Power linear Keller maps with ditto triangularizations

He Tong\textsuperscript{a,*}, Michiel de Bondt\textsuperscript{b,1}

\textsuperscript{a} Department of Mathematics, Beijing University of Aeronautics and Astronautics, Beijing, China
\textsuperscript{b} Department of Mathematics, Radbound University of Nijmegen, Nijmegen, Netherlands

Received 19 April 2006
Available online 28 December 2006
Communicated by Efim Zelmanov

Abstract

We show that power linear Keller maps $F = (x_1 + (A_1 x)^d, x_2 + (A_2 x)^d, \ldots, x_n + (A_n x)^d)$ are linearly triangularizable if (1) $\text{rk } A \leq 2$ or (2) $\text{corank } A \leq 2$ and $d \geq 3$ or (3) $\text{corank } A = 3$, $d \geq 5$ and the diagonal of $A$ is nonzero. Furthermore, we show that the triangularizations can be chosen power linear as well.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Jacobian Conjecture; Keller map; Linearly triangularizable

1. Introduction

The famous Jacobian Conjecture, which was first formulated by O.H. Keller in 1939, for short JC, asserts that for every $n \geq 1$ the following holds:

If $F = (F_1, F_2, \ldots, F_n)$ is a polynomial map over $\mathbb{C}$ with constant nontrivial Jacobian determinant, then $F$ is invertible.

In the 1980s, there are two famous reduction results. At first, it is shown that in order to prove the JC, it suffices to verify the JC for polynomial maps $F$ over $\mathbb{C}$ of special cubic homogeneous form:

$F = x + H = (x_1 + H_1, x_2 + H_2, \ldots, x_n + H_n)$

* Corresponding author.
E-mail addresses: hetong@ustc.edu (H. Tong), debondt@math.ru.nl (M. de Bondt).
1 Supported by the Netherlands Organisation of Scientific Research (NWO).
where each component $H_i$ of $H$ is either zero or homogeneous of degree 3, see [1]. Later, Ludwik Drużkowski showed in [8] that in addition, one may assume that each component $H_i$ of $H$ is a third power of a linear form:

$$F = x + (Ax)^3 = (x_1 + (A_1 x)^3, x_2 + (A_2 x)^3, \ldots, x_n + (A_n x)^3)$$

where $x = (x_1, x_2, \ldots, x_n)$, $A_i$ is the $i$th row of an $(n \times n)$-matrix $A$, and $A_i x$ is the matrix product

$$(A_{i1} A_{i1} \cdots A_{in}) \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

For the case $\deg F \leq 2$, S. Wang had already proved in 1980 that the JC is true over any field of characteristic $\neq 2$, see [17] and [1].

In 1993, David Wright showed that in case $n = 3$, the JC holds for maps $F$ having special cubic homogeneous form, see [18]. In particular $F$ is so-called ‘linearly triangularizable,’ see Definition 2.5. In 1994, the result of Wright was extended to the case $n = 4$ by Engelbert Hubbers, see [13], but for $n = 4$, maps of special cubic homogeneous form are not always linearly triangularizable. Hubbers used a (for those days) strong computer to get these results.

More than 10 years later, the result of Wright was extended in another direction: Arno van den Essen and the second author showed that in case $n = 3$ the JC holds for maps $F$ having special homogeneous form in general (not just cubic) in [2]. The main theorem of [2] asserts that $F$ is even linearly triangularizable, just as in the cubic case.

But let us focus on special cubic linear maps $x + (Ax)^3$ and, more generally, special power linear maps $x + (Ax)^d$, from now on. At the same time that Wright showed the case $n = 3$ for special homogeneous cubic maps, Drużkowski showed that for special cubic linear maps $F = x + (Ax)^3$ with $\text{rk } A \leq 2$ or $\text{cork } A \leq 2$, $F$ is invertible, see [9]. In particular, $F$ is tame.

Although the results of Drużkowski for degree $d = 3$ generalize to degree $d \geq 3$ in a straightforward manner, we have chosen to rewrite these results. The main reason for this is that the proofs of Drużkowski are very sketchy; at some points, one can better speak of ‘guidelines of how to prove.’

Furthermore, Drużkowski only proved tameness in [9], which is weaker than linear triangularizability, but for the case $\text{cork } A \leq 2$, his proof is powerful enough for linear triangularizability, as Charles Ching-An Cheng observes in [4]. In the same article, Cheng proves linear triangularizability for the case $\text{rk } A = 2$ and $d = 3$.

But this proof is quite long. Cheng presents a much shorter proof for the case $\text{rk } A = 2$ and $d$ arbitrary in [6], by showing the following result (Theorem 2 in [6]):

**Theorem 1.1.** Let $F = x + (Ax)^d$ be a power linear Keller map, $r = \text{rk } A$, and assume that all special homogeneous Keller maps of degree $d$ in dimension $r$ are linearly triangularizable. Then $F$ is linearly triangularizable as well.

Since it is a classical result that for $r = 2$, the conditions of this theorem are fulfilled (see [1], [2] or [6]), the case $\text{rk } A = 2$ and $d$ arbitrary follows. As mentioned above, the main result of [2]
was exactly the case \( r = 3 \) of the conditions of the above theorem for all \( d \), so the case \( \text{rk} \, A = 3 \) and \( d \) arbitrary follows as well, as mentioned in [2].

We shall show that power linear Keller maps \( F = (x_1 + (A_1x)^d, x_2 + (A_2x)^d, \ldots, x_n + (A_nx)^d) \) are linearly triangularizable in each of the following cases:

1. \( \text{rk} \, A \leq 2 \),
2. \( \text{cork} \, A \leq 2 \) and \( d \geq 3 \),
3. \( \text{cork} \, A = 3 \), \( d \geq 5 \) and the diagonal of \( A \) is nonzero.

Furthermore, we show that in all of the above cases, the triangularizations can be chosen power linear as well. For a significant part, our results are based on the work of Drużkowski in [9].

