MATHEMATICS OF COMPUTATION
Volume 83, Number 290, November 2014, Pages 3017–3037
S 0025-5718(2014)02809-3
Article electronically published on February 25, 2014

CONSTRAINTS ON COUNTEREXAMPLES TO THE CASAS-ALVERO CONJECTURE AND A VERIFICATION IN DEGREE 12

WOUTER CASTRYCK, ROBERT LATERVEER, AND MYRIAM OUNAÏES

Abstract. In the first (theoretical) part of this paper, we prove a number of constraints on hypothetical counterexamples to the Casas-Alvero conjecture, building on ideas of Graf von Bothmer, Labs, Schicho and van de Woestijne that were recently reinterpreted by Draisma and de Jong in terms of p-adic valuations. In the second (computational) part, we present ideas improving upon Diaz-Toca and Gonzalez-Vega's Gröbner basis approach to the Casas-Alvero conjecture. One application is an extension of the proof of Graf von Bothmer et al. to the cases $5p^k$, $6p^k$ and $7p^k$ (that is, for each of these cases, we determine the finite list of primes p to which their proof is not applicable). Finally, by combining both parts, we settle the Casas-Alvero conjecture in degree 12 (the smallest open case).

1. Introduction and overview

1.1. The subject of this article is the following intriguing conjecture [3]:

Conjecture 1 (The Casas-Alvero conjecture, 2001). *Let $f(x) \in \mathbb{C}[x]$ be of degree $d > 0$ and suppose that for each $j = 1, \ldots, d-1$ there exists an $a \in \mathbb{C}$ such that $f(a) = f^{(j)}(a) = 0$, where $f^{(j)}(x)$ denotes the jth derivative. Then $f(x)$ is the dth power of a linear polynomial.*

For each given degree d, proving Conjecture 1 (if true) boils down to a finite Gröbner basis computation. In 2006, this was used by Diaz-Toca and Gonzalez-Vega to verify the conjecture for $d \leq 7$ [6]. Shortly afterwards, Graf von Bothmer, Labs, Schicho and van de Woestijne [8] proved a theoretical result settling the cases $d = p^k$ and $d = 2p^k$ (where p is prime and $k \geq 0$ is an integer). The proof uses reduction-mod-p arguments in algebraic geometry. It was recently rewritten in the more elementary (and slightly more powerful) language of p-adic valuations, in a nice overview due to Draisma and de Jong [7].

1.2. Because of a lack of a general strategy, beyond the degree, we subdivide the set of hypothetical counterexamples $f(x)$ to the Casas-Alvero conjecture by

- their number of distinct roots #roots(f),
- their *type* type(f), which is the minimal number of "recycled" roots minus one; that is,

$$\min\left\{ \#S \;\middle|\; S \subset \mathbb{C} \text{ and } \forall j : \exists\, a \in S : f(a) = f^{(j)}(a) = 0 \right\} - 1,$$

Received by the editor August 27, 2012 and, in revised form, February 7, 2013 and February 12, 2013.

2010 *Mathematics Subject Classification.* Primary 12D10; Secondary 13Y05, 14Q99.

where j ranges over $\{1, \dots, d-1\}$;

- their *scenario* $\operatorname{scen}(f)$, which is

$$\min \left\{ (s_1, \dots, s_{d-1}) \in \mathbb{Z}_{\geq 0}^{d-1} \,\middle|\, \exists\, a_i\text{'s} \in \mathbb{C} : \forall j : f(a_{s_j}) = f^{(j)}(a_{s_j}) = 0 \right\}, \tag{1}$$

where the minimum is taken lexicographically and j ranges over $\{1, \dots, d-1\}$. Note that $\operatorname{type}(f)$ is the maximal entry of $\operatorname{scen}(f)$.

Example. Since it is conjecturally impossible to give examples over $\mathbb{C}$, consider $f(x) = x(x-1)^4(x-8)(x-18) \in \mathbb{F}_{23}[x]$. One checks that the common roots of f with $f^{(1)}, \dots, f^{(6)}$ are

$$\{1\}, \quad \{1, 18\}, \quad \{1\}, \quad \{0\}, \quad \{18\}, \quad \{1\},$$

respectively. So $\operatorname{type}(f) = 2$ and $\operatorname{scen}(f) = (0,0,0,1,2,0)$ (take $a_0 = 1, a_1 = 0, a_2 = 18$).

1.3. The scenario $(s_1, \dots, s_{d-1}) \in \mathbb{Z}_{\geq 0}^{d-1}$ of a degree d counterexample $f \in \mathbb{C}[x]$ to the Casas-Alvero conjecture always satisfies $s_1 = 0$ and $s_j \leq \max\{\, s_i \,|\, i < j \,\} + 1$ for all $j = 2, \dots, d-1$. A sequence of this form will therefore be called *a scenario for degree d*. In view of the above, the *type* of a scenario is defined to be its maximal entry—we denote it by $\operatorname{type}(s)$. The number of scenarios for a given degree d grows quickly with d. For example, in our main case of interest $d = 12$, we have

$$1, 1023, 28501, 145750, 246730, 179487, 63987, 11880, 1155, 55, 1$$

scenarios of type $0, \dots, 10$, respectively, amounting to a total of 678570.

1.4. Let $s = (s_1, \dots, s_{d-1})$ be a scenario for degree d, and let $t = \operatorname{type}(s)$. Let $f(x) \in \mathbb{C}[x]$ be a degree d counterexample to the Casas-Alvero conjecture. Then we say that $f(x)$ *matches with* s if there exist $a_0, \dots, a_t \in \mathbb{C}$ such that

- $f(x) = g(x) \cdot (x - a_0)(x - a_1) \cdots (x - a_t)$ for a degree $d - 1 - t$ polynomial $g(x) \in \mathbb{C}[x]$,
- $f(a_{s_j}) = f^{(j)}(a_{s_j}) = 0$ for all $j = 1, \dots, d-1$.

Clearly $f(x)$ matches with its own scenario $\operatorname{scen}(f)$, but it may also match with various other scenarios.

Example (continued). The polynomial $f(x) = x(x-1)^4(x-8)(x-18) \in \mathbb{F}_{23}[x]$ also matches with $(0,1,0,2,1,0)$ (and many more).

1.5. In Section 2, we prove a number of general constraints on these attributes. For instance, we find that

- $\#\operatorname{roots}(f) \geq 5$,
- $2 \leq \operatorname{type}(f) \leq d-3$ (the first inequality being due to Draisma and Knopper [7, Proposition 6]),
- if $\operatorname{type}(f) = d-3$, then no consecutive entries of $\operatorname{scen}(f)$ are equal.

The methods used here are classically flavoured (Gauss–Lucas, Newton, Rolle).

1.6. In Section 3, using the p-adic valuation approach, we prove additional constraints for certain special degrees. Our main results are on degrees of the form $p+1$:

Theorem 2. *Let p be prime and let $f(x)$ be a degree $d=p+1$ counterexample to the Casas-Alvero conjecture. Let c be the root of $f^{(d-1)}(x)$. Then $f^{(1)}(c)\neq 0$, and there exist at least two indices $2\leq j_1<j_2\leq d-2$ such that $f^{(j_1)}(c)=f^{(j_2)}(c)=0$. In particular,* $\operatorname{type}(f)\leq d-4$. *Moreover, if $j_1<\cdots<j_m$ are the indices between 2 and $d-2$ for which $f^{(d-j_1)}(c)=\cdots=f^{(d-j_m)}(c)=0$, then the determinant of*

$$\Delta_f=\begin{bmatrix} -1 & j_1 & 0 & 0 & \cdots & 0 \\ -1 & \binom{j_2-2}{j_1-2}j_2 & j_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & \binom{j_m-2}{j_1-2}j_m & \binom{j_m-2}{j_2-2}j_m & \binom{j_m-2}{j_3-2}j_m & \cdots & j_m \\ -1 & (-1)^{j_1} & (-1)^{j_2} & (-1)^{j_3} & \cdots & (-1)^{j_m} \end{bmatrix} \tag{2}$$

is a multiple of p.

Theorem 2 implies that every degree $d=p+1$ counterexample to the Casas-Alvero conjecture matches with an element of the strongly reduced list of scenarios $s=(s_1,\ldots,s_{d-1})$ for which

- $s_{d-1}\neq 0$,
- the set of indices $2\leq j\leq d-2$ for which $s_{d-j}=s_{d-1}$ satisfies the above determinant condition.

For $d=12$ ($p=11$), the list contains

$$0,48,1668,8172,11586,6298,1469,146,5,0,0 \tag{3}$$

scenarios of type $0,\ldots,10$, respectively, amounting to a total of 29392. In type 8, the five scenarios read

$$\begin{aligned} &(0,1,2,3,4,5,6,7,3,8,3),\\ &(0,1,2,3,4,5,5,6,7,8,5),\\ &(0,1,2,3,4,3,5,6,7,8,3),\\ &(0,1,2,3,4,2,5,6,7,8,2),\\ &(0,1,2,3,2,4,5,6,7,8,2); \end{aligned} \tag{4}$$

indeed, the only pairs (j_1,j_2) for which $\det\Delta_f\equiv 0 \bmod 11$ are $(3,8)$, $(5,6)$, $(6,8)$, $(6,9)$, $(7,9)$.

