GROWTH RATES OF FINITE ALGEBRAS

KEITH A. KEARNES, EMIL W. KISS, AND ÁGNES SZENDREI

ABSTRACT. We investigate the function $d_{\mathbf{A}}(n)$, which gives the size of a smallest generating set for $\mathbf{A}^n$, in the case where $\mathbf{A}$ is a finite algebra.

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 $\mathbf{A}^n$. These functions map positive integers to positive integers and satisfy the bi-implication

$$d_{\mathbf{A}}(n) \leq g \iff n \leq h_{\mathbf{A}}(g),$$

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 $d_{\mathbf{A}}, h_{\mathbf{A}} \colon \mathbb{Z}^+ \to \mathbb{Z}^+$ are monotone 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.

In the literature, the notation is slightly different. The least size of a generating set for $\mathbf{A}$ is denoted $d(\mathbf{A})$. The *growth sequence* for $\mathbf{A}$ is then defined to be the sequence

$$\mathbf{d}(\mathbf{A}) := (d(\mathbf{A}), d(\mathbf{A}^2), d(\mathbf{A}^3), \ldots).$$

The largest n such that $\mathbf{A}^n$ is g-generated is denoted $h(g, \mathbf{A})$. The relationship between the two notations is $\mathbf{d}(\mathbf{A}) = (d_{\mathbf{A}}(1), d_{\mathbf{A}}(2), \ldots)$ and $h(g, \mathbf{A}) = h_{\mathbf{A}}(g)$. In this paper we will use the notation $d_{\mathbf{A}}(n)$ and $h_{\mathbf{A}}(n)$ in place of $\mathbf{d}(\mathbf{A})$ and $h(n, \mathbf{A})$, and use phrases "growth rate" or "growth function" in place of "growth sequence".

The h function was studied by Philip Hall in [14]. Hall proved that if A is a finite, simple, nonabelian group, then

$$h_A(g) = \frac{1}{|\operatorname{Aut}(A)|} \sum_{H \leq A} \mu(H)|H|^g, \tag{1.1}$$

where μ is the Möbius function of the subgroup lattice of A. An estimate for the d function of A was derived from (1.1) by James Wiegold in [36]. Wiegold showed that $d_{\mathbf{A}}(n)$ is one of the three integers nearest $\log_{|A|}(n) + \log_{|A|}(|\operatorname{Aut}(A)|)$. Wiegold's paper

This material is based upon work supported by the Hungarian National Foundation for Scientific Research (OTKA) grants no. K77409 and K83219.

initiated a program of research into growth rates of groups [5, 6, 7, 8, 9, 10, 11, 23, 25, 29, 30, 34, 37, 38, 39, 41, 42]. The program expanded to include the investigation of growth rates of semigroups, in [31, 40], and later to include the investigation of growth rates of arbitrary algebraic structures, in [13, 33]. Our own investigations into growth rates of finite algebras, on which we are reporting here, was stimulated by [32] and [16].

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[1] 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 ($d_{\mathbf{A}}(n) \in \Theta(n)$), [38].

(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 ($d_{\mathbf{A}}(n) \in 2^{\Theta(n)}$), [40].

Herbert Riedel partially extended item (I) to congruence uniform varieties in [33] by proving that finite perfect algebras in such varieties have logarithmic growth rate. He adds that *"Unfortunately, Wiegold's results for the imperfect case do not hold even for modules."* It is not clear what this sentence means. It is not very difficult to show that any finite imperfect algebra in a congruence uniform variety has linear growth, so either Riedel's remark is a mistake or he means that Wiegold's sharp estimate on the leading coefficient of the linear growth function is not valid for every finite module. Riedel concludes his paper by asking whether his result about perfect algebras can be extended from congruence uniform varieties to congruence permutable varieties. We will see that it does.

The paper [32] by Martyn Quick and Nikola Ruškuc extends item (I) to any variety of rings, modules, k-algebras or Lie algebras, but also fall short of extending item (I) to arbitrary congruence uniform varieties.

The results from [32] 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 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 [32] 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},[2] then $\mathbf{A}$ has a logarithmic growth rate if it is perfect and has a linear growth rate if it is imperfect. Hence the arguments in [32] can be used to extend Wiegold's result (I) from finite groups to expansions of finite groups.

Our main results are also best expressed in the language of algebras realizing a set of identities. Call a term *basic* if it contains at most one nonnullary function

[1] G is perfect if $[G, G] = G$.

[2] The statement "$\mathbf{A}$ realizes Σ_{Grp}" is equivalent to what elsewhere might be expressed as "$\mathbf{A}$ is an expansion of a group" or "$\mathbf{A}$ has underlying group structure".

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 a $\mathcal{L}$-term $F(x_1, \ldots, x_m)$ is a *p-pointed, k-cube term* if there is a $k \times m$ matrix M consisting of variables and at most p distinct constant symbols, with every column of M containing a symbol different from x, such that

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

(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,

$$\Sigma \models B \begin{pmatrix} 1 & x \\ x & 1 \end{pmatrix} \approx \begin{pmatrix} x \\ x \end{pmatrix}, \tag{1.4}$$

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,

$$\Sigma \models M \begin{pmatrix} y & x & x \\ x & y & x \\ x & x & y \end{pmatrix} \approx \begin{pmatrix} x \\ x \\ x \end{pmatrix}, \tag{1.5}$$

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 in a language with only finitely many constant symbols.

(1) If Σ does not entail the existence of a pointed cube term, then Σ imposes no restriction on growth rates of finite algebras (Theorem 4.2.3). That is, for every finite algebra $\mathbf{A}$ there is a finite algebra $\mathbf{B}$ realizing Σ such that $d_{\mathbf{B}} = d_{\mathbf{A}}$.[3]
(2) If Σ entails the existence of an m-ary p-pointed k-cube term, then any finite algebra realizing Σ has growth rate that is bounded above by a polynomial of degree at most $\log_w(mp)$, where $w = 2k/(2k-1)$ (Theorem 4.3.1). Moreover, there exist finite algebras with 1-pointed cube terms whose growth rate

[3]We call Σ with this property *nonrestrictive*. Otherwise Σ is *restrictive*.

is asymptotically equivalent to a polynomial of any prescribed degree (Theorem 4.4.1).

(3) If Σ entails the existence of a 0-pointed cube term, and $\mathbf{A}$ realizes Σ, then the growth rate of $\mathbf{A}$ is logarithmic if $\mathbf{A}$ is perfect and linear if $\mathbf{A}$ is imperfect (Theorem 5.4.1).

Item (3) extends Wiegold's result (I) to a setting that includes, as special cases, any algebra with a Maltsev term or any algebra with a majority term. To further specialize,[4] this includes the case of finite algebras in congruence uniform varieties.

Concerning Wiegold's result (II), Remark 4.15 of [32] 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."* Item (2) on the list establishes the existence of such examples.

Item (1) on this list is perhaps the most striking of all. It implies, for example, that any function that arises as the d-function of an arbitrary finite algebra must also arise as the d-function of a finite algebra in a congruence distributive and congruence 3-permutable variety. This is the only item on this list that requires the assumption that only finitely many distinct constants appear in Σ. Example 3.2.4 shows that this assumption is necessary.

In addition to our main results, we give a new proof of Kelly's Completeness Theorem for basic identities (Theorem 3.1.1). 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 4.2.4). Some of our results are valid and interesting for infinite algebras. This is discussed in Section 6.

2. Preliminaries

2.1. **Notation.** $[n]$ denotes the set $\{1, \dots, n\}$. A tuple in A^n may be denoted $(a_1, \dots, a_n)$ or $\mathbf{a}$. A tuple $(a, a, \dots, a) \in A^n$ 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, e.g. $\mathbf{A}$, while the universe of a structure is denoted by the same character in italic font, e.g., A. The subuniverse of $\mathbf{A}$ generated by a subset $G \subseteq A$ is denoted $\langle G \rangle$.

We will use Big Oh 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 constants M and N such that $|f(x)| \leq M|g(x)|$ for all $x > N$. We write $f \in \Omega(g)$ and $f = \Omega(g)$ to mean that there are 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.

[4] A finite algebra in a congruence uniform variety has a Maltsev term, according to [28].

2.2. **Easy estimates.**

Theorem 2.2.1. *Let* $\mathbf{A}$ *be a finite 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 [32].) *If* $\mathbf{B}$ *is the full polynomial expansion of* $\mathbf{A}$*, 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 $d_{\mathbf{A}}(kn)$ count the number of elements in a smallest 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 generating set for $\mathbf{A}^n$, then $\varphi(G)$ is a generating set for $\mathbf{B}^n$. Hence $d_{\mathbf{B}}(n) \leq |G| = d_{\mathbf{A}}(n)$.

For (3), if $G \subseteq A^n$ is a smallest 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 generating set for $\mathbf{B}^n$ and let $H \subseteq A$ be a smallest 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, hence is generated from $G \cup \widehat{H}$ by term operations of $\mathbf{A}$ acting coordinatewise. This proves $d_{\mathbf{A}}(n) \leq |G| + |H| = d_{\mathbf{B}}(n) + d_{\mathbf{A}}(1)$, from which the left-hand inequality follows. □

Theorem 2.2.2. *If* $\mathbf{A}$ *is a finite algebra of more than one element, then*

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

and

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

Moreover,

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

Proof. It follows from Theorem 2.2.1 (3) that, among all algebras with universe A, the algebra equipped with no operations has the smallest d-function and the algebra equipped with all finitary operations has the largest d-function. These two algebras are also extremes for the h-function.

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

Now assume that $\mathbf{A}$ is equipped with all finitary operations, i.e., $\mathbf{A}$ is primal. The n-generated free algebra in the variety generated by $\mathbf{A}$ is isomorphic to $\mathbf{A}^{|A|^n}$

(Theorem 3 of [12]). 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, $d_{\mathbf{A}}(n) = \lceil \log_{|A|}(n) \rceil$, 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

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

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

$$h_{\mathbf{A}}(k) < n \iff k < d_{\mathbf{A}}(n), \tag{2.2}$$

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

Allow "$\forall N$" to stand for "for all large n". We have

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

because the monotonicity of g guarantees that $\lceil g(\lfloor g^{-1}(n) \rfloor) \rceil) \leq n$. The reverse implication is proved 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 2.2.4. *If $\mathbf{A}$ is a finite algebra and $f_{\mathcal{V}}$ is the free spectrum of the variety $\mathcal{V} = \mathcal{V}(\mathbf{A})$, then $h_{\mathbf{A}}(n) \leq \log_{|A|}(f_{\mathcal{V}}(n))$. In particular,*

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

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

If $f_{\mathcal{V}}(n) \in O(n^m)$ for some fixed $m \in \mathbb{Z}^+$, then $\log(f_{\mathcal{V}}(n)) \in O(\log(n))$, hence $h_{\mathbf{A}}(n) \in O(\log(n))$. It follows from Theorem 2.2.2 (3) that $d_{\mathbf{A}}(n) \in 2^{\Theta(n)}$.

If $f_{\mathcal{V}}(n) \in 2^{O(n)}$, then $\log(f_{\mathcal{V}}(n)) \in O(n)$, hence $h_{\mathbf{A}}(n) \in O(n)$. It follows from Theorem 2.2.2 (2) that $d_{\mathbf{A}}(n) \in \Omega(n)$. □

Corollary 2.2.5. *Let $\mathbf{A}$ be a finite algebra and let $\mathbf{B}$ be a homomorphic image of $\mathbf{A}^k$ for some k.*

1. *If $\mathbf{B}$ 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 (1), Theorem 5.3 of [20] proves that a finite strongly rectangular algebra generates a variety with free spectrum bounded above by a polynomial. By Theorem 2.2.4, $d_{\mathbf{A}}(n) \in 2^{\Theta(n)}$ in this case. The strong abelian property is more restrictive than the strong rectangular property by Lemma 2.2 (11) of [20].

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 [3], so Theorem 2.2.4 (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 iff $\mathbf{A}$ is abelian and has a Maltsev term iff $\mathbf{A}$ is abelian and has a Maltsev polynomial.

Theorem 2.2.6. *If* $\mathbf{A}$ *is a finite affine algebra of more than one element, then* $d_{\mathbf{A}}(n) \in \Theta(n)$.

Proof. If $\mathbf{M}$ is a module, then the set of tuples in $\mathbf{M}^n$ with exactly one nonzero entry is a generating set for $\mathbf{M}^n$ of size $\leq |M|n$. Hence $d_{\mathbf{M}}(n) \in O(n)$.

Theorem 2.2.1 (4) implies that if $d_{\mathbf{M}}(n) \in O(n)$, then any finite algebra polynomially equivalent to $\mathbf{M}$ has the same property.

It follows from Corollary 2.2.5 (2) that if $\mathbf{A}$ is abelian, then $d_{\mathbf{A}}(n) \in \Omega(n)$, so together with the previous conclusions we get that, if $\mathbf{A}$ is affine, then $d_{\mathbf{A}}(n) \in \Theta(n)$. □

3. Kelly's Completeness Theorem

In Subsection 3.1 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 Subsection 3.2 we construct a simpler model by modifying the construction from the Completeness Theorem. The simpler model is not adequate for proving the Completeness Theorem, but it is exactly we need for our investigation of growth rates. Strictly speaking, it would be possible and easier to construct only the simpler model, but the Completeness Theorem is important for testing claims about specific examples, and no proof of the theorem appears in the literature; hence we include one here.

3.1. 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 3.1.1. X is *large enough* if (a) X contains at least 2 variables, (b) $|X| \geq \text{arity}(F)$ 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 3.1.1. (David Kelly, [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.[5]

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 3.1.2. *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 $y \in X \setminus \{x\}$ (Rule (iv)); $(h \approx y)$ (Rule (ii)); and $(x \approx y)$ (Rule (iii)). □

Proof of Theorem 3.1.1. 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.

[5]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.

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: i.e., $s \equiv t$ 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 $F^{\mathbf{M}}([a_1], \ldots, [a_m]) = [F(a_1, \ldots, a_m)]$. However, this does not work, since $F(a_1, \ldots, a_m)$ will not be a basic term unless all the a_i's 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 $F^{\mathbf{M}}([a_1], \ldots, [a_m])$ 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 partial injective functions $\imath\colon M \to X \cup C$ that satisfy the following conditions:

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

According to Lemma 3.1.2, 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 $\imath \in \mathcal{I}$.

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 $([a_1], \ldots, [a_m]) \in M^m$, then choose $\imath \in \mathcal{I}$ that is defined on $S := \{[a_1], \ldots, [a_m]\}$. Note that $f := F(\imath[a_1], \ldots, \imath[a_m])$ is a basic term, since it is a function symbol applied to elements of $X \cup C$. We refer to this term to define $F^{\mathbf{M}}([a_1], \ldots, [a_m])$.