Although the results for \( \text{rk} \, A \leq 2 \) are valid for any \( d \), those for \( \text{cork} \, A \leq 2 \) apply only to the case \( d \geq 3 \). This restriction is not important for the JC, since it has already been proved for any polynomial map over \( \mathbb{C} \) with degree \( d \leq 2 \). On the other hand, the invertibility statement of the JC is weaker than linear triangularizability, so it is worth mentioning that in 2002, Cheng proved that quadratic linear Keller maps \( x + (Ax)^* \) with \( \text{cork} \, A = 1 \) are linearly triangularizable, see [5].

In the last section, we present a quadratic linear map in dimension 6 with \( \text{rk} \, A = \text{cork} \, A = 3 \), which is, as observed above, linearly triangularizable, but without a linear triangularization that is quadratic linear as well. So in our result for \( \text{cork} \, A = 3 \), the assumption \( d \geq 5 \) or at least some assumption on \( d \), is necessary. Another study of nilpotent Jacobians and the some linearizable problems can be found in [7,11,14,15]

2. Definitions and preliminaries

**Definition 2.1.** Write \( A^t \) for the transpose of a matrix \( A \). Now let \( A \) be an \((n \times n)\)-matrix. We write \( e_i \) for the \( i \)th standard basis vector over \( \mathbb{C}^n \). Viewing vectors as column matrices, the matrix product \( Ae_i \) evaluates to the \( i \)th column of \( A \) and \( e_i^tA \) evaluates to the \( i \)th row of \( A \). But we will just write \( A_i \) for the \( i \)th row of \( A \).

**Definition 2.2.** We call a map \( H \) power linear (of degree \( d \)) if \( H \) is of the form

\[
H = (Ax)^*d := ((A_1x)^d, (A_2x)^d, \ldots, (A_nx)^d)
\]

and a map \( F \) special power linear (of degree \( d \)) if \( F \) is of the form

\[
F = x + (Ax)^*d = (x_1 + (A_1x)^d, x_2 + (A_2x)^d, \ldots, x_n + (A_nx)^d).
\]

So \( H \) is power linear if and only if \( x + H \) is special power linear.

**Definition 2.3.** Let \( F \) be a polynomial map. We say that \( F \) is upper/lower triangular if its Jacobian \( \mathcal{J}F \) is upper/lower triangular. We call \( F \) triangular if it is either upper or lower triangular.

A triangular Keller map is tame and hence invertible.

**Definition 2.4.** Let \( F = x + H \) be a polynomial map. We call \( F \) special homogeneous (of degree \( d \)) if \( H \) is homogeneous (of degree \( d \)).
In [1, Lemma 4.1], it is shown that a special homogeneous map of degree \( d \geq 2 \) is a Keller map, if and only if \( JH \) is nilpotent.

**Definition 2.5.** Let \( F \) be a polynomial map over \( \mathbb{C} \). We call \( F \) *linearly triangularizable* if there exists a \( T \in \text{GL}_n(\mathbb{C}) \) such \( T^{-1} \circ F \circ T \) is triangular.

A linear triangularizable map can be triangularized to both an upper and a lower triangular map: take \( T = (x_n, x_{n-1}, \ldots, x_1) \) to get from lower to upper and vice versa.

**Proposition 2.6.** If \( F = x + H \) is a linearly triangularizable Keller map and the components of \( H \) do not have linear parts, then \( JH \) is nilpotent.

**Proof.** The proof is left as an exercise to the reader. A stronger result can be found in [10, Theorem 1.6]. \( \square \)

**Proposition 2.7.** If \( F = x + H \) is a triangular Keller map and the components of \( H \) do not have linear parts, then \( JH \) only has zeros on its diagonal.

**Proof.** From Proposition 2.6, it follows that \( JH \) is nilpotent. Since a nilpotent matrix over a reduced ring only has eigenvalue 0 and the diagonal of a triangular matrix is formed by its eigenvalues, it follows that \( JH \) only has zeros on its diagonal. \( \square \)

**Definition 2.8.** Let \( f \in \mathbb{C}[x] = \mathbb{C}[x_1, x_2, \ldots, x_n] \). We write \( \deg f \) for the total degree of \( f \). We write \( \deg_{x_i} f \) for the degree of \( f \), seen as a polynomial in \( x_i \) over \( \mathbb{C}[x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n] \). We write \( \deg_{x_i, x_j, x_k} f \) for the (total) degree of \( f \), seen as polynomial in \( x_i, x_j, x_k \).

### 3. Some results on linear dependence

**Lemma 3.1.** Let \( H := (Ax)^*d \) such that \( JH \) is nilpotent. Assume that the first \( r \) rows of \( A_1, A_2, \ldots, A_r \) of \( A \) are independent and the last \( n - r \) rows of \( A \) are dependent of \( A_{r-1} \) and \( A_r \) only. Assume a similar condition on the columns of \( A \), i.e. the last \( n - r \) columns of \( A \) are dependent of \( Ae_{r-1} \) and \( Ae_r \) only. Then the components of \( H := (Ax)^*d \) are linearly dependent.

**Proof.** Write \( Ae_{r+i} = \lambda_{r+i} Ae_{r+i} + \mu_{r+i} Ae_r \). Put

\[
L = \begin{pmatrix}
x_1 \\
\vdots \\
x_{r-2} \\
x_{r-1} - \lambda_{r+1} x_{r+1} - \cdots - \lambda_n x_n \\
x_r - \mu_{r+1} x_{r+1} - \cdots - \mu_n x_n \\
x_{r+1} \\
\vdots \\
x_n
\end{pmatrix}
\]
and let $B := A \cdot J L$. Then the last $n - r$ columns of $B$ and hence those of $J \tilde{H}$ are zero, where

$$\tilde{H} := L^{-1} \circ H \circ L = \begin{pmatrix}
(B_1 x)^d \\
\vdots \\
(B_{r-1} x)^d + \lambda_{r+1} (B_{r+1} x)^d + \cdots + \lambda_n (B_n x)^d \\
(B_r x)^d + \mu_{r+1} (B_{r+1} x)^d + \cdots + \mu_n (B_n x)^d \\
\vdots \\
(B_n x)^d
\end{pmatrix}.$$

