GROUP GRADINGS ON THE LIE ALGEBRA $\mathfrak{psl}_n$ IN POSITIVE CHARACTERISTIC

YURI BAHTURIN AND MIKHAIL KOCHETOV

Abstract. In this paper we describe all gradings by an abelian group G on the simple Lie algebra $\mathfrak{psl}_n(F)$ where F is an algebraically closed field of characteristic p different from 2 and dividing n.

1. Introduction

We study group gradings on finite-dimensional simple Lie algebras over an algebraically closed field F. In the case $\operatorname{char} F = 0$, all gradings on the classical simple Lie algebras (except of type $\mathcal{D}_4$) have been described in [2, 5, 3]. It turns out that the same description is valid if $\operatorname{char} F = p > 0$, $p \neq 2$, for the types $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$, except $\mathcal{A}_{pk-1}$ [1]. It is the latter case that we settle in this paper.

Let $R = M_n(F)$ where F is an algebraically closed field, $\operatorname{char} F = p > 0$, $p \neq 2$, and p divides n. Then $[R, R]$ contains the center $Z(R)$ and $L := \mathfrak{psl}_n(F) = [R, R]/Z(R)$ is a simple Lie algebra. We will show that, except when $n = p = 3$, all gradings on L can be obtained essentially in the same way as in the case of $\mathfrak{sl}_n(F)$, with $\operatorname{char} F = 0$ or $\operatorname{char} F = p$ not dividing n, from the gradings on the full matrix algebra R (Theorem 5.1) and thus can be completely described.

We recall the description of gradings on R in Section 2. The reduction of gradings on L to gradings on R is first done for the case of p-groups. We use duality (recalled in Section 3) to translate the problem to the action of a certain divided power Hopf algebra on L and prove in Section 4 that any such Hopf action can be lifted to an action on R, regarded as an associative algebra (Theorem 4.5). Then we extend these results in Section 5 to arbitrary finite abelian groups (Theorem 5.1).

2. Gradings on matrix algebras

First we recall the classification of gradings (up to isomorphism, i.e., conjugation by a nonsingular matrix) on the full matrix algebra $R = M_n(F)$ over an algebraically closed field F by an arbitrary group G [4]. There exist graded unital subalgebras $A \cong M_k(F)$ and $B \cong M_l(F)$ in R such that $R = A \otimes B$ (thus $kl = n$), A has a *"fine"* grading, i.e., $\dim A_g \leq 1$ for each $g \in G$, and B has an *elementary* grading defined by an l-tuple $(g_1, \ldots, g_l)$ of elements of G, i.e., $B_g = \operatorname{span}\{E_{ij} \mid g_i^{-1} g_j = g\}$ for each $g \in G$, where $\{E_{ij}\}$ is a basis of matrix units in B. For abelian G, all "fine" gradings have also been classified. In particular, the *support* of the "fine" grading on A (i.e., the set of all $g \in G$ such that $A_g \neq 0$) is a subgroup $H \subset G$ of order k^2, and $\operatorname{char} F \nmid k$ [4, Theorem 8]. In particular, when $\operatorname{char} F = p$ and the

2000 Mathematics Subject Classification: 17B70, 16W30.

The first author was partially supported by NSERC grant # 227060-04. The second author was supported by NSERC grant # 341792-07.

Keywords: graded algebra, simple Lie algebra, gradings.

torsion subgroup of G is a p-group, then all G-gradings on $M_n(F)$ are elementary. We will need this fact later.

Consider the case of an elementary grading on $R = M_n(F)$. Conjugating by a permutation matrix, we may assume that the n-tuple has the form:

$$(g_1^{(k_1)}, \dots, g_l^{(k_l)}) \text{ with } k_1 + \dots + k_l = n,$$

where $g_1, \dots, g_l$ are pairwise distinct and we have written $g^{(k)}$ for $\underbrace{g, \dots, g}_{k \text{ times}}$.

Consider the block decomposition of R induced by the partition $n = k_1 + \dots + k_l$. Then the identity component R_1 consists of the block diagonal matrices and hence is the direct product of the full matrix algebras $S_1 \cong M_{k_1}(F), \dots, S_l \cong M_{k_l}(F)$. We will also need this later.

Suppose $R = \bigoplus_{g\in G} R_g$ is any G-graded algebra. If $* : R \to R$ is an involution, then we say that $*$ is *preserves* the grading if $(R_g)^* = R_g$ for all $g \in G$. For any grading on $R = M_n(F)$ by an abelian group G, a complete description of involutions preserving the grading was given in [3] (over an algebraically closed field F of characteristic different from 2).

3. Duality between gradings and actions

Let G be a finite group, F an algebraically closed field. Let $H = FG$ be the group algebra of G viewed as a Hopf algebra with comultiplication $\Delta(g) = g \otimes g$, counit $\varepsilon(g) = 1$, and antipode $S(g) = g^{-1}$, for all $g \in G$. We will use Sweedler's notation: $\Delta(h) = \sum h_1 \otimes h_2$, for any $h \in H$. For basic facts on Hopf algebras the reader is referred to [11].

Let A be an algebra over F, not necessarily associative. It is well-known that a G-grading on A is equivalent to the structure of a right H-comodule algebra, i.e., a homomorphism of algebras $\rho : A \to A \otimes H$ such that $(\rho \otimes id)\rho = (id \otimes \Delta)\rho$ and $(id \otimes \varepsilon)\rho = id$. Namely, if $A = \bigoplus_{g\in G} A_g$ is a G-graded algebra, then the mapping ρ is defined on a homogeneous element a of degree g by $\rho(a) = a \otimes g$. Conversely, given $\rho : A \to A \otimes H$, one can define a G-grading on A by setting $A_g = \{a \in A \mid \rho(a) = a \otimes g\}$, for any $g \in G$.

Consider the dual Hopf algebra $K = H^*$. Let $\{p_g \mid g \in G\}$ be the basis of K dual to $\{g \mid g \in G\}$, i.e., $p_g \in K$ are such that $\langle p_g, h\rangle = \delta_{g,h}$ for any $h \in G$ (Kronecker's delta). Then the multiplication in K is given by $p_{g'} p_{g''} = \delta_{g',g''} p_{g'}$ and the comultiplication by $\Delta(p_g) = \sum_{g',g''\in G:\, g'g''=g} p_{g'} \otimes p_{g''}$.

The structure of an H-comodule is equivalent to the structure of a K-module in the usual way: K acts on an H-comodule A by $f \cdot a = (id \otimes f)\rho(a)$, which in our case reads $f \cdot a = \langle f, g\rangle a$ for all $a \in A_g$, $g \in G$. In particular, the elements p_g act as the projections on the respective homogeneous components. If A is an H-comodule algebra, then it becomes a K-module algebra, i.e., we have

$$k \cdot (ab) = \sum (k_1 \cdot a)(k_2 \cdot b) \text{ for all } k \in K,\ a, b \in A.$$

Conversely, if A is a K-module algebra, then there exists a homomorphism of algebras $\rho : A \to A \otimes H$ such that K acts on A by $f \cdot a = (id \otimes f)\rho(a)$.

If $f \in K$ is a group-like element, i.e., $\Delta(f) = f \otimes f$ (hence $S(f) = f^{-1}$), then f acts on A as an automorphism: $f \cdot (ab) = (f \cdot a)(f \cdot b)$ for any $a, b \in A$. The group-like elements of K are the algebra homomorphisms $H \to F$, so their set can be identified with the group $\widehat{G}$ of multiplicative characters of G. It follows that if

G is abelian and $\operatorname{char} F$ does not divide $|G|$, then $K \cong F\widehat{G}$ as Hopf algebras and thus in this case G-gradings on an algebra A are equivalent to $\widehat{G}$-actions on A by automorphisms.

If $f \in K$ is primitive, i.e., $\Delta(f) = f \otimes 1 + 1 \otimes f$, then f acts on A as a derivation: $f \cdot (ab) = (f \cdot a)b + a(f \cdot b)$ for any $a, b \in A$. It is easy to check that the primitive elements of K are precisely the additive characters of G.

For example, let $G = \langle a_1 \rangle_p \times \cdots \times \langle a_k \rangle_p$, an elementary abelian p-group. Then there exist k additive characters $\alpha_1, \ldots, \alpha_k$ such that $\alpha_i(a_j) = \delta_{i,j}$. The span of the elements α_i in K is an abelian p-Lie algebra $\mathfrak{g}$, with $(\alpha_i)^p = \alpha_i$, and K is isomorphic to the restricted enveloping algebra $u(\mathfrak{g})$. Thus in this case G-gradings on an algebra A are equivalent to $\mathfrak{g}$-actions on A by derivations.

