Diagonalizable Derivations of Finite-Dimensional Algebras I

DANIEL R. FARKAS
Department of Mathematics
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

CHRISTOF GEISS
Instituto de Matemáticas, UNAM
Ciuadad Universitaria, C.P. 04510
Mexico D. F., Mexico

EDWARD L. GREEN
Department of Mathematics
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

EDUARDO N. MARCOS
Departamento de Matemática, I.M.E.
Universidade de São Paulo, CP 66281
São Paulo, SP05389.970, Brasil

This study traces its ancestry to a venerable example. The polynomial algebra $k[X_1, \ldots, X_n]$ over the field k can be graded by total degree. The Euler derivation $E = \sum_{j=1}^{n} X_j \frac{\partial}{\partial X_j}$ has the remarkable property that $E(m) = \deg(m)m$ for every monomial m. More generally, suppose that G is a group and R is a G-graded k-algebra. Every additive character $\chi \in \mathrm{Hom}(G, k^+)$ can be assigned a derivation $D_\chi \in \mathrm{Der}(R)$ according to the rule

$$D_\chi(r) = \chi(g)r$$

for every $r \in R_g$, the homogeneous component of "degree" $g \in G$. Obviously, D_χ is diagonalizable. Conversely, if D is a diagonalizable derivation of the k-algebra S and H is the additive subgroup of k generated by the eigenvalues of D then, for the inclusion map $\iota : H \to k^+$, we have $D = D_\iota$.

In the case that R is a finite-dimensional k-elementary algebra, it can be presented as $k\Gamma/I$ where $k\Gamma$ is a path algebra. The graph Γ together with the ideal of relations I define a fundamental group; once certain parameters are specified, the fundamental group grades R. The diagonalizable derivations which arise make a mathematical appearance in several contexts, including computations of $H^1(R)$. These observations, which are discussed in [FGM], inspired a general project on diagonalizable derivations.

While trying to compute examples, it became clear that a better understanding of diagonalizable derivations was required. At first glance, it is not obvious that the set of diagonalizable derivations of a finite-dimensional algebra has any structure. However, it turns out that their span is a Lie ideal in the Lie algebra of all derivations over fields of characteristic zero and algebraically closed fields of positive characteristic. (It may be that there is no genuine restriction on the scalar field.) This result is Lie theoretic and occupies section 1 of the paper. It is fundamental to all that follows.

Section 2 is devoted to examples which are finite-dimensional algebras over a field of characteristic zero. Here are some highlights. A finite-dimensional commutative algebra can be identified with $k[X_1, \ldots, X_n]/I$ for some ideal I of finite codimension. When I is a monomial ideal the diagonalizable derivations span the Lie algebra of all derivations. The same result is true if I is a homogeneous ideal and $n = 1$ or $n = 2$ but it fails for $n \geq 3$. Diagonalizable derivations span in the parallel case of noncommutative monomial algebras, i.e., the free algebra or, more generally, a path algebra modulo word relations. The arguments we present depend on a rudimentary understanding of Lie algebras, in particular $\mathbf{sl}(2)$.

Our knowledge of derivations for algebras over fields of finite characteristic is spottier. In section 3, we show that the entire Lie algebra is spanned by diagonalizable ones for finite-dimensional images of the polynomial algebra in one variable and for commutative monomial algebras over an algebraically closed field different from 2. We establish the same result for finite-dimensional noncommutative monomial algebras only with the additional hypothesis that all long words (relative to the characteristic) are relations.

The last section handles a technical issue. There are some algebras for which there is more than one natural choice of "constants". For example, one might require that all appropriate derivations of path algebras vanish on the vertex idempotents. We show that the choice is irrelevant for understanding the span of diagonalizable derivations of images of path algebras and their relatives.

1 Lie ideals

For the remainder of the paper, k will always denote a field. Unless specified to the contrary, all algebras and vector spaces will have scalar field k. Suppose that $\mathcal{L}$ is a Lie algebra of endomorphisms of some (possibly infinite dimensional) vector space. We say that $x \in \mathcal{L}$ is *spanned-by-split* provided x is a sum of diagonalizable endomorphisms in $\mathcal{L}$. (One of the authors asserts that the correct pronunciation for this term is "splat".) The subspace of all spanned-by-split elements is denoted $\mathrm{SP}(\mathcal{L})$. We shall show that if $\mathcal{L}$ is a finite-dimensional Lie algebra and $\mathrm{char} k = 0$ or $\mathcal{L}$ is a finite-dimensional restricted Lie algebra and k is an algebraically closed field of positive characteristic then $\mathrm{SP}(\mathcal{L})$ is a Lie ideal of $\mathcal{L}$. Some restriction on k is turns out to be necessary for the second part.

Lemma 1.1 *Assume that* $\mathrm{char} k = 0$ *and that* $\mathcal{L}$ *is a Lie algebra of endomorphisms of a finite-dimensional space. If* $a, b \in \mathcal{L}$ *satisfy* $[a, b] = \mu b$ *for a nonzero scalar* μ *and* b *is nilpotent then*

$$\exp(\frac{1}{\mu}b)^{-1}\, a \exp(\frac{1}{\mu}b) = a + b \quad .$$

Proof: Since $\frac{1}{\mu}b$ and b commute,

$$[a, f(\frac{1}{\mu}b)] = f'(\frac{1}{\mu}b)b$$

for any polynomial $f \in k[X]$. The nilpotence of b implies that exp can be replaced with a polynomial when evaluating at b. Hence

$$[a, \exp(\frac{1}{\mu}b)] = \exp(\frac{1}{\mu}b)b$$

in the Lie algebra of all endomorphisms. The lemma now follows from a simple algebraic manipulation and the fact that $\exp(\frac{1}{\mu}b)$ is invertible. ■

Theorem 1.1 *Assume that* $\mathrm{char} k = 0$ *and that* $\mathcal{L}$ *is a Lie algebra of endomorphisms of a finite-dimensional space* V*. Then* $\mathrm{SP}(\mathcal{L})$ *is a Lie ideal of* $\mathcal{L}$*.*

Proof: It suffices to prove that if $a \in \mathcal{L}$ is diagonalizable then $[a, b] \in \mathrm{SP}(\mathcal{L})$ for all $b \in \mathcal{L}$. As indicated in 4.2 of [Hu], $\mathrm{ad}\, a$ is also diagonalizable. Write

$b = \sum_\mu b(\mu)$ where $b(\mu)$ is an eigenvector for $\mathrm{ad}\, a$ corresponding to eigenvalue μ. Then $[a, b] = \sum \mu b(\mu)$, where we may assume that $\mu \neq 0$.

We claim that $b(\mu)$ is nilpotent for $\mu \neq 0$. In the Lie algebra of all endomorphisms, $[a, b(\mu)^n] = n\mu b(\mu)^n$. Thus if $b(\mu)^n$ is nonzero for all positive integers n then $\mathrm{ad}_{\mathbf{gl}(V)}\, a$ has infinitely many eigenvalues, an impossibility. This establishes the claim. Alternatively, we may invoke the Jacobson Lemma ([He1]).

