PACIFIC JOURNAL OF MATHEMATICS
Vol. 117, No. 1, 1985

DERIVATIONS AND CAYLEY DERIVATIONS OF GENERALIZED CAYLEY-DICKSON ALGEBRAS

KEVIN MCCRIMMON

The Cayley-Dickson doubling process can be continued past the quaternions and octonions to obtain an infinite series of algebras of dimension 2^n. After $n = 3$ these algebras are no longer composition algebras. R. D. Schafer established the surprising result that the derivation algebras stop growing at $n = 3$. Schafer's proof assumed the scalars were a field of characteristic $\neq 2, 3$. In this paper we will give a different proof of his result which works for arbitrary rings of scalars, making use of the concept of a Cayley derivation.

Throughout, A denotes a unital nonassociative algebra over an arbitrary (unital, commutative, associative) ring of scalars Φ. We assume A has a *scalar involution* $*$ where all norms and traces are scalars,

$$xx^* = N(x)1, \qquad x + x^* = T(x)1, \tag{0.1}$$

$$N(x, y) = T(xy^*), \qquad T(xy) = T(yx). \tag{0.2}$$

If μ is a *cancellable* scalar ($\mu a = 0 \Rightarrow a = 0$) then we can construct a new algebra with scalar involution by the *Cayley-Dickson construction*

$$\mathbf{C}(A, \mu) = A \oplus Al, \tag{0.3}$$

$$(a + bl)(c + dl) = (ac + \mu b^*d) + (da + bc^*)l, \tag{0.4}$$

$$(a + bl)^* = a^* - bl. \tag{0.5}$$

The Cayley-Dickson process starts with $\Phi 1$ of dimension 1 and builds a $*$-extension $\Phi 1 + \Phi w$ of dimension 2 ($w + w^* = 1, 1 - 4N(w)$ cancellable — if $\frac{1}{2} \notin \Phi$ we must take this by fiat for the second stage), then a quaternion algebra of dimension 4, then an octonion algebra of dimension 8; the process continues to furnish algebras of dimension 2^n ($n \geq 4$), but these *generalized Cayley-Dickson algebras* are no longer alternative nor permit composition. Recall that the *commuter* Comm(A) consists of all elements commuting with A, the (left, middle, right) *nuclei* $N_i(A)$ consist of all elements associating with A, and the *center* $C(A)$ consists of all elements which commute and associate with A:

$$\begin{aligned}(0.6)\quad &\mathrm{Comm}(A) = \{c \mid [c, A] = 0\},\\ &N_l(A) = \{n \mid [n, A, A] = 0\}, \quad N_m(A) = \{n \mid [A, n, A] = 0\},\\ &N_r(A) = \{n \mid [A, A, n] = 0\}, \quad N(A) = N_l(A) \cap N_m(A) \cap N_r(A),\\ &C(A) = \mathrm{Comm}(A) \cap N(A),\end{aligned}$$

where the commutator is $[x, y] = xy - yx$ and the associator is $[x, y, z] = (xy)z - x(yz)$. For algebras with scalar involution we always have

$$N_l(A) = N_r(A),$$

and for the generalized Cayley-Dickson algebras of dimension 2^n we have

$$\begin{aligned}(0.7)\qquad &N(x, y) \text{ is nondegenerate} \quad \text{if } n \geq 1,\\ &\mathrm{Comm}(A) = C(A) = \Phi 1 \quad \text{if } n \geq 2,\\ &N(A) = \Phi 1 \quad \text{if } n \geq 3\end{aligned}$$

(see **[4]**, 6.8–9).

Although the linearized norm form $N(x, y) = T(xy^*)$ is by (0.7) usually nondegenerate, its radical will prove a nuisance later. The radical consists of the skew $*$-elements (cf. (1.1)),

$$(0.8)\qquad \mathrm{Rad}\, N(\cdot, \cdot) = \{z \mid z^* = -z,\ az = za^* \text{ for all } a \in A\}$$

since $N(z, 1) = 0$ iff $z^* = -z$, and then $N(z, a) = 0$ iff $za^* - az = 0$. Any such nuclear z's kill commutators,

$$(0.9)\qquad z \in \mathrm{Rad}\, N(\cdot, \cdot) \cap N(A) \Rightarrow [A, A]z = 0$$

since

$$[a, b]z = (ab)z - b(az) = z(b^*a^*) - b(za^*) = (zb^* - bz)a^* = 0$$

by nuclearity of z.

Any algebra with scalar involution has (generic) degree 2,

$$(0.10)\qquad x^2 - T(x)x + N(x)1 = 0,$$

$$(0.10')\qquad x \circ y - T(x)y - T(y)x + N(x, y)1 = 0 \quad (x \circ y = xy + yx).$$

A *derivation* of A into a unital bimodule M is a linear transformation D: $A \to M$ such that

$$(0.11)\qquad D(xy) = D(x)y + xD(y).$$

The *anti-derivations* of A into M are just the derivations of A into the *opposite bimodule* M^{op} (with $a \cdot_{\mathrm{op}} m = ma$, $m \cdot_{\mathrm{op}} a = am$), or from A^{op} into M,

$$(0.11')\qquad D(xy) = yD(x) + D(y)x.$$

We denote by $\mathrm{Der}(A, M)$ and $\mathrm{Der}^{\mathrm{op}}(A, M) = \mathrm{Der}(A, M^{\mathrm{op}}) = \mathrm{Der}(A^{\mathrm{op}}, M)$ the space of derivations and anti-derivations of A into M. In the special case of the regular bimodule $M = A$, we denote the derivations and anti-derivations of A by $\mathrm{Der}(A)$ and $\mathrm{Der}^{\mathrm{op}}(A)$.

Setting $x = y = 1$ in (0.11) or (0.11′) shows $D(1) = 2D(1)$, so

$$(0.12)\qquad D(1) = 0, \quad D(x^*) = -D(x) \qquad (D \in \mathrm{Der}^{\varepsilon}(A, M)).$$

If D is a derivation or antiderivation of a degree 2 algebra into itself, setting $x = y$ in (0.11) or (0.11′) shows

$$\begin{aligned} 0 &= D(x^2) - D(x)\circ x \\ &= D(T(x)x) - T(x)D(x) - T(D(x))x + N(D(x), x)1 \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{(by (0.10), (0.10′), (0.12))} \\ &= -T(D(x))x + N(D(x), x)1, \end{aligned}$$

so derivations are traceless and skew

$$(0.13)\qquad T(Dx) = N(Dx, x) = 0 \qquad (D \in \mathrm{Der}^{\varepsilon}(A), A \text{ rigid degree 2})$$

as long as A is *unitally faithful* and *rigid*

$$(0.14)\qquad \alpha A = 0 \Rightarrow \alpha = 0,$$

$$(0.15)\qquad F(x)x \in \Phi 1 \Rightarrow F = 0 \quad \text{if } F \text{ is a linear functional with } F(1) = 0.$$

(Note $F(x) = T(Dx)$ has $F(1) = 0$ by (0.12)). Assuming faithfulness (0.14) entails no loss of generality (pass to $\Phi/A^{\perp}$), and rigidity (0.15) holds in most reasonable cases (eg. if Φ has no nilpotent elements or A is unitally free as Φ-module ([4] 2.3)). From (0.3) we see $\mathbf{C}(A, \mu)$ is unitally faithful and rigid if A is, in particular all generalized Cayley-Dickson algebras are faithful and rigid.

We will formulate our results quite generally for general (not-necessarily-rigid) algebras with scalar involution and algebras obtained from them by the Cayley-Dickson construction. The proofs would simplify considerably if we restricted ourselves to the case of generalized Cayley Dickson algebras.

