Vol. 58, No. 3 DUKE MATHEMATICAL JOURNAL © June 1989

VARIATIONS OF HODGE STRUCTURE OF MAXIMAL DIMENSION

JAMES A. CARLSON, AZNIF KASPARIAN, AND DOMINGO TOLEDO

1. Introduction. A variation of Hodge structure is a holomorphic map with values in a Griffiths period domain which satisfies the differential equation

$$\partial F^p/\partial z_i \subset F^{p-1}. \tag{1.1}$$

The purpose of this paper is to give a general (and sharp) bound on the rank of such mappings. That a bound exists is clear from general principles. Equation (1.1) defines a subbundle T^h of the holomorphic tangent bundle of the period domain to which the image of a variation f is tangent, and its fiber dimension gives a first bound on the rank of f [8]. In general, however, the distribution defined by the horizontal tangent bundle T^h is nonintegrable, so that additional restrictions must hold. This is the case whenever D is not of hermitian type. In the simplest case (weight two with $h^{2,0} > 1$) one has the result of [1, 5]:

$$\text{rank } df \leqslant \tfrac{1}{2} \dim T^h, \tag{1.2}$$

or, more explicitly,

$$\text{rank } df \leqslant \tfrac{1}{2} h^{2,0} h^{1,1}. \tag{1.3}$$

The general bound is similar to this: it is given by a piecewise quadratic function of the Hodge numbers for domains of fixed Lie type.

To give a precise statement, fix a period domain D which classifies structures of weight w, let h^q stand for $h^{p,q}$, and set

$$\begin{aligned}
m &= [w/2] \\
m^* &= [(w-1)/2] \\
d^i &= h^i h^{i+1} \quad \text{for } i < m^* \\
d^{m^*} &= \tfrac{1}{2} h^{m^*}(h^{m^*} + 1) \quad \text{for } w \text{ odd (Type } C) \\
d^{m^*} &= h^{m^*}[h^{m^*+1}/2] + \varepsilon \text{ for } w \text{ even, (Types } B, D),
\end{aligned} \tag{1.4}$$

where $\varepsilon = 0$ if $h^{m,m} = h^{m^*+1}$ is even (Type D), $\varepsilon = 1$ if $h^{m,m}$ is odd (type B), and where $h^{m^*} \neq 1$. When $h^{m^*} = 1$, set $d^{m^*} = h^{m^*+1}$.

Received April 28, 1988. Research partially supported by the National Science Foundation and the Max Planck Institut für Mathematik.

Define

(1.5) $$q(L, \bar{h}) = \max\left\{\sum_{j \in I} d^j \mid j \leqslant m^*, j \in I \Rightarrow j+1 \notin I\right\},$$

where $\bar{h}$ is the 'vector' of Hodge numbers $h^{p,q}$, and where the formula for d^{m^*} depends on the Lie type L. Then we have:

THEOREM 1.6. *Let $f: X \to \Gamma\backslash D$ be a period mapping. Then rank $df \leqslant q(L, \bar{h})$.*

The theorem implies the bound (1.3) and in fact sharpens it for the case of $h^{1,1}$ odd:

$$\text{rank } df \leqslant \tfrac{1}{2}h^{2,0}(h^{1,1} - 1) + 1.$$

Since the proof relies on local arguments, the natures of X and Γ are irrelevant. X can be a polydisk and Γ may consist of the identity alone. The bound is sharp for mappings of polydisks and, except for one case, is also sharp for mappings of quasi-projective varieties (see §7).

Let us now sketch the proof. Denote by $\mathfrak{g}$ the complexified Lie algebra of infinitesimal isometries of D, fix a reference point $H \in D$, and define subspaces

(1.7) $$\mathfrak{g}^{-p,p} = \{\phi \in \mathfrak{g}_{\mathbf{C}} \mid \phi(H^{a,b}) \subset H^{a-p,b+p} \text{ for all } (a,b)\}$$

to obtain a Hodge structure [10]. Then the horizontality condition becomes $df(T^{1,0}) \subset \mathfrak{g}^{-1,1}$, where $T^{1,0}$ is the holomorphic tangent space, so that rank $df \leqslant \dim \mathfrak{g}^{-1,1}$. The integrability condition can then be expressed by saying that the image $\mathfrak{a} = df(T^{1,0})$ is abelian. Since $\mathfrak{g}^{-1,1}$ is generally nonabelian, this is a nontrivial restriction. Thus, if we define

(1.8) $$a(\mathfrak{g}^{-1,1}) = \max\{\dim \mathfrak{a} \mid \mathfrak{a} \subset \mathfrak{g}^{-1,1} \text{ is abelian}\},$$

then

(1.9) $$\text{rank } df \leqslant a(\mathfrak{g}^{-1,1}).$$

To compute this quantity, we shall construct a root space decomposition of $\mathfrak{g}$ which refines the Hodge decomposition (1.7). Next, we apply an argument of Malcev to show that to each abelian $\mathfrak{a} \subset \mathfrak{g}^{-1,1}$ there is an associated vector space $\lambda(\mathfrak{a}) \subset \mathfrak{g}^{-1,1}$, also abelian, satisfying $\dim \mathfrak{a} = \dim \lambda(\mathfrak{a})$, which splits as a direct sum of positive root spaces. It therefore suffices to bound the dimension of abelian subspaces of $\mathfrak{g}^{-1,1}$ which are spanned by a set of root vectors $\{r_\alpha \mid \alpha \in C\}$. The set of roots C is *commutative*, meaning that

$$\text{for all } \alpha, \beta \in C, \quad \alpha + \beta \text{ is not a root}.$$

Therefore

(1.10) $$a(\mathfrak{g}^{-1,1}) = \max\{\text{card } C \mid C \text{ is a commutative set of roots of type } (-1, 1)\}.$$

By the *type* of a root we shall mean the Hodge type of the associated root vector. The problem is therefore reduced to a purely combinatorial one. Sharpness is demonstrated by constructing suitable examples.

The root decomposition is calculated from a description of the Lie algebra in terms of block matrices. From this description we also obtain a distinguished system of simple roots, as in the figure below.

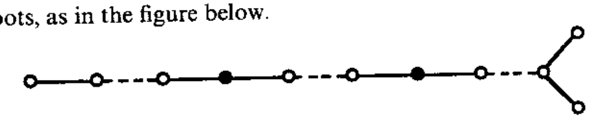

FIGURE 1

The white nodes correspond to compact simple roots, the black to noncompact ones. The former are of type (0, 0), and for gap-free period domains the latter are of type $(-1, 1)$. Here a gap means an integer p such that $h^{p+1,q-1} \neq 0$, and $h^{p,q} = 0$, with p not exceeding half the weight. For gap-free domains there are precisely $[l(l+1)/2]$ noncompact simple roots, where l is the level of the Hodge structure (if $[a, b]$ be the smallest interval such that $H^{p,q} = 0$ for $p \notin [a, b]$, then $l = b - a$; if there are no gaps, then $l = w$).

As a corollary of the preceding description of the simple roots one has the following:

THEOREM 1.11 *Let D be a gap-free period domain. Then any two points of D are accessible through the horizontal distribution.*

By accessibility we meant that any two points can be joined by a piecewise holomorphic horizontal curve. The proof is based on a theorem of Chow [7] which asserts that accessibility holds for a distribution if the ambient tangent space is generated under Lie bracket by vector fields belonging to the distribution. In the case at hand the holomorphic tangent space may be identified with $\mathfrak{g}^-$, the span of the root vectors of type $(p, -p)$, where p is negative. From the fact that all simple roots are of type (0, 0) or $(-1, 1)$, one dedutes that $\mathfrak{g}^{-1,1}$ generates $\mathfrak{g}^-$, so that Chow's theorem applies.

The authors would like to thank Nathan Jacobson for bringing Malcev's article to their attention, and would like to thank Dan Burns and Eduardo Cattani for useful conversations.

2. Abelian subspaces. In the introduction we claimed that the image under df of a tangent space to a variation of Hodge structure can be identified with an abelian subspace $\mathfrak{a}$ of $\mathfrak{g}^{-1,1}$, with $\mathfrak{a}$ abelian. To see this, consider the compact dual $\check{D}$ of D. If G is the (transitive) isometry group of D, then the associated complex group $G_{\mathbf{C}}$

acts transitively on $\check{D}$. Fix a reference Hodge structure $H_0 \in D$, let V and B be the corresponding isotropy subgroups of G and $G_{\mathbb{C}}$, and let $\mathfrak{v}$ and $\mathfrak{b}$ be the associated Lie algebras. Then

$$\mathfrak{v} \otimes \mathbb{C} = \mathfrak{g}^{0,0},$$

$$\mathfrak{b} = F^0\mathfrak{g} = \sum_{p \geqslant 0} \mathfrak{g}^{p,-p},$$

and if we set

$$\mathfrak{g}^- = \overline{F^1}\mathfrak{g} = \sum_{p<0} \mathfrak{g}^{p,-p},$$

then the Hodge-theoretic splitting

$$\mathfrak{g} = F^0 \oplus \overline{F^1} = \mathfrak{b} \oplus \mathfrak{g}^-$$

is a decomposition into subalgebras. We shall let G^- denote the unipotent group corresponding to g^-.