Apply the lemma to conclude that $a + b(\mu)$ is similar to a. In particular, $a + b(\mu)$ is diagonalizable. Hence

$$\mu b(\mu) = \mu((a + b(\mu)) - a) \in \mathrm{SP}(\mathcal{L}) \ .$$

Summing, we obtain $[a, b] \in \mathrm{SP}(\mathcal{L})$. ■

The theorem has a concrete interpretation which can be quite useful. Assume that $a \in \mathcal{L}$ is diagonalizable and let

$$V = \bigoplus \sum V_\lambda$$

be the corresponding eigenspace decomposition of the underlying space, with λ running over the eigenvalues. Let π_λ be the idempotent projection from V to V_λ. Every endomorphism b of V can now be written in block form with $\pi_\alpha b \pi_\beta$ in the (α, β)-block. If b happens to be an eigenvector of $\mathrm{ad}\, a$ for the eigenvalue μ then

$$b = \sum_{\alpha - \beta = \mu} \pi_\alpha b \pi_\beta \quad .$$

The nilpotence argument effectively says that, in characteristic zero, the eigenvalues of a can be ordered so that the b we have described can be written in strictly lower block triangular form. We can now avoid the "exp trick" if we wish. The assertion that $a + b$ is diagonalizable follows from a generalization of the well known exercise that a matrix with distinct eigenvalues is diagonalizable: *if a matrix is in lower block triangular form and the diagonal blocks have the form γI for distinct scalar values of γ then the matrix is diagonalizable.*

The technique of the theorem applies to Lie algebras over fields characteristic $p > 0$ when we are able to "avoid or break cycles". We can make this clear with a graph. Let V be a finite-dimensional vector space over the field k with $\mathrm{char} k = p > 0$. Assume that $\mathcal{L}$ is a Lie algebra of endomorphisms of V, that $a, b \in \mathcal{L}$, and that a is diagonalizable. We borrow the notation from the paragraphs above. Assume $\mu \neq 0$. The a-graph for $b(\mu)$ has as vertices the collection of nonzero $\pi_\sigma b\pi_\tau$ such that $\sigma - \tau = \mu$. There is an arrow from $\pi_\sigma b\pi_\tau$ to $\pi_\rho b\pi_\nu$ provided that $\tau = \rho$. (We emphasize that each $\pi_\sigma b\pi_\tau$ is an endomorphism of V which is an eigenvector for $\operatorname{ad} a$ but it need not lie in $\mathcal{L}$.) The a-graph for $b(\mu)$ is a disconnected union of line segments and cycles of length p.

The point is that if there are no cycles then $b(\mu)^p = 0$. In that case, the characteristic zero argument carries over with exp replaced by the truncated exponential,

$$\operatorname{texp}(x) = 1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{(p-1)!}x^{p-1} \ .$$

We shall see later that it is also possible to break a cycle. There may be some nonzero $\pi_\sigma b\pi_\tau$ which turns out to be in $\mathcal{L}$. If so, set $b(\mu)_1 = \pi_\sigma b\pi_\tau$ and $b(\mu)_2 = b(\mu) - b(\mu)_1$. Then $b(\mu)_i^p = 0$ and $[a, b(\mu)_i] = \mu b(\mu)_i$ for $i = 1$ and 2. Thus the truncated version of the first lemma holds. We conclude that both $b(\mu)_1$ and $b(\mu)_2$ lie in $\mathrm{SP}(\mathcal{L})$, whence $b(\mu)$ does as well.

Entirely different ideas are needed to handle a general Lie ideal theorem in positive characteristic. Here we borrow an identity from the elementary theory of restricted Lie algebras, see [J], 5.7 . If r and s lie in some associative algebra over a field of characteristic $p > 0$ then

$$(\operatorname{ad} s)^{p^t-1}(r) = \sum_{j=0}^{p^t-1} s^j r s^{p^t-1-j}$$

for all positive integers t.

Lemma 1.2 *Assume that k is an infinite field of characteristic p and that $\mathcal{L}$ is a Lie algebra of vector space endomorphisms which is closed under the p^{th}-power map. Fix a positive integer t. If $a, c \in \mathcal{L}$ are such that $(\lambda c + a)^{p^t} \in \mathrm{SP}(\mathcal{L})$ for all $\lambda \in k$ then*

$$(\operatorname{ad} a)^{p^t-1}(c) \in \mathrm{SP}(\mathcal{L}) \quad .$$

Proof: Write

$$(\lambda c + a)^{p^t} = \lambda^{p^t} c^{p^t} + a^{p^t} + \sum_{j=1}^{p^t-1} s_j(c,a)\lambda^j$$

in the algebra of endomorphisms. Specializing at $\lambda = 0$, we see that $a^{p^t} \in \mathrm{SP}(\mathcal{L})$. Since k is infinite, the formula holds for p^t values of λ and so a Vandermonde determinant argument allows us to solve for $s_j(c,a)$ in terms of $(\lambda c + a)^{p^t} - a^{p^t}$ over varying λ. Hence $s_j(c,a) \in \mathrm{SP}(\mathcal{L})$ for $j = 1, \ldots, p^t - 1$. Now $s_1(c,a) = \sum_{i+j=p^t-1} a^j c a^i$ and, by the identity, this sum equals $(\mathrm{ad}\, a)^{p^t-1}(c)$. ■

Theorem 1.2 *Assume that k is an algebraically closed field of positive characteristic p. Suppose that $\mathcal{L}$ is a Lie algebra of endomorphisms of a finite-dimensional space V which is closed under p^{th} powers. Then* $\mathrm{SP}(\mathcal{L})$ *is a Lie ideal of $\mathcal{L}$.*

Proof: Choose t large enough so that p^t exceeds the dimension of V. Since k is algebraically closed, we may use the Jordan Decomposition to write any endomorphism f of V as a sum of semisimple and nilpotent parts which commute. It follows from the choice of t that f^{p^t} is semisimple. In particular, if $f \in \mathcal{L}$ then $f^{p^t} \in \mathrm{SP}(\mathcal{L})$. Apply the lemma to $f = \lambda c + a$. For every $a, c \in \mathcal{L}$ we have

$$(\mathrm{ad}\, a)^{p^t-1}(c) \in \mathrm{SP}(\mathcal{L}) \quad .$$

Assume that $a \in \mathcal{L}$ is diagonalizable and $b \in \mathcal{L}$ is arbitrary. Write $b = \sum b(\mu)$, the decomposition of b as a sum of $\mathrm{ad}\, a$ eigenvectors. Then

$$\mu^{p^t-1} b(\mu) = (\mathrm{ad}\, a)^{p^t-1}(b(\mu)) \in \mathrm{SP}(\mathcal{L}) \ .$$

Hence if $\mu \neq 0$ we have $b(\mu) \in \mathrm{SP}(\mathcal{L})$. Since $[a,b] = \sum_{\mu \neq 0} \mu b(\mu)$ we see that $[a,b] \in \mathrm{SP}(\mathcal{L})$. ■

The extra requirement that k be algebraically closed when k has positive characteristic appears suspicious. It reflects a surprising phenomenon.

Proposition 1.1 *Assume that p is a prime and $k = \mathbf{GF}(p)$. Suppose that a and b are nonzero endomorphisms of a finite-dimensional vector space with the following properties:*

(i) $a^p = a$;

(ii) $b^p \neq 0$ *and* $b^{p^2} = 0$;

(iii) $[a,b] = b$.

Then a and b generate a three-dimensional Lie algebra $\mathcal{L}$ such that $\mathrm{SP}(\mathcal{L})$ *is not a Lie ideal of $\mathcal{L}$.*

Proof: First note that a is diagonalizable because it satisfies the polynomial $X^p - X$, which splits into distinct linear factors in $k[X]$. Second,

$$[a, b^p] = pb^p = 0$$

so b^p is central in $\mathcal{L}$. It follows that a, b, and b^p span a Lie algebra.

We claim that a, b, and b^p are linearly independent. Suppose $\alpha a + \beta b + \gamma b^p = 0$ for $\alpha, \beta, \gamma \in k$. Then

$$0 = [a, \alpha a + \beta b + \gamma b^p] = \beta[a, b] = \beta b \ .$$

Hence $\beta = 0$. Now $\alpha a + \gamma b^p = 0$ implies that the diagonalizable transformation αa equals the nilpotent transformation γb^p. Therefore $\alpha = \gamma = 0$.

