

Inner derivations of alternative algebras over commutative rings

Ottmar Loos, Holger P. Petersson
Fakultät für Mathematik und Informatik
FernUniversität in Hagen
D-58084 Hagen, Germany
Email: ottmar.loos@uibk.ac.at
holger.petersson@FernUni-Hagen.de

Michel L. Racine
Department of Mathematics and Statistics
University of Ottawa
Ottawa, Ontario, Canada K1N 6N5
Email: mracine@science.uottawa.ca

Erhard Neher zum 60. Geburtstag gewidmet

Abstract

We define Lie multiplication derivations of an arbitrary non-associative algebra A over any commutative ring and, following an approach due to K. McCrimmon, describe them completely if A is alternative. Using this description, we propose a new definition of inner derivations for alternative algebras, among which Schafer's standard derivations and McCrimmon's associator derivations occupy a special place, the latter being particularly useful to resolve difficulties in characteristic 3. We also show that octonion algebras over any commutative ring have only associator derivations.

2000 *Mathematics Subject Classification.* Primary 17D05; Secondary 17A36, 17A45, 17B40

Keywords. Inner derivations; Alternative algebras; Derivation functors; Composition algebras; Automorphisms

Introduction

There are many important properties satisfied by inner but not in general by all derivations of Lie, associative or (linear) Jordan algebras. A particularly important one may be described as follows.

Let $f: A \rightarrow B$ be a homomorphism of non-associative algebras and D a derivation

of A . We say a derivation D' of B is *f-related* to D if

$$f(D(a)) = D'(f(a))$$

for all $a \in A$. In general, there will be no such D' . The situation is better for inner derivations, which satisfy the following

Mapping Principle. *Given a homomorphism $f: A \rightarrow B$ of algebras (Lie, associative or Jordan), every inner derivation D of A admits an inner derivation D' of B that is f -related to D .*

Indeed, an inner derivation of A can be naturally expressed in terms of left and right multiplication operators. This suggests and yields an inner derivation D' of B in a natural way which is f -related to D .

Properties of this kind tie up nicely with the fact that, under suitable regularity conditions, all derivations of A are inner.

This satisfactory state of affairs has led Schafer [19] (see also [20, II, §3, p. 21]) to propose a notion of inner derivations for arbitrary non-associative algebras over a field F that reduces to the usual one when dealing with Lie or unital associative (resp. Jordan) algebras [19]. Moreover, inner derivations in his sense always form an ideal in the full derivation algebra, so it follows also in this generality that all derivations are inner provided (i) non-zero inner derivations exist and (ii) the derivation algebra is simple (as a Lie algebra).

While (i) is a harmless condition rarely causing any difficulties, (ii) is a much more delicate one. Moreover, it points to a strong link between Lie theory and non-associative algebras in general that has dominated the scene for decades. For example, the interest in derivations of *alternative* algebras grew out of the fundamental observation, due to various authors, most notably É. Cartan [6], Jacobson [9, 11], Bannow [2] and Alberca-Elduque-Martín-Navarro [1], that the derivations of an octonion (= Cayley) algebra C over F form a central simple Lie algebra of type G_2 if and only if F has characteristic not 3; in particular, all derivations of C are inner in this case. Going one step further, the proof of [1, Prop. 1] may be combined with a base field extension argument to show that an octonion algebra over any field (possibly of characteristic 3) has only inner derivations.

In spite of these remarkable advances, a particularly annoying deficiency of Schafer's approach remains: again in the setting of alternative algebras, inner derivations in his sense fail to satisfy the Mapping Principle. Already implicit in Schafer's own work on the subject (cf. [20, p. 78]), this deficiency comes into full view through their characteristic-free description in McCrimmon's unpublished monograph on alternative algebras [15] that only quite recently has been made accessible to the mathematical public.

In view of the preceding circumstances, E. Neher has suggested to relinquish altogether the idea of a universal definition for inner derivations of arbitrary non-associative algebras. Instead, he argued, they should be defined, as in the old days, for each relevant class of non-associative algebras individually, always taking into account the special requirements of the theory at hand. In the present paper, Neher's suggestions will be implemented for the class of alternative algebras over an arbitrary commutative ring k . The basic concepts and results of the paper may be summarized as follows.

A slight modification of Schafer's original approach will lead us in Section 1 to what we call *Lie multiplication derivations*, which turn out to be the same as inner derivations in the sense of Schafer when dealing with unital algebras (Remark to Prop. 1.4) but not in general (Example 1.5). The Lie multiplication derivations always form an ideal in the full derivation algebra and specialize to inner derivations (in the usual sense) of associative and linear Jordan algebras even when these fail to have a unit. We then proceed to show that the Lie multiplication algebra of a non-associative k -algebra A commutes with flat base change if A is finitely spanned as a k -module (Cor. 1.10). The same conclusion holds for the algebra of multiplication derivations if A is also projective as a k -module and its automorphism group is smooth as a group scheme (Cor. 1.12).

In Section 2, we follow McCrimmon [15, A5.2] to describe the Lie multiplication derivations of an alternative k -algebra A (Thm. 2.3). It follows immediately from this description that they do not in general satisfy the Mapping Principle. For this reason, we define *inner derivations* of A by a condition that is more restrictive than the one of just being a Lie multiplication derivation and automatically ensures the validity of the Mapping Principle (2.5). Adapting McCrimmon's terminology (loc. cit.) to the present set-up, we also introduce a few subclasses of inner derivations that turn out to be useful later on. Among them, *associator derivations* (2.5(a)), having the form $\sum [L_{a_i}, R_{b_i}]$ where $a_i, b_i \in A$ satisfy $\sum [a_i, b_i] = 0$, and *standard derivations* (2.5(b)), which are sums of operators $D_{a,b} = [L_a, L_b] + [L_a, R_b] + [R_a, R_b]$ for $a, b \in A$, seem to be of particular importance. Standard derivations made their first appearance in the work of Schafer [19] and, historically, constitute the oldest class of derivations known for arbitrary alternative algebras. Associator derivations, on the other hand, which can be transformed quite easily into standard ones if $3A = A$ (Prop. 2.7(b)), are apparently best suited for dealing with difficulties in characteristic 3, for which alternative algebras are notorious. The aforementioned definitions give rise to various ideals in the full derivation algebra that all commute with flat base change provided the algebra itself is finitely spanned as a k -module (Prop. 2.9).

Let again $f: A \rightarrow B$ be a homomorphism of algebras and D an inner derivation of A . The derivation D' of B furnished by the Mapping Principle will in general not be uniquely determined by D , so we don't have a natural map from inner derivations of A to those of B : the inner derivations of A do not depend functorially on A . In many examples, functoriality can be achieved at the cost of replacing the inner derivation algebra by a suitable central extension. This problem is addressed in Section 3. We introduce the notion of *derivation functor* and show that the ideals of standard, associator and commutator derivations are all induced by suitable derivation functors. These derivation functors commute with flat base change, and the standard derivation functor even with arbitrary base change (Proposition 3.12), without finiteness assumptions on the underlying algebra.

In Section 4, we take up the study of octonion algebras over commutative rings. They will be introduced here in a "rational" manner, i.e., without the need of changing scalars, along the lines of [17]. We define a *splitting* of any octonion algebra C over k as an isomorphism from Z onto C , where $Z = \text{Zor}(k)$ stands for the split octonion algebra of ordinary Zorn vector matrices over k . We then proceed to show, using the

notion of splitting datum (4.7), that the functor assigning to each unital commutative associative k -algebra R the set of splittings of $C \otimes_k R$ over R is a smooth affine torsor in the étale topology whose structure group is the automorphism group scheme of Z (Thm. 4.10). As immediate consequences, we conclude that octonion algebras, just like Azumaya algebras, become split after a faithfully flat (even étale) extension (Cor. 4.11) and that their automorphism group schemes are smooth (Cor. 4.12). In particular, our definition of octonion algebras is equivalent to the one given by Thakur [22] over base rings containing $\frac{1}{2}$.

In the final section of the paper, the preceding results are applied to show that an octonion algebra C over an arbitrary commutative ring has only associator derivations (Thm. 5.1). This theorem is new even when the base ring is a field. After a reduction to the case where C is reduced, the proof consists in a careful analysis of the $\mathbb{Z}/3\mathbb{Z}$ -grading given on the derivation algebra (Example 5.4, Prop. 5.5) by an elementary idempotent of C (cf. 4.4).

Notations. Throughout we fix an arbitrary commutative ring k . Unadorned tensor products will always be taken over k . We write $\text{Spec}(k)$ for prime spectrum of k , i.e., for the totality of all prime ideals in k , equipped with the Zariski topology. The category of commutative associative k -algebras with 1 will be denoted by $k\text{-alg}$. For $R \in k\text{-alg}$, a k -module M and $x \in M$, we abbreviate $M_R = M \otimes R$ as R -modules and $x_R = x \otimes 1_R \in M_R$; we also write f_R for the R -linear extension of a k -linear map f between k -modules. The standard terminology of non-associative algebras (including notation) will be used as in Schafer [20], except that (linear) operators will always act on the left so, e.g., the equations $L_a b = ab = R_b a$ describe left and right multiplications in a non-associative algebra A , and the associator of elements $a, b, c \in A$ (resp. the commutator of a, b) will be indicated by $[a, b, c] = (ab)c - a(bc)$ (resp. $[a, b] = ab - ba$). The symbols \mathbb{N} and \mathbb{Z} denote the positive natural numbers and the rational integers, respectively.

1. Lie multiplication derivations

In this section, we fix an arbitrary non-associative algebra A over k . We do not assume that A has a unit.

1.1. The Lie multiplication algebra. The Lie algebra defined on the k -module $\text{End}_k(A)$ by the usual commutator of linear maps will be denoted by $\mathfrak{gl}(A)$. The subalgebra of $\mathfrak{gl}(A)$ generated by all left and right multiplication operators of arbitrary elements in A is called the *Lie multiplication algebra* of A , denoted by $\mathfrak{L}(A)$. For example, if A is associative, then $\mathfrak{L}(A) = L_A + R_A$. Or, if A is a (linear) Jordan algebra over a ring containing $\frac{1}{2}$, then $\mathfrak{L}(A) = L_A + [L_A, L_A]$.

1.2. Derivations. Recall that a *derivation* of A is a linear map $D: A \rightarrow A$ satisfying one (hence all) of the following equivalent relations, for all $x, y \in A$:

$$D(xy) = (Dx)y + x(Dy),$$

$$[D, L_x] = L_{Dx}, \tag{1}$$

$$[D, R_y] = R_{Dy}. \tag{2}$$

The derivations of A form a Lie algebra (more precisely, a subalgebra of $\mathfrak{gl}(A)$), denoted by $\text{Der}(A)$. The elements of $\text{Der}(A)$ also act on commutators and associators in a derivation-like manner, i.e., we have

$$D([x, y]) = [Dx, y] + [x, Dy], \tag{3}$$

$$D([x, y, z]) = [Dx, y, z] + [x, Dy, z] + [x, y, Dz] \tag{4}$$

for all $x, y, z \in A$. By (1), (2), $\text{Der}(A)$ acts on $\mathfrak{L}(A)$ through the adjoint representation of $\mathfrak{gl}(A)$, i.e.,

$$[\text{Der}(A), \mathfrak{L}(A)] \subseteq \mathfrak{L}(A). \tag{5}$$

1.3. The ideal of Lie multiplication derivations. We write $\hat{A} = k1 \oplus A$ for the algebra obtained by adjoining a unit $1 = 1_{\hat{A}}$ to A and \hat{L}, \hat{R} for the left, right multiplication, respectively, of \hat{A} . The relations $\hat{L}_{\alpha 1+a} = \alpha \text{Id}_{\hat{A}} + \hat{L}_a$, $\hat{R}_{\alpha 1+a} = \alpha \text{Id}_{\hat{A}} + \hat{R}_a$ ($\alpha \in k, a \in A$) show $\mathfrak{L}(\hat{A}) = k\text{Id}_{\hat{A}} + \hat{\mathfrak{L}}(A)$, where $\hat{\mathfrak{L}}(A)$ stands for the subalgebra of $\mathfrak{gl}(\hat{A})$ generated by $\hat{L}_A \cup \hat{R}_A$. Observe that there are no natural maps $\mathfrak{L}(A) \rightarrow \mathfrak{L}(\hat{A})$ satisfying $L_a \mapsto \hat{L}_a$ (resp. $R_a \mapsto \hat{R}_a$) unless $aA = \{0\}$ (resp. $Aa = \{0\}$) implies $a = 0$. But since $A \subseteq \hat{A}$ is an ideal, we obtain the inclusions

$$\mathfrak{L}(\hat{A}) \subseteq \mathfrak{g} := \{f \in \mathfrak{gl}(\hat{A}) \mid f(A) \subseteq A\}, \quad \hat{\mathfrak{L}}(A) \subseteq \mathfrak{g}' := \{f \in \mathfrak{gl}(\hat{A}) \mid f(\hat{A}) \subseteq A\} \subseteq \mathfrak{g} \tag{6}$$

as subalgebras of $\mathfrak{gl}(\hat{A})$, and the restriction homomorphism $\rho: \mathfrak{g} \rightarrow \mathfrak{gl}(A)$ satisfies $\rho(\hat{L}_a) = L_a$, $\rho(\hat{R}_a) = R_a$ for all $a \in A$, hence $\rho(\hat{\mathfrak{L}}(A)) = \mathfrak{L}(A)$.

There is a natural embedding $\text{Der}(A) \rightarrow \text{Der}(\hat{A})$, $D \mapsto \hat{D}$ of Lie algebras, where \hat{D} stands for the unique linear extension of $D \in \text{Der}(A)$ to \hat{A} given by $\hat{D}1 = 0$. It follows from (5) applied to \hat{A} in place of A that

$$\text{LMDer}(A) := \{D \in \text{Der}(A) \mid \hat{D} \in \mathfrak{L}(\hat{A})\} \subseteq \text{Der}(A) \tag{7}$$

is an ideal. The elements of $\text{LMDer}(A)$ are called *Lie multiplication derivations* of A . Indeed, as we will now see, they all belong to the Lie multiplication algebra of A and may thus be expressed as Lie polynomials in left and right multiplication operators by suitable elements of A .

1.4. Proposition. *The inclusion*

$$\text{LMDer}(A) \subseteq \mathfrak{L}(A) \cap \text{Der}(A)$$

always holds; it may be strengthened to the equality

$$\text{LMDer}(A) = \mathfrak{L}(A) \cap \text{Der}(A)$$

if A has a unit.

Proof. For the first part of the proposition, we must show $D \in \mathfrak{L}(A)$ for all $D \in \text{LMDer}(A)$. To this end, using 1.3, we decompose $\hat{D} \in \mathfrak{L}(\hat{A})$ as $\hat{D} = \alpha \text{Id}_{\hat{A}} + D'$ with $\alpha \in k$, $D' \in \hat{\mathfrak{L}}(A)$ and obtain $0 = \hat{D}1_{\hat{A}} = \alpha 1_{\hat{A}} + D'1_{\hat{A}}$, where the second summand on the right by (6) belongs to A . This implies $\alpha = 0$, $\hat{D} = D' \in \hat{\mathfrak{L}}(A)$, hence $D = \rho(\hat{D}) \in \rho(\hat{\mathfrak{L}}(A)) = \mathfrak{L}(A)$, as claimed.

For the second part, we assume A has a unit 1_A , put $e = 1_{\hat{A}} - 1_A$ and conclude $\hat{A} = ke \oplus A$ as a direct sum of ideals. This implies $\mathfrak{L}(\hat{A}) = k\text{Id}_{ke} \oplus \mathfrak{L}(A)$, the right-hand side being diagonally embedded into

$$\mathfrak{gl}(\hat{A}) = \begin{pmatrix} k \cdot \text{Id}_{ke} & \text{Hom}_k(A, ke) \\ \text{Hom}_k(ke, A) & \mathfrak{gl}(A) \end{pmatrix}. \quad (8)$$

On the other hand, given $D \in \text{Der}(A)$, we obtain $\hat{D}e = 0$ since D kills 1_A , and this amounts to $\text{Der}(A)^\wedge = \{\hat{D} \mid D \in \text{Der}(A)\} = \{0\} \oplus \text{Der}(A)$, the right-hand side again being embedded diagonally into (8). The assertion follows by comparing the decompositions for $\mathfrak{L}(\hat{A})$ and $\text{Der}(A)^\wedge$. \square

Remark. Comparing Prop. 1.4 with [20, p. 21], we conclude that the Lie multiplication derivations of A and its inner derivations in the sense of Schafer are the same if A has a unit. In general, however, this need not be so, as may be seen from the following example.