The holomorphic tangent bundle of $\check{D}$ is homogeneous, given by the formula

$$T_{\check{D}} = G_{\mathbb{C}} \times_B \mathfrak{g}/\mathfrak{b}.$$

Since the Lie bracket is compatible with the Hodge decomposition,

$$[\mathfrak{g}^{p,-p}, \mathfrak{g}^{q,-q}] \subset \mathfrak{g}^{p+q,-p-q}, \tag{2.1}$$

the space $\mathfrak{g}^{-1,1} \oplus \mathfrak{b}$ is Ad_B-stable, and so the homogeneous bundle

$$T^h_{\check{D}} = G_{\mathbb{C}} \times_B (\mathfrak{g}^{-1,1} + \mathfrak{b})/\mathfrak{b}$$

is defined. This is Griffiths' horizontal tangent bundle [8].

Consider now the map $\exp\colon \mathfrak{g}^- \to \check{D}$ given by sending ξ in $\mathfrak{g}^-$ to $e^{\xi}F_0^*$ in $\check{D}$. Restricted to a small enough ball $\mathfrak{g}^-(\varepsilon)$ about the origin, this gives a univalent parametrization of a neighborhood $U(\varepsilon)$ in D of the reference structure. Consequently the map $e^{\xi}F_0^* \mapsto e^{\xi}$ gives a lifting $n\colon U(\varepsilon) \to G^-$, and the associated Maurer-Cartan form $\omega = n^{-1}\,dn$ takes values in $\mathfrak{g}^- \otimes \mathscr{E}^1$, where $\mathscr{E}^p$ denotes the space of smooth p-forms. Now let $f\colon V \to D$ be the inclusion for a variation of Hodge structure, and set

$$\mathfrak{a} = f^*\omega(T_oV).$$

We shall call this the *canonical lift* of T_o to $\mathfrak{g}$. Then we have

PROPOSITION 2.2. *Let V be a variation of Hodge structure, o a point of V, and $\mathfrak{a}$ the canonical lift of T_oV to $\mathfrak{g}$. Then $\mathfrak{a}$ is an abelian subspace of $\mathfrak{g}^{-1,1}$.*

Proof. The horizontality condition can be written as

$$f^*\omega(X) \in \mathfrak{g}^{-1,1} \otimes \mathscr{E}^0,$$

for any vector field X on $U \cap V$, so that $\mathfrak{a} \subset \mathfrak{g}^{-1,1}$. To show that $\mathfrak{a}$ is abelian, pull the integrability condition $d\omega - \omega \wedge \omega = 0$ back to U along f and evaluate on a pair of vector fields (X, Y) to get

$$Xf^*\omega(Y) - Yf^*\omega(X) - f^*\omega([X, Y]) - [f^*\omega(X), f^*\omega(Y)] = 0.$$

Since the first three terms are in $\mathfrak{g}^{-1,1} \otimes \mathscr{E}^0$, so must be the last. Therefore $[\mathfrak{a}, \mathfrak{a}] \subset \mathfrak{g}^{-1,1}$. But the relations $\mathfrak{a} \subset \mathfrak{g}^{-1,1}$, $[\mathfrak{g}^{-1,1}, \mathfrak{g}^{-1,1}] \subset \mathfrak{g}^{-2,2}$ force $[\mathfrak{a}, \mathfrak{a}] = 0$, since $\mathfrak{g}^{-1,1}$ and $\mathfrak{g}^{-2,2}$ are complementary. This completes the proof.

Proposition 2.2 may be thought of as lifting to $\mathfrak{g}$ of the corresponding commutativity condition for infinitesimal variations of Hodge structure [3, §2 formula (iv)].

3. Malcev's theorem. According to the last proposition, the problem of bounding the dimension of a variation of Hodge structure leads to that of bounding the dimension of a space of commuting matrices. The first result of this kind was obtained in 1905 [14] by I. Schur: a commutative space of $\mathfrak{gl}_n$ is dimension at most $1 + [n/2]$. In 1945 Malcev generalized this result to the case of abelian subalgebras of an arbitrary complex simple Lie algebra [12]. His method was quite different from that of Schur, and it is central to our treatment of the present problem.

The fundamental difficulty is that a general abelian subspace $\mathfrak{a}$ is invariant for *no* Cartan subalgebra of $\mathfrak{g}$, and therefore cannot be written as a direct sum of root spaces. To construct such examples it suffices to take spaces of commuting matrices for which the minimum rank of an element is large. Since root vectors have small rank, such a space cannot be a sum of root spaces. However, we have the following:

THEOREM 3.1 (MALCEV). *Let $\mathfrak{a}$ be an abelian subspace of a semisimple Lie algebra $\mathfrak{g}$ which consists entirely of nilpotent elements. Then there is a Cartan subalgebra $\mathfrak{h}$, a system Δ_+ of positive roots, and an auxiliary space $\lambda(\mathfrak{a})$, the space of leading roots, with the properties*

(a) *$\lambda(\mathfrak{a})$ is abelian,*
(b) *$\dim \lambda(\mathfrak{a}) = \dim \mathfrak{a}$, and*
(c) *$\lambda(\mathfrak{a})$ is a direct sum of positive root spaces.*

We may therefore write

$$\lambda(\mathfrak{a}) = \sum_{\alpha \in C} \mathfrak{g}^{\alpha}$$

for a commutative subset C of Δ_+.

Proof. The argument is quite simple. Let $\mathfrak{b}$ be a Borel subalgebra containing $\mathfrak{a}$, and let $\mathfrak{b} = \mathfrak{h} + \mathfrak{n}$, where $\mathfrak{h}$ is a Cartan subalgebra and $\mathfrak{n}$ is the nilradical of $\mathfrak{b}$. Write $X \in \mathfrak{a}$ as $H + N$, where $H \in \mathfrak{h}$ and $N \in \mathfrak{n}$. Since this is the decomposition into semisimple and nilpotent parts of X, $H = 0$. Therefore $\mathfrak{a} \subset \mathfrak{n}$. Now choose a system of positive roots $\Delta_+ = \{\alpha_1, \ldots, \alpha_n\}$ so that

$$\mathfrak{n} = \sum_{\alpha \in \Delta_+} \mathfrak{g}^\alpha,$$

and choose an ordering on Δ, compatible with addition, so that its set of positive roots is Δ_+. Let X_i be the root vector associated with α_i, let $\{U_1, \ldots U_k\}$ be a basis for $\mathfrak{a}$, and let $C = (C_{ij})$ be the matrix which expresses the U_j in terms of the X_i:

$$U_j = \sum_i C_{ij} X_i .$$

Apply Gaussian elimination to bring the matrix C to echelon form. Let $\lambda(U_j)$ be the first X_i in the ordering which appears in the expression of U_j with nonzero coefficient, i.e., let it be the *leading root* in the expression of U_j in terms of positive roots. Now consider the Lie bracket of two basis vectors, $[U_i, U_j]$, which we may write as a sum of positive root vectors X_k. The first potentially nonzero term is $[\lambda(U_i), \lambda(U_j)]$, and this term is less than the other root vectors in the expression under study. But $[U_i, U_j] = 0$, so that $[\lambda(U_i), \lambda(U_j)] = 0$ as well. Therefore the space $\lambda(\mathfrak{a}) = \text{span}\{\lambda(U_1), \ldots, \lambda(U_k)\}$ fulfills the requirements of the Theorem.

Let us now consider Malcev's theorem in the context of Hodge theory. According to the results of §5, $\mathfrak{g}$ has a Cartan subalgebra $\mathfrak{h}$ contained in $\mathfrak{g}^{0,0}$. It then follows (by relation (2.1)) that the $\mathfrak{g}^{-p,p}$ are $\mathfrak{h}$-stable, so that the root decomposition of $\mathfrak{g}$ relative to $\mathfrak{h}$ refines the Hodge decomposition. Moreover, there is an ordering of the roots compatible with addition such that $\mathfrak{g}^-$ is a sum of positive root spaces. Such an ordered system of roots will be called *compatible with the Hodge decomposition.* The next result shows that the construction $\mathfrak{a} \mapsto \lambda(\mathfrak{a})$ respects this structure:

THEOREM 3.2 *Let $\mathfrak{a}$ be an abelian subspace of $\mathfrak{g}^{-1,1}$ and let Δ be an ordered system of roots compatible with the Hodge decomposition. Then $\lambda(\mathfrak{a})$ exists and is a subspace of $\mathfrak{g}^{-1,1}$.*

Proof. In the light of the discussion above, it suffices to show that $\mathfrak{a}$ is a space of nilpotent transformations, so that Malcev's argument applies. But $\mathfrak{a} \subset \mathfrak{g}^-$, which (by (2.1)) is a space of nilpotents.