To complete the argument that a, b, and b^p span a three-dimensional restricted Lie algebra we must show that the span is closed under p^{th} powers. This is a consequence of a close look at the formula

$$(a + \lambda b)^p = a^p + \lambda^p b^p + \sum_{i=1}^{p-1} s_i(a, b)\lambda^i$$

for $\lambda \in k$. The expression $s_i(a, b)$ is a linear combination of left-bracketed Lie words in a and b, each of which has length p and i appearances of b. (See [S], page 7 for a discussion of these formulas.) Since $[a, b]$ and b commute in our context, we see that $s_i(a, b) = 0$ for $i \geq 2$. As indicated in the previous theorem, $s_1(a, b) = (\operatorname{ad} a)^{p-1}(b) = b$. Thus

$$\begin{aligned}(a + \lambda b)^p &= a^p + \lambda^p b^p + \lambda b \\ &= a + \lambda b + \lambda b^p \ .\end{aligned}$$

Therefore

$$\begin{aligned}(\alpha a + \beta b + \gamma b^p)^p &= (\alpha a + \beta b)^p + \gamma^p b^{p^2} \\ &= \alpha a + \alpha^{p-1}\beta b + \beta b^p \ .\end{aligned}$$

We have verified that $\mathcal{L}$ is a restricted Lie algebra with basis $\{a, b, b^p\}$.

Next, we use our formula for p^{th} powers to determine all of the diagonalizable members of $\mathcal{L}$. Certainly $\beta b + \gamma b^p$ is always nilpotent. Thus any possible diagonalizable endomorphism is a scalar multiple of $c = a + \beta b + \gamma b^p$. But

$$c^p = (a + \beta b + \gamma b^p)^p = a + \beta b + \beta b^p = c + (\beta - \gamma)b^p \ .$$

Thus c satisfies the polynomial $(X^p - X)^p$. Assuming c is diagonalizable, it must satisfy $X^p - X$. That is, $\beta = \gamma$. Conversely, $a + \beta b + \beta b^p$ does satisfy $X^p - X$, making it diagonalizable.

We conclude that $\mathrm{SP}(\mathcal{L})$ is the two-dimensional subspace of $\mathcal{L}$ with basis a and $b+b^p$. Since $[a,b]=b$ and $b \notin \mathrm{SP}(\mathcal{L})$, we see that $\mathrm{SP}(\mathcal{L})$ is not a Lie ideal of $\mathcal{L}$. ■

We now demonstrate that Theorem 1.2 fails when $k = \mathbf{GF}(p)$. Let $R = k[X,Y]/(X^2, Y^{p+1})$. The derivations $Y\frac{\partial}{\partial Y}$ and $XY\frac{\partial}{\partial X}$ of $k[X,Y]$ induce derivations a and b, respectively, of R. An easy computation shows that the hypotheses of the proposition hold for a and b.

We will apply our two Lie ideal theorems to Lie algebras of derivations. Suppose that R is an associative algebra. Let $\mathrm{Der}(R)$ denote the Lie algebra of all k-linear derivations from R to itself. The results of this section tell us that if R is a finite-dimensional k-algebra ($\mathrm{char} k = 0$ or k is algebraically closed of positive characteristic) then $\mathrm{SPDer}(R)$ is a Lie ideal of $\mathrm{Der}(R)$. This statement will be referred to as the Lie Ideal Theorem. We do not have a counterexample to the assertion that $\mathrm{SPDer}(R)$ is a Lie ideal of $\mathrm{Der}(R)$ over all fields.

The Lie Ideal Theorem has an antecedent in work of Herstein. The next lemma is based on Lemma 1.10 of [He2]. There is no restriction on the field. We denote the Lie ideal of $\mathrm{Der}(R)$ consisting of inner derivations of an algebra R by $\mathrm{Inn}(R)$; an alternate notation is $\mathrm{ad}\, R$.

Lemma 1.3 *Let e be an idempotent in the algebra R. If $D \in \mathrm{Der}(R)$ then*

$$e + D(e) - eD(e) \quad \textit{and} \quad e - D(e) + D(e)e$$

are idempotents. Hence $D(e)$ is a difference of two idempotents. ■

Theorem 1.3 *Let R be a finite-dimensional associative algebra over an algebraically closed field. Then $\mathrm{SP}(\mathrm{Inn}(R))$ is a Lie ideal of $\mathrm{Der}(R)$.*

Proof: First observe that if $e \in R$ is an idempotent then the classical decomposition

$$R = eRe + eR(1-e) + (1-e)Re + (1-e)R(1-e)$$

is also an eigenspace decomposition for $\mathrm{ad}\, e$. Hence $\mathrm{ad}\, e$ is always diagonalizable.

Let $r \in R$. By decomposing the commutative subalgebra generated by r, we can write $r = s + n$ where $s = \sum \lambda_i e_i$ with $\lambda_i \in k$ and the e_i are commuting idempotents, n is nilpotent, and $[s,n]=0$. It follows that $\mathrm{ad}\, s$

is diagonalizable. Suppose that $\operatorname{ad} r$ is diagonalizable. Since $\operatorname{ad} r$ and $\operatorname{ad} s$ commute, $\operatorname{ad} r - \operatorname{ad} s$ is diagonalizable. Therefore $\operatorname{ad} n$ is both diagonalizable and nilpotent. We conclude that $\operatorname{ad} r = \operatorname{ad} s$.

Thus, in order to prove the theorem, we need only show that $[\operatorname{ad} e, D] \in \mathrm{SP}(\operatorname{ad} R)$ for e idempotent in R and $D \in \mathrm{Der}(R)$. But $[\operatorname{ad} e, D] = -\operatorname{ad} D(e)$. We are done by the lemma. ■

2 Derivations in characteristic zero

We wish to compute $\mathrm{SPDer}(R)$ for familiar classes and examples of finite-dimensional associative algebras R. In this section, k will always be a field of characteristic zero. We will further restrict our attention to well-behaved positively graded algebras. For the time being,

$$S = S_0 \oplus S_1 \oplus \cdots \oplus S_t$$

will denote a finite-dimensional graded algebra such that $S_0 = k$ and S_1 generates S as an algebra.

The Euler derivation E associated with the grading is defined by requiring that $E(s) = ms$ for all $s \in S_m$. Set $\mathcal{D} = \mathrm{Der}(S)$ and set $\mathcal{D}_\lambda$ to be the eigenspace for $\operatorname{ad} E$ corresponding to the eigenvalue λ. If $x \in S_m$ and $D \in \mathcal{D}_\lambda$ then

$$ED(x) = [E, D](x) + DE(x) = (\lambda + m)D(x) \quad .$$

It follows that λ is an integer and $\lambda \geq -1$. Similarly,

$$\mathcal{D} = \mathcal{D}_{-1} \oplus \mathcal{D}_0 \oplus \mathcal{D}_1 \oplus \cdots \oplus \mathcal{D}_{t-1}$$

with $\mathcal{D}_i(S_{j)} \subseteq S_{i+j}$ and $[\mathcal{D}_i, \mathcal{D}_j] \subseteq \mathcal{D}_{i+j}$.

Lemma 2.1 $\mathcal{D}_{-1} = 0$.

Proof: Let $D \in \mathcal{D}_{-1}$ and $0 \neq x \in S_1$. There exists a positive integer m with $x^m = 0$ and $x^{m-1} \neq 0$. Since $D(x) \in S_0$ and $S_0 = k$, we have

$$0 = D(x^m) = mx^{m-1}D(x) \quad .$$

It follows that $D(x) = 0$. In other words, D restricted to S_1 is zero. Since S_1 generates S, we have $D = 0$. ■

We now keep in mind that

$$\mathcal{D} = \mathcal{D}_0 \oplus \mathcal{D}_1 \oplus \cdots \oplus \mathcal{D}_{t-1}$$

Lemma 2.2 *Let $H \in \mathcal{D}$ be diagonalizable. If $H = A + D$ with $A \in \mathcal{D}_0$ and $D \in \sum_{j \geq 1} \mathcal{D}_j$ then A is diagonalizable.*

Proof: It suffices to show that the restriction of A to S_1 is diagonalizable. For any $s \in S$ we write

$$s = s_0 + s_1 + \cdots + s_t$$

according to the grading. Suppose $H(s) = \lambda s$. Looking at the component in degree one, we obtain $A(s_1) = \lambda s_1$.

