The divergence of polynomial derivations

Andrzej Nowicki

N. Copernicus University, Faculty of Mathematics and Computer Science,
ul. Chopina 12–18, 87–100 Toruń, Poland
(e-mail: anow@mat.uni.torun.pl)

December 29, 1993

Abstract

Let \( k \) be a ring containing \( \mathbb{Q} \). A \( k \)-derivation \( d \) of \( k[X] = k[x_1, \ldots, x_n] \) is called special if the divergence of \( d \) is equal to zero. We prove the following three theorems (up to some additional assumption on \( k \)):

1. Locally nilpotent \( k \)-derivations of \( k[X] \) are special.
2. Every derivation of a commutative basis of \( \text{Der}_k(k[X]) \) is special.
3. If \( n = 2 \) and \( d \neq 0 \) is a primitive \( k \)-derivation of \( k[X] \) then the ring of constants of \( d \) is nontrivial if and only if there exists \( 0 \neq h \in k[X] \) such that the derivation \( hd \) is special.

Let \( k \) be a commutative ring containing the field \( \mathbb{Q} \) of rational numbers, let \( k[X] = k[x_1, \ldots, x_n] \) be the polynomial ring over \( k \), and let \( d : k[X] \rightarrow k[X] \) be a \( k \)-derivation of \( k[X] \). Denote by \( d^* \) the divergence of \( d \), that is,

\[
d^* = \frac{\partial d(x_1)}{\partial x_1} + \cdots + \frac{\partial d(x_n)}{\partial x_n}.
\]

The derivation \( d \) is said to be special if \( d^* = 0 \).

It is well known (see [1], [2], [9]) that if \( d \) is locally finite and \( k \) is reduced (i.e., \( k \) has no nonzero nilpotent elements), then \( d^* \) is an element from \( k \). In this note we shall show that if \( k \) is reduced and \( d \) is locally nilpotent then \( d \) is special.

Consider the \( k[X] \)-module \( \text{Der}_k(k[X]) \) of all \( k \)-derivations of \( k[X] \). This module is free and the set \( \{ \frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n} \} \) is a one of its bases. We say that a basis \( \{ d_1, \ldots, d_n \} \) of \( \text{Der}_k(k[X]) \) is locally nilpotent if each \( d_i \) is locally nilpotent. In [5] we proved that the Jacobian Conjecture is true in the \( n \)-variable case if and only if every commutative basis of \( \text{Der}_k(k[X]) \) is locally nilpotent. We shall show that if \( \{ d_1, \ldots, d_n \} \) is a commutative basis of \( \text{Der}_k(k[X]) \), then each \( d_i \) is special.

Assume now that \( n = 2 \) and \( k \) is a field (of characteristic zero). Denote by \( k[X]^d \) the ring of constants with respect to \( d \), that is, \( k[X]^d = \text{Ker} \, d \). It is obvious that if \( h \) is a nonzero element of \( k[X] \) then \( k[X]^d = k[X]^{hd} \). In Section 5 we prove that if \( d \) is nonzero and the polynomials \( d(x_1) \) and \( d(x_2) \) are relatively prime, then \( k[X]^d \neq k \) if and only if there exists a nonzero element \( h \in k[X] \) such that the \( k \)-derivation \( hd \) is special.
1 Some properties of the divergence.

Let us start from the following easy

**Proposition 1.1.** If \( d, \delta \in \text{Der}_k(k[X]) \) and \( r \in k[X] \), then:

1. \((d + \delta)^* = d^* + \delta^*\).
2. \((rd)^* = rd^* + d(r)\).
3. \([d, \delta]^* = d(\delta^*) - \delta(d^*)\). \(\square\)

The partial derivatives are special derivations. The above proposition implies that the set of all special derivations of \( k[X] \) is closed under the sum and the Lie product.

