J. Aust. Math. Soc. **101** (2016), 56–94
doi:10.1017/S144678871500052X

GROWTH RATES OF ALGEBRAS, I: POINTED CUBE TERMS

KEITH A. KEARNES[✉], EMIL W. KISS and ÁGNES SZENDREI

(Received 11 November 2013; accepted 7 January 2015; first published online 22 January 2016)

Communicated by B. Martin

Abstract

We investigate the function $d_{\mathbf{A}}(n)$, which gives the size of a least size generating set for $\mathbf{A}^n$.

2010 *Mathematics subject classification*: primary 08A40; secondary 08A55, 08B05.

Keywords and phrases: growth rate, basic identity, pointed cube term.

1. Introduction

For a finite algebra $\mathbf{A}$, write $d_{\mathbf{A}}(n) = g$ if g is the least size of a generating set for $\mathbf{A}^n$, and write $h_{\mathbf{A}}(g) = n$ if the largest power of $\mathbf{A}$ that is g-generated is $\mathbf{A}^n$. The functions $d_{\mathbf{A}}$ and $h_{\mathbf{A}}$ map natural numbers to natural numbers and are related by

$$d_{\mathbf{A}}(n) \leq g \iff \mathbf{A}^n \text{ is } g\text{-generated} \iff n \leq h_{\mathbf{A}}(g),$$

which asserts that $d_{\mathbf{A}}$ is the lower adjoint of $h_{\mathbf{A}}$ and $h_{\mathbf{A}}$ is the upper adjoint of $d_{\mathbf{A}}$. It follows that $d_{\mathbf{A}}, h_{\mathbf{A}} : \omega \to \omega$ are increasing functions, which are inverse bijections between their images:

$$\operatorname{im}(d_{\mathbf{A}}) \underset{d}{\overset{h}{\rightleftharpoons}} \operatorname{im}(h_{\mathbf{A}});$$

and, moreover, each determines the other. These functions make sense for partial algebras and infinite algebras, too.

The study of the functions $d_{\mathbf{A}}$ and $h_{\mathbf{A}}$ has a long history, which we briefly survey.

1.1. The ϕ-function of a group. In the 1936 paper [15], Hall generalized the Euler ϕ-function from number theory by defining $\phi_k(G)$ to be the number of k-tuples $\mathbf{t} = (t_1, \ldots, t_k)$ for which $\{t_1, \ldots, t_k\}$ is a generating set of the group G. The classical Euler ϕ-function is therefore $\phi(k) = \phi_1(\mathbb{Z}_k)$. Hall called two generating k-tuples $\mathbf{t}_1$ and

This material is based upon work supported by the Hungarian National Foundation for Scientific Research (OTKA), grant nos. K77409, K83219, and K104251.

$\mathbf{t}_2$ 'equivalent' if there is an automorphism α of G which applied coordinatewise to $\mathbf{t}_1$ yields $\mathbf{t}_2$. The automorphism group of G acts freely on generating k-tuples and hence the number of equivalence classes of generating k-tuples is $\phi_k(G)/|\mathrm{Aut}(G)|$. Hall denoted $\phi_k(G)/|\mathrm{Aut}(G)|$ by $d_k(G)$, an unfortunate conflict with more recent notation since $\phi_k(G)/|\mathrm{Aut}(G)|$ is closer to the h-function than to the d-function. Indeed, if G is a finite simple nonabelian group, then $h_G(k) = \phi_k(G)/|\mathrm{Aut}(G)|$.

Hall called the function $\phi_k(G)/|\mathrm{Aut}(G)|$ 'intrinsically more interesting' than $\phi_k(G)$, and derived a formula for it in the case where G is a finite simple nonabelian group, namely

$$h_G(k) = \frac{1}{|\mathrm{Aut}(G)|} \sum_{H \le G} \mu(H)|H|^k, \tag{1.1}$$

where μ is the Möbius function of the subgroup lattice of G. This calculation is the first result of our topic.

1.2. Non-Hopf kernels. A group is *Hopfian* if every surjective endomorphism is an isomorphism, and non-Hopfian otherwise. A group N is a *non-Hopf kernel* of G if it is isomorphic to the kernel of a surjective endomorphism of G that is not an isomorphism. In the 1969 paper [4], Dey investigated the problem of determining which groups are non-Hopf kernels. Dey noted that every nontrivial group is a non-Hopf kernel, since, for example, the kernel of the shift

$$N^\omega \to N^\omega : (n_0, n_1, n_2, \ldots) \mapsto (n_1, n_2, n_3 \ldots)$$

is isomorphic to N. Dey restricted attention to non-Hopf kernels of finitely generated groups, and noted the following: a finite complete group is not a non-Hopf kernel of a finitely generated group. (The group N is complete if it is centerless and $\mathrm{Aut}(N) = \mathrm{Inn}(N)$.) His reasoning goes like this: if N is complete and a non-Hopf kernel of G, then $C_G(N)$ is a normal complement to N. By the non-Hopf property, $C_G(N) \cong G$, so

$$G \cong N \times G \cong N^2 \times G \cong N^3 \times G \cong \cdots.$$

If G is finitely generated, say by g elements, then so are the quotient groups N^n for all finite n. But this contradicts the local finiteness of the variety $\mathcal{V}(N)$. Specifically, the g-generated groups in this variety have size at most $|N|^{|N|^g}$. Thus, Dey's paper drew attention to the (easy) fact that if N is finite, then the number of elements required to generate N^n goes to infinity as n goes to infinity. (In symbols, $\lim_{n\to\infty}(d_N(n)) = \infty$.)

1.3. Growth rates of groups. In the 1974 paper [31], Wiegold cited Dey's work on non-Hopf kernels as the inspiration for his investigation into the question 'What are the ways in which ... $[d_G(n)]$ ... can tend to infinity [when G is a finite group]?' Wiegold inverted Hall's formula (1.1) to show that, for $n > 0$, $d_G(n)$ is one of the three natural numbers nearest

$$\log_{|G|}(n) + \log_{|G|}(|\mathrm{Aut}(G)|)$$

when G is a finite simple nonabelian group, so in this case $d_G(n)$ is asymptotically equivalent to $\log(n)$. He showed that $d_G(n)$ has logarithmic upper and lower bounds whenever G is a finite perfect group. (The group G is perfect if $[G, G] = G$.) He showed

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

also that $d_G(n)$ agrees with a linear function for large n if G is a finite imperfect group. Thus, he established that $d_G(n)$ tends to infinity as a logarithmic or linear function when G is a finite group.

1.4. Growth rates of groups, semigroups, and group expansions. Wiegold's paper initiated a program of research into growth rates of groups including, for example, [5–11, 23–26, 30, 32–34, 36, 37]. The program expanded to include the investigation of growth rates of semigroups, in [16, 27, 35], and later to include the investigation of growth rates of more general algebraic structures, in [14, 29]. Some of the questions being investigated about growth rates of finite algebras are related to the following theorems of Wiegold:

(I) a finite perfect group has growth rate that is logarithmic ($d_{\mathbf{A}}(n) \in \Theta(\log(n))$), while a finite imperfect group has growth rate that is linear ($d_{\mathbf{A}}(n) \in \Theta(n)$);

(II) a finite semigroup with identity has growth rate that is logarithmic or linear, while a finite semigroup without identity has growth rate that is exponential ($d_{\mathbf{A}}(n) \in 2^{\Theta(n)}$) [35].

Riedel partially extended Item (I) to congruence uniform varieties in [29] by proving that finite algebras in such varieties that are perfect (in the sense of modular commutator theory) have logarithmic growth rate. The paper [28] by Quick and Ruškuc extended Item (I) to any variety of rings, modules, k-algebras, or Lie algebras, but also fell short of extending Item (I) to arbitrary congruence uniform varieties.

1.5. Our work. We got interested in growth rates of finite algebras after reading [28, Remark 4.15], which states that *At present no finite algebraic structure is known for which the d-sequence does not have one of logarithmic, linear or exponential growth.* We found some of these missing algebras (Theorem 5.10).

Our interest in growth rates was later strengthened upon learning about the paper [3] by Chen, which linked growth rates with the constraint satisfaction problem by giving a polynomial time reduction from the quantified constraint satisfaction problem to the ordinary constraint satisfaction problem for algebras with $d_{\mathbf{A}}(n) \in O(n^k)$ for some k. Our new algebras are relevant to this investigation.

Our work is currently a three-paper series, of which this is the first.

1.5.1. This paper. The results from [28], about growth rates in varieties of classical algebraic structures, can be presented in a stronger way. Let Σ be a set of identities. If $\mathbf{A}$ is an algebra in a language $\mathcal{K}$, then we say that $\mathbf{A}$ *realizes* Σ if there is a way to interpret the function symbols occurring in Σ as $\mathcal{K}$-terms in such a way that each identity in Σ holds in $\mathbf{A}$. What is really proved in [28] is that if Σ_{Grp} is the set of identities axiomatizing the variety of groups and $\mathbf{A}$ is a finite algebra realizing Σ_{Grp}, then $\mathbf{A}$ has a logarithmic growth rate if it is perfect and has a linear growth rate if it is imperfect. Although the results of [28] are stated for only a few specific varieties of group expansions, the results hold for *any* variety of group expansions.

The main results of this paper are also best expressed in the terminology of algebras realizing a set of identities. Call a term *basic* if it contains at most one nonnullary

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

function symbol. An identity $s \approx t$ is basic if the terms on both sides are. This paper is an investigation into the restrictions imposed on growth rates of finite algebras by a set Σ of basic identities. A new concept that emerges from this investigation is the notion of a pointed cube term. If Σ is a set of identities in a language $\mathcal{L}$, then an $\mathcal{L}$-term $F(x_1, \dots, x_m)$ is a *p-pointed k-cube term* for the variety axiomatized by Σ if there is a $k \times m$ matrix M consisting of variables and p distinct constant symbols, with every column of M containing a symbol different from x, such that

$$\Sigma \models F(M) \approx \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix}. \tag{1.2}$$

Equation (1.2) is meant to be a compact representation of a sequence of k row identities of a special kind. For example,

$$\Sigma \models m\begin{pmatrix} x & y & y \\ y & y & x \end{pmatrix} \approx \begin{pmatrix} x \\ x \end{pmatrix}, \tag{1.3}$$

which is the assertion that $\Sigma \models m(x, y, y) \approx x$ and $\Sigma \models m(y, y, x) \approx x$, witnesses that $m(x_1, x_2, x_3)$ is a 3-ary, 0-pointed, 2-cube term. The basic identities (1.3) define what is called a *Maltsev term*. For another example,

$$\Sigma \models B\begin{pmatrix} 1 & x \\ x & 1 \end{pmatrix} \approx \begin{pmatrix} x \\ x \end{pmatrix},$$

which is the assertion that $\Sigma \models B(1, x) \approx x$ and $\Sigma \models B(x, 1) \approx x$, witnesses that $B(x_1, x_2)$ is a 2-ary, 1-pointed, 2-cube term. As a final example,

$$\Sigma \models M\begin{pmatrix} y & x & x \\ x & y & x \\ x & x & y \end{pmatrix} \approx \begin{pmatrix} x \\ x \\ x \end{pmatrix},$$

which is the assertion that M is a majority term for the variety axiomatized by Σ, witnesses that $M(x_1, x_2, x_3)$ is a 3-ary, 0-pointed, 3-cube term.

To state our main results, let Σ be a set of basic identities. We show the following results.

(1) The growth rate of any partial algebra can be realized as the growth rate of a total algebra (Corollary 3.3). If the partial algebra is finite, then the total algebra can be taken to be finite.

(2) A function $D\colon \omega \to \omega^+$ arises as the d-function of a countably infinite algebra if and only if (i) D is increasing and satisfies (ii) $D(0) = 0$ or 1 and (iii) $D(2) > 0$ (Theorem 3.4).

(3) If Σ does not entail the existence of a pointed cube term, then Σ imposes no restriction on growth rates of nontrivial algebras (Theorem 5.3). That is, for every algebra $\mathbf{A}$ of more than one element, there is an algebra $\mathbf{B}$ realizing Σ such that $d_{\mathbf{B}}(n) = d_{\mathbf{A}}(n)$ for $n > 0$. The algebra $\mathbf{B}$ can be taken to be finite if $\mathbf{A}$ is finite and the set Σ involves only finitely many distinct constants.

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

(4) If Σ entails the existence of a p-pointed cube term, $p \geq 1$, then any algebra $\mathbf{A}$ realizing Σ such that $\mathbf{A}^{p+k-1}$ is finitely generated has growth rate that is bounded above by a polynomial (Theorem 5.5). This is a nontrivial restriction.

(5) There exist finite algebras with pointed cube terms whose growth rate is asymptotically equivalent to a polynomial of any prescribed degree (Theorem 5.10).

(6) Any function that arises as the growth rate of an algebra with a pointed cube term also arises as the growth rate of an algebra without a pointed cube term (Theorem 5.15).

In addition to these items, we give a new proof of Kelly's completeness theorem for basic identities (Theorem 4.1). We give a procedure, based on this theorem, for deciding if a finite set of basic identities implies the existence of a pointed cube term (Corollary 5.2).

1.5.2. Our second paper [19]. We investigate growth rates of algebras with a 0-pointed k-cube term, which we shall just call a 'k-cube term'. Such terms were first identified in [1] in connection with investigations into constraint satisfaction problems, while an equivalent type of term was identified independently in [21] in connection with investigations into compatible relations of algebras.

We show in [19] that if $\mathbf{A}$ has a k-cube term and $\mathbf{A}^k$ is finitely generated, then $d_{\mathbf{A}}(n) \in O(\log(n))$ if $\mathbf{A}$ is perfect, while $d_{\mathbf{A}}(n) \in O(n)$ if $\mathbf{A}$ is imperfect. One can strengthen 'Big O' to 'Big Theta' if $\mathbf{A}$ is finite. This extends Wiegold's result (I) for groups to a setting that includes, as special cases, any finite algebra with a Maltsev term (in particular, any finite algebra in a congruence uniform variety) or any finite algebra with a majority term.

1.5.3. Our third paper [20]. We investigate growth rates of finite solvable algebras. Our original aim was to show that the only growth rates exhibited by such algebras are linear or exponential functions. We do prove this for finite nilpotent algebras and we prove it for finite solvable algebras with a pointed cube term, but the general case of a finite solvable algebra without a pointed cube term remains open.

2. Preliminaries

2.1. Notation. We denote the set $\{1, \ldots, n\}$ by $[n]$. A tuple in A^n may be denoted $(a_1, \ldots, a_n)$ or $\mathbf{a}$, and may be viewed as a function $\mathbf{a}\colon [n] \to A$. A tuple $(a, a, \ldots, a) \in A^n$ with all coordinates equal to a may be denoted $\hat{a}$. The size of a set A, the length of a tuple $\mathbf{a}$, and the length of a string σ are denoted $|A|$, $|\mathbf{a}|$, and $|\sigma|$. Structures are denoted in bold face font, for example $\mathbf{A}$, while the universe of a structure is denoted by the same character in italic font, for example A. The subuniverse of $\mathbf{A}$ generated by a subset $G \subseteq A$ is denoted $\langle G \rangle$.

We will use Big O notation. If f and g are real-valued functions defined on some subset of the real numbers, then $f \in O(g)$ and $f = O(g)$ both mean that there are positive constants M and N such that $|f(x)| \leq M|g(x)|$ for all $x > N$. We write

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

$f \in \Omega(g)$ and $f = \Omega(g)$ to mean that there are positive constants M and N such that $|f(x)| \geq M|g(x)|$ for all $x > N$. Finally, $f \in \Theta(g)$ and $f = \Theta(g)$ mean that both $f \in O(g)$ and $f \in \Omega(g)$ hold.

2.2. Easy estimates.

THEOREM 2.1. *Let* $\mathbf{A}$ *be an algebra.*

(1) $d_{\mathbf{A}^k}(n) = d_{\mathbf{A}}(kn)$.
(2) *If* $\mathbf{B}$ *is a homomorphic image of* $\mathbf{A}$*, then* $d_{\mathbf{B}}(n) \leq d_{\mathbf{A}}(n)$.
(3) *If* $\mathbf{B}$ *is an expansion of* $\mathbf{A}$ *(equivalently, if* $\mathbf{A}$ *is a reduct of* $\mathbf{B}$*), then* $d_{\mathbf{B}}(n) \leq d_{\mathbf{A}}(n)$.
(4) (From [28]) *If* $\mathbf{B}$ *is the expansion of* $\mathbf{A}$ *obtained by adjoining all constants, then*

$$d_{\mathbf{A}}(n) - d_{\mathbf{A}}(1) \leq d_{\mathbf{B}}(n) \leq d_{\mathbf{A}}(n).$$

PROOF. For (1), both $d_{\mathbf{A}^k}(n)$ and $d_{\mathbf{A}}(kn)$ represent the number of elements in a smallest size generating set for $(\mathbf{A}^k)^n \cong \mathbf{A}^{kn}$.

For (2), if $\varphi\colon \mathbf{A} \to \mathbf{B}$ is surjective and $G \subseteq A^n$ is a smallest size generating set for $\mathbf{A}^n$, then $\varphi(G)$ is a generating set for $\mathbf{B}^n$. Hence, $d_{\mathbf{B}}(n) \leq |\varphi(G)| \leq |G| = d_{\mathbf{A}}(n)$.

For (3), if $G \subseteq A^n$ is a smallest size generating set for $\mathbf{A}^n$, then G is also a generating set for $\mathbf{B}^n$. Hence, $d_{\mathbf{B}}(n) \leq |G| = d_{\mathbf{A}}(n)$.

For (4), the right-hand inequality $d_{\mathbf{B}}(n) \leq d_{\mathbf{A}}(n)$ follows from (3). Now let $G \subseteq A^n$ be a smallest size generating set for $\mathbf{B}^n$ and let $H \subseteq A$ be a smallest size generating set for $\mathbf{A}$. For each $a \in H$, let $\hat{a} = (a, a, \ldots, a) \in A^n$ be the associated constant tuple, and let $\widehat{H}$ be the set of these. Every tuple of A^n is generated from G by polynomial operations of $\mathbf{A}$ acting coordinatewise and hence is generated from $G \cup \widehat{H}$ by term operations of $\mathbf{A}$ acting coordinatewise. This proves that $d_{\mathbf{A}}(n) \leq |G| + |H| = d_{\mathbf{B}}(n) + d_{\mathbf{A}}(1)$, from which the left-hand inequality follows. □

The next theorem will not be used later in the paper, except that in Section 6 one should know that the d-function of a finite algebra is bounded below by a logarithmic function and above by an exponential function.

THEOREM 2.2. *If* $\mathbf{A}$ *is a finite algebra of more than one element and* $n > 0$*, then*

$$\lceil \log_{|A|}(n) \rceil \leq d_{\mathbf{A}}(n) \leq |A|^n$$

and

$$\lfloor \log_{|A|}(n) \rfloor \leq h_{\mathbf{A}}(n) \leq |A|^n.$$

Hence, $d_{\mathbf{A}}(n), h_{\mathbf{A}}(n) \in \Omega(\log(n)) \cap 2^{O(n)}$*. Moreover:*

(1) $d_{\mathbf{A}}(n) \in O(\log(n))$ *if and only if* $h_{\mathbf{A}}(n) \in 2^{\Omega(n)}$*;*
(2) $d_{\mathbf{A}}(n) \in O(n)$ *if and only if* $h_{\mathbf{A}}(n) \in \Omega(n)$*, and* $d_{\mathbf{A}}(n) \in \Omega(n)$ *if and only if* $h_{\mathbf{A}}(n) \in O(n)$*;*
(3) $d_{\mathbf{A}}(n) \in 2^{\Omega(n)}$ *if and only if* $h_{\mathbf{A}}(n) \in O(\log(n))$.

