CARTAN SUBALGEBRAS OF SIMPLE LIE ALGEBRAS

BY

ROBERT LEE WILSON

Abstract. Let $L$ be a finite-dimensional simple Lie algebra over an algebraically closed field of characteristic $p > 7$. Let $H$ be a Cartan subalgebra of $L$, let $L = H + \sum_{\gamma \in \Gamma} L_\gamma$ be the Cartan decomposition of $L$ with respect to $H$, and let $\tilde{H}$ be the restricted subalgebra of $\text{Der} L$ generated by $\text{ad} \ H$. Let $T$ denote the maximal torus of $H$ and $I$ denote the nil radical of $H$. Then $\tilde{H} = T + I$. Consequently, each $\gamma \in \Gamma$ is a linear function on $H$.

Let $L$ be a finite-dimensional simple Lie algebra over an algebraically closed field $F$. Let $H$ be a Cartan subalgebra of $L$ and let $L = H + \sum_{\gamma \in \Gamma} L_\gamma$ be the Cartan decomposition of $L$ with respect to $H$. The main results on the structure of $H$, that $H$ is abelian and that $\text{ad} \ h$ is semisimple for each $h \in H$, which hold when $F$ is of characteristic zero are known [3, Satz 12] to fail when $F$ is of prime characteristic. The resulting lack of information about the structure of $H$ has been a severe handicap in the development of structure and classification theory for finite-dimensional simple Lie algebras of prime characteristic.

In this paper we assume that $F$ is of characteristic $p > 7$ and investigate the structure of $\tilde{H}$, the restricted subalgebra of $\text{Der} L$ generated by $\text{ad} \ H$. Our main result (Theorem 2.1) is that $\tilde{H} = T + I$ where $T$ is the maximal torus of $\tilde{H}$ and $I$ is the nil radical of $\tilde{H}$. (See §1 for definitions.) An immediate consequence (Corollary 2.2) is that each $\gamma \in \Gamma$ is a linear function on $H$. (Although in characteristic zero this is a trivial consequence of Lie's Theorem, the result is new in prime characteristic.)

This type of result was first proved by Schue [1] under the additional hypotheses that $\dim T = 1$ and that every proper subalgebra of $L$ is solvable.

Our proof begins with Schue's observation that if $\tilde{H} \neq T + I$ then there exists $b \in \tilde{H}$, $b \notin T + I$ such that $b^p \in I$ and $[b, \tilde{H}] \subseteq T + I$. We then let

$$S = \{(\gamma, \delta) \in \Gamma \times \Gamma | \gamma([b[L_\delta, L_{-\delta}]] \neq (0))\}.$$

Using Schue's techniques we show (§3) that there exist $\alpha, \beta \in \Gamma$ with $(\alpha, \beta) \in S$ and $(\beta, \alpha) \in S$. The argument then divides into two cases depending on
whether or not there exists $\gamma \in \Gamma$ with $(\gamma, \gamma) \in S$. We consider these two cases separately (§§4 and 5), showing that either one leads to a contradiction. These results have been announced in [4].

1. Preliminaries.

(1.1) If $K \supseteq L$ are Lie algebras let $N_K(L)$ denote the normalizer of $L$ in $K$ and let $C(K)$ denote the center of $K$.

Let $R$ be a restricted Lie algebra over a field $F$ of characteristic $p > 0$. If $X$ is a subset of $R$ let $\langle X \rangle$ denote the restricted subalgebra of $R$ generated by $X$. If $X$ is a subalgebra then $\langle X \rangle$ is clearly the $F$-span of \{ $x^p | x \in X, i > 0$ \}.

(1.2) Following [2, Chapter V.7] or [1, § 1] we say that $x \in R$ is semisimple if $x \notin \langle x^p \rangle$ and that $x \in R$ is nilpotent if $x^{p^n} = 0$ for some $n$. See also [5].

An ideal $J \subseteq R$ is said to be nil if every $x \in J$ is nilpotent. It is easily seen that a finite-dimensional restricted Lie algebra $R$ contains a nil ideal $I$ which contains every nil ideal. We call $I$ the nil radical of $R$. An abelian subalgebra $T \subseteq R$ is called a torus if every element of $T$ is semisimple.

(1.3) Proposition. Let $R$ be a finite-dimensional restricted Lie algebra over a perfect field $F$. Then:

(i) If $x, y \in R$ are nilpotent (respectively, semisimple) and $[x, y] = 0$ then every element of $\langle \{x, y\} \rangle$ is nilpotent (respectively, semisimple).

(ii) If $x \in R$ then there exist $x_s, x_n \in \langle x \rangle$ such that $x_s$ is semisimple, $x_n$ is nilpotent, and $x = x_s + x_n$.

(iii) If $x, y, z \in R$, $y$ semisimple, $z$ nilpotent, $[y, z] = 0$, and $x = y + z$, then $y = x_s$ and $z = x_n$.

(iv) If $R$ is nilpotent and $x \in R$ is semisimple then $x \in C(R)$.

(v) If $R$ is nilpotent $\{x_s | x \in R \}$ is the unique maximal torus of $R$.