Consider any $w \in S_1$ and expand $w = \sum_j \beta_j s(j)$ where the $s(j)$ are eigenvectors for H and the β_j are scalars. Projecting onto the degree one component,

$$w = \sum_j \beta_j s(j)_1 \quad .$$

That is, S_1 is spanned by eigenvectors for A. ∎

Regard $\mathcal{D}_0$ as a Lie algebra of endomorphisms of the finite-dimensional vector space S_1.

Theorem 2.1 SPDer(S)=SP($\mathcal{D}_0$) $\oplus \mathcal{D}_1 \oplus \cdots \oplus \mathcal{D}_{t-1}$.

Proof: Suppose that $D \in \mathcal{D}_j$ for some $j \neq 0$. By the Lie Ideal Theorem, $\frac{1}{j}[E, D] \in \mathrm{SPDer}(S)$. Hence $\mathcal{D}_j \subseteq \mathrm{SPDer}(S)$ for all nonzero j. As a consequence,

$$\mathrm{SPDer}(S) = (\mathrm{SPDer}(S) \cap \mathcal{D}_0) \oplus \mathcal{D}_1 \oplus \cdots \oplus \mathcal{D}_{t-1} \ .$$

We have $\mathrm{SP}(\mathcal{D}_0) \subseteq \mathrm{SPDer}(S) \cap \mathcal{D}_0$ because any derivation which is diagonalizable on S_1 is diagonalizable when extended to S. On the other hand, the lemma implies that if $B \in \mathcal{D}_0$ and $B = \sum_i (A_i + D_i)$ with $A_i + D_i$ diagonalizable in $\mathcal{D}$, $A_i \in \mathcal{D}_0$, and $D_i \in \sum_{j \geq 1} \mathcal{D}_j$ then

$$B = \sum_i A_i$$

and each A_i is diagonalizable. Thus $\mathrm{SPDer}(S) \cap \mathcal{D}_0 \subseteq \mathrm{SP}(\mathcal{D}_0)$. ∎

We apply the theorem to some examples. Notice that if the dimension of S_1 is 1 then $\mathcal{D}_0$ acts like scalars on S_1; all such endomorphisms are diagonalizable. Hence all derivations of S are spanned-by-split. In other words,

$$\mathrm{SPDer}(k[X]/(X^m)) = \mathrm{Der}(k[X]/(X^m))$$

whenever char$k = 0$. Inspired, we look at commutative graded algebras with $\dim(S_1) = 2$. Assume that k is algebraically closed and recall one of the standard notations for $\mathbf{sl}(2)$ (cf. [Hu]):

$$x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \; ; \; y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \; ; \; h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

Lemma 2.3 *Every element of* $\mathbf{sl}(2)$ *is conjugate to either* x *or a scalar multiple of* h*.*

Proof: If $a \in \mathbf{sl}(2)$ is nilpotent and not zero, it's Jordan Canonical Form is $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. If a is not nilpotent then it has a nonzero eigenvalue λ. Then $-\lambda$ is also an eigenvalue. Hence a is similar to $\begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}$. ■

Lemma 2.4 *Assume k is algebraically closed. Let V be a finite-dimensional irreducible* $\mathbf{sl}(2)$*-module. If W is a subspace of V and $x \cdot W \subseteq W$ then $h \cdot W \subseteq W$.*

Proof: We may assume that $\mathbf{sl}(2)$ acts nontrivially on V and that $W \neq 0$. Let $v_0, v_1, \ldots, v_m$ be the "push-up" basis of V as described in 2.7 of [Hu]. We claim that if j is the largest subscript such that v_j appears in the support of some vector in W then W has basis $v_0, v_1, \ldots, v_j$. It suffices to show that these basis vectors of V lie in W.

Say v_j appears in the support of w. Then $x^j \cdot w$ is a nonzero scalar multiple of v_0. Hence $v_0 \in W$. Next, $x^{j-1} \cdot w$ is a linear combination of v_0 and v_1 with the coefficient of v_1 nonzero. Hence $v_1 \in W$. Continuing in this manner, we verify the claim.

The lemma follows immediately because $h \cdot v_i$ is always a scalar multiple of v_i. ■

Theorem 2.2 *Assume k is algebraically closed and let S be a commutative graded algebra as described above. If the dimension of S_1 is 2 then* SPDer(S)=Der(S)*.*

Proof: Let R denote the polynomial algebra $k[X_1, X_2]$. It can be graded by total degree,

$$R = R_0 \oplus R_1 \oplus \cdots \quad .$$

We identify S with R/I for some homogeneous ideal I. Every derivation of S lifts to a derivation of R; indeed, the lifting respects the degree of derivations, as determined by the respective Euler derivations. According to the main theorem of this section, we need to show that $\mathcal{D}_0$ is spanned by diagonalizable endomorphisms. But we can now identify $\mathcal{D}_0$ with the stabilizer of I in $\mathbf{gl}(2)$. We are reduced to proving that the Lie algebra of matrices in $\mathbf{sl}(2)$ which stabilize I is spanned-by-split.

Clearly $\mathbf{sl}(2)$ stabilizes each homogeneous component R_j; as indicated in [Hu] 2.7, each R_j is actually an irreducible $\mathbf{sl}(2)$-module. Suppose that $a \in \mathbf{sl}(2)$. Then a stabilizes I if and only if it stabilizes each $I \cap R_j$. Assume I is stabilized by a.. By the next-to-last lemma, a is either similar to a scalar multiple of h or to x. In the first case, a is obviously diagonalizable. In the second case, we may assume that $a = x$ and invoke the previous lemma to conclude that h also stabilizes each $I \cap R_j$. The fact that x is spanned-by-split in the stabilizer can be seen explicitly:

$$x = (h + x) - h \quad \text{and} \quad h + x = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} . \quad \blacksquare$$

Let $\Re$ denote the ideal generated by $X_1, \ldots, X_n$ in the polynomial algebra $k[X_1, \ldots, X_n]$.

Theorem 2.3 *Assume that k is algebraically closed and let Q be a quadratic form in $k[X_1, \ldots, X_n]$. If $d \geq 3$ and $S = k[X_1, \ldots, X_n]/(Q, \Re^d)$ then* $\mathrm{SPDer}(S)=\mathrm{Der}(S)$.

Proof: By making a linear change of variables, we may assume that

$$Q = X_1^2 + \cdots + X_r^2$$

with $r \leq n$. In analogy with the previous theorem, we need only show that the stabilizer in $\mathbf{gl}(n)$ of the one-dimensional space $k \cdot Q$ is spanned by diagonalizable matrices. As a first approximation, it is easy to check that the stabilizer consists of all block matrices of the form

$$\begin{pmatrix} A & * \\ 0 & * \end{pmatrix}$$

where A runs over all matrices in $\mathbf{gl}(r)$ which stabilize Q. With a minimum of thought, we may now reduce to the case that $r = n$.

Write $A = (a_{ij})$ or, equivalently, $A = \sum_{i,j=1}^{r} a_{ij} X_i \frac{\partial}{\partial X_j}$. The restriction on A is that there exist a scalar λ such that

$$A(X_1^2 + \cdots + X_r^2) = \lambda(X_1^2 + \cdots + X_r^2) \ .$$

Calculating, this means $2a_{ii} = \lambda$ for $1 \leq i \leq r$ and $2a_{ij} + 2a_{ji} = 0$ for $1 \leq i < j \leq r$. In other words, $A = \lambda I_{r\times r} + K$ where $K^\tau = -K$. This space of matrices has as basis $I_{r\times r}$ and all differences of matrix units $e_{ij} - e_{ji}$ for $1 \leq i < j \leq r$. These are all diagonalizable over k because $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ has distinct eigenvalues in $\mathbf{Q}(\imath)$. $\blacksquare$

An easier argument along these lines shows that all derivations of $k[X_1, \ldots, X_n]/(g, \Re^d)$ are spanned-by-split for any field k of characteristic zero and for

$$g = X_1^s + X_2^s + \cdots + X_n^s$$

with $d > s \geq 3$. In this case, the stabilizer of $k \cdot g$ consists solely of scalar matrices. Of more interest is the example $Q = X_1^2 + X_2^2$ over the real numbers $\mathbf{R}$. The theorem still tells us that the stabilizing Lie algebra consists of all $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ but this matrix has real eigenvalues (i.e., is diagonalizable!) only when $b = 0$. So it is not true that all derivations are spanned-by-split when we drop the requirement that k be algebraically closed.

