TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 347, Number 8, August 1995

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

ALEXANDER PREMET

ABSTRACT. Let $\mathfrak{g}$ 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 $\lambda \subset G$ let $\mathfrak{g}(\lambda, i)$ denote the weight component of $\mathrm{Ad}(\lambda)$ corresponding to weight $i \in X(\lambda) \cong \mathbb{Z}$. It is proved in the paper that, for any nonzero nilpotent element $e \in \mathfrak{g}$, there is a one-dimentional torus $\lambda_e \subset G$ such that $e \in \mathfrak{g}(\lambda_e, 2)$ and $\mathrm{Ker}\,\mathrm{ad}\, e \subseteq \bigoplus_{i \geq 0} \mathfrak{g}(\lambda_e, i)$.

1. INTRODUCTION

Let G be a connected reductive group over an algebraically closed field K of characteristic $p > 0$ and $\mathfrak{g} = \mathrm{Lie}(G)$. The group G acts on $\mathfrak{g}$ via the adjoint representation Ad. Given a one-dimensional torus $\lambda \subset G$ denote by $\mathfrak{g}(\lambda, i)$ the weight component of $\mathrm{Ad}\,\lambda$ corresponding to weight $i \in X(\lambda) \cong \mathbb{Z}$. Throughout the paper we assume that $p = \mathrm{char}(K)$ 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 $\mathfrak{g}$ has a canonical $[p]$-operation invariant under the adjoint action of G. An element $x \in \mathfrak{g}$ is said to be *nilpotent* or *$[p]$-nilpotent* (resp., *semisimple* or *$[p]$-semisimple*) if $x^{[p]^e} = 0$ for some $e \in \mathbb{Z}_+$ (resp., if x lies in the p-envelope of $x^{[p]}$ in $\mathfrak{g}$). The group G acts on the set of all nilpotent elements of $\mathfrak{g}$. The orbits of this action are classified by Bala-Carter under the assumption that $p \gg 0$ (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 E_7 and E_8 for $p \leq 5$ though it follows from [20] that the number of nilpotent orbits of $\mathfrak{g}$ is finite for any $p > 0$.

Let $p \gg 0$ and e a nonzero nilpotent element of $\mathfrak{g}$. The Jacobson-Morozov theorem [21, III, 4.3] says that $\mathfrak{g}$ contains a subalgebra $\mathfrak{s}$ such that $\mathfrak{s} \cong sl(2)$, $e \in \mathfrak{s}$ and $\mathfrak{g}$ is a completely reducible $\mathfrak{s}$-module. Moreover, a standard Lie theory argument shows that there is a connected subgroup $S \subseteq G$ such that $\mathfrak{s} = \mathrm{Lie}(S)$ and any $\mathfrak{s}$-submodule of $\mathfrak{g}$ is S-stable. There exist a maximal unipotent subgroup $U_e \subset S$ and a one-dimensional torus $\lambda_e \subset S$ 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.

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

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 $\mathfrak{g}$ and the Kempf-Rousseau theory as exposed in [18].

The motive for this investigation originated in the representation theory of $\mathfrak{g}$. It is well known [6] that all irreducible representations of $\mathfrak{g}$ are of finite dimension. To each irreducible $\mathfrak{g}$-module V, one can assign in a canonical way a linear function $\chi \in \mathfrak{g}^*$ called the *p-character* of V. The ideal I_χ of the universal enveloping algebra $U(\mathfrak{g})$ generated by the central elements of the form $x^p - x^{[p]} - \chi(x)^p \cdot 1$, where $x \in \mathfrak{g}$, acts trivially on V. Given a restricted subalgebra $\mathfrak{a} \subseteq \mathfrak{g}$ denote by $u_\chi(\mathfrak{a})$ the associative subalgebra of $U(\mathfrak{g})/I_\chi$ generated by $\mathfrak{a}$. It follows from the PBW-theorem that $\dim u_x(\mathfrak{a}) = p^{\dim \mathfrak{a}}$.

In [23], Kac and Weisfeiler conjectured that if G is simple and $\mathfrak{g}$ admits a nondegenerate G-invariant trace form, then any irreducible $\mathfrak{g}$-module with p-character χ has dimension divisible by $p^{(\dim \Omega(\chi))/2}$ where $\Omega(\chi)$ is the orbit of χ under the coadjoint action of G. As I recently observed (see [15]), for any $\chi \in \mathfrak{g}^*$, there exists a restricted nilpotent subalgebra $\tilde{\mathfrak{m}}_\chi$ of $\mathfrak{g}$ such that $\dim \tilde{\mathfrak{m}}_\chi = \frac{1}{2}\dim \Omega(\chi)$ and any irreducible, faithful $\mathfrak{g}$-module with p-character χ is free over $u_\chi(\tilde{\mathfrak{m}}_\chi)$. This result proves the Kac-Weisfeiler conjecture. In constructing the subalgebra $\tilde{\mathfrak{m}}_\chi$, 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 $gl_p(K)/\mathfrak{z} = \mathrm{Lie}(G)$ where $G = PGL_p(K)$. It is easily seen that the preimage of $\mathrm{Ker}\,\mathrm{ad}\,e$ in $gl_p(K)$ acts irreducibly on the standard $gl_p(K)$-module of dimension p. It follows that $\mathrm{Ker}\,\mathrm{ad}\,e \not\subseteq \mathrm{Lie}(P)$ 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 $\mathfrak{g}$ into weight spaces under the adjoint action of T giving a Cartan decomposition

$$\mathfrak{g} = \mathfrak{t} \bigoplus \sum_{\alpha\in R} Ke_\alpha$$

where $\mathfrak{t} = \mathrm{Lie}(T)$. Let $B = \{\alpha_1, \alpha_2, \dots, \alpha_l\}$ be a basis of simple roots in R, R_+ the corresponding system of positive roots, $\{\overline{\omega}_1, \overline{\omega}_2, \dots, \overline{\omega}_l\}$ 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

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

and extending to arbitrary root by linearity. Denote

$$\mathfrak{g}_J(i) = \begin{cases} \sum_{\eta_J(\alpha)=i} K e_\alpha & \text{if } i \neq 0, \\ \mathfrak{t} \bigoplus \sum_{\eta_J(\alpha)=0} K e_\alpha & \text{if } i = 0. \end{cases}$$

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

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

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

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

$$\mathrm{Lie}(Z_G(x)) = \mathrm{Lie}(Z_G(x)^\circ) \subseteq \mathfrak{z}_\mathfrak{g}(x)$$

(the symbol H° stands for the connected component of a Zariski closed subgroup $H \subseteq G$). By [21, I, §5], $\mathrm{Lie}(Z_G(x)) = \mathfrak{z}_\mathfrak{g}(x)$ 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 $\mathfrak{p}$, then $Z_G(x)^\circ \subseteq P$ (see [5, Corollary 5.2.4]).

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

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

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

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

The Weyl group $W = N_G(T)/T$ acts on both $X_*(T)$ and $X^*(T)$. By fixing a W-invariant positively defined symmetric bilinear form $X_*(T) \times X_*(T) \to \mathbb{Z}$ one can identify the dual vector spaces $X^\mathbb{R} = X^*(T) \otimes_\mathbb{Z} \mathbb{R}$ and $X_\mathbb{R} = X_*(T) \otimes_\mathbb{Z} \mathbb{R}$.

To simplify notation we denote the scalar product on $X_{\mathbb{R}}$ by the above symbol $\langle\ ,\ \rangle$. Let $\|\ \|$ 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 $\|\ \|$ 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 $\mathrm{Int}(g)\circ\lambda \in X_*(T)$, then (by definition)

$$\|\lambda\| = \|\,\mathrm{Int}(g)\circ\lambda\|.$$

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

$$\mathrm{Lie}(L(\lambda)) = \mathfrak{t} \bigoplus \sum_{\langle\lambda,\alpha\rangle=0} Ke_\alpha, \quad \mathrm{Lie}(U(\lambda)) = \sum_{\langle\lambda,\alpha\rangle>0} Ke_\alpha.$$

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

$$V_i = \{v \in V \mid \rho(\lambda(t))(v) = t^i v,\ \forall t \in \mathbb{G}_m\}.$$

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

$$V_i = \bigoplus_{\langle\lambda,\chi\rangle=i} V_\chi.$$

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

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

$$m(v, \lambda) = \max\{i \in \mathbb{Z} \mid v \in V_{(i)}\}$$

and

$$\mathrm{Supp}_T(v) = \{\chi \in X^*(T) \mid v_\chi \neq 0\}.$$

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

A vector $v \in V$ is called *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 *semistable* with respect to H (or H-semistable). A one-dimensional torus $\lambda \in X_*(G)$ is called an *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 *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

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

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

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

(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 λ, the Levi subgroup $L(\lambda) \subset P(\lambda)$ preserves all graded components V_i, $i \in \mathbb{Z}$, of this $\mathbb{Z}$-grading. Denote

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

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* λ *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 $V = \mathfrak{g}$, $\rho = \mathrm{Ad}$. The adjoint action of $\lambda \in X_*(G)$ turns $\mathfrak{g}$ into a $\mathbb{Z}$-graded Lie algebra:

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

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

$$b(X, Y) = \mathrm{tr}\, d\rho(X) d\rho(Y) \qquad (X, Y \in \mathfrak{g})$$

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

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

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

Lemma 2.3 (Kac [9], Slodowy [18]).
(i) *The polynomial function* $f: \mathfrak{g}(2) \to K$ *is* $L_2(\lambda)$*-invariant.*
(ii) *If the map* $(\mathrm{ad}\, A)^2: \mathfrak{g}(-2) \to \mathfrak{g}(2)$ *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 $(\mathrm{ad}\, A)^2: \mathfrak{g}(-2) \to \mathfrak{g}(2)$ is a surjective map, then $\lambda \in X_*(G)$ is an optimal torus for A and so $Z_G(A) \subseteq P(\lambda)$. Since $\mathrm{Lie}(P(\lambda)) = \bigoplus_{i \geq 0} \mathfrak{g}(i)$, this forces $\mathfrak{z}_{\mathfrak{g}}(A) \subseteq \bigoplus_{i \geq 0} \mathfrak{g}(i)$ (see (2.2) for more detail).

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

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

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

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

The rest of the paper is devoted to proving the following

Theorem 2.5. *Any nonzero nilpotent element e of $\mathfrak{g}$ has at least one Dynkin torus.*

Let $\mathfrak{g}' = \mathrm{Lie}(G^{(1)})$. Clearly, $\mathfrak{g} = \mathfrak{t}+\mathfrak{g}'$. It is well known that the canonical $[p]$-operation of $\mathfrak{g}$ is bijective on $\mathfrak{t}$. Since $\mathfrak{g}'$ is a restricted ideal of $\mathfrak{g}$, Jacobson's identity [8, V, §7] yields that any nilpotent element of $\mathfrak{g}$ lies in $\mathfrak{g}'$. Suppose that a nilpotent element $e \in \mathfrak{g}'$ has a Dynkin torus $\lambda \in X_*(G^{(1)})$. Decompose $\mathfrak{g}$ into weight spaces under the adjoint action of λ: $\mathfrak{g} = \bigoplus_{i\in\mathbb{Z}} \mathfrak{g}(i)$. Let T_1 be a maximal torus of G containing λ and let $\mathfrak{t}_1 = \mathrm{Lie}(T_1)$. As $\mathfrak{t}_1 \subseteq \mathfrak{g}(0)$ and $\mathfrak{g} = \mathfrak{t}_1 + \mathfrak{g}'$, we have $\mathfrak{g} = \mathfrak{g}(0) + \mathfrak{g}'$. This yields that $\mathfrak{g}(i) \subset \mathfrak{g}'$ for each $i \neq 0$. As λ is a Dynkin torus for $e \in \mathfrak{g}'$, $\mathfrak{z}_{\mathfrak{g}'}(e) \subseteq \sum_{i\geq 0} \mathfrak{g}(i)$ and $e \in \mathfrak{g}(2)$. Hence λ preserves $\mathfrak{z}_{\mathfrak{g}}(e)$. But then

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

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

Assuming that $G = G^{(1)}$, denote by $G_1, G_2, \dots, G_s$ the simple (and simply connected) normal subgroups of G. Let $\mathfrak{g}_i = \mathrm{Lie}(G_i)$, $1 \leq i \leq s$. Clearly, $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_s$ and $[\mathfrak{g}_i, \mathfrak{g}_j] = 0$ if $i \neq j$. If $e = e_1 + \cdots + e_s$ where $e_i \in \mathfrak{g}_i$, then $\mathfrak{z}_{\mathfrak{g}}(e) = \mathfrak{z}_{\mathfrak{g}_1}(e_1) \oplus \cdots \oplus \mathfrak{z}_{\mathfrak{g}_s}(e_s)$. Suppose that each nonzero e_i has a Dynkin torus $\lambda_i \in X_*(G_i)$. Let $\mathfrak{p}_i$ be the parabolic subalgebra of $\mathfrak{g}_i$ associated with λ_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 $\mathfrak{p} = (\bigoplus_{i\leq s_0} \mathfrak{p}_i) \oplus (\bigoplus_{i>s_0} \mathfrak{g}_i)$ is the parabolic subalgebra of $\mathfrak{g}$ associated with $\lambda = \prod_i \lambda_i \in X_*(G)$. As $\mathfrak{z}_{\mathfrak{g}_i}(e_i) \subseteq \mathfrak{p}_i$ for all $1 \leq i \leq s_0$, then $\mathfrak{z}_{\mathfrak{g}}(e) \subseteq \mathfrak{p}$. Moreover,

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

Hence λ 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, $\mathrm{Lie}(Z_G(e)) = \mathfrak{z}_{\mathfrak{g}}(e)$ by [21, I, Corollary 5.2] and one can apply [21, IV, §§1.7, 2.23]. If $G = SL_n(K)$, then $G = \widetilde{G}^{(1)}$ where $\widetilde{G} = GL_n(K)$. Any nilpotent element $e \in \mathfrak{g}$ can be regarded as an element of $\tilde{\mathfrak{g}} = \mathrm{Lie}(\widetilde{G}) = gl_n(K)$. By [21, IV, §1], one can find a Dynkin torus $\lambda \in X_*(\widetilde{G})$ for $e \in \tilde{\mathfrak{g}}$ contained in G. Let $\mathfrak{p}$ (resp., $\tilde{\mathfrak{p}}$) denote the parabolic subalgebra of $\mathfrak{g}$ (resp., $\tilde{\mathfrak{g}}$) associated with $\lambda \in X_*(G) \subseteq X_*(\widetilde{G})$. Clearly, $\mathfrak{p} = \tilde{\mathfrak{p}} \cap \mathfrak{g}$. But then $\mathfrak{z}_{\mathfrak{g}}(e) = \mathfrak{z}_{\tilde{\mathfrak{g}}}(e) \cap \mathfrak{g} \subseteq \tilde{\mathfrak{p}} \cap \mathfrak{g} = \mathfrak{p}$. Hence $\lambda \in X_*(G)$ is a Dynkin torus for $e \in \mathfrak{g}$.