Each row $B_{r+i}$ with $i \geq 1$ is a linear combination of $B_{r-1}$ and $B_r$, for a similar statement holds for the rows of $A$. So $\hat{H} := (\tilde{H}_1, \ldots, \tilde{H}_{r-2}, \tilde{H}_{r-1}, \tilde{H}_r)$ is of the form

$$\hat{H} = \begin{pmatrix}
(B_1 x)^d \\
\vdots \\
(B_{r-1} x)^d + p(B_{r-1} x, B_r x) \\
p(B_{r-1} x, B_r x) \\
q(B_{r-1} x, B_r x)
\end{pmatrix}.$$

Furthermore, since the last $n - r$ columns of $J \tilde{H}$ are zero, the $(r \times r)$-matrix $J \hat{H}$ is nilpotent as well. In particular, $\det J \hat{H} = 0$. If $p(B_{r-1} x, B_r x)$ and $q(B_{r-1} x, B_r x)$ are algebraically independent, then all linear forms $B_i x$ with $i \leq r$ are algebraically dependent of the components of $\hat{H}$. So

$$\text{trdeg}_C \hat{H} = \text{trdeg}_C(B_1 x, \ldots, B_r x) = \text{trdeg}_C(A_1 x, \ldots, A_r x) = r$$

for the first $r$ rows of $A$ are linearly independent. This contradicts $\det J \hat{H} = 0$, so $p(B_{r-1} x, B_r x)$ and $q(B_{r-1} x, B_r x)$ are algebraically dependent. But with $p$ and $q$ homogeneous of the same degree $d$, this dependence relation refines to a linear relation, say that $v_1 p + v_2 q = 0$ with $v \neq 0$. Then

$$v_1 ((B_{r-1} x)^d + \lambda_{r+1} (B_{r+1} x)^d + \cdots + \lambda_n (B_n x)^d) + v_2 ((B_r x)^d + \mu_{r+1} (B_{r+1} x)^d + \cdots + \mu_n (B_n x)^d) = 0.$$ 

So the components of $(B x)^{sd}$, and hence those of $H = (Ax)^{sd}$ also, are linearly dependent. \qed

The preceding lemma is a special case of the following theorem:

**Theorem 3.2.** Let $H := (Ax)^{sd}$ such that $\mathcal{J} H$ is nilpotent. Assume that the first $r$ rows of $A_1, A_2, \ldots, A_r$ of $A$ are independent and the last $n - r$ rows of $A$ are dependent of $A_{r-1}$ and $A_r$ only. Then the components of $H := (Ax)^{sd}$ are linearly dependent.

**Proof.** Since the rows of $A$ are dependent, the columns are dependent as well. We distinguish two cases:
• There is an $i \leq r - 2$ such that column $Ae_i$ of $A$ is dependent of the other columns of $A$.

Then there is a vector $\lambda$ with $\lambda_i \neq 0$ for some $i \leq r - 2$ such that $A\lambda = 0$. Replacing $H$ by $P^{-1} \circ H \circ P$ for a suitable permutation $P$ within $x_1, x_2, \ldots, x_{r-2}$, we may assume that $\lambda_1 \neq 0$. Since $r - 2 \geq 1$, ‘variable’ $A_1x$ only appears in the first row of (1). So substituting $A_1x = 0$ in $JH$ just makes the first row of $JH$ zero. This substitution does not affect the condition $\det(TI_n + JH) = T^n$. So $JH$ is nilpotent, where $H := (0, H_2, \ldots, H_n)$. Next, let

$$\hat{H} := L^{-1} \circ H \circ L = \tilde{H} \circ L$$

where $L = x + \lambda_1^{-1}(0, \lambda_2, x_1, \ldots, \lambda_r, x_1)$. Now $x + \hat{H}$ is power linear of degree $d$ as well, but both the first row and the first column of $JH$ are zero. Hence $x + \hat{H}$ is essentially a power linear map in dimension $n - 1$, and the result follows by induction.

• For each $i \leq r - 2$, column $Ae_i$ of $A$ is independent of the other columns of $A$.

Since in particular the first $r - 2$ columns of $A$ are independent, there exists a basis of the column space of $A$ of the form $Ae_1, Ae_2, \ldots, Ae_{r-2}, Ae_i_1, Ae_i_2$. Furthermore, for each $j \geq r - 1$, column $Ae_j$ is a linear combination of $Ae_{i_1}$ and $Ae_{i_2}$ only. We shall show that we may assume that $i_1 = r - 1$ and $i_2 = r$, in order to be able to apply Lemma 3.1.

For that purpose let us look at the rows $A_{i_1}$ and $A_{i_2}$ of $A$. If both rows are dependent, then $H_{i_1}$ and $H_{i_2}$ are linearly dependent and we are done. So assume that $A_{i_1}$ and $A_{i_2}$ are independent. Since the last $n - r$ rows of $A$ are linear combinations of $A_{r-1}$ and $A_r$ and $i_1, i_2 \geq r - 1$, both $A_{i_1}$ and $A_{i_2}$ are linear combinations of $A_{r-1}$ and $A_r$. Hence the spaces $CA_{i_1} + CA_{i_2}$ and $CA_{r-1} + CA_r$ are equal.

Hence $A_{i_1}$ and $A_{i_2}$ can take the role of $A_{r-1}$ and $A_r$, i.e. the rows $A_1, A_2, \ldots, A_{r-2}, A_{i_1}, A_{i_2}$ are independent and each row $A_j$ with $j \geq r - 1$ is a linear combination of $A_{i_1}$ and $A_{i_2}$ only.

Replacing $H$ by $P^{-1} \circ H \circ P$ for a suitable permutation $P$ within $x_{r-1}, x_r, \ldots, x_n$, we may assume that $H$ satisfies the conditions of Lemma 3.1. So the components of $H$ are linearly dependent. \Box

The proof of Theorem 3.2 and its preceding lemma was essentially given by Drożkowski in [9], where he proved the case $r = n - 2$ of Theorem 3.2. The remaining theorems in this section show that under certain conditions, the components of $H$ are not only linearly dependent, but the linear dependence even restricts to two components of $H$, i.e. $H_i = sH_j$ for some $i \neq j$ and an $s \in \mathbb{C}$.