1.7. For the computational part of our paper, we turn back to the original reduction-mod-p setting used by Graf von Bothmer et al. Because of the interplay between characteristic 0 and characteristic $p>0$, the following general definition is convenient.

Definition 1. Let k be an algebraically closed field. We say that a degree d polynomial $f\in k[x]$ $(d>0)$ is a *Casas-Alvero polynomial* or *CA-polynomial* (over k) if f is not a power of a linear polynomial and if for each $j=1,\ldots,d-1$ there exists an $a\in k$ such that $f(a)=f_H^{(j)}(a)=0$.

Here, $f_H^{(j)}$ denotes the jth Hasse derivative (using Hasse derivatives makes the Casas-Alvero condition somewhat more restrictive; it makes no difference in characteristic 0 or $p>d-1$, where $f_H^{(j)}=\frac{1}{j!}f^{(j)}$). Then the main theorem of [8] reads:

Theorem 3 (Graf von Bothmer, Labs, Schicho, van de Woestijne)**.** *Let $d > 0$ be an integer and let p be a prime number. If no CA-polynomials of degree d exist over $\overline{\mathbb{F}}_p$, then the Casas-Alvero conjecture is true in degree dp^k for all integers $k \geq 0$.*

Since it is trivial that no CA-polynomials of degree 1 or 2 can exist (in any characteristic), the cases p^k and $2p^k$ follow. More generally, we call a prime p a *bad prime for degree d* if there exist CA-polynomials of degree d in characteristic p. Then it is easily verified that $p = 2$ is the sole bad prime for degree $d = 3$. De Jong and Draisma [7] proved that the bad primes for degree $d = 4$ are $p = 3, 5, 7$.

1.8. In Section 5 we present a Gröbner basis algorithm, the basic version of which takes as input an integer $d > 0$ and a prime number p (or $p = 0$), and outputs whether or not CA-polynomials of degree d exist in characteristic p. The basic idea is to classify all CA-polynomials by their scenario (the definitions in 1.2 straightforwardly generalize to arbitrary k—this was already used in the example after 1.2). We will see that scenarios of moderately low type t can be ruled out easily (if the Casas-Alvero conjecture is true). In characteristic 0, the computation seems feasible up to $d \cdot t \approx 50$, say. In small characteristic p, this can be pushed to about twice that value.

1.9. By running the algorithm in characteristic 0 and analyzing the prime factors appearing in certain resulting Nullstellensatz expansions, we can find the bad primes for d up to 7.

Theorem 4. *There are*

- 9 *bad primes for degree $d = 5$, namely,*

$$p = 2, 3, 7, 11, 131, 193, 599, 3541, \textit{ and } 8009;$$

- 53 *bad primes for degree $d = 6$, namely, the primes listed in Table* 1*,*
- 366 *bad primes for degree $d = 7$, namely, the primes listed in the file* `badprimes7.txt` *that accompanies this paper—the smallest non-bad prime (apart from $p = 7$) is* 127*—the largest bad prime is*

$$\begin{gathered}24984712021698392647916525667237483011737174983678606896870094983849\\90961418068252878569331239547247984884225516598909122297267921020 63\end{gathered}$$

(a 135*-digit number).*

It should be noted that the possibility of enumerating bad primes using Gröbner bases is already reported upon by Graf von Bothmer et al. [8], and that the bad primes for $d = 5$, resp., $d \in \{5, 6\}$ have been independently computed by Chellali and Salinier [4] and de Frutos [5], respectively.

1.10. Finally, in Section 6, we combine our theoretical and computational approaches. A naive run of our algorithm in degree 12 lies completely out of reach. But in view of Theorem 2 and certain reduction-mod-p considerations, it suffices to restrict to a limited list of scenarios and to run the algorithm in characteristic p. In doing so, the computation becomes feasible:

Theorem 5. *Conjecture* 1 *is true for $d = 12$.*

The margin is tight: Each of the five scenarios of (4) took approximately three weeks of computation and required about 90 GB of RAM. Pushing the analogous computation to $d = 20$, the next open case, lies out of reach (see 6.5).

TABLE 1. Bad primes for degree 6 (53 primes)

2	5	7	11
13	19	23	29
37	47	61	67
73	97	257	811
983	1069	1087	1187
1487	1499	1901	2287
3209	3877	3881	4019
4943	5471	6983	8699
9337	15131	15823	20771
21379	23993	150203	266587
547061	685177	885061	1030951
7783207	17250187	40362599	9348983563
70016757407	2610767527031	225833117528659	7390044713023799
51313000813080529			

1.11. The main computations have been carried out using Magma [2] version 2.18-2 on a computer called `matrix`, running Ubuntu 11.10 on a 6-core Intel Xeon 2.53 GHz processor with 96 GB RAM. Some additional calculations were executed using Magma version 2.15-12 on `kasparov`, running Debian GNU/Linux 6.0.4 on an 8-core x86-64 2.93 GHz processor with 64 GB RAM.[*]

1.12. We would like to thank Filip Cools, Jan Schepers, Fréderik Vercauteren and an anonymous referee for some helpful discussions and/or comments. We are also grateful to the Department of Electrical Engineering (KU Leuven), for allowing us to use `kasparov`.

2. GENERAL CONSTRAINTS ON COUNTEREXAMPLES

2.1. The following fact is easy to check and will be used throughout:

Lemma 6. *Let f be a CA-polynomial over k of degree $d > 0$, $\alpha_1, \alpha_2 \in k^*$ and $\beta \in k$. Then the polynomial $g(x) = \alpha_1 f(\alpha_2 x + \beta)$ is also CA.*

The polynomials f and g will be called *equivalent*. Note that the number of distinct roots, the type, the scenario, the matching or not with a given scenario, ... are all preserved by equivalence.

Another frequently used fact is:

Proposition 7. *Let $f \in \mathbb{C}[x]$ be a polynomial of degree d all of whose roots are on a line (when plotted in the complex plane). Then for all $j = 0, \ldots, d-2$:*

$$f^{(j)}(c) \neq 0, f^{(j+1)}(c) = 0 \Longrightarrow f^{(j+2)}(c) \neq 0.$$

[*] Four files accompany this paper: `CAbadprimes.m`, `CAbadprimes7test.m`, `badprimes7.txt` and `CAdeg12.m`. These can be downloaded at `https://perswww.kuleuven.be/~u0040935`.

Proof. Using a transformation of the above kind we may assume that f and its derivatives are real polynomials having real roots only. For $c \in \mathbb{R}$ and $j \in \{0, \dots, d-2\}$, denote by $m_j(c)$ the multiplicity of c as a zero of $f^{(j)}$, and by N_j the number of distinct roots of $f^{(j)}$. We have the relations

$$\sum_{c:f^{(j)}(c)=0} m_{j+1}(c) = \sum_{c:f^{(j)}(c)=0} (m_j(c) - 1) = d - j - N_j$$

$$\sum_{c:f^{(j)}(c)\neq 0} m_{j+1}(c) = d - j - 1 - \sum_{c:f^{(j)}(c)=0} m_{j+1}(c) = N_j - 1. \tag{5}$$

By Rolle's theorem, there is a zero of $f^{(j+1)}$ strictly between any pair of zeros of $f^{(j)}$; thus,

$$\#\{\, c \in \mathbb{R} \mid f^{(j)}(c) \neq 0, f^{(j+1)}(c) = 0 \,\} \geq N_j - 1.$$

From (5) we conclude that $m_{j+1}(c) = 1$ whenever $f^{(j)}(c) \neq 0, f^{(j+1)}(c) = 0$. □

Finally, thanks to the results of Graf von Bothmer et al. [8], we may assume that the degree of a counterexample to the Casas-Alvero conjecture is at least 12. This will also be used in some of the proofs below.

2.2. We begin with some considerations on the type:

Proposition 8. *Let $f \in \mathbb{C}[x]$ be a CA-polynomial of degree d and let Γ be the convex hull of the roots of f (when plotted in the complex plane). Let $m \geq 2$ be the maximum of the multiplicities of these roots, and let $\delta = 1$ if this maximum is attained by a non-vertex of Γ (let $\delta = 0$ otherwise). Let $\gamma \geq 2$ be the number of vertices of Γ. Then $2 \leq \operatorname{type}(f) \leq d + 1 - \gamma - m - \delta \leq d - 3$.*

Proof. For each vertex v of Γ we have:

- $f^{(j)}(v) \neq 0$ for all $j = 1, \dots, d-1$, or
- v has multiplicity at least 2

(by the Gauss–Lucas theorem). This means that among the d roots of f, counting multiplicities, at least γ of them are not needed to find a common root for each derivative. If $\delta = 1$, some non-vertex has multiplicity m, so $m - 1$ additional roots are superfluous. Therefore, at most $d - \gamma - (m-1)$ roots are needed. If $\delta = 0$, then the bound reads $d - (\gamma - 1) - (m - 1)$. In both cases, the upper bound for $\operatorname{type}(f)$ follows. The lower bound follows from an observation by Draisma and Knopper [7, Proposition 6]. □

Remark. Proposition 8 remains valid when we replace γ by the number of roots on the boundary of Γ (given that we adapt δ accordingly).