Proof. Parts (i)–(iii) are proved in Chapter V.7 of [2]. If $x$ is semisimple then $x \in \langle x^p \rangle$ for any $n > 1$. Since $R$ is nilpotent, $ad(x^p) = (ad x)^p = 0$ for sufficiently large $n$. Hence $x = 0$, proving part (iv). Since $\{x_s | x \in R \}$ contains every semisimple element it must contain every torus of $R$. By (i) and (iv) $\{x_s | x \in R \}$ is a torus, proving part (v).

(1.4) Assume that $F$ is algebraically closed, that $L$ is a finite-dimensional simple Lie algebra over $F$, that $H$ is a Cartan subalgebra of $L$, and that $L = H + \sum_{\gamma \in \Gamma} L_\gamma$ is the corresponding Cartan decomposition.

Identify $L$ with the isomorphic subalgebra $ad L$ of the restricted Lie algebra $Der L$. Let $H = \langle H \rangle \subseteq Der L$. (Thus $H$ is the $F$-span of $\{ h^p | h \in H, n > 0 \}$, where we write $h^p$ for $(ad h)^p$.) Let $\overline{L} = \overline{H} + L$.

Lemma. (i) $\overline{H}$ is a Cartan subalgebra of $\overline{L}$.

(ii) If $\overline{L} = \overline{H} + \sum_{\gamma \in \Gamma} L_\gamma$ is the Cartan decomposition of $\overline{L}$ with respect to $\overline{H}$ then the map $\overline{\gamma} \mapsto \overline{\gamma}|_H$ is a bijection of $\overline{\Gamma}$ onto $\Gamma$ and $L_\gamma = L_{\overline{\gamma}|_H}$ for all $\overline{\gamma} \in \overline{\Gamma}$.
Proof. Since \([x, y^p] = x(\text{ad}y)^p\) and since \(\overline{H}\) is the \(F\)-span of \(\{h^n| h \in H, n > 0\}\), we see by induction on \(i\) that \(\overline{H}^i = H^i\) for all \(i > 2\). Hence \(\overline{H}\) is nilpotent. If \(x \in L \cap N_r(\overline{H})\) then \(\{x, H\} \subseteq \overline{H} \cap L = H\). Thus \(x \in N_L(H) = H\). Hence \(\overline{H} = N_G(H)\), proving (i).

The map \(\overline{\gamma} \mapsto \overline{\gamma}|_H\) is clearly surjective, since for each \(\gamma \in \Gamma\), \(\overline{H}\) acts on \(L_\gamma\) and hence \(L_\gamma \subseteq \sum_{\gamma \in \Gamma} r_\gamma = L_\gamma\).

If \(\overline{\gamma} \in \overline{\Gamma}\) then there exists \(x \in \overline{H}\) such that \(\overline{\gamma}(x) \neq 0\). Then \(\text{ad} x : L_\gamma \to L_\gamma\) is surjective so \(L_\gamma = [x, L_\gamma] \subseteq [L_\gamma, L_\gamma] \subseteq L\). Thus \(L_\gamma \subseteq L_\gamma\).

Now suppose \(\overline{\gamma}, \overline{\eta} \in \overline{\Gamma}\) and \(\overline{\gamma}|_H = \overline{\eta}|_H \in \Gamma\). Then \(L_\gamma \otimes (L_\eta)^* \subseteq L_\gamma \otimes (L_\eta)^*\) (where \(V^*\) denotes the contragredient module to \(V\)). Now \(H\) acts on \(L_\gamma \otimes (L_\eta)^*\). Moreover, the only weight is \(\overline{\gamma} - \overline{\eta} = 0\). Hence \(H\) acts as a Lie algebra of nilpotent linear transformations on \(L_\gamma \otimes (L_\eta)^*\). By Engel's Theorem \(\overline{H}\) also acts as a Lie algebra of nilpotent linear transformations on \(L_\gamma \otimes (L_\eta)^*\). Thus the only weight of \(\overline{H}\) on \(L_\gamma \otimes (L_\eta)^*\) is 0. Since \(\overline{\gamma} - \overline{\eta}\) is a weight of \(\overline{H}\) on \(L_\gamma \otimes (L_\eta)^*\) we have \(\overline{\gamma} = \overline{\eta}\). Thus the map \(\overline{\gamma} \mapsto \overline{\gamma}|_H\) is bijective and hence \(L_\gamma = L_\gamma\).

(1.5) In view of Lemma 1.4(ii) we may identify \(\overline{\Gamma}\) and \(\Gamma\) and write \(\overline{L} = H + \sum_{\gamma \in \Gamma} L_\gamma\). Let \(T\) denote the maximal torus of \(\overline{H}\) and \(N\) denote the nil radical of \(\overline{H}\).

Lemma. Each \(\gamma \in \Gamma\) is linear on \(T + I\).

Proof. Since \(T\) is a torus each \(t \in T\) acts diagonally on each \(L_\gamma\) and hence \(\gamma\) is linear on \(T\).

If \(h \in H\) then by Lemma 1.3(ii) \(h = h_s + h_n\). Then \(h^{n'} = h_s^{n'} + h_n^{n'}\) for all \(i\), hence \(h^{n''} = h_s^{n''}\) for sufficiently large \(n\). Now if \(y \in L_\gamma\) then

\[0 = y(\text{ad} h - \gamma(h))^{n''} = y(\text{ad} h_s - \gamma(h))^{n''}\]

for sufficiently large \(n\). Hence \(\gamma(h) = \gamma(h_s)\) for all \(h \in H\). Now Lemma 1.3 shows that if \(h \in T + I\) then \(h_s \in T, h_n \in I,\) and \(h_s\) is a linear function of \(h\). Hence \(\gamma(h) = \gamma(h_s)\) is a linear function of \(h\).

