A FINITE BASIS THEOREM FOR DIFFERENCE-TERM VARIETIES WITH A FINITE RESIDUAL BOUND

KEITH KEARNES, ÁGNES SZENDREI, AND ROSS WILLARD

ABSTRACT. We prove that if $\mathcal{V}$ is a variety of algebras (i.e., an equationally axiomatizable class of algebraic structures) in a finite language, $\mathcal{V}$ has a difference term, and $\mathcal{V}$ has a finite residual bound, then $\mathcal{V}$ is finitely axiomatizable. This provides a common generalization of R. McKenzie's finite basis theorem for congruence modular varieties with a finite residual bound, and R. Willard's finite basis theorem for congruence meet-semidistributive varieties with a finite residual bound.

This paper is a contribution to an old problem in logic from the schools of A. Tarski and A. Maltsev: which finite algebraic structures $\mathbf{A}$ (*algebras* for short) have a *finite basis* for their identities? Equivalently, for which finite algebras $\mathbf{A}$ is the variety $\mathcal{V}(\mathbf{A})$ (the smallest equational class containing $\mathbf{A}$) finitely axiomatizable?

On the one hand, every finite group [38], finite ring [22, 28], finite commutative semigroup [40], finite lattice [33], or two-element algebra in a finite language [30] is known to be finitely based. On the other hand, the list of finite algebras which are not finitely based includes, in addition to pathological examples (e.g., [31, 37]), some finite semigroups [40], some finite non-associative K-algebras [41, 29, 12], and even a finite group with one non-identity element named by a constant [5]. In 1996 R. McKenzie [36] proved that the problem of determining whether a finite algebra is finitely based is undecidable, settling Tarski's *finite basis problem.* The evidence suggests that a full classification of finitely based finite algebras is beyond reach.

However, there are some remarkable partial results. In particular, in the early 1970s K. Baker [1, 2] proved the following: if $\mathbf{A}$ is a finite algebra in a finite language and $\mathcal{V}(\mathbf{A})$ is *congruence distributive* (i.e., for every $\mathbf{B} \in \mathcal{V}(\mathbf{A})$, the lattice of congruence relations of $\mathbf{B}$ is a distributive lattice), then $\mathbf{A}$ is finitely based. Two important ingredients in the proof were provided by B. Jónsson [13]: (1) a characterization, in terms of identities, of the condition that a variety be congruence distributive, and (2) a proof that if $\mathbf{A}$ is finite and $\mathcal{V}(\mathbf{A})$ is congruence distributive, then every subdirectly

Date: July 8, 2014.

2010 *Mathematics Subject Classification.* Primary 03C05; Secondary 08B05.

Key words and phrases. equational theory, variety, finite basis, residually finite, difference term.

This material is based upon work supported by the Hungarian National Foundation for Scientific Research (OTKA) grant no. K83219 and K104251, and by the Natural Sciences and Engineering Research Council of Canada.

irreducible member of $\mathcal{V}(\mathbf{A})$ has size at most $|A|$. This second ingredient is equivalent to the statement that $\mathcal{V}(\mathbf{A})$ is "residually $\leq |A|$"; that is, distinct elements in any member of $\mathcal{V}(\mathbf{A})$ can be separated by homomorphisms to algebras of cardinality at most the cardinality of $\mathbf{A}$.

A variety is said to *have a finite residual bound* if it is residually $\leq r$ for some $r < \omega$ (equivalently, if every subdirectly irreducible member of the variety has size at most r). In the mid-1970s Jónsson asked whether the existence of a finite residual bound, or some related assumption on the residual character of the variety, already implies that a variety is finitely based.[1] For definiteness, we focus on the following version of the problem:

Problem (Jónsson): If a variety in a finite language has a finite residual bound, must the variety be finitely based?

A conjectured positive answer is sometimes called *Park's Conjecture* (see e.g. [45]), named for R.E. Park, in whose 1976 PhD thesis the conjecture explictly appears [39].

Jónsson's problem remains open to this day. Baker's theorem establishes a positive answer for congruence distributive varieties. R. McKenzie gave a significant generalization of Baker's theorem in 1987 when he answered the problem affirmatively for congruence modular varieties [35]. R. Willard, in 2000, extended Baker's theorem in a different direction by giving a positive answer for congruence meet-semidistributive varieties [44]. Roughly speaking, McKenzie's theorem exploited the well-behaved *commutator theory* for congruence modular varieties, which describes abelianness and the centrality relation among congruences. Willard's theorem, following Baker, exploited the combinatorial properties of congruence generation which arise from the absence of abelianness in congruence meet-semidistributive varieties. Both McKenzie's and Willard's theorems have been further extended [35, p. 226], [3, 32], but until now no common generalization of the two theorems has been found.

[1]Several variations of this question have been attributed to Jónsson. In [43, p. 477], from lectures delivered in 1975, W. Taylor writes "**Problem**. If the smallest variety containing a finite algebra $\mathcal{A}$ has no s.i. algebras except homomorphic images of subalgebras of $\mathcal{A}$, then is $\mathrm{Eq}(\mathcal{A})$ finitely based? (Jónsson et al.)" In the abstract [46, p. 1] of his talk at an Oberwolfach workshop in August 1976, Baker mentions "the conjecture of Jónsson that states that if a variety contains only finitely many subdirectly irreducible members, all finite, then it must be finitely definable." In [34, p. 337], published in 1977, R. McKenzie states that Jónsson "once asked whether (1) alone [having a finite residual bound] is enough to ensure a finite basis." Ten years later, in [35, p. 226], McKenzie writes that "Jónsson wondered, in the early 1970's, whether it is the case that every finite algebra $\mathbf{A}$ belonging to a residually small variety of finite type has a finite equational base."

The variation which we can with greatest confidence attribute to Jónsson is the following [46, p. 25], which is Problem 39 in the report from the Oberwolfach workshop mentioned earlier: "Jónsson: Is it true for every variety $\mathcal{V}$ of algebras that if the class $\mathcal{V}_{\mathrm{FSI}}$ of all finitely subdirectly irreducible algebras of $\mathcal{V}$ is strictly elementary, then $\mathcal{V}$ is finitely based?"

In this paper we provide such a generalization. That is, we answer Jónsson's problem affirmatively for a "natural" class of varieties containing all congruence modular varieties and all congruence meet-semidistributive varieties. As this class is not so well known, we will introduce it here and define it precisely in the next section. Consider groups; they have a ternary term $p(x, y, z)$ satisfying the *Maltsev identities* $p(x, x, y) \approx y$ and $p(x, y, y) \approx x$. (Just take $p(x, y, z) := xy^{-1}z$.) In general, a variety $\mathcal{V}$ is said to be *Maltsev* if it possesses a term satisfying these identities. A deep result of commutator theory states that every congruence modular variety has a term approximating a Maltsev term; that is, there exists a term $p(x, y, z)$ which

- satisfies the first Maltsev identity $p(x, x, y) \approx y$ throughout $\mathcal{V}$, and
- satisfies at least those instances $p(a, b, b) = a$ of the second identity where a, b belong to a common block of an *abelian* congruence of an algebra in the variety.

Congruence meet-semidistributive varieties also possess such a term (trivially – they have no nontrivial abelian congruences, so $p(x, y, z) := z$ is such a term). In general, if $\mathcal{V}$ is a variety and $p(x, y, z)$ is a term satisfying the two properties displayed above, then $p(x, y, z)$ is called a *difference term* for the variety and $\mathcal{V}$ is said to *have a difference term.*

Varieties with a difference term have been studied in [16, 19, 23, 25, 26, 27], and have been revealed to form a reasonably natural class of varieties. The commutator theory of such varieties satisfies some (though not all) of the desirable features of the modular commutator [23, 16]. K. Kearnes showed that the class of locally finite varieties having a difference term has a simple characterization in the language of tame congruence theory [16, Theorem 3.9]. Finally, Kearnes and Á. Szendrei in [19] implicitly gave a characterization, in terms of identities, of the condition that a variety have a difference term. Their characterization formally resembles "congruence meet-semidistributive + Maltsev" in the same way that H.-P. Gumm's characterization of congruence modularity [9] resembles "congruence distributive + Maltsev."

In this paper we answer Jónsson's problem affirmatively for the class of varieties having a difference term. That is, we prove:

Theorem. *If $\mathcal{V}$ is a variety in a finite language, $\mathcal{V}$ has a difference term, and $\mathcal{V}$ has a finite residual bound, then $\mathcal{V}$ is finitely based.*

Roughly speaking, the plan of the proof is as follows. Suppose $\mathcal{V}$ is a locally finite variety in a finite language. For each $n < \omega$ let $\mathcal{V}^{(n)}$ denote the variety in the same language axiomatized by the set of all n-variable identities which hold in $\mathcal{V}$. Clearly $\mathcal{V}^{(0)} \supseteq \mathcal{V}^{(1)} \supseteq \cdots$ and $\mathcal{V} = \bigcap_{n=0}^{\infty} \mathcal{V}^{(n)}$. G. Birkhoff [4] observed in 1935 that $\mathcal{V}$ is finitely based if and only if $\mathcal{V} = \mathcal{V}^{(n)}$ for some sufficiently large n. Now suppose in addition that $\mathcal{V}$ has a finite residual bound. Then a necessary condition for $\mathcal{V}$ to be finitely based is that some $\mathcal{V}^{(n)}$ have a finite residual bound. Conversely, an easy argument shows that if some $\mathcal{V}^{(n)}$ has a finite residual bound, then $\mathcal{V} = \mathcal{V}^{(k)}$ for some

$k \geq n$ and hence $\mathcal{V}$ is finitely based. Thus our strategy is to prove that $\mathcal{V}^{(n)}$ has a finite residual bound for sufficiently large n, assuming $\mathcal{V}$ has a difference term.

To a significant extent, our execution of this strategy follows McKenzie's proof of his affirmative answer to Jónsson's problem in the congruence modular case. A key tool used by McKenzie is a characterization, due to R. Freese and McKenzie [7], of congruence modular varieties having a finite residual abound. In Proposition 3.2 we give a straightforward generalization of this characterization to varieties with a difference term. An important component of both characterizations is a commutator property called "C1." Following McKenzie, we show in Proposition 3.3 that, for locally finite varieties with a difference term, the property C1 can be further analyzed as the conjunction of a handful of technical properties about the behavior of the commutator operation on low congruences of members of the variety. Thus much of our work in this paper is to show, for sufficiently large n, that $\mathcal{V}^{(n)}$ is locally finite and satisfies these technical properties. The latter task is accomplished by showing that each technical property can be expressed in $\mathcal{V}^{(n)}$ (for large n) by a first-order sentence and then invoking the compactness theorem of first-order logic.

Along the way, we provide a new characterization of varieties with a difference term (Section 1), analyze the principal congruence centralizer relation in such varieties (Section 4), prove that this relation is first-order definable if the variety has a finite residual bound (Section 5), and extend a result of E. Kiss from the congruence modular setting to varieties having a difference term (Section 6). The reader may profit by first reading McKenzie's paper [35], and having references [16] and [19] at hand may be desirable. A working knowledge of tame congruence theory (see [11]) is needed to follow some of our arguments.

1. Varieties having a difference term

In this section we formally define a difference term, and give a characterization of varieties having a difference term. We begin by reviewing basic definitions from commutator theory.

Suppose $\mathbf{A}$ is an algebra and $\alpha, \beta, \theta, \delta \in \operatorname{Con} \mathbf{A}$. $C(\alpha, \beta; \theta)$ is the usual ("term-condition") centralizer relation (see e.g. [11, Definition 3.3]). That is, $C(\alpha, \beta; \theta)$ is the assertion that for all $m+n$-ary term operations $t(\mathbf{x}, \mathbf{y})$ of $\mathbf{A}$, all $(a_i, b_i) \in \alpha$ and $(c_j, d_j) \in \beta$, if $t(\mathbf{a}, \mathbf{c}) \equiv t(\mathbf{a}, \mathbf{d}) \pmod{\theta}$ then $t(\mathbf{b}, \mathbf{c}) \equiv t(\mathbf{b}, \mathbf{d}) \pmod{\theta}$. $[\alpha, \beta]$ denotes the least $\gamma \in \operatorname{Con} \mathbf{A}$ satisfying $C(\alpha, \beta; \gamma)$. More generally, if $\delta \leq \alpha \wedge \beta$, then $[\alpha, \beta]_\delta$ denotes the least $\gamma \geq \delta$ such that $C(\alpha, \beta; \gamma)$; this equals the unique congruence $\gamma \geq \delta$ satisfying $\gamma/\delta = [\alpha/\delta, \beta/\delta]$, where the last commutator is calculated in $\mathbf{A}/\delta$. α is *abelian* if $[\alpha, \alpha] = 0_A$. We use annihilator notation $(\gamma : \beta)$ to denote the greatest $\alpha \geq \gamma$ satisfying $C(\alpha, \beta; \gamma)$; if $\gamma \leq \beta$, then this equals the unique congruence $\alpha \geq \gamma$ satisfying $\alpha/\gamma = (0_{A/\gamma} : \beta/\gamma)$. If $a, b \in A$, we also write $\operatorname{ann}(a, b)$ for the annihilator $(0_A : \operatorname{Cg}^{\mathbf{A}}(a, b))$.

Definition 1.1. A term $p(x, y, z)$ is a *difference term* for a variety $\mathcal{V}$ if

(i) $\mathcal{V}$ satisfies the identity $p(x, x, y) \approx y$, and
(ii) $\mathcal{V}$ satisfies the property $p^{\mathbf{A}}(a, b, b) \equiv a \pmod{[\alpha, \alpha]}$ for all $\mathbf{A} \in \mathcal{V}$ and $a, b \in A$, where $\alpha = \mathrm{Cg}^{\mathbf{A}}(a, b)$.

This definition is easily seen to be equivalent to the definition given in the previous section.[2] For congruence modular varieties the last term in Gumm's characterization for congruence modularity is always a difference term [10, p. 53]. On the other hand, because $[\alpha, \alpha] = \alpha$ in congruence meet-semidistributive varieties, [19, Corollary 4.7], it follows that the term $p(x, y, z) := z$ is a difference term for any congruence meet-semidistributive variety.

Following [14, 6], given any congruences α, β, γ of an algebra, we canonically define $\beta_0 = \beta$, $\gamma_0 = \gamma$, $\beta_{n+1} = \beta \vee (\alpha \wedge \gamma_n)$, and $\gamma_{n+1} = \gamma \vee (\alpha \wedge \beta_n)$. Then $\beta = \beta_0 \leq \beta_1 \leq \cdots$ and $\gamma = \gamma_0 \leq \gamma_1 \leq \cdots$; furthermore, if $\beta_\omega = \bigvee_n \beta_n$ and $\gamma_\omega = \bigvee_n \gamma_n$, then $\beta_\omega \vee \gamma_\omega = \beta \vee \gamma$ and $\alpha \wedge \beta_\omega = \alpha \wedge \gamma_\omega =: \delta$. This implies $C(\beta_\omega, \alpha; \delta)$ and $C(\gamma_\omega, \alpha; \delta)$, so $(\delta : \alpha) \geq \beta_\omega \vee \gamma_\omega$, so $[\beta \vee \gamma, \alpha] = [\beta_\omega \vee \gamma_\omega, \alpha] \leq \delta$. (See [24, Lemma 1(ii)] for a similar claim.)

Kearnes and Szendrei [19, Corollary 4.7] and P. Lipparini [26, Theorem 4.1] proved that a variety is congruence meet-semidistributive if and only if it satisfies the congruence inclusion $\alpha \cap (\beta \circ \gamma) \subseteq \beta_n$ for some $n \geq 0$. This implies that congruence meet-semidistributivity can be characterized by identities, similar to Jónsson's and Gumm's characterizations of congruence distributivity and congruence modularity respectively. (This is worked out in detail in [44].) Kearnes and Szendrei noted in [19, p. 522] that a variety has a difference term if and only if it satisfies the congruence inclusion[3] $\alpha \cap (\beta \circ \gamma) \subseteq (\alpha \wedge \beta_n) \circ \gamma \circ \beta$ for some $n < \omega$. In principle, this implies a characterization by identities for varieties having a difference term. In the following theorem we establish a version of these remarks which will be a crucial tool in the present paper.

Theorem 1.2. *Suppose $\mathcal{V}$ is a variety, $p(x, y, z)$ is a 3-ary term, and $p(x, x, y) \approx y$ holds in $\mathcal{V}$. The following are equivalent.*

(1) *p is a difference term for $\mathcal{V}$.*
(2) *There exists $n \geq 0$ such that, for all $\mathbf{A} \in \mathcal{V}$, $a, b, c \in A$, and $\alpha, \beta, \gamma \in \mathrm{Con}\,\mathbf{A}$, if $(a, c) \in \alpha$, $(a, b) \in \beta$, and $(b, c) \in \gamma$, then*

$$a \equiv p(a, b, c) \pmod{(\alpha \wedge \beta_n) \circ \gamma}.$$

[2]We note that [8] defines a "difference term" by *three* properties, namely (i) and (ii) above along with a third property we will not introduce. It can be shown that a variety has a term satisfying all three properties if and only if it is congruence modular, and for congruence modular varieties the third property is equivalent to the conjunction of properties (i) and (ii) of Definition 1.1. We follow [20] in defining a difference term with only these two properties.

[3][19], following [25], gave $\alpha \cap (\beta \circ \gamma) \subseteq \gamma \circ \beta \circ (\alpha \wedge \gamma_n)$ for some n, which is equivalent.