4. Block decompositions. Fix a Hodge structure $H \in D$ and consider the subspaces

$$\mathfrak{g}_{(j)}^{-p,p} = \{\phi \in \mathfrak{g}^{-p,p} \mid \phi(H^{w-j,j}) = 0 \Rightarrow \phi = 0\} \tag{4.1}$$

for $p > 0$, and set

$$\mathfrak{g}_{(j)}^{p,-p} = \overline{\mathfrak{g}_{(j)}^{-p,p}} = \{\phi \in \mathfrak{g}^{-p,p} \mid \phi(H^{j,w-j}) = 0 \Rightarrow \phi = 0\}. \tag{4.2}$$

These give a refinement of the Hodge decomposition:

$$\mathfrak{g}^{-p,p} = \bigoplus_{j=0}^{n(p)} \mathfrak{g}_{(j)}^{-p,p}, \tag{4.3}$$

where

$$n(p) = [(w-p)/2], \tag{4.4}$$

and where the restriction on the range of j comes from the antisymmetry condition on elements of $\mathfrak{g}$. Let $\mathfrak{g}_{(j)} = \oplus_p \mathfrak{g}_{(j)}^{-p,p}$, and note that these give a real decomposition of $\mathfrak{g}$. The purpose of this section is to give an explicit description of the $\mathfrak{g}_{(j)}^{-p,p}$'s in terms of block matrices. From this one easily obtains the required root structure.

To begin, define an *adapted frame* of H to be a set of bases

$$B^q = \{B_j^q \mid j = 1, \ldots h^q\}$$

such that B^q spans $H^{p,q}$. The matrix of an endomorphism of $H_{\mathbb{C}}$ relative to such a frame has a natural decomposition $\sum A_{ij}$ where A_{ij} denotes the block in position (i, j) and the indices range from 0 to the weight w. A_{ij} represents a homomorphism from $H^{w-j,j}$ to $H^{w-i,i}$, and so is a block of size $h^i h^j$.

Define the *support* of a matrix A to be the set of indices $|A| = \{(i,j) | A_{ij} \neq 0\}$. Define sets $\mathscr{D}_p = \{(i, i+p)\}$ (the p-th diagonal), $\mathscr{D}' = \{(i, w-i)\}$ (the principal antidiagonal), and $\mathscr{M} = \{i,j) | i = w/2 \text{ or } j = w/2\}$ (the middle). Then matrices representing elements of $\mathfrak{g}^{-p,p}$ are supported in the $-p$-th diagonal $\mathscr{D}_{-p}$, and elements of $\mathfrak{g}_{(j)}^{-p,p}$ are supported in a pair of blocks j units from the extremities of the $-p$-th diagonal, each lying in positions symmetric relative to the principal antidiagonal. If $2j + p = w$ then this pair is degenerate, i.e., is supported in the principal antidiagonal $\mathscr{D}'$.

Denote by $C[ij]$ a matrix whose only nonzero block is C in the (i, j) position, and abbreviate $C[ii]$ to $C[i]$. The algebra of blocks follows the rules ${}^t(C[ij]) = {}^tC[ji]$ and $A[ij]B[kl] = \delta_{jk} AB[il]$. In this notation a typical element of $\mathfrak{g}_{(j)}^{-p,p}$ for $p > 0$ takes the form

$$X^{(jp)} = A[j+p, j] + B[w-j, w-j-p] \tag{4.5}$$

and the polarization becomes

$$i^w S = \sum_{k=0}^{w} (-1)^k S_k[k, w-k].$$

Since every $X^{(jp)}$ is a $Y^{(kp)}$ with $j + k = w - p$, we may assume $0 \leqslant j \leqslant n(p) = [(w-p)/2]$ for $p = 0, \ldots, w$.

When the S_k are identity matrices, the adapted frame is a so-called *Hodge frame*. In the case of even weight we also use this term if the middle component $M = S_{w/2}$ takes one of the forms below.

Case of $h^{m,m} = 2t$ *even*:

$$M = \begin{pmatrix} 0_t & I_t \\ I_t & 0_t \end{pmatrix}$$

Case of $h^{m,m} = 2t + 1$ *odd*:

$$M = \begin{pmatrix} 0_t & I_t & 0 \\ I_t & 0_t & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

In both cases the subscripts indicates the size of the given (square) matrix, and $t = [h^{m,m}/2]$ where $m = [w/2]$.

(4.6) *Example.* Consider a Hodge structure of weight 4. The polarization in a Hodge frame takes the form

$$S = \begin{bmatrix} 0 & 0 & 0 & 0 & I \\ 0 & 0 & 0 & -I & 0 \\ 0 & 0 & M & 0 & 0 \\ 0 & -I & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Consider also a general element $X \in \mathfrak{g}^-$. The support condition and the S-antisymmetry imply that X is block-lower triangular and 'graded symmetric' with respect to the principal antidiagonal:

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ A & 0 & 0 & 0 & 0 \\ C & B & 0 & 0 & 0 \\ E & D^- & {}^tB & 0 & 0 \\ F^- & {}^tE & -{}^tC & {}^tA & 0 \end{bmatrix},$$

where the superscript minus indicates antisymmetry. In the notation just introduced, the Hodge components of X take the form

$$X^{-1,1} = (A[1,0] + {}^tA[4,3]) + (B[2,1] + {}^tB[3,2]) \in \mathfrak{g}_{(0)}^{-1,1} \oplus \mathfrak{g}_{(1)}^{-1,1}$$

$$X^{-2,2} = (C[2,0] - {}^tC[4,2]) + D^-[3,1] \in \mathfrak{g}_{(0)}^{-2,2} \oplus \mathfrak{g}_{(1)}^{-2,2}, \quad \text{etc.},$$

with $X^{-p,p}$ on the $-p$-th diagonal. For Hodge structures of odd weight, the blocks on the principal antidiagonal are symmetric rather than antisymmetric.

To determine the general form of an element of $\mathfrak{g}_{(j)}^{-p,p}$ relative to an adapted frame, compute the relation ${}^tX^{(jp)}S + SX^{(jp)} = 0$ to obtain

$$(4.7) \qquad {}^tAS_{j+p} + (-1)^p S_j B = 0.$$

When the A and B blocks are distinct, which is to say not on the principal antidiagonal, there are no restrictions on A, so that $\dim \mathfrak{g}_{(j)}^{-p,p} = h^j h^{j+p}$. In the contrary case, A satisfies the graded symmetry condition ${}^tA = (-1)^{p+1} S_j^{-1} A S_{j+p}$, from which one calculates the dimension of $\dim \mathfrak{g}_{(j)}^{-p,p}$:

$$\dim \mathfrak{g}_{(j)}^{-p,p} = h^j h^{j+p} \quad \text{for } 2j + p \neq w,$$

$$\dim \mathfrak{g}_{(j)}^{-p,p} = \tfrac{1}{2} h^j(h^j - 1) \quad \text{for } 2j + p = w \text{ and } p \text{ even},$$

$$\dim \mathfrak{g}_{(j)}^{-p,p} = \tfrac{1}{2} h^j(h^j + 1) \quad \text{for } 2j + p = w \text{ and } p \text{ odd}.$$

In particular, one obtains

$$\begin{aligned} \dim \mathfrak{g}^{-1,1} &= \sum_{j=0}^{n_1} \dim \mathfrak{g}_{(j)}^{-1,1} \\ &= h^0 h^1 + h^1 h^2 + \cdots h^{m-1} h^m \quad (w \text{ even}) \\ &= h^0 h^1 + h^1 h^2 + \cdots h^{m-1} h^m + \tfrac{1}{2} h^m (h^m + 1) \quad (w \text{ odd}), \end{aligned}$$

where $m = [w/2]$.

5. Roots and the Hodge decomposition. A Hodge frame determines both a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ contained in $\mathfrak{g}^{0,0}$ and a distinguished basis of $\mathfrak{h}$. To see this, fix such a frame and define, for $j \neq m = [w/2]$, the elements

$$(5.1) \qquad E_k(j) \underset{\text{def}}{=} E_k[j] - E_k[w-j] \in \mathfrak{g}_{(j)}^{0,0},$$

where E_k is a matrix with a 1 in position (k, k) and with no other nonzero entries. Note that $E_k(j) = -E_k(w-j)$. For w even, set

$$(5.2) \qquad E_k(m) \underset{\text{def}}{=} E_k[m] - E_{t+k}[m],$$

where $t = [h^{m,m}/2]$. Let $I_j = \{1, \ldots, h_j\}$ for $j < m$, let $I_m = \{1, \ldots, t\}$, and define

$$\mathfrak{h}_{(j)} = \operatorname{span}\{E_k(j) \mid k \in I_j\},$$

$$\mathfrak{h} = \sum_{j=0}^{m} \mathfrak{h}_{(j)}.$$

Then $\mathfrak{h}$ is a Cartan subalgebra and $\mathfrak{h}_{(j)} = \mathfrak{h} \cap \mathfrak{g}_{(j)}$.

Because $\mathfrak{h}$ is of type (0, 0), it preserves the Hodge decomposition under the adjoint representation, and so the $\mathfrak{g}^{-p,p}$ decompose into root spaces. Moreover, the multiplicative properties of block matrices imply that the $\mathfrak{g}_{(j)}^{-p,p}$ are stable under the adjoint action of the $\mathfrak{h}_{(k)}$, so that they decompose as well. Consequently there is a partition

$$\Delta = \bigcup_{j,p} \Delta_{(j)}^{-p,p},$$

such that

$$\mathfrak{g}_{(j)}^{-p,p} = \sum_{\alpha \in \Delta_{(j)}^{-p,p}} \mathfrak{g}^{\alpha}$$

for $p \neq 0$, and such that

$$\mathfrak{g}^{0,0} = \mathfrak{h} \oplus \sum_{\alpha \in \Delta_{(j)}^{0,0}} \mathfrak{g}^{\alpha}$$

for $p = 0$. Moreover, in the natural ordering, relative to which upper triangular blocks correspond to positive roots, the root spaces for $\mathfrak{g}^{-p,p}$ are negative:

$$\Delta^{-p,p} \subset \Delta_- \quad \text{for } p > 0.$$

We shall say that α in $\Delta_{(j)}^{-p,p}$ is of

Type I if $2j + p \neq w$ and $j + p \neq m$

(5.3) Type I' if $2j + p \neq w$ and $j + p = m$, with w even

Type II if $2j + p = w$.

