Model Theory

no. 3 vol. 3 2024

$M \nvDash T$

msp

Model Theory

msp.org/mt

See inside back cover or msp.org/mt for submission instructions.

Model Theory (ISSN 2832-904X electronic, 2832-9058 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online.

MT peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY
mathematical sciences publishers
nonprofit scientific publishing
https://msp.org/

Model Theory
Vol. 3, No. 3, 2024
https://doi.org/10.2140/mt.2024.3.733

Noncommutative algebraic geometry
I: Monomial equations with a single variable

Zlil Sela

This paper is the first in a sequence on the structure of sets of solutions to systems of equations over a free associative algebra. We start by constructing a Makanin–Razborov diagram that encodes all the homogeneous solutions to a homogeneous monomial system of equations. Then we analyze the set of solutions to monomial systems of equations with a single variable.

Algebraic geometry studies the structure of sets of solutions to systems of equations usually over fields or commutative rings. The developments and the considerable abstraction that currently exist in the study of varieties over commutative rings still resists application to the study of varieties over nonabelian rings or over other nonabelian algebraic structures.

Since 1960 ring theorists such as P. M. Cohn [1971], G. M. Bergman [1969] and others have tried to study varieties over nonabelian rings, notably free associative algebras (and other free rings). However, the pathologies that they tackled and the lack of unique factorization that they study in detail [Cohn 1971, Chapters 3–4] prevented any attempt to prove or even speculate what can be the structure of varieties over free associative algebras.

In this sequence of papers we suggest studying varieties over free associative algebras using techniques and analogies of structural results from the study of varieties over free groups and semigroups. Over free groups and semigroups geometric techniques as well as low-dimensional topology play an essential role in the structure of varieties. These include Makanin's algorithm for solving equations, Razborov's analysis of sets of solutions over a free group, the concepts and techniques that were used to construct and analyze the JSJ decomposition, and the applicability of the JSJ machinery to study varieties over free groups and semigroups [Sela 2001; 2016]. Our main goal is to demonstrate that these techniques and concepts can be modified to be applicable over free associative algebras as well.

Furthermore, we believe that the concepts and techniques that proved to be successful over free groups and semigroups can be adapted to analyze varieties over

Partially supported by an Israel Academy of Sciences fellowship.
MSC2020: primary 16S10; secondary 03C68, 20F65.
Keywords: varieties, systems of equations, associative algebra, Makanin–Razborov diagram.

free objects in other noncommutative and at least "partially" associative algebraic structures. In that respect, we hope that it will be possible to use or even axiomatize the properties of varieties over the free objects in these algebraic structures, in order to set dividing lines between noncommutative algebraic structures, in analogy with classification theory (of first-order theories) in model theory [Shelah 1990].

We start the analysis of systems of equations over a free associative algebra with what we call monomial systems of equations. These are systems of equations over a free associative algebra in which every polynomial in the system contains two monomials. In Section 1 we analyze the case of homogeneous solutions to homogeneous monomial systems of equations. In this case it is possible to apply the techniques that were used in analyzing varieties over free semigroups [Sela 2016], and associate a Makanin–Razborov diagram that encodes all the homogeneous solutions to a homogeneous monomial system of equations.

In Section 2 we introduce *limit algebras*, which are a natural analogue of a *limit group*, and prove that such algebras are always embedded in (limit) division algebras (in analogy with the embeddings of limit semigroups in limit groups, that we termed *limit pairs* in [Sela 2016]). The automorphism (modular) groups of these division algebras are what is needed in the sequel in order to modify and shorten solutions to monomial systems of equations.

In Section 3 we present a combinatorial approach to (cases of) the celebrated Bergman's centralizer theorem [1969]. Finally, in the fourth section we use this combinatorial approach to analyze the set of solutions to a monomial system of equations with a single variable. The results that we obtain are analogous to the well known structure of the set of solutions to systems of equations with a single variable over a free group or semigroup. We prove all our results under the assumption that the top homogeneous parts of the coefficients in the equations are monomials with no periodicity, in order to simplify our arguments, but we believe that eventually this assumption can be dropped.