Let us denote by \([h_1, \ldots, h_n] \) the jacobian of \( h_1, \ldots, h_n \in k[X] \), that is,

\[
[h_1, \ldots, h_n] = \det \left[ \frac{\partial h_i}{\partial x_j} \right].
\]

**Proposition 1.2.** Let \( d : k[X] \to k[X] \) be a \( k \)-derivation and let \( h_1, \ldots, h_n \in k[X] \). Then

\[
d([h_1, \ldots, h_n]) = -[h_1, \ldots, h_n]d^* + \sum_{p=1}^{n} [h_1, \ldots, d(h_p), \ldots, h_n].
\]

**Proof.** Put \( f_i = d(x_i), f_{ij} = \frac{\partial f_i}{\partial x_j}, h_{ij} = \frac{\partial h_i}{\partial x_j} \), for all \( i, j \in \{1, \ldots, n\} \), and let \( S_n \) denote the group of all permutations of \( \{1, \ldots, n\} \). First observe that

\[
d(h_{\sigma(p)p}) = \frac{\partial}{\partial x_p} d(h_{\sigma(p)}) - \sum_{q=1}^{n} h_{\sigma(p)q} f_{qp}, \tag{1.1}
\]

for all \( \sigma \in S_n \) and \( p \in \{1, \ldots, n\} \). Observe also that

\[
\sum_{\sigma \in S_n} (-1)^{|\sigma|} h_{\sigma(1)1} \cdots h_{\sigma(p-1)(p-1)} h_{\sigma(p)q} h_{\sigma(p+1)(p+1)} \cdots h_{\sigma(n)n} = [h_1, \ldots, h_n] \delta_{pq}, \tag{1.2}
\]

for all \( p, q \in \{1, \ldots, n\} \), where \(|\sigma|\) is the sign of \( \sigma \), and \( \delta_{pq} \) is the Kronecker delta. Now, using these observations, we get:

\[
d([h_1, \ldots, h_n]) = \sum_{p=1}^{n} \sum_{\sigma \in S_n} (-1)^{|\sigma|} h_{\sigma(1)1} \cdots d(h_{\sigma(p)p}) \cdots h_{\sigma(n)n}
\]

\[
\overset{(1.1)}{=} \sum_{p=1}^{n} \sum_{\sigma \in S_n} (-1)^{|\sigma|} h_{\sigma(1)1} \cdots \left[ \frac{\partial}{\partial x_p} d(h_{\sigma(p)}) - \sum_{q=1}^{n} h_{\sigma(p)q} f_{qp} \right] \cdots h_{\sigma(n)n}
\]

\[
\overset{(1.2)}{=} \sum_{p=1}^{n} [h_1, \ldots, d(h_p), \ldots, h_n] - \sum_{p=1}^{n} \sum_{q=1}^{n} f_{pq} [h_1, \ldots, h_n] \delta_{pq}
\]

\[
= \sum_{p=1}^{n} [h_1, \ldots, d(h_p), \ldots, h_n] - \sum_{p=1}^{n} f_{pp} [h_1, \ldots, h_n]
\]

\[
= \sum_{p=1}^{n} [h_1, \ldots, d(h_p), \ldots, h_n] - [h_1, \ldots, h_n]d^*.
\]
This completes the proof. □

**Corollary 1.3.** If \( d \) is a special \( k \)-derivation of \( k[X] \) and \( h_1, \ldots, h_n \in k[X] \), then

\[
d([h_1, \ldots, h_n]) = \sum_{p=1}^{n} [h_1, \ldots, d(h_p), \ldots, h_n]. \quad □
\]

## 2 Automorphism \( E_d \)

Let \( A \) be a \( k \)-algebra (commutative with 1), let \( A[[t]] \) be the power series ring over \( A \), and let \( d \) be a \( k \)-derivation of \( A \). Denote by \( \tilde{d} \) the \( k[[t]] \)-derivation of \( A[[t]] \) defined by

\[
\tilde{d} \left( \sum_{p=0}^{\infty} a_p t^p \right) = \sum_{p=0}^{\infty} d(a_p) t^p,
\]

and set

\[
E_d(\varphi) = \sum_{p=0}^{\infty} \frac{1}{p!} \tilde{d}^p(\varphi) t^p
\]

for all \( \varphi \in A[[t]] \).

It is well known that \( E_d \) is a \( k[[t]] \)-automorphism of \( A[[t]] \), which is very useful in the differential algebra (see, for example: [7], [8], [4], [3]).

