Stably tame automorphisms

Engelbert Hubbers* David Wright†

Abstract

In this paper we explore two different methods to find a factorization of $F^{[m]}$ by triangular automorphisms, showing that it is tame, for all $F = X + H$ with $H \in \mathcal{H}_n(A)$ as described in [1]. Furthermore we give an explicit upperbound for this m in both methods.

1 Introduction

Throughout this paper X denotes the sequence $X_1, \ldots, X_n$ and A denotes an arbitrary commutative ring. Hence $A[X] := A[X_1, \ldots, X_n]$ denotes the polynomial ring in n variables over A. Definition 1.1 below defines $\mathcal{H}_n(A)$ for all $n \in \mathbb{N}$. Now if $H \in \mathcal{H}_k(A[X_{k+1}, \ldots, X_n])$ ($k \leq n$ and with respect to the variables $X_1, \ldots, X_k$) and $S = (S_1, \ldots, S_n)$ is a vector, we denote by $H(S)$ the vector H where each X_i is replaced by S_i. This looks like a composition $H \circ S$, but because of the different dimensions it is not exactly the regular composition: in fact it is a real composition in the variables $X_1, \ldots, X_k$ and a substitution in the scalars $X_{k+1}, \ldots, X_n$. Unless otherwise stated $\partial_i = \frac{\partial}{\partial X_i}$. In this paper we use '$\circ$' for composition of polynomial maps and '$*$' for matrix multiplication. In fact we often even omit this '$*$'. Furthermore if g is a vector, we denote the derivation $g_1\partial_1 + \cdots + g_n\partial_n$ by $\mathcal{D}(g; \partial_1, \ldots, \partial_n)$. And hence applying this derivation on h gives: $\mathcal{D}(g; \partial_1, \ldots, \partial_n)(h) = g_1\partial_1 h + \cdots + g_n\partial_n h$.

In order to make the contents of section 2 and 3 understandable we must recall some of the definitions and results from [1] and [3]. In [1] a new class of polynomial maps, denoted by $\mathcal{H}_n(A)$, was introduced and it was shown that for each $H \in \mathcal{H}_n(A)$ the Jacobian matrix JH is nilpotent and the polynomial map $F = X + H$ is invertible over A with $\det(JF) = 1$. We recall the definition of $\mathcal{H}_n(A)$.

Definition 1.1 If $n = 1$ then $\mathcal{H}_1(A) := A$. If $n \geq 2$ we define $\mathcal{H}_n(A)$ inductively: let $H \in A[X]^n$, then $H \in \mathcal{H}_n(A)$ if and only if there exist $T \in M_n(A)$, $c \in A^n$ and $\tilde{H} \in \mathcal{H}_{n-1}(A[X_n])$ such that

$$H = \mathrm{Adj}(T) \begin{pmatrix} \tilde{H} \\ 0 \end{pmatrix}_{|TX} + c.$$

*University of Nijmegen, The Netherlands, hubbers@sci.kun.nl
†Washington University, St. Louis, Missouri, U.S.A., wright@einstein.wustl.edu

In [3] the notion of $\mathcal{D}_n(A)$ was introduced, providing a tool to describe each $H \in \mathcal{H}_n(A)$ in terms of a sequence of matrices and vectors:

Definition 1.2 Let $n \geq 2$. Then $\mathcal{D}_n(A)$ is the set of $(2n-1)$-tuples

$$(T, c) := (T_2, \ldots, T_n, c_1, \ldots, c_n)$$

where $T_n \in M_n(A)$, $T_i \in M_i(A[X_{i+1}, \ldots, X_n])$ for all $2 \leq i \leq n-1$, $c_n \in A^n (= M_{n,1}(A))$ and $c_i \in M_{i,1}(A[X_{i+1}, \ldots, X_n])$ for all $1 \leq i \leq n-1$.

From certain points of view it is nice that these matrices and vectors all have different dimensions. However, in order to really work with these tuples, coercions are needed. As in [3] these coercions are made explicitly by writing $\tilde{T}_i$ and $\tilde{c}_i$:

$$\tilde{T}_i = \begin{pmatrix} T_i & 0 \\ 0 & I_{n-i} \end{pmatrix} \text{ and } \tilde{c}_i = \begin{pmatrix} c_i \\ 0 \end{pmatrix}$$

such that we have $n \times n$ matrices and n-dimensional vectors. Because these coercions are so obvious we often omit the tildes.

Now for $n \geq 2$ and $0 \leq p \leq n-2$ we define a mapping $E_{n,p} : \mathcal{D}_n(A) \to A[X]^n$ such that each $H \in \mathcal{H}_n(A)$ can be written as $c_n + \sum_{p=0}^{n-2} E_{n,p}(T, c)$. For the exact definition we refer to [3]. Here we'll only use proposition 1.3 below. We sometimes omit the (T, c) part and simply write $E_{n,p}$. This $E_{n,p}$-mapping is used to provide a link between polynomial maps and derivations. Normally we compute these $E_{n,p}$'s using proposition 1.3. This proposition uses a non-standard, associative, matrix multiplication denoted by 'Δ', which is defined by:

$$S \Delta T := S(T * X) * T$$

for all $S, T \in M_n(A[X])$.

Proposition 1.3 *Let $n \geq 2$, $0 \leq p \leq n-2$ and $(T, c) \in \mathcal{D}_n(A)$. Then*

$$E_{n,p}(T, c) = \operatorname{Adj}(\tilde{T}_{n-p} \Delta \cdots \Delta \tilde{T}_{n-1} \Delta T_n) * \left(\tilde{c}_{n-p-1} \,|_{(\tilde{T}_{n-p} \Delta \cdots \Delta \tilde{T}_{n-1} \Delta T_n) * X} \right)$$

See the proof of [3, proposition 1.5].

Remark 1.4 Note that proposition 1.3 means that $E_{n,p}(T, c)$ only depends on the tuple $(T_{n-p}, \ldots, T_n)$ and the vector c_{n-p-1}. Therefore we introduce a small modification to the notation of definition 1.2. We write

$$E_{n,p}(T; c_{n-p-1})$$

if we mean $E_{n,p}(T', c')$ for some $(T', c') \in \mathcal{D}_n(A)$ where $(T', c') = (T'_2, \ldots, T'_{n-p-1}, T_{n-p}, \ldots, T_n, c'_1, \ldots, c'_{n-p-2}, c_{n-p-1}, c'_{n-p}, \ldots, c'_n)$.

So the semicolon shows that we use a 'stripped' version of $(T, c) \in \mathcal{D}_n(A)$. We make further abuse of the notation by still saying $(T; c) \in \mathcal{D}_n(A)$.

We try to clarify these notions by giving an example.