In the next paper in the sequence we use the techniques that are presented in this paper to analyze monomial systems of equations that have more than a single variable, but have no quadratic (or surface) parts. In the third paper in the sequence we analyze the quadratic parts of monomial systems of equations. Eventually, we hope to use our analysis of sets of solutions to monomial systems of equations to the analysis of general varieties.

1. Homogeneous solutions of monomial equations

For simplicity, we will always assume that the free algebras that we consider are over the field with two elements GF_2. Let FA be a free associative algebra over GF_2: $FA = GF_2\langle a_1, \dots, a_k\rangle$. In order to study the structure of general varieties over the associative algebra *FA*, we start with varieties that are defined by monomial systems of equations. A system of equations Φ is called *monomial* if it is defined using a finite set of unknowns $x_1, \dots, x_n$, and a finite set of equations

$$u_1(c_1, \dots, c_\ell, x_1, \dots, x_n) = v_1(c_1, \dots, c_\ell, x_1, \dots, x_n),$$
$$\vdots$$
$$u_s(c_1, \dots, c_\ell, x_1, \dots, x_n) = v_s(c_1, \dots, c_\ell, x_1, \dots, x_n),$$

where the words u_i and v_i are monomials in the free algebra generated by the variables $x_1, \dots, x_n$ and coefficients $c_1, \dots, c_\ell$ from the algebra *FA*, i.e., a word in the free semigroup generated by these elements (note that the coefficients $c_1, \dots, c_\ell$ are general elements and not necessarily monomials). A monomial system of equations is called *homogeneous* if all the coefficients $c_1, \dots, c_\ell$ in the system are homogeneous elements in the free associative algebra *FA*.

We start by analyzing all the homogeneous solutions of a homogeneous monomial system, i.e., all the assignments of homogeneous elements in *FA* to the variables $x_1, \dots, x_n$ such that the equalities in a homogeneous monomial system of equations are valid.

Let $x_1^0, \dots, x_n^0$ be a homogeneous solution of the monomial system Φ. Substituting the elements $x_1^0, \dots, x_n^0$ in the monomials u_i and v_i, $1 \leq i \leq s$, we get a finite set of equalities in the free algebra *FA*. Since all the elements that appear in each of these equalities are homogeneous, for each index i we can associate a segment J_i of length that is equal to the degree of u_i and v_i after the substitution of $x_1^0, \dots, x_n^0$. We further add notation on the segment J_i for the beginnings and the ends of each of the elements $x_1^0, \dots, x_n^0$ and the coefficients $c_1, \dots, c_\ell$ of the system.

With the segments $J_1, \dots, J_s$, and the notation for the beginnings and ends of $x_1^0, \dots, x_n^0$ and $c_1, \dots, c_\ell$, we can naturally associate a generalized equation as in [Makanin 1977; Casals-Ruiz and Kazachkov 2011], or alternatively a band complex (bands are added for different appearances of the same variable) as it appears in [Bestvina and Feighn 1995]. All the lengths that appear in the band complex are integers, so the band complex must be simplicial. Note that all the operations that are used in the Rips machine, or in the Makanin procedure, to transfer the original complex into a standard band complex are valid in our context, i.e., it is possible to cut the elements $x_1^0, \dots, x_n^0$ and $c_1, \dots, c_\ell$ and represent them as multiplication of new elements according to the operations that are performed in modifying the band complexes (or the generalized equation) along the procedure.

To clarify the applicability of the Makanin moves, one can look at the band complex or the corresponding Makanin generalized equation differently. Given the homogeneous solution $x_1^0, \dots, x_n^0$, and substituting it in the homogeneous monomial system of equations, we can naturally associate with each side of a monomial equation a homogeneous tree. Since each of the trees is composed from homogeneous elements, there are no cancellations between paths (monomials) in each separate tree, so the monomial equation implies that the homogeneous trees that are associated with the two sides of the equation are identical.

