AN ANALOGUE OF THE JACOBSON-MOROZOV THEOREM
FOR LIE ALGEBRAS OF REDUCTIVE GROUPS
OF GOOD CHARACTERISTICS

ALEXANDER PREMET

Abstract. Let g be the Lie algebra of a connected reductive group G over an
algebraically closed field of characteristic p > 0. Suppose that G(1) is simply
connected and p is good for the root system of G. Given a one-dimensional
torus λ ⊂ G let g(λ, i) denote the weight component of Ad(λ) corresponding
to weight i ∈ X(λ) ∼= Z. It is proved in the paper that, for any nonzero
nilpotent element e ∈ g, there is a one-dimensional torus λe ⊂ G such that
e ∈ g(λe, 2) and Ker ad e ⊆ ⊕i≥0 g(λe, i).

1. Introduction

Let G be a connected reductive group over an algebraically closed field K of
characteristic p > 0 and g = Lie(G). The group G acts on g via the adjoint
representation Ad. Given a one-dimensional torus k ⊂ G denote by g(λ, i) the
weight component of Ad λ corresponding to weight i ∈ X(λ) ∼= Z. Throughout
the paper we assume that p = char(K) is a good prime number for G (see (2.1)
for a precise definition). Note that if p > 5, then p is good for any reductive
group over K.

The Lie algebra g has a canonical [p]-operation invariant under the adjoint
action of G. An element x ∈ g is said to be nilpotent or [p]-nilpotent (resp.,
semisimple or [p]-semisimple) if x^p^e = 0 for some e ∈ Z+ (resp., if x lies
in the p-envelope of x^p in g). The group G acts on the set of all nilpotent
elements of g. The orbits of this action are classified by Bala-Carter under the
assumption that p > 0 (see [1, 2]). Their results are extended by Pommerening
to the case when p is a good prime number for G (see [13, 14]). Nothing seems
to be published about nilpotent orbits of the Lie algebras of type E7 and E8
for p ≤ 5 though it follows from [20] that the number of nilpotent orbits of g
is finite for any p > 0.

Let p ≫ 0 and e a nonzero nilpotent element of g. The Jacobson-Morozov
theorem [21, III, 4.3] says that g contains a subalgebra s such that s ∼= sl(2),
e ∈ s and g is a completely reducible s-module. Moreover, a standard Lie
theory argument shows that there is a connected subgroup S ⊆ G such that
s = Lie(S) and any s-submodule of g is S-stable. There exist a maximal
unipotent subgroup Ue ⊂ S and a one-dimensional torus λe ⊂ S satisfying

Received by the editors August 15, 1994.
1991 Mathematics Subject Classification. Primary 17B10, 20G05.
Key words and phrases. Nilpotent element, Lie algebra, reductive group, prime characteristic.

©1995 American Mathematical Society
\[ \text{Lie}(U_e) = K e, \quad \lambda_e(t)U_e\lambda_e(t)^{-1} = U_e. \]
Let \( V \) be an irreducible \( s \)-submodule of \( g \). As \( V \) is \( S \)-stable and \( \lambda_e U_e \) is a Borel subgroup of \( S \), there exists \( k \in \mathbb{Z}_+ \) such that \( V \cap \text{Ker ad } e \subseteq g(\lambda_e, k) \). As \( g \) is a completely reducible \( s \)-module, we obtain that for any nonzero nilpotent element \( e \in g \), there exists a one-dimensional torus \( \lambda_e \subseteq G \) such that \( e \in g(\lambda_e, 2) \) and \( \text{Ker ad } e \subseteq \bigoplus_{i \geq 0} g(\lambda_e, i) \).

The purpose of this paper is to extend this result to the case of an arbitrary good \( p \). Note that in this setting the result is known to be true provided all simple components of \( G \) are groups of classical type (see [21, IV, §§1,2]). So in the sequel we mostly deal with the groups of exceptional types. Throughout the paper we assume that the derived subgroup of \( G \) is simply connected.

In proving our main theorem we crucially use Pommerening’s classification of nilpotent elements of \( g \) and the Kempf-Rousseau theory as exposed in [18]. The motive for this investigation originated in the representation theory of \( g \). It is well known [6] that all irreducible representations of \( g \) are of finite dimension. To each irreducible \( g \)-module \( V \), one can assign in a canonical way a linear function \( \chi \in g^* \) called the \( p \)-character of \( V \). The ideal \( I_\chi \) of the universal enveloping algebra \( U(g) \) generated by the central elements of the form \( x^p - x^p \chi(x)^p \cdot 1 \), where \( x \in g \), acts trivially on \( V \). Given a restricted subalgebra \( a \subseteq g \) denote by \( u_\chi(a) \) the associative subalgebra of \( U(g)/I_\chi \) generated by \( a \). It follows from the PBW-theorem that \( \dim u_\chi(a) = p^{\dim a} \).

In [23], Kac and Weisfeiler conjectured that if \( G \) is simple and \( g \) admits a nondegenerate \( G \)-invariant trace form, then any irreducible \( g \)-module with \( p \)-character \( \chi \) has dimension divisible by \( p^{(\dim \Omega(\chi))/2} \) where \( \Omega(\chi) \) is the orbit of \( \chi \) under the coadjoint action of \( G \). As I recently observed (see [15]), for any \( \chi \in g^* \), there exists a restricted nilpotent subalgebra \( \tilde{m}_\chi \) of \( g \) such that \( \dim \tilde{m}_\chi = \frac{1}{2} \dim \Omega(\chi) \) and any irreducible, faithful \( g \)-module with \( p \)-character \( \chi \) is free over \( u_\chi(\tilde{m}_\chi) \). This result proves the Kac-Weisfeiler conjecture. In constructing the subalgebra \( \tilde{m}_\chi \), I crucially use the main result of this paper.

Concluding the introduction, note that our main result is no longer true for some simple groups of adjoint type. Indeed, let \( e \) be the image of a nilpotent Jordan block of order \( p \) in \( gl_p(K)/\mathfrak{d} = \text{Lie}(G) \) where \( G = PGL_p(K) \). It is easily seen that the preimage of \( \text{Ker ad } e \) in \( gl_p(K) \) acts irreducibly on the standard \( gl_p(K) \)-module of dimension \( p \). It follows that \( \text{Ker ad } e \not\subseteq \text{Lie}(P) \) for any parabolic subgroup \( P \) of \( G \).

2. Dynkin tori for nilpotent elements

2.1. Let \( G \) be a connected reductive algebraic group over an algebraically closed field \( K \) of characteristic \( p > 0 \). We assume that \( p \) is good for \( G \); i.e. \( p \) is greater than any coefficient of any positive root of the root system \( R = R(G) \) relative to a basis of simple roots in \( R \).

Given a maximal torus \( T \) in \( G \) decompose \( g \) into weight spaces under the adjoint action of \( T \) giving a Cartan decomposition

\[ g = t \bigoplus_{\alpha \in R} K e_\alpha \]

where \( t = \text{Lie}(T) \). Let \( B = \{ \alpha_1, \alpha_2, \ldots, \alpha_l \} \) be a basis of simple roots in \( R \), \( R_+ \) the corresponding system of positive roots, \( \{ \bar{\omega}_1, \bar{\omega}_2, \ldots, \bar{\omega}_l \} \) the corresponding system of fundamental weights in the lattice of the rational characters.
of $T$. Everywhere below the indexing of the simple roots in $B$ corresponds to Bourbaki's tables [4, VI, Tables I-IX].

Given a subset $B_J \subset B$ one can define the standard parabolic subgroup $P_J$ of $G$ with Levi decomposition $P_J = U_J L_J$. Following Carter [5] define a function $\eta_J : R \to 2\mathbb{Z}$ by

$$
\eta_J(\alpha) = \begin{cases} 
0 & \text{if } \alpha \in B_J, \\
2 & \text{if } \alpha \in B \setminus B_J
\end{cases}
$$

and extending to arbitrary root by linearity. Denote

$$
g_J(i) = \begin{cases} 
\sum_{\eta_J(\alpha) = i} \lambda \alpha & \text{if } i \neq 0, \\
\bigoplus_{\eta_J(\alpha) = 0} K e_\alpha & \text{if } i = 0.
\end{cases}
$$

Then one has $g = \bigoplus_j g_J(i)$, $[g_J(i), g_J(k)] \subseteq g_J(i + k)$ and $\bigoplus_{i \geq 0} g_J(i) = \text{Lie}(P_J)$. It is well known [5, p. 166] that $\dim g_J(0) = \dim L_J$, $\dim g_J(2) = \dim U_J/U_J^{(1)}$ and $\dim L_J \geq \dim U_J/U_J^{(1)}$.

For $G$ semisimple, a parabolic subgroup $P$ is called distinguished if $\dim P/U_P = \dim U_P/U_P^{(1)}$ where $U_P$ is the unipotent radical of $P$. Any parabolic subgroup of $G$ is conjugate in $G$ with precisely one of the standard parabolic subgroups. A standard parabolic subgroup $P_J$ is distinguished if and only if $\dim g_J(0) = \dim g_J(2)$.

2.2. Let $p$ denote the Lie algebra of a parabolic subgroup $P$ of $G$. Set $n_P = \text{Lie}(U_P)$. An element $x \in n_P$ is called a Richardson element of $P$ if the orbit $(\text{Ad } P) \cdot x$ is dense in $n_P$. Clearly, all Richardson elements of $p$ are conjugate with respect to the adjoint action of $P$.

If $P = P_J$ for some $J \subseteq \{1, 2, \ldots, l\}$ we arrange $p = p_J$ and $n = n_J$. By [5, Proposition 5.8.5] any $U_J$-orbit containing a Richardson element of $p_J$ intersects with the graded subspace $g_J(2)$.

Given $x$ in $\mathfrak{g}$ denote by $Z_G(x)$ (resp., by $\mathfrak{z}_G(x)$) the centralizer of $x$ in $G$ (resp., in $\mathfrak{g}$). Clearly,

$$
\text{Lie}(Z_G(x)) = \text{Lie}(Z_G(x)^0) \subseteq \mathfrak{z}_G(x)
$$

(the symbol $H^0$ stands for the connected component of a Zariski closed subgroup $H \subset G$). By [21, I, \S 5], $\text{Lie}(Z_G(x)) = \mathfrak{z}_G(x)$ provided $\mathfrak{g}$ admits a nondegenerate trace form associated with a rational representation of $G$. If $x$ is a Richardson element of a parabolic subalgebra $p$, then $Z_G(x)^0 \subseteq P$ (see [5, Corollary 5.2.4]).

2.3. In the next three subsections we follow Slodowy's exposition [18].

Denote by $X_\ast(T) = \text{Hom}(\mathbb{G}_m, T)$ the group of all one-parameter subgroups of $T$ and by $X^\ast(T) = \text{Hom}(T, \mathbb{G}_m)$ the group of the rational characters of $T$. As $T \cong (\mathbb{G}_m)^l$, one has $X_\ast(T) \cong \mathbb{Z}^l \cong X^\ast(T)$. The pairing $X_\ast(T) \times X^\ast(T) \to \mathbb{Z}$ given by

$$
(\lambda, \omega) \mapsto \langle \lambda, \omega \rangle,
$$

$\omega(\lambda(t)) = t^{\langle \lambda, \omega \rangle}$, is nondegenerate. The set $X_\ast(G)$ of all one-dimensional tori in $G$ is the union $\bigcup_H X_\ast(H)$ where $H$ runs over all maximal tori of $G$.

The Weyl group $W = N_G(T)/T$ acts on both $X_\ast(T)$ and $X^\ast(T)$. By fixing a $W$-invariant positively defined symmetric bilinear form $X_\ast(T) \times X_\ast(T) \to \mathbb{Z}$ one can identify the dual vector spaces $X^\mathbb{R} = X^\ast(T) \otimes_{\mathbb{Z}} \mathbb{R}$ and $X_\mathbb{R} = X_\ast(T) \otimes_{\mathbb{Z}} \mathbb{R}$.
To simplify notation we denote the scalar product on \( X_\mathbb{R} \) by the above symbol \( \langle , \rangle \). Let \( \|\cdot\| \) denote the corresponding norm mapping: \( \|x\| = \sqrt{\langle x, x \rangle} \), \( x \in X_\mathbb{R} \).

Using the \( W \)-invariance of \( \langle , \rangle \) and the fact that

\[
X_*(G) = \bigcup_{g \in G} X_*(g^{-1}Tg)
\]

one can extend the norm \( \|\cdot\| \) up to a well-defined \( G \)-invariant mapping from \( X_*(G) \) into \( \mathbb{R} \). If \( \lambda \in X_*(G) \) and \( g \in G \) is such that \( \text{Int}(g) \circ \lambda \in X_*(T) \), then (by definition)

\[
\|\lambda\| = \|\text{Int}(g) \circ \lambda\|.
\]

To each one-dimensional torus \( \lambda \in X_*(G) \), one can assign a parabolic subgroup \( P(\lambda) \) with Levi decomposition \( P(\lambda) = U(\lambda)L(\lambda) \). If \( \lambda \in X_*(T) \), then

\[
\text{Lie}(L(\lambda)) = \bigoplus_{\{\lambda, \alpha\} = 0} \mathbb{K} e_\alpha, \quad \text{Lie}(U(\lambda)) = \bigoplus_{\{\lambda, \alpha\} > 0} \mathbb{K} e_\alpha.
\]

2.4. Let \( \rho: G \to GL(V) \) be a finite-dimensional rational representation of \( G \) in a vector space \( V \) over \( K \). If \( \lambda \in X_*(G) \), then the induced action \( \rho \circ \lambda: G_m \to GL(V) \) turns \( V \) into a \( \mathbb{Z} \)-graded vector space: \( V = \bigoplus_{i \in \mathbb{Z}} V_i \) where

\[
V_i = \{ v \in V | \rho(\lambda(t))(v) = t^i v, \forall t \in G_m \}.
\]

If \( \lambda \in X_*(T) \) and \( V = \bigoplus_{\chi \in X^*(T)} V_\chi \) is the weight space decomposition of \( V \) with respect to \( T \), then

\[
V_i = \bigoplus_{\{\lambda, \chi\} = i} V_\chi.
\]

It is easy to check that, for any \( \lambda \in X_*(G) \), the parabolic subgroup \( P(\lambda) \) defined in (2.3) preserves the subspaces \( V_{(i)} = \bigoplus_{j \geq i} V_j, \ i \in \mathbb{Z} \).

Let \( v \in V \) and \( v = \sum_{i \in \mathbb{Z}} v_i \). Set

\[
m(v, \lambda) = \max \{ i \in \mathbb{Z} | v \in V_{(i)} \}
\]

and

\[
\text{Supp}_T(v) = \{ \chi \in X^*(T) | v_\chi \neq 0 \}.
\]

By the above, \( m(v, \lambda) = \min_{\chi \in \text{Supp}_T(v)} \langle \lambda, \chi \rangle \).

A vector \( v \in V \) is called \textit{instable} with respect to a closed subgroup \( H \subset G \) (or \( H \)-instable) if the closure \( \overline{H \cdot v} \) of the orbit \( H \cdot v \subset V \) contains \( 0 \). If \( 0 \notin \overline{H \cdot v} \), then \( v \) is said to be \textit{semistable} with respect to \( H \) (or \( H \)-semistable). A one-dimensional torus \( \lambda \in X_*(G) \) is called an \textit{optimal} torus for a \( G \)-instable vector \( v \in V \) if

\[
\frac{m(v, \lambda)}{\|\lambda\|} \geq \frac{m(v, \mu)}{\|\mu\|}
\]

for any nonzero \( \mu \in X_*(G) \). An element \( \lambda \in X_*(G) \) is called \textit{primitive} if there is no \( \mu \in X_*(G) \) with \( \lambda = n \mu, \ n \in \mathbb{Z}, \ n \geq 2 \).

Given a \( G \)-instable vector \( v \in V \) define

\[
\Lambda_v = \{ \lambda \in X_*(G) | \lambda \text{ is primitive and optimal for } v \}.
\]

**Theorem 2.1** (Kempf [10], Rousseau [16]). Let \( v \in V \) be \( G \)-instable. Then

(i) \( \Lambda_v \neq \emptyset \) and there exists a parabolic subgroup \( P(v) \subset G \) such that \( P(v) = P(\lambda) \) for any \( \lambda \in \Lambda_v \).

