Aequat. Math. 93 (2019), 1127–1138

0001-9054/19/061127-12
published online January 3, 2019
https://doi.org/10.1007/s00010-018-0630-z

Aequationes Mathematicae

Derivations and Leibniz differences on rings: II

Bruce Ebanks

Abstract. In an earlier paper we discussed the composition of derivations of order 1 on a commutative ring R, showing that (i) the composition of n derivations of order 1 yields a derivation of order at most n, and (ii) under additional conditions on R the composition of n derivations of order exactly 1 forms a derivation of order exactly n. In the present paper we consider the composition of derivations of any orders on rings. We show that on any commutative ring R the composition of a derivation of order at most n with a derivation of order at most m results in a derivation of order at most $n+m$. If R is an integral domain of sufficiently large characteristic, then the composition of a derivation of order exactly n with a derivation of order exactly m results in a derivation of order exactly $n+m$. As in the previous paper, the results are proved using Leibniz difference operators.

Mathematics Subject Classification. 39B52, 13N15, 39B72.

Keywords. Commutative ring, Integral domain, Derivation, Derivations of higher order, Leibniz difference operator.

1. Introduction

This paper is a sequel to [2], in which the composition of derivations of order 1 was investigated on commutative rings. There it was proved that the composition of n derivations of order 1 results in a derivation of order at most n. If the ring satisfies some additional conditions, then the composition of n derivations of order exactly 1 forms a derivation of order exactly n. (See also [5] for similar results.) Here we consider the more general question of compositions of derivations of any orders. On commutative rings we show in Proposition 4.1 below that the composition of a derivation of order at most n with a derivation of order at most m gives a derivation of order at most $n+m$. We then prove in Theorem 4.3 that on integral domains of sufficiently large characteristic (including characteristic 0) the composition of a derivation of order exactly n with a derivation of order exactly m yields a derivation of order exactly $n+m$.

The main tool used is the Leibniz difference operator, which plays a fundamental role in the theory of derivations (of all orders).

Birkhäuser

Given a commutative ring S with subring R, a *derivation* from R into S is a function $d : R \to S$ which satisfies *additivity* and the *product rule* (or *Leibniz rule*), respectively

$$d(x+y) = d(x) + d(y) \qquad \text{and} \qquad d(xy) = xd(y) + d(x)y,$$

for all $x, y \in R$. A mapping $B : R \times R \to S$ is a *bi-derivation* (from R into S) if B is a derivation in each variable when the other variable is fixed.

We define derivations of other orders inductively as follows.

Definition 1.1. Let $R \subseteq S$ be commutative rings. The zero function is the only *derivation of order 0* from R into S, and there are no derivations of order less than 0. For each $n \in \mathbb{N}$, suppose we have defined derivations of order at most $n-1$. If $f : R \to S$ is additive, then f is said to be a *derivation of order at most* n from R into S if there exists a function $B : R \times R \to S$ such that B is a derivation of order at most $n-1$ in each variable and

$$f(xy) - xf(y) - f(x)y = B(x,y) \ \text{ for all } \ x, y \in R. \tag{1}$$

Such a function B is called a *bi-derivation of order at most* $n-1$ (from R into S). Let $\mathcal{D}_n(R,S)$ denote the set of all derivations of order at most n from R into S.

It is evident that the definition of derivation of order 1 agrees with the earlier definition of derivation.

We say that a function $d : R \to S$ is a derivation of order *exactly* 1 if $d \in \mathcal{D}_1(R,S)$ and $d \neq 0$, that is if $d \in \mathcal{D}_1(R,S) \setminus \mathcal{D}_0(R,S)$. Similarly, d is a derivation of order *exactly* n if $d \in \mathcal{D}_n(R,S) \setminus \mathcal{D}_{n-1}(R,S)$. For each $n \in \mathbb{N}$, let $\mathcal{D}_n^*(R,S) = \mathcal{D}_n(R,S) \setminus \mathcal{D}_{n-1}(R,S)$ be the set of derivations of order exactly n. If $R = S$ we abbreviate $\mathcal{D}_n(R,S)$ and $\mathcal{D}_n^*(R,S)$ by $\mathcal{D}_n(R)$ and $\mathcal{D}_n^*(R)$, respectively.

It was shown independently in [2,5] that for any commutative ring R, if $d_1, \ldots, d_n \in \mathcal{D}_1(R)$, then $d_n \circ \cdots \circ d_1 \in \mathcal{D}_n(R)$. Furthermore, if R is an integral domain of characteristic larger than $n!$ (including characteristic 0) and $d_1, \ldots, d_n \in \mathcal{D}_1^*(R)$, then $d_n \circ \cdots \circ d_1 \in \mathcal{D}_n^*(R)$. Here we generalize those results by considering compositions of derivations of all orders, not just those of order 1.