**Lemma 3.3.** Let $L_1, L_2, \ldots, L_r \in \mathbb{C}[x]$ be linear such that $2 \leq r \leq d + 1$ and

$$\lambda_1 L_1^d + \lambda_2 L_2^d + \cdots + \lambda_r L_r^d = 0$$

for some $\lambda = (\lambda_1, \ldots, \lambda_r) \neq 0$. Then there are $i \neq j$ and an $s \in \mathbb{C}$ such that $L_i = sL_j$. 

\[935\]
Proof. Assume the opposite. In particular, $L_1 \neq s L_r$ and $L_r \neq s L_1$ for all $s \in \mathbb{C}$, whence $L_1$ and $L_r$ are independent. There exists a linear basis $y_1, y_2, \ldots, y_n$ of $\mathbb{C}[x]$ with $y_1 = L_1$ and $y_2 = L_r$.

The case $d = 1$ is easy, so assume $d \geq 2$. Differentiating (2) with respect to $y_1$ gives

$$
\mu_1 L_1^{d-1} + \mu_2 L_2^{d-1} + \cdots + \mu_{r-1} L_{r-1}^{d-1} = 0
$$

for certain $\mu_i \in \mathbb{C}$. In particular, $\mu_1 = d \lambda_1$, whence not all $\mu_i$ are zero. Hence, the result follows by induction on $d$. \hfill \square

The following theorem generalizes Theorem 3.1 of [16] (the case cork $A = 3$ of this theorem). [16] is a co-production of Song Shuang and the first author.

**Theorem 3.4.** Assume $H$ is of the form $(Ax)^{sd}$ such that cork $A \leq d - 2$, $\text{tr} \, \mathcal{J} H = 0$, and the diagonal of $A$ is nonzero. Then there are $i \neq j$ and an $s \in \mathbb{C}$ such that $A_i = s A_j \neq 0$.

**Proof.** Since the diagonal of $\mathcal{J} H$ is nonzero, we can replace $H$ by $P^{-1} \circ H \circ P$ to get $A_{11} \neq 0$, where $P$ is a permutation. Similarly, we can make the first $r$ rows of $A$ independent in addition, where $r = \text{rk} A \geq n - (d - 2)$. Since $\text{tr} \, \mathcal{J} H = 0$, we have

$$
d A_{11}(A_1 x)^{d-1} + d A_{22}(A_2 x)^{d-1} + \cdots + d A_{nn}(A_n x)^{d-1} = 0.
$$

(3)

Since the first $r$ rows of $A$ are independent, there exists a basis $y$ of $\mathbb{C} x_1 + \mathbb{C} x_2 + \cdots + \mathbb{C} x_n$ such that $A_i x = y_i$ for all $i \leq r$. Differentiating (3) with respect to $y_1$ gives

$$
d (d - 1) A_{11}(A_1 x)^{d-2} + \lambda_{r+1} A_{r+1}(A_{r+1} x)^{d-2} + \cdots + \lambda_n (A_n x)^{d-2} = 0
$$

for certain $\lambda_i \in \mathbb{C}$. These are $n - r + 1 \leq d - 1$ linear powers. Now apply Lemma 3.3 to get $A_i = s A_j$ for some $i \neq j$ and $s \in \mathbb{C}$ with $i, j \in \{1, r + 1, r + 2, \ldots, n\}$. \hfill \square

**Theorem 3.5.** Assume $H$ is as in Theorem 3.2 and cork $A \leq d - 1$. Then there are $i \neq j$ and an $s \in \mathbb{C}$ such that $A_i = s A_j$.

**Proof.** From Theorem 3.2, it follows that there is a linear relation between the components of $H$. Similar to the proof of Theorem 3.4 (but with $\mathcal{J}$ instead of $d - 1$), one can show that this relation is of the form $H_i = \alpha H_j$ for some $i \neq j$. So $A_i = s A_j$ for some $s \in \mathbb{C}$. \hfill \square

We will use the above theorems in the next section.

**4. Linear triangularization to power linear maps**

The following lemma is crucial in both [9] and our study of power linear maps $(Ax)^{sd}$ where $A$ has a small corank. It can be found at the beginning of page 238 in [9].

**Lemma 4.1.** Let $H = (Ax)^{sd}$ such that $\mathcal{J} H$ is nilpotent. If $A$ has a principal minor of any size which determinant is nonzero, then there exists a relation $R \neq 0$ such that

$$R((A_1 x)^{d-1}, (A_2 x)^{d-1}, \ldots, (A_n x)^{d-1}) = 0$$

and $\deg_y R(y) \leq 1$ for all $i \leq n$. Furthermore, if $A_k = 0$ for some $k$, then $\deg_y R = 0$ as well.
Proof. Write
\[
\det \left( TI_n + d \begin{pmatrix} A_{11}y_1 & A_{12}y_1 & \cdots & A_{1n}y_1 \\ A_{21}y_2 & A_{22}y_2 & \cdots & A_{2n}y_2 \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1}y_n & A_{n2}y_n & \cdots & A_{nn}y_n \end{pmatrix} \right)
\]
\[= T^n + R_1(y)T^{n-1} + R_2(y)T^{n-2} + \cdots + R_{n-2}(y)T^2 + R_{n-1}(y)T + R_n(y).\]

Since $\mathcal{H}$ is nilpotent, $\det(T I_n + \mathcal{H}) = T^n$. It follows from (1) that the coefficient of $T^{n-j}$ of $\det(T I_n + \mathcal{H})$ equals
\[
R_j((A_1x)^{d-1}, (A_2x)^{d-1}, \ldots, (A_nx)^{d-1}) = 0
\]
for all $j \geq 1$. Furthermore, it follows from the definition of determinant that $\deg_{y_i} R_j \leq 1$ for all $i, j$. For some $j$, $A$ has a principal minor of size $j$ which determinant is $\alpha \neq 0$, say with rows $i_1, i_2, \ldots, i_j$. Then the coefficient of $y_{i_1}y_{i_2}\cdots y_{i_j}$ of $R_j$ equals $d\alpha$, whence $R_j \neq 0$.

If $A_k = 0$, then all minors with row $k$ of $A$ have determinant zero, whence $\deg_{y_k} R_j = 0$. \(\square\)

In all remaining lemmas in this section, relations $R$ between linear powers $L_1^d, L_2^d, \ldots, L_m^d$ with $\deg_{y_i} R \leq 1$ for all $i \leq m$ are studied. For such relations, conditions are formulated that imply $L_i = sL_j$ for some $i \neq j$ and an $s \in \mathbb{C}$.