(3) *There exist* 3*-ary terms* f_i, g_i ($i \in I$*,* I *a finite set*) *satisfying the following conditions throughout* $\mathcal{V}$*:*
   (a) $f_i(x,x,x) \approx x \approx g_i(x,x,x)$ *for all* $i \in I$*.*
   (b) $f_i(x,y,x) \approx g_i(x,y,x)$ *for all* $i \in I$*.*
   (c) $\bigwedge_{i\in I}[f_i(x,x,y) = g_i(x,x,y) \leftrightarrow f_i(x,y,y) = g_i(x,y,y)] \ \rightarrow\ p(x,y,y) = x$.

In particular, $\mathcal{V}$ *has a difference term if and only if* $\mathcal{V}$ *satisfies the congruence inclusion* $\alpha \cap (\beta \circ \gamma) \subseteq (\alpha \wedge \beta_n) \circ \gamma \circ \beta$ *for some* $n \geq 0$*.*

Proof. (1) $\Rightarrow$ (2). Define β_n, γ_n ($n \leq \omega$) and δ as in the paragraphs preceding the statement of the theorem. Also let $\theta = \alpha \wedge (\beta \vee \gamma)$. Then $(a,c) \in \theta$, so $(a, p(a,c,c)) \in [\theta,\theta]$ as p is a difference term for $\mathcal{V}$. But $[\theta,\theta] \leq [\beta \vee \gamma, \alpha] \leq \delta$ (see the remarks preceding the statement of the theorem), which proves $(a, p(a,c,c)) \in \alpha \wedge \beta_n$ for some $n < \omega$. We can get a uniform n by applying this argument in the usual way to $\mathbf{A} = \mathbb{F}_{\mathcal{V}}(x,y,z)$, $\alpha = \mathrm{Cg}^{\mathbf{A}}(x,z)$, $\beta = \mathrm{Cg}^{\mathbf{A}}(x,y)$, $\gamma = \mathrm{Cg}^{\mathbf{A}}(y,z)$, and $(a,b,c) = (x,y,z)$. Finally, $p(a,c,c) \equiv p(a,b,c) \pmod{\gamma}$, proving $(a,p(a,b,c)) \in (\alpha \wedge \beta_n) \circ \gamma$.

(2) $\Rightarrow$ (3). As in the previous paragraph, we apply (2) to $\mathbf{A} = \mathbb{F}_{\mathcal{V}}(x,y,z)$, $\alpha = \mathrm{Cg}^{\mathbf{A}}(x,z)$, $\beta = \mathrm{Cg}^{\mathbf{A}}(x,y)$, $\gamma = \mathrm{Cg}^{\mathbf{A}}(y,z)$, and $(a,b,c) = (x,y,z)$. By (2), there exists a 3-ary term q with $(x,q) \in \alpha \wedge \beta_n$ and $(q,p) \in \gamma$. Applying the argument in the proof of [44, Theorem 2.1(4$\Rightarrow$5)] to $(x,q) \in \alpha \wedge \beta_n$ produces

- a finite vertex-labelled planar tree T with vertex set I and root 0, such that every vertex is labelled by `b` or `g`, and a child always has the opposite label of its parent;
- a set $\{(s_i,t_i) : i \in I\}$ of pairs of 3-ary terms indexed by I;

such that the identities (1)–(10) named in the proof of [44, Theorem 2.1(4$\Rightarrow$5)] hold in $\mathcal{V}$, with one exception: in place of $t_0(x,y,z) \approx z$ we get $t_0(x,y,z) \approx q(x,y,z)$. These identities include $s_i(x,y,x) \approx t_i(x,y,x)$ and imply $s_i(x,x,x) \approx x \approx t_i(x,x,x)$ for all $i \in I$. The proof of [44, Theorem 2.1(5$\Rightarrow$6)] can now be adapted in the obvious way to show that if $\mathbf{B} \in \mathcal{V}$, $a, b \in B$ and $s_i(a,a,b) = t_i(a,a,b) \leftrightarrow s_i(a,b,b) = t_i(a,b,b)$ for all $i \in I$, then $a = q(a,b,b)$. As $(q,p) \in \gamma$ we have $q(a,b,b) = p(a,b,b)$, which establishes (3) (modulo renaming s_i, t_i as f_i, g_i).

(3) $\Rightarrow$ (1). Suppose $\mathbf{A} \in \mathcal{V}$, $a, b \in A$, and $\alpha = \mathrm{Cg}^{\mathbf{A}}(a,b)$. Let $\varepsilon = [\alpha,\alpha]$; we must show $p(a,b,b) \equiv a \pmod{\varepsilon}$. By passing to $\mathbf{A}/\varepsilon$ we can assume that $\varepsilon = 0_A$, so α is abelian, and under this assumption we must show that $p(a,b,b) = a$. By (3), it will suffice to prove $f_i(a,a,b) = g_i(a,a,b) \leftrightarrow f_i(a,b,b) = g_i(a,b,b)$ for all $i \in I$. Fix $i \in I$ and define

$$T(x_1,\ldots,x_9) = g_i(f_i(x_1,x_2,x_3), f_i(x_4,x_5,x_6), f_i(x_7,x_8,x_9)).$$

Assume $f_i(a,b,b) = g_i(a,b,b)$; then

$$T(a,b,b,a,\underline{b},\underline{b},a,b,b) \;=\; T(a,a,a,b,\underline{b},\underline{b},b,b,b).$$

Because $(a,b)\in\alpha$ and α is abelian, we can replace the underlined occurrences of b by a to get

$$T(a,b,b,a,\underline{a},\underline{a},a,b,b) = T(a,a,a,b,\underline{a},\underline{a},b,b,b). \tag{1.1}$$

Recall that $f_i(x,y,z)\approx g_i(x,y,x)$; hence

$$T(a,b,\underline{a},a,\underline{b},a,a,b,\underline{a}) = T(a,a,\underline{a},b,\underline{b},b,a,a,\underline{a}).$$

Again because $(a,b)\in\alpha$ and α is abelian, we deduce

$$T(a,b,\underline{b},a,\underline{a},a,a,b,\underline{b}) = T(a,a,\underline{b},b,\underline{a},b,a,a,\underline{b}). \tag{1.2}$$

Combining (1.1) and (1.2) gives

$$T(a,a,b,\underline{b},a,b,a,a,b) = T(a,a,a,\underline{b},a,a,b,b,b).$$

One more application of abelianness gives

$$T(a,a,b,\underline{a},a,b,a,a,b) = T(a,a,a,\underline{a},a,a,b,b,b),$$

i.e., $f_i(a,a,b)=g_i(a,a,b)$, proving the reverse implication. The forward implication is proved similarly.

To prove the final claim, suppose first p is a difference term for $\mathcal{V}$ and let n be given by (2). Assume $\mathbf{A}\in\mathcal{V}$, $\alpha,\beta,\gamma\in\operatorname{Con}\mathbf{A}$, and $(a,c)\in\alpha\cap(\beta\circ\gamma)$. Pick b with $(a,b)\in\beta$ and $(b,c)\in\gamma$; then (2) gives $(a,p(a,b,c))\in(\alpha\wedge\beta_n)\circ\gamma$, and clearly $c=p(a,a,c)\equiv p(a,b,c) \pmod{\beta}$. Thus $p(a,b,c)$ witnesses $(a,c)\in((\alpha\wedge\beta_n)\circ\gamma)\circ\beta$, proving the congruence inclusion.

Suppose next that $\mathcal{V}$ satisfies the congruence inclusion $\alpha\cap(\beta\circ\gamma)\subseteq(\alpha\wedge\beta_n)\circ\gamma\circ\beta$. Applying it in the usual way to $\mathbb{F}_{\mathcal{V}}(x,y,z)$ etc. yields 3-ary terms q,p satisfying the following: $q(x,y,y)\approx p(x,y,y)$, $p(x,x,y)\approx y$, and for all $\mathbf{A}\in\mathcal{V}$, $a,b,c\in A$, and $\alpha,\beta,\gamma\in\operatorname{Con}\mathbf{A}$ with $(a,c)\in\alpha$, $(a,b)\in\beta$, and $(b,c)\in\gamma$, we have $(a,q(a,b,c))\in\alpha\wedge\beta_n$. In particular, $q(a,b,c)\equiv q(a,c,c)=p(a,c,c)\equiv p(a,b,c) \pmod{\gamma}$. Thus $(a,p(a,b,c))\in(\alpha\wedge\beta_n)\circ\gamma$, so p is a difference term as it satisfies (2). □

We note in passing that, by arguing as in [19, Section 5] and using [16, Lemma 2.7], one can show that a locally finite variety has a difference term if and only if it satisfies $\alpha\cap(\beta\circ\gamma)\subseteq(\alpha\wedge\beta_2)\circ\gamma\circ\beta$.

2. Commutator properties of varieties with a difference term

In this section we gather some known properties of varieties with a difference term.

Lemma 2.1. *Let $\mathcal{V}$ be a variety with a difference term p, and suppose $\mathbf{A}\in\mathcal{V}$ and $\alpha,\beta,\alpha_i\in\operatorname{Con}\mathbf{A}$.*

(1) $[\alpha,\beta]=[\beta,\alpha]$.
(2) *If $[\alpha_i,\beta]=0_A$ for all $i\in I$, then $[\bigvee_i\alpha_i,\beta]=0_A$.*
(3) *If α is abelian, then p is a Maltsev operation on each α-block.*

Proof. (1) is proved in [16, Lemma 2.2]. (2) holds in all varieties, not just those with a difference term. (3) follows directly from the definition of a difference term. □

We call the properties in items (1), (2) above *symmetry* and *left semi-distributivity.*

Definition 2.2. Fix an algebra $\mathbf{A}$ and $a, b, c, d \in A$. Let $\alpha = \mathrm{Cg}^{\mathbf{A}}(a, b)$ and $\beta = \mathrm{Cg}^{\mathbf{A}}(c, d)$.

(1) $C(a, b, c, d)$ denotes the condition $C(\alpha, \beta; 0_A)$.
(2) $C_2(a, b, c, d)$ denotes the *two-term condition* for (α, β). That is, $C_2(a, b, c, d)$ iff for all $m, n \geq 1$, all $r_1, r_2 \in \mathrm{Pol}_{m+n}(\mathbf{A})$, all $(a_1, b_1), \ldots, (a_m, b_m) \in \alpha$, and all $(c_1, d_1), \ldots, (c_n, d_n) \in \beta$, if three of the pairs

$$(r_1(\mathbf{a}, \mathbf{c}), r_2(\mathbf{a}, \mathbf{c})),\ (r_1(\mathbf{a}, \mathbf{d}), r_2(\mathbf{a}, \mathbf{d})),\ (r_1(\mathbf{b}, \mathbf{c}), r_2(\mathbf{b}, \mathbf{c})),\ (r_1(\mathbf{b}, \mathbf{d}), r_2(\mathbf{b}, \mathbf{d}))$$

are in 0_A, then so is the fourth.

Lemma 2.3. *If $\mathcal{V}$ has a difference term, then for all $\mathbf{A} \in \mathcal{V}$ and all $a, b, c, d \in A$, $C(a, b, c, d) \iff C_2(a, b, c, d)$.*

Proof. Follows from [19, Corollary 4.5] and Lemma 2.1(1). □

Lemma 2.4. *Suppose $\mathcal{V}$ is a variety with a difference term, $\mathbf{A} \in \mathcal{V}$, and $\alpha, \beta, \delta \in \mathrm{Con}\,\mathbf{A}$.*

(1) $[\alpha, \beta] = 0_A$ *iff* $[\mathrm{Cg}^{\mathbf{A}}(a, b), \mathrm{Cg}^{\mathbf{A}}(c, d)] = 0_A$ *for all* $(a, b) \in \alpha$ *and* $(c, d) \in \beta$.
(2) *If* $\delta \leq \alpha \wedge \beta$, *then* $[\alpha, \beta]_\delta = [\alpha, \beta] \vee \delta$.
(3) *If* α *is abelian, then* $[\alpha, \beta] \leq \delta$ *iff* $C(\beta, \alpha; \delta)$.

Proof. (1) follows in the usual way from left semi-distributivity and symmetry. (2) is [16, Lemma 2.4]. (3) follows from [16, Lemma 2.3] and symmetry. □

We also need the following two facts from tame congruence theory.

Lemma 2.5 ([16, Theorem 3.9])**.** *A locally finite variety $\mathcal{V}$ has a difference term if and only if (i) $\mathcal{V}$ omits type* **1** *and (ii) every type* **2** *minimal set has empty tail.*

Lemma 2.6. *Suppose $\mathbf{A}$ is a finite algebra and θ/δ and θ'/δ' are perspective prime quotients in* $\mathrm{Con}\,\mathbf{A}$ *with* $\mathrm{typ}(\delta, \theta) \neq \mathbf{1}$. *Then* $(\delta : \theta) = (\delta' : \theta')$.

Proof. If θ/δ is non-abelian, then the claim follows from [11, Remark 5.13]. Assume $\mathrm{typ}(\delta, \theta) = \mathbf{2}$. Let U be a (δ, θ)-minimal set with $U = e(\mathbf{A})$ for some $e \in \mathrm{Pol}_1(\mathbf{A})$ satisfying $e^2 = e$. As perspective prime quotients have the same minimal sets [11, Lemma 6.2], it suffices to show that $(\delta : \theta)$ has an intrinsic characterization referencing only $\mathbf{A}$ and U. Define

$$\begin{aligned}\gamma \;=\; \{(a, b) \in A^2 : \forall f \in \mathrm{Pol}_2(\mathbf{A}),\ & ef(a, x){\restriction_U} \text{ is a permutation} \\ \iff\ & ef(b, x){\restriction_U} \text{ is a permutation}\}.\end{aligned}$$

γ is clearly an equivalence relation and is invariant under unary polynomials of $\mathbf{A}$, so is a congruence. Let N be a (δ,θ)-trace in U. The proof of [18, Lemma 3.4 Case 2] shows $C(\gamma, N^2;\delta)$, which implies $C(\gamma,\theta;\delta)$ by [15, Lemma 4.2]; hence $\gamma \leq (\delta:\theta)$. To prove equality, assume that $(\delta:\theta)\not\leq\gamma$ and choose $(a,b)\in(\delta:\theta)\setminus\gamma$. By definition there is $f\in \mathrm{Pol}_2(\mathbf{A})$ such that $ef(a,x){\restriction_U}$ is a permutation of U and $ef(b,x){\restriction_U}$ is not (or the same with a and b interchanged). Let 0 and 1 be elements of N that are not δ related. Then (by properties of minimal sets) $(ef(b,0),ef(b,1))\in\delta$ while $(ef(a,0),ef(a,1))\notin\delta$, which proves $C((\delta:\theta),\theta;\delta)$ fails, which is impossible. □

Remark. Lemma 2.6 was proved in the case $U=A$ in [21, Theorem 3.4].

Finally, we need a fact about abelian principal congruences.

Definition 2.7. Given a variety $\mathcal{V}$ with difference term p, $\mathbf{A}\in\mathcal{V}$, $a,b\in A$, and $r>0$, let $\Gamma_r(a,b)=\{(u,p(t(a,\mathbf{e}),t(b,\mathbf{e}),u)) : t\in \mathrm{Clo}_{r+1}(\mathbf{A}),\ \mathbf{e}\in A^r,\ u\in A\}$.

Lemma 2.8. *Suppose $\mathcal{V}$ is a variety with difference term p, $\mathbf{A}\in\mathcal{V}$, $a,b\in A$, and $r>0$. If $\mathrm{Cg}^{\mathbf{A}}(a,b)$ is abelian and $\mathrm{ann}(a,b)$ has index at most r, then*

(1) $\mathrm{Cg}^{\mathbf{A}}(a,b)=\Gamma_r(a,b)$.
(2) *Each block of* $\mathrm{Cg}^{\mathbf{A}}(a,b)$ *has size at most* $|\mathbb{F}_{\mathcal{V}}(r+1)|$.

Proof. For (1), we mimic McKenzie's proof of the same claim in the congruence modular case [35, Lemma 2.16]. Start with the observation that the set θ of all pairs $(f(a),f(b))$, where f is a unary polynomial of $\mathbf{A}$, is a reflexive compatible relation on $\mathbf{A}$. The facts that $\mathrm{Cg}^{\mathbf{A}}(a,b)$ is abelian and that p is Maltsev on classes of abelian congruences implies that θ is a congruence. Hence if $(u,v)\in\mathrm{Cg}^{\mathbf{A}}(a,b)$, then there exists a polynomial $f(x)=s(x,\mathbf{e})$, $s\in\mathrm{Clo}(\mathbf{A})$, such that $(v,u)=(s(a,\mathbf{e}),s(b,\mathbf{e}))$. Then

$$p(s(a,\underline{\mathbf{e}}),s(b,\underline{\mathbf{e}}),u)=p(v,u,u)=v=p(v,v,v)=p(s(a,\underline{\mathbf{e}}),s(a,\underline{\mathbf{e}}),v).$$

Since $(a,b),(u,v)\in\mathrm{Cg}^{\mathbf{A}}(a,b)$, we retain the equality of left and right hand sides if we simultaneously change all underlined occurrences of $\mathbf{e}$ to any tuple $\mathbf{e}'$ that is congruent to $\mathbf{e}$ modulo $\mathrm{ann}(a,b)$ coordinatewise. Thus

$$p(s(a,\underline{\mathbf{e}}'),s(b,\underline{\mathbf{e}}'),u)=p(s(a,\underline{\mathbf{e}}'),s(a,\underline{\mathbf{e}}'),v)=v,$$

where the last equality follows from the identity $p(x,x,y)\approx y$. We may choose $\mathbf{e}'$ so that it has at most $|A/\mathrm{ann}(a,b)|\leq r$ distinct entries, and write $\mathbf{e}''$ for a sequence of length r containing the distinct entries of $\mathbf{e}'$ (possibly with repeated entries). There is a $t\in\mathrm{Clo}_{r+1}(\mathbf{A})$ such that $s(x,\mathbf{e}')=t(x,\mathbf{e}'')$ holds for all $x\in A$. This shows that

$$(u,v)=(u,p(t(a,\underline{\mathbf{e}}''),t(b,\underline{\mathbf{e}}''),u))$$

for some $t\in\mathrm{Clo}_{r+1}(\mathbf{A})$, $\mathbf{e}''\in A^r$, $u\in A$. Hence $(u,v)\in\Gamma_r(a,b)$. Since $(u,v)\in\mathrm{Cg}^{\mathbf{A}}(a,b)$ was arbitrary, we get that $\mathrm{Cg}^{\mathbf{A}}(a,b)\subseteq\Gamma_r(a,b)$.