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 $\mathfrak{g}$ is said to be *distinguished* if it commutes with no nonzero semisimple element of $\mathfrak{g}$. Generalizing [1, 2], Pommerening proved (see [14, p. 377]) that any distinguished nilpotent element of $\mathfrak{g}$ is a Richardson element of a distinguished parabolic subalgebra of $\mathfrak{g}$. If G

is exceptional and p is good for G, then the Killing form of $\mathfrak{g}$ is nondegenerate. Applying (2.2), one can now easily observe that any distinguished nilpotent element of $\mathfrak{g}$ has at least one Dynkin torus. Hence in proving Theorem 2.5 we may assume that $e \in \mathfrak{g}$ is not distinguished.

Since $\mathfrak{z}_{\mathfrak{g}}(e) = \mathrm{Lie}(Z_G(e)^\circ)$ and e commutes with a nonzero semisimple element, the group $Z_G(e)^\circ$ 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 $\mathfrak{s} = \mathrm{Lie}(S)$. Combining [3, §9.2] and [5, §5.9] one can obtain that

$$\mathrm{Lie}(Z_G(S)) = \mathfrak{g}^S = \mathfrak{z}_{\mathfrak{g}}(\mathfrak{s}) = \mathfrak{t} \bigoplus \sum_{\alpha \in R_1} K e_\alpha.$$

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 Δ_J. Hence in what follows we may suppose that $R_1 = R_J = \{\gamma \in R \mid \gamma = \sum_{\alpha \in B_J} n_\alpha \alpha,\ n_\alpha \in \mathbb{Z}\}$ for some $J \subseteq \{1, 2, \ldots, l\}$. Thus $\mathfrak{z}_{\mathfrak{g}}(\mathfrak{s})$ coincides with the Levi subalgebra $\mathfrak{l}_J = \mathrm{Lie}(L_J)$ of the standard parabolic subalgebra $\mathfrak{p}_J$. It is immediate that the Killing form of $\mathfrak{g}$ is nondegenerate on $\mathfrak{l}_J$. By [21, II, §5], the semisimple group $L_J^{(1)}$ is simply connected. By Jacobson's identity [8, V, §7], $e \in \mathfrak{l}_J^{(1)} = \mathrm{Lie}(L_J^{(1)})$.

If $\mathfrak{z}(\mathfrak{l}_J^{(1)}) = 0$, then $\mathfrak{l}_J = \mathfrak{s} \oplus \mathfrak{l}_J^{(1)}$ and $[\mathfrak{s}, \mathfrak{l}_J] = 0$. The group $Z_{L_J^{(1)}}(e)^\circ$ is unipotent (otherwise S would be properly contained in a bigger torus in $Z_G(e)^\circ$ contradicting the maximality of S). Since $\mathfrak{s} \perp \mathfrak{l}_J^{(1)}$, $\mathfrak{l}_J^{(1)}$ admits a nondegenerate trace form. Hence $\mathrm{Lie}(Z_{L_J^{(1)}}(e)^\circ) = \mathfrak{z}_{\mathfrak{l}_J^{(1)}}(e)$. This implies that e is a distinguished nilpotent element of $\mathfrak{l}_J^{(1)}$.

If $\mathfrak{z}(\mathfrak{l}_J^{(1)}) \neq 0$, 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, $\mathfrak{l}_J^{(1)}$ is a direct sum of commuting ideals $\mathfrak{l}_i$ isomorphic to $sl_{r_i}(K)$ for some $r_i \leq 6$. A standard argument used above shows that $Z_{L_J^{(1)}}(e)^\circ$ is unipotent. Together with [21, IV, §1] this yields that e is a regular nilpotent element of $\mathfrak{l}_J^{(1)}$ (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 $\mathfrak{l}_J^{(1)} = \mathrm{Lie}(L_J^{(1)})$.

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 $\mathfrak{p}_I \cap \mathfrak{l}_J$ of the Levi subalgebra $\mathfrak{l}_J$. As $\mathfrak{l}_J = \mathfrak{g}_J(0)$, we have

$$\mathfrak{p}_I \cap \mathfrak{l}_J = \sum_{i \geq 0} \mathfrak{g}_J(0) \cap \mathfrak{g}_I(i).$$

By (2.2) we can also assume that $e \in \mathfrak{g}_J(0) \cap \mathfrak{g}_I(2)$.

We will use the W-invariant scalar product $(\ |\ )$ on $X^{\mathbb{R}} = \mathbb{R}\overline{\omega}_1 \oplus \cdots \oplus \mathbb{R}\overline{\omega}_l$ defined in [4, VI, Tables I–IX] via embedding $X^{\mathbb{R}}$ into a bigger Euclidean space.

Clearly, the $\mathbb{Q}$-span of B_J in $X^{\mathbb{R}}$ has basis $\{\omega_i^J | i \in J\}$ satisfying

$$(\omega_i^J | \alpha_k) = \delta_{ik}$$

for all $i, k \in J$. This implies that

$$2\sum_{i\notin I}\omega_i^J = \sum_{k\in J}\frac{2}{(\alpha_k|\alpha_k)}m_k\alpha_k$$

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

$$(\mathrm{Ad}\, h_\alpha(t))\cdot e_\beta = t^{\langle\beta,\alpha\rangle}e_\beta$$

for all $\alpha, \beta \in R$ (here $\langle\beta,\alpha\rangle = 2\frac{(\beta|\alpha)}{(\alpha|\alpha)}$). Put $h_i(t) = h_{\alpha_i}(t)$ for each $\alpha_i \in B$ and define $\lambda_e \in X_*(G)$ by setting

$$\lambda_e(t) = \prod_{k\in J} h_k(t^{m_k}) \qquad (t \in \mathbb{G}_m).$$

We intend to show that in most of the remaining cases λ_e is an optimal torus for e relative to the scalar product $(\ |\ )$. By construction,

$$(\mathrm{Ad}\,\lambda_e(t))\cdot e_\alpha = \begin{cases} e_\alpha & \text{if } \alpha \in B_I, \\ t^2 e_\alpha & \text{if } \alpha \in B_J\backslash B_I. \end{cases}$$

Hence $\lambda_e(t)$ acts on $\mathfrak{g}_J(0)\cap\mathfrak{g}_I(i)$ by multiplying each vector by t^i, $i\in\mathbb{Z}$. In particular, $(\mathrm{Ad}\,\lambda_e(t))\cdot e = t^2 e$. Since $\mathrm{Lie}(h_\alpha(t)) = K[e_\alpha, e_{-\alpha}]$ for any $\alpha \in R$, we have

$$\mathrm{Lie}(\lambda_e) \subset \sum_{k\in J} K[e_{\alpha_k}, e_{-\alpha_k}] \subseteq \mathfrak{t}\cap\mathfrak{l}_J^{(1)}.$$

Moreover, the Lie algebra $\mathrm{Lie}(\lambda_e)$ is spanned by $h \in \mathfrak{t}\cap\mathfrak{l}_J^{(1)}$ such that $[h, x] = ix$ for any $x \in \mathfrak{g}_J(0)\cap\mathfrak{g}_I(i)$, $i \in \mathbb{Z}$.

As e is a Richardson element of $\mathfrak{p}_I\cap\mathfrak{l}_J^{(1)}$, a distinguished parabolic subalgebra of $\mathfrak{l}_J^{(1)} = \mathrm{Lie}(L_J^{(1)})$, the map

$$\mathrm{ad}\, e\colon \mathfrak{l}_J^{(1)}\cap\mathfrak{g}_I(-2) \to \mathfrak{l}_J^{(1)}\cap\mathfrak{g}_I(0)$$

is bijective (for $\dim \mathfrak{l}_J^{(1)}\cap\mathfrak{g}_I(-2) = \dim \mathfrak{l}_J^{(1)}\cap\mathfrak{g}_I(0)$ and $\mathfrak{z}_{\mathfrak{l}_J}(e) \subseteq \sum_{i\geq 0}\mathfrak{g}_J(0)\cap\mathfrak{g}_I(i)$ as the Killing form of $\mathfrak{g}$ is nondegenerate on $\mathfrak{l}_J$). This implies that there exists $f \in \mathfrak{l}_J^{(1)}\cap\mathfrak{g}_I(-2)$ 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 $\mathfrak{g} = \bigoplus_{i\in\mathbb{Z}}\mathfrak{g}_i$ where

$$\mathfrak{g}_i = \{x \in \mathfrak{g} | (\mathrm{Ad}\,\lambda_e(t))\cdot x = t^i x \text{ for all } t \in \mathbb{G}_m\}.$$

As $\lambda_e \subset L_J$, it preserves $\mathfrak{g}_J(k)$ for any $k \in \mathbb{Z}$. Set

$$M_J = \bigoplus_{i\neq 0}\mathfrak{g}_J(i), \quad M_J^i = M_J\cap\mathfrak{g}_i \quad \text{and} \quad \mathfrak{g}_J^i(k) = \mathfrak{g}_J(k)\cap\mathfrak{g}_i.$$

Lemma 2.7. *If $M_J^{2(p-1)} = 0$ and B_J has no component of type A_{p-1}, then λ_e is a Dynkin torus for e.*

Proof. Since $e \in \mathfrak{g}_2$, it suffices to show that λ_e is an optimal torus for e with respect to the scalar product $(\,|\,)$. Hence, in view of Lemma 2.3, it suffices to check that the map $(\operatorname{ad} e)^2 \colon \mathfrak{g}_{-2} \to \mathfrak{g}_2$ is surjective. We have

$$\mathfrak{g}_{-2} = \mathfrak{g}_J^{-2}(0) \oplus M_J^{-2}; \quad \mathfrak{g}_2 = \mathfrak{g}_J^2(0) \oplus M_J^2.$$

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

$$(\operatorname{ad} e)^2 \colon \mathfrak{g}_J^{-2}(0) \to \mathfrak{g}_J^2(0)$$

is one-to-one.

Since $e, h, f \in \mathfrak{l}_J$, M_J is $\langle e, h, f \rangle$-stable. Our goal is to show that the map $(\operatorname{ad} e)^2 \colon M_J^{-2} \to M_J^2$ is bijective. We first check that $\operatorname{ad} e$ is injective on M_J^{-2}. Indeed, if $[e, v] = 0$ for some nonzero $v \in M_J^{-2}$, then

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

for any natural i. But then

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

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

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

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

for any natural i, we have

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

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

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

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

2.10. Denote by $m_i(e)$ the maximal weight of the $(\operatorname{Ad} \lambda_e)$-module $\mathfrak{g}_J(i)$ and set $m(e) = \max_{i \neq 0} m_i(e)$. It follows from the definition of λ_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 $\mathfrak{g}$ are both defined over $\mathbb{C}$.

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

$$(\operatorname{Ad}\lambda_e(t)) \cdot e_\gamma = t^{(\lambda_{I,J}|\gamma)} \cdot e_\gamma \qquad (\gamma \in R)$$

where $\lambda_{I,J} = 2\sum_{i \notin I} \omega_i^J$. It is well known that, for any nonzero i, the subspace $\mathfrak{g}_J(i)$ is completely irreducible as an $L_J^{(1)}$-module. Moreover, any nontrivial irreducible $L_J^{(1)}$-submodule of $\mathfrak{g}_J(i)$ 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 $\gamma \in R_J$, $\delta \in R\backslash R_J$).

Since the linear function $(\lambda_{I,J}|\ )$ is nonpositive on $R_J \cap (-R_+)$, the $L_J^{(1)}$-module $\mathfrak{g}_J(i)$ contains a highest weight vector e_{γ_i}, $\gamma_i \in R$, such that $m_i(e) = (\lambda_{I,J}|\gamma_i)$. Clearly, $X^{\mathbb{R}} = X_J \oplus X_J^\perp$ where X_J is the $\mathbb{R}$-span of B_J and $X_J^\perp$ is its orthogonal complement relative to $(\ |\ )$. Let $\gamma_i = \gamma_i^J + \gamma_i^\perp$ where $\gamma_i^J \in X_J$, $\gamma_i^\perp \in X_J^\perp$. Then $(\lambda_{I,J}|\gamma_i) = (\lambda_{I,J}|\gamma_i^J)$ and γ_i^J is a minimal weight of R_J. The vector γ_i^J 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 $\rho = \sum_{i=1}^l \overline{\omega}_i$. As $(\omega_i^J|\omega_j^J) \geq 0$ for each $i, j \in J$ (see [4, VI, Tables I–IX]), we have

$$(\lambda_{I,J}|\gamma_i^J) \leq (2\rho_J|\gamma_i^J).$$

Computing $(2\rho_J|\gamma_i^J)$ can be reduced to the corresponding problem for the irreducible components of R_J. Using [4, VI,Tables I–IX], one can check that

$$(2\rho|\overline{\omega}_k) = k(l-k+1)$$

if R is of type A_l, $1 \leq k \leq l$;

$$(2\rho|\overline{\omega}_1) = 2(l-1), \quad (2\rho|\overline{\omega}_{l-1}) = (2\rho|\overline{\omega}_l) = l(l-1)/2$$

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 $e \in \mathfrak{l}_J$ and the root system R_J has one of the following types:

$$A_1,\ A_1\times A_1,\ A_1\times A_1\times A_1,\ A_2,\ A_2\times A_1,\ A_2\times A_1\times A_1,\ A_2\times A_2,$$
$$A_2\times A_2\times A_1,\ A_3,\ A_3\times A_1,\ D_4$$

for $R \cong E_6, E_7$ or E_8, $p > 3$;

$$A_1\times A_1\times A_1\times A_1,\ A_2\times A_1\times A_1\times A_1,\ A_3\times A_1\times A_1,\ A_3\times A_2,$$
$$A_3\times A_2\times A_1,\ D_4\times A_1$$

for $R \cong E_7$ or E_8, $p > 3$;

$$A_2\times A_2\times A_1\times A_1,\ A_3\times A_3,\ A_4,\ A_4\times A_1,\ A_4\times A_1\times A_1,\ A_4\times A_2,$$
$$A_4\times A_2\times A_1,\ A_4\times A_3,\ D_4\times A_2,\ A_5,\ A_5\times A_1,\ D_5,\ D_5\times A_1$$

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

By Lemma 2.7, if R_J has one of the types listed above, then λ_e is a Dynkin torus for $e \in \mathfrak{l}_J$.