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
(ii) The stabilizer $G_v = \{ x \in G | \rho(x)(v) = v \}$ is contained in $P(v)$.

By (2.4), any $\lambda \in X_*(G)$ defines a $\mathbb{Z}$-grading of $V$. Being the centralizer of $\lambda$, the Levi subgroup $L(\lambda) \subset P(\lambda)$ preserves all graded components $V_i$, $i \in \mathbb{Z}$, of this $\mathbb{Z}$-grading. Denote

$$L_n(\lambda) = \{ g \in L(\lambda) | \det(\rho(g)(v_i)) = 1 \}.$$

**Proposition 2.2** (Kirwan [11], Ness [12]). Let $n > 0$, $v \in V_i$ and $v \neq 0$. If $v$ is $L_n(\lambda)$-semistable, then $\lambda$ is an optimal torus for $v$.

Note that any vector $v \in V_i$ must be $G$-instable if $i \neq 0$.

2.5. We will make use of Theorem 2.1 and Proposition 2.2 in the case $V = g$, $\rho = \text{Ad}$. The adjoint action of $\lambda \in X_*(G)$ turns $g$ into a $\mathbb{Z}$-graded Lie algebra:

$$g = \bigoplus_{i \in \mathbb{Z}} g(i), \quad [g(i), g(j)] \subseteq g(i+j).$$

If $A \in g(2)$, then $A$ is a nilpotent element of $g$ and $(\text{ad}A)^i$ maps $g(-i)$ into $g(i)$. Suppose that $g$ admits a nondegenerate trace form $b: g \times g \rightarrow K$ associated with a rational representation $\rho: G \rightarrow GL(V)$:

$$b(X, Y) = \text{tr} \rho(X)\rho(Y) \quad (X, Y \in g)$$

where $d\rho$ denotes the differential of the rational representation $\rho$. Given $X \in g(2)$ define a bilinear form $b_X: g(-2) \times g(-2) \rightarrow K$ by setting

$$b_X(Y, Y') = b([X, Y], [X, Y']) = -b(Y, (\text{ad}X)^2 \cdot Y').$$

Set $f(X) = \det(b_X)$.

**Lemma 2.3** (Kac [9], Slodowy [18]).

(i) The polynomial function $f: g(2) \rightarrow K$ is $L_2(\lambda)$-invariant.

(ii) If the map $(\text{ad}A)^2: g(-2) \rightarrow g(2)$ is surjective, then $f(A) \neq 0$. In particular, $A$ is semistable with respect to $L_2(\lambda)$.

Lemma 2.3 together with Proposition 2.2 and Theorem 2.1 (ii) implies that if $(\text{ad}A)^2: g(-2) \rightarrow g(2)$ is a surjective map, then $\lambda \in X_*(G)$ is an optimal torus for $A$ and so $Z_G(A) \subseteq P(\lambda)$. Since Lie($P(\lambda)$) = $\bigoplus_{i \geq 0} g(i)$, this forces $\mathfrak{z}_g(A) \subseteq \bigoplus_{i \geq 0} g(i)$ (see (2.2) for more detail).

2.6. From now on we suppose that Lie($G^{(1)}$) is isomorphic to the Lie algebra of a simply connected group isogeneous to $G^{(1)}$. Given $\lambda \in X_*(G)$ decompose $g = \text{Lie}(G)$ into weight spaces under the adjoint action of $\lambda$:

$$g = \bigoplus_{i \in \mathbb{Z}} g(i).$$

By (2.3), Lie($P(\lambda)$) = $\bigoplus_{i \geq 0} g(i)$, Lie($U(\lambda)$) = $\bigoplus_{i \geq 0} g(i)$ and Lie($L(\lambda)$) = $g(0)$.

**Definition 2.4.** A one-dimensional torus $\lambda \in X_*(G)$ is called a *Dynkin torus* for a nilpotent element $e \in g$ if $e \in g(2)$ and $\mathfrak{z}_g(e) \subseteq \bigoplus_{i \geq 0} g(i)$.

The rest of the paper is devoted to proving the following
Theorem 2.5. Any nonzero nilpotent element $e$ of $g$ has at least one Dynkin torus.

Let $g' = \text{Lie}(G'(1))$. Clearly, $g = t + g'$. It is well known that the canonical $[p]$-operation of $g$ is bijective on $t$. Since $g'$ is a restricted ideal of $g$, Jacobson's identity [8, V, §7] yields that any nilpotent element of $g$ lies in $g'$. Suppose that a nilpotent element $e \in g'$ has a Dynkin torus $\lambda \in X_*(G'(1))$. Decompose $g$ into weight spaces under the adjoint action of $\lambda$: $g = \bigoplus_{i \in \mathbb{Z}} B(i)$. Let $T_1$ be a maximal torus of $G$ containing $\lambda$ and let $t_1 = \text{Lie}(T_1)$. As $t_1 \subseteq g(0)$ and $g = t_1 + g'$, we have $g = g(0) + g'$. This yields that $g(i) \subseteq g'$ for each $i \neq 0$. As $\lambda$ is a Dynkin torus for $e \in g'$, $\delta_g(e) \subseteq \sum_{i \geq 0} g(i)$ and $e \in g(2)$. Hence $\lambda$ preserves $\delta_g(e)$. But then

$$\delta_g(e) = \delta_g(e) \cap g(0) \bigoplus \sum_{i \neq 0} \delta_g(e) \cap g(i) \subseteq g(0) + \delta_g(e) \cap g' \subseteq \sum_{i \geq 0} g(i)$$

showing that $\lambda$ is a Dynkin torus for $e \in g$. Thus, we may suppose that $G$ is semisimple and simply connected.

Assuming that $G = G^{(1)}$, denote by $G_1, G_2, \ldots, G_s$ the simple (and simply connected) normal subgroups of $G$. Let $g_i = \text{Lie}(G_i), 1 \leq i \leq s$. Clearly, $g = g_1 \oplus \cdots \oplus g_s$ and $[g_i, g_j] = 0$ if $i \neq j$. If $e = e_1 + \cdots + e_s$ where $e_i \in g_i$, then $\delta_g(e) = \delta_{g_1}(e_1) \oplus \cdots \oplus \delta_{g_s}(e_s)$. Suppose that each nonzero $e_i$ has a Dynkin torus $\lambda_i \in X_*(G_i)$. Let $p_i$ be the parabolic subalgebra of $g_i$ associated with $\lambda_i$. We may assume that $e_i \neq 0$ if $i \leq s_0 \leq s$ and $e_i = 0$ if $i > s_0$. Then $p = (\bigoplus_{i \leq s_0} p_i) \oplus (\bigoplus_{i > s_0} g_i)$ is the parabolic subalgebra of $g$ associated with $\lambda = \prod_i \lambda_i \in X_*(G)$. As $\delta_{g_i}(e_i) \subseteq p_i$ for all $1 \leq i \leq s_0$, then $\delta_g(e) \subseteq p$. Moreover,

$$(\text{Ad}\lambda(t)) \cdot e = \sum_i (\text{Ad}\lambda_i(t)) \cdot e_i = t^2 e.$$ 

Hence $\lambda$ is a Dynkin torus for $e$. Thus we may assume that $G$ is simple and simply connected.

If $R$ is of type $A_n, B_n, C_n$ or $D_n$, Theorem 2.5 follows immediately from the results of Springer and Steinberg (see [21, IV]). Indeed, if $G \cong GL_n(K)$ or $R$ is of type $B_n, C_n$ or $D_n$ and $p > 2$, then $G$ admits a nondegenerate trace form (by [21, I, Lemma 5.3]). Therefore, $\text{Lie}(Z_G(e)) = \delta_g(e)$ by [21, I, Corollary 5.2] and one can apply [21, IV, §5.1.7, 2.23]. If $G = SL_n(K)$, then $G = G^{(1)}$ where $G = GL_n(K)$. Any nilpotent element $e \in g$ can be regarded as an element of $g = \text{Lie}(G) = gl_n(K)$. By [21, IV, §5.1], one can find a Dynkin torus $\lambda \in X_*(G)$ for $e \in g$ contained in $G$. Let $p$ (resp., $\tilde{p}$) denote the parabolic subalgebra of $g$ (resp., $\tilde{g}$) associated with $\lambda \in X_*(G) \subseteq X_*(\tilde{G})$. Clearly, $p = \tilde{p} \cap g$. But then $\delta_g(e) = \delta_{\tilde{g}}(e) \cap g \subseteq p \cap g = p$. Hence $\lambda \in X_*(G)$ is a Dynkin torus for $e \in g$.

2.7. Considering the remaining case of exceptional groups we will use some classification results due to Bala-Carter [5, V] and Pommerening [13, 14].

Recall that a nilpotent element $x$ in $g$ is said to be distinguished if it commutes with no nonzero semisimple element of $g$. Generalizing [1, 2], Pommerening proved (see [14, p. 377]) that any distinguished nilpotent element of $g$ is a Richardson element of a distinguished parabolic subalgebra of $g$. If $G$
is exceptional and \( p \) is good for \( G \), then the Killing form of \( g \) is nondegenerate. Applying (2.2), one can now easily observe that any distinguished nilpotent element of \( g \) has at least one Dynkin torus. Hence in proving Theorem 2.5 we may assume that \( e \in g \) is not distinguished.

Since \( \mathfrak{z}_g(e) = \text{Lie}(Z_G(e)^o) \) and \( e \) commutes with a nonzero semisimple element, the group \( Z_G(e)^o \) contains a maximal torus \( S \) of positive dimension. No generality is lost by assuming \( S \subseteq T \).

Let \( R_1 \) denote the subsystem of roots vanishing on \( S \). Set \( s = \text{Lie}(S) \). Combining [3, §9.2] and [5, §5.9] one can obtain that

\[
\text{Lie}(Z_G(S)) = g^S = \mathfrak{z}_g(s) = t \bigoplus_{\alpha \in R_1} K e_\alpha.
\]

Moreover, by [17, p. 23], there exist a system of simple roots \( \Delta \subset R \) and a subset \( J \) with \( \Delta_J \subseteq \Delta \) such that any \( \gamma \in R_1 \) is an integer linear combination of the elements from \( \Delta_J \). Hence in what follows we may suppose that \( R_1 = R_J = \{ \gamma \in R \mid \gamma = \sum_{\alpha \in J} n_\alpha \alpha, n_\alpha \in \mathbb{Z} \} \) for some \( J \subseteq \{ 1, 2, \ldots, l \} \). Thus \( \mathfrak{z}_g(s) \) coincides with the Levi subalgebra \( I_J = \text{Lie}(L_J) \) of the standard parabolic subalgebra \( p_J \). It is immediate that the Killing form of \( g \) is nondegenerate on \( I_J \). By [21, II, §5], the semisimple group \( L_J^{(1)} \) is simply connected. By Jacobson's identity [8, V, §7], \( e \in I_J^{(1)} = \text{Lie}(L_J^{(1)}) \).

If \( \mathfrak{z}(l_J^{(1)}) = 0 \), then \( l_J = s \oplus I_J^{(1)} \) and \([s, l_J] = 0 \). The group \( Z_{L_J^{(1)}}(e)^o \) is unipotent (otherwise \( S \) would be properly contained in a bigger torus in \( Z_G(e)^o \) contradicting the maximality of \( S \)). Since \( s \perp I_J^{(1)} \), \( I_J^{(1)} \) admits a non-degenerate trace form. Hence \( \text{Lie}(Z_{L_J^{(1)}}(e)^o) = \mathfrak{z}(l_J^{(1)}(e)) \). This implies that \( e \) is a distinguished nilpotent element of \( I_J^{(1)} \).

If \( \mathfrak{z}(l_J^{(1)}) \neq 0 \), then either \( p = 5 \) and \( R_J \) has a component of type \( A_4 \) or \( p = 7 \) and \( R_J \) has a component of type \( A_6 \). As \( R \) is exceptional, this yields that all components of \( R_J \) have type \( A \). Therefore, \( l_J^{(1)} \) is a direct sum of commuting ideals \( I_i \) isomorphic to \( sl_{r_i}(K) \) for some \( r_i \leq 6 \). A standard argument used above shows that \( Z_{L_J^{(1)}}(e)^o \) is unipotent. Together with [21, IV, §1] this yields that \( e \) is a regular nilpotent element of \( I_J^{(1)} \) (see [21, III]).

Combining [14, p. 377] with [21, IV, §1] we obtain now that in both cases \( e \) is a Richardson element of a distinguished parabolic subalgebra of \( I_J^{(1)} = \text{Lie}(L_J^{(1)}) \).

2.8. In what follows we may (and will) assume that there exists \( I \subseteq J \) such that \( e \) is a Richardson element of the standard parabolic subalgebra \( p_I \cap I_J \) of the Levi subalgebra \( I_J \). As \( I_J = g_J(0) \), we have

\[
p_I \cap I_J = \sum_{i \geq 0} g_J(0) \cap g_I(i).
\]

By (2.2) we can also assume that \( e \in g_J(0) \cap g_I(2) \).

We will use the \( W \)-invariant scalar product \( \langle \ , \ \rangle \) on \( X^R = \mathbb{R} \omega_1 \oplus \cdots \oplus \mathbb{R} \omega_l \) defined in [4, VI, Tables I-IX] via embedding \( X^R \) into a bigger Euclidean space. Clearly, the \( \mathbb{Q} \)-span of \( B_J \) in \( X^R \) has basis \( \{ \omega_i \mid i \in J \} \) satisfying

\[
(\omega_i, \alpha_k) = \delta_{ik}.
\]
for all \( i, k \in J \). This implies that
\[
2 \sum_{i \notin I} \omega_i' = \sum_{k \in J} \frac{2}{\alpha_k \alpha_k} m_k \alpha_k
\]
for some \( m_k \in \mathbb{Q} \). A direct computation based on the Bala-Carter classification of the distinguished parabolic subgroups (see [5, pp. 174–177]) shows that all \( m_k \)'s are positive integers (note that the classification of the distinguished parabolic subgroups given in [1, 2] remains true for any good \( p \)).

By Steinberg [22, §5], the maximal torus \( T \) is generated by the one-parameter subgroups \( h_\alpha(t) \), \( \alpha \in R \), such that
\[
(\text{Ad } h_\alpha(t)) \cdot e_\beta = t^{(\beta, \alpha)} e_\beta
\]
for all \( \alpha, \beta \in R \) (here \( (\beta, \alpha) = 2 \frac{[\beta \alpha]}{\alpha \alpha} \)). Put \( h_i(t) = h_{\alpha_i}(t) \) for each \( \alpha_i \in B \) and define \( \lambda_e \in X_+(G) \) by setting
\[
\lambda_e(t) = \prod_{k \in J} h_k(t^{m_k}) \quad (t \in \mathbb{G}_m).
\]
We intend to show that in most of the remaining cases \( \lambda_e \) is an optimal torus for \( e \) relative to the scalar product \( (\; | \; ) \). By construction,
\[
(\text{Ad } \lambda_e(t)) \cdot e_\alpha = \begin{cases} 
 e_\alpha & \text{if } \alpha \in B_I, \\
 t^2 e_\alpha & \text{if } \alpha \in B_J \setminus B_I.
\end{cases}
\]
Hence \( \lambda_e(t) \) acts on \( g_I(0) \cap g_I(i) \) by multiplying each vector by \( t^i \), \( i \in \mathbb{Z} \). In particular, \( (\text{Ad } \lambda_e(t)) \cdot e = t^2 e \). Since \( \text{Lie}(h_\alpha(t)) = K[e_\alpha, e_{-\alpha}] \) for any \( \alpha \in R \), we have
\[
\text{Lie}(\lambda_e) \subset \sum_{k \in J} K[e_{\alpha_k}, e_{-\alpha_k}] \subseteq t \cap \mathfrak{l}_J^{(1)}.
\]
Moreover, the Lie algebra \( \text{Lie}(\lambda_e) \) is spanned by \( h \in t \cap \mathfrak{l}_J^{(1)} \) such that \( [h, x] = ix \) for any \( x \in g_I(0) \cap g_I(i) \), \( i \in \mathbb{Z} \).