Now let G be any finite abelian group and F an algebraically closed field of characteristic $p > 0$. We can write $G = G_0 \times G_1$ where G_0 is of order not divisible by p and G_1 is a p-group. This induces the following decompositions of H and K: $H = H_0 \otimes H_1$ where $H_0 = FG_0$ and $H_1 = FG_1$, and $K = K_0 \otimes K_1$ where $K_0 = (H_0)^*$ and $K_1 = (H_1)^*$. Therefore, the structure of a K-module algebra on A is equivalent to a pair of mutually commuting actions on A by K_0 and by K_1 that make A a K_0-module algebra, resp., K_1-module algebra.

More generally, let G be a finitely generated abelian group. Then we can write $G = G_0 \times G_1$ where G_1 is the p-torsion subgroup of G and G_0 has trivial p-torsion. Then a G_0-grading on a finite-dimensional algebra A is equivalent to an action of the algebraic group $\widehat{G_0}$ by automorphisms of A, i.e., to a homomorphism of algebraic groups $\widehat{G_0} \to \operatorname{Aut}(A)$. Consequently, a G-grading on A is equivalent to a pair of mutually commuting actions on A: namely, $\widehat{G_0}$ acts by automorphisms and $K_1 = (FG_1)^*$ acts in a way to make A a K_1-module algebra.

From the above discussion it follows that in the case when G_1 is elementary, any problem on G-gradings can be reformulated in terms of automorphisms and derivations. If G_1 is not elementary, however, the situation is more complicated and involves the so-called divided power algebras.

Consider the case of a cyclic group $G = \langle a \rangle_{p^N}$. Then $H = F[t]/(t^{p^N} - 1) = F[\xi]/(\xi^{p^N})$ where $\xi = t - 1$. Let $\{\delta^{(m)} \mid m = 0, \ldots, p^N - 1\}$ be the basis of K dual to $\{\xi^m \mid m = 0, \ldots, p^N - 1\}$. Then the comultiplication in K is given by

$$\Delta\delta^{(m)} = \sum_{i=0}^{m} \delta^{(i)} \otimes \delta^{(m-i)}. \tag{1}$$

Elements $\delta^{(m)}$ with coproduct of this form are sometimes called "divided powers". In particular, $\delta^{(0)} = 1$ and $\delta^{(1)}$ spans the space of primitive elements of K. One can also write an explicit formula for the product $\delta^{(i)}\delta^{(j)}$, but we will only need that

$$\delta^{(i)}\delta^{(j)} = \binom{i+j}{i} \delta^{(i+j)} \pmod{\operatorname{span}\{\delta^{(m)} \mid m < i + j\}}.$$

It follows (see e.g. [8, Chapter II, §2, 6]) that, for any $1 \le l \le N$, the subspace

$$K_l = \operatorname{Span}\{\delta^{(m)} \mid 0 \le m < p^l\}$$

is a subalgebra of K, which is generated by the elements $\delta^{(p^k)}$, $k = 0, \ldots, l - 1$. Moreover, the monomials

$$\left(\delta^{(p^0)}\right)^{m_0} \left(\delta^{(p^1)}\right)^{m_1} \cdots \left(\delta^{(p^{l-1})}\right)^{m_{l-1}} \text{ where } 0 \le m_k < p \quad \forall k = 0, \ldots, l - 1$$

form a basis of K_l. One also checks that

$$\left(\delta^{(m)}\right)^p = \delta^{(m)} \text{ for all } m = 0, \ldots, p^N - 1.$$

Indeed, $\langle (\delta^{(m)})^p\,,\, g\rangle = \langle \delta^{(m)} \otimes \cdots \otimes \delta^{(m)}\,,\, g \otimes \cdots \otimes g\rangle = \langle \delta^{(m)}, g\rangle^p = \langle \delta^{(m)}, g\rangle$ for all $g \in G$ (the last equality holds, because $\langle \delta^{(m)}, g\rangle$ lies in the prime field).

In particular, for $N > 1$ the algebra K is not generated by primitive elements and, consequently, we will have to consider operators with more complicated "product expansion laws" than the ordinary Leibniz rule. Namely, (1) implies that an algebra A with an action of K is a K-module algebra if and only if, for all $0 \le m < p^N$, we have

$$\delta^{(m)}(ab) = (\delta^{(m)} \cdot a)b + a(\delta^{(m)} \cdot b) + \sum_{k=1}^{m-1} (\delta^{(k)} \cdot a)(\delta^{(m-k)} \cdot b) \qquad \forall a, b \in A. \tag{2}$$

We will need the following technical lemma in the proof of our main result in Section 4. Let $R = M_n(F)$ where $\operatorname{char} F = p \neq 2$ and $p \mid n$, so $Z = Z(R)$ is contained in $[R, R]$. Assume that, if $p = 3$, then $n \neq 3$. Let $L = [R, R]/Z$. Let $\{E_{ij}\}$ be a basis of matrix units for R. We will use the notation $e_{ij} = E_{ij} + Z$. Now let $K = (FG)^*$ where, as above, $G = \langle a\rangle_{p^N}$. Let $q = p^{N-1}$.

Lemma 3.1. *Suppose the Lie algebra L is a K-module algebra in such a way that all e_{ij}, $i \neq j$, are eigenvectors for the action of $\delta^{(m)}$, $0 \le m < q$. Denote by $\sigma : L \to L$ the action of $\delta^{(q)}$. Then there exists $s \in R$ such that all e_{ij}, $i \neq j$, are eigenvectors for the operator $\sigma - \operatorname{ad} s$.*

Proof. The construction of s consists of a sequence of steps that are adaptations of the computations found in the proof of [6, Theorem 3.3]. Before we begin, we point out that (2), applied to the algebra L, allows us to expand the action on iterated commutators. For example, with $\delta^{(q)}$ and a commutator of degree three, we have

$$\begin{aligned} \sigma([[x,y],z]) &= [[\sigma(x),y],z] + [[x,\sigma(y)],z] + [[x,y],\sigma(z)] \\ &\quad + \sum_{\substack{m_1,m_2,m_3=0 \\ m_1+m_2+m_3=q}}^{q-1} [[\delta^{(m_1)} \cdot x, \delta^{(m_2)} \cdot y], \delta^{(m_3)} \cdot z] \end{aligned} \tag{3}$$

for all $x, y, z \in L$. Also note that since $p \neq 2$ and the case $n = p = 3$ is excluded, we have $n \ge 5$.

Step 1. Fix $i \neq j$. We claim that

$$\sigma(e_{ij}) = \alpha(e_{ii} - e_{jj}) + \sum_{l \neq i} \alpha_{il} e_{il} + \sum_{k \neq i,j} \alpha_{kj} e_{kj} \tag{4}$$

for some α's in F (depending on i and j).

Without loss of generality, assume $i = 1$ and $j = 2$. Applying σ to the identical relation

$$[[[x, e_{12}], e_{12}], e_{12}] = 0 \qquad \forall x \in L$$

and using the analog of (3) for four factors to expand, we obtain

$$\begin{aligned} &[[[x,\sigma(e_{12})],e_{12}],e_{12}] + [[[x,e_{12}],\sigma(e_{12})],e_{12}] + [[[x,e_{12}],e_{12}],\sigma(e_{12})] \\ &+ \sum_{\substack{m_1,m_2,m_3,m_4=0 \\ m_1+m_2+m_3+m_4=q}}^{q-1} [[[\delta^{(m_1)} \cdot x, \delta^{(m_2)} \cdot e_{12}], \delta^{(m_3)} \cdot e_{12}], \delta^{(m_4)} \cdot e_{12}] = 0 \qquad \forall x \in L. \end{aligned}$$

Now each term in the summation vanishes due to the assumption that e_{12} is an eigenvector for all $\delta^{(m)}$ with $m < q$. Set $a = \sigma(e_{12})$. Then we obtain

$$[[[x, a], e_{12}], e_{12}] + [[[x, e_{12}], a], e_{12}] + [[[x, e_{12}], e_{12}], a] = 0 \qquad \forall x \in L. \tag{5}$$

Write $a = \sum_{kl} \alpha_{kl} E_{kl} + Z$ and let $A = (\alpha_{kl}) \in R$. Lifting (5) to R, we obtain:

$$[[[X, A], E_{12}], E_{12}] + [[[X, E_{12}], A], E_{12}] + [[[X, E_{12}], E_{12}], A] \in Z \qquad \forall X \in [R, R].$$

Rewriting the commutators in terms of products then yields

$$\begin{aligned} &3AE_{12}XE_{12} - 3E_{12}XE_{12}A - 3E_{12}XAE_{12} + 3E_{12}AXE_{12} \\ &+XE_{12}AE_{12} - E_{12}AE_{12}X \in Z \qquad \forall X \in [R, R]. \end{aligned} \tag{6}$$

Substituting $X = E_{21}$ into (6) and evaluating yields

$$3\sum_k \alpha_{k1}E_{k2} - 3\sum_l \alpha_{2l}E_{1l} - 3\alpha_{11}E_{12} + 3\alpha_{22}E_{12} + \alpha_{21}E_{22} - \alpha_{21}E_{11} \in Z,$$

which gives $3\sum_{k\neq 1,2} \alpha_{k1}E_{k2} - 3\sum_{l\neq 1,2} \alpha_{2l}E_{1l} + 4\alpha_{21}(E_{22} - E_{11}) \in Z$. This implies that $\alpha_{21} = 0$.