The reverse inclusion, $\Gamma_r(a,b) \subseteq \mathrm{Cg}^{\mathbf{A}}(a,b)$, is an immediate consequence of the identity $p(x,x,y) \approx y$. Specifically, if $(u,v) = (u, p(t(a,\underline{\mathrm{e}}), t(b,\underline{\mathrm{e}}), u)) \in \Gamma_r(a,b)$, then

$$v = p(t(a,\mathbf{e}), t(b,\mathbf{e}), u) \overset{\mathrm{Cg}(a,b)}{\equiv} p(t(a,\mathbf{e}), t(a,\mathbf{e}), u) = u.$$

(2) follows from (1) and the fact that we can choose one fixed $\mathbf{e}$ in the definition of $\Gamma_r(a,b)$ (namely, any $\mathbf{e}$ containing a transversal for $\mathrm{ann}(a,b)$). Then for any $u \in A$ the function

$$t \mapsto p(t(a,\mathbf{e}), t(b,\mathbf{e}), u)$$

maps $\mathrm{Clo}_{r+1}(\mathbf{A})$ surjectively onto the $\mathrm{Cg}^{\mathbf{A}}(a,b)$-block of u. □

3. The commutator identity C1

C1 is the commutator identity $[\alpha \wedge \beta, \beta] = \alpha \wedge [\beta,\beta]$ (asserted for all congruences α, β of an algebra), or equivalently, the implication $\alpha \le [\beta,\beta] \implies [\alpha,\beta] = \alpha$. C1 was identified in [7] and named in [8]. If $\mathcal{V}$ is a locally finite variety, we say that $\mathcal{V}$ satisfies C1* if $(0_A : \mu)$ is abelian for every finite subdirectly irreducible algebra $\mathbf{A} \in \mathcal{V}$ with abelian monolith μ. In this section we collect the facts about C1 and C1* that we will need.

Proposition 3.1. *Suppose $\mathcal{V}$ is a locally finite variety with a difference term. $\mathcal{V}$ satisfies* C1 *if and only if it satisfies* C1*.

Proof sketch. The ($\Rightarrow$) implication is proved by applying C1 to the situation $\alpha = \mu$ and $\beta = (0_A : \mu)$.

For ($\Leftarrow$), assume that C1 fails. Then it fails in a finite algebra $\mathbf{A} \in \mathcal{V}$; say $\alpha, \beta \in \mathrm{Con}\,\mathbf{A}$ with $\alpha \le \beta$ and $[\alpha,\beta] < \alpha \wedge [\beta,\beta]$. Because $[\alpha,\beta]$ lies below all the relevant congruences and commutators in this witnessing failure, we can factor by $[\alpha,\beta]$ and obtain a parallel failure of C1 in $\mathbf{A}/[\alpha,\beta]$ (using Lemma 2.4(2)). Thus we may assume that $[\alpha,\beta] = 0_A$ and $\alpha \wedge [\beta,\beta] > 0_A$. Choose an atom γ below $\alpha \wedge [\beta,\beta]$. Let δ be a completely meet irreducible that is disjoint from γ, but whose upper cover θ contains γ. Then $(0_A : \gamma) = (\delta : \theta) =: \psi$ by Lemmas 2.5 and 2.6, and $\beta \le \psi$ because $[\beta,\alpha] = [\alpha,\beta] = 0_A$ and $\gamma \le \alpha$. Then in $\mathbf{A}/\delta$, $(0_{A/\delta} : \theta/\delta) = (\delta : \theta)/\delta = \psi/\delta$, while Lemma 2.4(2) yields $[\psi/\delta, \psi/\delta] = ([\psi,\psi] \vee \delta)/\delta \ne 0_{A/\delta}$ as $[\beta,\beta] \not\le \delta$ and $\beta \le \psi$. Hence $\mathbf{A}/\delta$ violates C1*. □

Proposition 3.2. *Suppose $\mathcal{V}$ is a locally finite variety omitting type* **1**.

1. *If $\mathcal{V}$ is residually small, then $\mathcal{V}$ satisfies* C1*.
2. *$\mathcal{V}$ has a finite residual bound if and only if it satisfies* C1* *and there exists a positive integer r such that for every finite subdirectly irreducible $\mathbf{A} \in \mathcal{V}$ with monolith μ, $(0_A : \mu)$ has index at most r.*

Proof. (1) follows from [17, Corollary 4.3].

(2) Necessity follows from (1). For sufficiency, assume that $\mathcal{V}$ satisfies C1* and that r satisfies the condition stated in (2). Let $m = |\mathbb{F}_{\mathcal{V}}(r+1)|$. It will suffice by Quackenbush's Theorem [42] to prove that every *finite* subdirectly irreducible $\mathbf{A} \in \mathcal{V}$ has size at most $r \cdot m^m$. Let μ be the monolith of $\mathbf{A}$ and let $\alpha = (0_A : \mu)$. The claim follows immediately if μ is nonabelian, so assume μ is abelian. Then $\mathrm{typ}(0_A, \mu) = \mathbf{2}$, and α is abelian by C1*. The proof of [17, Theorem 5.1] shows that each class of α has size at most m^m, which proves the claim. (Alternatively, one can mimic the proof of [7, Theorem 8] to show that each class of α has size at most $(m+1)!$.) $\square$

The next result is inspired by McKenzie's proof of his finite basis theorem [35].

Proposition 3.3. *Let $\mathcal{V}$ be a locally finite variety with a difference term. $\mathcal{V}$ satisfies* C1 *if and only if for all (or all finite) $\mathbf{A} \in \mathcal{V}$, all of the following conditions hold:*

(1) *If $\alpha, \beta \in \mathrm{Con}\,\mathbf{A}$ are abelian, then $\alpha \vee \beta$ is abelian.*
(2) *If β is a principal congruence of $\mathbf{A}$ and $[\beta, [\beta, \beta]] = 0_A$, then β is abelian.*
(3) *If $\alpha_0, \alpha_1, \beta_1, \beta_2 \in \mathrm{Con}\,\mathbf{A}$ with β_1, β_2 principal, $0_A \prec \alpha_0 \prec \alpha_1$, α_1 abelian, and $[\alpha_0, \beta_1] = [\alpha_0, \beta_2] = 0_A$, then there exists an abelian atom $\gamma \in \mathrm{Con}\,\mathbf{A}$ such that $[\alpha_1, \beta_1], [\alpha_1, \beta_2] \leq \gamma$.*
(4) *If $0_A \prec \alpha \prec \beta$ in $\mathrm{Con}\,\mathbf{A}$ with α abelian and $[\beta, \beta] = \beta$, then $[\alpha, \beta] = \alpha$.*

Proof. ($\Rightarrow$) Assume $\mathcal{V}$ satisfies C1 and $\mathbf{A} \in \mathcal{V}$. To prove (1), it suffices by semi-distributivity and symmetry to prove $[\alpha, \beta] = 0_A$. Let $\delta = [\alpha, \beta]$. Then $[\alpha, \delta] \leq [\alpha, \alpha] = 0_A$ and similarly $[\beta, \delta] = 0_A$. Thus if $\gamma = \alpha \vee \beta$ then $[\delta, \gamma] = [\gamma, \delta] = 0_A$ by semi-distributivity. But $\delta \leq [\gamma, \gamma]$, so with C1 this implies $\delta = 0_A$ as required. To prove (2), let $\alpha = [\beta, \beta]$ and apply the implicational version of C1. To prove (3), note first that the hypotheses give $[\alpha_0, \beta_1 \vee \beta_2] = 0_A$. Let $\delta := [\alpha_1, \beta_1 \vee \beta_2]$ and assume next that δ is not at height 0 or 1 in $\mathrm{Con}\,\mathbf{A}$. Note that as α_1 is abelian, the interval from 0_A to α_1 is solvable, hence is a modular sublattice of $\mathrm{Con}\,\mathbf{A}$ by Lemma 2.5 and [11, Lemma 6.5] and so has height 2. These facts and our assumption about the height of δ imply $\delta = \alpha_1$. But then $\alpha_1 \leq [\beta_1 \vee \beta_2, \beta_1 \vee \beta_2]$, so C1 implies $\alpha_0 = [\alpha_0, \beta_1 \vee \beta_2]$, contrary to a previous calculation. Thus δ has height 0 or 1. If $\delta = 0_A$ then we can take $\gamma = \alpha_0$, while if $\delta \neq 0_A$ then we can take $\gamma = \delta$, which proves (3). (4) is an immediate consequence of C1.

($\Leftarrow$) Assume that every finite $\mathbf{A} \in \mathcal{V}$ satisfies (1)–(4) but $\mathcal{V}$ fails to satisfy C1. Then C1 fails in some finite member of $\mathcal{V}$. Let $\mathbf{A}$ be a finite member of $\mathcal{V}$ of minimum cardinality in which C1 fails, and pick $\alpha, \beta \in \mathrm{Con}\,\mathbf{A}$ with $\alpha \leq [\beta, \beta]$ but $[\alpha, \beta] < \alpha$. Note that, by (1), $\mathbf{A}$ has a largest abelian congruence which we will denote by ν.

We first prove $(*)$ $[\theta, [\theta, \theta]] = [\theta, \theta]$ for all $\theta \in \mathrm{Con}\,\mathbf{A}$. Indeed, suppose $\delta := [\theta, [\theta, \theta]] < [\theta, \theta]$. Then in $\mathbf{A}/\delta$, $[\theta/\delta, \theta/\delta] = [\theta, \theta]/\delta \neq 0_{A/\delta}$ but $[\theta/\delta, [\theta/\delta, \theta/\delta]] = 0_{A/\delta}$, both by Lemma 2.4(2). Hence $\mathbf{A}/\delta$ fails to satisfy C1, so by minimality we have $\delta = 0_A$. Next observe that, since $[\theta, \theta] \neq 0_A$ we have $\theta \not\leq \nu$; pick $(a, b) \in \theta \setminus \nu$ and put $\beta' = \mathrm{Cg}^{\mathbf{A}}(a, b)$. Then β' violates condition (2).

In particular, $[\beta,\beta] = [[\beta,\beta],\beta]$. As $\alpha \leq [\beta,\beta]$ but $[\alpha,\beta] \neq \alpha$, there exist $\alpha \leq \alpha_0 \prec \alpha_1 \leq [\beta,\beta]$ such that $[\alpha_1,\beta] = \alpha_1$ but $[\alpha_0,\beta] < \alpha_0$. Because $[\alpha_0,\beta]$ is below all of the relevant congruences and commutators, we can factor by it and still preserve the above facts (by Lemma 2.4(2)); thus by minimality we have $[\alpha_0,\beta] = 0_A$. As $\alpha_0 \leq \alpha_1 = [\alpha_1,\beta] \leq \beta$, it follows that α_0 is abelian. Choose $0_A \leq \psi \prec \alpha_0$. Lemma 2.4(3) then gives $C(\beta,\alpha_0;\psi)$. Hence we can factor by ψ and preserve the relevant facts, so by minimality, $0_A \prec \alpha_0$.

In summary, we have $0_A \prec \alpha_0 \prec \alpha_1$, $[\alpha_1,\beta] = \alpha_1$, and $[\alpha_0,\beta] = 0_A$. This implies $\alpha_1 \leq \beta$, so $\alpha_0 \leq [\beta,\beta]$. The proof of Proposition 3.1 then shows **A** is subdirectly irreducible, by minimality, so α_0 is its monolith. Consider $\mu := [\alpha_1,\alpha_1]$. By the fact $(*)$ established two paragraphs back (with $\theta = \alpha_1$), $[\alpha_1,\mu] = \mu$, which with $[\alpha_0,\alpha_1] = 0_A$ implies $\mu \neq \alpha_0$. Hence $\mu \in \{0_A,\alpha_1\}$.

Case 1: $\mu = 0_A$, i.e., α_1 is abelian.

Assume that β is minimal among all congruences β' satisfying $[\alpha_0,\beta'] = 0_A$ and $[\alpha_1,\beta'] = \alpha_1$. As $[\beta,\alpha_1] = \alpha_1$ we have $\neg C(\beta,\alpha_1;0_A)$ and $\neg C(\beta,\alpha_1;\alpha_0)$. We can pick principal congruences $\beta_1,\beta_2 \leq \beta$ witnessing $\neg C(\beta_1,\alpha_1;0_A)$ and $\neg C(\beta_2,\alpha_1;\alpha_0)$. Then $\beta' := \beta_1 \vee \beta_2$ satisfies $[\alpha_0,\beta'] = 0_A$ and $[\alpha_1,\beta'] = \alpha_1$, so $\beta = \beta'$.

By condition (3) and subdirect irreducibility, we have $[\alpha_1,\beta_i] \leq \alpha_0$ for $i = 1,2$. If $[\alpha_1,\beta_1] = [\alpha_1,\beta_2]$ then $[\alpha_1,\beta] \leq \alpha_0$ by semi-distributivity, contradicting $[\alpha_1,\beta] = \alpha_1$. Hence $[\alpha_1,\beta_1] = \alpha_0$ and $[\alpha_1,\beta_2] = 0_A$. Hence $C(\beta_1,\alpha_1;\alpha_0)$, and as α_1 is abelian, we get $C(\beta_2,\alpha_1;\alpha_0)$ by Lemma 2.4(3). Thus $C(\beta,\alpha_1;\alpha_0)$, which contradicts $[\alpha_1,\beta] = \alpha_1$. This case is impossible.

Case 2: $\mu = \alpha_1$.

Then we have a violation of condition (4) (with α,β replaced by α_0,α_1). This case is also impossible. □

4. Characterizing the principal centralizer relation

Our goal in this section is to provide characterizations of $C(a,b,c,d)$ in varieties with a difference term, similar to the characterizations of $C(a,b,c,d)$ in congruence modular varieties provided in [35, Theorem 2.7] and [8, Chapter 6, Exercise 6].

Definition 4.1. Fix an algebra **A** and $a,b,c,d \in A$.

(1) If $r \in \mathrm{Pol}_2(\mathbf{A})$, $\vec{H}^r(a,b,c,d)$ is the implication $r(a,c) = r(a,d) \implies r(b,c) = r(b,d)$.

(2) $\vec{H}(a,b,c,d)$ iff $\vec{H}^r(a,b,c,d)$ for all $r \in \mathrm{Pol}_2(\mathbf{A})$.

(3) $H(a,b,c,d)$ iff $\vec{H}(a,b,c,d)$ & $\vec{H}(b,a,c,d)$.

Definition 4.2. Fix an algebra **A** and $a,b,c,d \in A$.

(1) If $r_1,r_2 \in \mathrm{Pol}_2(\mathbf{A})$, $\vec{H}_2^{r_1,r_2}(a,b,c,d)$ is the implication

$$[r_1(a,c) = r_2(a,c)\ \&\ r_1(a,d) = r_2(a,d)\ \&\ r_1(b,c) = r_2(b,c)] \implies r_1(b,d) = r_2(b,d).$$

The pair $(r_1(b,d), r_2(b,d))$ is called the *critical pair* of $\vec{H}_2^{r_1,r_2}(a,b,c,d)$ (whether $\vec{H}_2^{r_1,r_2}(a,b,c,d)$ holds or not).

(2) $\vec{H}_2(a,b,c,d)$ iff $\vec{H}_2^{r_1,r_2}(a,b,c,d)$ for all $r_1, r_2 \in \mathrm{Pol}_2(\mathbf{A})$.

(3) $H_2(a,b,c,d)$ iff $\vec{H}_2(a,b,c,d)$ & $\vec{H}_2(b,a,c,d)$ & $\vec{H}_2(a,b,d,c)$ & $\vec{H}_2(b,a,d,c)$.

Lemma 4.3. *For any algebra* $\mathbf{A}$ *and* $a,b,c,d \in A$*:*

(1) $C_2(a,b,c,d)$ *implies* $H_2(a,b,c,d)$.

(2) $H_2(a,b,c,d)$ *implies* $H(a,b,c,d)$.

The next condition is borrowed from McKenzie's relation K [35].

Definition 4.4. Let $\mathbf{A} \in \mathcal{V}$ and $a,b,c,d \in A$.

(1) If $r \in \mathrm{Pol}_2(\mathbf{A})$, $\vec{K}_p^r(a,b,c,d)$ is the following equation:

$$p(r(a,c), r(b,c), r(b,d)) = p(r(a,d), r(b,d), r(b,d)).$$

(2) $\vec{K}_p(a,b,c,d)$ iff $\vec{K}_p^r(a,b,c,d)$ for all $r \in \mathrm{Pol}_2(\mathbf{A})$.

(3) $K_p(a,b,c,d)$ iff $\vec{K}_p(a,b,c,d)$ & $\vec{K}_p(a,b,d,c)$ & $\vec{K}_p(b,a,c,d)$ & $\vec{K}_p(b,a,d,c)$.