As \( e \) is a Richardson element of \( p_I \cap \mathfrak{l}_J^{(1)} \), a distinguished parabolic subalgebra of \( \mathfrak{l}_J^{(1)} = \text{Lie}(L_J^{(1)}) \), the map
\[
\text{ad } e : \mathfrak{l}_J^{(1)} \cap g_I(-2) \rightarrow \mathfrak{l}_J^{(1)} \cap g_I(0)
\]
is bijective (for \( \dim \mathfrak{l}_J^{(1)} \cap g_I(-2) = \dim \mathfrak{l}_J^{(1)} \cap g_I(0) \) and \( \mathfrak{z}_J(e) \subseteq \sum_{i \geq 0} g_J(0) \cap g_I(i) \) as the Killing form of \( g \) is nondegenerate on \( \mathfrak{z}_J \)). This implies that there exists \( f \in \mathfrak{l}_J^{(1)} \cap g_I(-2) \) such that \( [e, f] = h \). Clearly, \( \langle e, h, f \rangle \) is an \( sl_2 \)-triple in \( \mathfrak{l}_J^{(1)} \) (see [14]).

**Remark 2.6.** By construction, \( h^{[p]} = h \) but it may happen for some small \( p \) that \( e^{[p]} \neq 0 \) or \( f^{[p]} \neq 0 \).

**2.9.** Let \( g = \bigoplus_{i \in \mathbb{Z}} g_i \) where
\[
g_i = \{ x \in g | (\text{Ad } \lambda_e(t)) \cdot x = t^i x \text{ for all } t \in \mathbb{G}_m \}.
\]
As \( \lambda_e \subset L_J \), it preserves \( g_J(k) \) for any \( k \in \mathbb{Z} \). Set
\[
M_J = \bigoplus_{i \neq 0} g_J(i), \quad M'_J = M_J \cap g_i \quad \text{and} \quad g'_J(k) = g_J(k) \cap g_i.
\]
Lemma 2.7. If $M^2_{i}(p^{-1}) = 0$ and $B_f$ has no component of type $A_{p-1}$, then $\lambda_e$ is a Dynkin torus for $e$.

Proof. Since $e \in g_2$, it suffices to show that $\lambda_e$ is an optimal torus for $e$ with respect to the scalar product ($\langle \cdot, \cdot \rangle$). Hence, in view of Lemma 2.3, it suffices to check that the map $(\text{ad} e)^2 : g_{-2} \to g_2$ is surjective. We have

$$g_{-2} = g^2_{j}(0) \oplus M^2_j; \quad g_2 = g^3_{j}(0) \oplus M^2_j.$$ 

Moreover, by (2.8), $g^2_{j}(0) = g^j_{j}(0) \cap g(\pm 2) = l^j_{(1)} \cap g(\pm 2)$. Therefore, the map $\text{ad} e : g^2_{j}(0) \to l^j_{(1)} \cap g(0)$ is bijective (see (2.8)). If $[e, l^j_{(1)} \cap g(0)] \neq g^2_{j}(0)$, then a nonzero subspace $N \subset g^2_{j}(0)$ is orthogonal to $[e, l^j_{(1)} \cap g(0)]$ with respect to the Killing form $k$ of $g$. But then $[e, N] \subset l^j_{(1)} \cap g(0)$ is orthogonal to $l^j_{(1)} \cap g(0)$. By our assumption, $\lambda(l^j_{(1)}) = 0$. Thus $k$ remains nondegenerate if restricted to $l^j_{(1)}$ (see (2.7)). This implies that $k$ is nondegenerate on $l^j_{(1)} \cap g(0)$ forcing $[e, N] = 0$. Summarizing we obtain that $[e, l^j_{(1)} \cap g(0)] = g^2_{j}(0)$ and so

$$(\text{ad} e)^2 : g^2_{j}(0) \to g^3_{j}(0)$$

is one-to-one.

Since $e, h, f \in l_f$, $M_j$ is $(e, h, f)$-stable. Our goal is to show that the map $(\text{ad} e)^2 : M^2_{j} \to M^3_{j}$ is bijective. We first check that $\text{ad} e$ is injective on $M^2_{j}$. Indeed, if $[e, v] = 0$ for some nonzero $v \in M^2_{j}$, then

$$[e, (\text{ad} f)^i(v)] = i(p - 1 - i)(\text{ad} f)^{i-1}(v)$$

for any natural $i$. But then

$$(\text{ad} e)^{p-2}(\text{ad} f)^{p-2}(v) = v \neq 0$$

yielding $(\text{ad} f)^{p-2}(v) \in M_{j}^{2(p-1)} \setminus \{0\}$. Since $k$ induces a nondegenerate pairing between $M^k_{j}$ and $M^{-k}_{j}$ for all $k \in \mathbb{Z}$, this forces $M^2_{j} \neq 0$ violating the assumption. Thus $\text{ad} e$ is injective on $M^2_{j}$.

Suppose that $[e, [e, x]] = 0$ for some nonzero $x \in M^{-2}_{j}$ and let $w = [e, x]$. Since

$$[e, (\text{ad} f)^i(w)] = i(p + 1 - i)(\text{ad} f)^{i-1}(w)$$

for any natural $i$, we have

$$(\text{ad} e)^{p-2}(\text{ad} f)^{p-1}(w) = [f, w].$$

As $(\text{ad} f)^{p-1}(w) \in M^{2(p-1)}_{j} = 0$, this yields $[f, w] = 0$. Using this fact it is easy to note that

$$(\text{ad} e)^{p-2}(\text{ad} f)^{p-2}(x) = x \neq 0.$$ 

But then $(\text{ad} f)^{p-2}(x) \neq 0$ contradicting the equality $M^{2(p-1)}_{j} = 0$. Therefore, the map $(\text{ad} e)^2 : M^{-2}_{j} \to M^3_{j}$ is injective. To complete the proof of the lemma it remains to note that $\dim M^{-2}_{j} = \dim M^3_{j}$. □

2.10. Denote by $m_i(e)$ the maximal weight of the $(\text{Ad} \lambda_e)$-module $g_f(i)$ and set $m(e) = \max_{i \neq 0} m_i(e)$. It follows from the definition of $\lambda_e$ that the numbers $m_i(e)$ do not depend on the characteristic of the ground field. Thus in computing $m_i(e)$'s we may assume that $G$ and $g$ are both defined over $\mathbb{C}$. License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
We first consider the case when all roots in $R$ have the same length. In this case $\alpha|\alpha| = 2$ for any $\alpha \in R$. This implies that
\[(\text{Ad}_\lambda(t)) \cdot e_\gamma = t^{\langle \lambda, \gamma \rangle} \cdot e_\gamma \quad (\gamma \in R)\]
where $\lambda_i = 2 \sum_{i \in R} \omega_i^\perp$. It is well known that, for any nonzero $i$, the subspace $g_i(i)$ is completely irreducible as an $L_j^{(1)}$-module. Moreover, any nontrivial irreducible $L_j^{(1)}$-submodule of $g_j(i)$ is generated by a highest weight vector that is a root element with respect to $T$ and corresponds to a minimal (minuscule) weight of the root system $R_j$ (indeed, as all roots of $R$ have the same length, it suffices to note that $\langle \gamma, \delta \rangle \in \{-1, 0, 1\}$ if $\gamma \in R_j$, $\delta \in R \setminus R_j$).

Since the linear function $\langle \lambda_i, \gamma \rangle$ is nonpositive on $R_j \cap (-R_+)$, the $L_j^{(1)}$-module $g_j(i)$ contains a highest weight vector $e_{\gamma_i}$, $\gamma_i \in R$, such that $m(e) = \langle \lambda_i, \gamma_i \rangle$. Clearly, $X^R = X_J \oplus X_J^\perp$ where $X_J$ is the $\mathbb{R}$-span of $B_j$ and $X_J^\perp$ is its orthogonal complement relative to $\langle \cdot, \cdot \rangle$. Let $\gamma_i = \gamma_i^+ + \gamma_i^-$ where $\gamma_i^+ \in X_J$, $\gamma_i^- \in X_J^\perp$. Then $\langle \lambda_i, \gamma_i \rangle = \langle \lambda_i, \gamma_i^+ \rangle$ and $\gamma_i^+$ is a minimal weight of $R_J$. The vector $\gamma_i^+$ is a sum of minimal weights of irreducible components of $R_J$. These, in turn, lie in the set $\{\omega_i^j | i \in J\}$.

Let $\rho_j = \sum_{i \in J} \omega_i^j$ and $\rho = \sum_{i = 1}^l \omega_i$. As $\langle \omega_i, \omega_j \rangle \geq 0$ for each $i, j \in J$ (see [4, VI, Tables I-IX]), we have
\[
\langle \lambda_i, \gamma_i^+ \rangle \leq \langle 2\rho, \gamma_i^+ \rangle.
\]
Computing $\langle 2\rho, \gamma_i^+ \rangle$ can be reduced to the corresponding problem for the irreducible components of $R_J$. Using [4, VI, Tables I-IX], one can check that
\[
(2\rho, \omega_k) = k(l - k + 1)
\]
if $R$ is of type $A_l$, $1 \leq k \leq l$;
\[
(2\rho, \omega_l) = 2(l - 1), \quad (2\rho, \omega_{l-1}) = (2\rho, \omega_l) = l(l - 1)/2
\]
if $R$ is of type $D_l$, $l \geq 4$.

These equations together with the above remarks yield that $m(e) < 2(p - 1)$ if $e \in l_j$ and the root system $R_j$ has one of the following types:

- for $R \cong E_6$, $E_7$ or $E_8$, $p > 3$:
  \[A_1, A_1 \times A_1, A_1 \times A_1 \times A_1, A_2, A_2 \times A_1, A_2 \times A_1 \times A_1, A_2 \times A_2, A_2 \times A_2 \times A_1, A_3, A_3 \times A_1, D_4\]
- for $R \cong E_7$ or $E_8$, $p > 3$:
  \[A_1 \times A_1 \times A_1 \times A_1, A_2 \times A_1 \times A_1 \times A_1, A_3 \times A_1 \times A_1 \times A_1, A_3 \times A_2, A_3 \times A_2 \times A_1, D_4 \times A_1\]
- for $R \cong E_8$, $p > 5$.

By Lemma 2.7, if $R_j$ has one of the types listed above, then $\lambda_e$ is a Dynkin torus for $e \in l_j$. 

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
2.11. Thus in what follows we may assume that $R_J$ has one of the following types:

$$A_4, A_4 \times A_1, A_5, D_5$$

for $R \cong E_6, p > 3$;

$$A_4, A_4 \times A_1, A_4 \times A_2, A_5, A_5 \times A_1, D_5, D_5 \times A_1, A_6, D_6, E_6$$

for $R \cong E_7, p > 3$;

$$D_5 \times A_2, A_6, A_6 \times A_1, E_6, E_6 \times A_1, A_6, D_6, A_7, D_7, E_7$$

for $R \cong E_8, p > 5$.

Note that if $e$ is a regular nilpotent element of $l_J$ we can always arrange that $e = \sum_{i \in J} e_{i}$. Denote $Q_+ = \{\sum_{i=1}^{l} n_i \alpha_i | n_i \in \mathbb{Z}^+\}$ and set $Q'_+ = Q_+ \cap X'_J$. Let $w_J$ be the element of maximal length in the Weyl group $W(R_J) \subseteq W$. Given $\eta = \sum_{i=1}^{l} m_i \alpha_i$ in $Q_+$ and $k \leq l$ define $\nu_k(\eta) = m_k$ and let $ht(\eta) = \nu_1(\eta) + \nu_2(\eta) + \ldots + \nu_l(\eta)$. We call the number $ht(\eta)$ the height of $\eta$. Set

$$Y(\eta) = \{\alpha_k \in B | \nu_k(\eta) \neq 0\}.$$ 

For $k \in \{1, 2, \ldots, l\} \setminus J$, define

$$\Gamma^k_J = \{\gamma \in R_+ | Y(\gamma) = B \cup \{\alpha_k\}, \nu_k(\gamma) = 1\}.$$ 

Let $\beta^k_J = w_J(\alpha_k)$. Clearly, $\beta^k_J \in \Gamma^k_J$.

**Lemma 2.8.** $\beta^k_J$ is the only element of maximal height in $\Gamma^k_J$.

**Proof.** Let $\delta \in \Gamma^k_J$. As $w_J$ acts on $\Gamma^k_J$, $\delta = w_J(\delta')$ for some $\delta' \in \Gamma^k_J$. We have $\delta' = \alpha_k + \sum_{i \in J} c_i e_{\alpha_i}$, where $c_i \in \mathbb{Z}^+$. As $-w_J$ acts on $B_J$, $\beta^k_J - \delta = w_J(\alpha_k - \delta') = \sum_{i \in J} c_i (-w_J e_{\alpha_i})$ yielding

$$ht(\beta^k_J - \delta) = \sum_{i \in J} c_i \geq 0.$$ 

Clearly, $\sum_{i \in J} c_i = 0$ implies $\alpha_k = \delta'$ forcing $\beta^k_J = \delta$ as desired. \qed

2.12. Let $M_{J,+} = \bigoplus_{i>0} g_J(i)$ and $M_{J,-} = \bigoplus_{i<0} g_J(i)$. Obviously, the $L_J$-modules $M_{J,+}^2$ and $M_{J,-}^2$ are isomorphic. Let $M_{J,\pm} = M_{J,\pm} \cap M_{J}^2$. If the map

$$(\text{ad } e)^2: M_{J,+}^2 \to M_{J,+}^2$$

is bijective, then so is the induced map

$$(\text{ad } e)^2|_{M_{J,+}}: (M_{J,+}^2)^* \to (M_{J,+}^2)^*$$

which can be identified with

$$(\text{ad } e)^2: M_{J,-}^2 \to M_{J,-}^2$$

via the above isomorphism.

Thus in order to show that $(\text{ad } e)^2: M_{J}^{-2} \to M_{J}^{-2}$ is bijective, it suffices to prove that so is $(\text{ad } e)^2: M_{J,+}^{-2} \to M_{J,+}^{-2}$. It is easy to check that $M_{J,+}^{-2}$ is spanned by $e_{\gamma}$ such that $\gamma \in R_+ \setminus R_J$ and $(\lambda_{I,J}|\gamma) = -2$. 
Lemma 2.9. Let \( b_j^j = \text{ht}\, \beta_j^j + 1, \, k \in J, \) and \( \gamma = \sum_{i=1}^l m_i \alpha_i \in R_+ \). Then
\[
(2\rho_J|\gamma) = 2\text{ht}\, \gamma - \sum_{i \notin J} m_i b_j^j.
\]

Proof. We have \( \rho - w_J \rho \in Q_+ \) and \( (\rho - w_J \rho|\alpha_i) = (\rho|\alpha_i - w_J \alpha_i) = 2 \) for each \( i \in J \). Since \( 2\rho_J \in X_J \) and \( (2\rho_J|\alpha_i) = 2 \) for each \( i \in J \), we have \( \rho - w_J \rho = 2\rho_J \). Hence
\[
(2\rho_J|\gamma) = (\rho - w_J \rho|\gamma) = (\rho|\gamma - w_J \gamma)
\]
\[
= \left( \rho \sum_{i \in J} m_i(\alpha_i - w_J \alpha_i) + \sum_{i \notin J} m_i(\alpha_i - \beta_j^j) \right)
\]
\[
= 2 \sum_{i \in J} m_i + \sum_{i \notin J} m_i(2 - b_j^j) = 2\text{ht}\, \gamma - \sum_{i \notin J} m_i b_j^j
\]
as required. \( \square \)

2.13. For any \( k \in \mathbb{Z}_+ \), \( \dim g_J^{-2}(k) = \dim g_J^2(k) \). To observe this one can assume that \( G \) and \( g \) are defined over \( \mathbb{C} \). In this case the statement follows from (2.8) and standard properties of the representations of \( SL_2(\mathbb{C}) \). If \( m_k(e) < 2(p - 1) \), then the argument used in proving Lemma 2.7 shows that \( (\text{ad} \, e)^2: g_J^{-2}(k) \to g_J^2(k) \) is one-to-one.

If \( R_J \) had rank \( l - 1 \), then \( B \setminus B_J = \{ \alpha_3 \} \). Set
\[
\Delta_3(a) = \{ \alpha \in R_+ | \nu_2(\alpha) = a, (2\rho_J|\alpha) = -2 \}.
\]
Clearly, \( g_J^{-2}(k) \) is spanned by \( \{ e_j | \gamma \in \Delta_3(k) \} \).