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

Proof. It follows from Theorem 2.1(3) that, among all algebras with universe A, the algebra with only projection operations for its term operations has the largest d-function and the algebra with all finitary operations as term operations has the smallest d-function. These two algebras are also extremes for the h-function.

If $\mathbf{A}$ has no nontrivial term operations, then every element of A^n is a required generator, so $d_{\mathbf{A}}(n) = |A|^n$. In this case, $h_{\mathbf{A}}(n) = \lfloor \log_{|A|}(n) \rfloor$ for $n > 0$, since h is the upper adjoint of d.

Now assume that $\mathbf{A}$ has all finitary operations as term operations. The n-generated free algebra in the variety generated by $\mathbf{A}$ is isomorphic to $\mathbf{A}^{|A|^n}$ [12, Theorem 3]. Since the largest n-generated algebra in this variety is a power of $\mathbf{A}$, it is also the largest n-generated power of $\mathbf{A}$ in the variety; we obtain that $h_{\mathbf{A}}(n) = |A|^n$. In this case, $d_{\mathbf{A}}(n) = \lceil \log_{|A|}(n) \rceil$ for $n > 0$, since d is the lower adjoint of h.

The fact that $d_{\mathbf{A}}$ is the lower adjoint of $h_{\mathbf{A}}$ suggests an asymmetry, in that

$$d_{\mathbf{A}}(n) \le k \iff n \le h_{\mathbf{A}}(k) \tag{2.1}$$

relates an upper bound of $d_{\mathbf{A}}$ to a lower bound of $h_{\mathbf{A}}$. But the fact that these functions are defined between totally ordered sets allows us to rewrite (2.1) as

$$h_{\mathbf{A}}(k) < n \iff k < d_{\mathbf{A}}(n),$$

which almost exactly reverses condition (2.1) on $d_{\mathbf{A}}$ and $h_{\mathbf{A}}$. Using this fact and the following claim, one easily verifies Items (1)–(3). □

Claim 2.3. *If $f, g\colon [a, \infty) \to \mathbb{R}$ are increasing functions that tend to infinity as x tends to infinity, then $\lfloor f(n) \rfloor < d_{\mathbf{A}}(n) \le \lceil g(n) \rceil$ holds for all large n if and only if $\lfloor g^{-1}(n) \rfloor \le h_{\mathbf{A}}(n) < \lceil f^{-1}(n) \rceil$ holds for all large n.*

Proof of Claim. To make the following lines slightly easier to read, allow '$\forall^+ n$' to stand for 'for all large n', that is, for '$(\exists N)(\forall n > N)$'. We have

$$\begin{aligned}\forall^+ n\,(d_{\mathbf{A}}(n) \le \lceil g(n) \rceil) &\implies \forall^+ n\,(n \le h_{\mathbf{A}}(\lceil g(n) \rceil))\\ &\implies \forall^+ n\,(\lfloor g^{-1}(n) \rfloor \le h_{\mathbf{A}}(\lceil g(\lfloor g^{-1}(n) \rfloor) \rceil))\\ &\implies \forall^+ n\,(\lfloor g^{-1}(n) \rfloor \le h_{\mathbf{A}}(n)),\end{aligned}$$

because the monotonicity of g guarantees that $\lceil g(\lfloor g^{-1}(n) \rfloor) \rceil \le n$. The reverse implication is proved in the same way, as are both implications in $\lfloor f \rfloor < d \Leftrightarrow h < \lceil f^{-1} \rceil$. □

Recall that the *free spectrum* of a variety $\mathcal{V}$ is the function $f_{\mathcal{V}}(n) := |F_{\mathcal{V}}(n)|$ whose value at n is the cardinality of the n-generated free algebra in $\mathcal{V}$.

Theorem 2.4. *If $\mathbf{A}$ is a nontrivial finite algebra and $f_{\mathcal{V}}$ is the free spectrum of the variety $\mathcal{V} = \mathcal{V}(\mathbf{A})$, then $h_{\mathbf{A}}(n) \le \log_{|A|}(f_{\mathcal{V}}(n))$ for $n > 0$. In particular:*

(1) *if $f_{\mathcal{V}}(n) \in O(n^k)$ for some fixed $k \in \mathbb{Z}^+$, then $d_{\mathbf{A}}(n) \in 2^{\Theta(n)}$;*
(2) *if $f_{\mathcal{V}}(n) \in 2^{O(n)}$, then $d_{\mathbf{A}}(n) \in \Omega(n)$.*

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

PROOF. Assume that $n > 0$.

The algebra $\mathbf{A}^{h_{\mathbf{A}}(n)}$ is n-generated and hence a quotient of the n-generated free algebra $\mathbf{F}_{\mathcal{V}}(n)$. This proves that $|A|^{h_{\mathbf{A}}(n)} \le f_{\mathcal{V}}(n)$ or $h_{\mathbf{A}}(n) \le \log_{|A|}(f_{\mathcal{V}}(n))$.

If $f_{\mathcal{V}}(n) \in O(n^k)$ for some fixed $k \in \mathbb{Z}^+$, then $\log(f_{\mathcal{V}}(n)) \in O(\log(n))$ and hence $h_{\mathbf{A}}(n) \in O(\log(n))$. Theorem 2.2 proves that $d_{\mathbf{A}}(n) \in 2^{\Omega(n)}$ holds when $h_{\mathbf{A}}(n)$ is bounded like this and that $d_{\mathbf{A}}(n) \in 2^{O(n)}$ holds just because $\mathbf{A}$ is finite, so $d_{\mathbf{A}}(n) \in 2^{\Theta(n)}$.

If $f_{\mathcal{V}}(n) \in 2^{O(n)}$, then $\log(f_{\mathcal{V}}(n)) \in O(n)$ and hence $h_{\mathbf{A}}(n) \in O(n)$. It follows from Theorem 2.2(2) that $d_{\mathbf{A}}(n) \in \Omega(n)$. □

COROLLARY 2.5. *Let* $\mathbf{A}$ *be a nontrivial finite algebra and let* $\mathbf{B}$ *be a nontrivial homomorphic image of* $\mathbf{A}^k$ *for some* k.

(1) *If b is strongly abelian (or even just strongly rectangular), then* $d_{\mathbf{A}}(n) \in 2^{\Theta(n)}$.

(2) *If* $\mathbf{B}$ *is abelian, then* $d_{\mathbf{A}}(n) \in \Omega(n)$.

PROOF. For the definitions and background about abelian, strongly abelian, and strongly rectangular algebras, we point the reader to [18, Chs 2 and 3] and to the paper [17]. This information is not needed to follow this proof. We only require knowledge about the free spectrum of certain varieties, and we cite the literature for this knowledge. (Note: the term 'rectangular' in [17] means the same thing as the phrase 'strongly rectangular' in [18], and this is the concept referred to in part (1) of this corollary.)

By Theorem 2.1(1) and (2), we have $d_{\mathbf{B}}(n) \le d_{\mathbf{A}}(kn)$, while by Theorem 2.2 we have $d_{\mathbf{A}}(kn) \in 2^{O(n)}$. Thus, to prove this corollary it suffices in (1) to prove that $d_{\mathbf{B}}(n) \in 2^{\Theta(n)}$, and it suffices in (2) to prove that $d_{\mathbf{B}}(n) \in \Omega(n)$. (That is, it suffices to prove the conclusions for $d_{\mathbf{B}}$ in place of $d_{\mathbf{A}}$.)

The proof really begins now. For (1), [17, Theorem 5.3] proves that a finite strongly rectangular algebra generates a variety with free spectrum bounded above by a polynomial. By Theorem 2.4, $d_{\mathbf{B}}(n) \in 2^{\Theta(n)}$ in this case. The strong abelian property is more restrictive than the strong rectangular property by [17, Lemma 2.2(11)].

For (2), any finite abelian algebra generates a variety $\mathcal{V}$ whose free spectrum satisfies $f_{\mathcal{V}}(n) \in 2^{O(n)}$, according to [2], so Theorem 2.4(2) completes the argument. □

Recall that an algebra is *affine* if it is polynomially equivalent to a module. It is known that $\mathbf{A}$ is affine if and only if $\mathbf{A}$ is abelian and has a Maltsev term if and only if $\mathbf{A}$ is abelian and has a Maltsev polynomial.

THEOREM 2.6. *If* $\mathbf{A}^2$ *is a finitely generated affine algebra, then* $d_{\mathbf{A}}(n) \in O(n)$. *If, moreover,* $\mathbf{A}$ *is finite and has more than one element, then* $d_{\mathbf{A}}(n) \in \Theta(n)$.

PROOF. The theorem is true under the weaker assumption that $\mathbf{A}$ (rather than $\mathbf{A}^2$) is finitely generated, provided $\mathbf{A}$ is a module rather than an arbitrary affine algebra. To see this, suppose that $\mathbf{M}$ is a module generated by a finite subset G. The set of tuples in $\mathbf{M}^n$ with exactly one nonzero entry, which is taken from G, is a generating set for $\mathbf{M}^n$ of size $\le |G| \cdot n$. Hence, $d_{\mathbf{M}}(n) \in O(n)$. If, moreover, $\mathbf{M}$ is finite and has more than one element, then Corollary 2.5(2) proves that $d_{\mathbf{M}}(n) \in \Omega(n)$, so $d_{\mathbf{M}}(n) \in \Theta(n)$.

It now follows from Theorem 2.1(4) that if $\mathbf{A}$ is an algebra that is polynomially equivalent to a finitely generated module, then $d_{\mathbf{A}}(n) \in O(n)$, and $d_{\mathbf{A}}(n) \in \Theta(n)$ if $\mathbf{A}$ is finite and nontrivial. Unfortunately, not every finitely generated affine algebra is polynomially equivalent to a finitely generated module. But if $\mathbf{A}$ is affine and $\mathbf{A}^2$ is finitely generated, then the linearization $\mathbf{A}^2/\Delta$ (see [13, page 114]) is also finitely generated and term equivalent to a reduct of the underlying module of $\mathbf{A}$. Hence, when $\mathbf{A}^2$ is finitely generated, then $\mathbf{A}$ is polynomially equivalent to a finitely generated module, and the conclusions of the theorem hold. □

3. General growth rates

3.1. Growth rates of partial algebras. A partial algebra is a set equipped with a set of partial operations. A total algebra is considered to be a partial algebra, but, of course, some partial algebras are not total.

The definitions of functions $d_{\mathbf{A}}$ and $h_{\mathbf{A}}$ make sense when $\mathbf{A}$ is a partial algebra, as does the problem of determining growth rates of partial algebras. Theorem 2.1(3), which relates the growth rate of an algebra to that of a reduct, holds in exactly the same form if a 'reduct of $\mathbf{B}$' is interpreted to mean an algebra $\mathbf{A}$ with the same universe as $\mathbf{B}$ whose basic partial operations are obtained from *some* of the term partial operations of $\mathbf{B}$ by possibly restricting their domains.

We will learn in this subsection that a function arises as the growth rate of a partial algebra if and only if it arises as the growth rate of a total algebra.

Definition 3.1. Let $\mathbf{A} = \langle A; P\rangle$ be a partial algebra with universe A and a set P of partial operations on A. The *one-point completion* of $\mathbf{A}$ is the total algebra whose universe is $A_0 := A \cup \{0\}$, where 0 is some element not in A, and whose operations $P_0 = \{p_0 \mid p \in P\} \cup \{\wedge\}$ are defined as follows.

(1) If $p \in P$ is a partial m-ary operation on A with domain $D \subseteq A^m$, then the total operation $p_0\colon (A_0)^m \to A_0$ is defined by

$$p_0(\mathbf{a}) = \begin{cases} p(\mathbf{a}) & \text{if } \mathbf{a} \in D, \\ 0 & \text{otherwise.} \end{cases}$$

(2) A meet operation $\wedge$ on A_0 is defined by

$$a \wedge b = \begin{cases} a & \text{if } a = b, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 3.2. *Let $\mathbf{A}$ be a partial algebra of more than one element, and let $\mathbf{A}_0$ be its one-point completion.*

(1) *Any generating set for $\mathbf{A}^n$ is a generating set for $\mathbf{A}_0^n$, and*
(2) *any generating set for $\mathbf{A}_0^n$ contains a generating set for $\mathbf{A}^n$.*

In particular, least size generating sets for $\mathbf{A}^n$ and $\mathbf{A}_0^n$ have the same size, and if $\mathbf{A}^n$ or $\mathbf{A}_0^n$ have any minimal generating sets, then they are the same.

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

Proof. In this paragraph we prove (1). If $G \subseteq A^n$ is a generating set for $\mathbf{A}^n$, then as a subset of $\mathbf{A}_0^n$ it will generate (in exactly the same manner) all tuples in A_0^n which have no 0's. If $\mathbf{z} \in A_0^n$ is an arbitrary tuple and $a, b \in A$ are distinct, let $\mathbf{z}_a$ and $\mathbf{z}_b$ be the tuples obtained from $\mathbf{z}$ by replacing all 0's with a and b, respectively. Then $\mathbf{z}_a, \mathbf{z}_b \in A^n$, so they are generated by G, and $\mathbf{z} = \mathbf{z}_a \wedge \mathbf{z}_b$, so $\mathbf{z}$ is also generated by G. Hence, G generates all of $\mathbf{A}_0^n$.

Now we prove (2). Assume that $H \subseteq A_0^n$ is a generating set for $\mathbf{A}_0^n$. If $\mathbf{a} \in A_0^n$, let $Z(\mathbf{a}) \subseteq [n]$ be the *zero set* of $\mathbf{a}$, by which we mean the set of coordinates where $\mathbf{a}$ is 0. It is easy to see that for any basic operation F of $\mathbf{A}_0$, it is the case that

$$Z(\mathbf{a}_1) \cup \cdots \cup Z(\mathbf{a}_m) \subseteq Z(F(\mathbf{a}_1, \ldots, \mathbf{a}_m)),$$

since 0 is absorbing for every basic operation. If the right-hand side is empty, then the left-hand side is empty as well; that is, tuples with empty zero sets can be generated only by tuples with empty zero sets. Said in a different way, if $H \subseteq A_0^n$ generates $\mathbf{A}_0^n$, then $H \cap A^n$ suffices to generate all tuples in A^n. If you consider how H generates elements of A^n in the algebra $\mathbf{A}_0^n$, it is clear that H generates those elements in the algebra $\mathbf{A}^n$ in exactly the same way, so H is a generating set for $\mathbf{A}^n$. □

Corollary 3.3. *If* $\mathbf{A}$ *is a partial algebra and* $\mathbf{A}_0$ *is its one-point completion, then* $d_{\mathbf{A}_0}(n) = d_{\mathbf{A}}(n)$ *for all* $n \in \omega$.

3.2. Growth rates of countably infinite algebras. In this section, we characterize the d-functions of countably infinite algebras. We will see that there are a few obvious properties that these functions have, and that any function $D\colon \omega \to \omega^+$ that has these properties may be realized as a d-function.

One obvious property of d-functions is that they are increasing: $m \le n$ implies $d_{\mathbf{A}}(m) \le d_{\mathbf{A}}(n)$. The d-function of a countably infinite algebra is an increasing function from the ordered set of natural numbers, ω, to the ordered set $\omega^+ = \omega \cup \{\omega\} = \{0, 1, \ldots, \omega\}$, where $d_{\mathbf{A}}(n) = \omega$ means that $\mathbf{A}^n$ is not finitely generated. The d-functions also have special initial values. The algebra $\mathbf{A}^0$ is a one-element algebra, so $\mathbf{A}^0$ is 0-generated if $\mathbf{A}$ has a nullary term and is 1-generated if $\mathbf{A}$ has no nullary term. Thus, $d_{\mathbf{A}}(0) = 0$ or 1, with the cases distinguished according to whether $\mathbf{A}$ has a nullary term. Finally, if $\mathbf{A}$ has more than one element, then $d_{\mathbf{A}}(2) > 0$, since any 0-generated subalgebra of $\mathbf{A}^2$ is contained in the diagonal and the diagonal is a proper subalgebra of $\mathbf{A}^2$ when $|A| > 1$. We now prove the following theorem.

Theorem 3.4. *If* $D\colon \omega \to \omega^+$*:*

(i) *is increasing;*

(ii) *satisfies* $D(0) = 0$ *or* 1*; and*

(iii) *satisfies* $D(2) > 0$,

then there is a countably infinite total algebra $\mathbf{A}$ *such that* $d_{\mathbf{A}}(n) = D(n)$ *for all* $n \in \omega$.

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

Proof. We construct a partial algebra $\mathbf{A}$ such that $d_{\mathbf{A}}(n) = D(n)$ for all $n \in \omega$. By Corollary 3.3, the one-point completion of $\mathbf{A}$ (Definition 3.1) will be a total algebra with the same growth rate.

First we describe the universe of our partial algebra. Start with a countably infinite set X. This set will be a subset of the universe of $\mathbf{A}$, and its main function is to ensure that the constructed algebra is infinite. Next, for any algebra $\mathbf{B}$, $d_{\mathbf{B}}(0) = 0$ happens exactly when $\mathbf{B}$ has a nullary term. Hence, if $D(0) = 0$ and we wish to represent D as $d_{\mathbf{A}}$ for some $\mathbf{A}$, then we must ensure that $\mathbf{A}$ has a nullary term. So, let $Y = \{y\}$ be a singleton set. If we need our algebra to have a nullary term, we will introduce a term with value y. Finally, for each nonzero $n \in \omega$, where $D(n)$ is finite, let $M^{(n)} = [z^{(n)}_{i,j}]$ be an $n \times D(n)$ matrix of elements such that all entries of all the $M^{(n)}$ are different from each other and are different from the elements of $X \cup Y$. Let $Z = \{z^{(n)}_{i,j}\}$ be the set of all entries appearing in these matrices, and take $A := X \cup Y \cup Z$ to be the universe of the partial algebra.

If $D(0) = 0$, then we introduce a nullary operation whose value is y. We may introduce more nullary operations later in the case $D(0) = 0$, but, if $D(0) = 1$, then we do not introduce any nullary operations throughout the construction.

For each nonzero $n \in \omega$, where $D(n)$ is finite, and for each tuple $\mathbf{b} \in A^n$, introduce a $D(n)$-ary partial operation $F_{\mathbf{b}}$ for which $F_{\mathbf{b}}(M^{(n)}) = \mathbf{b}$. This means that $F_{\mathbf{b}}$ has domain of size n, consisting of the n rows of $M^{(n)}$, and that $F_{\mathbf{b}}(z^{(n)}_{i,1}, \dots, z^{(n)}_{i,D(n)}) = b_i$ for each $i = 1, \dots, n$.