Lemma 4.5. *Suppose* $\mathcal{V}$ *is a variety with a difference term,* $\mathbf{A} \in \mathcal{V}$*, and* $a,b,c,d \in A$*. Let* $\alpha = \mathrm{Cg}^{\mathbf{A}}(a,b)$ *and* $\beta = \mathrm{Cg}^{\mathbf{A}}(c,d)$*.*

(1) $\vec{H}(a,b,c,d)$ *implies* $\vec{K}_p(a,b,c,d)$.

(2) *Suppose* $r_1, r_2 \in \mathrm{Pol}_2(\mathbf{A})$ *and define*

$$\widehat{r}_2(x,y) = r_2(y,x) \quad \text{and} \quad s(x,y) = p(r_1(x,y), r_2(x,d), r_2(b,d)).$$

Suppose $\vec{K}_p^{\widehat{r}_2}(c,d,a,b)$ *and* $\vec{K}_p^s(b,a,c,d)$ *hold but* $\vec{H}_2^{r_1,r_2}(a,b,c,d)$ *fails with critical pair* (u,v)*. Then* $p(p(u,v,v),v,v) = v$.

(3) *If* $K_p(a,b,c,d)$ *and* $K_p(c,d,a,b)$ *but* $\neg\vec{H}_2(a,b,c,d)$*, then* $\alpha \cap \beta$ *is not abelian.*

Proof. (1) Given $r \in \mathrm{Pol}_2(\mathbf{A})$, let $r'(x,y) = p(r(a,y), r(x,y), r(b,d))$. Then

$$\begin{aligned} r'(a,c) &= p(r(a,c), r(a,c), r(b,d)) = r(b,d) \\ r'(a,d) &= p(r(a,d), r(a,d), r(b,d)) = r(b,d) \\ r'(b,c) &= p(r(a,c), r(b,c), r(b,d)) \\ r'(b,d) &= p(r(a,d), r(b,d), r(b,d)). \end{aligned}$$

As $r'(a,c) = r'(a,d)$, $\vec{H}^{r'}(a,b,c,d)$ implies $r'(b,c) = r'(b,d)$, which is $\vec{K}_p^r(a,b,c,d)$.

(2) As $\neg\vec{H}_2^{r_1,r_2}(a,b,c,d)$, we have

$$\begin{aligned} r_1(a,c) &= r_2(a,c) \\ r_1(b,c) &= r_2(b,c) \\ r_1(a,d) &= r_2(a,d) := E \\ u = r_1(b,d) &\neq r_2(b,d) = v. \end{aligned}$$

$\vec{K}_p^{\widehat{r_2}}(c,d,a,b)$ gives $p(r_2(a,c),r_2(a,d),r_2(b,d)) = p(r_2(b,c),r_2(b,d),r_2(b,d))$, which can be rewritten as

$$p(r_1(a,c),r_2(a,d),r_2(b,d)) = p(r_1(b,c),r_2(b,d),r_2(b,d)). \tag{4.1}$$

Observe that equation (4.1) can be written as $s(a,c)=s(b,c)$. $\vec{K}_p^s(b,a,c,d)$ gives

$$p(s(b,c),s(a,c),s(a,d)) = p(s(b,d),s(a,d),s(a,d)),$$

which by the last observation is equivalent to

$$s(a,d) = p(s(b,d),s(a,d),s(a,d)). \tag{4.2}$$

Calculating, we find

$$\begin{aligned} s(a,d) &= p(r_1(a,d),r_2(a,d),r_2(b,d)) = p(E,E,v) = v \\ s(b,d) &= p(r_1(b,d),r_2(b,d),r_2(b,d)) = p(u,v,v). \end{aligned}$$

Thus equation (4.2) gives $p(p(u,v,v),v,v)=v$.

(3) As the hypotheses are symmetric, we may assume that $\neg\vec{H}_2^{r_1,r_2}(a,b,c,d)$ fails for some $r_1,r_2 \in \mathrm{Pol}_2(\mathbf{A})$. Define $u=r_1(b,d)$ and $v=r_2(b,d)$, so $u\neq v$. By item (2), we have $p(p(u,v,v),v,v)=v$, which with $u\neq v$ implies $p(u,v,v)\neq u$. But $(u,v)\in\alpha\cap\beta$, so p is not a Maltsev operation on $(\alpha\cap\beta)$-blocks, so $\alpha\cap\beta$ is not abelian by Lemma 2.1(3). □

The next definition is the first of two which addresses the operations in Theorem 1.2(3). For the remainder of this section, if $\mathcal{V}$ is a variety with a difference term p, then we assume that f_i, g_i $(i\in I)$ is a finite family of ternary terms witnessing Theorem 1.2(3) for $\mathcal{V}, p$.

Definition 4.6. Let $\mathcal{V}$ be a variety with a difference term, $\mathbf{A}\in\mathcal{V}$ and $a,b,c,d\in A$.

(1) If $s,t\in\mathrm{Pol}_1(\mathbf{A})$ and $i\in I$, then $\vec{K}_{fg}^{s,t,i}(a,b,c,d)$ is the implication

$$\begin{aligned} &f_i(s(a),t(c),s(b)) = g_i(s(a),t(c),s(b)) \\ &\quad\Longrightarrow\quad f_i(s(a),t(d),s(b)) = g_i(s(a),t(d),s(b)). \end{aligned}$$

(2) $\vec{K}_{fg}(a,b,c,d)$ iff $\vec{K}_{fg}^{s,t,i}(a,b,c,d)$ for all $s,t\in\mathrm{Pol}_1(\mathbf{A})$ and all $i\in I$.

(3) $K_{fg}(a,b,c,d)$ iff $\vec{K}_{fg}(a,b,c,d)$ & $\vec{K}_{fg}(b,a,c,d)$ & $\vec{K}_{fg}(a,b,d,c)$ & $\vec{K}_{fg}(b,a,d,c)$.

The final relation to be defined generalizes K_{fg}, but is less well-behaved.

Definition 4.7. Let $\mathcal{V}$ be a variety with a difference term, $\mathbf{A}\in\mathcal{V}$ and $a,b,c,d\in A$.

(1) If $s_1,s_2,t\in\mathrm{Pol}_1(\mathbf{A})$ and $i\in I$, then $\vec{L}_{fg}^{s_1,s_2,t,i}(a,b,c,d)$ is the implication

$$\begin{aligned} &[s_1(a)=s_2(a) \quad \& \quad f_i(s_1(b),t(c),s_2(b)) = g_i(s_1(b),t(c),s_2(b))] \\ &\qquad\Longrightarrow\quad f_i(s_1(b),t(d),s_2(b)) = g_i(s_1(b),t(d),s_2(b)). \end{aligned}$$

(2) $\vec{L}_{fg}(a,b,c,d)$ iff $\vec{L}_{fg}^{s_1,s_2,t,i}(a,b,c,d)$ for all $s_1,s_2,t\in\mathrm{Pol}_1(\mathbf{A})$ and all $i\in I$.

(3) $L_{fg}(a,b,c,d)$ iff $\vec{L}_{fg}(a,b,c,d)$ & $\vec{L}_{fg}(a,b,d,c)$ & $\vec{L}_{fg}(b,a,c,d)$ & $\vec{L}_{fg}(b,a,d,c)$.

Lemma 4.8. *Let $\mathcal{V}$ be a variety with a difference term, $\mathbf{A} \in \mathcal{V}$ and $a,b,c,d \in A$.*

(1) *$\vec{H}_2(a,b,c,d)$ implies $\vec{L}_{fg}(a,b,c,d)$.*

(2) *$\vec{L}_{fg}(a,b,c,d)$ implies $\vec{K}_{fg}(a,b,c,d)$.*

Proof. (1) Given $s_1, s_2, t \in \mathrm{Pol}_1(\mathbf{A})$ and $i \in I$, define $r_1(x,y) = f_i(s_1(x), t(y), s_2(x))$ and $r_2(x,y) = g_i(s_1(x), t(y), s_2(x))$. Then $\vec{H}_2^{r_1,r_2}(a,b,c,d)$ implies $\vec{L}_{fg}^{s_1,s_2,t,i}(a,b,c,d)$.

(2) Given $s, t \in \mathrm{Pol}_1(\mathbf{A})$, define $s_2 = s$ and $s_1(x) = s(a)$. Then $\vec{L}_{fg}^{s_1,s_2,t,i}(a,b,c,d) \equiv \vec{K}_{fg}^{s,t,i}(a,b,c,d)$. □

Lemma 4.9. *Let $\mathcal{V}$ be a variety with a difference term, $\mathbf{A} \in \mathcal{V}$ and $a,b,c,d \in A$.*

(1) *$H(c,d,a,b)$ and $L_{fg}(a,b,c,d)$ imply $C(a,b,c,d)$.*

(2) *$K_p(a,b,c,d)$ and $K_{fg}(a,b,c,d)$ imply $H(c,d,a,b)$.*

Proof. (1) Assume $H(c,d,a,b)$ and $L_{fg}(a,b,c,d)$ hold but $C(a,b,c,d)$ fails. Thus there exist $(c_1,d_1),\ldots,(c_n,d_n) \in \mathrm{Cg}^{\mathbf{A}}(c,d)$ and $r \in \mathrm{Pol}_{1+n}(\mathbf{A})$ such that, without loss of generality, $r(a,\mathbf{c}) = r(a,\mathbf{d})$ while $c' := r(b,\mathbf{c}) \neq r(b,\mathbf{d}) =: d'$. Define

$$r'(\mathbf{x},y) := p(r(y,\mathbf{x}), r(y,\mathbf{c}), r(b,\mathbf{c})).$$

Starting from $r'(\mathbf{c},a) = r'(\mathbf{c},b)$ and using $H(c,d,a,b)$ and Maltsev chains of polynomial images of $\{c,d\}$ connecting each c_i to d_i, we can deduce a succession of equations, the last of which is $r'(\mathbf{d},a) = r'(\mathbf{d},b)$, i.e.,

$$c' = p(d',c',c').$$

Since $c' \neq d'$, Theorem 1.2(3c) gives $i \in I$ such that

$$f_i(d',d',c') = g_i(d',d',c') \iff f_i(d',c',c') = g_i(d',c',c') \quad \text{fails.}$$

Assume with no loss of generality that $f_i(d',d',c') = g_i(d',d',c')$ while $f_i(d',c',c') \neq g_i(d',c',c')$. For $0 \leq j \leq n$ define

$$e_j \;=\; r(b,d_1,\ldots,d_j,c_{j+1},\ldots,c_n).$$

As $f_i(d',e_0,c') \neq g_i(d',e_0,c')$ but $f_i(d',e_n,c') = g_i(d',e_n,c')$, there exists $1 \leq j \leq n$ such that $f_i(d',e_{j-1},c') \neq g_i(d',e_{j-1},c')$ while $f_i(d',e_j,c') = g_i(d',e_j,c')$. Define

$$\begin{aligned} \sigma_1(x) &= r(x,\mathbf{d}) \\ \sigma_2(x) &= r(x,\mathbf{c}) \\ t(x) &= r(b,d_1,\ldots,d_{j-1},x,c_{j+1},\ldots,c_n). \end{aligned}$$

Observe that

$$\begin{aligned} \sigma_1(a) &= \sigma_2(a) \\ f_i(\sigma_1(b), t(c_j), \sigma_2(b)) &\neq g_i(\sigma_1(b), t(c_j), \sigma_2(b)) \\ f_i(\sigma_1(b), t(d_j), \sigma_2(b)) &= g_i(\sigma_1(b), t(d_j), \sigma_2(b)). \end{aligned}$$

As $(c_j, d_j) \in \mathrm{Cg}^{\mathbf{A}}(c, d)$, there exists a Maltsev chain $c_j = u_0, u_1, \ldots, u_m = d_j$ and unary polynomials $\lambda_1, \ldots, \lambda_m$ such that $\{\lambda_k(c), \lambda_k(d)\} = \{u_{k-1}, u_k\}$ for each k. Choose k such that

$$\begin{aligned} f_i(\sigma_1(b), t(u_{k-1}), \sigma_2(b)) &\neq g_i(\sigma_1(b), t(u_{k-1}), \sigma_2(b)) \\ f_i(\sigma_1(b), t(u_k), \sigma_2(b)) &= g_i(\sigma_1(b), t(u_k), \sigma_2(b)). \end{aligned}$$

Let $\lambda(x) = t(\lambda_k(x))$. Then

$$\begin{aligned} & f_i(\sigma_1(b), \lambda(c), \sigma_2(b)) = g_i(\sigma_1(b), \lambda(c), \sigma_2(b)) \\ \iff \quad & f_i(\sigma_1(b), \lambda(d), \sigma_2(b)) \neq g_i(\sigma_1(b), \lambda(d), \sigma_2(b)). \end{aligned}$$

This is a violation of $L_{fg}(a, b, c, d)$.

(2) Assume that $K_{fg}(a, b, c, d)$ and $K_p(a, b, c, d)$ hold but $\vec{H}(c, d, a, b)$ fails at $r_1 \in \mathrm{Pol}_2(\mathbf{A})$. Define $r(x, y) = r_1(y, x)$. Thus $r(a, c) = r(b, c)$ but $r(a, d) \neq r(b, d)$. Define

$$s(x) = r(x, d) \quad \text{and} \quad t(y) = r(a, y) \quad \text{and} \quad t'(y) = r(b, y).$$

Define

$$\begin{aligned} B &= r(a, c) = t(c) \\ B' &= r(b, c) = t'(c) \\ A &= r(a, d) = t(d) = s(a) \\ C &= r(b, d) = t'(d) = s(b). \end{aligned}$$

Applying $K_{fg}(a, b, c, d)$ at s, t yields

$$f_i(A, B, C) = g_i(A, B, C) \leftrightarrow f_i(A, A, C) = g_i(A, A, C),$$

while applying $K_{fg}(a, b, c, d)$ at s, t' yields

$$f_i(A, B', C) = g_i(A, B', C) \leftrightarrow f_i(A, C, C) = g_i(A, C, C).$$

As $B = B'$, we get

$$f_i(A, A, C) = g_i(A, A, C) \leftrightarrow f_i(A, C, C) = g_i(A, C, C) \quad \text{for all } i \in I.$$

Hence $p(A, C, C) = A$ by Theorem 1.2(3c). Now apply $K_p(a, b, c, d)$ at r to get

$$p(B, B', C) = p(A, C, C).$$

As $B = B'$, the identity $p(x, x, y) \approx y$ gives $p(A, C, C) = C$. This proves $A = C$. But that contradicts our assumptions. □

Corollary 4.10. *Let $\mathcal{V}$ be a variety with a difference term, $\mathbf{A} \in \mathcal{V}$, and $a, b, c, d \in A$. The following are equivalent:*

1. *$C(a, b, c, d)$.*
2. *$H_2(a, b, c, d)$.*
3. *$K_p(a, b, c, d)$ and $L_{fg}(a, b, c, d)$.*

Proof. (1) $\Rightarrow$ (2). By Lemma 2.3 and Lemma 4.3(1).

(2) $\Rightarrow$ (3). $H_2(a,b,c,d) \Rightarrow H(a,b,c,d) \Rightarrow K_p(a,b,c,d)$ by Lemma 4.3(2) and Lemma 4.5(1). $H_2(a,b,c,d) \Rightarrow L_{fg}(a,b,c,d)$ by Lemma 4.8(1).

(3) $\Rightarrow$ (1). $K_p(a,b,c,d)$ and $L_{fg}(a,b,c,d)$ imply $H(c,d,a,b)$ by Lemma 4.9(2). This with $L_{fg}(a,b,c,d)$ implies $C(a,b,c,d)$ by Lemma 4.9(1). □

5. Definability in varieties with a finite residual bound

In this section we study the relation $C(a,b,c,d)$ in varieties with a difference term and having a finite residual bound. For this purpose, we introduce more notation.

Definition 5.1. If $k \geq 1$ and $m \geq 0$, then $\mathrm{Pol}_k^{(m)}(\mathbf{A})$ denotes the set of $r \in \mathrm{Pol}_k(\mathbf{A})$ which can realized by a term operation of $\mathbf{A}$ using at most m parameters from A; that is, $r(\mathbf{x}) = t(\mathbf{x}, \mathbf{e})$ for some $t \in \mathrm{Clo}_{k+m}(\mathbf{A})$ and some $\mathbf{e} \in A^m$. Moreover,

(1) $K_p^{(m)}(a,b,c,d)$ indicates the restriction of $K_p(a,b,c,d)$ to $r \in \mathrm{Pol}_2^{(m)}(\mathbf{A})$.
(2) $K_{fg}^{(m)}(a,b,c,d)$ denotes the restriction of $K_{fg}(a,b,c,d)$ to $s,t \in \mathrm{Pol}_1^{(m)}(\mathbf{A})$.
(3) $L_{fg}^{(m)}(a,b,c,d)$ denotes the restriction of $L_{fg}(a,b,c,d)$ to $s_1,s_2,t \in \mathrm{Pol}_1^{(m)}(\mathbf{A})$.

Definition 5.2. Suppose $\mathbf{A}$ is an algebra, $a,b,c,d \in A$, and $m,k \geq 1$.

(1) $(a,b) \Rightarrow_{(m)} (c,d)$ iff $\{c,d\} = \{s(a),s(b)\}$ for some $s \in \mathrm{Pol}_1^{(m)}(\mathbf{A})$.
(2) $(a,b) \Rrightarrow_{(m)}^k (c,d)$ iff there exist $c = c_0, c_1, \ldots, c_k = d$ such that $(a,b) \Rightarrow_{(m)} (c_i, c_{i+1})$ for all $i < k$.

Observe that $(a,b) \Rrightarrow_{(m)}^k (c,d)$ implies $(c,d) \in \mathrm{Cg}^{\mathbf{A}}(a,b)$, and that $\Rrightarrow_{(m)}^k$ is first-order definable for each $m,k \geq 1$ in any locally finite variety.