Example 1.5 Take $F = X + H$ where

$$H = (X_4(X_3X_1 + X_4X_2)^2, -X_3(X_3X_1 + X_4X_2)^2, X_4^3, 0) \in \mathcal{H}_4(\mathbb{C})$$

Then we can find $(T, c) \in \mathcal{D}_4(\mathbb{C})$ such that $H = c_n + \sum_{p=0}^{n-2} E_{n,p}(T, c)$. For instance take

$$T := \left(\begin{pmatrix} 1 & 0 \\ X_3 & X_4 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right)$$

and

$$c := \left(\begin{pmatrix} X_2^2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ X_4^3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right)$$

Simple computations yield:

$$E_{4,0} = \begin{pmatrix} 0 \\ 0 \\ X_4^3 \\ 0 \end{pmatrix}, \quad E_{4,1} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad E_{4,2} = \begin{pmatrix} X_4(X_3X_1 + X_4X_2)^2 \\ -X_3(X_3X_1 + X_4X_2)^2 \\ 0 \\ 0 \end{pmatrix}$$

And from this we get the derivations:

$$\begin{aligned} \mathcal{D}(E_{4,0}(T, c); \partial_1, \ldots, \partial_4) &= X_4^3 \partial_3 \\ \mathcal{D}(E_{4,1}(T, c); \partial_1, \ldots, \partial_4) &= 0 \\ \mathcal{D}(E_{4,2}(T, c); \partial_1, \ldots, \partial_4) &= X_4(X_3X_1 + X_4X_2)^2\partial_1 - X_3(X_3X_1 + X_4X_2)^2\partial_2 \end{aligned}$$

The H in this example is the crucial part of the counterexamples to the Discrete Markus-Yamabe Problem and the Deng-Meisters-Zampieri Conjecture. See [2] for details.

2 The quick method

The first approach we present in this paper only deals with $\mathcal{H}_n(A)$. It acts on the level of polynomial maps; no $\mathcal{D}_n(A)$ or derivations are used. This method provides an explicit recipe to factor $F^{[\frac{n(n-1)}{2}]}$ by triangular automorphisms.

We start by looking at the two-dimensional case.

Example 2.1 Let $F = X + H$ with $H \in \mathcal{H}_2(A)$. From [1] it follows that F can be written as

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \mathrm{Adj}(T) \begin{pmatrix} f \\ 0 \end{pmatrix}_{|TX} + c = \begin{pmatrix} X_1 + a_2 f(a_1X_1 + a_2X_2) + c_1 \\ X_2 - a_1 f(a_1X_1 + a_2X_2) + c_2 \end{pmatrix}$$

where $T = \begin{pmatrix} 1 & 0 \\ a_1 & a_2 \end{pmatrix} \in M_2(A)$, $c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \in A^2$ and $f \in A[X_2]$. Now extend this F to $F^{[1]} := (F_1, F_2, X_3)$. Define $P := (X_1, X_2, X_3 + f(a_1X_1 + a_2X_2))$. Then

$$F^{[1]} \circ P = \begin{pmatrix} X_1 + a_2 f(a_1X_1 + a_2X_2) + c_1 \\ X_2 - a_1 f(a_1X_1 + a_2X_2) + c_2 \\ X_3 + f(a_1X_1 + a_2X_2) \end{pmatrix}$$

Define also $Q := (X_1 - c_1, X_2 - c_2, X_3)$ and $R := (X_1 - a_2X_3, X_2 + a_1X_3, X_3)$. Then

$$R \circ Q \circ F^{[1]} \circ P = \begin{pmatrix} X_1 - a_2X_3 \\ X_2 + a_1X_3 \\ X_3 + f(a_1X_1 + a_2X_2) \end{pmatrix}$$

And hence

$$\begin{aligned} R \circ Q \circ F^{[1]} \circ P \circ R^{-1} &= \begin{pmatrix} X_1 \\ X_2 \\ X_3 + f(a_1(X_1 + a_2X_3) + a_2(X_2 - a_1X_3)) \end{pmatrix} \\ &= \begin{pmatrix} X_1 \\ X_2 \\ X_3 + f(a_1X_1 + a_2X_2) \end{pmatrix} \\ &= P \end{aligned}$$

which is triangular.

If we take another look at the definition of P, Q and R, we see that we can describe these maps directly in terms of T, c and f:

$$P = \begin{pmatrix} X \\ X_3 + f(TX) \end{pmatrix}, \quad Q = \begin{pmatrix} X - c \\ X_3 \end{pmatrix}, \quad R = \begin{pmatrix} X - \mathrm{Adj}(T) \begin{pmatrix} X_3 \\ 0 \end{pmatrix} \\ X_3 \end{pmatrix}$$

And it is exactly this idea that gives us the quick method of factorization. Note that $Q \circ F^{[1]} = R^{-1} \circ P \circ R \circ P^{-1}$, a commutator.

Before we can present the main theorem of this section, we present the following lemma which is an extension of [1, Lemma 2.1].

Lemma 2.2 *Let $H(X_1, \ldots, X_n) \in \mathcal{H}_n(A)$, $d \in A$ and $a = (a_1, \ldots, a_n) \in A^n$. Then also $H(dX_1 + a_1, \ldots, dX_n + a_n) \in \mathcal{H}_n(A)$.*

Proof. Induction on n. If $n = 1$ we get $H(X_1) \in \mathcal{H}_1(A) = A$ and hence $H(X_1)$ is a constant and hence also $H(dX_1 + a_1) \in \mathcal{H}_1(A)$. Now assume $n \geq 2$. We want to show that $H(dX_1 + a_1, \ldots, dX_n + a_n) \in \mathcal{H}_n(A)$. Note that we can write this map as the composition of three maps: H, a translation Tr_a over a and a multiplication D_d with d, where

$$\begin{aligned} Tr_a &= (X_1 + a_1, \ldots, X_n + a_n) \\ D_d &= (dX_1, \ldots, dX_n) \end{aligned}$$

Hence we get:

$$\begin{aligned} & H(dX_1+a_1,\ldots,dX_n+a_n) = \\ = \; & H \circ Tr_a \circ D_d \\ = \; & \left(Tr_c \circ \operatorname{Adj}(T)\begin{pmatrix}\tilde{H}\\0\end{pmatrix} \circ TX\right) \circ Tr_a \circ D_d \\ = \; & Tr_c \circ \operatorname{Adj}(T)\begin{pmatrix}\tilde{H}\\0\end{pmatrix} \circ Tr_{\tilde{a}} \circ TX \circ D_d \qquad (\tilde{a} := Ta) \\ = \; & Tr_c \circ \operatorname{Adj}(T)\begin{pmatrix}\tilde{H}\\0\end{pmatrix} \circ Tr_{\tilde{a}} \circ D_d \circ TX \\ = \; & Tr_c \circ \operatorname{Adj}(T)\begin{pmatrix}\tilde{H}(dX_1+\tilde{a}_1,\ldots,dX_n+\tilde{a}_n)\\0\end{pmatrix} \circ TX \end{aligned}$$

