DERIVATIONS IN PRIME RINGS[1]

EDWARD C. POSNER

We prove two theorems that are easily conjectured, namely: (1) In a prime ring of characteristics not 2, if the iterate of two derivations is a derivation, then one of them is zero; (2) If d is a derivation of a prime ring such that, for all elements a of the ring, $ad(a)-d(a)a$ is central, then either the ring is commutative or d is zero.

DEFINITION. *A ring R is called prime if and only if $xay=0$ for all $a\in R$ implies $x=0$ or $y=0$.*

From this definition it follows that no nonzero element of the centroid has nonzero kernel, so that we can divide by the prime p, unless $px=0$ for all x in R, in which case we call R of characteristic p.

A known result that will be often used throughout this paper is given in

LEMMA 1. *Let d be a derivation of a prime ring R and a be an element of R. If $ad(x)=0$ for all $x\in R$, then either $a=0$ or d is zero.*

PROOF: In $ad(x)=0$ for all $x\in R$, replace x by xy. Then

$$ad(xy) = 0 = ad(x)y + axd(y) = axd(y) = 0$$

for all $x, y\in R$. If d is not zero, that is, if $d(y)\neq 0$ for some $y\in R$, then, by the definition of a prime ring, $a=0$.

The following lemma may have some independent interest.

LEMMA 2. *Let R be a prime ring, and let p, q, r be elements of R such that $paqar=0$ for all a in R. Then one, at least, of p, q, r is zero.*

PROOF. In $paqar=0$, replace a by $a+b$; using $paqar=pbqbr=0$, we find $paqbr+pbqar=0$, for all a, b in R. If now $pa=0$, then, for all b in R, $pbqar=0$, so that $p=0$, or else $qar=0$. But if $pa=0$, then $pat=0$ for all $t\in R$, so that $p=0$ or $qatr=0$ for all t in R; again $r=0$, or else $qa=0$. So $p=0$ or $r=0$ or qa is zero whenever pa is zero; replace a by $aqar$; since $p(aqar)=0$ for all $a\in R$, we see that $p=0$ or $r=0$ or $qaqar=0$ for all $a\in R$. Similarly, $p=0$ or $r=0$ or $qaqaq=0$ for all $a\in R$. Assuming therefore that $p\neq 0$, $r\neq 0$, replace a by $a+b$ in $qaqaq=0$ to find as before that $qaqbq+qbqaq=0$. In this equation, replace b by aqb to find

Presented to the Society February 23, 1957; received by the editors February 7, 1957.

[1] This paper was sponsored in part by the Office of Ordnance Research, United States Army, under contract DA-11-022-ORD-1571.

$(qaqaq)bq+qaqbqaq=0$, $(qaq)b(qaq)=0$, for all $b\in R$, for all $a\in R$. So $qaq=0$ for all $a\in R$, $q=0$ if $p\neq 0$, $r\neq 0$.

THEOREM 1. *Let R be a prime ring of characteristic not 2 and d_1, d_2 derivations of R such that the iterate d_1d_2 is also a derivation; then one at least of d_1, d_2 is zero.*

PROOF. d_1d_2 is a derivation, so

$$d_1d_2(ab) = d_1d_2(a)b + ad_1d_2(b).$$

However, d_1, d_2 are each derivations so

$$\begin{aligned} d_1d_2(ab) &= d_1(d_2(ab)) = d_1(d_2(a)b + ad_2(b)) \\ &= d_1d_2(a)b + d_2(a)d_1(b) + d_1(a)d_2(b) + ad_1d_2(b). \end{aligned}$$

But $d_1d_2(ab)=d_1d_2(a)b+ad_1d_2(b)$, so

$$d_2(a)d_1(b) + d_1(a)d_2(b) = 0 \qquad \text{for all } a, b \in R. \tag{1}$$

Replace a by $ad_1(c)$ in (1).

$$d_2(ad_1(c))d_1(b) + d_1(ad_1(c))d_2(b) = 0$$

for all a, b, $c\in R$.

$$d_2(a)d_1(c)d_1(b) + ad_2d_1(c)d_1(b) + d_1(a)d_1(c)d_2(b) + ad_1^2(c)d_2(b) = 0.$$

Now $a(d_2(d_1(c))d_1(b)+d_1(d_1(c))d_2(b))=0$, since $d_2(d_1(c))d_1(b)+d_1 \cdot (d_1(c))d_2(b)=0$, which is merely equation (1) with a replaced by $d_1(c)$. We are left, then, with

$$d_2(a)d_1(c)d_1(b) + d_1(a)d_1(c)d_2(b) = 0 \qquad \text{for all } a, b, c \in R. \tag{2}$$

