The following full text is a postprint version which may differ from the publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/35521

Please be advised that this information was generated on 2020-01-26 and may be subject to change.
Power linear Keller maps with ditto triangularizations

He Tong Michiel de Bondt*

October 21, 2014

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) $\text{cork} A \leq 2$ and $d \geq 3$ or (3) $\text{cork} A = 3$, $d \geq 5$ and the diagonal of $A$ is nonzero. Furthermore, we show that the triangularizations can be chosen power linear as well.

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 1980’s, 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:

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

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

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

*Supported by the Netherlands Organisation of Scientific Research (NWO).

2000 Mathematical Subject Classification: 14R15, 14R10.

Key words and phrases: Jacobian Conjecture, Keller map, linearly triangularizable.
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 $\rk A \leq 2$ or $\cork A \leq 2$, $F$ 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 $\cork A \leq 2$, 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 $\rk A = 2$ and $d = 3$.

But this proof is quite long. Cheng presents a much shorter proof for the case $\rk A = 2$ 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, $r = \rk A$, 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 $\rk A = 2$ 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 $\rk A = 3$ 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. $\rk A \leq 2$,
2. $\cork A \leq 2$ and $d \geq 3$,
3. $\cork A = 3$, $d \geq 5$ 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 Družkowski 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 \( x + (Ax)^2 \) 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.

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 \( JF \) 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, Th. 1.6].

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

**Proof.** From proposition 2.6, it follows that $JH$ is nilpotent. Since a nilpotent matrix over a reduced ring has only eigenvalue zero and the diagonal of a triangular matrix is formed by its eigenvalues, it follows that $JH$ has only zeros on its diagonal.

**Definition 2.8.** Let $f \in \mathbb{C}[x] = \mathbb{C}[x_1, x_2, \ldots, x_n]$. We write $\text{deg} f$ for the total degree of $f$. We write $\text{deg}_{x_i} f$ 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 $\text{deg}_{x_i, x_j, x_k} f$ 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 $Ae_{r+i} = \lambda_{r+i} Ae_{r-1} + \mu_{r+i} Ae_r$. 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 $B := A \cdot JL$. Then the last $n - r$ columns of $B$ and hence those of $JH$
are zero, where

\[ \tilde{H} := L^{-1} \circ H \circ L = \begin{pmatrix} (B_1 x)^d \\ \vdots \\ (B_{r-2} x)^d \\ (B_{r-1} x)^d + \lambda_{r+1}(B_{r+1} x)^d + \cdots + \lambda_n(B_n x)^d \\ (B_r x)^d + \mu_{r+1}(B_{r+1} x)^d + \cdots + \mu_n(B_n x)^d \\ \vdots \\ (B_r x)^d \end{pmatrix} \]

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

\[ \hat{H} = \begin{pmatrix} (B_1 x)^d \\ \vdots \\ (B_{r-2} x)^d \\ p(B_{r-1} x, B_r x) \\ q(B_{r-1} x, B_r x) \end{pmatrix} \]

Furthermore, since the last \( n - r \) columns of \( J\tilde{H} \) are zero, the \((r \times r)\)-matrix \( J\tilde{H} \) is nilpotent as well. In particular, \( \det J\tilde{H} = 0 \). 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

\[ \text{trdeg}_C \tilde{H} = \text{trdeg}_C(B_1 x, \ldots, B_r x) = \text{trdeg}_C(A_1 x, \ldots, A_r x) = r \]

for the first \( r \) rows of \( A \) are linearly independent. This contradicts \( \det J\tilde{H} = 0 \), 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 \( \nu_1 p + \nu_2 q = 0 \) with \( \nu \neq 0 \). Then

\[ \nu_1((B_{r-1} x)^d + \lambda_{r+1}(B_{r+1} x)^d + \cdots + \lambda_n(B_n x)^d) + \\
\nu_2((B_r x)^d + \mu_{r+1}(B_{r+1} x)^d + \cdots + \mu_n(B_n x)^d) = 0 \]

So the components of \((B x)^d\), and hence those of \( H = (A x)^d \) also, are linearly dependent.