1. Cayley derivations. If A is an algebra with involution, a *$*$-module* is a bimodule M consisting entirely of *$*$-elements* m

$$(1.1)\qquad am = ma^* \quad \text{for all } a \in A.$$

These are precisely the bimodules which become skew $*$-bimodules ($(am)^* = m^*a^*, (ma)^* = a^*m^*$) under $m^* = -m$. There is a 1-1 correspondence between left, right, and $*$-modules for A. A *Cayley derivation* of A into a

$*$-module M is a linear map C such that

$$C(xy) = C(x)y^* + C(y)x, \tag{1.2}$$

and a *Cayley anti-derivation* is a Cayley derivation of A into M^{op}

$$C(xy) = C(x)y + C(y)x^*. \tag{1.2'}$$

We denote by $\mathrm{Cayder}(A)$ and $\mathrm{Cayder}^{\mathrm{op}}(A)$ the spaces of Cayley derivations and anti-derivations of A into itself (regarded as the regular right module). The archetypal example of a $*$-module is the *Cayley-Dickson bimodule* $\mathrm{Cay}(A) = Al$ as in (0.3); the importance of Cayley derivations is

$$C \in \mathrm{Cayder}(A) \quad \text{iff } D \in \mathrm{Der}(A, \mathrm{Cay}(A)) \quad \text{for } D(a) = C(a)l. \tag{1.3}$$

Again setting $x = y = 1$ in (1.2) or (1.2′) shows $C(1) = 2C(1)$,

$$C(1) = 0, \quad C(x^*) = -C(x) \qquad (C \in \mathrm{Cayder}^{\varepsilon}(A)). \tag{1.4}$$

Unlike the derivation case (0.13), a Cayley derivation need not be traceless. We say C is *tracial* if it has a *trace element* $c = t(C)$ such that

$$T(C(x)) = T(cx) = T(xc). \tag{1.5}$$

THE TRACE ELEMENT IS UNIQUELY DETERMINED ONLY IF $N(x, y) = T(xy^*)$ IS NONDEGENERATE; in general it is determined only up to an element of $\mathrm{Rad}\, N(\cdot, \cdot)$, which by (0.8) means up to a skew $*$-element. For tracial C any *conjugate*

$$\hat{C} = C - R_c \qquad (c = t(C), C \in t\,\mathrm{Cayder}(A)) \tag{1.6}$$

has traceless range by (1.5),

$$T(\hat{C}(x)) = 0, \qquad \hat{C}(x)^* = -\hat{C}(x) \tag{1.7}$$

(by (1.4) $\hat{C}$ is never a Cayley derivation unless $t(C) = 0$). Note that if $N(x, y)$ is nondegenerate over a field and A is finite-dimensional, then all Cayley derivations are uniquely tracial: the linear functional $T(C(x))$ must be represented by a vector c^*, $T(C(x)) = N(x, c^*) = T(xc)$. We denote by $t\,\mathrm{Cayder}(A)$ the space of tracial Cayley derivations of A, and by $n\,\mathrm{Cayder}(A)$ the space of Cayley derivations having a nuclear trace element $t(C) \in N(A)$.

1.8. EXAMPLE. The *standard skew Cayley map*

$$S(x) = x^* - x$$

is a (tracial) Cayley derivation of A iff $3[A, A] = 0$. Indeed,

$$\begin{aligned} S(xy) - S(x)y^* - S(y)x &= \{(xy)^* - xy\} - \{x^* - x\}y^* - \{y^* - y\}x \\ &= y^*x^* - xy - x^*y^* + xy^* - y^*x + yx \\ &= [y^*, x^*] - [y^*, x] + [y, x] = 3[y, x] \end{aligned}$$

vanishes for all x, y iff $3[A, A] = 0$. We call it skew because its range is skew, so it has trace element 0: $T(S(x)) = T(x^* - x) = 0$. If A is commutative then S is both a Cayley derivation and antiderivation, and if A has characteristic 2 then S is scalar-valued, $S(x) = x^* + x = T(x)1 \in \Phi 1$. □

1.9. PROPOSITION. *If A with scalar involution $*$ either has* (i) *dimension ≥ 3, or is unitally rigid with* (ii) $\mathrm{Comm}(A) = \Phi 1$ *or* (iii) $A = \Phi 1 + [A, A]$, *then it admits no scalar-valued Cayley derivations or anti-derivations*: $C(A) \subset \Phi 1 \Rightarrow C = 0$.

Proof. If $C(a) = F(a)1$ for a linear functional F satisfies (1.2) then $F(1) = 0$, $F(ab)1 = F(a)b^* + F(b)a$. In case (i), if $1, a, b$ are independent we get $F(a) = F(b) = 0$, so $F = C = 0$. Applying $[a, \cdot]$ gives $0 = [F(a)a, b^*]$, $F(a)a \in \mathrm{Comm}(A)$, so in case (ii) $F = C = 0$ by unital rigidity (0.15). In case (iii) we apply $T(\cdot)$ to see $2F(ab) = F(a)T(b) + F(b)T(a) = F(T(b)a + T(a)b) = (a \circ b + N(a, b)1)$ (by (0.10′)) $= F(ab + ba)$, so $F([a, b]) = 0$ and F vanishes on $[A, A]$ as well as $\Phi 1$, hence on $\Phi 1 + [A, A] = A$ and again $F = C = 0$. A similar argument applies to antiderivations. □

1.10. PROPOSITION. *If $3\Phi = 0$ and C is a Cayley derivation with $(C - \gamma)(A) \subset \Phi 1$ for some γ, then $C = \gamma S$ is a multiple of the standard skew Cayley map if either* (i) *A has dimension ≥ 3, or is unitally rigid with* (ii) $\mathrm{Comm}(A) = \Phi 1$ *or* (iii) $A = \Phi 1 + [A, A]$.

Proof. The condition (1.2) for $C(x) = \gamma x - F(x)1$ (F a linear functional with $F(1) = \gamma$) becomes $\gamma xy - F(xy)1 = \gamma xy^* - F(x)y^* + \gamma yx - F(y)x$, i.e. $0 = \gamma\{T(y)x - xy\} - F(x)\{T(y)1 - y\} + \gamma\{x \circ y - xy\} - F(y)x - \gamma xy + F(xy)1 = \{2\gamma T(y) - F(y)\}x + \{\gamma T(y) + F(x)\}y + \{-\gamma T(xy^*) + F(xy) - F(x)T(y)\}1 - 3\gamma xy$ (by (0.10′)), so using $3\gamma = 0$ we see

$$(1.11) \quad -H(y)x + H(x)y - H(xy^*)1 = 0 \quad (H(x) = \gamma T(x) + F(x)).$$

Whenever $H = 0$ we have $F(x) = -\gamma T(x)$, $C(x) = \gamma x + \gamma(x + x^*) = \gamma(x^* + 2x) = \gamma(x^* - x) = \gamma S(x)$, and $C = \gamma S$. (i) If $\dim A \geq 3$ we take $x, y, 1$ independent in (1.11) to see $H = 0$. (ii) We commute (1.11) with x to get $H(x)[x, y] = 0$, $H(x)x \in \mathrm{Comm}(A) = \Phi 1$ with $H(1) = 3\gamma = 0$, so if A is unitally rigid as in (0.15) we see $H = 0$. (iii) Since $\frac{1}{2} = -1$ in characteristic 3 we can write $A = \Phi 1 \oplus A_0$ $(T(A_0) = 0)$, so $-H(y_0)x_0 + H(x_0)y_0 = -H(x_0y_0)1$ in (1.11) implies $H(x_0y_0) = 0$, $H(x_0)y_0 = H(y_0)x_0$ for x_0, y_0 in A_0, in particular

$$H(A)[A, A] = H(A_0)[A_0, A_0] = H([A_0, A_0])A_0 = 0$$

since $[A, A] \subset A_0$ by (0.2), so $H([A, A]) = H([A_0, A_0]) = 0$ and already $H(1) = 1$, therefore $H(A) = 0$ and $H = 0$. □