But $d_1(c)d_2(b)=-d_2(c)d_1(b)$ by (1) with c replacing a. Then (2) becomes $d_2(a)d_1(c)d_1(b)-d_1(a)d_2(c)d_1(b)=0$; factoring out $d_1(b)$ on the right, we have $(d_2(a)d_1(c)-d_1(a)d_2(c))d_1(b)=0$ for all $b\in R$, for all a, $c\in R$. Lemma 1 is just what we need to tell us that $d_2(a)d_1(c)-d_1(a)d_2(c)=0$ for all a, $c\in R$, unless d_1 is zero. But (1) with c replacing b tells us that instead $d_2(a)d_1(c)+d_1(a)d_2(c)=0$ for all a, $c\in R$. Adding these last two equations, we find that $2d_2(a)d_1(c)=0$, $d_2(a)d_1(c)=0$, (since R is not of characteristic 2), for all a, $c\in R$, or else d_1 is zero. Using Lemma 1 again with $d_2(a)$ replacing a, we find that d_1 is zero or else $d_2(a)=0$ for all $a\in R$, i.e. $d_1=0$ or $d_2=0$.

In order to prove Theorem 2, we find it necessary to prove the following lemma.

LEMMA 3. *Let R be a prime ring, and d a derivation of R such that $ad(a)-d(a)a=0$ for all $a\in R$. Then R is commutative, or d is zero.*

PROOF. $(a+b)d(a+b)-(d(a+b))(a+b)=0$ for all $a, b\in R$; subtracting $ad(a)-d(a)a+bd(b)-d(b)b=0$ from this, we arrive at $ad(b)+bd(a)-d(a)b-d(b)a=0$ for all $a, b\in R$. Write this as

$$ad(b) - d(a)b = d(b)a - bd(a).$$

Add to this $ad(b)+d(a)b=d(ab)$ to find

$$2ad(b) = d(b)a - bd(a) + d(ab) \qquad \text{for all } a, b \in R. \tag{3}$$

In (3), replace b by ax

$$2ad(ax) = d(ax)a - axd(a) + d(a^2x),$$

or

$$2ad(a)x + 2a^2d(x) = d(a)xa + ad(x)a - axd(a) + 2ad(a)x + a^2d(x),$$

since $d(a^2)=2ad(a)$; or

$$a^2d(x) = d(a)xa + ad(x)a - axd(a) \qquad \text{for all } a, x \in R. \tag{4}$$

In (3), replace b by xa, and find similarly

$$d(x)a^2 = ad(x)a + axd(a) - d(a)xa, \qquad \text{for all } a, x \in R. \tag{5}$$

Add (4) and (5).

$$a^2d(x) + d(x)a^2 = 2ad(x)a \qquad \text{for all } a, x \in R, \tag{6}$$

or

$$a(d(x)a - ad(x)) = (d(x)a - ad(x))a \qquad \text{for all } a, x \in R. \tag{7}$$

Replace in (7) a by $a+d(x)$; we find that $d(x)$ commutes with $d(x)a-ad(x)$, for all $a\in R$, for all x in R; this says that the square of the inner derivation by x is zero, for all $x\in R$. Let R not be of characteristic 2. Then Theorem 1 says that $d(x)$ is central, for all x in R; let a be an element of R, and A denote inner derivation by a. $ad(x)=d(x)a$, or $Ad(x)=0$ for all $x\in R$. Theorem 1 again shows that $d=0$ or, if not, then A is zero, every a in R is central, R is commutative. But if R is of characteristic 2, (6) says that for all $x\in R$, $d(x)$ commutes with all squares of elements of R. Let R be a prime ring of characteristic 2, and let $e\in R$ commute with a^2, for all $a\in R$.

$$a^2e = ea^2 \qquad \text{for all } a \in R. \tag{8}$$

Replace a by $a+b$ and use $ea^2=a^2e$, $eb^2=b^2e$.

$$(ab + ba)e = e(ab + ba) \qquad \text{for all } a, b \in R. \tag{9}$$

In (9), replace b by ae and commute e and a^2; then $a^2e^2+aeae=ea^2e+eaea$; $a^2e^2=ea^2e$, so

(10) $$aeae = eaea \qquad \text{for all } a \in R.$$

In (9), replace b by e; then $ae^2+eae=eae+e^2a$,

(11) e^2 is in the center of R.