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

**Theorem 3.2.** Let \( H := (A x)^d \) 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. Then the components of \( H := (A x)^d \) are linearly dependent.

**Proof.** Since the rows of \( A \) are dependent, the columns are dependent as well. We distinguish two cases:

\[ \Box \]
• 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

$$JH = d \begin{pmatrix} A_{11}(A_1x)^{d-1} & A_{12}(A_1x)^{d-1} & \cdots & A_{1n}(A_1x)^{d-1} \\ A_{21}(A_2x)^{d-1} & A_{22}(A_2x)^{d-1} & \cdots & A_{2n}(A_2x)^{d-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1}(A_nx)^{d-1} & A_{n2}(A_nx)^{d-1} & \cdots & A_{nn}(A_nx)^{d-1} \end{pmatrix}$$

(1)

the expression $\det(TI_n + JH)$, which is $T^n$ on account of the nilpotence of $JH$, can be seen as a polynomial in the transcendent ‘variables’ $A_1x, A_2x, \ldots, A_rx$. 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 $\hat{H} := (0, H_2, \ldots, H_n)$. Next, let

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

where $L = x + \lambda_1^{-1}(0, \lambda_2x_1, \ldots, \lambda_nx_1)$. 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 $Ae_1, A_2e_2, \ldots, A_{r-2}e_{r-2}, Ae_{i_1}, Ae_{i_2}$. 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 $i_1, i_2 \leq r - 1$, both $A_{i_1}$ and $A_{i_2}$ are linear combinations of $A_{r-1}$ and $A_r$. Hence the spaces $\mathbb{C}A_{i_1} + \mathbb{C}A_{i_2}$ and $\mathbb{C}A_{r-1} + \mathbb{C}A_r$ 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. \qed
The proof of theorem 3.2 and its preceding lemma was essentially given by Druzkowski 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

\[
\lambda_1 L_1^d + \lambda_2 L_2^d + \ldots + \lambda_r L_r^d = 0
\]

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

**Proof.** Assume the opposite. In particular, \( L_1 \neq sL_r \) and \( L_r \neq sL_1 \) 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

\[
\mu_1 L_1^{d-1} + \mu_2 L_2^{d-1} + \ldots + \mu_{r-1} L_{r-1}^{d-1} = 0
\]

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 \( d \). \( \square \)

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 \((Ax)^d\) such that cork \( A \leq d - 2 \), \( \text{tr}JH = 0 \), and the diagonal of \( A \) is nonzero. Then there are \( i \neq j \) and an \( s \in \mathbb{C} \) such that \( A_i = sA_j \neq 0 \).

**Proof.** Since the diagonal of \( JH \) 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 \( \text{tr}JH = 0 \), we have

\[
dA_{11}(A_1x)^{d-1} + dA_{22}(A_2x)^{d-1} + \cdots + dA_{nn}(A_ax)^{d-1} = 0 \tag{3}
\]

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 \( A_ix = y_i \) for all \( i \leq r \). Differentiating (3) with respect to \( y_1 \) gives

\[
d(d-1)A_{11}(A_1x)^{d-2} + \lambda_{r+1}(A_{r+1}x)^{d-2} + \cdots + \lambda_n(A_nx)^{d-2} = 0
\]

for certain \( \lambda_i \in \mathbb{C} \). These are \( n - r + 1 \leq d - 1 \) linear powers (powers of linear forms). Now apply lemma 3.3 to get \( A_i = sA_j \) for some \( i \neq j \) and \( s \in \mathbb{C} \) with \( i, j \in \{1, r+1, r+2, \ldots, n\} \). \( \square \)

**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 \( A_i = sA_j \).
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 $d$ 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 $A_i = sA_j$ for some $s \in \mathbb{C}$.

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 $(Ax)^*d$ where $A$ has a small corank. It can be found at the beginning of page 238 in [9].