It is worth mentioning how to interpret the instructions of the previous paragraph in the case where $n = 1$ and $D(n) = 0$. Here $M^{(n)}$ is defined to be a 1×0 matrix and, for each $\mathbf{b} \in A^1 = A$, we are instructed to add a partial operation $F_{\mathbf{b}}$ with the property that $F_{\mathbf{b}}(M) = \mathbf{b}$. One should view $F_{\mathbf{b}}$ as a nullary partial operation with range $\mathbf{b}$. Hence, in the case $(n, D(n)) = (1, 0)$, we are to add nullary operations naming each element of A. (Consider how one might interpret the instructions of the previous paragraph in the case where $n = 2$ and $D(n) = 0$, if such were permitted by the assumptions on D. We would be instructed to add nullary partial operations to $\mathbf{A}$ with range $\mathbf{b}$ for each $\mathbf{b} \in A^2$. Such nullary operations do not exist for those $\mathbf{b} \in A^2$ off of the diagonal, so we would be unable to adhere to the instructions if we allowed $D(2) = 0$. This is the place in our construction where we make use of the assumption that $D(2) > 0$.)

Our partial algebra is A equipped with all partial operations of the type described in the previous three paragraphs.

Observe that $d_{\mathbf{A}}(0) = 0$ if and only if $\mathbf{A}$ has a nullary term if and only if $D(0) = 0$, so $d_{\mathbf{A}}(0) = D(0)$.

Observe that if $D(n) = \omega$ for some $n > 0$, then none of the partial operations has n distinct elements of A in its image. Hence, every tuple $\mathbf{b} \in A^n$ with distinct coordinates must appear in any generating set for $\mathbf{A}^n$. This proves that $d_{\mathbf{A}}(n) = \omega$ whenever $D(n) = \omega$.

Observe that if $D(1) = 0$, then we have added nullary operations to $\mathbf{A}$ naming each element of A, so $d_{\mathbf{A}}(1) = 0$, too.

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

Now we consider generating sets for $\mathbf{A}^n$ when $n > 0$ and $D(n)$ is finite and positive. In this case, $F_{\mathbf{b}}(M^{(n)}) = \mathbf{b}$ whenever $\mathbf{b} \in A^n$, so the columns of $M^{(n)}$ form a generating set of size $D(n)$ for $\mathbf{A}^n$. The following claim will help us to prove that there is no smaller generating set for $\mathbf{A}^n$.

Claim 3.5. *If $n > 0$ and a subset $G \subseteq A^n$ has fewer than $D(n)$ tuples whose coordinates are distinct, then the same is true for $\langle G \rangle$.*

Proof of Claim. If the claim is not true, then it must be possible to generate in one step a tuple $\mathbf{c} \in A^n$ whose coordinates are all distinct using other tuples, where fewer than $D(n)$ of these other tuples have the property that their coordinates are all distinct. If the partial operation used is some $F_{\mathbf{b}}$, $\mathbf{b} \in A^m$ for some m, and the tuples used to generate are $\mathbf{x}_1, \ldots, \mathbf{x}_{D(m)}$, then the following row equations must be satisfied

$$F_{\mathbf{b}}(\mathbf{x}_1, \ldots, \mathbf{x}_{D(m)}) = F_{\mathbf{b}}\left(\begin{bmatrix} x_{1,1} \\ \vdots \\ x_{n,1} \end{bmatrix}, \ldots, \begin{bmatrix} x_{1,D(m)} \\ \vdots \\ x_{n,D(m)} \end{bmatrix}\right) = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{c}.$$

Considering the definition of $F_{\mathbf{b}}$, it is clear that the (distinct) entries of $\mathbf{c}$ are among the entries of $\mathbf{b}$, so $m = |\mathbf{b}| \geq |\mathbf{c}| = n$. Moreover, the row equations $F_{\mathbf{b}}(x_{i,1}, \ldots, x_{i,D(m)}) = c_i$ can be solved in only one way, namely by using the appropriate row of $M^{(m)}$. This forces all entries of $[x_{i,j}]$ to be distinct. But this means there are $D(m)$ columns, $\mathbf{x}_j$, whose coordinates are distinct, and we assumed that there were fewer than $D(n)$ such columns. Altogether this yields that $m \geq n$ and $D(m) < D(n)$, contradicting the monotonicity of $D(n)$. The claim is proved. □

The claim shows that $d_{\mathbf{A}}(n) = D(n)$ when $n > 0$ and $D(n)$ is finite and positive, since a subset $G \subseteq A^n$ of size less than $D(n)$ must have fewer than $D(n)$ tuples whose coordinates are distinct. Such a set cannot generate $\mathbf{A}^n$, since the generated subuniverse $\langle G \rangle$ contains fewer than $D(n)$ tuples whose coordinates are distinct while A^n contains infinitely many such tuples. □

The construction in this proof may be modified to give some information about d-functions of finite algebras. Namely, suppose that $D\colon \{0, 1, \ldots, k\} \to \omega$ is (i) increasing and satisfies (ii) $D(0) = 0$ or 1 and (iii) $D(2) > 0$. If one modifies the construction in the proof by omitting the inclusion of the set X in the universe of $\mathbf{A}$ and then adding only the partial operations that are nullary or of the form $F_{\mathbf{b}}(M^{(n)}) = \mathbf{b}$, where $n \in \{0, 1, 2, \ldots, k\}$, then the proof shows that there is an algebra of size $|Y \cup Z| = 1 + \sum_{j=0}^{k} j \cdot D(j)$ (finite) such that $d_{\mathbf{A}}(n) = D(n)$ for $n \in \{0, 1, 2, \ldots, k\}$. Thus, there is no special behavior of d-functions of finite algebras on initial segments of ω.

4. Kelly's completeness theorem

In Section 4.1, we give a new proof of Kelly's completeness theorem for basic identities. The proof involves the construction of a model of a set of basic identities. In Section 4.2, we construct a simpler model by modifying the construction from

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

the completeness theorem. The simpler model is not adequate for proving the completeness theorem, but it is exactly what we need for our investigation of growth rates.

4.1. The completeness theorem for basic identities. Let $\mathcal{L}$ be an algebraic language. Recall that an $\mathcal{L}$-term is *basic* if it contains at most one nonnullary function symbol. An $\mathcal{L}$-identity $s \approx t$ is basic if both s and t are basic terms. If $\Sigma \cup \{\varphi\}$ is a set of basic identities, then φ is a *consequence* of Σ, written $\Sigma \models \varphi$, if every model of Σ is a model of φ.

Let C be the set of constant symbols of $\mathcal{L}$ and let X be a set of variables. The *weak closure of* Σ *in the variables* X is the smallest set $\overline{\Sigma}$ of basic identities containing Σ for which:

(i) $(t \approx t) \in \overline{\Sigma}$ for all basic $\mathcal{L}$-terms t with variables from X;
(ii) if $(s \approx t) \in \overline{\Sigma}$, then $(t \approx s) \in \overline{\Sigma}$;
(iii) if $(r \approx s) \in \overline{\Sigma}$ and $(s \approx t) \in \overline{\Sigma}$, then $(r \approx t) \in \overline{\Sigma}$;
(iv) if $(s \approx t) \in \overline{\Sigma}$ and $\gamma \colon X \to X \cup C$ is a function, then $(s[\gamma] \approx t[\gamma]) \in \overline{\Sigma}$, where $s[\gamma]$ denotes the basic term obtained from s by replacing each variable $x \in X$ with $\gamma(x) \in X \cup C$;
(v) if t is a basic $\mathcal{L}$-term and $(c \approx d) \in \overline{\Sigma}$ for $c, d \in C$, then $(t \approx t') \in \overline{\Sigma}$, where t' is the basic term obtained from t by replacing one occurrence of c with d.

These closure conditions may be interpreted as the inference rules of a proof calculus for basic identities. Therefore, write $\Sigma \vdash_X \varphi$ if φ belongs to the weak closure of Σ in the variables X. If the set X is large enough, the relation $\vdash_X$ captures $\models$ for basic identities, as we will prove in Theorem 4.1. We define X to be *large enough* if:

(a) X contains at least two variables;
(b) $|X| \geq \operatorname{arity}(F)$ for any function symbol F occurring in Σ; and
(c) $|X|$ is at least as large as the number of distinct variables occurring in any identity in $\Sigma \cup \{\varphi\}$.

Call Σ *inconsistent relative to* X if $\Sigma \vdash_X x \approx y$ for distinct $x, y \in X$ and large enough X. Otherwise Σ is *consistent relative to* X.

THEOREM 4.1 [22]. *Let $\Sigma \cup \{\varphi\}$ be a set of basic identities and X be a set of variables that is large enough. If Σ is consistent relative to X, then $\Sigma \vdash_X \varphi$ if and only if $\Sigma \models \varphi$.*

Kelly's theorem is a natural restriction of Birkhoff's completeness theorem for equational logic to the special case of basic identities. However, it is in general undecidable for finite $\Sigma \cup \{\varphi\}$ whether $\Sigma \vdash \varphi$ using Birkhoff's inference rules, while it is decidable for basic identities using Kelly's restricted rules[1].

[1]The reason that $\Sigma \vdash_X \varphi$ is decidable with Kelly's inference rules when $\Sigma \cup \{\varphi\}$ is finite is that deciding $\Sigma \vdash_X \varphi$ amounts to generating $\overline{\Sigma}$. If $\mathcal{L}$ is the language whose function and constant symbols are those occurring in $\Sigma \cup \{\varphi\}$, X is a minimal (finite) set of variables that is large enough, and $\mathcal{T}$ is defined to be the set of basic $\mathcal{L}$-terms in the variables X, then generating $\overline{\Sigma}$ amounts to generating an equivalence relation on the finite set $\mathcal{T}$ using Kelly's inference rules.

In the proof we use a variation of Kelly's Rule (iv): rather than use functions $\gamma\colon X \to X \cup C$ for substitutions we will use functions $\Gamma\colon X \cup C \to X \cup C$ whose restriction to C is the identity. (That is, we replace γ with $\Gamma := \gamma \cup \mathrm{id}|_C$.)

Lemma 4.2. *If $\Sigma \vdash_X x \approx h$ for some basic term h in which x does not occur, then Σ is inconsistent relative to any set X containing a variable other than x.*

Proof. Append to a Σ-proof of $x \approx h$ the formulas $(y \approx h)$ for some $y \in X\backslash\{x\}$ (Rule (iv)); $(h \approx y)$ (Rule (ii)); and $(x \approx y)$ (Rule (iii)). □

Proof of Theorem 4.1. Kelly's inference rules are sound, since they are instances of Birkhoff's inference rules for equational logic. Hence, $\Sigma \vdash_X \varphi$ implies $\Sigma \models \varphi$ for any X.

Now assume that $\Sigma \nvdash_X \varphi$, where X is large enough and Σ is consistent relative to X. We construct a model of $\Sigma \cup \{\neg\varphi\}$ to show that $\Sigma \not\models \varphi$. Let $\mathcal{T}$ be the set of basic $\mathcal{L}$-terms in the variables X, and let $\equiv$ be the equivalence relation on $\mathcal{T}$ defined by Kelly provability: that is, $s \equiv t$ if and only if $\Sigma \vdash_X s \approx t$. Write $[t]$ for the $\equiv$-class of t. Now extend $\mathcal{T}$ to a set $\mathcal{T}_0 = \mathcal{T} \cup \{0\}$, where 0 is a new symbol, and extend $\equiv$ to this set by taking the equivalence class of 0 to be $\{0\}$.

The universe of the model will be the set $M := \mathcal{T}_0/\!\equiv$ of equivalence classes of $\mathcal{T}_0$ under $\equiv$. We interpret a constant symbol c as the element $c^{\mathbf{M}} := [c] \in M$. Now let F be an m-ary function symbol for some $m > 0$. The natural idea for interpreting F as an m-ary operation on this set is to define $F^{\mathbf{M}}([a_1], \dots, [a_m]) = [F(a_1, \dots, a_m)]$. However, this does not work, since $F(a_1, \dots, a_m)$ will not be a basic term unless all the a_i belong to $X \cup C$. Nevertheless, we shall follow this idea as far as it takes us, and when we cannot apply it to assign a value to $F^{\mathbf{M}}([a_1], \dots, [a_m])$ we shall assign the value $[0]$.

Choose and fix a well-order $<$ of the set C of constant symbols of $\mathcal{L}$. Let $\mathcal{I}$ be the set of injective partial functions $\iota\colon M \to X \cup C$ that satisfy the following conditions:

(1) if a class $[t] \in M$ in the domain of ι contains a constant symbol, $c \in C$, then $\iota[t] = d$, where $d \in C$ is the least element in $[t] \cap C$ under $<$;
(2) if a class $[t]$ in the domain of ι contains a variable, $x \in X$, then $\iota[t] = x$;
(3) if a class $[t]$ in the domain of ι fails to contain a variable or constant symbol, then $\iota[t] \in X$.

According to Lemma 4.2, the consistency of Σ implies that any class $[t]$ contains at most one variable and, if $[t]$ contains a constant symbol, then $[t]$ contains no variable. Hence, there is no ambiguity in conditions (1) and (2).

If $S \subseteq M$ has size at most $|X|$, then S is the domain of some $\iota \in \mathcal{I}$.

If $S \subseteq M$ and a class $[t] \in S$ contains a variable x, then call x a *fixed* variable of S. Any other variable is an *unfixed* variable of S.

Now we define how to interpret an m-ary function symbol F as an m-ary operation on the set M. Choose any $([a_1], \dots, [a_m]) \in M^m$ and then choose $\iota \in \mathcal{I}$ that is defined on $S := \{[a_1], \dots, [a_m]\}$. Note that $f := F(\iota[a_1], \dots, \iota[a_m])$ is a basic term, since it is a function symbol applied to elements of $X \cup C$. We refer to this term f to define $F^{\mathbf{M}}([a_1], \dots, [a_m])$.

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

Case 1. (The class $[f]$ contains a term h whose only variables are among the fixed variables of S.) Define $F^{\mathbf{M}}([a_1], \ldots, [a_m]) = [f]$.

Case 2. ($[f]$ contains a variable.) If x is a variable in $[f]$, then $\Sigma \vdash_X f \approx x$. Since Σ is consistent, Lemma 4.2 proves that x must occur in f, that is, $x = \imath[a_k]$ for some k. Hence,

$$\Sigma \vdash_X F(\imath[a_1], \ldots, \imath[a_m]) \approx \imath[a_k]$$

for some k. In this case, we define $F^{\mathbf{M}}([a_1], \ldots, [a_m]) = [a_k]$ $(= \imath^{-1}(x).)$

Case 3. (The remaining cases.) Define $F^{\mathbf{M}}([a_1], \ldots, [a_m]) = [0]$.

Before proceeding, we point out that there is overlap in Cases 1 and 2, but no conflict in the definition of $F^{\mathbf{M}}([a_1], \ldots, [a_m])$. If $[f]$ contains a term h whose variables are fixed variables of S and $[f]$ also contains a variable x, then $\Sigma \vdash_X f \approx x$ and $\Sigma \vdash_X h \approx x$. The consistency of Σ forces x to be a common variable of f and h and (since only fixed variables of S occur in h) to be a fixed variable of S. In this situation, Case 1 defines $F^{\mathbf{M}}([a_1], \ldots, [a_m]) = [f] = [x]$ while Case 2 defines $F^{\mathbf{M}}([a_1], \ldots, [a_m]) = \imath^{-1}(x) = [x]$.

Claim 4.3. *The function $F^{\mathbf{M}} : M^m \to M$ is a well-defined function.*

Proof of Claim. Choose $([a_1], \ldots, [a_m]) \in M^m$ and define $S = \{[a_1], \ldots, [a_m]\}$. There exist elements of $\mathcal{I}$ defined on S, because this set has size $\leq \operatorname{arity}(F) \leq |X|$. Suppose that $\imath, \jmath \in \mathcal{I}$ are both defined on this set. Let $f = F(\imath[a_1], \ldots, \imath[a_m])$ and $g = F(\jmath[a_1], \ldots, \jmath[a_m])$. To show that $F^{\mathbf{M}}([a_1], \ldots, [a_m])$ is uniquely defined, it suffices to show that the same value is assigned whether we refer to the term f or the term g.

In all cases of the definition of $F^{\mathbf{M}}([a_1], \ldots, [a_m])$, the assigned value depends only on the term $f = F(\imath[a_1], \ldots, \imath[a_m]) = F(\imath|_S[a_1], \ldots, \imath|_S[a_m])$. Thus, to complete the proof of Claim 4.3, we may replace both $\imath$ and $\jmath$ by $\imath|_S$ and $\jmath|_S$ and assume that $\imath$ and $\jmath$ have domain S. Now $\imath$ and $\jmath$ are injective functions from S into $X \cup C$, and $\imath[t] = \jmath[t]$ whenever $[t] \in S$ and $[t]$ contains a constant symbol or a fixed variable of S. When $\imath[t] \neq \jmath[t]$, then both are unfixed variables of S. In this situation, there is a function $\Gamma : X \cup C \to X \cup C$ that is the identity on C and on the fixed variables of S for which $\jmath = \Gamma \circ \imath$. Hence, $f[\Gamma] = g$ and, if h is a term whose only variables are fixed variables of S, then $h[\Gamma] = h$.

Case 1. ($[f]$ contains a term h whose only variables are among the fixed variables of S.) Here $\Sigma \vdash_X f = F(\imath[a_1], \ldots, \imath[a_m]) \approx h$. Append to a Σ-proof of $f \approx h$ the formula $f[\Gamma] \approx h[\Gamma]$ (Rule (iv)). Since $f[\Gamma] = g$ and $h[\Gamma] = h$, this is a proof of $g \approx h$. Next append $h \approx g$ (Rule (ii)) and $f \approx g$ (Rule (iii)). We conclude that $[f] = [g]$, so the value $[f]$ assigned to $F^{\mathbf{M}}([a_1], \ldots, [a_m])$ using $\imath$ is the same as the value $[g]$ assigned using $\jmath$.

Case 2. ($[f]$ contains a variable.) If $x \in X$ is a variable in $[f]$, then $x = \imath[a_k]$ for some k and $\Sigma \vdash_X F(\imath[a_1], \ldots, \imath[a_m]) \approx \imath[a_k]$ for this k. Append to a Σ-proof of $f \approx x$ the formula $f[\Gamma] \approx x[\Gamma]$ (Rule (iv)). Since $f[\Gamma] = g$ and $x[\Gamma] = \jmath[a_k]$, we conclude

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

that $\Sigma \vdash_X F(\jmath[a_1], \dots, \jmath[a_m]) \approx \jmath[a_k]$ for the same k. Whether we use $\imath$ or $\jmath$, we get $F^{\mathbf{M}}([a_1], \dots, [a_m]) = [a_k]$.

Case 3. (The remaining cases.) In Case 1 we showed that $[f] = [g]$ while in Case 2 we showed that if x is a variable in $[f]$, then $x[\Gamma]$ is a variable in $[g]$; together these show that if $[g]$ does not contain a variable nor a term whose only variables are among the fixed variables of S, then the same is true of $[f]$. This argument works with f and g interchanged, so the remaining cases are those where both $[f]$ and $[g]$ contain no variables nor terms whose only variables are among the fixed variables of S. Whether we use $\imath$ or $\jmath$, we get $F^{\mathbf{M}}([a_1], \dots, [a_m]) = [0]$. □

M is defined. We now argue that **M** is a model of Σ. Choose an identity $(s \approx t) \in \Sigma$. If s is an n-ary function symbol F followed by a sequence $\alpha\colon [n] \to X \cup C$ of length n consisting of variables and constant symbols, then let $F[\alpha]$ be an abbreviation for s. If s is a variable or constant symbol, then s determines a function $\alpha\colon [1] \to X \cup C\colon 1 \mapsto s$, so abbreviate s by $\diamond[\alpha]$. We will, in fact, write s as $F[\alpha]$ in either case, but will remember that F may equal the artificially introduced symbol $\diamond$. The identity $s \approx t$ takes the form $F[\alpha] \approx G[\beta]$.