Now fix $u \neq 1, 2$ and $v \neq 1, 2$ such that $u \neq v$. Applying σ to the identical relation

$$[[[x, e_{12}], e_{12}], e_{uv}] = 0 \qquad \forall x \in L$$

and making the same computation as the one leading to (5), we obtain

$$[[[x, a], e_{12}], e_{uv}] + [[[x, e_{12}], a], e_{uv}] + [[[x, e_{12}], e_{12}], b] = 0 \qquad \forall x \in L, \tag{7}$$

where $a = \sigma(e_{12})$, as before, and $b = \sigma(e_{uv})$. Take $A, B \in R$ such that $a = A + Z$ and $b = B + Z$. Then lifting (7) to R and rewriting the commutators in terms of products yields

$$\begin{aligned} &XE_{12}AE_{uv} - E_{uv}XE_{12}A + AE_{12}XE_{uv} - E_{uv}AE_{12}X \\ &+E_{12}AXE_{uv} - E_{uv}XAE_{12} - 2E_{12}XAE_{uv} + 2E_{uv}AXE_{12} \\ &-2E_{12}XE_{12}B + 2BE_{12}XE_{12} \in Z \qquad \forall X \in [R, R]. \end{aligned} \tag{8}$$

Now pick $w \neq 1, 2, u, v$. Substituting $X = E_{w1}$ into (8) and evaluating yields

$$\alpha_{2u}E_{wv} + 2\alpha_{vw}E_{u2} \in Z,$$

so $\alpha_{2u} = 0$ and $\alpha_{vw} = 0$. Similarly, substituting $x = E_{2w}$ into (8) yields

$$-\alpha_{v1}E_{uw} - 2\alpha_{wu}E_{1v} \in Z,$$

so $\alpha_{v1} = 0$ as well.

We have so far established that $\alpha_{21} = 0$, $\alpha_{k1} = \alpha_{2k} = 0$ for all $k \neq 1, 2$, and $\alpha_{kl} = 0$ for all $k, l \neq 1, 2$, $k \neq l$. It remains to deal with the diagonal entries. Let $c = \sigma(e_{21})$. Applying σ to the equation

$$[[e_{12}, e_{21}], e_{12}] = 2e_{12},$$

we obtain

$$[[a, e_{21}], e_{12}] + [[e_{12}, c], e_{12}] + [[e_{12}, e_{21}], a] + \lambda e_{12} = 2a \tag{9}$$

for some $\lambda \in F$. Lifting (9) to R and rewriting the commutators in terms of products, we obtain

$$\begin{aligned} &2E_{11}A + 2AE_{22} - AE_{11} - E_{22}A - E_{12}AE_{21} - E_{21}AE_{12} \\ &+2E_{12}CE_{12} + \lambda E_{12} = 2A + \mu I \end{aligned} \tag{10}$$

for some $\mu \in F$. Looking at the coefficient of E_{kk} for $k \neq 1,2$, we see that $2\alpha_{kk} + \mu = 0$. Looking at the coefficient of E_{11}, we obtain $\alpha_{11} - \alpha_{22} = 2\alpha_{11} + \mu$. Thus $\alpha_{kk} = \frac{1}{2}(\alpha_{11} + \alpha_{22}) = -\frac{\mu}{2}$ for $k \neq 1,2$. It follows that

$$a = \alpha(e_{11} - e_{22}) + \sum_{l\neq 1} \alpha_{1l} e_{1l} + \sum_{k\neq 1,2} \alpha_{k2} e_{k2}, \tag{11}$$

where $\alpha = \frac{1}{2}(\alpha_{11} - \alpha_{22})$. The claim has been proved.

Step 2. We want to show that there exists $s' \in R$ such that $\sigma(e_{1j}) = (\operatorname{ad} s')e_{1j}$ for all $j \neq 1$.

As before, let $a = \sigma(e_{12})$. By Step 1, a is given by (11). Set $s_2 = -\alpha E_{21} - \sum_{l\neq 1} \alpha_{1l} E_{2l} + \sum_{k\neq 1,2} \alpha_{k2} E_{k1}$. Then $a = (\operatorname{ad} s_2)e_{12}$. Define $\sigma' = \sigma - \operatorname{ad} s_2$. Since $\operatorname{ad} s_2$ is a derivation of L, expansion rules like (3) still hold when σ is replaced with σ'. We also have $\sigma'(e_{12}) = 0$.

Now let $b = \sigma'(e_{13})$ and $c = \sigma'(e_{23})$. By Step 1 applied to σ', we have

$$\begin{aligned} b &= \beta(e_{11} - e_{33}) + \sum_{l\neq 1} \beta_{1l} e_{1l} + \sum_{k\neq 1,3} \beta_{k3} e_{k3}, \\ c &= \gamma(e_{22} - e_{33}) + \sum_{l\neq 2} \gamma_{2l} e_{2l} + \sum_{k\neq 2,3} \gamma_{k3} e_{k3}. \end{aligned} \tag{12}$$

Applying σ' to the equation $e_{13} = [e_{12}, e_{23}]$, we obtain $b = [e_{12}, c] + \lambda e_{13}$ for some $\lambda \in F$. Substituting the above expressions for b and c and evaluating, we obtain

$$\beta(e_{11} - e_{33}) + \sum_{l\neq 1} \beta_{1l} e_{1l} + \sum_{k\neq 1,3} \beta_{k3} e_{k3} = \gamma e_{12} + \sum_{l\neq 1,2} \gamma_{2l} e_{1l} + \gamma_{21}(e_{11} - e_{22}) + \lambda e_{13}.$$

Comparing the coefficients on both sides, we see that $\beta = 0$ and $\beta_{k3} = 0$ for all $k \neq 1,3$. Thus $b = \sum_{l\neq 1} \beta_{1l} e_{1l}$. Setting $s_3 = -\sum_{l\neq 1} \beta_{1l} E_{3l}$, we see that $b = (\operatorname{ad} s_3)e_{13}$ and also $(\operatorname{ad} s_3)e_{12} = 0$. Therefore, replacing σ' by $\sigma - \operatorname{ad} s_2 - \operatorname{ad} s_3$, we have $\sigma'(e_{12}) = \sigma'(e_{13}) = 0$.

Continuing this process, we obtain $\sigma' = \sigma - \operatorname{ad} s_2 - \cdots - \operatorname{ad} s_n$ such that $\sigma'(e_{1j}) = 0$ for all $j \neq 1$. Thus we can set $s' = \sum_{j\neq 1} s_j$.

Step 3. Let s' be as in Step 2 and $\sigma' = \sigma - \operatorname{ad} s'$. We claim that for all $i, j \neq 1$ with $i \neq j$, we have

$$\sigma'(e_{i1}) = \xi(i)(e_{ii} - e_{11}) + \sum_{l\neq i} \xi_{il} e_{il}, \tag{13}$$

$$\sigma'(e_{ij}) = \mu(i,j)e_{1j} + \gamma_{ij} e_{ij} \tag{14}$$

for some ξ's, γ's and μ's in F.

Without loss of generality, we assume $i = 2$ and $j = 3$. Let $c = \sigma'(e_{23})$, as before, and $x = \sigma'(e_{21})$. Then by Step 1 (applied to σ'), c has the form as in (12) and x has the form $x = \xi(e_{22} - e_{11}) + \sum_{l\neq 2} \xi_{2l} e_{2l} + \sum_{k\neq 1,2} \xi_{k1} e_{k1}$. Applying σ' to the equation $e_{23} = [e_{21}, e_{13}]$, we get $c = [x, e_{13}] + \lambda e_{23}$ for some $\lambda \in F$. Substituting the expressions for c and x and evaluating, we obtain

$$\begin{aligned} &\gamma(e_{22} - e_{33}) + \sum_{l\neq 2} \gamma_{2l} e_{2l} + \sum_{k\neq 2,3} \gamma_{k3} e_{k3} \\ &= -\xi e_{13} + \xi_{21} e_{23} + \sum_{k\neq 1,2,3} \xi_{k1} e_{k3} + \xi_{31}(e_{33} - e_{11}) + \lambda e_{23}. \end{aligned} \tag{15}$$

Comparing the coefficients on both sides, we see that $\gamma = 0$ and $\xi_{31} = 0$. A similar argument using the equation $e_{2k} = [e_{21}, e_{1k}]$ shows that $\xi_{k1} = 0$ for all $k \neq 1, 2$. Thus $x = \xi(e_{22} - e_{11}) + \sum_{l\neq 2} \xi_{2l} e_{2l}$ and (13) has been established for $i = 2$. Looking again at (15), we conclude that $\gamma_{k3} = 0$ for all $k \neq 1, 2, 3$ and also $\gamma_{2l} = 0$ for $l \neq 2, 3$. Hence $c = \gamma_{13}e_{13} + \gamma_{23}e_{23}$, which completes the proof of (14) for $i = 2$ and $j = 3$.