Let $\{e_k(j)\}$ be the basis dual to $\{E_k(j)\}$. Then one has the following description of the roots:

THEOREM 5.4. *Let $\mathfrak{g}$ be the Lie algebra of a period domain. Then all type I roots are given by*

$$\Delta_{(j)}^{-p,p} = \{-e_a(j) + e_b(j+p) \mid a \in I_j,\ b \in I_{j+p}\},$$

where $p > 0$ and $j \leqslant [(w-p)/2]$. The remaining roots, of types I' and II, depend on the Lie type:

Lie Type C_n. *The remaining roots are of type II and take the form*

$$\Delta_{(j)}^{-p,p} = \{-e_a(j) - e_b(j) \mid a, b \in I_j\},$$

Lie Type D_n. *The remaining roots are of types I' and II and take the form*

$$\Delta_{(j)}^{-p,p} = \{-e_a(j) \pm e_b(m) \mid a \in I_j,\ b \in I_m\} \quad (type\ I')$$

$$\Delta_{(j)}^{-p,p} = \{-e_a(j) - e_b(j) \mid a, b \in I_j, \text{ where } a \neq b\} \quad (type\ II).$$

Lie Type B_n. *The remaining roots are of types I' and II and take the form*

$$\Delta_{(j)}^{-p,p} = \{-e_a(j) \pm e_b(m) \mid a \in I_j,\ b \in I_m\} \cup \{-e_a(j) \mid a \in I_j\} \quad (type\ I')$$

$$\Delta_{(j)}^{-p,p} = \{-e_a(j) - e_b(j) \mid a, b \in I_j, \text{ where } a \neq b\}, \quad (type\ II)$$

The complex conjugation operator of $\mathfrak{g}$ acts on the roots by change of sign (see [9, 3.6]).

REMARK 5.5. For $j \neq m$, $e_a(j) = -e_a(w-j)$, so that for $j + p > m$ one has

$$\Delta_{(j)}^{-p,p} = \{-e_a(j) - e_b(w-j-p) \mid a \in I_j,\ b \in I_{w-j-p}\},$$

for the type I roots.

Proofs. The necessary verifications are straightforward: we exhibit the root vectors and calculate the adjoint action of the $E_a(j)$. To this end let E_{ab} be the matrix with a 1 in position (a, b) and zeros elsewhere, and let $E_{ab}(jp)$ be the matrix whose block in position $(j, j+p)$ is E_{ab} and whose complementary block is determined by (4.7). These give a basis of $\mathfrak{g}_{(j)}^{-p,p}$ consisting of root vectors. Viewing the conditions (5.1) on roots as conditions on root vectors, we see that a vector of type I has support outside the principal antidiagonal $\mathscr{D}'$ and outside the middle row and column $\mathscr{M}$, a vector of type I' has support outside the principal antidiagonal but inside $\mathscr{M}$, and a vector of type II has support in the principal antidiagonal. Let us check that the roots are as claimed.

Roots of Type I. In this case $E_{ab}(jp)$ consists of a pair of blocks off the principal antidiagonal:

$$E_{ab}(jp) = E_{ab}[j+p, j] + (-1)^{p+1} E_{ab}[w-j, w-j-p].$$

Then

$$[E_k(j), E_{ab}(jp)] = -\delta_{kb} E_{ab}(jp),$$

$$[E_k(j+p), E_{ab}(jp)] = \delta_{ka} E_{ab}(jp),$$

and

$$[E_k(s), E_{ab}(jp)] = 0 \quad \text{for} \quad s \neq j, j+p,$$

so that

$$E_{ab}(jp) \text{ is a root vector with root } -e_a(j) + e_b(j+p).$$

It remains to consider the root vectors of types I' and II for each of the Lie types B, C, D.

Lie type C_n: Odd weight. The remaining root vectors are of type II:

$$E_{ab}(jp) = (E_{ab} + E_{ba})[j+p, j].$$

The only nontrivial action by elements in the Cartan subalgebra is by those in $\mathfrak{h}_{(j)}$, for which we obtain

$$[E_k(j), E_{ab}(jp)] = -(\delta_{ka} + \delta_{kb})E_{ab}(jp),$$

so that

$$E_{ab}(jp) \text{ is a root vector with root } -e_a(j) - e_b(j).$$

Lie Type D_n: Even weight, $h^{m,m}$ even. There are two remaining root types, I' and II. The type II root vectors are of the form

$$E_{ab}(jp) = (E_{ab} - E_{ba})[j+p, j],$$

and a calculation similar to that for Lie type C_n shows its root is also $-e_a(j) - e_b(j)$. However, since the block $E_{ab} - E_{ba}$ is antisymmetric instead of symmetric, there is no root vector when $a = b$, and hence no root in this case. The type I' root vectors are of the form

$$E_{ab}(jp) = E_{ab}[j+p, j] + (-1)^{p+1} M E_{ba}[w-j, w-j-p].$$

Since the basis elements for $\mathfrak{h}_{(m)}$ are slightly different (see (5.2)), one finds

$$[E_k(m), E_{ab}(jp)] = (\delta_{k,b} - \delta_{t+k,b})E_{ab}(jp).$$

Then for $b \in I_m = \{1, \ldots, t\}$, $E_{ab}(jp)$ has root $-e_a(j) + e_b(m)$ and $E_{a,b+m}(jp)$ has root $-e_a(j) - e_b(m)$, as required.

Lie Type B_n: Even weight, $h^{m,m}$ odd. The analysis is similar to that of the preceding case. However, one new kind of type I' root appears. Since the Cartan elements $E_k(m)$, $k = 1, \ldots, t$, act trivially on $E_{a,2t+1}(jp)$, these latter have $-e_a(j)$ as root.

6. Proof of the bound. We can now prove Theorem (1.6). According to the discussion of the introduction, it suffices to compute the quantity

$$\max\{\text{card } C \mid C \subset \mathfrak{g}^{-1,1} \text{ is a commutative set of roots}\}.$$

This we shall do using the root structure given by the results of the preceding section.

Type C_n (odd weight). The 'horizontal roots' are given by

$$\Delta_{(j)}^{-1,1} = \{-e_a(j) + e_b(j+1) \mid a \in I_j, b \in I_{j+1} \quad \text{for } j = 0, \ldots, m-1.\}$$

and

$$\Delta_{(m)}^{-1,1} = \{-e_a(j) - e_b(j) \mid a, b \in I_m\}.$$

Fix a commutative set of roots $C \subset \mathfrak{g}^{-1,1}$ and define the following: For $j = 0, \ldots, m-1$, set

$$I_j^- = \{a \in I_j \mid -e_s(j) + e_b(j+1) \in C \text{ for some } b\}.$$

For $j = 1, \ldots, m$, set

$$I_j^+ = \{b \in I_j \mid -e_a(j-1) + e_b(j) \in C \text{ for some } a\},$$

and set $I_0^+ = \varnothing$. Finally, set

$$I_m^- = \{a \in I_m \mid -e_a(m) - e_b(m) \in C \text{ for some } b\}.$$

The condition that C be commutative is then

$$I_j^+ \cap I_j^- = \varnothing \quad \text{for } j = 0, \ldots, m.$$

(For $j = m$ we use remark (5.5)). By enlarging C if necessary we may assume that

$$I_j = I_j^+ \cup I_j^-$$

for $j = 0, \ldots, m$. Again enlarging C if necessary, we have a bijection

$$C \leftrightarrow I_0^- \times I_1^+ \cup I_1^- \times I_2^+ \cup \cdots \cup I_{m-1}^- \times I_m^+ \cup Sym^2(I_m^-),$$

where $Sym^2(A)$ denotes the set of unordered pairs of elements of A. Let $x_j = \text{card } I_j^-$, and set

$$f(x_1, \ldots, x_m) = h^0(h^1 - x_1) + \sum_{i=1}^{m-1} x_i(h^{i+1} - x_{i+1}) + \frac{1}{2}x_m(x_m + 1).$$

Then $f(x)$ is the cardinality of C for given values of the parameters, and so

$$a = \max\{f(x) | x \in J_1 \times \cdots \times J_m, x \text{ has integer entries}\},$$

where $J_k = [0, \ldots, h^j]$. Let us consider the maximum value problem for all values of x in the given rectangle. The function f is quadratic with never negative semidefinite Hessian, and so has an interior maximum only if it is negative definite, which it is not (consider x_m). Therefore the maximum must occur on one of the faces. But the restriction to a face is a function of similar form: it is either (a) linear, or (b) quadratic with nondegenerate but nonnegative Hessian. Applying this argument repeatedly to faces, one finds that the maximum value must occur at a vertex, in particular, at a point with integer coordinates. From the form of f we conclude that $f(vertex)$ is a sum of nonconsecutive elements of the sequence

$$h^0h^1, h^1h^2, \ldots, h^{m-1}h^m, \tfrac{1}{2}h^m(h^m + 1).$$

Therefore *max card* C is given by the function a defined above, as claimed.