We believe that the proofs of our main results, especially Proposition 4.2 and Theorem 4.3, are actually simpler than the corresponding proofs given in [2] for the special case of compositions of derivations of order 1. Also the condition on the characteristic of R is sharpened.

The organization of the paper is as follows. We begin in the next section with some preliminaries concerning Leibniz differences and their applicability to the theory of derivations. Section 3 contains additional preliminaries, notation and definitions. The main results are in Sect. 4, and the final section contains some examples.

2. The Leibniz difference operator and derivations

Our fundamental tool in these investigations is the Leibniz difference operator. In this section we define Leibniz differences of various orders and state some of their relevant properties. In particular we import some results from [2] and [3] illustrating their close relationship with derivations.

Definition 2.1. Let $R \subseteq S$ be commutative rings. For any function $f : R \to S$ and any $y \in R$ we define the function $\Lambda_y f : R \to S$, called the *Leibniz difference (of order 1) of* f *with increment* y, by

$$\Lambda_y f(x) := f(xy) - xf(y) - f(x)y \quad \text{for all} \quad x \in R.$$

We call Λ_y a *Leibniz difference operator of order 1* on S^R. We also define *Leibniz difference operators of order* n on S^R by

$$\Lambda_{y_1,\ldots,y_n} := \Lambda_{y_n} \circ \cdots \circ \Lambda_{y_1} \qquad \text{for all} \qquad y_1, \ldots, y_n \in R,$$

for each $n \in \mathbb{N}$. For any $f \in S^R$ we say $\Lambda_{y_1,\ldots,y_n} f$ is a *Leibniz difference of order* n *of* f. For the special case $y_1 = \cdots = y_n = y$ we define

$$\Lambda_y^n := \Lambda_{y,\ldots,y}.$$

Note that if f is a derivation (of any order), then the function $(x, y) \mapsto \Lambda_y f(x)$ is identical to the bi-derivation B appearing in (1).

It follows from the definition of Leibniz differences that for any $f \in S^R$ and any $y_1, \ldots, y_{n+1} \in R$ we have

$$\begin{aligned}\Lambda_{y_1,\ldots,y_n} f(y_{n+1}) =& f(y_1 y_2 \cdots y_{n+1}) - \sum_{i=1}^{n+1} y_i f(y_1 \cdots \widehat{y_i} \cdots y_{n+1}) \\ &+ \sum_{1 \le i < j \le n+1} y_i y_j f(y_1 \cdots \widehat{y_i} \cdots \widehat{y_j} \cdots y_{n+1}) \\ &+ \ldots + (-1)^n \sum_{i=1}^{n+1} y_1 \cdots \widehat{y_i} \cdots y_{n+1} f(y_i),\end{aligned} \tag{2}$$

where we use the hat symbol $\widehat{\ }$ over a variable to indicate that that variable is omitted. Letting $|I| =$ the cardinality of a set $I \subseteq \mathbb{N}$, this can be written more concisely as

$$\Lambda_{y_1,\ldots,y_n} f(y_{n+1}) = \sum_{j=0}^{n} (-1)^j \sum_{|I|=j} \left(\prod_{i \in I} y_i \right) f \left(\prod_{r \in \{1,\ldots,n+1\} \setminus I} y_r \right) \tag{3}$$

for all $y_1, \ldots, y_{n+1} \in R$, where the second summation is taken over all subsets I of cardinality j of the index set $\{1, \ldots, n+1\}$. In particular, taking $y_i = y$ for $1 \le i \le n+1$ we have

$$\Lambda_y^n f(y) = \sum_{j=0}^{n} (-1)^j \binom{n+1}{j} y^j f(y^{n+1-j}). \tag{4}$$

Since our rings are commutative, Eqs. (2) and (3) exhibit symmetry with respect to all variables. Hence we have the following, where S_m for any positive integer m denotes the *symmetric group* of all permutations on the set $\{1,\ldots,m\}$.

Lemma 2.2. *Let $R \subseteq S$ be commutative rings, let $n \in \mathbb{N}$ and let $f : R \to S$. For any $y_1,\ldots,y_{n+1} \in R$ and any permutation $\pi \in S_{n+1}$ we have*

$$\Lambda_{y_1,\ldots,y_n} f(y_{n+1}) = \Lambda_{\pi(y_1),\ldots,\pi(y_n)} f(\pi(y_{n+1})).$$

In particular, taking only permutations in S_{n+1} that leave y_{n+1} fixed we get

$$\Lambda_{y_1,\ldots,y_n} f = \Lambda_{\pi(y_1),\ldots,\pi(y_n)} f \quad \text{for all} \quad \pi \in S_n.$$

The next observation is Proposition 2.3 in [2]. (Its final sentence is also a restatement of Proposition 4.5 in [1]; see Eq. (3) above.)