Refining to the level of scenarios, we find:

Proposition 9. *Let $d > 2$ be an integer and let $s = (s_1, s_2, \dots, s_{d-1})$ be a scenario for degree d. If*

1. *$\operatorname{type}(s) \in \{0, 1, d-2\}$, or*
2. *$\operatorname{type}(s) \leq d - 3$, the first $d - 2 - \operatorname{type}(s)$ entries of s are zero, and among $s_{d-1-\operatorname{type}(s)}, \dots, s_{d-1}$ there is a zero or two consecutive entries that are equal,*

then there are no CA-polynomials $f \in \mathbb{C}[x]$ for which $\operatorname{scen}(f) = s$.

Proof. The first part is an immediate corollary to Proposition 8. As for the second statement, suppose on the contrary that f is a CA-polynomial for which $\operatorname{scen}(f) = s$, with $t = \operatorname{type}(s) \leq d-3$, and that the first $d-2-t$ entries of s are equal to zero. Let $a_0, \ldots, a_t \in \mathbb{C}$ be as in (1). Then a_0 is a root with multiplicity at least $d-1-t$. Let Γ be the convex hull of the roots of f and let γ be its number of vertices. Using Proposition 8, we conclude that $\gamma = 2$ and that a_0 is a vertex. Then if another 0 were to appear in $s = \operatorname{scen}(f)$, by Gauss–Lucas we would conclude that the multiplicity of a_0 is strictly bigger than $d-1-t$, which would contradict Proposition 8. On the other hand, if two consecutive entries would be equal, some high-order derivative of $f(x)$ would have a double root. But since $\gamma = 2$, $f(x)$ is equivalent to a real-root polynomial, so Rolle's theorem (see Proposition 7) would imply that this double root is actually a root of $f(x)$ with multiplicity strictly bigger than $d-t$, again contradicting Proposition 8. □

Remark. Let s be as in the statement of Proposition 9. Then one cannot merely conclude (without using new arguments, that is) the stronger statement that there are no CA-polynomials $f \in \mathbb{C}[x]$ that *match* with s.

2.3. As immediate corollaries to the lower bound $2 \leq \operatorname{type}(f)$, we get the following three easy facts: if f is a CA-polynomial (over $\mathbb{C}$) of degree d, then

(1) $f^{(2)}(x)$ cannot be the $(d-2)$th power of a linear polynomial,
(2) f cannot have a root of multiplicity at least $d-1$,
(3) f has at least three distinct roots

(note that these statements can be proved in various other ways, see e.g. [13, Proposition 2.2]). In the next two propositions, we will go a step further in directions (1) and (2). Later on (Proposition 13 and Theorem 14), we will go two steps further in direction (3).

Proposition 10. *If $f \in \mathbb{C}[x]$ is a CA-polynomial of degree d, then $f^{(3)}(x)$ cannot be the $(d-3)$th power of a linear polynomial.*

Proof. Suppose on the contrary that $f^{(3)}(x)$ is the $(d-3)$th power of a linear polynomial. Thanks to Lemma 6, we may assume $f^{(3)}(x) = \frac{d!}{(d-3)!}x^{d-3}$. Assume that $f^{(1)}(0) \neq 0$, so that f has a root of multiplicity at least 2 which is different from 0. Then again by Lemma 6, we may assume $f(1) = f^{(1)}(1) = 0$. Thus

$$f(x) = x^d - (d-1)x^2 + (d-2)x; \quad f^{(2)}(x) = (d-1)\left(dx^{d-2} - 2\right).$$

Solving $f(x) = f^{(2)}(x) = 0$, we get $x = \frac{d}{d+1}$: a contradiction because rational roots of monic integer polynomials are necessarily integral. We conclude that $f^{(1)}(0) = 0$. Then, for some constant c, $f^{(2)}(x) = d(d-1)x^{d-2} + 2c$ and $f(x) = x^d + cx^2$. Solving $f(x) = f^{(2)}(x) = 0$, we get that $c = 0$. □

Proposition 11. *Let $f \in \mathbb{C}[x]$ be a CA-polynomial of degree d, then f cannot have a root of multiplicity at least $d-2$.*

Proof. Suppose that 0 is such a root. If $f^{(d-1)}(0) \neq 0$, then we may assume that $f(1) = f^{(d-1)}(1) = 0$ and

$$f(x) = x^{d-2}(x^2 - dx + d - 1), \quad f^{(d-2)}(x) = \frac{(d-1)!}{2}(dx^2 - 2dx + 2).$$

Solving $f(x) = f^{(d-2)}(x) = 0$, we get $x = \frac{d+1}{d}$; again a contradiction. We conclude that we necessarily have $f^{(d-1)}(0) = 0$. Then, for some constant c, $f(x) = x^d + cx^{d-2}$ and $f^{(d-2)}(x) = \frac{d!}{2}x^2 + c$. Solving $f(x) = f^{(d-2)}(x) = 0$, we get $c = 0$. □

We have chosen to present an elementary proof of Proposition 11, though we also can see it as a direct consequence of the forthcoming Proposition 13.

2.4. Let us recall some basic properties of the elementary symmetric polynomials. Let a polynomial f and its derivatives be of the form

$$f^{(j)}(x) = \frac{d!}{(d-j)!}\left(x^{d-j} + \binom{d-j}{1}a_1x^{d-j-1} + \binom{d-j}{2}a_2x^{d-j-2} + \cdots + a_{d-j}\right)$$

(here by convention $f = f^{(0)}$). Let $\sigma_m(j)$ be the sum of the mth powers of the roots of $f^{(j)}$, for $j = 0, \ldots, d-1$. Then Newton's formulas applied to each $f^{(j)}$ give the following relations (see for example [11] for more details on Newton formulas):

Lemma 12.

$$\sum_{k=1}^{r} \sigma_k(j)\binom{d-j}{r-k}a_{r-k} = -r\binom{d-j}{r}a_r$$

for $0 \leq j \leq d-1$, $1 \leq r \leq d-j$. *(It is understood that* $a_0 = 1$.*)*

In particular, for $r = 1$, we have that

$$\frac{\sigma_1(j)}{d-j} = \frac{\sigma_1(0)}{d}$$

for $j = 0, \ldots, d-1$, which means that the center of mass of the roots of the derivatives is fixed. As obviously

$$\sigma_1(d-1) = \frac{\sigma_1(0)}{d} = -a_1$$

is the only root of $f^{(d-1)}$, we see that whenever f is a CA-polynomial over $\mathbb{C}$, the center of mass of its roots $\frac{\sigma_1(0)}{d}$ is itself a root of f. As a direct consequence, the number of distinct roots of a CA-polynomial cannot be two. Actually, we can say more: If f has more than two distinct roots, then at least one of them (the center of mass) has to be in the interior of the convex hull of the roots. This fact also follows immediately from the Gauss–Lucas theorem, and can be pushed further:

Proposition 13. *Let* $f \in \mathbb{C}[x]$ *be a CA-polynomial. Then* f *has at least two distinct roots in the interior of the convex hull of the roots, when plotted in the complex plane. In particular,* f *has at least four distinct roots.*

Proof. Assume that f has exactly one root, say 0, in the interior. Let ζ be among the roots of f located on the boundary with maximal multiplicity m. Then by Gauss–Lucas, $f^{(m)}(0) = f^{(m+1)}(0) = \cdots = f^{(d-1)}(0) = 0$ which means that for $j = m, \ldots, d-1$:

$$f^{(j)}(x) = \frac{d!}{(d-j)!}x^{d-j}.$$

Taylor expansion at $x = \zeta$ gives

$$f(0) = \sum_{j=m}^{d} \frac{f^{(j)}(\zeta)}{j!}(-\zeta)^j = \zeta^d \sum_{j=m}^{d}(-1)^j\binom{d}{j} = \zeta^d(-1)^m\binom{d-1}{m-1}.$$

As $f(0) = 0$, we get $\zeta = 0$, which is a contradiction. □

Note that Proposition 13 can also be deduced directly from $2 \leq \text{type}(f)$.

2.5. We now prove the main result of this section:

Theorem 14. *Let f be a CA-polynomial over $\mathbb{C}$, then f has at least five distinct roots.*

Proof. Assume that f has four distinct roots. Then by the previous proposition, it has at least two distinct roots in the interior of its Gauss–Lucas hull. This implies that the four roots are on a line. By Lemma 6, we may assume that this is the real line. We denote by m the maximal multiplicity of the roots of f. By Proposition 11, we know that $2 \leq m \leq d-3$.

