot\lambda=\lambda\circ\pi_{\sigma,\nu}(m)^{-1}$. Then $Wh_\psi^\infty(I_{P,\sigma,\nu}^\infty)$ is equivalent with a subrepresentation of the representation contragredient to $(\sigma_{|M_\psi},H_\sigma)$ as an $M_\psi$-module.*

*Proof.* We first note that if $T\in(C_c^\infty(N))^{'}$ (continuous linear functionals in the usual topology) is such that $T\circ R_x=\psi(x)^{-1}T$ $(R_xf(y)=f(yx))$. Then

$$T(f)=c(T)\int_N f(n)\psi(n)dn.$$

This can be seen as follows. Let $\{X_j\}$ be a basis of $Lie(N)$ (thought of as left invariant vector fields). Then we have

$$X_jT=-\psi(X_j)T$$

in the sense of distributions. Thus

$$\Bigl(\sum_j X_j^2-\sum_j(\psi(X_j)^2)T=0.$$

The elliptic regularity theorem implies that $T$ is given by integration against real analytic $n=\dim N$-form on $N$. The transformation law now easily implies that the form is $\psi(n)\omega$ with $\omega$ invariant.

If $f\in C_c^\infty(N)$ and $v\in H_\sigma$ then we define $S(f\bigotimes v)(ps_1n)=\sigma_\nu(p)f(n)v$ for $p\in P$, $n\in N$. If $\lambda\in Wh_\psi^\infty(I_{P,\sigma,\nu}^\infty)$ then we have $T_{\lambda,v}(f)=\lambda(S(f\bigotimes v))$ defines a distribution on $N$ and $T_{\lambda,v}(R_xf)=\psi(x)T_{\lambda,v}(f)$ for all $x\in N$. Thus we have a $\mathbb{C}$-bilinear pairing defined on $Wh_\psi^\infty(I_{P,\sigma,\nu}^\infty)\times H_\sigma$ defined by

$$T_{\lambda,v}(f)=\langle\lambda,v\rangle\int_N f(n)\psi(n)dn.$$

A direct calculation shows that

$$\langle m\cdot\lambda,v\rangle=\left\langle\lambda,\sigma(s_1ms_1^{-1})^{-1}v\right\rangle.$$

These observations combined with the previous theorem complete the proof. □

We now come to the main result of this section which is now proved by exactly the same method as that of Theorem 7.2 of [Wa5].

**Theorem 13.** *Let $\psi$ be a generic unitary character of $N$ then $Wh_\psi^\infty(I_{P,\sigma,\nu}^\infty)$ is equivalent as an $M_\psi$ with the representation contragredient to $(\sigma_{|M_\psi}, H_\sigma)$.*

## 9. Heisenberg–Whittaker vectors for quaternionic discrete series

The purpose of this section is to prove the theorem for Heisenberg–Whittaker vectors for quaternionic discrete series that is analogous to Theorem 5 for abelian Whittaker vectors for holomorphic discrete series. We will first recall a few results from [Wa5]. If $V$ is a $\mathfrak{g}$-module and $\psi : \mathfrak{n} \longrightarrow \mathbb{C}$ is a Lie algebra homomorphism then we say that $V$ is a Whittaker module with respect to $\psi$ if for each $v \in V$ there exists $m > 0$ such that $(X - \psi(X))^m v = 0$ for all $X \in \mathfrak{n}$. We denote by $\mathbb{C}_\psi$ the $\mathfrak{n}$-module that has action given by $\psi$. We set $W_\psi(V) = \{v \in V | Xv = \psi(X)v, X \in \mathfrak{n}\}$. We recall the two results of [Wa5] that we will be using (we note that the assertion about the action of $M_\psi$ in the second result is a consequence of the construction of the map $T$).