Given the evidence so far, one might hope that SPDer(S)=Der(S) for finite-dimensional commutative graded algebras S over an algebraically closed field of characteristic zero. Unfortunately, the stabilizer in $\mathbf{gl}(3)$ of the one-dimensional space spanned by

$$4(X_1^2X_3^2 - 4X_1X_2^2X_3 + X_2^4) + (2X_1^3X_3 - X_1^2X_2^2) + X_1^4$$

is the space of all matrices

$$\begin{pmatrix} a & b & 0 \\ 0 & a & b \\ 0 & 0 & a \end{pmatrix} .$$

(A tedious but routine computation is required.) This algebra obviously cannot be spanned by diagonalizables over any field.

It is possible to prove that SPDer(S)=Der(S) for commutative monomial algebras S. Since the argument is the same as the one we shall present for noncommutative monomial algebras, we change direction for the remainder of this section and study noncommutative examples. Let $k\langle X_1,\ldots,X_n\rangle$ denote the free algebra on the letters $X_1,\ldots,X_n$. We grade it by length of words. Its Lie algebra of degree zero derivations can then be identified with $\mathbf{gl}(n)$; in essence, the formal derivative $X_i\frac{\partial}{\partial X_j}$ makes sense as the unique derivation which extends the map on letters "replace X_j with X_i and send all other letters to 0".

Observe that if w is any word and $X_i\frac{\partial}{\partial X_j}(w)\neq 0$ then we can recover i and j from knowing both expressions: compare letter counts in w and in any word in the support of $X_i\frac{\partial}{\partial X_j}(w)$. The same observation holds for commutative words in the polynomial algebra.

Theorem 2.4 *Let $S=k\langle X_1,\ldots,X_n\rangle/I$ be a finite-dimensional monomial algebra, so that I is generated by words and $I\cap\mathrm{span}\{X_1,\ldots,X_n\}=0$. Then* SPDer($S$)=Der($S$). *(The same assertion is true with the polynomial algebra replacing the free algebra.)*

Proof: As we have done earlier, identify $\mathcal{D}_0$ with the space of degree zero derivations of the free algebra which stabilize I. If θ is such a derivation, and w is a fixed word in I then the observation implies that the supports of each $X_i\frac{\partial}{\partial X_j}(w)$, for those partial derivatives which appear in θ with nonzero coefficient, are pairwise disjoint. If a linear combination of words lies in I then the monomial property forces each word in the support to lie in I. We conclude that when $X_i\frac{\partial}{\partial X_j}$ appears in θ and $X_i\frac{\partial}{\partial X_j}(w)\neq 0$ then

$$X_i\frac{\partial}{\partial X_j}(w)\in I\ .$$

Of course, $X_i\frac{\partial}{\partial X_j}(w)\in I$ when the partial derivative vanishes on w as well. Since w is arbitrary in I, we conclude that this choice of $X_i\frac{\partial}{\partial X_j}$ also stabilizes I. In summary, the stabilizer of I has a basis which consists of some of the $X_i\frac{\partial}{\partial X_j}$.

Notice that $X_i\frac{\partial}{\partial X_i}$ always stabilizes words in I for $i=1,\ldots,n$. Thus the stabilizer in $\mathbf{gl}(n)$ of I has a basis consisting of some of the matrix units and includes all diagonal matrices. Let Δ denote a diagonal matrix with distinct scalars down the diagonal. Then $\Delta+e_{ij}$ is diagonalizable provided $i\neq j$. It follows that each off-diagonal matrix unit in the stabilizer is the difference of diagonalizable members of the stabilizer. Thus Der(S) is spanned by diagonalizables. ∎

We modify our discussion in this section so that it applies to images of path algebras. Recall that if Γ is a finite directed multigraph and k is any field then we may form an algebra $k\Gamma$ with a basis consisting of paths and multiplication given by concatenation. Here a vertex is regarded as a path of length zero and the product of two paths is zero if they do not concatenate. By definition, the vertices become idempotents in $k\Gamma$. This leads us to expand our class of graded algebras.

Consider graded algebras $S = S_0 \oplus \cdots \oplus S_t$ for which

$$S_0 = k \cdot e_1 \oplus \cdots \oplus k \cdot e_n$$

with orthogonal idempotents e_j and with S generated as an algebra by S_0 and S_1. Denote by $\mathcal{D}^*$ the Lie algebra of derivations of S which vanish on S_0. It remains true that the Euler derivation E lies in $\mathcal{D}^*$ and that

$$\mathcal{D}^* = \mathcal{D}^*_{-1} \oplus \mathcal{D}^*_0 \oplus \cdots \oplus \mathcal{D}^*_{t-1} \ .$$

We show that $\mathcal{D}^*_{-1} = 0$. Since $1 = e_1 + \cdots + e_n$ we have $S_1 = \sum_{i,j} e_i S_1 e_j$. Thus we need only show that $D(x) = 0$ for $D \in \mathcal{D}^*_{-1}$ and $0 \neq x \in e_i S_1 e_j$. Since D vanishes on the idempotents,

$$D(x) = D(e_i x e_j) = e_i D(x) e_j \ .$$

Now $D(x) \in S_0$ yields $D(x) = 0$ whenever $i \neq j$. For the case that $i = j$, notice that $D(x) \in k \cdot e_i$ and $x = e_i x e_i$; the elements $D(x)$ and x commute. The original argument that $\mathcal{D}_{-1} = 0$ can now be mimicked.

We will also need to know that if $D \in \mathcal{D}^*_0$ is diagonalizable when restricted to S_1 then it is globally diagonalizable. But this is true because the idempotents e_i are eigenvectors for D (lying, in fact, in the null space), so S is generated as an algebra by eigenvectors. Hence

$$\mathcal{D}^* = \mathcal{D}^*_0 \oplus \cdots \oplus \mathcal{D}^*_{t-1} \ .$$

We have established that Theorem 2.1 holds for graded algebras with $S_0 \simeq k^n$, provided we replace $\mathcal{D}$ with $\mathcal{D}^*$.

We wish to use the path algebra in the role of the free algebra, to generalize the previous theorem. Grade $k\Gamma$ by the length of paths. Denote the vertices by $e_1, \ldots, e_n$ and the span of the arrows by $\mathcal{A}$. If D is any degree zero derivation of $k\Gamma$ which vanishes on vertices then the calculation two paragraphs above gives

$$D(e_i \mathcal{A} e_j) \subseteq e_i \mathcal{A} e_j \ .$$

As an example, if a and b are arrows which both begin at the same vertex and both end at the same vertex (though the origin and terminus need not coincide) then the formal derivative $a\frac{\partial}{\partial b}$ is such a derivation. (This is the obvious map on arrows and extends uniquely to a derivation of the full algebra.) It is now immediate that the degree zero derivations of $k\Gamma$ which vanish on vertices can be identified with the span of all linear transformations $T : \mathcal{A} \to \mathcal{A}$ such that $T(e_i\mathcal{A}e_j) \subseteq e_i\mathcal{A}e_j$ for all i and j. It can be identified, equivalently, with the span of all $a\frac{\partial}{\partial b}$ where a and b arrows sharing origin and terminus.

We have provided all of the pieces to adapt the monomial theorem without further ado. The star in the statement below refers to the requirement that derivations vanish on S_0; we will be able to remove the star after Theorem 4.1 .