Now let \( A = k[X] = k[x_1, \ldots, x_n] \) and let \( J \) be the jacobian of \( E_d(x_1), \ldots, E_d(x_n) \), that is,

\[
J = [E_d(x_1), \ldots, E_d(x_n)] = \det \left[ \frac{\partial E_d(x_i)}{\partial x_j} \right].
\]

The jacobian \( J \) is an element of \( k[X][[t]] \). Let us set \( J = \sum_{p=0}^{\infty} \frac{1}{p!} B_p t^p \), where each \( B_p \) is in \( k[X] \).

Assume that \( n = 2 \). Put \( x = x_1, y = x_2 \) and let \( f_x, f_y \) denote \( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \), respectively, for any \( f \in k[x, y] \). In such a case we have:

\[
J = E_d(x) E_d(y) - E_d(x) E_d(y) = E_d(x(x) E_d(y) x)
\]

\[
= \left( \sum_{p=0}^{\infty} \frac{1}{p!} d^p(x) x t^p \right) \left( \sum_{p=0}^{\infty} \frac{1}{p!} d^p(y) y t^p \right) - \left( \sum_{p=0}^{\infty} \frac{1}{p!} d^p(x) y t^p \right) \left( \sum_{p=0}^{\infty} \frac{1}{p!} d^p(y) x t^p \right)
\]

\[
= \sum_{p=0}^{\infty} \left( \sum_{i+j=p} \frac{1}{i! j!} \left( d^i(x) x d^j(y) y - d^i(x) y d^j(y) x \right) \right) t^p
\]

\[
= \sum_{p=0}^{\infty} \left( \sum_{i+j=p} \frac{1}{i! j!} [d^i(x), d^j(y)] \right) t^p
\]

Thus, we see that if \( n = 2 \), then

\[
B_p = \sum_{i+j=p} \langle i, j \rangle [d^i(x), d^j(y)],
\]

for every \( p \geq 0 \). If \( n \) is arbitrary then, repeating the same argument, we get
Proposition 2.1. For every \( p \geq 0 \),
\[
B_p = \sum_{i_1+\ldots+i_n=p} \langle i_1, \ldots, i_n \rangle [d^{i_1}(x_1), \ldots, d^{i_n}(x_n)],
\]
where
\[
\langle i_1, \ldots, i_n \rangle = \frac{(i_1 + \cdots + i_n)!}{i_1! \cdots i_n!}.
\]

Let us continue are calculations for \( n = 2 \). From Propositions 2.1 and 1.2 we obtain
the following generalization of this equality:
\[
d(B_p) = \sum_{i+j=p} \langle i, j \rangle d([d^i(x), d^j(y)])
= \sum_{i+j=p} \langle i, j \rangle ([d^{i+1}(x), d^j(y)] + [d^i(x), d^{j+1}(y)] - [d^i(x), d^j(y)]d^*)
= -B_p d^* + \sum_{i+j=p} \langle i, j \rangle ([d^{i+1}(x), d^j(y)] + [d^i(x), d^{j+1}(y)]).
\]

Now, by standard formulas and Proposition 2.1, we get:
\[
d(B_p) + B_p d^* = [d^{p+1}(x), y] + [x, d^{p+1}(y)]
+ \sum_{i=0}^{p-1} \langle i, p - i \rangle [d^{p+1}(x), d^{p-i}(y)] + \sum_{i=1}^{p} \langle i, p - i \rangle [d^{i}(x), d^{p+1-i}(y)]
= [d^{p+1}(x), y] + [x, d^{p+1}(y)]
+ \sum_{i=1}^{p} (\langle i - 1, p + 1 - i \rangle + \langle i, p - i \rangle) [d^{i}(x), d^{p+1-i}(y)]
= \sum_{i+j=p+1} \langle i, j \rangle [d^{i}(x), d^{j}(y)]
= B_{p+1}.
\]

We used the well known equality: \( \langle i - 1, j \rangle + \langle i, j \rangle = \langle i, j + 1 \rangle \) (where \( i \geq 1 \)). There exists
the following generalization of this equality:
\[
\langle i_1, \ldots, i_n \rangle + \sum_{j=1}^{n-1} \langle i_1, \ldots, i_j \rangle (\langle i_1, \ldots, i_{j-1}, i_{j+1}, \ldots, i_n \rangle - \langle i_1, \ldots, i_{n-1}, i_n \rangle - \langle i_1, \ldots, i_n \rangle) = \langle i_1, \ldots, i_n \rangle + 1, \quad (2.1)
\]
for \( i_1, \ldots, i_{n-1} \geq 1 \).