Now we are done if we can show that $\tilde{H}(dX_1+\tilde{a}_1,\ldots,dX_n+\tilde{a}_n) \in \mathcal{H}_{n-1}(A[X_n])$. Note that $X_1,\ldots,X_{n-1}$ are variables and X_n is a constant in $\mathcal{H}_{n-1}(A[X_n])$. By this notion we can apply [1, Lemma 2.1] with the substitution homomorphism

$$\begin{aligned} A[X_n] & \rightarrow A[X_n] \\ X_n & \mapsto dX_n+\tilde{a}_n \end{aligned}$$

to see that $\hat{H} = \tilde{H}(dX_n+\tilde{a}_n) \in \mathcal{H}_{n-1}(A[X_n])$. But now we can apply the induction hypothesis on $\hat{H}$ and the ring $A[X_n]$ to get that $\hat{H}(dX_1+\tilde{a}_1,\ldots,dX_{n-1}+\tilde{a}_{n-1}) \in \mathcal{H}_{n-1}(A[X_n])$. And hence

$$\begin{aligned} & \tilde{H}(dX_1+\tilde{a}_1,\ldots,dX_n+\tilde{a}_n) = \\ = \; & \hat{H}(dX_1+\tilde{a}_1,\ldots,dX_{n-1}+\tilde{a}_{n-1}) \in \mathcal{H}_{n-1}(A[X_n]). \end{aligned}$$

This proves the lemma. □

Now we can formulate the main theorem of this section:

Theorem 2.3 *Let $F = X + H$ with $H \in \mathcal{H}_n(A)$. Then there exist tame automorphisms U and V of $A[X_1,\ldots,X_n,Y_1,\ldots,Y_{n-1}]$ such that $U \circ F^{[n-1]} \circ V$ is of the form $(X, Y+H')$ where $H' \in \mathcal{H}_{n-1}(A[X_1,\ldots,X_n])$ (with respect to the variables $Y_1,\ldots,Y_{n-1}$).*

Proof. Let

$$H = \operatorname{Adj}(T)\begin{pmatrix}\tilde{H}\\0\end{pmatrix}_{|TX} + c$$

with $T \in M_n(A)$, $c \in A^n$ and $\tilde{H} \in \mathcal{H}_{n-1}(A[X_n])$ (with respect to the variables $X_1,\ldots,X_n$). Put $d = \det(T)$ and $TX = (L_1,\ldots,L_n)$. Define

$$P := \begin{pmatrix}X\\Y+\tilde{H}(TX)\end{pmatrix}, \quad Q := \begin{pmatrix}X-c\\Y\end{pmatrix}, \quad R := \begin{pmatrix}X-\operatorname{Adj}(T)\begin{pmatrix}Y\\0\end{pmatrix}\\Y\end{pmatrix}$$

$$U := R \circ Q, \quad V := P \circ R^{-1}$$

Then

$$U \circ F^{[n-1]} \circ V = \begin{pmatrix} X \\ Y + \tilde{H}\left(TX + d\begin{pmatrix} Y \\ 0 \end{pmatrix}\right) \end{pmatrix}$$

And if we can show that $H' := \tilde{H}(TX + d(Y,0)) \in \mathcal{H}_{n-1}(A[X_1,\ldots,X_n])$ with respect to the variables $Y_1,\ldots,Y_{n-1}$, this theorem is proved. To do this, note that $\tilde{H}(X_1,\ldots,X_n) \in \mathcal{H}_{n-1}(A[X_n])$ with respect to the variables $X_1,\ldots,X_{n-1}$. In particular X_n is not a variable but a scalar. So obviously

$$\begin{array}{rcl} A[X_n][X_1,\ldots,X_{n-1}] & \to & A[X_n][Y_1,\ldots,Y_{n-1}] \\ X_n & \mapsto & X_n \\ X_i & \mapsto & Y_i \quad \text{for} \quad i < n \end{array}$$

shows that $\tilde{H}(Y_1,\ldots,Y_{n-1}) \in \mathcal{H}_{n-1}(A[X_n])$ with respect to the variables $Y_1,\ldots,Y_{n-1}$.

Now consider the homomorphism $\varphi : A[X_n] \to A[X_1,\ldots,X_n]$ with $\varphi(X_n) = L_n$. Apply [1, Lemma 2.1] and see that $\hat{H} := \tilde{H}(Y_1,\ldots,Y_{n-1},L_n) \in \mathcal{H}_{n-1}(A[X_1,\ldots,X_n])$ with respect to the variables $Y_1,\ldots,Y_{n-1}$. Finally apply lemma 2.2 to $\hat{H}(Y_1,\ldots,Y_{n-1})$ and the ring $A[X_1,\ldots,X_n]$ to conclude that

$$\begin{array}{rcl} H' & = & \tilde{H}(TX + d(Y,0)) \\ & = & \tilde{H}(dY_1 + L_1,\ldots,dY_{n-1} + L_{n-1}, L_n) \\ & = & \hat{H}(Y_1,\ldots,Y_{n-1})(dY_1 + L_1,\ldots,dY_{n-1} + L_{n-1}) \\ & \in & \mathcal{H}_{n-1}(A[X_1,\ldots,X_n]) \end{array}$$

with respect to the variables $Y_1,\ldots,Y_{n-1}$. This completes the proof. □

Corollary 2.4 *Let F be as in theorem 2.3. Then $F^{\left[\frac{n(n-1)}{2}\right]}$ is tame.*

Proof. With induction on n. If $n = 1$ all is clear. So assume $n \geq 2$. By theorem 2.3 we have that there exist tame automorphisms U and V such that

$$G = U \circ F^{[n-1]} \circ V = (X, Y + H')$$

with $H' \in \mathcal{H}_{n-1}(A[X_1,\ldots,X_n])$. Now by induction we know that

$$G^{\left[\frac{(n-1)(n-2)}{2}\right]}$$

is tame and hence

$$F^{\left[n-1+\frac{(n-1)(n-2)}{2}\right]}$$

is tame. Obviously $n - 1 + \frac{(n-1)(n-2)}{2} = \frac{n(n-1)}{2}$, which proves the corollary. □

We conclude this section by applying this quick method to the map of example 1.5. Although this is basically a simple example, the computations are already pretty complicated. Hence all computations are done using Maple. In order to save some space, we display the vectors horizontally.