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 3.1.2 proves that x must occur in f, i.e., $x = \imath[a_k]$ 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 $F^{\mathbf{M}}([a_1], \dots, [a_m]) = [0]$.

Before proceeding, we point out that there is overlap in Cases 1 and 2, but no conflict in the definition of $F^{\mathbf{M}}([a_1], \dots, [a_m])$. 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 $F^{\mathbf{M}}([a_1], \dots, [a_m]) = [f] = [x]$ while Case 2 defines $F^{\mathbf{M}}([a_1], \dots, [a_m]) = \imath^{-1}(x) = [x]$.

Claim 3.1.3. *$F^{\mathbf{M}} \colon M^m \to M$ is a well-defined function.*

Choose $([a_1], \dots, [a_m]) \in M^m$ and define $S = \{[a_1], \dots, [a_m]\}$. There exist elements of $\mathcal{I}$ defined on S, because this set has size $\leq \text{arity}(F) \leq |X|$. Suppose that $\imath, \jmath \in \mathcal{I}$ are both defined on this set. Let $f = F(\imath[a_1], \dots, \imath[a_m])$ and $g = F(\jmath[a_1], \dots, \jmath[a_m])$. To show that $F^{\mathbf{M}}([a_1], \dots, [a_m])$ 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 $F^{\mathbf{M}}([a_1], \dots, [a_m])$, the assigned value depends only on the term $f = F(\imath[a_1], \dots, \imath[a_m]) = F(\imath|_S[a_1], \dots, \imath|_S[a_m])$. Thus, to complete the proof of Claim 3.1.3, 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 $\Gamma \colon X \cup C \to X \cup C$ 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 $\Sigma \vdash_X f = F(\imath[a_1], \dots, \imath[a_m]) \approx h$. 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 $F^{\mathbf{M}}([a_1], \dots, [a_m])$ 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 $\Sigma \vdash_X F(\imath[a_1], \dots, \imath[a_m]) \approx \imath[a_k]$ 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 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 $F^{\mathbf{M}}([a_1], \ldots, [a_m]) = [0]$.

$\mathbf{M}$ is defined. We now argue that $\mathbf{M}$ 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 $\diamondsuit[\alpha]$. We will, in fact, write s as $F[\alpha]$ in either case, but will remember that F may equal the artificially introduced symbol $\diamondsuit$. The identity $s \approx t$ takes the form $F[\alpha] \approx G[\beta]$.

A valuation in $\mathbf{M}$ is a function $v\colon X \cup C \to M$ satisfying $v(c) = c^{\mathbf{M}} = [c]$ for each $c \in C$. To show that $\mathbf{M}$ satisfies $F[\alpha] \approx G[\beta]$ we must show that $F^{\mathbf{M}}[v \circ \alpha] \approx G^{\mathbf{M}}[v \circ \beta]$ for any valuation v. Choose $\imath \in \mathcal{I}$ that is defined on the set $\operatorname{im}(v \circ \alpha) \cup \operatorname{im}(v \circ \beta)$. 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 $F^{\mathbf{M}}[v \circ \alpha]$ and $G^{\mathbf{M}}[v \circ \beta]$ are defined in reference to the terms $f := F[\imath \circ v \circ \alpha]$ and $g := G[\imath \circ v \circ \beta]$ respectively.

Claim 3.1.4. $[f] = [g]$.

Observe that $(\imath \circ v)(c) = \imath[c] = d$, 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 $(i \circ v)$ on the variables in $\operatorname{im}(\alpha) \cup \operatorname{im}(\beta)$, but is the identity on C, then applications of Rule (v) show that $\Sigma \vdash_X F[\imath \circ v \circ \alpha] \approx F[\Gamma \circ \alpha]$ and $\Sigma \vdash_X G[\imath \circ v \circ \beta] \approx G[\Gamma \circ \beta]$. From Rule (iv), the fact that $\Sigma \vdash_X F[\alpha] \approx G[\beta]$ implies that $\Sigma \vdash_X F[\Gamma \circ \alpha] \approx G[\Gamma \circ \beta]$. Hence

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

from which we get $[f] = [g]$.

We conclude the argument that $\mathbf{M}$ 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 $F^{\mathbf{M}}[v \circ \alpha] = [f] = [g] = G^{\mathbf{M}}[v \circ \beta]$.

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

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

To complete the proof of the theorem we must show that $\mathbf{M}$ does not satisfy φ. Suppose φ has the form $F[\alpha] \approx G[\beta]$. Define the canonical valuation to be

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

Choose $\imath \in \mathcal{I}$ that is defined on $\operatorname{im}(v \circ \alpha) \cup \operatorname{im}(v \circ \beta)$. It follows from the definitions that $\imath \circ v\colon X \cup C \to X \cup C$ fixes every variable in $\operatorname{im}(\alpha) \cup \operatorname{im}(\beta)$. If Γ is the identity function on $X \cup C$, then Γ agrees with $\imath \circ v$ on the variables in $\operatorname{im}(v \circ \alpha) \cup \operatorname{im}(v \circ \beta)$, while $(\imath \circ v)(c) = d$ is the $<$-least constant symbol in the class of $c = \Gamma(c)$. Just as in the proof of Claim 3.1.4, we obtain $\Sigma \vdash_X f = F[\imath \circ v \circ \alpha] \approx F[\Gamma \circ \alpha] = F[\alpha]$ and

$\Sigma \vdash_X g = G[\imath \circ v \circ \beta] \approx G[\Gamma \circ \beta] = G[\beta]$. 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 $F^{\mathbf{M}}(v \circ \alpha) = [F[\alpha]]$, and similarly $G^{\mathbf{M}}(v \circ \beta) = [G[\beta]]$. 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, v witnesses that $\mathbf{M}$ does not satisfy φ. □

Theorem 3.1.1 establishes that if X and Y are two sets of variables that are large enough, then $\Sigma \vdash_X \varphi$ holds iff $\Sigma \vdash_Y \varphi$, 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.

3.2. **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 3.1.1 may be infinite, so we explain how to produce finite models.

Definition 3.2.1. 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 $\mathbf{M}$ constructed in the proof of Theorem 3.1.1 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 $\{[c] \mid c \in C\}$.

Let F be an m-ary function symbol of $\mathcal{L}$. If $([a_1], \ldots, [a_m]) \in V^n$, then let $a_i' = a_i$ if $a_i \in Y \cup C$ and $a_i' = z$ if $a_i = 0$. Define $F^{\mathbf{V}}([a_1], \ldots, [a_m]) = [t]$ if there exists $t \in Y \cup C$ such that $\Sigma \vdash F(a_1', \ldots, a_m') \approx t$, and define $F^{\mathbf{V}}([a_1], \ldots, [a_m]) = [0]$ if there is no such t.

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

Theorem 3.2.2. $\mathbf{V}$ *is a model of* Σ.

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

The function v is also a valuation in $\mathbf{M}$, because $V \subseteq M$. Since $\mathbf{M}$ is a model of Σ, we get $F^{\mathbf{M}}[v \circ \alpha] = G^{\mathbf{M}}[v \circ \beta]$. Choose $\imath \in \mathcal{I}$ defined on the set $\mathrm{im}(v \circ \alpha) \cup \mathrm{im}(v \circ \beta)$ such that $\imath[0] = z$, if $[0]$ is in this set. Let $f = F[\imath \circ v \circ \alpha]$ and $g = G[\imath \circ v \circ \beta]$. As in the proof of Claim 3.1.4, if $\Gamma \colon X \cup C \to X \cup C$ is the identity on C and agrees with $\imath \circ v$ 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 $F(a_1', \ldots, a_m')$ of Definition 3.2.1 is none other than f. $F^{\mathbf{V}}[v \circ \alpha] = [t]$ for some $t \in Y \cup C$ if and only if $\Sigma \vdash f = F(a_1', \ldots, a_m') \approx t$. But since $\Sigma \vdash f \approx g$ we also get $G^{\mathbf{V}}[v \circ \beta] = [t]$. This shows that $F^{\mathbf{V}}[v \circ \alpha]$ and $G^{\mathbf{V}}[v \circ \beta]$ are equal when at least one of them is not $[0]$. Of course, they are also equal when both of them equal $[0]$, so $F^{\mathbf{V}}(v \circ \alpha) = G^{\mathbf{V}}(v \circ \beta)$. □

Corollary 3.2.3. *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 3.2.2. □

Corollary 3.2.3 is close to the best possible result about sizes of models of a set of basic identities, as the next example shows.

Example 3.2.4. Let C be a set of constant symbols and let $\mathcal{B} = \{B_{c,d} \mid c, d \in C\}$ be a set of binary function symbols. Let

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

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 $|C|$.

In the case where C is infinite, Σ does restrict growth rates of finite algebras for the artificial reason that Σ has no nontrivial finite models. On the other hand, Σ does not entail the existence of a pointed cube term according to the criterion of Lemma 4.2.1.

4. Restrictive Σ

Recall (from footnote 3) that a set Σ of identities is called nonrestrictive if, whenever $\mathbf{A}$ is a finite algebra, there is a finite algebra $\mathbf{B}$ realizing Σ such that $d_{\mathbf{B}}(n) = d_{\mathbf{A}}(n)$. Otherwise Σ is restrictive. At first glance this definition might seem too mild to be of any use. Surely the realization of any nontrivial Σ ought to affect growth rates? In fact, this is not so. Only relatively strong Σ will enforce any kind of restriction on growth rates. In Subsection 4.1 we prove that the growth rates of finite partial algebras are the same as the growth rates for finite total algebras. In Subsection 4.2 we prove that if Σ is restrictive, then it entails the existence of a pointed cube term. The converse is proved in Subsection 4.3, by showing that a finite algebra with a pointed cube term has growth rate the 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 Subsection 4.4 we describe an example of a 3-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 Subsection 4.5 we classify the finite abelian algebras that have pointed cube terms. In Subsection 4.6 give a way of recognizing when 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.

4.1. **Growth rates of partial algebras.** A partial algebra is a set equipped with a set of partial operations. 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. We will learn in this subsection that a function arises as the growth function of a partial algebra if and only if it arises as the growth function of a total algebra.

Definition 4.1.1. 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

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

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

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

Theorem 4.1.2. *If $\mathbf{A}$ is a partial algebra of more than one element, then $\mathbf{A}^n$ and $\mathbf{A}_0^n$ have the same minimal generating sets. Hence $d_{\mathbf{A}_0} = d_{\mathbf{A}}$.*

Proof. It suffices to prove the following two statements: (i) any generating set for $\mathbf{A}^n$ is a generating set for $\mathbf{A}_0^n$, and (ii) any *minimal* generating set for $\mathbf{A}_0^n$ is a subset of A^n and a generating set for $\mathbf{A}^n$.

In this paragraph we prove (i). 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 (ii). Assume that $H \subseteq A_0^n$ is a minimal 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

$$Z(\mathbf{a}_1) \cup \cdots \cup Z(\mathbf{a}_m) \subseteq Z(F(\mathbf{a}_1, \ldots, \mathbf{a}_m)), \tag{4.1}$$

since 0 is absorbing for every basic operation. If the right-hand side is empty, then the left-hand side is empty as well; i.e., tuples with empty zero sets can be generated

only by tuples with empty zero sets. Said 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, and, by (i), these tuples suffice to generate $\mathbf{A}_0^n$; hence $H \cap A^n$ generates $\mathbf{A}_0^n$. Therefore, if $H \subseteq A_0^n$ is a minimal generating set for $\mathbf{A}_0^n$, then $H \subseteq A^n$. 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 $\mathbf{A}^n$. □

4.2. **Restrictive Σ forces a pointed cube term.** Let Σ be a set of basic identities in a language $\mathcal{L}$ whose set C of constant symbols is finite. Given an algebra $\mathbf{A}$ in an arbitrary language, 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 3.2.1. 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 Subsection 4.1 $p+1$ times to $\mathbf{A}$ to produce a sequence $\mathbf{A}$, $\mathbf{A}_{z_1}$, $\mathbf{A}_{z_1,z_2}$, …, 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 3.2.1. 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 3.2.1. 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

$$F^{\mathbf{A}_\Sigma}(x_1, \ldots, x_m) := \varphi(F^{\mathbf{V}}(\varphi^{-1}(x_1), \ldots, \varphi^{-1}(x_m))) \tag{4.2}$$

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 and all operations of the form (4.2). Under this definition the function φ is an isomorphism from $\mathbf{V}$ to the $\mathcal{L}$-reduct of $\mathbf{A}_\Sigma$.

Lemma 4.2.1. *Let $\mathbf{A}$ be a finite algebra with more than one element and let Σ be a set of basic identities involving finitely many constant symbols. Let $\mathcal{V}$ be the variety axiomatized by Σ. The following statements about a positive integer k are equivalent.*

1. *$\mathcal{V}$ has a pointed k-cube term.*
2. *For $n \geq k$, the family of minimal generating subsets of $\mathbf{A}^n$ is different from the family of minimal generating subsets of $\mathbf{A}_\Sigma^n$.*
3. *$\mathcal{V}$ has a pointed k-cube term of the form $F(x_1, \ldots, x_m)$, where $m > 1$, F is a function symbol occurring in Σ, and the variables $x_1, \ldots, x_m$ are distinct.*

Proof. [(1) ⇒ (2)] Let $\mathsf{P}(x_1, \ldots, x_m)$ 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

$$\mathsf{P}(\mathbf{y}_1, \dots, \mathbf{y}_m) = \mathsf{P}\left(\begin{bmatrix} y_{1,1} \\ \vdots \\ y_{k,1} \end{bmatrix}, \cdots, \begin{bmatrix} y_{1,m} \\ \vdots \\ y_{k,m} \end{bmatrix} \right) \approx \begin{bmatrix} x \\ \vdots \\ x \end{bmatrix}. \tag{4.3}$$

Let $G \subseteq \mathbf{A}^n$ be a minimal generating set for some $n \geq k$, and let $\mathbf{a} \in G$ be a tuple. Using the row identities of (4.3), solve the equation $\mathsf{P}(\mathbf{b}_1, \dots, \mathbf{b}_m) = \mathbf{a}$ for the $\mathbf{b}_i$'s, row by row, according to the following rules. In the i-th 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, $\mathbf{b}_i \in A^n_\Sigma$ for all i and $\mathsf{P}(\mathbf{b}_1, \dots, \mathbf{b}_m) = \mathbf{a}$. Moreover, since each column $\mathbf{y}_i$ in (4.3) 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 $\{z_1, \dots, z_p, 0\}$. Hence $\mathbf{b}_i \in A^n_\Sigma \setminus A^n$ for all i.

