§9 Derivations

In this section we will give a stepwise construction of derivations paralleling the stepwise construction of algebras by the Cayley-Dickson process. Later (Ch. IV) we will give another construction, in terms of inner derivations, valid for split Cayley algebras over an arbitrary ring. In this section we assume we are working over a <u>field</u> Φ.

We begin with some generalizations about derivations. The basic bimodules for a Cayley-Dickson process algebra A are the regular and Cayley bimodules A and $A\ell$ ($\subset \mathfrak{C}(A,\mu)$). A derivation $A \xrightarrow{D} \text{reg } A$ satisfies $D(ab) = D(a)\cdot b + a\cdot D(b)$ where the bimodule operation $\cdot$ is just ordinary multiplication in A, so that D is nothing but an ordinary derivation of A as algebra. A derivation $A \xrightarrow{D} \text{cay } A$ will be called a Cayley derivation of A; the derivation condition on $D: a \to C(a)\ell$ is $C(ab)\ell = C(a)\ell\cdot b + a\cdot C(b)\ell$. We can also identify cay A with A under the birepresentation $\lambda_a = R_a$, $\rho_a = R_{\bar{a}}$; then the derivation condition on C becomes

$$(9.1) \qquad C(ab) = C(a)\bar{b} + C(b)a\ .$$

In general we will call a linear map $A \xrightarrow{C} A$ a <u>Cayley derivation</u> of A if it satisfies (9.1) (such a map need <u>not</u> be a derivation of A).

From (9.1) we see $C(1) = 0$ and $C(a^2) = C(a)(a+\bar{a})$; in the degree 2 case $a^2 = t(a)a - n(a)1$ shows $C(a^2) = t(a)C(a) = C(a)(a+\bar{a})$ automatically, so by linearization

$$(9.1') \qquad \{C(ab) - C(a)\bar{b} - C(b)a\} + \{C(ba) - C(b)\bar{a} - C(a)b\} = 0$$

is also automatic. This means in practice that we need only verify (9.1) for pairs (a,b) with $a \neq b$, and then only for one of (a,b) or (b,a).

A derivation, of course, satisfies $D(1) = 0$, but for degree 2 algebras (over a field) it is also traceless:

$$t(Da) = 0 \quad n(Da,a) = 0 \ . \tag{9.2}$$

Indeed, $0 = D(a^2 - t(a)a + n(a)1) = D(a)\circ a - t(a)D(a) = \{t(Da)a + t(a)D(a) - n(Da,a)1\} - t(a)D(a) = t(Da)a - n(Da,a)1$. If 1, a are linearly independent both coefficients must vanish, $t(Da) = n(Da,a) = 0$, while if $a = \alpha 1$ then trivially both vanish since $Da = \alpha D(1) = 0$.

Unlike the derivation case, a Cayley derivation C need not be traceless: we needn't have $t(Ca) = 0$. Now $a \to t(Ca)$ is a linear functional, and if $t(x,y) = n(x,\bar{y})$ is a nondegenerate bilinear form (i.e. $n(x,y)$ is nondegenerate, which is true in all standard composition algebras, failing only for inseparable fields of characteristic 2) then by finite-dimensionality there must be a unique vector $c \in A$ giving rise to that linear functional:

$$t(c,a) = t(Ca) \ . \tag{9.3}$$

We call this the <u>trace element</u> of C, written (somewhat misleadingly) as $c = t(C)$. Since $C(1) = 0$ we have $t(c) = t(c,1) = t(C1) = 0$, so the trace element c is always skew.

We now attempt to describe the derivations and Cayley derivations of an algebra $\mathbb{C}(A,\mu)$, obtained from A by the Cayley-Dickson process, in terms of derivations and Cayley derivations of the original algebra A.