(1.6) Lemma. Let \(X\) be a subset of \(H\) and \(E = \{\gamma \in \Gamma| \gamma(X) \neq 0\}\). If \(E \neq \emptyset\) then \(H = \sum_{\gamma \in E} [L_\gamma, L_{-\gamma}]\).

Proof. This is a special case of (4.2) of [1]. (The hypothesis in (4.2) of [1] that \(L\) is semirestricted can be dropped here, since we assume \(X \subseteq H\).)

2. Statement of results. Our main result is

(2.1) Theorem. Let \(L\) be a finite-dimensional simple Lie algebra over an algebraically closed field \(F\) of characteristic \(p > 7\). Let \(H\) be a Cartan subalgebra of \(L\), \(\overline{H}\) be the restricted subalgebra of \(\text{Der} L\) generated by \(\text{ad} H, T\) be the maximal torus of \(\overline{H}\), and \(I\) be the nil ideal of \(\overline{H}\). Then \(\overline{H} = T + I\).
(2.2) Let \( L = H + \sum_{\gamma \in \Gamma} \mathcal{L}_\gamma \) be the Cartan decomposition of \( L \) with respect to \( H \). By Lemma 1.5 each \( \gamma \in \Gamma \) is linear on \( T + I \). By Theorem 2.1 \( \mathcal{H} = T + I \). Hence we have

**Corollary.** Each \( \gamma \in \Gamma \) is a linear function on \( H \).

3. Action of \( \mathcal{H} \) on \( L \).

(3.1) Let \( F, L, \) and \( H \) be as in Theorem 2.1 and assume \( \mathcal{H} \neq T + I \). We will eventually derive a contradiction, thus proving Theorem 2.1. We begin with an analysis of the structure of \( \mathcal{H} \) modeled on §3.5 of [1].

**Lemma.** If \( \mathcal{H} \neq T + I \) then there exists \( b \in \mathcal{H} \) such that

(i) \( b \notin T + I \),
(ii) \( [b, H] \subseteq T + I \),
(iii) \( [b[H, H]] \subseteq I \),
(iv) \( [b, x^p] \in I \) for all \( x \in \mathcal{H} \),
(v) \( b^p \in I \), and
(vi) \( [b, H] \not\subseteq I \).

**Proof.** Since \( T \) is central and \( I \) is an ideal, \( T + I \) is a proper ideal of \( \mathcal{H} \). Then \( \mathcal{H}/(T + I) \) is a nonzero \( \mathcal{H} \)-module. As \( \mathcal{H} \) is nilpotent, there exists \( b \in \mathcal{H}, b \notin T + I \), such that \( [b, H] \subseteq T + I \). Since \( b = b_s + b_n \) and \( b_s \in T \) we see that \( b_n \in \mathcal{H}, b_n \notin T + I \). Thus we may assume that \( b = b_n \) is nilpotent. Now

\[
[b^p, \mathcal{H}] = [b, \mathcal{H}](ad b)^p-1 \subseteq (T + I)(ad b)^p-1 \subseteq I
\]

so \( b^pF + I \) is a nil ideal containing \( I \). Thus by the maximality of \( I, b^p \in I \). Now,

\[
[b[H, H]] \subseteq [b, H]H \subseteq [T + I, H] \subseteq I
\]

and

\[
[b, x^p] = [b, x](ad x)^p-1 \in (T + I)(ad x)^p-1 \subseteq I
\]

for all \( x \in \mathcal{H} \). If \( [b, H] \subseteq I \) then \( [b, H] \subseteq I \), and so \( bF + I \) is a nil ideal containing \( I \). This contradicts the maximality of \( I \), so \( [b, H] \not\subseteq I \).

(3.2) We continue to assume that \( L \) has Cartan decomposition \( L = H + \sum_{\gamma \in \Gamma} \mathcal{L}_\gamma \).

**Lemma.** There exists \( \gamma \in \Gamma \) such that \( \gamma([b, H]) \neq (0) \).

**Proof.** If \( t \in T \) then \( \gamma(t, t) = \gamma(t)\gamma(t) \) for all \( y \in L_{\gamma} \). Hence \( \gamma(t) = 0 \) implies \( [L_{\gamma}, t] = 0 \) and \( \gamma(t) = 0 \) for all \( \gamma \in \Gamma \) implies \( t \in C(L) = (0) \).

Now if \( t \in T, n \in I \) then we have seen (in (1.5)) that \( \gamma(t + n) = \gamma(t) \). Thus \( \gamma(t + n) = 0 \) for all \( \gamma \in \Gamma \) implies \( t + n \in I \). Since, by Lemma 3.1 (vi), \( [b, H] \not\subseteq I \), we have \( \gamma([b, H]) \neq (0) \) for some \( \gamma \in \Gamma \).
(3.3) Let \( S = \{ (\gamma, \delta) \in \Gamma \times \Gamma | \gamma([b_L, L_\gamma]) \neq (0) \} \).