Lemma 5.3. *Suppose $\mathcal{V}$ is a variety with a difference term and having residual bound m. Let $\mathbf{A} \in \mathcal{V}$ and $a,b,c,d \in A$, and put $\alpha = \mathrm{Cg}^{\mathbf{A}}(a,b)$.*

(1) $\vec{K}_p(a,b,c,d) \equiv \vec{K}_p^{(m)}(a,b,c,d)$.
(2) *If* $0_A \prec \alpha$, *then* $\vec{K}_{fg}(a,b,c,d) \equiv \vec{K}_{fg}^{(m)}(a,b,c,d)$.
(3) *If* $0_A \prec \alpha$, *then* $\vec{L}_{fg}(a,b,c,d)$ *iff* $\vec{L}_{fg}^{s_1,s_2,t,i}(a,b,c,d)$ *for all* $s_1,s_2 \in \mathrm{Pol}_1(\mathbf{A})$, *all* $t \in \mathrm{Pol}_1^{(m)}(\mathbf{A})$, *and all* $i \in I$.

Proof. (1) We follow the proof of Lemma 3.5 in [35]. Assume that $\vec{K}_p^{(m)}(a,b,c,d)$ holds, and let $r \in \mathrm{Pol}_2(\mathbf{A})$. We must prove

$$p(r(a,c), r(b,c), r(b,d)) = p(r(a,d), r(b,d), r(b,d)),$$

and to do that it suffices to show that

$$p(r(a,c), r(b,c), r(b,d)) \stackrel{\theta}{\equiv} p(r(a,d), r(b,d), r(b,d)) \tag{5.1}$$

for all $\theta \in \mathrm{Con}\,\mathbf{A}$ of index at most m. Fix such θ and let T be a transversal for θ; thus $T \subseteq A$, $|T| \leq m$, and T intersects each θ-class in exactly one element. Pick a term $t(x, y, \mathbf{z})$ and parameters $\mathbf{e}$ from A so that $r(x, y) = t^{\mathbf{A}}(x, y, \mathbf{e})$. Define $\mathbf{u}$ so that u_i is the unique element of $e_i/\theta \cap T$ and define $r'(x, y) = t^{\mathbf{A}}(x, y, \mathbf{u})$. Then

- $r(x, y) \overset{\theta}{\equiv} r'(x, y)$ for all $x, y \in A$.
- $r' \in \mathrm{Pol}_2^{(m)}(\mathbf{A})$.

As $\vec{K}_p^{(m)}(a, b, c, d)$ holds by assumption, we have

$$p(r'(a, c), r'(b, c), r'(b, d)) = p(r'(a, d), r'(b, d), r'(b, d)),$$

which implies (5.1).

(2) We follow the main idea of the proof of Lemma 2 in [3]. Let $\theta \in \mathrm{Con}\,\mathbf{A}$ be meet-irreducible and satisfying $\theta \cap \alpha = 0_A$. Then θ has index at most m. Assume that $\vec{K}_{fg}^{(m)}(a, b, c, d)$ holds but $\vec{K}_{fg}(a, b, c, d)$ fails at $s, t \in \mathrm{Pol}_1(\mathbf{A})$ and $i \in I$. Thus

$$\begin{aligned} f_i(s(a), t(c), s(b)) &= g_i(s(a), t(c), s(b)) \\ u := f_i(s(a), t(d), s(b)) &\neq g_i(s(a), t(d), s(b)) =: v. \end{aligned}$$

Note that $(u, v) \in \alpha$, so $(u, v) \notin \theta$. As in the proof of (1), we can find $s', t' \in \mathrm{Pol}_1^{(m)}(\mathbf{A})$ such that $s(x) \overset{\theta}{\equiv} s'(x)$ and $t(x) \overset{\theta}{\equiv} t'(x)$ for all $x \in A$. Thus

$$\begin{aligned} f_i(s'(a), t'(c), s'(b)) &\overset{\theta}{\equiv} g_i(s'(a), t'(c), s'(b)) \\ f_i(s'(a), t'(d), s'(b)) &\overset{\theta}{\not\equiv} g_i(s'(a), t'(d), s'(b)). \end{aligned}$$

In addition, Theorem 1.2(3b) implies

$$\begin{aligned} f_i(s'(a), t'(c), s'(b)) &\overset{\alpha}{\equiv} f_i(s'(a), t'(c), s'(a)) \\ &= g_i(s'(a), t'(c), s'(a)) \\ &\overset{\alpha}{\equiv} g_i(s'(a), t'(c), s'(b)). \end{aligned}$$

As $\theta \cap \alpha = 0_A$, this proves

$$f_i(s'(a), t'(c), s'(b)) = f_i(s'(a), t'(c), s'(a))$$

which contradicts $\vec{K}_{fg}^{(m)}(a, b, c, d)$.

(3) The proof is similar to the proof of item (2). □

Theorem 5.4. *Suppose $\mathcal{V}$ is a variety with a difference term and having a finite residual bound m. $C(x, y, z, w)$ is equivalent in $\mathcal{V}$ to the following condition:*

($*$) For all x_1, y_1, z_1, w_1, if $(x, y) \Rrightarrow^2_{(m+3)} (x_1, y_1)$ and $(z, w) \Rrightarrow^2_{(m+3)} (z_1, w_1)$, then $K_p^{(m)}(x_1, y_1, z_1, w_1)$ & $K_p^{(m)}(z_1, w_1, x_1, y_1)$ & $K_{fg}^{(m)}(z_1, w_1, x_1, y_1)$.

Proof. Suppose $\mathbf{A} \in \mathcal{V}$ and $a, b, c, d \in A$. Clearly $C(a, b, c, d)$ implies the above condition. For the remainder of the proof, assume ($*$) holds and yet $\neg C(a, b, c, d)$; we will find a contradiction. We may assume that $\mathbf{A}$ is finite.

Let $\alpha = \mathrm{Cg}^{\mathbf{A}}(a,b)$ and $\beta = \mathrm{Cg}^{\mathbf{A}}(c,d)$. By condition $(*)$ and Lemma 5.3(1), we have $K_p(a,b,c,d)$ and $K_p(c,d,a,b)$. Hence by Corollary 4.10 and Lemma 4.5(3), $\alpha\cap\beta$ is not abelian. Choose $\gamma \in \mathrm{Con}\,\mathbf{A}$ with $0_A \prec \gamma \leq [\alpha\cap\beta, \alpha\cap\beta]$. Let θ be a maximal congruence satisfying $\gamma \not\leq \theta$, and let μ be the unique upper cover of θ. Also let $\nu = (0_A : \gamma)$, so $\nu = (\theta : \mu)$ by Lemmas 2.5 and 2.6. Observe that $\gamma \cap \theta = 0_A$ implies $\theta \leq \nu$, which with $C(\nu, \mu; \theta)$ implies $[\nu/\theta, \mu/\theta] = 0_{A/\theta}$.

Because $\mathcal{V}$ has a difference term and is residually small, it satisfies C1 by Propositions 3.1 and 3.2(1). Applied to $\mathbf{A}/\theta$ and the previous commutator fact, this gives $\mu/\theta \not\leq [\nu/\theta, \nu/\theta]$, so $[\nu/\theta, \nu/\theta] = 0_{A/\theta}$, which implies $[\nu, \nu] \leq \theta$.

Observe that if α centralized γ, then we would have $\alpha \leq \nu$ and hence

$$\gamma \leq [\alpha, \alpha] \leq [\nu, \nu] \leq \theta,$$

which is false. This proves that α does not centralize ν. Similarly, β does not centralize ν.

Pick $(u, v) \in \gamma \setminus 0_A$. By what we have just proved and symmetry of the centralizer relation, we have $\neg C(u, v, a, b)$. Thus by Corollary 4.10, at least one of $K_p(u, v, a, b)$ or $L_{fg}(u, v, a, b)$ must fail. Suppose first that $K_p(u, v, a, b)$ fails; then $K_p^{(m)}(u, v, a, b)$ fails by Lemma 5.3(1). Pick $r \in \mathrm{Pol}_2^{(m)}(\mathbf{A})$ witnessing the failure; thus

$$a_1 := p(r(u, a), r(v, a), r(v, b)) \neq p(r(u, a), (v, b), (v, b)) =: b_1.$$

Note that $(a_1, b_1) \in \gamma \setminus 0_A$ and $(a, b) \Rightarrow_{(m+3)} (a_1, b_1)$ witnessed by the polynomial $s(x) = p(u', r(v, x), v')$ where $u' = r(u, a)$ and $v' = r(v, b)$. Suppose instead that $L_{fg}(u, v, a, b)$ fails; pick $s_1, s_2, t \in \mathrm{Pol}_1(\mathbf{A})$ and $i \in I$ such that, without loss of generality,

$$\begin{aligned} s_1(u) &= s_2(u) \\ f_i(s_1(v), t(a), s_2(v)) &= g_i(s_1(v), t(a), s_2(v)) \\ a_1 := f_i(s_1(v), t(b), s_2(v)) &\neq g_i(s_1(v), t(b), s_2(v)) =: b_1. \end{aligned}$$

By Lemma 5.3(3), we may assume that $t \in \mathrm{Pol}_1^{(m)}(\mathbf{A})$. Note that $(a_1, b_1) \in \gamma \setminus 0_A$ and $(a, b) \Rrightarrow^2_{(m+2)} (a_1, b_1)$ witnessed by the polynomials $f_i(u', t(x), v')$ and $g_i(u', t(x), v')$ where $u' = s_1(v)$ and $v' = s_2(v)$.

Thus in either case, we have established the existence of $(a_1, b_1) \in \gamma \setminus 0_A$ with $(a, b) \Rrightarrow^2_{(m+3)} (a_1, b_1)$. A similar argument proves the existence of $(c_1, d_1) \in \gamma \setminus 0_A$ with $(c, d) \Rrightarrow^2_{(m+3)} (c_1, d_1)$. Choose and fix such a_1, b_1, c_1, d_1.

By condition $(*)$, we have both $K_p^{(m)}(a_1, b_1, c, d)$ and $K_p^{(m)}(c, d, a_1, b_1)$. Hence by Lemma 5.3(1) we have $K_p(a_1, b_1, c, d)$ and $K_p(c, d, a_1, b_1)$. As $C(c, d, a_1, b_1)$ fails (because β does not centralize γ), Lemma 4.5(3) and Corollary 4.10 imply that $\beta \cap \gamma = \gamma$ is nonabelian. Let U be a $(0, \gamma)$-minimal set with trace $\{0, 1\}$. Let $r \in \mathrm{Pol}_2(\mathbf{A})$ be a polynomial whose range is U and whose restriction to $\{0, 1\}$ is the meet semilattice operation. As $(a_1, b_1), (c_1, d_1) \in \gamma \setminus 0_A$, there exist $s, t \in \mathrm{Pol}_1(\mathbf{A})$

such that $\{s(a_1), s(b_1)\} = \{t(c_1), t(d_1)\} = \{0, 1\}$. Define $r'(x, y) = r(s(x), t(y))$. Then three of $r'(a_1, c_1), r'(a_1, d_1), r'(b_1, c_1), r'(b_1, d_1)$ equal 0 while the fourth equals 1. Hence $H(a_1, b_1, c_1, d_1)$ fails at r'. By Lemma 4.9(2), one of $K_p(c_1, d_1, a_1, b_1)$ or $K_{fg}(c_1, d_1, a_1, b_1)$ must fail. Lemma 5.3 then implies that one of $K_p^{(m)}(c_1, d_1, a_1, b_1)$ or $K_{fg}^{(m)}(c_1, d_1, a_1, b_1)$ must fail, contradicting condition $(*)$. □

6. The Kiss 4-ary term

Throughout this section, $\mathcal{V}$ is a variety having a difference term p. Define the associated Kiss 4-ary term by Lipparini's Formula $q(x, y, z, w) := p(p(x, z, z), p(y, w, z), z)$.[4] Following Kiss we call (a, b, c, d) an *α, β-rectangle* if $(a, b), (c, d) \in \alpha$ and $(a, c), (b, d) \in \beta$, and we let $R(\alpha, \beta)$ be the set of these. $R(\alpha, \beta)$ is a subuniverse of $\mathbf{A}^4$.

Lemma 6.1 ([27]). *If $\mathbf{A} \in \mathcal{V}$ and $\alpha, \beta \in \mathrm{Con}(\mathbf{A})$, then*

(1) $\mathcal{V} \models q(x, y, x, y) \approx x$,
(2) $\mathcal{V} \models q(x, x, y, y) \approx y$, *and*
(3) $q(a, b, c, d) \equiv_{[\beta,\alpha]} q(a, b, c', d)$ *if* $(a, b, c, d), (a, b, c', d) \in R(\alpha, \beta)$. □

Lemma 6.2. *If $\mathbf{A} \in \mathcal{V}$ and $\alpha, \beta \in \mathrm{Con}(\mathbf{A})$, then $[\alpha, \beta] = 0$ iff*

(i) $q\colon R(\alpha, \beta) \to \mathbf{A}$ *is a homomorphism, and*
(ii) q *is independent of its third variable on* $R(\alpha, \beta)$.

Proof. In the case where $\mathcal{V}$ is congruence modular this lemma is Theorem 3.8 (iii) of [20]. The proof below follows the argument from page 472 of [20].

Let $\Delta_{\alpha,\beta}$ be the congruence on $\mathbf{A} \times_\alpha \mathbf{A}$ generated by the β-diagonal. Kiss argues that if $(a_i, b_i, c_i, d_i) \in R(\alpha, \beta)$, then for any term s we have

(6.1)
$$(q(s(\mathbf{a}), s(\mathbf{b}), s(\mathbf{c}), s(\mathbf{d})), s(\mathbf{d})) \equiv_{\Delta_{\alpha,\beta}} (s(\mathbf{a}), s(\mathbf{b})) \equiv_{\Delta_{\alpha,\beta}} (s(\overline{q(a_i, b_i, c_i, d_i)}), s(\mathbf{d})).$$

The argument he gives works under our hypotheses. Kiss then uses a property of the modular commutator to derive from (6.1) that

$$q(s(\mathbf{a}), s(\mathbf{b}), s(\mathbf{c}), s(\mathbf{d})) \equiv_{[\alpha,\beta]} s(\overline{q(a_i, b_i, c_i, d_i)}).$$

A justification that this step works under our hypotheses is required.

If $\mathcal{V}$ has a difference term, then it satisfies a nontrivial idempotent Maltsev condition. Lemma 4.4 of [19] shows (with a slight change of notation) that if $[\alpha, \beta] = 0$, then on $\mathbf{A} \times_\alpha \mathbf{A}$ it is the case that

$$\beta_1 \wedge 0_2 \wedge \Delta_\alpha = 0.$$

Here β_1 is the congruence on $\mathbf{A} \times_\alpha \mathbf{A}$ that relates pairs whose first coordinates are β-related, 0_2 is the congruence on $\mathbf{A} \times_\alpha \mathbf{A}$ that relates pairs whose second coordinates

[4] Lipparini's difference term [27] has its variables in the reverse order of Kiss's difference term. Kiss's convention agrees with ours, so this formula looks different from the one in Lipparini's paper.

are equal, and Δ_α is the largest congruence on $\mathbf{A} \times_\alpha \mathbf{A}$ which relates no diagonal pair to any off-diagonal pair.

The two sides of (6.1) are equal in the second coordinate, hence are 0_2-related. Since $\Delta_{\alpha,\beta} \subseteq \Delta_\alpha$, as a consequence of $[\beta, \alpha] = [\alpha, \beta] = 0$, we get that the two sides of (6.1) are Δ_α-related. Since

$$q(s(\mathbf{a}), s(\mathbf{b}), s(\mathbf{c}), s(\mathbf{d})) \equiv_\beta q(s(\mathbf{a}), s(\mathbf{b}), s(\mathbf{a}), s(\mathbf{b})) = s(\mathbf{a}),$$

and similarly $s(\overline{q(a_i, b_i, c_i, d_i)}) \equiv_\beta s(\overline{q(a_i, b_i, a_i, b_i)}) = s(\mathbf{a})$, we get that the two sides of (6.1) are β_1-related. Altogether we get the desired conclusion, that

$$q(s(\mathbf{a}), s(\mathbf{b}), s(\mathbf{c}), s(\mathbf{d})) = s(\overline{q(a_i, b_i, c_i, d_i)})$$

when $[\alpha, \beta] = 0$. This is the property that $q\colon R(\alpha, \beta) \to \mathbf{A}$ is a homomorphism, hence item (i) holds if $[\alpha, \beta] = [\beta, \alpha] = 0$. We get that item (ii) also holds from Lemma 6.1 (3).

Now we prove that (i) and (ii) force $[\beta, \alpha] = 0$. Define

$$\Delta = \{((a, b), (q(a, b, c, d), d)) \mid (a, b, c, d) \in R(\alpha, \beta)\}.$$

Kiss shows that Δ is a congruence on $\mathbf{A} \times_\alpha \mathbf{A}$ that contains $\Delta_{\alpha,\beta}$. If the first pair in the pair of pairs, $((a, b), (q(a, b, c, d), d)) \in \Delta$, lies on the diagonal (a complicated way of writing "if $a = b$"), then

$$q(a, b, c, d) = q(a, a, c, d) = q(a, a, d, d) = d, \tag{6.2}$$

and the second pair in the pair of pairs also lies on the diagonal. (In the middle equality of (6.2) we are using that q is independent of its third variable on α, β-rectangles.) Since $\Delta_{\alpha,\beta} \subseteq \Delta$, and Δ relates no diagonal pair of $\mathbf{A} \times_\alpha \mathbf{A}$ to an off-diagonal pair, we derive that $[\beta, \alpha] = 0$ holds. □

7. The finite basis theorem

In this section we prove the main result of our paper:

Theorem 7.1. *If $\mathcal{V}$ is a variety in a finite language, $\mathcal{V}$ has a difference term, and $\mathcal{V}$ has a finite residual bound, then $\mathcal{V}$ is finitely based.*

Our strategy is to mimic McKenzie's argument [35, Section 4] for the congruence modular case, to the extent that that is possible. Parenthetical references are to the corresponding results from [35]. Some technical issues in McKenzie's argument become easier here because of our use of the Kiss term. We are forced to give an entirely new proof of the final step in establishing C1 (i.e., property (4) of Proposition 3.3).