Theorem 2.5 *Let $S = k\Gamma/I$ be a finite-dimensional path-monomial algebra, so that I is an ideal generated by words of length at least 2. Then* SPDer$^*(S)$=Der$^*(S)$. ■

3 Derivations in positive characteristic

We begin with an anomaly. Assume that char$k = 2$ and consider the algebra $R = k[X]/(X^4)$. Any $D \in$Der(R) is determined by its value on X; conversely, any assignment

$$X \mapsto a_0 + a_1X + a_2X^2 + a_3X^3$$

defines a derivation because the obvious lift to a derivation of $k[X]$ vanishes on X^4. For which values of the a_i is the derivation diagonalizable? The matrix for the associated derivation D with respect to the basis $1, X, X^2, X^3$ is

$$\begin{pmatrix} 0 & a_0 & 0 & 0 \\ 0 & a_1 & 0 & 0 \\ 0 & a_2 & 0 & a_0 \\ 0 & a_3 & 0 & a_1 \end{pmatrix} .$$

Its characteristic polynomial is $T^2(T - a_1)^2$. If $a_1 = 0$ then D is diagonalizable only when $D = 0$. Suppose $a_1 \neq 0$. Then D is diagonalizable if and only if $D^2 - a_1D = 0$. This holds if and only if $a_3 = 0$. It follows that

$$\mathrm{SPDer}(R) = \{D \in \mathrm{Der}(R) \mid D(X) = a_0 + a_1X + a_2X^2\} .$$

Obviously, this is not all of $\mathrm{Der}(R)$. It turns out that we have simply encountered some characteristic 2 misbehavior.

To explain this comment, we look again at the polynomial algebra. We begin with no restrictions on k. If m is a monomial in $k[X_1, \ldots, X_n]$ then $\deg_i(m)$ is the number of appearances of X_i in m; we will view the degree as an integer even when k has positive characteristic. If $S = k[X_1, \ldots, X_n]/I$ for an ideal I then every derivation of S is the image of a derivation D of the polynomial algebra which stabilizes I. We will write $D \in \mathrm{Der}(S) \pmod{I}$. The next lemma generalizes the observation preceding the description of SPDer in the monomial case, Theorem 2.4. It is false for noncommutative monomial algebras.

Lemma 3.1 *Let $S = k[X_1, \ldots, X_n]/I$ with I a monomial ideal. Then*

$$\{m\frac{\partial}{\partial X_t} \in \mathrm{Der}(S) \pmod{I} \mid m \text{ is a monomial } \}$$

spans $\mathrm{Der}(S)$.

Proof: Suppose that a, b, and w are monomials and

$$a\frac{\partial}{\partial X_i}(w) = b\frac{\partial}{\partial X_j}(w) \ .$$

We claim that either $a = b$ and $i = j$ or that $a\frac{\partial}{\partial X_i}(w)$ is a scalar multiple of w. We may assume that $\frac{\partial}{\partial X_i}(w) \neq 0$. If $i = j$ then $a = b$. If $a\frac{\partial}{\partial X_i}(w)$ is not a multiple of w then $\deg_i a = 0$, so

$$\deg_i a\frac{\partial}{\partial X_i}(w) = (\deg_i w) - 1 \ .$$

But if $i \neq j$ then

$$\deg_i b\frac{\partial}{\partial X_j}(w) \geq \deg_i w \ .$$

The contradiction forces $a\frac{\partial}{\partial X_i}(w)$ to be a multiple of w as claimed.

Assume that $D \in \mathrm{Der}(S) \pmod{I}$ and expand

$$D = \sum_{i,j} \lambda_{ij} m_{ij} \frac{\partial}{\partial X_j} \ .$$

Here λ_{ij} is a nonzero scalar and m_{ij} is a monomial. Let w be a monomial relation in I. According to the claim, the $m_{ij}\frac{\partial}{\partial X_j}(w)$ which are not multiples of w are distinct monomials. They each lie in I because I is monomial. We conclude that $m_{ij}\frac{\partial}{\partial X_j}(w) \in I$ for each i,j in the sum defining D. That is, $m_{ij}\frac{\partial}{\partial X_j}$ stabilizes I. ■

Theorem 3.1 *Assume that k is an algebraically closed field of characteristic $p > 2$. If $S = k[X_1, \ldots, X_n]/I$ is finite-dimensional with I a monomial ideal then* SPDer(S)=Der(S).

Proof: We may assume that no indeterminate X_j lies in I. By virtue of the lemma, it suffices to prove for monomials m that $m\frac{\partial}{\partial X_t}$ is spanned-by-split whenever $m\frac{\partial}{\partial X_t} \in$Der$(S) \pmod I$.

It is easy to check that $X_j\frac{\partial}{\partial X_j}$ stabilizes I and is diagonalizable (monomials being eigenvectors) for $j = 1, \ldots, n$. It follows from the Lie Ideal Theorem, with apologies for the abuse of notation, that

$$[X_t\frac{\partial}{\partial X_t}, m\frac{\partial}{\partial X_t}] \in \mathrm{SPDer}(S) \pmod I$$

for all $m\frac{\partial}{\partial X_t} \in \mathrm{Der}(S) \pmod I$. Hence

$$(\deg_t m - 1)m\frac{\partial}{\partial X_t} \in \mathrm{SPDer}(S) \pmod I\ .$$

It follows that

$$m\frac{\partial}{\partial X_t} \in \mathrm{SPDer}(S) \pmod I$$

unless $\deg_t m \equiv 1 \pmod p$. So assume the latter situation. We sudivide into two cases: either $\deg_t m > 1$ or $\deg_t m = 1$.

Consider the first case and set $\widetilde{m} = \frac{m}{X_t}$. This derivation stabilizes I. Notice that $\frac{\partial}{\partial X_t}(\widetilde{m}) = 0$. Our computation, with X_t^2 playing the role of m, yields

$$X_t^2\frac{\partial}{\partial X_t} \in \mathrm{SPDer}(S) \pmod I\ .$$

Hence $[X_t^2\frac{\partial}{\partial X_t}, \widetilde{m}\frac{\partial}{\partial X_t}] \in \mathrm{SPDer}(S) \pmod I$. But

$$[X_t^2\frac{\partial}{\partial X_t}, \widetilde{m}\frac{\partial}{\partial X_t}] = -\widetilde{m}(2X_t)\frac{\partial}{\partial X_t} = -2m\frac{\partial}{\partial X_t}\ .$$

Since $p \neq 2$, we see that

$$m\frac{\partial}{\partial X_t} \in \mathrm{SPDer}(S) \pmod I\ .$$

We must still study the possibility that $\deg_t m = 1$. If $j \neq t$ then

$$[X_j \frac{\partial}{\partial X_j}, m\frac{\partial}{\partial X_t}] = (\deg_j m) m \frac{\partial}{\partial X_t} \ .$$

The Lie Ideal Theorem implies that $m\frac{\partial}{\partial X_t} \in \mathrm{SPDer}(S)$ whenever there is some j distinct from t with $\deg_j m$ not congruent to $0 \pmod p$. We are reduced to the assumptions that

$$\deg_t m = 1 \ \text{ and } \ \deg_j m \equiv 0 \pmod p \ \text{ for all } \ j \neq t \ .$$

If $\deg_j m = 0$ for all $j \neq t$ then $m\frac{\partial}{\partial X_t} = X_t \frac{\partial}{\partial X_t}$, which is clearly diagonalizable. Otherwise, there is some $i \neq t$ such that $\deg_i m \geq 2$. Echoing the earlier construction, set $\widetilde{m} = \frac{m}{X_i}$. Because X_t appears in $\widetilde{m}$, the derivation $\widetilde{m}\frac{\partial}{\partial X_t}$ stabilizes I. Now

$$[\widetilde{m}\frac{\partial}{\partial X_t}, X_i^2 \frac{\partial}{\partial X_i}] = m\frac{\partial}{\partial X_t} \ .$$

(Successful calculation requires that X_i appear in $\widetilde{m}$.) Since $X_i^2 \frac{\partial}{\partial X_i} \in \mathrm{SPDer}(S)$, we conclude, once again, that $m\frac{\partial}{\partial X_t} \in \mathrm{SPDer}(S)$. ■

Corollary 3.1 *Assume that k is an algebraically closed field with* $\mathrm{char} k \neq 2$. *If $f \in k[X]$ is a nonconstant polynomial then*

$$\mathrm{SPDer}(k[X]/(f)) = \mathrm{Der}(k[X]/(f)) \quad .$$

Proof: By factoring f and exploiting the compatibility of Der and SPDer with cartesian products, we reduce to the case that f is a power of a linear polynomial. By a change of variable, we may assume that $f = X^d$ for some positive integer d. This case has been handled in Theorems 2.4 and 3.1 . ■

The requirement that k be algebraically closed in the previous theorem stems from our use of the Lie Ideal Theorem. The much more modest result that $\mathrm{SPDer}(k[X]/(X^d))$ is a Lie ideal of $\mathrm{Der}(k[X]/(X^d))$ for all k is established below without the restriction. The argument is a good illustration of "cycle avoiding".