1.5. Example. Equality does not always hold in Prop. 1.4. To see this, suppose A is associative. We first claim

$$\text{LMDer}(A) = \text{InDer}_{\text{ass}}(A) = \{L_a - R_a \mid a \in A\},$$

i.e., that the Lie multiplication derivations of A and its inner derivations (in the usual sense) are the same. As inner derivations of A obviously belong to $\text{LMDer}(A)$, we need only worry about the converse, so let $D \in \text{LMDer}(A)$. Observing $\hat{D} \in \mathfrak{L}(\hat{A}) = k1_{\hat{A}} + \hat{L}_A + \hat{R}_A$ and $\hat{D}(1_{\hat{A}}) = 0$, we obtain $\hat{D} = \alpha 1_{\hat{A}} + \hat{L}_a + \hat{R}_b$ for some $\alpha \in k$, $a, b \in A$, hence $0 = \alpha 1_{\hat{A}} + (a + b)$, which yields $\alpha = 0$, $b = -a$, $D = L_a - R_a$, as claimed.

On the other hand, a derivation of A belonging to $\mathfrak{L}(A) = L_A + R_A$ need not be inner. To see this, suppose A is also commutative. Then there are no inner derivations other than zero, while $L_z \in L_A \subseteq \mathfrak{L}(A)$ for $z \in A$ is easily seen to be a derivation if and only if $AzA = \{0\}$, which in the absence of a unit element does not imply $L_z = 0$.

Remark. An analogous argument also works for a linear Jordan algebra J over k (with $\frac{1}{2} \in k$) since $D \in \text{LMDer}(J)$ implies $\hat{D} \in \mathfrak{L}(\hat{J}) = \hat{L}_J + [\hat{L}_J, \hat{L}_J] = k\text{Id}_J + \hat{L}_J + [\hat{L}_J, \hat{L}_J]$, $\hat{D}(1_{\hat{J}}) = 0$, hence $\hat{D} = \alpha \text{Id}_J + \hat{L}_a + \sum[\hat{L}_{a_i}, \hat{L}_{b_i}]$ for some $\alpha \in k$, $a, a_i, b_i \in J$, and

from $0 = \hat{D}(1_J) = \alpha 1_J + a$ we conclude $D = \sum [L_{a_i}, L_{b_i}]$. Thus the Lie multiplication derivations of J are just the inner ones in the usual sense:

$$\text{LMDer}(J) = \text{InDer}_{\text{Jord}}(J) = [L_J, L_J].$$

Not so, however, in the case of Lie algebras. The idea of defining Lie multiplication derivations by passing to the algebra \hat{A} seems to work well only when dealing with varieties of algebras that are stable under adjoining a unit.

Our principal objective in the present section will be to show that Lie multiplication derivations are well behaved under suitable scalar extensions. We begin by treating the analogous question for the Lie multiplication algebra $\mathfrak{L}(A)$.

1.6. Flat k -algebras. Let $R \in k\text{-alg}$ be a *flat* k -algebra, so R is flat as a k -module, equivalently, the assignment $M \mapsto M_R$ gives an exact functor from k -modules to R -modules. For a k -module M and a k -submodule $N \subseteq M$ with inclusion $i: N \rightarrow M$, we can and always will identify $N_R \subseteq M_R$ as an R -submodule via the injection $i_R: N_R \rightarrow M_R$.

The following easy lemma collects a few properties of flat k -algebras that are surely well known but seem to lack a convenient reference.

1.7. Lemma. *Conventions being as in 1.6, let $f: M \rightarrow M'$ be a k -linear map of k -modules and let $N, P \subseteq M, N' \subseteq M'$ be arbitrary k -submodules.*

- (a) $\text{Ker}(f)_R = \text{Ker}(f_R), \quad \text{Im}(f)_R = \text{Im}(f_R)$.
- (b) $f(N)_R = f_R(N_R), \quad f^{-1}(N')_R = f_R^{-1}(N'_R)$.
- (c) $(N \cap P)_R = N_R \cap P_R$.
- (d) *For every family $(N_\alpha)_{\alpha \in I}$ of k -submodules in M we have*

$$\left(\sum_{\alpha \in I} N_\alpha \right)_R = \sum_{\alpha \in I} (N_\alpha)_R \subseteq M_R.$$

- (e) *If N is generated as a k -module by a family $(x_\alpha)_{\alpha \in I}$ of elements in M , then $N_R \subseteq M_R$ is generated as an R -module by the family $(x_{\alpha R})_{\alpha \in I}$ of elements in M_R .*

Proof. By flatness, the functor $- \otimes R$ preserves kernels and co-kernels, which yields (a). The first (resp. second) part of (b) follows by applying (a) to $f|_N: N \rightarrow M'$ (resp. to $\pi \circ f: M \rightarrow M'/N', \pi: M' \rightarrow M'/N'$ being the canonical projection). In (c) we apply (b) with $N' = P$ to the natural embedding $i: N \rightarrow M$. For (d), we consider the canonical map $\bigoplus_{\alpha \in I} N_\alpha \rightarrow M$ determined by the inclusions $N_\alpha \rightarrow M$ and apply (a). Finally, in (e), we let M^0 be a free k -module with basis $(e_\alpha)_{\alpha \in I}$ and apply (a) to the k -linear map $M^0 \rightarrow M, e_\alpha \mapsto x_\alpha, \alpha \in I$. \square

1.8. Proposition. *Conventions being as in 1.6, let $R \in k\text{-alg}$ be a flat k -algebra and write B for the k -subalgebra of A generated by a family $(x_\alpha)_{\alpha \in I}$ of elements in A . Then B_R is the R -subalgebra of A_R generated by the family $(x_{\alpha R})_{\alpha \in I}$ of elements in A_R .*

Proof. We denote by $(y_\beta)_{\beta \in J}$ the family of non-associative monomials built in A over the family $(x_\alpha)_{\alpha \in I}$. Then B is generated as a k -module by $(y_\beta)_{\beta \in J}$. By Lemma 1.7(e), $B_R \subseteq A_R$ is therefore generated as an R -module by the family $(y_{\beta R})_{\beta \in J}$, which consists precisely of the non-associative monomials built in A_R over the family $(x_{\alpha R})_{\alpha \in I}$. Thus B_R is generated as an R -algebra by $(x_{\alpha R})_{\alpha \in I}$. \square

1.9. Finitely generated modules: base change of endomorphisms. Let M be a k -module and $S \in k\text{-alg}$ a unital commutative associative k -algebra. Then the natural map

$$\text{End}_k(M) \longrightarrow \text{End}_S(M_S), \quad f \longmapsto f_S,$$

extends to a homomorphism $\text{End}_k(M)_S \rightarrow \text{End}_S(M_S)$ of S -algebras, which is injective if S is a flat k -algebra and M is finitely generated [3, I, §2, Prop. 11]; we will then identify $\text{End}_k(M)_S \subseteq \text{End}_S(M_S)$ as an S -subalgebra accordingly. Under this identification, we even have equality $\text{End}_k(M)_S = \text{End}_S(M_S)$ if M is also projective [4, II, §5, Prop. 7]; in fact, equality then holds for *any* $S \in k\text{-alg}$.

1.10. Corollary. *If A is finitely generated as a k -module, its Lie multiplication algebra is stable under flat base change: For all flat k -algebras $S \in k\text{-alg}$, we have $\mathfrak{L}(A)_S = \mathfrak{L}(A_S)$ after the identifications of 1.9.*

Proof. $\mathfrak{L}(A)$ is generated by $L_A \cup R_A$ as a k -algebra. But $(L_a)_S = L_{a_S}$, $(R_a)_S = R_{a_S}$ for all $a \in A$. Hence, by Prop. 1.8, $\mathfrak{L}(A)_S$ and $\mathfrak{L}(A_S)$ are both generated as S -algebras by $L_{A_S} \cup R_{A_S}$. \square

Remark. In this generality, Cor. 1.10 is due to E. Neher (oral communication), who also pointed out that exactly the same argument yields exactly the same conclusion for the ordinary multiplication algebra in place of the Lie multiplication algebra $\mathfrak{L}(A)$.

1.11. Affine group schemes. Writing \mathbf{grp} for the category of groups, we let \mathbf{G} be an affine group scheme over k [7, II, §1, n° 1], so $\mathbf{G} : k\text{-alg} \rightarrow \mathbf{grp}$ is a functor represented by some commutative associative k -algebra with 1. We write $\text{Lie}(\mathbf{G})$ for its Lie algebra [7, II, §4, 4.8] and recall from loc. cit. that, if \mathbf{G} is smooth [7, I, §4, n° 4], $\text{Lie}(\mathbf{G})$ commutes with base change, so $\text{Lie}(\mathbf{G})_R \cong \text{Lie}(\mathbf{G}_R)$ canonically, for all $R \in k\text{-alg}$.

In this paper, we will be interested in the following special case. Assume A is finitely generated projective as a k -module and consider its automorphism group scheme by defining

$$\mathbf{Aut}(A) : k\text{-alg} \longrightarrow \mathbf{grp}, \quad R \longmapsto \mathbf{Aut}(A)(R) := \text{Aut}(A_R).$$

Then its Lie algebra is $\text{Der}(A)$ [7, II, §4, 2.3], so assuming that $\mathbf{Aut}(A)$ is smooth forces $\text{Der}(A)$ to commute with base change: $\text{Der}(A)_R = \text{Der}(A_R)$ for all $R \in k\text{-alg}$.

1.12. Corollary. *If A is finitely generated and projective as a k -module and $\mathbf{Aut}(A)$ is smooth as an affine group scheme, then $\mathbf{LMDer}(A)$ commutes with flat base change: $\mathbf{LMDer}(A)_R = \mathbf{LMDer}(A_R)$ for all flat k -algebras $R \in k\text{-alg}$.*

Proof. This follows immediately from (7), Lemma 1.7, Cor. 1.10 and 1.11. \square

1.13. Nucleus and centre. We close this section by reminding the reader of the *nucleus* of A , which is defined by

$$\mathbf{Nuc}(A) := \{x \in A \mid [x, A, A] = [A, x, A] = [A, A, x] = \{0\}\}. \quad (9)$$

It is an associative subalgebra of A and even a unital one if A contains an identity element. By (4), the nucleus is stable under derivations, i.e.,

$$\mathbf{Der}(A) \mathbf{Nuc}(A) \subseteq \mathbf{Nuc}(A).$$

Recall also that the *centre* of A , denoted by $\mathbf{Cent}(A)$, consists of those elements x in the nucleus satisfying $[A, x] = 0$. It is a commutative associative subalgebra of A but may collapse to zero unless A is unital and not zero.

1.14. Proposition. *If A is finitely generated as a k -module, then its nucleus and its centre both commute with flat base change: $\mathbf{Nuc}(A)_R = \mathbf{Nuc}(A_R)$, $\mathbf{Cent}(A)_R = \mathbf{Cent}(A_R)$ for all flat k -algebras $R \in k\text{-alg}$.*

Proof. Assume that the elements a_1, \dots, a_m span A as a k -module and, for $1 \leq i, j \leq m$, consider the linear maps $L_{ij}, M_{ij}, R_{ij}, C_i: A \rightarrow A$ defined respectively by $x \mapsto [a_i, a_j, x], [a_i, x, a_j], [x, a_i, a_j], [a_i, x]$. Intersecting the kernels of the L_{ij}, M_{ij}, R_{ij} gives the nucleus of A , whose intersection with the kernels of the C_i in turn gives the centre of A . Hence the assertion follows from Lemma 1.7(a),(c). \square

2. Alternative algebras: inner derivations

We now specialize A to a (possibly non-unital) alternative algebra over k .

2.1. Some useful identities. A is alternative if and only if the associator $[x, y, z] = (xy)z - x(yz)$ is an alternating function of its arguments. Hence an element $x \in A$ belongs to the nucleus if and only if one of the relations defining the nucleus (1.9) is fulfilled, and we have the left and right alternative laws

$$x(xy) = x^2y, \quad (1)$$

$$(yx)x = yx^2 \quad (2)$$

as well as flexibility

$$x(yx) = (xy)x =: xyx, \quad (3)$$

for all $x, y \in A$. We also recall the left, middle and right Moufang identities

$$x(y(xz)) = (xyx)z, \quad (xy)(zx) = x(yz)x, \quad ((zx)y)x = z(xyx) \quad (4)$$

for all $x, y, z \in A$.

We now derive a number of identities that will play an important role in the explicit description of Lie multiplication derivations. The following relations hold for all $a, b, c, x, y \in A$:

$$[L_a, L_b] = L_{[a,b]} - 2[L_a, R_b], \quad (5)$$

$$[R_a, R_b] = -R_{[a,b]} - 2[L_a, R_b], \quad (6)$$

$$[[L_a, R_b], L_c] = L_{[a,b,c]} - [L_{[a,b]}, R_c], \quad (7)$$

$$[[L_a, R_b], R_c] = R_{[a,b,c]} - [L_c, R_{[a,b]}], \quad (8)$$

$$L_a(xy) = ((L_a + R_a)x)y - x(L_a y), \quad (9)$$

$$R_a(xy) = -(R_a x)y + x((L_a + R_a)y), \quad (10)$$

$$(L_a - R_a)(xy) = ((L_a - R_a)x)y + x((L_a - R_a)y) + [x, 3a, y], \quad (11)$$

$$[L_a, R_b]x = [a, b, x] = (ab)x - a(bx) = b(ax) - (ba)x = x(ba) - (xb)a, \quad (12)$$

$$[L_a, R_b](xy) = ([L_a, R_b]x)y + x([L_a, R_b]y) + [x, [a, b], y]. \quad (13)$$

Proof. Identities (5), (6) may be found in Schafer [20, (3.68), (3.67)]. While his proof is carried out over fields, it works equally well over the commutative ring k . Ignoring (7), (8) for the moment, (9) (resp. (10)) follows immediately by linearizing (1) (resp. (2)). Subtracting (10) from (9) yields (11). To establish (12), one simply observes $[L_a, R_b]x = -[a, x, b] = [a, b, x] = -[b, a, x] = -[x, b, a]$ by alternativity. (13) is slightly more troublesome. By (12), the left-hand side may be written as

$$[L_a, R_b](xy) = (ab)(xy) - a(b(xy)). \quad (14)$$

Linearizing (1), the first term on the right becomes

$$(ab)(xy) = ((ab)x + x(ab))y - x((ab)y). \quad (15)$$

To the second term, we apply the linearized left Moufang identity and obtain

$$a(b(xy)) = ((ab)x)y + ((xb)a)y - x(b(ay)). \quad (16)$$

Subtracting (16) from (15) and observing (14), (12) implies

$$\begin{aligned} [L_a, R_b](xy) &= (x(ab) - (xb)a)y - x((ab)y - b(ay)) \\ &= (x(ba) - (xb)a)y + (x[a, b])y + x(b(ay) - (ba)y) - x([a, b]y) \\ &= ([L_a, R_b]x)y + x([L_a, R_b]y) + [x, [a, b], y], \end{aligned}$$

which is (13). To prove (7), we set $x = c$ in (13), view the result as a linear map in y and observe $[c, [a, b], y] = -[[a, b], c, y] = -[L_{[a,b]}, R_c]y$ by (12). Finally, (8) follows by reading (7) in the opposite algebra of A .

2.2. Proposition (McCrimmon [15, A5, 2.15]). *The Lie multiplication algebra of A is*

$$\mathfrak{L}(A) = L_A + R_A + [L_A, R_A].$$

Proof. By (5)–(8) above, it suffices to show $[[L_A, R_A], [L_A, R_A]] \subseteq L_A + R_A + [L_A, R_A]$, which follows from the Jacobi identity by applying (7) and (8) twice. \square

Remark. Prop. 2.2 is due to Schafer [19, Thm. 5] if k is a field of characteristic not 2.

2.3. Theorem (cf. McCrimmon [15, A5, 2.16]). *D is a Lie multiplication derivation of A if and only if it has the form*

$$D = L_a - R_a + \sum_{i=1}^m [L_{a_i}, R_{b_i}] \quad (17)$$

for some $m \in \mathbb{N}$, $a, a_i, b_i \in A$ ($1 \leq i \leq m$) satisfying

$$3a + \sum_{i=1}^m [a_i, b_i] \in \text{Nuc}(A). \quad (18)$$

Proof. If D has the form (17), then

$$D(xy) - (Dx)y - x(Dy) = \left[x, 3a + \sum_{i=1}^m [a_i, b_i], y \right]$$

for all $x, y \in A$ by (11), (13), so (18) is equivalent to D being a derivation of A .