Proposition 2.3. *Let $R \subseteq S$ be commutative rings, let $f : R \to S$ be additive, let $n \in \mathbb{N}$ and $j \in \{1,\ldots,n\}$. Then $f \in \mathcal{D}_n(R,S)$ if and only if $\Lambda_{y_1,\ldots,y_j} f \in \mathcal{D}_{n-j}(R,S)$ for all $y_1,\ldots,y_j \in R$. In particular, $f \in \mathcal{D}_n(R,S)$ if and only if all Leibniz differences of f of order n vanish.*

From this follows the nesting property

$$\{0\} = \mathcal{D}_0(R,S) \subseteq \mathcal{D}_1(R,S) \subseteq \cdots \subseteq \mathcal{D}_n(R,S) \subseteq \cdots$$

for spaces of derivations on commutative rings.

The following result further illustrates the close connection between Leibniz differences and derivations. In view of Eq. (4) above it is simply a restatement of Theorem 5 in [3] (see also Corollary 2 in [4]) in terms of Leibniz differences.

Proposition 2.4. *Let $n \in \mathbb{N}$, let $R \subseteq S$ be commutative rings, and suppose $f : R \to S$ is additive. In addition suppose multiplication by $(n+1)!$ is either injective in S or surjective in R. Then $f \in \mathcal{D}_n(R,S)$ if and only if $\Lambda_y^n f(y) = 0$ for all $y \in R$.*

3. Further notation, definitions, and observations

We introduce some special notation for the treatment of Leibniz differences. Suppose we are given a list of elements $y_1,\ldots,y_n \in R$ which are to serve as increments for Leibniz differences. Although the order of the increments is immaterial in a Leibniz difference operator $\Lambda_{y_1,\ldots,y_n}$ because of symmetry, it is not correct to think of the list of increments as a set. This is because some of the increments y_i may be used more than once, so we need to count them with their multiplicities. Thus the list of increments is actually a *multiset*. For this reason we use the notation

$$Y = [y_1, \ldots, y_n] \sqsubseteq R$$

to indicate that Y is an unordered list of n increments (some of which may be repeated) taken from R. For such a list we write $y_i \in Y$ for $1 \leq i \leq n$.

Definition 3.1. Suppose $n \in \mathbb{N}$, R is a commutative ring, $Y = [y_1, \ldots, y_n] \sqsubseteq R$, $f : R \to S$, and $j \in \{1 \ldots, n\}$.

(a) Let $P_j(Y)$ denote the collection of unordered sublists of Y with length j.
(b) For any $X = [x_1, \ldots, x_j] \in P_j(Y)$, define $\Lambda_X f := \Lambda_{x_1,\ldots,x_j} f$. Define $\Lambda_\emptyset f := f$.
(c) For any $X = [x_1, \ldots, x_j] \in P_j(Y)$, let $\Lambda_{Y\setminus X} f$ denote the Leibniz difference of f of order $n - j$ with increments from the unordered list $Y \setminus X$ that remains after the increments $x_1, \ldots, x_j$ are deleted from the unordered list Y.
(d) If $j \geq 2$ and $X = [x_1, \ldots, x_j] \in P_j(Y)$, define $\Lambda f(X) := \Lambda_{x_1,\ldots,x_{j-1}} f(x_j)$. If $X = [x_1]$, define $\Lambda f(X) := f(x_1)$.
(e) For any two unordered lists $U = [u_1, \ldots, u_j]$ and $V = [v_1, \ldots, v_k]$, let $U \sqcup V$ denote the concatenated unordered list

$$U \sqcup V := [u_1, \ldots, u_j, v_1, \ldots, v_k].$$

(f) For any $X \in P_j(Y)$, write $|X| = j$ to indicate that X is an unordered list of length j.

The following observation is a simple consequence of Eqs. (2) or (3) when any increment is 0, together with the fact that $f(0) = 0$ for any additive function f.

Lemma 3.2. *Let $R \subseteq S$ be commutative rings, let $f : R \to S$ be any function such that $f(0) = 0$, and let X be an unordered list of elements of R containing 0. Then $\Lambda_X f = 0$.*

In particular, if $0 \in X$ and f is additive then $\Lambda_X f = 0$.

For a commutative ring R, let $char(R)$ denote the *characteristic* of R. When we declare that $char(R) > m$ for some $m \in \mathbb{N}$ we also include the possibility that $char(R) = 0$.

An *integral domain* is a commutative ring R in which for any $x, y \in R$ the equation $xy = 0$ implies that $x = 0$ or $y = 0$. The characteristic of an integral domain is either 0 or a prime p. Hence for an integral domain R and any positive integer m we have that $char(R) > m$ if and only if multiplication by p is injective in R for every prime $p \leq m$. It follows that $char(R) > m$ if and only multiplication by $m!$ is injective in R.