Type D_n *(even weight,* $h^{m,m}$ *even).* The set $\Delta_{(j)}^{-1,1}$ of horizontal roots is the same as in the preceding case for $j = 0, \ldots m-2$. For $j = m-1$ we have

$$\Delta_{(m-1)}^{-1,1} = \{-e_a(m-1) \pm e_b(m) | a \in I_{m-1}, b \in I_m\}.$$

The sets I_j are as before, except that $I_m = \{1, \ldots, t\}$, where $t = h^{m,m}/2$. Define sets I_j^- and I_j^+ as before for $j = 0, \ldots, m-1$, and set

$$I_m^{\pm} = \{b \in I_m | -e_a(m-1) \pm e_b(m) \in C \text{ for some } a\}.$$

Arguing as before, and enlarging C if necessary, we obtain a bijection

$$C \leftrightarrow I_0^- \times I_1^+ \cup I_1^- \times I_2^+ \cup \cdots \cup I_{m-2}^- \times I_{m-1}^+ \cup I_{m-1}^- \times I_m^+ \cup I_{m-1}^- \times I_m^-,$$

where the sets $I_j^{\pm}$ partition I_j. The cardinality of such a set is given by the function

$$f(x) = h^0(h^1 - x_1) + \sum_{i=1}^{m-2} x_i(h^{i+1} - x_{i+1}) + \frac{1}{2}x_{m-1}h^m,$$

since card $I_m = h^{m,m}/2$. Considerations of quadratic programming analogous to those of the preceding section show that the maximum occurs at a vertex, from which we obtain the required computation of *max card* C.

Type B_n *(even weight,* $h^{m,m}$ *odd).* The sets $\Delta_{(j)}^{-1,1}$ of horizontal roots in this case are the same as in the preceding one for $j = 0, \ldots m-2$. For $j = m-1$ we have

$$\Delta_{(m-1)}^{-1,1} = \{-e_a(m-1) \pm e_b(m) | a \in I_{m-1}, b \in I_m\} \cup \{-e_a(j) | a \in I_{m-1}\},$$

where as before $I_m = \{1, \ldots, t\}$, with $t = [h^{m,m}/2]$. Arguing as above, one finds

$$C \leftrightarrow I_0^- \times I_1^+ \cup I_1^- \times I_2^+ \cup \cdots \cup I_{m-2}^- \times I_{m-1}^+ \cup I_{m-1}^- \times I_m^- \cup I_{m-1}^- \times I_m^+ \cup \{a\},$$

where the last term is a singleton corresponding to the root $-e_a(m-1)$. The cardinality of such a set is given by:

$$f(x) = h^0(h^1 - x_1) + \sum_{i=1}^{m-2} x_i(h^{i+1} - x_{i+1}) + \frac{1}{2}x_{m-1}(h^m - 1) + 1,$$

and the usual quadratic programming argument completes the proof.

7. Sharpness. We shall now describe certain basic variations of Hodge structure from which variations of maximal dimension are constructed. Each of the basic variations is highly degenerate in the sense that all but one part of the Hodge filtration is constant. We shall denote this part by F^a, and we shall refer to the associated variation as $V(a)$. The variations of type II are classical, those of types I and III were introduced in [1], and that of type IV, which is the only one not parametrized by a hermitian symmetric space, is new.

Type I. $a > m + 2$. Let H_0 be a reference Hodge structure of type $\{(a, b), (a-1, b+1), (b, a), (b+1, a-1)\}$, where $a + b = w$. By 'type' we mean a subset of $\mathbb{Z} \times \mathbb{Z}$ such that $H^{p,q} = 0$ for $(p, q) \notin S$. Consider the set $V(a)$ of all subspaces F^a of $T = F_o^{a-1}$ which are of dimension $h^{a,b}$ and are positive for the indefinite Hermitian form

$$h(x, y) = S(x, \bar{y}).$$

Since F^{a-1} is constant, $dF^a/dt \subset F^{a-1}$ for any one-parameter family $F^a(t)$ in $V(a)$. This makes sense because if $U(t) = \text{span}\{u_1(t), \ldots, u_n(t)\}$ is a family of subspaces, then the subspace $dU/dt = \text{span}\{\dot{u}_1(t), \ldots, \dot{u}_n(t)\}$ makes sense modulo $U(t)$. Horizontality is therefore satisfied, so that $V(a)$ is a variation of Hodge structure.

Let G be the isometry group of the ambient period domain D, let $N(T)$ be the normalizer of T in G, let $Z(T)$ be the centralizer, and let $G(T)$ be the quotient group. Then $V(a)$ is a $G(T)$-orbit in D. Since one may identify $G(T)$ with $SU(p, q)$, where $p = h^{a,b}$ and $q = h^{a+1,b-1}$, the parameter space for the variation is identified with the Hermitian symmetric domain

$$B_{pq} = \{Z | Z \text{ a } p \times q \text{ matrix with } Z^*Z < I\},$$

where Z^* denotes the hermitian conjugate.

Type II. $a = m + 1$ and $w = 2m + 1$ odd. Let $\mathscr{H}_g$ be the Siegel upper half-space of genus g, and let $V(1)$ be the natural variation of weight one Hodge structures, with $h^{1,0} = g$. Let A_m be any Hodge structure of dimension one and type (m, m).

Then the map

$$H \in \mathscr{H}_g \mapsto H \otimes A_m$$

defines the required basic variation $V(m+1)$ of weight $w = 2m+1$.

Type III. $a = m+1$, $w = 2m$ even, and $h^{m,m}$ even. Consider first the case of $m = 1$. Choose a decomposition $H_{\mathbb{C}} = I \oplus \bar{I}$ with I isotropic and contained in F_0^1. Let $V(2)$ be the set of all h-positive subspaces F^2 of I of dimension $h^{2,0}$. Let $F^2(t)$ be a curve in $V(2)$ and let $\phi(t)$ be a vector-valued function with $\phi(t) \in F^2(t)$. Because of the relation

$$(7.1) \qquad F^2 \subset I = I^{\perp} \subset (F^2)^{\perp} = F^1,$$

one has $d\phi/dt \in I \subset F^1$, so that horizontality holds. Therefore $V(2)$ is a variation. By considerations similar to those of the first case, one sees that it arises as an $SU(p, q)$-orbit, where $p = h^{2,0}$ and $q = h^{1,1}/2$, and so can be identified with the Hermitian symmetric domain B_{pq}. Now define $V(m+1)$ by sending $H \in V(2)$ to $H \otimes A_{m-1}$.

Type IV. $a = m+1$, $w = 2m$ even, and $h^{m,m}$ odd. As in the preceding case, it suffices to define $V(2)$, since $V(m+1)$ can be obtained as a tensor product of $V(2)$ with a trivial variation. To begin the construction, fix a reference structure of the form $H_o = A \oplus B$, where

$$B = \operatorname{span}\{\alpha, \beta, \bar{\alpha}\},$$

where α has type (2, 0) with $\langle \alpha, \bar{\alpha} \rangle = 1$, and β is real and of type (1, 1) and $\langle \beta, \beta \rangle = -1$. Set

$$\alpha(t) = \alpha + t\beta + \frac{t^2}{2}\bar{\alpha},$$

and note that

$$\langle \alpha(t), \alpha(t) \rangle = 0$$

$$\langle \alpha(t), \bar{\alpha}(t) \rangle = (1 - |t|^2/2)^2.$$

One verifies that $F_B^2(t) = \langle \alpha(t) \rangle$ defines a variation of Hodge structure over the disk $|t| < \sqrt{2}$. Since $\dim A$ is even, there is a subspace J_A such that $A = J_A \oplus \bar{J}_A$, $J_A = J_A^{2,0} \oplus J_A^{1,1}$, and $J_A^{\perp} = J_A$, so that J_A is maximal isotropic. Define a pencil of maximal isotropic subspaces of $H_{\mathbb{C}}$ by setting

$$J(t) = J_A \oplus \langle \alpha(t) \rangle.$$

Because J_A is fixed and $\alpha(t)$ defines a variation of Hodge structure, one has

$$dJ/dt \subset J_A \oplus \langle \alpha(t) \rangle^{\perp} = J^{\perp}.$$

Let $V(2)$ be the set of all Hodge filtrations with $F^2 \subset J(t)$ for some t. This condition is equivalent to the requirement that F^2 be a subspace of some $J(t)$ of dimension $p = h^{2,0}$ which is h-positive. To see that $V(2)$ is a variation, we note that

$$F^2 \subset J(t) \subset J(t)^{\perp} \subset (F^2)^{\perp} = F^1,$$

generalizing (7.1). If $F^2(x)$ is a one-parameter family of filtrations contained in $J(t(x))$ for some function $t(x)$, then we have

$$dF^2/dx \subset dJ/dx \subset J^{\perp} \subset F^1,$$

as required.