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

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

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

Note that if e is a regular nilpotent element of $\mathfrak{l}_J$ we can always arrange that $e=\sum_{i\in J}e_{\alpha_i}$.

Denote $Q_+=\{\sum_{i=1}^{l}n_i\alpha_i|n_i\in\mathbb{Z}_+\}$ and set $Q_+^J=Q_+\cap X^J$. Let w_J be the element of maximal length in the Weyl group $W(R_J)\subset W$. Given $\eta=\sum_{i=1}^{l}m_i\alpha_i$ in Q_+ and $k\le l$ define $\nu_k(\eta)=m_k$ and let $\mathrm{ht}(\eta)=\nu_1(\eta)+\nu_2(\eta)+\cdots+\nu_l(\eta)$. We call the number $\mathrm{ht}(\eta)$ the *height* of η. Set

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

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

$$\Gamma_J^k=\{\gamma\in R_+|Y(\gamma)=B_J\cup\{\alpha_k\}\,,\ \nu_k(\gamma)=1\}.$$

Let $\beta_J^k=w_J(\alpha_k)$. Clearly, $\beta_J^k\in\Gamma_J^k$.

Lemma 2.8. *β_J^k is the only element of maximal height in Γ_J^k.*

Proof. Let $\delta\in\Gamma_J^k$. As w_J acts on Γ_J^k, $\delta=w_J(\delta')$ for some $\delta'\in\Gamma_J^k$. We have $\delta'=\alpha_k+\sum_{i\in J}c_ie_{\alpha_i}$ where $c_i\in\mathbb{Z}_+$. As $-w_J$ acts on B_J,

$$\beta_J^k-\delta=w_J(\alpha_k-\delta')=\sum_{i\in J}c_i(-w_J\alpha_i)$$

yielding

$$\mathrm{ht}(\beta_J^k-\delta)=\sum_{i\in J}c_i\ge 0.$$

Clearly, $\sum_{i\in J}c_i=0$ implies $\alpha_k=\delta'$ forcing $\beta_J^k=\delta$ as desired. □

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

$$(\mathrm{ad}\,e)^2\colon M_{J,+}^{-2}\to M_{J,+}^2$$

is bijective, then so is the induced map

$$((\mathrm{ad}\,e)^2|_{M_{J,+}})^*\colon (M_{J,+}^2)^*\to(M_{J,+}^{-2})^*$$

which can be identified with

$$(\mathrm{ad}\,e)^2\colon M_{J,-}^{-2}\to M_{J,-}^2$$

via the above isomorphism.

Thus in order to show that $(\mathrm{ad}\,e)^2\colon M_J^{-2}\to M_J^2$ is bijective, it suffices to prove that so is $(\mathrm{ad}\,e)^2\colon M_{J,+}^{-2}\to M_{J,+}^2$. It is easy to check that $M_{J,+}^{-2}$ is spanned by e_γ such that $\gamma\in R_+\backslash R_J$ and $(\lambda_{I,J}|\gamma)=-2$.

Lemma 2.9. *Let* $b_J^k = \operatorname{ht}\beta_J^k + 1$, $k \in J$, *and* $\gamma = \sum_{i=1}^l m_i\alpha_i \in R_+$. *Then*

$$(2\rho_J|\gamma) = 2\operatorname{ht}\gamma - \sum_{i\notin J} m_i b_J^i.$$

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

$$\begin{aligned}(2\rho_J|\gamma) &= (\rho - w_J\rho|\gamma) = (\rho|\gamma - w_J\gamma)\\ &= \left(\rho|\sum_{i\in J} m_i(\alpha_i - w_J\alpha_i) + \sum_{i\notin J} m_i(\alpha_i - \beta_J^i)\right)\\ &= 2\sum_{i\in J} m_i + \sum_{i\notin J} m_i(2 - b_J^i) = 2\operatorname{ht}\gamma - \sum_{i\notin J} m_i b_J^i\end{aligned}$$

as required. □

2.13. For any $k \in \mathbb{Z}_+$, $\dim \mathfrak{g}_J^{-2}(k) = \dim \mathfrak{g}_J^2(k)$. To observe this one can assume that G and $\mathfrak{g}$ 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 $(\operatorname{ad} e)^2\colon \mathfrak{g}_J^{-2}(k) \to \mathfrak{g}_J^2(k)$ is one-to-one.

If R_J had rank $l-1$, then $B\backslash B_J = \{\alpha_s\}$. Set

$$\Delta_s(a) = \{\alpha \in R_+ | \nu_s(\alpha) = a,\ (2\rho_J|\alpha) = -2\}.$$

Clearly, $\mathfrak{g}_J^{-2}(k)$ is spanned by $\{e_\gamma | \gamma \in \Delta_s(k)\}$.

Let R_J be of type $D_5 \times A_2 \subset E_8$. Then $B_J = B\backslash\{\alpha_6\}$. It is straightforward that the $L_J^{(1)}$-modules $\mathfrak{g}_J(1)$, $\mathfrak{g}_J(2)$, $\mathfrak{g}_J(3)$ and $\mathfrak{g}_J(4)$ are irreducible and have highest weights $\omega_2^J + \omega_8^J$, $\omega_1^J + \omega_7^J$, ω_5^J and ω_8^J respectively. By (2.10), $(2\rho_J|\omega_2^J + \omega_8^J) = 12$, $(2\rho_J|\omega_1^J + \omega_7^J) = 10$, $(2\rho_J|\omega_5^J) = 10$ and $(2\rho_J|\omega_8^J) = 2$. If $e \in \mathfrak{l}_J$ is not regular, then $I \neq \emptyset$ and so $(\lambda_{I,J}|\omega_2^J + \omega_8^J) < 12 \le 2(p-1)$. Therefore, in this case e satisfies the conditions of Lemma 2.7.

Suppose that e is regular in $\mathfrak{l}_J^{(1)}$. By our previous remark the map $(\operatorname{ad} e)^2\colon \mathfrak{g}_J^{-2}(k) \to \mathfrak{g}_J^2(k)$ is bijective if $k > 1$ (if $p > 7$, it is bijective for all $k \ge 0$). To show that $(\operatorname{ad} e)^2\colon \mathfrak{g}_J^{-2}(1) \to \mathfrak{g}_J^2(1)$ is bijective, observe that $\beta_J^6 = \substack{1232111\\2}$, $b_J^6 = 14$. Using Lemma 2.10 and [4, VI, Table VII] we obtain

$$\Delta_1(6) = \left\{\substack{0011111\\1},\ \substack{0111110\\1},\ \substack{0111111\\1},\ \substack{1111110\\0},\ \substack{1111100\\1},\ \substack{0121100\\1}\right\}.$$

Now it is not difficult to verify that $\operatorname{ad} e = \sum_{i\neq 6} \operatorname{ad} e_{\alpha_i}$ sends $\mathfrak{g}_J^{-2}(1)$ onto the subspace spanned by e_β where

$$\beta \in \left\{\substack{0111111\\1},\ \substack{1111111\\0},\ \substack{1111110\\1},\ \substack{0121110\\1},\ \substack{0122100\\1},\ \substack{1121100\\1}\right\}.$$

This, in turn, is mapped by $\operatorname{ad} e$ onto the span of e_γ where

$$\gamma \in \left\{\substack{1111111\\1},\ \substack{0121111\\1},\ \substack{1121110\\1},\ \substack{0122110\\1},\ \substack{1122100\\1},\ \substack{1221100\\1}\right\}$$

(note that this is true for any p).

Thus $(\operatorname{ad} e)^2\colon M_{J,+}^{-2} \to M_{J,+}^2$ 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 $B_J = B\backslash\{\alpha_2\}$, $\beta_J^2 = {}^{1233321}_{1}$, $b_J^2 = 17$. Using [4, VI, Table VII] and Lemma 2.9 we get $\Delta_2(k) = \varnothing$ if $k \neq 2$ and $\Delta_2(2) = \{\gamma_1, \gamma_2, \gamma_3\}$ where

$$\gamma_1 = {}^{1233221}_{2}, \quad \gamma_2 = {}^{1243211}_{2}, \quad \gamma_3 = {}^{1343210}_{2}.$$

Set

$$\delta_1 = {}^{1343221}_{2}, \quad \delta_2 = {}^{1243321}_{2}, \quad \delta_3 = {}^{2343211}_{2}.$$

We may (and will) assume that the root elements e_γ, $\gamma \in R$, 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

$$(\operatorname{ad} e)^2(e_{\gamma_1}) = [e_{\alpha_4}, [e_{\alpha_6}, e_{\gamma_1}]] + [e_{\alpha_3}, [e_{\alpha_4}, e_{\gamma_1}]] + [e_{\alpha_6}, [e_{\alpha_4}, e_{\gamma_1}]] = \pm e_{\delta_1} \pm 2e_{\delta_2}.$$

Similarly,

$$\begin{aligned}(\operatorname{ad} e)^2(e_{\gamma_2}) &= [e_{\alpha_3}, [e_{\alpha_7}, e_{\gamma_2}]] + [e_{\alpha_6}, [e_{\alpha_7}, e_{\gamma_2}]] + [e_{\alpha_7}, [e_{\alpha_3}, e_{\gamma_2}]] + [e_{\alpha_1}, [e_{\alpha_3}, e_{\gamma_2}]] \\ &= \pm 2e_{\delta_1} \pm e_{\delta_2} \pm e_{\delta_3}\end{aligned}$$

and

$$(\operatorname{ad} e)^2(e_{\gamma_3}) = [e_{\alpha_8}, [e_{\alpha_1}, e_{\gamma_3}]] + [e_{\alpha_1}, [e_{\alpha_8}, e_{\gamma_3}]] + [e_{\alpha_7}, [e_{\alpha_8}, e_{\gamma_3}]] = \pm e_{\delta_1} \pm 2e_{\delta_3}.$$

Since

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

for $p > 3$, we conclude that $(\operatorname{ad} e)^2 \colon M_{J,+}^{-2} \to M_{J,+}^2$ is one-to-one. Since any distinguished parabolic subalgebra of $\mathfrak{l}_J$ 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 $\mathfrak{g}_J(k) = 0$ if $k > 3$ and the $L_J^{(1)}$-modules $\mathfrak{g}_J(1)$, $\mathfrak{g}_J(2)$ and $\mathfrak{g}_J(3)$ are irreducible and have highest weights $\omega_1^J + \omega_8^J$, ω_6^J and ω_8^J respectively. The Lie algebra $\mathfrak{l}_J^{(1)}$ has three distinguished nilpotent conjugacy classes under the adjoint action of L_J (Table 1):

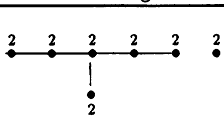
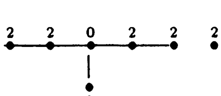
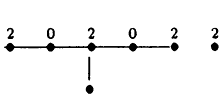
Bala-Carter diagram	Type	Representative
	$E_6 \times A_1$	$e_{\alpha_1} + e_{\alpha_2} + e_{\alpha_3} + e_{\alpha_4} + e_{\alpha_5} + e_{\alpha_6} + e_{\alpha_8}$
	$E_6(a_1) \times A_1$	$e_{\alpha_1} + e_{\alpha_3} + e_{\alpha_5} + e_{\alpha_6} + e_{\alpha_2+\alpha_4} + e_{\alpha_3+\alpha_4} + e_{\alpha_8}$
	$E_6(a_3) \times A_1$	$e_{\alpha_1} + e_{\alpha_6} + e_{\alpha_1+\alpha_3} + e_{\alpha_2+\alpha_4} + e_{\alpha_5+\alpha_6} + e_{\alpha_3+\alpha_4+\alpha_5} + e_{\alpha_8}$

TABLE 1

(see, for example, [19]).

If e is regular in $\mathfrak{l}_J^{(1)}$, then applying Lemma 2.9 yields $\Delta_7(1) = \Delta_7(3) = \varnothing$, $\Delta_7(2) = \{\gamma_1, \gamma_2\}$, where $\gamma_1 = {}^{1343221}_{\quad 2}$, $\gamma_2 = {}^{1243321}_{\quad 2}$ (one should take into account that $\beta_J^7 = {}^{2343211}_{\quad 2}$, $b_J^7 = 19$). Let $\delta_1 = {}^{1344321}_{\quad 2}$, $\delta_2 = {}^{2343321}_{\quad 2}$. Without loss of generality we may assume that $[e_{\alpha_3}, e_{\gamma_2}] = [e_{\alpha_6}, e_{\gamma_1}]$, $e_{\delta_1} = [e_{\alpha_1}, [e_{\alpha_6}, e_{\gamma_1}]]$, $e_{\delta_2} = [e_{\alpha_5}, [e_{\alpha_6}, e_{\gamma_1}]]$. Then

$$(\operatorname{ad} e)^2(e_{\gamma_1}) = [e_{\alpha_6}, [e_{\alpha_1}, e_{\gamma_1}]] + [e_{\alpha_1}, [e_{\alpha_6}, e_{\gamma_1}]] + [e_{\alpha_5}, [e_{\alpha_6}, e_{\gamma_1}]] = 2e_{\delta_1} + e_{\delta_2},$$

and

$$(\operatorname{ad} e)^2(e_{\gamma_2}) = [e_{\alpha_1}, [e_{\alpha_3}, e_{\gamma_2}]] + [e_{\alpha_5}, [e_{\alpha_3}, e_{\gamma_2}]] + [e_{\alpha_3}, [e_{\alpha_5}, e_{\gamma_2}]] = e_{\delta_1} + 2e_{\delta_2}.$$

Since

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

if $p > 3$, the map $(\operatorname{ad} e)^2\colon M_{J,+}^{-2} \to M_{J,+}^2$ is one-to-one.