Let \( R_J \) be of type \( D_5 \times A_2 \subset E_8 \). Then \( B_J = B \setminus \{ \alpha_6 \} \). It is straightforward that the \( L_J^{(1)} \)-modules \( g_J(1), g_J(2), g_J(3) \) and \( g_J(4) \) are irreducible and have highest weights \( \omega_2^J + \omega_4^J, \omega_1^J + \omega_4^J, \omega_4^J \) and \( \omega_2^J \) respectively. By (2.10),
\[
(2\rho_J|\omega_2^J + \omega_4^J) = 12, \quad (2\rho_J|\omega_1^J + \omega_4^J) = 10, \quad (2\rho_J|\omega_4^J) = 10 \quad \text{and} \quad (2\rho_J|\omega_2^J) = 2.
\]
If \( e \in l_J \) is not regular, then \( I \neq \emptyset \) and so \( (\lambda_{e, J}\omega_2^J + \omega_4^J) < 12 \leq 2(p - 1) \).

Therefore, in this case \( e \) satisfies the conditions of Lemma 2.7.

Suppose that \( e \) is regular in \( l_J^{(1)} \). By our previous remark the map \( (\text{ad} \, e)^2: g_J^{-2}(k) \to g_J^2(k) \) is bijective if \( k > 1 \) (if \( p > 7 \), it is bijective for all \( k \geq 0 \)). To show that \( (\text{ad} \, e)^2: g_J^{-2}(1) \to g_J^2(1) \) is bijective, observe that \( \beta_j^j = 12321111, \quad b_j^j = 14 \). Using Lemma 2.10 and [4, VI, Table VII] we obtain
\[
\Delta_1(6) = \left\{ \begin{array}{cccccccc}
0011111, & 0111110, & 0111111, & 1111110, & 1111100, & 0121100, & 111100, & 011100
\end{array} \right\}.
\]

Now it is not difficult to verify that \( ad \, e = \sum_{i \neq \beta} \text{ad} \, e_i \) sends \( g_J^{-2}(1) \) onto the subspace spanned by \( e_\beta \) where
\[
\beta \in \left\{ \begin{array}{cccccccc}
0111111, & 1111111, & 1111111, & 1111110, & 0121110, & 0122100, & 1121100, & 1121100
\end{array} \right\}.
\]
This, in turn, is mapped by \( \text{ad} \, e \) onto the span of \( e_\gamma \) where
\[
\gamma \in \left\{ \begin{array}{cccccccc}
1111111, & 0121111, & 1121110, & 0122110, & 1122100, & 1221100
\end{array} \right\}.
\]

(note that this is true for any \( p \)).

Thus \( (\text{ad} \, e)^2: M_J^{-2} \to M_J^2 \) is one-to-one and we can exclude the subsystem of type \( D_5 \times A_2 \) from our list.
2.14. Let $R_j$ be of type $A_7 \subset E_8$. Then $B_j = B \setminus \{\alpha_2\}$, $\beta_j = \frac{1}{2} 1233213$, $b_j^2 = 17$. Using [4, VI, Table VII] and Lemma 2.9 we get $\Delta_2(k) = \emptyset$ if $k \neq 2$ and $\Delta_2(2) = \{\gamma_1, \gamma_2, \gamma_3\}$ where

$$\gamma_1 = \frac{1233213}{2}, \quad \gamma_2 = \frac{1243211}{2}, \quad \gamma_3 = \frac{1343210}{2}.$$ 

Set

$$\delta_1 = \frac{1234321}{2}, \quad \delta_2 = \frac{12432321}{2}, \quad \delta_3 = \frac{2343211}{2}.$$ 

We may (and will) assume that the root elements $e_\gamma$, $\gamma \in R$, belong to a Chevalley basis of $\mathfrak{g}$:

$$[e_\alpha, e_\beta] = \pm e_{\alpha + \beta} \quad \text{if } \alpha, \beta, \alpha + \beta \in R.$$ 

We have

$$(\text{ad}\ e)^2(e_{\gamma_1}) = [e_{a_4}, [e_{a_6}, e_{\gamma_1}]] + [e_{a_5}, [e_{a_4}, e_{\gamma_1}]] + [e_{a_6}, [e_{a_4}, e_{\gamma_1}]] = \pm e_{\delta_1} \pm 2e_{\delta_2}.$$ 

Similarly,

$$(\text{ad}\ e)^2(e_{\gamma_2}) = [e_{a_3}, [e_{a_7}, e_{\gamma_2}]] + [e_{a_6}, [e_{a_7}, e_{\gamma_2}]] + [e_{a_7}, [e_{a_3}, e_{\gamma_2}]] + [e_{a_1}, [e_{a_3}, e_{\gamma_2}]] = \pm 2e_{\delta_1} \pm e_{\delta_2} \pm e_{\delta_3}$$

and

$$(\text{ad}\ e)^2(e_{\gamma_3}) = [e_{a_8}, [e_{a_1}, e_{\gamma_3}]] + [e_{a_1}, [e_{a_8}, e_{\gamma_3}]] + [e_{a_7}, [e_{a_8}, e_{\gamma_3}]] = \pm e_{\delta_1} \pm 2e_{\delta_3}.$$ 

Since

$$\begin{pmatrix} \pm 1 & \pm 2 & \pm 1 \\ \pm 2 & \pm 1 & 0 \\ 0 & \pm 1 & \pm 2 \end{pmatrix} \neq 0$$

for $p > 3$, we conclude that $(\text{ad}\ e)^2: M_{j_1^-}^{j_1^+} \to M_{j_1^-}^{j_1^+}$ is one-to-one. Since any distinguished parabolic subalgebra of $L_j$ is a Borel subalgebra (see [5, p. 174]), the subsystem of type $A_7$ can be excluded from our list.

2.15. Let $R_j \cong E_6 \times A_1 \subset E_8$. In this case $g_j(k) = 0$ if $k > 3$ and the $L_j^{(1)}$-modules $g_j(1), g_j(2)$ and $g_j(3)$ are irreducible and have highest weights $\omega_1^j + \omega_2^j$, $\omega_6^j$ and $\omega_2^j$ respectively. The Lie algebra $I_j^{(1)}$ has three distinguished nilpotent conjugacy classes under the adjoint action of $L_j$ (Table 1):

<table>
<thead>
<tr>
<th>Bala-Carter diagram</th>
<th>Type</th>
<th>Representative</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\begin{pmatrix} 2 &amp; 2 &amp; 2 &amp; 2 &amp; 2 &amp; 2 \ 1 &amp; \end{pmatrix}$</td>
<td>$E_6 \times A_1$</td>
<td>$e_{a_1} + e_{a_2} + e_{a_3} + e_{a_4} + e_{a_5} + e_{a_6} + e_{a_8}$</td>
</tr>
<tr>
<td>$\begin{pmatrix} 2 &amp; 2 &amp; 0 &amp; 2 &amp; 2 &amp; 2 \ 1 &amp; \end{pmatrix}$</td>
<td>$E_6(a_1) \times A_1$</td>
<td>$e_{a_1} + e_{a_2} + e_{a_3} + e_{a_4} + e_{a_5} + e_{a_6} + e_{a_3} + e_{a_4} + e_{a_8}$</td>
</tr>
<tr>
<td>$\begin{pmatrix} 2 &amp; 0 &amp; 2 &amp; 0 &amp; 2 &amp; 2 \ 0 &amp; \end{pmatrix}$</td>
<td>$E_6(a_3) \times A_1$</td>
<td>$e_{a_1} + e_{a_2} + e_{a_1} + e_{a_3} + e_{a_4} + e_{a_5} + e_{a_6} + e_{a_3} + e_{a_4} + e_{a_5} + e_{a_6}$</td>
</tr>
</tbody>
</table>

Table 1

(see, for example, [19]).
If \( e \) is regular in \( l^{(1)} \), then applying Lemma 2.9 yields \( \Delta_7(1) = \Delta_7(3) = \emptyset \), \( \Delta_7(2) = \{ \gamma_1, \gamma_2 \} \), where \( \gamma_1 = \frac{134^2321}{2} \), \( \gamma_2 = \frac{124^3321}{2} \) (one should take into account that \( b^7_l = 234^3321, b^7_l = 19 \)). Let \( \delta_1 = \frac{134^4321}{2}, \delta_2 = \frac{234^3321}{2} \).

Without loss of generality we may assume that \( [e_{a_3}, e_{y_2}] = [e_{a_6}, e_{y_1}] \), \( e_{\delta_1} = \{ [e_{a_1}, [e_{a_6}, e_{y_1}]], e_{\delta_2} = [e_{a_5}, [e_{a_6}, e_{y_1}]] \} \). Then

\[
(ad e)^2(e_y) = [e_{a_1}, [e_{a_1}, e_{y_1}]] + [e_{a_3}, [e_{a_6}, e_{y_1}]] + [e_{a_5}, [e_{a_6}, e_{y_1}]] = 2e_{\delta_1} + e_{\delta_2},
\]
and

\[
(ad e)^2(e_{y_2}) = [e_{a_1}, [e_{a_1}, e_{y_2}]] + [e_{a_3}, [e_{a_6}, e_{y_2}]] + [e_{a_5}, [e_{a_6}, e_{y_2}]] = e_{\delta_1} + 2e_{\delta_2}.
\]

Since

\[
\begin{vmatrix}
2 & 1 & 1 \\
1 & 2 & 1 \\
1 & 2 & 0
\end{vmatrix}
\neq 0
\]
if \( p > 3 \), the map \( (ad e)^2 : M_{j, k}^2 \to M_{j, k}^2 \) is one-to-one.

Suppose that \( e \) has type \( E_6(a_1) \times A_1 \). Then \( \lambda_{l, j} = 2\rho_j - 2\omega_j^l \). Since \( \omega_j^l = \frac{2464200}{3} \), we have \( (\omega_j^l | \alpha_7) = -2, (\omega_j^l | \alpha_8) = 0 \). Hence

\[
(\lambda_{l, j} | \beta) = (2\rho_j | \beta) - (2\omega_j^l | \beta) = 2 \text{ht} \beta - \nu_7(\beta)b^7_l - 2 \nu_4(\beta) + 4 \nu_7(\beta).
\]

Since the number \( b^7_l \) is odd, this implies that all weights of \( \lambda_e \) on \( g_{j}(1) \) and \( g_{j}(3) \) are odd. Therefore, \( M_{j, k}^2 = g_{j}^{-2}(2) \).

If \( \nu_7(\beta) = 2 \), then \( (\lambda_{l, j} | \beta) = -2 \) forces

\[
\text{ht} \beta = 14 + \nu_4(\beta).
\]

Using [4, VI, Table VII] it is now easy to see that \( g_{j}^{-2}(2) \) is spanned by \( e_{\beta_1}, e_{\beta_2} \) and \( e_{\beta_3} \) where

\[
e_1 = \frac{134^4321}{2}, \quad \eta_2 = \frac{234^3321}{2}, \quad \eta_3 = \frac{1354321}{2}.
\]

Then

\[
(ad e)^2(e_{\beta_1}) = [e_{a_1}, [e_{a_3+a_4}, e_{\beta_1}]] + [e_{a_5}, [e_{a_3+a_4}, e_{\beta_1}]] = \pm e_{\eta_1} \pm e_{\eta_2}.
\]

Similarly,

\[
(ad e)^2(e_{\beta_2}) = [e_{a_6}, [e_{a_1}, e_{\beta_1}]] + [e_{a_1}, [e_{a_6}, e_{\beta_1}]] + [e_{a_5}, [e_{a_6}, e_{\beta_1}]] = \pm e_{\eta_1} \pm 2e_{\eta_2}
\]
and

\[
(ad e)^2(e_{\beta_3}) = [e_{a_1}, [e_{a_3}, e_{\beta_3}]] + [e_{a_3}, [e_{a_3}, e_{\beta_3}] + [e_{a_3+a_4}, [e_{a_3}, e_{\beta_3}]]
\]
\[
= \pm 2e_{\eta_1} \pm e_{\eta_1} \pm e_{\eta_3}.
\]

Since

\[
\begin{vmatrix}
\pm 1 & \pm 1 & \pm 2 \\
\pm 1 & \pm 2 & \pm 1 \\
0 & 0 & \pm 1
\end{vmatrix}
\neq 0
\]
if \( p \neq 3 \), the map \( (ad e)^2 : M_{j, k}^2 \to M_{j, k}^2 \) is one-to-one.
Let $e$ be of type $E_6(a_3) \times A_1$. Then $\lambda_{l,J} = 2\rho_J - 2(\omega_2^J + \omega_3^J + \omega_5^J)$. By [4, VI, Table V],
\[ \omega_2^J + \omega_3^J + \omega_5^J = \frac{48118400}{6}. \]

Hence
\[ (\lambda_{l,J}|\beta) = (2\rho_J|\beta) - 2(\nu_2(\beta) + \nu_3(\beta) + \nu_5(\beta)) + 8\nu_7(\beta) \]
\[ = 2ht\beta - \nu_2(\beta)\delta_3^J - 2(\nu_2(\beta) + \nu_3(\beta) + \nu_5(\beta)) + 8\nu_7(\beta) \]
\[ = 2(\nu_1(\beta) + \nu_4(\beta) + \nu_6(\beta) + \nu_8(\beta)) - 9\nu_7(\beta). \]

This implies $M_{J,+}^{-2} = g_{J}^{-2}(2)$. If $\nu_7(\beta) = 2$, then $\nu_8(\beta) = 1$ and $(\lambda_{l,J}|\beta) = -2$ forces
\[ \nu_1(\beta) + \nu_4(\beta) + \nu_6(\beta) = 7. \]

Using [4, VI, Table VII] one can find out that $g_{J}^{-2}(2)$ is spanned by $e_{\gamma_1}, e_{\gamma_2}, e_{\gamma_3}$ and $e_{\gamma_4}$ where
\[ \gamma_1 = \frac{1233321}{2}, \quad \gamma_2 = \frac{1233321}{2}, \quad \gamma_3 = \frac{1243221}{2}, \quad \gamma_4 = \frac{1343221}{2}. \]

Let
\[ \delta_1 = \frac{2343321}{2}, \quad \delta_2 = \frac{1354321}{2}, \quad \delta_3 = \frac{1354321}{3}, \quad \delta_4 = \frac{2344321}{2}. \]

Without loss of generality we may assume that
\[ e_{\gamma_4} = [e_{a_3}, e_{\gamma_1}], \quad [e_{a_6}, e_{\gamma_1}] = [e_{a_2+a_4}, e_{\gamma_1}], \]
\[ [e_{a_3+a_5}, e_{\gamma_2}] = [e_{a_3+a_6}, e_{\gamma_4}], \quad [e_{a_1}, e_{a_7}] = e_{a_1+a_3}, \]
\[ e_{\delta_1} = [e_{a_1+a_3}, [e_{a_6}, e_{\gamma_3}]], \quad e_{\delta_2} = [e_{a_3+a_5}, [e_{a_2+a_4}, e_{\gamma_1}]], \]
\[ e_{\delta_3} = [e_{a_2+a_4}, [e_{a_3+a_5+a_6}, e_{\gamma_2}]], \quad e_{\delta_4} = [e_{a_1}, [e_{a_3+a_6}, e_{\gamma_1}]]. \]

Direct computation shows that
\[
(ad e)^2(e_{\gamma_1}) = [e_{a_1+a_3}, [e_{a_2+a_4}, e_{\gamma_1}]] + [e_{a_3+a_4+a_5}, [e_{a_2+a_4}, e_{\gamma_1}]]
\]
\[ = [e_{a_1+a_3}, [e_{a_6}, e_{\gamma_3}]] + e_{\delta_2} = e_{\delta_1} + e_{\delta_2} \]

and
\[
(ad e)^2(e_{\gamma_2}) = [e_{a_1}, [e_{a_3+a_4+a_5}, e_{\gamma_2}]] + [e_{a_2+a_4}, [e_{a_3+a_4+a_5}, e_{\gamma_2}]]
\]
\[ = [e_{a_1}, [e_{a_3+a_6}, e_{\gamma_4}]] + e_{\delta_4} = e_{\delta_1} + e_{\delta_4}. \]

Also
\[
(ad e)^2(e_{\gamma_3}) = [e_{a_1+a_3}, [e_{a_6}, e_{\gamma_3}]] + [e_{a_3+a_4+a_5}, [e_{a_6}, e_{\gamma_3}]] + [e_{a_6}, [e_{a_1+a_3}, e_{\gamma_3}]]
\]
\[ + [e_{a_1+a_3}, [e_{a_1+a_1}, e_{\gamma_3}]] + [e_{a_1+a_1}, [e_{a_3+a_6}, e_{\gamma_3}]]
\]
\[ = 2e_{\delta_1} + [e_{a_3+a_4+a_5}, [e_{a_2+a_4}, e_{\gamma_1}]] + 2[e_{a_3+a_6}, [e_{a_1}, [e_{a_3}, e_{\gamma_3}]]
\]
\[ = 2e_{\delta_1} + e_{\delta_2} + 2[e_{a_1}, [e_{a_3+a_6}, e_{\gamma_4}]] = 2e_{\delta_1} + e_{\delta_2} + 2e_{\delta_4} \]

and
\[
(ad e)^2(e_{\gamma_4}) = [e_{a_6}, [e_{a_1}, e_{\gamma_4}]] + [e_{a_5+a_6}, [e_{a_1}, e_{\gamma_4}]] + [e_{a_1}, [e_{a_6}, e_{\gamma_4}]]
\]
\[ + [e_{a_1}, [e_{a_5+a_6}, e_{\gamma_4}]] + [e_{a_2+a_4}, [e_{a_3+a_6}, e_{\gamma_4}]]
\]
\[ = 2[e_{a_6}, [e_{a_1}, [e_{a_3}, e_{\gamma_3}]]] + 2e_{\delta_4} + [e_{a_2+a_4}, [e_{a_3+a_4+a_5}, e_{\gamma_2}]]
\]
\[ = 2e_{\delta_1} + e_{\delta_3} + 2e_{\delta_4}. \]
Since

\[
\begin{pmatrix}
1 & 0 & 2 & 2 \\
1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 \\
0 & 1 & 2 & 2
\end{pmatrix}
\begin{pmatrix}
= -3,
\end{pmatrix}
\]

the map \((\text{ad } e)^2: M_{J,+,} \to M_{J,+,}^2\) is bijective.