Since G is a generating set for $\mathbf{A}^n$, it is a generating set for $\mathbf{A}^n_\Sigma$. Since $\mathbf{a}$ can be generated from $\{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ using P, it follows that

$$G' = (G \setminus \{\mathbf{a}\}) \cup \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$$

is a generating set for $\mathbf{A}_\Sigma$. Hence there exists a minimal generating subset $G'' \subseteq G'$ for $\mathbf{A}^n_\Sigma$. If $\mathbf{A}^n$ and $\mathbf{A}^n_\Sigma$ had the same minimal generating sets, then G'' would be a subset of $G' \cap A^n = G \setminus \{\mathbf{a}\}$. But this is not so, since $G \setminus \{\mathbf{a}\}$ is a proper subset of the minimal generating set G of $\mathbf{A}^n$, hence it is not itself a generating set.

[(2) $\Rightarrow$ (3)] By repeated application of Theorem 4.1.2 we know that $\mathbf{A}^n$ has the same minimal generating sets as $\mathbf{A}^n_{z_1,\dots,z_p,0}$. Our assumption is that these minimal generating sets are not the same as the minimal generating sets of the expansion $\mathbf{A}^n_\Sigma$ of $\mathbf{A}^n_{z_1,\dots,z_p,0}$. Since every generating set for $\mathbf{A}^n_{z_1,\dots,z_p,0}$ is a generating set for the expansion $\mathbf{A}^n_\Sigma$, there must be a generating set G for $\mathbf{A}^n_\Sigma$ that is not a generating set for $\mathbf{A}^n_{z_1,\dots,z_p,0}$. The subuniverse $S = \langle G \rangle$ of $\mathbf{A}^n_{z_1,\dots,z_p,0}$ cannot contain all of A^n, since A^n is a generating set for $\mathbf{A}^n_{z_1,\dots,z_p,0}$, and G is not.

Claim 4.2.2. *The set $H = S \cup (A^n_{z_1,\dots,z_p,0} \setminus A^n)$ is a proper subuniverse of $\mathbf{A}^n_{z_1,\dots,z_p,0}$.*

Choose tuples $\mathbf{h}_1, \dots, \mathbf{h}_r \in H$ and operate on them to produce some element $\mathbf{g} = E^{\mathbf{A}^n_{z_1,\dots,z_p,0}}(\mathbf{h}_1, \dots, \mathbf{h}_r)$. If some $\mathbf{h}_i$ belongs to $A^n_{z_1,\dots,z_k,0} \setminus A^n$, then so does $\mathbf{g}$, by (4.1). Otherwise all $\mathbf{h}_i$ belong to S, in which case $\mathbf{g}$ does, since S is a subuniverse. $H = S \cup (A^n_{z_1,\dots,z_p,0} \setminus A^n)$ is proper, since $H \cap A^n = S \cap A^n \neq A^n$.

Since the proper subuniverse H of $\mathbf{A}^n_{z_1,\dots,z_p,0}$ contains the set G, which generates the algebra $\mathbf{A}^n_\Sigma$, and contains the interpretations of the $\mathcal{L}$-constants, it cannot be closed under the the interpretations of the function symbols of $\mathcal{L}$. Hence there is a

tuple $\mathbf{a} \notin H$, an m-ary function symbol F, and m tuples $\mathbf{b}_1, \dots, \mathbf{b}_m \in H$ such that $F^{\mathbf{A}_\Sigma}(\mathbf{b}_1, \dots, \mathbf{b}_m) = \mathbf{a}$. Necessarily $\mathbf{a} \in A^n$.

Using the isomorphism φ from $\mathbf{V}$ to the $\mathcal{L}$-reduct of $\mathbf{A}_\Sigma$, we obtain that there is a tuple $\mathbf{y} = \varphi^{-1}(\mathbf{a}) \in \varphi^{-1}(A^n) = [Y]^n$ and tuples $\mathbf{v}_i = \varphi^{-1}(\mathbf{b}_i) \neq \mathbf{y}$ such that $F^{\mathbf{V}}(\mathbf{v}_1, \dots, \mathbf{v}_m) = \mathbf{y}$. Since $\mathbf{v}_1 \neq \mathbf{y}$, there is a coordinate ℓ where these tuples differ. In the ℓ-th coordinate we have $F^{\mathbf{V}}([v_{\ell,1}], \dots, [v_{\ell,m}]) = [y_\ell]$ 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 $\mathbf{V}$,

$$\Sigma \vdash F(v'_{\ell,1}, \dots, v'_{\ell,m}) \approx y_\ell, \tag{4.4}$$

where $v'_{\ell,j} = v_{\ell,j}$ when $v_{\ell,j} \in Y \cup C$ and $v'_{\ell,j} = z$ 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, (4.4) can be rewritten as

$$\Sigma \vdash F(m_{1,1}, \dots, m_{1,m}) \approx x,$$

where each $m_{1,j}$ is a variable or constant and $m_{1,1} \neq x$. Similarly, for each i, the fact that $\mathbf{v}_i \neq \mathbf{y}$ produces an identity

$$\Sigma \vdash F(m_{i,1}, \dots, m_{i,m}) \approx x,$$

where each $m_{i,j}$ is a variable or constant and $m_{i,i} \neq x$. Thus, it is a consequence of Σ that the row identities of

$$F([m_{i,j}]) \approx \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix}$$

hold. Since the diagonal elements of $[m_{i,j}]$ are not x, these identities make F a pointed cube term for $\mathcal{V}$.

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

The next result is the main result of this subsection.

Theorem 4.2.3. *Let Σ be a 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.*

Proof. Recall that "Σ is nonrestrictive" means that for every finite algebra $\mathbf{A}$ there is a finite algebra $\mathbf{B}$ realizing Σ such that $d_{\mathbf{B}} = d_{\mathbf{A}}$, "$\Sigma$ is restrictive" means the opposite.

If Σ is restrictive, then there must exist a finite algebra $\mathbf{A}$ with the property that $d_{\mathbf{B}} \neq d_{\mathbf{A}}$ whenever $\mathbf{B}$ is finite and realizes Σ. For $\mathbf{B} = \mathbf{A}_\Sigma$ the fact that $d_{\mathbf{B}} \neq d_{\mathbf{A}}$ implies that $\mathbf{A}^n$ and $\mathbf{A}^n_\Sigma$ do not have the same minimal generating sets for some n, so the theorem follows from Lemma 4.2.1 (1)$\Leftrightarrow$(2). □

One unplanned consequence of Lemma 4.2.1 is a procedure to decide if a strong Maltsev condition involving only basic identities implies the existence of a pointed cube term.

Corollary 4.2.4. *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 $F(x_1, \dots, x_m)$ is a pointed k-cube term, where $m > 1$, F is a function symbol occurring in Σ, and the variables $x_1, \dots, x_m$ are distinct.*

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 4.2.1 (1)$\Leftrightarrow$(3). □

That the property in the theorem statement can be decided follows from Theorem 3.1.1.

4.3. **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. As the title of the subsection suggests, we shall prove that if $\mathbf{A}$ is a finite algebra with a pointed cube term, then $d_{\mathbf{A}}(n)$ is bounded above by a polynomial. Since not all finite algebras have polynomially bounded growth rates, this suffices to show that Σ is restrictive when it entails the existence of a pointed cube term.

Theorem 4.3.1. *If $\mathbf{A}$ is a finite algebra with an m-ary, p-pointed, k-cube term, with $p \geq 1$, then $d_{\mathbf{A}}$ is bounded above by a polynomial of degree at most $\log_w(mp)$, where $w = 2k/(2k-1)$.*

Proof. Here is a coarse outline of the proof. Let $F(x_1, \dots, x_m)$ be a p-pointed k-cube term for $\mathbf{A}$. We shall describe a way to "process" a randomly selected tuple $\mathbf{a} \in A^n$ which accomplishes the following things. If $\mathbf{a}$ is not already "fully processed", then processing $\mathbf{a}$ produces tuples $\mathbf{b}_1, \dots, \mathbf{b}_m \in A^n$ with the properties that (i) $F^{\mathbf{A}}(\mathbf{b}_1, \dots, \mathbf{b}_m) = \mathbf{a}$ and (ii) each of $\mathbf{b}_1, \dots, \mathbf{b}_m$ is recognizably "more processed" than $\mathbf{a}$. It will follow that $\mathbf{A}^n$ can be generated by the set of fully processed tuples. A count will show that there are only polynomially many fully processed tuples.

Let's fix some notation for the proof. Suppose that the fact that $F(x_1, \dots, x_m)$ is a p-pointed k-cube term (with $p \geq 1$) is witnessed by identities

$$F(M) \approx \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix},$$

where M is a $k \times m$ matrix of variables and constant symbols, where each column contains a symbol that is not x. Since $p \geq 1$, we may assume that the only variable appearing in M is x, since we may choose a constant symbol c appearing in M, replace all instances of other variables in M by c (Kelly's Rule (iv)), thereby obtaining 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 rows identities, $F(R) \approx [x, \dots, x]^{\mathsf{T}}$, is fixed once and for all.

For example, if the original cube term identities were

$$F(M) = F\begin{pmatrix} 1 & x & 2 \\ x & y & 3 \end{pmatrix} \approx \begin{pmatrix} x \\ x \end{pmatrix},$$

then after replacing y with, say, the constant symbol 2, we obtain

$$F(R) = F\begin{pmatrix} 1 & x & 2 \\ x & 2 & 3 \end{pmatrix} \approx \begin{pmatrix} x \\ x \end{pmatrix}. \tag{4.5}$$

Finally we fix the order of the identities. For this example, once and for all, we declare the 'first' cube identity to be $F(1, x, 2) \approx x$ and the 'second' cube identity to be $F(x, 2, 3) \approx x$. There will be k such identities for a k-cube term.

We will need a function $\lambda\colon [m] \to [k]$ from the column indices to the row indices with the property that if $\lambda(j) = i$, then the matrix R has a constant symbol in the i, j-th position. Such λ exists because every column of R contains at least one constant symbol. For the example in (4.5), one could take $\lambda\colon [3] \to [2]$ to be the function $\lambda(1) = \lambda(3) = 1, \lambda(2) = 2$. (This is not the only choice.)

The 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 $[m] = \{1, \dots, m\}$. 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 $\mathbf{n}_{\sigma 1}, \dots, \mathbf{n}_{\sigma m}$.

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 $U = \{u_1, \dots, u_r\} \subseteq [n]$, consider it to be a linearly ordered set $u_1 < \ldots < u_r$ under the order inherited from $[n]$. Now, define $\pi(U) = (U_1, \dots, U_k)$ to be the ordered partition of U into k consecutive nonempty intervals

that are as equal sized as possible. That is, let

$$(U_1, \dots, U_k) = (\{u_1, u_2, \dots, u_{i_1}\}, \{u_{i_1+1}, \dots, u_{i_2}\}, \dots, \{u_{i_{k-1}+1}, \dots, u_{i_k} = u_r\}),$$

where

$$u_1 < u_2 < \cdots < u_{i_1} < u_{i_1+1} < \cdots < u_{i_2} < u_{i_{k-1}+1} < \cdots < u_{i_k} = u_r$$

(i.e., the cells of the partition are consecutive nonempty intervals) and

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

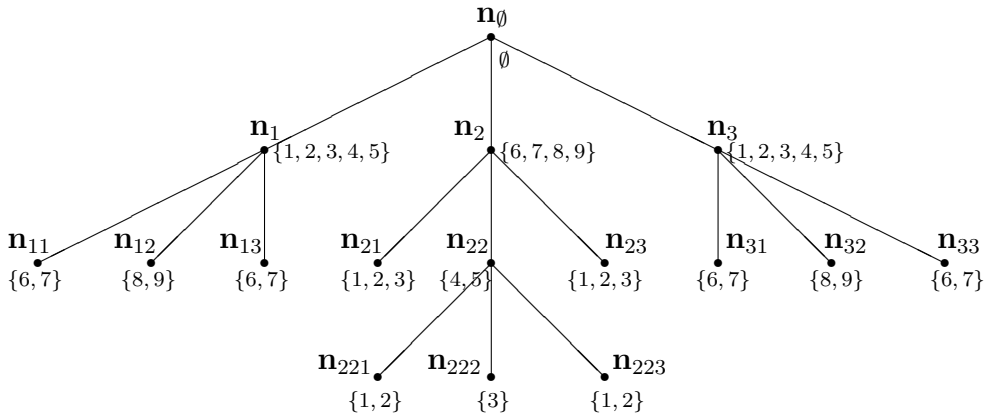
(i.e., the cells are as equal sized as possible). The k appearing here as the number of cells of the partition is the same k as the one in the assumption that F is a k-cube term. In order for $\pi(U)$ to be defined, it is necessary that $|U| \geq k$.

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 $\pi([n]\setminus V) = (U_1, \dots, U_k)$, then $\ell(\mathbf{n}_{\sigma i}) = U_{\lambda(i)}$.

In (2), if $[n] \setminus V$ 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's illustrate our progress with the example started back at (4.5). In this case, the reduced matrix R for F is 2×3, so $k = 2$ and $m = 3$. The following picture depicts the processing template in the case $[n] = [9] = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$.



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 i-th 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.