We begin with the case of a derivation D of $\mathbb{C}(A,\mu)$. Write D as a matrix

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

relative to the decomposition $\mathfrak{C}(A,\mu) = A \oplus A\ell$. The condition $D(1) = 0$ becomes

$$D_{11}(1) = D_{21}(1) = 0 \tag{9.4}$$

while the alternating condition (9.2) $n(Dx,x) = 0$ for $x = a + b\ell$ becomes $n(D_{11}(a) + D_{12}(b),\ a) - \mu n(D_{21}(a) + D_{22}(b),\ b) = 0$. (Since A, $A\ell$ are orthogonal and $n(a\ell,\ b\ell) = n(a,b)n(\ell)$ and $n(\ell) = -\mu$):

$$\text{(9.5)} \quad \begin{aligned} n(D_{11}a,\ a) &= n(D_{22}b,\ b) = 0 \\ n(a,\ D_{12}(b)) &= \mu n(D_{21}(a),b)\ . \end{aligned}$$

The derivation condition $D(xy) = D(x)y + xD(y)$ for pairs $x,y \in A$, $x \in A$ and $y \in A\ell$, $x,y \in A\ell$ become

$$D(ab) = D(a)b + aD(b) \tag{9.6}$$

$$D(a\cdot b\ell) = D(a)(b\ell) + aD(b\ell) \tag{9.7}$$

$$D(a\ell\cdot b\ell) = D(a\ell)(b\ell) + (a\ell)D(b\ell) \tag{9.8}$$

or

$$D_{11}(ab) + D_{21}(ab)\ell = [D_{11}(a)b+\{D_{21}(a)\bar{b}\}\ell] + [aD_{11}(b)+\{D_{21}(b)a\}\ell] \tag{9.6'}$$

$$D_{12}(ba) + D_{22}(ba)\ell = [\{bD_{11}(a)\}\ell+\mu\bar{b}D_{21}(a)] + [aD_{12}(b)+\{D_{22}(b)a\}\ell] \tag{9.7'}$$

$$\mu D_{11}(\bar{b}a) + \mu D_{21}(\bar{b}a)\ell = [\{bD_{12}(a)\}\ell+\mu\bar{b}D_{22}(a)] + [\{a\overline{D_{12}(b)}\}\ell+ \mu\overline{D_{22}(b)}a]\ . \tag{9.8'}$$

Separating these into components in A and $\hat{A}$,

(9.6a) $$D_{11}(ab) = D_{11}(a)b + aD_{11}(b)$$

(9.6b) $$D_{21}(ab) = D_{21}(a)\bar{b} + D_{21}(b)a$$

(9.7a) $$D_{12}(ba) = \mu\bar{b}D_{21}(a) + aD_{12}(b)$$

(9.7b) $$D_{22}(ba) = bD_{11}(a) + D_{22}(b)a$$

(9.8a) $$\mu D_{11}(\bar{b}a) = \mu\bar{b}D_{22}(a) + \mu\overline{D_{22}(b)a}$$

(9.8b) $$\mu D_{21}(\bar{b}a) = bD_{12}(a) + a\overline{D_{12}(b)}$$

9.9 (Derivation Condition for $\mathbb{C}(A,\mu)$) D is a derivation of a composition algebra $\mathbb{C}(A,\mu)$ iff

(i) D_{11} is a derivation D_o of A

(ii) $D_{22} = D_o + L_{d_o}$ where d_o is skew in the nucleus of A

(iii) $D_{21} = C_o$ is a Cayley derivation of A whose trace element c_o is skew in the nucleus of A

(iv) $D_{12} = \mu\{C_o - R_{c_o}\}$.

In this case $D = \hat{D}_o + \hat{L}_{d_o} + \hat{C}_o = \begin{pmatrix} D_o & \mu\{C_o - R_{c_o}\} \\ C_o & D_o + L_{d_o} \end{pmatrix}$.

Proof. Conditions (9.4) - (9.8) are necessary and sufficient conditions that D be a derivation. (9.6a) is the condition that D_{11} be a derivation of A, which automatically yields $D_{11}(1) = 0$ in (9.4) and $n(D_{11}a,a) = 0$ in (9.5).

This is (i).