Example 2.5 Let F and $(T, c) \in \mathcal{D}_4(\mathbb{C})$ be as in example 1.5. Then by theorem 2.3 we can define:

$$\begin{aligned} P_1 &:= \Big(X_1, X_2, X_3, X_4, X_5 + X_4\,(X_3\,X_1 + X_4\,X_2)^2, \\ &\quad X_6 - X_3\,(X_3\,X_1 + X_4\,X_2)^2, X_7 + {X_4}^3 \Big) \\ Q_1 &:= (X_1, X_2, X_3, X_4, X_5, X_6, X_7) \\ R_1 &:= (X_1 - X_5, X_2 - X_6, X_3 - X_7, X_4, X_5, X_6, X_7) \end{aligned}$$

Composition of the maps in the appropriate order gives:

$$\begin{aligned} &R_1 \circ Q_1 \circ F^{[3]} \circ P_1 \circ R_1^{-1} = \\ &\Big(X_1, X_2, X_3, X_4, \\ &X_5 + X_4\,(X_7\,X_1 + X_3\,X_1 + X_6\,X_4 + X_4\,X_2 + X_5\,X_7 + X_5\,X_3)^2, \\ &X_6 - (X_3 + X_7)\,(X_7\,X_1 + X_3\,X_1 + X_6\,X_4 + X_4\,X_2 + X_5\,X_7 + X_5\,X_3)^2, \\ &X_7 + {X_4}^3\Big) \end{aligned}$$

Now if we restrict ourselves to the last three components, we can find a describing tuple $(T', c') \in \mathcal{D}_3(\mathbb{C}[X_1, X_2, X_3, X_4])$ for this three-dimensional map:

$$T' := \left(\begin{pmatrix} 1 & 0 \\ X_3 + X_7 & X_4 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right)$$

$$\begin{aligned} c' := \Big(&(X_6\,(2\,X_4\,X_2 + 2\,X_3\,X_1 + 2\,X_7\,X_1 + X_6)), \\ &\big(X_7\,X_4\,X_1\,(2\,X_3\,X_1 + X_7\,X_1 + 2\,X_4\,X_2), -X_7\,(4\,X_3\,X_1\,X_4\,X_2 \\ &+2\,X_7\,X_4\,X_1\,X_2 + 3\,{X_3}^2\,{X_1}^2 + 3\,X_7\,X_3\,{X_1}^2 + {X_4}^2\,{X_2}^2 + {X_7}^2\,{X_1}^2)\big), \\ &\big(X_4\,(X_3\,X_1 + X_4\,X_2)^2, -X_3\,(X_3\,X_1 + X_4\,X_2)^2, {X_4}^3\big)\Big) \end{aligned}$$

Now applying theorem 2.3 on this tuple (T', c') gives the maps[1] P_2, Q_2 and R_2 such that the last two components of

$$R_2 \circ Q_2 \circ \left(R_1 \circ Q_1 \circ F^{[3]} \circ P_1 \circ R_1^{-1}\right)^{[2]} \circ P_2 \circ R_2^{-1}$$

[1] In order to save some space we do not present all details in this example. The computations can be checked easily by a computer.

can be seen as an element of $\mathcal{H}_2(\mathbb{C}[X_1, X_2, X_3, X_4, X_5, X_6, X_7])$ and hence we can find a describing tuple $(T'', c'') \in \mathcal{D}_2(\mathbb{C}[X_1, X_2, X_3, X_4, X_5, X_6, X_7])$. We use this tuple to define the last couple of automorphisms following the scheme of theorem 2.3: P_3, Q_3 and R_3. And combining all these automorphisms one gets:

$$R_3 \circ Q_3 \circ \left(R_2 \circ Q_2 \circ \left(R_1 \circ Q_1 \circ F^{[3]} \circ P_1 \circ R_1^{-1} \right)^{[2]} \circ P_2 \circ R_2^{-1} \right)^{[1]} \circ P_3 \circ R_3^{-1} =$$

$$\Big(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10} + (X_3\, X_8 + X_7\, X_8 + X_4\, X_9)$$
$$(X_3\, X_8 + X_7\, X_8 + X_4\, X_9 + 2\, X_5\, X_7 + 2\, X_7\, X_1 + 2\, X_6\, X_4$$
$$+ 2\, X_4\, X_2 + 2\, X_3\, X_1 + 2\, X_5\, X_3)\Big)$$

which is a triangular map and hence tame. In correspondence with corollary 2.4 we have added 6 new variables.

3 The stronger method

The quick method in the previous section is based on the $\mathcal{H}_n(A)$-structure. The method in this section also uses the notion of $\mathcal{D}_n(A)$ and derivations. The benefit of adding this extra structure lies in the fact that we get a sharper upper bound if $n \geq 3$ for the number of extra variables needed compared to corollary 2.4: $n-1$ instead of $\frac{n(n-1)}{2}$. Therefore we named it the stronger method.

We recall the main theorem of the paper [3].

Theorem 3.1 *Let $F = X + H$, where $H = \sum_{p=0}^{n-2} E_{n,p}(T, c) + c_n$, for some $(T, c) \in \mathcal{D}_n(A)$. Then*

$$F = \exp(\mathcal{D}\,(c_n; \partial_1, \ldots, \partial_n)) \prod_{p=0}^{n-2} \exp(\mathcal{D}\,(E_{n,p}(T, c); \partial_1, \ldots, \partial_n)).$$

From remark 3.2 it follows that we can restrict ourselves to the automorphism $X + E_{n,p}(T, c)$ with $p = n-2$.

Remark 3.2 If $(\exp(a_i D_i), t_1, \ldots, t_{m_i})$ is a tame automorphism for $i = 1, \ldots, k$ and $m_i, k \in \mathbb{N}$, then

$$\prod_{i=1}^{k} (\exp(a_i D_i), t_1, \ldots, t_m) = (\prod_{i=1}^{k} \exp(a_i D_i), t_1, \ldots, t_m)$$

where $m = \max\{m_1, \ldots, m_k\}$, is a tame automorphism.

So if we can show that we can reduce $X + E_{n,p}(T, c)$ by adding $p+1$ new variables we have accomplished our goal.

Proposition 3.3 *If $F = X + E_{n,p}(T,c)$ then $F^{[p+1]}$ is tame.*

The proof is split into several parts.

By remark 1.4 we see that we can focus on $E_{n,p}(T;c_{n-p-1})$ instead of $E_{n,p}(T,c)$. Now write c_{n-p-1} as $f = (f_1,\dots,f_{n-p-1})$ where each $f_i \in A[X_{n-p},\dots,X_n]$. The next lemma provides another reduction without loss of generality.