1.12 EXAMPLE. If A is associative with scalar involution, then

$$C(x) = \sum a_i x b_i + T(x)d$$

is a Cayley derivation if $d = \Sigma a_i b_i^*$ and $0 = \Sigma a_i(b_i + 2b_i^*)$, in which case C has trace element $t(C) = \Sigma b_i a_i + T(d)1$. As a special case, if $(ab)^* = ab^*$ (eg. if b and ab are skew) then

$$C(x) = [a, x]b \qquad ((ab)^* = ab^*)$$

is a Cayley derivation with trace element $t(C) = [b, a]$ Indeed, if $\Delta F(x, y) = F(xy) - F(x)y^* - F(y)x$ measures how far F is from being a Cayley derivation, then for $E_{a,b}(x) = axb$ and $T_d(x) = T(x)d$ we have

$$\Delta E_{a,b}(x, y) = ab^*[x, y] - abxy^*, \quad \Delta T_d(x, y) = -d[x, y] - 2dxy^*,$$

so $C = \Sigma E_{a_i,b_i} + T_d$ has

$$\Delta C(x, y) = \left\{\sum a_i b_i^* - d\right\}[x, y] - \left\{\sum a_i b_i + 2d\right\}xy^*. \qquad \square$$

2. Derivations of $\mathbf{C}(A, \mu)$. In this section we show how the derivations of $\mathbf{C}(A, \mu)$ are built out of derivations, Cayley derivations, and skew nuclear elements of A. An immediate calculation from the definition (0.11) of derivation and the definition (0.4) of the product on $\mathbf{C}$ shows that every $*$-derivation of A $(D(a^*) = D(a)^*)$ extends to one of $\mathbf{C}$. The $*$-condition is just that $D(a)^* = -D(a)$, i.e. that D be traceless $T(D(a)) = 0$. We noted in (0.13) that all derivations are traceless in the unitally rigid case. Our calculations could be simplified if we assumed unital rigidity.

2.1 LEMMA. *A map $D_+(a + bl) = D_{11}(a) + D_{22}(b)l$ is a derivation of $\mathbf{C}(A, \mu)$ iff*

(i) *$D_{11} = D_0$ is a traceless derivation of A*

(ii) *$D_{22} = D_0 + L_{d_0}$ for a skew element d_0 in the nucleus of A.*

Proof. D_+ restricts to a derivation of A into $A + Al$, hence its projection D_{11} into the submodule of A must be a derivation D_0 of A.

Then the derivation condition $D(x_1x_2) = D(x_1)x_2 + x_1D(x_2)$ for $x_i = a_i + b_il$ reduces to

$$\begin{aligned} &D_0(a_1a_2 + \mu b_2^*b_1) + D_{22}(b_1a_2^* + b_2a_1)l \\ &\quad = \{D_0(a_1)a_2 + a_1D_0(a_2)\} + \mu\{b_2^*D_{22}(b_1) + D_{22}(b_2)^*b_1\} \\ &\qquad + \{b_2D_0(a_1) + D_{22}(b_2)a_1\}l + \{D_{22}(b_1)a_2^* + b_1D_0(a_2)^*\}l, \end{aligned}$$

i.e.

(1) $$D_{22}(b_1a_2^*) = D_{22}(b_1)a_2^* + b_1D_0(a_2)^*$$

(2) $$D_{22}(b_2a_1) = D_{22}(b_2)a_1 + b_2D_0(a_1)$$

(3) $$D_0(b_2^*b_1) = b_2^*D_{22}(b_1) + D_{22}(b_2)^*b_1.$$

Setting $d_0 = D_{22}(1)$, we see by (2) that $D_{22}(a) = d_0a + D_0(a)$, so $D_{22} = D_0 + L_{d_0}$ as in (ii), and (2) reduces to left nuclearity $d_0(ba) = (d_0b)a$ of d_0. (1) + (2) reduces to $D_{22}(bT(a)) = D_{22}(b)T(a) + bT(D_0(a))$, which is just tracelessness of D_0 as in (i). Hence (3) reduces to $0 = b_2^*(d_0b_1) + (b_2^*d_0^*)b_1$, which for $b_1 = b_2 = 1$ yields skewness $d_0 + d_0^* = 0$, and therefore (3) is middle nuclearity $0 = -[b_2^*, d_0, b_1]$ of d_0. □

2.2. Lemma. *$D\_(a + bl) = D_{12}(b) + D_{21}(a)l$ is a derivation of $C(A, \mu)$ iff*

(i) *$D_{21} = C_0$ is a Cayley derivation of A such that μC_0 has a skew nuclear trace element c_0*

(ii) $D_{12} = \widehat{\mu C_0} = \mu C_0 - R_{c_0}$.

Proof. If $D\_$ is a derivation of C then its restriction to A and projection into Al gives a derivation of A into Al, so by (1.3) D_{21} is a Cayley derivation C_0. The derivation condition (0.11) for $x_i = a_i + b_il$ becomes

$$\begin{aligned} &D_{12}(b_1a_2^* + b_2a_1) + C_0(a_1a_2 + \mu b_2^*b_1)l \\ &\quad = \{D_{12}(b_1)a_2 + \mu b_2^*C_0(a_1)\} + \{C_0(a_1)a_2^* + b_2D_{12}(b_1)\}l \\ &\qquad + \{a_1D_{12}(b_2) + \mu C_0(a_2)^*b_1\} + \{C_0(a_2)a_1 + b_1D_{12}(b_2)^*\}l, \end{aligned}$$

i.e.,

(1) $$D_{12}(b_2a_1) = \mu b_2^*C_0(a_1) + a_1D_{12}(b_2)$$

(2) $$D_{12}(b_1a_2^*) = D_{12}(b_1)a_2 + \mu C_0(a_2)^*b_1$$

(3) $$\mu C_0(b_2^*b_1) = b_2D_{12}(b_1) + b_1D_{12}(b_2)^*.$$

If we let $c_0 = -D_{12}(1)$, $C = \mu C_0$ then $b_2 = 1$ in (1) yields $D_{12}(a) = C(a) - ac_0$, so $D_{12} = C - R_{c_0}$ as in (ii). Setting $a_2 = b_2 = 1$ in (3) yields $c_0 + c_0^* = 0$ (using (1.4)), so $b_1 = 1$ in (2) yields

$$\begin{aligned} 0 &= C(a^*) - a^*c_0 + c_0a - C(a)^* \\ &= -C(a) + a^*c_0^* + c_0a - C(a)^* \quad \text{(by (1.4))} \\ &= T(c_0a - C(a)), \end{aligned}$$

so C is tracial with trace element c_0 as in (1.5) and conjugate $\hat{C} = D_{12}$ as in (1.6). Thus (2) becomes

$$0 = C(ba^*) + (ba^*)c_0^* - C(b)a + (bc_0)a - T(c_0a)b - C(a^*)b$$

$$\text{(using skewness of } c_0, (1.5), (1.4))$$

$$\begin{aligned} &= [b, a^*, c_0^*] + b(a^*c_0^*) + [b, c_0, a] + b(c_0a) - bT(c_0a) \\ &= [b, a, c_0] + [b, c_0, a], \end{aligned}$$

(1) + (2) becomes

$$\begin{aligned} 0 &= T(a)\hat{C}(b) - a \circ \hat{C}(b) - N(C(a), b)1 \\ &= N(a, \hat{C}(b)) - N(C(a), b) \quad \text{(by (0.10'), (1.7))} \\ &= -T([a, b, c_0]) \end{aligned}$$

(see (2.10) below), (1) + (3) becomes

$$\begin{aligned} 0 &= \{\hat{C}(ba) - b^*C(a) - a\hat{C}(b)\} + \{C(ba^*) - b\hat{C}(a) - a\hat{C}(b)^*\} \\ &= C(T(b)a) - (ba)c_0 - T(b)C(a) + b(ac_0) \quad \text{(by (1.7))} \\ &= -[b, a, c_0]. \end{aligned}$$

Therefore (1)–(3) hold iff $D_{12} = \hat{C}$ where the trace c_0 of C is right = left nuclear and middle nuclear, i.e. c_0 is skew nuclear. □