A valuation in **M** is a function $\nu\colon X \cup C \to M$ satisfying $\nu(c) = c^{\mathbf{M}} = [c]$ for each $c \in C$. To show that **M** satisfies $F[\alpha] \approx G[\beta]$, we must show that $F^{\mathbf{M}}[\nu \circ \alpha] = G^{\mathbf{M}}[\nu \circ \beta]$ for any valuation ν. Choose $\imath \in \mathcal{I}$ that is defined on the set $\mathrm{im}(\nu \circ \alpha) \cup \mathrm{im}(\nu \circ \beta)$. This is possible, since we assume that $|X|$ is at least as large as the number of distinct variables in the identity $F[\alpha] \approx G[\beta] \in \Sigma$. The values of $F^{\mathbf{M}}[\nu \circ \alpha]$ and $G^{\mathbf{M}}[\nu \circ \beta]$ are defined in reference to the terms $f := F[\imath \circ \nu \circ \alpha]$ and $g := G[\imath \circ \nu \circ \beta]$, respectively.

Claim 4.4. $[f] = [g]$.

Proof of Claim. Observe that $(\imath \circ \nu)(c) = \imath[c] = d$, where $d \in C$ is the $<$-least constant symbol in the class $[c]$. If $\Gamma\colon X \cup C \to X \cup C$ is a function that agrees with $(\imath \circ \nu)$ on the variables in $\mathrm{im}(\alpha) \cup \mathrm{im}(\beta)$, but is the identity on C, then applications of Rule (v) show that $\Sigma \vdash_X F[\imath \circ \nu \circ \alpha] \approx F[\Gamma \circ \alpha]$ and $\Sigma \vdash_X G[\imath \circ \nu \circ \beta] \approx G[\Gamma \circ \beta]$. From Rule (iv), the fact that $(F[\alpha] \approx G[\beta]) \in \Sigma$ implies that $\Sigma \vdash_X F[\Gamma \circ \alpha] \approx G[\Gamma \circ \beta]$. Hence,

$$\Sigma \vdash_X f = F[\imath \circ \nu \circ \alpha] \approx F[\Gamma \circ \alpha] \approx G[\Gamma \circ \beta] \approx G[\imath \circ \nu \circ \beta] = g,$$

from which we get $[f] = [g]$. □

We conclude the argument that **M** satisfies $F[\alpha] \approx G[\beta]$ as follows.

Case 1. ($[f] = [g]$ contains a term h whose only variables are among the fixed variables of S.) In this case, $F^{\mathbf{M}}[\nu \circ \alpha] = [f] = [g] = G^{\mathbf{M}}[\nu \circ \beta]$.

Case 2. ($[f] = [g]$ contains a variable.) If $[f] = [x] = [g]$, then $F^{\mathbf{M}}[\nu \circ \alpha] = \imath^{-1}(x) = G^{\mathbf{M}}[\nu \circ \beta]$.

Case 3. (The remaining cases with $[f] = [g]$.) $F^{\mathbf{M}}[\nu \circ \alpha] = [0] = G^{\mathbf{M}}[\nu \circ \beta]$.

Hence, **M** is a model of Σ.

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

To complete the proof of the theorem, we must show that **M** does not satisfy φ. Suppose that φ has the form $F[\alpha] \approx G[\beta]$. Let ν be the canonical valuation

$$X \cup C \to M\colon x \mapsto [x], \quad c \mapsto [c].$$

Choose $\iota \in \mathcal{I}$ that is defined on $\mathrm{im}(\nu \circ \alpha) \cup \mathrm{im}(\nu \circ \beta)$. It follows from the definitions that $\iota \circ \nu\colon X \cup C \to X \cup C$ fixes every variable in $\mathrm{im}(\alpha) \cup \mathrm{im}(\beta)$, while $(\iota \circ \nu)(c) = d$ is the <-least constant symbol in the class of c. If Γ is the identity function on $X \cup C$, then, just as in the proof of Claim 4.4, we obtain $\Sigma \vdash_X f = F[\iota \circ \nu \circ \alpha] \approx F[\Gamma \circ \alpha] = F[\alpha]$ and $\Sigma \vdash_X g = G[\iota \circ \nu \circ \beta] \approx G[\Gamma \circ \beta] = G[\beta]$. Now $[f]$ contains a term $h := F[\alpha]$ whose only variables are among the fixed variables of $S = \mathrm{im}(\alpha)$, so we are in Case 1 of the definition of $F^{\mathbf{M}}$. Hence, $F^{\mathbf{M}}(\nu \circ \alpha) = [F[\alpha]]$ and similarly $G^{\mathbf{M}}(\nu \circ \beta) = [G[\beta]]$. Part of our assumption about $\varphi = (F[\alpha] \approx G[\beta])$ is that $\Sigma \nvdash_X \varphi$, so $[F[\alpha]]$ and $[G[\beta]]$ are distinct elements of M. Therefore, ν witnesses that **M** does not satisfy φ. □

Theorem 4.1 establishes that if X and Y are two sets of variables that are large enough, then $\Sigma \vdash_X \varphi$ holds if and only if $\Sigma \vdash_Y \varphi$ and hence Σ is consistent relative to X if and only if it is consistent relative to Y. Now that the theorem is proved, we drop the subscript in $\vdash_X$ and the phrase 'relative to X' when writing about provability.

4.2. The model V. Later in the paper we prove theorems about finite algebras realizing a set Σ of basic identities. For this, we need to be able to construct finite models of Σ. The model constructed in Theorem 4.1 may be infinite, so we explain how to produce finite models.

Definition 4.5. Let Σ be a set of basic identities in a language $\mathcal{L}$ whose set of constant symbols is C. Let Y be a set of variables, z a variable not in Y, and X a large enough set of variables containing $Y \cup \{z\}$. Let V be the subset of the model **M** constructed in the proof of Theorem 4.1 consisting of

$$\{[y] \mid y \in Y\} \cup \{[c] \mid c \in C\} \cup \{[0]\}.$$

Write $[Y]$ for $\{[y] \mid y \in Y\}$ and $[C]$ for $\{[c] \mid c \in C\}$.

As in the proof of Theorem 4.1, let < be a well-ordering of C. If F is an m-ary function symbol of $\mathcal{L}$ and $([a_1], \ldots, [a_m]) \in V^m$, then let:

(1) $\iota[a_k] = d$ if $[a_k] \in [C]$ and d is the <-least element of $[a_k] \cap C$;
(2) $\iota[a_k] = y$ if $[a_k] = [y] \in [Y]$; and
(3) $\iota[a_k] = z$ if $[a_k] = [0]$.

Define $F^{\mathbf{V}}([a_1], \ldots, [a_m]) = [t]$ if there exists $t \in Y \cup C$ such that

$$\Sigma \vdash F(\iota[a_1], \ldots, \iota[a_m]) \approx t, \tag{4.1}$$

and define $F^{\mathbf{V}}([a_1], \ldots, [a_m]) = [0]$ if there is no such t.

The algebra **V** is the algebra with universe V equipped with all operations of the form $F^{\mathbf{V}}$.

Theorem 4.6. *The algebra* **V** *is a model of* Σ.

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

Proof. Let $F[\alpha] \approx G[\beta]$ be an identity in Σ and let $\nu\colon X \cup C \to V$ be a valuation. We must show that $F^{\mathbf{V}}(\nu \circ \alpha) = G^{\mathbf{V}}(\nu \circ \beta)$.

The function ν is also a valuation in $\mathbf{M}$, because $V \subseteq M$. Since $\mathbf{M}$ is a model of Σ, we get $F^{\mathbf{M}}[\nu \circ \alpha] = G^{\mathbf{M}}[\nu \circ \beta]$. Choose $\iota \in \mathcal{I}$ defined on the set $\mathrm{im}(\nu \circ \alpha) \cup \mathrm{im}(\nu \circ \beta)$ such that $\iota[0] = z$, if $[0]$ is in this set. Let $f = F[\iota \circ \nu \circ \alpha]$ and $g = G[\iota \circ \nu \circ \beta]$. As in the proof of Claim 4.4, if $\Gamma\colon X \cup C \to X \cup C$ is the identity on C and agrees with $\iota \circ \nu$ on the variables in $\mathrm{im}(\alpha) \cup \mathrm{im}(\beta)$, then $\Sigma \vdash f \approx F[\Gamma \circ \alpha] \approx G[\Gamma \circ \beta] \approx g$.

The term $F(\iota[a_1], \ldots, \iota[a_m])$ from line (4.1) is none other than f. $F^{\mathbf{V}}[\nu \circ \alpha] = [t]$ for some $t \in Y \cup C$ if and only if $\Sigma \vdash f = F(\iota[a_1], \ldots, \iota[a_m]) \approx t$. But, since $\Sigma \vdash f \approx g$, we also get $G^{\mathbf{V}}[\nu \circ \beta] = [t]$. This shows that $F^{\mathbf{V}}[\nu \circ \alpha]$ and $G^{\mathbf{V}}[\nu \circ \beta]$ are equal when at least one of them is not $[0]$. Of course, they are also equal when both of them equal $[0]$, so $F^{\mathbf{V}}(\nu \circ \alpha) = G^{\mathbf{V}}(\nu \circ \beta)$. □

Corollary 4.7. *If Σ is a consistent set of basic identities in a language whose set of constant symbols is C, then Σ has models of every cardinality strictly exceeding $|C|$.*

Proof. Vary the size of Y in the definition of $\mathbf{V}$, and use Theorem 4.6. □

Corollary 4.7 is close to the best possible result about sizes of models of a set of basic identities, as the next example shows.

Example 4.8. Let C be a set of constant symbols and let $\mathcal{B} = \{B_{c,d} \mid c, d \in C, c \neq d\}$ be a set of binary function symbols. Let

$$\Sigma = \{B_{c,d}(c, x) \approx x, B_{c,d}(d, x) \approx d \mid c, d \in C, c \neq d\}.$$

The set Σ is a consistent set of basic identities, since if A is any set containing C we can interpret each $c \in C$ in A as itself and each $B_{c,d}$ on A by letting $B^{\mathbf{A}}_{c,d}(c, y) = y$ and $B^{\mathbf{A}}_{c,d}(x, y) = x$ if $x \neq c$.

If $\mathbf{M}$ is any model of Σ and $c^{\mathbf{M}} = d^{\mathbf{M}}$ for some $c, d \in C$, then the identity function $B^{\mathbf{M}}_{c,d}(c^{\mathbf{M}}, x)$ equals the constant function $B^{\mathbf{M}}_{c,d}(d^{\mathbf{M}}, x)$, so $|M| = 1$. Thus, elements of C must have distinct interpretations in any nontrivial model of Σ, implying that nontrivial models have size at least $|C|$.

5. Restrictive Σ

Call a set Σ of basic identities *nonrestrictive* if, whenever $\mathbf{A}$ is an algebra of more than one element, there is an algebra $\mathbf{B}$ realizing Σ such that $d_{\mathbf{B}}(n) = d_{\mathbf{A}}(n)$ for all $n > 0$. Otherwise Σ is *restrictive*.

Call Σ *nonrestrictive for finite algebras* if, whenever $\mathbf{A}$ is a finite algebra of more than one element, there is a finite algebra $\mathbf{B}$ realizing Σ such that $d_{\mathbf{B}}(n) = d_{\mathbf{A}}(n)$ for all $n > 0$. Otherwise Σ is *restrictive for finite algebras*.

It is possible for Σ to be nonrestrictive, yet restrictive for finite algebras. The set Σ from Example 4.8 has this property when the set of constants is infinite (cf. Remark 5.4). But the concepts defined in the two preceding paragraphs are close enough that the arguments of this section apply equally well to both of them. We

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

will see that if only finitely many constant symbols appear in Σ, then Σ is restrictive if and only if it is restrictive for finite algebras. Both are equivalent to the property that Σ entails the existence of a pointed cube term.

Recall from the introduction that an m-ary p-pointed k-cube term for the variety axiomatized by Σ is an m-ary term $F(x_1, \ldots, x_m)$ for which there is a $k \times m$ matrix $M = [y_{i,j}]$ of variables and constant symbols, where every column contains a symbol different from x, such that Σ proves the identities

$$F(\mathbf{y}_1, \ldots, \mathbf{y}_m) = F\left(\begin{bmatrix} y_{1,1} \\ \vdots \\ y_{k,1} \end{bmatrix}, \ldots, \begin{bmatrix} y_{1,m} \\ \vdots \\ y_{k,m} \end{bmatrix}\right) \approx \begin{bmatrix} x \\ \vdots \\ x \end{bmatrix}.$$

If Σ is consistent, the parameters are constrained by $m, k \geq 2$ and $p \geq 0$.

In Section 5.1, we prove that if Σ is restrictive, then it entails the existence of a pointed cube term. The converse is proved in Section 5.2, by showing that an algebra with a pointed cube term whose d-function assumes only finite values has growth rate that is bounded above by a polynomial. In particular, it is shown that a finite algebra $\mathbf{A}$ with a 1-pointed k-cube term satisfies $d_{\mathbf{A}}(n) \in O(n^{k-1})$. In Section 5.3, we give an example of a three-element algebra with a 1-pointed k-cube term whose growth rate satisfies $d_{\mathbf{A}}(n) \in \Theta(n^{k-1})$, showing that the preceding estimate is sharp. In Section 5.4, we show that any function $D\colon \mathbf{Z}^+ \to \mathbf{Z}^{\geq 0}$ that occurs as the d-function of an algebra with a pointed cube term also occurs as the d-function of an algebra that does not have a pointed cube term. In Subsection 5.5, we describe one way of showing that an algebra has an exponential growth rate, and we use it to exhibit a variety containing a chain of finite algebras $\mathbf{A}_1 \leq \mathbf{A}_2 \leq \cdots$, each one a subalgebra of the next, where $\mathbf{A}_i$ has logarithmic growth when i is odd and exponential growth when i is even.

5.1. Restrictive Σ forces a pointed cube term. Let Σ be a consistent set of basic identities in a language $\mathcal{L}$ whose set C of constant symbols is finite. Given an algebra $\mathbf{A}$ in a language disjoint from Σ, we construct another algebra $\mathbf{A}_\Sigma$ which realizes Σ, where $\mathbf{A}_\Sigma$ is finite if $\mathbf{A}$ is.

For the first step, let $[C] = \{[c_1], \ldots, [c_p]\}$ be the same set of equivalence classes denoted by $[C]$ in Definition 4.5. These classes represent the different Σ-provability classes of constant symbols. If there are p such classes, then apply the one-point completion construction of Section 3.1 $p + 1$ times to $\mathbf{A}$ to produce a sequence $\mathbf{A}$, $\mathbf{A}_{z_1}$, $\mathbf{A}_{z_1,z_2}, \ldots,$ ending at $\mathbf{A}_{z_1,\ldots,z_p,0}$. This is an algebra whose universe is the disjoint union of A and $\{z_1, \ldots, z_p, 0\}$.

$\mathbf{A}_\Sigma$ will be an expansion of $\mathbf{A}_{z_1,\ldots,z_p,0}$ obtained by merging the latter algebra with the model $\mathbf{V}$ introduced in Definition 4.5. Let Y be a set of variables satisfying $|Y| = |A|$, and let $[Y] = \{[y] \mid y \in Y\}$ be the set of equivalence classes also denoted by $[Y]$ in Definition 4.5. The consistency of Σ guarantees that each equivalence class $[y]$ is a singleton. The universe of $\mathbf{V}$ is the disjoint union $V = [Y] \cup [C] \cup \{[0]\}$.

Let $\varphi\colon [Y] \to A$ be a bijection. Extend this to a bijection from $V = [Y] \cup [C] \cup \{[0]\}$ to $A \cup \{z_1, \ldots, z_p\} \cup \{0\}$ by defining $\varphi([c_i]) = z_i$ and $\varphi([0]) = 0$. Now φ is a bijection from the universe of $\mathbf{V}$ to the universe of $\mathbf{A}_{z_1,\ldots,z_p,0}$. Use this bijection to transfer the operations of $\mathbf{V}$ over to $\mathbf{A}_{z_1,\ldots,z_p,0}$ to create $\mathbf{A}_\Sigma$. Specifically, the interpretation of the constant symbol c_i in $\mathbf{A}_\Sigma$ will be z_i and, if F is an m-ary function symbol of $\mathcal{L}$, then

$$F^{\mathbf{A}_\Sigma}(x_1,\ldots,x_m) := \varphi(F^{\mathbf{V}}(\varphi^{-1}(x_1),\ldots,\varphi^{-1}(x_m))) \tag{5.1}$$

will be the interpretation of the symbol F in $\mathbf{A}_\Sigma$. $\mathbf{A}_\Sigma$ is the expansion of $\mathbf{A}_{z_1,\ldots,z_p,0}$ by all constant operations $c_i^{\mathbf{A}_\Sigma}$ and all operations of the form (5.1). Under this definition, the function φ is an isomorphism from $\mathbf{V}$ to the $\mathcal{L}$-reduct of $\mathbf{A}_\Sigma$.

Lemma 5.1. *Let $\mathbf{A}$ be an algebra with more than one element and let Σ be a consistent set of basic identities involving finitely many constant symbols. Let $\mathcal{V}$ be the variety axiomatized by Σ. The following statements about an integer $k \geq 2$ are equivalent.*

(1) *$\mathcal{V}$ has a pointed k-cube term.*
(2) *$A_\Sigma^k \backslash A^k$ generates $\mathbf{A}_\Sigma^k$.*
(3) *There is a generating set $G(k)$ of $\mathbf{A}_\Sigma^k$ such that $G(k) \cap A^k$ does not generate $\mathbf{A}^k$.*
(4) *$\mathcal{V}$ has a pointed k-cube term of the form $F(x_1,\ldots,x_m)$, where $m \geq 2$, F is a function symbol occurring in Σ, and the variables $x_1,\ldots,x_m$ are distinct.*

Proof. [(1) ⇒ (2)] Let $F(x_1,\ldots,x_m)$ be a pointed k-cube term of the variety axiomatized by Σ. There is a $k \times m$ matrix $M = [y_{i,j}]$ of variables and $\mathcal{L}$-constant symbols, where every column contains a symbol different from x, such that Σ proves the identities

$$F(\mathbf{y}_1,\ldots,\mathbf{y}_m) = F\left(\begin{bmatrix} y_{1,1} \\ \vdots \\ y_{k,1}\end{bmatrix},\ldots,\begin{bmatrix} y_{1,m} \\ \vdots \\ y_{k,m}\end{bmatrix}\right) \approx \begin{bmatrix} x \\ \vdots \\ x\end{bmatrix}. \tag{5.2}$$

Choose any tuple $\mathbf{a} \in A_\Sigma^k$. Using the row identities of (5.2), solve the equation $F(\mathbf{b}_1,\ldots,\mathbf{b}_m) = \mathbf{a}$ for the $\mathbf{b}_i$, row by row, according to the following rules. In the ith row:

(a) if $y_{i,j} = x$, then let $b_{i,j} = a_i$;
(b) if $y_{i,j} = c_r$ is a constant symbol, then let $b_{i,j} = z_r$ be its interpretation in $\mathbf{A}_\Sigma$;
(c) if $y_{i,j}$ is a variable different from x, then let $b_{i,j} = 0$.