Now suppose D satisfies (17), (18). Then $D \in \text{Der}(A)$ and $\hat{D} = \hat{L}_a - \hat{R}_a + \sum [\hat{L}_{a_i}, \hat{R}_{b_i}]$ since the right-hand side kills 1. This implies $\hat{D} \in \mathfrak{L}(\hat{A})$, hence $D \in \text{LMDer}(A)$ by (1.7). Conversely, let D be a Lie multiplication derivation of A . By Prop. 2.2,

$$\hat{D} = \hat{L}_{\hat{a}} + \hat{R}_{\hat{b}} + \sum_{i=1}^m [\hat{L}_{\hat{a}_i}, \hat{R}_{\hat{b}_i}]$$

for some $m \in \mathbb{N}$, $\hat{a}, \hat{a}_i, \hat{b}_i \in \hat{A}$ ($1 \leq i \leq m$), where $\hat{D}1 = 0$ implies $\hat{b} = -\hat{a}$, hence (17) with a, a_i, b_i being the A -components of $\hat{a}, \hat{a}_i, \hat{b}_i$, respectively. But then (18) drops out automatically since D was assumed to be a derivation. \square

When it comes to applications of Thm. 2.3, the following more concise description of Lie multiplication derivations turns out to be useful.

2.4. Derivations and exterior powers. We introduce the notation

$$W(A) := A \oplus \bigwedge^2 A.$$

Our description will be based on two linear maps defined on $W(A)$. The first one is

$$s = s_A: W(A) \longrightarrow A, \quad s(a \oplus (b \wedge c)) := 3a + [b, c] \quad (a, b, c \in A).$$

To define the second one, we note that the flexible law (3) makes the bilinear expression $[L_a, R_b]$ alternating in $a, b \in A$ and thus leads to a linear map

$$\Delta = \Delta_A: W(A) \longrightarrow \mathfrak{gl}(A), \quad \Delta_{a \oplus (b \wedge c)} := \Delta(a \oplus (b \wedge c)) := L_a - R_a + [L_b, R_c]$$

for $a, b, c \in A$. With these notations, Thm. 2.3 implies

$$\text{LMDer}(A) = \{\Delta_x \mid x \in W(A), s(x) \in \text{Nuc}(A)\}. \quad (19)$$

Now observe that every $g \in \mathfrak{gl}(A)$ induces a linear map

$$g^\dagger: W(A) \longrightarrow W(A), \quad g^\dagger(a \oplus (b \wedge c)) = g(a) \oplus (g(b) \wedge c + b \wedge g(c)) \quad (20)$$

for $a, b, c \in A$. Clearly, the assignment $g \mapsto g^\dagger$ determines an embedding $\mathfrak{gl}(A) \rightarrow \mathfrak{gl}(W(A))$ of Lie algebras, and (1.1), (1.2), (20) are easily seen to imply

$$[D, \Delta_x] = \Delta_{D^\dagger x}, \quad [D^\dagger, \Delta_x^\dagger] = \Delta_{D^\dagger x}^\dagger \quad (21)$$

for $D \in \text{Der}(A)$, $x \in W(A)$, while (1.3) and (20) imply that the diagram

$$\begin{array}{ccc} W(A) & \xrightarrow{D^\dagger} & W(A) \\ s \downarrow & & \downarrow s \\ A & \xrightarrow{D} & A \end{array} \quad (22)$$

commutes.

Remark. Lie multiplication derivations of alternative algebras as described in Thm. 2.3 do not in general satisfy the Mapping Principle of the introduction since a homomorphism $A \rightarrow B$ may not map the nucleus of A into the nucleus of B , so if (18) holds for elements $a, a_i, b_i \in A$, it may no longer do so for their images in B . For this reason, we will introduce *inner* derivations of alternative algebras as a special type of Lie multiplication derivations where such unpleasantness can be ruled out. McCrimmon [15, A5.2, p. 24] gets around this difficulty in a slightly different manner, by means of his notion of *strictly inner* derivations.

2.5. Classes of inner derivations. With the terminology of 2.4, the elements of

$$\text{InDer}_{\text{alt}}(A) := \{\Delta_x \mid x \in W(A), s(x) = 0\} \quad (23)$$

are called *inner derivations* of A . Thus the difference to (19) is that $s(x)$ is required to be zero instead of in the nucleus. In more explicit terms, the inner derivations of A are precisely the linear maps

$$L_a - R_a + \sum_{i=1}^m [L_{a_i}, R_{b_i}], \quad (24)$$

where $m \in \mathbb{N}$ and $a, a_i, b_i \in A$ ($1 \leq i \leq m$) satisfy the relation

$$3a + \sum_{i=1}^m [a_i, b_i] = 0. \quad (25)$$

Inner derivations obviously satisfy the Mapping Principle. Adapting the terminology of McCrimmon [15] to the present set-up, and identifying A and $\bigwedge^2 A$ canonically with submodules of $W(A) = A \oplus \bigwedge^2 A$ throughout the rest of the paper, we now introduce the following three classes of inner derivations.

(a) *Associator derivations.* These are the elements of

$$\text{AssDer}(A) := \{\Delta_u \mid u \in \bigwedge^2 A, s(u) = 0\} \subseteq \text{InDer}_{\text{alt}}(A). \quad (26)$$

They have the form $\sum [L_{a_i}, R_{b_i}]$, where $a_i, b_i \in A$ satisfy $\sum [a_i, b_i] = 0$, so by (12), they act on $x \in A$ as $x \mapsto \sum [a_i, b_i, x]$, i.e., as a sum of associators, hence the name. In particular, $\text{AssDer}(A) = \{0\}$ if A is associative. On the other hand, as we shall see in Thm. 5.1 below, associator derivations play an important role in octonion algebras.

(b) *Standard derivations.* These are the elements of

$$\text{StanDer}(A) := \{\Delta_{s(u) \oplus (-3u)} \mid u \in \bigwedge^2 A\} \subseteq \text{InDer}_{\text{alt}}(A). \quad (27)$$

As a k -module, $\text{StanDer}(A)$ is spanned by the elements

$$D_{a,b} = L_{[a,b]} - R_{[a,b]} - 3[L_a, R_b] = [L_a, L_b] + [L_a, R_b] + [R_a, R_b] \quad (28)$$

for $a, b \in A$, the last equation being a consequence of (5), (6). Standard derivations have the advantage of being parametrized by the full k -module $\bigwedge^2 A$, with no further constraints on the parameters involved. On the other hand, $\text{StanDer}(A) = \{0\}$ if A is commutative, since this is well known to imply $3[A, A, A] = \{0\}$, hence $D_{a,b} = 0$ for all $a, b \in A$ by (12) and (28); cf. Prop. 2.7 below for a more precise statement.

(c) *Commutator derivations.* These are the elements of

$$\text{ComDer}(A) := \{\Delta_a \mid a \in A, 3a = 0\} = \{L_a - R_a \mid a \in A, 3a = 0\} \subseteq \text{InDer}_{\text{alt}}(A).$$

They appear only in the presence of 3-torsion. Note $\Delta_a x = [a, x]$ for $a, x \in A$, justifying the chosen terminology.

2.6. Proposition. *In the terminology of 2.5,*

$$3\Delta_x = \Delta_{s(x)} + \Delta_{s(-u) \oplus 3u} \quad (29)$$

for all $x = a \oplus u \in W(A)$ satisfying $s(x) \in \text{Nuc}(A)$. In particular,

$$3\text{LMDer}(A) \subseteq \{L_a - R_a \mid a \in \text{Nuc}(A)\} + \text{StanDer}(A) \subseteq \text{LMDer}(A). \quad (30)$$

Proof. (29) is obvious by (19) and the definition of s ; it immediately implies (30) since the first summand on the right of (29) is the nuclear derivation $L_{a'} - R_{a'}$, $a' \in \text{Nuc}(A)$, while the second one by (27) is a standard derivation.

Remark. If $3A = A$, we obtain

$$\text{LMDer}(A) = \{L_a - R_a \mid a \in \text{Nuc}(A)\} + \text{StanDer}(A),$$

hence Schafer's classical description [20, pp. 76–78] of Lie multiplication derivations of unital alternative algebras over fields of characteristic not 2 or 3, see also McCrimmon [15, A5, 2.17].

2.7. Proposition. *In the terminology of 2.5, the following statements hold.*

- (a) $\text{InDer}_{\text{alt}}(A)$, $\text{AssDer}(A)$, $\text{StanDer}(A)$, $\text{ComDer}(A)$ are all ideals in the full derivation algebra of A .
- (b) $\text{AssDer}(A)A \subseteq [A, A, A]$, $3\text{AssDer}(A) \subseteq \text{StanDer}(A)$.
- (c) $(\text{StanDer}(A) + \text{ComDer}(A))A \subseteq [A, A]$.
- (d) If $3[A, A] = [A, A]$, then

$$\text{InDer}_{\text{alt}}(A) = \text{AssDer}(A) + \text{StanDer}(A) + \text{ComDer}(A).$$

- (e) If $3A = A$, then

$$\text{InDer}_{\text{alt}}(A) = \text{StanDer}(A) + \text{ComDer}(A).$$

- (f) If $\frac{1}{3} \in k$, then

$$\text{InDer}_{\text{alt}}(A) = \text{StanDer}(A), \quad \text{ComDer}(A) = \{0\}.$$

- (g) If $3A = \{0\}$, then

$$\text{InDer}_{\text{alt}}(A) = \text{AssDer}(A) + \text{ComDer}(A).$$

Proof. (a) follows immediately from (21), (22).

(b) The first part has already been observed in 2.5(a), while the second one follows from (29) in the special case $a = s(u) = 0$.

(c) The relation $\text{ComDer}(A)A \subseteq [A, A]$ is obvious. Since (28) and (12) imply $\text{StanDer}(A)A \subseteq [A, A] + 3[A, A, A]$, it remains to show $3[A, A, A] \subseteq [A, A]$ which is probably known; we include a proof for convenience. Modulo $[A, A]$ we have by (11) that $3[x, a, y] \equiv [x, a]y + x[y, a] \equiv y[x, a] + x[y, a]$, which is $\equiv 0$, being the bilinearization of $x[x, a] = x(xa) - x(ax) = x(xa) - (xa)x$ (by (3)) $\in [A, A]$.

(d), (e) Suppose $x = a \oplus u \in W(A)$ satisfies $s(x) = 0$ (cf. (23)). Since $s(\bigwedge^2 A) = [A, A]$, the hypothesis in (d) leads to an element $w \in \bigwedge^2 A$ such that $s(u) = 3s(w)$, so $v := u - 3w$ satisfies $s(v) = 0$. On the other hand, the hypothesis in (e) leads to an element $w \in \bigwedge^2 A$ such that $u = 3w$, so again $s(u) = 3s(w)$, but this time even $v := u - 3w = 0$. In any event, setting $b = a + s(w)$, we conclude $3b = 3a + 3s(w) = 3a + s(u) = s(x) = 0$. Moreover,

$$\Delta_x = \Delta_v + \Delta_{(-s(w)) \oplus 3w} + \Delta_b \in \text{AssDer}(A) + \text{StanDer}(A) + \text{ComDer}(A).$$

Hence (d) and (e) hold. Now (f) follows immediately from (e) since $\text{ComDer}(A) = \{0\}$ in the absence of 3-torsion, and (g) is a consequence of (23). \square

2.8. Example. Let A be an associative k -algebra. Using (24), (25), one checks easily that

$$\text{InDer}_{\text{alt}}(A) \subseteq \text{InDer}_{\text{ass}}(A).$$

However, equality does not hold in general. To see this, suppose k contains $\frac{1}{3}$. Then $\text{InDer}_{\text{alt}}(A) = \text{StanDer}(A)$ by Prop. 2.7, while (28) reduces to $D_{a,b} = L_{[a,b]} - R_{[a,b]}$ for all $a, b \in A$. Hence we obtain

$$L_x - R_x \in \text{InDer}_{\text{ass}}(A) \setminus \text{InDer}_{\text{alt}}(A) \quad (31)$$

for any $x \in A$ that does not belong to $Z + [A, A]$, Z being the centre of A . More specifically, let k be a field of characteristic $p > 0$, $p \neq 3$, and put $A = \text{Mat}_p(k)$, the algebra of $p \times p$ matrices with entries in k . Then $[A, A]$, being the kernel of the trace, contains $Z = k \cdot 1_A$, so any $x \in A$ with non-zero trace will satisfy (31).

In the presence of the alternative law, we can improve and expand Cor. 1.12 considerably.

2.9. Proposition. *If A is finitely generated as a k -module, then the Lie algebras $\text{LMDer}(A)$, $\text{InDer}_{\text{alt}}(A)$, $\text{AssDer}(A)$, $\text{StanDer}(A)$, $\text{ComDer}(A)$ all commute with flat base change: For all flat k -algebras $R \in k\text{-alg}$, we have*

$$\text{LMDer}(A)_R = \text{LMDer}(A_R), \quad (32)$$

$$\text{InDer}_{\text{alt}}(A)_R = \text{InDer}(A_R), \quad (33)$$

$$\text{AssDer}(A)_R = \text{AssDer}(A_R), \quad (34)$$

$$\text{StanDer}(A)_R = \text{StanDer}(A_R), \quad (35)$$

$$\text{ComDer}(A)_R = \text{ComDer}(A_R). \quad (36)$$

Proof. Since taking exterior powers commutes with flat (even arbitrary) base change [4, III, §7, Prop. 8], so do the linear maps Δ_A and s_A . Furthermore, (19),(23) yield

$$\text{LMDer}(A) = \Delta_A \left(s_A^{-1}(\text{Nuc}(A)) \right), \quad \text{InDer}_{\text{alt}}(A) = \Delta_A(\text{Ker}(s_A)).$$

Hence (32),(33) follow from Lemma 1.7(a),(b) and Prop. 1.14, while an analogous argument yields (36). After the identifications of 1.9, we obtain $(D_{a,b})_R = D_{a_R, b_R}$ for all $a, b \in A$, and (35) follows from Lemma 1.7(e). It remains to prove (34). To do so, we put $K_A := \text{Ker}(s_A) \cap \bigwedge^2 A$ (which commutes with flat base change since s_A does), $T_A = \text{Ker}(\Delta_A)$, note that Δ by (26) restricts to a linear surjection $\Delta_0: K_A \rightarrow \text{AssDer}(A)$,

and obtain a commutative diagram

$$\begin{array}{ccccccc}
& & & 0 & & 0 & \\
& & & \downarrow & & \downarrow & \\
0 & \longrightarrow & T_A & \longrightarrow & K_A & \xrightarrow{\Delta_0} & \text{AssDer}(A) \longrightarrow 0 \\
& & & \downarrow & & \downarrow & \\
& & & \bigwedge^2 A & \xrightarrow{\Delta} & \text{End}_k(A) & \\
& & & \downarrow s & & & \\
& & & [A, A] & & & \\
& & & \downarrow & & & \\
& & & 0 & & &
\end{array}$$

with exact rows and columns. Tensoring with R , we end up with (34). \square

3. Making inner derivations functorial

In this section, we address the lack of functoriality of the inner derivation algebra mentioned in the introduction. Before delving into the general categorical setup, it may be helpful to consider the following example.

3.1. Associative algebras. Every associative algebra A defines a Lie algebra A^- having underlying k -module A and Lie product $[x, y] = xy - yx$, and A^- depends functorially on A : every homomorphism $f: A \rightarrow B$ induces a Lie algebra homomorphism $f^- = f: A^- \rightarrow B^-$. The inner derivations of A are the derivations $D_x = L_x - R_x$ and the map sending $x \in A^-$ to $D_x \in \text{Der}(A)$ is a Lie algebra homomorphism $\alpha: A^- \rightarrow \text{Der}(A)$, i.e., an action of A^- on A by derivations, whose kernel is central in (indeed, equals the centre of) A^- . Moreover, the formula

$$f(D_x(a)) = D_{f^-(x)}(f(a)) \quad (1)$$

holds for all $x \in A^-$, $a \in A$, so the derivation $D_{f^-(x)}$ of B is f -related to the derivation D_x of A . But note $D_x = 0$ for central $x \in A$, yet $D_{f^-(x)}$ need not be zero if $f^-(x)$ is not central in B , showing why $D_x \mapsto D_{f^-(x)}$ is in general not a well defined map from $\text{InDer}_{\text{ass}}(A)$ to $\text{InDer}_{\text{ass}}(B)$.

3.2. Categories of algebras. Abstracting from the previous example, we replace associative algebras by an arbitrary category of non-associative algebras. To discuss questions of base change, it is convenient to consider not only algebras over a fixed base ring, but over all possible commutative rings, more generally, over a subcategory \mathcal{R} of all commutative rings. Thus let \mathcal{A} be a category of *algebras over* \mathcal{R} in the following sense: objects of \mathcal{A} are pairs (k, A) where $k \in \mathcal{R}$ and A is a non-associative k -algebra. Morphisms

$(k, A) \rightarrow (l, B)$ of \mathcal{A} are pairs (τ, f) where $\tau: k \rightarrow l$ is a morphism of \mathcal{R} and $f: A \rightarrow B$ is a τ -semilinear map from the k -module A to the l -module B preserving the algebraic structure. Examples of this situation abound: associative, alternative or Jordan algebras, unital or not, Lie algebras, etc.