Now, using (2.1), Propositions 2.1, 1.2 and repeating the same arguments, one can
easily deduce the following

Theorem 2.2. Let \( d \) be a \( k \)-derivation of \( k[X] = k[x_1, \ldots, x_n] \) and let
\[
[E_d(x_1), \ldots, E_d(x_n)] = \sum_{p=0}^{\infty} \frac{1}{p!} B_p d^p,
\]
where \( B_p \in k[X] \). Then \( B_0 = 1, B_1 = d^* \) and \( B_{p+1} = B_p d^* + d(B_p) \) for all \( p \geq 0 \). □
Corollary 2.3. Let \( d \) be a \( k \)-derivation of \( k[X] \). If \( d^* \in k \) then
\[
[E_d(x_1), \ldots, E_d(x_n)] = \sum_{p=0}^{\infty} \frac{1}{p!} b^p t^p = e^{bt},
\]
where \( b = d^* \). \( \square \)

Corollary 2.4. If \( d \) is a special \( k \)-derivation of \( k[X] \), then
\[
[E_d(x_1), \ldots, E_d(x_n)] = 1. \quad \square
\]

3 The divergence of locally finite derivations.

Let us recall that if \( A \) is a commutative \( k \)-algebra and \( d \) is a \( k \)-derivation of \( A \), then \( d \) is called locally nilpotent if for each \( r \in A \) there exists a natural number \( s \) such that \( d^s(r) = 0 \), and is called locally finite if for any \( r \in A \) there exists a finite generated \( k \)-module \( M \subseteq A \) such that \( r \in M \) and \( d(M) \subseteq M \).

It is easy to check the following two propositions

Proposition 3.1. If \( d \) is a \( k \)-derivation of \( k[X] \) then, \( d \) is locally nilpotent if and only if \( E_d(k[t][X]) \subseteq k[t][X] \). \( \square \)

Proposition 3.2. If \( d \) is a \( k \)-derivation of \( k[X] \) then the following conditions are equivalent:

1. \( d \) is locally finite;
2. there exists a natural \( s \) such that \( \deg d^p(x_i) \leq s \), for all \( p \geq 0 \) and \( i = 1, \ldots, n \);
3. \( E_d(k[[t]][X]) \subseteq k[[t]][X] \). \( \square \)

The following result is due to H. Bass, G. Meisters [1] and B. Coomes, V. Zurkowski [2]. An another proof of this fact is given in [9].

Theorem 3.3. Let \( k \) be a reduced ring containing \( \mathbb{Q} \) and let \( k[X] = k[x_1, \ldots, x_n] \) be a polynomial ring over \( k \). If \( d \) is a locally finite \( k \)-derivation of \( k[X] \) then \( d^* \), the divergence of \( d \), is an element of \( k \).

Proof ([1]). Since \( d \) is locally finite, the series \( E_d(x_1), \ldots, E_d(x_n) \) are elements of \( k[[t]][X] \) (Proposition 3.2). By the same way, since the derivation \( -d \) is locally finite, the series \( E_{-d}(x_1), \ldots, E_{-d}(x_n) \) are also elements of \( k[[t]][X] \). Hence, the mapping \( E_d \) is a \( k[[t]] \)-automorphism of the polynomial ring \( k[[t]][X] \), and hence \( J = [E_d(x_1), \ldots, E_d(x_n)] \), the jacobian of \( E_d \), is an invertible element of \( k[[t]][X] \). But \( k \) is reduced, so \( J \in k[[t]] \).