Now, the identical trees that are associated with the two sides of a monomial equation admit two product structures that are associated with the two sides of the equation. Therefore, the tree that is associated with a monomial equation admits a product structure which is the common refinement of the product structure coming from the two sides of the equation. Each band in the band complex, or alternatively each pair of bases in the Makanin generalized equation that is associated with the system, indicates that a certain part in this refined product structure of the tree that is associated with one monomial equation is identical to another part in the product structure of a tree that is associated with another (possibly the same) monomial equation. Alternatively, homogeneous elements in a free associative algebra have the unique factorization property. Hence, given two factorizations of a homogeneous element, there is a common refinement of the two factorizations.

Furthermore, each of the basic Makanin moves that can be performed on generalized equations can be performed in an identical way on the homogeneous trees that are associated with homogeneous monomial equations using their refined product structure. This means that the entire Makanin process to analyze solutions to systems of equations over a free semigroup, which is composed from sequences of basic moves, can be applied to the product structures of homogeneous trees that are associated with homogeneous monomial systems of equations.

The ability to apply the Makanin basic moves to the generalized equation or the band complex that is associated with a homogeneous system of monomial equations implies that it is possible to associate with such a system of equations a Makanin–Razborov diagram, using the construction of such a diagram for a system of equations over a free semigroup as it appears in [Sela 2016]. As in a free semigroup, the constructed diagram encodes all the homogeneous solutions to the homogeneous system of equations in the free algebra *FA*.

Let G_Φ be the semigroup that is generated by copies of $x_1, \dots, x_n$ and the coefficients $c_1, \dots, c_\ell$ modulo the relations

$$u_i(c_1, \dots, c_\ell, x_1, \dots, x_n) = v_i(c_1, \dots, c_\ell, x_1, \dots, x_n)$$

for $1 \le i \le s$, where the monomials u_i and v_i are interpreted as words in a free semigroup. With a (homogeneous) solution of the system Φ it is possible to associate a homomorphism from G_Φ into the free semigroup that is generated by a free generating set of *FA*.

Conversely, given a semigroup homomorphism of G_Φ into a free semigroup that fits with a decomposition of the constants $c_1, \dots, c_\ell$ into a product of homogeneous elements (there are finitely many possible ways to represent each of the coefficients $c_1, \ldots, c_\ell$ as such a product), it is possible to associate with such a product a family of solutions of the systems Φ.

Therefore, the study of homogeneous solutions of a homogeneous monomial system of equations over an associative free algebra is reduced to the study of a collection of semigroup homomorphisms from a given f.g. semigroup into a free semigroup. By [Sela 2016] with this collection of semigroup homomorphisms it is possible to associate canonically a finite collection of pairs $(S_1, L_1), \ldots, (S_m, L_m)$, where each of the groups L_j is a limit group, and each of the semigroups S_j is a f.g. subsemigroup that generates L_j. Furthermore, with G_Φ and its collection of homomorphisms it is possible to associate (noncanonically) a Makanin–Razborov diagram that encodes all its homomorphisms into free semigroups. By our observation, this Makanin–Razborov diagram of pairs encodes all the homogeneous solutions of the system Φ in the algebra *FA*.

Theorem 1.1. *With a homogeneous monomial system of equations over the free associative algebra FA it is possible to associate* (*noncanonically*) *a Makanin–Razborov diagram that encodes all its homogeneous solutions.*

As a corollary of the encoding of homogeneous solutions of a system of homogeneous monomial equations by pairs of limit groups and their subsemigroups we get the following.