The projection onto the first component $\Pi: \mathcal{A} \rightarrow \mathcal{R}$ is a functor, and we denote by \mathcal{A}_k the fibre of Π over $k \in \mathcal{R}$, that is, the subcategory of \mathcal{A} with objects algebras over the fixed ring k , and morphisms k -linear maps, thus of the form $(\text{Id}_k, f): (k, A) \rightarrow (k, B)$. It is often inconvenient to indicate explicitly the base ring k of an object of \mathcal{A} . Thus we frequently write simply $A \in \mathcal{A}_k$ or even $A \in \mathcal{A}$ and $k = \Pi(A)$ instead of $(k, A) \in \mathcal{A}$, or employ a phrase like “let A be a k -algebra in \mathcal{A} ”. Similarly, a morphism of \mathcal{A} will often be written as $f: A \rightarrow B$, with $\Pi(f) = \tau: \Pi(A) = k \rightarrow \Pi(B) = l$ the corresponding homomorphism of the respective base rings.

3.3. The derivation category. Let \mathcal{A} be a category of (non-associative) algebras over \mathcal{R} as before and let $\mathcal{L}ie$ be the category of Lie algebras over \mathcal{R} . We define a category $\mathcal{D}er(\mathcal{A})$ over \mathcal{A} as follows.

A *derivation action* of a Lie algebra $\mathfrak{g} \in \mathcal{L}ie_k$ on an algebra $A \in \mathcal{A}_k$ is a homomorphism $\alpha: \mathfrak{g} \rightarrow \text{Der}(A)$ of k -Lie algebras. We write a derivation action as a quadruple $(k, A, \mathfrak{g}, \alpha)$ or simply as (A, α) . (Since A is not uniquely determined by $\text{Der}(A)$, the algebra A must be explicitly indicated. On the other hand, \mathfrak{g} is the domain of definition of α , so a derivation action is determined by A and α). Now construct a category $\mathcal{D}er(\mathcal{A})$, called the *derivation category of \mathcal{A}* , whose objects are the derivation actions, and whose morphisms are defined as follows.

Let $f: A \rightarrow B$ be a morphism of \mathcal{A} and $\tau = \Pi(f): k \rightarrow l$ the corresponding morphism of \mathcal{R} . An *f -derivation from A to B* is a τ -semilinear map $d: A \rightarrow B$ such that

$$d(aa') = d(a)f(a') + f(a)d(a'),$$

for all $a, a' \in A$. Denote the set of f -derivations from A to B by $\text{Der}_f(A, B)$. This is in a natural way an l -module by defining $(sd)(a) = sd(a)$ for $a \in A$ and $s \in l$. Given derivations $D \in \text{Der}(A)$ and $D' \in \text{Der}(B)$, the maps $f_*(D) := f \circ D$ and $f^*(D') := D' \circ f$ both belong to $\text{Der}_f(A, B)$.

A *morphism* from (A, α) to (B, β) is now defined as a pair (f, φ) where $f: A \rightarrow B$ and $\varphi: \mathfrak{g} \rightarrow \mathfrak{h}$ are morphisms of \mathcal{A} and $\mathcal{L}ie$, respectively, satisfying $\Pi(f) = \Pi(\varphi)$, and making the following diagram commutative:

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\varphi} & \mathfrak{h} \\ \alpha \downarrow & & \downarrow \beta \\ \text{Der}(A) & \xrightarrow{f_*} \text{Der}_f(A, B) \xleftarrow{f^*} & \text{Der}(B) \end{array}$$

Explicitly, this means:

$$f(\alpha(X) \cdot a) = \beta(\varphi(X)) \cdot f(a),$$

for all $X \in \mathfrak{g}$ and all $a \in A$. There is again a functor $\mathcal{D}er(\mathcal{A}) \rightarrow \mathcal{R}$ given by $(A, \alpha) \mapsto \Pi(A)$ and $(f, \varphi) \mapsto \Pi(f)$. Moreover, the projections $(A, \alpha) \mapsto A$ and $(f, \varphi) \mapsto f$ define a functor

$P_1: \mathcal{D}er(\mathcal{A}) \rightarrow \mathcal{A}$. Similarly, the projections $(A, \alpha) \mapsto \mathfrak{g}$ and $(f, \varphi) \mapsto \varphi$ define a functor $P_2: \mathcal{D}er(\mathcal{A}) \rightarrow \mathcal{L}ie$.

3.4. Derivation functors. Let \mathcal{A} be a category of algebras over \mathcal{R} as before. A *derivation functor* is a functor $F: \mathcal{A} \rightarrow \mathcal{D}er(\mathcal{A})$ commuting with the projections onto \mathcal{R} which is a section of the projection $P_1: \mathcal{D}er(\mathcal{A}) \rightarrow \mathcal{A}$ in the sense that $P_1 \circ F = \text{Id}_{\mathcal{A}}$.

In more detail, this means the following: for every $A \in \mathcal{A}_k$, we have a derivation action ρ_A of a k -Lie algebra $\mathfrak{d}(A)$ on A , and for every homomorphism $f: A \rightarrow B$ of algebras in \mathcal{A} we have a morphism $\mathfrak{d}(f): \mathfrak{d}(A) \rightarrow \mathfrak{d}(B)$ of Lie algebras, semilinear with respect to $\Pi(f)$, and compatible with the actions in the sense that the diagram

$$\begin{array}{ccc} \mathfrak{d}(A) & \xrightarrow{\mathfrak{d}(f)} & \mathfrak{d}(B) \\ \rho_A \downarrow & & \downarrow \rho_B \\ \mathcal{D}er(A) & \xrightarrow{f_*} \mathcal{D}er_f(A, B) \xleftarrow{f^*} & \mathcal{D}er(B) \end{array} \quad (2)$$

is commutative. As before, this means

$$f(\rho_A(X) \cdot a) = \rho_B(\mathfrak{d}(f)(X)) \cdot f(a) \quad (X \in \mathfrak{d}(A), a \in A).$$

In particular, $\mathfrak{d}: \mathcal{A} \rightarrow \mathcal{L}ie$ is a functor from \mathcal{A} to $\mathcal{L}ie$. It is tempting to say that ρ is a natural transformation from \mathfrak{d} to the ‘‘functor’’ $\mathcal{D}er$, but $\mathcal{D}er(A)$ does not depend functorially on A . As a substitute, for every morphism $f: A \rightarrow B$ in \mathcal{A} , commutativity of (2) is equivalent to $(f, \mathfrak{d}(f)): (A, \rho_A) \rightarrow (B, \rho_B)$ being a morphism of $\mathcal{D}er(\mathcal{A})$. By abuse of notation, we will often write $F = (\mathfrak{d}, \rho)$ for a derivation functor.

Suppose $F = (\mathfrak{d}, \rho)$ and $F' = (\mathfrak{d}', \rho')$ are derivation functors. A *morphism* from F to F' is a natural transformation of functors. This amounts to Lie algebra homomorphisms $h_A: \mathfrak{d}(A) \rightarrow \mathfrak{d}'(A)$ for all $A \in \mathcal{A}$ defining a natural transformation $\mathfrak{d} \rightarrow \mathfrak{d}'$ and making the diagrams

$$\begin{array}{ccc} \mathfrak{d}(A) & \xrightarrow{h_A} & \mathfrak{d}'(A) \\ \rho_A \searrow & & \swarrow \rho'_A \\ & \mathcal{D}er(A) & \end{array}$$

commutative.

A derivation functor F is called *inner* if ρ_A maps $\mathfrak{d}(A)$ into the Lie multiplication algebra of A , and *central* if the kernel of ρ_A is central in $\mathfrak{d}(A)$, for all $A \in \mathcal{A}$.

In the special case where $f = g \in \text{Aut}(A)$, the functoriality of F implies that $\text{Aut}(A)$ acts on the Lie algebra $\mathfrak{d}(A)$ by automorphisms and the map ρ_A is equivariant with respect to this action on the one hand, and with respect to conjugation of $\text{Aut}(A)$ on $\mathcal{D}er(A)$ on the other, because (2) now says

$$g \circ \rho_A(X) \circ g^{-1} = \rho_A(\mathfrak{d}(g)(X)),$$

for all $g \in \text{Aut}(A)$ and $X \in \mathfrak{d}(A)$.

If we assume that F commutes with flat base change (see 3.10), which holds in all standard examples, then by extending k to the dual numbers $k(\varepsilon)$ and specializing $g = \text{Id} + \varepsilon D$ for $D \in \text{Der}(A)$, we obtain an action of $\text{Der}(A)$ on $\mathfrak{d}(A)$ by derivations and ρ_A is equivariant with respect to this action and the adjoint representation of $\text{Der}(A)$ on itself. In particular, $\mathfrak{d}(A)$ is then an ideal in $\text{Der}(A)$.

3.5. Remarks and examples. (a) The definition of a derivation functor does not tie the Lie algebras $\mathfrak{d}(A)$ very closely to the derivations of A : the kernels of ρ_A can be arbitrarily big. For example, let \mathfrak{l} be a fixed Lie algebra over \mathbb{Z} and define F by $\mathfrak{d}(A) = \mathfrak{l} \otimes_{\mathbb{Z}} k$ and $\rho_A = 0$ for all $A \in \mathcal{A}_k$ and $k \in \mathcal{R}$. The requirement that F be central cuts down the kernels to some extent; it is satisfied in all the examples treated below. The same is true of the condition of innerness.

(b) The example of associative algebras treated in 3.1 yields a derivation functor F with $\mathfrak{d}(A) = A^-$ and $\rho_A(x) = D_x$. The commutativity of (2) is formula (1), and F is inner and central.

(c) Let $\mathcal{A} = \mathcal{L}ie$. Then a natural choice of F is $\mathfrak{d} = \text{Id}_{\mathcal{L}ie}$ and ρ the adjoint representation. Again F is inner and central.

We show next that the classes of inner derivations of alternative algebras introduced in 2.5 come from derivation functors as well. The construction rests on the following simple lemma which is essentially contained in [13, 2.1]. We include a proof for the convenience of the reader.

3.6. Lemma. *Let \mathfrak{g} be a k -Lie algebra, M a k - \mathfrak{g} -module, and let $\phi: M \rightarrow \mathfrak{g}$ be a homomorphism of left k - \mathfrak{g} -modules, where \mathfrak{g} acts on itself by the adjoint representation. Write the action of an element $x \in \mathfrak{g}$ on $u \in M$ as $x_M \cdot u$, and define a non-associative product $\{u, v\}$ on M by*

$$\{u, v\} := \phi(u)_M \cdot v.$$

(a) *The map $\phi: M \rightarrow \mathfrak{g}$ is a homomorphism of non-associative algebras:*

$$\phi(\{u, v\}) = [\phi(u), \phi(v)]. \quad (3)$$

(b) *\mathfrak{g} acts by derivations of the product $\{-, -\}$. The Jacobi identity holds in the following form on M :*

$$\{u, \{v, w\}\} - \{v, \{u, w\}\} = \{\{u, v\}, w\}, \quad (4)$$

so M is a left Leibniz algebra [13].

(c) *Let Q be the k -linear span of all squares $\{u, u\}$, $u \in M$, and let $Z = \text{Ker}(\phi) \subseteq M$. Then Q and Z are stable under the action of \mathfrak{g} , and*

$$Q \subseteq Z, \quad \{Z, M\} = 0, \quad \{M, Z\} \subseteq Q. \quad (5)$$

(d) The product $\{-, -\}$ induces a Lie algebra structure on $\mathfrak{h} := M/Q$ and ϕ induces a Lie algebra homomorphism $\bar{\phi} : \mathfrak{h} \rightarrow \mathfrak{g}$ whose kernel Z/Q is central in \mathfrak{h} . The action of \mathfrak{g} on M induces an action of \mathfrak{g} on \mathfrak{h} by derivations.

Proof. (a) Put $x = \phi(u) \in \mathfrak{g}$. Since ϕ is a homomorphism of \mathfrak{g} -modules, $\phi(\{u, v\}) = \phi(\phi(u)_M \cdot v) = \phi(x_M \cdot v) = [x, \phi(v)] = [\phi(u), \phi(v)]$.

(b) Let $x \in \mathfrak{g}$ and $v, w \in M$. Then

$$\begin{aligned} x_M \cdot \{v, w\} &= x_M \cdot (\phi(v)_M \cdot w) = [x_M, \phi(v)_M] \cdot w + \phi(v)_M \cdot (x_M \cdot w) \\ &= [x, \phi(v)]_M \cdot w + \phi(v)_M \cdot (x_M \cdot w) = \phi(x_M \cdot v)_M \cdot w + \{v, x_M \cdot w\} \\ &= \{x_M \cdot v, w\} + \{v, x_M \cdot w\}. \end{aligned}$$

Now (4) follows by specializing $x = \phi(u)$.

(c) Since ϕ is a homomorphism of \mathfrak{g} -modules, it is clear that Z is stable under \mathfrak{g} , and (b) implies that Q is stable under \mathfrak{g} as well. The inclusion $Q \subseteq Z$ follows from (3) and the fact that the Lie product in \mathfrak{g} is alternating. Let $u \in Z$ and $v \in M$. Then $\phi(u) = 0$, hence also $\{u, v\} = \phi(u)_M \cdot v = 0$ which proves $\{Z, M\} = 0$. Moreover, $\{v, u\} = \{u, v\} + \{v, u\} = \{u + u, v + v\} - \{u, u\} - \{v, v\} \in Q$, so $\{M, Z\} \subseteq Q$.

(d) Since $Q \subseteq Z$, (5) implies $\{Q, M\} + \{M, Q\} \subset Q$, so Q is an ideal of $\{-, -\}$. As it contains all squares, the product induced on M/Q is alternating. Now (4) shows that $\mathfrak{h} = M/Q$ is a Lie algebra, and $\bar{\phi}$ is a Lie algebra homomorphism by (3). Finally, $\{Z, M\} = 0$ implies that $\text{Ker}(\bar{\phi}) = Z/Q$ is central in \mathfrak{h} , and it follows from (b) that \mathfrak{g} acts on \mathfrak{h} by derivations.

3.7. The inner derivation functor of alternative algebras. Let A be an alternative algebra over k . Recall from 2.4 that $W(A) := A \oplus \wedge^2 A$ is a $\mathfrak{gl}(A)$ -module under the action

$$\mathfrak{gl}(A) \times W(A) \longrightarrow W(A), \quad (g, x) \longmapsto g \cdot x := g^\dagger(x).$$

For a τ -semilinear homomorphism $f: A \rightarrow B$ of alternative algebras over k and l , respectively, the map $W(f) := f \oplus \wedge^2 f: W(A) \rightarrow W(B)$ is again τ -semilinear, and one checks easily that the relations

$$f \circ s = s \circ W(f), \quad W(f) \circ \Delta_x^\dagger = \Delta_{W(f)(x)}^\dagger \circ W(f), \quad f \circ \Delta_x = \Delta_{W(f)(x)} \circ f \quad (6)$$

hold for all $x \in W(A)$.

We consider the k -submodule

$$W_{\text{in}}(A) := \text{Ker}(s) = \{x \in W(A) \mid s(x) = 0\}.$$

of the $\mathfrak{gl}(A)$ -module $W(A)$, which by (2.22) remains stable under $\text{Der}(A)$, hence may be regarded canonically as a $\text{Der}(A)$ -module. Define $\phi: W_{\text{in}}(A) \rightarrow \text{Der}(A)$ by $\phi(x) := \Delta_x$ for $x \in W_{\text{in}}(A)$. Then (2.21) shows that ϕ is a homomorphism of $\text{Der}(A)$ -modules.

Applying Lemma 3.6 therefore yields a Leibniz algebra $W_{\text{in}}(A)$, a Lie algebra $\mathfrak{d}_{\text{in}}(A) = W_{\text{in}}(A)/Q$ and a homomorphism $\rho_{\text{in},A} = \bar{\phi}: \mathfrak{d}_{\text{in}}(A) \rightarrow \text{Der}(A)$ whose image, by (2.23), is precisely $\text{InDer}_{\text{alt}}(A)$.

Returning to the τ -semilinear homomorphism $f: A \rightarrow B$, we conclude from (6) that $W(f)$ sends $W_{\text{in}}(A)$ to $W_{\text{in}}(B)$ and hence induces a τ -semilinear map $W_{\text{in}}(f): W_{\text{in}}(A) \rightarrow W_{\text{in}}(B)$, which, again by (6), is a homomorphism of Leibniz algebras, inducing canonically a Lie algebra homomorphism $\mathfrak{d}_{\text{in}}(f): \mathfrak{d}_{\text{in}}(A) \rightarrow \mathfrak{d}_{\text{in}}(B)$. The commutativity of (2) is a consequence of (6), so we have defined a derivation functor F_{in} for alternative algebras, the *inner derivation functor*.