**Lemma 4.2.** Let $d \geq 2$ and $R$ be a nonzero relation with $\deg_{y_i} R \leq 1$ such that
\[
R(x_1^d, x_2^d, \ldots, x_r^d, (\lambda_1x_1 + \lambda_2x_2 + \cdots + \lambda_rx_r)^d) = 0.
\]
Then $\lambda = \lambda_ie_i$ for some $i$.

**Proof.** Since $x_1^d, x_2^d, \ldots, x_r^d$ are algebraically independent, it follows that $R$ has a term of the form
\[
\alpha \cdot y_1^{t_1} \cdots y_r^{t_r} \cdot y_{r+1}
\]
with $\alpha \neq 0$ and $0 \leq t_i \leq 1$ for all $i$. The coefficient of $x_1^{dt_1}x_2^{dt_2} \cdots x_r^{dt_r}x_j^{d-1}x_k$ in (4) equals $(d-1)\alpha \lambda_j \lambda_k = 0$, so $\lambda_j \lambda_k = 0$ for all $j \neq k$. It follows that $\lambda$ has at most one nonzero coordinate, i.e. $\lambda = \lambda_ie_i$ for some $i$. \(\square\)

**Lemma 4.3.** Let $d \geq 2$ and $R$ be a nonzero relation with $\deg_{y_i} R \leq 1$ such that
\[
R(x_1^d, x_2^d, \ldots, x_r^d, (\lambda_1x_1 + \lambda_2x_2 + \cdots + \lambda_rx_r)^d, (\mu_1x_1 + \mu_2x_2 + \cdots + \mu_rx_r)^d) = 0.
\]
Assume further that $\lambda_i = \mu_i = 0$ for at most $r - 3$ $i$'s. Then either $\lambda = \lambda_ie_i$ for some $i$ or $\mu = \mu_ie_i$ for some $i$ or $\lambda$ and $\mu$ are dependent.
Proof. Assume that \( \lambda \) and \( \mu \) are independent. Without loss of generality, we assume that \((\lambda_1, \lambda_2)\) and \((\mu_1, \mu_2)\) are independent. The cases \( \deg_{y_{r+1}} R = 0 \) and \( \deg_{y_{r+2}} R = 0 \) follow from Lemma 4.2. So assume the opposite.

(i) Suppose first that \( \lambda_1 = \mu_2 = 0 \). Then \( \lambda_2 \mu_1 \neq 0 \). Since \( \deg_{y_{r+2}} R = 1 \), \( R \) has a term of the form

\[ \alpha y_1^{t_1/2} \cdots y_r^{t_r/2} y_{r+1}^{t_{r+1}} y_{r+2} \]

with \( 0 \leq t_i \leq 1 \) for all \( i \). If \( t_{r+1} = 0 \), then by looking at the term

\[ x_1^{d_1} x_2^{d_2} \cdots x_r^{d_r} \cdot (x_1^{d-1} x_m) \]

of (5), we see that \( \mu_m = 0 \) for all \( m \neq 1 \), i.e. \( \mu = \mu_1 e_1 \). So assume \( t_{r+1} = 1 \). Looking at the term

\[ x_1^{d_1} x_2^{d_2} \cdots x_r^{d_r} \cdot x_2^{d-1} x_1^{2d-1} \]

of (5), we see that \( \lambda_i \mu_i = 0 \) for all \( i \geq 3 \) and the left-hand side of (5) has degree 3 with respect to \( x_1 \); contradiction. Since \( \deg_{y_{r+1}} R \neq 0 \), \( s \geq 1 \). So two cases remain:

- \( s = 1 \):

  We can write

  \[ R = R_1 y_1 + R_2 y_{r+1} + R_3 y_{r+2} + R_4 \]

  with \( R_i \in \mathbb{C}[y_2, \ldots, y_r] \). Looking at the coefficient of \( x_1^{d-1} \) in (5) gives

  \[ R_2(x_2^{d}, \ldots, x_r^{d}) L = -R_3(x_2^{d}, \ldots, x_r^{d}) M. \]
Assume $R_2 \neq 0$. Notice that $d \geq 2$. Reduction modulo $x_i^d - y_i$ for all $i$ gives $R_2 L = -R_3 M$. Next, a generic substitution into the $y_i$’s gives $L = \alpha M$ for some $\alpha \in \mathbb{C}$. So $L$ and $M$ are linearly dependent. This contradicts the independence of $(\lambda_2, \lambda_3)$ and $(\mu_2, \mu_3)$, so $R_2 = R_3 = 0$. Looking at the coefficient of $x_1^d$ in (5) gives $R_1 = 0$. So $R = R_4$. This contradicts $s = 1$.

$s = 2$:

We can write

$$R = R_1y_{r+1}y_{r+2} + R_2y_1y_{r+2} + R_3y_1y_{r+1} + R_4$$

with $R_i \in \mathbb{C}[y_2, \ldots, y_r]$ for all $i \leq 3$ and $\deg_{y_1,y_{r+1},y_{r+2}} R_4 \leq 1$. Looking at the coefficient of $x_1^{2d-1}$ in (5) gives

$$(R_1 + R_3)(x_2^d, \ldots, x_r^d)L = -(R_1 + R_2)(x_2^d, \ldots, x_r^d)M$$

and $(R_1 + R_3) = (R_1 + R_2) = 0$ follows similar as $R_2 = R_3 = 0$ in the case $s = 1$. Looking at the coefficient of $x_1^{2d}$ in (5) gives $R_1 + R_2 + R_3 = 0$, so $R_2 = R_3 = 0$ and also $R_1 = 0$. So $R = R_4$. This contradicts $s = 2$. □

**Theorem 4.4.** Assume $A$ is a matrix of corank 2 at most, $d \geq 3$ and $H = (Ax)^*d$ such that $JH$ is nilpotent. Then there exists a $T \in \text{GL}_n(\mathbb{C})$ and a lower triangular matrix $B$ such that

$$T^{-1} \circ (Ax)^*d \circ T = (Bx)^*d.$$ 

**Proof.** Assume first that every principal minor of $A$ has determinant zero. From [9, Lemma 1.2] (see also [12, Proposition 6.3.9]), it follows that there is a permutation $P$ such that $P^{-1}AP$ is lower triangular. So take $T = P$.