Writing $D'_{22} = D_{22} - D_{11}$ the conditions (9.7b), (9.8a), and $n(D_{22}b,b) = 0$ in (9.5) become $D'_{22}(ba) = D'_{22}(b)a$, $0 = \bar{b}D'_{22}(a) + D'_{22}(b)a$, and $n(D'_{22}b,b) = 0$ respectively. Setting $b = 1$ in the first of these shows $D'_{22}(a) = D'_{22}(1)a = d_o a$ or $D'_{22} = L_{d_o}$ for $d_o = D'_{22}(1)$; then for general b the first shows $d_o(ba) = (d_o b)a$, so d_o is (left) nuclear. The third of these becomes $0 = n(d_o b,b) = t(d_o)n(b)$ (see 2.0) for all b, so $t(d_o) = 0$ and d_o is skew. Whenever d_o is skew and nuclear the second of these is automatic: $\bar{b}(d_o a) + \overline{(d_o b)}\, a = \bar{b}\, d_o a + \bar{b}\, \bar{d}_o\, a = \bar{b}t(d_o)a = 0$. This is (ii).

(9.6b) is the condition that D_{21} be a Cayley derivation C_o of A, which automatically yields $D_{21}(1) = 0$ in (9.4). Setting $b = 1$ in (9.7a) shows $D_{12}(a) = \mu D_{21}(a) - \mu a c_o$ for $D_{12}(1) = -\mu c_o$. This yields $D_{12} = \mu\{C_o - R_{c_o}\}$. Setting $a = 1$ in (9.5), and recalling $D_{21}(1) = 0$, shows $n(1,D_{12}(b)) = 0$, and D_{12} is skew: $t(D_{12}(b)) = 0$ for all b. Writing $D_{12} = \mu\{C_o - R_{c_o}\}$, skewness implies $0 = t(C_o b - bc_o) = t(C_o b) - t(bc_o)$, so that c_o is the trace element; in particular, c_o is skew.

So far we have all of (iii) and (iv) but nuclearity of c_o. Subtracting (9.86) (with b replaced by $\bar{b}$) from (9.7a) yields $-(ba)c_o = \bar{b}(ac_o) - t(b)ac_o$ since $D_{12} - \mu D_{21} = -\mu R_{c_o}$, $\mu D_{21} - D_{12} = R_{c_o}$, $D_{12}(b) - \overline{D_{12}(\bar{b})} = D_{12}(b) + D_{12}(\bar{b}) = D_{12}(t(b)1) = t(b)D_{12}(1) = -\mu t(b)c_o$ by skewness of D_{12}. Thus $-(ba)c_o = \{\bar{b} - t(b)\}ac_o = -b(ac_o)$ and c_o is (right) nuclear.

Conversely, when (iii) and (iv) hold we get 9.7a (hence 9.8b too by nuclearity of c_o): since $\overline{D_{12}(x)} = -D_{12}(x)$ by skewness, applying the involution to 9.7a gives an equivalent condition $-D_{12}(ba) = \mu\overline{D_{21}(a)}b - D_{12}(b)\bar{a}$, or $-\mu C_o(ba) + \mu(ba)c_o = \mu\overline{C_o(a)}b - \mu C_o(b)\bar{a} + \mu(bc_o)\bar{a}$; cancelling

μ and using the fact that C_o is a Cayley derivation with $t(C_o(a)) = t(ac_o)$, we can rewrite this condition as $-C_o(b)\bar{a} - C_o(a)b + bac_o = t(ac_o)b - C_o(a)b - C_o(b)\bar{a} + bc_o\bar{a}$, which is automatic from skewness and nuclearity of c_o: $bt(ac_o) + bc_o\bar{a} = b\{t(ac_o) - \bar{c}_o\bar{a}\} = b\{t(ac_o) - \overline{ac_o}\} = b\{ac_o\}$.

Finally, (iii) and (ii) also imply the last relation in (9.5): $n(a,D_{12}(b)) - \mu n(D_{21}(a),b) = t(a\overline{D_{12}(b)}) - \mu t(D_{21}(a)\bar{b})$ (recall $n(x,y) = t(x\bar{y})$ by 2.0) $= -t(aD_{12}(b)) - \mu t(D_{21}(ab) - D_{21}(b)a) = t(a\{-D_{12}(b) + \mu D_{21}(b)\}) - \mu t(D_{21}(ab)) = t(a\{\mu bc_o\}) - \mu t((ab)c_o) = 0$. ∎