Under these choices, $\mathbf{b}_i \in A_\Sigma^k$ for all i and $F(\mathbf{b}_1,\ldots,\mathbf{b}_m) = \mathbf{a}$. Moreover, since each column $\mathbf{y}_i$ in (5.2) has a symbol different from x, it follows from (a)–(c) that each $\mathbf{b}_i$ has a coordinate value that is in the set $\{z_1,\ldots,z_p,0\}$. Hence, $\mathbf{b}_i \in A_\Sigma^n \backslash A^k$ for all i. This shows that the arbitrarily chosen tuple $\mathbf{a} \in A_\Sigma^k$ lies in the subalgebra $\mathbf{A}_\Sigma^k$ that is generated by $A_\Sigma^k \backslash A^k$.

[(2) ⇒ (3)] Let $G(n) = A_\Sigma^k \backslash A^k$.

[(3) ⇒ (4)] Let $G(k)$ be the generating set for $\mathbf{A}_\Sigma^k$ that is guaranteed by Item (3). Since $G(k) \cap A^k$ does not generate $\mathbf{A}^n$, it follows from Theorem 3.2 that $(A_\Sigma^k \backslash A^k) \cup G(k)$ is not a generating set for $\mathbf{A}^k_{z_1,\ldots,z_p,0}$. Let S be the proper subuniverse of $\mathbf{A}^k_{z_1,\ldots,z_p,0}$ that is generated by $(A_\Sigma^k \backslash A^k) \cup G(n)$.

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

Since S contains $G(k)$, which generates $\mathbf{A}_\Sigma^k$, and contains the interpretations of the $\mathcal{L}$-constants, it cannot be closed under the interpretations of the function symbols of $\mathcal{L}$. Hence, there are a tuple $\mathbf{a} \notin S$, an m-ary function symbol F, and m tuples $\mathbf{b}_1, \ldots, \mathbf{b}_m \in S$ such that $F^{\mathbf{A}_\Sigma}(\mathbf{b}_1, \ldots, \mathbf{b}_m) = \mathbf{a}$. Necessarily $\mathbf{a} \in A^k$.

Using the isomorphism φ from $\mathbf{V}$ to the $\mathcal{L}$-reduct of $\mathbf{A}_\Sigma$, we obtain that there are a tuple $\mathbf{y} = \varphi^{-1}(\mathbf{a}) \in \varphi^{-1}(A^k) = [Y]^k$ and tuples $\mathbf{v}_i = \varphi^{-1}(\mathbf{b}_i) \neq \mathbf{y}$ such that $F^{\mathbf{V}}(\mathbf{v}_1, \ldots, \mathbf{v}_m) = \mathbf{y}$. Since $\mathbf{v}_1 \neq \mathbf{y}$, there is a coordinate ℓ where these tuples differ. In the ℓth coordinate, we have $F^{\mathbf{V}}([v_{\ell,1}], \ldots, [v_{\ell,m}]) = [y_\ell]$ for some variable $y_\ell \in Y$ and some elements $v_{\ell,j} \in Y \cup C \cup \{0\}$ with $[v_{\ell,1}] \neq [y_\ell]$. By the definition of $\mathbf{V}$,

$$\Sigma \vdash F(\imath[v_{\ell,1}], \ldots, \imath[v_{\ell,m}]) \approx y_\ell, \tag{5.3}$$

where $\imath[v_{\ell,j}] = v_{\ell,j}$ when $v_{\ell,j} \in Y$, $\imath[v_{\ell,j}]$ is a constant Σ-provably equivalent to $v_{\ell,j}$ when $v_{\ell,j} \in C$, and $\imath[v_{\ell,j}] = z$ is a variable not in Y when $v_{\ell,j} = 0$. Since $[v_{\ell,1}] \neq [y_\ell]$, we have $v_{\ell,1} \neq y_\ell$. After renaming variables, (5.3) can be rewritten as

$$\Sigma \vdash F(w_{\ell,1}, \ldots, w_{\ell,m}) \approx x,$$

where each $w_{\ell,j}$ is a variable or constant and $w_{\ell,1} \neq x$.

Similarly, for each $j \leq m$, the fact that $\mathbf{v}_j \neq \mathbf{y}$ produces an identity

$$\Sigma \vdash F(w_{i,1}, \ldots, w_{i,m}) \approx x,$$

where each $w_{p,q}$ is a variable or constant and $w_{i,j} \neq x$. Each of these identities may be read off of one of the rows of $F^{\mathbf{A}_\Sigma}(\mathbf{b}_1, \ldots, \mathbf{b}_m) = \mathbf{a}$, leading to a $k \times m$ matrix identity

$$F(\mathbf{w}_1, \ldots, \mathbf{w}_m) = F\left(\begin{bmatrix} w_{1,1} \\ \vdots \\ w_{k,1} \end{bmatrix}, \ldots, \begin{bmatrix} w_{1,m} \\ \vdots \\ w_{k,m} \end{bmatrix}\right) \approx \begin{bmatrix} x \\ \vdots \\ x \end{bmatrix},$$

where each $w_{p,q}$ is a variable or constant, and each $\mathbf{w}_j$ contains a symbol different from x. These identities make F a pointed k-cube term for $\mathcal{V}$.

[(4) $\Rightarrow$ (1)] This is a tautology. □

If Item (1) of Lemma 5.1 holds for some k, then it holds for each $k' \geq k$, since a pointed k-cube term is also a pointed k'-cube term. (If M is a $k \times m$ matrix of variables and constants with each column containing a symbol different from x, and F is such that $\Sigma \vdash F(M) \approx [x, \ldots, x]^T$, then one can construct a similar $k' \times m$ matrix M' for F by duplicating some rows in M.) It follows that if any of the items of Lemma 5.1 holds for some k, then all hold for al $k' \geq k$.

One consequence of Lemma 5.1 is a procedure to decide if a strong Maltsev condition involving only basic identities implies the existence of a pointed cube term.

Corollary 5.2. *A strong Maltsev condition defined by a set Σ of basic identities entails the existence of a pointed k-cube term if and only if it is possible to prove from Σ that some term of the form $F(x_1, \ldots, x_m)$ is a pointed k-cube term, where $m \geq 2$, F is a function symbol occurring in Σ, and the variables $x_1, \ldots, x_m$ are distinct.*

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

Proof. A strong Maltsev condition defined by a set Σ of identities entails the existence of a pointed k-cube term if and only if the variety axiomatized by Σ has a pointed k-cube term, so the corollary follows from Lemma 5.1(1) $\Leftrightarrow$ (4). □

That the property in the statement of the corollary can be decided follows from Theorem 4.1.

The next result is the main one of the subsection.

Theorem 5.3. *Let Σ be a consistent set of basic identities involving finitely many constant symbols. If Σ does not entail the existence of a pointed cube term, then Σ is nonrestrictive (and also nonrestrictive for finite algebras).*

Proof. Recall that 'Σ is nonrestrictive' means that for every algebra $\mathbf{A}$ of more than one element, there is an algebra $\mathbf{B}$ realizing Σ such that $d_{\mathbf{B}}(n) = d_{\mathbf{A}}(n)$ for $n > 0$. The phrase 'Σ is restrictive' means the opposite.

Assume that Σ fails to entail the existence of a pointed cube term. Choose any $\mathbf{A}$ of more than one element and let $\mathbf{B} = \mathbf{A}_{\Sigma}$. $\mathbf{B}$ realizes Σ because $\mathbf{V}$ is a reduct of $\mathbf{B}$ and a model of Σ. We argue that $d_{\mathbf{B}}(n) = d_{\mathbf{A}}(n)$ for $n > 0$.

First assume that $n \geq 2$. Choose a generating set G for $\mathbf{B}^n$ such that $|G| = d_{\mathbf{B}}(n)$. By Lemma 5.1(3) $\Rightarrow$ (4), we get that $G \cap A^n$ is a generating set for $\mathbf{A}^n$, so $d_{\mathbf{A}}(n) \leq |G \cap A^n| \leq d_{\mathbf{B}}(n)$.

The argument just given also works when $n = 1$. To see this, re-examine the proof of Lemma 5.1(3) $\Rightarrow$ (4) in the case $n = k = 1$. We get that if Lemma 5.1(3) holds in this case, then $\mathcal{V}$ has a basic term involving a nonnullary operation symbol F from Σ, which is a '1-cube term'. We take the quoted phrase to mean a term where $\Sigma \vdash F(w_1, \ldots, w_m) \approx x$, where each w_i is a constant or a variable not equal to x. This can only happen if Σ is inconsistent. Thus, we can use our assumption that Σ is consistent to derive the nonexistence of '1-cube terms', and thereby derive the failure of Lemma 5.1(3) in the case $n = k = 1$. Hence, if G is a generating set for $\mathbf{B}$, then $G \cap A$ is also a generating set for $\mathbf{A}$. Now the proof that $d_{\mathbf{A}}(n) \leq d_{\mathbf{B}}(n)$ in the case $n = 1$ is the same as the one above for the cases $n \geq 2$.

For the reverse inequality, choose a generating set H for $\mathbf{A}^n$ such that $|H| = d_{\mathbf{A}}(n)$. Repeated use of Theorem 3.2(1) shows that H generates $\mathbf{A}^n_{z_1,\ldots,z_p,0}$ and hence also generates $\mathbf{A}^n_{\Sigma} = \mathbf{B}$. This shows that $d_{\mathbf{B}}(n) \leq |H| = d_{\mathbf{A}}(n)$. □

Remark 5.4. In the third paragraph of this section, we stated that the set Σ from Example 4.8 is nonrestrictive, yet restrictive for finite algebras when Σ involves infinitely many constants. Here we explain why this remark is true, and also explain to what degree we may remove the assumption of finitely many constants in Theorem 5.3.

Let Σ be as in Example 4.8 with C an infinite set of constants. Let $\mathbf{A}$ be any finite algebra. There is no finite $\mathbf{B}$ that realizes Σ and hence none that realizes Σ and satisfies $d_{\mathbf{B}}(n) = d_{\mathbf{A}}(n)$ for $n > 0$, since any nontrivial model of Σ has cardinality at least $|C|$.

On the other hand, Σ does not entail the existence of a pointed cube term. Without attempting to give the full argument for this, we indicate only that if Σ entailed the existence of a pointed cube term, then (i) one would have the form $B_{c,d}(x_1, x_2)$, by the

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

appropriate generalization of Lemma 5.1, (ii) it could not be a projection, so we would have to have $\Sigma \vdash B_{c,d}(e, x) \approx x$ and $\Sigma \vdash B_{c,d}(x, f) \approx x$ for some constants e and f, and (iii) $\{B_{c,d}(x, f)\}$ is a singleton class of the weak closure of Σ; hence, we do not have $\Sigma \vdash B_{c,d}(x, f) \approx x$ after all.

Finally, we sketch how to modify the proof of Theorem 5.3 to eliminate the restriction to finitely many constants in the case where the algebras may be infinite.

Recall that we started with an algebra $\mathbf{A}$, enlarged it to $\mathbf{A}_{z_1,\ldots,z_p,0}$ by iterating the one-point completion construction, and then merged it with the model $\mathbf{V}$ of Σ to create $\mathbf{A}_\Sigma$, which realized Σ and had the same growth rate as $\mathbf{A}$. In this construction, we used the one-point completion construction p times, where p was the number of equivalence classes of constant symbols under Σ-provable equivalence. The only thing different here is that we may not have finitely many equivalence classes of constant symbols. However, we may well-order the equivalence classes of constants (say, by stipulating that $[c] < [d]$ if the least constant in class $[c]$ is smaller than the least constant in $[d]$ under the well-order from the proof of Kelly's theorem). Now, rather than using the one-point completion construction p times, we use the idea of the construction exactly once to adjoin a well-ordered set $\{0\} \cup Z$ to $\mathbf{A}$ to create $\mathbf{A}_{Z,0}$. Here the well-order is $0 < z_1 < z_2 < \cdots$, with 0 the least element, and $\langle Z; < \rangle$ is a well-ordered set for which there is a bijection $\varphi\colon [C] \to Z$ from the set of equivalence classes of constants. The algebra has universe $A_{Z,0}$ equal to the disjoint union of A, Z, and 0. If F is a function symbol in the language of $\mathbf{A}$, then it is defined on $A_{Z,0}$ by

$$F^{\mathbf{A}_{Z,0}}(\mathbf{a}) = \begin{cases} F^{\mathbf{A}}(\mathbf{a}) & \text{if } \mathbf{a} \in A^n, \\ \min\{\{a_1, \ldots, a_n\} \cap (\{0\} \cup Z)\} & \text{else.} \end{cases}$$

We also define binary operations corresponding to the operation $x \wedge y$ of the one-point completion, namely $x \wedge_z y$ for $z \in Z \cup \{0\}$. Here

$$x \wedge_z y = \begin{cases} x & \text{if } x = y, \\ z & \text{if } x \neq y \text{ and } x, y \in A \cup [z), \\ \min\{\{x, y\} \cap (\{0\} \cup Z)\} & \text{else.} \end{cases}$$

(Here $[z)$ is the principal filter in $\langle Z; < \rangle$ that is generated by z.) Arguments similar to those in Theorem 3.2 show that a generating set for $\mathbf{A}^n$ also generates $\mathbf{A}^n_{Z,0}$ and any generating set for $\mathbf{A}^n_{Z,0}$ contains a generating set for $\mathbf{A}^n$. We can merge $\mathbf{A}_{Z,0}$ with a model $\mathbf{V}$ of Σ from Definition 4.5 to obtain a model $\mathbf{A}_\Sigma$, as we did for the proof of Theorem 5.3. Using the same arguments as before, it can be shown that $\mathbf{A}_\Sigma$ has the same growth rate as $\mathbf{A}$ unless Σ entails the existence of a pointed cube term.

5.2. Pointed cube terms enforce polynomially bounded growth. In the preceding subsection, we proved that if Σ is restrictive, then Σ entails the existence of a pointed cube term. We now prove the converse by showing that if $\mathbf{A}$ is an algebra with a pointed cube term and sufficiently many of the small powers of $\mathbf{A}$ are finitely generated, then all finite powers of $\mathbf{A}$ are finitely generated and $d_{\mathbf{A}}(n)$ is bounded above by a polynomial.

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

Theorem 5.5. *Let* $\mathbf{A}$ *be an algebra with an m-ary p-pointed k-cube term, with at least one constant symbol appearing in the cube identities (so* $p \geq 1$*). If* $\mathbf{A}^{p+k-1}$ *is finitely generated, then all finite powers of* $\mathbf{A}$ *are finitely generated and* $d_{\mathbf{A}}(n)$ *is bounded above by a polynomial of degree at most* $\log_w(m)$*, where* $w = 2k/(2k-1)$.

The proof rests on the fact that a cube term, like

$$F\begin{pmatrix} 1 & x & 2 \\ x & 2 & 3 \end{pmatrix} \approx \begin{pmatrix} x \\ x \end{pmatrix}, \tag{5.4}$$

may be used to 'factor' a typical tuple $\mathbf{a} \in A^n$ into simpler tuples:

$$F\left(\begin{bmatrix} \mathbf{1} \\ \mathbf{a}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{2} \end{bmatrix}, \begin{bmatrix} \mathbf{2} \\ \mathbf{3} \end{bmatrix}\right) = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{bmatrix}.$$

Here the n-tuple $\mathbf{a}$ has been split into two blocks of coordinates of roughly equal size, $\mathbf{a} = \left[\begin{smallmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{smallmatrix}\right]$, and then factored into $\left[\begin{smallmatrix} \mathbf{1} \\ \mathbf{a}_2 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} \mathbf{a}_1 \\ \mathbf{2} \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} \mathbf{2} \\ \mathbf{3} \end{smallmatrix}\right]$, which are simpler than $\mathbf{a}$ in the sense that some of the coordinates have been replaced by constants. This factorization process can be iterated until the final factors have at most $k-1$ coordinate entries that have not been replaced by elements from the set of constants. The proof of the theorem develops such a factorization scheme under which there are only polynomially many different types of final factors, and the collection of all final factors of a given type lie in a subalgebra of $\mathbf{A}^n$ isomorphic to $\mathbf{A}^j$ for some $j \leq p+k-1$. The set consisting of the generators of all of these subalgebras is a polynomial-size generating set for $\mathbf{A}^n$.

Proof. Suppose that the fact that $F(x_1, \ldots, x_m)$ is a p-pointed k-cube term (with $p \geq 1$) is witnessed by identities $F(M) \approx [x, \ldots, x]^{\mathsf{T}}$, where M is a $k \times m$ matrix of variables and constant symbols, with at least one constant symbol, where each column of M contains a symbol that is not x. Choose a constant symbol c appearing in M and replace all instances of variables in M that are not x by c. This produces another matrix R with no variables other than x which also witnesses that F is a p-pointed k-cube term. The order of the k row identities, $F(R) \approx [x, \ldots, x]^{\mathsf{T}}$, is fixed once and for all.

We will refer to the function $\lambda\colon [m] \to [k]$ from the column indices to the row indices defined by the property that $\lambda(j) = i$ exactly when i is the least index such that R has a constant symbol in its (i, j)th position. Such λ exists because every column of R contains at least one constant symbol. (For the cube term in the example immediately following the theorem statement, $\lambda\colon [3] \to [2]$ is the function $\lambda(1) = \lambda(3) = 1$, $\lambda(2) = 2$.)

The factoring, or 'processing', of tuples in A^n will make use of an m-ary tree, which we refer to as the *(processing) template*. We refer to nodes of the template by their *addresses*, which are finite strings in the alphabet $[m] = \{1, \ldots, m\}$. The root node has empty address, and is denoted $\mathbf{n}_\emptyset$. If $\mathbf{n}_\sigma$ is the node at address σ, then its children are the nodes $\mathbf{n}_{\sigma 1}, \ldots, \mathbf{n}_{\sigma m}$.

Each node $\mathbf{n}$ of the template is labeled by a subset $\ell(\mathbf{n}) \subseteq [n]$. (Recall that n is the number appearing in the exponent of A^n.) To define the labeling function ℓ, we first specify a fixed method for partitioning some subsets $U \subseteq [n]$. Given a subset

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

$U = \{u_1, \ldots, u_r\} \subseteq [n]$, consider it to be a linearly ordered set $u_1 < \cdots < u_r$ under the order inherited from $[n]$. Define $\pi(U) = (U_1, \ldots, U_k)$ to be the ordered partition of U into k consecutive nonempty intervals that are as equal in size as possible. In more detail, let

$$\pi(U) = (U_1, \ldots, U_k) = (\{u_1, \ldots, u_{i_1}\}, \{u_{i_1+1}, \ldots, u_{i_2}\}, \ldots, \{u_{i_{k-1}+1}, \ldots, u_{i_k} = u_r\}),$$

where

$$u_1 < \cdots < u_{i_1} < u_{i_1+1} < \cdots < u_{i_2} < u_{i_{k-1}+1} < \cdots < u_{i_k} = u_r$$

(that is, the cells of the partition are consecutive nonempty intervals) and

$$|U_1| \geq \cdots \geq |U_k| \geq |U_1| - 1$$

(that is, the cells are as equal sized as possible). The k appearing here as the number of cells of the partition is the same k as the one in the assumption that F is a k-cube term. In order for $\pi(U)$ to be defined, it is necessary that $|U| \geq k$.