2.3. COROLLARY. *$D(a + bl) = bz_0$ is a derivation of $\mathbf{C}(A, \mu)$ iff $z_0 \in$ Rad $N(\cdot, \cdot) \cap N(A)$ is radical in the nucleus, i.e. a skew nuclear $*$-element.*

Proof. This is the case $C_0 = C = 0$ of the Lemma, $c_0 = z_0$ is any skew nuclear trace element for 0: $T(z_0x) = 0$ for all x, i.e. (by (0.2)) $N(z_0, A) = 0$ and we apply the characterization (0.8). □

2.4. DERIVATION THEOREM. *If A is an algebra with scalar involution and μ a cancellable scalar, then the derivations of $\mathbf{C}(A, \mu) = A \oplus Al$ are precisely all*

$$(2.5) \qquad D = \begin{pmatrix} D_0 & C_0 - R_{c_0} - R_{z_0} \\ C_0 & D_0 + L_{d_0} \end{pmatrix} = \tilde{D}_0 + \tilde{C}_0 + \tilde{d}_0 + \tilde{z}_0$$

where

(i) D_0 *is a traceless derivation of* A

(ii) $d_0 \in N_0(A)$ *is skew nuclear*

(iii) C_0 *is a Cayley derivation of* A *such that* μC_0 *has skew nuclear trace element* $c_0 \in N_0(A)$

(iv) $z_0 \in N_{\text{rad}}(A)$ *is radical in the nucleus (skew nuclear* $*$*-element).*

We have the Schafer decomposition

$$\mathrm{Der}(\mathbf{C}(A,\mu)) = \mathscr{D}_0 \oplus \mathscr{N}_0 \oplus \mathscr{C}_0 \oplus \mathscr{Z}_0 \tag{2.6}$$

where

$$D_0 \to \tilde{D}_0 = \begin{pmatrix} D_0 & 0 \\ 0 & D_0 \end{pmatrix}$$

is a bijection of $\mathrm{Der}_0(A)$ onto $\mathscr{D}_0$,

$$d_0 \to \tilde{d}_0 = \begin{pmatrix} 0 & 0 \\ 0 & L_{d_0} \end{pmatrix}$$

is a bijection of $N_0(A)$ onto $\mathscr{N}_0$,

$$C_0 + z_0 \to \tilde{C}_0 + \tilde{z}_0 = \begin{pmatrix} 0 & C_0 - R_{c_0} - R_{z_0} \\ C_0 & 0 \end{pmatrix}$$

is a bijection of $\mathrm{Cayder}_\mu(A) + N_{\text{rad}}(A)$ onto $\mathscr{C}_0 \oplus \mathscr{Z}_0$, with multiplication rules

$$\begin{aligned}
&[\tilde{D}_0, \tilde{D}_0'] = \left[\widetilde{D_0, D_0'}\right] \in \mathscr{D}_0, && [\tilde{d}_0, \tilde{d}_0'] = \left[\widetilde{d_0, d_0'}\right] \in \mathscr{N}_0,\\
&[\tilde{D}_0, \tilde{C}_0] = \left[\widetilde{D_0, C_0}\right] \in \mathscr{C}_0, && [\tilde{d}_0, \tilde{C}_0] = \widetilde{L_{d_0}C_0} \in \mathscr{C}_0,\\
&[\tilde{D}_0, \tilde{d}_0] = \widetilde{D_0(d_0)} \in \mathscr{N}_0, && [\tilde{d}_0, \tilde{z}_0] = \widetilde{-d_0 z_0} \in \mathscr{Z}_0,
\end{aligned} \tag{2.7}$$

$$[\tilde{C}_0, \tilde{C}_0'] = \tilde{D}_0 + \tilde{d}_0 \qquad (D_0 = \widetilde{\mu C_0'}C_0' - \widetilde{\mu C_0'}C_0 \in \mathrm{Der}(A),$$
$$d_0 = C_0'(c_0) - C_0(c_0') \in N_0(A))$$

$$[\tilde{C}_0, \tilde{z}_0] = \tilde{D}_0 + \tilde{d}_0 \qquad \big(D_0 = -R_{z_0}C_0 \in \mathrm{Der}(A),\, d_0 = C_0(z_0) \in N_0(A)\big).$$

Proof. $\mathbf{C}(A,\mu) = A \oplus Al$ is a $\mathbf{Z}_2$-graded algebra, and

$$D = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}$$

is a derivation iff its even and odd parts

$$D_+ = \begin{pmatrix} D_{11} & 0 \\ 0 & D_{22} \end{pmatrix} \qquad (D_+(a + bl) = D_{11}(a) + D_{22}(b)l)$$

and

$$D_- = \begin{pmatrix} 0 & D_{12} \\ D_{21} & 0 \end{pmatrix} \qquad (D_-(a + bl) = D_{12}(b) + D_{21}(a)l)$$

are derivations. Here D_+ is a derivation iff $D_+ = \tilde{D}_0 + \tilde{d}_0$ as in (i), (ii) by Lemma 2.1, and D_- is a derivation iff $D_- = \tilde{C}_0 + \tilde{z}_0$ as in (iii), (iv) by Lemma 2.2 (any two nuclear trace elements for μC_0 differ by a nuclear radical element z_0) and Corollary 2.3. The multiplication rules (2.7) follow from direct matrix calculation, noting that if

$$D = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}$$

then $d_0 = D_{22}(1)$, $c_0 + z_0 = D_{12}(1)$, $D_0 = D_{11}$, $C_0 = D_{21}$ in (2.5) (so, for example ,

$$[\tilde{D}_0, \tilde{z}_0] = \begin{pmatrix} 0 & [D_0, R_{z_0}] \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & R_{D_0(z_0)} \\ 0 & 0 \end{pmatrix}$$

must have the form

$$\begin{pmatrix} 0 & R_{w_0} \\ 0 & 0 \end{pmatrix} \quad \text{in } \mathscr{Z}_0,$$

so $w_0 = D_0(z_0) \in N_{\text{rad}}(A)$). □

We call 2.6 the *Schafer decomposition* of Der(C(A, μ)) since it was first noticed by R. D. Schafer [5, p. 66] for the case when A is a quaternion algebra and C(A, μ) a Cayley algebra, and was used to analyze the Lie algebra Der(C) of type G_2.

When A has no Cayley derivations, C(A, μ) has essentially the same derivations as A.

2.8. COROLLARY. *If A has $N(A) = \Phi 1$ and $N(x, y)$ nondegenerate, then*

(i) *when A has no 2 or 3-torsion and Cayder(A) = 0 we have*

$$\mathrm{Der}(\mathbf{C}(A, \mu)) = \widetilde{\mathrm{Der}_0(A)}$$

(ii) *when* $2A = 0$ *and* $\mathit{Cayder}(A) = 0$ *we have*

$$\mathrm{Der}(\mathbf{C}(A, \mu)) = \widetilde{\mathrm{Der}_0(A)} \boxplus \Phi Z \quad (Z(a + bl) = bl \text{ central})$$

(iii) *when* $3A = 0$ *and* $\mathit{Cayder}(A) = \Phi S$ *we have*

$$\mathrm{Der}(\mathbf{C}(A, \mu)) = \widetilde{\mathrm{Der}_0(A)} \boxplus \Phi W$$
$$(W(a + bl) = \mu S(b) + S(a)l \text{ central}, S(a) = a^* - a).$$

Proof. We know any $D \in \mathrm{Der}(\mathbf{C}(A, \mu))$ has the form $\tilde{D}_0 + \tilde{d}_0 + \tilde{C}_0 + \tilde{z}_0$; if $N(x, y)$ is nondegenerate then $N_{\mathrm{rad}}(A) = 0$, $\tilde{z}_0 = 0$; if $N(A) = \Phi 1$ then $d_0 \in \Phi 1$ is skew iff $2d_0 = 0$, so $d_0 = 0$ when A has no 2-torsion and $d_0 = \delta 1$ is arbitrary if $2A = 0$, with $[\tilde{D}_0, \tilde{d}_0] = \widetilde{D_0(d_0)} = \widetilde{D_0(1)} = 0$ by (2.7), (0.12). By hypothesis $C_0 = 0$ in the first two cases and $C_0 = \gamma S$ in case (iii), with $[\tilde{D}_0, \tilde{S}] = \widetilde{[D_0, S]} = 0$ by (2.7) and $[D_0, S](a) = D_0(a^* - a) - \{D_0(a)^* - D_0(a)\} = 0$ if D_0 is traceless. □