3.8. The associator and commutator derivation functors. Consider the submodule $W_{\text{ass}}(A) = W_{\text{in}}(A) \cap \bigwedge^2 A$ of $W_{\text{in}}(A)$, which is in fact a $\text{Der}(A)$ -stable subalgebra of the Leibniz algebra $W_{\text{in}}(A)$, so the construction of 3.7 can be performed mutatis mutandis on $W_{\text{ass}}(A)$ and yields a derivation functor F_{ass} , called the *associator derivation functor* of alternative algebras. The inclusions $W_{\text{ass}}(A) \rightarrow W_{\text{in}}(A)$ induce homomorphisms $\mathfrak{d}_{\text{ass}}(A) \rightarrow \mathfrak{d}_{\text{in}}(A)$ of Lie algebras (in general no longer injective) which are compatible with the representations ρ and ρ_{ass} and with morphisms of \mathcal{A} . Thus we have a natural transformation $F_{\text{ass}} \rightarrow F_{\text{in}}$ of derivation functors.

Similarly, let $W_{\text{com}}(A) = W_{\text{in}}(A) \cap A = {}_3A$, the 3-torsion elements of A . As before, $W_{\text{com}}(A)$ is a subalgebra of $W_{\text{in}}(A)$ and in fact is already a Lie algebra, because $\{a, a\} = \Delta_a \cdot a = [a, a] = 0$. We obtain a derivation functor $F_{\text{com}} = (\mathfrak{d}_{\text{com}}, \rho_{\text{com}})$, the *commutator derivation functor*, where $\mathfrak{d}_{\text{com}}(A) = {}_3A$ with Lie bracket $[x, y] = xy - yx$, and $\rho_{\text{com}}: {}_3A \rightarrow \text{Der}(A)$ given by the commutator: $x \mapsto (a \mapsto [x, a])$. Again, there is a natural transformation $F_{\text{com}} \rightarrow F_{\text{in}}$ induced from the inclusions $W_{\text{com}}(A) \rightarrow W_{\text{in}}(A)$.

3.9. Standard derivation functors. By definition (2.5(b)), the standard derivations of an alternative algebra A are of the form $\Delta_{s(u) \oplus (-3u)}$, $u \in \bigwedge^2 A$. Thus they can be parametrized by all of $\bigwedge^2 A$ or by the image in $W(A)$ of $\bigwedge^2 A$ under the map $\zeta: u \mapsto s(u) \oplus (-3u)$. This gives rise to two standard derivation functors as follows.

First let $M := \bigwedge^2 A$ and define $\phi: M \rightarrow \text{Der}(A)$ by $\phi(a \wedge b) = [L_a, L_b] + [L_a, R_b] + [R_a, R_b]$ as in (2.28). It follows from the formulas in 2.4 that ϕ is equivariant with respect to the action of $\text{Der}(A)$ on M and on itself by the adjoint representation. Hence Lemma 3.6 yields a Lie algebra $\mathfrak{d}_{\text{st}}(A) = M/Q$ and a homomorphism $\rho_{\text{st}}(A): \mathfrak{d}_{\text{st}}(A) \rightarrow \text{Der}(A)$ with central kernel and image $\text{StanDer}(A)$. Also, for $f: A \rightarrow B$ a homomorphism of alternative algebras, we have

$$f \circ \phi(u) = \phi\left(\bigwedge^2 f(u)\right) \circ f.$$

Hence $\bigwedge^2 f$ is a homomorphism of Leibniz algebras, and induces a homomorphism $\mathfrak{d}_{\text{st}}(f): \mathfrak{d}_{\text{st}}(A) \rightarrow \mathfrak{d}_{\text{st}}(B)$, compatible with the representations ρ_{st} . This defines a derivation functor $F_{\text{st}} = (\mathfrak{d}_{\text{st}}, \rho_{\text{st}})$, the *standard derivation functor* of alternative algebras.

Next, imitating the procedure of 3.8, we have a derivation functor induced from

$$W_{\text{st}}(A) = \{\zeta(u) \mid u \in \bigwedge^2 A\} \subseteq W_{\text{in}}(A),$$

denoted $\bar{F}_{\text{st}} = (\bar{\mathfrak{d}}_{\text{st}}, \bar{\rho}_{\text{st}})$. As before, the inclusions $W_{\text{st}}(A) \rightarrow W_{\text{in}}(A)$ induce a morphism $\bar{F}_{\text{st}} \rightarrow F_{\text{in}}$. The map ζ induces a morphism $h: F_{\text{st}} \rightarrow \bar{F}_{\text{st}}$ with the property that $h_A: \mathfrak{d}_{\text{st}}(A) \rightarrow \bar{\mathfrak{d}}_{\text{st}}(A)$ is always surjective. In general, however, h_A is not injective. For example, if A is commutative then $\text{StanDer}(A) = \{0\}$, $\mathfrak{d}_{\text{st}}(A) = \bigwedge^2 A$ and $\bar{\mathfrak{d}}_{\text{st}}(A) = W_{\text{st}}(A) \cong 3 \bigwedge^2 A$ (abelian Lie algebras), and h_A is multiplication by -3 . Thus \bar{F}_{st} is closer to the standard derivations in the sense that the kernels of $\bar{\rho}_{\text{st}}(A)$ are smaller. On the other hand, F_{st} commutes with arbitrary base change, whereas \bar{F}_{st} does so only for flat base change, see Prop. 3.12 below for details.

3.10. Base change. Let \mathcal{A} be a category of algebras over \mathcal{R} as in 3.2. We say that \mathcal{A} admits base change if for every $A \in \mathcal{A}_k$ and every homomorphism $\tau: k \rightarrow R$ of \mathcal{R} the R -algebra $A_R = A \otimes_k R$ (with the naturally extended algebraic structure) belongs to \mathcal{A}_R . (In more precise categorical language, this says that \mathcal{A} is a co-fibred category over \mathcal{R} .) This is true for all the examples considered in this paper, in particular for the category of Lie algebras. If \mathcal{A} admits base change then so does $\mathcal{D}er(\mathcal{A})$: Indeed, for a morphism $\tau: k \rightarrow R$ of \mathcal{R} and a derivation action α of $\mathfrak{g} \in \mathcal{L}ie_k$ on $A \in \mathcal{A}_k$, it is easily seen that

$$\alpha_R := \text{can} \circ (\rho \otimes \text{Id}_R): \mathfrak{g} \otimes_k R \rightarrow \text{Der}(A) \otimes_k R \rightarrow \text{Der}(A \otimes_k R).$$

is a derivation action of $\mathfrak{g}_R = \mathfrak{g} \otimes_k R$ on A_R , called the *base change of α with respect to $\tau: k \rightarrow R$* .

Suppose \mathcal{A} admits base change and $F: \mathcal{A} \rightarrow \mathcal{D}er(\mathcal{A})$ is a derivation functor. We say F commutes with base change if for all morphisms $\tau: k \rightarrow R$ of \mathcal{R} there are natural isomorphisms

$$g(\tau): \mathfrak{d}(A)_R \xrightarrow{\cong} \mathfrak{d}(A_R), \quad (7)$$

making the diagrams

$$\begin{array}{ccc} \mathfrak{d}(A)_R & \xrightarrow[\cong]{g(\tau)} & \mathfrak{d}(A_R) \\ & \searrow^{(\rho_A)_R} & \swarrow_{\rho_{A_R}} \\ & \text{Der}(A_R) & \end{array} \quad (8)$$

commutative. (Naturality means that the $g(\tau)$ behave in the expected way with respect to composition of morphisms in \mathcal{R} and the usual canonical isomorphisms between repeated tensor products. A more precise formulation would require the formalism of fibred categories.)

One sees immediately that the usual inner derivation functor of associative algebras, see 3.5(b), commutes with arbitrary base change. We will now show that the standard derivation functor F_{st} of alternative algebras commutes with arbitrary base change, and that the other derivation functors of alternative algebras introduced earlier commute with flat base change. Let us emphasize that this does not improve Proposition 2.9 since we are not dealing with the algebras $\rho_A(\mathfrak{d}(A))$ of inner derivations of the respective type themselves, but with the more abstractly defined Lie algebras $\mathfrak{d}(A)$. We begin with a lemma.

3.11. Lemma. *Let M, N be k -modules and let $q: M \rightarrow N$ be a quadratic map. Let $R \in k\text{-alg}$ and let $q_R: M_R \rightarrow N_R$ be the extension of q to a quadratic map of R -modules, cf. [18, Proposition 2.1]. Let $q(M) \subseteq N$ be the k -linear span of $\{q(x) : x \in M\}$ and define $q_R(M_R) \subseteq N_R$ analogously. Finally, let $\iota: q(M) \rightarrow N$ be the inclusion map. Then the base extension $\iota_R: q(M) \otimes R \rightarrow N_R$ has image $q_R(M_R)$.*

Proof. For $y = \sum x_i \otimes r_i \in M_R$, we have

$$q_R(y) = \sum q(x_i) \otimes r_i^2 + \sum_{i < j} q(x_i, x_j) \otimes r_i r_j, \quad (9)$$

where $q(-, -)$ is the polar map of q . On the other hand, let us denote a typical spanning element of $q(M) \otimes R$ by $q(x) \tilde{\otimes} r$, to distinguish the tensor product in $q(M) \otimes R$ (where $q(M)$ is taken as a k -module in its own right) from the tensor product in $N \otimes R$. Then $\iota_R(q(x) \tilde{\otimes} r) = q(x) \otimes r = q_R(x \otimes 1_R)r$. This implies that indeed $\iota_R(q(M) \otimes R) \subseteq q_R(M_R)$. Moreover, (9) shows that every $q_R(y)$ belongs to the image of ι_R , proving the lemma.

3.12. Proposition. (a) *The standard derivation functor F_{st} of alternative algebras commutes with arbitrary base change.*

(b) *The derivation functors $F_{\text{in}}, F_{\text{ass}}, F_{\text{com}}$ and \bar{F}_{st} commute with flat base change.*

Proof. (a) Let $\tau: k \rightarrow R$ be a ring homomorphism, so R is a k -algebra. It is well known that $\eta: (\wedge^2 A) \otimes R \xrightarrow{\cong} \wedge^2(A_R)$, sending $(a \wedge b)_R \mapsto a_R \wedge b_R$, is an isomorphism of R -modules.

Let $M := \wedge^2 A$, considered as a Leibniz algebra over k as in 3.6 and 3.9, and let $M' = \wedge^2(A_R)$ be the analogously defined Leibniz algebra over R for A_R . The multiplication on A_R is just the R -linear extension of the multiplication on A . Hence $\eta: M_R \rightarrow M'$ is an isomorphism of Leibniz algebras over R .

Let $q: M \rightarrow M$ be the quadratic map $x \mapsto \{x, x\}$ and define $q': M' \rightarrow M'$ in the same way. Then by definition of $\mathfrak{d}_{\text{st}}(A)$ in 3.9, we have an exact sequence

$$0 \longrightarrow q(M) \xrightarrow{\iota} \wedge^2 A \xrightarrow{\pi} \mathfrak{d}_{\text{st}}(A) \longrightarrow 0,$$

which upon tensoring with R yields the first row of the following commutative diagram with exact rows:

$$\begin{array}{ccccccc} q(M) \otimes R & \xrightarrow{\iota_R} & M \otimes R & \xrightarrow{\pi_R} & \mathfrak{d}_{\text{st}}(A) \otimes R & \longrightarrow & 0 \\ \varphi \downarrow & & \cong \downarrow \eta & & \downarrow \psi & & \\ 0 \longrightarrow & q'(M') & \xrightarrow{\iota'} & M' & \xrightarrow{\pi'} & \mathfrak{d}_{\text{st}}(A_R) & \longrightarrow 0 \end{array}$$

In the second row, $q'(M')$ corresponds under η to the image of the quadratic map q_R as in Lemma 3.11. This yields the homomorphism φ and the commutativity of the left hand square. Exactness of the second row is clear from the definition of $\mathfrak{d}_{\text{st}}(A_R)$. Finally, ψ is

the unique map making the right hand square commutative. We complete this diagram by adding the kernels and co-kernels of the vertical maps and obtain:

$$\begin{array}{ccccccc}
\text{Ker}(\varphi) & \longrightarrow & 0 & \longrightarrow & \text{Ker}(\psi) & & \\
\downarrow & & \downarrow & & \downarrow & & \\
q(M) \otimes R & \longrightarrow & M \otimes R & \longrightarrow & \mathfrak{d}_{\text{st}}(A) \otimes R & \longrightarrow & 0 \\
\downarrow \varphi & & \cong \downarrow \eta & & \downarrow \psi & & \\
0 \longrightarrow & q'(M') & \longrightarrow & M' & \longrightarrow & \mathfrak{d}_{\text{st}}(A_R) & \longrightarrow 0 \\
\downarrow & & \downarrow & & \downarrow & & \\
\text{Coker}(\varphi) & \longrightarrow & 0 & \longrightarrow & \text{Coker}(\psi) & &
\end{array}$$

Now the Snake Lemma [5, §1, No. 2, Prop. 2] yields an isomorphism $\text{Ker}(\psi) \cong \text{Coker}(\varphi)$. Lemma 3.11 implies that φ is surjective, so ψ is injective. But ψ is surjective as well, because η and π' are surjective. This establishes the isomorphisms $g(\tau) = \psi$ of (7), and (8) is easily verified.

(b) Suppose R is a flat k -algebra. We consider first the inner derivation functor F_{in} defined in 3.7. Here $M = W_{\text{in}}(A)$ is the kernel of the map $s: W(A) \rightarrow A$, and $M' = W_{\text{in}}(A_R)$ is similarly defined. Since R is flat over k , Lemma 1.7(a) yields an isomorphism $\eta: M \otimes R \rightarrow M'$ of Leibniz algebras over R . Now the argument in the proof of (a) can be repeated with \mathfrak{d}_{in} in place of \mathfrak{d}_{st} and yields an isomorphism $\mathfrak{d}_{\text{in}}(A) \otimes R \cong \mathfrak{d}_{\text{in}}(A_R)$. The proof of the other cases follows the same pattern. The details are left to the reader.

3.13. Remarks. The argumentation in the proof of Proposition 3.12(b) made use of flatness only to ensure that there is an isomorphism $\eta: M \otimes R \rightarrow M'$. This can be used to prove base change results for arbitrary R , by restricting the category \mathcal{A} . For example, let \mathcal{A} be the category octonion algebras (see 4.1 below for the definition), with morphisms unital homomorphisms of algebras, and consider associator derivations. Here $M = W_{\text{ass}}(A)$ is the kernel of the commutator map $s: \bigwedge^2 A \rightarrow A$. The linear span of all commutators $[a, b]$ in an octonion algebra A is precisely the kernel of the trace $t_A: A \rightarrow k$, and the trace is surjective. Hence $\text{Ker}(t_A) = \text{Im}(s)$ is a finitely generated and projective module (of rank 7). It follows that the exact sequence

$$0 \longrightarrow M \longrightarrow \bigwedge^2 A \xrightarrow{s} [A, A] \longrightarrow 0$$

splits and therefore remains exact (and split) upon tensoring with an arbitrary $R \in k\text{-alg}$. Hence the natural map $\eta: M \otimes R \rightarrow M' = \text{Ker}(s_R)$ is an isomorphism. It follows that the functor F_{ass} commutes with arbitrary base change for octonion algebras.

4. Octonion algebras: basic properties

In this section, we prepare the ground for describing derivations of octonion algebras over arbitrary commutative rings.

4.1. The concept of an octonion algebra. Following [17, 1.8], a non-associative algebra C over k is called an *octonion algebra* if it is finitely generated projective of rank 8 as a k -module, contains an identity element and admits a *norm*, i.e., a quadratic form $n_C: C \rightarrow k$ uniquely determined by the following two conditions:

(i) n_C is *non-singular*, so its induced symmetric bilinear form $n_C(x, y) = n_C(x + y) - n_C(x) - n_C(y)$ defines a linear isomorphism from the k -module C onto its dual C^* by $x \mapsto n_C(x, -)$.

(ii) n_C *permits composition*, i.e., the relation

$$n_C(xy) = n_C(x)n_C(y) \quad (1)$$

holds for all $x, y \in C$.

We then call $t_C = n_C(1_C, -)$ the *trace* of C . Since the rank of C is everywhere positive, $1_C \in C$ is a *unimodular vector* [14, 0.3], i.e., $k1_C$ is a free k -module of rank 1 and a direct summand of C (as a k -module).

Octonion algebras are alternative (but not associative) and invariant under base change. They also descend from faithfully flat base change: If $R \in k\text{-alg}$ is faithfully flat over k and C is a k -algebra such that C_R is an octonion algebra over R then C is an octonion algebra over k . This follows from faithfully flat descent and the fact that the norm and the unit element of an octonion algebra are uniquely determined.