Consider $(ae+ea)^2=aeae+eaea+ae^2a+ea^2e$. But $aeae+eaea=0$ by (10), $ae^2a+ea^2e=e^2a^2+e^2a^2=0$ by (11) and (8). We have

(12) $$(ae + ea)^2 = 0 \qquad \text{for all } a \in R.$$

Let x, y now be elements of R with $xy=0$. By (9), $(xy+yx)e=e(xy+yx)$, so

(13) $xy = 0$ implies $yxe = eyx$.

Now $x^2y=0$, so (13) becomes also $yx^2e=eyx^2$; $yx^2e=yex^2$ since e commutes with all squares. Thus

(14) $xy = 0$ implies $(ye + ey)x^2 = 0$.

But $(ax)y=0$ for all $a\in R$; then we can replace x by ax in (14), to obtain $(ye+ey)axax=0$ for all $a\in R$, whenever $xy=0$. Lemma 2 now says $x=0$ or $ye+ey=0$; in fact, since $x(yv)=0$ for all $v\in R$, Lemma 2 even says $x=0$ or $yve+(ey)v=0$ for all $v\in R$. Since $ye=ey$ if $x\neq 0$, then $x=0$ or $yve+yev=0$ for all $v\in R$, $y(ve+ev)=0$ for all $v\in R$. Lemma 1 applied to the inner derivation by e shows that either $x=0$, $y=0$, or e is central. But by (12) $(ae+ea)(ae+ea)=0$, for all $a\in R$; putting $x=ae+ea$, $y=ae+ea$, we find that for all $a\in R$, $ae+ea=0$, or e is central. That is, for all $a\in R$, $ae+ea=0$, e is central if e commutes with all squares in R.

For all $x\in R$, then, $d(x)$ commutes with all squares in R, $d(x)$ is central for all $x\in R$. Let $d(b)=0$; for all $a\in R$, $d(ab)=d(a)b+ad(b)=d(a)b$; $d(ab)$ is central, so $d(a)b$ is central for all a in R if $d(b)=0$. Now if d is not zero, so that $d(a)\neq 0$ for some $a\in R$, we have $d(a)bx=xd(a)b$; $d(a)$ is central so $xd(a)b=d(a)xb$, whence $d(a)(bx+xb)=0$ for all $x\in R$, if $d(b)=0$. But as previously remarked, no nonzero element of the centroid of R has nonzero kernel; since we are assuming $d(a)\neq 0$, and since $d(a)$ is central, we have proved that b is central whenever $d(b)=0$. But for all $c\in R$, $d(c^2)=d(c)c+cd(c)=2d(c)c=0$, so c^2 commutes with all x in R, for all $c\in R$. Referring back to the conclusion of the previous paragraph with x for e shows x central for all $x\in R$, if d is not the zero derivation.

The following lemma may also be of independent interest.

LEMMA 4.[2] *Let A be a Lie ring, I an ideal of A, d an element of A such*

[2] An oral communication from Professor Kaplansky.

that $dx\cdot x=0$ *for all* $x\in I$. *Then for all* $a\in R$, $(da\cdot x)x=0$ *for all* $x\in I$ (*i.e. the set of* d *satisfying* $dx\cdot x=0$ *for all* $x\in I$ *is an ideal of* A).

PROOF. Let R_a denote right multiplication by a. We want to prove $d(R_aR_x^2)=0$ for all $a\in A$, $x\in I$. The Jacobi identity for a Lie ring may be written as $R_{ax}=R_aR_x-R_xR_a$. Furthermore, since I is an ideal, it contains ax, and $x+ax$, for all $a\in A$, so that $(d\cdot ax)ax=0$, $(d(x+ax))\cdot(x+ax)=0$ for all $a\in A$. From these two equations, and from $dx\cdot x=0$, we get $dx\cdot ax+(d\cdot ax)\cdot x=0$ for all $a\in A$, $x\in I$, or, in the other notation, $d(R_xR_{ax}+R_{ax}R_x)=0$. But from

$$d(R_xR_{ax}+R_{ax}R_x)=d(R_x(R_aR_x-R_xR_a)+(R_aR_x-R_xR_a)R_x)$$
$$=d(R_xR_aR_x-R_x^2R_a+R_aR_x^2-R_xR_aR_x)=d(R_aR_x^2-R_x^2R_a),$$

$d(R_aR_x^2-R_x^2R_a)=0$ for all $a\in A$, $x\in I$. By hypothesis, $d(R_x^2)=0$, so that $d(R_aR_x^2)=0$ for all $a\in A$, $x\in I$. This is exactly what we had to prove.