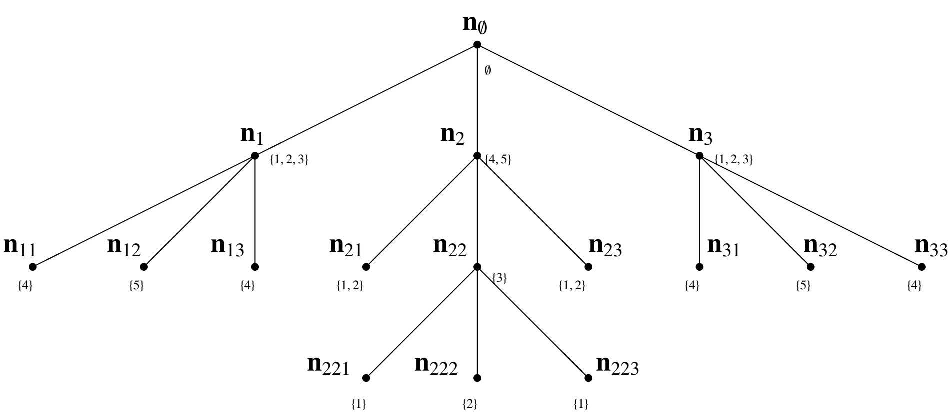
As mentioned earlier, the label on node $\mathbf{n}_\sigma$ will be some subset $\ell(\mathbf{n}_\sigma) \subseteq [n]$. Recursively define the labels as follows:

(1) $\ell(\mathbf{n}_\emptyset) = \emptyset$;

(2) if all nodes between $\mathbf{n}_\sigma$ and $\mathbf{n}_\emptyset$ are labeled, V is the union of labels occurring between $\mathbf{n}_\sigma$ and the root $\mathbf{n}_\emptyset$, and $\pi([n] \backslash V) = (U_1, \ldots, U_k)$, then $\ell(\mathbf{n}_{\sigma i}) = U_{\lambda(i)}$.

In (2), if $[n] \backslash V$ has fewer than k elements, then it is impossible to partition it into k nonempty intervals, in which case there do not exist sufficiently many labels for potential children. In this case, we do not include any descendants of $\mathbf{n}_\sigma$ in the template.

Let us illustrate our progress with the example started back at line (5.4). The following picture depicts the processing template in the case $[n] = [5] = \{1, 2, 3, 4, 5\}$. (Recall that $\lambda(1) = \lambda(3) = 1, \lambda(2) = 2$.)



https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

Now we define precisely what is meant by processing. Let $P = \{c_1, \dots, c_p\}$ be the constant symbols appearing in the cube identities for F. A tuple $\mathbf{a} \in A^n$ is *processed for node* $\mathbf{n}_\sigma$ if there is a constant symbol $c \in P$ such that the ith coordinate of $\mathbf{a}$ is $c^{\mathbf{A}}$ for all $i \in \ell(\mathbf{n}_\sigma)$. A tuple $\mathbf{a}$ is *fully processed* if there is a path through the template from the root to a leaf such that $\mathbf{a}$ is processed for each node in the path.

The processing template describes, in reverse order, a particular way to generate tuples in $\mathbf{A}^n$. Given a tuple $\mathbf{a} \in A^n$, we assign it to the root $\mathbf{n}_\emptyset$ and denote it $\mathbf{a}_\emptyset$. This tuple $\mathbf{a} = \mathbf{a}_\emptyset$ is already processed for $\mathbf{n}_\emptyset$, since this is an empty requirement. Now, for each address σ of a node in the template, we will construct $\mathbf{a}_{\sigma 1}, \dots, \mathbf{a}_{\sigma m}$ from $\mathbf{a}_\sigma$ so that (i) $F^{\mathbf{A}}(\mathbf{a}_{\sigma 1}, \dots, \mathbf{a}_{\sigma m}) = \mathbf{a}_\sigma$ and (ii) each $\mathbf{a}_{\sigma i}$ is processed at all nodes between $\mathbf{n}_{\sigma i}$ and $\mathbf{n}_\emptyset$. Assign $\mathbf{a}_{\sigma i}$ to $\mathbf{n}_{\sigma i}$. The original tuple $\mathbf{a}$ can be generated via $F^{\mathbf{A}}$ by the fully processed tuples derived from $\mathbf{a}$ in this way. The following claim is the heart of this argument.

Claim 5.6. *Suppose that* $\mathbf{n}_\sigma$ *is an internal node of the processing template. Given an arbitrary tuple* $\mathbf{a} \in A^n$*, there exist tuples* $\mathbf{b}_1, \dots, \mathbf{b}_m$ *such that:*

(1) $F^{\mathbf{A}}(\mathbf{b}_1, \dots, \mathbf{b}_m) = \mathbf{a}$*;*

(2) $\mathbf{b}_i$ *is processed for node* $\mathbf{n}_{\sigma i}$ *for* $i = 1, \dots, m$*;*

(3) *if* $\mathbf{n}$ *is a node between* $\mathbf{n}_\sigma$ *and* $\mathbf{n}_\emptyset$*, and* $\mathbf{a}$ *is processed for* $\mathbf{n}$*, then each* $\mathbf{b}_i$ *is also processed for* $\mathbf{n}$ *for* $i = 1, \dots, m$*.*

Proof of Claim. Let V be the union of labels on nodes between $\mathbf{n}_\sigma$ and $\mathbf{n}_\emptyset$. If $\pi([n]\backslash V) = (U_1, \dots, U_k)$, then $\{V, U_1, \dots, U_k\}$ is a partition of $[n]$ (with V possibly empty). For simplicity of expression, re-order coordinates so that $\mathbf{a}$ and $\mathbf{b}_i$ can be written $[\mathbf{a}_V, \mathbf{a}_{U_1}, \dots, \mathbf{a}_{U_k}]^\top$ and $[\mathbf{b}_{i,V}, \mathbf{b}_{i,U_1}, \dots, \mathbf{b}_{i,U_k}]^\top$, with coordinates from V or U_j grouped together. Given $\mathbf{a}$, we need to solve for $\mathbf{b}_{i,V}$ and $\mathbf{b}_{i,U_j}$ in

$$F^{\mathbf{A}}(\mathbf{b}_1, \dots, \mathbf{b}_m) = F^{\mathbf{A}}\left(\begin{bmatrix}\mathbf{b}_{1,V}\\ \mathbf{b}_{1,U_1}\\ \vdots\\ \mathbf{b}_{1,U_k}\end{bmatrix}, \dots, \begin{bmatrix}\mathbf{b}_{m,V}\\ \mathbf{b}_{m,U_1}\\ \vdots\\ \mathbf{b}_{m,U_k}\end{bmatrix}\right) = \begin{bmatrix}\mathbf{a}_V\\ \mathbf{a}_{U_1}\\ \vdots\\ \mathbf{a}_{U_k}\end{bmatrix} = \mathbf{a}$$

in order to satisfy Item (1) of the claim. We shall do so using the first cube identity in the V-coordinates and the U_1-coordinates, and the ith cube identity in the U_i-coordinates.

Whether $W = V$ or $W = U_i$, to solve $F^{\mathbf{A}}(\mathbf{b}_{1,W}, \dots, \mathbf{b}_{m,W}) = \mathbf{a}_W$ for the $\mathbf{b}_{i,W}$ using a particular cube identity, take $\mathbf{b}_{i,W} = \mathbf{a}_W$ if there is an x in the ith place of F in the cube identity, and take $\mathbf{b}_{i,W} = [c^{\mathbf{A}}, \dots, c^{\mathbf{A}}]^\top$ if there is a c in the ith place of the cube identity. It is not hard to see that this works, and so (1) holds.

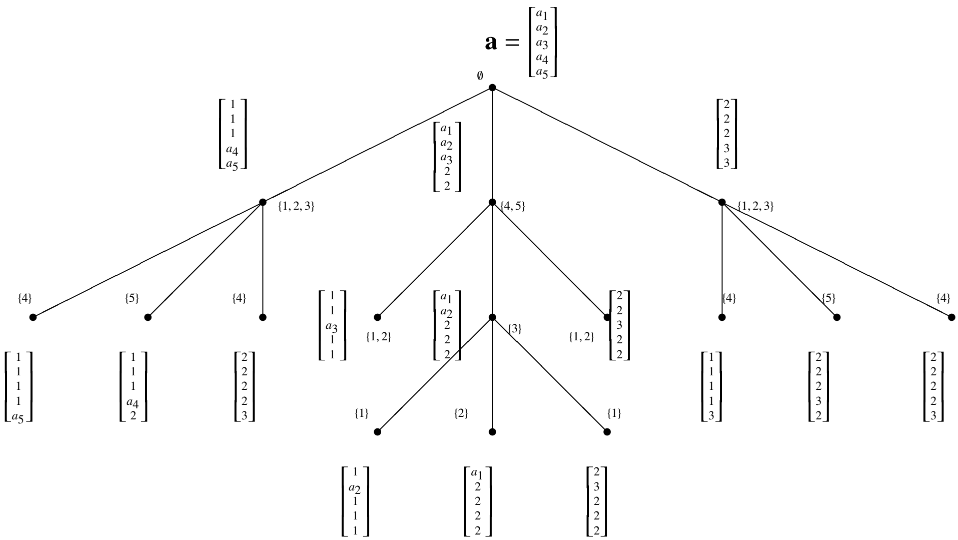
The label on node $\mathbf{n}_{\sigma i}$ is $U_{\lambda(i)}$. The element $\lambda(i) \in [k]$ is the number of the first cube identity that has some constant symbol $c \in P$ in the ith place of F. Hence, $\mathbf{b}_{i,U_{\lambda(i)}} = [c^{\mathbf{A}}, \dots, c^{\mathbf{A}}]^\top$. Thus, $\mathbf{b}_i$ is processed for node $\mathbf{n}_{\sigma i}$, establishing (2).

If, in the first cube identity, there is an x in the ith place of F, then $\mathbf{b}_{i,V} = \mathbf{a}_V$. If there is a constant symbol $c \in P$ in the ith place of F, then $\mathbf{b}_{i,V} = [c^{\mathbf{A}}, \dots, c^{\mathbf{A}}]^\top$. In the

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

latter case, $\mathbf{b}$ is processed at all coordinates in V and hence at all nodes between $\mathbf{n}_\sigma$ and $\mathbf{n}_\emptyset$. In the former case, $\mathbf{b}_i$ is processed at any node between $\mathbf{n}_\sigma$ and $\mathbf{n}_\emptyset$ where $\mathbf{a}$ is processed, since $\mathbf{b}_{i,V} = \mathbf{a}_V$. In either case, (3) holds. The claim is proved. □

The claim shows that we can attach any tuple $\mathbf{a} \in A^n$ to the root node and then process it down the tree using the cube identities until we have attached to the leaves the fully processed tuples associated to $\mathbf{a}$. Here we indicate the processing of a tuple $\mathbf{a} \in A^5$ using the example template given earlier.



Each leaf of the template determines a *type* of fully processed tuples. Two fully processed tuples $\mathbf{u}$ and $\mathbf{v}$ of the same type have the same processed coordinates, and the same constant entries in the processed coordinates. They differ only in the unprocessed coordinates. For any given type, there is a partition of the n coordinates into at most $p + k - 1$ cells where each unprocessed coordinate is a singleton cell (there are at most $k - 1$ of these cells) and all processed coordinates with a given constant entry form a cell (there are at most p of these cells). The collection of all tuples of this type lie in the subalgebra of all tuples constant on these cells, and this subalgebra is isomorphic to $\mathbf{A}^j$ for some $j \leq p + k - 1$. The assumption of the theorem is that $\mathbf{A}^{p+k-1}$ is finitely generated, say by g elements. This paragraph explains why $\mathbf{A}^n$ has a subalgebra generated by $\leq g$ elements (and isomorphic to $\mathbf{A}^j$ for some $j \leq p + k - 1$) which contains all fully processed tuples of a given type.

For example, the fully processed tuple $\begin{bmatrix} 1 \\ 1 \\ 1 \\ a_4 \\ 2 \end{bmatrix}$ from the preceding figure lies in the

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

subalgebra of all tuples of the form $\begin{bmatrix} x \\ x \\ x \\ y \\ z \end{bmatrix}$, which is isomorphic to $\mathbf{A}^3$. Here $3 \leq p + k - 1 = 3 + 2 - 1 = 4$.

Now let us count the number of types. Since the template is an m-ary tree, and the types are determined by the leaves, the number of types is at most m^r, where r is an upper bound on the length of the longest branch in the processing template. We must estimate r.

Let $V_0 = \ell(\mathbf{n}_\emptyset) = \emptyset$. This represents the set of coordinate positions that have been processed before the processing begins, that is, no coordinate positions. As we progress down a branch in the template, $\mathbf{n}_\emptyset, \mathbf{n}_i, \mathbf{n}_{ij}, \ldots, \mathbf{n}_\sigma$, we may construct sets $V_{\sigma i} = V_\sigma \cup \ell(\mathbf{n}_{\sigma i})$, where V_σ represents the set of coordinate positions that have been processed along this branch from $\mathbf{n}_\emptyset$ to $\mathbf{n}_\sigma$. The unprocessed coordinate positions, $[n]\backslash V_\sigma$, are then divided evenly, $\pi([n]\backslash V_\sigma) = (U_1, \ldots, U_k)$, to appear as labels of the children of $\mathbf{n}_\sigma$. Thus, $|V_\emptyset| = 0$ and

$$|V_{\sigma i}| = |V_\sigma \cup \ell(\mathbf{n}_{\sigma i})| = |V_\sigma| + |\ell(\mathbf{n}_{\sigma i})|. \tag{5.5}$$

The useful parameter is the number $u_\sigma := |[n]\backslash V_\sigma| = n - |V_\sigma|$ of nodes that remain unprocessed after reaching $\mathbf{n}_\sigma$. This parameter satisfies $u_\emptyset = |[n]\backslash V_\emptyset| = n$ and, from (5.5),

$$u_{\sigma i} = (n - |V_{\sigma i}|) = (n - |V_\sigma|) - |\ell(\mathbf{n}_{\sigma i})| = u_\sigma - |\ell(\mathbf{n}_{\sigma i})|. \tag{5.6}$$

Since $\pi([n]\backslash V_\sigma) = (U_1, \ldots, U_k)$ is an even division of $[n]\backslash V_\sigma$ into k sets, and $\ell(\mathbf{n}_{\sigma i}) = U_{\lambda(i)}$,

$$|\ell(\mathbf{n}_{\sigma i})| = |U_{\lambda(i)}| \geq \lfloor (n - |V_\sigma|)/k \rfloor = \lfloor u_\sigma/k \rfloor. \tag{5.7}$$

Combining (5.6) and (5.7),

$$u_{\sigma i} \leq u_\sigma - \lfloor u_\sigma/k \rfloor = \left\lceil \left(\frac{k-1}{k}\right) u_\sigma \right\rceil.$$

In order to avoid considering truncation error, we use the following fact, whose proof we leave to the reader.

Claim 5.7. *If $u \geq k \geq 1$, then $\lceil ((k-1)/k)u \rceil \leq ((2k-1)/2k)u$.* □

Hence,

$$u_{\sigma i} \leq \left(\frac{2k-1}{2k}\right) u_\sigma$$

for each σ and therefore

$$u_\sigma \leq \left(\frac{2k-1}{2k}\right)^{|\sigma|} u_\emptyset = \left(\frac{2k-1}{2k}\right)^{|\sigma|} n$$

for each σ. If, for some r, it happens that $((2k-1)/2k)^r n < k$, then there are fewer than k unprocessed nodes at address σ for any σ satisfying $|\sigma| \geq r$. Such an r is an upper bound on the length of paths through the template.

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

Solving $((2k-1)/2k)^r n < k$ for r, we obtain that any $r > \log_w(n/k)$, $w = 2k/(2k-1)$, is an upper bound on the length of paths in the template; hence, $r = \log_w(n/k) + 1$ is such a bound. Hence, the number of types of fully processed tuples is no more than

$$m^r = m^{\log_w(n/k)+1} = m^{\log_w(n/k)} m = (n/k)^{\log_w(m)} m \in O(n^{\log_w(m)}).$$

Recall that for each type, the set of fully processed tuples lies in a g-generated subalgebra of $\mathbf{A}^n$. Collecting these generators yields a set of size $O(n^{\log_w(m)})$ which generates all fully processed tuples and hence generates $\mathbf{A}^n$. □

This theorem deals only with the case $p \geq 1$. We describe next how to refine the estimate in the case $p = 1$ and how to derive the result for $p = 0$ from the $p = 1$ case.

Corollary 5.8. *If $\mathbf{A}^k$ is a finitely generated algebra with a 0-pointed or 1-pointed k-cube term, then $d_{\mathbf{A}}(n) \in O(n^{k-1})$.*

Proof. Suppose that $\mathbf{A}$ has a 1-pointed k-cube term, and that c is the one constant that appears among the cube identities. Then a fully processed tuple $\mathbf{a}$ has c in every processed coordinate position, and has at most $k-1$ unprocessed coordinate positions. Hence, the set of tuples with a c in all but at most $k-1$ positions contains all the fully processed tuples and therefore is a generating set for $\mathbf{A}^n$.

Suppose that $\mathbf{A}^k$ is g-generated. If $U \subseteq [n]$ has size $k-1$, then the subalgebra $\mathbf{A}[U]$ of tuples in $\mathbf{A}^n$ that are constant off of U is isomorphic to $\mathbf{A}^k$ and so is also g-generated. This subalgebra contains all tuples that have entry c off of U. If we collect the g generators for $\mathbf{A}[U]$ for each $(k-1)$-element subset $U \subseteq [n]$, we obtain a set of size $\binom{n}{k-1}g$ which generates $\mathbf{A}^n$. Therefore, $d_{\mathbf{A}}(n) \leq \binom{n}{k-1}g \in O(n)$.

Now suppose that $F(x_1, \ldots, x_m)$ is a 0-pointed k-cube term of $\mathbf{A}$ and that the cube identities are

$$F(M) \approx \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix}. \tag{5.8}$$

Expand $\mathbf{A}$ to an algebra $\mathbf{B}$ by adjoining a single constant, say c. Replace all variables other than x in (5.8) with c to obtain identities witnessing that $F(x_1, \ldots, x_m)$ is a 1-pointed k-cube term for $\mathbf{B}$. Hence, $d_{\mathbf{B}}(n) \in O(n^{k-1})$ by the earlier part of the argument. Now $d_{\mathbf{A}}(n) \in O(n^{k-1})$ by Theorem 2.1(4). □

In [19], we improve this result by showing that finite algebras with a 0-pointed k-cube term have logarithmic or linear growth.

Let us combine the results of this subsection with the results of the previous subsection.