Let $\mathcal{V}$ be a variety in a finite language $\mathcal{L}$ with a difference term p, and assume that $\mathcal{V}$ has residual bound $r < \omega$. By Theorem 1.2, there exists a finitely based variety $\mathcal{V}_0$ containing $\mathcal{V}$ and having the same difference term. For each $j \geq 3$ let $\mathcal{V}^{(j)}$ be the subvariety of $\mathcal{V}_0$ axiomatized by the j-variable identities of $\mathcal{V}$. Let $X = (\mathsf{x}_1, \mathsf{x}_2, \ldots)$ be a fixed infinite sequence of variables. Define the *height* of a term in some standard way, so that for each $n, h \geq 0$, the set $\mathrm{Trm}_n^{(h)}(\mathcal{L})$ of $\mathcal{L}$-terms over $\{\mathsf{x}_1, \ldots, \mathsf{x}_n\}$ of height at most h is a finite set closed under subterms. Let $h_{\mathcal{V}} : \omega \to \omega$ be a function so that for all $n \geq 0$, every $\mathcal{L}$-term over $\{\mathsf{x}_1, \ldots, \mathsf{x}_n\}$ is equivalent modulo $\mathcal{V}$ to a term in $\mathrm{Trm}_n^{(h_{\mathcal{V}}(n))}(\mathcal{L})$. For simplicity, we denote $\mathrm{Trm}_n^{(h_{\mathcal{V}}(n))}(\mathcal{L})$ by $\mathrm{Trm}_n(\mathcal{V})$. For each $n > 0$ let σ_n be a sentence asserting $\forall \mathbf{x}[f(s_1(\mathbf{x}), \ldots, s_k(\mathbf{x})) = t(\mathbf{x})]$ for all k-ary fundamental operation symbols of $\mathcal{L}$ and all $s_1, \ldots, s_k, t \in \mathrm{Trm}_n(\mathcal{V})$ such that $\mathcal{V} \models f(s_1, \ldots, s_k) \approx t$. Thus σ_j is a finite axiomatization of $\mathcal{V}^{(j)}$ relative to $\mathcal{V}_0$.

Lemma 7.2 (Lemma 4.1)**.** *There exists a first-order formula $\Omega(x, y, z, w)$ such that:*

(1) $\mathcal{V} \models C(x, y, z, w) \leftrightarrow \Omega(x, y, z, w)$.
(2) $\mathcal{V}_0 \models C(x, y, z, w) \to \Omega(x, y, z, w)$.
(3) *There exists $m > 0$ such that for all sufficiently large j,*

$$\mathcal{V}^{(j)} \models \Omega(x, y, z, w) \leftrightarrow [K_p^{(m)}(x, y, z, w) \;\&\; L_{fg}^{(m)}(x, y, z, w)].$$

(4) *There exists an existential first-order formula $W(u, v, x, y, z, w)$ satisfying*
  (a) $\mathcal{V}_0 \models W(u, v, x, y, z, w) \to$ "$(u, v) \in [\mathrm{Cg}(x, y), \mathrm{Cg}(z, w)]$".
  (b) *For all sufficiently large j,*

$$\mathcal{V}^{(j)} \models \Omega(x, y, z, w) \leftrightarrow \forall u, v[W(u, v, x, y, z, w) \to u = v].$$

Proof. Start with the condition $(*)$ expressed in Theorem 5.4 (with m replaced by r). Modulo σ_{r+4}, $(x, y) \Rrightarrow^3_{(r+3)} (u, v)$ is equivalent to its restriction to unary polynomials defined from terms in $\mathrm{Trm}_{r+4}(\mathcal{V})$. Similarly, modulo σ_{r+2}, $K_p^{(r)}(x, y, z, w)$ is equivalent to its restriction to binary polynomials defined from terms in $\mathrm{Trm}_{r+2}(\mathcal{V})$, and modulo σ_{r+1}, $K_{fg}^{(r)}(x, y, z, w)$ is equivalent to its restriction to unary polynomials defined from terms in $\mathrm{Trm}_{r+1}(\mathcal{V})$. Hence in models of σ_{r+1} & σ_{r+2} & σ_{r+4} (in particular, in $\mathcal{V}$), the condition $(*)$ can be expressed by a first-order formula $\Omega(x, y, z, w)$. This proves (1). As $\Omega(x, y, z, w)$ is a special case of $(*)$, which in turn is implied by $C(x, y, z, w)$ in $\mathcal{V}_0$, we get (2).

For $m, h > 0$ let $K_p^{(m,h)}(x, y, z, w)$ denote the restriction of $K_p^{(m)}(x, y, z, w)$ to binary polynomials definable from terms in $\mathrm{Trm}_{m+2}^{(h)}(\mathcal{L})$, and let $L_{fg}^{(m,h)}(x, y, z, w)$ be the restriction of $L_{fg}^{(m)}(x, y, z, w)$ to unary polynomials definable from terms in $\mathrm{Trm}_{m+1}^{(h)}(\mathcal{L})$. Because

$$(\dagger) \qquad \mathcal{V}_0 \models C(x, y, z, w) \leftrightarrow \bigwedge_{m,h>0} (K_p^{(m,h)}(x, y, z, w) \;\&\; L_{fg}^{(m,h)}(x, y, z, w))$$

by Corollary 4.10, and because $K_p^{(m,h)}(x, y, z, w)$ and $L_{fg}^{(m,h)}(x, y, z, w)$ are expressible by a first-order formulas for each fixed $m, h > 0$, the compactness theorem with (2) and $(\dagger)$ imply the existence of $m, h > 0$ such that $\mathcal{V}_0 \models (K_p^{(m,h)} \;\&\; L_{fg}^{(m,h)}) \to \Omega$. Thus

$\mathcal{V} \models \Omega \leftrightarrow (K_p^{(m,h)} \,\&\, L_{fg}^{(m,h)})$, so again by the compactness theorem,

$$(\ddagger) \qquad \mathcal{V}^{(j)} \models \Omega \leftrightarrow (K_p^{(m,h)} \,\&\, L_{fg}^{(m,h)}) \quad \text{for all sufficiently large } j.$$

We may assume $h \geq h_{\mathcal{V}}(m+2)$. By the compactness theorem, for all sufficiently large j, $\mathcal{V}^{(j)}$ models σ_{m+1} & σ_{m+2} and hence satisfies $K_p^{(m,h)} \equiv K_p^{(m)}$ and $L_{fg}^{(m,h)} \equiv L_{fg}^{(m)}$. This and (‡) prove (3).

Recall that $K_p^{(m,h)}(x,y,z,w)$ is a conjunction of equations, while $L_{fg}^{(m,h)}(x,y,z,w)$ is a conjunction of quasi-equations. Given $\mathbf{A} \in \mathcal{V}_0$ and $a,b,c,d,u,v \in A$, call (u,v) an (m,h)-*critical pair for* (a,b,c,d) if there exists an equation in $K_p^{(m,h)}(a,b,c,d)$ whose left and right sides are u,v respectively, or there exists a quasi-equation in $L_{fg}^{(m,h)}(a,b,c,d)$ whose conclusion is the equation with left and right sides u,v respectively. We can take $W(u,v,x,y,z,w)$ to be a first-order sentence which asserts that (u,v) is an (m,h)-critical pair for (x,y,z,w). (4a) is then obvious, and (4b) follows from (‡). □

Definition 7.3. Given $\mathbf{A} \in \mathcal{V}_0$ and $a,b \in A$, let

$$\begin{aligned} \Omega(a,b) &:= \{(x,y) : \Omega(x,y,a,b)\} \\ \Omega_{\text{op}}(a,b) &:= \{(z,w) : \Omega(x,y,z,w) \text{ for all } (x,y) \in \Omega(a,b)\} \end{aligned}$$

Lemma 7.4 (Lemma 4.4)**.** *For all sufficiently large j, all $\mathbf{A} \in \mathcal{V}^{(j)}$, and all $a,b \in A$, $\Omega(a,b)$ and $\Omega_{\text{op}}(a,b)$ are congruences.*

Proof. In $\mathcal{V}$, $\Omega(a,b) = \text{ann}(a,b)$ and $\Omega_{\text{op}}(a,b) = (0_A : \text{ann}(a,b))$. Hence the claim is true in $\mathcal{V}$, and as it can be expressed by a first-order sentence, is true in $\mathcal{V}^{(j)}$ for all sufficiently large j by the compactness theorem. □

Lemma 7.5. *For all sufficiently large j, $\mathcal{V}^{(j)} \models C(x,y,z,w) \leftrightarrow \Omega(x,y,z,w)$.*

Proof. By Lemma 7.2, $C(x,y,u,v) \to \Omega(x,y,z,w)$ holds in $\mathcal{V}_0$.

By Lemma 7.4, the relations $\Omega(a,b)$ and $\Omega_{\text{op}}(a,b)$ are congruences for any (a,b) in $\mathcal{V}^{(j)}$ for j sufficiently large, and their definitions yield that $\Omega(x,y,z,w)$ holds for any $(x,y) \in \Omega(a,b)$ and any $(z,w) \in \Omega_{\text{op}}(a,b)$. We can write a first-order sentence that asserts (in an algebra $\mathbf{A}$) that for all $a,b \in A$, (i) $q\colon R(\Omega(a,b),\Omega_{\text{op}}(a,b)) \to \mathbf{A}$ is a homomorphism and (ii) q is independent of its third variable on $R(\Omega(a,b),\Omega_{\text{op}}(a,b))$. This sentence is true in $\mathcal{V}$ by Lemma 7.2(1) and Lemma 6.2, so is true in $\mathcal{V}^{(j)}$ for sufficiently large j. Hence by Lemma 6.2, $[\Omega(z,w),\Omega_{\text{op}}(z,w)] = 0$ holds in $\mathcal{V}^{(j)}$ for sufficiently large j. But then in $\mathcal{V}^{(j)}$, we must have $\Omega(x,y,z,w) \to C(x,y,z,w)$, because if $\Omega(x,y,z,w)$ holds, then $(x,y) \in \Omega(z,w)$, while $(z,w) \in \Omega_{\text{op}}(z,w)$ always holds. □

Definition 7.6. Let $\mu(x,y)$ be the formula $\Omega(x,y,x,y)$.

Corollary 7.7 (Lemma 4.5)**.** *For all sufficiently large j, $\mathbf{A} \in \mathcal{V}^{(j)}$, and $a,b,c \in A$,*

(1) $\mathbf{A} \models \mu(a,b)$ *iff* $\mathrm{Cg}^{\mathbf{A}}(a,b)$ *is abelian.*
(2) $\Omega(a,b) = \mathrm{ann}(a,b)$.
(3) $\Omega_{\mathrm{op}}(a,b) = (0_A : \mathrm{ann}(a,b))$.

Proof. Follows easily from Lemma 7.5. □

Lemma 7.8 (Lemma 4.9)**.** *For all sufficiently large j, if $\mathbf{A} \in \mathcal{V}^{(j)}$ and $a, b, e_0, \dots, e_r \in A$ with $a \neq b$, then there exists (c,d) satisfying:*

(1) $(c,d) \in \mathrm{Cg}^{\mathbf{A}}(a,b) \setminus 0_A$.
(2) $C(c,d,e_i,e_j)$ *for some* $0 \leq i < j \leq r$.

Proof sketch. The argument is a little different than in the congruence modular case as we haven't established [35, Lemma 4.8].

Assume j is large enough to satisfy the claims in Lemmas 7.2, 7.4, 7.5 and Corollary 7.7. Let $(u_0, v_0) = (a,b)$. If $C(u_0, v_0, e_0, e_1)$, then we're done. Otherwise, we have $\neg\Omega(u_0, v_0, e_0, e_1)$, and by Lemma 7.2(4) this is witnessed by a critical pair (u_1, v_1) satisfying $u_1 \neq v_1$ and $W(u_1, v_1, u_0, v_0, e_0, e_1)$. This implies that $(u_1, v_1) \in [\mathrm{Cg}(u_0, v_0), \mathrm{Cg}(e_0, e_1)]$ and hence that $(u_1, v_1) \in \mathrm{Cg}(a,b) \cap \mathrm{Cg}(e_0, e_1)$. Next, check whether $C(u_1, v_1, e_0, e_2)$; again if true we're done, while if false then the failure gives a critical pair (u_2, v_2) with $u_2 \neq v_2$ and $W(u_2, v_2, u_1, v_1, e_0, e_2)$. This again implies $(u_2, v_2) \in \mathrm{Cg}(u_1, v_1) \cap \mathrm{Cg}(e_0, e_2)$. We can proceed in this way through all $M := \binom{r+1}{2}$ pairs (e_i, e_j). As r is fixed, if we never find what we want, we end up with a system of short proofs of $(u_{t+1}, v_{t+1}) \in \mathrm{Cg}(u_t, v_t)$ for $1 \leq t \leq M$, so that $u_t \neq v_t$ for all t, and for all $0 \leq i < j \leq r$ there exists t and a short proof of $(u_t, v_t) \in \mathrm{Cg}(e_i, e_j)$. This is a first-order definable configuration. Any algebra in which it occurs has a subdirectly irreducible quotient of cardinality greater than r. This cannot occur in $\mathcal{V}$, so by the compactness theorem, it cannot occur in $\mathcal{V}^{(j)}$ for sufficiently large j. □

Lemma 7.9 (Lemma 4.10)**.** *For all sufficiently large j, if $\mathbf{A} \in \mathcal{V}^{(j)}$ is finitely generated and $a, b \in A$ with $a \neq b$, then there exists $(c,d) \in \mathrm{Cg}^{\mathbf{A}}(a,b) \setminus 0_A$ such that $|A/\mathrm{ann}(c,d)| \leq r$.*

Proof. Identical to the proof of [35, Lemma 4.10]. □

Corollary 7.10 (Lemma 4.13; cf. Lemma 4.19)**.** *For all sufficiently large j:*

(1) *If $\mathbf{A} \in \mathcal{V}^{(j)}$ and $0_A \prec \alpha \in \mathrm{Con}\,\mathbf{A}$, then $|A/(0_A : \alpha)| \leq r$.*
(2) *There exist first-order formulas $\mathrm{AbAt}(x,y)$ and $\theta(u,v,x,y)$, not depending on j, such that for all $\mathbf{A} \in \mathcal{V}^{(j)}$ and all $c, d \in A$, letting $\alpha = \mathrm{Cg}^{\mathbf{A}}(c,d)$:*
   (a) $\mathbf{A} \models \mathrm{AbAt}(c,d)$ *iff α is an abelian atom in* $\mathrm{Con}\,\mathbf{A}$.
   (b) *if α is an abelian atom, then $\alpha = \{(a,b) \in A^2 : \mathbf{A} \models \theta(a,b,c,d)\}$.*

Proof. (1) Suppose $0_A \prec \alpha = \mathrm{Cg}^{\mathbf{A}}(a,b)$. Let $\gamma = \mathrm{ann}(a,b)$ and suppose $|A/\gamma| > r$. Pick $e_0, \dots, e_r \in A$ so that no two are related by γ. By Lemma 7.8 there exists

$(c,d) \in \mathrm{Cg}^{\mathbf{A}}(a,b)$ with $c \neq d$ and $C(c,d,e_i,e_j)$ for some $i<j$. But then $\mathrm{Cg}^{\mathbf{A}}(c,d) = \alpha$ and $(e_i,e_j) \in \mathrm{ann}(c,d) = \gamma$, a contradiction. Thus $|A/\gamma| \leq r$, which proves (1).

(2) Let $\theta(u,v,x,y)$ be the following formula:

$$\theta(u,v,x,y) \quad : \bigvee_{t \in \mathrm{Trm}_{r+1}(\mathcal{V})} \exists e_1 \cdots e_r [p(t(x,\mathbf{e}), t(y,\mathbf{e}), u) = v].$$

We can assume $\mathcal{V}^{(j)} \models \sigma_{r+1}$; hence if $\mathbf{A} \in \mathcal{V}^{(j)}$ and $c,d \in A$, then the set $\{(a,b) : \mathbf{A} \models \theta(a,b,c,d)\}$ coincides with $\Gamma_r(c,d)$ from Definition 2.7. Thus (2b) follows from Lemma 2.8(1). Now let $\mathrm{AbAt}(x,y)$ be a formula expressing the following:

$$x \neq y \ \& \ \mu(x,y) \ \& \ \text{"}\{(a,b) : \theta(a,b,x,y)\} \text{ is a congruence containing } (x,y)\text{"}$$
$$\& \ \forall u,v[(\theta(u,v,x,y) \ \& \ u \neq v) \ \rightarrow \ \theta(x,y,u,v)].$$

That $\mathrm{AbAt}(x,y)$ has the claimed property follows from (1), Corollary 7.7, Lemma 2.8, and the fact that $\Gamma_r(c,d) \subseteq \mathrm{Cg}^{\mathbf{B}}(c,d)$ for any $\mathbf{B} \in \mathcal{V}_0$ and $c,d \in B$. □

Lemma 7.11 (Lemma 4.15)**.** *For all sufficiently large j, $\mathcal{V}^{(j)}$ is locally finite.*

Proof. If not, we can find a finitely generated infinite algebra $\mathbf{A} \in \mathcal{V}^{(j)}$ such that every nonzero congruence of $\mathbf{A}$ has finite index. Using Lemma 7.9, there exists a nonzero congruence β such that $(0_A : \beta)$ has finite index. Then $\beta \cap (0_A : \beta)$ also has finite index so is nonzero. As $\beta \cap (0_A : \beta)$ is abelian, this proves the existence of a nonzero *abelian* congruence α. Using Lemma 7.9 again, we get $(a,b) \in \alpha \setminus 0_A$ with $\mathrm{ann}(a,b)$ having index at most r. But Lemma 2.8(2) then says that each $\mathrm{Cg}^{\mathbf{A}}(a,b)$-block is finite, which is impossible. □

Next, we work towards establishing that $\mathcal{V}^{(j)}$ satisfies the commutator identity C1. (Recall that $\mathcal{V}$ itself satisfies C1 by Propositions 3.1 and 3.2(1).) Our strategy will be to verify each of the conditions in Proposition 3.3.