We are now ready for Theorem 2.

THEOREM 2. *Let* R *be a prime ring and* d *a derivation of* R *such that, for all* $a\in R$, $ad(a)-d(a)a$ *is in the center of* R. *Then, if* d *is not the zero derivation,* R *is commutative.*

PROOF. Let A be the Lie ring of derivations of R and I the ideal of A consisting of inner derivations. Let, for $a\in R$, I_a denote inner derivation by a. Let $[d_1, d_2]$ for $d_1, d_2\in A$ denote the (commutator) product of derivations in A. We are assuming $[(d, I_a), I_a]=0$. By the preceding lemma, for all $x\in R$, that is, for all $I_x\in I$, $[[[d, I_x]I_a]I_a]=0$ for all $a\in R$. That is, $a(ad(x)-d(x)a)-(ad(x)-d(x)a)a$ is central for all $x, a\in R$,

(15) $\quad a^2d(x)+d(x)a^2-2ad(x)a$ is central for all $x, a\in R$.

Commute (15) with a.

$$3ad(x)a^2+a^3d(x)=3a^2d(x)a+d(x)a^3. \tag{16}$$

Suppose R is of characteristic 3. Then for all $a\in R$, $I_a{}^3d=0$. Theorem 1 says that d is zero, or else every a^3 is in the center of R; if this is the case, then for all $a, b\in R$, $(a+b)^3-a^3-b^3=a^2b+aba+ba^2+b^2a+bab+ab^2$ is central; replace a by $-a$ to find $a^2b+aba+ba^2-(b^2a+bab+ab^2)$ central for all $a, b\in R$; adding these last two and dividing by 2, we see that $a^2b+aba+ba^2$ is central, for all $a, b\in R$. Replace b by ab: $a^3b+a^2ba+aba^2=a(a^2b+aba+ba^2)$ is central; if $a^2b+aba+ba^2$ is not zero, given a, for some b, then, since it is central, we can divide by it, whence a would be central. So assume that R has the property that for all $a, b \in R$, $a^2b+aba+ba^2=0$. This reads, since R is of characteristic 3, as $a(ab-ba)-(ab-ba)a=0$ for all $b \in R$, $I_a^2=0$; by Theorem 1, a is central, R is commutative.

Suppose now that R is of characteristic different from 3. Write $d(x)=x'$. In (16), replace x by a: $3aa'a^2+a^3a'-3a^2a'a-a'a^3=0$, or $a^3a'-a'a^3=3a^2a'a-3aa'a^2=3a(aa'-a'a)a$. Since $aa'-a'a$ is central by the hypothesis of this theorem, we find

(17) $$a^3a' - a'a^3 = 3(aa' - a'a)a^2, \qquad \text{for all } a \in R.$$

Furthermore, $(aa'-a'a)a=aa'a-a'a^2$. But $(aa'-a'a)a=a(aa'-a'a)$ $=a^2a'-aa'a$; adding these last two, we obtain

(18) $$2(aa' - a'a)a = a^2a' - a'a^2.$$

In (16), replace x by ax'.

$$3a^2x''a^2 + a^4x'' - 3a^3x''a - ax''a^3 + 3aa'x'a^2 + a^3a'x' - 3a^2a'x'a - a'x'a^3 = 0.$$

However,

$$3a^2x''a^2 + a^4x'' - 3a^3x''a - ax''a^3 = a(3ax''a^2 + a^3x'' - 3a^2x''a - x''a^3) = 0,$$

as is seeen from (16) by replacing x by x'. So

(19) $$3aa'x'a^2 + a^3a'x' - 3a^2a'x'a - a'x'a^3 = 0 \qquad \text{for all } x, a \in R.$$

Multiply (16) on the left by a'.

(20) $$3a'ax'a^2 + a'a^3x' - 3a'a^2x'a - a'x'a^3 = 0.$$

Subtract (20) from (19):

$$3(aa' - a'a)x'a^2 + (a^3a' - a'a^3)x' - 3(a^2a' - a'a^2)x'a = 0 \qquad \text{for all } x, a \in R.$$

Using (17) and (18), we arrive at, after dividing by 3,

$$(aa' - a'a)(x'a^2 + a^2x' - 2ax'a) = 0 \qquad \text{for all } x, a \in R.$$

If $aa'-a'a \neq 0$ for some a, then for that a, and all x,

(21) $$x'a^2 + a^2x' - 2ax'a = 0.$$

Replace x by ax in (21):

$$ax'a^2 + a^3x' - 2a^2x'a + a'xa^2 + a^2a'x - 2aa'xa = 0;$$

since

$$ax'a^2 + a^3x' - 2a^2x'a = a(x'a^2 + a^2x' - 2ax'a) = 0$$

by (21), we have

$$(22) \qquad a'xa^2 + a^2a'x - 2aa'xa = 0 \qquad \text{for all } x \in R.$$

Now in (21) replace x by a: $a'a^2+a^2a'-2aa'a=0$. Multiply this on the right by x.