- First case: $m \leq d-5$. Again, using Lemma 6, we may assume without loss of generality that the roots of f are as follows: $a < 0 < 1 < b$ and $f^{(d-1)}(0) = 0$. Then a and b cannot be zeros of $f^{(j)}$ for $d-5 \leq j \leq d-1$. Moreover, by Rolle's theorem (see Proposition 7), each zero of $f^{(j)}$ is simple. Then we necessarily have $f^{(d-2)}(1) = 0$, $f^{(d-3)}(0) = 0$, $f^{(d-4)}(1) = 0$, $f^{(d-5)}(0) = 0$. Integrating five times the expression $f^{(d-1)}(x) = d!x$ and taking into account these constraints, we get $f^{(d-5)}(x) = \frac{d!}{5!}x(x^2-5)^2$. But this contradicts the fact that the roots are simple.
- Second case: $m = d-4$. In view of Lemma 6, we arrange the roots as $a < 0 < b < 1$ and we assume that $f^{(d-1)}(0) = 0$. Denote by m_a, m_0, m_b, m_1 their respective multiplicities. Then again we must have $f^{(d-2)}(b) = 0$, $f^{(d-3)}(0) = 0$, $f^{(d-4)}(b) = 0$. As in the first case, computing the last derivatives, we get

$$f^{(d-1)}(x) = d!x, \quad 2!f^{(d-2)}(x) = d!(x^2-b^2),$$
$$3!f^{(d-3)}(x) = d!x(x^2-3b^2), \quad 4!f^{(d-4)}(x) = d!(x^2-5b^2)(x^2-b^2).$$

Obviously, as $f^{(d-4)}(b) = 0$, we have $m_b \leq d-5$. From the Gauss–Lucas theorem, we deduce that $a < -\sqrt{5}b$. Now we apply Lemma 12 with $j=0$, $r=1$ and with $j=0$, $r=3$ to obtain

$$m_a a + m_b b + m_1 = m_a a^3 + m_b b^3 + m_1 = 0. \tag{6}$$

We deduce that $m_a a(a^2-1) = -m_b b(b^2-1)$ and, looking at the sign, we see that $-a < 1$. Then $m_a > -am_a = m_b b + m_1 > m_1$, which implies that $m_a \geq 2$ and $m_1 \leq d-5$. In the case where $m_a = 2, m_1 = m_b = 1$, (6) gives $a(a+1)^2 = 0$. Thus this case cannot occur. We can readily deduce that $m_0 \leq d-5$. The only possibility left is $m_a = m = d-4$.

From the relation $-(d-4)a(1-a^2) = m_b b(1-b^2)$, we deduce that $\phi(-a) \leq \phi(b)$ where we put $\phi(t) = t(1-t^2)$. But ϕ is increasing on $[0, 1/\sqrt{3}]$ and we know that $-a > b > 0$. Thus we have $-a > 1/\sqrt{3}$. Along with the linear equation in (6) this implies:

$$d-4 = m_b\frac{b}{-a} + m_1\frac{1}{-a} < \frac{m_b}{\sqrt{5}} + m_1\sqrt{3} < 4.$$

Since the Casas-Alvero conjecture is true for $d \leq 7$, this is a contradiction.
- Third case: $m = d-3$. We proceed as in the previous case. We have

$$f^{(d-1)}(x) = d!x, \quad 2!f^{(d-2)}(x) = d!(x^2-b^2), 3!f^{(d-3)}(x) = d!x(x^2-3b^2).$$

From Gauss–Lucas we deduce that $a < -\sqrt{3}b$. Again, we obtain that $m_a \geq 2$. Thus we necessarily have: $m_a = m$, $m_0 = m_1 = m_b = 1$. The linear equation in (6) gives

$$d - 3 = \frac{b}{-a} + \frac{1}{-a} < \frac{1}{\sqrt{3}} + \sqrt{3} < 3,$$

again a contradiction. □

3. Additional constraints for special degrees

3.1. We now turn our attention to certain special instances of d, in each case involving a prime number p. Inspired by Draisma and de Jong's approach [7], we use p-adic valuations. Most of the proofs below have straightforward analogs in the original reduction-mod-p setting of Graf von Bothmer et al. But at some points, the valuation language does seem slightly more powerful. Our starting point is the existence of a map

$$v_p\,:\,\mathbb{C} \to \mathbb{Q} \cup \{+\infty\}$$

satisfying

- $v_p(a) = +\infty$ if and only if $a = 0$,
- $v_p(ab) = v_p(a) + v_p(b)$ for all $a, b \in \mathbb{C}$,
- $v_p(a + b) \geq \min\{v_p(a), v_p(b)\}$ for all $a, b \in \mathbb{C}$,

and extending the usual p-adic valuation on $\mathbb{Z}$ (i.e. if $n = p^r \cdot n'$ with n' prime to p, then $v_p(n) = r$). See e.g. [12, Chapter 4, Theorem 1]. Note that the last property implies $v_p(a + b) = \min\{v_p(a), v_p(b)\}$ if $v_p(a) \neq v_p(b)$: we will make a frequent use of this fact.

3.2. The p-adic valuations of binomial coefficients are well-understood. A formula due to Legendre [10] states that for any $n \in \mathbb{Z}_{>0}$ and any $j \in \{0, \ldots, n\}$ one has

$$v_p\binom{n}{j} = \frac{s_p(j) + s_p(n-j) - s_p(n)}{p-1},$$

where $s_p(\cdot)$ denotes the sum of the p-adic digits. Note that $s_p(j)+s_p(n-j)-s_p(n)$ is a measure for the number of carries when adding $n-j$ to j in base p. In particular,

$$v_p\binom{n}{j} = 0 \quad \text{iff} \quad \text{there are no carries.}$$

It follows that:

Lemma 15. *Let $n \in \mathbb{Z}_{>0}$ and $k \in \mathbb{Z}_{\geq 0}$. If $j \in \{0, 1, 2, \ldots, np^k\}$ is not a multiple of p^k, then*

$$v_p\binom{np^k}{j} > 0.$$

If $n = p^r + 1$ for some $r \in \mathbb{Z}_{\geq 0}$, then the same conclusion holds under the weaker assumption that $j \notin \{0, p^k, (n-1)p^k, np^k\}$.

Proof. According to Legendre's formula

$$v_p\binom{np^k}{j} = \frac{s_p(j) + s_p(np^k - j) - s_p(np^k)}{p-1}.$$

Let q and $\rho \neq 0$ be the quotient and remainder of j when divided by p^k. Then $s_p(np^k) = s_p(n)$, $s_p(j) = s_p(q) + s_p(\rho)$, and

$$s_p(np^k - j) = s_p((n-q-1)p^k + (p^k - \rho)) \geq s_p(n-q) - 1 + 1,$$

from which

$$v_p\binom{np^k}{j} \geq v_p\binom{n}{q} + \frac{s_p(\rho)}{p-1} > 0.$$

A similar argument proves the second statement. □

3.3. We use this to prove:

Proposition 16. *Let $n \in \mathbb{Z}_{>0}$ and $k \in \mathbb{Z}_{\geq 0}$ be integers, and let $f \in \mathbb{C}[x]$ be a CA-polynomial of degree $d = np^k$. Then*

$$f, f^{(p^k)}, f^{(2p^k)}, \ldots, f^{(d-p^k)}$$

do not share a common root. If $n = p^r + 1$ for some integer $r \geq 0$, one even has that

$$f, f^{(p^k)}, f^{(d-p^k)}$$

do not share a common root. As a consequence, if $s = (s_1, \ldots, s_{d-1})$ is a scenario for degree d and $s_{p^k} = s_{2p^k} = \cdots = s_{d-p^k}$ (resp. $s_{p^k} = s_{d-p^k}$), then there are no CA-polynomials that match with s.

Proof. We only prove the first statement (the second assertion can be proved in a similar way). Suppose on the contrary that f is a CA-polynomial such that $f, f^{(p^k)}, \ldots, f^{(d-p^k)}$ do have a common root. We may assume without loss of generality, using Lemma 6, that f is of the form

$$f(x) = x^d + \binom{d}{1}a_1x^{d-1} + \binom{d}{2}a_2x^{d-2} + \cdots + \binom{d}{d-1}a_{d-1}x, \tag{7}$$

that the assumed common root of $f, f^{(p^k)}, \ldots, f^{(d-p^k)}$ is 0, and that

$$\min\{v_p(x_i) \mid i = 1, \ldots, d\} = 0,$$

where we have denoted by $x_1, x_2, \ldots, x_d$ the zeros of f.