Step 4. We want to show that there exists $s'' \in R$ such that $(\operatorname{ad} s'')e_{1k} = 0$ for all $k \neq 1$ and $\sigma'(e_{ij}) = (\operatorname{ad} s'')e_{ij} + \gamma_{ij}e_{ij}$ for all $i, j \neq 1$, $i \neq j$.

First we show that in (14) the scalar $\mu(i, j)$ does not depend on j. Fix $i \neq 1$ and $u, v \neq 1, i$ with $u \neq v$. Applying σ' to the equation $e_{iv} = [e_{iu}, e_{uv}]$ and using (14), we obtain

$$\mu(i, v)e_{1v} + \gamma_{iv}e_{iv} = [\mu(i, u)e_{1u} + \gamma_{iu}e_{iu}, e_{uv}] + [e_{iu}, \mu(u, v)e_{1v} + \gamma_{uv}e_{uv}] + \lambda e_{iv}$$

for some $\lambda \in F$. Simplifying, we get $\mu(i, v)e_{1v} = \mu(i, u)e_{1v} + (\gamma_{iu} + \gamma_{uv} - \gamma_{iv} + \lambda)e_{iv}$, which implies that $\mu(i, u) = \mu(i, v)$. (By a similar argument, one can show that in fact $\mu(i, j)$ does not depend on i either, but we do not need this fact here.)

Now write $\mu(i, j) = \mu(i)$ and set $s'' = \sum_{k\neq 1} \mu(k)e_{1k}$. Then $(\operatorname{ad} s'')e_{ij} = \mu(i)e_{1j}$ and thus $\sigma'(e_{ij}) = (\operatorname{ad} s'')e_{ij} + \gamma_{ij}e_{ij}$. Clearly, $(\operatorname{ad} s'')e_{1k} = 0$.

Step 5. Set $s = s' + s''$ and $\widetilde{\sigma} = \sigma - \operatorname{ad} s$. We claim that each e_{ij} with $i \neq j$ is an eigenvector for $\widetilde{\sigma}$.

By Step 2 and Step 4, we already have $\widetilde{\sigma}(e_{1k}) = 0$ for all $k \neq 1$ and $\widetilde{\sigma}(e_{ij}) = \gamma_{ij}e_{ij}$ for $i, j \neq 1$, $i \neq j$. It remains to show that each e_{k1}, $k \neq 1$, is also an eigenvector of $\widetilde{\sigma}$. Fix $k \neq 1$ and $u \neq 1, k$. By (13) applied to $\widetilde{\sigma}$ and e_{k1}, we have

$$\widetilde{\sigma}(e_{k1}) = \widetilde{\xi}(k)(e_{kk} - e_{11}) + \sum_{l\neq k} \widetilde{\xi}_{kl}e_{kl}.$$

Now applying $\widetilde{\sigma}$ to the equation $[e_{k1}, e_{1u}] = e_{ku}$, we obtain

$$\left[\widetilde{\xi}(k)(e_{kk} - e_{11}) + \sum_{l\neq k} \widetilde{\xi}_{kl}e_{kl}, e_{1u}\right] + \lambda e_{ku} = \gamma_{ku}e_{ku}$$

for some $\lambda \in F$, which yields $-\widetilde{\xi}(k)e_{1u} = (\gamma_{ku} - \widetilde{\xi}_{k1} - \lambda)e_{ku}$. Hence $\widetilde{\xi}(k) = 0$ for all $k \neq 1$.

Finally, applying $\widetilde{\sigma}$ to the equation $e_{k1} = [e_{ku}, e_{u1}]$, we obtain

$$\sum_{l\neq k} \widetilde{\xi}_{kl}e_{kl} = \left[e_{ku}, \sum_{l\neq u} \widetilde{\xi}_{ul}e_{ul}\right] + \lambda e_{k1}$$

for some $\lambda \in F$, which yields $\sum_{l\neq k,u}(\widetilde{\xi}_{kl} - \widetilde{\xi}_{ul})e_{kl} + \widetilde{\xi}_{ku}e_{ku} = \widetilde{\xi}_{uk}(e_{kk} - e_{uu}) + \lambda e_{k1}$. Hence $\widetilde{\xi}_{ku} = 0$ for all $u \neq 1, k$ and $\widetilde{\sigma}(e_{k1}) = \widetilde{\xi}_{k1}e_{k1}$.

The proof of Lemma 3.1 is complete. □

4. Gradings by a p-group

Let $R = M_n(F)$ where F is a field of characteristic $p > 0$ (not necessarily algebraically closed). Let G be an abelian p-group. We want to describe all G-gradings on the simple Lie algebra $L = [R, R]/Z$ where $Z = [R, R] \cap Z(R)$. Replacing G with the subgroup generated by the support of the grading, we can assume that G is *finite*. As in the previous section, set $H = FG$ and $K = H^*$. Then a G-grading on R, resp. L, is equivalent to a K-module algebra structure on R, resp. L.

First suppose that G is an elementary abelian p-group of rank k. Then $K = u(\mathfrak{g})$ where $\mathfrak{g} = \langle \alpha_1, \ldots, \alpha_k \rangle$ is the abelian p-Lie algebra of dimension k corresponding to G, and any element $\delta \in \mathfrak{g}$ acts as a derivation of R, resp. L.

Theorem 4.1 ([6, Theorem 3.3]). *Let $R = M_n(F)$, $n \geq 2$, where* $\operatorname{char} F \neq 2$ *and, in the case $n = 3$, also* $\operatorname{char} F \neq 3$. *Let $Z = [R,R] \cap Z(R)$ and $L = [R,R]/Z$. If $d : L \to L$ is a Lie derivation, then there exists a derivation $D : R \to R$ such that $d(x + Z) = D(x) + Z$ for all $x \in [R,R]$.*

Suppose a G-grading $L = \bigoplus_{g\in G} L_g$ is given. Then L is a $u(\mathfrak{g})$-module algebra. Applying the above result to each of the Lie derivations $d_i : L \to L$ defined by $x \mapsto \alpha_i \cdot x$, we obtain derivations $D_i : R \to R$. As is well-known, all derivations of R are inner, so $D_i = \operatorname{ad} s_i$ for some $s_i \in R$. The proof of the following lemma is straightforward and thus omitted.

Lemma 4.2. *Let $R = M_n(F)$,* $\operatorname{char} F \neq 2$. *Let $s \in R$. If $(\operatorname{ad} s)x \in Z(R)$ for all $x \in [R,R]$, then $\operatorname{ad} s = 0$.* $\square$

Remark 4.3. Lemma 4.2 implies that D in Theorem 4.1 is uniquely determined by d.

The operators d_i, $i = 1, \ldots, k$, commute with each other, so we have $[D_i, D_j](x) \in Z(R)$ for all i, j and $x \in [R,R]$. Applying Lemma 4.2 to $s = [s_i, s_j]$, we see that in fact $[D_i, D_j] = 0$.

Lemma 4.4. *Let $R = M_n(F)$,* $\operatorname{char} F = p \neq 2$. *Let $s \in R$. If $(\operatorname{ad} s)^p x - (\operatorname{ad} s)x \in Z(R)$ for all $x \in [R,R]$, then $(\operatorname{ad} s)^p = \operatorname{ad} s$.*

Proof. Since $(\operatorname{ad} s)^p = \operatorname{ad} s^p$, we have $(\operatorname{ad}(s^p - s))x \in Z(R)$ for all $x \in [R,R]$. By Lemma 4.2 applied to $s^p - s$, we obtain $\operatorname{ad}(s^p - s) = 0$ and thus $(\operatorname{ad} s)^p = \operatorname{ad} s$. $\square$

Now $(d_i)^p = d_i$, $i = 1, \ldots, k$, so we have $(D_i)^p x - D_i x \in Z(R)$ for all $x \in [R,R]$. Applying Lemma 4.4 to s_i, we see that in fact $(D_i)^p = D_i$.

It now follows that R is a $u(\mathfrak{g})$-module algebra via $\alpha_i \circ r = D_i(r)$ for all $r \in R$, $i = 1, \ldots, k$. Therefore, there exists a (unique) G-grading $R = \bigoplus_{g\in G} R_g$ such that $L_g = (R_g \cap [R,R]) + Z$ for all $g \in G$. We wish to extend this result to the case where G is not necessarily elementary.