Assume next that $A$ has an invertible principal minor. From Lemma 4.1, it follows that there exists a nonzero relation $R$ such that

$$R((A_1x)^{d-1}, (A_2x)^{d-1}, \ldots, (A_nx)^{d-1}) = 0.$$ 

Let $r := \text{rk} A \geq n - 2$. After a suitable permutation, we have that the rows $A_1, A_2, \ldots, A_r$ are independent,

$$A_{r+1} = \lambda_1 A_1 + \lambda_2 A_2 + \cdots + \lambda_r A_r$$

and, in case $r = n - 2$,

$$A_{r+2} = \mu_1 A_1 + \lambda_2 A_2 + \cdots + \mu_r A_r.$$ 

We first show that $A_i = sA_j$ for some $i \neq j$ and $s \in \mathbb{C}$. The case $r = n - 1$ follows from Lemma 4.2, so assume that $r = n - 2$. The case $\lambda_i = \mu_i = 0$ for at most $r - 3$ $i$’s follows from Lemma 4.3, so assume $\lambda_i = \mu_i = 0$ for at least $r - 2$ $i$’s. Replacing $A$ by $P^{-1}AP$ for a suitable permutation $P$, we get that $\lambda_i = \mu_i = 0$ for all $i \leq r - 2$, and Theorem 3.5 applies. So $A_i = sA_j$ for some $i \neq j$ and $s \in \mathbb{C}$.
So the components of $H$ are linearly dependent. Replacing $H$ by $T^{-1} \circ H \circ T$ for a suitable linear transformation $T$, we get $H_1 = 0$ and hence $A_1 = 0$. This transformation may make all principal minor determinants zero, but then, again by [9, Lemma 1.2], there is a permutation matrix $P$ such that $P^{-1}AP$ is lower triangular. So we may assume that there is still a nonzero principal minor determinant in $A$. From Lemma 4.1 it follows that there exists a nonzero relation $R_1$ such that

$$R_1((A_2x)^{d-1}, \ldots, (A_nx)^{d-1}) = 0.$$  

After a suitable permutation, we have that the rows $A_2, A_3, \ldots, A_{r+1}$ are independent and

$$A_{r+2} = \lambda_2 A_2 + \lambda_3 A_3 + \cdots + \lambda_{r+1} A_{r+1}.$$  

Applying Lemma 4.2 again gives $A_i = s A_j$ for some $i \neq j$ with $i, j \neq 1$ and $s \in \mathbb{C}$, i.e. a linear relation between $(A_2x)^d, \ldots, (A_nx)^d$. So after a suitable linear transformation, we have $A_2 = 0$ as well.

Since cok $A \leq 2$, $(A_3x)^{d-1}, \ldots, (A_nx)^{d-1}$ are algebraically independent. It follows from Lemma 4.1 that all principal minor determinants of $A$ are zero. So again we can take for $T$ a suitable permutation matrix $P$. □

The proof of the above theorem was essentially given by Drużkowski in [9]. Drużkowski observed something more or less similar to Lemma 4.3, but found it unnecessary to prove that in full detail.

Lemma 4.5. Let $d \geq 3$ and $R$ be a nonzero relation with $\deg_y R \leq 1$ such that

$$R(x_1^d, x_2^d, \ldots, x_r^d, (\lambda_1 x_1 + \lambda_2 x_1 + \cdots + \lambda_r x_r)^d, (\mu_1 x_1 + \mu_2 x_1 + \cdots + \mu_r x_r)^d) = 0. \quad (6)$$  

Then either $\lambda = \lambda_i e_i$ for some $i$ or $\mu = \mu_i e_i$ for some $i$ or $\lambda$ and $\mu$ are dependent.

Proof. The cases $\deg_{x_1^d} R = 0$ and $\deg_{x_2^d} R = 0$ follow from Lemma 4.2, so assume the opposite. The case $\lambda_i = \mu_i = 0$ for at most $r-3$ $i$’s follows from Lemma 4.3, so assume without loss of generality that $\lambda_i = \mu_i = 0$ for all $i \geq 3$.

Similar as in the proof of Lemma 4.3, we assume that $\lambda_1 = \mu_1 = 1$ and write $\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_r x_r = x_1 + L$ and $\mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_r x_r = x_1 + M$.

Put $s := \deg_{x_1^d, x_1^d, x_2^d} R$. If $s \geq 3$, then $s = 3$ and the left-hand side of (6) has degree $3d$ in $x_1$; contradiction. Since $\deg_{x_r^d} R \neq 0$, $s \geq 1$. So two cases remain:

- $s = 1$:

Since $\lambda_i = \mu_i = 0$ for all $i \geq 3$, $R$ is in fact a relation between $x_1^d, x_2^d, (x_1 + L)^d$ and $(x_1 + M)^d$, say

$$R_0(x_1^d, x_2^d, (x_1 + L)^d, (x_1 + M)^d) = 0$$

for some homogeneous $R_0 \neq 0$ with $\deg_{x_1^d, x_2^d} R_0 = s$ and $\deg_{x_2^d} R_0 = 1$. If $R_0$ is linear, then it follows from Lemma 3.3 and $d \geq 3$ that $L = 0, M = 0$ or $L = M$. If $R_0$ is not linear, then it fol-
allows from $s = 1$ that $R_0$ is quadratic and $y_2 | R_0$, for $R_0$ is homogeneous. Hence, $R_0$ decomposes into linear factors and can be chosen linear instead.

- $s = 2$:

Write

$$R = R_1 y_{r+1} y_{r+2} + R_2 y_1 y_{r+2} + R_3 y_1 y_{r+1} + R_4$$

with $R_i \in \mathbb{C}[y_2, \ldots, y_r]$ for all $i \leq 3$ and $\deg_{y_1, y_{r+1}, y_{r+2}} R_4 \leq 1$. Looking at the coefficient of $x_1^{2d-1}$ in (6) gives

$$(R_1 + R_3)(x_2^d, \ldots, y_r^d)L = -(R_1 + R_2)(x_2^d, \ldots, y_r^d)M.$$  

Looking at the coefficient of $x_1^{2d}$ in (6), gives $R_1 + R_2 + R_3 = 0$, which implies $-R_2 L = R_3 M$.