2.16. Let \(R_J \cong D_7 \subset E_8\). In this case \(g_J(k) = 0\) for \(k > 2\) and \(g_J(1)\) and \(g_J(2)\) are irreducible \(L_J^{(1)}\)-modules with highest weights \(\omega_J^2\) and \(\omega_J^3\) respectively (note that \(B_J = B\backslash\{\alpha_1\}\)).

We have \(\beta_J^1 = 1354321,\ b_J^1 = 23\) and so

\[
(2\rho_J|\gamma) = 2 \text{ht } \gamma - 23\nu_1(\gamma).
\]

This yields \(\Delta_J(1) = \emptyset,\ \Delta_J(2) = \{\gamma \in R|\nu_1(\gamma) = 2,\ \text{ht } \gamma = 22\} = \{\beta\}\) where \(\beta = 23^{54321}\). If \(e = \sum_{i=2}^8 e_{a_i}\), then

\[
(\text{ad } e)^2(e_{\beta}) = [e_{\alpha_2}, [e_{\alpha_3}, e_{\beta}]] + [e_{\alpha_3}, [e_{\alpha_2}, e_{\beta}]] = 2e_\eta
\]

where \(\eta = 245321^{34}\). Hence \((\text{ad } e)^2: M_{J,+,}^2 \to M_{J,+,}^2\) is bijective.

The Lie algebra \(l_J^{(1)}\) has two nonregular distinguished nilpotent classes under the adjoint action of \(L_J\). Their Bala-Carter diagrams are given in Table 2.

<table>
<thead>
<tr>
<th>Bala-Carter diagram</th>
<th>Type</th>
</tr>
</thead>
</table>
| \[\begin{array}{cccccc}
2 & 0 & 2 & 2 & 2 \\
\end{array}\] | \(D_7(a_1)\) |
| \[\begin{array}{cccccc}
2 & 0 & 2 & 2 & 2 \\
\end{array}\] | \(D_7(a_2)\) |

Table 2

If \(e\) has type \(D_7(a_2)\), then \(\lambda_J, \gamma = 2\rho_J - 2(\omega_4^J + \omega_6^J)\). Clearly,

\[
2\omega_4^J = \frac{05108642}{5}, \quad 2\omega_6^J = \frac{0366642}{3}.
\]

Hence

\[
(\lambda_J, \gamma) = (2\rho_J|\gamma) - 2\nu_4(\gamma) - 2\nu_6(\gamma) + 8\nu_1(\gamma) = 2 \text{ht } \gamma - 2\nu_4(\gamma) - 2\nu_6(\gamma) - 15\nu_1(\gamma).
\]

This implies that \(\Delta_J(1) = \emptyset\) and \((\lambda_J, \gamma|\alpha) = 8 < 2(p - 1)\) where \(\alpha\) is the highest root of \(R_+\). As \((\lambda_J, \gamma|\delta) \leq (\lambda_J, \gamma|\alpha)\) for any \(\delta \in R\) with \(\nu_1(\delta) = 2\), we conclude that \(m_2(e) < 2(p - 1)\) and so \((\text{ad } e)^2: M_{J,+,}^2 \to M_{J,+,}^2\) is one-to-one (see our remark in (2.13)).

Let \(e\) be of type \(D_7(a_1)\). Direct verification based on the fact that \(e\) is a Richardson element of \(p_J \cap l_J\) shows that no generality is lost by assuming

\[
e = e_{\alpha_1} + e_{\alpha_5} + e_{\alpha_6} + e_{\alpha_7} + e_{\alpha_8} + e_{\alpha_2 + \alpha_4} + e_{\alpha_4 + \alpha_5}.
\]
As \( \lambda_l, j = 2 \rho_j - 2 \omega'_j \), one computes

\[(\lambda_l, j | \gamma) = 2ht \gamma - 2\nu_4(\gamma) + 5\nu_1(\gamma) - b_j \nu_1(\gamma) = 2(ht \gamma - \nu_4(\gamma) - 9\nu_1(\gamma)).\]

Hence \( m_2(e) = (\lambda_l, j | \alpha) = 10 < 2(p - 1) \). This implies that \((\text{ad} e)^2 : g_j^{-2}(2) \to g_j^{2}(2)\) is one-to-one. If \( \nu_1(\gamma) = 1 \), then \((\lambda_l, j | \gamma) = -2\) forces \( \text{ht} \gamma = \nu_4(\gamma) + 8 \).

Using [4, VI, Table VII] we obtain that \( g_j^{-2}(1) \) is spanned by \( e_{\gamma_i}, 1 \leq i \leq 6 \), where

\[
\begin{align*}
\gamma_1 &= \frac{1221111}{1}, & \gamma_2 &= \frac{1122111}{1}, & \gamma_3 &= \frac{1122210}{1} \\
\gamma_4 &= \frac{1222110}{1}, & \gamma_5 &= \frac{1232110}{1}, & \gamma_6 &= \frac{1232100}{2}
\end{align*}
\]

Let

\[
\begin{align*}
\delta_1 &= \frac{1122221}{1}, & \delta_2 &= \frac{1222221}{1}, & \delta_3 &= \frac{1232211}{1} \\
\delta_4 &= \frac{1233210}{1}, & \delta_5 &= \frac{1233211}{2}, & \delta_6 &= \frac{1232210}{2}
\end{align*}
\]

and \( \sigma = \frac{1122210}{1} \). A suitable transformation of the form \( e_\alpha \mapsto (-1)^{\sigma(\alpha)}e_\alpha \), \( \alpha \in R \), allows one to assume that

\[
\begin{align*}
e_{\alpha_4 + \alpha_5} &= [e_{\alpha_4}, e_{\alpha_5}], & e_{\gamma_2} &= [e_{\alpha_4}, e_\sigma], & e_{\gamma_3} &= [e_{\alpha_5}, e_\sigma], & e_{\gamma_4} &= [e_{\alpha_3}, e_\sigma], \\
e_{\gamma_5} &= [e_{\alpha_4}, e_{\gamma_2}], & [e_{\alpha_3}, e_{\gamma_3}] &= [e_{\alpha_5}, e_{\gamma_4}], & [e_{\alpha_3}, e_{\gamma_5}] &= [e_{\alpha_3 + \alpha_4}, e_{\gamma_4}], \\
e_{\delta_1} &= [e_{\alpha_7}, [e_{\alpha_6}, e_{\gamma_2}]], & e_{\delta_2} &= [e_{\alpha_6}, [e_{\alpha_5}, e_{\gamma_2}]], & e_{\delta_3} &= [e_{\alpha_5}, [e_{\alpha_3 + \alpha_4}, e_{\gamma_2}]], \\
e_{\delta_4} &= [e_{\alpha_4 + \alpha_5}, [e_{\alpha_3}, e_{\gamma_3}]], & e_{\delta_5} &= [e_{\alpha_2 + \alpha_4}, [e_{\alpha_3}, e_{\gamma_3}]], & e_{\delta_6} &= [e_{\alpha_2 + \alpha_4}, [e_{\alpha_3}, e_{\gamma_4}]].
\end{align*}
\]

Computations show that

\[
\begin{align*}
(\text{ad} e)^2(e_{\gamma_1}) &= [e_{\alpha_6}, [e_{\alpha_3}, e_{\gamma_1}]] + [e_{\alpha_4 + \alpha_5}, [e_{\alpha_3}, e_{\gamma_1}]] + [e_{\alpha_6}, [e_{\alpha_4 + \alpha_5}, e_{\gamma_1}]] \\
&= e_{\delta_1} + e_{\delta_2} + e_{\delta_3},
\end{align*}
\]

\[
\begin{align*}
(\text{ad} e)^2(e_{\gamma_2}) &= [e_{\alpha_6}, [e_{\alpha_3}, e_{\gamma_2}]] + [e_{\alpha_4 + \alpha_5}, [e_{\alpha_3}, e_{\gamma_2}]] + [e_{\alpha_6}, [e_{\alpha_4 + \alpha_5}, e_{\gamma_2}]] \\
&= e_{\delta_6} + 2[e_{\alpha_6}, [e_{\alpha_5}, e_\sigma]] + [e_{\alpha_4}, [e_{\alpha_3}, e_\sigma]] \\
&= e_{\delta_6} + 2[e_{\alpha_6}, [e_{\alpha_5}, e_\sigma]] + [e_{\alpha_4}, [e_{\alpha_3}, e_\sigma]] = e_{\delta_6} + 2e_{\delta_5} + e_{\delta_3},
\end{align*}
\]

\[
\begin{align*}
(\text{ad} e)^2(e_{\gamma_3}) &= [e_{\alpha_8}, [e_{\alpha_3}, e_{\gamma_1}]] + [e_{\alpha_4 + \alpha_5}, [e_{\alpha_3}, e_{\gamma_1}]] + [e_{\alpha_4 + \alpha_5}, [e_{\alpha_3}, e_{\gamma_3}]] + [e_{\alpha_7}, [e_{\gamma_3}]] \\
&= e_{\delta_1} + 2[e_{\alpha_6}, [e_{\alpha_5}, e_{\gamma_1}]] + [e_{\alpha_4 + \alpha_5}, [e_{\alpha_3}, e_{\gamma_1}]] \\
&= e_{\delta_1} + 2[e_{\alpha_6}, [e_{\alpha_5}, e_{\gamma_3}]] + [e_{\alpha_4 + \alpha_5}, [e_{\alpha_3}, e_{\gamma_3}]] = e_{\delta_1} + 2e_{\delta_2} + e_{\delta_3} + e_{\delta_4} + e_{\delta_5},
\end{align*}
\]

\[
\begin{align*}
(\text{ad} e)^2(e_{\gamma_4}) &= [e_{\alpha_8}, [e_{\alpha_6}, e_{\gamma_1}]] + [e_{\alpha_4 + \alpha_5}, [e_{\alpha_6}, e_{\gamma_1}]] + [e_{\alpha_4 + \alpha_5}, [e_{\alpha_6}, e_{\gamma_4}]] \\
&= [e_{\alpha_7}, [e_{\alpha_5}, e_{\gamma_1}]] + 2[e_{\alpha_8}, [e_{\alpha_6}, e_\sigma]] + e_{\delta_4} + e_{\delta_5} \\
&= e_{\delta_1} + 2[e_{\alpha_6}, [e_{\alpha_5}, e_{\gamma_1}]] + e_{\delta_4} + e_{\delta_6} = e_{\delta_1} + 2e_{\delta_2} + e_{\delta_4} + e_{\delta_5} + e_{\delta_6},
\end{align*}
\]

\[
\begin{align*}
(\text{ad} e)^2(e_{\gamma_5}) &= [e_{\alpha_8}, [e_{\alpha_6}, e_{\gamma_4}]] + [e_{\alpha_4 + \alpha_5}, [e_{\alpha_6}, e_{\gamma_4}]] + [e_{\alpha_4 + \alpha_5}, [e_{\alpha_6}, e_{\gamma_4}]] \\
&= [e_{\alpha_7}, [e_{\alpha_5}, e_{\gamma_4}]] + 2[e_{\alpha_8}, [e_{\alpha_6}, e_\sigma]] + 2[e_{\alpha_7}, [e_{\alpha_6}, e_{\gamma_4}]] + e_{\delta_4} + e_{\delta_5} \\
&= e_{\delta_1} + 2[e_{\alpha_6}, [e_{\alpha_5}, e_{\gamma_4}]] + e_{\delta_4} + e_{\delta_6} = e_{\delta_1} + 2e_{\delta_2} + e_{\delta_4} + e_{\delta_5} + e_{\delta_6},
\end{align*}
\]

\[
\begin{align*}
(\text{ad} e)^2(e_{\gamma_6}) &= [e_{\alpha_8}, [e_{\alpha_6}, e_{\gamma_4}]] + [e_{\alpha_4 + \alpha_5}, [e_{\alpha_6}, e_{\gamma_4}]] + [e_{\alpha_4 + \alpha_5}, [e_{\alpha_6}, e_{\gamma_4}]] \\
&= [e_{\alpha_7}, [e_{\alpha_5}, e_{\gamma_4}]] + 2[e_{\alpha_8}, [e_{\alpha_6}, e_\sigma]] + 2[e_{\alpha_7}, [e_{\alpha_6}, e_{\gamma_4}]] + 2[e_{\alpha_4 + \alpha_5}, [e_{\alpha_3}, e_{\gamma_1}]] \\
&= [e_{\alpha_7}, [e_{\alpha_5}, e_{\gamma_4}]] + 2[e_{\alpha_8}, [e_{\alpha_6}, e_\sigma]] + 2[e_{\alpha_7}, [e_{\alpha_6}, e_{\gamma_4}]] + 2[e_{\alpha_4 + \alpha_5}, [e_{\alpha_3}, e_{\gamma_1}]] \\
&= 2e_{\delta_1} + 2[e_{\alpha_2 + \alpha_4}, [e_{\alpha_3}, e_{\gamma_1}]] + 2e_{\delta_2} + 2[e_{\alpha_2 + \alpha_4}, [e_{\alpha_3}, e_{\gamma_3}]] \\
&= 2e_{\delta_1} + 2e_{\delta_2} + 2e_{\delta_4} + 2e_{\delta_5} + 2e_{\delta_6}.
\end{align*}
\]
\[(\text{ad } e)^2(e_{\gamma_5}) = [e_{a_3}, [e_{a_6}, e_{\gamma_5}]] + [e_{a_8}, [e_{a_6}, e_{\gamma_5}]] + [e_{a_6}, [e_{a_8}, e_{\gamma_5}]] = 2[e_{a_6}, [e_{a_3}, [e_{a_6}, e_{\gamma_5}]]] + [e_{a_8}, [e_{a_6}, e_{\gamma_5}]] = 2[e_{a_6}, [e_{a_3}, e_{\gamma_5}]] + [e_{a_8}, [e_{a_6}, e_{\gamma_5}]] = 2e_{\delta_3} + [e_{a_6}, [e_{a_8}, e_{\gamma_5}]] = 2e_{\delta_3} - e_{\delta_4},
\]
\[(\text{ad } e)^2(e_{\gamma_6}) = [e_{a_6}, [e_{a_7}, e_{\gamma_6}]] + [e_{a_8}, [e_{a_7}, e_{\gamma_6}]] = [e_{a_6}, [e_{a_7}, e_{\gamma_6}]] + [e_{a_8}, [e_{a_6}, e_{\gamma_6}]] = [e_{a_6}, [e_{a_7}, e_{\gamma_6}]] + [e_{a_8}, [e_{a_6}, e_{\gamma_6}]] = [e_{a_6}, [e_{a_7}, e_{\gamma_6}]] + [e_{a_8}, [e_{a_6}, e_{\gamma_6}]] = 2e_{\delta_3} + [e_{a_6}, [e_{a_8}, e_{\gamma_6}]] = 2e_{\delta_3} - e_{\delta_4}.
\]

Since
\[
\begin{bmatrix}
0 & 1 & 1 & 0 & 0 & 0 \\
1 & 2 & 2 & 2 & 0 & 0 \\
1 & 0 & 0 & 0 & 2 & 0 \\
0 & 0 & 1 & 1 & -1 & 0 \\
1 & 1 & 0 & 2 & 0 & 1 \\
0 & 0 & 1 & 2 & 0 & 1
\end{bmatrix} = 4,
\]
the map \((\text{ad } e)^2: M_{j,+.}^2 \to M_{j,+.}^2\) is bijective.