Corollary 1.2. *The collections of sets of homogeneous solutions to homogeneous monomial systems of equations is Noetherian*, *i.e.*, *every descending sequence of such sets terminates after a finite time.*

Proof. Follows immediately from the descending chain condition for limit groups [Sela 2001], or the Noetherianity of varieties over free groups and semigroups [Guba 1986]. □

Theorem 1.1 associates a Makanin–Razborov diagram with the set of homogeneous solutions to a homogeneous monomial systems of equations. Our main goal in this sequence of papers is to associate a Makanin–Razborov diagram with the set of (not necessarily homogeneous) solutions of a general monomial system of equations, at least in the minimal rank case, i.e., in the case in which the Makanin–Razborov diagram that is associated with the homogeneous system that is associated with top homogeneous part of the nonhomogeneous system contains no free products.

2. Limit algebras, their division algebras and modular groups

The construction of the Makanin–Razborov diagram of a system of equations over a free group uses extensively the (modular) automorphism groups of the limit groups that are associated with its nodes. These modular groups, defined in [Sela 2001, Definition 5.2], enable one to proceed from a limit group to maximal shortening quotients of it that are always proper quotients.

The semigroups that appear in the construction of the Makanin–Razborov diagram of a system of equations over a free semigroup do not have a large automorphism group in general, e.g., a finitely generated free semigroup has a finite automorphism group. Hence, to study homomorphisms from a given f.g. semigroup S to the free semigroup FS_k we did the following in [Sela 2016].

Given a f.g. semigroup S we can naturally associate a group with it. Given a presentation of S as a semigroup, we set the f.g. group $\mathrm{Gr}(S)$ to be the group with the presentation of S interpreted as a presentation of a group. Clearly, the semigroup S is naturally mapped to the group $\mathrm{Gr}(S)$ and the image of S in $\mathrm{Gr}(S)$ generates $\mathrm{Gr}(S)$. We set $\eta_S : S \to \mathrm{Gr}(S)$ to be this natural homomorphism of semigroups.

The free semigroup FS_k naturally embeds into a free group F_k. By the construction of the group $\mathrm{Gr}(S)$, every homomorphism of semigroups $h : S \to FS_k$ extends to a unique homomorphism of groups $h_G : \mathrm{Gr}(S) \to F_k$ such that $h = h_G \circ \eta_S$.

By construction, every homomorphism (of semigroups) $h : S \to FS_k$ extends to a homomorphism (of groups) $h_G : \mathrm{Gr}(S) \to F_k$. Therefore, the study of the structure of $\mathrm{Hom}(S, FS_k)$ is equivalent to the study of the structure of the collection of homomorphisms of groups $\mathrm{Hom}(\mathrm{Gr}(S), F_k)$ that restrict to homomorphisms of (the semigroup) S into the free semigroup (the *positive cone*) FS_k.

By (canonically) associating a finite collection of maximal limit quotients with the set of homomorphisms $\mathrm{Hom}(\mathrm{Gr}(S), F_k)$ that restrict to (semigroup) homomorphisms from S to FS_k, we are able to (canonically) replace the pair $(S, \mathrm{Gr}(S))$ with a finite collection of limit quotients $(S_1, L_1), \dots, (S_m, L_m)$, where each of the groups L_i is a limit group. Limit groups have rich modular groups, and these are later used to proceed to the next levels of the Makanin–Razborov diagram of the given system of equations over the free semigroup FS_k.

In studying sets of solutions to systems of equations over a free associative algebra, we need to study homomorphisms: $h : A \to FA_k$, where A is a f.p. algebra and FA_k is the free associative algebra of rank k. As in the case of groups and semigroups, to study such homomorphisms we pass to convergent sequences of homomorphisms $\{h_n : A \to FA_k\}$, and look at the *limit algebras* LA that are associated with such convergent sequences. Algebras, and in particular limit algebras, have automorphisms, but these are not the automorphisms that will be needed in the sequel to modify and shorten homomorphisms.