At last, the coefficient of $x_1^{2d-2}$ in (6) implies that the following is zero:

$$2dR_1LM + (d - 1)(R_1 + R_3)L^2 + (d - 1)(R_1 + R_2)M^2$$

$$= 2dR_1LM - (d - 1)R_2L^2 - (d - 1)R_3M^2$$

$$= 2dR_1LM + (d - 1)R_3LM + (d - 1)R_2LM$$

$$= (d + 1)R_1LM.$$  

So $LM = 0$ or $R_1 = 0$. So assume $R_1 = 0$. Then $-R_2 = R_3$ due to $R_1 + R_2 + R_3 = 0$. From $-R_2 = R_3$ and $-R_2L = R_3M$, it follows that either $R = R_4$, which contradicts $s = 2$, or $L = M$. \(\square\)

**Theorem 4.6.** If $H$ is as in Theorem 3.4 and $\text{cork } A = 3$, then there exists a $T \in \text{GL}_n(\mathbb{C})$ and a lower triangular matrix $B$ such that

$$T^{-1} \circ (Ax)^{sd} \circ T = (Bx)^{sd}.$$  

**Proof.** Since the proof of Theorem 4.6 is more or less similar to that of Theorem 4.4, we only give a sketch of it.

From Theorem 3.4 or [16, Theorem 3.1], it follows that $A_i = sA_j$ for some $i \neq j$ and $s \in \mathbb{C}$, i.e. the components of $H$ are linearly dependent. So we may assume that the first row of $A$ is zero. Assume $A$ has a nonzero principal minor determinant. The conditions of Theorem 3.4 imply that $3 = \text{cork } A \leq d - 2$, so $d \geq 5$. So it follows from Lemmas 4.1 and 4.5 that we may assume that the first two rows of $A$ are zero. Next, it follows from Lemmas 4.1 and 4.2 that we may assume that the first three rows of $A$ are zero. Since $\text{cork } A = 3$, all principal minors of $A$ have determinant zero. So $B$ as above exists. \(\square\)

Observe that in the proofs of Theorems 4.4 and 4.6, the process of triangularization is as follows: first, all occurrences of $A_i = sA_j$ with $i \neq j$ and $s \in \mathbb{C}^*$ are eliminated by linear transformations ‘within $\mathbb{C}[x_i, x_j]$.’ After that, $A$ is made triangular by a permutation transformation. This result does not follow from the methods of Drużkowski.
The above observation does not hold for power linear maps \((Ax)^sd\) with \(\text{rk} A = 2\), but still there exist a triangularization of \((Ax)^sd\) that is power linear as well. The following theorem, which is in fact a closer look on what happens in the proof of Theorem 1 of [6], shows this result not only for \(d \geq 3\), but for any \(d \geq 1\).

**Theorem 4.7.** Assume \(A\) is a matrix of rank 2 at most and \(\mathcal{J}(Ax)^sd\) is nilpotent. Then there exists a \(T \in \text{GL}_n(\mathbb{C})\) and a lower triangular matrix \(B\) such that

\[
T^{-1} \circ (Ax)^sd \circ T = (Bx)^sd.
\]

**Proof.** The case \(\text{rk} A = 1\) was already done by Drużkowski in [9]. So assume that \(\text{rk} A = 2\). Then there are two rows \(A_{i_1}\) and \(A_{i_2}\) of \(A\) such that all other rows of \(A\) are linear combinations of \(A_{i_1}\) and \(A_{i_2}\). There are \(n - 2\) distinct unit vectors \(e_{k_3}, \ldots, e_{k_n}\) such that the rows \(A_{i_1}, A_{i_2}, e_{k_3}, \ldots, e_{k_n}\) are independent. Replacing \(A\) by \(P^{-1}AP\) for a suitable permutation \(P\) makes that the rows \(A_{j_1}, A_{j_2}, e_{3}^l, \ldots, e_{n}^l\) are independent.

Hence the matrix with those \(n\) rows is invertible. So set

\[
T := \begin{pmatrix}
A_{j_1} \\
A_{j_2} \\
e_{3}^l \\
\vdots \\
e_{n}^l
\end{pmatrix}^{-1}.
\]

Then the last \(n - 2\) rows of \(T\) are \(e_{3}^l, \ldots, e_{n}^l\) as well. Put \(\tilde{H} = T^{-1} \circ H \circ T\), where \(H = (Ax)^d\). The components \(\tilde{H}_3, \ldots, \tilde{H}_n\) of \(\tilde{H}\) are clearly linear powers.

Write \(A_i = \lambda_i A_{j_1} + \mu_i A_{j_2}\) for all \(i\). Then

\[
A = \begin{pmatrix}
\lambda_1 & \mu_1 & 0 & \cdots & 0 \\
\lambda_2 & \mu_2 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\lambda_n & \mu_n & 0 & \cdots & 0
\end{pmatrix} \cdot T^{-1}.
\]

So the last \(n - 2\) columns of \(A \cdot T\) are zero. It follows that \(\tilde{H}_i \in \mathbb{C}[x_1, x_2]\) for each \(i\). Hence \((x_1, x_2) + (\tilde{H}_1, \tilde{H}_2)\) is a homogeneous Keller map in dimension 2. Such maps are classified in e.g. [1]: we have either \(\tilde{H}_1 = \tilde{H}_2 = 0\), in which case \(\tilde{H}\) is already of the form \((Bx)^sd\) with \(B\) triangular, or

\[
\begin{pmatrix}
\tilde{H}_1 \\
\tilde{H}_2
\end{pmatrix} = S^{-1} \circ \begin{pmatrix} 0 \\ x_1^d \end{pmatrix} \circ S.
\]

Now \((S, x_3, \ldots, x_n)^{-1} \circ \tilde{H} \circ (S, x_3, \ldots, x_n)\) is of the form \((Bx)^sd\) with \(B\) triangular. \(\Box\)

In case \(\text{rk} A = 1\), Drużkowski found a matrix \(B\) with \(n - 1\) zero rows, but an argument similar as above would give a matrix \(B\) with \(n - 1\) zero columns.
5. Some final remarks

At first, we like to mention that in [5], Cheng proves that in case cork $A = 1$, $A_i = sA_j$ for some $i \neq j$ and $s \in \mathbb{C}$, also in the quadratic case. So the conclusion of Theorem 4.4 holds for this case as well: see the proof of Theorem 4.4.