Theorem 5.9. *The following are equivalent for a consistent set Σ of basic identities in which only finitely many constant symbols occur.*

(1) *Σ is restrictive. (That is, the class of d-functions of nontrivial algebras is not equal to the class of d-functions of algebras that realize Σ.)*

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

(2) *Σ is restrictive for nontrivial finite algebras. (The class of d-functions of finite algebras is not equal to the class of d-functions of finite algebras that realize Σ.)*
(3) *The variety axiomatized by Σ has a pointed cube term.*
(4) *The variety axiomatized by Σ has a pointed cube term of the form $F(x_1, \dots, x_m)$, where $m \geq 2$, F is a function symbol occurring in Σ, and the variables $x_1, \dots, x_m$ are distinct.*
(5) *If $\mathbf{A}$ is an algebra realizing Σ and $d_{\mathbf{A}}(n)$ is finite for all n, then $d_{\mathbf{A}}(n)$ is bounded above by a polynomial.*
(6) *There is no (finite) algebra $\mathbf{A}$ realizing Σ such that $d_{\mathbf{A}}(n) = 2^n$ for all n.*

PROOF. [(1) $\Rightarrow$ (3) and (2) $\Rightarrow$ (3)] Theorem 5.3.

[(3) $\Leftrightarrow$ (4)] Lemma 5.1.

[(3) $\Rightarrow$ (5)] Theorem 5.5 and Corollary 5.8.

[(5) $\Rightarrow$ (6)] $d_{\mathbf{A}}(n) = 2^n$ is not bounded above by a polynomial.

[(6) $\Rightarrow$ (1) and (6) $\Rightarrow$ (2)] If (1) or (2) fails, then Σ is nonrestrictive (for finite algebras). Thus, there exists a (finite) algebra $\mathbf{A}$ realizing Σ with the same growth rate $d_{\mathbf{A}}(n) = 2^n$ as the two-element set equipped with no operations. Hence, (6) fails. □

5.3. Finite algebras with polynomial growth. In this subsection, we prove that the bound on growth rates for finite algebras with 1-pointed k-cube terms, established in Corollary 5.8, is sharp.

THEOREM 5.10. *For each $k \geq 2$, there is a finite algebra with a 1-pointed k-cube term whose growth rate satisfies $d_{\mathbf{A}}(n) \in \Theta(n^{k-1})$.*

PROOF. We shall first construct a partial algebra with the desired growth rate and then modify it slightly to obtain a total algebra satisfying the hypotheses of the theorem.

The universe of the partial algebra will be $A = \{a_1, \dots, a_q, 1\}$. We equip this set with a partial k-ary operation F which satisfies

$$F^{\mathbf{A}}(1, x, \dots, x, x) = F^{\mathbf{A}}(x, 1, \dots, x, x) = \cdots = F^{\mathbf{A}}(x, x, \dots, x, 1) = x$$

for each $x \in A$, and which is undefined otherwise. Thus, $F^{\mathbf{A}}$ is a partial near-unanimity operation that is defined only on the nearly unanimous tuples, where the lone dissenter is 1, and on the tuple whose entries are unanimously 1. Set $\mathbf{A} = \langle A; F \rangle$.

We shall prove the exact formula

$$d_{\mathbf{A}}(n) = \binom{n}{0} + q\binom{n}{1} + q^2\binom{n}{2} + \cdots + q^{k-1}\binom{n}{k-1} \tag{5.9}$$

for this partial algebra, which is a polynomial in n of degree $k - 1$, since $k = \text{arity}(F)$ and $q = |A| - 1$ are fixed. This will show that $\mathbf{A}$ is a $(q + 1)$-element partial algebra with $d_{\mathbf{A}}(n) \in \Theta(n^{k-1})$.

Choose and fix n. Define the *support* of a tuple $\mathbf{a} \in A^n$ to be the subset $\text{supp}(\mathbf{a}) \subseteq [n]$ consisting of indices s, where $a_s \neq 1$. The proof involves showing that the set of all tuples whose support has size at most $k - 1$ is the unique minimal generating set for $\mathbf{A}^n$. To set up language for the argument, call a tuple $\mathbf{b} \in A^n$ an *essential generator* if it is contained in any generating set for $\mathbf{A}^n$.

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

Claim 5.11. *If $S \subseteq [n]$ and $G \subseteq A^n$, then let G_S denote the set of tuples in G that have support contained in S. If $\mathbf{a} \in \langle G \rangle$ has support in S, then $\mathbf{a} \in \langle G_S \rangle$.*

Proof of Claim. In $\mathbf{A}$,

$$F^{\mathbf{A}}(x_1, \dots, x_k) = 1 \iff x_1 = x_2 = \cdots = x_k = 1.$$

Hence, in $\mathbf{A}^n$, if $F^{\mathbf{A}^n}(\mathbf{g}_1, \dots, \mathbf{g}_k)$ is defined and equal to $\mathbf{b}$, then $i \notin \mathrm{supp}(\mathbf{b})$ if and only if $i \notin \mathrm{supp}(\mathbf{g}_i)$ for any $\mathbf{g}_i$. Equivalently,

$$\mathrm{supp}(F^{\mathbf{A}^n}(\mathbf{g}_1, \dots, \mathbf{g}_k)) = \bigcup_{i=1}^{k} \mathrm{supp}(\mathbf{g}_i) \tag{5.10}$$

whenever $F^{\mathbf{A}^n}(\mathbf{g}_1, \dots, \mathbf{g}_k)$ is defined. Now let $G(0) = G$, $G_S(0) = G_S$, $G(j+1) = G(j) \cup F^{\mathbf{A}^n}(G(j), \dots, G(j))$, and $G_S(j+1) = G_S(j) \cup F^{\mathbf{A}^n}(G_S(j), \dots, G_S(j))$. By induction on j, using (5.10), it can be shown that any tuple in $G(j)$ that has support in S lies in $G_S(j)$. Since $\langle G \rangle = \bigcup_j G(j)$ and $\langle G_S \rangle = \bigcup_j G_S(j)$, any tuple in $\langle G \rangle$ with support in S lies in $\langle G_S \rangle$. □

Claim 5.12. *The tuple $\hat{1} = [1, 1, \dots, 1]^{\mathsf{T}}$ of empty support is an essential generator.*

Proof of Claim. This follows immediately from Claim 5.11. □

Claim 5.13. *Any tuple whose support has size at most $k - 1$ is an essential generator of $\mathbf{A}^n$.*

Proof of Claim. Let $\mathbf{b} \in A^n$ be a tuple of support S, where $1 \leq |S| \leq k - 1$. Without loss of generality, $S = [\ell] = \{1, \dots, \ell\}$ for some $1 \leq \ell \leq k - 1$. In order to obtain a contradiction to the claim, assume that $\mathbf{b}$ is not an essential generator. Then $\mathbf{b}$ can be generated by elements different from $\mathbf{b}$, so the equation $F^{\mathbf{A}^n}(\mathbf{x}_1, \dots, \mathbf{x}_k) = \mathbf{b}$ can be solved for the $\mathbf{x}_i$ in such a way that $\mathbf{b} \notin \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$. Moreover, by (5.10), the $\mathbf{x}_i$ must be taken from the tuples whose support is contained in S. The equation to be solved is therefore

$$F^{\mathbf{A}^n}(\mathbf{x}_1, \dots, \mathbf{x}_k) = F^{\mathbf{A}^n}\left(\begin{bmatrix} x_{1,1} \\ \vdots \\ x_{\ell,1} \\ \hline 1 \\ \vdots \\ 1 \end{bmatrix}, \dots, \begin{bmatrix} x_{1,k} \\ \vdots \\ x_{\ell,k} \\ \hline 1 \\ \vdots \\ 1 \end{bmatrix} \right) = \begin{bmatrix} b_1 \\ \vdots \\ b_\ell \\ \hline 1 \\ \vdots \\ 1 \end{bmatrix} = \mathbf{b}.$$

We have introduced horizontal segments as dividers separating the coordinates in $S = [\ell]$ from the remaining coordinates in order to make the argument clearer. Since $F^{\mathbf{A}^n}(\mathbf{x}_1, \dots, \mathbf{x}_k)$ is defined, every row above the dividers is a nearly unanimous row with exactly one 1. Hence, there are exactly ℓ 1's above the dividers. This means that there are at most ℓ columns which contain a 1 above the dividers. Since there are k

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

such columns, and $k > \ell$, there is a column $\mathbf{x}_j$ that contains no 1 above the dividers. Since the ith row above the dividers is nearly unanimous with majority value b_i, the column $\mathbf{x}_j$ which contains no 1's above the dividers is exactly $\mathbf{b}$. This contradicts the assumption that $\mathbf{b} \notin \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$, showing that $\mathbf{b}$ is indeed an essential generator. □

CLAIM 5.14. *$\mathbf{A}^n$ is generated by the tuples whose support has size at most $k-1$.*

PROOF OF CLAIM. It is enough to show that if $\mathbf{b}$ has support S of size $\ell \geq k$, then $\mathbf{b}$ can be generated from tuples whose support is properly contained in S. It is enough to prove this in the case where $S = [\ell]$. For this, we must explain how to solve

$$F^{\mathbf{A}^n}(\mathbf{x}_1, \ldots, \mathbf{x}_k) = F^{\mathbf{A}^n}\left(\begin{bmatrix} x_{1,1} \\ \vdots \\ x_{\ell,1} \\ \hline 1 \\ \vdots \\ 1 \end{bmatrix}, \ldots, \begin{bmatrix} x_{1,k} \\ \vdots \\ x_{\ell,k} \\ \hline 1 \\ \vdots \\ 1 \end{bmatrix}\right) = \begin{bmatrix} b_1 \\ \vdots \\ b_\ell \\ \hline 1 \\ \vdots \\ 1 \end{bmatrix} = \mathbf{b}$$

when $\ell \geq k$ in such a way that every column contains at least one 1 above the dividers and the ith row above the dividers is nearly unanimously equal to b_i. This is easy to do. Set $x_{1,1} = \cdots = x_{k,k} = 1$, then put exactly one 1 arbitrarily in each of rows $k+1$ to ℓ, and then fill in the remaining entries above the dividers so that the ith row above the dividers is nearly unanimously equal to b_i. □

We have established up to this point that the set of tuples of support of size at most $k-1$ is the unique minimal generating set for $\mathbf{A}^n$. To complete the proof that the partial algebra $\mathbf{A}$ has the specified growth rate, observe that the number of tuples with support S is $(|A|-1)^{|S|} = q^{|S|}$, so the number of tuples whose support has size i is $q^i\binom{n}{i}$. This yields the formula $d_{\mathbf{A}}(n) = \sum_{i=0}^{k-1} q^i\binom{n}{i}$.

The one-point completion, $\mathbf{A}_0$, is a total algebra with the same growth rate as $\mathbf{A}$. Let $\mathbf{B}$ be the expansion of $\mathbf{A}_0$ by one constant symbol 1 whose interpretation is $1^{\mathbf{B}} = 1$. The operation $F^{\mathbf{B}}$ still satisfies

$$F^{\mathbf{B}}(1, x, \ldots, x, x) = F^{\mathbf{B}}(x, 1, \ldots, x, x) = \cdots = F^{\mathbf{B}}(x, x, \ldots, x, 1) = x$$

for each $x \in A_0$, so it is a 1-pointed k-cube term for $\mathbf{B}$.

By Theorem 3.2 $\mathbf{A}^n$ and $\mathbf{A}_0^n$ have the same unique minimal generating set, G, which is the set of all tuples with support at most $k-1$; this set contains $\hat{1}$. The algebra $\mathbf{B}$ must also have a unique minimal generating set, namely the set obtained from G by deleting $\hat{1} = 1^{\mathbf{B}^n}$. Thus, $d_{\mathbf{B}}(n) = d_{\mathbf{A}}(n) - 1 = \sum_{i=1}^{k-1} q^i\binom{n}{i} \in O(n^{k-1})$. □

5.4. Pointed cube polynomials can be avoided. We have established that if $\mathbf{A}$ is an algebra whose d-function assumes only finite values, and $\mathbf{A}$ has a pointed cube term (or pointed cube polynomial for that matter), then $d_{\mathbf{A}}(n)$ is bounded above by a polynomial function of n. The same growth rate can be obtained without a pointed cube term (or polynomial), as we show next.

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

Theorem 5.15. *Let* $\mathbf{A}$ *be an algebra with* $|A| > 1$ *whose* d*-function assumes only finite values. There is an algebra* $\mathbf{B}$ *such that* $d_{\mathbf{B}}(n) = d_{\mathbf{A}}(n)$ *for all* n *and:*

(1) *the universe of* $\mathbf{B}$ *is* $B := A \cup \{0, z\}$*, where* $0 \neq z$ *and* $0, z \notin A$*;*

(2) $\mathbf{B}$ *has a meet semilattice term operation,* $\wedge$*, with respect to which* $\mathbf{B}$ *has height one and least element* 0*; and*

(3) *if* $p(x, \mathbf{y})$ *is an* m*-ary polynomial of* $\mathbf{B}$ *in which* x *actually appears and* $p(z, \mathbf{b}) = z$ *for some* $\mathbf{b} \in B^{m-1}$*, then either* $p(x, \mathbf{y}) \approx x$ *or else* $p(x, \mathbf{y}) \approx x \wedge q(\mathbf{y})$ *for some polynomial* q *in which* x *does not appear.*

In particular, $\mathbf{B}$ *does not have a pointed cube polynomial.*

Proof. Let $G(n) := \{\mathbf{g}_{n,1}, \ldots, \mathbf{g}_{n,d(n)}\}$ be a least size generating set for $\mathbf{A}^n$. Let $\mathbf{A}_z$ be the one-point completion of $\mathbf{A}$ with the element z ($\notin A$) taken to be the new point added. According to Theorem 3.2, the set $G(n)$ is also a least size generating set for $\mathbf{A}_z$. Next, copying the idea of the construction in Theorem 3.4, for each $\mathbf{a} \in (A_z)^n$ introduce a partial operation $F_{\mathbf{a}}(x_1, \ldots, x_{d(n)})$ on A_z with the properties that (i) the vector equation

$$F_{\mathbf{a}}(\mathbf{g}_{n,1}, \ldots, \mathbf{g}_{n,d(n)}) = \mathbf{a} \tag{5.11}$$

holds coordinatewise and (ii) $F_{\mathbf{a}}$ is defined only on those tuples required to make this equation hold. Let $\overline{\mathbf{A}}_z$ be the set A_z equipped with these partial operations. The partial algebra $\overline{\mathbf{A}}_z$ is a reduct of $\mathbf{A}_z$, so in passing from $\mathbf{A}_z$ to $\overline{\mathbf{A}}_z$ we may have lost but not gained some generating subsets of powers. On the other hand, our choice of the partial operations guarantees that $G(n)$ still generates the nth power of $\overline{\mathbf{A}}_z$. This implies that $G(n)$ is a least size generating set for $\overline{\mathbf{A}}_z^n$ for each n and hence that the d-functions of $\overline{\mathbf{A}}_z$ and $\mathbf{A}_z$ are the same. Finally, let $\mathbf{B} = (\overline{\mathbf{A}}_z)_0$ be the one-point completion of $\overline{\mathbf{A}}_z$ with the element 0 ($\notin A \cup \{z\}$) taken to be the new point added. With this choice, the universe of $\mathbf{B}$ is $B = A \cup \{0, z\}$. Again citing Theorem 3.2, we see that $G(n)$ is a least size generating set for $\mathbf{B}^n$.

At this point we have that $d_{\mathbf{B}}(n) = d_{\mathbf{A}}(n)$ for all n and also, by construction, that Items (1) and (2) hold. (Here the meet operation referred to in Item (3) is the one introduced in the second one-point completion, the one used to construct $\mathbf{B}$ from $\overline{\mathbf{A}}_z$.)

Let us prove that Item (3) holds. Our argument depends on a key fact: z does not appear in any coordinate of any tuple in $G(n)$ for any n and hence z does not appear in any tuple in the domain of any partial operation of the form $F_{\mathbf{a}}$. This implies that any basic operation of $\mathbf{B}$ of the form $(F_{\mathbf{a}})_0$ (Definition 3.1) assigns the value 0 to any tuple containing a z (or a 0).

We first prove that if $p(x, \mathbf{y})$ is an m-ary polynomial of $\mathbf{B}$ in which x appears and $\mathbf{b} \in B^{m-1}$, then $p(z, \mathbf{b}) \in \{0, z\}$. Arguing by induction on the complexity of p, we need to consider the cases where p is a constant, a variable, or of the form

$$p(x, \mathbf{y}) = F(p_1(x, \mathbf{y}), \ldots, p_\ell(x, \mathbf{y})), \tag{5.12}$$

where $F = (F_{\mathbf{a}})_0$ or $F = \wedge$. The polynomial p cannot be a constant, since x appears in p. If p is a variable, it must be x, since x appears in p. In this case, $p(z, \mathbf{b}) = z \in \{0, z\}$, as

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

claimed. If (5.12) holds in the case where $F = (F_{\mathbf{a}})_0$, then by the induction hypothesis we have $p_i(z, \mathbf{b}) \in \{0, z\}$ for at least one i; hence, by the key fact,

$$p(z, \mathbf{b}) = (F_{\mathbf{a}})_0(p_1(z, \mathbf{b}), \ldots, p_\ell(z, \mathbf{b})) = 0 \in \{0, z\},$$

as claimed. If (5.12) holds in the case where $F = \wedge$, then by the induction hypothesis we have $p_i(z, \mathbf{b}) \in \{0, z\}$ for at least one i; hence, $p_i(z, \mathbf{b}) \leq z$. It follows that $p(z, \mathbf{b}) = p_1(z, \mathbf{b}) \wedge p_2(z, \mathbf{b}) \leq z$, so, since $\langle B; \wedge \rangle$ has height one, we get that $p(z, \mathbf{b}) \in \{0, z\}$.

Now we prove Item (3) by induction on the complexity of p. Under the assumptions of Item (3), the polynomial p cannot be a constant, since x appears in p. If p is a variable, it must be x, since x appears in p, in which case $p(x, \mathbf{y}) = x$ for all x and $\mathbf{y}$, and Item (3) holds. Now assume that (5.12) holds in the case where $F = (F_{\mathbf{a}})_0$, and fix a tuple $\mathbf{b} \in B^{m-1}$ satisfying $p(z, \mathbf{b}) = z$ (the existence of such a $\mathbf{b}$ is assumed in Item (3)). Since x appears in p, it also appears in some $p_i(x, \mathbf{y})$. By the preceding paragraph, we get that $p_i(z, \mathbf{b}) \in \{0, z\}$. But, as noted two paragraphs ago, any basic operation of $\mathbf{B}$ of the form $(F_{\mathbf{a}})_0$ assigns the value 0 to any tuple containing a 0 or a z, so we obtain the right-hand equality (the only nontrivial equality) in

$$z = p(z, \mathbf{b}) = (F_{\mathbf{a}})_0(p_1(z, \mathbf{b}), \ldots, p_\ell(z, \mathbf{b})) = 0.$$

This contradiction shows that this case cannot occur. Finally, if $p(x, \mathbf{y}) = p_1(x, \mathbf{y}) \wedge p_2(x, \mathbf{y})$ and $\mathbf{b} \in B^{m-1}$ is such that $p(z, \mathbf{b}) = z$, then $p_i(z, \mathbf{b}) = z$ for $i = 1, 2$, since z is meet irreducible in $\langle B; \wedge \rangle$. If x appears in both $p_1(x, \mathbf{y})$ and $p_2(x, \mathbf{y})$, then by induction both have the form x or $x \wedge q_i(\mathbf{y})$. Hence, $p(x, \mathbf{y})$ has the form

$$x \wedge x, \quad x \wedge (x \wedge q_2(\mathbf{y})), \quad (x \wedge q_1(\mathbf{y})) \wedge x, \quad \text{or} \quad (x \wedge q_1(\mathbf{y})) \wedge (x \wedge q_2(\mathbf{y})),$$

each of which has the form x or $x \wedge q(\mathbf{y})$ for some polynomial q. A similar conclusion is reached if x appears in one of the polynomials $p_i(x, \mathbf{y})$ but not the other. Hence, Item (3) holds.