Proposition 3.1 *Assume* $\mathrm{char} k > 0$. *Then* $\mathrm{SPDer}(k[X]/(X^d))$ *is a Lie ideal of* $\mathrm{Der}(k[X]/(X^d))$.

Proof: Set $S = k[X]/(X^d)$ and set S_0 to be the span of all X^{pt} for integers t with $pt < d$. (In this argument, we will freely identify X^n with its image in S when $n < d$.) The bulk of the proof will be devoted to showing that if A is a nonzero diagonalizable derivation of S then

$$S_0 = \{g \in S \mid A(g) = 0\}$$

and the eigenvalues of A consist of integer multiples of a single scalar. Call the scalar λ. Let's see how to finish the proposition with this information. We must prove that $[A, D] \in \mathrm{SPDer}(S)$ for any $D \in \mathrm{Der}(S)$. It is enough to show that the A-graph for the eigenderivation $D(\mu)$ has no cycles. Now $\mu \neq 0$ and all eigenvalues of $\operatorname{ad} A$ are differences of eigenvalues of A; thus $\mu = j\lambda$ for some integer j not congruent to 0 mod p. It follows that there can be at most one cycle and it goes through $\pi_\mu D \pi_0$. But $\pi_0(S) = S_0$ and every derivation vanishes on S_0. In other words, $\pi_\mu D \pi_0 = 0$, so it is not really in the graph after all.

We proceed to the heart of the proof. Suppose A is a nonzero diagonalizable derivation. There must be some $s \in S$ which has X in its support and satisfies $A(s) = \lambda s$ for some scalar λ. We first argue that $\lambda \neq 0$. Otherwise, $A(s) = 0$. Since $A(1) = 0$, we could assume that X is the lowest term of s. Then X^h is the lowest term of s^h for $h = 1, 2, \ldots, d-1$ and $A(s^h) = 0$ for these values of h. It follows that $\{s^h \mid h = 0, 1, \ldots, d-1\}$ constitutes a basis for S, forcing $A = 0$. (Equivalently, we may argue that if s has lowest term X then s generates the radical of S.)

So far, we know that $A(s) = \lambda s$ with $\lambda \neq 0$. For $m = 1, 2, \ldots, p-1$ we have $A(s^m) = (m\lambda)s^m$ and X^m lies in the support of s^m. Fix one such m and consider $t \geq 0$ with $pt + m < d$. Then $A(X^{pt}s^m) = (m\lambda)(X^{pt}s^m)$ and X^{pt+m} is the smallest power of X in the support of $X^{pt}s^m$ whose exponent is congruent to $m \pmod p$. It follows that

$$\{X^{pt}s^m \mid pt + m < d\}$$

is a linearly independent set of eigenvectors for the eigenvalue $m\lambda$. We add up the number of independent eigenvectors for A we have explicitly produced. There is one X^{pt} in the eigenspace corresponding to zero for each non-negative multiple of p smaller than d. For each $1 \leq m \leq p-1$ we have listed one eigenvector corresponding to $m\lambda$ for each positive integer congruent to $m \pmod p$ which is less than d. The total number of linearly independent eigenvectors we have constructed is obviously d. Therefore we have an explicit basis of eigenvectors. In particular, the eigenspace corresponding to 0 is contained in S_0. Since the opposite containment is trivial, we see that S_0 is the entire eigenspace, as we had hoped. ■

We turn next to noncommutative monomial algebras. Although little is understood in this case, we are able to show that all derivations are spanned-by-split when long words do not survive. The next lemma prepares us for cycle breaking.

Lemma 3.2 *Assume that* $\mathrm{char} k = p > 0$. *Let* $S = S_0 \oplus S_1 \oplus \cdots \oplus S_{p-1}$ *be a graded* k*-algebra and have* π_m *denote the projection of* S *on* S_m. *Suppose that* S *is generated by* S_0 *and* S_1 *with* $S_1^p = 0$. *If* $0 < \mu \leq p-1$ *and* D *is any derivation of* S *then there exists an* n *such that* $\pi_{\mu+n} D \pi_n$ *is a derivation, the subscripts read modulo* p.

Proof: First consider $2 \leq \nu \leq p-1$. Then

$$D(S_\nu) = D(S_1^\nu) \subseteq \sum_{\alpha=0}^{\nu-1} (S_1^\alpha) S (S_1^{\nu-1-\alpha}) \subseteq \sum_{\beta=\nu-1}^{p-1} S_\beta \ .$$

It follows that $\pi_0 D \pi_\nu = 0$. This argument takes care of the case that $\mu \neq p-1$ by default.

We complete the proof by demonstrating that $\pi_{p-1} D \pi_0$ is a derivation. It suffices to check the product rule on uv with $u \in S_i$ and $v \in S_j$. If $i \neq 0$ then either $uv = 0$ or $uv \in S_{i+j}$ with $0 < i+j \leq p-1$. Thus

$$\pi_0(uv) = 0 \quad \text{and} \quad u(\pi_{p-1} D \pi_0)(v) + (\pi_{p-1} D \pi_0)(u) v \in S_i S_{p-1} = 0 \ .$$

Together with the mirror argument, we have

$$0 = (\pi_{p-1} D \pi_0)(uv) = u(\pi_{p-1} D \pi_0)(v) + (\pi_{p-1} D \pi_0)(u) v = 0$$

unless both u and v lie in S_0. However, if $u, v \in S_0$ then

$$\begin{aligned} (\pi_{p-1} D \pi_0)(uv) &= \pi_{p-1}\left(uD(v) + D(u)v\right) \\ &= u(\pi_{p-1} D)(v) + (\pi_{p-1} D)(u) v \\ &= u(\pi_{p-1} D \pi_0)(v) + (\pi_{p-1} D \pi_0)(u) v \ . \end{aligned}$$ ■

Theorem 3.2 *Assume that* $\mathrm{char} k = p > 0$. *Let* $S = k\langle X_1, \ldots, X_n \rangle / I$ *denote a finite-dimensional noncommutative monomial algebra. Assume, further, that* I *contains all words with* p *appearances of the same letter but no singleton letters. Then* $\mathrm{SPDer}(S) = \mathrm{Der}(S)$.

The same result holds for path-monomial algebras, without any added technicalities.

Proof: If w is a word let $\deg_t w$ denote the number of appearances of X_t in w. We use the degree to give the Lie algebra of derivations for the free algebra $R = k\langle X_1, \dots, X_n\rangle$ a group grading relative to $(\mathbf{Z}/p\mathbf{Z})^n$. (We will only study the grading as a vector space.) For a derivation $w\frac{\partial}{\partial X_t}$ such that w is a word, set $\delta(w\frac{\partial}{\partial X_t})$ to be the n-tuple with $\deg_t w - 1 \pmod p$ in the t^{th} coordinate and $\deg_t w \pmod p$ in the j^{th} coordinate for $j \neq t$. For $\mathbf{a} \in (\mathbf{Z}/p\mathbf{Z})^n$, let $\mathcal{D}_{\mathbf{a}}$ denote the vector space spanned by all such $w\frac{\partial}{\partial X_t}$ with $\delta(w\frac{\partial}{\partial X_t}) = \mathbf{a}$. The motivating identity is

$$[X_j\frac{\partial}{\partial X_j}, w\frac{\partial}{\partial X_t}] = a_j w\frac{\partial}{\partial X_t}$$

where a_j is the j^{th} coordinate of $\delta(w\frac{\partial}{\partial X_t})$.