By [16], given an octonion algebra C over k , the relations

$$n_C(1_C) = 1, \quad t_C(1_C) = 2, \quad (2)$$

$$x^2 - t_C(x)x + n_C(x)1_C = 0, \quad (3)$$

$$t_C(xy) = t_C(x)t_C(y) - n_C(x, y) \quad (4)$$

hold for all $x, y \in C$, and t_C is an *associative* linear form in the sense that it vanishes on all commutators and associators of the algebra. Moreover, the *conjugation* of C , i.e., the linear map $\iota_C: C \rightarrow C$, $x \mapsto \bar{x} := t_C(x)1_C - x$, is an algebra involution satisfying $x\bar{x} = n_C(x)1_C$, $x + \bar{x} = t_C(x)1_C$, and

$$xyx = n_C(x, \bar{y})x - n_C(x)\bar{y} \quad (5)$$

for all $x, y \in C$. In particular, x is invertible in C if and only if $n_C(x)$ is a unit in k , in which case $x^{-1} = n_C(x)^{-1}\bar{x}$. Recall that octonion algebras over fields are simple [23, Chap. 2, Lemma 3]. As a consequence, octonion algebras over rings share with Azumaya algebras the property that a *unital homomorphism* $f: C \rightarrow C'$ of octonion algebras is an isomorphism. Indeed, localizing if necessary, we may assume that k is a local ring, with residue field K . Then the kernel of the induced homomorphism $f_K: C_K \rightarrow C'_K$ is an ideal $\neq C_K$, hence $\{0\}$. Thus f_K is injective and therefore bijective, because both algebras have dimension 8. It follows that f is an isomorphism by [3, II, §3.2, Cor. of Prop. 6].

Remark. The same argument leads to the same conclusion in the more general setting of arbitrary unital non-associative k -algebras C, C' that are finitely generated projective of the same rank as k -modules and have C_K simple for all fields $K \in k\text{-alg}$.

We now proceed to describe particularly simple and useful examples of octonion algebras.

4.2. Zorn vector matrices and split octonions. There are various formally different but equivalent ways of defining an octonion algebra structure on the k -module

$$Z := \text{Zor}(k) := \begin{bmatrix} k & k^3 \\ k^3 & k \end{bmatrix}$$

of *Zorn vector matrices* over k , i.e., of 2×2 matrices with diagonal entries in k and off-diagonal ones in column space k^3 over k . The normalization chosen here is due to Zorn [24] and turns out to be the most convenient for our subsequent computations. Accordingly, we define

$$\begin{bmatrix} \alpha_1 & u \\ x & \alpha_2 \end{bmatrix} \begin{bmatrix} \beta_1 & v \\ y & \beta_2 \end{bmatrix} = \begin{bmatrix} \alpha_1\beta_1 - u^t y & \alpha_1 v + \beta_2 u + x \times y \\ \beta_1 x + \alpha_2 y + u \times v & -x^t v + \alpha_2 \beta_2 \end{bmatrix} \quad (6)$$

for $\alpha_i, \beta_i \in k$, ($i = 1, 2$), $u, v, x, y \in k^3$, where $u^t v$ and $u \times v$ stand for the ordinary scalar and vector product, respectively, of $u, v \in k^3$. Then $\text{Zor}(k)$ becomes an octonion algebra under the multiplication (6). Its unit element, norm, and trace are given by the formulas

$$1_Z = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad n_Z(a) = \alpha_1 \alpha_2 + u^t x, \quad t_Z(a) = \alpha_1 + \alpha_2$$

for $a = \begin{bmatrix} \alpha_1 & u \\ x & \alpha_2 \end{bmatrix} \in Z$. Note that this definition has the advantage of yielding complete symmetry in the indices 1,2 but is not consistent with the usual definition of matrix multiplication of 2×2 -matrices in the following sense: Let $u, x \in k^3$ such that $u^t x = 1$. Then $\begin{bmatrix} k & k \cdot u \\ k \cdot x & k \end{bmatrix}$ is a subalgebra of $\text{Zor}(k)$ isomorphic to $\text{Mat}_2(k)$ under the map $\begin{bmatrix} \alpha & \beta \cdot u \\ \gamma \cdot x & \delta \end{bmatrix} \mapsto \begin{pmatrix} \alpha & \beta \\ -\gamma & \delta \end{pmatrix}$, and $n_Z \begin{bmatrix} \alpha & \beta \cdot u \\ \gamma \cdot x & \delta \end{bmatrix} = \det \begin{pmatrix} \alpha & \beta \\ -\gamma & \delta \end{pmatrix}$. The square brackets (instead of the usual round brackets) serve to indicate this fact.

Let u_i ($i = 1, 2, 3$) be the standard basis of k^3 . It is evident that Z is free of rank eight as a k -module with basis

$$\mathbf{b}_s = (E_1, X_1, X_2, X_3; E_2, Y_1, Y_2, Y_3)$$

given by

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad X_i = \begin{bmatrix} 0 & u_i \\ 0 & 0 \end{bmatrix}, \quad Y_i = \begin{bmatrix} 0 & 0 \\ u_i & 0 \end{bmatrix}.$$

We call \mathbf{b}_s the *standard basis* of Z . It satisfies the following relations:

$$E_1 E_2 = E_2 E_1 = 0, \quad E_1^2 = E_1, \quad E_2^2 = E_2, \quad (7)$$

$$E_1 X_i = X_i = X_i E_2, \quad E_2 X_i = 0 = X_i E_1, \quad (8)$$

$$E_2 Y_i = Y_i = Y_i E_1, \quad E_1 Y_i = 0 = Y_i E_2, \quad (9)$$

$$X_i X_j = \text{sgn}(i, j) Y_l, \quad Y_i Y_j = \text{sgn}(i, j) X_l, \quad (10)$$

$$X_i Y_j = -\delta_{ij} E_1, \quad Y_i X_j = -\delta_{ij} E_2, \quad (11)$$

$$E_1 + E_2 = 1_Z. \quad (12)$$

Here $\text{sgn}(i, j)$ is zero for $i = j$ and equals the sign of the permutation (i, j, l) (sending 1 to i , 2 to j , 3 to l) if $i \neq j$ and l is the missing index.

Our next step consists in introducing twisted versions of Zorn vector matrices.

4.3. Reduced octonion algebras. An octonion algebra over k is said to be *reduced* if it is isomorphic to an algebra $\text{Zor}(M, \theta)$, defined as follows ([17, 3.2, 3.3])¹: Let M be a finitely generated projective module of rank 3 over k . Writing $M^* = \text{Hom}_k(M, k)$ for the dual of M and $\langle \cdot, \cdot \rangle: M^* \times M \rightarrow k$ for the natural pairing, we identify $\bigwedge^3 M^* = (\bigwedge^3 M)^*$ canonically by means of the formula

$$\langle \alpha_1 \wedge \alpha_2 \wedge \alpha_3, x_1 \wedge x_2 \wedge x_3 \rangle = \det(\langle \alpha_i, x_j \rangle)$$

for $\alpha_i \in M^*$, $x_j \in M$, $1 \leq i, j \leq 3$. Now suppose we are given a *volume element* of M , i.e., an isomorphism $\theta: \bigwedge^3 M \xrightarrow{\sim} k$ of k -modules (which may not exist but if it does is unique up to an invertible factor in k). Then $\theta^*: k = k^* \xrightarrow{\sim} \bigwedge^3 M^*$, the dual of θ , gives rise to the volume element θ^{*-1} of M^* , and we obtain two associated vector products

$$\times_\theta: M \times M \longrightarrow M^*, \quad \times_\theta: M^* \times M^* \longrightarrow M$$

by means of the formulas

$$\langle x \times_\theta y, z \rangle = \theta(x \wedge y \wedge z), \quad \langle \zeta, \xi \times_\theta \eta \rangle = \theta^{*-1}(\zeta \wedge \xi \wedge \eta) \quad (13)$$

for all $x, y, z \in M$, $\xi, \eta, \zeta \in M^*$. Note that both vector products are alternating and induce isomorphisms $\bigwedge^2 M \xrightarrow{\sim} M^*$, $\bigwedge^2 M^* \xrightarrow{\sim} M$. To simplify notations, we write \times instead of \times_θ whenever the context is clear. Furthermore, to make matters more symmetric, we identify $M \cong M^{**}$ canonically, put $M^+ := M$, $M^- := M^*$ and then have two dualizing bilinear forms $\langle \cdot, \cdot \rangle: M^\pm \times M^\mp \rightarrow k$ satisfying the relation $\langle x^+, x^- \rangle = \langle x^-, x^+ \rangle$ for all $x^\pm \in M^\pm$. Now the k -module

$$C := \text{Zor}(M, \theta) = \begin{bmatrix} k & M^+ \\ M^- & k \end{bmatrix}$$

becomes an octonion algebra over k under the multiplication

$$\begin{bmatrix} \alpha_1 & u \\ x & \alpha_2 \end{bmatrix} \begin{bmatrix} \beta_1 & v \\ y & \beta_2 \end{bmatrix} = \begin{bmatrix} \alpha_1 \beta_1 - \langle u, y \rangle & \alpha_1 v + \beta_2 u + x \times y \\ \beta_1 x + \alpha_2 y + u \times v & -\langle x, v \rangle + \alpha_2 \beta_2 \end{bmatrix} \quad (14)$$

for $\alpha_i, \beta_i \in k$ ($i = 1, 2$), $u, v \in M^+$, $x, y \in M^-$, whose unit element, norm, trace are given by

$$1_C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad n_C(a) = \alpha_1 \alpha_2 + \langle u, x \rangle, \quad t_C(a) = \alpha_1 + \alpha_2 \quad (15)$$

for $a = \begin{bmatrix} \alpha_1 & u \\ x & \alpha_2 \end{bmatrix} \in C$. If $M = k^3$ is free, then $C = \text{Zor}(k)$ is the *split* octonion algebra of ordinary Zorn vector matrices over k .

4.4. Elementary idempotents. We claim: *For an element e in an octonion algebra C over k to be an idempotent different from $0, 1$ in all scalar extensions (so $e^2 = e$ and $e_R \neq 0, 1_{C_R}$ for all $R \in k\text{-alg}$, $R \neq \{0\}$) it is necessary and sufficient that $t_C(e) = 1$*

¹We deviate from the terminology in [17], where these algebras are called split.

and $n_C(e) = 0$. The condition is clearly sufficient, by (2), (3). To prove necessity, we may assume that k is a local ring, hence, in particular, connected. Then $n_C(e)$, being an idempotent in k by (1), satisfies $n_C(e) = 0$ or $n_C(e) = 1$. In the latter case, e would be invertible, forcing the contradiction $e = 1_C$. Hence $n_C(e) = 0$, and (3) yields $e = e^2 = t_C(e)e$. Taking traces, we conclude that $t_C(e) \in k$ is an idempotent which cannot be zero since $e \neq 0$. Thus $t_C(e) = 1$.

Elements of C satisfying the equivalent conditions above are called *elementary idempotents*. If e is such and $e_1 := e$, $e_2 := 1_C - e = \bar{e}_1$, then e_2 is an elementary idempotent as well and (e_1, e_2) is a hyperbolic pair of the quadratic space (C, n_C) . Moreover, using (5) as well as (1) and its bi-linearizations, the Peirce components $C_{ij} := C_{ij}(e)$ ($i, j = 1, 2$) of C relative to e [20, III, §2] are easily seen to satisfy the relations ($i, j = 1, 2$, $i \neq j$)

$$C_{ii} = ke_i, \quad n_C(C_{ij}) = n_C(e_i, C_{12} + C_{21}) = t_C(C_{12} + C_{21}) = \{0\}. \quad (16)$$

Since e_i is a unimodular vector, $C_{ii} \cong k$ as k -algebras. Also, (16) implies that the k -modules C_{12} and C_{21} are dually paired by $n_C(-, -)$, so the decomposition $C = \bigoplus_{i,j \in \{1,2\}} C_{ij}$ together with $\text{rk } C = 8$ shows that they are both finitely generated projective of rank 3.

4.5. Schemes. In a slightly more general vein than 1.11, we follow [7] and view schemes over k as special covariant set-valued functors on $k\text{-alg}$. Then the affine scheme \mathbf{X} defined by a fixed k -algebra A is the functor $\mathbf{X}(R) = \text{Hom}_{k\text{-alg}}(A, R)$ ($R \in k\text{-alg}$); i.e., the affine schemes are precisely the representable functors. For example, if M is a k -module, we denote by $M_{\mathbf{a}}$ the functor defined by $M_{\mathbf{a}}(R) := M \otimes R$ for all $R \in k\text{-alg}$. If M is finitely generated and projective then $M_{\mathbf{a}}$ is an affine k -scheme represented by the symmetric algebra over the dual M^* of M . We say that a k -scheme \mathbf{X} is *faithful* if it has non-empty geometric fibres: $\mathbf{X}(K) \neq \emptyset$ for all algebraically closed fields $K \in k\text{-alg}$. In case \mathbf{X} is affine and represented by a finitely presented k -algebra A , this is equivalent to the canonical map $\text{Spec}(A) \rightarrow \text{Spec}(k)$ of the prime spectra being surjective.

Given any k -scheme \mathbf{X} , we will make use of the following facts:

- (i) ([8, (17.16.2)]) If \mathbf{X} is fppf (=flat, faithful and finitely presented), there exists an fppf extension R of k such that $\mathbf{X}(R) \neq \emptyset$.
- (ii) ([8, (17.16.3)]) If \mathbf{X} is smooth and faithful, we may choose R as in (i) to be even étale.
- (iii) ([8, (17.1.1), (17.3.1)], [7, I, §4, 4.6]) \mathbf{X} is smooth if and only if it is finitely presented and, for all $R \in k\text{-alg}$ and all ideals $I \subseteq R$ satisfying $I^2 = \{0\}$, the natural map $\mathbf{X}(R) \rightarrow \mathbf{X}(R/I)$ is surjective. A smooth scheme is flat.
- (iv) ([8, (17.7.3)]) If $R \in k\text{-alg}$ is faithfully flat, then for \mathbf{X} to be smooth over k it is necessary and sufficient that its base change \mathbf{X}_R from k to R be smooth over R .

4.6. Splittings and splitting bases. A *splitting* of an octonion algebra C over k is an isomorphism $f: Z = \text{Zor}(k) \rightarrow C$. We denote by $\text{Isom}(Z, C)$ the (possibly empty) set of splittings of C and define a functor $\mathbf{X} = \mathbf{Isom}(Z, C): k\text{-alg} \rightarrow \text{set}$ by

$$\mathbf{X}(R) = \text{Isom}(\text{Zor}(R), C_R) \quad (R \in k\text{-alg}).$$

Let $\mathbf{G} = \mathbf{Aut}(Z)$ be the automorphism group scheme of Z . If $\mathbf{X}(R) \neq \emptyset$ then it is immediately seen that the group $\mathbf{G}(R)$ acts simply transitively on the right on $\mathbf{X}(R)$ by composition.

A *splitting basis* of C is an octuple $\mathbf{b} = (e_1, x_1, x_2, x_3; e_2, y_1, y_2, y_3) \in C^8$ satisfying the relations (7)–(12), with upper case letters replaced by lower case ones. Thus by its definition, a splitting basis is not required to be a basis of the k -module C but in fact is, as will be seen now.

Given a splitting $f: Z \xrightarrow{\cong} C$ of C , it is clear that the image $f(\mathbf{b}_s)$ of the standard basis of Z is a splitting basis of C . We claim that this establishes a bijection between $\text{Isom}(Z, C)$ and the set of splitting bases of C . Indeed, since f is linear and \mathbf{b}_s is in particular a basis of Z as a k -module, f is uniquely determined by its values on \mathbf{b}_s so the map $f \mapsto f(\mathbf{b}_s)$ is injective. To prove surjectivity, let \mathbf{b} be a splitting basis of C . The defining relations (7)–(12) of a splitting basis say precisely that the linear map $f: Z \rightarrow C$ defined by $f(\mathbf{b}_s) = \mathbf{b}$ is a unital homomorphism of octonion algebras and therefore an isomorphism, as remarked in 4.1. In particular, \mathbf{b} is a basis of C as a k -module.

An essential step in the proof of the main result of this section is to show that \mathbf{X} is a smooth k -scheme. The proof will be facilitated by introducing the following concept.

4.7. Splitting data. Let C be an octonion algebra over k . A *splitting datum* for C is a quadruple $\mathbf{d} = (e, x_1, x_2, x_3) \in C^4$ satisfying the following conditions:

$$e \text{ is an elementary idempotent,} \quad (17)$$

$$\text{the } x_i \text{ belong to the Peirce space } C_{12}(e), \quad (18)$$

$$x_1(x_2x_3) = -e. \quad (19)$$

Let \mathbf{b}_s be the standard basis of $Z = \text{Zor}(k)$ as in 4.2. It is clear from 4.2 that $\mathbf{d}_s = (E_1, X_1, X_2, X_3)$ is a splitting datum of the split algebra Z , called the *standard splitting datum*.