For $j = 1, \ldots, d-1$, we have:

$$\frac{j!}{d!}f^{(d-j)}(x) = x^j + \binom{j}{1}a_1x^{j-1} + \binom{j}{2}a_2x^{j-2} + \cdots + \binom{j}{j-1}a_{j-1}x + a_j. \tag{8}$$

Using (8) with $j = 1, \ldots, d-1$, each time plugging in a common root of $f^{(d-j)}$ and f (taking 0 if j is a multiple of p^k), one proves by induction on j that

$$\begin{cases} v_p(a_j) \geq 0 & \text{for all } \; j = 1, \ldots, d-1, \\ a_j = 0 & \text{as soon as } \; p^k \mid j. \end{cases} \tag{9}$$

Let x_j be such that $v_p(x_j) = 0$. Then taking valuations of both sides of the equality

$$x_j^d = -\binom{d}{1}a_1x_j^{d-1} - \binom{d}{2}a_2x_j^{d-2} - \cdots - \binom{d}{d-2}a_{d-2}x_j^2 - \binom{d}{d-1}a_{d-1}x_j$$

yields a contradiction with (9) and Lemma 15. □

Note that if $n = 1$ or $n = 2$ the conclusion that f, resp., f and $f^{(p^k)}$ do not share a common root is trivially impossible. Hence the cases p^k and $2p^k$ tautologically follow from Proposition 16. If $d = p^r + 1$, the proposition implies that the root of $f^{(d-1)}(x)$ must be a simple root of $f(x)$. If $p \geq 3$, this in turn can be seen as a limiting case of the following statement:

Proposition 17. *If $d = p^r + 1$, then the root of $f^{(d-1)}(x)$ cannot be the mean of two distinct roots of $f(x)$.*

Proof. Using Lemma 6 we can assume that $f(x)$ is of the form (7) with $a_1 = 0$ (i.e. the root of $f^{(d-1)}(x)$ is 0), and that again all roots $x_1, \ldots, x_d$ have non-negative valuation, with minimum 0. Let x_j be such that $v_p(x_j) = 0$. Then the equality

$$da_{d-1}x_j = -x_j^d - \binom{d}{2}a_2x_j^{d-2} - \cdots - \binom{d}{d-2}a_{d-2}x_j^2$$

implies that $v_p(a_{d-1}) = 0$. Let $w \in \mathbb{C}^*$ be such that $f(w) = f(-w) = 0$. Then $0 = f(w) - f(-w)$ gives

$$da_{d-1}w = -\binom{d}{3}a_3w^{d-3} - \binom{d}{5}a_5w^{d-5} - \cdots - \binom{d}{d-3}a_{d-3}w^3.$$

Taking valuations yields a contradiction. □

The same argument can be used to show that the root of $f^{(d-1)}(x)$ cannot be the mean of two distinct roots of $f^{(1)}(x)$.

3.4. From now on, we focus on the special case $d = p+1$. Using once again Lemma 6, we may assume that

$$(10) \qquad \begin{cases} f(x) = x^d + da_1x^{d-1} + \binom{d}{2}a_2x^{d-2} + \cdots + \binom{d}{d-2}a_{d-2}x^2, \\ \min\{v_p(x_j) \,|\, j = 1, \ldots, d\} = 0, \end{cases}$$

where we have denoted by $x_1, \ldots, x_{d-3}, x_{d-2} = x_{d-1} = 0, x_d = -a_1$ the roots of f. For $j = 1, \ldots, d-2$, we then again have that expression (8) holds. Observe that $v_p(a_1) \geq 0$ because $-a_1$ is one of the roots of f. As before, using equality (8) with $j = 2, \ldots, d-2$, each time plugging in a common root of $f^{(d-j)}$ and f, we prove by induction on j that

$$(11) \qquad v_p(a_j) \geq 0 \quad \text{for all } \ j = 1, \ldots, d-2.$$

Let x_j be such that $v_p(x_j) = 0$. The equality

$$-da_1x_j^{d-1} = x_j^d + \binom{d}{2}a_2x_j^{d-2} + \cdots + \binom{d}{d-2}a_{d-2}x_j^2$$

shows that $v_p(a_1) = 0$. Therefore, we may assume without loss of generality that $a_1 = -1$. Then we can write $f(x) = (x-1)g(x)$ where

$$\begin{aligned} g(x) = x^{d-1} - (d-1)x^{d-2} &+ \left(\binom{d}{2}a_2 - (d-1)\right)x^{d-3} \\ &+ \left(\binom{d}{3}a_3 + \binom{d}{2}a_2 - (d-1)\right)x^{d-4} \\ &+ \cdots + \left(\binom{d}{d-3}a_{d-3} + \cdots + \binom{d}{2}a_2 - (d-1)\right)x^2. \end{aligned}$$

In view of (11) and Lemma 15, all roots of g have strictly positive valuations (actually greater than $1/(d-3)$). As a consequence, we see that 1 is a simple root of f (a fact already implied by Proposition 16) and that $v_p(x_j) > 0$ for $j = 1, \ldots, d-3$. Now whenever $f^{(d-j)}(1) \neq 0$, the Casas-Alvero property implies that $f^{(d-j)}(x_j) = 0$ with $v_p(x_j) > 0$ and from equality (8) we get $v_p(a_j) > 0$. But as

$$f(1) = 1 - d + \binom{d}{2} a_2 + \cdots + \binom{d}{d-2} a_{d-2} = 0,$$

there is at least one index $2 \leq j \leq d-2$ such that $v_p(a_j) = 0$. In other words, at least one of the derivatives $f^{(d-j)}(1) = 0$. If we put this together with Proposition 11 and the observations following Lemma 12, we get:

Lemma 18. *Let f be a CA-polynomial over $\mathbb{C}$ of degree $d = p+1$, where p is prime. Let c be the center of mass of the roots of f. Then the following conditions are satisfied:*

- $f^{(1)}(c) \neq 0, f^{(d-1)}(c) = 0,$
- $f^{(j)}(c) \neq 0$ *for at least one* $j \in \{2, \ldots, d-2\}$,
- $f^{(j)}(c) = 0$ *for at least one* $j \in \{2, \ldots, d-2\}$.

3.5. Let us now go further into the investigation of the orders of the derivatives having the center of mass as a root, in order to prove Theorem 2. We may again assume that f is of the form (10) and that $a_1 = -1$. We will use the notation $x \equiv y$ if $v_p(x-y) > 0$. In view of Lemma 18, let $j_1 < j_2 < \cdots < j_m$ be the indices between 2 and $d-2$ such that $f^{(d-j_i)}(1) = 0$ for $i = 1, \ldots, m$. As observed previously, for all $j \in \{2, \cdots, d-2\}$, we have $v_p(a_j) \geq 0$. If, moreover, $j \notin \{j_1, \cdots, j_m\}$, then $a_j \equiv 0$. From (8) with $x = 1$ and $j = j_1, j_2, \ldots, j_m$, we get

$$(12) \qquad \begin{cases} 1 - j_1 + a_{j_1} \equiv 0, \\ 1 - j_2 + \binom{j_2}{j_1} a_{j_1} + a_{j_2} \equiv 0, \\ \vdots \\ 1 - j_m + \binom{j_m}{j_1} a_{j_1} + \binom{j_m}{j_2} a_{j_2} + \cdots + a_{j_m} \equiv 0. \end{cases}$$

Using the equation $\frac{f(1)}{p} = 0$ and the inequalities $v_p\binom{d}{j} \geq 1$ for $j = 2, \ldots, d-2$, we obtain

$$(13) \qquad -1 + \frac{\binom{d}{j_1}}{p} a_{j_1} + \cdots + \frac{\binom{d}{j_m}}{p} a_{j_m} \equiv 0.$$

Observe that for all $2 \leq j \leq d-2$ we have:

$$\begin{aligned} \frac{\binom{d}{j}}{p} &= \frac{d(d-2)(d-3)\cdots(d-(j-1))}{j!} \\ &= \frac{(p+1)(p-1)(p-2)\cdots(p-(j-2))}{j!} \\ &= \frac{1}{j!}(p^{j-1} + \alpha_{j-2}p^{j-2} + \cdots + \alpha_1 p) + \frac{(-1)^{j-2}(j-2)!}{j!}, \end{aligned}$$

where $\alpha_1, \ldots, \alpha_{j-2}$ are integers. Therefore,

$$\frac{\binom{d}{j}}{p} \equiv \frac{(-1)^j}{j(j-1)}.$$

Putting equations (12) and (13) together and defining $\tilde{a}_{j_i} = \frac{a_{j_i}}{j_i(j_i-1)}$, we obtain:

$$\begin{cases} -1 + j_1\tilde{a}_{j_1} \equiv 0, \\ -1 + \binom{j_2-2}{j_1-2} j_2\tilde{a}_{j_1} + j_2\tilde{a}_{j_2} \equiv 0, \\ \quad \vdots \\ -1 + \binom{j_m-2}{j_1-2} j_m\tilde{a}_{j_1} + \binom{j_m-2}{j_2-2} j_m\tilde{a}_{j_2} + \cdots + j_m\tilde{a}_{j_m} \equiv 0, \\ -1 + (-1)^{j_1}\tilde{a}_{j_1} + (-1)^{j_2}\tilde{a}_{j_2} + \cdots + (-1)^{j_m}\tilde{a}_{_m} \equiv 0. \end{cases} \tag{14}$$