4. Compositions of derivations of various orders

Now we show that in the setting of commutative rings, if $f \in \mathcal{D}_j(R,S)$ and $g \in \mathcal{D}_k(S,T)$ then $g \circ f \in \mathcal{D}_{j+k}(R,T)$. In other words, the composition of derivations is additive with respect to their orders. This result generalizes Proposition 3.1 in [2].

Proposition 4.1. *Let $m \in \mathbb{N}$ with $m \geq 2$.*

(a) *Let $R \subseteq S \subseteq T$ be commutative rings, and let $j \in \mathbb{N}$ with $1 \leq j \leq m-1$. Suppose $f \in \mathcal{D}_j(R,S)$ and $g \in \mathcal{D}_{m-j}(S,T)$. Then*

$$\Lambda_y(g \circ f) = g(y)f + f(y)\tilde{g} + \Lambda_{f(y)}\tilde{g} + (\Lambda_y g) \circ f + g \circ (\Lambda_y f) \tag{5}$$

for all $y \in R$, where $\tilde{g} := g|_R$ is the restriction of g to R. Moreover $g \circ f \in \mathcal{D}_m(R,T)$.

(b) *Let $R_1 \subseteq R_2 \subseteq \cdots \subseteq R_{n+1}$ be commutative rings, and let $n_1, \ldots, n_m \in \mathbb{N}$. If $f_i \in \mathcal{D}_{n_i}(R_i, R_{i+1})$ for $1 \leq i \leq m$, then $f_n \circ \cdots \circ f_1 \in \mathcal{D}_{n_1+\cdots+n_m}(R_1, R_{n+1})$.*

Proof. We start by proving Eq. (5). For any $x, y \in R$ we have

$$\begin{aligned}
\Lambda_y(g \circ f)(x) &= (g \circ f)(xy) - x(g \circ f)(y) - (g \circ f)(x)y \\
&= [g(xf(y) + yf(x) + \Lambda_y f(x))] - xg(f(y)) - yg(f(x)) \\
&= [g(xf(y)) - xg(f(y))] + [g(yf(x)) - yg(f(x))] + g(\Lambda_y f(x)) \\
&= [g(x)f(y) + \Lambda_{f(y)}g(x)] + [g(y)f(x) + \Lambda_y g(f(x))] + g(\Lambda_y f(x)),
\end{aligned}$$

which is (5).

Next we show that $g \circ f \in \mathcal{D}_m(R,T)$ by induction on m. For the initial step $m = 2$ we have $j = 1$, so $f \in \mathcal{D}_1(R,S)$ and $g \in \mathcal{D}_1(S,T)$. By (5) for $j = 1$ and $m = 2$ we have

$$\Lambda_y(g \circ f) = g(y)f + f(y)\tilde{g},$$

since in this case $\Lambda_{f(y)}\tilde{g} = \Lambda_y g = \Lambda_y f = 0$. Thus we see that $\Lambda_y(g \circ f) \in \mathcal{D}_1(R,T)$ for all $y \in R$. Therefore $g \circ f \in \mathcal{D}_2(R,T)$ and the statement is true for $m = 2$.

Now let $M \geq 3$ and suppose our statement is true for all $2 \leq m \leq M-1$. We have to prove that it is true for $m = M$, so suppose that $f \in \mathcal{D}_j(R,S)$ and $g \in \mathcal{D}_{M-j}(S,T)$, with $1 \leq j \leq M-1$. Consider the right hand side of (5) term by term for a given element $y \in R$.

(i) $g(y)f \in \mathcal{D}_j(R,T) \subseteq \mathcal{D}_{M-1}(R,T)$.
(ii) $f(y)\tilde{g} \in \mathcal{D}_{M-j}(R,T) \subseteq \mathcal{D}_{M-1}(R,T)$.
(iii) $\Lambda_{f(y)}\tilde{g} \in \mathcal{D}_{M-j-1}(R,T) \subseteq \mathcal{D}_{M-1}(R,T)$.
(iv) $(\Lambda_y g) \circ f$ is the composition of $\Lambda_y g \in \mathcal{D}_{M-j-1}(S,T)$ with $f \in \mathcal{D}_j(R,S)$. By the induction hypothesis the composite function belongs to $\mathcal{D}_{M-1}(R,T)$.

(v) $g \circ (\Lambda_y f)$ is the composition of $g \in \mathcal{D}_{M-j}(S,T)$ with $\Lambda_y f \in \mathcal{D}_{j-1}(R,S)$. By the induction hypothesis the composite function belongs to $\mathcal{D}_{M-1}(R,T)$.

By the linearity of Eq. (5) we conclude that $\Lambda_y(g \circ f) \in \mathcal{D}_{M-1}(R,T)$, and thus $(g \circ f) \in \mathcal{D}_M(R,T)$. Therefore part (a) is proved.

Part (b) follows from part (a) by a simple inductive argument. □

Our ultimate goal is to show that under certain conditions on the underlying rings the same type of additivity holds for derivations with respect to *exact* orders, that is for the sets $\mathcal{D}_n^*$. The key to reaching that goal is the following proposition.