Lemma 3.4 *Let $(T;c),(T;d) \in \mathcal{D}_n(A)$. Then*

$$(X + E_{n,p}(T;c)) \circ (X + E_{n,p}(T;d)) = X + E_{n,p}(T;c+d)$$

Proof. The key step here is to prove that

$$E_{n,p}(T;c) \circ (X + E_{n,p}(T;d)) = E_{n,p}(T;c) \quad (1)$$

For $c = (c_1,\dots,c_{n-p-1},0,\dots,0)$ and $d = (d_1,\dots,d_{n-p-1},0,\dots,0)$. Both c and d are in $A[X_{n-p},\dots,X_n]^n$. The proof of (1) goes by induction on p.

- $p = 0$. By definition we have

  $$E_{n,p}(T;c) = \mathrm{Adj}(T_n)\left((c_1,\dots,c_{n-1},0)_{|T_nX}\right)$$

  where $c_i \in A[X_n]$. The same holds for $E_{n,p}(T;d)$. Since the coefficients of T_n are scalars, this can be written in the following way as the composition of polynomial maps

  $$[\mathrm{Adj}(T_n)X] \circ (c_1,\dots,c_{n-1},0) \circ T_nX$$

  Let $\delta = \det(T_n) \in A$. Then

  $$\begin{aligned} & E_{n,0}(T;c) \circ (X + E_{n,0}(T;d)) \\ = \;& [\mathrm{Adj}(T_n)X] \circ (c_1,\dots,c_{n-1},0) \circ T_nX \\ & \circ (X + [\mathrm{Adj}(T_n)X] \circ (d_1,\dots,d_{n-1},0) \circ T_nX) \\ = \;& [\mathrm{Adj}(T_n)X] \circ (c_1,\dots,c_{n-1},0) \circ (T_nX + (\delta X) \circ (d_1,\dots,d_{n-1},0) \circ T_nX) \end{aligned} \quad (2)$$

  We note that the n^{th} coordinate function of $(\delta X) \circ (d_1,\dots,d_{n-1},0) \circ T_nX$ is 0, and since $c_1,\dots,c_{n-1}$ only involve X_n, the composition $(c_1,\dots,c_{n-1},0) \circ (T_nX + (\delta X) \circ (d_1,\dots,d_{n-1},0) \circ T_nX)$ is equal to $(c_1,\dots,c_{n-1},0) \circ T_nX$. Thus the composition of (2) is equal to

  $$[\mathrm{Adj}(T_n)X] \circ (c_1,\dots,c_{n-1},0) \circ T_nX = E_{n,0}(T;c)$$

  as desired.

- $p > 0$. By the inductive definition we have

  $$E_{n,p}(T;c) = \mathrm{Adj}(T_n)\left((E_{n-1,p-1}(T';c'),0)_{|T_nX}\right)$$

(in the notation of [3]), which, again since T_n is a scalar, can be written as the polynomial composition:

$$[\mathrm{Adj}(T_n)X] \circ (E_{n-1,p-1}(T';c'),0) \circ T_nX$$

And hence

$$\begin{aligned}
&E_{n,p}(T;c) \circ (X + E_{n,p}(T;d)) \\
&\quad= [\mathrm{Adj}(T_n)X] \circ (E_{n-1,p-1}(T',c'),0) \circ T_nX \\
&\qquad \circ (X + [\mathrm{Adj}(T_n)X] \circ (E_{n-1,p-1}(T';d'),0) \circ T_nX) \\
&\quad= [\mathrm{Adj}(T_n)X] \circ (E_{n-1,p-1}(T',c'),0) \circ \\
&\qquad (T_nX + \underbrace{\delta(E_{n-1,p-1}(T';d'),0)}_{=(E_{n-1,p-1}(T';\delta d'),0)\text{ (prop. 1.3)}} \circ T_nX) \\
&\quad= [\mathrm{Adj}(T_n)X] \circ (E_{n-1,p-1}(T',c'),0) \circ (T_nX + (E_{n-1,p-1}(T';\delta d'),0) \circ T_nX) \\
&\quad= [\mathrm{Adj}(T_n)X] \circ \underbrace{(E_{n-1,p-1}(T',c'),0) \circ (X + (E_{n-1,p-1}(T';\delta d'),0))}_{=E_{n-1,p-1}(T',c')\text{ (induction)}} \circ T_nX \\
&\quad= [\mathrm{Adj}(T_n)X] \circ (E_{n-1,p-1}(T',c'),0) \circ T_nX \\
&\quad= E_{n,p}(T;c)
\end{aligned}$$

(where again $\delta = \det(T_n) \in A$).

It now easily follows that $(X + E_{n,p}(T;c)) \circ (X + E_{n,p}(T;d)) = X + E_{n,p}(T;c+d)$ using (1) and proposition 1.3. □

In some sense this lemma says that $X + E_{n,p}(T;c)$ is additive in c. The impact is that we can split our $X + E_{n,p}(T;f)$ into

$$\begin{aligned}
&(X + E_{n,p}(T;(f_1,0,\ldots,0)))\circ \\
&\quad (X + E_{n,p}(T;(0,f_2,0,\ldots,0))) \circ \cdots \circ (X + E_{n,p}(T;(0,\ldots,0,f_{n-p-1})))
\end{aligned}$$

And for the purpose of reducing to a tame automorphism this means that we can restrict to one general $X + E_{n,p}(T;(0,\ldots,0,f_i,0,\ldots,0))$ for some $1 \le i \le n-p-1$.

The next step in the process is the observation:

Lemma 3.5 *$X + E_{n,p}(T;(0,\ldots,0,f_i,0,\ldots,0)) = \exp(hD)$ where D is a locally nilpotent derivation and $h \in A[X_1,\ldots,X_n]$.*

Proof. Proposition 1.3 shows

$$E_{n,p}(T,f) =$$

$$\mathrm{Adj}(T_{n-p}\Delta T_{n-p+1}\Delta\cdots\Delta T_n)\begin{pmatrix} 0 \\ \vdots \\ 0 \\ f_i(X_{n-p},\ldots,X_n) \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{|(T_{n-p}\Delta T_{n-p+1}\Delta\cdots\Delta T_n)X}$$

It is obvious that we can split this object into two smaller parts:

$$g := \mathrm{Adj}(T_{n-p}\Delta T_{n-p+1}\Delta\cdots\Delta T_n)e_i$$

where e_i is the i-th unit vector and

$$h := f_i(X_{n-p},\ldots,X_n)_{|(T_{n-p}\Delta T_{n-p+1}\Delta\cdots\Delta T_n)X}$$

Multiplying these two factors gives back the complete result. Lemma 3.6 below shows that

$$\begin{aligned} h &= f_i(X_{n-p},\ldots,X_n)_{|(T_{n-p}X\circ T_{n-p+1}X\circ\cdots\circ T_nX)} \\ & f_i(T_{n-p}X\circ T_{n-p+1}X\circ\cdots\circ T_nX) \end{aligned}$$