By a classical construction of [Malcev 1948; Neumann 1949], and by different constructions of Amitsur [1966] and others, the free associative algebra FA_k can be embedded into a division algebra $\mathrm{Div}(FA_k)$ (note that there are various different division algebras into which FA_k embeds). Given a convergent sequence

$\{h_n : A \to FA_k\}$ with an associated limit algebra LA, it is straightforward to get an embedding $LA \to \mathrm{Div}(LA)$, where $\mathrm{Div}(LA)$ is a division algebra that is also obtained from the convergent sequence and from the embedding $FA_k \to \mathrm{Div}(FA_k)$.

In the sequel we will use (a subgroup of) the group of automorphisms of the division algebra $\mathrm{Div}(LA)$ in order to modify (shorten) the homomorphisms $h : A \to FA_k$ that we need to study. These will be the modular groups that are associated with limit algebras that appear along the nodes of the Makanin–Razborov diagrams of the given systems of equations over the free associative algebra FA_k.

An important example is (a special case of) what we call *surface* (or *quadratic*) *algebra*:

$$SA = \langle x_1, \dots, x_n \mid x_1 \cdots x_n = x_{\sigma(1)} \cdots x_{\sigma(n)} \rangle$$

for an appropriate permutation $\sigma \in S_n$. Such a surface algebra is a limit algebra. Hence, it is embedded in a division algebra $\mathrm{Div}(SA)$. For appropriate convergent sequences, the modular group of $\mathrm{Div}(SA)$ contains the automorphism group of a corresponding surface. Therefore, we call the modular group of $\mathrm{Div}(SA)$ the *Bergman modular group* of a surface algebra, since it contains (or is generated by) generalized Dehn twists that are inspired by Bergman's centralizer theorem [1969]. These modular groups generalize the mapping class groups of surfaces; they will be defined in the sequel, and they play an essential role in constructing Makanin–Razborov diagrams for monomial systems of equations over a free associative algebra.

3. A combinatorial approach to Bergman's theorem

In the first section we studied homogeneous solutions to homogeneous monomial systems of equations. In this section we start the study of nonhomogeneous solutions to arbitrary monomial systems of equations. We start by studying the centralizers of elements in a free associative algebra, i.e., we give combinatorial proof to Bergman's theorem, and then use these techniques to study related systems of equations. We start with the following theorem, which can be proved easily by a direct induction, but we also present a proof that uses techniques that we will use in the sequel.

Theorem 3.1. *Let $u \in FA$ be an element for which its top degree homogeneous part is a monomial, and suppose that this top degree monomial has no nontrivial roots. Then the centralizer of u in FA is precisely the elements in the* (*one variable*) *algebra that is generated by u.*

Proof. Suppose that x is a (nontrivial) element that satisfies $xu = ux$. By our analysis of homogeneous elements, the top degree homogeneous part of x must be a monomial which is a power of the top degree monomial in u. Hence, the top degree monomial of x has to be identical to the top degree monomial of u^m for some m. Therefore, $\deg(x+u^m)<\deg(x)$ and $u(x+u^m)=(x+u^m)u$, so the theorem follows by induction on the degree of x. □

For later applications we present a different proof.

Proof. First, note that $xu=ux$ if and only if $x(u+1)=(u+1)x$. Hence, we may assume that the monomials in u do not include the one corresponding to the identity.

In the sequel we denote by G^m the (additive) abelian group that is generated by all the monomials of degree at most m in FA. Given $x, y \in FA$, we write $x=y \mod G^m$ if x and y have the same monomials of degree bigger than m. If $x \in FA$, we say that a monomial w is a monomial of codegree m in x if $\deg(w)+m=\deg(x)$, and w is in the support of x (i.e., w appears in writing x as a sum of monomials).

Lemma 3.2. *Suppose that* $\deg(u)\geq 2$, *the top degree homogeneous part of u is a monomial and has no periodicity, and that* $\deg(x)\geq 2\deg(u)$. *There exists an element* $w \in FA$, *such that*

$$x=uw=wu \qquad \mod G^{\deg(x)-2},$$
$$ux=xu=uwu \qquad \mod G^{\deg(ux)-2}.$$

Proof. We analyze the codegree 1 monomials in the two sides of the equation $xu=ux$. Let u_0, x_0 be the top monomials, and let u_1, x_1 be the codegree 1 element in u, x. Clearly, $x_0u_1+x_1u_0$ is the codegree 1 element in xu, and $u_0x_1+u_1x_0$ is the codegree 1 element in ux.