With Δ_f as in the statement of Theorem 2, we see that necessarily $\det \Delta_f \equiv 0$; otherwise inverting (14) we would get that $1 \equiv 0$. To conclude the proof of Theorem 2 we show:

Lemma 19. *Let $f \in \mathbb{C}[x]$ be a CA-polynomial of degree $d = p + 1$ and let c be the center of mass of its roots. Then there are at least two indices $2 \leq j_1 < j_2 \leq d - 2$ such that $f^{(j_1)}(c) = f^{(j_2)}(c) = 0$.*

Proof. If not, in virtue of Lemma 18 there exists a unique index $2 \leq j \leq d-2$ such that $f^{(d-j)}(c) = 0$. We can assume without loss of generality that f is of the form (10) with $a_1 = -1$ and construct Δ_f as above:

$$\det \Delta_f = \begin{vmatrix} -1 & j \\ -1 & (-1)^j \end{vmatrix} = j - (-1)^j. \tag{15}$$

Observe that $1 \leq j - (-1)^j \leq j + 1 \leq d - 2$ for $j \in 2, \ldots, d - 3$. Also, $d - 2 - (-1)^{d-2} = d - 3$ because d is even (indeed, $p \neq 2$ since the Casas-Alvero conjecture is true for degree 3). Thus there is no way for p to divide $\det \Delta_f$. □

3.6. Theorem 2 implies that every CA-polynomial of degree $d = p+1$ matches with a scenario $s = (s_1, \ldots, s_{d-1})$ for which $s_{d-1} \neq 0$ and the index set

$$\operatorname{ind}(s) = \{\, j \mid 2 \leq j \leq d-2 \text{ and } s_{d-j} = s_{d-1} \,\}$$

satisfies the corresponding determinant condition. We remark, however, that this does not necessarily imply that *the* scenario of a CA-polynomial satisfies these conditions. Indeed, imagine a CA-polynomial $f \in \mathbb{C}[x]$ of degree 12 for which

$$\operatorname{scen}(f) = s = (0, 1, 2, 3, 4, 2, 5, 6, 4, 7, 4),$$

i.e., there exist $a_1, \ldots, a_7 \in \mathbb{C}$ such that $f(a_{s_j}) = f^{(j)}(a_{s_j}) = 0$ for $j = 1, \ldots, d-1$. Then $\operatorname{ind}(s) = \{3, 7\}$ does not satisfy the determinant condition. However, it might a priori be that $f^{(6)}(x)$ has both a_2 and a_4 as a root. Then $f(x)$ also matches with the scenario $(0, 1, 2, 3, 4, 4, 5, 6, 4, 7, 4) \neq \operatorname{scen}(f)$. In this example, the index set $\{3, 6, 7\}$ satisfies the determinant condition.

3.7. We end our study of the degree $p + 1$ case with the following observation.

Proposition 20. *Let p be a prime number. Then there is no CA-polynomial of degree $d = p + 1$ all of whose roots are rational.*

Proof. Using the notation and the results found in the proof of Lemma 18, we may assume that f is of the form

$$\begin{aligned} f(x) =& x^d - dx^{d-1} + \binom{d}{2} x^{d-2} \\ &+ \cdots + (-1)^{k-1}\binom{d}{k-1} x^{d-k+1} + \binom{d}{k} a_k x^{d-k} + \cdots + \binom{d}{d-2} a_{d-2} x^2, \end{aligned}$$

with $v_p(x_j) \geq 1$ for $j = 1, \ldots, d-3$. Here, we have denoted by k the smallest index between 2 and $d-2$ such that $f^{(d-k)}(1) \neq 0$ (we know from Lemma 18 that such a k exists). We introduce the notation

$$S_m = \sum_{j=1}^{d-3} x_j^m.$$

Then we have $v_p(S_1) = v_p(d-1) = 1$, and $v_p(S_j) \geq 2$ for $j = 2, \ldots, d-2$. Using Newton's formulas (see Lemma 12 applied to $j = 0$), we obtain

$$\begin{aligned} -k\binom{d}{k}a_k &= \sum_{j=0}^{k-1}(-1)^j(1+S_{k-j})\binom{d}{j} \\ &= \sum_{j=0}^{k-1}(-1)^j\binom{d}{j} + \sum_{j=0}^{k-1}(-1)^j S_{k-j}\binom{d}{j} \\ &= (-1)^{k-1}\binom{d-1}{k-1} + \sum_{j=0}^{k-1}(-1)^j S_{k-j}\binom{d}{j}. \end{aligned}$$

Note that $v_p\left(\binom{d}{k}a_k\right) > 1$, which will lead to a contradiction:

- If $k = 2$, then the last equality becomes

$$-2\binom{d}{2}a_2 = -(d-1) + S_2 - dS_1 = -(d-1) + S_2 - d(d-1) = -(d+1)(d-1) + S_2.$$

  The valuation of the right-hand term is 1.
- If $3 \leq k \leq d-2$, then the right-hand term is

$$(-1)^{k-1}\binom{d-1}{k-1} + \sum_{j=0}^{k-2}(-1)^j S_{k-j}\binom{d}{j} + (-1)^{k-1}S_1\binom{d}{k-1}.$$

  But $v_p(S_{k-j}) \geq 2$ for $j = 0, \ldots, k-2$, and $v_p(S_1\binom{d}{k-1}) = 2$, so the valuation of the right-hand term is $v_p\binom{d-1}{k-1} = 1$. □

We remark that the proof of Proposition 20 in fact implies that there are no CA-polynomials of degree $p+1$ all of whose roots are contained in a number field in which p does not ramify. Indeed, this ensures that the valuations of the x_j are integers. Hence we can still conclude that $v_p(x_j) \geq 1$.

4. Algebraic varieties of counterexamples

4.1. Let k be an algebraically closed field and let $d > 0$ be an integer. The set of equivalence classes (in the sense of Lemma 6) of CA-polynomials of degree d will be denoted by $\mathrm{CA}_k(d)$.

4.2. We have a surjective map

$$\Phi_k(d, d-2) : V_k(d, d-2) \to \mathrm{CA}_k(d) : (p_1, \ldots, p_{d-2}) \mapsto x^2(x-p_1)\cdots(x-p_{d-2}),$$

where $V_k(d, d-2) \subset \mathbb{P}_k^{d-3}$ is the projective variety defined by the ideal

$$I_k(d, d-2) = \left(\mathrm{Res}_x(F, F_H^{(j)}) \;\middle|\; j = 2, \ldots, d-1 \right),$$

with $F = x^2(x-P_1)\dots(x-P_{d-2}) \in k[P_1,\dots,P_{d-2}][x]$. Therefore, in order to prove that no CA-polynomials exist in degree d, it suffices to show that $V_k(d,d-2) = \varnothing$. Note that $V_k(d,d-2)$ is invariant under coordinate permutations, so it is sufficient to show that $V_k(d,d-2)$ does not contain any points of the form $(p_1,\dots,p_{d-3},1)$. Setting $P_{d-2} = 1$ in $I_k(d,d-2)$, we obtain an ideal of $k[P_1,\dots,P_{d-3}]$ that is equal to the unit ideal if and only if $V_k(d,d-2) = \varnothing$. This can be checked using a finite Gröbner basis computation, which is exactly the approach of [6].

4.3. We also have a surjective map

$$\Phi_k(d,0) : V_k(d,0) \to \mathrm{CA}_k(d) :$$
$$(a_1,\dots,a_{d-2}) \mapsto x^2(x^{d-2} + a_1x^{d-3} + \cdots + a_{d-2}),$$

where now $V_k(d,0) \subset \mathbb{P}_k(d-2; d-1; \dots; 2; 1)$ is the weighted projective variety defined by the ideal

$$I_k(d,0) = \left(\operatorname{Res}_x(F, F_H^{(j)}) \,\middle|\, j = 2,\dots,d-1 \right),$$

with $F = x^2(x^{d-2} + A_1x^{d-3} + \cdots + A_{d-2}) \in k[A_1,\dots,A_{d-2}][x]$. Again, in order to show that no Casas-Alvero polynomials can exist in degree d, it is sufficient to prove that $V_k(d,0) = \varnothing$. This was used in the theoretical approach of [8].