We follow a similar procedure for Cayley derivations D of $\mathbb{C}(A,\mu)$. Recalling the remarks after (9.1'), it suffices to establish the Cayley derivation condition $D(xy) = D(x)\bar{y} + D(y)x$ only for $x,y \in A$ and $x \in A$, $y \in A\ell$ and $x,y \in A\ell$:

(9.10) $$D(ab) = D(a)\bar{b} + D(b)a$$

(9.11) $$D(a\cdot b\ell) = -D(a)(b\ell) + D(b\ell)a$$

(9.12) $$D(a\ell\cdot b\ell) = -D(a\ell)(b\ell) + D(b\ell)(a\ell)\ .$$

In terms of

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

these become

(9.10') $$D_{11}(ab) + D_{21}(ab)\ell = [D_{11}(a)\bar{b} + \{D_{21}(a)b\}\ell] + [D_{11}(b)a + \{D_{21}(b)\bar{a}\}\ell]$$

(9.11') $$D_{12}(ba) + D_{22}(ba)\ell = -[\{bD_{11}(a)\}\ell + \mu\bar{b}D_{21}(a)] + [D_{12}(b)a + \{D_{22}(b)\bar{a}\}\ell]$$

$$(9.12') \quad \mu D_{11}(\bar{b}a) + \mu D_{21}(\bar{b}a)\ell = - [\{bD_{12}(a)\}\ell + \mu\bar{b}D_{22}(a)] + [\{aD_{12}(b)\}\ell + \mu\bar{a}D_{22}(b)]$$

or componentwise

$$(9.10a) \qquad D_{11}(ab) = D_{11}(a)\bar{b} + D_{11}(b)a$$

$$(9.10b) \qquad D_{21}(ab) = D_{21}(a)b + D_{21}(b)\bar{a}$$

$$(9.11a) \qquad D_{12}(ba) = - \mu\bar{b}D_{21}(a) + D_{12}(b)a$$

$$(9.11b) \qquad D_{22}(ba) = - bD_{11}(a) + D_{22}(b)\bar{a}$$

$$(9.12a) \qquad \mu D_{11}(\bar{b}a) = - \mu\bar{b}D_{22}(a) + \mu\bar{a}D_{22}(b)$$

$$(9.12b) \qquad \mu D_{21}(\bar{b}a) = - bD_{12}(a) + aD_{12}(b) \ .$$

Recall that if $a \in A$ commutes with A it lies in the center of A. This is of course trivial when A is associative, but in 3.0 we also saw it held when A is a Cayley algebra.

9.13 (Cayley Derivation Condition for $\mathbb{C}(A,\mu)$) D is a Cayley derivation of $\mathbb{C}(A,\mu)$ iff

(i) D_{11} is a Cayley derivation D_o of A such that the values $3D_o(A)$ are all contained in the center of A

(ii) $D_{22}(a) = d_o\bar{a} - D_o(a)$ where d_o and all $D_o(a) - d_oa$ are central

(iii) $D_{21}(a) = \overline{C_o(a)}$ where C_o is a Cayley derivation with values $C_o(A)$ in the center

(iv) $D_{12}(a) = {}_o a - \mu \overline{C_o(a)}$ where ${}_o A$ is central.

Proof. (9.10a) is the condition that D_{11} be a Cayley derivation D_o of A. Setting $b = 1$ in (9.12a) yields $\mu D_{11}(a) = -\mu D_{22}(a) + \mu \bar{a} D_{22}(1)$, so cancelling μ shows $D_{22}(a) = \bar{a}\, d_o - D_o(a)$ for $d_o = D_{22}(1)$. On the other hand, setting $b = 1$ in (9.11b) yields $D_{22}(a) = -D_o(a) + d_o \bar{a}$, so $\bar{a}\, d_o = d_o \bar{a}$. Since d_o commutes with everything, it is central by our above recollection. Once we have this, (9.11b) becomes $d_o(\overline{ba}) - D_o(ba) = -bD_o(a) + d_o \bar{b}\bar{a} - D_o(b)\bar{a}$ (d_o is central) or $d_o[\bar{a},\bar{b}] = [D_o(a),b]$ (D_o is Cayley) or $[D_o(a) - d_o a, b] = 0$ (d_o is central and $[\bar{a},\bar{b}] = [-a,-b] = [a,b]$).