Suppose that e has type $E_6(a_1) \times A_1$. Then $\lambda_{I,J} = 2\rho_J - 2\omega_4^J$. Since $\omega_4^J = {}^{2464200}_{\quad 3}$, we have $(\omega_4^J|\alpha_7) = -2$, $(\omega_4^J|\alpha_8) = 0$. Hence

$$(\lambda_{I,J}|\beta) = (2\rho_J|\beta) - (2\omega_4^J|\beta) = 2\operatorname{ht}\beta - \nu_7(\beta)b_J^7 - 2\nu_4(\beta) + 4\nu_7(\beta).$$

Since the number b_J^7 is odd, this implies that all weights of λ_e on $\mathfrak{g}_J(1)$ and $\mathfrak{g}_J(3)$ are odd. Therefore, $M_{J,+}^{-2} = \mathfrak{g}_J^{-2}(2)$.

If $\nu_7(\beta) = 2$, then $(\lambda_{I,J}|\beta) = -2$ forces

$$\operatorname{ht}\beta = 14 + \nu_4(\beta).$$

Using [4, VI, Table VII] it is now easy to see that $\mathfrak{g}_J^{-2}(2)$ is spanned by e_{β_1}, e_{β_2} and e_{β_3} where

$$\beta_1 = {}^{1233321}_{\quad 2}, \quad \beta_2 = {}^{1343221}_{\quad 2}, \quad \beta_3 = {}^{1243321}_{\quad 2}.$$

Set

$$\eta_1 = {}^{1344321}_{\quad 2}, \quad \eta_2 = {}^{2343321}_{\quad 2}, \quad \eta_3 = {}^{1354321}_{\quad 2}.$$

Then

$$(\operatorname{ad} e)^2(e_{\beta_1}) = [e_{\alpha_1}, [e_{\alpha_3+\alpha_4}, e_{\beta_1}]] + [e_{\alpha_5}, [e_{\alpha_3+\alpha_4}, e_{\beta_1}]] = \pm e_{\eta_1} \pm e_{\eta_2}.$$

Similarly,

$$(\operatorname{ad} e)^2(e_{\beta_2}) = [e_{\alpha_6}, [e_{\alpha_1}, e_{\beta_2}]] + [e_{\alpha_1}, [e_{\alpha_6}, e_{\beta_2}]] + [e_{\alpha_5}, [e_{\alpha_6}, e_{\beta_2}]] = \pm e_{\eta_1} \pm 2e_{\eta_2}$$

and

$$\begin{aligned}(\operatorname{ad} e)^2(e_{\beta_3}) &= [e_{\alpha_1}, [e_{\alpha_3}, e_{\beta_3}]] + [e_{\alpha_5}, [e_{\alpha_3}, e_{\beta_3}]] \\ &\quad + [e_{\alpha_3}, [e_{\alpha_5}, e_{\beta_3}]] + [e_{\alpha_3+\alpha_4}[e_{\alpha_5}, e_{\beta_3}]] \\ &= \pm 2e_{\eta_1} \pm e_{\eta_2} \pm e_{\eta_3}.\end{aligned}$$

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 $(\operatorname{ad} e)^2\colon M_{J,+}^{-2} \to M_{J,+}^2$ is one-to-one.

Let e be of type $E_6(a_3)\times A_1$. Then $\lambda_{I,J}=2\rho_J-2(\omega_2^J+\omega_3^J+\omega_5^J)$. By [4, VI, Table V],

$$\omega_2^J+\omega_3^J+\omega_5^J=\begin{smallmatrix}4&8&11&8&4&0&0\\&&6&&&&\end{smallmatrix}.$$

Hence

$$\begin{aligned}(\lambda_{I,J}|\beta)&=(2\rho_J|\beta)-2(\nu_2(\beta)+\nu_3(\beta)+\nu_5(\beta))+8\nu_7(\beta)\\&=2\,\mathrm{ht}\,\beta-\nu_7(\beta)b_J^7-2(\nu_2(\beta)+\nu_3(\beta)+\nu_5(\beta))+8\nu_7(\beta)\\&=2(\nu_1(\beta)+\nu_4(\beta)+\nu_6(\beta)+\nu_8(\beta))-9\nu_7(\beta).\end{aligned}$$

This implies $M_{J,+}^{-2}=\mathfrak{g}_J^{-2}(2)$. If $\nu_7(\beta)=2$, then $\nu_8(\beta)=1$ and $(\lambda_{I,J}|\beta)=-2$ forces

$$\nu_1(\beta)+\nu_4(\beta)+\nu_6(\beta)=7.$$

Using [4, VI, Table VII] one can find out that $\mathfrak{g}_J^{-2}(2)$ is spanned by e_{γ_1}, e_{γ_2}, e_{γ_3} and e_{γ_4} where

$$\gamma_1=\begin{smallmatrix}1233321\\1\end{smallmatrix},\quad \gamma_2=\begin{smallmatrix}1233321\\2\end{smallmatrix},\quad \gamma_3=\begin{smallmatrix}1243221\\2\end{smallmatrix},\quad \gamma_4=\begin{smallmatrix}1343221\\2\end{smallmatrix}.$$

Let

$$\delta_1=\begin{smallmatrix}2343321\\2\end{smallmatrix},\quad \delta_2=\begin{smallmatrix}1354321\\2\end{smallmatrix},\quad \delta_3=\begin{smallmatrix}1354321\\3\end{smallmatrix},\quad \delta_4=\begin{smallmatrix}2344321\\2\end{smallmatrix}.$$

Without loss of generality we may assume that

$$\begin{gathered}e_{\gamma_4}=[e_{\alpha_3},e_{\gamma_3}],\quad [e_{\alpha_6},e_{\gamma_3}]=[e_{\alpha_2+\alpha_4},e_{\gamma_1}],\\ [e_{\alpha_3+\alpha_4+\alpha_5},e_{\gamma_2}]=[e_{\alpha_5+\alpha_6},e_{\gamma_4}],\quad [e_{\alpha_1},e_{\alpha_3}]=e_{\alpha_1+\alpha_3},\\ e_{\delta_1}=[e_{\alpha_1+\alpha_3},[e_{\alpha_6},e_{\gamma_3}]],\quad e_{\delta_2}=[e_{\alpha_3+\alpha_4+\alpha_5},[e_{\alpha_2+\alpha_4},e_{\gamma_1}]],\\ e_{\delta_3}=[e_{\alpha_2+\alpha_4},[e_{\alpha_3+\alpha_4+\alpha_5},e_{\gamma_2}]],\quad e_{\delta_4}=[e_{\alpha_1},[e_{\alpha_5+\alpha_6},e_{\gamma_4}]].\end{gathered}$$

Direct computation shows that

$$\begin{aligned}(\operatorname{ad}e)^2(e_{\gamma_1})&=[e_{\alpha_1+\alpha_3},[e_{\alpha_2+\alpha_4},e_{\gamma_1}]]+[e_{\alpha_3+\alpha_4+\alpha_5},[e_{\alpha_2+\alpha_4},e_{\gamma_1}]]\\&=[e_{\alpha_1+\alpha_3},[e_{\alpha_6},e_{\gamma_3}]]+e_{\delta_2}=e_{\delta_1}+e_{\delta_2}\end{aligned}$$

and

$$\begin{aligned}(\operatorname{ad}e)^2(e_{\gamma_2})&=[e_{\alpha_1},[e_{\alpha_3+\alpha_4+\alpha_5},e_{\gamma_2}]]+[e_{\alpha_2+\alpha_4},[e_{\alpha_3+\alpha_4+\alpha_5},e_{\gamma_2}]]\\&=[e_{\alpha_1},[e_{\alpha_5+\alpha_6},e_{\gamma_4}]]+e_{\delta_3}=e_{\delta_3}+e_{\delta_4}.\end{aligned}$$

Also

$$\begin{aligned}(\operatorname{ad}e)^2(e_{\gamma_3})&=[e_{\alpha_1+\alpha_3},[e_{\alpha_6},e_{\gamma_3}]]+[e_{\alpha_3+\alpha_4+\alpha_5},[e_{\alpha_6},e_{\gamma_3}]]+[e_{\alpha_6},[e_{\alpha_1+\alpha_3},e_{\gamma_3}]]\\&\quad+[e_{\alpha_5+\alpha_6},[e_{\alpha_1+\alpha_3},e_{\gamma_3}]]+[e_{\alpha_1+\alpha_3},[e_{\alpha_5+\alpha_6},e_{\gamma_3}]]\\&=2e_{\delta_1}+[e_{\alpha_3+\alpha_4+\alpha_5},[e_{\alpha_2+\alpha_4},e_{\gamma_1}]]+2[e_{\alpha_5+\alpha_6},[e_{\alpha_1},[e_{\alpha_3},e_{\gamma_3}]]]\\&=2e_{\delta_1}+e_{\delta_2}+2[e_{\alpha_1},[e_{\alpha_5+\alpha_6},e_{\gamma_4}]]=2e_{\delta_1}+e_{\delta_2}+2e_{\delta_4}\end{aligned}$$

and

$$\begin{aligned}(\operatorname{ad}e)^2(e_{\gamma_4})&=[e_{\alpha_6},[e_{\alpha_1},e_{\gamma_4}]]+[e_{\alpha_5+\alpha_6},[e_{\alpha_1},e_{\gamma_4}]]+[e_{\alpha_1},[e_{\alpha_6},e_{\gamma_4}]]\\&\quad+[e_{\alpha_1},[e_{\alpha_5+\alpha_6},e_{\gamma_4}]]+[e_{\alpha_2+\alpha_4},[e_{\alpha_5+\alpha_6},e_{\gamma_4}]]\\&=2[e_{\alpha_6},[e_{\alpha_1},[e_{\alpha_3},e_{\gamma_3}]]]+2e_{\delta_4}+[e_{\alpha_2+\alpha_4},[e_{\alpha_3+\alpha_4+\alpha_5},e_{\gamma_2}]]\\&=2e_{\delta_1}+e_{\delta_3}+2e_{\delta_4}.\end{aligned}$$

Since

$$\begin{vmatrix} 1 & 0 & 2 & 2 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 2 & 2 \end{vmatrix} = -3,$$

the map $(\operatorname{ad} e)^2 \colon M_{J,+}^{-2} \to M_{J,+}^{2}$ is bijective.

2.16. Let $R_J \cong D_7 \subset E_8$. In this case $\mathfrak{g}_J(k) = 0$ for $k > 2$ and $\mathfrak{g}_J(1)$ and $\mathfrak{g}_J(2)$ are irreducible $L_J^{(1)}$-modules with highest weights ω_2^J and ω_8^J respectively (note that $B_J = B \backslash \{\alpha_1\}$).

We have $\beta_J^1 = {}^{1354321}_{\;\;3}$, $b_J^1 = 23$ and so

$$(2\rho_J | \gamma) = 2 \operatorname{ht} \gamma - 23\nu_1(\gamma).$$

This yields $\Delta_1(1) = \varnothing$, $\Delta_1(2) = \{\gamma \in R | \nu_1(\gamma) = 2, \operatorname{ht} \gamma = 22\} = \{\beta\}$ where $\beta = {}^{2354321}_{\;\;2}$. If $e = \sum_{i=2}^{8} e_{\alpha_i}$, then

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

where $\eta = {}^{2454321}_{\;\;3}$. Hence $(\operatorname{ad} e)^2 \colon M_{J,+}^{-2} \to M_{J,+}^{2}$ is bijective.

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

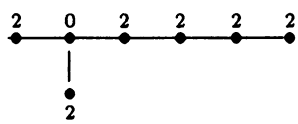
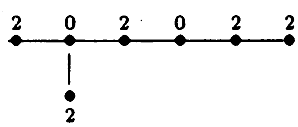
Bala-Carter diagram	Type
	$D_7(a_1)$
	$D_7(a_2)$

TABLE 2

If e has type $D_7(a_2)$, then $\lambda_{I,J} = 2\rho_J - 2(\omega_4^J + \omega_6^J)$. Clearly,

$$2\omega_4^J = {}^{05108642}_{\;\;\;5}, \quad 2\omega_6^J = {}^{0366642}_{\;\;3}.$$

Hence

$$(\lambda_{I,J} | \gamma) = (2\rho_J | \gamma) - 2\nu_4(\gamma) - 2\nu_6(\gamma) + 8\nu_1(\gamma) = 2 \operatorname{ht} \gamma - 2\nu_4(\gamma) - 2\nu_6(\gamma) - 15\nu_1(\gamma).$$

This implies that $\Delta_1(1) = \varnothing$ and $(\lambda_{I,J} | \tilde{\alpha}) = 8 < 2(p-1)$ where $\tilde{\alpha}$ is the highest root of R_+. As $(\lambda_{I,J} | \delta) \le (\lambda_{I,J} | \tilde{\alpha})$ for any $\delta \in R$ with $\nu_1(\delta) = 2$, we conclude that $m_2(e) < 2(p-1)$ and so $(\operatorname{ad} e)^2 \colon M_{J,+}^{-2} \to M_{J,+}^{2}$ 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 $\mathfrak{p}_I \cap \mathfrak{l}_J$ shows that no generality is lost by assuming

$$e = e_{\alpha_3} + e_{\alpha_5} + e_{\alpha_6} + e_{\alpha_7} + e_{\alpha_8} + e_{\alpha_2+\alpha_4} + e_{\alpha_4+\alpha_5}.$$

As $\lambda_{I,J} = 2\rho_J - 2\omega_4^J$, one computes

$$(\lambda_{I,J}|\gamma) = 2\,\mathrm{ht}\,\gamma - 2\nu_4(\gamma) + 5\nu_1(\gamma) - b_J^1\nu_1(\gamma) = 2(\mathrm{ht}\,\gamma - \nu_4(\gamma) - 9\nu_1(\gamma)).$$

Hence $m_2(e) = (\lambda_{I,J}|\tilde{\alpha}) = 10 < 2(p-1)$. This implies that $(\mathrm{ad}\,e)^2 \colon \mathfrak{g}_J^{-2}(2) \to \mathfrak{g}_J^2(2)$ is one-to-one. If $\nu_1(\gamma) = 1$, then $(\lambda_{I,J}|\gamma) = -2$ forces $\mathrm{ht}\,\gamma = \nu_4(\gamma) + 8$. Using [4, VI, Table VII] we obtain that $\mathfrak{g}_J^{-2}(1)$ is spanned by e_{γ_i}, $1 \le i \le 6$, where

$$\gamma_1 = \begin{smallmatrix}1221111\\1\end{smallmatrix}, \quad \gamma_2 = \begin{smallmatrix}1122111\\1\end{smallmatrix}, \quad \gamma_3 = \begin{smallmatrix}1122210\\1\end{smallmatrix}$$
$$\gamma_4 = \begin{smallmatrix}1222110\\1\end{smallmatrix}, \quad \gamma_5 = \begin{smallmatrix}1232110\\1\end{smallmatrix}, \quad \gamma_6 = \begin{smallmatrix}1232100\\2\end{smallmatrix}.$$

