ON $\sigma$-DERIVATIONS ON FINITE MATRIX ALGEBRAS

ANA P. MADRID

Abstract. Our aim is to characterize the class $D(\sigma)$ of $\sigma$-derivations on finite matrix algebras. Nevertheless the full description of $\sigma$-derivations is quite difficult, even in the finite dimensional case, the problem has its own interest. We will prove that $D(\sigma)$ consists of the solution of some linear system of matrix equations. Some posed questions concerning the structure of $\sigma$-derivations will also be considered.

1. Introduction

Let $A$ be an algebra over a field $K$. A linear mapping $\delta: A \to A$ is called a derivation if it satisfies the Leibniz rule $\delta(ab) = \delta(a)b + a\delta(b)$ for all $a, b \in A$. For a fixed element $a_0$ of $A$ the mapping $\delta_{a_0}: A \to A$, $\delta_{a_0}(a) = a_0a - aa_0$, is called the inner derivation implemented by $a_0$. The concept of $(\sigma, \tau)-$derivation on an algebra generalizes the usual notion of derivation and presently constitutes a matter of intensive research. Let $\sigma, \tau: A \to A$ be two linear mappings. A linear mapping $d: A \to A$ is called a $(\sigma, \tau)-$derivation if $d(ab) = \sigma(a)d(b) + d(a)\sigma(b)$ for all $a, b \in A$. If $\sigma = \tau$ we will simple say that $d$ is a $\sigma-$derivation. For instance, if $\sigma$ and $\tau$ are morphisms the mapping $d(a) = a_0\sigma(a) - \tau(a)a_0$ is called the $(\sigma, \tau)$-inner derivation implemented by $a_0 \in A$. For studies of inner $\sigma$-derivations the reader can see [2], [3]. For investigations of $(\sigma, \tau)$-derivations in Banach algebras, automatic continuity of $\sigma$-derivations in $C^*$-algebras or their ultraweak continuity on von Neumann algebras see [6], [5] or [7] respectively.

Given a $(\sigma, \tau)$-derivation $d$ it is of interest to decide:

(A) $\sigma$ and $\tau$ must be necessarily morphisms?

(B) There exists some derivation $\delta$ to that $d = \delta \circ \sigma$? (Cf. [4]).

Both problems concerns to the structure of $\sigma$-derivations and barring elementary situations it seems natural to do a first look in a finite dimensional setting. The problem is still interesting since it is difficult to handle even in
this frame. However, it will suffice for our purposes to characterize the class of finite dimensional $\sigma$-derivations and to give their full description in some concrete case (see Theorem 4).

2. $\sigma$-derivations on finite matrix algebras

2.1. First remarks. Throughout this section, if $n \in \mathbb{N}$ then $M_n(K)$ will be the algebra of $n \times n$ matrices over $K$. Let us consider the mapping

$$M_n(M_n(K)) \times M_n(K) \to M_n(K),$$

$$(a, x) \to a \cdot x = \sum_{k,h=1}^{n} x_{k,h} a_{k,h}.$$

If $\{e^{i,j}\}_{1 \leq i,j \leq n}$ is the canonical basis of $M_n(K)$, $x \in M_n(K)$ and $\sigma \in L(M_n(K))$ we see that

$$\sigma(x) = \sum_{i,j=1}^{n} \sigma_{i,j}(x)e^{i,j}$$

$$= \sum_{i,j=1}^{n} \sum_{k,h=1}^{n} x_{k,h} \sigma_{i,j}(e^{k,h})e^{i,j}$$

$$= \sum_{k,h=1}^{n} x_{k,h} \sum_{i,j=1}^{n} \sigma_{i,j}(e^{k,h})e^{i,j}$$

$$= \left[\sigma(e^{i,j})\right]_{1 \leq i,j \leq n} \cdot x.$$

The mapping $F(\sigma) = [\sigma(e^{i,j})]_{1 \leq i,j \leq n}$ defines a linear isomorphism between $L(M_n(K))$ and $M_n(M_n(K))$. So, if $\sigma, \mu \in L(M_n(K))$ then

$$F(\sigma \circ \mu) = [F(\sigma) \cdot \mu^{i,j}]_{1 \leq i,j \leq n}.$$

For, by (2) and (1) if $1 \leq i, j \leq n$ then

$$F(\sigma \circ \mu)_{i,j} = (\sigma \circ \mu)(e^{i,j}) = \sigma(\mu(e^{i,j}))$$

$$= \sigma(\sum_{k,h=1}^{n} \mu_{k,h}^{i,j} e^{k,h}) = \sum_{k,h=1}^{n} \mu_{k,h}^{i,j} \sigma(e^{k,h})$$

$$= \sum_{k,h=1}^{n} F(\sigma)^{k,h} \mu_{k,h}^{i,j} = F(\sigma) \cdot \mu^{i,j}.$$