**Proposition.** Either

(3.3.1) \((\alpha, \alpha) \in S\) for some \( \alpha \in \Gamma \),

or

(3.3.2) for all \( \gamma \in \Gamma \), \((\gamma, \gamma) \notin S\) but there exist \( \alpha, \beta \in \Gamma \) with \((\alpha, \beta) \in S\) and \((\beta, \alpha) \in S\).

**Proof.** By Lemma 3.2 there exist \( x \in H \) and \( \gamma \in \Gamma \) such that \( \gamma([b, x]) \neq 0 \). Then by Lemma 1.6

\[
H = \sum_{\gamma \in \Gamma, \gamma([b, x]) \neq 0} [L_\gamma, L_{-\gamma}].
\]

Now by Lemma 3.1(vi) we have \([b, H] \subseteq I\). Thus there exists \( \alpha \in \Gamma \) such that \( \alpha([b, x]) \neq 0 \) and \([b[L_\alpha, L_{-\alpha}]\] \( \subseteq I \). Hence for some root \( \delta \), \( \delta([b[L_\alpha, L_{-\alpha}]] \neq (0) \). Again by Lemma 1.6 we have

\[
H = \sum_{\gamma \in \Gamma, \gamma([b[L_\alpha, L_{-\alpha}]]) \neq (0)} [L_\gamma, L_{-\gamma}].
\]

Since \( x \in H \) and \( \alpha([b, x]) \neq 0 \) there exists \( \beta \in \Gamma \) such that \( \beta([b[L_\alpha, L_{-\alpha}]] \neq (0) \) and \( \alpha([b[L_{\beta}, L_{-\beta}]]) \neq (0) \). Thus either (3.3.1) or (3.3.2) holds.

(3.4) For \( \gamma \in \Gamma \) define \( \overline{H}_\gamma = \{ x \in \overline{H} | \gamma([b, x]) = 0 \} \). By Lemma 1.5 \( \overline{H}_\gamma \) is a subspace of \( \overline{H} \). By (iii) and (iv) of Lemma 3.1, \( \overline{H}_\gamma \) is a restricted ideal of \( \overline{H} \). If \( \overline{H} \neq \overline{H}_\gamma \) fix an element \( c_\gamma \in \overline{H} \) with \( \gamma([b, c_\gamma]) = 1 \).

Let \( V \) be a finite-dimensional irreducible restricted \( \overline{H} \)-module. Then \( V \) is a submodule and, since \( I \) is nil, \( V \neq \overline{H} \). Assume that \( \nu t = \gamma(t)\nu \) for all \( \nu \in V \), \( t \in T \). Let \( V_0 = \{ \nu \in V | \nu b = 0 \} \). Let \( \{ v_1, \ldots, v_n \} \) be a basis for \( V_0 \). Let \( C = C_\gamma \).

**Lemma.** (i) For each \( k \), \( 0 < k < p - 1 \), the \( F \)-span of \( \{ v_i c^i | 1 < i < n, 0 < j < k \} \) is an \( \overline{H}_\gamma \)-subspace of \( V \).

(ii) \( \{ v_i c^i | 1 < i < n, 0 < j < p - 1 \} \) is a basis for \( V \).

**Proof.** Since \( T \cap \ker \gamma \) is an ideal of \( \overline{H} \) of codimension 1 in \( T \) it is sufficient to prove the result under the assumption \( \dim T = 1 \). This is done in \$3.7 \) of [1].

(3.5) Fix \( \alpha \in \Gamma \) as in Lemma 3.3. For \( \gamma \in \Gamma \) define a bilinear form on \( L_\gamma \times L_{-\gamma} \) by

\[
(x, y) = \alpha([b[x, y]]) \quad \text{for } x \in L_\gamma, y \in L_{-\gamma}.
\]

In view of Lemma 3.1(iii) and the Jacobi identity we have

\[
([x, h], y) = (x, [h, y]) \quad \text{for all } x \in L_\gamma, y \in L_{-\gamma}, h \in \overline{H}.
\]
If $X$ is an ad $\overline{H}$ invariant subspace of $L_\gamma$ then $X^\perp = \{y \in L_{-\gamma} \mid \langle X, y \rangle = 0 \}$ is an ad $\overline{H}$ invariant subspace of $L_{-\gamma}$. Define $K_{-\gamma} = L_{-\gamma}^\perp$ for all $\gamma \in \Gamma$ and define $n_\gamma = \dim L_\gamma/K_\gamma$.

Now if $(\alpha, \alpha) \in S$ and if for some $i$, $1 < i < p - 1$, we have $n_{\alpha i} \neq 0$, then $(i\alpha, i\alpha) \in S$. Thus, replacing $\alpha$ by $i\alpha$ if necessary, we may assume that the root $\alpha \in \Gamma$ in (3.3.1) satisfies

\[ n_\alpha > n_{\alpha i} \quad \text{for all } i, 1 < i < p - 1. \]  

Similarly, if (3.3.2) holds and if $n_{\beta + i\alpha} \neq 0$ for some $i$, $0 < i < p - 1$, then we have $(\alpha, \beta + i\alpha) \in S$ and $(\beta + i\alpha, \alpha) \in S$. Hence, replacing $\beta$ by $\beta + i\alpha$ if necessary, we may assume that the pair $(\alpha, \beta)$ in (3.3.1) satisfies

\[ n_\beta > n_{\beta + i\alpha} \quad \text{for all } i, 0 < i < p - 1. \]  

(3.6) We will complete the proof of Theorem 2.1 in the next two sections by showing that either conclusion in Lemma 3.3 leads to a contradiction. In §4 we will show that if (3.3.1) holds then $n_{2\alpha} + n_{3\alpha} > 2n_\alpha$, contradicting (3.5.1), and in §5 we will show that if (3.3.2) holds then $n_{\beta + \alpha} + n_{\beta - \alpha} > 2n_\beta$, contradicting (3.5.2).