Assume that I is an ideal of the free algebra which satisfies the hypotheses of the theorem. Let D be a derivation of the free algebra which stabilizes I. Certainly $X_j\frac{\partial}{\partial X_j}$ stabilizes I because the ideal is monomial. Let S_i be the span of the images of all words v such that $\deg_j v \equiv i \pmod p$. Then $S = S_0 \oplus S_1 \oplus \cdots \oplus S_{p-1}$, the algebra S is generated by S_0 and S_1, and the restriction on letter appearance implies that $S_1^p = 0$. The grading we have described is the grading on S induced by the image of the diagonalizable derivation $X_j\frac{\partial}{\partial X_j}$. Now the eigenspace decomposition of D relative to $X_j\frac{\partial}{\partial X_j}$ is $D = D_0 + \cdots + D_{p-1}$ where

$$D_m \in \sum_{a_j = m} \mathcal{D}_{\mathbf{a}} \ .$$

Here we have used the identity given above. Each D_m stabilizes I (e.g., via a Vandermonde determinant argument) so the same eigenspace decomposition holds modulo I. The lemma, in conjunction with cycle breaking, implies that

$$D - D_0 \in \mathrm{SPDer}(S) \pmod I \ .$$

First apply the previous paragraph to D with $j = 1$, i.e., with the diagonalizable derivation $X_1\frac{\partial}{\partial X_1}$. It says that the image of D in $\mathrm{Der}(S)$ lies in $\mathrm{SPDer}(S)$ if and only if the image of the projection of D in $\sum_{a_1=0}\mathcal{D}_{\mathbf{a}}$ lies in $\mathrm{SPDer}(S)$. So we may assume that

$$D \in \sum_{a_1 = 0} \mathcal{D}_{\mathbf{a}} \ .$$

Since $X_2 \frac{\partial}{\partial X_2}$ stabilizes $\sum_{a_1=0} \mathcal{D}_{\mathbf{a}}$, we may repeat our reduction with $j = 2$ and assume that

$$D \in \sum_{a_1=a_2=0} \mathcal{D}_{\mathbf{a}} \ .$$

Iterating, we need only prove that if D stabilizes I and $D \in \mathcal{D}_{\mathbf{0}}$ then $D \in \mathrm{SPDer}(S) \pmod{I}$.

Any derivation in $\mathcal{D}_{\mathbf{0}}$ is a linear combination of derivations $w\frac{\partial}{\partial X_t}$ for words w with

$$\deg_t w \equiv 1 \pmod{p} \ \text{ and } \ \deg_j w \equiv 0 \pmod{p}, \ j \neq t \ .$$

If D stabilizes I we may discard all of those w with p appearances of any particular letter. This means we can limit our attention to linear combinations of $w\frac{\partial}{\partial X_t}$ where

$$\deg_t w = 1 \ \text{ and } \ \deg_j w = 0 \ \text{ for } j \neq t \ .$$

In other words, we may assume that D is a linear combination of the $X_t\frac{\partial}{\partial X_t}$. Such a derivation is diagonalizable; its image lies in $\mathrm{SPDer}(S)$. ■

4 Change of constants and inner derivations

In dealing with path algebras, we considered derivations whose "constants" comprised a larger associative subalgebra than k. We begin this section by analyzing the discrepancy between such a Lie algebra and the entire algebra of derivations.

We will say that a finite-dimensional algebra R is *hypersplit* if $R/\mathrm{rad}(R)$ is a finite product of full matrix algebras over k. For example, basic algebras are hypersplit. If R is hypersplit and we denote $R/\mathrm{rad}(R)$ by S then S is a separable algebra. Hence there exists a copy of S in R with $R = S \oplus \mathrm{rad}(R)$. We will write $\mathrm{Der}_S(R)$ for the Lie algebra of derivations of R which vanish on S.

Lemma 4.1 *Assume that R is a finite-dimensional hypersplit algebra and S is a semisimple complement of the radical of R inside R. If V is the span of all idempotents in R and $C(S)$ is the centralizer of S in R then $R = V + C(S)$.*

Proof: We begin with a simple observation in any ring. If e and f are idempotents such that $fe = 0$ then for any a, the element $e + eaf$ is idempotent.

Write $S = S_1 \times \cdots \times S_q$ where each S_i is a full matrix algebra over k; the corresponding central idempotents are $\eta_1, \ldots, \eta_q$. If we look at R as an S-bimodule then it decomposes as a direct sum of simple $S_i \otimes S_j$-modules for $1 \leq i, j \leq q$. If $i \neq j$ then each element a in a simple $S_i \otimes S_j$ summand satisfies

$$a = \eta_i a \eta_j \ ;$$

the initial comment in our proof implies that $a \in V$. If $S_i \simeq \mathrm{Mat}_{d\times d}(k)$ then a simple $S_i \otimes S_i$ summand M looks like $\mathrm{Mat}_{d\times d}(k)$ with the usual bimodule action. Let e_{ij} denote the matrix unit in $\mathrm{Mat}_{d\times d}(k)$ with 1 in the (i, j) position and 0 in all other places. Assume that $i \neq j$ and let b be the element of M identified with e_{ij}. Then $e_{ii} \cdot b \cdot e_{jj} = b$. It follows as above that $b \in V$. Notice that when $d \geq 2$ and $i \neq j$ we have orthogonal idempotents $e_{ii} + e_{ji}$ and $-e_{ji} + e_{jj}$ in S. Let b be the element of M already identified. Then

$$(e_{ii} + e_{ji}) \cdot b \cdot (-e_{ji} + e_{jj}) \in V \ .$$

It follows that the member of M identified with $-e_{ii} + e_{jj}$ lies in V. In summary, all of those elements of M identified with $\mathbf{sl}(d)$ belong to V.

We have shown that $M = (M \cap V) + (k \cdot c)$ where c is the member of M identified with the identity matrix. Certainly $c \in C(S)$. Putting together the summands, we obtain the result. ■

Theorem 4.1 *Assume that R is a finite-dimensional hypersplit algebra and S is a semisimple complement of the radical of R inside R. Then*

$$\mathrm{Der}(R) = \mathrm{Der}_S(R) + \mathrm{SP}(\mathrm{Inn}(R)) \ .$$

Proof: Let $D \in \mathrm{Der}(R)$. The restriction of D to S maps into $H^1(S, R)$. But $H^1(S, R) = 0$ by separability. Hence there is an element $r \in R$ such that $D - \mathrm{ad}\, r \in \mathrm{Der}_S(R)$. Use the lemma to write $r = (\sum \lambda_i e_i) + c$ where the λ_i are scalars, the e_i are idempotents, and c centralizes S. As indicated in Theorem 1.3, each derivation $\mathrm{ad}\, e_i$ is diagonalizable and so $\mathrm{ad}\,(\sum \lambda_i e_i)$ belongs to $\mathrm{SP}(\mathrm{Inn}(R))$. Obviously, $\mathrm{ad}\, c \in \mathrm{Der}_S(R)$. ■

Corollary 4.1 *Assume that R is a finite-dimensional hypersplit algebra and S is a semisimple complement of the radical of R inside R. If all derivations in $\mathrm{Der}_S(R)$ are spanned-by-split then all derivations in $\mathrm{Der}(R)$ are spanned-by-split.* ■

REFERENCES

[FGM] D. R. Farkas, E. L. Green, and E. N. Marcos, Diagonalizable derivations of finite-dimensional algebras II, preprint.

[He1] I. N. Herstein, *Topics in Algebra*, John Wiley, New York, 1964.

[He2] I. N. Herstein, *Topics in Ring Theory*, Chicago Lect. in Math., U. of C. Press, Chicago, 1969.

[Hu] J.E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, GTM 9, Springer-Verlag, New York, 1980.

[J] N. Jacobson, *Lie Algebras*, Dover, New York, 1962.

[S] G. B. Seligman, *Modular Lie Algebras*, Erg. der Math. Vol 40, Springer-Verlag, Berlin, 1967.