2.9 Remark. As a consequence of the multiplication rules (2.7) we immediately obtain

(i) If $D_0 \in \mathrm{Der}_0(A)$, $C_0 \in n\,\mathrm{Cayder}(A)$ then $[D_0, C_0] \in n\,\mathrm{Cayder}(A)$ with $\widetilde{[D_0, C_0]} = [D_0, \hat{C}_0]$, $t([D_0, C_0]) = D_0(t(C_0))$;

(ii) If $D_0 \in \mathrm{Der}_0(A)$, $d_0 \in N_0(A)$, $z_0 \in N_{\mathrm{rad}}(A)$ then $D_0(d_0) \in N_0(A)$, $D_0(z_0) \in N_{\mathrm{rad}}(A)$;

(iii) If $d \in N_0(A)$, $C \in n\,\mathrm{Cayder}(A)$ then $L_dC \in n\,\mathrm{Cayder}(A)$ with $\widetilde{L_dC} = -\hat{C}L_d$, $t(L_dC) = \hat{C}(d) = C(d) - dt(C) \in N_0(A)$;

(iv) If $d \in N_0(A)$, $z \in N_{\mathrm{rad}}(A)$ then $R_zL_d = R_{dz}$ for $dz \in N_{\mathrm{rad}}(A)$;

(v) If C, $C' \in n\,\mathrm{Cayder}(A)$ then $D = \hat{C}C' - \hat{C}'C \in \mathrm{Der}(A)$ and $C\hat{C}' - C'\hat{C} = D + L_d$ for $d = C'(t(C)) - C(t(C')) \in N_0(A)$;

(vi) If $C \in n\,\mathrm{Cayder}(A)$, $z \in N_{\mathrm{rad}}(A)$ then $D = R_zC \in \mathrm{Der}(A)$ with $R_zC + CR_z = L_d$ for $d = C(z) \in N_0(A)$.

These of course can all be proven directly from the definitions, though at the expense of considerable effort. Direct calculation often yields slightly stronger statements, as we will now indicate. For computing trace elements it will be convenient to note that $\hat{C}$ is the adjoint of C when $t(C)$ is nuclear; more generally

$$N(Cx, y) - N(x, \hat{C}(y)) = T([x, y, t(C)]) \quad (C \in t\,\mathrm{Cayder}(A)) \tag{2.10}$$

since

$$\begin{aligned} &T(C(x)y^*) - T(C(y)x^*) + T(x^*(yc)) \quad (\text{using } (0.2), c = t(C)) \\ &\quad = T((x^*y)c - [x^*, y, c] - C(x^*y)) \quad (\text{by } (1.4)) \\ &\quad = T([x, y, c]) \quad (\text{by } (1.5)). \end{aligned}$$

Improving on (i), from the definitions (0.11) and (1.2) we see

(i′) If $D_0 \in \mathrm{Der}_0(A)$, $C \in \mathrm{Cayder}(A)$ then $[D_0, C] \in \mathrm{Cayder}(A)$; if C is tracial so is $[D_0, C]$ with $\widehat{[D_0, C]} = [D_0, \hat{C}]$ and $t([D_0, C]) = D_0(t(C))$

(we need D_0 traceless for it to be a $*$-derivation); note that if $c = t(C)$ then $T([D, C]x) = -T(C(Dx)) = -T(cD(x)) = T(D(c)x - D(cx)) = T(D(c)x)$ when D is traceless. Here $D(c)$ is nuclear if c is since derivations preserve nuclei, is skew if D is traceless, and is a $*$-element if c is (as in (ii)). For the conjugate (1.6), note $[D, R_c] = R_{D(c)}$ for any derivatives by (0.11).

Improving upon (iii), we have directly from (1.2) that

(iii′) If $d \in N_l(A)$, $C \in \mathrm{Cayder}(A)$ then $L_d C \in \mathrm{Cayder}(A)$; if d is nuclear and C is tracial so is $L_d C$, with $\widehat{L_d C} = \hat{C} L_{d^*}$, $t(L_d C) = -\hat{C}(d^*)$.

For the trace, when d is nuclear $T(dC(x)) = N(d^*, C(x)) = N(\hat{C}(d^*), x)$ (by (2.10)) $= T(\hat{C}(d^*)^* x)$ (by (0.2)) $= -T(\hat{C}(d^*)x)$ (by (1.7)) with $-\hat{C}(d^*) = -C(d^*) + d^*c = C(d) + d^*c$ by (1.4). For the conjugate,

$$\{\hat{C}L_{d^*} - \widehat{L_d C}\}(x) = \hat{C}(d^*x) + \widehat{L_d C}(x)^* \quad \text{(by (1.7))}$$

$$= C(d^*x) + \{dC(x) - x(cd - C(d)^*)\}^*$$

$$= C(d^*x) - d^*xc + C(x)^*d^* - d^*c^*x^* + C(d)x^* \quad (d^* \text{ is left nuclear})$$

$$= T(C(x))d^* - d^*(xc + c^*x^*) \quad \text{(by (1.2), (1.4))}$$

$$= 0 \quad \text{(by (1.5))}.$$

If c is nuclear so is $\hat{C}(d)$, i.e. $C(d)$ is, by the nontrivial calculation in Lemma 2.11 below.

Improving on (iv), we have

(iv′) If z is a skew $*$-element then $R_z L_d = L_d R_z = R_{dz}$ for dz a skew $*$-element when either d or z is nuclear; if both are nuclear, so is dz.

Indeed, by (0.9) $dxz = xdz$ where $T(dz) = 0$, $adz = zd^*a^* = dza^*$ if one of d, z is nuclear.

To see why the first part of (v) should hold, if $c_i = t(C_i)$ are nuclear then

$$\hat{C}_1C_2(xy) - \hat{C}_1C_2(x)y - x\hat{C}_1C_2(y)$$

$$= \hat{C}_1\{C_2(x)y^* + C_2(y)x\} - \hat{C}_1C_2(x)y + x\{\hat{C}_1C_2(y)\}^* \quad \text{(by (1.2), (1.7))}$$

$$= \{C_1C_2(x)y - C_1(y)C_2(x) - C_2(x)y^*c_1\}$$
$$+\{C_1C_2(y)x^* + C_1(x)C_2(y) - C_2(y)xc_1\}$$
$$-\hat{C}_1C_2(x)y - \hat{C}_1C_2(y)x^* + N(x, \hat{C}_1C_2(y)) \quad \text{(by (1.2), (1.4))}$$
$$= C_1(x)C_2(y) - C_1(y)C_2(x) + C_2(x)(c_1y + y^*c_1^*)$$
$$-C_2(y)(xc_1 + c_1^*x^*) + N(C_2(x), C_2(y))$$
$$\text{(by (1.6), skewness of } c_1\text{, and (2.10))}$$
$$= \{T(c_1y) - C_1y\}C_2(x) - \{T(xc_1) - C_1(x)\}C_2(y)$$
$$+C_1(x)^*C_2(y) + C_2(y)^*C_1(x)$$
$$= C_1(y)^*C_2(x) + C_2(y)^*C_1(x)$$

is symmetric in the indices 1 and 2, therefore $D = \hat{C}_1C_2 - \hat{C}_2C_1$ has $D(xy) - D(x)y - xD(y) = 0$ and D is a derivation. For the second part of (v), note that

$$(C_1\hat{C}_2 - C_2\hat{C}_1) - (\hat{C}_1C_2 - \hat{C}_2C_1)$$
$$= (-C_1R_{c_2} + C_2R_{c_1}) - (-R_{c_1}C_2 + R_{c_2}C_1)$$
$$= (C_2R_{c_1} + R_{c_1}C_2) - (C_1R_{c_2} + R_{c_2}C_1) = L_{C_2(c_1)} - L_{C_1(c_2)}$$

for skew c_i by (1.2).