Lemma 4.1. Let $H = (Ax)^*d$ 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

$$\det(TI_n + d)
\begin{pmatrix}
A_{11}y_1 & A_{12}y_1 & \cdots & A_{1n}y_1 \\
A_{21}y_2 & A_{22}y_2 & \cdots & A_{2n}y_2 \\
\vdots & \vdots & \ddots & \vdots \\
A_{n1}y_n & A_{n2}y_n & \cdots & A_{nn}y_n
\end{pmatrix}
= T^n + R_1(y)T^{n-1} + R_2(y)T^{n-2} + \cdots + R_{n-2}(y)T^2 + R_{n-1}(y)T + R_n(y)$$

Since $\mathcal{J}H$ is nilpotent, $\det(TI_n + \mathcal{J}H) = T^n$. It follows from (1) that the coefficient of $T^{n-j}$ of $\det(TI_n + \mathcal{J}H)$ equals

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

for all $j \geq 1$. Furthermore, it follows from the definition of determinant that $\deg_y R_j \leq 1$ for all $i, j$. For some $j$, $A$ has a principal minor of size $j$ which determinant is $\alpha \neq 0$, say with rows and columns $i_1, i_2, \ldots, i_j$. Then the coefficient of $y_{i_1}y_{i_2}\cdots y_{i_j}$ of $R_j$ equals $da$, whence $R_j \neq 0$.

If $A_k = 0$, then all minors with row $k$ of $A$ have determinant zero, whence $\deg_y R_j = 0$.

In all remaining lemmas in this section, relations $R$ between linear powers $L_i^d, L_2^d, \ldots, L_m^d$ with $\deg_y R \leq 1$ 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
\[
R(x_1^d, x_2^d, \ldots, x_r^d, (\lambda_1 x_1 + \lambda_2 x_1 + \cdots + \lambda_r x_r)^d) = 0
\]
(4)

Then \( \lambda = \lambda_i e_i \) 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
\[
\alpha y_1^{t_1} \cdots y_r^{t_r} y_{r+1}
\]
with \( \alpha \neq 0 \) and \( 0 \leq t_i \leq 1 \) for all \( i \). The coefficient of \( x_1^{d_1} x_2^{d_2} \cdots x_r^{d_r} x_j^{d-1} x_k \) 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. \( \lambda = \lambda_i e_i \) for some \( i \).

Lemma 4.3. Let \( d \geq 2 \) and \( R \) be a nonzero relation with \( \deg y_i, R \leq 1 \) such that
\[
R(x_1^d, x_2^d, \ldots, x_r^d, (\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_r x_r)^d, (\mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_r x_r)^d) = 0
\]
(5)

Assume further that \( \lambda_i = \mu_i = 0 \) for at most \( r - 3 \) \( i \)'s. 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. 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
\[
\alpha y_1^{t_1} y_2^{t_2} \cdots y_r^{t_r} y_{r+1} y_{r+2}
\]
with \( 0 \leq t_i \leq 1 \) for all \( i \). If \( t_{r+1} = 0 \), then by looking at the term
\[
x_1^{d_1} x_2^{d_2} \cdots x_r^{d_r} (x_1^{d-1} x_m)
\]
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
\[
x_1^{d_1} x_2^{d_2} \cdots x_r^{d_r} x_2^{d-1} x_1 x_1^{d-1}
\]
of (5), we see that \( \lambda_l \mu_t = 0 \) for all \( l \geq 3 \). Assume \( \lambda \neq \lambda_2 e_2 \). Then there is an \( l \geq 3 \) such that \( \lambda_l \neq 0 \). So \( \mu_l = 0 \). Looking at the term
\[
x_1^{d_1} x_2^{d_2} \cdots x_r^{d_r} x_2^{d-1} x_1 x_m x_1^{d-1}
\]
gives \( \mu_m = 0 \) for all \( m \geq 3 \). So \( \mu = \mu_1 e_1 \).