4. Dimension arguments. I.

(4.1) We continue to let $F$, $L$, and $H$ be as in Theorem 2.1 and to assume $\overline{H} \neq T + I$. In addition, we assume that (3.3.1) holds. Our object is to show that $n_{2\alpha} + n_{3\alpha} > 2n_\alpha$, thus contradicting (3.5.1).

(4.2) Let $L_\alpha \supseteq N_\alpha \supseteq M_\alpha \supseteq K_\alpha$, where $K_\alpha$ is as in (3.5), $N_\alpha$ and $M_\alpha$ are ad $\overline{H}$ invariant subspaces of $L_\alpha$, and $N_\alpha/M_\alpha$ is an irreducible $\overline{H}$-module (necessarily restricted).

For $X = M$ or $N$ and for $i = 2, 3$ define

\[ X'_i = \{x \in L_\alpha \mid x(\operatorname{ad} L_{-\alpha})^{i-1} \subseteq X_\alpha \} \]

and

\[ X_{i\alpha} = X'_i + K_{i\alpha}. \]

Then $X_{i\alpha}$ is an ad $\overline{H}$ submodule of $L_{i\alpha}$ and $L_{i\alpha} \supseteq N_{i\alpha} \supseteq M_{i\alpha} \supseteq K_{i\alpha}$ for $i = 2, 3$. By the Jacobi identity, $[X_\alpha, X_\alpha] \subseteq [X_{2\alpha}, X_{2\alpha}]$ and $[[X_\alpha, X_\alpha], X_\alpha] \subseteq X_{3\alpha}$.

**Lemma.** $[X_{2\alpha}^\perp, X_{2\alpha}^\perp] \subseteq X_{2\alpha}^\perp$ and $[[X_{2\alpha}^\perp, X_{2\alpha}^\perp], X_{3\alpha}^\perp] \subseteq X_{3\alpha}^\perp$.

**Proof.**

\[
([X_{2\alpha}^\perp, X_{2\alpha}^\perp], X_{2\alpha}) \subseteq (\{ [X_{2\alpha}^\perp, X_{2\alpha}^\perp], X_{2\alpha}^\perp \} + ([X_{2\alpha}^\perp, X_{2\alpha}^\perp], K_{2\alpha}) \]

\[
= (\{ [X_{2\alpha}^\perp, X_{2\alpha}^\perp], X_{2\alpha}^\perp \} \subseteq (X_{2\alpha}^\perp, [X_{2\alpha}^\perp, X_{2\alpha}]) \]

\[
\subseteq (X_{2\alpha}^\perp, [L_{-\alpha}, X_{2\alpha}^\perp]) \subseteq (X_{2\alpha}^\perp, X_{\alpha}) = (0) \]

so $[X_{2\alpha}^\perp, X_{2\alpha}^\perp] \subseteq X_{2\alpha}^\perp$. The other result is similar.

(4.3) For $v \in L$ write $vC$ for $\langle v, c_\alpha \rangle$. Let $V_0 = \{ v \in N_\alpha \mid \langle v, b \rangle \in M_\alpha \}$. (Thus
Choose \( v_1, \ldots, v_n \in N_\alpha \) so that \( \{v_1 + M_\alpha, \ldots, v_n + M_\alpha\} \) is a basis for \( V_0/M_\alpha \). Define \( V_{i+1} = V_i + V_iC \) for \( 0 < i < p - 2 \). Then by Lemma 3.4 each \( V_i, 0 < i < p - 1 \), is an \( \text{ad} \ H_\alpha \) subspace of \( L_\alpha, N_\alpha = V_{p-1} \), and \( N_\alpha/M_\alpha \) has basis \( \{v_i C^j + M_\alpha | 1 < i < n, 0 < j < p - 1\} \).

Thus

\[(4.3.1) \quad [V_j, H_\alpha] \subseteq V_j \quad \text{for } 0 < j < p - 1\]

and, since \( H = c_\alpha F + H_\alpha \),

\[(4.3.2) \quad [V_i, H] \subseteq V_{i+1} \quad \text{for } 0 < j < p - 2.\]

Taking annihilators gives

\[(4.3.3) \quad L_{-\alpha} \supseteq M_\alpha \supseteq V_0 \supseteq V_1 \supseteq \cdots \supseteq V_{p-1} = N_\alpha \supseteq K_{-\alpha},\]

\[(4.3.4) \quad [V_j, H_\alpha] \subseteq V_j \quad \text{for } 0 < j < p - 1\]

and

\[(4.3.5) \quad [V_j, H] \subseteq V_{j+1} \quad \text{for } 1 < j < p - 1.\]

(4.4) Since \( H = c_\alpha F + H_\alpha \), if \( x \in L_\beta, y \in L_{-\beta} \) we have \([x, y] = c_\alpha u + h \) for some \( u \in F, h \in H_\alpha \). Then \( (x, y) = \alpha([b[x, y]]) = u \). Hence we have

\[(4.4.1) \quad [x, y] \in c_\alpha (x, y) + H_\alpha \quad \text{for all } x \in L_\beta, y \in L_{-\beta}.\]

(4.5) Choose \( d_j \in M_\alpha \) for \( 1 < i < n, 0 < j < p - 1, \) so that

\[(4.5.1) \quad (v_i C^j, d_j) = \delta_{ir} \delta_{js}.\]

Then \( V_{j+1}/N_\alpha \) has basis \( \{d_{ik} + N_\alpha | 1 < i < n, j < k < p - 1\} \).