Before proceeding, we refine the coarse outline given in the first paragraph of this proof. 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 4.3.2. *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 $\mathbf{b}_1, \ldots, \mathbf{b}_m$ such that*

1. $F^{\mathbf{A}}(\mathbf{b}_1, \ldots, \mathbf{b}_m) = \mathbf{a}$.
2. *$\mathbf{b}_i$ is processed for node $\mathbf{n}_{\sigma i}$ for $i = 1, \ldots, m$.*
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 $i = 1, \ldots, m$.*

Let V be the union of labels on nodes between $\mathbf{n}_\sigma$ and $\mathbf{n}_\emptyset$. If $\pi([n] \setminus V) = (U_1, \ldots, U_k)$, then $\{V, U_1, \ldots, U_k\}$ is a partition of $[n]$ (with V possibly empty). For simplicity of expression, reorder coordinates so that $\mathbf{a}$ and $\mathbf{b}_i$ can be written $[\mathbf{a}_V, \mathbf{a}_{U_1}, \ldots, \mathbf{a}_{U_k}]^\mathsf{T}$ and $[\mathbf{b}_{i,V}, \mathbf{b}_{i,U_1}, \ldots, \mathbf{b}_{i,U_k}]^\mathsf{T}$, 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

$$F^{\mathbf{A}}(\mathbf{b}_1, \ldots, \mathbf{b}_m) = F^{\mathbf{A}}\left(\begin{bmatrix} \mathbf{b}_{1,V} \\ \mathbf{b}_{1,U_1} \\ \vdots \\ \mathbf{b}_{1,U_k} \end{bmatrix}, \ldots, \begin{bmatrix} \mathbf{b}_{m,V} \\ \mathbf{b}_{m,U_1} \\ \vdots \\ \mathbf{b}_{m,U_k} \end{bmatrix}\right) = \begin{bmatrix} \mathbf{a}_V \\ \mathbf{a}_{U_1} \\ \vdots \\ \mathbf{a}_{U_k} \end{bmatrix} = \mathbf{a} \tag{4.6}$$

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

Whether $W = V$ or $W = U_i$, to solve $F^{\mathbf{A}}(\mathbf{b}_{1,W}, \ldots, \mathbf{b}_{m,W}) = \mathbf{a}_W$ for the $\mathbf{b}_{i,W}$'s using a particular cube identity, take $\mathbf{b}_{i,W} = \mathbf{a}_W$ if there is an x in the i-th place of F in the cube identity, and take $\mathbf{b}_{i,W} = [c^{\mathbf{A}}, \ldots, c^{\mathbf{A}}]^\mathsf{T}$ if there is a c in the i-th 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 a cube identity which has some constant symbol $c \in P$ in the i-th place of F. Hence $\mathbf{b}_{i,U_{\lambda(i)}} = [c^{\mathbf{A}}, \ldots, c^{\mathbf{A}}]^\mathsf{T}$. 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 i-place of F, then $\mathbf{b}_{i,V} = \mathbf{a}_V$. If there is a constant symbol $c \in P$ in the i-th place of F, then $\mathbf{b}_{i,V} = [c^{\mathbf{A}}, \ldots, c^{\mathbf{A}}]^\mathsf{T}$. In the latter case, $\mathbf{b}$ is processed at all coordinates in V, 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.

Let's estimate the number of fully processed tuples by estimating the number of descriptions of such tuples. 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. For $\mathbf{a}$ to be processed at a node $\mathbf{n}$, there must be a constant symbol $c \in P$ such that the i-th coordinate of $\mathbf{a}$ is $c^{\mathbf{A}}$ for all $i \in \ell(\mathbf{n})$. So, beginning at the root, descend through the template choosing a new child $\mathbf{n}_{\sigma i}$ and a new constant symbol to assign to the coordinates in $\ell(\mathbf{n}_{\sigma i})$ at each step. Each such step can be accomplished in mp different ways, since each node has m children and there are p constant symbols to choose from. If the length of a longest path in the template is r, we will produce at

most $(mp)^r$ partial descriptions of fully processed tuples when we arrive at the leaves of the template. But there may be $k-1$ elements of $[n]$ that were not encountered in the labels of the maximal path, so all but perhaps $k-1$ coordinates of a typical fully processed tuple have been described. By filling in the last $k-1$ coordinates randomly, one sees that there are at most $(mp)^r|A|^{k-1}$ fully processed tuples. What remains is to estimate r, the length of the longest branch in the processing template.

Let $V_0 = \ell(\mathbf{n}_\emptyset) = \emptyset$. This represents the set of coordinate positions that have been processed before the processing begins, i.e., no coordinate positions. As we progress down a longest 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, $[n] \setminus V_\sigma$ are then divided evenly, $\pi([n] \setminus V_\sigma) = (U_1, \ldots, U_k)$, to appear as labels of the children of $\mathbf{n}_\sigma$. Thus, $|V_\emptyset| = 0$ and

$$|V_{\sigma i}| = |V_\sigma \cup \ell(\mathbf{n}_{\sigma i})| = |V_\sigma| + |\ell(\mathbf{n}_{\sigma i})|. \tag{4.7}$$

The useful parameter is the number $u_\sigma := |[n] \setminus V_\sigma| = n - |V_\sigma|$ of nodes that remain unprocessed after reaching $\mathbf{n}_\sigma$. This parameter satisfies $u_\emptyset = |[n] \setminus V_\emptyset| = n$ and, from (4.7),

$$u_{\sigma i} = (n - |V_{\sigma i}|) = (n - |V_\sigma|) - |\ell(\mathbf{n}_{\sigma i})| = u_\sigma - |\ell(\mathbf{n}_{\sigma i})|. \tag{4.8}$$

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

$$|\ell(\mathbf{n}_{\sigma i})| = |U_{\lambda(i)}| \geq \lfloor (n - |V_\sigma|)/k \rfloor = \lfloor u_\sigma / k \rfloor. \tag{4.9}$$

Combining (4.8) and (4.9) we have

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

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

Claim 4.3.3. *If $u \geq k \geq 1$, then $\left\lceil \left(\frac{k-1}{k}\right) u \right\rceil \leq \left(\frac{2k-1}{2k}\right) u$.*

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 $\left(\frac{2k-1}{2k}\right)^r n < k$, 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.

Solving $\left(\frac{2k-1}{2k}\right)^r n < k$ for r we obtain that any $r > \log_w(n/k)$, $w = \frac{2k}{2k-1}$, is an upper bound on the length of paths in the template; hence $r = \log_w(n/k) + 1$ is such a bound. Combining this with our earlier number $(mp)^r|A|^{k-1}$ estimating the number of fully processed tuples, we get that the number of such tuples is no more than

$$\begin{aligned}(mp)^{\log_w(n/k)+1}|A|^{k-1} &= (mp)^{\log_w(n/k)}(mp)|A|^{k-1}\\ &= (n/k)^{\log_w(mp)}(mp)|A|^{k-1}\\ &= n^{\log_w(mp)}(k^{-\log_w(mp)}(mp)|A|^{k-1}) \in O(n^{\log_w(mp)}).\end{aligned}$$

□

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 4.3.4. *If* $\mathbf{A}$ *is a finite 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 $c^{\mathbf{A}}$ in every processed coordinate position, and has at most $k-1$ unprocessed coordinate positions. Hence the set of tuples with a $c^{\mathbf{A}}$ in all but at most $k-1$ positions contains all the fully processed tuples, and therefore is a generating set for $\mathbf{A}^n$. The number of such tuples is

$$\binom{n}{k-1}|A|^{k-1} \in O(n^{k-1}).$$

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

$$F(M) \approx \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix}. \tag{4.10}$$

If $\mathbf{B}$ is the polynomial expansion of $\mathbf{A}$, then we can choose an element $c \in B = A$ and replace all variables other than x in (4.10) 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 2.2.1 (4). □

In Section 5 we shall improve this result by showing that Wiegold dichotomy holds for algebras with a 0-pointed k-cube term. In particular, this means that growth rates are at most linear for such algebras.

Let's combine the results of this subsection with the results of the previous subsection.

Theorem 4.3.5. *The following are equivalent for a set* Σ *of basic identities in which only finitely many constant symbols occur.*

(1) Σ *is restrictive.*
(2) *The variety axiomatized by* Σ *has a pointed cube term.*
(3) *The variety axiomatized by* Σ *has a pointed cube term of the form* $F(x_1, \dots, x_m)$*, where* $m > 1$*,* F *is a function symbol occurring in* Σ*, and the variables* $x_1, \dots, x_m$ *are distinct.*
(4) *If* $\mathbf{A}$ *is a finite algebra realizing* Σ*, then* $d_{\mathbf{A}}(n)$ *is bounded above by a polynomial.*
(5) *There is no finite algebra* $\mathbf{A}$ *realizing* Σ *such that* $d_{\mathbf{A}}(n) = 2^n$ *for all* n*.*

Proof. [(1) $\Rightarrow$ (2)] Theorem 4.2.3.
[(2) $\Leftrightarrow$ (3)] Lemma 4.2.1.
[(2) $\Rightarrow$ (4)] Theorem 4.3.1 and Corollary 4.3.4.
[(4) $\Rightarrow$ (5)] $d_{\mathbf{A}}(n) = 2^n$ is not bounded above by a polynomial.
[(5) $\Rightarrow$ (1)] There exists a finite algebra $\mathbf{A}$ with $d_{\mathbf{A}}(n) = 2^n$, namely the 2-element set equipped with no operations. □

4.4. **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, which we established in Corollary 4.3.4, is sharp.

Theorem 4.4.1. *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 with the desired growth rate, 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

$$d_{\mathbf{A}}(n) = \binom{n}{0} + q\binom{n}{1} + q^2\binom{n}{2} + \cdots + q^{k-1}\binom{n}{k-1} \tag{4.11}$$

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 an $(q+1)$-element partial algebra with $d_{\mathbf{A}}(n) \in \Theta(n^{k-1})$. (When $q = 1$ we will obtain a 2-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 $\mathbf{A}^n$.

Claim 4.4.2. *If $S \subseteq [n]$ and $G \subseteq A^n$, then let G_S denote the tuples in G that have support contained in S. If $\mathbf{a} \in \langle G \rangle$, then $\mathbf{a} \in \langle G_S \rangle$.*

In $\mathbf{A}$, we have

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

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 $i \notin \operatorname{supp}(\mathbf{b})$ if and only if $i \notin \operatorname{supp}(\mathbf{g}_i)$ for any $\mathbf{g}_i$. Equivalently,

$$\operatorname{supp}(F^{\mathbf{A}^n}(\mathbf{g}_1, \dots, \mathbf{g}_k)) = \bigcup_{i=1}^{k} \operatorname{supp}(\mathbf{g}_i) \tag{4.12}$$

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 (4.12), 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 4.4.3. *The tuple $\hat{1} = [1, 1, \dots, 1]^{\mathsf{T}}$ of empty support is an essential generator.*

This follows immediately from Claim 4.4.2.

Claim 4.4.4. *Any tuple whose support has size at most $k-1$ is an essential generator of $\mathbf{A}^n$.*

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 (4.12), the $\mathbf{x}_i$'s must be taken from the tuples whose support is contained in S. The equation to be solved is therefore:

$$F^{\mathbf{A}^n}(\mathbf{x}_1, \dots, \mathbf{x}_k) = F^{\mathbf{A}^n}\left(\begin{bmatrix} x_{1,1} \\ \vdots \\ x_{\ell,1} \\ \hline 1 \\ \vdots \\ 1 \end{bmatrix}, \dots, \begin{bmatrix} x_{1,k} \\ \vdots \\ x_{\ell,k} \\ \hline 1 \\ \vdots \\ 1 \end{bmatrix}\right) = \begin{bmatrix} b_1 \\ \vdots \\ b_\ell \\ \hline 1 \\ \vdots \\ 1 \end{bmatrix} = \mathbf{b}. \tag{4.13}$$

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

such columns, and $k > \ell$, there is a column $\mathbf{x}_j$ that contains no 1 above the dividers. Since the i-th 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 $\mathbf{b} \notin \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$, showing that $\mathbf{b}$ is indeed an essential generator.

Claim 4.4.5. *$\mathbf{A}^n$ is generated by the tuples whose support has size at most $k-1$.*

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

$$F^{\mathbf{A}^n}(\mathbf{x}_1, \dots, \mathbf{x}_k) = F^{\mathbf{A}^n}\left(\begin{bmatrix} x_{1,1} \\ \vdots \\ x_{\ell,1} \\ \hline 1 \\ \vdots \\ 1 \end{bmatrix}, \dots, \begin{bmatrix} x_{1,k} \\ \vdots \\ x_{\ell,k} \\ \hline 1 \\ \vdots \\ 1 \end{bmatrix}\right) = \begin{bmatrix} b_1 \\ \vdots \\ b_\ell \\ \hline 1 \\ \vdots \\ 1 \end{bmatrix} = \mathbf{b} \tag{4.14}$$

when $\ell \geq k$ in such a way that every column contains at least one 1 above the dividers and the i-th 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, then fill in the remaining entries above the dividers so that the i-th 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 $\mathbf{A}$ by one constant symbol 1 whose interpretation is $1^{\mathbf{B}} = 1$. The operation $F^{\mathbf{B}}$ still satisfies

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

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

$\mathbf{A}^n$ 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 $d_{\mathbf{B}}(n) = d_{\mathbf{A}}(n) - 1 = \sum_{i=1}^{k-1} q^i\binom{n}{i} \in O(n^{k-1})$. □

4.5. **Abelian algebras with pointed cube terms are affine.** Theorem 4.3.5 shows that a set Σ of basic identities is restrictive iff it forces polynomially bounded growth rates iff it entails the existence of a pointed cube term. In this subsection

we classify the finite *abelian* algebras that have pointed cube terms by showing that they are exactly the finite affine algebras.

In fact, we do somewhat more than this, although somewhat less than fully determine the growth rates of finite abelian algebras. It is plausible that the following five properties are equivalent for finite abelian algebras:

(i) $\mathbf{A}$ has a Maltsev term.
(ii) $\mathbf{A}$ has a pointed cube term.
(iii) $d_{\mathbf{A}}(n) \in \Theta(n)$.
(iv) $d_{\mathbf{A}}(n) \notin 2^{\Omega(n)}$.
(v) No finite power $\mathbf{A}^n$ has a nontrivial strongly abelian homomorphic image.

It is clear that for abelian algebras item (i) is the strongest of these properties and item (v) is the weakest. We have (i)⇒(ii) since Maltsev terms are cube terms, (i)⇒(iii) by Theorem 2.2.6, (ii)⇒(iv) by Theorem 4.3.5, (iii)⇒(iv) (trivial), and (iv)⇒(v) by Corollary 2.2.5 (1). In this subsection we shall prove in addition that (i) and (ii) are equivalent for abelian algebras (Corollary 4.5.3), and that (iii), (iv) and (v) are equivalent for algebras that generate abelian varieties (Theorem 4.5.4).

The equivalence of (i) and (ii) for abelian algebras is a corollary to the following theorem about nilpotent algebras.

Theorem 4.5.1. *A finite left nilpotent algebra has a Maltsev polynomial iff it has a pointed cube polynomial.*

Proof. [⇒] A Maltsev polynomial is a 3-ary, 0-pointed, 2-cube polynomial.

[⇐] This part of the proof uses tame congruence theory, [15].

Let $\mathbf{A}$ be a finite left nilpotent algebra. Replace $\mathbf{A}$ by its polynomial expansion. In this setting, our goal is to show that if $\mathbf{A}$ has a pointed cube term, then it has a Maltsev term. In general, if $\mathbf{A}$ is a finite solvable algebra and $\mathbf{B} \leq \mathbf{A}^k$ is a subalgebra of a finite power of $\mathbf{A}$, then $\mathbf{B}$ has typeset contained in $\{\mathbf{1}, \mathbf{2}\}$. A finite solvable algebra $\mathbf{A}$ has a Maltsev term iff the typeset of any such $\mathbf{B}$ is contained in $\{\mathbf{2}\}$. Since left nilpotent algebras are solvable, we can prove this theorem by showing that any subalgebra of a finite power of $\mathbf{A}$ omits type $\mathbf{1}$.