So assume \( (\lambda_1, \mu_3, \ldots) \neq 0 \) for \( i = 1, 2 \). Since \( (\lambda_1, \lambda_2) \) and \( (\mu_1, \mu_2) \) are independent, at least three of their four coordinates are nonzero. Assume without loss of generality that \( \lambda_1 \lambda_2 \mu_1 \neq 0 \). If \( \mu_2 = 0 \), then we may assume that \( \mu_3 \neq 0 \) on account of the assumption \( \mu \neq \mu_1 e_1 \).
If $\mu_2 \neq 0$, then $\lambda_1 \lambda_2 \mu_1 \mu_2 \neq 0$. From the assumption $\lambda_i = \mu_i = 0$ for at most $r - 3$ $i$’s, it follows that $\lambda_i \neq 0$ or $\mu_i \neq 0$ for some $i \geq 3$. So without loss of generality, we may assume $\mu_3 \neq 0$. So assume $\mu_3 \neq 0$ regardless of whether $\mu_2 = 0$ or not.

Assume that $(\lambda_2, \lambda_3)$ and $(\mu_2, \mu_3)$ are dependent. Then $\mu_2 \mid \lambda_2 \mu_3 \neq 0$, so $\lambda_2 \mu_2 \neq 0$. If we interchange $(\lambda_1, \mu_1)$ and $(\lambda_2, \mu_2)$, which can be realized by flipping $x_1$ and $x_2$, $(\lambda_2, \lambda_3)$ and $(\mu_2, \mu_3)$ get independent but the condition $\lambda_1\mu_1 \neq 1$ is not affected. So we may assume that $(\lambda_2, \lambda_3)$ and $(\mu_2, \mu_3)$ are independent and in addition $\lambda_1\mu_1 \neq 0$.

ii) We show that the above assumptions lead to a contradiction. Replacing $R$ by $R(y_1, y_2, \ldots, y_r, \lambda_3^2 y_{r+1}, \mu_3^2 y_{r+2})$, we may assume that $\lambda_1 = \mu_1 = 1$. Write $\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_r x_r = x_1 + L$ and similarly $\mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_r x_r = x_1 + M$.

Let $s := \deg_{y_{r+1}, y_{r+2}} R$. Notice that $\deg_{y_i} R \leq 1$ for all $i$. If $s \geq 3$, then $s = 3$ and the left hand side of (5) has degree $3d$ 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
  
  $$R = R_1 y_{r+1} y_{r+2} + R_2 y_1 y_{r+1} + R_3 y_1 y_{r+2} + R_4$$
  
  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
  
  $$(R_1 + R_3)(x_2^d, \ldots, x_r^d) L = -(R_1 + R_3)(x_2^d, \ldots, x_r^d) M$$
  
  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 $s = 2$.  

\[\square\]
Applying lemma 4.2 again gives $A \in \mathcal{s}$ linear transformation, we have $A$ are independent and after a suitable permutation, we have that the rows $P$ again by [9, lemma 1.2], there is a permutation matrix $P$. This transformation may make all principal minor determinants zero, but then, for a suitable linear transformation $T$ such that $P^{-1}AP$ is lower triangular. So we may assume that there is still a nonzero principal minor determinant in $A$ 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 $R$ 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 $A_i = sA_j$ 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 $\text{cork}A \leq 2$, $(A_3x)^{d-1}, \ldots, (A_nx)^{d-1}$ 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 $P$. \hfill \Box

\textbf{Theorem 4.4.} Assume $A$ is a matrix of corank 2 at most, $d \geq 3$ and $H = (Ax)^d$ such that $\mathcal{J}H$ is nilpotent. Then there exists a $T \in \text{GL}_n(\mathbb{C})$ and a lower triangular matrix $B$ such that $T^{-1} \circ (Ax)^d \circ T = (Bx)^d$.