Let \( W \) denote the linear span of \( \{v_i C^j | 1 < i < n, 1 < j < p - 1\} \). Define

\[\Phi: W \wedge W \to [W, W] \subseteq [N_\alpha, N_\alpha] \subseteq N_{2\alpha}\]

by

\[(w_1 \wedge w_2)\Phi = [w_1, w_2].\]

Define

\[\Psi: W \wedge W \to [[W, W]N_\alpha] \subseteq [[[N_\alpha, N_\alpha]N_\alpha] \subseteq N_{2\alpha}\]

by

\[(w_1 \wedge w_2)\Psi = [[w_1, w_2] v_1].\]

Let \( \overline{\Phi}: W \wedge W \to N_{2\alpha}/M_{2\alpha} \) denote the composition of \( \Phi \) with the canonical epimorphism and \( \overline{\Psi}: W \wedge W \to N_{3\alpha}/M_{3\alpha} \) denote the composition of \( \Psi \) with the canonical epimorphism.

(4.6) Let \( w \in W \wedge W \) and \( e_1, e_2 \in V_{j+1} \). Then by the Jacobi identity
\[(w\Psi, [[e_1, d_{11}], e_2]) = ([w\Phi, v_1], [[e_1, d_{11}], e_2]) = A + B + D + E\]

where

\[A = (w\Phi, [v_1, [[e_1, d_{11}], e_2]]), \quad B = ([[w\Phi, [e_1, d_{11}]], e_2], v_1),\]

\[D = ([e_1, [d_{11}, [w\Phi, e_2]]], v_1), \quad E = ([[e_1, [w\Phi, e_2]], d_{11}], v_1).\]

Now

\[[v_1, [[e_1, d_{11}], e_2]] \subseteq [L_\alpha[[M_\alpha^+, M_\alpha^+] M_\alpha^+]] \subseteq [M_\alpha^+, M_\alpha^+] \subseteq M_\alpha^+\]

by Lemma 4.2. Thus \(A \in (w\Phi, M_2^+).\)

Since \([w\Phi, [e_1, d_{11}]] \subseteq \overline{H}\) we have

\[B \subseteq ([\overline{H}, V_0^+], V_0) \subseteq (V_0^+, V_0) = (0) \quad \text{by (4.3.4)}.
\]

Similarly \([d_{11}, [w\Phi, e_2]] \subseteq \overline{H}\) so \(D = 0.\)

Finally, since (by (4.4.1))

\[[e_1, [w\Phi, e_2]] \subseteq (e_1, [w\Phi, e_2])e_\alpha + \overline{H_\alpha},\]

and since

\[[[\overline{H_\alpha}, d_{11}], v_1] \subseteq ([\overline{H_\alpha}, V_0^+], V_0) \subseteq (V_0^+, V_0) = (0)\]

(by (4.3.3)), while

\[[[e_\alpha, d_{11}], v_1] = -(d_{11}, [e_\alpha, v_1]) = -(e_\alpha C, d_{11}) = -1,\]

we have \(E = ([[w\Phi, e_2], e_1]).\)

Let \(J = \{(s, j, r, i) \in \mathbb{Z}^4 | 1 < r < s < p - 1, 1 < i, j < n, \text{ and } i < j \text{ if } r = s\}.\) Then \(W \wedge W\) has basis \([v_j C' \wedge v_j C'|(s, j, r, i) \in J].\) Order \(J\) lexicographically. Then we have

\[(4.7) \text{LEMMA. Let} \]

\[G = ([[v_j C' \wedge v_j C'], \Phi, d_\alpha], d_{j,s+1})\]

where \((s, j, r, i)\) and \((s', j', r', i')\) \(\in J.\) Then \(G = -1\) if \((s, j, r, i) = (s', j', r', i')\)

and \(G = 0\) if \((s, j, r, i) > (s', j', r', i').\)

PROOF. By the Jacobi identity

\[G = ([[v_j C', v_j C'], d_\alpha], d_{j,s+1})\]

\[= ([[v_j C', d_\alpha], v_j C'], d_{j,s+1}) + ([[v_j C', v_j C', d_\alpha], d_{j,s+1}]).\]

Assume \((s, j, r, i) > (s', j', r', i').\) Since (by (4.3.1)) \([\overline{H_\alpha}, v_j C'] \subseteq V_\alpha'\) and \([v_j C', \overline{H_\alpha}] \subseteq V_\alpha' \subseteq V_\alpha',\) we see from (4.4.1) that

\[G = -(v_j C', d_\alpha)(v_j C'+1, d_{j,s+1}) + (v_j C', d_\alpha)(v_j C'+1, d_{j,s+1}).\]
The first summand is $-1$ if $(s, j, r, i) = (s', j', r', i')$ and is zero otherwise. If the second summand is nonzero then $s + 1 = r' + 1 < s' + 1 = r + 1$ so $r = s = r' = s'$. But then $j = i' < j' = i < j$, a contradiction. Thus the second summand is zero, proving the lemma.