Let

$$\delta_1 = \begin{smallmatrix}1122221\\1\end{smallmatrix}, \quad \delta_2 = \begin{smallmatrix}1222211\\1\end{smallmatrix}, \quad \delta_3 = \begin{smallmatrix}1232211\\1\end{smallmatrix},$$
$$\delta_4 = \begin{smallmatrix}1233210\\1\end{smallmatrix}, \quad \delta_5 = \begin{smallmatrix}1232111\\2\end{smallmatrix}, \quad \delta_6 = \begin{smallmatrix}1232210\\2\end{smallmatrix}$$

and $\sigma = \begin{smallmatrix}1122110\\1\end{smallmatrix}$. A suitable transformation of the form $e_\alpha \mapsto (-1)^{\varepsilon(\alpha)}e_\alpha$, $\alpha \in R$, allows one to assume that

$$e_{\alpha_4+\alpha_5} = [e_{\alpha_4}, e_{\alpha_5}], \quad e_{\gamma_2} = [e_{\alpha_8}, e_\sigma], \quad e_{\gamma_3} = [e_{\alpha_6}, e_\sigma], \quad e_{\gamma_4} = [e_{\alpha_3}, e_\sigma],$$
$$e_{\gamma_5} = [e_{\alpha_4}, e_{\gamma_4}], \quad [e_{\alpha_5}, e_{\gamma_1}] = [e_{\alpha_8}, e_{\gamma_4}], \quad [e_{\alpha_7}, e_{\gamma_6}] = [e_{\alpha_2+\alpha_4}, e_{\gamma_4}],$$
$$e_{\delta_1} = [e_{\alpha_7}, [e_{\alpha_6}, e_{\gamma_2}]], \quad e_{\delta_2} = [e_{\alpha_6}, [e_{\alpha_5}, e_{\gamma_1}]], \quad e_{\delta_3} = [e_{\alpha_6}, [e_{\alpha_4+\alpha_5}, e_{\gamma_1}]],$$
$$e_{\delta_4} = [e_{\alpha_4+\alpha_5}, [e_{\alpha_3}, e_{\gamma_3}]], \quad e_{\delta_5} = [e_{\alpha_2+\alpha_4}, [e_{\alpha_5}, e_{\gamma_1}]], \quad e_{\delta_6} = [e_{\alpha_2+\alpha_4}, [e_{\alpha_3}, e_{\gamma_3}]].$$

Computations show that

$$\begin{aligned}
(\mathrm{ad}\,e)^2(e_{\gamma_1}) &= [e_{\alpha_6}, [e_{\alpha_5}, e_{\gamma_1}]] + [e_{\alpha_2+\alpha_4}, [e_{\alpha_5}, e_{\gamma_1}]] + [e_{\alpha_6}, [e_{\alpha_4+\alpha_5}, e_{\gamma_1}]] \\
&= e_{\delta_2} + e_{\delta_3} + e_{\delta_5}, \\
(\mathrm{ad}\,e)^2(e_{\gamma_2}) &= [e_{\alpha_6}, [e_{\alpha_3}, e_{\gamma_2}]] + [e_{\alpha_2+\alpha_4}, [e_{\alpha_3}, e_{\gamma_2}]] \\
&\quad + [e_{\alpha_3}, [e_{\alpha_6}, e_{\gamma_2}]] + [e_{\alpha_7}, [e_{\alpha_6}, e_{\gamma_2}]] \\
&= e_{\delta_1} + 2[e_{\alpha_3}, [e_{\alpha_6}, [e_{\alpha_8}, e_\sigma]]] + [e_{\alpha_2+\alpha_4}, [e_{\alpha_3}, [e_{\alpha_8}, e_\sigma]]] \\
&= e_{\delta_1} + 2[e_{\alpha_6}, [e_{\alpha_8}, e_{\gamma_4}]] + [e_{\alpha_2+\alpha_4}, [e_{\alpha_8}, e_{\gamma_4}]] \\
&= e_{\delta_1} + 2[e_{\alpha_6}, [e_{\alpha_5}, e_{\gamma_1}]] + [e_{\alpha_2+\alpha_4}, [e_{\alpha_5}, e_{\gamma_1}]] = e_{\delta_1} + 2e_{\delta_2} + e_{\delta_5}, \\
(\mathrm{ad}\,e)^2(e_{\gamma_3}) &= [e_{\alpha_8}, [e_{\alpha_3}, e_{\gamma_3}]] + [e_{\alpha_2+\alpha_4}, [e_{\alpha_3}, e_{\gamma_3}]] + [e_{\alpha_4+\alpha_5}, [e_{\alpha_3}, e_{\gamma_3}]] \\
&\quad + [e_{\alpha_3}, [e_{\alpha_8}, e_{\gamma_3}]] + [e_{\alpha_7}, [e_{\alpha_8}, e_{\gamma_3}]] \\
&= [e_{\alpha_7}, [e_{\alpha_8}, [e_{\alpha_6}, e_\sigma]]] + 2[e_{\alpha_3}, [e_{\alpha_8}, [e_{\alpha_6}, e_\sigma]]] + e_{\delta_4} + e_{\delta_6} \\
&= 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_6}, \\
(\mathrm{ad}\,e)^2(e_{\gamma_4}) &= [e_{\alpha_8}, [e_{\alpha_6}, e_{\gamma_4}]] + [e_{\alpha_2+\alpha_4}, [e_{\alpha_6}, e_{\gamma_4}]] + [e_{\alpha_4+\alpha_5}, [e_{\alpha_6}, e_{\gamma_4}]] \\
&\quad + [e_{\alpha_6}, [e_{\alpha_8}, e_{\gamma_4}]] \\
&\quad + [e_{\alpha_2+\alpha_4}, [e_{\alpha_8}, e_{\gamma_4}]] + [e_{\alpha_6}, [e_{\alpha_2+\alpha_4}, e_{\gamma_4}]] + [e_{\alpha_8}, [e_{\alpha_2+\alpha_4}, e_{\gamma_4}] \\
&= 2[e_{\alpha_6}, [e_{\alpha_5}, e_{\gamma_1}]] + 2[e_{\alpha_2+\alpha_4}, [e_{\alpha_6}, [e_{\alpha_3}, e_\sigma]]] + 2[e_{\alpha_2+\alpha_4}, [e_{\alpha_5}, e_{\gamma_1}]] \\
&\quad + [e_{\alpha_4+\alpha_5}, [e_{\alpha_6}, [e_{\alpha_3}, e_\sigma]]] \\
&= 2e_{\delta_2} + 2[e_{\alpha_2+\alpha_4}, [e_{\alpha_3}, e_{\gamma_3}]] + 2e_{\delta_5} + [e_{\alpha_4+\alpha_5}, [e_{\alpha_3}, e_{\gamma_3}]] \\
&= 2e_{\delta_2} + e_{\delta_4} + 2e_{\delta_5} + 2e_{\delta_6},
\end{aligned}$$

$$
\begin{aligned}
(\operatorname{ad} e)^2(e_{\gamma_5}) &= [e_{\alpha_5}, [e_{\alpha_6}, e_{\gamma_5}]] + [e_{\alpha_8}, [e_{\alpha_6}, e_{\gamma_5}]] + [e_{\alpha_6}, [e_{\alpha_8}, e_{\gamma_5}]] \\
&= 2[e_{\alpha_6}, [e_{\alpha_8}, [e_{\alpha_4}, e_{\gamma_4}]]] + [e_{\alpha_5}, [e_{\alpha_6}, [e_{\alpha_4}, e_{\gamma_4}]]] \\
&= 2[e_{\alpha_6}, [e_{\alpha_4}, [e_{\alpha_5}, e_{\gamma_1}]]] + [e_{\alpha_5}, [e_{\alpha_6}, [e_{\alpha_4}, [e_{\alpha_3}, e_\sigma]]]] \\
&= 2[e_{\alpha_6}, [[e_{\alpha_4}, e_{\alpha_5}], e_{\gamma_1}]] + [e_{\alpha_5}, [e_{\alpha_4}, [e_{\alpha_3}, e_{\gamma_3}]]] \\
&= 2e_{\delta_3} + [[e_{\alpha_5}, e_{\alpha_4}], [e_{\alpha_3}, e_{\gamma_3}]] = 2e_{\delta_3} - e_{\delta_4}, \\
(\operatorname{ad} e)^2(e_{\gamma_6}) &= [e_{\alpha_6}, [e_{\alpha_7}, e_{\gamma_6}]] + [e_{\alpha_8}, [e_{\alpha_7}, e_{\gamma_6}]] \\
&= [e_{\alpha_6}, [e_{\alpha_2+\alpha_4}, e_{\gamma_4}]] + [e_{\alpha_8}, [e_{\alpha_2+\alpha_4}, e_{\gamma_4}]] \\
&= [e_{\alpha_2+\alpha_4}, [e_{\alpha_6}, [e_{\alpha_3}, e_\sigma]]] + [e_{\alpha_2+\alpha_4}, [e_{\alpha_8}, e_{\gamma_4}]] \\
&= [e_{\alpha_2+\alpha_4}, [e_{\alpha_3}, e_{\gamma_3}]] + [e_{\alpha_2+\alpha_4}, [e_{\alpha_5}, e_{\gamma_1}]] = e_{\delta_5} + e_{\delta_6}.
\end{aligned}
$$

Since

$$
\begin{vmatrix}
0 & 1 & 1 & 0 & 0 & 0 \\
1 & 2 & 2 & 2 & 0 & 0 \\
1 & 0 & 0 & 0 & 2 & 0 \\
0 & 0 & 1 & 1 & -1 & 0 \\
1 & 1 & 0 & 2 & 0 & 1 \\
0 & 0 & 1 & 2 & 0 & 1
\end{vmatrix} = 4,
$$

the map $(\operatorname{ad} e)^2 \colon M_{J,+}^{-2} \to M_{J,+}^{2}$ is bijective.

2.17. We now suppose that $R_J \cong E_7 \subset E_8$. Clearly, $B_J = B \backslash \{\alpha_8\}$. Therefore, $\mathfrak{g}_J(k) = 0$ if $k > 2$. Moreover, the $L_J^{(1)}$-module $\mathfrak{g}_J(2)$ is trivial and $\mathfrak{g}_J(1)$ is irreducible over $L_J^{(1)}$ and has highest weight ω_7^J.

By [5, p. 176], any standard distinguished parabolic subalgebra $\mathfrak{p}_I \cap \mathfrak{l}_J^{(1)}$ of $\mathfrak{l}_J^{(1)}$ has the following property:

$$7 \notin I \text{ and either } \{2, 5\} \subset I \text{ or } \{2, 5\} \cap I = \varnothing.$$

Using [4, VI, Table VI] it is easy to note that $\omega_1^J, \omega_3^J, \omega_4^J, \omega_6^J, \omega_2^J + \omega_5^J \in Q_+^J$ and $\nu_7(2\omega_7^J) = 3$. This implies that, for any $\gamma \in R$ with $\nu_8(\gamma) = 1$,

$$(\lambda_{I,J}|\gamma) \equiv (\lambda_{I,J}|\alpha_8) \equiv (2\omega_7^J|\alpha_8) \equiv 1 \pmod 2.$$

Thus all $(\operatorname{Ad} \lambda_e)$-weights of $\mathfrak{g}_J(1)$ are odd. But then $M_J^{2(p-1)} = 0$ and Lemma 2.7 applies. Therefore, λ_e is a Dynkin torus for $e \in \mathfrak{p}_I \cap \mathfrak{l}_J^{(1)}$.

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 $B_J = B \backslash \{\alpha_3\}$ and $e = \sum_{i \neq 3} e_{\alpha_i}$. It is immediate from [4, VI, Table VI] that $\mathfrak{g}_J(k) = 0$ if $k > 3$, $\beta_J^3 = {}^{112221}_{\ \ 1}$, $b_J^3 = 11$. Hence $(2\rho_J|\gamma) = 2\operatorname{ht}\gamma - 11\nu_3(\gamma)$ (see Lemma 2.9). This implies that $\Delta_3(k) = \varnothing$ if $k \neq 2$ and $\Delta_3(2) = \{\beta_1, \beta_2\}$ where $\beta_1 = {}^{122211}_{\ \ 1}$, $\beta_2 = {}^{123210}_{\ \ 1}$. Set

$$\delta_1 = {}^{123221}_{\ \ 1}, \quad \delta_2 = {}^{123211}_{\ \ 2}, \quad \eta = {}^{122210}_{\ \ 1}.$$

We may and do assume that

$$
\begin{aligned}
e_{\beta_1} &= [e_{\alpha_7}, e_\eta], & e_{\beta_2} &= [e_{\alpha_4}, e_\eta], \\
e_{\delta_1} &= [e_{\alpha_6}, [e_{\alpha_4}, e_{\beta_1}]], & e_{\delta_2} &= [e_{\alpha_2}, [e_{\alpha_7}, e_{\beta_2}]].
\end{aligned}
$$

We have

$$
\begin{aligned}
(\operatorname{ad} e)^2(e_{\beta_1}) &= [e_{\alpha_2}, [e_{\alpha_4}, e_{\beta_1}]] + [e_{\alpha_6}, [e_{\alpha_4}, e_{\beta_1}]] + [e_{\alpha_4}, [e_{\alpha_6}, e_{\beta_1}]] \\
&= 2e_{\delta_1} + [e_{\alpha_2}, [e_{\alpha_4}, [e_{\alpha_7}, e_\eta]]] = 2e_{\delta_1} + [e_{\alpha_2}, [e_{\alpha_7}, e_{\beta_2}]] = 2e_{\delta_1} + e_{\delta_2}
\end{aligned}
$$

and

$$\begin{aligned}(\operatorname{ad} e)^2(e_{\beta_2}) &= [e_{\alpha_7}, [e_{\alpha_2}, e_{\beta_2}]] + [e_{\alpha_2}, [e_{\alpha_7}, e_{\beta_2}]] + [e_{\alpha_6}, [e_{\alpha_7}, e_{\beta_2}]] \\ &= 2e_{\delta_2} + [e_{\alpha_6}, [e_{\alpha_7}, [e_{\alpha_4}, e_{\eta}]]] = 2e_{\delta_2} + [e_{\alpha_6}, [e_{\alpha_4}, e_{\beta_1}]] = e_{\delta_1} + 2e_{\delta_2}.\end{aligned}$$

Since

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

if $p > 3$, the map $(\operatorname{ad} e)^2 \colon M_{J,+}^{-2} \to M_{J,+}^{2}$ is bijective.