**Theorem 14.** *If $\psi$ is an admissible character of $\mathfrak{n}$ and if $V$ is a Whittaker module for $\mathfrak{g}$ with respect to $\psi$ then*

$$H^i(\mathfrak{n}, V \bigotimes \mathbb{C}_{-\psi}) = 0$$

*for $i > 0$.*

**Theorem 15.** *Let $F$ be a finite dimensional $\mathfrak{g}$-module, let $\psi$ be an admissible character of $\mathfrak{n}$ and let $V$ be a Whittaker module for $\mathfrak{g}$ with respect to $\psi$ that is also a compatible $M_\psi$ module ($m \cdot X \cdot v = (\mathrm{Ad}(m)X) \cdot m \cdot v$). Then there exists an element $T \in U(\mathfrak{g})^{M_\psi}$ depending only on the dimension of $F$ such that the action of $T$ on $V \bigotimes F$ induces an $M_\psi$-module isomorphism of*

$$W_\psi(V) \bigotimes F \to W_\psi(V \bigotimes F).$$

In addition we will need the following simple consequence of Theorem 15 [Wa5, Proposition 1.3]. First we need a bit of notation. If $V$ is an $\mathfrak{n}$-module then we set $V[\psi] = \{v \in V | (x - \psi(x))^k v = 0$ for all $x \in \mathfrak{n}$ and some $k\}$. If $V$ is a $\mathfrak{g}$ module then since the adjoint action of $\mathfrak{n}$ on $\mathfrak{g}$ is nilpotent we see that $V[\psi]$ is a $\mathfrak{g}$-submodule of $V$. If $W$ is a vector space then we define an $\mathfrak{n}$-module structure on $\mathrm{Hom}_{\mathbb{C}}(U(\mathfrak{n}), W)$ as follows: $xf(g) = f(gx)$. We set $P(W) = \mathrm{Hom}_{\mathbb{C}}(U(\mathfrak{n}), W)[\psi]$. If $W$ is a compatible module $M_\psi$ then we can define a compatible action of $M_\psi$ on $P(W)$ by setting $mf(g) = m \cdot f(\mathrm{Ad}(m)^{-1} g)$.

**Proposition 7.** *If $V$ is a Whittaker module for a generic character $\psi$ that is a compatible $M_\psi$-module then $V$ is isomorphic with $P(W_\psi(V))$ as a compatible $M_\psi$ and $\mathfrak{n}$-module.*

Before we can state the main result of this paper we need to recall several properties of quaternionic discrete series representations. We will use the notation of section 6. Let $(\Pi_\Lambda, L^\Lambda)$ be as in that section. Then the general theory of discrete series implies that the irreducible representation, $(\tau_\lambda, V_\lambda)$, of $K$ with highest weight $\lambda = \Lambda+\rho-2\rho_c$ ($\rho_c$ the half sum of the positive roots of $K$) occurs with multiplicity 1 in $L^\Lambda$ and it is the lowest $K$-type in a sense unimportant to us in this context (see [H-S]). Since $K$ is locally isomorphic with $K_1 \times K_2$ the representation $V_\lambda$ is isomorphic with a tensor product $F_{k(\lambda)} \bigotimes W_{\lambda_1}$ with $F_k$ the $k+1$ dimensional irreducible representation of $K_1 \cong SU(2)$. Let $T_1$ be the maximal compact torus of $K_1$ such that there exists a maximal compact torus of $K_2$ such that $T_1 \times T_2$ corresponds to our chosen torus for $K$. Let $L_\mathbb{C}$ denote the complexification of the subgroup of $G$ corresponding to $T_1 \times K_2$. As in the holomorphic case the group $L_\mathbb{C}$ is isomorphic with the subgroup $M_\mathbb{C} = \{g \in G_\mathbb{C} | \mathrm{Ad}(g)H = H\}$ (see the previous section for $H$). There is also an isomorphism of $L_\mathbb{C}$ onto $M_\mathbb{C}$ given by an appropriate Cayley transform (see [Gr-W]). We may thus look upon $\mathbb{C}_{k\beta} \bigotimes W_{\lambda_1}$ as a representation of $M_\mathbb{C}$ which we denote by the symbol $\widetilde{W_\lambda}$.