Now let $D = \mathcal{D}(g;\partial_1,\ldots,\partial_n)$. Then hD is the same derivation as presented in corollary 3.4 in [3]. And there it is shown that this derivation is locally nilpotent and $X + E_{n,p}(T,c) = \exp(hD)$. □

Lemma 3.6 *The 'Δ' operator has the property:*

$$(S_1\Delta S_2\Delta\cdots\Delta S_k)X = S_1X\circ S_2X\circ\cdots\circ S_kX \text{ for all } k\geq 2$$

Proof. The proof goes by induction on k. Note that $(S_1\Delta S_2)X = (S_1(S_2X)*S_2)*X = S_1(S_2X)*S_2*X = S_1(S_2X)*S_2X = S_1X\circ S_2X$, which proves the case $k=2$. Now for $k>2$ we have:

$$\begin{aligned} (S_1\Delta S_2\Delta\cdots\Delta S_k)X &= (S_1\Delta(S_2\Delta\cdots\Delta S_k))X \\ &= S_1X\circ(S_2\Delta\cdots\Delta S_k)X \\ &= S_1X\circ S_2X\circ\cdots\circ S_kX \end{aligned}$$

which proves the lemma. □

At this point we introduce a new set of matrices. We use it in the next step of the proof of proposition 3.3.

$$S_{n-r} = \begin{cases} r=0: & T_n \\ r>0: & T_{n-r}(S_{n-(r-1)}\cdots S_nX) \end{cases}$$

Note that the matrix S_{n-r} has the form

$$\begin{pmatrix} S & 0 \\ 0 & I_r \end{pmatrix}$$

for some S because T_{n-r} has this form. Furthermore this means that

$$\mathrm{Adj}(S_{n-r}) = \begin{pmatrix} \mathrm{Adj}(S) & 0 \\ 0 & \delta I_r \end{pmatrix}$$

where $\delta = \det(S)$.

Lemma 3.7 $T_{n-p}\Delta T_{n-p+1}\Delta \cdots \Delta T_n = S_{n-p} \cdot S_{n-p+1} \cdots S_n$, *where* S_n *is defined as above.*

Proof. The proof is with induction on p. As usual, the case $p = 0$ is clear, hence assume $p > 0$. Then

$$\begin{aligned} & T_{n-p}\Delta T_{n-p+1}\Delta \cdots \Delta T_n \\ &= T_{n-p}\Delta(T_{n-p+1}\Delta \cdots \Delta T_n) \\ &= T_{n-p}((T_{n-p+1}\Delta \cdots \Delta T_n)X) \cdot (T_{n-p+1}\Delta \cdots \Delta T_n)) \\ &= T_{n-p}(S_{n-p+1} \cdots S_n X) \cdot (S_{n-p+1} \cdots S_n) \\ &= S_{n-p} S_{n-p+1} \cdots S_n \end{aligned}$$

□

Now that we have written $X + E_{n,p}(T; f)$ as $\exp(hD)$, the next step in the proof is using Martha Smith's result of [4]. From her paper it follows that if $a \in \ker(D)$ and $\rho = (X, t + a)$ then

$$(\exp(aD), t) = \exp(tD)\rho \exp(-tD)\rho^{-1}$$

and hence we are reduced to factoring $\exp(tD)$. In order to exploit this step we have to show that $h \in \ker(D)$.

Lemma 3.8 $D(h) = 0$.

Proof. Let $(H_1, \ldots, H_n)$ be the coordinate functions of the map $(T_{n-p}\Delta \cdots \Delta T_n)X$. We have

$$\begin{aligned} h &= f_i(X_{n-p}, \ldots, X_n)_{|(T_{n-p}\Delta\cdots\Delta T_n)X} \\ &= f_i(H_{n-p}, \ldots, H_n) \end{aligned}$$

So it suffices to show that D kills $H_{n-p}, \ldots, H_n$.

- $D(H_n) = 0$. Since $(T_{n-p}\Delta \cdots \Delta T_n)X = T_{n-p}X \circ \cdots \circ T_n X$ (lemma 3.6) and since $T_{n-p}X, \ldots, T_{n-1}X$ fix X_n, it is clear that $H_n = a_{n,1}X_1 + \cdots + a_{n,n}X_n$ where $T_n = (a_{i,j}) \in M_n(A)$. Let $(b_{i,j}) = \mathrm{Adj}(T_n) = \mathrm{Adj}(S_n)$. Then

$$\begin{aligned} D(H_n) &= \mathcal{D}(\mathrm{Adj}(T_{n-p}\Delta \cdots \Delta T_n)e_i; \partial_1, \ldots, \partial_n)(H_n) \\ &= \mathcal{D}(\mathrm{Adj}(S_{n-p} \cdots S_n)e_i; \partial_1, \ldots, \partial_n)(H_n) \\ &= \mathcal{D}(\mathrm{Adj}(S_n)\,\mathrm{Adj}(S_{n-p} \cdots S_{n-1})e_i; \partial_1, \ldots, \partial_n)(H_n) \\ &= (\partial_1 H_n, \ldots, \partial_n H_n)\,\mathrm{Adj}(S_n)\,\mathrm{Adj}(S_{n-p} \cdots S_{n-1})e_i \\ &= (a_{n,1}, \ldots, a_{n,n})(b_{i,j})\,\mathrm{Adj}(S_{n-p} \cdots S_{n-1})e_i \\ &= (0, \ldots, 0, \delta)\,\mathrm{Adj}(S_{n-p} \cdots S_{n-1})e_i \end{aligned}$$

where $\delta = \det(T_n)$. We have seen before that $\mathrm{Adj}(S_{n-p} \cdots S_{n-1})$ has the form

$$\begin{pmatrix} * & \ldots & * & 0 \\ \vdots & & \vdots & \vdots \\ * & \ldots & * & 0 \\ 0 & \ldots & 0 & * \end{pmatrix}$$

and therefore

$$\begin{aligned} D(H_n) &= (0, \ldots, 0, \delta)\,\mathrm{Adj}(S_{n-p} \cdots S_{n-1})e_i \\ &= (0, \ldots, 0, *)e_i \\ &= 0 \end{aligned}$$

since $i \le n - p - 1 < n$.

- $D(H_r) = 0$ for $n - p \le r < n$. Let $G \in A[X_n]^{n-1}$ where

$$G = (G_1, \ldots, G_{n-1}) = (T_{n-p}\Delta \cdots \Delta T_{n-1}) \begin{pmatrix} X_1 \\ \vdots \\ X_{n-1} \end{pmatrix}$$

Let $L = (L_1, \ldots, L_n) = T_n X$. Then

$$\begin{aligned} (H_1, \ldots, H_n) &= (G, X_n) \circ L \\ &= (G_1(L), \ldots, G_{n-1}(L), L_n) \end{aligned}$$

So in particular $H_r = G_r(L)$.