2.19. Let R_J be of type $D_5 \times A_1$. In this case $B_J = B \backslash \{\alpha_6\}$. First suppose that e is regular in $\mathfrak{l}_J^{(1)}$. By [4, VI, Table VI], $\mathfrak{g}_J(k) = 0$ for $k > 2$, $\beta_J^6 = \begin{smallmatrix}123211\\2\end{smallmatrix}$ and $b_J^6 = 13$. Hence $(2\rho_J|\gamma) = 2\operatorname{ht}\gamma - 13\nu_6(\gamma)$ yielding $\Delta_6(1) = \varnothing$, $\Delta_6(2) = \{\gamma\}$ where $\gamma = \begin{smallmatrix}123221\\1\end{smallmatrix}$.

As

$$\begin{aligned}(\operatorname{ad} e)^2(e_\gamma) &= \left(\sum_{i \neq 6} e_{\alpha_i}\right)^2 (e_\gamma) = [e_{\alpha_5}, [e_{\alpha_2}, e_\gamma]] + [e_{\alpha_2}, [e_{\alpha_5}, e_\gamma]] \\ &= 2[e_{\alpha_5}, [e_{\alpha_2}, e_\gamma]] \neq 0,\end{aligned}$$

we conclude that $(\operatorname{ad} e)^2 \colon M_{J,+}^{-2} \to M_{J,+}^{2}$ is one-to-one.

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

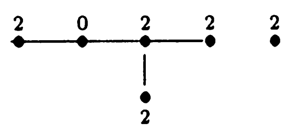
Bala-Carter diagram	Type
	$D_5(a_1) \times A_1$

TABLE 3

A simple checking shows that no generality is lost by assuming

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

Since $\lambda_{I,J} = 2\rho_J - 2\omega_4^J$ and $2\omega_4^J = \begin{smallmatrix}246300\\3\end{smallmatrix}$, we have

$$(\lambda_{I,J}|\gamma) = 2\operatorname{ht}\gamma - 13\nu_6(\gamma) + 3\nu_6(\gamma) - 2\nu_4(\gamma) = 2(\operatorname{ht}\gamma - \nu_4(\gamma) - 5\nu_6(\gamma)).$$

In particular, $(\lambda_{I,J}|\tilde{\alpha}) = 6 < 2(p-1)$. Since $(\lambda_{I,J}|\beta) \leq (\lambda_{I,J}|\tilde{\alpha})$ 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 $(\operatorname{ad} e)^2 \colon \mathfrak{g}_J^{-2}(2) \to \mathfrak{g}_J^2(2)$ is bijective.

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

$$\beta_1 = \begin{smallmatrix}111110\\0\end{smallmatrix}, \quad \beta_2 = \begin{smallmatrix}011110\\1\end{smallmatrix}, \quad \beta_3 = \begin{smallmatrix}001111\\1\end{smallmatrix},$$
$$\beta_4 = \begin{smallmatrix}011111\\0\end{smallmatrix}, \quad \beta_5 = \begin{smallmatrix}012110\\1\end{smallmatrix}.$$

Set

$$\delta_1 = {}^{111111}_{\ \ 1}, \quad \delta_2 = {}^{112111}_{\ \ 1}, \quad \delta_3 = {}^{112210}_{\ \ 1},$$
$$\delta_4 = {}^{122110}_{\ \ 1}, \quad \delta_5 = {}^{012211}_{\ \ 1}, \quad \eta = {}^{011110}_{\ \ 0}.$$

We may assume that

$$[e_{\alpha_5}, e_{\alpha_5}] = e_{\alpha_4+\alpha_5}, \quad e_{\beta_1} = [e_{\alpha_1}, e_\eta], \quad e_{\beta_4} = [e_{\alpha_7}, e_\eta],$$
$$e_{\beta_5} = [e_{\alpha_2+\alpha_4}, e_\eta], \quad [e_{\alpha_4}, e_{\beta_2}] = e_{\beta_5}, \quad [e_{\alpha_7}, e_{\beta_2}] = [e_{\alpha_3}, e_{\beta_3}],$$
$$e_{\delta_1} = [e_{\alpha_1}, [e_{\alpha_7}, e_{\beta_2}]], \quad e_{\delta_2} = [e_{\alpha_7}, [e_{\alpha_2+\alpha_4}, e_{\beta_1}]], \quad e_{\delta_3} = [e_{\alpha_5}, [e_{\alpha_2+\alpha_4}, e_{\beta_1}]],$$
$$e_{\delta_4} = [e_{\alpha_3}, [e_{\alpha_2+\alpha_4}, e_{\beta_1}]], \quad e_{\delta_5} = [e_{\alpha_4+\alpha_5}, [e_{\alpha_7}, e_{\beta_2}]].$$

Then

$$\begin{aligned}
(\operatorname{ad} e)^2(e_{\beta_1}) &= [e_{\alpha_3}, [e_{\alpha_2+\alpha_4}, e_{\beta_1}]] + [e_{\alpha_5}, [e_{\alpha_2+\alpha_4}, e_{\beta_1}]] + [e_{\alpha_7}, [e_{\alpha_2+\alpha_4}, e_{\beta_1}]] \\
&\quad + [e_{\alpha_2+\alpha_4}, [e_{\alpha_7}, e_{\beta_1}]] \\
&= 2e_{\delta_2} + e_{\delta_3} + e_{\delta_4}, \\
(\operatorname{ad} e)^2(e_{\beta_2}) &= [e_{\alpha_4+\alpha_5}, [e_{\alpha_1}, e_{\beta_2}]] + [e_{\alpha_7}, [e_{\alpha_1}, e_{\beta_2}]] + [e_{\alpha_1}, [e_{\alpha_4+\alpha_5}, e_{\beta_2}]] \\
&\quad + [e_{\alpha_7}, [e_{\alpha_4+\alpha_5}, e_{\beta_2}]] + [e_{\alpha_1}, [e_{\alpha_7}, e_{\beta_2}]] + [e_{\alpha_4+\alpha_5}, [e_{\alpha_7}, e_{\beta_2}]] \\
&= 2e_{\delta_1} + 2e_{\delta_5} - 2[e_{\alpha_1}, [e_{\alpha_5}, [e_{\alpha_4}, e_{\beta_2}]]] \\
&= 2e_{\delta_1} + 2e_{\delta_5} - 2[e_{\alpha_1}, [e_{\alpha_5}, e_{\beta_5}]] \\
&= 2e_{\delta_1} + 2e_{\delta_5} - 2[e_{\alpha_5}, [e_{\alpha_1}, [e_{\alpha_2+\alpha_4}, e_\eta]]] \\
&= 2e_{\delta_1} + 2e_{\delta_5} - 2[e_{\alpha_5}, [e_{\alpha_2+\alpha_4}, e_{\beta_1}]] = 2(e_{\delta_1} - e_{\delta_3} + e_{\delta_5}), \\
(\operatorname{ad} e)^2(e_{\beta_3}) &= [e_{\alpha_1}, [e_{\alpha_3}, e_{\beta_3}]] + [e_{\alpha_4+\alpha_5}, [e_{\alpha_3}, e_{\beta_3}]] \\
&= [e_{\alpha_1}, [e_{\alpha_7}, e_{\beta_2}]] + [e_{\alpha_4+\alpha_5}, [e_{\alpha_7}, e_{\beta_2}]] = e_{\delta_1} + e_{\delta_5}, \\
(\operatorname{ad} e)^2(e_{\beta_4}) &= [e_{\alpha_2+\alpha_4}, [e_{\alpha_1}, e_{\beta_4}]] + [e_{\alpha_1}, [e_{\alpha_2+\alpha_4}, e_{\beta_4}]] + [e_{\alpha_5}, [e_{\alpha_2+\alpha_4}, e_{\beta_4}]] \\
&= 2[e_{\alpha_2+\alpha_4}, [e_{\alpha_1}, [e_{\alpha_7}, e_\eta]]] + [e_{\alpha_5}, [e_{\alpha_2+\alpha_4}, [e_{\alpha_7}, e_\eta]]] \\
&= 2[e_{\alpha_2+\alpha_4}, [e_{\alpha_7}, e_{\beta_1}]] + [e_{\alpha_5}, [e_{\alpha_7}, e_{\beta_5}]] \\
&= 2e_{\delta_2} + [e_{\alpha_7}, [e_{\alpha_5}, [e_{\alpha_4}, e_{\beta_2}]]] = 2e_{\delta_2} + [e_{\alpha_7}, [[e_{\alpha_5}, e_{\alpha_4}], e_{\beta_2}]] \\
&= 2e_{\delta_2} - e_{\delta_5}, \\
(\operatorname{ad} e)^2(e_{\beta_5}) &= [e_{\alpha_3}, [e_{\alpha_1}, e_{\beta_5}]] + [e_{\alpha_5}, [e_{\alpha_1}, e_{\beta_5}]] \\
&\quad + [e_{\alpha_7}, [e_{\alpha_1}, e_{\beta_5}]] + [e_{\alpha_1}, [e_{\alpha_5}, e_{\beta_5}]] \\
&\quad + [e_{\alpha_7}, [e_{\alpha_5}, e_{\beta_5}]] + [e_{\alpha_1}, [e_{\alpha_7}, e_{\beta_5}]] + [e_{\alpha_5}, [e_{\alpha_7}, e_{\beta_5}]] \\
&= 2[e_{\alpha_7}, [e_{\alpha_1}, [e_{\alpha_2+\alpha_4}, e_\eta]]] + 2[e_{\alpha_1}, [e_{\alpha_5}, [e_{\alpha_2+\alpha_4}, e_\eta]]] \\
&\quad + [e_{\alpha_3}, [e_{\alpha_1}, [e_{\alpha_2+\alpha_4}, e_\eta]]] + 2[e_{\alpha_7}, [e_{\alpha_5}, [e_{\alpha_4}, e_{\beta_2}]]] \\
&= 2[e_{\alpha_7}, [e_{\alpha_2+\alpha_4}, e_{\beta_1}]] + 2[e_{\alpha_5}, [e_{\alpha_2+\alpha_4}, e_{\beta_1}]] + [e_{\alpha_3}, [e_{\alpha_2+\alpha_4}, e_{\beta_1}]] \\
&\quad + 2[e_{\alpha_7}, [[e_{\alpha_5}, e_{\alpha_4}], e_{\beta_2}]] \\
&= 2e_{\delta_2} + 2e_{\delta_3} + e_{\delta_4} - 2e_{\delta_5}.
\end{aligned}$$

Since

$$\begin{vmatrix} 0 & 2 & 1 & 0 & 0 \\ 2 & 0 & 0 & 2 & 2 \\ 1 & -2 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 2 & 1 & -1 & -2 \end{vmatrix} = -8,$$

we can exclude $D_5 \times A_1$ from our list.

2.20. Let $R_J \cong A_6 \subset E_7$. Then $B_J = B \backslash \{\alpha_2\}$, $\beta_J^2 = {}^{123321}_{\ \ 1}$, $b_J^2 = 14$ and $e = \sum_{i \neq 2} e_{\alpha_i}$. Clearly, $\mathfrak{g}_J(k) = 0$ 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

$$\beta_1 = {}^{112100}_{\ \ 1}, \quad \beta_2 = {}^{111110}_{\ \ 1}, \quad \beta_3 = {}^{012110}_{\ \ 1},$$
$$\beta_4 = {}^{011111}_{\ \ 1}, \quad \gamma = {}^{123221}_{\ \ 2}.$$

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

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

It can be easily seen that we may suppose that

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

$$[e_{\alpha_4}, e_{\beta_4}] = [e_{\alpha_7}, e_{\beta_3}], \quad e_{\delta_1} = [e_{\alpha_3}, [e_{\alpha_6}, e_{\beta_1}]], \quad e_{\delta_2} = [e_{\alpha_5}, [e_{\alpha_6}, e_{\beta_1}]],$$
$$e_{\delta_3} = [e_{\alpha_7}, [e_{\alpha_6}, e_{\beta_1}]], \quad e_{\delta_4} = [e_{\alpha_5}, [e_{\alpha_7}, e_{\beta_3}]].$$

Calculations show that

$$\begin{aligned}
(\operatorname{ad} e)^2(e_{\beta_1}) &= [e_{\alpha_6}, [e_{\alpha_3}, e_{\beta_1}]] + [e_{\alpha_3}, [e_{\alpha_6}, e_{\beta_1}]] \\
&\quad + [e_{\alpha_5}, [e_{\alpha_6}, e_{\beta_1}]] + [e_{\alpha_7}, [e_{\alpha_6}, e_{\beta_1}]] \\
&= 2e_{\delta_1} + e_{\delta_2} + e_{\delta_3}, \\
(\operatorname{ad} e)^2(e_{\beta_2}) &= [e_{\alpha_3}, [e_{\alpha_4}, e_{\beta_2}]] + [e_{\alpha_5}, [e_{\alpha_4}, e_{\beta_2}]] \\
&\quad + [e_{\alpha_7}, [e_{\alpha_4}, e_{\beta_2}]] + [e_{\alpha_4}, [e_{\alpha_7}, e_{\beta_2}]] \\
&= e_{\delta_1} + e_{\delta_2} + 2e_{\delta_3}, \\
(\operatorname{ad} e)^2(e_{\beta_3}) &= [e_{\alpha_3}, [e_{\alpha_1}, e_{\beta_3}]] + [e_{\alpha_5}, [e_{\alpha_1}, e_{\beta_3}]] \\
&\quad + [e_{\alpha_7}, [e_{\alpha_1}, e_{\beta_3}]] + [e_{\alpha_1}, [e_{\alpha_5}, e_{\beta_3}]] + [e_{\alpha_7}, [e_{\alpha_5}, e_{\beta_3}]] \\
&\quad + [e_{\alpha_1}, [e_{\alpha_7}, e_{\beta_3}]] + [e_{\alpha_5}, [e_{\alpha_7}, e_{\beta_3}]] \\
&= e_{\delta_1} + 2e_{\delta_2} + 2e_{\delta_3} + 2e_{\delta_4}, \\
(\operatorname{ad} e)^2(e_{\beta_4}) &= [e_{\alpha_4}, [e_{\alpha_1}, e_{\beta_4}]] + [e_{\alpha_7}, [e_{\alpha_4}, e_{\beta_4}]] + [e_{\alpha_5}, [e_{\alpha_4}, e_{\beta_4}]] \\
&= 2[e_{\alpha_1}, [e_{\alpha_7}, e_{\beta_3}]] + [e_{\alpha_5}, [e_{\alpha_7}, e_{\beta_3}]] \\
&= 2[e_{\alpha_7}, [e_{\alpha_6}, e_{\beta_1}]] + e_{\delta_4} = 2e_{\delta_3} + e_{\delta_4}.
\end{aligned}$$

Since

$$\begin{vmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 1 \end{vmatrix} = -7,$$

the map $(\operatorname{ad} e)^2 \colon M_{J,+}^{-2} \to M_{J,+}^{2}$ is one-to-one if $p \neq 7$. Note that if $p = 7$, then R_J has type A_{p-1}. We shall handle this case later.