By introducing explicit parameters for F^2, one may present $V(2)$ as a domain in $\mathbb{C}^N$, one of whose connected components fibers over over the disk $|t| < \sqrt{2}$. To see this, let C denote the Weil operator, which acts on $H^{p,q}$ by multiplication by i^{p-q}, and let $h_C(x, y) = \langle Cx, \bar{y} \rangle$ be the positive Hermitian form defined by the polarization (and the choice of Hodge structure). Let $\phi_1, \ldots, \phi_r$ be an h_C-unitary basis of $J_A^{2,0}$ and let $\psi_1, \ldots, \psi_s$ be an h_C-unitary basis of $J_A^{1,1}$. Then one may write

$$F^2(t, y, z) = \operatorname{span}\left\{\alpha(t) + \sum_j y_j\psi_j,\ \phi_i + \sum_j z_{ij}\psi_j\right\}.$$

The condition that h be positive on F^2 gives a system of inequalities that exhibits as a domain in complex (t, y, z)-space. One may also present $V(2)$ as a Siegel domain of type III.

The infinitesimal character of $V(2)$, i.e., the nature of the abelian subspaces $\mathfrak{a}$, is easily determined. To this end note that the partial derivatives relative to the parameters t, y_i, z_{ij} define velocity vectors in a space of filtrations, and so are represented by homomorphisms. We calculate these and observe that they can be identified with root vectors:

Tangent Vector	Homomorphism	Root Vector
$\partial/\partial t$	$\alpha \mapsto \beta$	$-e_{p+1}(0)$
$\partial/\partial y_i$	$\alpha \mapsto \psi_i$	$-e_{p+1}(0) + e_i(1)$
$\partial/\partial z_{ij}$	$\phi_i \mapsto \psi_j$	$-e_i(0) + e_j(1)$

The ordering used for the Hodge frame is

$$\{\phi_1, \ldots, \phi_p, \alpha; \psi_1, \ldots, \psi_q, \bar{\psi}_1, \ldots, \bar{\psi}_q, \beta; \bar{\phi}_1, \ldots, \bar{\phi}_p, \bar{\alpha}\}.$$

Similar calculations show that tangent spaces to all of the basic variations lift canonically to direct sums of root spaces. For the a type III variation of weight two, for example, the roots spanning $\mathfrak{a}$ are those appearing in the last row of the table above. The present case, however, is distinguished by the fact that the complexified tangent space of V, which may be identified with $\mathfrak{a} \oplus \bar{\mathfrak{a}}$, is not closed under triple brackets. To see this, set $\rho_1 = -e_p(0) + e_1(1)$, $\rho_2 = -e_{p+1}(0) + e_1(1)$, and $\rho_3 = -e_{p+1}(0)$. Denote by $\bar{\rho} = -\rho$ the conjugate root, and observe that although the partial sums in the expression below are roots, the total sum is not in $\mathfrak{a} \oplus \bar{\mathfrak{a}}$:

$$(\rho_1 + \bar{\rho}_2) + \rho_3 = -e_p(0).$$

Indeed, the only short roots in this space are $e_{p+1}(0)$ and $-e_{p+1}(0)$. It follows that type IV variations are not equivariantly imbedded hermitian symmetric domains, as are those of types I, II, and III. One can show more: the parameter space for a variation of type IV is a bounded but nonsymmetric domain. This is because the singularity structure and Levi form of the boundary does not match that of any of the symmetric domains.

Let us examine in detail the type IV parameter space in the lowest-dimensional case, that with $h^{2,0} = 2$, $h^{1,1} = 3$, for which $\dim V = 3$ and

$$F^2 = \operatorname{span}\{\gamma_1, \gamma_2\} = \operatorname{span}\{\alpha(t) + y\psi, \phi + z\psi\}. \tag{7.2}$$

The condition that F^2 define a polarized structure is given by the positivity of the matrix of hermitian inner products of the basis vectors: $G = h_C(\gamma_i, \bar{\gamma}_j) > 0$, which is equivalent to $4 \det G = (1 - |z|^2)(|t|^2 - 2)^2 - 4|y|^2 > 0$. The connected component of the solution set which contains the origin is a domain in $\mathbb{C}^3$ bounded by $|t| < \sqrt{2}$, $|y| < 1$, and $|z| < 1$. It can be shown that this domain has no quasi-projective quotients. This follows from a computation of the automorphism group, shown to us by Dan Burns.

The geometry of this domain is more easily understood in an unbounded model, which we now describe. The essence of this presentation is to write the filtration in terms of vectors corresponding to a mixed Hodge structure in the boundary of the period domain. To this end, let $E = \mathbb{Z}\{e_1, e_0\}$ be the standard symplectic lattice with $\langle e_1, e_0 \rangle = 1$. Set $\omega(x) = e_1 + xe_0$ and note that $i\langle \omega(x), \bar{\omega}(x) \rangle = 2\operatorname{Im}(x)$, so that $F^1(x) = \mathbb{C}\omega(x)$ defines a polarized variation of Hodge structure for $\operatorname{Im} x > 0$—the simplest SL_2-orbit, i.e., the one associated to the unique irreducible representation of dimension 2. Note that the e_i frame a mixed Hodge structure, with e_p of weight (p, p). Standard linear algebra constructions give variations on the symmetric and tensor squares:

$$F^2(T)S^2E = \mathbb{C}(\omega(T))^2$$

$$F^2(Z)E \otimes E = \mathbb{C}\omega(Z) \otimes \omega(i).$$

Since $(d/dZ)\omega(Z) \otimes \omega(i) = e_0 \otimes \omega(i)$, the space $span\{\omega(Z) \otimes \omega(i), (d/dZ)\omega(Z) \otimes \omega(i)\} = E \otimes \omega(i)$ is maximal isotropic in $E \otimes E$. One can therefore define a variation as above (7.2):

$$F^2 = \operatorname{span}\{(\omega(T))^2 + Ye_0 \otimes \omega(i), \omega(Z) \otimes \omega(i)\}. \tag{7.3}$$

The associated matrix of hermitian inner products is

$$G = 2\begin{pmatrix} 4(\operatorname{Im} T)^2 & -iY \\ i\bar{Y} & 2\operatorname{Im} Z \end{pmatrix},$$

and $G > 0$ is equivalent to

$$\operatorname{Im} Z - \frac{1}{8}\frac{|Y|^2}{(\operatorname{Im} T)^2} > 0.$$

If one replaces $T \in \{\operatorname{Im} T > 0\}$ by $\tau \in \{|\tau| < 1\}$ where $T = -i((\tau + i)/(\tau - i))$ replaces Y by $Y/(T + i)^2$, then the last expression becomes

$$\operatorname{Im} Z - \frac{1}{8}\left(\frac{1 + |\tau|^2}{1 - |\tau|^2}\right)^2 |Y|^2 > 0.$$

These are the defining inequalities of a Siegel domain of type III [13, page 32].

To relate the bounded and unbounded models, set $v = \omega(i)$, and write $\omega(T) = c(v + t\bar{v})$, where $t = -(T - i)/(T + i)$ and $c(T) = (T + i)/2i$. Then

$$F^2 = \operatorname{span}\{v^2 + 2tv\bar{v} + t^2\bar{v}^2 + y\bar{v} \otimes v, v \otimes v + z\bar{v} \otimes v\},$$

where $y = -c^{-2}(T)Y(1 + z)/2i$ and Z stands in the same relation to z as T does to t. This is essentially the family of subspaces defined in (7.2). To make the correspondence precise, replace t by $t/\sqrt{2}$, set $v^2 = \alpha$, $\sqrt{2}v\bar{v} = \beta$, etc.

From the basic variations just described one can form composite variations. To describe their construction, consider a Hodge structure in D which splits as

$$H = \bigoplus_i^{\ell} H_i \oplus T,$$

where the type of H_i is $\mathscr{H}_i = \{(w - i, i), (w - i - 1, i + 1), (i + 1, w - i - 1), (i, w - i)\}$. Let V_i be a basic variation on H_i and set

$$V = \bigoplus_i V_i \oplus T.$$

For an appropriate choice of split reference structure V realizes the bound of Theorem 1.6. Let $\sigma = (d^a, \ldots, d^b)$ be the sequence which attains the maximum in the definition of the function $q(L, \overline{h})$, let $S = \{i | d^i \in \sigma\}$ be the corresponding set of indices, and let H be a splitting with summands H_i whose Hodge numbers are $\{h^{w-i,i}, h^{w-i-1,i+1}, h^{i+1,w-i-1}, h^{i,w-i}\}$. Then $V = V_S$ is of dimension $q(L, h)$.

Note that the variations described above are, with the exception of those which are pure type II, degenerate in the sense that they admit fixed subspaces. For example, if S is an admissible sequence consisting of even integers for a variation V_S of odd weight, then F^p is constant for p odd. Naturally occurring variations of Hodge structure, e.g., those of hypersurfaces, tend to be nondegenerate, with rank $df \ll \dim \mathfrak{g}^{-1,1}$. Nonetheless, these are almost always maximal in the sense of inclusion [2].

As noted above, the basic variations $V(b)$ of types I, II, III have a hermitian symmetric parameter space. In fact, they occur, for suitable choice of the reference filtration, as 'Shimura varieties', i.e., as an irreducible constituent of the variation defined by the w-th cohomology of a family of abelian varieties with special symmetry properties [4]. For type IV variations we do not believe that an algebro-geometric realization is possible. We note, however, that geometric variations of dimension one less than maximal do exist (the fibers $t = const$ of the type IV variations are $SU(p, q)$-homogeneous).