So, if $a, b \in M_n(M_n(K))$ we set $a \ast b = F(F^{-1}(a) \circ F^{-1}(b))$. Provided with this multiplication $M_n(M_n(K))$ is an associative algebra over $K$, $F$ becomes an algebraic homomorphism and $M_n(K)$ becomes a left $M_n(M_n(K))$-module.
2.2. Characterization of $\sigma$-derivations.

**Theorem 1.** The class $D(\sigma)$ is defined as the set of solutions of a linear system of matricial equations of the following type:

\[
\delta_{j,k}d_{i,k} = d_{i,j}\sigma_{k,h} + \sigma_{i,j}d_{k,h}, \quad 1 \leq i, j, k, h \leq n,
\]

where $\delta_{j,k}$ denotes the usual Kronecker symbol.

**Proof.** Let $d \in D(\sigma), x, y \in M_n(K), 1 \leq i, j \leq n$. Then

\[
d_{i,j}(xy) = [d(x)\sigma(y) + \sigma(x)d(y)]_{i,j}
\]
\[
= \sum_{l=1}^{n} [d_{i,l}(x)\sigma_{l,j}(y) + \sigma_{i,l}(x)d_{l,j}(y)]
\]
\[
= \sum_{l=1}^{n} \left[ \sum_{1 \leq r, s, t, u \leq n} d_{r,s}^{i,l} \sigma_{t,u}^{l,j} x_{r,s} y_{t,u} + \sigma_{r,s}^{i,l} d_{t,u}^{l,j} x_{r,s} y_{t,u} \right]
\]

For $1 \leq m, p \leq n$ let $e_{i,j}^{m,p} = \{ \delta_{i,j}^{m,p} \}_{1 \leq i, j \leq n}$ be the canonical matrices of $M_n(K)$.

By the previous relations, if $1 \leq p, q, m, w, i, j \leq n$ we have

\[
(\delta_{p,q}d_{m,w})_{i,j} = \delta_{p,q}d_{i,j}^{m,w} = \delta_{p,q}d_{i,j}(e_{m,w})
\]
\[
= d_{i,j}(e_{m,p}^{m,p} e_{q,w}^{q,w}) = \sum_{l=1}^{n} \left( d_{i,l}^{m,p} \sigma_{l,j}^{q,w} + \sigma_{l,i}^{m,p} d_{l,j}^{q,w} \right)
\]
\[
= (d_{i,j}^{m,p} \sigma_{q,w}^{q,w} + \sigma_{m,p}^{m,p} d_{q,w}^{q,w})_{i,j}
\]

and (3) holds. This completes the first part of our theorem.

Now, we prove the second part of our theorem. Given a solution $\{d_{k,h}\}_{1 \leq k, h \leq n}$ of (3) and $x \in M_n(K)$ let us write

\[
d(x) = \left\{ \sum_{1 \leq k, h \leq n} d_{i,j}^{k,h} x_{k,h} \right\}_{1 \leq i, j \leq n}.
\]
Clearly $d \in L(M_n(K))$ and for $x, y \in M_n(K)$ and $1 \leq i, j \leq n$ we see that

$$[d(x)\sigma(y) + \sigma(x)d(y)]_{i, j} = \sum_{l=1}^{n} [d_{l, i}(x)\sigma_{l, j}(y) + \sigma_{l, i}(x)d_{l, j}(y)]$$

$$= \sum_{l=1}^{n} \left[ \sum_{u, v=1}^{n} d_{l, i}^{u, v} x_{u, v} \sum_{s, t=1}^{n} \sigma_{l, j}^{s, t} y_{s, t} + \sum_{u, v=1}^{n} \sigma_{l, i}^{u, v} x_{u, v} \sum_{s, t=1}^{n} d_{l, j}^{s, t} y_{s, t} \right]$$