**Theorem 16.** *Let $\psi$ be a generic unitary character of $N$. If $(\Pi_\Lambda, L^\Lambda)$ is a quaternionic discrete series representation of $G$ then $Wh_\psi^\infty(L^\Lambda) = 0$ unless $\psi$ is admissible. If $\psi$ is admissible then as an $M_\psi$-module $Wh_\psi^\infty(L^\Lambda)$ is isomorphic with the representation contragredient to $\widetilde{W_\lambda}$.*

*Proof.* We set $\sigma_\Lambda$ equal to the irreducible representation of $M$ gotten by restriction of $|\chi|\widetilde{W_\lambda}$ to $M$ (see Theorem 9). We will denote by $I(\Lambda)$ the corresponding (unnormalized) parabolic induced representation induced from $P = MN$. Then $\Lambda \to L^\Lambda$ and $\Lambda \to I(\Lambda)$ form coherent families (see the discussion preceding Theorem 6 section 4 for the special case of $SU(2)$). Theorem 9 implies that $\dim \mathrm{Hom}_{\mathfrak{g},K}((L^\Lambda)_K, I(\Lambda)_K) = 1$ if $\Lambda = \Lambda_k$ for $k \geq d+1$. Coherence of the family implies that the same result is true for dominant integral $\Lambda$. We set $\widetilde{I}(\Lambda)$ equal to the admissible, smooth conjugate dual of $I(\Lambda)$. Then $\Lambda \to \widetilde{I}(\Lambda)$ is also a coherent family of parabolic induced representations and since the representations $L^\Lambda$ are unitary we have $\dim \mathrm{Hom}_{\mathfrak{g},K}(\widetilde{I}(\Lambda)_K, (L^\Lambda)_K) = 1$. The theorem of Casselman and the author (cf. [Wa4]) implies that up to scalar multiple there exists a unique, nonzero continuous intertwining operator $T_\Lambda : \widetilde{I}(\Lambda)^\infty \to (L^\Lambda)^\infty$ which is surjective (by the exactness of this globalization). The adjoint map $T'_\Lambda : ((L^\Lambda)^\infty)' \to (\widetilde{I}(\Lambda)^\infty)'$ is injective. The intertwining condition implies that $T'_\Lambda(Wh_\psi^\infty(L^\Lambda)) \subset Wh_\psi^\infty(\widetilde{I}(\Lambda))$. If $\Lambda = \Lambda_k$ with $k \geq d+1$ then Corollary 2 of section 8 and Theorem 11 of section 7 imply the assertions of the theorem. To complete the proof of the theorem we will need to study coherence properties of generic Whittaker vectors in these coherent families.

We note that the injectivity of the adjoint map implies that if $\psi$ is a generic character of $N$ then $Wh_\psi^\infty(L^\Lambda)$ is equivalent as an $M_\psi$ module with a subrepresentation of the contragredient module to $\widetilde{W_\lambda}$ (here note that the inducing module for $\widetilde{I}(\Lambda)$ is equivalent with that for $I(\Lambda)$ as an $M_\psi$-module). We now consider the module $((L^\Lambda)^\infty)'[\psi]$. If $F$ is a finite dimensional $G$-module then $((L^\Lambda \bigotimes F)^\infty)'[\psi] = ((L^\Lambda)^\infty)'[\psi] \bigotimes F^*$. Thus from the definition of the Zuckerman translation functors we have $((L^{\Lambda+\mu})^\infty)'[\psi]$ is a compatible $M_\psi$ and $\mathfrak{g}$ submodule of $((L^\Lambda)^\infty)'[\psi] \bigotimes F^*$ and $((L^\Lambda)^\infty)'[\psi]$ is a compatible $M_\psi$ and $\mathfrak{g}$ submodule of $((L^{\Lambda+\mu})^\infty)'[\psi] \bigotimes F$ for $\mu$ dominant integral. Applying Theorem 16 and Proposition 7 we find that $Wh_\psi^\infty(L^{\Lambda+\mu})$ is isomorphic with an $M_\psi$ submodule of $Wh_\psi^\infty(L^{\Lambda+\mu}) \bigotimes F^*$ and $Wh_\psi^\infty(L^\Lambda)$ is isomorphic with an $M_\psi$ submodule of $Wh_\psi^\infty(L^\Lambda) \bigotimes F$. This implies that if $Wh_\psi^\infty(L^\Lambda) = 0$ for one element of the family then $Wh_\psi^\infty(L^\Lambda) = 0$ for all elements of the family. Thus Theorem 11 implies that $Wh_\psi^\infty(L^\Lambda) \neq 0$ for some $\Lambda$ only if $\psi$ is admissible. To complete the proof of the theorem we must show that if $\psi$ is admissible then $Wh_\psi^\infty(L^\Lambda)$ is isomorphic with $Wh_\psi^\infty(\widetilde{I}(\Lambda))$ as an $M_\psi$-module. For this we note that (iii) in the proof of Theorem 7.2 of [Wa5] proves that if $\mu$ is dominant integral then