4.8. Lemma. *Let C be an octonion algebra over k . Then the map $\phi: f \mapsto f(\mathbf{d}_s)$ is a bijection between $\text{Isom}(Z, C)$ and the set of splitting data of C .*

Proof. If $f: Z \rightarrow C$ is an isomorphism then it is clear that $f(\mathbf{d}_s)$ is a splitting datum of C , so the map ϕ is well-defined. To prove ϕ injective we have to show that an $f \in \text{Isom}(Z, C)$ is uniquely determined by its values on \mathbf{d}_s . Since f is a homomorphism of unital algebras, the relations (10) and (12) show that $f(Y_i) = f(X_jX_l) = f(X_j)f(X_l)$ (where (i, j, l) is a cyclic permutation of $(1, 2, 3)$), and $f(E_2) = f(1 - E_1) = 1 - f(E_1)$. Hence $f(\mathbf{d}_s)$ determines the values of f on \mathbf{b}_s and therefore f , by 4.6.

Again by 4.6, ϕ surjective means every splitting datum \mathbf{d} of C extends to a splitting basis \mathbf{b} . Thus let $\mathbf{d} = (e, x_1, x_2, x_3)$ be a splitting datum of C and define $\mathbf{b} = (e_1, x_1, x_2, x_3; e_2, y_1, y_2, y_3)$ by

$$e_1 := e, \quad e_2 := 1_C - e_1, \quad y_1 := x_2x_3, \quad y_2 := x_3x_1, \quad y_3 := x_1x_2.$$

We verify the relations (7)–(12) for \mathbf{b} . Here (7) and (12) are clear, and (8) holds by (18). The equations (9) just say $y_i \in C_{21}$, which follows from $C_{12}^2 \subseteq C_{21}$ (by the Peirce rules) and the definition of y_i .

Before continuing, we make the following remarks. Since e is an elementary idempotent by (17), we have $t_C(x) = n_C(x) = 0$ for $x \in C_{12} \cup C_{21}$ by (16). This implies $x^2 = t_C(x)x - n_C(x)1 = 0$, so the multiplication of C restricted to C_{12} and to C_{21} is alternating. Moreover:

The trilinear expressions $x(yz)$ and $(xy)z$ (where $x, y, z \in C_{12}$) are alternating. (20)

Indeed, their difference is the associator which is alternating, so it suffices to prove that $x(yz)$ is alternating. But this follows immediately from $xy^2 = 0 = x^2y = x(xy)$ and the fact that a multilinear map is alternating as soon as it vanishes when two adjacent arguments are equal.

The equations $x_i x_j = \text{sgn}(i, j) y_l$ of (10) now hold by definition of the y_i and the alternating character of the product. The latter also allows us to assume, in proving the second group of equations $y_i y_j = \text{sgn}(i, j) x_l$, that (i, j, l) is a cyclic permutation. Then the middle Moufang identity (cf. (2.4)), the alternating nature of the product together with (20) and (19) imply

$$y_i y_j = (-x_l x_j)(-x_i x_l) = (x_l(x_j x_i))x_l = -(x_l(x_2 x_3))x_l = e_1 x_l = x_l.$$

To prove the first group of relations $x_i y_j = -\delta_{ij} e_1$ of (11), write $y_j = x_l x_m$ where (j, l, m) is cyclic, so that $x_i y_j = x_i(x_l x_m)$. If $i = j$ this is $x_j(x_l x_m) = x_1(x_2 x_3) = -e_1$ by (19) and (20). If $i \neq j$ then either $i = l$ or $i = m$, and hence $x_i(x_l x_m) = 0$, again by (20).

The remaining equations $y_i x_j = -\delta_{ij} e_2$ follow by applying the involution and observing that $t_C(x) = 0$ implies $\bar{x} = -x$ for $x \in C_{12} + C_{21}$.

4.9. Torsors. Let \mathbf{X} be a k -scheme and \mathbf{G} a k -group scheme acting on \mathbf{X} on the right in a simply transitive manner; i.e., for all $R \in k\text{-alg}$ and all $x, y \in \mathbf{X}(R)$ there exists exactly one $g \in \mathbf{G}(R)$ such that $y = xg$. Note that $\mathbf{X}(R)$ may well be empty. Then \mathbf{X} is said to be a *torsor in the flat topology with structure group \mathbf{G}* if there exists a fppf $S \in k\text{-alg}$ such that $\mathbf{X}(S) \neq \emptyset$ [7, III, §4]. If S can be chosen in addition étale then \mathbf{X} is called a *torsor in the étale topology*. Fixing an element $x_0 \in \mathbf{X}(S)$, we have an isomorphism $\mathbf{G}_S \xrightarrow{\sim} \mathbf{X}_S$ by $g \mapsto x_0 g$. Consequently, by faithfully flat descent, properties of \mathbf{X} and \mathbf{G} correspond to each other. In particular, \mathbf{X} is smooth if and only if \mathbf{G} is smooth, cf. 4.5(iv).

4.10. Theorem. *Let C be an octonion algebra over k and $Z = \text{Zor}(k)$ the algebra of Zorn vector matrices. Then $\mathbf{X} = \mathbf{Isom}(Z, C)$ is an affine smooth torsor in the étale topology with structure group $\mathbf{G} = \mathbf{Aut}(Z)$.*

Proof. Let \mathbf{Y} be the functor assigning to $R \in k\text{-alg}$ the set of splitting data of C_R . The map ϕ of Lemma 4.8 is compatible with arbitrary base changes and thus induces an isomorphism $\phi: \mathbf{X} \rightarrow \mathbf{Y}$ of functors. The conditions (17)–(19) show that $\mathbf{Y} \subset C_{\mathbf{a}}^4$ is defined by finitely many polynomial equations, so \mathbf{Y} and therefore \mathbf{X} is an affine finitely presented k -scheme. Hence to prove smoothness of \mathbf{X} , we may use 4.5(iii), and have to show: If $R \in k\text{-alg}$ and $I \subset R$ is an ideal of square zero then every splitting datum of $C_{R/I}$ over R/I can be lifted to a splitting datum of C_R over R .

We may assume $R = k$, replacing C by C_R if necessary. Write $\pi: C \rightarrow C' := C/IC = C \otimes (k/I)$ for the canonical map and let $\mathbf{d}' = (e', x'_1, x'_2, x'_3)$ be a splitting datum of C' . Denote norm and trace of C and C' by n, t and n', t' , respectively. As $IC \subseteq C$ is a nil ideal, it is a standard fact that e can be lifted to an idempotent e of C . We have $n'(e') = 0$, $t'(e') = 1$ by 4.4, so we conclude from (1) that $n(e)$ is a nilpotent idempotent in k . Hence $n(e) = 0$ and $e = t(e)e$ by (3). Applying t , we obtain $t(e)^2 = t(e)$, and $t'(e') = 1$ shows $t(e) \equiv 1 \pmod{I}$, whence $t(e)$ is an invertible idempotent in k . Thus $t(e) = 1$, and we have shown that e is elementary.

Let C_{ij} and C'_{ij} denote the Peirce spaces of C and C' relative to e and e' , respectively. Since π is a surjective algebra homomorphism mapping e to e' , we have $\pi(C_{12}) = C'_{12}$. Hence the elements $x'_i \in C'_{12}$ can be lifted to elements $x_i \in C_{12}$ ($1 \leq i \leq 3$). Now $x'_1(x'_2x'_3) = -e'_1$ implies $x_1(x_2x_3) = (-1+\alpha)e_1$ for some $\alpha \in I$. As $(1-\alpha)^{-1} = 1+\alpha$ (recall that $\alpha^2 = 0$ since I squares to zero), we see that (19) holds for $\mathbf{d} = (e, x_1, x_2, (1+\alpha)x_3)$, so \mathbf{d} is the desired lift of \mathbf{d}' to a splitting datum of C .

As noted in 4.6, \mathbf{G} acts simply transitively on \mathbf{X} . Thus to prove that \mathbf{X} is a torsor in the étale topology, it remains to show that $\mathbf{X}(S) \neq \emptyset$ for some étale faithfully flat $S \in k\text{-alg}$. It is a standard fact that an octonion algebra over an algebraically closed field K is split [21, Thm. 1.8.1, 1.10(i)]. Hence $\mathbf{X}(K) \neq \emptyset$, so \mathbf{X} is a faithful k -scheme. Now the existence of S follows from (ii) of 4.5. \square

The following corollaries are now an immediate consequence of the theorem and 4.5.

4.11. Corollary. *For C to be an octonion algebra over k it is necessary and sufficient that C be a k -algebra and there exist a faithfully flat étale k -algebra R such that $C_R \cong \text{Zor}(R)$ is a split octonion algebra over R .*

4.12. Corollary. *Let C be an octonion algebra over k . Then $\mathbf{Aut}(C)$ is a smooth group scheme.*

5. Octonion algebras: derivations

We are now ready for our main result on derivations of octonion algebras.

5.1. Theorem. *Every derivation of an octonion algebra C over k is an associator derivation: $\text{AssDer}(C) = \text{Der}(C)$, so every derivation of C has the form*

$$\sum_{i=1}^m [L_{a_i}, R_{b_i}]$$

where $m \in \mathbb{N}$ and $a_i, b_i \in C$ ($1 \leq i \leq m$) satisfy the relation $\sum [a_i, b_i] = 0$.

If the base ring contains $\frac{1}{3}$, every inner derivation of C is standard (Prop. 2.7(f)), so Thm. 5.1 yields the following extension of Schafer's theorem ([19, Thm. 6], [20, Cor. 3.29] or, more generally, [1, Prop. 1]) to commutative rings.

5.2. Corollary. *If $\frac{1}{3} \in k$, then every derivation of an octonion algebra over k is standard, i.e., it is a sum of derivations $D_{u,v}$, $u, v \in C$. \square*

Remarks. (a) Without the hypothesis on k , e.g., over fields of characteristic 3, Cor. 5.2 is false; see [1] for details.

(b) Derivations of Azumaya algebras are always inner [12, III, Thm. 1.4, Thm. 5.1]. Thm. 5.1 may be regarded as an analogue of this result for octonion algebras.

Proof of Thm. 5.1, step 1. Our proof of Thm. 5.1 proceeds in two steps, the first one combining Cor. 4.11 with the fact that the full derivation algebra of C by 1.11 and Cor. 4.12 (resp. the ideal of associator derivations by Prop. 2.9) commutes with arbitrary (resp. flat) base change. Hence we may assume that C is split.

In the sequel, we will work under the less restrictive assumption that C be reduced, and we do so for two reasons. For one, the proof will sometimes become more natural in this slightly more general setting. For another, we will be able to derive a number of intermediate results of independent interest that retain their validity for reduced rather than just split octonions.

The second step of the proof will be preceded by a digression into graded modules and algebras.

5.3. Graded modules. Let Γ be a *finite* additive abelian group and $M = \bigoplus_{\gamma \in \Gamma} M_\gamma$ a Γ -graded k -module [4, II, §11, n° 2]. Since Γ is finite, $\text{End}_k(M)$ becomes a Γ -graded k -algebra whose γ -homogeneous component, $\text{End}_k(M)_\gamma$, $\gamma \in \Gamma$, consists of all graded homomorphisms $f: M \rightarrow M$ of degree γ , so f is k -linear and satisfies $f(M_\delta) \subseteq M_{\gamma+\delta}$ for all $\delta \in \Gamma$. Clearly, since $\text{End}_k(M)$ is a Γ -graded k -algebra, so is $\mathfrak{gl}(M)$. Moreover, if we are given a non-associative Γ -graded k -algebra structure A on M , one checks easily that $\text{Der}(A)$ is a graded subalgebra of $\mathfrak{gl}(A)$.

5.4. Example. Let C be a unital alternative k -algebra and $e \in C$ an idempotent. Writing $e_1 := e$, $e_2 := 1_C - e$, the multiplication rules for the Peirce components $C_{ij} = C_{ij}(e)$ ($i, j = 1, 2$) [20, III, §2] imply that

$$C = C_0 \oplus C_1 \oplus C_2, \quad C_0 := C_{00} \oplus C_{11}, \quad C_1 := C_{12}, \quad C_2 := C_{21}$$

gives a $\mathbb{Z}/3\mathbb{Z}$ -grading of C as a k -algebra, called the *e-grading* of C . We write

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_1 \oplus \mathfrak{g}_2$$

for the corresponding $\mathbb{Z}/3\mathbb{Z}$ -grading of the derivation algebra $\mathfrak{g} := \text{Der}(C)$ in the sense of 5.3 and call this the *e-grading* of \mathfrak{g} . Fixing $i, j \in \{1, 2\}$, $i \neq j$ and $u_{ij} \in C_{ij} = C_i$, it is straightforward to check, using (2.28), that the derivation

$$D_i(u_{ij}) := -D_{e_i, u_{ij}} = D_{e_j, u_{ij}} \in \text{StanDer}(C)$$

satisfies the relations

$$D_i(u_{ij})x_{ii} = x_{ii}u_{ij}, \quad (1)$$

$$D_i(u_{ij})x_{jj} = -u_{ij}x_{jj}, \quad (2)$$

$$D_i(u_{ij})x_{ij} = u_{ij}x_{ij}, \quad (3)$$

$$D_i(u_{ij})x_{ji} = -[u_{ij}, x_{ji}] \quad (4)$$

for all $x_{\lambda,\mu} \in C_{\lambda,\mu}$, $\lambda, \mu = 1, 2$.

5.5. Proposition. *Assumptions and notations being as in 5.4, the e -grading of $\mathfrak{g} = \text{Der}(C)$ is given by $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_1 \oplus \mathfrak{g}_2$, where*

$$\mathfrak{g}_0 = \{D \in \mathfrak{g} \mid De = 0\}, \quad \mathfrak{g}_i = \{D_i(u_{ij}) \mid u_{ij} \in C_{ij}\} \quad (\{i, j\} = \{1, 2\}).$$

Moreover, the maps $u_{ij} \mapsto D_i(u_{ij})$ are k -module isomorphisms $C_{ij} \cong \mathfrak{g}_i$ for $i \neq j$.

Proof. Setting $\mathfrak{g}'_0 := \{D \in \mathfrak{g} \mid De = 0\}$ and $\mathfrak{g}'_i := \{D_i(u_{ij}) \mid u_{ij} \in C_{ij}\}$ ($\{i, j\} = \{1, 2\}$), we first claim

$$\mathfrak{g} = \mathfrak{g}'_0 + \mathfrak{g}'_1 + \mathfrak{g}'_2. \quad (5)$$

Given $D \in \mathfrak{g}$, let $D(e) = u_{11} + u_{12} + u_{21} + u_{22}$ be the Peirce decomposition of $D(e)$. Then $D(e) = D(e^2) = D(e) \cdot e + e \cdot D(e)$, and comparing Peirce components yields $D(e) = u_{12} + u_{21}$. This implies $D_0 := D - D_1(u_{12}) + D_2(u_{21}) \in \mathfrak{g}'_0$ by (1), (2), which completes the proof of (5). But now, consulting (1)–(4) and the Peirce rules, we conclude $\mathfrak{g}'_i \subseteq \mathfrak{g}_i$ for $i = 0, 1, 2$, hence obtain equality by comparing (5) with the e -grading of \mathfrak{g} . Finally, $D_i(u_{ij}) = 0$ implies $D_i(u_{ij}) \cdot e_i = u_{ij} = 0$ by (1), proving the last statement. \square

The elements of $\mathfrak{g}_1, \mathfrak{g}_2$ are obviously standard derivations. But one can do better than that by deriving the following result, which is due to the referee.