Let $\mathbf{B}$ be a subalgebra of a finite power of $\mathbf{A}$ and assume that $\alpha \prec \beta$ is a covering of type $\mathbf{1}$ in $\mathbf{Con}(\mathbf{B})$. Let U be an $\langle \alpha, \beta \rangle$-minimal set, let e be an idempotent unary polynomial with range U, let T be a trace in U, and let 1 be an element of T. Let $F(x_1, \ldots, x_m)$ be a pointed cube term for the variety generated by $\mathbf{A}$. Let M be a $k \times m$ matrix of variables and constants, where each column contains an entry different from x, such that $F(M) \approx [x, \ldots, x]^{\mathsf{T}}$ in the variety generated by $\mathbf{A}$. As shown in the proof of Theorem 4.3.1, we may (and do) assume that the only variable appearing in M is x. Let N be the matrix obtained from M be replacing each constant with 1. Thus, N is a matrix of x's and 1's, and each column of N has at least one 1.

Claim 4.5.2. *In the variety generated by* $\mathbf{A}$, $F(N) \approx [\pi_1(x), \dots, \pi_k(x)]^\mathsf{T}$ *holds, where each* π_i *is a unary term of finite order. (I.e., for some positive integer* r *the identity* $\pi_i^r(x) \approx x$ *holds.) Moreover, each polynomial* $e\pi_i^{\mathbf{B}}(x)$ *is a permutation of* U*.*

The i-th cube identity is of the form $F(z_1, \dots, z_m) \approx x$ where each z_i is either a constant symbol or is x. Rewrite it as $t_i(x, \mathbf{c}_i) \approx x$, where $\mathbf{c}_i$ is the sequence of constant symbols that appear and $t_i(x, \mathbf{y})$ is the appropriate term. Let $\hat{1}_i = (1, 1, \dots, 1)$ have the same length as $\mathbf{c}_i$. The polynomials $t_i^{\mathbf{B}}(x, \mathbf{c}_i)$ and $t_i^{\mathbf{B}}(x, \hat{1}_i)$ are *twins*, meaning that they are derived from the same term with different constants as parameters. Theorem 3.9 and Corollary 3.8 of [19] establish that idempotent twins polynomials of a finite left nilpotent algebra have the ranges of the same cardinality. $\mathbf{B}$ is finite and left nilpotent since it is a subalgebra of a finite power of $\mathbf{A}$ and $\mathbf{A}$ has these properties. The polynomial $t_i^{\mathbf{B}}(x, \mathbf{c}_i) = F^{\mathbf{B}}(z_1, \dots, z_m) = x$ is idempotent with range of size $|B|$, so the idempotent iterate of its twin $t_i^{\mathbf{B}}(x, \hat{1}_i)$ also has range of size $|B|$. This forces $\pi_i(x) := t_i^{\mathbf{B}}(x, \hat{1}_i)$ to be a permutation of B, necessarily of finite order. Since $\mathbf{A}$ and $\mathbf{B}$ generate the same variety, there is an identity of the form $\pi_i^r(x) \approx x$ that holds in this variety.

If one repeats this argument with the polynomials $et_i^{\mathbf{B}}(x, \mathbf{c}_i) = e(x)$ and $et_i^{\mathbf{B}}(x, \mathbf{1}_i) = e\pi_i(x)$, then one obtains that $e\pi_i(x)$ is a permutation of U.

Now we consider the behavior of the polynomial $eF^{\mathbf{B}}(x_1, \dots, x_m)$ when its arguments are restricted to the trace T. Since the type of this trace is $\mathbf{1}$, the function

$$eF^{\mathbf{B}} \colon T^m \to U$$

is essentially unary modulo α, which means that the induced function

$$eF^{\mathbf{B}/\alpha} \colon (T/\alpha)^m \to U$$

depends on at most one variable. Each column of N contains an entry equal to 1, so some row of N contains a 1 in its first position. For this row, the fact that $eF^{\mathbf{B}}(N) = [e\pi_1(x), \dots, e\pi_m(x)]^\mathsf{T}$ where each $e\pi_i^{\mathbf{B}}(x)$ is a permutation of U yields that $eF^{\mathbf{B}}(x_1, \dots, x_m)$, restricted to T, depends modulo α on some variable *other than* x_1. The same type of conclusion holds for each variable of $eF^{\mathbf{B}}(x_1, \dots, x_m)$, which is impossible if this polynomial is essentially unary modulo α on T. □

Corollary 4.5.3. *A finite abelian algebra has a pointed cube polynomial iff it is affine.*

Not every finite abelian algebra generates an abelian variety, but for those that do we can prove that the growth rate is linear or exponential.

Theorem 4.5.4. *If* $\mathcal{V}$ *is an abelian variety and* $\mathbf{A} \in \mathcal{V}$ *is finite, then the following are equivalent.*

(iii) $d_{\mathbf{A}}(n) \in \Theta(n)$.
(iv) $d_{\mathbf{A}}(n) \notin 2^{\Omega(n)}$.

(v) *No finite power* $\mathbf{A}^n$ *has a nontrivial strongly abelian homomorphic image.*

Proof. We explained why (iii)⇒(iv)⇒(v) in the paragraphs before Theorem 4.5.1. Here we prove that (v)⇒(iii).

No assumption or conclusion is altered if we replace $\mathbf{A}$ with its polynomial expansion and the replace $\mathcal{V}$ by the variety this algebra generates, so henceforth we assume that every element of $\mathbf{A}$ is the interpretation of a constant term and that $\mathcal{V}$ is locally finite.

Claim 4.5.5. *Let* $\mathcal{U}$ *be a set containing exactly one member from each polynomial isomorphism class of type* **2** *minimal sets of* $\mathbf{A}$*. If* $\alpha \prec \beta$ *is a covering of type* **2** *in* $\mathbf{Con}(\mathbf{A}^n)$*, then there is an* $\langle \alpha, \beta \rangle$*-minimal set of the form* U^n *for some* $U \in \mathcal{U}$*.*

Since $\mathbf{Con}(\mathbf{A}^n)$ is modular modulo the strong solvability congruence, the covering $\alpha \prec \beta$ is projective via a sequence of coverings to a covering $\gamma \prec \delta$ which lies above a coordinate projection kernel, η. Thus the $\langle \alpha, \beta \rangle$-minimal sets are the same as the $\langle \gamma, \delta \rangle$-minimal sets. Identifying $\mathbf{A}$ with $\mathbf{A}^n/\eta$, choose some $U \in \mathcal{U}$ that is a $\langle \gamma/\eta, \delta/\eta \rangle$-minimal set. Then (a) U^n is the image of an idempotent unary polynomial of $\mathbf{A}^n$, (b) $\mathbf{A}^n|_{U^n}$ is an E-minimal algebra of type **2** (since $\mathbf{A}^n|_{U^n}$ is polynomially equivalent to $(\mathbf{A}|_U)^n$ and powers of solvable E-minimal algebras are E-minimal, according to Lemma 4.10 of [18]), and (c) $\gamma|_{U^n} \neq \delta|_{U^n}$. Items (a)–(c) are enough to show that U^n is a minimal set for $\langle \gamma, \delta \rangle$ and hence for $\langle \alpha, \beta \rangle$.

If $U \in \mathcal{U}$, then $\mathbf{A}|_U$ is affine, so there is a constant c_U such that $(\mathbf{A}|_U)^n$ can be generated by a set G_U of size at most $c_U \cdot n$. Let $G = \bigcup_{U \in \mathcal{U}} G_U$, which is a set of size at most $(\sum_{U \in \mathcal{U}} c_U)n$, and let $\mathbf{B} = \langle G \rangle$ be the subalgebra generated by this set. Since $\mathcal{V}$ is a locally finite abelian variety it is Hamiltonian (cf. [24]), which means that subalgebras are congruence classes. Let θ be a congruence on $\mathbf{A}^n$ that has $\mathbf{B}$ as a class.

Claim 4.5.6. $\mathbf{A}^n/\theta$ *is strongly abelian.*

Since $\mathcal{V}$ is abelian, to show that the algebra $\mathbf{A}^n/\theta \in \mathcal{V}$ is strongly abelian it suffices to prove that it has no type **2** covering. If it did, then there would be congruences $\alpha \prec \beta$ above θ in $\mathbf{Con}(\mathbf{A}^n)$ such that $\langle \alpha, \beta \rangle$ has type **2**. By Claim 4.5.5, there is some $\langle \alpha, \beta \rangle$-minimal set of the form U^n where $U \in \mathcal{U}$. By the definition of a minimal set, $\alpha|_{U^n}$ is properly contained in $\beta|_{U^n}$, so $\alpha|_{U^n}$ is not the universal equivalence relation on U^n. Since $\theta \leq \alpha$, $\theta|_{U^n}$ is also not the universal equivalence relation on U^n. But $U^n \subseteq B$, and B is a θ-class, so $\theta|_{U^n}$ must be the universal equivalence relation on U^n. Our assumption that $\mathbf{A}^n/\theta$ has a type **2** quotient has been contradicted, so this algebra is indeed strongly abelian.

To complete the proof, assume that (v) holds, i.e., no finite power $\mathbf{A}^n$ has a nontrivial strongly abelian homomorphic image. Claim 4.5.6 then implies that θ is the universal congruence on $\mathbf{A}^n$. Since θ has B as a class, we get $B = A^n$. But $\mathbf{B}$

was generated by at most $(\sum_{U\in\mathcal{U}} c_U)n \in O(n)$, proving that $d_{\mathbf{A}} \in O(n)$. By Corollary 2.2.5 (2) we get that $d_{\mathbf{A}} \in \Theta(n)$. □

4.6. **Exponential growth.** If $\mathbf{A}$ has exponential growth and $\mathbf{B}$ has arbitrary growth, then $\mathbf{A}\times\mathbf{B}$ has exponential growth. 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 2-element implication algebra, $\langle\{0,1\};\to\rangle$, contains a chain of finite algebras $\mathbf{A}_1 \le \mathbf{A}_2 \le \cdots$, 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 very 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 $A^n\setminus\{\mathbf{t}\}$ 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 $A\setminus\{t\}$ behaves like an 'ideal'. For this to work it is enough that $A\setminus\{t\}$ behaves like a 1-sided semigroup-theoretic ideal, so we introduce a definition that captures this notion for an arbitrary algebraic signature.

Definition 4.6.1. Let $\sigma=(F,\alpha)$ be an algebraic signature. I.e., let F be a set (of operation symbols) and let $\alpha\colon F\to\omega$ be a function (assigning arity). Let $F_0\subseteq F$ be the set consisting of those $f\in F$ such that $\alpha(f)>0$. (F_0 is the set of nonnullary symbols.) A *selector* for σ is a function $\phi\colon F_0\to\omega$ such that $1\le\phi(f)\le\alpha(f)$ for each $f\in F_0$. (ϕ selects one of the places of the function symbol f.)

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 $\phi(f)=i$ one has

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

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 $f^{\mathbf{A}}(a_1,\dots,a_n)\in I$.

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 4.6.2. *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 $d_{\mathbf{A}}(n)\ge 2^n$.*

Proof. The union of ϕ-ideals is again a ϕ-ideal, so if $\mathbf{A}$ is the union of $k\ge 2$ proper ϕ-ideals then it can be expressed as the union $I\cup J$ of 2 proper ϕ-ideals. The complements $I':=A\setminus I$ and $J':=A\setminus J$ 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.

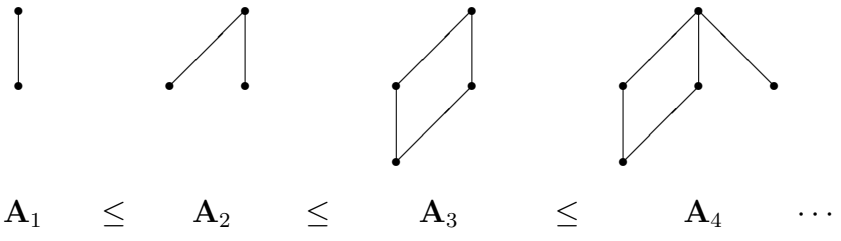
Each such set must contain at least one element of any generating set, since the ϕ-irreducibility of T implies that $A^n \setminus T$ is a subuniverse of $\mathbf{A}^n$. Since there are 2^n products of the form $X_1 \times \cdots \times X_n$ with $X_i = I'$ or J', and they are pairwise disjoint, any generating set for $\mathbf{A}^n$ must contain at least 2^n elements. □

Example 4.6.3. In this example, $\mathbf{2}$ is the 2-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 right ideal in $\mathbf{A}$ with respect to the operation $\to$. By Theorem 4.6.2, 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, i.e., $\mathbf{A} = (2^\circ)^k$. In this situation $\mathbf{A}$ is polynomially equivalent to the Boolean algebra $\mathbf{2}^k$, so it shares logarithmic growth rate with $\mathbf{2}^k$.

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 XXX shows how the chain might begin.



5. Wiegold dichotomy holds for algebras with a cube term

In this section we investigate growth rates of finite algebras with a (0-pointed) cube term. Unlike the previous section, where we focused primarily on upper bounds on growth rates, we will prove that growth rates of finite algebras with a cube term must be one of two kinds: if $\mathbf{A}$ is a finite algebra with a cube term, then $d_{\mathbf{A}}(n) \in \Theta(\log(n))$ if $\mathbf{A}$ is perfect and $d_{\mathbf{A}}(n) \in \Theta(n)$ if $\mathbf{A}$ is imperfect.

Our approach will be through the analysis of the maximal subuniverses of $\mathbf{A}^n$. The relevance of maximal subuniverses may be recognized from the fact that a subset $G \subseteq A^n$ is a generating set for $\mathbf{A}^n$ if and only if $G \not\subseteq M$ for any maximal subuniverse M of $\mathbf{A}^n$. Our analysis of a maximal subuniverse M of $\mathbf{A}^n$ involves finding surjective homomorphisms $\varphi\colon \mathbf{A}^n \to \mathbf{B}$ such that $\varphi(M) \neq B$. When φ is such a homomorphism, $M_\varphi := \varphi(M)$ will be a maximal subuniverse of $\mathbf{B}$ and M will equal $\varphi^{-1}(M_\varphi)$ (since $\varphi^{-1}(M_\varphi)$ is a proper subuniverse of $\mathbf{A}^n$ that contains M). We say that M is *induced* by φ in this situation.

We will consider two types of homomorphisms. The first type of homomorphism considered is the projection $\pi_U\colon \mathbf{A}^n \to \mathbf{A}^U$ onto the coordinates in U for some subset $U \subseteq [n]$. The second type will be the homomorphism $\eta\colon \mathbf{A}^n \to (\mathbf{A}/[1,1])^n$ induced by the abelianization map in each coordinate.