$$\Phi^\mu((\widetilde{I}(\Lambda)^\infty)'[\psi]) \cong (\widetilde{I}(\Lambda+\mu)^\infty)'[\psi]$$

and

$$\Phi_{-\mu}((\widetilde{I}(\Lambda+\mu)^\infty)'[\psi]) \cong (\widetilde{I}(\Lambda)^\infty)'[\psi]$$

as a compatible $\mathfrak{g}$ and $M_\psi$-modules. Now $T'_{\Lambda_k}$ defines an isomorphism of $Wh_\psi^\infty(L^{\Lambda_k})$ with $Wh_\psi^\infty(\widetilde{I}(\Lambda_k))$ thus by Proposition 7 it defines an $M_\psi$-module isomorphism of $((L^{\Lambda_k})^\infty)'[\psi]$ onto $(\widetilde{I}(\Lambda_k)^\infty)'[\psi]$. This combined with the above translation formulas implies that the compatible $\mathfrak{g}$ and $M_\psi$-modules $((L^{\Lambda_k+\mu})^\infty)'[\psi]$ and $(\widetilde{I}(\Lambda_k+\mu)^\infty)'[\psi]$ are isomorphic for all dominant integral $\mu$. This implies by using the $\Phi_{-\sigma}$ for appropriate $\sigma$ that $((L^\Lambda)^\infty)'[\psi]$ and $(\widetilde{I}(\Lambda)^\infty)'[\psi]$ are isomorphic as compatible $\mathfrak{g}$ and $M_\psi$-modules for all $\Lambda$ that are dominant integral and regular. In particular, this implies that $Wh_\psi^\infty(L^\Lambda)$ and $Wh_\psi^\infty(\widetilde{I}(\Lambda))$ are isomorphic as $M_\psi$-modules. The proof is now complete. □

## 10. The main result for holomorphic representations

In this section we will give an analysis of generalized Whittaker vectors for holomorphic representations. We will use results and notation in section 3. We will confine our attention to the case when $G$ is simply connected. We first need some results on canonical completions (in the sense of Casselman and the author) of Harish-Chandra modules for general semisimple Lie groups. This theory was (unfortunately) developed with the assumption that $G$ has finite center. The theory goes through unchanged if we confine our attention to the class of $(\mathfrak{g}, K)$-modules, $V$ that have a $K$-invariant pre-Hilbert space structure. That is $V$ splits into a direct sum of irreducible finite dimensional representations of $K$ each equivalent with a unitary representation. We will now give a rapid tour through some of the necessary changes for the context at hand. Let $W$ be a finite dimensional $(\mathfrak{q}, K)$-module unitary as a $K$-module and let $N(W) = U(\mathfrak{g})\bigotimes_{U(\mathfrak{q})}W$ with $K$-acting as in section 3. Then $N(W)$ is in the class that we had singled out above. Let $\widetilde{W}$ be the $M$-module corresponding to $W$ as in that section. As a $(\mathfrak{q}_1, M \cap K)$ module $N(W) \cong U(\mathfrak{n})\bigotimes\widetilde{W}$ (we will treat this isomorphism as equality). This implies that $N(W)/\mathfrak{n}^k N(W) \cong (U(\mathfrak{n})/\mathfrak{n}^k U(\mathfrak{n}))\bigotimes\widetilde{W}$ as a $(\mathfrak{q}_1, M \cap K)$-module. This module integrates to a representation of $Q_1 = MN$. In general if $V$ is a finite dimensional smooth representation of $Q_1$ then we denote by $I(V)$ the space of all $f : G \to V$ of class $C^\infty$ and such that $f(qx) = qf(x)$ for all $x \in G$, $q \in Q_1$. We endow $I(V)$ with the $C^\infty$-topology and have $G$ act by the right regular representation $(gf(x) = f(xg))$. Then $I(V)$ is an admissible, smooth, Fréchet $G$-module. We can now apply the methods of 4.2.2 and 4.2.3 in [Wa3, pp. 112–113] to see that the intertwining operator $T_k : N(W) \to I((U(\mathfrak{n})/\mathfrak{n}^k U(\mathfrak{n}))\bigotimes\widetilde{W})$ given by Frobenius reciprocity (4.2.2 [op. cit]) is injective for $k$ sufficiently large. If $k, l$ are sufficiently large then Lemma 11.5.5 (or at least its proof) of [Wa4, p. 87] implies that the closure of Image$(T_k)$ is isomorphic with the closure of Image$(T_l)$. We will denote this smooth Fréchet $G$-module by $\overline{N(W)}$. This the canonical completion (globalization) as in the work of Casselman and the author in this context. The arguments in Chapters 10 and 11 of [Wa4] apply. Using the fact that $W \to N(W)$ defines an exact functor from the category of finite dimensional $(\mathfrak{q}, K)$-modules unitary as $K$-modules to the category of $(\mathfrak{g}, K)$-modules we have delineated we have