8. Simple roots. In this section we shall describe simple root systems for period domains which are adapted to a Hodge structure on the Lie algebra. For the weight two case, see [6]. We shall assume that period domains are normalized in the sense that $h^{p,q} = 0$ for p or q negative, and $h^{w,0} \neq 0$. This is not a significant restriction, since one may always shift weights to obtain a normalized structure. The basic result is then:

THEOREM 8.1. *Let D be a normalized gap-free period domain of weight w and let $\mathfrak{g}$ be its Lie algebra of infinitesimal isometries, endowed with a Hodge structure relative to a reference point in D. Then there is a set of simple roots $\Delta_s = \{\alpha_1, \ldots, \alpha_n\}$ and a decomposition*

$$\Delta_s = \Delta_s^c \cup \Delta_s^{nc}$$

into compact and noncompact roots such that
(a) $\Delta_s^c \subset \Delta^{0,0}$,
(b) $\Delta_s^{nc} \subset \Delta^{-1,1}$,
(c) $|\Delta_s^{nc}| = [(w+1)/2]$.

The Theorem fails for period domains with gaps: Consider structures $H = H^{n,0} \oplus H^{0,n}$ with $n > 1$, for which the associated classifying spaces are $D = SO(2n)/U(n)$ for n even and $D = Sp(n, \mathbb{R})/U(n)$ for n odd. Both are Hermitian symmetric. For n even, D is compact, so there are no noncompact roots, and for n odd the Dynkin diagram (see figure 5) has a single noncompact root of type $(-n, n)$. In both cases the horizontal distribution is zero. Thus neither the statement about the number of noncompact roots nor about their type holds.

Below are three typical cases of such simple root systems. White dots indicate compact roots, and black dots, noncompact roots (the β_i to be defined later).

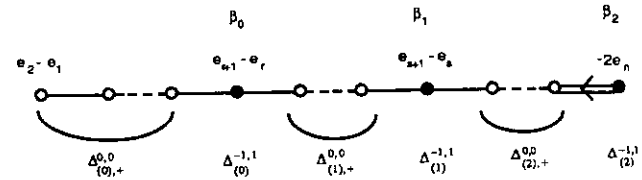

FIGURE 2. Type C_n (weight 5).

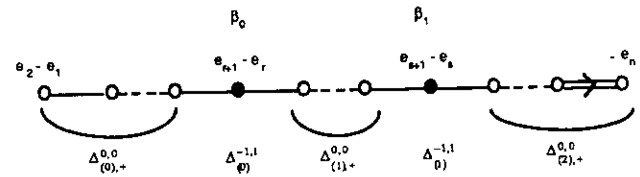

FIGURE 3. Type B_n (weight 4, $h^{2,2}$ odd).

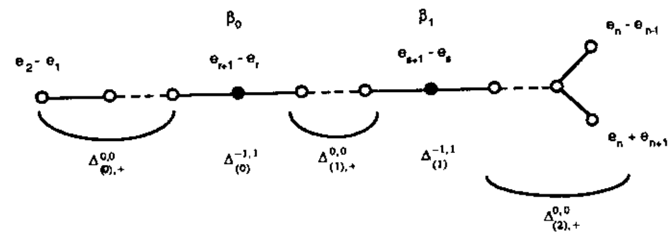

FIGURE 4. Type D_n (weight 4, $h^{2,2}$ even).

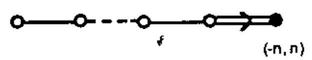

FIGURE 5. Degenerate period domain.

As the figures suggest, more detailed information is available. Write

$$\Delta_s^{nc} = \{\beta_0, \ldots, \beta_{m^*}\},$$

where $m^* = [(w-1)/2]$, and where the order defined by the subscripts is the same as that of the Dynkin diagram. For notational convenience set $\beta_{-1} = 0$ and let β_{m^*+1}

be a vector greater than any positive root. Then

$$\Delta^{0,0}_{(j),+} = \Delta^{0,0}_{(j)} \cap \Delta_+$$

is the set of positive roots corresponding to the j-th canonical factor of the isotropy subgroup. Define $\Delta^{0,0}_{(j),s} \subset \Delta^{0,0}_{(j),+}$ to be the corresponding set of simple roots. Then we have the following:

PROPOSITION 8.2. *The noncompact roots define a partition of the compact roots which determines the canonical summands of the isotropy subalgebra:*

$$\Delta^{0,0}_{(j),s} = \{\alpha \in \Delta_s | \beta_j < \alpha < \beta_{j+1}\}.$$

Proof. We shall give a full argument of both the theorem and the proposition for Lie algebras of type C_n; one treats the remaining cases in the same way. Begin by writing the standard basis for the dual of the Cartan subalgebra as

$$\{e_1(0), \ldots, e_{h^0}(0); e_1(1), \ldots, e_{h^1}(1); \ldots\} = \{e_{p_0}, \ldots, e_{q_0}; e_{p_1}, \ldots, e_{q_1}; \ldots\},$$

where $e_{p_0} = e_1$ and $e_{q_m} = e_n$ are the first and last basis vectors. The ordering determined by this basis (hence, by the Hodge frame) determines a system of simple roots,

$$\Delta_s = \{e_2 - e_1, \ldots, e_n - e_{n-1}, -2e_n\}.$$

Define a partition of the interval $K = [0, n]$, where $n = \dim H/2$, by $K_i = [p_i, q_i]$, with the p_i and q_i as above. Define a corresponding partition of Δ_s into subsets

$$\Delta_s(j) = \{e_{i+1} - e_i | i \in K_j\}$$

and

$$\Delta_s' = \{\beta_0, \ldots, \beta_m\}$$

where

$$\beta_j = e_{q_j+1} - e_{q_j}$$

for $0 \leqslant j < m$, and

$$\beta_m = -2e_{q_m} = -2e_n.$$

Write the e_i's in terms of the $e_b(j)$'s to obtain

$$e_r - e_s = e_b(j) - e_a(j) \in \Delta^{0,0}_{(j)}$$

for any $r, s \in K_j$ and for suitable a and b. Proceed in the same way to obtain

$$\beta_j = e_{p_j+1} - e_{q_j} = e_1(j+1) - e_{h^j}(j) \in \Delta^{-1,1}_{(j)}$$

for $j < m$ and

$$\beta_m = -2e_{h^m}(m) \in \Delta^{-1,1}_{(m)}.$$

Then the roots in the $\Delta_s(j)$'s—those with 'support' in a single subinterval K_j—are of type (0, 0) and the roots in Δ_s'—those with support in a pair of adjacent subintervals—are of type $(-1, 1)$.

According to [10], a (complexified) Cartan decomposition of $\mathfrak{g}$ is given by

$$\mathfrak{k}_{\mathbb{C}} = \sum_{p \text{ even}} \mathfrak{g}^{-p,p}$$

$$\mathfrak{p}_{\mathbb{C}} = \sum_{p \text{ odd}} \mathfrak{g}^{-p,p}.$$

Consequently the sets $\Delta_s(j)$ and Δ_s' consist of compact and noncompact roots, respectively, so that

$$\Delta_s^c = \bigcup \Delta_s(j)$$

and

$$\Delta_s^{nc} = \Delta_s'.$$

Moreover, $\Delta_s(j)$ is a basis for $\Delta^{0,0}_{(j)} = \{e_r - e_s | r, s \in K_j\}$, and the β's separate the $\Delta_s(j)$'s, as required.

Remark 8.3. *From the proof of the theorem in the case C_n, it is apparent that the greatest noncompact simple root β_{m^*} must be the right-most node of the Dynkin diagram. In case D_n at most one of the roots to the right of the Y-junction is noncompact. There are no other restrictions on the positions of the noncompact roots.*

There is a further decomposition $\mathfrak{p}_{\mathbb{C}} = \mathfrak{p}^- \oplus \mathfrak{p}^+$ into mutually conjugate subspaces. One such is according to the sign of the root:

$$\mathfrak{p}^- = \sum_{p<0 \text{ odd}} \mathfrak{g}^{p,-p}_{(j)}.$$

If D is nonhermitian, then this space is nonabelian. In the hermitian symmetric case, it coincides with $\mathfrak{g}^{-1,1}$ and is abelian. Another such decomposition is obtained by twisting according to the parity of j:

$$\mathfrak{p}^- = \sum_{\substack{p<0 \text{ odd},\\ j \text{ even}}} \mathfrak{g}_{(j)}^{p,-p} \oplus \sum_{\substack{p>0 \text{ odd},\\ j \text{ odd}}} \mathfrak{g}_{(j)}^{p,-p} \underset{def}{=} \mathfrak{p}_{ev}^- \oplus \mathfrak{p}_{odd}^- . \tag{8.5}$$

The properties of this decomposition are given by the next result:

PROPOSITION 8.6. *Let $\mathfrak{p} = \mathfrak{p}^- \oplus \mathfrak{p}^+$ be the twisted decomposition defined by* (8.5), *and let $X = G/K$ be the symmetric space associated to D. If the weight is even, then $\mathfrak{p}^-$ is generally nonabelian. If the weight is odd then $\mathfrak{p}^-$ is abelian and defines the holomorphic tangent bundle of X.*

Proof. The assertion to be proved in the case of even weight follows from proposition (9.1) below. In the odd case one must show that $\mathfrak{p} = \mathfrak{p}^- \oplus \mathfrak{p}^+$ is a $\mathfrak{k}$-invariant decomposition into mutually conjugate abelian subspaces, for then $\mathfrak{p}^-$ defines an integrable almost complex structure.