Theorem 4.5. *Let $R = M_n(F)$, $n \geq 2$, where* $\operatorname{char} F = p \neq 2$ *and, in the case $n = 3$, also $p \neq 3$. Let $Z = [R,R] \cap Z(R)$ and $L = [R,R]/Z$. Let G be a finite abelian p-group and $K = (FG)^*$. If $K \otimes L \to L$ is a K-module algebra structure on L sending $k \otimes a$ to $k \cdot a$, then there exists a unique K-module algebra structure $K \otimes R \to R$ on R sending $k \otimes r$ to $k \circ r$ such that*

$$k \cdot (x + Z) = (k \circ x) + Z \text{ for all } k \in K \text{ and } x \in [R,R].$$

Equivalently, if $L = \bigoplus_{g\in G} L_g$ is a G-grading, then there exists a unique G-grading $R = \bigoplus_{g\in G} R_g$ such that

$$L_g = (R_g \cap [R,R]) + Z \text{ for all } g \in G.$$

Proof. In the case when n is not divisible by p, i.e., $Z = 0$, the theorem has been proved in [1, Corollary 4.4]. Here we will consider the case when n is divisible by p. We will proceed by induction on $|G|$. We start by separating one cyclic factor: $G = \langle a \rangle_{p^N} \times \widetilde{G}$, hence $H = F\langle a \rangle \otimes \widetilde{H}$ and $K = (F\langle a \rangle)^* \otimes \widetilde{K}$. We introduce $\delta^{(m)}$

in the first factor as discussed in Section 3 for the case of a cyclic group. Let $\overline{K}$ be the subalgebra of K generated by $\widetilde{K}$ and $\delta^{(p^k)}$, $k=0,\ldots,N-2$. Then $\overline{K}=(F\overline{G})^*$ where $\overline{G}=G/\langle a^p\rangle$ is a group of smaller order.

By inductive hypothesis, R is a $\overline{K}$-module algebra via $k\otimes r\mapsto k\circ r$ for $k\in\overline{K}$ and $r\in R$. This $\overline{K}$-action on R induces the original $\overline{K}$-action on L and is uniquely determined by this property. Now consider the grading $R=\bigoplus_{\bar g\in\overline{G}}R_{\bar g}$ that corresponds to the $\overline{K}$-action. Since $\overline{G}$ is a p-group, this grading is elementary. Fix a basis of matrix units $\{E_{ij}\}$ in R such that

$$E_{ij}\in R_{\bar g_i^{-1}\bar g_j}\text{ where }\bar g_1,\ldots,\bar g_n\in\overline{G}.$$

The correspondence between gradings and actions gives $k\circ E_{ij}=\langle k,\bar g_i^{-1}\bar g_j\rangle E_{ij}$ for all $k\in\overline{K}$. It follows that, for $m=1,\ldots,p^{N-1}-1$, we have

$$\delta^{(m)}\circ E_{ii}=0\quad\text{ and }\quad\delta^{(m)}\circ E_{ij}=\lambda_{ij}^{(m)}E_{ij}\text{ for }i\neq j,\tag{16}$$

where $\lambda_{ij}^{(m)}=\langle\delta^{(m)},\bar g_i^{-1}\bar g_j\rangle\in F$.

Let $q=p^{N-1}$ and consider the operator $\sigma:L\to L$ defined by $\sigma(a)=\delta^{(q)}\cdot a$. Since L is a K-module algebra, (2) implies

$$\sigma([a,b])=[\sigma(a),b]+[a,\sigma(b)]+\sum_{k=1}^{q-1}[\delta^{(k)}\cdot a,\delta^{(q-k)}\cdot b]\qquad\forall a,b\in L.\tag{17}$$

Our first goal is to show that σ can be uniquely lifted to $\Sigma:R\to R$ such that

$$\Sigma(xy)=\Sigma(x)y+x\Sigma(y)+\sum_{k=1}^{q-1}(\delta^{(k)}\circ x)(\delta^{(q-k)}\circ y)\qquad\forall x,y\in R.\tag{18}$$

This will be done in three steps.

Step 1. In order to construct Σ, we will first "approximate" the operator σ with $\operatorname{ad}s$ for some appropriately chosen $s\in R$. Namely, we fix $s\in R$ as in Lemma 3.1 (clearly, such s is determined up to a diagonal matrix) and write for $\widetilde{\sigma}=\sigma-\operatorname{ad}s$:

$$\widetilde{\sigma}(e_{ij})=\gamma_{ij}e_{ij}\text{ for all }i\neq j.\tag{19}$$

Since $\operatorname{ad}s$ is a derivation of L, the operator $\widetilde{\sigma}$ also satisfies the expansion rule for commutators (17). We lift $\widetilde{\sigma}:L\to L$ to an operator $\widetilde{\Sigma}:R\to R$ by setting

$$\widetilde{\Sigma}(E_{ii})=0\text{ and }\widetilde{\Sigma}(E_{ij})=\gamma_{ij}E_{ij}\text{ for all }i\neq j.\tag{20}$$

Step 2. We show that $\widetilde{\Sigma}$ above satisfies the expansion rule for products (18).

We will need to know how $\widetilde{\sigma}$ acts on the elements $e_{ii}-e_{jj}$, $i\neq j$. Applying $\widetilde{\sigma}$ to the equation $e_{ii}-e_{jj}=[e_{ij},e_{ji}]$ and using (19), we obtain

$$\widetilde{\sigma}(e_{ii}-e_{jj})=[\gamma_{ij}e_{ij},e_{ji}]+[e_{ij},\gamma_{ji}e_{ji}]+\lambda[e_{ij},e_{ji}]$$

for some $\lambda\in F$. It follows that $e_{ii}-e_{jj}$ is also an eigenvector for $\widetilde{\sigma}$:

$$\widetilde{\sigma}(e_{ii}-e_{jj})=\beta_{ij}(e_{ii}-e_{jj}).\tag{21}$$

We claim that in fact $\beta_{ij}=0$ for all $i\neq j$. Indeed, fix $k\neq i,j$ and apply $\widetilde{\sigma}$ to the equation $e_{jk}=[e_{jk},e_{ii}-e_{jj}]$ using (19) and (21):

$$\gamma_{jk}e_{jk}=[\gamma_{jk}e_{jk},e_{ii}-e_{jj}]+[e_{jk},\beta_{ij}(e_{ii}-e_{jj})]+\sum_{l=1}^{q-1}[\delta^{(l)}\cdot e_{jk},\delta^{(q-l)}\cdot(e_{ii}-e_{jj})].$$

Now by (16) the summation on the right-hand side vanishes, so we have $\gamma_{jk}e_{jk} = (\gamma_{jk}+\beta_{ij})e_{jk}$ which gives $\beta_{ij}=0$, as desired. To summarize,

$$\widetilde{\sigma}(e_{ii}-e_{jj}) = 0 \text{ and } \widetilde{\sigma}(e_{ij}) = \gamma_{ij}e_{ij} \text{ for all } i \neq j. \tag{22}$$

We will need to know more about the relations among the scalars γ_{ij}. Fix $i \neq j$. Let $k \neq i, j$. Then applying $\widetilde{\sigma}$ to the equation $e_{ik} = [e_{ij}, e_{jk}]$ and using (22) and (16), we obtain

$$\gamma_{ik}e_{ik} = [\gamma_{ij}e_{ij}, e_{jk}] + [e_{ij}, \gamma_{jk}e_{jk}] + \sum_{l=1}^{q-1}[\lambda_{ij}^{(l)}e_{ij}, \lambda_{jk}^{(q-l)}e_{jk}],$$

which gives

$$\gamma_{ik} = \gamma_{ij} + \gamma_{jk} + \Lambda(i,j,k) \text{ for all distinct } i,j,k, \tag{23}$$

where

$$\Lambda(i,j,k) = \sum_{k=1}^{q-1}\lambda_{ij}^{(l)}\lambda_{jk}^{(q-l)}.$$

Note that the above definition of $\Lambda(i,j,k)$ makes sense when $i = k \neq j$. We also set $\Lambda(i,j,k) = 0$ if $i = j$ or $j = k$ — compare this convention with (16). Thus $\Lambda(i,j,k)$ is defined for all i,j,k (not necessarily distinct). Observe also that $\Lambda(i,j,k)$ depend only on the action of $\overline{K}$ (in other words, on the grading by $\overline{G}$).

Now applying $\widetilde{\sigma}$ to the equation $e_{ii} - e_{jj} = [e_{ij}, e_{ji}]$ and using (22) and (16), we obtain

$$0 = [\gamma_{ij}e_{ij}, e_{ji}] + [e_{ij}, \gamma_{ji}e_{ji}] + \sum_{l=1}^{q-1}[\lambda_{ij}^{(l)}e_{ij}, \lambda_{ji}^{(q-l)}e_{ji}],$$

which gives

$$\gamma_{ij} + \gamma_{ji} + \Lambda(i,j,i) = 0 \text{ for all } i \neq j. \tag{24}$$

We set for convenience $\gamma_{ii} = 0$ for all i. Then (23) and (24) can be combined into one formula:

$$\gamma_{ik} = \gamma_{ij} + \gamma_{jk} + \Lambda(i,j,k) \text{ for all } i,j,k. \tag{25}$$

We are now ready to prove that $\widetilde{\Sigma}$ defined by (20) satisfies (18). By linearity, it suffices to check (18) for $x = E_{ij}$ and $y = E_{lk}$. With our convention $\gamma_{ii} = 0$, (20) becomes

$$\widetilde{\Sigma}(E_{ij}) = \gamma_{ij}E_{ij} \text{ for all } i,j. \tag{26}$$

This observation along with (25) allow us to reduce the number of cases that need to be considered separately to just two.