Furthermore let $S = S_{n-p} \cdots S_n$, and $S' = S'_{n-p} \cdots S'_{n-1} = T_{n-p}\Delta \cdots \Delta T_{n-1}$ and note that

$$S = S'(L) \cdot T_n \tag{3}$$

As an $n \times n$ matrix, S' has the form $\begin{pmatrix} S' & 0 \\ 0 & 1 \end{pmatrix}$. (We assume the reader won't be disturbed by the slight abuse of notation in the double use of S'.) Hence $\mathrm{Adj}(S')$ has the form

$$(s_{i,j}) = \begin{pmatrix} \mathrm{Adj}(S') & 0 \\ 0 & d \end{pmatrix}$$

where $d = \det(S')$ and therefore $\mathrm{Adj}(S'(L))$ has the form

$$\begin{pmatrix} \mathrm{Adj}(S')(L) & 0 \\ 0 & d(L) \end{pmatrix}$$

Now

$$\begin{aligned} D(H_r) &= D(G_r(L)) \\ &= \mathcal{D}(\mathrm{Adj}(T_{n-p}\Delta \cdots \Delta T_n)e_i; \partial_1, \ldots, \partial_n)(G_r(L)) \\ &= \mathcal{D}(\mathrm{Adj}(S_{n-p} \cdots S_n)e_i; \partial_1, \ldots, \partial_n)(G_r(L)) \\ &= \mathcal{D}(\mathrm{Adj}(T_n)\,\mathrm{Adj}(S'(L))e_i; \partial_1, \ldots, \partial_n)(G_r(L)) \qquad \text{(by (3))} \\ &= (\partial_1 G_r(L), \ldots, \partial_n G_r(L))\,\mathrm{Adj}(T_n)\,\mathrm{Adj}(S'(L))e_i \end{aligned}$$

As before let $T_n = (a_{i,j})$ and let $\mathrm{Adj}(T_n) = (b_{i,j})$. The above is then equal to:

$$\begin{aligned} & \sum_{j=1}^{n}\sum_{u=1}^{n} \partial_j(G_r(L))b_{j,u}s_{u,i}(L) \\ = \; & \sum_{j=1}^{n}\sum_{u=1}^{n}\sum_{v=1}^{n}(\partial_j L_v)(\partial_v G_r)(L)b_{j,u}s_{u,i}(L) \qquad \text{(by chain rule)} \\ = \; & \sum_{j=1}^{n}\sum_{u=1}^{n}\sum_{v=1}^{n} a_{v,j}b_{j,u}(\partial_v G_r)(L)s_{u,i}(L) \qquad (L_v = a_{v,1}X_1 + \cdots + a_{v,n}X_n) \\ = \; & \sum_{u=1}^{n}\sum_{v=1}^{n} \delta_{v,u}\delta(\partial_v G_r)(L)s_{u,i}(L) \qquad (\delta = \det(T_n);\ \delta_{v,u} \text{ is Kronecker delta}) \\ = \; & \delta \sum_{u=1}^{n}(\partial_u G_r)(L)s_{u,i}(L) \\ = \; & \delta \sum_{u=1}^{n-1}(\partial_u G_r)(L)s_{u,i}(L) \qquad (\text{because } s_{n,i} = 0, \text{ since } i \le n-p-1 < n) \\ = \; & \delta D'(G_r)(L) \end{aligned}$$

where $D' = \mathcal{D}(\mathrm{Adj}(S'_{n-p} \cdots S'_{n-1})e_i; \partial_1, \ldots, \partial_{n-1})$. By induction on $n-r$, we know that $D'(G_r) = 0$, hence $D'(G_r)(L) = 0$, and $D(H_r) = 0$ as desired.

Hence $D(h) = 0$. □

The last step in the proof of proposition 3.3 is given by lemma 3.9. We have already added one new variable in order to get this $\exp(tD)$, hence if we can show that $\exp(tD)^{[p]}$ is tame, we have shown that $(X + E_{n,p}(T; f))^{[p+1]}$ is tame, the claim of proposition 3.3.

Lemma 3.9 *The map* $\exp(tD)^{[p]}$ *is tame.*

Proof. If we consider $A[t]$ to be the new base ring, we see that $\exp(tD) = X + E_{n,p}(T;q)$ where q is the $(n-p-1)$-tuple $(0,\ldots,0,t,0,\ldots,0)$. By using proposition 1.3 again we see that the middle part of $E_{n,p}(T;q)$ is given by the composition

$$\operatorname{Adj}(T_{n-p})(0,\ldots,0,t,0,\ldots,0,0,\ldots,0)_{|T_{n-p}X}$$

However since t is in the base ring, the substitution has no effect. Only the product remains and we get:

$$(t\operatorname{Adj}(T_{n-p})_{1,i},\ldots,t\operatorname{Adj}(T_{n-p})_{n-p,i},0,\ldots,0)$$

Now let g be the tuple of the first $n-p$ entries. If $p=0$ then g is an n-tuple over $A[t]$ and $X + E_{n,p}(T;q)$ is clearly tame and we have reduced the original $X + E_{n,p}(T;f)$ using one new variable. If $p > 0$ then we have $E_{n,p}(T;q) = E_{n,p-1}(T;g)$. And this expression can be factored using p new variables by induction. □

Theorem 3.10 *Let $F = X + H$ with $H \in \mathcal{H}_n(A)$. Then $F^{[n-1]}$ is tame.*

Proof. Proposition 3.3 shows that $(X + E_{n,p}(T;c))^{[p+1]}$ is tame. Theorem 3.1 and remark 3.2 show that it suffices to prove that $(X+E_{n,p}(T;c))^{[n-1]}$ is tame for $p \le n-2$. Because $p+1 \le n-1$ this is obviously the case. □

We show on our running example that this method really works:

Example 3.11 Take F as in example 1.5 and use the same $(T,c) \in \mathcal{D}_4(\mathbb{C})$:

$$F = X + (X_4(X_3X_1 + X_4X_2)^2, -X_3(X_3X_1 + X_4X_2)^2, X_4^3, 0)$$

The first step is splitting F into

$$F = (X + c_4) \circ (X + E_{4,0}(T,c)) \circ (X + E_{4,1}(T,c)) \circ (X + E_{4,2}(T,c))$$