**Theorem 17.** *The functor $W \to \overline{N(W)}$ defines an exact functor from the category of all finite dimensional $(\mathfrak{q}, K)$-modules that are unitary as $K$-modules to the category of smooth Fréchet representations of $G$. Furthermore, if $N(W)$ is the underlying $(\mathfrak{g}, K)$-module of a (strongly continuous) representation $(\pi, H)$ of $G$ ($H$ a Hilbert space). Then $\overline{N(W)}$ is equivalent as a smooth Fréchet module with the representation of $G$ on the $C^\infty$ vectors of $H$.*

We are now ready to state and prove our main result in this context.

**Theorem 18.** *Let $(\pi, H)$ be a unitary representation of $G$ such that $H_K \cong N(W)$ for an appropriate irreducible unitary representation, $W$, of $K$. If $\psi$ is a unitary character of $N$ that is not $0$ or positive then $Wh_\psi^\infty(H) = 0$. If $\psi$ is positive and generic then $Wh_\psi^\infty(H)$ is isomorphic as an $M_\psi$ with the restriction of the contragredient $M$-module to $\widetilde{W}$.*

*Proof.* We note that if $F$ is a finite dimensional $G$-module and if $W$ is a unitary finite dimensional unitary representation of $K$ then $W \bigotimes F$ is a finite dimensional $K$-representation that admits an invariant Hilbert space structure. We have $N(W \bigotimes F) \cong N(W) \bigotimes F$. This implies that $\overline{N(W \bigotimes F)} \cong \overline{N(W)} \bigotimes F$. Let $C$ denote the center of $K$. Then $C$ is isomorphic with the additive group of $\mathbb{R}$. Let $J \in Lie(C)$ be such that $\exp(\mathbb{R}J) = C$ and $ad(J)$ acts on $\mathfrak{p}^+$ by multiplication by $i$. If $\lambda \in \mathbb{R}$ let $\chi_\lambda$ be the unitary character given by $\chi_\lambda(\exp(tJ)) = e^{i\lambda t}$. If $W$ is a finite dimensional, irreducible, unitary representation of $K$ then there exists $\lambda = \lambda(W)$ such that the action of $C$ on $W$ is given by multiplication by $\chi_\lambda$. We also denote by $W_0$ the unitary $K$-representation $W$ restricted to the commutator subgroup of $K$. If $\lambda(W) < c(W_0)$ then Harish-Chandra's criterion implies that $N(W)$ is the underlying $(\mathfrak{g}, K)$-module of an element of the relative discrete series and thus Theorem 5 in section 3 implies the result in this case.

In general, the vanishing assertion of the theorem follows from Lemma 1 in section 3 as in the beginning of the proof of Theorem 5. Thus to complete the proof we must show that if $\psi$ is a generic, positive unitary character of $N$ and that if $W$ is an irreducible unitary representation of $K$ then in the notation of the previous section $W_\psi(\overline{N(W)}')$ is isomorphic as an $M_\psi$ module with the restriction of the contragredient $M$-module to $\widetilde{W}$. (Here $\overline{N(W)}'$ denotes the space of all continuous linear functionals on $\overline{N(W)}$). We will denote by $\widehat{W}$ the contragredient $M$-module to $\widetilde{W}$ restricted to $M_\psi$. We observe that Corollary 1 in section 3 implies that $W_\psi(\overline{N(W)}')$ is equivalent to a submodule of $\widehat{W}$. We will now follow the line of argument in the proof of Theorem 7.2 in [Wa5].