$$= \sum_{u, v, s, t=1}^{n} x_{u, v} y_{s, t} \sum_{l=1}^{n} [d_{l, i}^{u, v} \sigma_{l, j}^{s, t} + \sigma_{l, i}^{u, v} d_{l, j}^{s, t}]$$

$$= \sum_{u, v, s, t=1}^{n} [d_{l, i}^{u, v} \sigma_{l, j}^{s, t} + \sigma_{l, i}^{u, v} d_{l, j}^{s, t}]_{i, j} x_{u, v} y_{s, t}$$

$$= \sum_{u, v, s, t=1}^{n} \delta_{v, s} d_{l, i}^{u, v} x_{u, v} y_{s, t}$$

$$= \sum_{u, t=1}^{n} d_{l, i}^{u, t} (xy)_{u, t}$$

$$= d_{l, i}^{i, j} (xy),$$

i.e. $d$ is a $\sigma$–derivation. \hfill \Box

2.3. $\sigma$-derivations on $M_2(K)$. By § 2.2 we must seek for the solutions of the matricial linear system (3). In what follows we shall do an alternative description of (3) on $M_2(M_2(K))$. This system is rather difficult to handle and the natural point of view of § 2.1 is no longer applicable. So, we point out that the products in the system of equations (4) in Theorem 2 will be the usual formal product of block matrices. In the context of the proof it will be clear what happens. We shall use the following notation:

Let $x \in M_2(K)$. Then $a_1, a_2, a_3, a_4 \in L(M_2(K))$ will be given as

$$a_1(x) = x_{11}e_{11} + x_{22}e_{21}, \quad a_2(x) = x_{11}e_{21} + x_{12}e_{22},$$

$$a_3(x) = x_{12}e_{12} + x_{22}e_{22}, \quad a_4(x) = x_{21}e_{21} + x_{22}e_{22}.$$  

If $a, b \in M_2(K)$ we also introduce $\Delta_{a, b}, m \in L(M_2(K))$ as $\Delta_{a, b}(x) = ax + xb$ and

$$m(x) = \begin{bmatrix} x_{22} & x_{21} \\ x_{12} & x_{11} \end{bmatrix}.$$  

**Theorem 2.** The system of equations (3) is equivalent to the following system on $M_2(M_2(K))$:

\[
\begin{align*}
(a) & \quad \sigma d + d\sigma = 2d \\
(b) & \quad m(\sigma)d + m(d)\sigma = 0 \\
(c) & \quad \Delta_{a_1(\sigma), a_2(\sigma)} = 0 \\
(d) & \quad \Delta_{a_3(\sigma), a_4(\sigma)} = d
\end{align*}
\]
Proof. By (3), with $j = k = 1, 2$ we set

(5) \[ d^{1,1} = d^{1,1}\sigma^{1,1} + \sigma^{1,1}d^{1,1} \]

(6) \[ d^{1,2} = d^{1,1}\sigma^{1,2} + \sigma^{1,1}d^{1,2} \]

(7) \[ d^{2,1} = d^{1,2}\sigma^{2,1} + \sigma^{1,2}d^{2,1} \]

(8) \[ d^{2,2} = d^{1,2}\sigma^{2,2} + \sigma^{1,2}d^{2,2} \]

(9) \[ d^{2,1} = d^{2,1}\sigma^{1,1} + \sigma^{2,1}d^{1,1} \]

(10) \[ d^{2,2} = d^{2,1}\sigma^{1,2} + \sigma^{2,1}d^{1,2} \]

(11) \[ d^{2,1} = d^{2,2}\sigma^{2,1} + \sigma^{2,2}d^{2,1} \]

(12) \[ d^{2,2} = d^{2,2}\sigma^{2,2} + \sigma^{2,2}d^{2,2} \]

and if $j \neq k$ in \{1, 2\} we get

(13) \[ 0 = d^{1,1}\sigma^{2,1} + \sigma^{1,1}d^{2,1} \]

(14) \[ 0 = d^{1,1}\sigma^{2,2} + \sigma^{1,1}d^{2,2} \]

(15) \[ 0 = d^{1,2}\sigma^{1,1} + \sigma^{1,2}d^{1,1} \]