Our first main result, Theorem 5.2.4, states that if $\mathbf{A}$ has a 0-pointed k-cube term, then every maximal subuniverse is induced by either (i) a projection π_U onto a set of coordinates of size $|U| < \max\{3, k\}$, or (ii) the abelianization homomorphism η. The second main result, Theorem 5.4.1, is that there exists a set $G_\pi \subseteq A^n$ of size $O(\log(n))$ that is contained in no maximal subuniverse which can be induced by a type (i) map, and there exists a set $G_\eta \subseteq A^n$ of size $O(n)$ that is contained in no maximal subuniverse which can be induced by a type (ii) map. It follows that $G_\pi \cup G_\eta$ is contained in no maximal subuniverse at all, so it is a generating set of size $O(n)$ for $\mathbf{A}^n$. When $\mathbf{A}$ is imperfect, the $O(n)$-estimate is asymptotically optimal, since it is proved in Corollary 2.2.5 (2) that $d_{\mathbf{A}}(n) \in \Omega(n)$ when $\mathbf{A}$ is imperfect. When $\mathbf{A}$ is perfect, η is constant, hence no maximal subuniverse of $\mathbf{A}^n$ can be induced by η. It follows that G_π is a generating set of size $\Theta(\log(n))$ for $\mathbf{A}^n$, which is asymptotically optimal by Theorem 2.2.2.

5.1. **Maximal subuniverses of powers.** In this subsection we relate arbitrary maximal subuniverses of $\mathbf{A}^n$ to critical maximal subuniverses.

Definitions 5.1.1. [21] A *compatible n-ary relation* of $\mathbf{A}$ is a subuniverse of $\mathbf{A}^n$.

A compatible relation R is *critical* if it is completely $\cap$-irreducible in the subalgebra lattice of $\mathbf{A}^n$ and directly indecomposable as a relation. (The latter means that R is not of the form $S \times T$ for subsets $S \subseteq A^U$ and $T \subseteq A^V$, where $\{U, V\}$ is a partition of $[n]$ into two cells.)

Any maximal subuniverse M of $\mathbf{A}^n$ is completely $\cap$-irreducible in the subalgebra lattice of $\mathbf{A}^n$, so a critical maximal subuniverse of $\mathbf{A}^n$ is just a maximal subuniverse that is directly indecomposable as a relation.

Definition 5.1.2. If M is a subuniverse of $\mathbf{A}^n$, then a *support* of M is a subset $U \subseteq [n]$ such that $\pi_U(M) \neq A^U$.

Lemma 5.1.3. *If M is a maximal subuniverse of $\mathbf{A}^n$, then M has a unique minimal support. If U is the minimal support of M, then $M_U := \pi_U(M)$ is a critical maximal*

subuniverse of $\mathbf{A}^U$, $M = M_U \times A^V$ *for* $V = [n]\setminus U$, *and* M *is induced by the projection* π_U. *In particular,* M *itself is critical if and only if its unique support is* $[n]$.

Proof. Assume that $U, V \subseteq [n]$ are distinct minimal supports of the maximal subuniverse $M \leq \mathbf{A}^n$. U and V must be incomparable under inclusion. Let $M_U = \pi_U(M)$ and $M_V = \pi_V(M)$.

We shall view elements of $\mathbf{A}^n$ as functions from $[n]$ to A. In this language, M is a proper subset of the set of all functions, M_U is the set of restrictions to U of the functions in M, and M_V is the set of restrictions to V of the functions in M. Since $M_U \neq A^U$, there is a function $f\colon U \to A$ that is not in M_U. Since V is a minimal support and $U \cap V$ is properly contained in V, it follows that every function from $U \cap V$ to A is the restriction of some function in M. In particular, $f|_{U\cap V} = g|_{U\cap V}$ for some $g \in M$. Let $h \in A^n$ be any function that agrees with f on U and g on V. Then $h|_U = f \notin M_U$, so $h \notin M$. Yet $h|_V = g|_V \in M_V$, so $h \in M$, a contradiction. This shows that M has a unique minimal support.

If U is the minimal support of M and $V = [n] \setminus U$, then $M = M_U \times A^V$ is induced by π_U. (This does not require minimality of U, only that U is a support and that M is a maximal subuniverse.) To show that M_U is a critical maximal subuniverse of $\mathbf{A}^U$, observe that $M_U \leq \mathbf{A}^U$ is a maximal subuniverse since it is the image of one under the surjective homomorphism π_U. If $M_U = S \times T$, where $S \leq \mathbf{A}^X$ and $T \leq \mathbf{A}^Y$ for some partition $\{X, Y\}$ of U, then either $A^X \neq \pi_X^{\mathbf{A}^U}(M_U) = \pi_X^{\mathbf{A}^n}(M)$, or $A^Y \neq \pi_Y^{\mathbf{A}^U}(M_U) = \pi_Y^{\mathbf{A}^n}(M)$. Either way, one obtains that X or Y is a proper subset of U that is a support of M, contradicting the minimality of U.

For the final statement of the lemma, if the minimal support of M is $[n]$, then $\pi_{[n]}(M) = M$ is critical by the second statement of the lemma. Conversely, assume that M is critical and $U \subseteq [n]$ is its minimal support. Since $M = M_U \times A^V$ and M is directly indecomposable as a relation, we get $V = \emptyset$, equivalently $[n] = U$. □

5.2. **The parallelogram property for critical relations.** In the preceding subsection we showed that all maximal subuniverses of $\mathbf{A}^n$ are induced by critical maximal subuniverses on projections $\mathbf{A}^U$ of $\mathbf{A}^n$. In this section we show that the critical maximal subuniverses of $\mathbf{A}^U$ have a special structure when $\mathbf{A}$ has a 0-pointed k-cube term.

Definition 5.2.1. [21] Given a partition $\{S, T\}$ of $[n]$ into two cells, write $\mathbf{xy}$ for a tuple in A^n to mean that $\mathbf{x} \in A^S$ and $\mathbf{y} \in A^T$. A compatible n-ary relation R satisfies the *parallelogram property* if, for any partition $\{S, T\}$ of $[n]$, $\mathbf{au}, \mathbf{av}, \mathbf{bv} \in R$ implies $\mathbf{bu} \in R$.

Theorem 3.5 and Theorem 3.6 (3) of [21] together prove the following theorem.

Theorem 5.2.2. *A variety* $\mathcal{V}$ *has a* 0*-pointed* k*-cube term if and only if every member* $\mathbf{A} \in \mathcal{V}$ *has the property that any critical relation of* $\mathbf{A}$ *of arity at least* k *has the parallelogram property.*

It follows from this theorem that if $\mathbf{A}$ has a 0-pointed k-cube term, and M is a maximal subuniverse of $\mathbf{A}^n$, then either M is induced by a relation of arity less than k or $M = M_U \times A^V$ is induced by a critical maximal subuniverse $M_U \leq \mathbf{A}^U$ that has the parallelogram property. (In the latter case, M itself will also have the parallelogram property.) Our next step is to investigate the structure of maximal subuniverses with the parallelogram property.

The paper [21] analyzes arbitrary compatible relations with the parallelogram property in congruence modular varieties. It is shown in [2] that any algebra with a 0-pointed k-cube term generates a congruence modular variety, so the results of [21] apply here. The first step in the analysis is the "reduction" of a relation, which we describe next.

Suppose that $R \leq \mathbf{A}^n$ is a compatible relation with the parallelogram property; as a special case, suppose that $M \leq \mathbf{A}^n$ is a maximal critical subuniverse with the parallelogram property. For the first step in the reduction, realize R as a subdirect product $R \leq_{\mathrm{sd}} \prod_{i=1}^n \mathbf{A}_i$, where $\mathbf{A}_i := \pi_i(R) \leq \mathbf{A}$. In the special case involving the maximal subuniverse M we will have $\mathbf{A}_i = \pi_i(M) = \mathbf{A}$ unless the projection of M onto one single coordinate is not surjective. This happens only if M has a support of size one, which, by criticality, implies that M is a unary relation. We henceforth consider only M of arity at least two, so that in our special case $\pi_i(M) = A$ for all i. Thus, in the first step in reduction, nothing happens if M is maximal and of arity greater than one.

Second, define relations, called *coordinate kernels* in [21],

$$\theta_i = \{(a, b) \in A_i^2 \mid \exists \mathbf{c} \in \prod_{j \neq i} \mathbf{A}_j \ (a\mathbf{c} \in R \ \& \ b\mathbf{c} \in R)\}.$$

It is proved in Lemma 2.3 of [21] that (i) each θ_i is a congruence on $\mathbf{A}_i$, and (ii) R is induced by the homomorphism $\psi\colon \prod \mathbf{A}_i \to \prod \mathbf{A}_i/\theta_i$ that is the natural map in each coordinate. (This means: if $\overline{R} = \psi(R)$, then $R = \psi^{-1}(\overline{R})$.) The relation $\overline{R}$ is the *reduction* of R.

In our special case $M \leq \mathbf{A}^n$, where M is critical and maximal, we observe that $\overline{M} = \psi(M)$ is a maximal subuniverse of $\prod \mathbf{A}/\theta_i$. For, if $M < M' < \prod \mathbf{A}/\theta_i$, then $M = \psi^{-1}(M) < \psi^{-1}(M') < \mathbf{A}^n$, contradicting the maximality of M.

The next result is a specialization of (some parts of) Theorem 2.5 of [21] to the case where M is a critical maximal subuniverse of $\mathbf{A}^n$ and n is greater than one. We maintain the numbering of [21], but omit the unused parts of the theorem.

Theorem 5.2.3. *Let M be a critical maximal subuniverse of $\mathbf{A}^n$ that satisfies the parallelogram property, and let $\overline{M} \leq \prod \mathbf{A}/\theta_i$ be its reduction. If $n > 1$ and $\mathbf{A}$ lies in a congruence modular variety, then the following hold.*

(1) *$\overline{M} \leq \prod \mathbf{A}/\theta_i$ is a representation of $\overline{M}$ as a subdirect product of subdirectly irreducible algebras.*

(5)* *If $n > 2$, then the monolith of $\mathbf{A}/\theta_i$ is the total relation; i.e. $\mathbf{A}/\theta_i$ is simple.*
(7)* *If $n > 2$, then each simple algebra $\mathbf{A}/\theta_i$ is abelian.*

Here, items (5) and (7) are marked with asterisks, because we have altered the statement of (5) from [21] in order to take into account that $\overline{M}$ is a maximal subuniverse of $\prod \mathbf{A}/\theta_i$ and we have altered the statement of (7) in order to take into account the conclusion from (5)* that $\mathbf{A}/\theta_i$ is simple.

We explain what this theorem contributes to our current investigation. Suppose that $\mathbf{A}$ has a 0-pointed k-cube term. Suppose also that $M \leq \mathbf{A}^n$ is maximal, U is the minimal support of M, and $M = M_U \times A^V$ is induced by $\pi_U \colon \mathbf{A}^n \to \mathbf{A}^U$. If $|U|$ is at least as large as $\max\{3, k\}$, then the theorem proves that M_U is induced by a homomorphism $\psi \colon \mathbf{A}^U \to \prod_U \mathbf{A}/\theta_i$ where each factor $\mathbf{A}/\theta_i$ is a simple abelian algebra. Thus, M itself is induced by the composition of the surjective homomorphisms

$$\mathbf{A}^n \xrightarrow{\pi_U} \mathbf{A}^U \longrightarrow (\mathbf{A}/[1,1])^U \longrightarrow \prod_U \mathbf{A}/\theta_i,$$

where the last two maps are a factorization of the map $\psi \colon \mathbf{A}^U \to \prod_U \mathbf{A}/\theta_i$ which induces M_U, and these two maps are defined coordinatewise by the natural maps $\mathbf{A} \to \mathbf{A}/[1,1] \to \mathbf{A}/\theta_i$. (We have $\theta_i \geq [1,1]$, since $\mathbf{A}/\theta_i$ is abelian.) Hence M is induced by the sub-composition $\mathbf{A}^n \xrightarrow{\pi_U} \mathbf{A}^U \longrightarrow (\mathbf{A}/[1,1])^U$, which may be factored another way as $\mathbf{A}^n \xrightarrow{\eta} (\mathbf{A}/[1,1])^n \xrightarrow{\pi_U} (\mathbf{A}/[1,1])^U$. Hence M is induced by the single map η, which maps $\mathbf{A}^n$ onto its abelianization. Altogether this proves the desired result:

Theorem 5.2.4. *Assume that $\mathbf{A}$ has a 0-pointed k-cube term. If $M \leq \mathbf{A}^n$ is a maximal subuniverse, then either*

(π) *M is induced by a projection $\pi_U \colon \mathbf{A}^n \to \mathbf{A}^U$ for some subset $U \subseteq [n]$ satisfying $|U| < \max\{3, k\}$, or*
(η) *M is induced by $\eta \colon \mathbf{A}^n \to (\mathbf{A}/[1,1])^n$.*

5.3. **A solution to a combinatorial problem.** The problem considered here is: If A is a finite set and $n \geq k > 1$ are integers, then how small can a set $G \subseteq A^n$ be if its projection onto any subset of k coordinates is surjective?

If A is finite, $G \subseteq A^n$ and $|G| = g$, then G can be linearly ordered and taken to be the sequence of rows of a $g \times n$ matrix of elements of A, say $[a_{i,j}]$. If

$$\sigma \colon \quad 1 \leq j(1) < \cdots < j(k) \leq n$$

is a selection of k numbers between 1 and n, then the projection of G onto the coordinates in σ is the set of row vectors $(a_{1,j(1)}, \ldots, a_{1,j(k)}), \ldots, (a_{g,j(1)}, \ldots, a_{g,j(k)})$ which occur as the set of rows of the $g \times k$ minor of $[a_{i,j}]$ whose column indices are the indices in σ. G projects surjectively onto each k coordinates of A^n if and only if, for each choice σ of k column indices, the set of row vectors of the corresponding

$g \times k$ minor of $[a_{i,j}]$ exhausts A^k. Therefore, call a $g \times k$ matrix of elements of A a *bad minor* (or *bad matrix*) if its rows fail to exhaust A^k. The desired property of G is that its associated matrix has no bad minors.