Proposition 4.2. *Let $R \subseteq S \subseteq T$ be commutative rings, let $m \in \mathbb{N}$ with $m \geq 2$, and let $j \in \mathbb{N}$ with $1 \leq j \leq m-1$. Suppose $f \in \mathcal{D}_j(R,S)$ and $g \in \mathcal{D}_{m-j}(S,T)$, and let $Y = [y_1, \ldots, y_{m-1}]$ be an unordered list of $m-1$ increments taken from R. Then*

$$\Lambda_Y(g \circ f) = \sum_{j=1}^{m-1} \sum_{X \in P_j(Y)} \left[\Lambda f(X) \Lambda_{Y \setminus X} \tilde{g} + \Lambda g(X) \Lambda_{Y \setminus X} f\right], \qquad (6)$$

where $\tilde{g} := g|_R$ is the restriction of g to R.

Proof. The proof is by induction on m. For $m = 2$ and $Y = [y]$, Eq. (6) states that

$$\Lambda_y(g \circ f) = f(y)\tilde{g} + g(y) f$$

for any $f \in \mathcal{D}_1(R,S)$ and $g \in \mathcal{D}_1(S,T)$, which agrees with (5) for the case $j = 1, m = 2$. Therefore (6) holds for $m = 2$.

Now let $M \geq 2$ and suppose (6) is valid for $2 \leq m \leq M$. We will prove (6) is valid for $m = M+1$. To this end, let $Y = [y_1, \ldots, y_{M-1}]$ be an unordered list of $M-1$ increments taken from R, let $Y' = Y \sqcup [y_M]$ be the unordered list $[y_1, \ldots, y_M]$, and suppose $f \in \mathcal{D}_j(R,S)$ and $g \in \mathcal{D}_{M+1-j}(S,T)$ for some $j \in \{1, \ldots, M\}$. Then using (5) from Proposition 4.1 we calculate that

$$\begin{aligned}
\Lambda_{Y'}(g \circ f) &= \Lambda_{Y \sqcup [y_M]}(g \circ f) = \Lambda_Y \circ \Lambda_{y_M}(g \circ f) \\
&= \Lambda_Y [g(y_M) f + f(y_M)\tilde{g} + \Lambda_{f(y_M)}\tilde{g} + (\Lambda_{y_M} g) \circ f + g \circ (\Lambda_{y_M} f)] \\
&= g(y_M)\Lambda_Y f + f(y_M)\Lambda_Y \tilde{g} + \Lambda_{Y \sqcup [f(y_M)]}\tilde{g} \\
&\quad + \Lambda_Y [(\Lambda_{y_M} g) \circ f + g \circ (\Lambda_{y_M} f)] \\
&= g(y_M)\Lambda_Y f + f(y_M)\Lambda_Y \tilde{g} + \Lambda_Y [(\Lambda_{y_M} g) \circ f + g \circ (\Lambda_{y_M} f)]. \qquad (7)
\end{aligned}$$

In the last step we used $\Lambda_{Y \sqcup [f(y_M)]}\tilde{g} = 0$ because either $f(y_M) = 0$ or not. In the first case we apply Lemma 3.2, and in the second case we combine two facts: (i) $\Lambda_{Y \sqcup [f(y_M)]}\tilde{g}$ is a Leibniz difference of $\tilde{g}$ of order M, and (ii) $g \in \mathcal{D}_{M+1-j}(S,T)$ implying $\tilde{g} \subseteq \mathcal{D}_M(R,T)$ since $j \geq 1$.

Next observe that $(\Lambda_{y_M} g) \circ f$ is the composition of $\Lambda_{y_M} g \in \mathcal{D}_{M-j}(S,T)$ with $f \in \mathcal{D}_j(R,S)$. In the case $j = M$ we have $\Lambda_{y_M} g = 0$ so $(\Lambda_{y_M} g) \circ f = 0$; otherwise $j \leq M-1$ and we can apply the inductive hypothesis to the term $(\Lambda_{y_M} g) \circ f$. Also, $g \circ (\Lambda_{y_M} f)$ is the composition of $g \in \mathcal{D}_{M+1-j}(S,T)$ with $\Lambda_{y_M} f \in \mathcal{D}_{j-1}(R,S)$. In the case $j = 1$ we have $\Lambda_{y_M} f = 0$ so $g \circ (\Lambda_{y_M} f) = 0$, and in this case we also have $\Lambda_{Y\setminus X} f = 0$ for $X \in P_1(Y)$ in the second line of the calculation below. Otherwise $j \geq 2$ and we can apply the inductive hypothesis to the term $g \circ (\Lambda_{y_M} f)$. The result for (7) is that