(16) \[ 0 = d^{1,2}\sigma^{1,2} + \sigma^{1,2}d^{1,2} \]

(17) \[ 0 = d^{2,1}\sigma^{2,1} + \sigma^{2,1}d^{2,1} \]

(18) \[ 0 = d^{2,1}\sigma^{2,2} + \sigma^{2,1}d^{2,2} \]

(19) \[ 0 = d^{2,2}\sigma^{1,1} + \sigma^{2,2}d^{1,1} \]

(20) \[ 0 = d^{2,2}\sigma^{1,2} + \sigma^{2,2}d^{1,2} \]

Hence, by (5) and (7), (6) and (8), (9) and (11), (10) and (12) we have

\[ \sigma^{1,1}d^{1,1} + \sigma^{1,2}d^{2,1} = 2d^{1,1} - (d^{1,1}\sigma^{1,1} + d^{1,2}\sigma^{2,1}) \]

\[ \sigma^{1,1}d^{1,2} + \sigma^{1,2}d^{2,2} = 2d^{1,2} - (d^{1,1}\sigma^{1,2} + d^{1,2}\sigma^{2,2}) \]

\[ \sigma^{2,1}d^{1,1} + \sigma^{2,2}d^{2,1} = 2d^{2,1} - (d^{2,1}\sigma^{1,1} + d^{2,2}\sigma^{2,1}) \]

\[ \sigma^{2,1}d^{1,2} + \sigma^{2,2}d^{2,2} = 2d^{2,2} - (d^{2,1}\sigma^{1,2} + d^{2,2}\sigma^{2,2}) \]

i.e. $\sigma d + d\sigma = 2d$. Besides, by (13) and (15), (14) and (16), (17) and (19), (18) and (20):

(21) \[ \sigma^{1,2}d^{1,1} + \sigma^{1,1}d^{2,1} = -d^{1,2}\sigma^{1,1} - d^{1,1}\sigma^{2,1} \]

(22) \[ \sigma^{1,2}d^{1,2} + \sigma^{1,1}d^{2,2} = -d^{1,2}\sigma^{1,2} - d^{1,1}\sigma^{2,2} \]

(23) \[ \sigma^{2,2}d^{1,1} + \sigma^{2,1}d^{2,1} = -d^{2,2}\sigma^{1,1} - d^{2,1}\sigma^{2,1} \]
The equation corresponding to 
\( d \) is 
\[
\sigma^{2,1}d^{1,2} + \sigma^{2,1}d^{2,2} = -\sigma^{2,2}d^{1,2} - d^{2,1}\sigma^{2,2}
\]
that is \( m(\sigma)d + m(d)\sigma = 0 \). Now, \( \Delta_{a_1(\sigma),a_2(\sigma)}(d) = 0 \) by (15), (16), (19) and (20) while that \( \Delta_{a_3(\sigma),a_4(\sigma)}(d) \) is \( d \) by (7), (8), (11) and (12). On the other hand, a solution \( d = [d^j] \) of (4) satisfies (15), (16), (19), (20) and (7), (8), (11), (12) because (c) and (d) hold. As a consequence of (a) we get (5), (6), (9) and (10). Finally, by (b) are checked (21) to (24), which together with (15), (16), (19) and (20) implies (13), (14), (17) and (18). \( \square \)

Remark 3. The lack of symmetry in the system (4) of Th.2 is apparent. For, 
(b) is equivalent to the equation \( \sigma m(d) + dm(\sigma) = 0 \). Likewise, (c) and (d) can be replaced by the equations \( \Delta_{a_1(\sigma),a_2(\sigma)}(d) = 0 \) and \( \Delta_{a_3(\sigma),a_4(\sigma)}(d) = 0 \), where
\[
a^1(x) = x_{1,1}e^{1,2} + x_{2,1}e^{2,2}, \quad a^2(x) = x_{2,1}e^{1,1} + x_{2,2}e^{1,2},
\]
\[
a^3(x) = x_{1,1}e^{1,1} + x_{1,2}e^{1,2}, \quad a^4(x) = x_{1,1}e^{1,1} + x_{2,1}e^{2,1}.
\]

2.4. A full concrete description of \( D(\sigma) \).

Theorem 4. Let \( \sigma \in L(M_n(K)) \) be defined by \( \sigma(x) = \sum_{i=1}^nx_{i,i}e^{i,i} \) if \( x \in M_n(K) \). There are only trivial \( \sigma \)-derivations only if \( n = 2 \), when for any \( \sigma \)-derivation \( d \) there are non trivial constants \( k_1, k_2 \in K \) so that \( d(x) = k_1x_{1,2}e^{1,2} + k_2x_{2,1}e^{2,1} \) for all \( x \in M_2(K) \).