Let $F$ be a finite dimensional irreducible representation of $G$ such that $F^{\mathfrak{p}^+}$ is one dimensional and the action if $C$ is given by $\chi_\mu$ with $\mu > 0$. It is well known that such a module, $F$, exists. Then as a $(\mathfrak{q}, K)$-module we have $0 = F_0 \subset F_1 \subset \cdots \subset F_d = F$ with $F_i/F_{i-1}$ irreducible. We may assume that $F_1 = F^{\mathfrak{p}^+}$. We assume that if $W_0$ is an irreducible representation of the commutator group, $[K, K]$, of $K$ and that if $W$ restricted to $[K, K]$ agrees with $W_0$ and satisfies $\lambda(W) < s$ then $W_\psi(\overline{N(W)}')$ is isomorphic as an $M_\psi$-module with $\widehat{W}$. We will show that if $\lambda(W) < s$ then $W_\psi(\overline{N(W \bigotimes F^{\mathfrak{p}^+})}')$ is isomorphic with $\widehat{W \bigotimes F^{\mathfrak{p}^+}}$ as an $M_\psi$-module. Since the assumption is true for $\lambda(W) < c(W_0)$ it will imply the result for $\lambda(W) < c(W_0) + \mu$, then for $\lambda(W) < c(W_0) + 2\mu, \ldots$ Hence it will imply the result. Also by the above we need only show that $\dim W_\psi(\overline{N(W \bigotimes F^{\mathfrak{p}^+})}') = \dim W$. Then we can apply the results in the previous section to see that

$$\overline{N(W)}'[\psi] \bigotimes F^* \cong \overline{N(W \bigotimes F)}'[\psi]$$

and

$$W_\psi(\overline{N(W)}' \bigotimes F^*) \cong W_\psi(\overline{N(W \bigotimes F)}').$$

If we apply the analog of Theorem 16 of the previous section to this context we see that $W_\psi(\overline{N(W \bigotimes F)}')$ is equivalent as an $M_\psi$-module with $\widehat{W} \bigotimes F^*$. Theorem

18 implies that we have a $G$-module filtration

$$0 \subset \overline{N(W \bigotimes F_1)} \subset \overline{N(W \bigotimes F_2)} \subset \cdots \subset \overline{N(W \bigotimes F_d)} = \overline{N(W)} \bigotimes F$$

with each of the spaces closed in the largest and

$$\overline{N(W \bigotimes F_i)} / \overline{N(W \bigotimes F_{i-1})} \cong \overline{N(W \bigotimes F_i / F_{i-1})}$$

as smooth Fréchet modules. Now we observe that

$$W_\psi(\overline{N(W \bigotimes F)}')_{|\overline{N(W \bigotimes F_1)}} \subset W_\psi(\overline{N(W \bigotimes F_1)}').$$

Thus

$$\dim W_\psi(\overline{N(W \bigotimes F)}')_{|\overline{N(W \bigotimes F_1)}} \leq \dim W.$$

Set $X_j = \{\lambda \in W_\psi(\overline{N(W \bigotimes F)}') | \lambda_{|\overline{N(W \bigotimes F_l)}} = 0\}$. Then

$$\dim X_1 \geq \dim W \dim F - \dim W.$$

Now $X_1$ restricted to $\overline{N(W \bigotimes F_2)}$ pushes down to a subspace of $W_\psi(\overline{N(W \bigotimes F_2/F_1)}')$ whose dimension is at most $\dim W \dim(F_2/F_1)$. Thus

$$\dim X_2 \geq \dim W \dim F - \dim W \dim F_2.$$

Continuing in this way we eventually find that $\dim X_{d-1} \geq \dim W \dim F - \dim W \dim F_{d-1}$. But $X_{d-1}$ is a subspace of $W_\psi(\overline{N(W \bigotimes F/F_{d-1})}')$ whose dimension is at most $\dim W \dim(F/F_{d-1})$. Thus the inequality is an equality. We now have the exact sequence

$$0 \to X_{d-1} \to W_\psi(\overline{N(W \bigotimes F)}') \to W_\psi(\overline{N(W \bigotimes F_{d-1})}')$$

with the last arrow given by restriction. This now implies that

$$\dim W_\psi(\overline{N(W \bigotimes F_{d-1})}') \geq \dim W \dim F_{d-1}$$

so we have equality. We now do the same argument with $W_\psi(\overline{N(W \bigotimes F_{d-1})}')$. Continuing in this way we eventually have $\dim W_\psi(\overline{N(W \bigotimes F_1)}') \geq \dim W$. This implies the result. □

## References


[Bour] N. Bourbaki, *Groups et Algébres de Lie*, Chapitres 4, 5, et 6, Hermann, Paris, 1968.