(4.8) Lemma. $\ker \Phi \cap \ker \Psi = (0)$.

Proof. Suppose $w \in \ker \Phi \cap \ker \Psi$. Since $w\Psi = 0$ we have $w\Psi \in M_{3a}$, so by Lemma 4.2, if $e_1, e_2 \in V^+$, we have
\[(w\Psi, [[e_1, d_{11}], e_2]) \in (M_{3a}, M_{3a}^+) = (0).\]
Also, since $w\Phi = 0$ we have $w\Phi \in M_{2a}$. Thus in (4.6) $A \in (M_{2a}, M_{2a}^+) = (0)$. Thus
\[0 = (w\Psi, [[e_1, d_{11}], e_2]) = ([w\Phi, e_2], e_1).\]

If $w \neq 0$ then
\[w = u(v_i C' \wedge v_j C^+) \sum_{I \in J, I < (s, j, r, i)} w_I,
\]
where $0 \neq u \in F$ and where $w_I$ is a scalar multiple of the basis element corresponding to $I \in J$. But then Lemma 4.7 shows that $0 = [[w\Phi, d_r], d_{r+1}] = -u$. This contradiction shows $w = 0$.

(4.9) Corollary. $\dim N_{2a}/M_{2a} + \dim N_{3a}/M_{3a} > 2(\dim N_{a}/M_{a})$.

Proof. Since, by Lemma 4.8, $\Phi \oplus \Psi$ injects $W \wedge W$ into $N_{2a}/M_{2a} \oplus N_{3a}/M_{3a}$, it is sufficient to show that $\dim W \wedge W > 2(\dim N_{a}/M_{a})$. Now we have $\dim N_{a}/M_{a} = np$ (by (4.3)) and $\dim W = n(p - 3)$. Hence the corollary holds if $n(p - 3)(n(p - 3) - 1)/2 > 2np$ or, equivalently, if $n(p - 3)^2 > 5p - 3$. Since $p > 7$ this is valid for all $n > 1$.

(4.10) Corollary. $n_{2a} + n_{3a} > 2n_a$.

Proof. Apply Corollary 4.9 to each quotient in a composition series from $K_a$ to $L_a$.

5. Dimension arguments. II.

(5.1) We continue to let $F$, $L$, and $H$ be as in Theorem 2.1 and to assume $H \neq T + I$. In addition, we assume that (3.3.2) holds. Our object is to show that $n_\beta + n_{\beta - a} > 2n_\beta$, contradicting (3.5.2).

(5.2) If $v \in L_\beta$ write $vC = [v, c_\beta]$. Let
\[L_\beta = W_{t} \supseteq W_{t-1} \supseteq \cdots \supseteq W_1 \supseteq W_0 = K_\beta,
\]
where each $W_i$ is a $H$-submodule of $L_\beta$ and each $W_{i+1}/W_i$ is an irreducible $H$-module. Let $W_{i, 0} = \{w \in W_i | [b, w] \in W_{i-1}\}$ for $1 < i < t$ and $W_{0, 1} = W_{1,0} + W_{1,0}C$ for $0 < j < p - 2$. Then by Lemma 3.4 we have a chain of
ad $\overline{H}_\beta$ invariant subspaces:

$$W_i = W_{i,p-1} \supseteq W_{i,p-2} \supseteq \cdots \supseteq W_{i,1} \supseteq W_{i,0} \supseteq W_{i-1}.$$  

Furthermore, if $\{v_{ij} + W_{i-1}|1 < j < n_i\}$ is a basis for $W_{i,0}/W_{i-1}$, then

$$\{v_{ij} C^k + W_{i-1}|1 < j < n_i, 0 < k < p - 1\}$$

is a basis for $W_i/W_{i-1}$, so that

$$\{v_{ij} C^k + K_\beta|1 < i < t, 1 < j < n_i, 0 < k < p - 1\}$$

is a basis for $L_\beta/K_\beta$. Thus, if $n = \sum_{i=1}^t n_i$, we have $n_\beta = pn$. Finally, again by Lemma 3.4, we have

$$[W_{i,j}, \overline{H}_\beta] \subseteq W_{i,j} \text{ for all } i,j.$$

For $1 < i < t, 1 < j < n_i, 0 < k < p - 1$ choose $d_{i,j,k} \in L_\beta$ so that

$$(v_{i,j} C^k, d_{i,j,k}) = \delta_{ir} \delta_{js} \delta_{sk}.$$  

(5.3) Lemma. There exist elements $w_1, \ldots, w_p \in L_\alpha$ and $u_1, \ldots, u_p \in L_{-\alpha}$ such that

$$\beta \left( [b[w_i, u_j]] \right) = \delta_{ij}$$

and, hence,

$$[w_i, u_j] \in \delta_{ij} c_\beta + \overline{H}_\beta.$$

Proof. Let $Q_\alpha = \{x \in L_\alpha|\beta(b[x, L_{-\alpha}]) = (0)\}$. Then $Q_\alpha$ is an $\overline{H}$-submodule of $L_\alpha$. Applying Lemma 3.4 to an irreducible submodule of $L_\alpha/Q_\alpha$ (which is nonzero since $(\beta, \alpha) \in S$) gives the result.