2.21. Let R_J be of type $D_6 \subset E_7$. Then $B_J = B\backslash\{\alpha_1\}$, $\beta_J^1 = \genfrac{}{}{0pt}{}{134321}{2}$, $b_J^1 = 17$ and $\mathfrak{g}_J(k) = 0$ for $k > 2$. By [5, p. 175], there are two nonempty subsets $I \subset J$ for which $\mathfrak{p}_I \cap \mathfrak{l}_J^{(1)}$ is distinguished in $\mathfrak{l}_J^{(1)}$, namely, $\{4\}$ and $\{4, 6\}$. As

$$2\omega_4^J = \genfrac{}{}{0pt}{}{048642}{4}, \quad 2\omega_6^J = \genfrac{}{}{0pt}{}{024442}{2},$$

we have $(\lambda_{I,J}|\gamma) \equiv (2\rho_J|\gamma) \pmod 2$ for any $\gamma \in R$. But $(2\rho_J|\gamma) = 2\operatorname{ht}\gamma - 17$ is odd provided $\nu_1(\gamma) = 1$. It follows that all weights of λ_e on $\mathfrak{g}_J(1)$ are odd. Since $\mathfrak{g}_J(2)$ is a trivial $L_J^{(1)}$-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 $B_J = B\backslash\{\alpha_1\}$ and $M_{J,+} = \mathfrak{g}_J(1)$ is an irreducible $L_J^{(1)}$-module with highest weight ω_1^J.

Let $R_{\tilde{J}}$ be of type $E_6 \times A_1$ in E_8 (see 2.15). We suppose that $R_J \subset R_{\tilde{J}}$, $\mathfrak{l}_J^{(1)} \subset \mathfrak{l}_{\tilde{J}}^{(1)}$ and $L_J^{(1)} \subset L_{\tilde{J}}^{(1)}$. Denote by $\tilde{\mathfrak{g}}$ a Lie algebra of type E_8 that contains $\mathfrak{l}_{\tilde{J}}^{(1)}$. By (2.15), $\tilde{\mathfrak{g}}_{\tilde{J}}(2)$ is irreducible over $L_{\tilde{J}}^{(1)}$ and has highest weight $\omega_6^{\tilde{J}}$. Let $\tilde{e}$ denote a nilpotent element from Table 1. Then $\tilde{e} = e + e_{\alpha_8}$ where $e \in \mathfrak{l}_J^{(1)}$. Clearly $\lambda_{\tilde{e}}(t) = \lambda_e(t)h_{\alpha_8}(t)$ for each $t \in \mathbb{G}_m$. The $L_J^{(1)}$-modules $\tilde{\mathfrak{g}}_{\tilde{J}}(2)$ and $\mathfrak{g}_J(1)$ are dual to each other and e_{α_8} and $h_{\alpha_8}(t)$ both act trivially on $\tilde{\mathfrak{g}}_{\tilde{J}}(2)$. Therefore, we can apply a computation presented in (2.15) to conclude that $(\operatorname{ad} e)^2\colon \mathfrak{g}_J^{-2}(-1) \to \mathfrak{g}_J^2(-1)$ is bijective if $p \neq 3$ (recall that $\mathfrak{g}_J(-1) \cong \mathfrak{g}_J(1)^*$). By our remark in (2.12), it follows that

$$(\operatorname{ad} e)^2\colon M_{J,+}^{-2} \to M_{J,+}^2$$

is one-to-one.

2.23. Let R_J be of type A_5 in $R \cong E_6$. Then e is regular in $\mathfrak{l}_J^{(1)}$, $B_J = B\backslash\{\alpha_2\}$, $\beta_J^2 = \genfrac{}{}{0pt}{}{12321}{1}$ and $b_J^2 = 11$. Using this information and [4, VI, Table V] it is now easy to observe that $\Delta_2(k) = \varnothing$ 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 $M_{J,+} = \mathfrak{g}_J(1)$. By conjugating R_J by $w_0 \in W$ if necessary we obtain $B_J = B\backslash\{\alpha_1\}$. We have $\beta_J^1 = \tilde{\alpha}$, $b_J^1 = 12$, $(2\rho_J|\tilde{\alpha}) = 10$. If e is not regular in $\mathfrak{l}_J^{(1)}$, then $e \in \mathfrak{p}_I \cap \mathfrak{l}_J^{(1)}$ where $I = \{4\}$ (see [5, p. 175]). Since $2\omega_4^J = \genfrac{}{}{0pt}{}{03642}{3}$, then $m_1(e) = (\lambda_{I,J}|\tilde{\alpha}) = 7 < 2(p-1)$ 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 $\gamma_1 = \genfrac{}{}{0pt}{}{11111}{0}$, $\gamma_2 = \genfrac{}{}{0pt}{}{11110}{1}$. Set $\delta_1 = \genfrac{}{}{0pt}{}{11211}{1}$, $\delta_2 = \genfrac{}{}{0pt}{}{12210}{1}$. Then

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

$$(\operatorname{ad} e)^2(e_{\gamma_2}) = [e_{\alpha_3}, [e_{\alpha_4}, e_{\gamma_2}]] + [e_{\alpha_6}[e_{\alpha_4}, e_{\gamma_2}]] + [e_{\alpha_4}, [e_{\alpha_6}, e_{\gamma_2}]] = \pm 2e_{\delta_1} \pm e_{\delta_2}.$$

It follows that $(\operatorname{ad} e)^2\colon M_{J,+}^{-2} \to M_{J,+}^2$ is bijective for any prime p.

2.24. We now turn to the case when R_J has rank $l-2$ in R. If $B_J = B\backslash\{\alpha_i, \alpha_j\}$ and $k > 0$, then $\mathfrak{g}_J(k)$ is spanned by all e_γ such that $\nu_i(\gamma)+\nu_j(\gamma) = k$. Given $a, b \in \mathbb{Z}_+$ denote

$$\Delta_{i,j}(a, b) = \{\gamma \in R_+ | \nu_i(\gamma) = a, \nu_j(\gamma) = b, (2\rho_J|\gamma) = -2\}.$$

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 $B_J = B\backslash\{\alpha_2, \alpha_7\}$. Using [4, VI, Table VI] we get $\beta_J^2 = {}^{123210}_{\ \ 1}$, $\beta_J^7 = {}^{111111}_{\ \ 0}$, $b_J^2 = 11$, $b_J^7 = 7$. Applying Lemma 2.9 yields $\Delta_{2,7}(1, 0) = \Delta_{2,7}(2, 1) = \Delta_{2,7}(2, 0) = \varnothing$ and $\Delta_{2,7}(1, 1) = \{\beta_1, \beta_2\}$ where

$$\beta_1 = {}^{112111}_{\ \ 1}, \quad \beta_2 = {}^{012211}_{\ \ 1}.$$

Put $\delta_1 = {}^{12221}_{\ \ 1}$, $\delta_2 = {}^{112221}_{\ \ 1}$, $e_{\delta_1} = [e_{\alpha_3}, [e_{\alpha_5}, e_{\beta_1}]]$, $e_{\delta_2} = [e_{\alpha_6}, [e_{\alpha_5}, e_{\beta_1}]]$. Without loss of generality we may assume that $e = \sum_{i \neq 2,7} e_{\alpha_i}$ and $[e_{\alpha_5}, e_{\beta_1}] = [e_{\alpha_1}, e_{\beta_2}]$. Then

$$(\operatorname{ad} e)^2(e_{\beta_1}) = [e_{\alpha_5}, [e_{\alpha_3}, e_{\beta_1}]] + [e_{\alpha_3}, [e_{\alpha_5}, e_{\beta_1}]] + [e_{\alpha_6}, [e_{\alpha_5}, e_{\beta_1}]] = 2e_{\delta_1} + e_{\delta_2},$$

$$(\operatorname{ad} e)^2(e_{\beta_2}) = [e_{\alpha_3}, [e_{\alpha_1}, e_{\beta_2}]] + [e_{\alpha_6}, [e_{\alpha_1}, e_{\beta_2}]] + [e_{\alpha_1}, [e_{\alpha_6}, e_{\beta_2}]] = e_{\delta_1} + 2e_{\delta_2}.$$

Since

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

if $p \neq 3$, we conclude that $(\operatorname{ad} e)^2\colon M_{J,+}^{-2} \to M_{J,+}^2$ 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 $B_J = B\backslash\{\alpha_6, \alpha_7\}$. Let $\bar{\mathfrak{g}}$ denote the subalgebra of type E_6 generated by $e_{\pm\alpha_i}$, $i \leq 6$. Then $\mathfrak{g}_J(1) = \bar{\mathfrak{g}}_J(1) \oplus Ke_{\alpha_7}$. Clearly, Ke_{α_7} is a trivial $L_J^{(1)}$-module. Using [4, VI, Table VI] it is easy to check that the $L_J^{(1)}$-modules $\bar{\mathfrak{g}}_J(1)$ and $\mathfrak{g}_J(2)$ are isomorphic. Applying the second part of (2.22) we obtain now that

$$(\operatorname{ad} e)^2\colon \mathfrak{g}_J^{-2}(k) \to \mathfrak{g}_J^2(k)$$

is bijective for $k = 1, 2$. Clearly, $\beta_J^6 = {}^{123210}_{\ \ 2}$, $\beta_J^7 = \alpha_7$, $b_J^6 = 12$, $b_J^7 = 2$.

The subspace $\mathfrak{g}_J(3)$ is spanned by all e_γ with $\nu_6(\gamma) = 2$ and $\nu_1(\gamma) = 1$. By Lemma 2.9, $(2\rho_J|\tilde{\alpha}) = 34 - 2 \cdot 12 - 2 = 8 \leq 2(p-1)$. It follows that $m_3(e) = (\lambda_{I,J}|\tilde{\alpha}) < 2(p-1)$ provided $I \neq \varnothing$. Hence we may assume that $e = \sum_{i \leq 6} e_{\alpha_i}$. By Lemma 2.9, $\Delta_{6,7}(2, 1) = \{\gamma\}$ where $\gamma = {}^{123221}_{\ \ 1}$. Since $(\operatorname{ad} e)^2(e_\gamma) = 2[e_{\alpha_2}, [e_{\alpha_5}, e_\gamma]] \neq 0$, we can exclude D_5 from our list.

Let R_J be of type D_6 in $R \cong E_8$. In this case $B_J = B\backslash\{\alpha_1, \alpha_8\}$. For any $k > 0$, the $L_J^{(1)}$-module $\mathfrak{g}_J(k)$ is completely reducible. Moreover, the highest weights of the irreducible submodules of $\mathfrak{g}_J(k)$ lie in the set $\{0, \omega_2^J, \omega_3^J, \omega_7^J\}$ (see (2.10)). By (2.10), $(2\rho_J|\omega_2^J) = (2\rho_J|\omega_3^J) = 15$, $(2\rho_J|\omega_7^J) = 10 < 2(p-1)$. Reasoning as in (2.21) it can now be easily seen that $(\lambda_{I,J}|\omega_2^J) = (\lambda_{I,J}|\omega_3^J)$ is odd for any $I \subset J$ such that $\mathfrak{p}_I \cap \mathfrak{l}_J^{(1)}$ is distinguished in $\mathfrak{l}_J^{(1)}$. Summarizing we obtain that each $(\operatorname{Ad}\lambda_e)$-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, λ_e is a Dynkin torus for $e \in \mathfrak{p}_I \cap \mathfrak{l}_J^{(1)}$.

If R_J has type E_6 in $R \cong E_8$, then $B_J = B\backslash\{\alpha_7, \alpha_8\}$. 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 $s \in X(\lambda_e) \cong \mathbb{Z}$ under the action of $\lambda_e \subset L_J^{(1)}$ on $E_J(\omega)$. The Lie algebra $\mathfrak{l}_J^{(1)} = \operatorname{Lie}(L_J^{(1)})$ acts on $E_J(\omega)$ via the differential $d\rho$ of a rational representation $\rho\colon L_J^{(1)} \to GL(E_J(\omega))$.

If V is a nontrivial irreducible $L_J^{(1)}$-submodule of M_J, then either $V \cong E_J(\omega_1^J)$ or $V \cong E_J(\omega_6^J)$ (see 2.10)). Combining (2.22) with the computation in (2.15) one easily sees that the map $(dp(e))^2\colon E_J(\omega_i^J)_{-2} \to E_J(\omega_i^J)_2$ is a bijection if $i = 1, 6$, $p \neq 3$ (note that $E_J(\omega_1^J)$ is contragradient to $E_J(\omega_6^J)$). From this it is immediate that $(\operatorname{ad} e)^2\colon M_J^{-2} \to M_J^2$ 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 $(\operatorname{ad} e)^2\colon M_J^{-2} \to M_J^2$ is a bijection, then so is $(\operatorname{ad} e)^2\colon \mathfrak{g}_{-2} \to \mathfrak{g}_2$. Applying Lemma 2.3 shows now that in all examined cases λ_e is a Dynkin torus for e.

It remains to consider the following subsystems $R_J \subset R$:

$$\begin{array}{ll} A_4,\ A_4 \times A_1,\ A_4 \times A_2 \text{ for } R \cong E_6 \text{ or } E_7, & p \geq 5; \\ A_6 \text{ for } R \cong E_7 \text{ or } E_8,\ A_6 \times A_1 \text{ for } R \cong E_8, & p \geq 7. \end{array}$$

In all these cases we can suppose that $e = \sum_{i \in J} e_{\alpha_i}$. 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}, λ_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 $H = \operatorname{diag}(t, t^{-1})$ 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 (θ, X, Y) where θ is a rational representation of $\mathbb{G}_m$ in V and X, Y are endomorphisms of V such that

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

where $d\theta$ is the differential of θ.