4.4. We will make use of a hybrid version of the above maps. Namely, for each $t \in \{0,\dots,d-2\}$ we have a surjective map

$$\Phi_k(d,t) : V_k(d,t) \to \mathrm{CA}_k(d) :$$
$$(p_1,\dots,p_t,a_1,\dots,a_{d-2-t}) \mapsto x^2(x-p_1)\cdots(x-p_t)(x^{d-2-t}+a_1x^{d-3-t}+\cdots+a_{d-2-t}),$$

where $V_k(d,t) \subset \mathbb{P}_k(1;\dots;1;d-2-t;d-3-t;\dots;2;1)$ is the weighted projective variety defined by the ideal

$$I_k(d,t) = \left(\operatorname{Res}_x(F, F_H^{(j)}) \,\middle|\, j = 2,\dots,d-1 \right)$$

with

$$F = x^2(x-P_1)\cdots(x-P_t)(x^{d-2-t} + A_1x^{d-3-t} + \cdots + A_{d-2-t})$$

in $k[P_1,\dots,P_t,A_1,\dots,A_{d-2-t}][x]$. Once more it is sufficient to show that $V_k(d,t) = \varnothing$ (for any value of t) in order to prove that no Casas-Alvero polynomials of degree d exist over k.

4.5. To each scenario s for degree d of type t, we associate the variety

$$V_k(s) \subset V_k(d,t)$$

defined by the ideal

$$I_k(s) = \left(F_H^{(j)}(P_{s_j}) \,\middle|\, j = 2,\dots,d-1 \right) \subset k[P_1,\dots,P_t,A_1,\dots,A_{d-2-t}],$$

where

$$F = x^2(x-P_1)\cdots(x-P_t)(x^{d-2-t} + A_1x^{d-3-t} + \cdots + A_{d-2-t})$$

and $P_0 = 0$. Then it is clear that $V_k(s)$ parameterizes the CA-polynomials that match with s. Recall that every CA-polynomial matches with at least one scenario (e.g., its own scenario $\mathrm{scen}(f)$). Thus, if one wants to show that no CA-polynomials of degree d exist over k, it suffices to show that $V_k(s) = \varnothing$ for each scenario s for degree d. This is essentially the “more primary decomposition” that was mentioned in [8, Section 4], but in Section 5 below we will see that there is a significant amount of computational gain to be expected when viewing the set of CA-polynomials that match with s as a subvariety of $V_k(d,t)$ rather than of $V_k(d,d-2)$. Moreover, if $k=\mathbb{C}$, in view of the theoretical results obtained in Sections 2 and 3, it is actually sufficient to check whether $V_{\mathbb{C}}(s)=\varnothing$ for a restricted set of scenarios. We will elaborate the details of this for $d=12$ in Section 6.

5. Revisiting the computational approach

5.1. We now describe the basic version of our algorithm, discarding the theoretical results of Sections 2 and 3. The input is a field characteristic p (either 0 or a prime number) along with an integer $d>2$. The output is `yes` or `no`, depending on whether Casas-Alvero polynomials exist in degree d and characteristic p or not.

Step 1. Create a list L (of length $d-1$) of lists, such that $L[t]$ contains all scenarios for type t (for $t=0,\dots,d-2$). This can be done easily using $d-2$ nested for-loops. Let k be the field of rational numbers if $p=0$, and let k be the field with p elements otherwise. Set `answer := no`.

Step 2. For t going from 1 to $d-2$ do:

- Initiate the following variables/structures:
    * $R=k[P_1,\dots,P_{t-1},A_1,\dots,A_{d-2-t}]$,
    * $S=R[x]$,
    * $P_0=0$ and $P_t=1$,
    * $F(x)=x^2(x-P_1)\cdots(x-P_t)(x^{d-2-t}+A_1x^{d-3-t}+\cdots+A_{d-2-t})$,
    * $\prec$ = a monomial ordering that first eliminates $A_1,\dots,A_{d-2-t}$ and that behaves like `grevlex` on the remaining variables $P_1,\dots,P_{t-1}$.
- For s in $L[t]$ do:
    * Let $I_k^{\mathrm{aff}}(s)\subset R$ be the ideal generated by $F_H^{(j)}(P_{s_j})$ for $j=2,\dots,d-1$. Check whether or not $I_k^{\mathrm{aff}}(s)=R$ by checking if the reduced Gröbner basis (w.r.t. $\prec$) of $I_k^{\mathrm{aff}}(s)$ equals $\{1\}$. If it does not, set `answer := yes` and quit the loops.

Step 3. Output `answer`.

5.2. Modulo a base change to the algebraic closure of k, $I_k^{\mathrm{aff}}(s)$ is obtained from $I_k(s)$ (as described in 4.5) by setting $P_t=1$, so it only describes an affine part of $V_k(s)$. However, it suffices to verify that this affine part is empty. Indeed, the type of a CA-polynomial corresponding to a point $(p_1,\dots,p_t,a_1,\dots,a_{d-2-t})\in V_k(s)$ with $p_t=0$ is strictly smaller than t, so we would have encountered it already.

5.3. The variables $A_1,\dots,A_{d-2-t}$ appear linearly in the defining polynomials $F_H^{(j)}(P_{s_j})$. Therefore, they can be eliminated easily. (In fact, the corresponding linear system is in echelon form, so the A_i's could also be eliminated bottom-up by hand.) The lower the type, the more variables can be eliminated and the easier the Gröbner basis computation becomes (in the extreme case $t=1$, one obtains a linear system in $d-3$ variables). This is the main reason for our use of the hybrid varieties $V_k(d,t)$.

5.4. It is theoretically possible to avoid Gröbner basis computations and use linear algebra instead. Indeed, $I_k^{\mathrm{aff}}(s)=R$ is equivalent to the solvability of

$$1=g_1\cdot F_H^{(2)}(P_{s_2})\;+\;\cdots\;+\;g_{d-2}\cdot F_H^{(d-1)}(P_{s_{d-1}}) \tag{16}$$

in terms of polynomials $g_i \in R$. If such polynomials exist, by the effective Nullstellensatz they can be chosen such that their degree is bounded by d^d (e.g., see [9]). So in principle, one could use indeterminate coefficients to translate the solvability of (16) into the solvability of some linear system of equations. But this system is so huge that no gain is to be expected (although maybe this deserves a deeper analysis).

5.5. One can speed up the algorithm slightly by noting the following. If $s_2 = 0$, then the first defining polynomial is

$$F_H^{(2)}(0) = (-1)^t \cdot P_1 \cdots P_{t-1} \cdot A_{d-2-t}.$$

But Casas-Alvero polynomials corresponding to $P_1 \cdots P_{t-1} = 0$ are of strictly lower type than t, so they would have been encountered already. Therefore, our defining polynomial can be replaced by A_{d-2-t}. If in addition $s_3 = 0$, then similarly the second defining polynomial can be replaced by A_{d-3-t}, and so on. Suppose that the first non-zero entry of s appears at position j. Then after substituting $A_{d-2-t} = \cdots = A_{d-j+1-t} = 0$ (no substitutions if $j = 2$), one finds that

$$F_H^{(j)}(P_{s_j}) = F_H^{(j)}(P_1)$$

is a multiple of P_1. For the same reason, this factor can be removed.

5.6. The above algorithm can be used straightforwardly to find all bad primes for a given degree d (given that we know that the Casas-Alvero conjecture is true in degree d):

(1) Initialize a set of candidate bad primes $C = \{\,\}$.
(2) First run the basic algorithm with $p = 0$, but instead of just checking whether the reduced Gröbner basis of $I_{\mathbb{Q}}^{\mathrm{aff}}(s)$ equals $\{1\}$, compute polynomials $g_1, \ldots, g_{d-2} \in R$ for which (16) holds. Then add every prime factor appearing in the denominators of the g_j to C.
(3) If a prime p is not in C, it cannot be a bad prime because each of the expansions (16) can be reduced mod p. To find which candidate bad primes are actually bad primes, we run the basic algorithm for each $p \in C$.

An implementation of this method can be found in `CAbadprimes.m`.

5.7. The hardest part is step 2, because one computes in characteristic 0. Note that it is possible to give an upper bound for the elements of C purely in terms of d, so that step 2 could, in principle, be avoided. Indeed, see the discussion following (16)—the denominators of the solutions of the linear system can be bounded using Cramer's rule. But the bound one obtains is too large to be of any practical use.

5.8. We have executed the algorithm for $d = 5$, $d = 6$ and $d = 7$. In case of $d = 5$, the total time needed was less than 0.03 seconds. For $d = 6$, the computer needed less than 3 seconds. A naive run of the algorithm for $d = 7$ is not expected to end in a reasonable amount of time, because the denominators become very hard to factor. But by using several monomial orders and computing greatest common divisors, one can make the case $d = 7$ feasible in Magma (apart from the factorization of one composite 119-digit number, for which we used the `CADO-NFS` package [1]). The file `CAbadprimes7test.m` contains Magma code proving the correctness of our output. The case $d = 8$ lies out of reach. Of course, exhaustive lists of bad primes for increasing degrees become less and less interesting. But it would be good to have

TABLE 2. The smallest non-bad prime p that does not divide d

d	2	3	4	5	6	7	8	9	10
p	-	-	11	13	17	127	419	941	3803

an idea of the growth rate of the largest bad prime, or on the number of bad primes. Such lists can also be helpful in detecting patterns (we could not observe any). By just repeating our basic algorithm for increasing values of p, it is feasible to find the smallest non-bad prime (that does not divide d), for d up to 10. We have put the outcomes in Table 2.