Elements of $\mathfrak{p}_{ev}^-$ are represented by block matrices with support $\{(j+p, j), (w-j, w-j-p)\}$, where p is odd and j is even. If w is odd, then both blocks appear in positions of the form (*odd, even*). Elements of $\mathfrak{p}_{odd}^-$ are represented by block matrices with support $\{(j, j+p), (w-j-p, w-j)\}$, where p is odd and j is odd. If w is odd, then both blocks appear in positions of the form (*odd, even*), as before. Let A and B be any two block matrices with support in the set $\mathscr{S}^- = \{(odd, even)\}$. Then $AB = 0$. Therefore $\mathfrak{p}^-$ is abelian.

Observe next that $\mathfrak{p}^+$, since it consists of matrices conjugate to those in $\mathfrak{p}^-$, has support in the set $\mathscr{S}^+ = \{(even, odd)\}$. Therefore $\mathfrak{k}$ must have support in $\mathscr{S}^0 = \{(even, even)\} \cup \{(odd, odd)\}$. Now if A and B have support in $\mathscr{S}^-$ and $\mathscr{S}^0$, then AB and BA have support in $\mathscr{S}^-$, so that $[\mathfrak{k}, \mathfrak{p}^-] \subset \mathfrak{p}^-$, as required.

9. Accessibility. Using the root structure given in the preceding section, we shall verify the hypothesis of Chow's theorem, thereby proving Theorem (1.11). We note that the theorem fails if the Hodge type of D has a gap. Indeed, if p is a gap, then $F^{p+1} = F^p$, so that F^{p+1} is constant, by horizontality. Therefore Hodge structures which differ at level p cannot be joined by a piecewise horizontal curve: accessibility fails. An extreme example of this degenerate behavior occurs for structures of the form $H = H^{n,0} \oplus H^{0,n}$ with $n > 1$. Because the horizontal distribution is zero, there are no nonconstant variations. The result to be established is:

PROPOSITION 9.1. *The horizontal component $\mathfrak{g}^{-1,1}$ generates the holomorphic tangent space*

$$\mathfrak{g}^- = \sum_{p>0} g^{-p,p}.$$

Note that this Proposition also fails for domains with gaps. For the proof we require a small fact about roots of the classical Lie algebras. Call a sequence $\sigma = (\alpha_1, \ldots, \alpha_m)$ *good* if all consecutive partial sums formed from σ are roots. Then we have:

LEMMA 9.2. *Let Δ_s be a system of simple roots for a classical Lie algebra. Then every positive root α can be written as a sum of at most two good sequences in Δ_s. For SL_n one good sequence suffices.*

Proof. The argument is an easy exercise based on the standard form of the root system. For Lie type C_n one has $\Delta = \{\pm e_i \pm e_j\}$ with $\Delta_s = \{\alpha_i = e_i - e_j | i < n\} \cup \{\alpha_n = e_n\}$. Then

$$e_i - e_j = \alpha_i + \cdots + \alpha_{j-1}$$

$$e_i + e_j = (\alpha_i + \cdots + \alpha_n) + (\alpha_{n-1} + \cdots + \alpha_j)$$

The calculations for types B and D are similar, and that for SL is contained in the first of the above pair of equations.

Proof of the proposition. Let α be a root of type $(-p, p)$. Consider first the case in which α is the sum of a single good sequence containing k noncompact simple roots. Partition the sequence into k consecutive subsequences s_i with sum γ_i. By the lemma the γ_i are roots, as are all consecutive sums of the γ's. Let X be the root vector corresponding to α, and let X_i be the root vector corresponding to γ_i, each containing a single noncompact simple root. Then X_i is of type $(-1, 1)$, and X is, up to a constant, a k-fold bracket of vectors of type $(-1, 1)$: $X = c[\ldots[[X_1 X_2]X_3] \ldots X_k]$. Therefore $k = p$, and the proposition holds.

For the remaining case apply the preceding argument to each of the two good sequences σ_i associated to X. Let Y_i be the corresponding root vectors, and let k_i denote the number of noncompact roots in σ_i. Then Y_i is of type $(-k_i, k_i)$ and $X = c[Y_1, Y_2]$. If both k_1 and k_2 are positive, then we are done, since $p = k_1 + k_2$ must hold. If $k_2 = 0$, then $X = [[\ldots[X_1[X_2, X_3]\ldots]Y_2]$. Rewrite the right-hand side using the Jacobi identity so that when Y_2 occurs, it occurs in a factor $[X_i, Y_2]$. Since this last expression is of type $(-1, 1)$, X is of type $(-k_1, k_1) = (-p, p)$, and we are done. The case of $k_1 = 0$ is treated in the same way.

REFERENCES

1. J. A. CARLSON, *Bounds on the dimension of a variation of Hodge structure*, Trans. Amer. Math. Soc. **294** (1986) 45–64, Erratum Trans. Amer. Math. Soc. **299** (1987), 429.
2. ——— and RON DONAGI, *Hypersurface variations are maximal*, Invent. Math. **89** (1987), 371–374.
3. ——— and P. GRIFFITHS, *Infinitesimal variations of Hodge structure and the global Torelli problem*, Géometrie algébrique, Angers 1979, Sijthoff & Noordhoof, Alphen an den Rijn (1980), 51–76.
4. ——— and C. SIMPSON, "Shimura varieties of weight two Hodge structures" in *Hodge Theory*, Springer Verlag, Lecture Notes in Mathematics **1246** (1987), 1–15.
5. ——— and D. TOLEDO, *Variations of Hodge structure, Legendre submanifolds, and accessibility*, to appear in Trans. Amer. Math. Soc.
6. EDUARDO H. CATTANI and AROLDO G. KAPLAN, *Horizontal SL_2-orbits in flag domains*, Math. Ann. **235** (1978), 17–35.
7. W. L. CHOW, *Über systeme von linearen partiellen differential-gleichungen erster ordnung*, Math. Ann. **117** (1939), 98–105.

8. P. A. GRIFFITHS, *Periods of integrals on algebraic manifolds*, III, Publ. Math. I.H.E.S. **38** (1970), 125–180.
9. ——— and WILFRIED SCHMID, *Locally homogeneous complex manifolds*, Acta Math **123** (1969), 253–302.
10. ——— and ———, "Recent developments in Hodge theory: a discussion of techniques and results," *Proceedings of the Bombay Colloquium on Discrete Subgroups of Lie Groups*, Tata Institute for Fundamental Research, Bombay (1973).
11. ——— and L. TU, "Variation of Hodge Structure," *Topics in Transcendental Algebraic Geometry* (P. A. GRIFFITHS, ed.), Am. of Math. Studies **106**, Princeton University Press (1984).
12. A. I. MALCEV, Commutative subalgebras of semisimple Lie algebras, Izvestia Ak. Nauk USSR (Russian) **9** (1945), 125–133; English: Amer. Math. Sre. *Translations No.* 40 (1951).
13. I. I. PYATETSKII-SHAPIRO, *Automorphic Functions and the Geometry of Classical Domains*, Gordan and Breach, London, 1969, pp 264.
14. I. SCHUR, *Zur Theorie der vertauschbaren Matrizen*, J. Reine und Angew. Math. **130** (1905), 66–76.

CARLSON AND TOLEDO: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF UTAH, SALT LAKE CITY, UTAH 84112

KASPARIAN: BULGARIAN ACADEMY OF SCIENCES, P. O. BOX 373, SOFIA 1090, BULGARIA

Vol. 58, No. 3 DUKE MATHEMATICAL JOURNAL © June 1989

POINCARÉ SERIES FOR $GL(3, \mathbf{R})$-WHITTAKER FUNCTIONS

ERIC STADE

CONTENTS

Introduction. C. L. Siegel first suggested the construction of a Poincaré series for $GL(2, \mathbf{R})$ that is both automorphic for $GL(2, \mathbf{Z})$ and an eigenfunction of the Laplacian but is not an Eisenstein series. Such a series is built from the $GL(2, \mathbf{R})$ "Whittaker function" $M_v(z)$ and, in its domain of convergence, defines a holomorphic function of the complex variable v.

The properties of Siegel's function were originally studied by Neunhöffer [11] and Niebur [12]. They determined the necessary criteria for convergence and then obtained, among other results, a meromorphic continuation and functional equation for the series in the variable v. Specifically, this equation states that a certain linear combination of these Poincaré series is equal to the $GL(2, \mathbf{Z})$-Eisenstein series $E(z; v)$.

The theory of Neunhöffer and Niebur has recently been extended by Miatello and Wallach [10] to general rank-one groups. On the other hand, when one tries to generalize this theory to $GL(n, \mathbf{R})$ ($n > 2$), a fundamental obstruction occurs, namely, the relevant Poincaré series do not ever converge! It does not even seem possible to define these series as distributions.

Nevertheless it *is* possible, at least in the case of $GL(3, \mathbf{R})$, to give a meaning to Siegel's construction as the analytic continuation and specialization (with respect to a new variable μ) of a linear functional defined on any finite-dimensional space of cusp forms. In this context there then exist a meromorphic continuation and a functional equation analogous to those obtained on $GL(2, \mathbf{R})$. We wish, in this paper, to develop completely the $GL(3, \mathbf{R})$ theory.

Received June 27, 1988. Revision received October 3, 1988.