[Ca] W. Casselman, Canonical extensions of Harish–Chandra modules, *Can. J. Math.* **41** (1989), 315–438.

[E-W] T. Enright and N. Wallach, Embeddings of unitary highest weight representations and generalized Dirac operators, *Math. Ann.* **307** (1997), no. 4, 627–646.

[G-W] R. Goodman and N. Wallach, *Representations and invariants of the classical groups*, Encyclopedia of Mathematics and its Applications, 68. Cambridge University Press, Cambridge, 1998.

[Gr-D] K. Gross and H. Ding, Operator-valued Bessel functions on Jordan algebras, *J. Reine Angew. Math.* **435** (1993), 157–196.

[Gr-W] B. Gross and N. Wallach, On quaternionic discrete series representations, and their continuations, *J. Reine Angew. Math.* **481** (1996),73–123.

[H-C] Harish-Chandra, Representations of semisimple Lie groups VI, *Amer. J. Math.* **78** (1956), 564–628.

[He-S] H. Hecht and W. Schmid, A proof of Blattner's conjecture, *Inventiones math.* **31** (1975), 129–154.

[He] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, New York, 1978.

[J] H. Jacquet, Fonctiones de Whittaker associees aux groupes de Chevalley, *Bull. Soc. Math. France* **95** (1967), 243–309.

[J-W] K. Johnson and N. Wallach, Composition series and intertwining operators for the spherical principal series, I, *Trans. Amer. Math. Soc.* **229** (1977), 137–173.

[K-V] A. Knapp and D. Vogan, *Cohomological Induction and Unitary Representations*, Princeton University Press, Princeton, 1995.

[K-W] A. Knapp and N. Wallach, Szegő kernels associated with discrete series, *Invent. Math.* **34** (1976), no. 3, 163–200.

[Ko-V] Johan A. C. Kolk and V. S. Varadarajan, On the transversal symbol of vectorial distributions and some applications to harmonic analysis, *Indag. Mathem. N.S.* **7** (1996), 67–96.

[Ko1] B. Kostant, On Whittaker vectors and representation theory, *Inventiones math.* **48** (1978), 101–184.

[Ko2] B. Kostant, On Laguerre polynomials, Bessel functions, Hankel transform and a series in the unitary dual of the simply-connected covering group of $SL(2,\mathbb{R})$, *Represent. Theory* **4** (2000), 181–224.

[Wa1] Nolan R. Wallach, On the Selberg trace formula in the case of compact quotient, *Bull. Amer. Math. Soc.* **82** (1976), no. 2, 171–195.

[Wa2] Nolan R. Wallach, The asymptotic behavior of holomorphic representations. Harmonic analysis on Lie groups and symmetric spaces (Kleebach, 1983), *Mém. Soc. Math. France (N.S.)* No. 15, (1984), 291–305.

[Wa3] Nolan R. Wallach, *Real reductive groups. I*, Pure and Applied Mathematics, 132. Academic Press, Inc., Boston, MA, 1988.

[Wa4] Nolan R. Wallach, *Real reductive groups. II*, Pure and Applied Mathematics, 132-II. Academic Press, Inc., Boston, MA, 1992.

[Wa5] Nolan R. Wallach, Lie algebra cohomology and holomorphic continuation of generalized Jacquet integrals, in: *Representations of Lie groups, Kyoto, Hiroshima*, 1986, 123–151, Adv. Stud. Pure Math., 14, Academic Press, Boston, MA, 1988.


N. R. Wallach
University of California
Department of Mathematics
San Diego, La Jolla, CA 92093
USA
e-mail: nwallach@ucsd.edu

(Received: October 24, 2000)