2.17. We now suppose that \(R_j \cong E_7 \subset E_8\). Clearly, \(B_j = B_j \setminus \{\alpha_8\}\). Therefore, \(g_j(k) = 0\) if \(k > 2\). Moreover, the \(L_j^{(1)}\)-module \(g_j(2)\) is trivial and \(g_j(1)\) is irreducible over \(L_j^{(1)}\) and has highest weight \(\omega_7^J\).

By [5, p. 176], any standard distinguished parabolic subalgebra \(p_j \cap l_j^{(1)}\) of \(l_j^{(1)}\) has the following property:

\[
7 \notin I \text{ and either } \{2, 5\} \subseteq I \text{ or } \{2, 5\} \cap I = \emptyset.
\]

Using [4, VI, Table VI] it is easy to note that \(\omega_1^J, \omega_3^J, \omega_4^J, \omega_5^J, \omega_6^J, \omega_7^J \in Q_+^J\) and \(\nu_7(2\omega_7^J) = 3\). This implies that, for any \(\gamma \in R\) with \(\nu_8(\gamma) = 1\),

\[
(\lambda_l, J)(\gamma) \equiv (\lambda_l, J)(\alpha_8) \equiv (2\omega_7^J|\alpha_8) \equiv 1 \pmod{2}.
\]

Thus all \((\text{Ad } \lambda_e)\)-weights of \(g_j(1)\) are odd. But then \(M_j^{2(p-1)} = 0\) and Lemma 2.7 applies. Therefore, \(\lambda_e\) is a Dynkin torus for \(e \in p_j \cap l_j^{(1)}\).

2.18. We now deal with \(R \cong E_7\). Let \(R_j\) be of type \(A_5 \times A_1\). Clearly, we may assume that \(B_j = B_j \setminus \{\alpha_3\}\) and \(e = \sum_i \epsilon_i \epsilon_a_i\). It is immediate from [4, VI, Table VI] that \(g_j(k) = 0\) if \(k > 3\), \(\beta_j = 1^{12221}_1, b_j = 11\). Hence \((2\rho_j|\gamma) = 2\text{ht }\gamma - 11\nu_3(\gamma)\) (see Lemma 2.9). This implies that \(\Delta_3(k) = \emptyset\) if \(k \neq 2\) and \(\Delta_3(2) = \{\beta_1, \beta_2\}\) where \(\beta_1 = 1^{12221}_1, \beta_2 = 1^{23210}_1\). Set

\[
\delta_1 = 1^{12221}_1, \quad \delta_2 = 1^{23211}_2, \quad \eta = 1^{22210}_1.
\]

We may and do assume that
\[
e_{\beta_1} = [e_{a_7}, e_{\eta}], \quad e_{\beta_2} = [e_{a_8}, e_{\eta}],
e_{\delta_1} = [e_{a_6}, [e_{a_8}, e_{\delta_1}]], \quad e_{\delta_2} = [e_{a_6}, [e_{a_7}, e_{\delta_2}]].
\]

We have
\[(\text{ad } e)^2(e_{\beta_1}) = [e_{a_2}, [e_{a_4}, e_{\beta_1}]] + [e_{a_6}, [e_{a_4}, e_{\beta_1}]] + [e_{a_8}, [e_{a_4}, e_{\beta_1}]] = 2e_{\delta_1} + [e_{a_2}, [e_{a_4}, e_{\delta_1}]] + [e_{a_6}, [e_{a_7}, e_{\delta_1}]] = 2e_{\delta_1} + [e_{a_2}, [e_{a_7}, e_{\eta}]] = 2e_{\delta_1} + [e_{a_2}, [e_{a_7}, e_{\eta}]] = 2e_{\delta_1} + [e_{a_2}, [e_{a_7}, e_{\eta}]] = 2e_{\delta_1} + [e_{a_2}, [e_{a_7}, e_{\eta}]] = 2e_{\delta_1} + [e_{a_2}, [e_{a_7}, e_{\eta}]] = 2e_{\delta_1} + [e_{a_2}, [e_{a_7}, e_{\eta}]] = 2e_{\delta_1} + [e_{a_2}, [e_{a_7}, e_{\eta}]] = 2e_{\delta_1} + [e_{a_2}, [e_{a_7}, e_{\eta}]] = 2e_{\delta_1} + [e_{a_2}, [e_{a_7}, e_{\eta}]].
\]
and

\[(\text{ad } e)^2(e_{\beta_1}) = [e_{\alpha_7}, [e_{\alpha_2}, e_{\beta_1}]] + [e_{\alpha_2}, [e_{\alpha_7}, e_{\beta_1}]] + [e_{\alpha_6}, [e_{\alpha_7}, e_{\beta_1}]] = 2e_{\delta_2} + [e_{\alpha_6}, [e_{\alpha_7}, e_{\eta}]] = 2e_{\delta_2} + [e_{\alpha_6}, e_{\beta_1}] = e_{\delta_1} + 2e_{\delta_2}.
\]

Since

\[
\begin{vmatrix}
2 & 1 \\
1 & 2
\end{vmatrix} \neq 0
\]

if \( p > 3 \), the map \((\text{ad } e)^2: M_j^{-2} \to M_j^2 \) is bijective.

2.19. Let \( R_j \) be of type \( D_5 \times A_1 \). In this case \( B_j = B\{\alpha_6\} \). First suppose that \( e \) is regular in \( t^{(1)}_j \). By [4, VI, Table VI], \( g_j(k) = 0 \) for \( k > 2 \), \( \beta_6^j = \frac{123211}{2} \) and \( b_j^6 = 13 \). Hence \((2\rho_j|\gamma) = 2 \text{ht } \gamma - 13\nu_6(\gamma) \) yielding \( \Delta_6(1) = \emptyset \), \( \Delta_6(2) = \{\gamma\} \)

where \( \gamma = \frac{123221}{1} \).

As

\[(\text{ad } e)^2(e_7) = [e_{\alpha_7}, (e\gamma)] = [e_{\alpha_7}, [e_{\alpha_2}, e\gamma]] + [e_{\alpha_2}, [e_{\alpha_7}, e\gamma]] = 2[e_{\alpha_5}, [e_{\alpha_2}, e\gamma]] \\
= 0,
\]

we conclude that \((\text{ad } e)^2: M_j^{-2} \to M_j^2 \) is one-to-one.

Now we suppose that \( e \) is not regular. Then \( e \) has the data given in Table 3:

<table>
<thead>
<tr>
<th>Bala-Carter diagram</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 3

A simple checking shows that no generality is lost by assuming

\[e = e_{\alpha_1} + e_{\alpha_3} + e_{\alpha_7} + e_{\alpha_2 + \alpha_4} + e_{\alpha_4 + \alpha_5}.
\]

Since \( \lambda_{I,j} = 2\rho_j - 2\omega_4' \) and \( 2\omega_4' = \frac{246300}{3} \), we have

\[(\lambda_{I,j}|\gamma) = 2 \text{ht } \gamma - 13\nu_6(\gamma) + 3\nu_6(\gamma) - 2\nu_4(\gamma) = 2(\text{ht } \gamma - \nu_4(\gamma) - 5\nu_6(\gamma)).
\]

In particular, \((\lambda_{I,j}|\alpha) = 6 < 2(p - 1)\). Since \((\lambda_{I,j}|\beta) \leq (\lambda_{I,j}|\alpha)\) for any \( \beta \in R \) with \( \nu_6(\beta) = 2 \), we conclude that \( m_2(e) < 2(p - 1) \). By our remark in (2.13), it follows that \((\text{ad } e)^2: g_j^{-2}(2) \to g_j^2(2)\) is bijective.

Using [4, VI, Table VI] one can check that \( g_j^{-2}(1) \) is spanned by \( e_{\beta_i}, 1 \leq i \leq 5 \), where

\[
\begin{align*}
\beta_1 &= \begin{pmatrix} 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 1 \end{pmatrix}, \\
\beta_2 &= \begin{pmatrix} 0 & 1 & 1 & 1 \\
1 & 1 & 0 & 1 \end{pmatrix}, \\
\beta_3 &= \begin{pmatrix} 0 & 0 & 1 & 1 \\
1 & 1 & 0 & 1 \end{pmatrix}, \\
\beta_4 &= \begin{pmatrix} 0 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 \end{pmatrix}, \\
\beta_5 &= \begin{pmatrix} 0 & 1 & 2 & 1 \\
1 & 1 & 1 & 1 \end{pmatrix}.
\end{align*}
\]
Set
\[
\begin{align*}
\delta_1 &= \frac{111111}{1}, \quad \delta_2 = \frac{112111}{1}, \quad \delta_3 = \frac{112210}{1}, \\
\delta_4 &= \frac{122110}{1}, \quad \delta_5 = \frac{011211}{1}, \quad \eta = \frac{011110}{0}.
\end{align*}
\]

We may assume that
\[
\begin{align*}
[e_{a_3}, e_{a_2}] &= e_{a_3+a_2}, \quad e_{\beta_1} = [e_{a_1}, e_\eta], \quad e_{\beta_4} = [e_{a_7}, e_\eta], \\
[e_{a_3}, e_{a_4}] &= e_{a_3+a_4}, \quad [e_{a_1}, e_{\beta_2}] = e_{a_1}, \quad [e_{a_7}, e_{\beta_2}] = [e_{a_3}, e_{\beta_1}], \\
e_{\delta_1} &= [e_{a_1}, [e_{a_7}, e_{\beta_3}]], \quad e_{\delta_2} = [e_{a_7}, [e_{a_3+a_4}, e_{\beta_1}]], \quad e_{\delta_3} = [e_{a_3}, [e_{a_3+a_4}, e_{\beta_1}]], \\
e_{\delta_4} &= [e_{a_3}, [e_{a_3+a_4}, e_{\beta_1}]], \quad e_{\delta_5} = [e_{a_4+a_5}, [e_{a_7}, e_{\beta_2}]].
\end{align*}
\]

Then
\[
\begin{align*}
(\text{ad} e)^2(e_{\beta_1}) &= [e_{a_1}, [e_{a_3+a_4}, e_{\beta_1}]] + [e_{a_3}, [e_{a_3+a_4}, e_{\beta_1}]] + [e_{a_7}, [e_{a_3+a_4}, e_{\beta_1}]] \\
&= 2[e_{a_1}, [e_{a_7}, e_{\beta_1}]],
\end{align*}
\]
\[
\begin{align*}
(\text{ad} e)^2(e_{\beta_4}) &= [e_{a_3+a_4}, [e_{a_1}, e_{\beta_2}]] + [e_{a_1}, [e_{a_7}, e_{\beta_2}]] + [e_{a_3}, [e_{a_3+a_4}, e_{\beta_1}]] \\
&= 2[e_{a_3+a_4}, [e_{a_7}, e_{\beta_1}]],
\end{align*}
\]
\[
\begin{align*}
(\text{ad} e)^2(e_{\beta_2}) &= [e_{a_3+a_4}, [e_{a_1}, e_{\beta_2}]] + [e_{a_1}, [e_{a_7}, e_{\beta_2}]] + [e_{a_3}, [e_{a_3+a_4}, e_{\beta_1}]] \\
&= 2[e_{a_3+a_4}, [e_{a_7}, e_{\beta_1}]],
\end{align*}
\]
\[
\begin{align*}
(\text{ad} e)^2(e_{\beta_3}) &= [e_{a_3}, [e_{a_3+a_4}, e_{\beta_1}]] + [e_{a_3}, [e_{a_3+a_4}, e_{\beta_1}]] + [e_{a_7}, [e_{a_3+a_4}, e_{\beta_1}]] \\
&= 2[e_{a_3}, [e_{a_3+a_4}, e_{\beta_1}]],
\end{align*}
\]
\[
\begin{align*}
(\text{ad} e)^2(e_{\beta_5}) &= [e_{a_3+a_4}, [e_{a_1}, e_{\beta_2}]] + [e_{a_1}, [e_{a_7}, e_{\beta_2}]] + [e_{a_3}, [e_{a_3+a_4}, e_{\beta_1}]] \\
&= 2[e_{a_3+a_4}, [e_{a_7}, e_{\beta_1}]],
\end{align*}
\]
\[
\begin{align*}
(\text{ad} e)^2(e_{\beta_5}) &= [e_{a_1}, [e_{a_3+a_4}, e_{\beta_1}]] + [e_{a_3}, [e_{a_3+a_4}, e_{\beta_1}]] + [e_{a_7}, [e_{a_3+a_4}, e_{\beta_1}]] \\
&= 2[e_{a_1}, [e_{a_3+a_4}, e_{\beta_1}]],
\end{align*}
\]
\[
\begin{align*}
(\text{ad} e)^2(e_{\beta_4}) &= [e_{a_1}, [e_{a_3+a_4}, e_{\beta_1}]] + [e_{a_3}, [e_{a_3+a_4}, e_{\beta_1}]] + [e_{a_7}, [e_{a_3+a_4}, e_{\beta_1}]] \\
&= 2[e_{a_1}, [e_{a_3+a_4}, e_{\beta_1}]],
\end{align*}
\]
\[
\begin{align*}
(\text{ad} e)^2(e_{\beta_5}) &= [e_{a_1}, [e_{a_3+a_4}, e_{\beta_1}]] + [e_{a_3}, [e_{a_3+a_4}, e_{\beta_1}]] + [e_{a_7}, [e_{a_3+a_4}, e_{\beta_1}]] \\
&= 2[e_{a_1}, [e_{a_3+a_4}, e_{\beta_1}]],
\end{align*}
\]

Since
\[
\begin{pmatrix}
0 & 2 & 1 & 0 & 0 \\
2 & 0 & 0 & 2 & 2 \\
1 & -2 & 0 & 0 & 2 \\
1 & 0 & 0 & 0 & 1 \\
0 & 2 & 1 & -1 & -2
\end{pmatrix} = -8,
\]

we can exclude \(D_5 \times A_1\) from our list.
2.20. Let $R_J \cong A_6 \subset E_7$. Then $B_J = B\setminus\{\alpha_2\}$, $\beta_J^2 = 123321$, $b_J^2 = 14$ and $e = \sum_{i \neq 2} e_{\alpha_i}$. Clearly, $g_J(k) = 0$ if $k > 2$. By using [4, VI, Table VI] and Lemma 2.9 one gets

$$\Delta_2(1) = \{\beta_1, \beta_2, \beta_3, \beta_4\}, \quad \Delta_2(2) = \{\gamma\},$$

where

$$\beta_1 = \begin{array}{c} \begin{array}{c} 112110 \\ 1 \end{array} \end{array}, \quad \beta_2 = \begin{array}{c} \begin{array}{c} 111110 \\ 1 \end{array} \end{array}, \quad \beta_3 = \begin{array}{c} \begin{array}{c} 012110 \\ 1 \end{array} \end{array},$$

$$\beta_4 = \begin{array}{c} \begin{array}{c} 011111 \\ 1 \end{array} \end{array}, \quad \gamma = \frac{123221}{2}.$$

It is straightforward that $(ad e)^2(e_{\gamma}) = \pm e_{\eta}$ where $\eta = \frac{124321}{2}$. Denote

$$\delta_1 = \frac{122110}{1}, \quad \delta_2 = \frac{112210}{1}, \quad \delta_3 = \frac{112111}{1}, \quad \delta_4 = \frac{012211}{1}.$$