Theorem 5.3.1. *Let A be a finite set of size $|A| = a > 1$. Let $n \geq k > 1$ be natural numbers, and set $u = a^k/(a^k - 1)$. If $g \geq k \log_u(n) + \log_u\left(a^k/k!\right)$, then there is a matrix in $A^{g\times n}$ with no bad minors.*

Proof. This is a probabilistic proof. Our sample space is the set $A^{g\times n}$ of all $g \times n$ matrices of elements of A. Our probability distribution is the uniform one, so each individual matrix $M \in A^{g\times n}$ has probability $P(M) = |A^{g\times n}|^{-1} = a^{-gn}$. For each matrix $M \in A^{g\times n}$ and each sequence of k column indices,

$$\sigma\colon \quad 1 \leq j(1) < \cdots < j(k) \leq n,$$

let M_σ denote the $g \times k$ minor of M whose column indices are those enumerated by σ (called the *σ-minor* of M). Let X_σ be the random variable whose value at the element $M \in A^{g\times n}$ is 1 if M_σ is a bad minor and 0 otherwise, i.e., X_σ is the indicator variable for bad σ-minors.

Claim 5.3.2. *For any σ, the expected value of X_σ satisfies*

$$E(X_\sigma) \leq a^k \left(a^k - 1\right)^g a^{-gk}. \tag{5.1}$$

The expectation is computed

$$\begin{aligned} E(X_\sigma) &= \textstyle\sum_{M\in A^{g\times n}} (X_\sigma(M) \cdot P(M)) \\ &= \textstyle\sum_{M\in A^{g\times n}} (X_\sigma(M) \cdot a^{-gn}) \\ &= \left(\textstyle\sum_{M\in A^{g\times n}} X_\sigma(M)\right) a^{-gn}, \end{aligned}$$

where the sum $\sum_{M\in A^{g\times n}} X_\sigma(M)$ on the last line represents the number matrices in $A^{g\times n}$ whose σ-minor is bad. By definition, a $g \times k$ matrix is bad if some tuple $\mathbf{a} \in A^k$ does not appear among its rows. So, for each $\mathbf{a} \in A^k$, let $\mathcal{U}_\mathbf{a}$ denote the set of all $g \times k$ matrices where $\mathbf{a}$ does not appear among the rows. $|\mathcal{U}_\mathbf{a}|$ can be computed by noting that the g rows of a matrix in $\mathcal{U}_\mathbf{a}$ may be freely chosen from the set $A^k - \{\mathbf{a}\}$, which has size $a^k - 1$, so $|\mathcal{U}_\mathbf{a}| = (a^k - 1)^g$. The bad $g \times k$ matrices are those from $\bigcup_{\mathbf{a}\in A^k} \mathcal{U}_\mathbf{a}$. Since the cardinality of the union is no more than the sum of the individual cardinalities, and these summands have the same size, we get that the number of bad $g \times k$ matrices is no more than $|A^k| \cdot |\mathcal{U}_\mathbf{a}| = a^k(a^k - 1)^g$. Each bad $g \times k$ matrix N can be extended in $a^{g(n-k)}$ ways to a matrix $M \in A^{g\times n}$ whose σ-minor satisfies $M_\sigma = N$, so the number of matrices in $A^{g\times n}$ with a bad σ-minor is no more than $a^k(a^k - 1)^g a^{g(n-k)}$. Hence

$$E(X_\sigma) = \left(\sum_{M\in A^{g\times n}} X_\sigma(M)\right) a^{-gn} \leq a^k(a^k - 1)^g a^{g(n-k)} a^{-gn} = a^k(a^k - 1)^g a^{-gk},$$

as claimed.

If $X := \sum_\sigma X_\sigma$ is the sum of all X_σ as σ ranges over all $\binom{n}{k}$ choices of k column indices and $M \in A^{g\times n}$, then $X(M)$ equals the number of bad $g \times k$ minors of M. Since expectation is linear, and since $\binom{n}{k} < n^k/k!$ when $n \geq k > 1$, we get from (5.1) that

$$E(X) = \sum_\sigma E(X_\sigma) \leq \binom{n}{k} a^k (a^k - 1)^g a^{-gk} < n^k (a^k/k!)(a^k - 1)^g a^{-gk}.$$

If it is the case that

$$n^k (a^k/k!)(a^k - 1)^g a^{-gk} \leq 1, \tag{5.2}$$

then we will have $E(X) < 1$, meaning that the expected number of bad minors in an element of $A^{g\times n}$ is strictly less than 1. This can happen only if matrices without bad minors exist. Rewriting (5.2) as

$$n^k \leq \left(\frac{a^k}{(a^k - 1)}\right)^g (a^k/k!)^{-1} = u^g (a^k/k!)^{-1},$$

using the definition $u = a^k/(a^k - 1)$, we can solve for g to get

$$g \geq k \log_u(n) + \log_u(a^k/k!). \tag{5.3}$$

When this inequality holds we get that (5.2) holds, so a matrix with no bad minors exists. This is exactly the statement of the theorem. □

Corollary 5.3.3. *Let A be a finite set of size $|A| = a > 1$. Let $n \geq k > 1$ be natural numbers, and set $u = a^k/(a^k - 1)$.*

1. *If $g = \lceil k \log_u(n) + \log_u\left(a^k/k!\right)\rceil$, then there exists a subset $G \subseteq A^n$ of size g whose projection onto any k coordinates of A^n is surjective.*
2. *Let $G \subseteq A^n$ be a subset.*
   - (i) *If the projections $\mathrm{pr}_i \colon G \to A$ and $\mathrm{pr}_j \colon G \to A$ onto distinct coordinates are different functions for all i and j, then $|G| \geq \log_a(n)$.*
   - (ii) *If no two projections $\mathrm{pr}_i \colon G \to A$ and $\mathrm{pr}_j \colon G \to A$ differ by a permutation of A (i.e., if it is not the case that $\pi \circ \mathrm{pr}_i = \mathrm{pr}_j$ for any i and j or $\pi \in Sym(A)$), then $|G| \geq \log_a(n) + \log_a(a!)$.*
   - (iii) *If all projections of G onto pairs of coordinates are surjective, then $|G| \geq \log_a(n) + \log_a(a!)$.*

When reading the statement of this corollary one should imagine that a and k (and hence u) are fixed while n ranges. Item (1) implies that there is a subset $G \subseteq A^n$ of size $O(\log(n))$ that projects surjectively onto any k coordinates. Item (2) implies that any set $G \subseteq A^n$ that projects surjectively onto each set of k coordinates must have size $\Omega(\log(n))$ if k is at least 2. Taken together these statements show that, as a function of n, the least size of a set $G \subseteq A^n$ that projects onto each k coordinates is $\Theta(\log(n))$.

Proof of Corollary 5.3.3. Part (1) is an immediate consequence of Theorem 5.3.1.

For (2)(i), suppose that $G \subseteq A^n$ is a subset whose projections onto distinct coordinates are different functions for different coordinates. Then the set of projections onto distinct coordinates constitute n distinct elements of the function space A^G, which has cardinality a^g. It follows that $n \leq a^g$, or $|G| = g \geq \log_a(n)$.

The argument for (2)(ii) is essentially the same as for (2)(i). The assumptions imply that the set of all functions of the form $\pi \circ \mathrm{pr}_i \colon G \to A$, $\pi \in \mathrm{Sym}(A)$, are distinct. This yields $|\mathrm{Sym}(A)| \cdot n = a! \cdot n$ functions from the set A^G, which has size a^g, hence $a! \cdot n \leq a^g$. Solving for g yields $|G| = g \geq \log_a(n) + \log_a(a!)$.

For (2)(iii), if two projections differ by a permutation, say $\pi \circ \mathrm{pr}_i = \mathrm{pr}_j$, then the projection of G onto the i-th and j coordinates maps G into the the set of pairs of the form $(x, \pi(x))$. All such pairs lie on the graph of π, which is a proper subset of A^2 when $|A| > 1$. Thus, if all projections of G onto pairs of coordinates are surjective, no two single coordinate projections can differ by a permutation. This shows that item (2)(iii) follows from (2)(ii). □

Remark 5.3.4. It is possible to refine the estimate in Corollary 5.3.3 (1) when we are dealing with algebras instead of just sets. Suppose that there is a subset $B \subseteq A$ such that B^k generates $\mathbf{A}^k$. Then we can apply the arguments above to obtain a set $G \subseteq B^n$ whose projection onto any k coordinates is B^k. According to the arguments, we can find such a G of size

$$|G| = \lceil k \log_v(n) + \log_v\left(b^k/k!\right) \rceil,$$

where $b = |B|$ and $v = b^k/(b^k - 1)$. This is a larger base for the logarithm and a smaller constant on the right, so an overall smaller estimate for G. The subalgebra of $\mathbf{A}^n$ generated by G will project surjectively onto each k coordinates, since its projection contains the generating set B^k. Note here that there is no longer any need to assume that $\mathbf{A}$ is finite, only that B is finite.

Another refinement can be made on top of the preceding one. Suppose that there is a subset $P \subseteq B^k$ of size $|P| = p$ that generates $\mathbf{A}^k$. (Here we do not assume that P is the k-th power of a subset of A.) The argument we have given does not generalize to produce a subset $G \subseteq A^n$ whose projection onto any k coordinates is exactly P, but it does generalize to produce a subset $G \subseteq A^n$ whose projection onto any k coordinates will be contained in B^k and will contain P. This enough to ensure that the subalgebra of $\mathbf{A}^n$ that is generated by G projects surjectively onto any k coordinates. How much of an improvement do we get in our estimate of $|G|$?

The assumption that P ($\subseteq B^k$) generates $\mathbf{A}^k$ affects the estimate in Claim 5.3.2 as follows. For each $\mathbf{b} \in P$, let $\mathcal{U}_{\mathbf{b}}$ denote the set of all matrices in $B^{g \times k}$ where $\mathbf{b}$ does not appear among the rows. Then $|\mathcal{U}_{\mathbf{b}}| = (b^k - 1)^g$, so $|\bigcup_{\mathbf{b} \in P} \mathcal{U}_{\mathbf{b}}| \leq |P|(b^k - 1)^g = p(b^k - 1)$. Each of these $g \times k$ matrices can be extended in $b^{g(n-k)}$ ways to a matrix in $B^{g \times n}$, so we obtain the estimate $E(X_\sigma) \leq p(b^k - 1)b^{-gk}$. This allows us to choose G so that it

is slightly smaller, namely

$$|G| = \lceil k \log_v(n) + \log_v(p/k!) \rceil.$$

Let's summarize the two refinements. Let $\mathbf{A}$ be a (possibly infinite) algebra. Suppose that $P \subseteq A^k$ is a finite generating set for $\mathbf{A}^k$. Let B be the smallest subset of A for which $P \subseteq B^k$, i.e., let $B \subseteq A$ be the set of elements that appear in the coordinates of the tuples in P. Let $b = |B|$, $p = |P|$, and $v = b^k/(b^k - 1)$. Assertion: There is a set $G \subseteq A^n$ of size $\lceil k \log_v(n) + \log_v(p/k!) \rceil$ that generates a subalgebra of $\mathbf{A}^n$ which projects surjectively onto any k coordinates.

5.4. **Growth rates for algebras with a cube term.** In this subsection we combine the preceding results to obtain the following.

Theorem 5.4.1. *Suppose that* $\mathbf{A}$ *has a (0-pointed)* k*-cube term. If* $\mathbf{A}$ *is imperfect, then* $d_{\mathbf{A}}(n) \in \Theta(n)$*. If* $\mathbf{A}$ *perfect, then* $d_{\mathbf{A}}(n) \in \Theta(\log(n))$*.*

Proof. According to Theorem 5.2.4, if $M \leq \mathbf{A}^n$ is a maximal subuniverse, then either

- (π) M is induced by a projection $\pi_U \colon \mathbf{A}^n \to \mathbf{A}^U$ for some subset $U \subseteq [n]$ satisfying $|U| < \max\{3, k\}$, or
- (η) M is induced by $\eta \colon \mathbf{A}^n \to (\mathbf{A}/[1,1])^n$.

For each n, choose a subset $G_\pi \subseteq A^n$ of size $O(\log(n))$ whose projection onto any subset of $\max\{3, k\}$ coordinates is surjective. The existence of such a set is guaranteed by Corollary 5.3.3. Clearly G_π is contained in no maximal subuniverse of $\mathbf{A}^n$ that is induced by a projection onto any subset of $\max\{3, k\}$ coordinates.

The algebra $\mathbf{A}/[1,1]$ is abelian and has a cube term, so according to Corollary 4.5.3 $\mathbf{A}$ is affine. According to Theorem 2.2.6, $(\mathbf{A}/[1,1])^n$ contains a set of generators of size $O(n)$. For each n, choose a set $G_\eta \subseteq A^n$ of size $O(n)$ such that $\eta(G_\eta)$ generates $(\mathbf{A}/[1,1])^n$. Then G_η is contained in no maximal subuniverse of $\mathbf{A}^n$ induced by η.

We now have that $G_\pi \cup G_\eta$ is a set of size $O(n)$ that is contained in no maximal subuniverse of $\mathbf{A}^n$, hence $G_\pi \cup G_\eta$ is a generating set for $\mathbf{A}^n$ of size $O(n)$.

When $\mathbf{A}$ is imperfect, then $d_{\mathbf{A}}(n) \in \Omega(n)$ by Corollary 2.2.5, so the generating sets we have found for the powers of $\mathbf{A}$ are asymptotically optimal in size. When $\mathbf{A}$ is perfect, then $\mathbf{A}^n$ has no maximal subuniverses induced by η, so G_π is a generating set for $\mathbf{A}^n$ of size $O(\log(n))$. By Theorem 2.2.2 $d_{\mathbf{A}}(n) \in \Omega(\log(n))$ for any nontrivial algebra, so the generating sets we have for the powers of $\mathbf{A}$ in this case are also asymptotically optimal in size. □

6. Extensions and problems

6.1. **Growth rates of infinite algebras.** Any set Σ of basic identities that does not entail the existence of a pointed cube term is also nonrestrictive for infinite algebras. This can be shown in essentially the same way we showed it for finite algebras, but the situation is clearer for infinite algebras because we can say explicitly which growth

rates are possible. In Example 6.1.1 we show that if $d\colon \mathbb{Z}^+ \to \mathbb{Z}^+$ is an arbitrary monotone function, then there is a countably infinite algebra whose growth rate is d. In Example 6.1.3 we show how to modify the example, without changing its growth rate, so that it realizes a given set Σ of basic identities, provided Σ does not entail the existence of a pointed cube term. The cardinality of this second example can be any infinite cardinal that is at least as large as the number of constant symbols appearing in Σ.

Example 6.1.1. Let $d\colon \mathbb{Z}^+ \to \mathbb{Z}^+$ be an arbitrary monotone function. We construct a countably infinite partial algebra $\mathbf{A}$ such that $d_{\mathbf{A}}(n) = d(n)$ for all n. The one-point completion of $\mathbf{A}$ (Definition 4.1.1) will be a total algebra with the same growth rate.