It is well known (see, for example, [7]) that for any $n = 0, 1, \dots, p-1$ 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 Π 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 nth symmetric power of the standard representation of $SL_2(K)$, the weights of $V_{n,k}$ are $n+kp$, $n-2+kp, \dots, -n+2+kp$, $-n+kp$.

For any $k \in \mathbb{Z}$, the module $V_{p-1,k}$ is projective. For any $n = 0, 1, \dots, p-2$ 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 $\theta(\mathbb{G}_m)$. 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* $\operatorname{Ker} X \subset \sum_{s\geq 0} M_s$.

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

$$\operatorname{Ker} X \subset \sum_{s\geq 0} M_s$$

as desired. □

2.27. Since the triple $(\lambda_e, \operatorname{ad} e, \operatorname{ad} f)$ 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 $R_J \cong A_6 \times A_1$, then $B_J = B\backslash\{\alpha_3\}$. Hence $M_{J,+} = \mathfrak{g}_J(1) \oplus \mathfrak{g}_J(2) \oplus \mathfrak{g}_J(3) \oplus \mathfrak{g}_J(4)$. Looking over [4, VI, Table VII] one obtains that $\mathfrak{g}_J(1)$, $\mathfrak{g}_J(2)$, $\mathfrak{g}_J(3)$ and $\mathfrak{g}_J(4)$ are irreducible over $L_J^{(1)}$ and have highest weights $\omega_1^J + \omega_7^J$, ω_5^J, $\omega_1^J + \omega_2^J$ and ω_8^J 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:

$$\mathfrak{g}_J(1) \cong N_2 \otimes (\overset{2}{\wedge} N_7)^*, \quad \mathfrak{g}_J(2) \cong \overset{3}{\wedge} N_7, \quad \mathfrak{g}_J(3) \cong N_2 \otimes N_7, \quad \mathfrak{g}_J(4) \cong N_7^*.$$

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 σ_2 (resp., σ_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 $Ke \oplus Kh \oplus Kf \subset \operatorname{Lie}(L_J^{(1)})$ together with the rational representation $\sigma_2 \circ \lambda_e$ (resp., $\sigma_7 \circ \lambda_e$) of $\mathbb{G}_m$ defines a representation $\theta_2 = (\sigma_2 \circ \lambda_e, d\sigma_2(e), d\sigma_2(f))$ (resp., $\theta_7 = (\sigma_7 \circ \lambda_e, d\sigma_7(e), d\sigma_7(f))$) of the group scheme $(SL_2)_{(H)}$ in the vector space N_2 (resp., N_7). It is immediate from the above remarks that

$$(\lambda_e, \operatorname{ad} e, \operatorname{ad} f)|_{M_{J,+}} \cong \theta_2 \otimes (\overset{2}{\wedge}\theta_7)^* \oplus \overset{3}{\wedge}\theta_7 \oplus \theta_2 \otimes \theta_7 \oplus \theta_7^*.$$

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 $\overset{2}{\wedge}\theta_7$ and $\overset{3}{\wedge}\theta_7$ 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 $m_1(e) = 11$, $m_2(e) = 12$, $m_3(e) = 7$ and $m_4(e) = 6$. Hence $s < 2p-1 = 13$ for any

$s \in X(M_{J,+})$. Applying Lemma 2.10 we get

$$\mathfrak{z}_{\mathfrak{g}}(e) \cap M_{J,+} \subset \sum_{i\geq 0} \mathfrak{g}_i.$$

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

$$\mathfrak{z}_{\mathfrak{g}}(e) \cap M_J \subset \sum_{i\geq 0} \mathfrak{g}_i.$$

By construction, $\mathfrak{l}_J \cap \sum_{i\geq 0} \mathfrak{g}_i = \mathfrak{p}_I \cap \mathfrak{l}_J$. Since e is a Richardson element of $\mathfrak{p}_I \cap \mathfrak{l}_J$, then $\mathfrak{z}_{\mathfrak{g}}(e) \cap \mathfrak{l}_J \subset \sum_{i\geq 0} \mathfrak{g}_i$ in view of (2.2) (recall that $\mathfrak{l}_J$ admits a nondegenerate trace form). Therefore, $\mathfrak{z}_{\mathfrak{g}}(e) \subset \sum_{i\geq 0} \mathfrak{g}_i$ and so λ_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 $p > 3$,

$$\mathfrak{z}_{\mathfrak{g}}(e_{\alpha_i}) \subseteq \mathfrak{t} \bigoplus \sum_{(\gamma|\alpha_i)\geq 0} Ke_\gamma = \sum_{i\geq 0} \mathfrak{g}_i$$

whence λ_e is a Dynkin torus for $e = e_{\alpha_i}$.

If R is of type F_4, then $G \cong \mathrm{Aut}(\mathfrak{g})$. We regard $\mathfrak{g}$ as a subalgebra of a Lie algebra $\tilde{\mathfrak{g}}$ of type E_6. Let σ denote the outer automorphism of $\tilde{\mathfrak{g}}$ defined by extending

$$\sigma(e_{\pm\alpha_1}) = e_{\pm\alpha_6}, \quad \sigma(e_{\pm\alpha_6}) = e_{\pm\alpha_1}, \quad \sigma = (e_{\pm\alpha_3}) = e_{\pm\alpha_5},$$
$$\sigma(e_{\pm\alpha_5}) = e_{\pm\alpha_3}, \quad \sigma(e_{\pm\alpha_2}) = e_{\pm\alpha_2}, \quad \sigma(e_{\pm\alpha_4}) = e_{\pm\alpha_4}.$$

It is well known that $\mathfrak{g}$ is isomorphic to the subalgebra $\tilde{\mathfrak{g}}^\sigma = \{x \in \tilde{\mathfrak{g}} | x^\sigma = x\}$. 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 $\mathfrak{l}_J^{(1)} \subset \mathfrak{g}$ 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 $\widetilde{J} \subset \{1, 2, \dots, 6\}$ such that $e = \sum_{i\in\widetilde{J}} e_{\alpha_i}$. Therefore, e is a regular nilpotent element of the standard Levi subalgebra of $\tilde{\mathfrak{g}}$ associated to the subset $\widetilde{J}$.

We may assume that $\tilde{\mathfrak{g}} = \mathrm{Lie}(\widetilde{G})$ where $\widetilde{G}$ is a simply connected group of type E_6. It has already been proved that $\lambda_e \subset X_*(\widetilde{G})$ is a Dynkin torus for $e \in \tilde{\mathfrak{g}}$. The automorphism σ is induced by the nontrivial symmetry of the Dynkin diagram of type E_6. Clearly, the subset $B_{\widetilde{J}} = \{\alpha_i | i \in \widetilde{J}\}$ is σ-stable. Since the scalar product $(\ |\ )$ is σ-stable as well, σ acts on the set $\{\omega_i^{\widetilde{J}} | i \in \widetilde{J}\}$. It follows that $\sigma(\rho_{\widetilde{J}}) = \rho_{\widetilde{J}}$. As

$$(\mathrm{Ad}\,\lambda_e(t)) \cdot e_\gamma = t^{(2\rho_{\widetilde{J}}|\gamma)} e_\gamma$$

for each $\gamma \in \{\pm\alpha_i | 1 \leq i \leq 6\}$ we conclude that

$$\sigma(\mathrm{Ad}\,\lambda_e(t))\sigma^{-1} = \mathrm{Ad}\,\lambda_e(t)$$

for each $t \in \mathbb{G}_m$. Therefore, $\operatorname{Ad}\lambda_e$ acts on $\tilde{\mathfrak{g}}^\sigma = \mathfrak{g}$. Let $\overline{\lambda}_e\colon \mathbb{G}_m \to \operatorname{Aut}\mathfrak{g}$ denote the homomorphism induced by restricting $\operatorname{Ad}\lambda_e$ to $\tilde{\mathfrak{g}}^\sigma$. Clearly, $\overline{\lambda}_e \in X_*(G)$ and $\overline{\lambda}_e(t)\cdot e = t^2 e$ for any $t \in \mathbb{G}_m$. Let $\tilde{\mathfrak{g}}_i$ (resp., $\mathfrak{g}_i$) be the weight component of $\operatorname{Ad}\lambda_e$ (resp., $\operatorname{Ad}\overline{\lambda}_e$) corresponding to weight $i \in \mathbb{Z}$. Obviously, $\mathfrak{g}_i = \tilde{\mathfrak{g}}_i \cap \mathfrak{g}$. But then

$$\mathfrak{z}_{\mathfrak{g}}(e) = \mathfrak{z}_{\tilde{\mathfrak{g}}}(e) \cap \mathfrak{g} \subseteq \left(\sum_{i\geq 0} \tilde{\mathfrak{g}}_i\right) \cap \mathfrak{g} = \sum_{i\geq 0} \mathfrak{g}_i.$$

Therefore, $\overline{\lambda}_e$ is a Dynkin torus for $e \in \mathfrak{g}$.

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 $\mathfrak{g}$ that contains a nonregular distinguished nilpotent element. Let e be such an element. By (2.2) and [5, p. 174], we may assume that $e \in \mathfrak{l}_{J_0} \cap \mathfrak{g}_{I_0}(2)$ where $I_0 = \{3\}$.

Using [4, VI, Table VIII] we get $\alpha_2 = \varepsilon_3 - \varepsilon_4$, $\alpha_3 = \varepsilon_4$, $\alpha_4 = \frac{1}{2}(\varepsilon_1 - \varepsilon_2 - \varepsilon_3 - \varepsilon_4)$. Computation shows that $\omega_2^J = \frac{3}{2}\alpha_2 + 2\alpha_3 + \alpha_4$ and $\omega_4^J = \alpha_2 + 2\alpha_3 + 2\alpha_4$. Therefore, $\lambda_{I_0, J_0} = 2(\omega_2^J + \omega_4^J) = 5\alpha_2 + 8\alpha_3 + 6\alpha_4 = m_2\alpha_2 + 2m_3\alpha_3 + 2m_4\alpha_4$ (as $(\alpha_2|\alpha_2) = 2$ and $(\alpha_3|\alpha_3) = (\alpha_4|\alpha_4) = 1$). 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 $\gamma \in R_+$, $M_{J_0,+} = \mathfrak{g}_{J_0}(1) \oplus \mathfrak{g}_{J_0}(2)$. It is immediate from [4, VI, Table VIII] that $\mathfrak{g}_{J_0}(2)$ is trivial over $L_{J_0}^{(1)}$ and the $L_{J_0}^{(1)}$-module $\mathfrak{g}_{J_0}(1)$ is generated by the highest weight vector e_γ where $\gamma = 1\,3\,4\,2$.

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

$$(\operatorname{Ad}\lambda_e(t))\cdot e_\gamma = t^5 e_\gamma.$$

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 λ_e is a Dynkin torus for $e \in \mathfrak{l}_{J_0}$.

The proof of Theorem 2.5 is now complete.

REFERENCES

1. P. Bala and R. W. Carter, *Classes of unipotent elements in simple algebraic groups*. I, Math. Proc. Cambridge Philos. Soc. **79** (1976), 401–425.
2. ——, *Classes of unipotent elements in simple algebraic groups*. II, Math. Proc. Cambridge Philos. Soc. **80** (1976), 1–18.
3. A. Borel, *Linear algebraic groups*, Benjamin, New York, 1969.
4. N. Bourbaki, *Groupes et algèbres de Lie*, Chapitres IV–VI, Hermann, Paris, 1968.
5. R. W. Carter, *Finite groups of Lie type, conjugacy classes and complex characters*, Wiley, Chichester and New York, 1985.
6. C. W. Curtis, *Noncommutative extensions of Hilbert rings*, Proc. Amer. Math. Soc. **4** (1953), 945–955.
7. S. Gelfand and D. Kazhdan, *Examples of tensor categories*, Invent. Math. **109** (1992), 595–617.
8. N. Jacobson, *Lie algebras*, Interscience, New York, 1962.
9. V. G. Kac, *Some remarks on nilpotent orbits*, J. Algebra **64** (1980), 190–213.
10. G. R. Kempf, *Instability in invariant theory*, Ann. of Math. (2) **108** (1978), 299–316.

11. F. C. Kirwan, *Cohomology of quotients in symplectic and algebraic geometry*, Math. Notes, no. 31, Princeton Univ. Press, Princeton, NJ, 1984.
12. L. Ness, *A stratification of the null-cone via the moment map*, Amer. J. Math. **106** (1984), 1281–1329.
13. K. Pommerening, *Über die unipotenten Klassen reduktiver Gruppen*, J. Algebra **49** (1977), 525–536.
14. ——, *Über die unipotenten Klassen reduktiver Gruppen.* II, J. Algebra **65** (1980), 373–398.
15. A. Premet, *Irreducible representations of Lie algebras of reductive groups and the Kac-Weisfeiler conjecture*, Invent. Math. (to appear).
16. G. Rousseau, *Immeubles sphériques et théorie des invariants*, C. R. Acad. Sci. Paris **286** (1978), 247–250.
17. P. Slodowy, *Simple singularities and simple algebraic groups*, Lecture Notes in Math., vol. 815, Springer, Berlin and Heidelberg, 1980.
18. ——, *Die Theorie der optimalen Einparametergruppen für instabile Vektoren*, Algebraic Transformation Groups and Invariant Theory, Birkhäuser, Basel, Boston, and Berlin, pp. 115–131.
19. N. Spaltenstein, *On unipotent and nilpotent elements of groups of type* E_6, J. London Math. Soc. (2) **27** (1983), 413–420.
20. T. A. Springer, *The Steinberg function of a finite Lie algebra*, Invent. Math. **58** (1980), 211–215.
21. T. A. Springer and R. Steinberg, *Conjugacy classes*, Seminar on Algebraic Groups and Related Topics, Lecture Notes in Math., vol. 131, Springer, Berlin, Heidelberg, and New York, 1970, pp. 167–266.
22. R. Steinberg, *Lectures on Chevalley groups*, Yale Univ., New Haven, Conn., 1968.
23. B. Ju. Weisfeiler and V. G. Kac, *Irreducible representations of Lie p-algebras*, Funct. Anal. Appl. **5** (1971), 28–36.

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

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

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