5.6. Proposition. *Notations and assumptions being as in 5.4, let $i, j \in \{1, 2\}$ be distinct and $u_{ji}, v_{ji} \in C_{ji}$. Using the formalism of 2.4, the element*

$$u := 2e_i \wedge (u_{ji}v_{ji}) - u_{ji} \wedge v_{ji} \in \bigwedge^2 C \subseteq W(C) = C \oplus \bigwedge^2 C$$

satisfies $s(u) = 0$ and

$$D_i(u_{ji}v_{ji}) = \Delta_u \in \text{AssDer}(C). \quad (6)$$

Proof. From 2.4 and the Peirce rules we conclude $s(u) = 2e_i(u_{ji}v_{ji}) - 2(u_{ji}v_{ji})e_i - u_{ji}v_{ji} + v_{ji}u_{ji} = 2u_{ji}v_{ji} - 2u_{ji}v_{ji} = 0$. Hence the proposition will follow once we have shown

$$D_i(u_{ji}v_{ji})a = Da, \quad D := \Delta_u = 2[L_{e_i}, R_{u_{ji}v_{ji}}] - [L_{u_{ji}}, R_{v_{ji}}] \quad (7)$$

for $a = x_{lm} \in C_{lm}$, $l, m \in \{i, j\}$. The Peirce rules combine with (1) and the linearization of left alternativity (2.1) to yield

$$\begin{aligned} Dx_{ii} &= 2e_i[x_{ii}(u_{ji}v_{ji})] - 2(e_ix_{ii})(u_{ji}v_{ji}) - u_{ji}(x_{ii}v_{ji}) + (u_{ji}x_{ii})v_{ji} \\ &= (u_{ji}x_{ii} + x_{ii}u_{ji})v_{ji} = u_{ji}(x_{ii}v_{ji}) + x_{ii}(u_{ji}v_{ji}) \\ &= D_i(u_{ji}v_{ji})x_{ii}; \end{aligned}$$

hence (7) holds for $a = x_{ii}$. Similarly, invoking (2.2) and (2), we obtain

$$\begin{aligned} Dx_{jj} &= 2e_j[x_{jj}(u_{ji}v_{ji})] - 2(e_jx_{jj})(u_{ji}v_{ji}) - u_{ji}(x_{jj}v_{ji}) + (u_{ji}x_{jj})v_{ji} \\ &= -u_{ji}(x_{jj}v_{ji} + v_{ji}x_{jj}) = -(u_{ji}x_{jj})v_{ji} - (u_{ji}v_{ji})x_{jj} \\ &= D_i(u_{ji}v_{ji})x_{jj}; \end{aligned}$$

hence (7) holds for $a = x_{jj}$. But the Peirce rules also combine with (4) and the linearization of $s_{ji}(s_{ji}t_{ji}) = s_{ji}^2t_{ji} = 0 = t_{ji}s_{ji}^2 = (t_{ji}s_{ji})s_{ji}$ for $s_{ji}, t_{ji} \in C_{ji}$ to yield $Dx_{ji} = 2e_i[x_{ji}(u_{ji}v_{ji})] - 2(e_ix_{ji})(u_{ji}v_{ji}) - u_{ji}(x_{ji}v_{ji}) + (u_{ji}x_{ji})v_{ji} = x_{ji}(u_{ji}v_{ji}) - (u_{ji}v_{ji})x_{ji} = -[u_{ji}v_{ji}, x_{ji}] = D_i(u_{ji}v_{ji})x_{ji}$, hence (7) for $a = x_{ji}$ as well. Finally, for $a = x_{ij}$, we linearize (2.1), (2.2) and obtain

$$\begin{aligned} Dx_{ij} &= 2e_i[x_{ij}(u_{ji}v_{ji})] - 2(e_ix_{ij})(u_{ji}v_{ji}) - u_{ji}(x_{ij}v_{ji}) + (u_{ji}x_{ij})v_{ji} \\ &= -2x_{ij}(u_{ji}v_{ji}) - u_{ji}(x_{ij}v_{ji}) + (u_{ji}x_{ij})v_{ji} \\ &= -2(x_{ij}u_{ji})v_{ji} - 2(u_{ji}x_{ij})v_{ji} + 2u_{ji}(x_{ij}v_{ji}) - u_{ji}(x_{ij}v_{ji}) + (u_{ji}x_{ij})v_{ji} \\ &= u_{ji}(x_{ij}v_{ji}) - (u_{ji}x_{ij})v_{ji} = u_{ji}(v_{ji}x_{ij} + x_{ij}v_{ji}) - (u_{ji}x_{ij})v_{ji} = (u_{ji}v_{ji})x_{ij} \\ &= D_i(u_{ji}v_{ji})x_{ij} \end{aligned}$$

by (3), as desired. □

5.7. Corollary. *If $C_{ij} = C_{ji}^2$, then $\mathfrak{g}_i \subseteq \text{AssDer}(C)$.* □

5.8. Setting the stage. In order to continue with the proof of Thm. 5.1, we fix once and for all a reduced octonion algebra $C = \text{Zor}(M, \theta)$ over k as in 4.3. Our aim will be to describe in more detail the zero component of the derivation algebra $\mathfrak{g} = \text{Der}(C)$ relative to its e -grading, where e is one of the two standard idempotents

$$e = e_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \in C.$$

To this end, we first note

$$C_0 = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix} = ke_1 \oplus ke_2, \quad C_1 = C_{12} = \begin{bmatrix} 0 & M^+ \\ 0 & 0 \end{bmatrix}, \quad C_2 = C_{21} = \begin{bmatrix} 0 & 0 \\ M^- & 0 \end{bmatrix} \quad (8)$$

by 5.4. The properties of the vector product assembled in 4.3 ensure the relations

$$C_{ji}^2 = C_{ij} \quad (i, j \in \{1, 2\} \text{ distinct}). \quad (9)$$

As usual, we will write $\mathrm{SL}(M)$ for the special linear group of $M = M^+$, and $\mathbf{SL}(M)$ for the affine group scheme over k given by

$$\mathbf{SL}(M)(R) := \mathrm{SL}(M_R) \quad (R \in k\text{-alg}),$$

whose Lie algebra is

$$\mathfrak{sl}(M) := \{g \in \mathfrak{gl}(M) \mid \mathrm{tr}(g) = 0\}.$$

The centralizer in the sense of [7, II, §1, 3.4] of e in $\mathbf{G} := \mathbf{Aut}(C)$ will be denoted by \mathbf{G}_0 . Then \mathbf{G}_0 acts like the identity on C_0 in all extensions since $e_2 = 1 - e_1$ and \mathbf{G} fixes the unit element. \mathbf{G}_0 is a subgroup scheme of \mathbf{G} whose Lie algebra is the subalgebra \mathfrak{g}_0 of \mathfrak{g} described in Prop. 5.5.

For $g \in \mathrm{End}(M)$ let $g^* \in \mathrm{End}(M^*)$ be defined by $\langle g^*(\xi), x \rangle = \langle \xi, g(x) \rangle$, for all $x \in M$, $\xi \in M^*$. Then formula (4.13) implies

$$\langle g(x) \times g(y), g(z) \rangle = \theta(g(x) \wedge g(y) \wedge g(z)) = \det(g) \cdot \theta(x \wedge y \wedge z) = \det(g) \cdot \langle x \times y, z \rangle$$

for all $x, y, z \in M$. Now let $g \in \mathrm{GL}(M)$ and replace z by $g^{-1}(z)$. Then

$$\langle g(x) \times g(y), z \rangle = \det(g) \cdot \langle x \times y, g^{-1}(z) \rangle = \det(g) \cdot \langle g^{*-1}(x \times y), z \rangle$$

holds for all z , so

$$g(x) \times g(y) = \det(g) \cdot g^{*-1}(x \times y) \quad (x, y \in M, g \in \mathrm{GL}(M)). \quad (10)$$

The following theorem generalizes [10, Thm. 4] from fields to commutative rings. It says that the group scheme \mathbf{G}_0 defined above is of type A_2 , i.e., a twisted form of \mathbf{SL}_3 .

5.9. Theorem. *With the notations of 5.8, there is an isomorphism*

$$\Phi : \mathbf{SL}(M) \xrightarrow{\sim} \mathbf{G}_0$$

of group schemes where $\Phi(g)$ is given by

$$\Phi(g) \cdot \begin{bmatrix} \alpha_1 & u \\ x & \alpha_2 \end{bmatrix} := \begin{bmatrix} \alpha_1 & g(u) \\ g^{*-1}(x) & \alpha_2 \end{bmatrix}$$

for all $g \in \mathrm{SL}(M_R)$, $\alpha_1, \alpha_2 \in R$, $u \in M_R^+$, $x \in M_R^-$, $R \in k\text{-alg}$.

Proof. After a suitable base extension, it suffices to prove this for $R = k$. Clearly, Φ is a group monomorphism from $\mathrm{SL}(M)$ to $\mathrm{GL}(C)$. Given $g \in \mathrm{SL}(M)$, a straightforward verification shows, using (10) and (4.14), that $\Phi(g)$ as defined above is an automorphism of C fixing e_1 . It therefore remains to show that $\Phi: \mathrm{SL}(M) \rightarrow \mathbf{G}_0(k)$ is surjective, so let $\varphi \in \mathbf{G}_0(k)$ be an automorphism of C fixing e_1 . Then φ also fixes $e_2 = 1_C - e_1$ and stabilizes the C_{ij} . Hence (8) yields elements $g \in \mathrm{GL}(M^+)$, $h \in \mathrm{GL}(M^-)$ such that

$$\varphi \left(\begin{bmatrix} \alpha_1 & u \\ x & \alpha_2 \end{bmatrix} \right) = \begin{bmatrix} \alpha_1 & g(u) \\ h(x) & \alpha_2 \end{bmatrix} \quad (\alpha_i \in k, u \in M^+, x \in M^-).$$

Since φ leaves the norm of C invariant, (4.15) implies $h = g^{*-1}$, so it remains to prove $\det(g) = 1$. Now $\begin{bmatrix} 0 & u \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & v \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \alpha_1 & 0 \\ u \times v & 0 \end{bmatrix}$ by (4.14), so φ multiplicative implies $g^{*-1}(u \times v) = g(u) \times g(v)$. But the linear map $\wedge^2 M^+ \rightarrow M^-$ determined by the vector product is surjective, forcing $\det(g) = 1$ by comparison with (10), as desired. \square

Passing to the level of Lie algebras in Thm. 5.9, we obtain

5.10. Corollary. *Notations being as in 5.8, the map*

$$\mathfrak{sl}(M) \xrightarrow{\sim} \mathfrak{g}_0, \quad f \mapsto D_0(f)$$

determined by

$$D_0(f) \begin{bmatrix} \alpha_1 & u \\ x & \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 & f(u) \\ f^\vee(x) & 0 \end{bmatrix}$$

for $f \in \mathfrak{sl}(M)$, $\alpha_i \in k$ ($i = 1, 2$), $u \in M^+$, $x \in M^-$, where $f^\vee := -f^*$, is an isomorphism of Lie algebras. \square

Remark. Combining Cor. 5.10 with Prop. 5.5, we see that $\text{Der}(C)$ is a finitely generated projective module of rank 14, as it should be.

Proof of Thm. 5.1, step 2. We are now in a position to finish the proof of Thm. 5.1. By Prop. 5.5 and Cor. 5.7 combined with (9), it suffices to show that \mathfrak{g}_0 consists entirely of associator derivations, i.e.,

$$\mathfrak{g}_0 \subseteq \text{AssDer}(C). \quad (11)$$

Inspired by the proof of [1, Prop. 1], we do so by identifying $M^+ \otimes M^- = \text{End}_k(M^+)$ canonically via

$$(u \otimes x)(v) = \langle x, v \rangle u \quad (u, v \in M^+, x \in M^-). \quad (12)$$

Observe

$$\text{tr}(u \otimes x) = \langle u, x \rangle \quad (u \in M^+, x \in M^-), \quad (13)$$

let $u, v \in M^+$, $x \in M^-$ and put

$$a := \begin{bmatrix} 0 & u \\ 0 & 0 \end{bmatrix}, \quad b := \begin{bmatrix} 0 & 0 \\ x & 0 \end{bmatrix}, \quad c := \begin{bmatrix} 0 & v \\ 0 & 0 \end{bmatrix} \in C. \quad (14)$$

A direct computation, involving (2.12), (4.14) and (12) yields

$$[a, b] = \langle u, x \rangle (e_2 - e_1), \quad [L_a, R_b]e_1 = 0, \quad [L_a, R_b]c = \begin{bmatrix} 0 & h(v) \\ 0 & 0 \end{bmatrix}, \quad (15)$$

where

$$h = u \otimes x - \langle u, x \rangle \text{Id}_{M^+}. \quad (16)$$

By Cor. 5.10, every element of \mathfrak{g}_0 has the form $D_0(f)$ for some $f \in \mathfrak{sl}(M^+)$. Write $f = \sum u_i \otimes x_i$, $u_i \in M^+$, $x_i \in M^-$, and put

$$a_i := \begin{bmatrix} 0 & u_i \\ 0 & 0 \end{bmatrix}, \quad b_i := \begin{bmatrix} 0 & 0 \\ x_i & 0 \end{bmatrix} \in C.$$

Then (13)–(16) combine with Cor. 5.10 to show $\sum [a_i, b_i] = 0$ and $D_0(f) = \sum [L_{a_i}, R_{b_i}] \in \text{AssDer}(C)$. This completes the proof of (11). \square

Acknowledgements

We thank Stefan Gille and Angelika Welte for useful comments. We are particularly grateful to Erhard Neher, who made important suggestions for improving an earlier version of the paper. He also drew our attention to McCrimmon’s unpublished monograph [15] and was instrumental in making it accessible online to the international mathematical community. Kevin McCrimmon’s kind permission to include his results on derivations of alternative algebras [15, A5.2] in the present investigation is also gratefully acknowledged. The first-named author is indebted to the Department of Theoretical Physics at the University of Innsbruck for its hospitality during the preparation of this paper.

Finally, we are especially grateful to the referee, who made numerous useful suggestions for improving an earlier version of the paper. In particular, Prop. 5.6 below, which greatly simplifies our original proof of Thm. 5.1, is due to him. More important still, the referee’s penetrating analysis illuminating the intuitive background of inner derivations induced us to re-examine our own preconceptions of the subject, leading to the notion of derivation functor expounded in Section 3.

References

- [1] P. Alberca, A. Elduque, C. Martín, and F. J. Navarro, *On the Cartan-Jacobson theorem*, J. Algebra **250** (2002), 397–407.
- [2] E. Bannow, *Die Automorphismengruppen der Cayley-Zahlen*, Abh. Math. Sem. Univ. Hamburg **13** (1940), 240–256.
- [3] N. Bourbaki, *Elements of mathematics. Commutative algebra*, Hermann, Paris, 1972, Translated from the French.
- [4] ———, *Elements of mathematics. Algebra, Part I: Chapters 1-3*, Hermann, Paris, 1974, Translated from the French.
- [5] ———, *Algèbre homologique*, Masson, Paris, 1980.
- [6] É. Cartan, *Les groupes réels simples, finis et continus*, Ann. Sci. École Norm. Sup. **31** (1914), 263–355.
- [7] M. Demazure and P. Gabriel, *Groupes algébriques. Tome I*, Masson & Cie, Paris, 1970.
- [8] A. Grothendieck, *Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de schémas IV*, Inst. Hautes Études Sci. Publ. Math. (1967), no. 32, 361.
- [9] N. Jacobson, *Cayley numbers and normal simple Lie algebras of type G*, Duke Math. J. **5** (1939), 775–783.

- [10] ———, *Composition algebras and their automorphisms*, Rend. Circ. Mat. Palermo (2) **7** (1958), 55–80.
- [11] ———, *Exceptional Lie algebras*, Lecture Notes in Pure and Applied Mathematics, vol. 1, Marcel Dekker Inc., New York, 1971.
- [12] M.-A. Knus and M. Ojanguren, *Théorie de la descente et algèbres d’Azumaya*, Springer-Verlag, Berlin, 1974, Lecture Notes in Mathematics, Vol. 389.
- [13] J.-L. Loday, *Une version non commutative des algèbres de Lie: les algèbres de Leibniz*, Enseign. Math. (2) **39** (1993), no. 3-4, 269–293.
- [14] O. Loos, *Generically algebraic Jordan algebras over commutative rings*, J. Algebra **297** (2006), 474–529.
- [15] K. McCrimmon, *Alternative algebras*, <http://www.mathstat.uottawa.ca/~neher/Papers/alternative/>, 1980.
- [16] ———, *Nonassociative algebras with scalar involution*, Pacific J. Math. **116** (1985), 85–109.
- [17] H. P. Petersson, *Composition algebras over algebraic curves of genus zero*, Trans. Amer. Math. Soc. **337** (1993), 473–493.
- [18] N. Roby, *Lois polynomes et lois formelles en théorie des modules*, Ann. Sci. École Norm. Sup. (3) **80** (1963), 213–348.
- [19] R. D. Schafer, *Inner derivations of non-associative algebras*, Bull. Amer. Math. Soc. **55** (1949), 769–776.
- [20] ———, *An introduction to nonassociative algebras*, Pure and Applied Mathematics, Vol. 22, Academic Press, New York, 1966.
- [21] T. A. Springer and F. D. Veldkamp, *Octonions, Jordan algebras and exceptional groups*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2000.
- [22] M. L. Thakur, *Cayley algebra bundles on \mathbf{A}_K^2 revisited*, Comm. Algebra **23** (1995), 5119–5130.
- [23] K. A. Zhevlakov, A. M. Slin’ko, I. P. Shestakov, and A. I. Shirshov, *Rings that are nearly associative*, Pure and Applied Mathematics, Vol. 104, Academic Press, New York, 1982.
- [24] M. Zorn, *Alternativkörper und quadratische Systeme*, Abh. Math. Sem. Univ. Hamburg **9** (1933), 395–402.