To complete the proof of the theorem, we argue that $\mathbf{B}$ does not have a pointed cube polynomial. By way of contradiction, assume that $p(x_1, \ldots, x_m)$ is such a polynomial and that M is a $k \times m$ matrix of variables and constants such that $p(M) \approx [x, \ldots, x]^{\mathsf{T}}$ and every column of M contains at least one entry that is not x. In fact, as we have seen before, by substituting constants for the variables different from x we may assume that the entries of M are constants or x and that each column contains at least one constant. We may also assume that p depends on all of its variables and hence that each of $x_1, \ldots, x_m$ appears in p.

Here are some elementary consequences of our assumptions.

(a) Each row of M must contain at least one x, since otherwise we may derive from the associated cube identity that $x \approx y$ holds in $\mathbf{B}$. By permuting columns of M (hence re-ordering the variables of p), we assume that the first entry of the first row is x.

(b) The first column of M contains a constant, which cannot be in the first row. By permuting the later rows of M (hence re-ordering the cube identities), we assume that the first entry of the second row of M is a constant. There is an x somewhere on the second row, by (a), and permuting the later columns we may assume that it is in the second position of the second row.

These consequences mean that the first two cube identities look like $p(x, b_2, \mathbf{b}) \approx x$ and $p(c_1, x, \mathbf{c}) \approx x$, where all $b_i, c_j \in B \cup \{x\}$ and c_1 is constant. If we substitute z for each x in these equations, we get $p(z, b_2', \mathbf{b}') = z$ and $p(c_1', z, \mathbf{c}') = z$, where the primes on elements and tuples indicate that the x in the string have been replaced by z and constants remain the same. Applying Item (3) of this theorem to these equalities,

$$p(x_1, x_2, \mathbf{y}) = x_1 \wedge q_1(x_2, \mathbf{y}) = x_2 \wedge q_2(x_1, \mathbf{y}),$$

where x_i does not appear in q_i. By meeting p with itself,

$$p(x_1, x_2, \mathbf{y}) = (x_1 \wedge q_1(x_2, \mathbf{y})) \wedge (x_2 \wedge q_2(x_1, \mathbf{y})).$$

Now the second cube identity may be written

$$x = p(c_1, x, \mathbf{c}) = (c_1 \wedge q_1(x, \mathbf{c})) \wedge (x \wedge q_2(c_1, \mathbf{c})) \leq c_1.$$

This implies $x \leq c_1$ for all $x \in B$ and therefore that the element $c_1 \in B$ is the largest element of $\langle B; \wedge\rangle$. But this semilattice has no largest element, since it has at least four elements and has height one. This contradiction proves that $\mathbf{B}$ has no pointed cube polynomial. □

5.5. Exponential growth. If $\mathbf{A}$ has exponential growth and $\mathbf{B}$ has arbitrary growth, then $\mathbf{A} \times \mathbf{B}$ has exponential growth according to Theorem 2.1(2). Hence, it is probably unrealistic to expect any meaningful classification of algebras with exponential growth. This subsection will therefore be limited to identifying one property that forces exponential growth. We will use the property to show that the variety generated by the two-element implication algebra, $\langle\{0, 1\}; \rightarrow\rangle$, contains a chain of finite algebras $\mathbf{A}_1 \leq \mathbf{A}_2 \leq \cdots$, each one a subalgebra of the next, where $\mathbf{A}_i$ has logarithmic growth when i is odd and exponential growth when i is even.

We explore a simple idea: suppose that $\mathbf{A}$ is finite and u and v are distinct elements of A. If every element of $\{u, v\}^n$ is an essential generator of $\mathbf{A}^n$ for each n, then the growth rate of $\mathbf{A}$ must be at least 2^n. A way to force some tuple $\mathbf{t} \in \{u, v\}^n$ to be an essential generator of $\mathbf{A}^n$ is to arrange that $A^n \backslash \{\mathbf{t}\}$ is a subuniverse of $\mathbf{A}^n$. This can be accomplished by imposing an irreducibility condition on each coordinate t of $\mathbf{t}$, or equivalently by requiring that the complementary set $A \backslash \{t\}$ behaves like an ideal. For this to work, it is enough that $A \backslash \{t\}$ behaves like a one-sided semigroup-theoretic ideal, so we introduce a definition that captures this notion for an arbitrary algebraic signature.

Definition 5.16. Let $\sigma = (F, \alpha)$ be an algebraic signature. That is, let F be a set (of operation symbols) and let $\alpha\colon F \rightarrow \omega$ be a function (assigning arity). Let $F_0 \subseteq F$ be the set consisting of those $f \in F$ such that $\alpha(f) > 0$. (The set F_0 is the set of nonnullary symbols.) A *selector* for σ is a function $\phi\colon F_0 \rightarrow \omega$ such that $1 \leq \phi(f) \leq \alpha(f)$ for each $f \in F_0$. (The function ϕ selects one of the places of the function symbol f.)

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

If ϕ is a selector for σ and $\mathbf{A}$ is an algebra of signature σ, then a *ϕ-irreducible* subset of $\mathbf{A}$ is a subset $U \subseteq A$ such that whenever $\alpha(f) = n$ and $\phi(f) = i$,

$$f^{\mathbf{A}}(a_1, \ldots, a_n) \in U \Rightarrow a_i \in U.$$

The complement of a ϕ-irreducible subset is called a *ϕ-ideal*. Explicitly, $I \subseteq A$ is a ϕ-ideal if whenever $\alpha(f) = n$, $\phi(f) = i$, and $a_i \in I$, then $f^{\mathbf{A}}(a_1, \ldots, a_n) \in I$.

In this terminology, a left ideal of a semigroup with multiplication represented by the symbol m would be a ϕ-ideal for the function $\phi\colon \{m\} \to \{1, 2\}\colon m \mapsto 2$, while a right ideal would be a ϕ-ideal for the function $\phi\colon \{m\} \to \{1, 2\}\colon m \mapsto 1$.

Theorem 5.17. *Let $\mathbf{A}$ be an algebra of signature σ and let ϕ be a selector for σ. If $\mathbf{A}$ is the union of finitely many proper ϕ-ideals, then $d_{\mathbf{A}}(n) \geq 2^n$.*

Proof. The union of ϕ-ideals is again a ϕ-ideal, so if $\mathbf{A}$ is the union of $k \geq 2$ proper ϕ-ideals then it can be expressed as the union $I \cup J$ of two proper ϕ-ideals. The complements $I' := A \backslash I$ and $J' := A \backslash J$ are disjoint ϕ-irreducible sets. Any product $T := X_1 \times \cdots \times X_n$, with $X_i = I'$ or J' for all i, is a ϕ-irreducible subset of A^n. Each such set must contain at least one element of any generating set, since the ϕ-irreducibility of T implies that $A^n \backslash T$ is a subuniverse of $\mathbf{A}^n$. Since there are 2^n products of the form $X_1 \times \cdots \times X_n$ with $X_i = I'$ or J', and they are pairwise disjoint, any generating set for $\mathbf{A}^n$ must contain at least 2^n elements. □

Example 5.18. In this example, $\mathbf{2}$ is the two-element Boolean algebra and $\mathbf{2}^\circ = \langle\{0, 1\}; \to\rangle$ is the reduct of $\mathbf{2}$ to the operation $x \to y = x' \vee y$. The variety $\mathcal{V}$ generated by $\mathbf{2}^\circ$ is called the variety of implication algebras. This variety is congruence distributive and has $\mathbf{2}^\circ$ as its unique subdirectly irreducible member. Each finite algebra in $\mathcal{V}$ may be viewed as an order filter in a finite Boolean algebra: if $\mathbf{A} \in \mathcal{V}_{\text{fin}}$, then an irredundant subdirect representation $\mathbf{A} \leq (\mathbf{2}^\circ)^k$ may be viewed as a representation of $\mathbf{A}$ as a subset of $\mathbf{2}^k$ closed under $\to$; such subsets of $\mathbf{2}^k$ are order filters.

Considering an algebra $\mathbf{A} \in \mathcal{V}_{\text{fin}}$ to be an order filter in $\mathbf{2}^k$, each order filter contained within $\mathbf{A}$ is a left ideal in $\mathbf{A}$ with respect to the operation $\to$. By Theorem 5.17, if $\mathbf{A}$ is the union of its proper order filters, its growth rate is exponential. This case must occur unless $\mathbf{A}$ itself is a principal order filter in $\mathbf{2}^k$. Since we represented $\mathbf{A}$ irredundantly, $\mathbf{A}$ is a principal order filter in $\mathbf{2}^k$ only when it is the improper filter, that is, $\mathbf{A} = (2^\circ)^k$. In this situation, $\mathbf{A}$ is polynomially equivalent to the Boolean algebra $\mathbf{2}^k$. It follows from Theorem 2.1(1) and the fact that $\mathbf{2}$ is primal that $\mathbf{2}^k$ has logarithmic growth rate. In summary, a finite implication algebra has logarithmic growth rate if it has a least element and has exponential growth rate otherwise.

Now it is easy to produce a chain of implication algebras $\mathbf{A}_1 \leq \mathbf{A}_2 \leq \cdots$, each one a subalgebra of the next, where $\mathbf{A}_i$ has logarithmic growth when i is odd and exponential growth when i is even. One simply chooses larger and larger Boolean order filters which are principal only when i is odd. Figure 1 below shows how the chain might begin.

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press

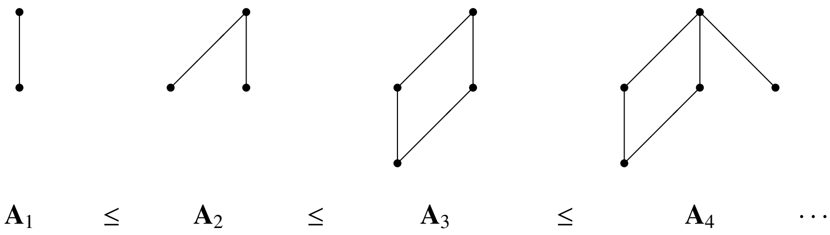


FIGURE 1. A chain of implication algebras.

6. Problems

In this paper, we have filled in one gap in knowledge about the spectrum of possible growth rates of finite algebras by producing examples with superlinear polynomial growth rates. There is an interesting gap in knowledge that remains between logarithmic and linear growth rates.

PROBLEM 6.1. Is there a finite algebra $\mathbf{A}$ where $d_{\mathbf{A}}(n) \notin \Omega(n)$ and $d_{\mathbf{A}}(n) \notin O(\log(n))$?

A special case that might be tractable is the following question.

PROBLEM 6.2. Is there a two-element partial algebra $\mathbf{A}$ where $d_{\mathbf{A}}(n) \notin \Omega(n)$ and $d_{\mathbf{A}}(n) \notin O(\log(n))$?

We know that no finite algebra with a 0-pointed cube term can have growth rate between logarithmic and linear, but do not know the situation for pointed cube terms. The following seems to be the most interesting special case.

PROBLEM 6.3. Is it true that a finite algebra with a two-sided unit for some binary term has logarithmic or linear growth?

There is also a possible gap near the exponential end of the spectrum.

PROBLEM 6.4. Is there a finite algebra $\mathbf{A}$ where $d_{\mathbf{A}}(n) \notin 2^{\Omega(n)}$ and $d_{\mathbf{A}}(n) \notin O(n^k)$ for any k?

Acknowledgement

The authors gratefully acknowledge the helpful advice of the referee.

References

[1] J. Berman, P. Idziak, P. Marković, R. McKenzie, M. Valeriote and R. Willard, 'Varieties with few subalgebras of powers', *Trans. Amer. Math. Soc.* **362**(3) (2010), 1445–1473.

[2] J. Berman and R. McKenzie, 'Clones satisfying the term condition', *Discrete Math.* **52** (1984), 7–29.

[3] H. Chen, 'Quantified constraint satisfaction and the polynomially generated powers property', in: *ICALP 2008, Part II*, Lecture Notes in Computer Science, 5126 (eds. L. Aceto *et al.*) (Springer, Berlin–Heidelberg, 2008), 197–208.
[4] I. M. S. Dey, 'Embeddings in non-Hopf groups', *J. Lond. Math. Soc.* (2) **1** (1969), 745–749.
[5] A. Erfanian, 'A problem on growth sequences of groups', *J. Aust. Math. Soc. Ser. A* **59**(2) (1995), 283–286.
[6] A. Erfanian, 'A note on growth sequences of alternating groups', *Arch. Math. (Basel)* **78**(4) (2002), 257–262.
[7] A. Erfanian, 'A note on growth sequences of $\mathrm{PSL}(m, q)$', *Southeast Asian Bull. Math.* **29**(4) (2005), 697–713.
[8] A. Erfanian, 'On the growth sequences of free product of $\mathrm{PSL}(m, q)$', *Ital. J. Pure Appl. Math.* **22** (2007), 19–26.
[9] A. Erfanian, 'Growth sequence of free product of alternating groups', *Int. J. Contemp. Math. Sci.* **2**(13–16) (2007), 685–691.
[10] A. Erfanian and R. Rezaei, 'On the growth sequences of $\mathrm{PSp}(2m, q)$', *Int. J. Algebra* **1**(1–4) (2007), 51–62.
[11] A. Erfanian and J. Wiegold, 'A note on growth sequences of finite simple groups', *Bull. Aust. Math. Soc.* **51**(3) (1995), 495–499.
[12] A. L. Foster, 'On the finiteness of free (universal) algebras', *Proc. Amer. Math. Soc.* **7** (1956), 1011–1013.
[13] R. Freese and R. McKenzie, *Commutator Theory for Congruence Modular Varieties*, London Mathematical Society Lecture Note Series, 125 (Cambridge University Press, Cambridge, 1987).
[14] A. M. W. Glass and H. H. J. Riedel, 'Growth sequences—a counterexample', *Algebra Universalis* **21**(2–3) (1985), 143–145.
[15] P. Hall, 'The Eulerian functions of a group', *Q. J. Math.* **7** (1936), 134–151.
[16] J. T. Hyde, N. J. Loughlin, M. Quick, N. Ruskuc and A. R. Wallis, 'On the growth of generating sets for direct powers of semigroups', *Semigroup Forum* **84** (2012), 116–130.
[17] K. Kearnes and E. Kiss, 'Finite algebras of finite complexity', *Discrete Math.* **207** (1999), 89–135.
[18] K. Kearnes and E. Kiss, 'The shape of congruence lattices', *Mem. Amer. Math. Soc.* **222**(1046) (2013).
[19] K. Kearnes, E. Kiss and Á. Szendrei, 'Growth rates of algebras, II: Wiegold dichotomy', *Internat. J. Algebra Comput.* **25**(4) (2015), 555–566.
[20] K. Kearnes, E. Kiss and Á. Szendrei, 'Growth rates of algebras, III: finite solvable algebras', *Algebra Universalis*, to appear.
[21] K. Kearnes and Á. Szendrei, 'Clones of algebras with parallelogram terms', *Internat. J. Algebra Comput.* **22** (2012).
[22] D. Kelly, 'Basic equations: word problems and Mal'cev conditions, Abstract 701-08-04', *Notices Amer. Math. Soc.* **20** (1972), A-54.
[23] W. Kimmerle, 'Growth sequences relative to subgroups', in: *Groups – St. Andrews 1981 (St. Andrews, 1981)*, London Mathematical Society Lecture Note Series, 71 (Cambridge University Press, Cambridge–New York, 1982), 252–260.
[24] J. C. Lennox and J. Wiegold, 'Generators and killers for direct and free products', *Arch. Math. (Basel)* **34**(4) (1980), 296–300.
[25] D. Meier and J. Wiegold, 'Growth sequences of finite groups. V', *J. Aust. Math. Soc. Ser. A* **31**(3) (1981), 374–375.
[26] V. N. Obraztsov, 'Growth sequences of 2-generator simple groups', *Proc. Roy. Soc. Edinburgh Sect. A* **123**(5) (1993), 839–855.
[27] G. Pollák, 'Growth sequence of globally idempotent semigroups', *J. Aust. Math. Soc. Ser. A* **48**(1) (1990), 87–88.
[28] M. Quick and N. Ruškuc, 'Growth of generating sets for direct powers of classical algebraic structures', *J. Aust. Math. Soc.* **89** (2010), 105–126.
[29] H. H. J. Riedel, 'Growth sequences of finite algebras', *Algebra Universalis* **20**(1) (1985), 90–95.

[30] A. G. R. Stewart and J. Wiegold, 'Growth sequences of finitely generated groups. II', *Bull. Aust. Math. Soc.* **40**(2) (1989), 323–329.
[31] J. Wiegold, 'Growth sequences of finite groups', *J. Aust. Math. Soc.* **17** (1974), 133–141; Collection of articles dedicated to the memory of Hanna Neumann, VI.
[32] J. Wiegold, 'Growth sequences of finite groups. II', *J. Aust. Math. Soc.* **20**(2) (1975), 225–229.
[33] J. Wiegold, 'Growth sequences of finite groups. III', *J. Aust. Math. Soc. Ser. A* **25**(2) (1978), 142–144.
[34] J. Wiegold, 'Growth sequences of finite groups. IV', *J. Aust. Math. Soc. Ser. A* **29**(1) (1980), 14–16.
[35] J. Wiegold, 'Growth sequences of finite semigroups', *J. Aust. Math. Soc. Ser. A* **43**(1) (1987), 16–20.
[36] J. Wiegold and J. S. Wilson, 'Growth sequences of finitely generated groups', *Arch. Math. (Basel)* **30**(4) (1978), 337–343.
[37] D. T. Wise, 'The rank of a direct power of a small-cancellation group', *Proc. Conf. Geometric and Combinatorial Group Theory, Part I (Haifa, 2000)*, *Geom. Dedicata* **94** (2002), 215–223.

KEITH A. KEARNES, Department of Mathematics,
University of Colorado, Boulder, CO 80309-0395, USA
e-mail: keith.kearnes@colorado.edu

EMIL W. KISS, Department of Algebra and Number Theory,
Loránd Eötvös University, Pázmány Péter stny 1/c.,
H-1117 Budapest, Hungary
e-mail: ewkiss@cs.elte.hu

ÁGNES SZENDREI, Department of Mathematics,
University of Colorado, Boulder, CO 80309-0395, USA
e-mail: agnes.szendrei@colorado.edu

https://doi.org/10.1017/S144678871500052X Published online by Cambridge University Press