Let $M^{(1)}, M^{(2)}, \ldots$ be a sequence of matrices with the following properties.

(1) $M^{(n)} = [a^{(n)}_{i,j}]$ is an $n \times d(n)$ matrix.
(2) All elements of all matrices are different from one another.

Let $A = \{a^{(n)}_{i,j}\}$ be the set of all entries appearing in these matrices. Our aim is to equip A with partial operations ensuring that the $d(n)$ columns of $M^{(n)}$ form a smallest size generating set for $\mathbf{A}^n$. If we achieve this, then we will have $d_{\mathbf{A}}(n) = d(n)$ for all n.

For each $n \in \mathbb{Z}^+$ and each $\mathbf{b} \in A^n$, 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 $F_{\mathbf{b}}(a^{(n)}_{i,1}, \ldots, a^{(n)}_{i,d(n)}) = b_i$ for each $i = 1, \ldots, n$.

Our partial algebra is A equipped with all partial operations of the type described in the previous paragraph. Since $F_{\mathbf{b}}(M^{(n)}) = \mathbf{b}$ whenever $\mathbf{b} \in A^n$, 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 6.1.2. *If a subset $G \subseteq A^n$ has fewer than $d(n)$ tuples whose coordinates are distinct, then the same is true for $\langle G \rangle$.*

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.

$$F_{\mathbf{b}}(\mathbf{x}_1, \ldots, \mathbf{x}_{d(m)}) = F_{\mathbf{b}}\left(\begin{bmatrix} x_{1,1} \\ \vdots \\ x_{n,1} \end{bmatrix}, \ldots, \begin{bmatrix} x_{1,d(m)} \\ \vdots \\ x_{n,d(m)} \end{bmatrix}\right) = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{c}. \tag{6.1}$$

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 completes the argument, 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 $\mathbf{A}^n$ has infinitely many such tuples.

Example 6.1.3. Here we explain how to modify the algebra from Example 6.1.1, without changing its growth rate, so that it realizes a given set Σ of basic identities. The only requirement on Σ is that it does not entail the existence of a pointed cube term.

The construction is like the one in Subsection 4.2, so we only outline it. 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

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

Arguments similar to those in Theorem 4.1.2 show that $\mathbf{A}$ and $\mathbf{A}_{Z,0}$ have the same minimal generating sets, so the same growth rate. We can merge this example with a model $\mathbf{V}$ from Definition 3.2.1 to obtain a model $\mathbf{A}_\Sigma$, as we did in Subsection 4.2.

Using the same arguments as before, it can be shown that this model will have the same growth rate as $\mathbf{A}$ unless Σ entails the existence of a pointed cube term.

In particular, it is possible to find infinite algebras generating congruence distributive and congruence 3-permutable varieties whose growth rate is any prescribed monotone function.

We showed in Subsection 4.3 that a finite algebra with a p-pointed k-cube term has growth that is bounded above by a polynomial. When $p \geq 1$, the constant $|A|^{k-1}$ appears in the polynomial, so the argument requires $\mathbf{A}$ to be finite. However, some version of the argument holds when $\mathbf{A}$ is infinite.

Let P be the set of p (≥ 1) constant symbols that appear in the cube identities for some k-cube term. Say that a term operation $t^{\mathbf{A}}$ *centralizes* P if $t^{\mathbf{A}}(c^{\mathbf{A}}, \dots, c^{\mathbf{A}}) = c^{\mathbf{A}}$ for all $c \in P$. The argument we gave in Subsection 4.3 may be extended to prove:

Theorem 6.1.4. *Let $\mathbf{A}$ be an algebra with an m-ary, p-pointed, k-cube term, and let P be the set of constant symbols appearing in the cube identities. Let $\widehat{\mathbf{A}}$ be the reduct of $\mathbf{A}$ to the term operations that centralize P. If the reduct $\widehat{\mathbf{A}}^{k-1}$ is finitely generated, then the growth rate of $\mathbf{A}$ is bounded above by a polynomial of degree at most $\log_w(mp)$, for $w = 2k/(2k-1)$.*

Proof. The proof is the same as the proof of Theorem 4.3.1, except that we apply one more step after fully processing all tuples. Let $F \subseteq A$ be a finite subset with the property that F^{k-1} generates $\widehat{\mathbf{A}}^{k-1}$.

Recall that a fully processed tuple $\mathbf{a}$ has the structure that all but $k-1$ of its coordinates may be divided into a small number of intervals, and in each interval the coordinate value is the interpretation of a constant symbol from P. Write such a tuple as $\mathbf{a} = \mathbf{pu}$, splitting it into its processed part and its unprocessed part. If we choose tuples $\mathbf{u}_i \in F^{k-1}$ so that $\mathbf{u} = t(\mathbf{u}_1, \dots, \mathbf{u}_r)$ for some term t which centralizes P, then $\mathbf{a} = \mathbf{pu} = t(\mathbf{pu}_1, \dots, \mathbf{pu}_r)$. Hence our previously fully processed tuple $\mathbf{pu}$ may be processed one more step into fully processed $\mathbf{pu}_1, \dots, \mathbf{pu}_r$, where the unprocessed part lies in F^{k-1}. Now the finite number $|F|^{k-1}$ may replace our use of the number $|A|^{k-1}$ in the proof of Theorem 4.3.1, yielding a polynomial upper bound on growth rate. □

6.2. **Problems.** Theorem 6.1.4 motivates the following problem.

Problem 6.2.1. Is it true that, for an arbitrary infinite algebra $\mathbf{A}$ with a pointed k-cube term, if $\mathbf{A}^{k-1}$ is finitely generated, then $d_{\mathbf{A}}(n)$ is bounded above by a polynomial? Is it even true that $\mathbf{A}^{k-1}$ being finitely generated implies that $\mathbf{A}^n$ is finitely generated for all n?

The first statement of the problem has an affirmative answer for 0-pointed k-cube terms, as one sees by examining the proof of Theorem 5.4.1 and noting that finitely generated infinite modules have growth rate bounded by a linear function.

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

Problem 6.2.2. Is there a finite algebra $\mathbf{A}$ where $d_{\mathbf{A}}(n) \notin O(\log(n))$ and $d_{\mathbf{A}}(n) \notin \Omega(n)$?

We know that no 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 6.2.3. Is it true that a finite algebra with a 2-sided unit for some binary term has logarithmic or linear growth?

Despite the results of Theorem 2.2.6, Corollary 4.5.3, and Theorem 4.5.4, we still do not know if the growth rate of a finite abelian algebra must be linear or exponential. We pose the following problem.

Problem 6.2.4. Let $\mathbf{A}$ be a finite abelian algebra. What is the relationship between the following properties?

(i) $\mathbf{A}$ has a Maltsev term.
(ii) $\mathbf{A}$ has a pointed cube term.
(iii) $d_{\mathbf{A}}(n) \in \Theta(n)$.
(iv) $d_{\mathbf{A}}(n) \notin 2^{\Omega(n)}$.
(v) No finite power $\mathbf{A}^n$ has a nontrivial strongly abelian homomorphic image.

Are they equivalent? What if $\mathbf{A}$ generates an abelian variety? What if $\mathbf{A}$ is simple?

References

[1] Berman, Joel, Idziak, Paweł, *Generative complexity in algebra.* Mem. Amer. Math. Soc. **175** (2005).

[2] Berman, Joel, Idziak, Paweł, Marković, Petar, McKenzie, Ralph, Valeriote, Matthew, Willard, Ross, *Varieties with few subalgebras of powers.* Trans. Amer. Math. Soc. **362** (2010), no. 3, 1445–1473.

[3] Berman, Joel, McKenzie, Ralph, *Clones satisfying the term condition.* Discrete Math. **52** (1984), 7–29.

[4] Chen, Hubie, *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, BerlinHeidelberg, 2008), pp. 197–208.

[5] Erfanian, Ahmad, *On the growth sequences of free product of* $\mathrm{PSL}(m, q)$. Ital. J. Pure Appl. Math. No. **22** (2007), 19–26.

[6] Erfanian, Ahmad, *Growth sequence of free product of alternating groups.* Int. J. Contemp. Math. Sci. **2** (2007), no. 13-16, 685–691.

[7] Erfanian, Ahmad, *A note on growth sequences of* $\mathrm{PSL}(m, q)$. Southeast Asian Bull. Math. **29** (2005), no. 4, 697–713.

[8] Erfanian, Ahmad, *A note on growth sequences of alternating groups.* Arch. Math. (Basel) **78** (2002), no. 4, 257–262.

[9] Erfanian, Ahmad, *A problem on growth sequences of groups.* J. Austral. Math. Soc. Ser. A **59** (1995), no. 2, 283–286.
[10] Erfanian, A., Rezaei, R., *On the growth sequences of* $\mathrm{PSp}(2m, q)$. Int. J. Algebra **1** (2007), no. 1-4, 51–62.
[11] Erfanian, Ahmad, Wiegold, James, *A note on growth sequences of finite simple groups,* Bull. Austral. Math. Soc. **51** (1995), no. 3, 495–499.
[12] Foster, Alfred L., *On the finiteness of free (universal) algebras.* Proc. Amer. Math. Soc. **7** (1956), 10111013.
[13] Glass, A. M. W., Riedel, Herbert H. J., *Growth sequences–a counterexample.* Algebra Universalis 21 (1985), no. 2-3, 143–145.
[14] Hall, P., *The Eulerian functions of a group.* Quart. J. Math. **7** 1936., 134–151.
[15] Hobby, David, McKenzie, Ralph, *The structure of finite algebras.* Contemporary Mathematics, **76**. American Mathematical Society, Providence, RI, 1988.
[16] Hyde, J. T., Loughlin, N. J., Quick, M., Ruskuc, N., Wallis, A. R., *On the growth of generating sets for direct powers of semigroups.* Semigroup Forum **84** (2012), 116–130.
[17] Jónsson, Bjarni, *Algebras whose congruence lattices are distributive.* Math. Scand. **21** (1967), 110121.
[18] Kearnes, Keith, *An order–theoretic property of the commutator.* Internat. J. Algebra Comput. **3** (1993), 491–533.
[19] Kearnes, Keith, *A Hamiltonian property for nilpotent algebras.* Algebra Universalis **37** (1997), 403–421.
[20] Kearnes, Keith, Kiss, Emil, *Finite algebras of finite complexity.* Discrete Math. **207** (1999), 89–135.
[21] Kearnes, Keith, Szendrei, Ágnes, *Clones of algebras with parallelogram terms.* Internat. J. Algebra Comput. **22**, (2012).
[22] Kelly, David, *Basic equations: word problems and Mal'cev conditions.* Abstract 701-08-04, AMS Notices **20** (1972) A-54.
[23] Kimmerle, W., *Growth sequences relative to subgroups.* Groups–St. Andrews 1981 (St. Andrews, 1981), pp. 252–260, London Math. Soc. Lecture Note Ser., 71, Cambridge Univ. Press, Cambridge-New York, 1982.
[24] Kiss, Emil W., Valeriote, Matthew A., *Abelian algebras and the Hamiltonian property.* J. Pure Appl. Algebra **87** (1993), no. 1, 37–49.
[25] Lennox, John C., Wiegold, James, *Generators and killers for direct and free products.* Arch. Math. (Basel) **34** (1980), no. 4, 296–300.
[26] Lucchini, Andrea, *A bound on the presentation rank of a finite group.* Bull. London Math. Soc. **29** (1997), no. 4, 389–394.
[27] Maltsev, A. I., *On the general theory of algebraic systems (in Russian).* Mat. Sb. N. S. 35(77) (1954), 320.
[28] McKenzie, Ralph, *Narrowness implies uniformity.* Algebra Universalis, **15** (1982) 67–85.
[29] Meier, D., Wiegold, James, *Growth sequences of finite groups. V.* J. Austral. Math. Soc. Ser. A **31** (1981), no. 3, 374–375.
[30] Obraztsov, V. N., *Growth sequences of* 2*-generator simple groups.* Proc. Roy. Soc. Edinburgh Sect. A **123** (1993), no. 5, 839–855.
[31] Pollák, György, *Growth sequence of globally idempotent semigroups.* J. Austral. Math. Soc. Ser. A **48** (1990), no. 1, 87–88.
[32] Quick, Martyn, Ruškuc, Nik, *Growth of generating sets for direct powers of classical algebraic structures.* J. Austral. Math. Soc. **89** (2010), 105–126.

[33] Riedel, Herbert H. J., *Growth sequences of finite algebras.* Algebra Universalis **20** (1985), no. 1, 90–95.
[34] Stewart, A. G. R., Wiegold, James, *Growth sequences of finitely generated groups. II.* Bull. Austral. Math. Soc. **40** (1989), no. 2, 323–329.
[35] Thévenaz, Jacques, *Maximal subgroups of direct products.* J. Algebra **198** (1997), no. 2, 352–361.
[36] Wiegold, James, *Growth sequences of finite groups.* Collection of articles dedicated to the memory of Hanna Neumann, VI. J. Austral. Math. Soc. **17** (1974), 133–141.
[37] Wiegold, James, *Growth sequences of finite groups. II.* J. Austral. Math. Soc. **20** (1975), part 2, 225–229.
[38] Wiegold, James, *Growth sequences of finite groups. III.* J. Austral. Math. Soc. Ser. A **25** (1978), no. 2, 142–144.
[39] Wiegold, James, *Growth sequences of finite groups. IV.* J. Austral. Math. Soc. Ser. A **29** (1980), no. 1, 14–16.
[40] Wiegold, James, *Growth sequences of finite semigroups.* J. Austral. Math. Soc. Ser. A **43** (1987), no. 1, 16–20.
[41] Wiegold, James, Wilson, John S., *Growth sequences of finitely generated groups.* Arch. Math. (Basel) **30** (1978), no. 4, 337–343.
[42] Wise, Daniel T., *The rank of a direct power of a small-cancellation group.* Proceedings of the Conference on Geometric and Combinatorial Group Theory, Part I (Haifa, 2000). Geom. Dedicata **94** (2002), 215–223.

(Keith Kearnes) Department of Mathematics, University of Colorado, Boulder, CO 80309-0395, USA
E-mail address: Keith.Kearnes@Colorado.EDU

(Emil W. Kiss) Eötvös University, Department of Algebra and Number Theory, 1088 Budapest, Múzeum krt. 6–8, Hungary
E-mail address: ewkiss@cs.elte.hu

(Ágnes Szendrei) Department of Mathematics, University of Colorado, Boulder, CO 80309-0395, USA
E-mail address: Agnes.Szendrei@Colorado.EDU