Definition 7.12. Given $\mathbf{A} \in \mathcal{V}_0$, let

$$\mu^{\mathbf{A}} := \{(x,y) \in A^2 : \mu(x,y)\}$$

where $\mu(x,y)$ is the formula from Definition 7.6.

Lemma 7.13 (Lemma 4.17(1))**.** *For all sufficiently large j and all $\mathbf{A} \in \mathcal{V}^{(j)}$, $\mu^{\mathbf{A}}$ is the largest abelian congruence of $\mathbf{A}$.*

Proof. As $\mu^{\mathbf{A}}$ contains every abelian congruence by Corollary 7.7(1), it suffices to prove that $\mu^{\mathbf{A}}$ is itself abelian. This property is first-order by Lemma 6.2, so it suffices to prove this latter claim for $\mathbf{A} \in \mathcal{V}$. Fix $\mathbf{A} \in \mathcal{V}$ and suppose α, β are abelian congruences. Let $\gamma = \alpha \vee \beta$ and $\delta = [\alpha,\beta]$. We have $\delta \leq [\gamma,\gamma]$ by monotonicity, so $[\gamma,\delta] = \delta$ by C1. On the other hand, $[\alpha,\delta] \leq [\alpha,\alpha] = 0_A$ so $\alpha \leq (0_A : \delta)$, and similarly $\beta \leq (0_A : \delta)$. Thus $\gamma \leq (0_A : \delta)$, which means $[\gamma,\delta] = 0_A$. This proves $[\alpha,\beta] = 0_A$

whenever α, β are abelian congruences. Now use [16, Lemma 2.8] to deduce that the join of all abelian congruences of $\mathbf{A}$ is itself abelian; call it $\alpha_{\max}$. Hence

$$\mu^{\mathbf{A}} \;=\; \bigvee_{(a,b)\in\mu^{\mathbf{A}}} \mathrm{Cg}^{\mathbf{A}}(a,b) \;\subseteq\; \bigvee_{\alpha \text{ abelian}} \alpha \;=\; \alpha_{\max}$$

which proves $\mu^{\mathbf{A}}$ is abelian. □

Recall that m is fixed satisfying Lemma 7.2(3). The next definition simply gives notation for the set of critical pairs for $K_p^{(6m+1)}(a,b,a,b)$.

Definition 7.14. For $\mathbf{A} \in \mathcal{V}_0$ and $a, b \in A$, define

$$\begin{aligned} G_m(a,b) \;=\; \{\, &(p(r(x,z), r(y,z), r(y,w)),\, p(r(x,w), r(y,w), r(y,w))) : \\ &r \in \mathrm{Pol}_2^{(6m+1)}(\mathbf{A}),\ \{x,y\} = \{z,w\} = \{a,b\}\,\}. \end{aligned}$$

Definition 7.15. Let $\mathbf{A} \in \mathcal{V}_0$ and $u, v \in A$. We say that (u,v) is a *p-snag* if $u \neq v$ and $p(p(u,v,v),v,v) = v$.

Lemma 7.16. *Suppose $\mathbf{A} \in \mathcal{V}_0$, (u,v) is a p-snag, and $\gamma = \mathrm{Cg}^{\mathbf{A}}(u,v)$.*

1. *$[\gamma,\gamma] = \gamma$.*
2. *(u,v) is not contained in any solvable congruence of $\mathbf{A}$.*

Proof. Let $\delta = [\gamma,\gamma]$. Then in $\mathbf{A}/\delta$ we have $\overline{\gamma} := \gamma/\delta$ is abelian, so p is Maltsev on $\overline{\gamma}$-blocks. Thus $\overline{u} = p(p(\overline{u},\overline{v},\overline{v}),\overline{v},\overline{v}) = \overline{v}$, implying $(u,v) \in \delta$, so $\delta = \gamma$. This proves (1), which obviously implies (2). □

Lemma 7.17. *For all sufficiently large j:*

1. *If $\mathbf{A} \in \mathcal{V}^{(j)}$, $a, b \in A$, $\alpha = \mathrm{Cg}^{\mathbf{A}}(a,b)$ is not abelian, and $K_p^{(6m+1)}(a,b,a,b)$, then there exists a p-snag in $[\alpha,\alpha]$.*
2. *$\mathcal{V}^{(j)} \models G_m(a,b) \subseteq \Omega(a,b) \to K_p^{(6m+1)}(a,b,a,b)$.*

Proof. (1) If α is not abelian then we have $\neg\Omega(a,b,a,b)$ by Lemma 7.5. If in addition $K_p^{(6m+1)}(a,b,a,b)$, then $\neg L_{fg}^{(m)}(a,b,a,b)$ by Lemma 7.2(3). The proof of Lemma 4.8(1) then gives a failure of $H_2(a,b,a,b)$ at some $r_1, r_2 \in \mathrm{Pol}_2^{(3m)}(\mathbf{A})$. $K_p^{(6m+1)}(a,b,a,b)$ and Lemma 4.5(2) then give a p-snag in $[\alpha,\alpha]$.

(2) Arguing as in the proof of Lemma 7.2(4), we can assume that $\mathcal{V}^{(j)} \models K_p^{(6m+1)} \leftrightarrow K_p^{(6m+1,h_{\mathcal{V}}(6m+3))}$. By the same device, the set $G_m(a,b)$ can be defined (uniformly in $\mathbf{A} \in \mathcal{V}^{(j)}$ and $a, b \in A$, for sufficiently large j) by a first-order formula. Hence the claim to be established can be expressed by a first-order sentence, so it suffices by the compactness theorem to prove that it holds for $\mathbf{A} \in \mathcal{V}$. Let $\alpha = \mathrm{Cg}^{\mathbf{A}}(a,b)$ and $\delta = \mathrm{Cg}^{\mathbf{A}}(G_m(a,b))$, and observe that $\Omega(a,b) = \mathrm{ann}(a,b)$ by Corollary 7.7(2). It should be clear that $\delta \leq [\alpha,\alpha]$, so by C1 we have $[\alpha,\delta] = \delta$. However the hypothesis implies $[\alpha,\delta] = 0_A$, so $\delta = 0_A$, implying $G_m(a,b) \subseteq 0_A$, which means $K_p^{(6m+1)}(a,b,a,b)$. □

Lemma 7.18 (Lemma 4.7(2))**.** *For all sufficiently large j, if $\mathbf{A} \in \mathcal{V}^{(j)}$ and $\beta \in \operatorname{Con}\mathbf{A}$, then $[\beta,[\beta,\beta]] = [\beta,\beta]$.*

Proof. We first show that if $a, b \in A$, $\alpha = \mathrm{Cg}^{\mathbf{A}}(a,b)$, and $[\alpha,[\alpha,\alpha]] = 0_A$, then $[\alpha,\alpha] = 0_A$. Observe that the hypothesis implies $G_m(a,b) \subseteq \Omega(a,b)$, so $K_p^{(6m+1)}(a,b,a,b)$ by Lemma 7.17(2). The hypothesis also implies that α is solvable, so α is abelian by Lemmas 7.16 and 7.17(1), giving $[\alpha,\alpha] = 0_A$ as claimed.

Now suppose $\beta \in \operatorname{Con}\mathbf{A}$ and $[\beta,[\beta,\beta]] < [\beta,\beta]$. Let $\delta = [\beta,[\beta,\beta]]$. For $\theta \in \operatorname{Con}\mathbf{A}$ satisfying $\theta \geq \delta$ let $\overline{\theta} = \theta/\delta \in \operatorname{Con}\mathbf{A}/\delta$. Then by Lemma 2.4(2), $[\overline{\beta},\overline{\beta}] = \overline{[\beta,\beta]} > 0_{A/\delta}$ and $[\overline{\beta},[\overline{\beta},\overline{\beta}]] = \overline{[\beta,[\beta,\beta]]} = 0_{A/\delta}$. Thus by passing to $\mathbf{A}/\delta$ we can assume that $\delta = 0_A$. Observe next that $[\beta,\beta] \neq 0_A$ implies $\beta \not\leq \mu^{\mathbf{A}}$; pick $(a,b) \in \beta \setminus \mu^{\mathbf{A}}$ and put $\alpha = \mathrm{Cg}^{\mathbf{A}}(a,b)$. Then we still have $[\alpha,\alpha] > 0_A$ but $[\alpha,[\alpha,\alpha]] \leq [\beta,[\beta,\beta]] = 0_A$, contradicting the previous paragraph. □

Lemma 7.19 (cf. the proof of Lemma 4.20)**.** *For all sufficiently large j, if $\mathbf{A} \in \mathcal{V}^{(j)}$ and $\alpha_0, \alpha_1, \beta_1, \beta_2 \in \operatorname{Con}\mathbf{A}$ with β_1, β_2 principal, $0_A \prec \alpha_0 \prec \alpha_1$, α_1 abelian, and $[\alpha_0,\beta_1] = [\alpha_0,\beta_2] = 0_A$, then there exists an abelian atom $\gamma \in \operatorname{Con}\mathbf{A}$ such that $[\alpha_1,\beta_1], [\alpha_1,\beta_2] \leq \gamma$.*

Proof. We essentially follow the proof of [35, Lemma 4.20]. As $\mathcal{V}$ satisfies C1, the claim is true in $\mathcal{V}$ by Proposition 3.3. Thus it will suffice to show that the claim can be formulated as a first-order sentence. The claim is equivalent to the following statement:

> For all $a_0, b_0, a_1, b_1, c_1, d_1, c_2, d_2$, letting $\alpha_0 = \mathrm{Cg}^{\mathbf{A}}(a_0,b_0)$, $\overline{\alpha}_0 = \mathrm{Cg}^{\mathbf{A}}(a_1,b_1)$, $\alpha_1 = \alpha_0 \vee \overline{\alpha}_0$, $\beta_1 = \mathrm{Cg}^{\mathbf{A}}(c_1,d_1)$, and $\beta_2 = \mathrm{Cg}^{\mathbf{A}}(c_2,d_2)$, if:
> (1) α_0 is an abelian atom;
> (2) α_1 is abelian;
> (3) α_1/α_0 is an abelian atom in $\operatorname{Con}(\mathbf{A}/\alpha_0)$;
> (4) $[\alpha_0,\beta_1] = [\alpha_0,\beta_2] = 0_A$;
> then there exist e, f such that, setting $\gamma = \mathrm{Cg}^{\mathbf{A}}(e,f)$,
> (5) γ is an abelian atom;
> (6) $[\alpha_1,\beta_1] \leq \gamma$ and $[\alpha_1,\beta_2] \leq \gamma$.

(1) and (5) are first-order by Corollary 7.10(2a), (2) is equivalent to $\mu(a_0,b_0)$ & $\mu(a_1,b_1)$ by Lemma 7.13, and (4) is equivalent to $\Omega(a_0,b_0,c_1,d_1)$ & $\Omega(a_0,b_0,c_2,d_2)$. Since $|A/\operatorname{ann}(a_0,b_0)| \leq r$ by Corollary 7.10(1), Lemma 2.8 implies that α_0 is definable by the formula $\theta(x,y,a_0,b_0)$. (3) can now be stated by asserting $\mathrm{AbAt}(a_1,b_1)$ "mod α_0." By this we mean taking the formula $\mathrm{AbAt}(x,y)$ and replacing every occurrence of an equality $u = v$ with $\theta(u,v,a_0,b_0)$.

It remains to show that (6) can be formulated as a first-order statement. Assume $a_0, b_0, a_1, b_1, \alpha_0, \overline{\alpha}_0, \alpha_1$ are given as above (in particular, satisfying (1)–(3)), let $\beta = \mathrm{Cg}^{\mathbf{A}}(c,d)$ be a principal congruence of $\mathbf{A}$ satisfying $[\alpha_0,\beta] = 0_A$ and let $\gamma = \mathrm{Cg}^{\mathbf{A}}(e,f)$ be an abelian atom.

Claim: $[\alpha_1, \beta] \leq \gamma$ iff

(a) $[\overline{\alpha}_0, \beta] = 0_A$, or
(b) in $\mathbf{A}/\gamma$, $C(a_0/\gamma, b_0/\gamma, c/\gamma, d/\gamma)$ & $C(a_1/\gamma, b_1/\gamma, c/\gamma, d/\gamma)$.

Proof of Claim. ($\Leftarrow$) If (a) holds, then $[\alpha_1, \beta] = [\alpha_0 \vee \overline{\alpha}_0, \beta] = 0_A$ by semi-distributivity. If (b) holds, then $C(\alpha_0 \vee \gamma, \beta \vee \gamma; \gamma)$ and $C(\overline{\alpha}_0 \vee \gamma, \beta \vee \gamma; \gamma)$ hold (this is equivalent to (b)), so $C(\alpha_0 \vee \overline{\alpha}_0, \beta; \gamma)$, so $[\alpha_1, \beta] \leq \gamma$.

($\Rightarrow$) Assume $[\alpha_1, \beta] \leq \gamma$. Then either $[\alpha_1, \beta] = 0_A$ or $[\alpha_1, \beta] = \gamma$. If $[\alpha_1, \beta] = 0_A$ then (a) holds. Assume $[\alpha_1, \beta] = \gamma$. Then $\gamma \leq \beta$ and $C(\alpha_1 \vee \gamma, \beta; \gamma)$. These facts imply $C(\alpha_0 \vee \gamma, \beta \vee \gamma; \gamma)$ and $C(\overline{\alpha}_0 \vee \gamma, \beta \vee \gamma; \gamma)$, so (b) holds, proving the Claim. □

Returning to the proof of (6), observe that γ (like α_0 considered above) is definable by the formula $\theta(x, y, e, f)$. It follows from the Claim that we can express $[\alpha_1, \beta_i] \leq \gamma$ by asserting

$$\Omega(a_1, b_1, c_i, d_i) \text{ or } [\ \Omega(a_0, b_0, c_i, d_i) \text{ “mod } \gamma\text{” and } \Omega(a_1, b_1, c_i, d_i) \text{ “mod } \gamma\text{”}]$$

where by $\Omega(x, y, z, w)$ “mod γ” we mean the formula obtained from $\Omega(x, y, z, w)$ by replacing each occurrence of an equality $u = v$ with $\theta(u, v, e, f)$. This shows that (6) is expressible as a first-order statement, and completes the proof of the Lemma. □

The remainder of the proof departs from McKenzie's proof for the congruence modular case.

Lemma 7.20. *For all sufficiently large j, if $\mathbf{A} \in \mathcal{V}^{(j)}$, α is an abelian atom in* $\operatorname{Con} \mathbf{A}$*, β is a principal congruence, and $[\alpha, \beta] = 0_A$, then:*

(1) *If $\lambda \in \operatorname{Con} \mathbf{A}$ satisfies $[\alpha, \lambda] = 0_A$ and $C(\lambda, \alpha \vee \beta; \alpha)$, then $[\lambda, \beta] = 0_A$.*
(2) *$|A/(0_A : \beta)| \leq r^2$ if $\alpha \prec \alpha \vee \beta$.*

Proof. (1) It suffices to prove the claim under the assumption that λ is principal. Let $\alpha = \operatorname{Cg}^{\mathbf{A}}(a, b)$, $\beta = \operatorname{Cg}^{\mathbf{A}}(c, d)$, and $\lambda = \operatorname{Cg}^{\mathbf{A}}(u, v)$. The claim is then equivalent to the following:

If $\operatorname{AbAt}(a, b)$, $\Omega(a, b, c, d)$, $\Omega(a, b, u, v)$ and “$\Omega(u, v, c, d)$ mod α,” then $\Omega(u, v, c, d)$.

All but the last of the hypotheses is clearly first-order, and the last (“$\Omega(u, v, c, d)$ mod α”) can also be expressed by a first-order formula since α is a definable congruence. Hence it suffices to prove the claim in $\mathcal{V}$. Assume $\mathbf{A} \in \mathcal{V}$. The hypotheses imply $[\alpha, \beta \vee \lambda] = 0_A$ and $[\lambda, \beta] \leq \alpha$. Thus if $[\lambda, \beta] \neq 0_A$ then $[\lambda, \beta] = \alpha$, so $\alpha \leq [\beta \vee \lambda, \beta \vee \lambda]$, so $\alpha = [\alpha, \beta \vee \lambda]$ by C1, contradiction.