Case 1: $j \neq l$. Then (18) is equivalent to the equation

$$0 = (\gamma_{ij}E_{ij})E_{lk} + E_{ij}(\gamma_{lk}E_{lk}) + \lambda E_{ij}E_{lk} \qquad (\lambda \in F),$$

which is obviously true.

Case 2: $j = l$. Then (18) is equivalent to the equation

$$\gamma_{ik}E_{ik} = (\gamma_{ij}E_{ij})E_{jk} + E_{ij}(\gamma_{jk}E_{jk}) + \Lambda(i,j,k)E_{ij}E_{jk},$$

which holds by (25).

Step 3. Regarding $\operatorname{ad} s$ as an operator $R \to R$ (rather than $L \to L$), we now set

$$\Sigma = \widetilde{\Sigma} + \operatorname{ad} s.$$

Then $\Sigma : R \to R$ is a lifting of $\sigma : L \to L$. Since $\widetilde{\Sigma}$ satisfies (18) and $\operatorname{ad} s$ is a derivation of R, we see that Σ also satisfies (18). To show the uniqueness of Σ, suppose there exists $\Sigma' : R \to R$ that is a lifting of σ and satisfies (18). Then $\Sigma' - \Sigma$ is a derivation of R and thus $\Sigma' - \Sigma = \operatorname{ad} r$ for some $r \in R$. We also know that $\operatorname{ad} r$ vanishes when regarded as an operator $L \to L$. By Lemma 4.2 it follows that $\operatorname{ad} r = 0$ and thus $\Sigma' = \Sigma$.

Now we can define the desired action $K \otimes R \to R : k \otimes r \mapsto k \circ r$ and thereby complete the proof of Theorem 4.5. We already have the action of $\overline{K} \otimes R \to R : k \otimes r \mapsto k \circ r$, so we only have to extend it to $k \in K$. Recalling the discussion in Section 3, we see that K is generated by $\overline{K}$ and $\delta^{(q)}$. Moreover, K is a free $\overline{K}$-module with basis $\{\left(\delta^{(q)}\right)^m \mid m = 0, \dots, p-1\}$, and also $\left(\delta^{(q)}\right)^p = \delta^{(q)}$. We set

$$\delta^{(q)} \circ r = \Sigma(r) \text{ for all } r \in R. \tag{27}$$

This will define a structure of K-module on R provided we verify that Σ commutes with the action of $\overline{K}$ and $\Sigma^p = \Sigma$.

Recall the grading $R = \bigoplus_{\bar{g} \in \overline{G}} R_{\bar{g}}$ that corresponds to the $\overline{K}$-action. Write $s = s' + s''$ where $s' \in R_{\bar{1}}$ and $s'' \in \bigoplus_{\bar{g} \neq \bar{1}} R_{\bar{g}}$. Since σ and $\widetilde{\sigma}$ preserve the induced grading $L = \bigoplus_{\bar{g} \in \overline{G}} L_{\bar{g}}$, so does $\operatorname{ad} s = \sigma - \widetilde{\sigma}$, hence $\operatorname{ad} s''$ acts trivially on L. By Lemma 4.2 it follows that $\operatorname{ad} s''$ acts trivially on R, hence we can replace s with s' and assume without loss of generality that $s \in R_{\bar{1}}$. Now since $\widetilde{\Sigma}$ and $\operatorname{ad} s$ preserve the grading $R = \bigoplus_{\bar{g} \in \overline{G}} R_{\bar{g}}$, so does $\Sigma = \widetilde{\Sigma} + \operatorname{ad} s$, which means that Σ commutes with the action of $\overline{K}$ on R, as desired.

In order to show that $\Sigma^p = \Sigma$, consider the identity component $R_{\bar{1}}$ in more detail. Recall from Section 2 that $R_{\bar{1}}$ consists of block diagonal matrices and is isomorphic to the direct product of full matrix algebras $S_1 \times S_2 \times \cdots \times S_l$. Observe that since $\overline{K}$ acts trivially on $R_{\bar{1}}$, the sum on the right-hand side of (18) vanishes for $x, y \in R_{\bar{1}}$, which implies that the restriction of Σ to $R_{\bar{1}}$, is a derivation of $R_{\bar{1}}$. It follows that $\Sigma \mid_{R_{\bar{1}}} = \operatorname{ad} r$ where $r = \operatorname{diag}(r_1, \dots, r_l)$ with $r_1 \in S_1, \dots, r_l \in S_l$. Now $\sigma^p = \sigma$ implies that $(\operatorname{ad} r)^p x - (\operatorname{ad} r) x \in Z(R)$ for all $x \in [R, R] \cap R_{\bar{1}}$. It follows that for each $i = 1, \dots, l$, we have $(\operatorname{ad} r_i)^p x - (\operatorname{ad} r_i) x \in Z(R) \cap S_i$ for all $x \in [S_i, S_i]$. By Lemma 4.4 we conclude that $(\operatorname{ad} r_i)^p = \operatorname{ad} r_i$ for all $i = 1, \dots, l$ and thus $\left(\Sigma \mid_{R_{\bar{1}}}\right)^p = \Sigma \mid_{R_{\bar{1}}}$. It remains to consider $\Sigma \mid_{R_{\bar{g}}}$ for $\bar{g} \neq \bar{1}$. Since $\sigma^p = \sigma$ and $R_{\bar{g}} \subset [R, R]$, we see that $(\Sigma^p - \Sigma)(x) \in Z(R) \subset R_{\bar{1}}$ for all $x \in R_{\bar{g}}$. On the other hand, we proved that Σ preserves $R_{\bar{g}}$. Therefore, $(\Sigma^p - \Sigma)(x) = 0$ for all $x \in R_{\bar{g}}$.

Thus we have defined the structure of a K-module on R: $k \otimes r \mapsto k \circ r$. Since K is generated by $\overline{K}$ and $\delta^{(q)}$, (18) and (27) imply that this is a structure of a K-module algebra. We also have $k \cdot (x + Z) = (k \circ x) + Z$ for all $x \in [R, R]$ when $k \in \overline{K}$ or $k = \delta^{(q)}$, which implies that $k \cdot (x + Z) = (k \circ x) + Z$ for all $x \in [R, R]$ and $k \in K$.

The proof of Theorem 4.5 is complete. □

We recall the notation $R^{(-)}$ for the Lie algebra structure given by commutator on an associative algebra R.

Corollary 4.6. *Let R and G be as in Theorem 4.5. Let $R = \bigoplus_{g \in G} R_g$ be a grading on the Lie algebra $R^{(-)}$. Then it is a grading on the associative algebra R if and only if the identity element of R is in the component R_1. Moreover, the latter condition is always satisfied in the case $p \mid n$.*

Proof. The case $p \nmid n$ is [1, Corollary 4.5]. Suppose $p \mid n$, so $Z = Z(R) \subset [R,R]$. Let $L = [R,R]/Z$. Then the G-grading on R induces a grading $L = \bigoplus_{g\in G} L_g$ where $L_g = (R_g \cap [R,R]) + Z$. Applying Theorem 4.5, we find a grading $R = \bigoplus_{g\in G} R'_g$ on the associative algebra R such that $L_g = (R'_g \cap [R,R]) + Z$. We claim that $R'_g = R_g$ for all $g \in G$. As in the proof of the theorem, we proceed by induction on $|G|$ and separate one cyclic factor: $G = \langle a\rangle_{p^N} \times \widetilde{G}$. Consider the dual actions of $K = (FG)^*$ on R corresponding to the two gradings. Let $\overline{K} = (F\overline{G})^* \subset K$ where $\overline{G} = G/\langle a^p\rangle$. By induction hypothesis, the two actions of the subalgebra $\overline{K}$ on R coincide. Let $\Sigma : R \to R$ and $\Sigma' : R \to R$ be the two actions of $\delta^{(p^{N-1})}$. Then $\Sigma - \Sigma'$ is a derivation of $R^{(-)}$. Hence $\Sigma - \Sigma' = \operatorname{ad} s + \zeta$ where $s \in R$ and $\zeta : R \to Z \subset R$ is a linear map such that $\zeta([R,R]) = 0$ [10, Theorem 2]. Since the grading $R = \bigoplus_{g\in G} R'_g$ is elementary, the identity component R'_1 contains the matrix unit E_{11} (relative to some basis). It follows that $\Sigma'(R) \subset [R,R]$. Also $\Sigma'(Z) = 0$, because $1 \in R'_1$. Hence the compositions $\zeta\Sigma'$ and $\Sigma'\zeta$ are both zero. By Lemma 4.2, $\operatorname{ad} s$ is also zero. Therefore, $\Sigma = \Sigma^p = (\Sigma' + \zeta)^p = (\Sigma')^p + \zeta^p = \Sigma'$, as desired. Finally, the identity matrix is in R'_1, and $R'_1 = R_1$. ☐

The following example shows that the restriction $n \neq 3$ if $p = 3$ in Theorem 4.5 cannot be omitted.