It can be easily seen that we may suppose that

$$[e_{\alpha_6}, e_{\beta_1}] = [e_{\alpha_4}, e_{\beta_2}] = [e_{\alpha_1}, e_{\beta_3}],$$

$$[e_{\alpha_4}, e_{\beta_1}] = [e_{\alpha_7}, e_{\beta_3}], \quad e_{\delta_1} = [e_{\alpha_3}, [e_{\alpha_6}, e_{\beta_1}]], \quad e_{\delta_2} = [e_{\alpha_5}, [e_{\alpha_6}, e_{\beta_1}]],$$

$$e_{\delta_3} = [e_{\alpha_7}, [e_{\alpha_6}, e_{\beta_1}]], \quad e_{\delta_4} = [e_{\alpha_5}, [e_{\alpha_7}, e_{\beta_1}]].$$

Calculations show that

$$(ad e)^2(e_{\beta_1}) = [e_{\alpha_6}, [e_{\alpha_3}, e_{\beta_1}]] + [e_{\alpha_3}, [e_{\alpha_6}, e_{\beta_1}]]$$

$$+ [e_{\alpha_5}, [e_{\alpha_6}, e_{\beta_1}]] + [e_{\alpha_7}, [e_{\alpha_6}, e_{\beta_1}]]$$

$$= 2e_{\delta_1} + e_{\delta_2} + e_{\delta_3},$$

$$(ad e)^2(e_{\beta_2}) = [e_{\alpha_3}, [e_{\alpha_4}, e_{\beta_1}]] + [e_{\alpha_4}, [e_{\alpha_3}, e_{\beta_1}]]$$

$$+ [e_{\alpha_7}, [e_{\alpha_4}, e_{\beta_1}]] + [e_{\alpha_4}, [e_{\alpha_7}, e_{\beta_1}]]$$

$$= e_{\delta_1} + e_{\delta_2} + 2e_{\delta_3},$$

$$(ad e)^2(e_{\beta_3}) = [e_{\alpha_3}, [e_{\alpha_1}, e_{\beta_1}]] + [e_{\alpha_1}, [e_{\alpha_3}, e_{\beta_1}]]$$

$$+ [e_{\alpha_7}, [e_{\alpha_1}, e_{\beta_1}]] + [e_{\alpha_1}, [e_{\alpha_7}, e_{\beta_1}]]$$

$$+ [e_{\alpha_1}, [e_{\alpha_7}, e_{\beta_1}]] + [e_{\alpha_7}, [e_{\alpha_1}, e_{\beta_1}]]$$

$$= e_{\delta_1} + 2e_{\delta_2} + 2e_{\delta_3} + 2e_{\delta_4},$$

$$(ad e)^2(e_{\beta_4}) = [e_{\alpha_4}, [e_{\alpha_1}, e_{\beta_1}]] + [e_{\alpha_1}, [e_{\alpha_4}, e_{\beta_1}]] + [e_{\alpha_5}, [e_{\alpha_4}, e_{\beta_1}]]$$

$$= 2[e_{\alpha_1}, [e_{\alpha_7}, e_{\beta_3}]] + [e_{\alpha_5}, [e_{\alpha_7}, e_{\beta_1}]]$$

$$= 2[e_{\alpha_7}, [e_{\alpha_6}, e_{\beta_1}]] + e_{\delta_4} = 2e_{\delta_3} + e_{\delta_4}.$$
2.21. Let $R_j$ be of type $D_6 \subset E_7$. Then $B_j = B \setminus \{\alpha_1\}$, $\beta_j = 134321$, $b_j = 17$ and $g_j(k) = 0$ for $k > 2$. By [5, p. 175], there are two nonempty subsets $I \subset J$ for which $p_I \cap I_j$ is distinguished in $i_j$, namely, $\{4\}$ and $\{4, 6\}$. As

$$2\omega_4' = \frac{048642}{4}, \quad 2\omega_6' = \frac{024442}{2},$$

we have $(\lambda_j, r_I(\gamma)) \equiv (2\rho_j | \gamma) \pmod{2}$ for any $\gamma \in R$. But $(2\rho_j | \gamma) = 2ht \gamma - 17$ is odd provided $\nu_I(\gamma) = 1$. It follows that all weights of $\lambda_j$ on $g_j(1)$ are odd. Since $g_j(2)$ is a trivial $L_j$-module, we derive that $M_j^{2(p-1)} = 0$. Thus $D_6 \subset E_7$ can be excluded by Lemma 2.7.

2.22. Now let $R_j$ be of type $E_6 \subset E_7$. In this case $B_j = B \setminus \{\alpha_1\}$ and $M_{j,+} = g_j(1)$ is an irreducible $L_j^{(1)}$-module with highest weight $\omega_j'$. Let $R_j$ be of type $E_6 \times A_1$ in $E_8$ (see 2.15). We suppose that $R_j \subset R_j$, $i_j \subset i_j$ and $L_j \subset L_j$. Denote by $\tilde{g}$ a Lie algebra of type $E_8$ that contains $i_j$. By (2.15), $g_j(2)$ is irreducible over $L_j^{(1)}$ and has highest weight $\omega_j'$. Let $\tilde{e}$ denote a nilpotent element from Table 1. Then $\tilde{e} = e + e_{\alpha_8}$ where $e \in i_j$. Clearly $\lambda_{\tilde{e}}(t) = \lambda_e(t)h_{\alpha_8}(t)$ for each $t \in G_m$. The $L_j^{(1)}$-modules $\tilde{g}_j(2)$ and $g_j(1)$ are dual to each other and $e_{\alpha_8}$ and $h_{\alpha_8}(t)$ both act trivially on $\tilde{g}_j(2)$. Therefore, we can apply a computation presented in (2.15) to conclude that $(ad e)^2: g_j^{(1)}(2) \to g_j(1)$ is bijective if $p \neq 3$ (recall that $g_j(-1) \cong g_j(1)^{1)}$).

By our remark in (2.12), it follows that

$$(ad e)^2: M_j^{2} \to M_j^{2}$$

is one-to-one.

2.23. Let $R_j$ be of type $A_5$ in $R \cong E_6$. Then $e$ is regular in $i_j$, $B_j = B \setminus \{\alpha_2\}$, $\beta_j = 12321$ and $b_j = 11$. Using this information and [4, VI, Table V] it is now easy to observe that $\Delta_2(k) = \emptyset$ for all $k > 0$. Therefore, this case can be excluded by applying Lemma 2.7.

If $R_j$ is of type $D_5$ in $R \cong E_6$, then $M_{j,+} = g_j(1)$. By conjugating $R_j$ by $w_0 \in W$ if necessary we obtain $B_j = B \setminus \{\alpha_1\}$. We have $\beta_j = \alpha_1$, $b_j = 12$, $(2\rho_j | \alpha_1) = 10$. If $e$ is not regular in $i_j$, then $e \in p_I \cap i_j$ where $I = \{4\}$ (see [5, p. 175]). Since $2\omega_4' = \frac{03642}{3}$, then $m_1(e) = (\lambda_l, j | \alpha_1) = 7 < 2(p-1)$ and Lemma 2.7 applies. Thus, we can suppose that $e = \sum_{i>1} e_{\alpha_i}$.

Using [4, VI, Table V] we get $\Delta_1(1) = \{\gamma_1, \gamma_2\}$ where $\gamma_1 = \begin{pmatrix} 11111 \\ 0 \end{pmatrix}$, $\gamma_2 = \begin{pmatrix} 11110 \\ 1 \end{pmatrix}$. Set $\delta_1 = \begin{pmatrix} 11211 \\ 1 \end{pmatrix}$, $\delta_2 = \begin{pmatrix} 12210 \\ 1 \end{pmatrix}$. Then

$$(ad e)^2(e_{\gamma_1}) = [e_{\alpha_4}, [e_{\alpha_2}, e_{\gamma_1}]] = \pm e_{\delta_1},$$

$$(ad e)^2(e_{\gamma_2}) = [e_{\alpha_3}, [e_{\alpha_4}, e_{\gamma_2}]] + [e_{\alpha_2}, [e_{\alpha_4}, e_{\gamma_2}]] = \pm 2e_{\delta_1} \pm e_{\delta_2}.$$
Let $R_j$ be of type $A_5$ in $R \cong E_7$. By conjugating $R_j$ by a suitable $w \in W$ one obtains $B_j = B \{ a_2, a_7 \}$. Using [4, VI, Table VI] we get $\beta_1^2 = \frac{123210}{1}$, $\beta_2^2 = \frac{111111}{1}$, $b_1^2 = 11$, $b_2^2 = 7$. Applying Lemma 2.9 yields $\Delta_2,\gamma(1,0) = \Delta_2,\gamma(2,1) = \Delta_2,\gamma(2,0) = \emptyset$ and $\Delta_2,\gamma(1,1) = \{ \beta_1, \beta_2 \}$ where

$$\beta_1 = \frac{111111}{1}, \quad \beta_2 = \frac{012211}{1}.$$

Put $\delta_1 = \frac{122222}{1}$, $\delta_2 = \frac{112222}{1}$, $e_{\delta_1} = [e_{a_3}, [e_{a_1}, e_{\beta_1}]]$, $e_{\delta_2} = [e_{a_6}, [e_{a_3}, e_{\beta_1}]]$. Without loss of generality we may assume that $e = \sum_{i \neq 2,7} e_{a_i}$ and $[e_{a_3}, e_{\beta_1}] = [e_{a_1}, e_{\beta_2}]$. Then

$$\begin{align*}
(ad e)^2(e_{\beta_1}) &= [e_{a_3}, [e_{a_3}, e_{\beta_1}]] + [e_{a_3}, [e_{a_1}, e_{\beta_1}]] + [e_{a_6}, [e_{a_3}, e_{\beta_1}]] = 2e_{\delta_1} + e_{\delta_2}, \\
(ad e)^2(e_{\beta_2}) &= [e_{a_3}, [e_{a_1}, e_{\beta_2}]] + [e_{a_6}, [e_{a_1}, e_{\beta_2}]] + [e_{a_1}, [e_{a_6}, e_{\beta_2}]] = e_{\delta_1} + 2e_{\delta_2}.
\end{align*}$$

Since

$$\begin{vmatrix}
2 & 1 \\
1 & 2
\end{vmatrix} \neq 0$$

if $p \neq 3$, we conclude that $(ad e)^2 : M^2_{j,+} \to M^2_{j,+}$ is bijective.

Now let $R_j$ be of type $D_5$ in $R \cong E_7$. By conjugating $R_j$ by a suitable $w \in W$ we get $B_j = B \{ a_6, a_7 \}$. Let $\bar{g}$ denote the subalgebra of type $E_6$ generated by $e_{\pm a_i}$, $i < 6$. Then $g_j(1) = \bar{g}_j(1) \oplus K e_{\alpha_7}$. Clearly, $K e_{\alpha_7}$ is a trivial $L_j^{(1)}$-module. Using [4, VI, Table VI] it is easy to check that the $L_j^{(1)}$-modules $\bar{g}_j(1)$ and $g_j(2)$ are isomorphic. Applying the second part of (2.22) we obtain now that

$$(ad e)^2 : g_j^{-2}(k) \to g_j^{2}(k)$$

is bijective for $k = 1, 2$. Clearly, $\beta_1^2 = \frac{123210}{2}$, $\beta_2^2 = \alpha_7$, $b_1^2 = 12$, $b_2^2 = 2$.

The subspace $g_j(3)$ is spanned by all $e_y$ with $\nu_6(y) = 2$ and $\nu_1(y) = 1$. By Lemma 2.9, $(2p_j|\gamma) = 34 - 2 \cdot 12 - 2 = 8 \leq 2(p - 1)$. It follows that $m_3(e) = (\lambda_{I,J}|\gamma) < 2(p - 1)$ provided $I \neq \emptyset$. Hence we may assume that $e = \sum_{i < 6} e_{a_i}$. By Lemma 2.9, $\Delta_6,\gamma(2,1) = \{ \gamma \}$ where $\gamma = \frac{123222}{1}$. Since $(ad e)^2(e_{\gamma}) = 2[e_{a_3}, [e_{a_6}, e_{\gamma}]] \neq 0$, we can exclude $D_5$ from our list.

Let $R_j$ be of type $D_6$ in $R \cong E_8$. In this case $B_j = B \{ a_1, a_8 \}$. For any $k > 0$, the $L_j^{(1)}$-module $g_j(k)$ is completely reducible. Moreover, the highest weights of the irreducible submodules of $g_j(k)$ lie in the set $\{ 0, \omega_2^2, \omega_3^2, \omega_4^2 \}$ (see (2.10)). By (2.10), $(2p_j|\omega_2^2) = (2p_j|\omega_3^2) = 15$, $(2p_j|\omega_4^2) = 10 < 2(p - 1)$. Reasoning as in (2.21) it can now be easily seen that $(\lambda_{I,J}|\omega_2^2) = (\lambda_{I,J}|\omega_3^2)$ is odd for any $I \subset J$ such that $p_I \cap l_{J}^{(1)}$ is distinguished in $l_{J}^{(1)}$. Summarizing we obtain that each $(Ad \lambda_{e})$-weight of $M_{j,+}$ is either odd or less than $2(p - 1)$. But then $M_j^{2(p-1)} = 0$ and, by Lemma 2.7, $\lambda_{e}$ is a Dynkin torus for $e \in p_I \cap l_{J}^{(1)}$.

If $R_j$ has type $E_8$ in $R \cong E_8$, then $B_j = B \{ a_7, a_8 \}$. Clearly, $M_j$ is a completely reducible $L_j^{(1)}$-module. Let $E_j(\omega)$ denote the (unique) irreducible $L_j^{(1)}$-module with highest weight $\omega = \sum_{i \in J} a_i \omega_i^J$ where $a_i \in \mathbb{Z}_+$. Let $E_j(\omega)_s$ be the weight space of $E_j(\omega)$ corresponding to weight $s \in X(\lambda_{\omega}) \cong \mathbb{Z}$. Let $\lambda_{e} \in L_j^{(1)}$ on $E_j(\omega)$. The Lie algebra $l_j^{(1)} = \text{Lie}(L_j^{(1)})$ acts on $E_j(\omega)$ via the differential $d\rho$ of a rational representation $\rho : L_j^{(1)} \to GL(E_j(\omega))$.

If $V$ is a nontrivial irreducible $L_j^{(1)}$-submodule of $M_j$, then either $V \cong E_j(\omega_1^J)$ or $V \cong E_j(\omega_5^J)$ (see (2.10)). Combining (2.22) with the computation
in (2.15) one easily sees that the map \((dp(e))^2: E_j(\omega_j^i)_{-2} \to E_j(\omega_j^i)_2\) is a bijection if \(i = 1, 6, p \neq 3\) (note that \(E_j(\omega_j^i)\) is contragradient to \(E_j(\omega_j^i)\)). From this it is immediate that \((ad e)^2: M_j^{-2} \to M_j^2\) is bijective.

2.25. An argument employed in proving Lemma 2.7 shows that, if \(R_j\) has no components of type \(A_{p-1}\) and \((ad e)^2: M_j^{-2} \to M_j^2\) is a bijection, then so is \((ad e)^2: g_{-2} \to g_2\). Applying Lemma 2.3 shows now that in all examined cases \(\lambda_e\) is a Dynkin torus for \(e\).

It remains to consider the following subsystems \(R_j \subset R\):

\[
\begin{align*}
A_4, & A_4 \times A_1, A_4 \times A_2 \text{ for } R \cong E_6 \text{ or } E_7, \quad p \geq 5; \\
A_6 & \text{ for } R \cong E_7 \text{ or } E_8, A_6 \times A_1 \text{ for } R \cong E_8, \quad p \geq 7.
\end{align*}
\]

In all these cases we can suppose that \(e = \sum_{i \in J} e_i\). If \(p > 5\) (resp., \(p > 7\)) and \(R_j\) is from the first line (resp., from the second line), then \(M_j^{2(p-1)} = 0\) (to obtain this one can argue as in (2.10)). Since in this case \(R_j\) has no components of type \(A_{p-1}\), \(\lambda_e\) is a Dynkin torus for \(e\) by Lemma 2.7.

Thus, in what follows we may assume that \(p = 5\) (resp., \(p = 7\)) for the subsystems from the first line (resp., from the second line). Note that, in any event, \(e^{[p]} = f^{[p]} = 0\).

2.26. Let \(H = \text{diag}(t, t^{-1})\) be the standard Cartan subgroup of the algebraic group \(SL_2\) over \(K\). Let \(F\) denote the ideal of \(H\) in the algebra \(A\) of all regular functions on \(SL_2\). The infinitesimal neighborhood of \(H\) in \(SL_2\) is defined as the group scheme \((SL_2(H))\) corresponding to the algebra \(A/F^p\).