\textbf{Proof.} Assume first that every principal minor of $A$ has determinant zero. From [9, lemma 1.2] (see also [12, prop. 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 $A_i = sA_j$ 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 $\text{cork}A \leq 2$, $(A_3x)^{d-1}, \ldots, (A_nx)^{d-1}$ 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 $P$. \hfill \Box

11
Lemma 4.5. Let \( d \geq 3 \) and \( R \) be a nonzero relation with \( \deg y_1 R \leq 1 \) 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 \( \deg y_{i+1} R = 0 \) and \( \deg y_{i+2} R = 0 \) 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
\[
\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_r x_r = x_1 + L \text{ and } \mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_r x_r = x_1 + M.
\]
Put \( s := \deg y_{i+1} y_{i+2} R \). If \( s \geq 3 \), then \( s = 3 \) and the left hand side of (6) has degree 3d in \( x_1 \); contradiction. Since \( \deg y_{i+1} R \neq 0 \), \( s \geq 1 \). 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 \( \deg y_{i+3} R_0 \leq s \) and \( \deg y_{i+2} R_0 \leq 1 \). 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 follows from \( s = 1 \) that \( R_0 \) is quadratic and \( y_2 \mid R_0 \), for \( R_0 \) is homogeneous. Hence, \( R_0 \) decomposes into linear factors and can be chosen linear instead.

- \( s = 2 \):
  
  Write
  
  \[
  R = R_1 y_{i+1} y_{i+2} + R_2 y_1 y_{i+2} + R_3 y_1 y_{i+1} + R_4
  \]
  with \( R_i \in \mathbb{C}[y_1, \ldots, y_r] \) for all \( i \leq 3 \) and \( \deg y_{i+1} y_{i+2} R_4 \leq 1 \). 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:
  
  \[
  2dR_1 LM + (d-1)(R_1 + R_3)L^2 + (d-1)(R_1 + R_2)M^2
  = 2dR_1 LM - (d-1)R_2 L^2 - (d-1)R_3 M^2
  = 2dR_1 LM + (d-1)R_3 LM + (d-1)R_2 LM
  = (d+1)R_1 LM
  \]
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

\[
T^{-1} \circ (Ax)^s d \circ T = (Bx)^s d
\]

**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, Th. 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 \((Ax)^s d\) with \( \text{rk}A = 2\), but still there exist a triangularization of \((Ax)^s d\) 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 \( J(Ax)^s d \) is nilpotent. Then there exists a \( T \in \text{GL}_n(\mathbb{C}) \) and a lower triangular matrix \( B \) such that

\[
T^{-1} \circ (Ax)^s d \circ T = (Bx)^s d
\]

**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 \( A_{i_1}, A_{i_2}, e_{k_3}^t, \ldots, e_{k_n}^t \) are independent. Replacing \( A \) by \( P^{-1}AP \) for a suitable permutation \( P \) makes that the rows \( A_{j_1}, A_{j_2}, e_{k_3}^t, \ldots, e_{k_n}^t \) are independent.

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

\[
T := \begin{pmatrix} A_{j_1} \\ A_{j_2} \\ e_{k_3}^t \\ \vdots \\ e_{k_n}^t \end{pmatrix}^{-1}
\]

13
Then the last \( n - 2 \) rows of \( T \) are \( e_3^T, \ldots, e_n^T \) 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 \((Bx)^d\) 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 \((Bx)^d\) with \( B \) triangular.

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 \( \text{cork} A = 1 \), \( A_i = s A_j \) 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 \( H = (Ax)^d \) be lower triangular. Then there exists an \( r \) and a \( G = (Bx)^d \) 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 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.

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

\[ G := E^{-1} \circ H \circ E = \begin{pmatrix} 0 \\ 0 \\ 8x_1 x_2 \\ (x_1 - x_2 + x_3)^2 \\ (x_1 - x_2 - x_3)^2 \\ (x_1 + x_2 - x_3)^2 \end{pmatrix} \]
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 \( \lambda^t G = 0 \). 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 \( G(Tx) \) as well, are linearly independent, it follows from \( 0 = \tilde{G}_1 = (T^{-1})_1 G(Tx) \) 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 \( \tilde{G}_3 \in \mathbb{C}[x_1, x_2] \), 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 \( \tilde{G}_3 \) 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 \( \tilde{G}_3 \) is not the square of a linear form.

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

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

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