Example 4.7. Let $R = M_3(F)$, $\operatorname{char} F = 3$, and $L = [R,R]/Z(R)$. Then $e_{11} - e_{22} = e_{22} - e_{33}$, e_{12}, e_{13}, e_{23}, e_{21}, e_{31}, e_{32} form a basis of L. In [6, Example 2] it is shown that $d : L \to L$ defined by

$$\begin{array}{lll} & e_{11} - e_{22} \mapsto 0, & \\ e_{13} \mapsto 0, & e_{12} \mapsto e_{23}, & e_{23} \mapsto 0, \\ e_{31} \mapsto 0, & e_{32} \mapsto -e_{21}, & e_{21} \mapsto 0, \end{array}$$

is a derivation of L that cannot be lifted to a derivation of R. However, since $d^3 = 0$, d does not correspond to a grading on L. We consider another derivation

$$\bar{d} : L \to L : x \mapsto d(x) + (\operatorname{ad} E_{22})x.$$

Clearly, $\bar{d}$ cannot be lifted to R either, but now we have $\bar{d}^3 = \bar{d}$, because

$$\begin{array}{ll} \bar{d}\,|_{\operatorname{Span}\{e_{11}-e_{22},e_{13},e_{31}\}} = 0 & \text{and} \\ \bar{d}\,|_{\operatorname{Span}\{e_{23},e_{12}\}} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}, & \bar{d}\,|_{\operatorname{Span}\{e_{21},e_{32}\}} = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}. \end{array}$$

Hence $\bar{d}$ corresponds to a grading on L by the cyclic group $\langle a\rangle_3$ that is not induced by a grading on R. Namely,

$$\begin{array}{ll} L_1 = \operatorname{Span}\{e_{11} - e_{22}, e_{13}, e_{31}\} & \text{and} \\ L_a = \operatorname{Span}\{e_{21}, e_{23}\}, & L_{a^{-1}} = \operatorname{Span}\{e_{12} + e_{23},\, e_{32} - e_{21}\}. \end{array}$$

Remark 4.8. The above example is not completely satisfactory, because the C_3-grading on L, though not liftable to R, is conjugate to the grading

$$\begin{array}{ll} L'_1 = \operatorname{Span}\{e_{11} - e_{22}, e_{13}, e_{31}\} & \text{and} \\ L'_a = \operatorname{Span}\{e_{21}, e_{23}\}, & L'_{a^{-1}} = \operatorname{Span}\{e_{12}, e_{32}\} \end{array}$$

by the automorphism of L (not liftable to R) given in [6, Example 1]. The latter grading is obviously induced by the elementary grading on R that corresponds to the triple $(a, 1, a)$.

Remark 4.9. It was pointed out to us by A. Elduque that all gradings on $\mathfrak{psl}_3(F)$ in the case $\operatorname{char} F = 3$ can be obtained if one uses, instead of 3×3 matrices, the realization of $\mathfrak{psl}_3(F)$ as the algebra of zero-trace octonions. By [9, Theorem 9], all gradings on this algebra come from gradings on the algebra of octonions. The latter gradings are completely described in [9].

5. Gradings by an arbitrary abelian group

The gradings on the Lie algebra $L = \mathfrak{sl}_n(F)$ over an algebraically closed field F of characteristic zero have been described in [5]. Namely, every grading $L = \bigoplus_{g\in G} L_g$ by an abelian group G arises from a grading on $R = M_n(F)$ in one of the following two ways:

I : $L_g = R_g$ for $g \neq 1$ and $L_1 = R_1 \cap L$ where $R = \bigoplus_{g\in G} R_g$ is a G-grading on R;

II : $L_g = \mathcal{K}(R_g, *) \oplus \mathcal{H}(R_{gh}, *)$ if $g \neq h$ and $L_h = \mathcal{K}(R_h, *) \oplus (\mathcal{H}(R_1, *) \cap L)$ where $R = \bigoplus_{g\in G} R_g$ is a G-grading on R, $*$ is an involution that preserves the grading, and $h \in G$ is an element of order 2.

Here $\mathcal{H}(R, *)$ and $\mathcal{K}(R, *)$ stand, respectively, for the subspaces of symmetric and skew-symmetric elements relative to $*$.

As shown in [1], the same holds in the case $\operatorname{char} F = p > 0$ if $p \neq 2$ and $p \nmid n$. (The group G was assumed finite in [1], but this assumption is not necessary — see the proof of Theorem 5.1 below.) In the case when $p \neq 2$ divides n, one has to modify the above slightly: $L = \mathfrak{psl}_n(F)$, $Z = Z(R)$, and

I′ : $L_g = R_g + Z$ for $g \neq 1$ and $L_1 = (R_1 + Z) \cap L$ where $R = \bigoplus_{g\in G} R_g$ is a G-grading on R;

II′ : $L_g = (\mathcal{K}(R_g, *) + Z) \oplus (\mathcal{H}(R_{gh}, *) + Z)$ if $g \neq h$ and $L_h = (\mathcal{K}(R_h, *) + Z) \oplus ((\mathcal{H}(R_1, *) + Z) \cap L)$ where $R = \bigoplus_{g\in G} R_g$ is a G-grading on R, $*$ is an involution that preserves the grading, and $h \in G$ is an element of order 2.

We state the result in such a way that it includes both cases: $p \mid n$ and $p \nmid n$.

Theorem 5.1. *Let $R = M_n(F)$, $n \geq 2$, where F is an algebraically closed field, $\operatorname{char} F = p \neq 2$ and, in the case $n = 3$, also $p \neq 3$. Let $Z = [R, R] \cap Z(R)$ and $L = [R, R]/Z$. Let G be an abelian group. Then any G-grading on L is either of type I' or of type II' above.*

The proof is similar to the one given in [1] for the case $p \nmid n$. Before we start, we state a result that allows us to lift automorphisms from L to R (quoted in [6, Theorem 3.1]).

Theorem 5.2 ([7, Theorem 6.1])**.** *Let $S = M_m(E)$, $R = M_n(F)$, $n > 1$, E and F fields with isomorphism $\gamma : F \to E$. Assume that $\operatorname{char} E \neq 2$, and $m \neq 3$ if $\operatorname{char} E = 3$. Suppose there is a γ-semilinear Lie isomorphism $\alpha : \overline{[R, R]} \to \overline{[S, S]}$ where $\overline{[R, R]} = [R, R]/[R, R] \cap Z(R)$ and $\overline{[S, S]} = [S, S]/[S, S] \cap Z(S)$. Then $n = m$ and there exists a γ-semilinear map $\sigma : R \to S$ such that σ is either an isomorphism or the negative of an anti-isomorphism and such that $\overline{\sigma(x)} = \alpha(\overline{x})$ for all $x \in [R, R]$.*

In our case, $E = F$, $\gamma = \mathrm{id}$, and $R = S$. Define a homomorphism of algebraic groups $\theta : \operatorname{Aut}^\sim(R) \to \operatorname{Aut}(L)$ by $\sigma \mapsto \alpha$ where $\operatorname{Aut}^\sim(R)$ denotes the group consisting of the automorphisms and the negatives of the automorphisms of R (both are automorphisms of the Lie algebra $R^{(-)}$).

Lemma 5.3. *If $p \neq 2$ and $(p,n) \neq (3,3)$, then $\theta : \mathrm{Aut}^{\sim}(R) \to \mathrm{Aut}\,(L)$ is an isomorphism of algebraic groups.*

Proof. By Theorem 5.2, θ is surjective. To prove injectivity, we have to show that σ is uniquely determined by α. This is obvious in the case $p \nmid n$ and can be shown in the case $p \mid n$ as follows. We have to verify that, if an automorphism σ of R induces the identity map on L, then $\sigma = id$. Now $\sigma(x) = sxs^{-1}$ for some invertible $s \in R$. Hence we have $sxs^{-1} - x \in Z(R)$ for all $x \in [R,R]$. Write $R = A \otimes B$ where $A \cong M_k(F)$, $B \cong M_l(F)$, k is a power of p and $p \nmid l$. Since $1 \otimes b \in [R,R]$ for all $b \in B$, we see that σ restricts to the identity map on the subalgebra B. Since $a \otimes 1 \in [R,R]$ for all $a \in [A,A]$, and $[A,A]$ generates A as an associative algebra, we conclude that σ preserves the subalgebra A and induces the identity map on $[A,A]/Z(A)$. Thus we are reduced to the case when n is a power of p. Now, taking $x = I + \lambda E_{ij}$ with $i \neq j$ and $\lambda \in F$, we obtain $sxs^{-1} = x + \mu I$ for some $\mu \in F$. Hence $sxs^{-1}x^{-1} = (\mu+1)I - \lambda\mu E_{ij}$. Evaluating the determinant of both sides, we obtain $1 = (\mu+1)^n$, which implies $\mu = 0$. Since the elements $I + \lambda E_{ij}$ generate $\mathrm{SL}_n(F)$, we conclude that $\sigma = id$.