$$\begin{aligned}
\Lambda_{Y'}(g \circ f) &= g(y_M)\Lambda_Y f + f(y_M)\Lambda_Y \tilde{g} + \Lambda_Y[(\Lambda_{y_M} g) \circ f + g \circ (\Lambda_{y_M} f)] \\
&= g(y_M)\Lambda_Y f + f(y_M)\Lambda_Y \tilde{g} + \sum_{j=1}^{M-1} \sum_{X \in P_j(Y)} \big[\Lambda f(X)\Lambda_{Y\setminus X}(\Lambda_{y_M}\widetilde{g}) \\
&\quad + \Lambda(\Lambda_{y_M} g)(X)\Lambda_{Y\setminus X} f\big] \\
&\quad + \sum_{j=1}^{M-1} \sum_{X \in P_j(Y)} \big[\Lambda(\Lambda_{y_M} f)(X)\Lambda_{Y\setminus X}\tilde{g} + \Lambda g(X)\Lambda_{Y\setminus X}(\Lambda_{y_M} f)\big] \\
&= g(y_M)\Lambda_Y f + f(y_M)\Lambda_Y \tilde{g} + \sum_{j=1}^{M-1} \sum_{X \in P_j(Y'), y_M \notin X} \big[\Lambda f(X)\Lambda_{Y'\setminus X}\tilde{g} \\
&\quad + \Lambda g(X)\Lambda_{Y'\setminus X} f\big] \\
&\quad + \sum_{j=1}^{M-1} \sum_{X \in P_j(Y)} \big[\Lambda(\Lambda_{y_M} g)(X)\Lambda_{Y\setminus X} f + \Lambda(\Lambda_{y_M} f)(X)\Lambda_{Y\setminus X}\tilde{g}\big], \quad (8)
\end{aligned}$$

where we have used in the last step that when $X \in P_j(Y)$, the quantities $\Lambda_{Y\setminus X}(\Lambda_{y_M}\widetilde{g})$ and $\Lambda_{Y\setminus X}(\Lambda_{y_M} f)$ can be replaced respectively by $\Lambda_{Y'\setminus X}\tilde{g}$ and $\Lambda_{Y'\setminus X} f$ for $X \in P_j(Y')$ with $y_M \notin X$.

Let us examine the last two summands more closely. Using Definition 3.1 (d) for $X = [x_1, \ldots, x_j] \in P_j(Y)$ and for all $1 \leq j \leq M-1$, we get

$$\begin{aligned}
\Lambda(\Lambda_{y_M} f)(X) &= [\Lambda_{x_1,\ldots,x_{j-1}} \circ (\Lambda_{y_M} f)](x_j) = \Lambda_{x_1,\ldots,x_{j-1},y_M} f(x_j) \\
&= \Lambda f(X \sqcup [y_M]),
\end{aligned}$$

and similarly

$$\Lambda(\Lambda_{y_M} g)(X) = \Lambda g(X \sqcup [y_M]).$$

Using these simplifications in (8), we find that

$$\begin{aligned}\Lambda_{Y'}(g\circ f) =& g(y_M)\Lambda_Y f + f(y_M)\Lambda_Y \tilde{g} + \sum_{j=1}^{M-1} \sum_{X\in P_j(Y'), y_M\notin X} \big[\Lambda f(X)\Lambda_{Y'\setminus X}\tilde{g} \\ &+\Lambda g(X)\Lambda_{Y'\setminus X} f\big] \\ &+\sum_{j=1}^{M-1}\sum_{X\in P_j(Y)} \big[\Lambda g(X\sqcup[y_M])\Lambda_{Y\setminus X} f + \Lambda f(X\sqcup[y_M])\Lambda_{Y\setminus X}\tilde{g}\big] \\ =&\sum_{j=1}^{M-1}\sum_{X\in P_j(Y'), y_M\notin X} \big[\Lambda f(X)\Lambda_{Y'\setminus X}\tilde{g} + \Lambda g(X)\Lambda_{Y'\setminus X} f\big] \\ &+ g(y_M)\Lambda_Y f + f(y_M)\Lambda_Y\tilde{g} \\ &+\sum_{j=2}^{M}\sum_{X'\in P_j(Y'), y_M\in X'} \big[\Lambda g(X')\Lambda_{Y'\setminus X'} f + \Lambda f(X')\Lambda_{Y'\setminus X'}\tilde{g}\big] \\ =&\sum_{j=1}^{M-1}\sum_{X\in P_j(Y'), y_M\notin X} \big[\Lambda f(X)\Lambda_{Y'\setminus X}\tilde{g} + \Lambda g(X)\Lambda_{Y'\setminus X} f\big] \\ &+\sum_{j=1}^{M}\sum_{X'\in P_j(Y'), y_M\in X'} \big[\Lambda g(X')\Lambda_{Y'\setminus X'} f + \Lambda f(X')\Lambda_{Y'\setminus X'}\tilde{g}\big] \\ =&\sum_{j=1}^{M}\sum_{X\in P_j(Y')} \big[\Lambda f(X)\Lambda_{Y'\setminus X}\tilde{g} + \Lambda g(X)\Lambda_{Y'\setminus X} f\big].\end{aligned}$$

In the last line we used the fact that for $j = M$ the only element of $P_M(Y')$ is the complete list Y' itself, which includes y_M.