Proof. It is easy to check that \( \sigma^{i,j} = \delta_{i,j}e^{i,j} \) for \( 1 \leq i, j \leq n \). In a first place, we observe that \( D(\sigma) = \{0\} \) if \( n \neq 2 \). For, if \( n = 1 \) any \( Id_K \)-derivation becomes a derivation and in this context any derivation is null. If \( n \geq 3 \) and \( d \in D(\sigma) \) by (3) \( d^{i,h} = d^{i,j}\sigma^{j,h} + \sigma^{i,j}d^{i,h} \) for all \( 1 \leq i,j,h \leq n \). But, given indices \( i,h \) we can choose \( j \in \{1,\ldots,n\} - \{i,h\} \) to conclude that \( d^{i,h} = 0 \). Hence, we focus our attention to the case \( n = 2 \). Let us consider (3) with \( j = 1 \) and \( k = 2 \). We address the following four cases:

(i) If \( i = h = 1 \), \( \sigma^{1,1}d^{2,1} = 0 \). Hence \( d^{2,1}_{1,1} = d^{2,1}_{1,2} = 0 \).
(ii) If \( i = 1 \) and \( h = 2 \),
\[
0 = d^{1,1}\sigma^{2,2} + \sigma^{1,1}d^{2,2} = d^{2,2}_{1,1}e^{1,1} + (d^{1,2}_{1,2} + d^{2,2}_{1,2})e^{1,2} + d^{2,1}_{1,2}e^{2,2}.
\]
Then \( d^{2,2}_{1,1} = d^{2,2}_{1,2} = 0 \) and \( d^{1,2}_{1,2} + d^{2,2}_{1,2} = 0 \).
(iii) The equation corresponding to \( i = 2 \) and \( h = 1 \) is trivial since \( \sigma^{2,1} = 0 \).
(iv) If \( i = h = 2 \),
\[
0 = d^{2,1}\sigma^{2,2} + \sigma^{1,1}d^{2,2} = d^{2,2}_{1,1}e^{1,1} + d^{2,1}_{1,2}e^{1,2} + d^{2,1}_{1,2}e^{2,2}.
\]
Thus \( d^{2,1}_{1,2} = d^{2,2}_{1,2} = 0 \).

Now, with \( j = 2 \) and \( k = 1 \) in (3) we get the following four cases:
(v) If \( i = h = 1 \),
\[
0 = d^{1,2}\sigma^{1,1} + \sigma^{1,2}d^{1,1} = d^{1,1}_{1,1}e^{1,1} + d^{1,2}_{1,2}e^{1,1}.
\]
Thus, \( d^{1,1}_{1,1} = d^{1,2}_{1,2} = 0 \).
(vi) The equation corresponding to \( i = 1 \) and \( h = 2 \) is trivial since \( \sigma^{1,2} = 0 \).
(vii) If \( i = 2 \) and \( h = 1 \),
\[
0 = d^{2,2}_i \sigma^{1,1} + \sigma^{2,2} d^{1,1} = d^{2,2}_{1,1} e^{1,1} + (d^{2,2}_{2,1} + d^{1,1}_{2,1}) e^{2,1} + d^{1,2}_{2,2} e^{2,2}.
\]
Then \( d^{2,2}_{1,1} = d^{1,1}_{2,1} = 0 \) and \( d^{2,2}_{2,1} + d^{1,1}_{2,1} = 0 \).

(viii) If \( i = h = 2 \),
\[
0 = d^{2,2}_i \sigma^{1,2} + \sigma^{2,2} d^{1,2} = d^{1,2}_{2,1} e^{1,1} + d^{1,2}_{2,2} e^{2,2}.
\]
Then \( d^{1,2}_{2,1} = d^{1,2}_{2,2} = 0 \).