Improving on (vi), a direct calculation using (0.8), (0.9) and nuclearity of z shows

(vi′) If C is a Cayley derivation and z is a skew nuclear $*$-element, then $D = R_zC$ is a derivation with $CR_z = L_{C(z)} - D$; if C is tracial then $C(z)$ is skew nuclear.

Note by (1.2) we have $R_zC + CR_z = L_{C(z)}$ whenever z is skew. Certainly $C(z)$ is skew if C is tracial, $T(C(z)) = T(cz) = 0$ for radical z, and $C(z)$ is nuclear by the following Lemma 2.11. □

It seems to be difficult to prove directly that Cayley derivations preserve the nucleus.

2.11. Lemma. *If $d \in N(A)$ is nuclear and C a Cayley derivation, then $C(d) \in N_l(A) = N_r(A)$ is outer-nuclear. If C has a nuclear trace element then $C(d) \in N(A)$ is also middle-nuclear.*

Proof. For any Cayley derivation we have

$$C([x, y, y]) = [C(x), y, y] - [C(y), x, y] \tag{2.12}$$

since

$$\begin{aligned}C(-[x, y, y^*]) &+ [C(x), y^*, y] + [C(y), x, y]\\ &= -C((xy)y^*) + C(xN(y)) + \{C(x)y^*\}y - C(x)N(y)\\ &\quad + \{C(y)x\}y + C(y^*)(xy) \quad (\text{by } (1.4))\\ &= \{-C(xy) + C(x)y^* + C(y)x\}y = 0 \quad (\text{by } (1.2)).\end{aligned}$$

Linearizing $y \to y, d$ for nuclear d shows $0 = -[C(d), x, y]$, i.e. $C(d) \in N_l(A)$.

When C has nuclear trace element c we have

$$\begin{aligned}&(2.13) \qquad \hat{C}(xy) = \hat{C}(x)y^* - C(y)^*x\\ &(2.14) \qquad \hat{C}(xy) = x^*C(y) + y\hat{C}(x)\end{aligned} \qquad (t(C) \in N(A))$$

since direct calculation from (1.2), (1.6), (1.7) shows that

$$\begin{aligned}\hat{C}(xy) - \hat{C}(x)y^* + C(y)^*x &= -xyc + xcy^* + T(C(y))x\\ &= x\{T(yc) - yc - c^*y^*\} = 0\end{aligned}$$

and $\{\hat{C}(x)y^* - C(y)^*x\} - \{x^*C(y) + y\hat{C}(x)\} = T(y)\hat{C}(x) - y \circ \hat{C}(x) - N(C(y), x) = 0$ by (0.10′), (1.7), (2.10). Then

$$\begin{aligned}[x^*, C(d), y] &= \{x^*C(d)\}y - x^*\{C(yd) - C(y)d^*\} \quad (\text{by } (1.2))\\ &= \{\hat{C}(xd) - d\hat{C}(x)\}y + x^*C(y)d^* + x^*C(d^*y^*) \quad (\text{by } (2.14), (1.4))\\ &= \{\hat{C}(xdy^*) + C(y^*)^*xd\} - d\hat{C}(x)y + \{x^*C(y)\}d^*\\ &\quad + \{\hat{C}(xd^*y^*) - d^*y^*\hat{C}(x)\} \quad (\text{by } (2.14))\\ &= T(d)\{\hat{C}(xy^*) - y^*\hat{C}(x) - x^*C(y^*)\} + d\{-\hat{C}(x)y + y^*\hat{C}(x)\}\\ &\quad + \{-C(y)^*x - x^*C(y)\}d \quad (\text{by } (1.4))\\ &= 0 + d\{T(y)\hat{C}(x) - \hat{C}(x) \circ y - N(C(y), x)\} \quad (\text{by } (2.14))\\ &= 0 \quad (\text{by } (0.10'), (1.7), (2.10)).\end{aligned}$$

□

3. Cayley derivations of $\mathbf{C}(A, \mu)$. In this section we describe how Cayley derivations of $\mathbf{C}(A, \mu)$ are built out of Cayley derivations and anti-derivations of A. In most cases of dimension ≥ 8 there are no Cayley derivations at all.

3.1. LEMMA. *A map $D_+(a + bl) = D_{11}(a) + D_{22}(b)l$ is a Cayley derivation of $\mathbf{C}(A, \mu)$ iff*

(i) *$D_{11} = D_0$ is a Cayley derivation of A*

(ii) $D_{22}(a) = d_0a^* - D_0(a)$

(iii) $[D_0(a), b] = d_0(ab) - (d_0b)a$

(iv) $d_0 \in \mathrm{Comm}(A)$ *has* $d_0[a, b] = [a, d_0, b]$ *and* $3d_0[A, A] = 0$ *for all* $a, b \in A$.

Proof. For $x_i = a_i + b_il$ we have

$$D(x_1x_2) = D_{11}(a_1a_2 + \mu b_2^*b_1) + D_{22}(b_1a_2^* + b_2a_1)l$$

and

$$\begin{aligned} D(x_1)x_2^* + D(x_2)x_1 &= \{D_{11}(a_1) + D_{22}(b_1)l\}\{a_2^* - b_2l\} \\ &\quad + \{D_{11}(a_2) + D_{22}(b_2)l\}\{a_1 + b_1l\} \\ &= \{D_{11}(a_1)a_2^* + D_{11}(a_2)a_1\} + \mu\{-b_2^*D_{22}(b_1) + b_1^*D_{22}(b_2)\} \\ &\quad + \{-b_2D_{11}(a_1) + D_{22}(b_2)a_1^*\}l \\ &\quad + \{D_{22}(b_1)a_2 + b_1D_{11}(a_2)\}l, \end{aligned}$$

so (using cancellability of μ) D_+ is a Cayley derivation iff $D_{11} = D_0$ is a Cayley derivation of A and for all $a, b \in A$

$$D_0(b_2^*b_1) = -b_2^*D_{22}(b_1) + b_1^*D_{22}(b_2) \tag{1}$$

$$D_{22}(b_2a_1) = D_{22}(b_2)a_1^* - b_2D_0(a_1) \tag{2}$$

$$D_{22}(b_1a_2^*) = D_{22}(b_1)a_2 + b_1D_0(a_2). \tag{3}$$

Here (3) is superfluous in the presence of (2): if we replace b_i by b, a_i by a then (2) + (3) becomes $D_{22}(bT(a)) = D_{22}(b)T(a)$, which holds automatically. Setting $b_2 = 1$ in (2) yields (ii) for $d_0 = D_{22}(1)$. Setting $b_1 = 1$ in (1) yields $D_0(b^*) = -b^*d_0 + d_0b^* - D_0(b) = [d_0, b^*] + D_0(b^*)$ (by (1.4)), and $[d_0, A] = 0$ is the definition (0.6) of d_0 belonging to $\mathrm{Comm}(A)$. Condition (2) reduces to

$$\begin{aligned} 0 &= \{d_0b^* - D_0(b)\}a^* - bD_0(a) - \{d_0(ba)^* - D_0(ba)\} \\ &= (d_0b^*)a^* - d_0(a^*b^*) + [D_0(a), b] \quad \text{(by (1.2))} \\ &= (d_0b^*)a^* - d_0(a^*b^*) + [D_0(a^*), b^*] \quad \text{(by (1.4))}, \end{aligned}$$

which is just (iii). Then (1) + (2) reduces to

$$\begin{aligned} 0 &= \{D_0(b^*a) + b^*(d_0a^* - D_0(a)) - a^*(d_0b^* - D_0(b))\} \\ &\quad + \{d_0(b^*a)^* - D_0(b^*a) - (d_0b - D_0(b^*))a^* + b^*D_0(a)\} \\ &= b^*(d_0a^*) - a^*(d_0b^*) + [a^*, D_0(b)] + d_0(a^*b) \\ &\quad - (d_0b)a^* \quad \text{(by (1.4))} \end{aligned}$$