Therefore (6) holds for $m = M + 1$ and the proof is finished. □

Now we are ready for the main result. (For the reason why we restrict to the case $R = S$ here, see Example 5.3 below.)

Theorem 4.3. *Let $R \subseteq T$ be integral domains, suppose $m, k \in \mathbb{N}$ with $m \geq 2$ and $1 \leq k \leq m-1$, and suppose char$(R) > m$. If $f \in \mathcal{D}^*_k(R)$ and $g \in \mathcal{D}^*_{m-k}(R,T)$, then $g \circ f \in \mathcal{D}^*_m(R,T)$.*

Proof. For a contradiction, suppose there exist an $f \in \mathcal{D}^*_k(R)$ and a $g \in \mathcal{D}^*_{m-k}(R,T)$ such that $g \circ f \in \mathcal{D}_{m-1}(R,T)$. Let $Y = [y_1, \ldots, y_{m-1}]$ be an unordered list of $m-1$ increments taken from R. Then by Proposition 4.2 we have (note that $\tilde{g} = g$ since $S = R$)

$$0 = \Lambda_Y(g\circ f) = \sum_{n=1}^{m-1}\sum_{X\in P_n(Y)} \big[\Lambda f(X)\Lambda_{Y\setminus X} g + \Lambda g(X)\Lambda_{Y\setminus X} f\big].$$

Considering the terms $\Lambda f(X)\Lambda_{Y\setminus X}g$ for $X=[x_1,\ldots,x_n]\in P_n(Y)$, we have

(i) $\Lambda f(X)=\Lambda_{x_1,x_2,\ldots,x_{n-1}}f(x_n)=0$ for $n-1\geq k$, since $f\in\mathcal{D}_k(R)$; and
(ii) $\Lambda_{Y\setminus X}g=0$ if length $|Y|-|X|=m-1-n\geq m-k$, since $g\in\mathcal{D}_{m-k}(R,T)$.

Hence these terms vanish when $n\geq k+1$ or $n\leq k-1$, and the only surviving terms are the ones for $n=k$. A similar analysis of the terms $\Lambda g(X)\Lambda_{Y\setminus X}f$ shows that they all vanish except for the ones with $n=m-k$. Thus we are left with

$$0=\sum_{X\in P_k(Y)}\Lambda f(X)\Lambda_{Y\setminus X}g\ +\sum_{X\in P_{m-k}(Y)}\Lambda g(X)\Lambda_{Y\setminus X}f. \tag{9}$$

By Proposition 2.4, since $f\notin\mathcal{D}_{k-1}(R)$, there exists an element $p\in R$ such that $a=\Lambda_p^{k-1}f(p)\neq 0$. Putting $y_i=p$ for all $1\leq i\leq m-1$ in (9) we get

$$0=\binom{m-1}{k}a\Lambda_p^{m-1-k}g\ +\binom{m-1}{m-k}\Lambda_p^{m-1-k}g(p)\Lambda_p^{k-1}f.$$

Multiplying by $k!(m-k)!$, this becomes

$$0=(m-1)![(m-k)a\Lambda_p^{m-1-k}g\ +k\Lambda_p^{m-1-k}g(p)\Lambda_p^{k-1}f]. \tag{10}$$

Evaluating this equation at p we have

$$\begin{aligned}0&=(m-1)![(m-k)a\Lambda_p^{m-1-k}g(p)\ +k\Lambda_p^{m-1-k}g(p)\Lambda_p^{k-1}f(p)]\\&=m!a\Lambda_p^{m-1-k}g(p).\end{aligned}$$

Since $char(R)>m$ (thus $char(T)>m$), as noted above multiplication by $m!$ is injective in T. Since $a\neq 0$, this implies $\Lambda_p^{m-1-k}g(p)=0$. Inserting this into (10) we get

$$0=(m-1)!(m-k)a\Lambda_p^{m-1-k}g.$$

But multiplication by $(m-1)!(m-k)$ is also injective in T because $1\leq k<m$, hence we arrive at the conclusion $\Lambda_p^{m-1-k}g=0$, which contradicts the fact that $g\in\mathcal{D}^*_{m-k}(R)$.

This completes the proof of the theorem. □

The theorem just proved contains the following corollary, which improves Theorem 3.6 in [2] and also includes Theorem 1.2 in [5].