The structure of an \((SL_2(H))\)-module in a finite-dimensional vector space \(V\) is given by a triple \((\theta, X, Y)\) where \(\theta\) is a rational representation of \(G_m\) in \(V\) and \(X, Y\) are endomorphisms of \(V\) such that

\[
\begin{align*}
X^p &= Y^p = 0, \\
\theta(t)X\theta(t)^{-1} &= t^2X, \\
\theta(t)Y\theta(t)^{-1} &= t^{-2}Y, \\
[X, Y] &= d\theta,
\end{align*}
\]

where \(d\theta\) is the differential of \(\theta\).

It is well known (see, for example, [7]) that for any \(n = 0, 1, \ldots, p-1\) and any \(k \in \mathbb{Z}\) there exists a unique irreducible \((SL_2(H))\)-module \(V_{n,k}\) with highest weight \(n + kp\). Moreover, \(V_{n,k} \cong V_{n,0} \otimes \Pi^k\) where \(\Pi\) is the one-dimensional \((SL_2(H))\)-module corresponding to the triple \((t^p, 0, 0)\). Any simple \((SL_2(H))\)-module is isomorphic to one of \(V_{n,k}\). Since the action of \((SL_2(H))\) on \(V_{n,0}\) is induced by the \(n\)th symmetric power of the standard representation of \(SL_2(K)\), the weights of \(V_{n,k}\) are \(n + kp, n-2+ kp, \ldots, -n+2+ kp, -n+k p\).

For any \(k \in \mathbb{Z}\), the module \(V_{p-1,k}\) is projective. For any \(n = 0, 1, \ldots, p-2\) and any \(k \in \mathbb{Z}\) there exists a \(2p\)-dimensional projective indecomposable \((SL_2(H))\)-module \(P_{n,k}\) whose socle and cosocle are both isomorphic to \(V_{n,k}\). The highest (resp., lowest) weight of \(P_{n,k}\) is equal to \((k+1)p + p - n - 2\) (resp., \((k-1)p - (p - n - 2))\). Any projective \((SL_2(H))\)-module is isomorphic to a direct sum of indecomposable projective modules listed above (see [7] for more detail).
Given an \((SL_2)(H)\)-module \(M\) denote by \(X(M)\) the set of weights of \(M\) relative to \(\theta(G_m)\). Let \(M_s\) denote the weight component of \(M\) corresponding to weight \(s \in X(M)\).

**Lemma 2.10.** Suppose that \(M\) is a projective \((SL_2)(H)\)-module such that \(X(M) = -X(M)\) and \(s < 2p - 1\) for each \(s \in X(M)\). Then \(\text{Ker} X \subset \sum_{s \geq 0} M_s\).

**Proof.** Since \(M\) is projective, it is isomorphic to a direct sum of some of \(V_{p-1,k}\)'s and \(P_{n,k}\)'s (with multiplicities). If \(V_{p-1,r}\) or \(P_{n,r}\) with \(r > 0\) (resp., \(r < 0\)) has nonzero multiplicity in \(M\), then there is \(d \in X(M)\) with \(d \geq 2p - 1\) (resp., \(d \leq 1 - 2p\)). As \(X(M)\) is symmetric, this violates the assumption that \(s < 2p - 1\) for each \(s \in X(M)\). Therefore, any indecomposable direct summand of \(M\) is isomorphic either to \(V_{p-1,0}\) or to \(P_{m,0}\) where \(0 \leq m \leq p - 2\).

Clearly, \(V_{p-1,0} \cap \text{Ker} X = (V_{p-1,0})_{p-1}\). Using [7, p. 600] one sees that \(P_{m,0} \cap \text{Ker} X \subset (P_{m,0})_{2p-m-2} \oplus (P_{m,0})_{2p-m-2}\). This implies that

\[
\text{Ker} X \subset \sum_{s \geq 0} M_s
\]
as desired. \(\Box\)

**2.27.** Since the triple \((\lambda_e, \text{ad} e, \text{ad} f)\) restricted to \(M_{J,+}\) satisfies the conditions (2.26 (1)), we may regard \(M_{J,+}\) as an \((SL_2)(H)\)-module. By (2.13), \(X(M_{J,+}) = -X(M_{J,+})\).

If \(R_J \cong A_4 \times A_1\), then \(B_J = B \setminus \{\alpha_3\}\). Hence \(M_{J,+} = g_J(1) \oplus g_J(2) \oplus g_J(3) \oplus g_J(4)\). Looking over [4, VI, Table VII] one obtains that \(g_J(1), g_J(2), g_J(3)\) and \(g_J(4)\) are irreducible over \(L_J^{(1)}\) and have highest weights \(\omega_i^1 + \omega_i^2, \omega_i^3 + \omega_i^4\) and \(\omega_i^5\) respectively.

Let \(N_r\) denote the standard \(SL_r(K)\)-module of dimension \(r\). As \(L_J^{(1)} \cong SL_2(K) \times SL_7(K)\), one has the following module isomorphisms:

\[
g_J(1) \cong N_2 \otimes (\Lambda^2 N_7)^*, \quad g_J(2) \cong \Lambda^2 N_7, \quad g_J(3) \cong N_2 \otimes N_7, \quad g_J(4) \cong N_7^*.
\]

One can view \(N_2\) (resp., \(N_7\)) as a natural \(L_J^{(1)}\)-module via the trivial action of the second (resp., the first) component of \(L_J^{(1)} \cong SL_2(K) \times SL_7(K)\). Let \(\sigma_2\) (resp., \(\sigma_7\)) denote the corresponding representation of \(L_J^{(1)}\). The differential \(d\sigma_2\) (resp., \(d\sigma_7\)) restricted to the principal \(sl_2\)-triple \(Ke \oplus Kh \oplus Kf \subset \text{Lie}(L_J^{(1)})\) together with the rational representation \(\sigma_2 \circ \lambda_e, \sigma_7 \circ \lambda_e\) of \(G_m\) defines a representation \(\theta_2 = (\sigma_2 \circ \lambda_e, d\sigma_2(e)), \quad \theta_7 = (\sigma_7 \circ \lambda_e, d\sigma_7(e), d\sigma_7(f))\) of the group scheme \((SL_2)(H)\) in the vector space \(N_2\) (resp., \(N_7\)). It is immediate from the above remarks that

\[
(\lambda_e, \text{ad} e, \text{ad} f)|_{M_{J,+}} \cong \theta_2 \otimes (\Lambda^2 \theta_7)^* + \Lambda^3 \theta_7 + \theta_2 \otimes \theta_7 + \theta_7.
\]

Since \(N_7\) is an irreducible \((SL_2)(H)\)-module and \(\dim N_7 = p\), we conclude that \(N_7 \cong V_{p-1,k}\) for some \(k \in \mathbb{Z}\) (see (2.26)). But then \(N_7\) and \(N_7^*\) are both projective as \((SL_2)(H)\)-modules. This implies that \(M_{J,+}\) is projective over \((SL_2)(H)\) (bear in mind that \(\Lambda^2 \theta_7\) and \(\Lambda^3 \theta_7\) are direct summands of \(\theta_7 \otimes \theta_7\) and \(\theta_7 \otimes \theta_7 \otimes \theta_7\) respectively).

An easy calculation based on our remarks in (2.10) shows that \(m_1(e) = 11, \quad m_2(e) = 12, \quad m_3(e) = 7\) and \(m_4(e) = 6\). Hence \(s < 2p - 1 = 13\) for any
Applying Lemma 2.10 we get
\[ \delta_g(e) \cap M_{J,+} \subset \sum_{i \geq 0} g_i. \]
Since \( M_{J,-} \) is contragradient to \( M_{J,+} \) in the category of finite-dimensional \( (SL_2)_H \)-modules, Lemma 2.10 applies to \( M_{J,-} \) as well yielding
\[ \delta_g(e) \cap M_{J,-} \subset \sum_{i \geq 0} g_i. \]
By construction, \( I_J \cap \sum_{i \geq 0} g_i = p_I \cap I_J \). Since \( e \) is a Richardson element of \( p_I \cap I_J \), then \( \delta_g(e) \cap I_J \subset \sum_{i \geq 0} g_i \) in view of (2.2) (recall that \( I_J \) admits a nondegenerate trace form). Therefore, \( \delta_g(e) \subset \sum_{i \geq 0} g_i \) and so \( \lambda_e \) is a Dynkin torus for \( e \).

2.28. One can analyze the remaining four cases repeating almost verbatim the argument from (2.27). Details are left to the reader.

If \( R \) is of type \( G_2 \), then \( B_J = \{ \alpha_i \} \) where \( i \in \{1, 2\} \). In this case \( \lambda_e(t) = h_i(t) \) for each \( t \in \mathbb{G}_m \). As \( p > 3 \),
\[ \delta_g(e_{\alpha_i}) \subset \tau \bigoplus_{(\gamma|\alpha_i) \geq 0} Ke_\gamma = \sum_{i \geq 0} g_i \]
whence \( \lambda_e \) is a Dynkin torus for \( e = e_{\alpha_i} \).

If \( R \) is of type \( F_4 \), then \( C_J = \text{Aut}(g) \). We regard \( g \) as a subalgebra of a Lie algebra \( \tilde{g} \) of type \( E_6 \). Let \( \sigma \) denote the outer automorphism of \( \tilde{g} \) defined by extending
\[
\begin{align*}
\sigma(e_{\pm\alpha_1}) &= e_{\pm\alpha_6}, & \sigma(e_{\pm\alpha_2}) &= e_{\pm\alpha_4}, & \sigma(e_{\pm\alpha_3}) &= e_{\pm\alpha_5}, \\
\sigma(e_{\pm\alpha_4}) &= e_{\pm\alpha_3}, & \sigma(e_{\pm\alpha_4}) &= e_{\pm\alpha_2}, & \sigma(e_{\pm\alpha_4}) &= e_{\pm\alpha_4}.
\end{align*}
\]
It is well known that \( g \) is isomorphic to the subalgebra \( \tilde{g}^{\sigma} = \{ x \in \tilde{g} | x^\sigma = x \} \). Moreover, the elements \( e_1 = e_{\alpha_2}, e_2 = e_{\alpha_4}, e_3 = e_{\alpha_3} + e_{\alpha_5} \) and \( e_4 = e_{\alpha_1} + e_{\alpha_6} \) can be viewed as root elements corresponding to the simple roots of \( R \) in Bourbaki's indexing.

If \( e \) is a regular nilpotent element of \( \mathfrak{t}_J^{(1)} \subset \mathfrak{g} \) where \( J \subset \{1, 2, 3, 4\} \), then, up to conjugacy in \( G \), \( e = \sum_{i \in J} e_i \). It is clear from the above that there exists \( \tilde{J} \subset \{1, 2, \ldots, 6\} \) such that \( e = \sum_{i \in \tilde{J}} e_{\alpha_i} \). Therefore, \( e \) is a regular nilpotent element of the standard Levi subalgebra of \( \tilde{g} \) associated to the subset \( \tilde{J} \).

We may assume that \( \tilde{g} = \text{Lie}(\tilde{G}) \) where \( \tilde{G} \) is a simply connected group of type \( E_6 \). It has already been proved that \( \lambda_e \subset X_*(\tilde{G}) \) is a Dynkin torus for \( e \in \tilde{g} \). The automorphism \( \sigma \) is induced by the nontrivial symmetry of the Dynkin diagram of type \( E_6 \). Clearly, the subset \( B_{\tilde{J}} = \{ \alpha_i | i \in \tilde{J} \} \) is \( \sigma \)-stable.

Since the scalar product \( (\cdot | \cdot) \) is \( \sigma \)-stable as well, \( \sigma \) acts on the set \( \{ \omega_{\tilde{J}}^i | i \in \tilde{J} \} \).

It follows that \( \sigma(\rho_{\tilde{J}}) = \rho_{\tilde{J}} \).

As
\[
(\text{Ad} \lambda_e(t)) \cdot e_\gamma = t^{(2\rho_{\tilde{J}}^\gamma)} e_\gamma
\]
for each \( \gamma \in \{ \pm \alpha_i | 1 \leq i \leq 6 \} \) we conclude that
\[
\sigma(\text{Ad} \lambda_e(t))\sigma^{-1} = \text{Ad} \lambda_e(t)
\]
for each \( t \in \mathbb{G}_m \). Therefore, \( \text{Ad} \lambda_e \) acts on \( \tilde{g}^\circ = g \). Let \( \tilde{\lambda}_e : \mathbb{G}_m \to \text{Aut} g \) denote the homomorphism induced by restricting \( \text{Ad} \lambda_e \) to \( \tilde{g}^\circ \). Clearly, \( \tilde{\lambda}_e \in X_*(G) \) and \( \tilde{\lambda}_e(t) \cdot e = t^2 e \) for any \( t \in \mathbb{G}_m \). Let \( \tilde{g}_i \) (resp., \( g_i \)) be the weight component of \( \text{Ad} \lambda_e \) (resp., \( \text{Ad} \tilde{\lambda}_e \)) corresponding to weight \( i \in \mathbb{Z} \). Obviously, \( \tilde{g}_i = \tilde{g}_i \cap g \). But then

\[
\tilde{g}(e) = \tilde{g}(e) \cap g \subseteq \left( \sum_{i \geq 0} \tilde{g}_i \right) \cap g = \sum_{i \geq 0} g_i.
\]

Therefore, \( \tilde{\lambda}_e \) is a Dynkin torus for \( e \in g \).

By [5, pp. 174, 175], any distinguished parabolic subalgebra of a Lie algebra of type \( A_1, A_1 \times A_1, A_2 \times A_1, B_2 \) or \( B_3 \) is a Borel subalgebra. Looking over [4, VI, Table VIII] we conclude now that the Levi subalgebra \( \mathfrak{l}_{J_0} \) corresponding to the subset \( J_0 = \{2, 3, 4\} \) is the only standard Levi subalgebra of \( g \) that contains a nonregular distinguished nilpotent element. Let \( e \) be such an element. By (2.2) and [5, p. 174], we may assume that \( e \in \mathfrak{l}_{J_0} \cap g_{J_0}(2) \) where \( J_0 = \{3\} \).

Using [4, VI, Table VIII] we get \( \omega_2^J = \frac{3}{2} \omega_2 + 2 \omega_3 + \omega_4 \) and \( \omega_4^J = \omega_2 + 2 \omega_3 + 2 \omega_4 \). Therefore, \( \lambda_{J_0} = 2(\omega_2 + \omega_4) = 5 \alpha_2 + 8 \alpha_3 + 6 \alpha_4 = m_2 \alpha_2 + m_3 \alpha_3 + m_4 \alpha_4 \) (as \( \alpha_2 \mid \alpha_2 = 2 \) and \( \alpha_3 \mid \alpha_3 = \alpha_4 \mid \alpha_4 = 1 \)). This implies

\[
\lambda_e(t) = h_2(t^{m_2}) h_3(t^{m_3}) h_4(t^{m_4}) = h_2(t^5) h_3(t^4) h_4(t^3).
\]

As \( \nu_1(\gamma) \leq 2 \) for each \( \gamma \in R_+ \), \( M_{J_0, +} = g_{J_0}(1) \oplus g_{J_0}(2) \). It is immediate from [4, VI, Table VIII] that \( g_{J_0}(2) \) is trivial over \( L_{J_0}^{(1)} \) and the \( L_{J_0}^{(1)} \)-module \( g_{J_0}(1) \) is generated by the highest weight vector \( e_\gamma \) where \( \gamma = 1 3 4 2 \).

Since \( \langle \gamma | \alpha_3 \rangle = \langle \gamma | \alpha_4 \rangle = 0 \) and \( \langle \gamma, \alpha_2 \rangle = 1 \), we obtain

\[
(\text{Ad} \lambda_e(t)) \cdot e_\gamma = t^5 e_\gamma.
\]

Consequently, \( m_1(e) = 5 \), \( m_2(e) = 0 \). But then \( M_{J_0}^{2(p-1)} = 0 \). Applying Lemma 2.7 we obtain that \( \lambda_e \) is a Dynkin torus for \( e \in \mathfrak{l}_{J_0} \).

The proof of Theorem 2.5 is now complete.

References


**Department of Mathematics, University of California at Riverside, Riverside, California 92521**

*Current address*: Department of Mathematics, The University of Manchester, Oxford Road, Manchester M1 39PL, U.K.

*E-mail address*: sashap@ma.man.ac.uk