$$(23) \qquad a'a^2x + a^2a'x - 2aa'ax = 0 \qquad \text{for all } x \in R.$$

Subtract (23) from (22).

$$(24) \qquad a'(xa^2 - a^2x) - 2aa'(xa - ax) = 0 \qquad \text{for all } x \in R.$$

Replace x by ax in (24).

$$(25) \qquad a'a(xa^2 - a^2x) - 2aa'a(xa - ax) = 0 \qquad \text{for all } x \in R.$$

Multiply (24) by a on the left.

$$(26) \qquad aa'(xa^2 - a^2x) - 2a^2a'(xa - ax) = 0 \qquad \text{for all } x \in R.$$

Subtract now (25) from (26):

$$(aa' - a'a)(xa^2 - a^2x) - 2a(aa' - a'a)(xa - ax) = 0 \quad \text{for all } x \in R.$$

Since $aa'-a'a\neq 0$,

$$(27) \quad xa^2 - a^2x - 2a(xa - ax) = 0 \qquad \text{for all } x \in R \text{ if } aa' - a'a \neq 0.$$

So $xa^2+a^2x-2axa=0$, $a(ax-xa)=(ax-xa)a$, $I_a^2=0$. That is, a is central by Theorem 1 or else $aa'=a'a$, if R is of characteristic different from 2. So when R is of characteristic not 2, $aa'=a'a$ for all $a\in R$; Lemma 3 now finishes the proof. Let R finally be of characteristic 2. (27) says $aa'=a'a$ or else a^2 is central, for all $a\in R$. If $aa'\neq a'a$ for some $a \in R$, a^2 is central and not zero. For if $a^2 = 0$ then $(a^2)' = aa'+a'a=0$, $aa'=a'a$. Then a is not a divisor of zero, since if $ya=0$, $ya^2=0$, $y=0$. Let $x\in R$; we shall prove that aa' commutes with x^2. Either $(axa)^2$ is central, or $(axa)(axa)'=(axa)'(axa)$. If $(axa)^2$ is central, axa^2xa is in the center of R. Then ax^2a is in the center of R, since a^2 is; call it c. Then $aca=a^2c$ is in the center of R, and equals $a^2x^2a^2$. So $a^2x^2a^2$ is in the center of R, and so is x^2, whence x^2 commutes with aa' if $(axa)^2$ is central. On the other hand, if x^2 is not central, then $xx'=x'x$ and $(axa)(axa)'=(axa)'(axa)$. Then $(axa) \cdot (a'xa+ax'a+axa')=(a'xa+ax'a+axa')axa$, or

$$axaa'xa + axa^2x'a + axa^2xa' = a'xa^2xa + ax'a^2xa + axa'axa.$$

Now a^2 is central, whence

$$ax(aa' + a'a)xa + (a(xx' + x'x)a + ax^2a' + a'x^2a)a^2 = 0.$$

But $xx'+x'x=0$, and $aa'+a'a$ is central so that

$$(aa' + a'a)ax^2a + (ax^2a' + a'x^2a)a^2 = 0.$$

Since a is not a right zero divisor,

$$(aa' + a'a)ax^2 + (ax^2a' + a'x^2a)a = 0,$$

$$ax^2(aa' + a'a) + (ax^2a' + a'x^2a)a = 0,$$

$$ax^2aa' + ax^2a'a + ax^2a'a + a'x^2a^2 = 0.$$

Thus $ax^2aa'+a'x^2a^2=0$; a^2 is central so $ax^2aa'+a^2a'x^2=0$; a is not a left divisor of zero so $x^2aa'+aa'x^2=0$, for any x such that x^2 is not central, hence, for all $x\in R$, as promised; otherwise $aa'=a'a$. Recourse to the latter part of Lemma 3 shows a^2 central and aa' central or else $aa'=a'a$. But in the former case, $a\cdot aa'=aa'\cdot a$; since a is not a zero divisor, $aa'=a'a$, for all $a\in R$. Lemma 3 completes the proof.

THE UNIVERSITY OF CHICAGO AND
BELL TELEPHONE LABORATORIES, NEW YORK