We know, by Theorem 2.2, that \( J = 1 + d^* t^1 + \cdots \). Therefore \( d^* \in k \). \( \square \)

If \( k \) is non-reduced then the above property does not hold, in general.

Example 3.4. Let \( k = \mathbb{Q}[y]/(y^2) \) and let \( d \) be the \( k \) derivation of \( k[x] \) (a polynomial ring in a one variable) defined by \( d(x) = a x^2 \), where \( a = y + (y^2) \). Since \( d^2(x) = 2ax^3 = 0 \), \( d \) is locally finite. But \( d^* = 2ax \notin k \). \( \square \)
Now we are ready to prove the main result of this section.

**Theorem 3.5.** Let $k$ be a reduced ring containing $\mathbb{Q}$. If $d$ is a locally nilpotent $k$-derivation of $k[X]$ then $d^* = 0$.

**Proof.** Put $b = d^*$ and let $J$ be the jacobian of $E_d(x_1), \ldots, E_d(x_n)$. We know, by Theorem 3.3 and Corollary 2.3, that $b \in k$ and $J = e^{tb} = \sum_{p=0}^{\infty} \frac{1}{p!} b^p$. Since $d$ is locally nilpotent, the series $E_d(x_1), \ldots, E_d(x_n)$ are polynomials from $k[X][t]$ and hence $J \subseteq k[X][t]$. Thus $b^p = 0$, for some $p$, and consequently (since $k$ is reduced), $b = 0$. □

The derivation $d$ from Example 3.4 is locally nilpotent. This means that if $k$ is non-reduced then there exist locally nilpotent $k$-derivations of $k[X]$ with a nonzero divergence.

### 4 Commutative bases of derivations.

**Theorem 4.1.** Let $k$ be a reduced ring containing $\mathbb{Q}$ and let $k[X] = k[x_1, \ldots, x_n]$ be the polynomial ring over $k$. If $\{d_1, \ldots, d_n\}$ is a commutative basis of $\text{Der}_k(k[X])$ then $d_i^* = 0$, for all $i = 1, \ldots, n$.

**Proof.** Denote by $A$ the matrix $[d_i(x_j)]$ and let $w = \det(A)$. Then $w$ is an invertible element of $k[X]$ ([5] Proposition 1) and so, since $k$ is reduced, $w$ is an invertible element of $k$.

Let $[b_{ij}]$ be the matrix such that $\frac{\partial}{\partial x_i} = \sum_{j=1}^{n} b_{ij}d_j$, for $i = 1, \ldots, n$. Then $\sum_{j=1}^{n} b_{ij}d_j(x_i) = \delta_{ij}$. Thus, $[b_{ij}] = A^{-1}$ and we have: $b_{ij} = (-1)^{i+j}A_{ij}w^{-1}$, where $A_{pq}$ is the determinant of the $(n-1) \times (n-1)$ matrix obtained from $A$ by removing in $A$ the $p$-th row and the $q$-th column. Set

$$D = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} \quad \text{and} \quad D(r) = \begin{bmatrix} d_1(r) \\ \vdots \\ d_n(r) \end{bmatrix}, \quad \text{for} \ r \in R.$$ 

Then we have $A = [D(x_1), \ldots, D(x_n)]$ and

$$\frac{\partial}{\partial x_i} = w^{-1} \det[D(x_1), \ldots, D(x_{i-1}), D, D(x_{i+1}), \ldots, D(x_n)],$$

for all $i = 1, \ldots, n$. Therefore, if $p \in \{1, \ldots, n\}$ then

$$0 = w^{-1}d_p(w)$$
$$= w^{-1}d_p(\det[D(x_1), \ldots, D(x_n)])$$
$$= w^{-1}\sum_{i=1}^{n} d_p \det[D(x_1), \ldots, d_pD(x_i), \ldots, D(x_n)]$$
$$= w^{-1}\sum_{i=1}^{n} d_p \det[D(x_1), \ldots, D(d_p(x_i)), \ldots, D(x_n)]$$
$$= \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(d_p(x_i))$$
$$= d_p^*.$$ 