$$(9.11b') \qquad D_o(a) - d_o a \in \Gamma(A).$$

Cancelling μ in (9.12a) (with b replaced by $\bar{b}$) and adding to (9.11b) yields $d_o(\overline{ba}) = -bd_o\bar{a} + D_{22}(b)\bar{a} + \bar{a}D_{22}(b) = -d_o b\bar{a} + [D_{22}(b),\bar{a}] + t(b)\bar{a}d_o$ (recall $d_o = D_{22}(1)$), or $[D_{22}(b),\bar{a}] = d_o(\bar{a}\bar{b} + b\bar{a} - t(b)\bar{a}) = d_o[\bar{a}\bar{b}] = -d_o[\bar{a},b] = +d_o[b,\bar{a}]$ or $[D_{22}(b) - d_o b, \bar{a}] = 0$:

$$(9.12a') \qquad D_{22}(b) - d_o b \in \Gamma(A) .$$

Adding (9.11b') and (9.12a') gives $d_o\bar{b} - 2d_o b \in \Gamma(A)$ or $t(b)d_o - 3d_o b \in \Gamma(A)$. Since d_o is already in $\Gamma(A)$, in the presence of (9.11b') the condition (9.12a') is equivalent to

$$(9.12a'') \qquad 3d_o b \in \Gamma(A) .$$

Thus the Cayley derivation D_o and the central element d_o must be related by all $D_o a - d_o a$ and all $3d_o a$ being central; this implies all $3(D_o a -$

$d_o a\} + 3d_o a = 3D_o(a)$ are central. These are the conditions of (i) and (ii).

We argue in a similar way for the off-diagonal entries D_{12} and D_{21}. Setting $b = 1$ in (9.11a) and (9.12b) yields $D_{12}(a) = -\mu D_{21}(a) + c_o a$ ($c_o = D_{12}(1)$) and $\mu D_{21}(a) = -D_{12}(a) + ac_o$. Once again c_o commutes with everything and therefore is central, and $D_{12}(a) = c_o a - \mu D_{21}(a)$. Thus if we multiply (9.10b) (with a, b interchanged) by μ and add it to (9.11a) we get $c_o(ba) = (c_o b)a + \mu[D_{21}(a),\bar{b}]$; since c_o is central this reduces to $[D_{21}(a),\bar{b}] = 0$, so $D_{21}(a)$ is central:

(9.11a') $$D_{21}(a) \in \Gamma(A) .$$

Using this commutativity, we can apply the involution to (9.10b) to obtain an equivalent condition $\overline{D_{21}(ab)} = \overline{D_{21}(a)}\,\bar{b} + \overline{D_{21}(b)}\,a$, which is the condition that $C_o = \overline{D_{21}}$ be a Cayley derivation. Since the values of D_{21} lie in the center, so do those of C_o. This is condition (iii).

Replacing b by $\bar{b}$ in (9.12b) and adding to (9.11a) yields $c_o(ba) = -\bar{b}(c_o a) + aD_{12}(\bar{b}) + D_{12}(b)a = -c_o(\bar{b}a) + t(b)aD_{12}(1) + [D_{12}(b),a]$ or $[D_{12}(b),a] = c_o\{ba + \bar{b}a - t(b)a\} = 0$: once more $D_{12}(b)$ are all central,

(9.12b') $$D_{12}(b) \in \Gamma(A) .$$

Since $D_{12}(b) = \mu D_{21}(b) - c_o b$ where $D_{21}(b)$ is already central by (9.11a'), we can replace (9.12b') by

(9.12b") $$c_o b \in \Gamma(A) .$$

This is condition (iv). ∎

The Cayley derivation D takes the form $\begin{pmatrix} D_o & L_{c_o} - \mu\overline{C_o} \\ \overline{C_o} & L_{d_o}\circ * - D_o \end{pmatrix}$.