The following quadratic linear map $(Ax)^{*2}$ in dimension 6 with $\text{rk } A = \text{cork } A = 3$, which is, as observed in the introduction, linearly triangularizable, but without a linear triangularization that is quadratic linear as well:

$$H = \begin{pmatrix}
0 \\
0 \\
(x_1 + x_2 + x_3 - x_4 - x_5 + x_6)^2 \\
(x_1 - x_2 + x_3 - x_4 - x_5 + x_6)^2 \\
(x_1 - x_2 - x_3 + x_4 + x_5 - x_6)^2 \\
(x_1 + x_2 - x_3 + x_4 + x_5 - x_6)^2
\end{pmatrix}.$$

In order to prove that the above quadratic linear $H$ has no ditto linear triangularization, we need the following normalization principle for triangular power linear maps.

**Proposition 5.1.** Let $H = (Ax)^{sd}$ be lower triangular. Then there exists an $r$ and a $G = (Bx)^{sd}$ which is lower triangular as well, such that $G_1 = G_2 = \cdots = G_r = 0$ and $G_{r+1}, G_{r+2}, \ldots, G_n$ are linearly independent over $\mathbb{C}$.

**Proof.** Assume

$$\lambda_1 H_1 + \lambda_2 H_2 + \cdots + \lambda_s H_s$$

is a linear dependence relation between the components of $H$ with $\lambda_s \neq 0$. After a suitable linear transformation that does not affect the fact that $H$ is lower triangular, we have $H_s = 0$. Repeating this argument, we can get that all linear relations between the components of $H$ are determined by zero components of $H$.

Next, if $H_s = 0$, but $H_i = 0$ does not hold for all $i \leq s$, then the map $P^{-1} \circ H \circ P$ with $P = (x_2, \ldots, x_s, x_1, x_{s+1}, \ldots, x_n)$, which is lower triangular as well, has more zero components at the beginning than $H$ has, and the result follows by induction. \[\square\]

Now let $E = (x_1, x_2, x_3 + x_4 + x_5 - x_6, x_4, x_5, x_6)$, then

$$G := E^{-1} \circ H \circ E = \begin{pmatrix}
0 \\
0 \\
8x_1x_2 \\
(x_1 - x_2 + x_3)^2 \\
(x_1 - x_2 - x_3)^2 \\
(x_1 + x_2 - x_3)^2
\end{pmatrix}$$

is a triangularization of $H$. In order to prove that $H$ has no triangularization that is quadratic linear as well, we show that $\tilde{G} = T^{-1} \circ G \circ T$ cannot be both lower triangular just as $G$ and quadratic linear just as $H$. 
Assume $\lambda^4 G = 0$. Looking at $(\frac{\partial}{\partial x_1})^2 G_i$ for all $i$, we see that $\lambda_4 + \lambda_5 + \lambda_6 = 0$. Looking at $(\frac{\partial}{\partial x_2})^2 G_i$ and $(\frac{\partial}{\partial x_3})^2 G_i$ for all $i$ as well, we see that $\lambda_4 = \lambda_5 = \lambda_6 = 0$. Since $G_1 = G_2 = 0$, $\lambda_3 = 0$ and the last four components of $G$ are linearly independent.

Assume that $\tilde{G}$ is lower triangular. From Proposition 5.1, it follows that we may assume that $G_1 = G_2 = 0$. Since the last four components of $G$, and hence those of $G(T x)$ as well, are linearly independent, it follows from $0 = G_3 = (T^{-1})_1 G(T x)$ that the last four coordinates of $(T^{-1})_1$ are zero. Similarly, the last four coordinates of $(T^{-1})_2$ are zero. Since $\tilde{G}$ is lower triangular, we have $G_3 \in \mathbb{C}[x_1, x_2]$, whence $(T^{-1} G)_3 = \tilde{G}_3 (T^{-1} x) \in \mathbb{C}[x_1, x_2]$ as well.

Looking at $\frac{\partial}{\partial x_3} G_i$ for all $i$, it follows that $(T^{-1} G)_3 \in \mathbb{C}[x_1, x_2]$, if and only if $(T^{-1})_3$ is of the form

$$T_3^{-1} = (\mu_1 \mu_2 \mu_3 0 0 0).$$

Assume $G_3$ is the square of a linear form. Then $(T^{-1} G)_3$ is such a square as well. This requires $\mu_3 = 0$, so the first three rows of $T^{-1}$ are dependent. Contradiction, so $G_3$ is not the square of a linear form.

In [12, Theorem 8.4.2], a special cubic linear map is given that is not linearly triangularizable; the proof follows from [12, Theorem 7.4.4] and [12, Theorem 8.3.2]. Another power linear map that is not linearly triangularizable is

$$H = \begin{pmatrix} 0 \\ 0 \\ (x_1 + x_5 - x_6 + x_7 - x_9)^2 \\ (x_2 + x_5 - x_6 + x_7 - x_9)^2 \\ (x_2 + x_3 - x_8)^2 \\ (x_3 - x_8)^2 \\ (x_4 - x_8)^2 \\ (x_5 - x_6 + x_7 - x_9)^2 \\ (x_1 + x_4 - x_8)^2 \end{pmatrix}.$$

The proof that this quadratic linear map cannot linearly be triangularized at all uses the same techniques as above, and is left as an exercise to the reader.

Since for a triangular special homogeneous map $x + H$, either the first or the last component of $H$ is zero, triangularizability of a power linear map $H$ implies that its components are linearly dependent over $\mathbb{C}$. So one can ask whether the components of $H$ need to be linearly dependent. This is not the case: in [3], the second author shows that there exists a cubic linear counterexample to this linear dependence problem in dimension 53.

References