Finally, the homomorphism of the tangent algebras corresponding to θ is injective by Lemma 4.2. It follows that θ is an isomorphism of algebraic groups. $\square$

Proof of Theorem 5.1. Without loss of generality, G is finitely generated. Write $G = G_0 \times G_1$ where G_0 has no p-torsion and G_1 is a finite p-group. Recall from Section 3 that G-gradings on an algebra A are equivalent to pairs of mutually commuting actions on A where $\widehat{G_0}$ acts by automorphisms and the Hopf algebra $K_1 = (FG_1)^*$ in such a way that A is a K_1-module algebra. By Lemma 5.3, we can lift the map $f : \widehat{G_0} \to \mathrm{Aut}\,(L)$ associated to the action of $\widehat{G_0}$ on L and obtain a homomorphism of algebraic groups $\tilde{f} : \widehat{G_0} \to \mathrm{Aut}^{\sim}(R)$ by setting $\tilde{f} = \theta^{-1} \circ f$. By Theorem 4.5, we can lift the K_1-action on L (denoted $\cdot$) to an action on the associative algebra R (denoted $\circ$). The actions of $\widehat{G_0}$ and K_1 on R commute with each other. Indeed, fix $g \in G_0$. Then $k \otimes x \mapsto \tilde{f}(g)[k \circ (\tilde{f}(g^{-1})x)]$ determines a K_1-action on R, which induces the same action $\cdot$ on L. It follows from uniqueness in Theorem 4.5 that $\tilde{f}(g)(k \circ x) = k \circ (\tilde{f}(g)x)$ for all $k \in K_1$ and $x \in R$. Hence we obtain a G-grading on the Lie algebra $R^{(-)}$, $R = \bigoplus_{g \in G} R_g$, which induces the original G-grading on L.

Now set $\Lambda = \tilde{f}^{-1}(\mathrm{Aut}\,(R))$. This is a subgroup in $\widehat{G_0}$ of index at most 2 that acts by automorphisms on R. Set $H = \Lambda^{\perp}$ in G_0. Then $H = \langle h \rangle$ where $h \in G_0$ is of order at most 2. Let $\overline{G} = G/H$ and consider the corresponding $\overline{G}$-grading, which is a coarsening the G-grading on the Lie algebra $R^{(-)}$ constructed in the previous paragraph. By definition of H, the $\overline{G}$-grading is a grading on the associative algebra R. Note that the elements of Λ and of K_1 act trivially on the component $R_{\bar{1}}$; they act by scalar multiplication on any other component $R_{\bar{g}}$, $\bar{g} \in \overline{G}$.

If $\Lambda = \widehat{G_0}$, then we are done: we have a type I$'$ grading on L. Otherwise $\widehat{G_0}$ is generated over Λ by an element χ such that $f(\chi) = -\varphi$ where φ is an anti-automorphism of R. Since $f(\chi)$ preserves each component R_g, so does φ. Moreover, $\chi^2 \in \Lambda$ implies that φ^2 acts trivially on the identity component $R_{\bar{1}}$ of the $\overline{G}$-grading. Thus we can apply (for $\overline{G}$) the following result:

Proposition 5.4 ([5, Proposition 6.4])**.** *Let $R = M_n(F)$ be graded by an abelian group G. Let φ be an anti-automorphism of R that preserves the grading and acts* *as an involution on the component R_1. Then there exists an automorphism ψ of R that also preserves the grading such that φ commutes with ψ and $\varphi^2 = \psi^2$.*

Now we can define a new $\widehat{G_0}$-action on R by making χ act as ψ (instead of $-\varphi$). This defines a new G-grading $R = \bigoplus_{g\in G} \widetilde{R}_g$, which is a refinement of the $\overline{G}$-grading. By construction, the new G-grading is a grading on the associative algebra R. Moreover, $* = \psi^{-1}\varphi$ is an involution on R that preserves both gradings $R = \bigoplus_{g\in G} R_g$ and $R = \bigoplus_{g\in G} \widetilde{R}_g$. It remains to apply the following "Exchange Formula" in order to express R_g in terms of $\widetilde{R}_g$.

Lemma 5.5 ([1, Lemma 5.4])**.** *Let G be a group. Let R be a vector space with two compatible gradings $R = \bigoplus_{g\in G} R_g$ and $R = \bigoplus_{g\in G} \widetilde{R}_g$, i.e., $\widetilde{R}_g = \bigoplus_{x\in G}(R_x \cap \widetilde{R}_g)$, or, equivalently, $R_g = \bigoplus_{x\in G}(\widetilde{R}_x \cap R_g)$, for all $g \in G$. Suppose $H \triangleleft G$ is such that the two factor-gradings by G/H coincide. Set $R^h = \bigoplus_{g\in G}(\widetilde{R}_g \cap R_{gh})$. Then*

$$R_g = \bigoplus_{h\in H} (\widetilde{R}_{gh^{-1}} \cap R^h).$$

Moreover, if R is a (nonassociative) algebra equipped with two such gradings and $H \subset Z(G)$, then $R = \bigoplus_{h\in H} R^h$ is an algebra grading.

In our case, $R^1 = \bigoplus_{g\in G}(\widetilde{R}_g \cap R_g) = \bigoplus_{g\in G} \mathcal{K}(\widetilde{R}_g, *) = \mathcal{K}(R, *)$ and also $R^h = \bigoplus_{g\in G}(\widetilde{R}_g \cap R_{gh}) = \bigoplus_{g\in G} \mathcal{H}(\widetilde{R}_g, *) = \mathcal{H}(R, *)$. Therefore,

$$R_g = (\widetilde{R}_g \cap R^1) \oplus (\widetilde{R}_{gh} \cap R^h) = \mathcal{K}(\widetilde{R}_g, *) \oplus \mathcal{H}(\widetilde{R}_{gh}, *).$$

Hence the grading $R = \bigoplus_{g\in G} R_g$ induces a grading of type II$'$ on L. □

Corollary 5.6. *Let F be an algebraically closed field,* $\operatorname{char} F = p \neq 2$. *Let G be an abelian group. Let R and L be as in Theorem 5.1. If G has no 2-torsion, then any G-grading on L is of type I$'$. If the torsion subgroup of G is a p-group, then the grading on L is induced by an* elementary *G-grading on R.*

References

[1] Bahturin, Y.; Kochetov, M.; Montgomery, S. *Group gradings on simple Lie algebras in positive characteristic.* Proc. Amer. Math. Soc., **137** (2009), 1245–1254.

[2] Bahturin, Y.; Shestakov, I.; Zaicev, M. *Gradings on Simple Jordan and Lie Algebras.* J. Algebra, **283** (2005), 849–868.

[3] Bahturin, Y.; Zaicev, M. *Involutions on graded matrix algebras.* J. Algebra, **315** (2007), 527–540.

[4] Bahturin, Y.; Zaicev, M. *Graded algebras and graded identities.* Polynomial identities and combinatorial methods (Pantelleria, 2001), 101–139, Lecture Notes in Pure and Appl. Math., 235, Dekker, New York, 2003.

[5] Bahturin, Y.; Zaicev, M. *Gradings on Simple Lie Algebras of Type "A".* J. Lie Theory, **16** (2006), 719-742.

[6] Beidar, K. I.; Brešar, M.; Chebotar, M. A.; Martindale 3rd, W. S. *On Herstein's Lie map conjectures, III.* J. Algebra, **249** (2002), 59–94.

[7] Blau, P. S.; Martindale 3rd, W. S. *Lie isomorphisms in ∗-prime GPI rings with involution.* Taiwanese J. Math., **4** (2000), 215–252.

[8] Dieudonné, J. *Introduction to the theory of formal groups.* Pure and Applied Mathematics, 20. Marcel Dekker, Inc., New York, 1973.

[9] Elduque, A. *Gradings on octonions.* J. Algebra, **207** (1998), 342–354.

[10] Martindale, W. S., III. *Lie derivations of primitive rings.* Michigan Math. J. **11** (1964), 183–187.

[11] Montgomery, S. *Hopf algebras and their actions on rings*. CBMS Regional Conference Series in Mathematics, 82. American Mathematical Society, Providence, RI, 1993.

Department of Mathematics and Statistics, Memorial University of Newfoundland, St. John's, NL, A1C5S7, Canada
E-mail address: bahturin@mun.ca

Department of Mathematics and Statistics, Memorial University of Newfoundland, St. John's, NL, A1C5S7, Canada
E-mail address: mikhail@mun.ca