Because this F has a simple structure the first three components of this composition are already tame. Hence we only have to look at $X + E_{4,2}(T,c)$. We have seen that

$$\begin{aligned} X + E_{4,2}(T,c) &= X + E_{4,2}(T;c_1) \\ &= X + E_{4,2}(T;(X_2^2,0,0,0)) \\ &= \exp(h_1D_1) \end{aligned}$$

where $h_1 = (X_3X_1 + X_4X_2)^2$ and $D_1 = X_4\partial_1 - X_3\partial_2$. Obviously $D_1({}_1h) = 0$. Now Smith tells us that

$$(\exp(h_1D_1), X_5) = \exp(X_5D_1)\rho_1\exp(-X_5D_1)\rho_1^{-1}$$

where $\rho_1 = (X_1, X_2, X_3, X_4, X_5 + (X_3X_1 + X_4X_2)^2)$ and $\exp(X_5D_1) = (X_1 + X_4X_5, X_2 - X_3X_5, X_3, X_4, X_5)$. In this example we see that both ρ_1 and $\exp(X_5D_1)$ are compositions of elementary maps and hence tame. This means that we can stop here and claim that F can be factored into tame automorphisms by adding only one variable.[2] However, for the sake of the argument, we continue with this algorithm to show that it really ends after $n-1=3$ steps. However we will not display all details. Note that

$$\begin{aligned} \exp(X_5D_1) &= X + E_{4,1}(T;\mathrm{Adj}(T_2)(X_5,0,0,0)) \\ &= \exp(h_2D_2)\circ\exp(h_3D_3) \end{aligned}$$

where $h_2 = X_4X_5$, $D_2 = \partial_1$, $h_3 = -X_3X_5$ and $D_3 = \partial_2$. Using Smith's result again we can write

$$\begin{aligned} (\exp(h_2D_2), X_6) &= \exp(X_6D_2)\rho_2\exp(-X_6D_2)\rho_2^{-1} \\ (\exp(h_3D_3), X_6) &= \exp(X_6D_3)\rho_3\exp(-X_6D_3)\rho_3^{-1} \end{aligned}$$

where $\rho_2 = (X_1,X_2,X_3,X_4,X_6+X_4X_5)$, $\exp(X_6D_2) = (X_1+X_6,X_2,X_3,X_4,X_6)$, $\rho_3 = (X_1,X_2,X_3,X_4,X_6-X_3X_5)$ and $\exp(X_6D_3) = (X_1,X_2+X_6,X_3,X_4,X_6)$.

The final step gives $\exp(X_6D_2) = \exp(h_4D_4)$ and $\exp(X_6D_3) = \exp(h_5D_5)$ where $h_4 = h_5 = X_6$, $D_4 = \partial_1$ and $D_5 = \partial_2$. Then

$$\begin{aligned} (\exp(h_4D_4), X_7) &= \exp(X_7D_4)\rho_4\exp(-X_7D_4)\rho_4^{-1} \\ (\exp(h_5D_5), X_7) &= \exp(X_7D_5)\rho_5\exp(-X_7D_5)\rho_5^{-1} \end{aligned}$$

where $\rho_4 = \rho_5 = (X_1,X_2,X_3,X_4,X_7+X_6)$, $\exp(X_7D_4) = (X_1+X_7,X_2,X_3,X_4,X_7)$ and $\exp(X_7D_5) = (X_1,X_2+X_7,X_3,X_4,X_7)$.
And now we have $\exp(X_7D_4), \exp(X_7D_5) \in \mathbb{C}[X_7]$ and hence the algorithm ends. Coercing to seven-dimensional mappings in the logical way gives:

$$\begin{aligned} \rho_1 &= (X_1,X_2,X_3,X_4,X_5+(X_3X_1+X_4X_2)^2,X_6,X_7) \\ \rho_2 &= (X_1,X_2,X_3,X_4,X_5,X_6+X_4X_5,X_7) \\ \rho_3 &= (X_1,X_2,X_3,X_4,X_5,X_6-X_3X_5,X_7) \\ \rho_4=\rho_5 &= (X_1,X_2,X_3,X_4,X_5,X_6,X_7+X_6) \\ \exp(X_5D_1) &= (X_1+X_4X_5,X_2-X_3X_5,X_3,X_4,X_5,X_6,X_7) \\ \exp(X_6D_2) &= (X_1+X_6,X_2,X_3,X_4,X_5,X_6,X_7) \\ \exp(X_6D_3) &= (X_1,X_2+X_6,X_3,X_4,X_5,X_6,X_7) \\ \exp(X_7D_4) &= (X_1+X_7,X_2,X_3,X_4,X_5,X_6,X_7) \\ \exp(X_7D_5) &= (X_1,X_2+X_7,X_3,X_4,X_5,X_6,X_7) \end{aligned}$$

And with these tame automorphisms we can write down the factorization:

$$\begin{aligned} &(X+E_{4,2}(T,c),X_5,X_6,X_7) \\ &= \exp(X_7D_4)\rho_4\exp(-X_7D_4)\rho_4^{-1}\rho_2\rho_4\exp(X_7D_4)\rho_4^{-1}\exp(-X_7D_4)\rho_2^{-1} \\ &\quad \exp(X_7D_5)\rho_5\exp(-X_7D_5)\rho_5^{-1}\rho_3\rho_5\exp(X_7D_5)\rho_5^{-1}\exp(-X_7D_5)\rho_3^{-1}\rho_1 \\ &\quad \rho_3\exp(X_7D_5)\rho_5\exp(-X_7D_5)\rho_5^{-1}\rho_3^{-1}\rho_5\exp(X_7D_5)\rho_5^{-1}\exp(-X_7D_5)\rho_2 \\ &\quad \exp(X_7D_4)\rho_4\exp(-X_7D_4)\rho_4^{-1}\rho_2^{-1}\rho_4\exp(X_7D_4)\rho_4^{-1}\exp(-X_7D_4)\rho_1^{-1} \end{aligned}$$

[2]This is a consequence of the fact that we can view the first two components of F as $(X_1,X_2)+(H_1,H_2)$ with $(H_1,H_2)\in\mathcal{H}_2(\mathbb{C}[X_3,X_4])$.

We see that $(X + E_{4,2}(T, c))^{[3]}$ is tame and hence $F^{[3]}$ is tame.

Acknowledgments

The first author thanks the mathematics department of Washington University, in St. Louis, for its hospitality during his visit, during which the quick method presented in section 2 was found. He also thanks Shell Travel Services for funding his trip.

References

[1] A.R.P. van den Essen and E.-M.G.M. Hubbers, *A new class of invertible polynomial maps*, J. of Algebra **187** (1997), 214–226.

[2] A.R.P. van den Essen and E.-M.G.M. Hubbers, *Chaotic Polynomial Automorphisms; counterexamples to several conjectures*, Advances in Applied Mathematics **18** (1997), 382–388.

[3] A.R.P. van den Essen and E.-M.G.M. Hubbers, $\mathcal{D}_n(A)$ *for polynomial automorphisms*, J. of Algebra **192** (1997), 460–475.

[4] M.K. Smith, *Stably tame automorphisms*, J. of Pure and Applied Algebra **58** (1989), 209–212.