(2) By Corollary 7.10(1), $(0_A : \alpha)$ and $(\alpha : \alpha \vee \beta)$ both have index at most r. Thus it will suffice to show $(0_A : \alpha) \cap (\alpha : \alpha \vee \beta) \subseteq (0_A : \beta)$. Let λ be a principal congruence contained in $(0_A : \alpha) \cap (\alpha : \alpha \vee \beta)$; it suffices to prove $\lambda \leq (0_A : \beta)$. We did this in (1). □

Recall [11, Definition 7.1] that a 2-*snag* in an algebra $\mathbf{A}$ is a pair $(c,d) \in A^2$ with $c \neq d$ for which there exists $f \in \mathrm{Pol}_2(\mathbf{A})$ satisfying $f(c,d) = f(d,c) = f(c,c) = c$ and $f(d,d) = d$. Such an f is called a *pseudo-meet operation* for the 2-snag. Note that if (c,d) is a 2-snag and $\beta = \mathrm{Cg}^{\mathbf{A}}(c,d)$, then $[\beta,\beta] = \beta$.

Lemma 7.21. *For all sufficiently large j, if $\mathbf{A} \in \mathcal{V}^{(j)}$ is finite, $\alpha, \beta \in \mathrm{Con}\,\mathbf{A}$ with $\alpha \prec \beta$, and β/α is non-abelian, then $\beta \setminus \alpha$ contains a 2-snag having a pseudo-meet operation in $\mathrm{Pol}_2^{(r)}(\mathbf{A})$.*

Proof. Let $\gamma = (\alpha : \beta)$. Then $\gamma/\alpha = (0_{A/\alpha} : \beta/\alpha)$, so γ has index at most r by Lemma 7.10(1). Moreover, γ is the largest congruence of $\mathbf{A}$ containing α but not β (because β/α is nonabelian). Let θ be its unique upper cover. Then (γ,θ) is perspective to (α,β), so θ/γ is non-abelian. Pick $(a,b) \in \beta \setminus \alpha$ and $e_3,\dots,e_r \in A$ so that $\{a,b,e_3,\dots,e_r\}$ contains a transversal for γ. Let $\mathbf{B} = \mathrm{Sg}^{\mathbf{A}}(a,b,e_3,\dots,e_r)$. The key observation is that the map

$$\mathbf{B} \overset{\mathsf{incl}}{\hookrightarrow} \mathbf{A} \overset{\mathsf{nat}}{\twoheadrightarrow} \mathbf{A}/\gamma \quad \text{given by} \quad x \mapsto x/\gamma$$

is surjective, since $\mathbf{B}$ contains a transversal for γ. This map has kernel $\gamma{\restriction}_B$, so the induced map $\mathbf{B}/\gamma{\restriction}_B \to \mathbf{A}/\gamma$ given by $x/\gamma{\restriction}_B \mapsto x/\gamma$ is an isomorphism. Hence $\gamma{\restriction}_B \prec \theta{\restriction}_B$ in $\mathrm{Con}\,\mathbf{B}$ and $\theta{\restriction}_B$ is nonabelian over $\gamma{\restriction}_B$.

Observe that we still have $(a,b) \notin \gamma{\restriction}_B$ but $(a,b) \in \beta' := \mathrm{Cg}^{\mathbf{B}}(a,b) \leq \beta{\restriction}_B \leq \theta{\restriction}_B$. Let $\psi \in \mathrm{Con}\,\mathbf{B}$ be an upper cover of $\alpha' := \alpha{\restriction}_B \cap \beta'$ below β'. Then (α',ψ) and $(\gamma{\restriction}_B, \theta{\restriction}_B)$ are perspective, so ψ is non-abelian over α'. Let (c,d) be a 2-snag of $\mathbf{B}$ in $\psi \setminus \alpha'$ (this exists by tame congruence theory; see [11, Exercise 5.11(1)]); then (c,d) satisfies the claim. □

Definition 7.22. Given a variety $\mathcal{V}$ with a difference term p, an algebra $\mathbf{A} \in \mathcal{V}$, and $a,b \in A$, we call (a,b) a *Maltsev pair* (for p) if $p(a,b,b) = a$. Given an algebra $\mathbf{A}$, $a,b,c,d \in A$, and $f \in \mathrm{Pol}_1(\mathbf{A})$, we write $(c,d) \xrightarrow{f} (a,b)$ to mean $f(c) = a$ and $f(d) = b$.

Lemma 7.23. *Assume that $\mathcal{V}$ is a variety with difference term p, $\mathbf{A} \in \mathcal{V}$, $a,b,c,d \in A$, (a,b) is a Maltsev pair, and $\mathrm{ann}(c,d)$ has finite index k. If $(c,d) \xrightarrow{f} (a,b)$ for some polynomial f, then $(c,d) \xrightarrow{g} (a,b)$ for some polynomial $g \in \mathrm{Pol}_1^{(k+3)}(\mathbf{A})$.*

Proof. Choose a term $t(x,y_1,\dots,y_m)$ and parameters $\mathbf{u} \in A^m$ so that $f(x) = t^{\mathbf{A}}(x,\mathbf{u})$. Let T be a transversal for $\mathrm{ann}(c,d)$. For each u_i let e_i be the unique member of T which is $\mathrm{ann}(c,d)$-related to u_i and let $f'(x) = t^{\mathbf{A}}(x,\mathbf{e})$. Then $f' \in \mathrm{Pol}_1^{(k)}(\mathbf{A})$.

We have

$$p(f(c),f(c),f(d)) = f(d) = p(f'(c),f'(c),f(d)),$$

so

$$p(f(c),f(d),f(d)) = p(f'(c),f'(d),f(d)),$$

i.e.,

$$a = p(a,b,b) = p(f'(c), f'(d), f(d)).$$

It follows that

$$g(x) := p(a, p(f'(c), f'(x), f(d)), b) = p(a, p(f'(c), f'(x), b), b)$$

witnesses that $(c,d) \xrightarrow{g} (a,b)$. Since the polynomial g involves only the k parameters of f' along with the three parameters $a, b, f'(c)$ we get that $g \in \mathrm{Pol}_1^{(k+3)}(\mathbf{A})$. □

Lemma 7.24. *For all sufficiently large j, if $\mathbf{A} \in \mathcal{V}^{(j)}$ is finite and $0_A \prec \alpha \prec \beta$ in* Con **A** *with α abelian and $[\beta,\beta] = \beta$, then $[\alpha,\beta] = \alpha$.*

Proof. Assume instead that $[\alpha,\beta] = 0_A$. If there exists $\gamma \prec \beta$ with $\gamma \neq \alpha$, then β/γ would be perspective with $\alpha/0_A$, so would be abelian, implying $[\beta,\beta] \leq \gamma$ which is false. Thus α is the unique lower cover of β.

As β/α is non-abelian, there exists a 2-snag $(c,d) \in \beta$ having a pseudo-meet operation $h \in \mathrm{Pol}_2^{(r)}(\mathbf{A})$, by Lemma 7.21. Clearly $(c,d) \notin \alpha$. Thus $\mathrm{Cg}^{\mathbf{A}}(c,d) = \beta$.

Let U be a $(0_A,\alpha)$-minimal set, let $e(x) \in \mathrm{Pol}_1(\mathbf{A})$ satisfy $e^2(x) = e(x)$ and $e(A) = U$, let V be a trace in U, and choose $(a,b) \in V^2 \setminus 0_V$. Because $(a,b) \in \mathrm{Cg}^{\mathbf{A}}(c,d)$, there exist $a = a_0, a_1, \ldots, a_n = b \in U$ and $f_1, \ldots, f_n \in \mathrm{Pol}_1(\mathbf{A})$ with $\{ef_i(c), ef_i(d)\} = \{a_{i-1}, a_i\}$ for each i. Let $T = \{(ef(c), ef(d)) : f \in \mathrm{Pol}_1(\mathbf{A})\}$. T is a reflexive subalgebra of $(\mathbf{A}|_U)^2$ and $\mathbf{A}|_U$ has a Maltsev operation (because $\mathrm{typ}(0_A,\alpha) = \mathbf{2}$ and U has empty tail by Lemma 2.5), so T is an equivalence relation on U. These facts imply $(a,b) \in T$ and hence $(c,d) \xrightarrow{f} (a,b)$ for some $f \in \mathrm{Pol}_1(\mathbf{A})$. As $\mathrm{ann}(c,d)$ has index at most r^2 by Lemma 7.20, we have $(c,d) \xrightarrow{g} (a,b)$ for some $g \in \mathrm{Pol}_1^{(r^2+3)}(\mathbf{A})$, by Lemma 7.23.

In summary, we have elements a,b,c,d in an algebra $\mathbf{A} \in \mathcal{V}^{(j)}$ satisfying:

- $a \neq b$;
- (c,d) is a 2-snag having a pseudo-meet operation in $\mathrm{Pol}_1^{(r)}(\mathbf{A})$;
- $(c,d) \xrightarrow{g} (a,b)$ for some $g \in \mathrm{Pol}_1^{(r^2+3)}(\mathbf{A})$;
- $\Omega(a,b,c,d)$.

Modulo σ_{r+1} & σ_{r^2+4}, this configuration is first-order definable. It cannot exist in $\mathcal{V}$ (for if $\alpha := \mathrm{Cg}^{\mathbf{A}}(a,b)$ and $\beta := \mathrm{Cg}^{\mathbf{A}}(c,d)$, then the configuration implies $0_A < \alpha \leq \beta = [\beta,\beta]$ and $[\alpha,\beta] = 0_A$, which violates C1). Hence it cannot exist in $\mathcal{V}^{(j)}$ for sufficiently large j. □

Corollary 7.25. $\mathcal{V}^{(j)}$ *satisfies* C1 *for all sufficiently large j.*

Proof. Proposition 3.3 with Lemmas 7.11, 7.13, 7.18, 7.19, and 7.24. □

We can now prove that $\mathcal{V}$ is finitely based. Choose j large enough so that all of the foregoing claims about $\mathcal{V}^{(j)}$ are satisfied. Then $\mathcal{V}^{(j)}$ is locally finite, satisfies C1, and is such that for every $\mathbf{A} \in \mathcal{V}^{(j)}$ and atom $0_A \prec \alpha$, the index of $(0_A : \alpha)$ is at most r.

It follows by Propositions 3.1 and 3.2(2) that $\mathcal{V}^{(j)}$ has a finite residual bound. Then in the usual way we can argue that $\mathcal{V}$ is finitely axiomatizable relative to $\mathcal{V}^{(j)}$. Since $\mathcal{V}^{(j)}$ is finitely based, so is $\mathcal{V}$.

Acknowledgment. We thank George McNulty for help in tracing the early history of Jónsson's Problem, and the referee for his/her careful scrutiny of the manuscript.

References

[1] K.A. Baker, Equational bases for finite algebras, *Notices Amer. Math. Soc.* **19** (1972), 691-08-02, p. A-44.
[2] K.A. Baker, Finite equational bases for finite algebras in a congruence-distributive equational class, *Adv. Math.* **24** (1977), 207–243.
[3] K.A. Baker, G. McNulty and J. Wang, An extension of Willard's finite basis theorem: congruence meet-semidistributive varieties of finite critical depth, *Algebra Universalis* **52** (2004), 289–302.
[4] G. Birkhoff, On the structure of abstract algebras, *Proc. Camb. Philos. Soc.* **31** (1935), 433–454.
[5] R. Bryant, The laws of finite pointed groups, *Bull. London Math. Soc.* **14** (1982), 119–123.
[6] G. Czédli, A characterization of congruence semi-distributivity, *Universal algebra and lattice theory (Proceedings of Conference, Puebla, 1982)*, Springer Lecture Notes, no. 1004, 1983, pp. 104–110.
[7] R. Freese and R. McKenzie, Residually small varieties with modular congruence lattices, *Trans. Amer. Math. Soc.* **264** (1981), 419–430.
[8] R. Freese and R. McKenzie, *Commutator theory for congruence modular varieties*, LMS Lecture Note Series, 125, Cambridge Univ. Press, Cambridge, 1987,
[9] H.-P. Gumm, Congruence modularity is permutability composed with distributivity, *Arch. Math. (Basel)* **36** (1981), 569–576.
[10] H.-P. Gumm, Geometrical methods in congruence modular algebras, *Mem. Amer. Math. Soc.* **45** (1983), no. 286.
[11] D. Hobby and R. McKenzie, *The Structure of Finite Algebras*, Contemporary Mathematics vol. 76, Amer. Math. Soc., Providence, 1988.
[12] I.M. Isaev, Inherently non-finitely based varieties of algebras, *Sibirsk. Mat. Ž.* **30** (1989), 75–77. English translation: *Siberian Math. J.* **30** (1989), 892–894.
[13] B. Jónsson, Algebras whose congruence lattices are distributive, *Math. Scand.* **21** (1967), 110–121.
[14] B. Jónsson and I. Rival, Lattice varieties covering the smallest nonmodular variety, *Pacific J. Math.* **82** (1979), 463–478.
[15] K. Kearnes, An order-theoretic property of the commutator, *Internat. J. Algebra Comput.* **3** (1993), 491–533.
[16] K. Kearnes, Varieties with a difference term, *J. Algebra* **177** (1995), 926–960.
[17] K. Kearnes, Cardinality bounds for subdirectly irreducible algebras, *J. Pure Appl. Algebra* **112** (1996), 293–312.
[18] K. Kearnes, A Hamiltonian property for nilpotent algebras, *Algebra Universalis* **37** (1997), 403–421.
[19] K. Kearnes and Á. Szendrei, The relationship between two commutators, *Internat. J. Algebra Comput.* **8** (1998), 497–531.

[20] E.W. Kiss, Three remarks on the modular commutator, *Algebra Universalis* **29** (1992), 455–476.
[21] E.W. Kiss, An easy way to minimal algebras, *Internat. J. Algebra Comput.* **7** (1997), 55–75.
[22] R. Kruse, Identities satisfied by a finite ring, *J. Algebra* **26** (1973), 298–318.
[23] P. Lipparini, Commutator theory without join-distributivity, *Trans. Amer. Math. Soc.* **346** (1994), 177–202.
[24] P. Lipparini, n-Permutable varieties satisfy non trivial congruence identities, *Algebra Universalis* **33** (1995), 159–168.
[25] P. Lipparini, A characterization of varieties with a difference term, *Canad. Math. Bull.* **39** (1996), 308–315.
[26] P. Lipparini, A characterization of varieties with a difference term. II. Neutral = meet semi-distributive, *Canad. Math. Bull.* **41** (1998), 318–327.
[27] P. Lipparini, A Kiss 4-difference term from a ternary term, *Algebra Universalis* **42** (1999), 153–154.
[28] I.V. L'vov, Varieties of associative rings, I. (Russian), *Algebra i Logika* **12** (1973), 269–297. English translation: *Algebra and Logic* **12** (1973), 150–167.
[29] I.V. L'vov, Finite-dimensional algebras with infinite bases of identities. (Russian), *Sibirsk. Mat. Ž.* **19** (1978), 91–99, 237. English translation: *Siberian Math. J.* **19** (1978), 63–39.
[30] R. Lyndon, Identities in two-valued calculi, *Trans. Amer. Math. Soc.* **3** (1951), 457–465.
[31] R. Lyndon, Identities in finite algebras, *Proc. Amer. Math. Soc.* **5** (1954), 8–9.
[32] M. Maróti and R. McKenzie, Finite basis problems and results for quasivarieties, *Studia Logica* **78** (2004), 293–320.
[33] R. McKenzie, Equational bases for lattice theories, *Math. Scand.* **27** (1970), 24–38.
[34] R. McKenzie, Para primal varieties: A study of finite axiomatizability and definable principal congruences in locally finite varieties, *Algebra Universalis* **8** (1978), 336–348.
[35] R. McKenzie, Finite equational bases for congruence modular varieties, *Algebra Universalis* **24** (1987), 224-250.
[36] R. McKenzie, Tarski's finite basis problem is undecidable, *Internat. J. Algebra Comput.* **6** (1996), 49–104.
[37] G. McNulty and C. Shallon, Inherently nonfinitely based finite algebras, in *Lecture Notes in Math.* vol. 1004, eds. R. Freese and O.C. Garcia, Springer, 1983, pp. 206–231.
[38] S. Oates and M.B. Powell, Identical relations in finite groups, *J. Algebra* **1** (1964), 11–39.
[39] R.E. Park, *Equational Classes of Non-Associative Ordered Algebras*, PhD dissertation, University of California (Los Angeles), 1976.
[40] P. Perkins, Bases for equational theories of semigroups, *J. Algebra* **11** (1969), 298–314.
[41] S.V. Polin, Identities of finite algebras. (Russian), *Sibirsk. Mat. Ž.* **17** (1976), 1356–1366, 1439. English translation: *Siberian Math. J.* **17** (1976), 992–999 (1977).
[42] R. Quackenbush, Equational classes generated by finite algebras, *Algebra Universalis* **1** (1971), 265–266.
[43] W. Taylor, Equational logic, *Contributions to universal algebra (Colloq, Jószef Attila Univ., Szeged, 1975)*, Colloq. Math. Soc. Janos Bolyai, vol. 17, North-Holland, 1977, pp. 465–501.
[44] R. Willard, A finite basis theorem for residually finite, congruence meet-semidistributive varieties, *J. Symbolic Logic* **65** (2000), 187–200.
[45] R. Willard, The finite basis problem, *Contributions to general algebra*, vol. 15, Heyn, Klagenfurt, 2004, pp. 199–206.
[46] Workshop Report, Mathematisches Forschungsinstitut Oberwolfach, workshop no. 35 held Aug. 15–21, 1976, ed. H.P. Gumm, unpublished.

(K. Kearnes) Department of Mathematics, University of Colorado at Boulder, Campus Box 395, Boulder, CO 80309-0385, USA

(Á. Szendrei) Department of Mathematics, University of Colorado at Boulder, Campus Box 395, Boulder, CO 80309-0385, USA, and Bolyai Institute, Aradi vértanúk tere 1, H-6720 Szeged, Hungary

(R. Willard) Pure Mathematics Department, University of Waterloo, Waterloo, ON N2L 3G1, Canada