We are mainly interested in Cayley algebras, and for these we have

9.14 (Cayley Derivation Theorem) A Cayley algebra over a field Φ has no Cayley derivations if the characteristic $\neq 3$, while in characteristic 3 the only Cayley derivations are the scalar multiples of the basic skew Cayley derivation

$$D(x) = x + t(x)1 .$$

Proof. Here the center is $\Gamma(A) = \Phi 1$, so $c_o A \subset \Phi 1$ implies $c_o = 0$. If $C_o(a) = \gamma(a)1$ is a Cayley derivation with values in the center $\Phi 1$ then $\gamma(ab)1 = \gamma(a)\bar{b} + \gamma(b)a$ for independent $1,a,b$, implies γ is zero, $C_o = 0$. Thus $D_{12} = D_{21} = 0$.

In characteristic $\neq 3$, $3D_o(A) \subset \Phi 1$ implies $D_o(A) \subset \Phi 1$ and again $D_o = 0$. Also $3d_o A \subset \Phi 1$ implies $d_o A \subset \Phi 1$ so $d_o = 0$, and $D_{11} = D_{22} = 0$. Thus in characteristic $\neq 3$ the only Cayley derivation is $D = 0$.

In characteristic 3, $3D_o(A) \subset \Phi 1$ and $3d_o A \subset \Phi 1$ are automatic, so the only restrictions on D_o and d_o are that $D_o a - d_o a$ are all central: $D_o(a) = \delta_o a + \delta(a)1$ for $d_o = \delta_o 1$. In terms of δ_o and the functional δ, the Cayley derivation condition condition becomes $\delta_o ab + \delta(ab)1 = D_o(ab)$ $= D_o(a)\bar{b} + D_o(b)a = \delta_o(a\bar{b} + ba) + \delta(a)\bar{b} + \delta(b)a$ or $\delta(b)a + \delta(a)\bar{b} -$ $\delta(ab)1 = \delta_o\{ab - ba - a\bar{b}\} = \delta_o\{2ab - a\circ b - t(b)a + ab\} = \delta_o\{3ab - t(a)b$ $- 2t(b)a + n(a,b)1\} = \delta_o\{t(b)a - t(a)b + n(a,b)1\}$ (characteristic 3!). Choosing $1,a,b$ independent we can identify coefficients of a to see

$\delta(b) = \delta_o t(b)$. Thus

$$D_{11}(a) = D_o(a) = \delta_o\{a + t(a)1\}$$

$$D_{22}(a) = \delta_o \bar{a} - D_o(a) = \delta_o\{-2a\} = \delta_o a$$

so $D(x) = D(a + b\ell) = D_{11}(a) + D_{22}(b)\ell = \delta_o\{a + t(a)1 + b\ell\} = \delta_o\{x + t(x)1\}$. ▩

This is not too surprising, since a Cayley derivation of a Cayley algebra $\mathfrak{C}$ is a derivation of $\mathfrak{C}$ into the bimodule cay $\mathfrak{C}$, and cay $\mathfrak{C}$ is no longer an alternative bimodule.

Exercises

9.1 Show the space Cayder(A) of Cayley derivations of an arbitrary degree 2 algebra is a module for the Lie algebra Der(A).

9.2 Show the map $C \to t(C)$ of Cayder(A) into A is a homomorphism of Der(A)-modules: $t([D,C]) = D(t(C))$.

9.3 Show that if C is a Cayley derivation of A and d a nuclear element of A then $C_d = L_d \circ C$ is a new Cayley derivation. (For a composition algebra of dimension 8 the nucleus is just $\Phi 1$, so $C_d = C_{\alpha 1}$ is just αC). In this case show $t(C_d) = t(C)d - \overline{C(d)} = C(d) + \bar{d}t(C)$.

9.4 In the situation of 9.3 show $L_d C - R_{t(C_d)} = \{R_{t(C)} - C\}d$.

9.5 If C, C' are Cayley derivations show $R_{t(C')}C - R_{t(C)}C' = C'R_{t(C)} - CR_{t(C')} + L_{C(t(C'))-C't(C)}$.

9.6 If C, C' are Cayley derivations show $D = \{C - R_{t(C)}\}C' - \{C' - R_{t(C')}\}C$ is an ordinary derivation. Find t(D).