With \( j = k = 1 \) in (3) we get the following four cases:

(ix) If \( i = h = 1 \),
\[
d^{1,1} = d^{1,1}_i \sigma^{1,1} + \sigma^{1,1} d^{1,1} = 2d^{1,1}_{1,1} e^{1,1} + d^{1,1}_{1,2} e^{1,2} + d^{1,1}_{2,1} e^{2,1}.
\]
Then \( d^{1,1}_{1,1} = d^{1,1}_{2,1} = 0 \).

(x) If \( i = 1 \) and \( h = 2 \),
\[
d^{1,2} = d^{1,2}_i \sigma^{1,2} + \sigma^{1,1} d^{1,2} = d^{1,2}_{1,1} e^{1,1} + d^{1,2}_{1,2} e^{1,2}.
\]
Then \( d^{1,2}_{2,1} = d^{1,2}_{2,2} = 0 \).

(xi) If \( i = 2 \) and \( h = 1 \),
\[
d^{2,1} = d^{2,1}_i \sigma^{1,1} + \sigma^{2,1} d^{1,1} = d^{2,1}_{1,1} e^{1,1} + d^{2,1}_{2,1} e^{2,1}.
\]
Then \( d^{2,1}_{1,2} = d^{2,1}_{2,2} = 0 \).

(xii) If \( i = h = 2 \), \( d^{2,2} = d^{2,2}_i \sigma^{1,2} + \sigma^{2,1} d^{1,2} = 0 \).

If \( j = k = 2 \) in (3) we have the following four cases:

(xiii) If \( i = h = 1 \), \( d^{1,1} = d^{1,2}_i \sigma^{2,1} + \sigma^{1,2} d^{2,1} = 0 \).

(xiv) If \( i = 1 \) and \( h = 2 \),
\[
d^{1,2} = d^{1,2}_i \sigma^{2,2} + \sigma^{1,2} d^{2,2} = d^{1,2}_{1,1} e^{1,2} + d^{1,2}_{2,2} e^{2,2}.
\]
Then \( d^{1,2}_{2,1} = d^{1,2}_{2,2} = 0 \).

(xv) If \( i = 2 \) and \( h = 1 \),
\[
d^{2,1} = d^{2,1}_i \sigma^{2,1} + \sigma^{2,2} d^{2,1} = d^{2,1}_{1,1} e^{2,1} + d^{2,1}_{2,2} e^{2,2}.
\]
Then \( d^{2,1}_{1,2} = d^{2,1}_{2,2} = 0 \).

(xvi) If \( i = h = 2 \),
\[
d^{2,2} = d^{2,2}_i \sigma^{2,2} + \sigma^{2,2} d^{2,2} = d^{2,2}_{1,2} e^{1,2} + d^{2,2}_{2,2} e^{2,2} + 2 d^{2,2}_{2,2} e^{2,2}.
\]
Then \( d^{2,2}_{1,1} = d^{2,2}_{2,2} = 0 \).

Finally, the claim follows as soon as we collect and integrate the conclusions (i)-(xvi).
2.5. **Mirzavaziri’s problems.** We continue the notation of Prop. 4. Clearly \(\sigma\) is not a homomorphism on \(M_2(K)\) but, \(D(\sigma)\) is not trivial. So, the answer to the question (A) posed in \(\S\) 1 is negative. Let us look at the question (B) considering the \(\sigma\)-derivation \(d(x) = x_{1,2} e^{1,2} + x_{2,1} e^{2,1}\) for all \(x \in M_2(K)\). Let us suppose that there is a derivation \(\delta\) on \(M_2(K)\) so that \(d = \delta \circ \sigma\). It is well known that \(\delta\) must be inner, even in a more general setting (cf. [1]). So, if \(a \in M_2(K)\) is so that \(\delta = \delta_a\) we would have

\[
x_{1,2} e^{1,2} + x_{2,1} e^{2,1} = d(x) = a\sigma(x) - \sigma(x)a = (a_{1,2} e^{1,2} - a_{2,1} e^{2,1}) (x_{2,2} - x_{1,1})
\]

for all \(x \in M_2(K)\), which is clearly impossible. The answer to question (B) is also negative.

**References**


Received July 6, 2012

CONICET- UNCPBA, Facultad de Ciencias Exactas, NUCOMPA. 
E-mail address: apmadrid@exa.unicen.edu.ar