(continues)

$$= T(b)[d_0, a^*] - b(d_0a^*) - a^*(d_0b) + \{-d_0(ba^*) + (d_0a^*)b\}$$
$$+ d_0(a^*b) - (d_0b)a^* \quad \text{(by (iii))}$$
$$= d_0[a^*, b] + [d_0a^*, b] + T(b)[d_0, a^*] - [d_0b, a^*]$$
$$= -d_0[a^*, b^*] - [d_0a^*, b^*] + [d_0b^*, a^*],$$

so replacing a by a^*, b by b^* it becomes

(iv)′ $d_0[a, b] + [d_0a, b] + [a, d_0b] = 0.$

Now (iii) can be rewritten (using $[d_0, A] = 0$) as

(iii)′ $[D_0(a), b] = d_0[a, b] - [d_0, b, a] = [a, b, d_0] - [d_0, b, a].$

Since $[x, y]^* = [y^*, x^*] = [y, x] = -[x, y]$, $[x, y, z]^* = -[z^*, y^*, x^*] = [z, y, x]$ for any scalar involution, we see $[a, b, d_0] = [D_0(a), b] + [d_0b, a]$ is skew, so

$$d_0[a, b] - [a, d_0b] = [a, b, d_0] + [d_0, b, a] \quad \text{(by (iii)}')$$
$$= T([a, b, d_0]) = 0,$$

hence $d_0[a, b] - [d_0a, b] = 0$ too by skewness in a, b, so

(iva) $d_0[a, b] = [d_0a, b] = [a, d_0b]$

and (iv)′ becomes

(ivb) $3d_0[a, b] = 0.$

Thus (1)–(3) are equivalent to (i)–(iv). □

3.2. LEMMA. *$D_-(a + bl) = D_{12}(b) + D_{21}(a)l$ is a Cayley derivation of $\mathbf{C}(A, \mu)$ iff*

(i) *$D_{21} = C_0$ is a Cayley anti-derivation of A*

(ii) $D_{12} = L_{c_0} - \mu C_0$

(iii) $\mu[C_0(a), b] = [c_0, b, a]$

(iv) *$c_0 \in \mathrm{Comm}(A)$ has $c_0[a, b] = [c_0a, b] + [a, c_0b]$ for all $a, b \in A$.*

Proof. D_- restricts and projects to a Cayley derivation of A into Al; since $al \to a$ is an isomorphism $Al \to A^{\mathrm{op}}$ of right A-modules, we see D_{21}: $A \to A^{\mathrm{op}}$ is a Cayley derivation, i.e. D_{21} is a Cayley anti-derivation C_0 of A. Then $D_-(x_1x_2) = D_{12}(b_1a_2^* + b_2a_1) + C_0(a_1a_2 + \mu b_2^*b_1)l$ and

$$D_-(x_1)x_2^* + D_-(x_2)x_1$$
$$= \{D_{12}(b_1) + C_0(a_1)l\}\{a_2^* - b_l\} + \{D_{12}(b_2) + C_0(a_2)l\}\{a_1 + b_1l\}$$
$$= \{D_{12}(b_1)a_2^* + \mu b_1^*C_0(a_2)\} + \{-\mu b_2^*C_0(a_1) + D_{12}(b_2)a_1\}$$
$$+ \{-b_2D_{12}(b_1) + b_1D_{12}(b_2)\}l + \{C_0(a_1)a_2 + C_0(a_2)a_1^*\}l,$$

so D_- is a Cayley derivation iff

$$\mu C_0(b_2^*b_1) = -b_2D_{12}(b_1) + b_1D_{12}(b_2) \tag{1}$$

(2) $$D_{12}(b_1 a_2^*) = D_{12}(b_1)a_2^* + \mu b_1^* C_0(a_2)$$

(3) $$D_{12}(b_2 a_1) = -\mu b_2^* C_0(a_1) + D_{12}(b_2)a_1.$$

Here (2) and (3) are equivalent since (2) + (3) is $D_{12}(bT(a)) = D_{12}(b)T(a)$, which holds automatically. If we set $c_0 = D_{12}(1)$ then (3) implies $D_{12}(a) = c_0(a) - \mu C_0(a)$, i.e. $D_{12} = L_{c_0} - \mu C_0$ as in (ii). Thus (3) reduces to

$$\begin{aligned} 0 &= c_0(ba) - \mu C_0(ba) + \mu b^* C_0(a) - (c_0 b)a + \mu C_0(b)a \\ &= -[c_0, b, a] - \mu[C_0(a), b^*] = \mu[C_0(a), b] - [c_0, b, a] \end{aligned}$$

as in (iii). Setting $b_2 = 1$ in (1) shows $c_0 b_1 = b_1 c_0$, i.e. $c_0 \in \operatorname{Comm}(A)$, so (1) + (3) reduces to

$$\begin{aligned} 0 &= \{\mu C_0(ba) + b^*(c_0 a) - \mu b^* C_0(a) - a(c_0 b^*) + \mu a C_0(b^*)\} \\ &\quad + \{c_0(ba) - \mu C_0(ba) + \mu b^* C_0(a) - (c_0 b)a + \mu C_0(b)a\} \\ &= \mu[C_0(b), a] - b(c_0 a) + a(c_0 b) - [c_0, b, a] \\ &\quad + T(b)(c_0 a - a c_0) \quad \text{(using (1.4))} \\ &= [c_0, a, b] - b(a c_0) + [a, c_0 b] + (ba)c_0 + 0 \quad \text{(using (iii))} \\ &= [c_0 a, b] + [a, c_0 b] - c_0[a, b] \end{aligned}$$

as in (iv). □

Putting these two pieces together, we get

3.3. Cayley Derivation Theorem. *When A is an algebra with scalar involution and μ a cancellable scalar, the Cayley derivations of $\mathbf{C}(A, \mu)$ are precisely all*

(3.4) $$D = \begin{pmatrix} D_0 & L_{c_0} - C_0 \\ C_0 & L_{d_0}J - D_0 \end{pmatrix}$$

where $J(x) = x^*$ and for all $a, b \in A$

(3.5)
- (i) *D_0 is a Cayley derivation of A*
- (ii) $[D_0(a), b] = d_0(ab) - (d_0 b)a$
- (iii) *$d_0 \in \operatorname{Comm}(A)$ has $d_0[a, b] = [a, d_0 b]$ and $3d_0[A, A] = 0$*
- (iv) *C_0 is a Cayley anti-derivation of A*
- (v) $\mu[C_0(a), b] = [c_0, b, a]$
- (vi) *$c_0 \in \operatorname{Comm}(A)$ has $c_0[a, b] = [c_0 a, b] + [a, c_0 b]$.*

Proof. As in (2.3), and

$$D = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}$$

is graded, where

$$D_+ = \begin{pmatrix} D_{11} & 0 \\ 0 & D_{22} \end{pmatrix}$$

is a derivation iff $D_{11} = D_0, D_{22} = L_{d_0}J - D_0$ as in (i)–(iii) by Lemma 3.1, and

$$D_- = \begin{pmatrix} 0 & D_{12} \\ D_{21} & 0 \end{pmatrix}$$

is a derivation iff $D_{21} = C_0$, $D_{12} = L_{c_0} - \mu C_0$ as in (iv)–(vi) by Lemma 3.2. □

3.4. Corollary. *If A is unitally rigid and*

(i) $\mathrm{Comm}(A) = \Phi 1$

(ii) $\lambda A \subset \Phi 1 \Rightarrow \lambda = 0$ (*e.g. if A has cancellable commutators*)

then when Φ has no 3-torsion there are no Cayley derivations

$$\mathrm{Cayder}(\mathbf{C}(A, \mu)) = 0$$

while if $3\Phi = 0$ *then*

$$\mathrm{Cayder}(\mathbf{C}(A, \mu)) = \Phi S \qquad (S \textit{ standard skew Cayley map}).$$