(5.4) Let $J'$ denote $\{(r, i, j, k)|1 < r < p, 1 < i < t, 1 < j < n_i, 0 < k < p - 1\}$ and $J = \{(r, i, j, k) \in J'|k \neq p - 1\}$.

Lemma. The $n p^2$ by $n p^2$ matrix

$$\left( \left( [w_r, u_r] v_{i,j} C^k, d_{i',j',k'} \right) \right)_{(r,i,j,k),(r',i',j',k') \in J'}$$

has rank $\geq n(p - 1)$.

Proof. It is sufficient to show that the rows corresponding to $(r, i, j, k) \in J$ are linearly independent. Thus assume that for each $(r, i, j, k) \in J$ we have $a_{r,i,j,k} \in F$ such that

$$0 = \sum_j a_{r,i,j,k} \left( [w_r, u_r] v_{i,j} C^k, d_{i',j',k'} \right)$$

for all $(r', i', j', k') \in J'$. We must show that all $a_{r,i,j,k} = 0$.

Assume that $1 < q < t$ and $0 < u < p - 2$ and that $a_{r,i,j,k} = 0$ for all $(r, i, j, k) \in J$ with $i > q$ or $i = q$ and $k > u$. (This condition is vacuous if $q = t, u = p - 2$.) We will show $a_{r,q,i,u} = 0$ for all $r$ and $j$ and, hence, by
induction that $a_{r,i,j,k} = 0$ for all $(r, i, j, k) \in J$.

We have, for all $r', 1 < r' < p$, and all $j', 1 < j' < n_q$,

$$0 = \sum_j a_{r,i,j,k} \left( \left[ \left[ w_r, u_r \right] v_{i,j} C^k \right], d_{q,j,u+1} \right).$$

Since $a_{r,i,j,k} = 0$ if $i > q$ and

$$\left( \left[ \left[ w_r, u_r \right] v_{i,j} C^k \right], d_{q,j,u+1} \right) \in (W_i, d_{q,j,u+1}) = (0)$$

if $i < q$, we have

$$0 = \sum_{r,i,j,k} a_{r,i,j,k} \left( \left[ \left[ w_r, u_r \right] v_{q,j} C^k \right], d_{q,j,u+1} \right).$$

Since, by (5.2) and (5.3),

$$\left[ \left[ w_r, u_r \right] v_{q,j} C^k \right] \in [\delta_{r,c}, v_{q,j} C^k] + [H_{c}, W_{q,k}] \subseteq -\delta_{r,c} v_{q,j} C^{k+1} + W_{q,k}$$

and since, by the induction assumption, $a_{r,i,j,k} = 0$ if $k > u$, we have

$$0 = -\sum_j a_{r,i,j,k} (v_{q,j} C^{u+1}, d_{q,j,u+1}) = -a_{r,i,j,u}$$

as required.

(5.5) Let $x_a \in L_\alpha, x_\beta \in L_\beta, y_a \in L_{-\alpha},$ and $y_\beta \in L_{-\beta}.$ Then

$$\left[ \left[ x_a, x_\beta \right], \left[ y_a, y_\beta \right] \right] + \left[ \left[ x_a, y_a \right], \left[ x_\beta, y_\beta \right] \right]$$

$$= \left( \left[ \left[ x_a, x_\beta \right] x_a \right], y_\beta \right) + \left( y_a, \left[ \left[ x_a, x_\beta \right] y_\beta \right] \right)$$

$$+ \left( \left[ \left[ x_\beta, y_\beta \right] x_\beta \right], y_\beta \right) + \left( x_\beta, \left[ \left[ x_\beta, y_\beta \right] y_\beta \right] \right)$$

$$= \left( \left[ \left[ x_a, x_\beta \right] y_a \right], y_\beta \right) + \left( \left[ \left[ x_\beta, y_\beta \right] x_\beta \right], y_\beta \right) \quad \text{(since } \alpha, \alpha \in S \text{)}$$

$$= \left( \left[ \left[ x_a, y_a \right] x_\beta \right], y_\beta \right).$$

Now

$$n_{\beta+a} = \dim(L_{\beta+a}/K_{\beta+a}) \geq \text{rank}([x_a, x_\beta], [y_a, y_\beta]),$$

where $x_a, x_\beta, y_a, y_\beta$ run over subsets of the appropriate $L$'s. A similar remark holds for $n_{\beta-a}.$ Also, if $A$ and $B$ are matrices then

$$\text{rank } A + \text{rank } B \geq \text{rank}(A + B).$$

Then setting $x_a = w_r, y_a = u_r, x_\beta = v_{i,j} C^k,$ and $y_\beta = d_{i,j,k},$ where $(r, i, j, k)$ and $(r', i', j', k') \in J'$, we see from Lemma 5.4 that $n_{\beta+a} + n_{\beta-a} > n(p - 1)p.$ Since $p - 1 > 2$ we have $n_{\beta+a} + n_{\beta-a} > 2np = 2n_\beta,$ as required.

REFERENCES


Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903