This completes the proof. □

The following example shows that the divergence of non-commutative basis is not a constant, in general.
Example 4.2. Let

\[ d_1 = \frac{\partial}{\partial x_1} + x_2^2 \frac{\partial}{\partial x_2}, \quad d_2 = \frac{\partial}{\partial x_2}, \ldots, \quad d_n = \frac{\partial}{\partial x_n}. \]

Then \( \{d_1, \ldots, d_n\} \) is a basis of \( \text{Der}_k(k[x_1, \ldots, x_n]) \) (since \( \det[d_i(x_j)] = 1 \)), and \( d_1^* = 2x_2 \notin k. \) □

5 The ring of constants in two variables

Let \( d \) be a \( k \)-derivation of the polynomial ring \( k[x, y] \). Put \( d(x) = P, \ d(y) = Q \) and consider \( k[x, y]^d \), the ring of constants of \( k[x, y] \) with respect to \( d \). It would be of considerable interest to find necessary and sufficient conditions on \( P \) and \( Q \) for \( k[x, y]^d \) to be nontrivial (that is, \( k[x, y]^d \neq k \)). The problem seems to be difficult. We may of course assume that the polynomials \( P \) and \( Q \) are relatively prime. Moreover, it is clear that \( k[x, y]^d = k[x, y]^d h \) for any nonzero \( h \in k[x, y] \).

The following two theorems give certain information concerning to our problem.

Theorem 5.1. Let \( k \) be a field of characteristic zero and let \( d \) be a nonzero \( k \)-derivation of \( k[x, y] \). Assume that the polynomials \( d(x) \) and \( d(y) \) are relatively prime. Then the following two conditions are equivalent:

1. \( k[x, y]^d \neq k; \)
2. there exists a nonzero polynomial \( h \in k[x, y] \) such that \( (hd)^* = 0 \).

Proof. (1) \( \Rightarrow \) (2). Let \( F \in k[x, y]^d \setminus k \). Put \( d(x) = P, \ d(y) = Q \) and \( h = \gcd(F_x, F_y) \). Then \( PF_x + QF_y = 0, h \neq 0 \) and there exist relatively prime polynomials \( A, B \in k[x, y] \) such that \( F_x = Ah \) and \( F_y = Bh \). Hence \( AP = -BQ \) and hence, \( A \mid Q, \ Q \mid A, \ B \mid P \) and \( P \mid B \). This implies that there exists an element \( a \in k \sim \{0\} \) such that \( A = aQ \) and \( B = -aP \). Thus, we have:

\[
(hd)^* = (hP)x + (hQ)y = -(a^{-1}hB)x + (a^{-1}hA)y = -a^{-1}F_{yx} + a^{-1}F_{xy} = 0.
\]

(2) \( \Rightarrow \) (1). Let \( 0 \neq h \in k[x, y] \) with \( (hd)^* = 0 \) and let \( \delta = hd, \delta(x) = P, \delta(y) = Q \). We may assume that \( P \neq 0 \) and \( Q \neq 0 \) (if \( P = 0 \) or \( Q = 0 \) then \( k[x, y]^d \neq k \)). Put \( f = Q \) and \( g = -P \). Then \( f_y = q_x \) and hence, by [5] Lemma 3, there exists \( H \in k[x, y] \) such that \( H_x = f \) and \( H_y = g \). It is clear that \( H \notin k \). Now observe that \( \delta(H) = 0 \). Indeed,

\[
\delta(H) = H_x\delta(x) + H_y\delta(y) = fP + gQ = QP - PQ = 0.
\]

Thus, \( H \in k[x, y]^d = k[x, y]^d \) and \( H \notin k. \) □

Using the above theorem and the results of [6] it is not difficult to prove the following

Theorem 5.2. Let \( k \) be a field of characteristic zero and let \( d, \delta \) be \( k \)-derivations of \( k[x, y] \) such that \( k[x, y]^d \neq k \) and \( k[x, y]^\delta \neq k \). Then the following two conditions are equivalent:

1. \( k[x, y]^d = k[x, y]^d; \)
2. there exist nonzero polynomials \( f, g \in k[x, y] \) such that \( fd = g\delta \). □
References