6. The Casas-Alvero conjecture in degree 12

6.1. A direct application of the basic algorithm for degree $d = 12$ and characteristic $p = 0$ does not seem to be realistic. Two observations lead to a crucial speed-up:

- as remarked in 4.5, in view of the theoretical results obtained in Sections 2 and 3, it suffices to show that $V_{\mathbb{C}}(s) = \varnothing$ for a restricted set of scenarios s,
- for each such s, it actually suffices to show that $V_{\overline{\mathbb{F}}_p}(s) = \varnothing$ for a single prime p, because the varieties are projective and take equations over $\mathbb{Z}$.

6.2. As for the first speed-up, by Theorem 2 and Proposition 16 it suffices to prove that $V_{\mathbb{C}}(s) = \varnothing$ for all scenarios $s = (s_1, \ldots, s_{11})$ for which

- $s_1 = 0 \neq s_{11}$,
- $s_3 \neq s_9$,
- $s_4 \neq s_8$,
- ind(s) satisfies the determinant condition mentioned in the énoncé of Theorem 2.

(We omit the contribution of Proposition 9 to this discussion, because the arguments involved are rather subtle, whereas the computational gain is limited.) Let L_{res} be obtained from L (as introduced in 5.1) by restricting to these scenarios. Then $\mathrm{L}_{\mathrm{res}}$ contains

$$0, 6, 718, 5210, 8918, 5404, 1352, 141, 5, 0, 0$$

scenarios of type $0, \ldots, 10$, respectively (this is less than was mentioned in (3), where only the determinant condition was taken into account). However, for the algorithm to work rigorously, the list L_{res} should be slightly enlarged again, so that it becomes closed under taking *descendants*, in the following sense.

Definition 2. Let $d > 0$ be an integer and let $s = (s_1, \ldots, s_{d-1})$ be a scenario for degree d. Let $t = \mathrm{type}(s)$. Then we say that $s' = (s'_1, \ldots, s'_{d-1})$ is a *descendant* of s if there exists a $1 \leq j \leq t$ such that for all $i = 1, \ldots, d-1$:

- $s'_i = s_i$ if $s_i < j$,
- $s'_i = 0$ if $s_i = j$,
- $s'_i = s_i - 1$ if $s_i > j$.

This ensures that working in the affine subvariety $P_t = 1$ (see 5.2) and speeding up the algorithm (as in 5.5) are still justified. Note that if s' is a descendant of s, then $\mathrm{type}(s') = \mathrm{type}(s) - 1$. By closing L_{res} under taking descendants, one obtains a list $L_{\mathrm{res}}^{\mathrm{cl}}$ containing

$$1, 279, 3892, 12073, 13661, 6685, 1491, 146, 5, 0, 0$$

Table 3. Approximate time and memory requirements for settling $d = 12$, as if the algorithm were executed on a single core. In practice, types 6 and 7 were spread among multiple cores. In case of type 8, this was not possible due to memory limitations.

type	# scenarios	time	memory
1	279	0.1 secs	$\ll$ 0.1 GB
2	3892	43 secs	$\ll$ 0.1 GB
3	12073	2 mins	$<$ 0.1 GB
4	13661	40 mins	0.1 GB
5	6685	20 hours	0.2 GB
6	1491	2 weeks	1.3 GB
7	146	16 weeks	10 GB
8	5	15 weeks	90 GB

scenarios of type $0, \ldots, 10$, respectively. This might appear as a significant increase in the number of scenarios. However, recall that scenarios of low type can be eliminated very easily.

6.3. As for the second speed-up, based on the experimentally observed distribution of bad primes in degrees $d \leq 7$, any prime p which is "not too small" is most likely to work. If, nevertheless, the computation breaks down and a `yes` is printed, one can redo the computation using a different value of p. (In principle, it is possible to give a lower bound on p so that it is guaranteed to work, but this bound is much too large to be of any practical use—recall from Theorem 4 that the largest bad prime for $d = 7$ has already 135 decimal digits.) Our first try was $p = 10^7 + 17$ and it immediately worked. It is convenient to use the same p for all scenarios listed in $L_{\text{res}}^{\text{cl}}$. At least, if a scenario s is treated modulo some p, then all of its subsequent descendants should be treated modulo the same p. Indeed, this enables us to conclude that the *projective* variety $V_{\overline{\mathbb{F}}_p}(s)$ is empty, and hence that $V_{\mathbb{C}}(s) = \varnothing$.

6.4. Magma code implementing the above method can be found in the file `CAdeg12.m`. We have executed the algorithm and the outcome was affirmative (i.e., the Casas-Alvero conjecture is true in degree 12, thereby proving Theorem 5). Approximate time and memory requirements can be found in Table 3.

6.5. The computation fills in the smallest open entry in the list of degrees for which the Casas-Alvero conjecture is known to hold. To our knowledge, the list of degrees $d \leq 100$ for which the conjecture is still open is

$$20, 24, 28, 30, 35, 36, 40, 42, 45, 48, 55, 56, 60, 63, 66, 70, 72, 77, 78, 80, 84, 88, 90, 91, 98, 99, 100.$$

Our algorithm can in principle be generalized to higher degrees (note, in particular, that the two next open cases $d = 20$ and $d = 24$ are also of the form $p + 1$). But without new theoretical ingredients, an implementation of this is expected to demand astronomical amounts of time and memory.

References

[1] Shi Bai, Pierrick Gaudry, Alexander Kruppa, François Morain, Emmanuel Thomé and Paul Zimmerman, *CADO-NFS 1.1*, available at `http://cado-nfs.gforge.inria.fr/`.

[2] Wieb Bosma, John Cannon, and Catherine Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), no. 3-4, 235–265, DOI 10.1006/jsco.1996.0125. Computational algebra and number theory (London, 1993). MR1484478

[3] Eduardo Casas-Alvero, *Higher order polar germs*, J. Algebra **240** (2001), no. 1, 326–337, DOI 10.1006/jabr.2000.8727. MR1830556 (2002e:14003)

[4] Mustapha Chellali and Alain Salinier, *La conjecture de Casas-Alvero pour les degrés* $5p^e$, An. Univ. Dunărea de Jos Galaţi, Fasc. II Mat. Fiz. Mec. Teor. 4(35) (2012), no. 1-2, 54–62. MR3136558

[5] Rosa de Frutos, *Perspectivas aritméticas para la conjetura de Casas-Alvero*, Ph.D. thesis, Universidad de Valladolid, (2013).

[6] Gema M. Diaz-Toca and Laureano Gonzalez-Vega, *On analyzing a conjecture about univariate polynomials and their roots by using maple*, Proceedings of the Maple Conference 2006, Waterloo (Canada), July 23-26, 2006, pp. 81–98 (2006).

[7] Jan Draisma and Johan P. de Jong, *On the Casas-Alvero conjecture*, Eur. Math. Soc. Newsl. **80** (2011), 29–33. MR2848893 (2012g:12004)

[8] Hans-Christian Graf von Bothmer, Oliver Labs, Josef Schicho, and Christiaan van de Woestijne, *The Casas-Alvero conjecture for infinitely many degrees*, J. Algebra **316** (2007), no. 1, 224–230, DOI 10.1016/j.jalgebra.2007.06.017. MR2354861 (2009a:13048)

[9] János Kollár, *Sharp effective Nullstellensatz*, J. Amer. Math. Soc. **1** (1988), no. 4, 963–975, DOI 10.2307/1990996. MR944576 (89h:12008)

[10] Adrien-Marie Legendre, *Théorie des nombres*, Firmin Didot Frères, Paris (1830).

[11] Victor V. Prasolov, *Polynomials*, Algorithms and Computation in Mathematics, vol. 11, Springer-Verlag, Berlin, 2010. Translated from the 2001 Russian second edition by Dimitry Leites; Paperback edition [of MR2082772]. MR2683151 (2011g:12001)

[12] Paulo Ribenboim, *The theory of classical valuations*, Springer Monographs in Mathematics, Springer-Verlag, New York, 1999. MR1677964 (2000d:12007)

[13] Hendrik Verhoek, *Some remarks about a polynomial conjecture of Casas-Alvero*, Séminaire Bourbakettes, Paris (2009).

Departement Wiskunde, KU Leuven, Celestijnenlaan 200B, 3001 Leuven (Heverlee), Belgium

E-mail address: `wouter.castryck@wis.kuleuven.be`

Institut de Recherche Mathématique Avancée, Université de Strasbourg, 7 Rue René Descartes, 67084 Strasbourg CEDEX, France

E-mail address: `robert.laterveer@math.unistra.fr`

Institut de Recherche Mathématique Avancée, Université de Strasbourg, 7 Rue René Descartes, 67084 Strasbourg CEDEX, France

E-mail address: `myriam.ounaies@unistra.fr`