Proof. If $\mathrm{Comm}(A) = \Phi 1$ then in (3.3) $d_0 = \delta 1$, $c_0 = \gamma 1 \in \Phi 1$ and (ii) becomes $[D_0(a), b] = \delta[a, b]$, $D_0(a) - \delta a \in \mathrm{Comm}(A) = \Phi 1$, so $(D_0 - \delta)(A) \subset \Phi 1$; similarly (3.3)(iii) becomes $3\delta[A, A] = 0$ (i.e. $3\delta A \subset \Phi 1$), (v) becomes $[C_0(a), b] = 0$ (cancelling μ) so $C_0(a) \in \mathrm{Comm}(A) = \Phi 1$ and $C_0(A) \subset \Phi 1$, while (vi) becomes $\gamma[a, b] = 2\gamma[a, b]$, so $\gamma[A, A] = 0$ (i.e. $\gamma A \subset \Phi 1$). If A is unitally rigid we have $C_0 = 0$ by (1.9)(ii). If 3.4(ii) holds we see $\gamma = 3\delta = 0$, so if Φ has no 3-torsion then $\delta = 0$ too; then $D_0(A) \subset \Phi 1$ forces $D_0 = 0$ by (1.9)(ii), therefore when A has no 3-torsion $D = 0$. When $3\Phi = 0$ we know by (1.8), (1.10)(ii) that $D_0 = \delta S_0$,

$$\begin{aligned} D(x) &= D(a + bl) = \delta S_0(a) + \delta(b^* - S_0(b))l \\ &= \delta\{a^* - a + (b^* - (b^* - b))l\} \\ &= \delta\{a^* - a - 2bl\} \quad (\text{since } 3b = 0) \\ &= \delta\{(a^* - bl) - (a + bl)\} = \delta\{x^* - x\} = \delta S(x). \end{aligned}$$

□

4. Derivations of generalized Cayley-Dickson algebras. Our calculations simplify greatly in the case of generalized Cayley-Dickson algebras. These algebras are always unitally rigid, and by (0.7) have cancellable commutators and $\mathrm{Comm}(A) = \Phi 1$ and $N(x, y)$ nondegenerate in dimension ≥ 4, have $N(A) = \Phi 1$ in dimension ≥ 8, and all derivations are traceless by (0.13). The description of derivations given in 2.4 simplifies to

4.1. SCHAFER DERIVATION THEOREM [5]. *Let* $\mathbf{C}_n = \mathbf{C}^{n-3}(C)$ *be a generalized Cayley-Dickson algebra of dimension* 2^n $(n \geq 3)$ *over* Φ *obtained from a Cayley algebra* C. *Then if* Φ *has no* 2- *or* 3-*torsion we have*

$$\text{(i)} \qquad \mathrm{Cayder}(\mathbf{C}_n) = 0, \qquad \mathrm{Der}(\mathbf{C}_n) = \widetilde{\mathrm{Der}(C)}$$

while if $2\Phi = 0$ *then*

$$\text{(ii)} \qquad \mathrm{Cayder}(\mathbf{C}_n) = 0,$$

$$\mathrm{Der}(\mathbf{C}_n) = \widetilde{\mathrm{Der}(C)} \boxplus \Phi Z_4 \boxplus \cdots \boxplus \Phi Z_n \text{ for central } Z_i$$

and if $3\Phi = 0$ *then*

$$\text{(iii)} \qquad \mathrm{Cayder}(\mathbf{C}_n) = \Phi S_n,$$

$$\mathrm{Der}(\mathbf{C}_n) = \widetilde{\mathrm{Der}(C)} \boxplus \Phi W_4 \boxplus \cdots \boxplus \Phi W_n \text{ for central } W_i.$$

Proof. Let $\mathbf{C}_n = \mathbf{C}(A, \mu)$ for $A = \mathbf{C}_{n-1}$ generalized Cayley-Dickson of dimenson $2^{n-1} \geq 4$. By Corollary 3.4 we have Cayley $(\mathbf{C}_n) = 0$ if Φ has no 3-torsion, and $\mathrm{Cayder}(\mathbf{C}_n) = \Phi \underline{S_n}$ $(\underline{S_n}(x) = x^* - x)$ if $3\Phi = 0$. The derivation statement $\mathrm{Der}(\mathbf{C}_n) = \widetilde{\mathrm{Der}(C)}$ is trivial if $n = 3$ $(\mathbf{C}_n = C)$, so assume $n \geq 4$, $\dim A \geq 8$. Then $N(A) = \Phi 1$, $N(x, y)$ is nondegenerate; if Φ has no 3-torsion then $\mathrm{Cayder}(A) = 0$ by the above, so we can apply Corollary 2.8 to see $\mathrm{Der}(\mathbf{C}_n) = \widetilde{\mathrm{Der}_0(A)} = \widetilde{\mathrm{Der}(A)}$ (hence $\mathrm{Der}(\mathbf{C}_n) = \widetilde{\mathrm{Der}(C)}$ as in (i) by induction) if Φ also has no 2-torsion, whereas if $2\Phi = 0$ then $\mathrm{Der}(\mathbf{C}_n) = \widetilde{\mathrm{Der}(A)} \boxplus \Phi Z_n$ (hence $\mathrm{Der}(\mathbf{C}_n) = \widetilde{\mathrm{Der}(C)} \boxplus \Phi Z_4 \boxplus \cdots \boxplus \Phi Z_n$ as in (ii) by induction) for central $Z_n(a + bl) = bl$. If $3\Phi = 0$ we have $\mathrm{Cayder}(A) = \Phi S$ by the above, so Corollary 2.8 says $\mathrm{Der}(\mathbf{C}_n) = \widetilde{\mathrm{Der}(A)} \boxplus \Phi W_n$ (hence $\mathrm{Der}(\mathbf{C}_n) = \widetilde{\mathrm{Der}(C)} \boxplus \Phi W_4 \boxplus \cdots \boxplus \Phi W_n$ as in (iii) by induction) for central $W_n(a + bl) = \mu S(b) + S(a)l$. □

The natural matrix form for the Z_i in (ii) is a string of 2^{n+1-i} blocks down the diagonal, alternating between the $2^{i-1} \times 2^{i-1}$ zero block and the $2^{i-1} \times 2^{i-1}$ identity block. The natural matrix for the W_i in (iii) is a string of 2^{n+1-i} blocks down the superdiagonal and subdiagonal; on the subdiagonal the blocks are alternatingly the $2^{i-1} \times 2^{i-1}$ matrix of $S = I + J$ and the $2^{i-1} \times 2^{i-1}$ zero block, while the superdiagonal is just μ_i times the subdiagonal $(\mathbf{C}_{i-3} = \mathbf{C}(\mathbf{C}_{i-4}, \mu_i))$.

Over a field of characteristic $\neq 3$, $\mathrm{Der}(C)$ is a simple Lie algebra of type G_2.

REFERENCES

[1] A. A. Albert, *Quadratic forms permitting composition*, Annals. of Math., **43** (1942), 161–177.

[2] R. B. Brown, *On generalized Cayley-Dickson algebras*, Pacific J. Math., **20** (1967), 415–422.

[3] R. Erdmann, *Über verallgemeinerte Cayley-Dickson Algebren*, J. für Reine und Angew Math., **250** (197), 153–181.

[4] K. McCrimmon, *Nonassociative algebras with scalar involution*, to appear.

[5] R. D. Schafer, *On the algebras formed by the Cayley-Dickson process*, Amer. J. Math., **76** (1954), 435–446.

[6] ______, *On the simplicity of the Lie algebra E_7 and E_8*, Indag. Math., **28** (1966), 64–69.

Received April 18, 1983. The author wishes to thank the NSF and NRC for financial support during the preparation of this paper, and the University of British Columbia for its hospitality.

UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VA 22903