Corollary 4.4. *Let R be an integral domain, suppose $n\in\mathbb{N}$ with $n\geq 2$, and suppose $char(R)>n$. If $d_1,\ldots,d_n\in\mathcal{D}^*_1(R)$, then $d_n\circ\cdots\circ d_1\in\mathcal{D}^*_n(R)$.*

Proof. In Theorem 4.3 take $k=1$, $f=d_1$ and successively $m=2,\ldots,n$ with $g=d_m\circ\cdots\circ d_2$ at each step. □

Another simple inductive argument leads to the following extension of Theorem 4.3.

Corollary 4.5. *Let R be an integral domain, suppose $m \in \mathbb{N}$ with $m \geq 2$, let $n_1, \ldots, n_m \in \mathbb{N}$, and suppose $char(R) > (n_1 + \cdots + n_m)$. If $f_i \in \mathcal{D}^*_{n_i}(R)$ for each $1 \leq i \leq m$, then $f_m \circ \cdots \circ f_1 \in \mathcal{D}^*_{n_1+\cdots+n_m}(R)$.*

5. Examples

We close the paper with three examples illustrating the need for our assumptions in Theorem 4.3. The first shows the necessity of $char(R) > m$, even if R is an integral domain.

Example 5.1. Let p be a prime number. Then the polynomial ring $R = \mathbb{Z}_p[x]$ is an integral domain with characteristic p. Let $d : R \to R$ be the derivative function defined by $d(f) := f'$ for $f \in R$, and define higher derivatives $d^2, d^3, \ldots$ by iteration. Then for $1 \leq j \leq p-1$ we have $d^j \in \mathcal{D}^*_j(R)$. On the other hand for any polynomial $f \in R$, say $f(x) = \sum_j a_j x^j$ with finitely many nonzero $a_j \in \mathbb{Z}_p$, we have

$$d^p(f) = d^p\left(\sum_j a_j x^j\right) = \sum_{j \geq p} p! a_j x^{j-p} = 0,$$

so $d^p = 0$.

The second example shows the necessity of R being an integral domain, even if $char(R) = 0$.

Example 5.2. Let R be the quotient ring $\mathbb{Z}[x, y]/(xy)$, where (xy) is the ideal generated by xy. Let $m \in \mathbb{N}$ with $m \geq 2$, and suppose $j \in \{1 \ldots, m-1\}$. Define $d_1, d_2 : R \to R$ by

$$d_1(p) := \frac{\partial^j p}{\partial x^j}, \qquad d_2(p) := \frac{\partial^{m-j} p}{\partial y^{m-j}}.$$

It is easy to see that $d_1 \in \mathcal{D}^*_j(R)$ and $d_2 \in \mathcal{D}^*_{m-j}(R)$, yet $d_2 \circ d_1 = 0$.

Finally, one might ask why we restrict to the case $R = S$ in Theorem 4.3 rather than working on a "tower" of integral domains $R \subseteq S \subseteq T$. The next example shows that the corresponding statement on such a tower is not true if $S \neq R$, not even for fields of characteristic 0.

Example 5.3. Let $R = \mathbb{Q}(x)$ be the field of rational functions in x with coefficients from $\mathbb{Q}$, and let $S = \mathbb{Q}(x, y)$ be the field (extension of R) of rational functions in two variables x, y. Define $D_1 : R \to S$ and $D_2 : S \to S$ by

$$D_1(f) := f' \quad \text{and} \quad D_2(g) := \frac{\partial g}{\partial y},$$

for all rational functions $f = f(x) \in \mathbb{Q}(x)$ and $g = g(x, y) \in \mathbb{Q}(x, y)$. Clearly $D_1 \in \mathcal{D}^*_1(R, S)$ and $D_2 \in \mathcal{D}^*_1(S, S)$. Nevertheless $D_2 \circ D_1 = 0$.

Acknowledgements

I am very grateful to the reviewer for several comments and suggestions which have clarified and improved the presentation of these results.

References

[1] Ebanks, B.: Characterizing ring derivations of all orders via functional equations: results and open problems. Aequ. Math. **89**, 685–718 (2015)

[2] Ebanks, B.: Derivations and Leibniz differences on rings. Aequ. Math. (2018) https://doi.org/10.1007/s00010-018-0601-4

[3] Ebanks, B.: Functional equations characterizing derivations: a synthesis. Results Math **73**(3), 13 (2018). https://doi.org/10.1007/s00025-018-0881-y. Art. 120

[4] Gselmann, E., Kiss, G., Vincze, C.: On functional equations characterizing derivations: methods and examples. Results Math. **73**(2), 27 (2018). https://doi.org/10.1007/s00025-018-0833-6. Art. 74

[5] Kiss, G., Laczkovich, M.: Derivations and differential operators on rings and fields. Ann. Univ. Sci. Budapest. Sect. Comput. **48**, 31–43 (2018)

Bruce Ebanks
Department of Mathematics
University of Louisville
Louisville KY40292
USA
e-mail: ebanks1950@gmail.com

Received: September 1, 2018
Revised: December 7, 2018