Suppose that there are no cancellations between the codegree 1 monomials (that are obtained using the distributive law) in each of the two sides of the equation. In that case, since u_0 was assumed to have no periodicity, monomials in x_0u_1 cannot be monomials in u_1x_0, since otherwise x_0 overlaps with itself in a shift of a single place, and x_0 has a period which is u_0 that has degree at least 2 by the assumption of the lemma. Hence, monomials in x_0u_1 have to be monomials in u_0x_1. Similarly, monomials in u_1x_0 have to be monomials in x_1u_0.

Hence, if $x_0=u_0^m$ and $w_0=u_0^{m-1}$, then from the right side of the equation $xu=ux$, $x_1=u_1w_0+u_0\hat{w}_1$. From the left side of the equation, $x_1=w_0u_1+\hat{w}_2u_0$. So if we consider elements of codegree at most 1, $x=uw_1=w_2u$. From the equation $xu=ux$, we get that for elements of codegree at most 1, $uw_1u=uw_2u$, so $u(w_1-w_2)u$=0, so $w_1=w_2$, and $x=wu=uw$, for elements of codegree at most 1 (for some element w of degree $\deg(w)=\deg(x)-\deg(u)$).

Suppose that there are cancellations between codegree 1 monomials in the left-hand side xu. In that case monomials in x_0u_1 cancel with monomials in x_1u_0. Let v_1 be the codegree 1 suffix of u_0, and y_1 be the unique monomial in x_1 for which $x_0v_1=y_1u_0$. In that case $y_1=u_0\widetilde{w}$. Hence, the monomial y_1 has a product structure which is similar with the other codegree 1 monomials in x_1. Therefore, as in the previous case, when analyzing monomials of codegree at most 1 in x, x can be described both as $w_1 u$ and $u w_2$ (from the two sides of the equation), and the same argument that was used in case there are no cancellations works. □

We continue the proof of Theorem 3.1, by iteratively uncovering the homogeneous parts in an element x that is in the centralizer of u from top to bottom. Since $x = t_1 u \mod G^{\deg(x)-2}$, if $\deg(x) \geq 2\deg(u)$, it follows that $xu = t_1 u^2$ and $ux = ut_1 u \mod G^{\deg(xu)-2}$. Hence, if $\deg(t_1) \geq 2\deg(u)$, then $t_1 = t_2 u = ut_2 \mod G^{\deg(t_1)-2}$. Applying these arguments iteratively we get that $x = tu^m \mod G^{\deg(x)-2}$, for some t that satisfies $\deg(t) = \deg(u)$.

Therefore, $xu = tu^m u = ux = utu^m \mod G^{\deg(xu)-2}$, which means that $tu = ut \mod G^{\deg(tu)-2}$ and $\deg(t) = \deg(u)$.

In this case, in which $tu = ut$ and $\det(t) = \deg(u)$, the top degree monomials of t and u are identical, and we denote this monomial u_0. Suppose that s, v are monomials of codegree 1 in either t or u, and suppose that $su_0 = u_0 v$. In that case v is the suffix of u_0 and s is the prefix of u_0. Since u_0 is not a proper power, vu_0 cannot be presented as $u_0 w$ for any codegree 1 monomial w, and $u_0 s$ cannot be presented as wu_0 for any codegree 1 monomial w.

Hence, if s, v are codegree 1 monomials in either t or u, and $su_0 = u_0 v$, then both $u_0 s$ and vu_0 can be presented uniquely in each of the two products tu and ut, which implies that s and v must be codegree 1 monomials in both u and t. Therefore, the codegree 1 monomials of t and u must be identical, so $t = u \mod G^{\deg(t)-2}$, and $x = u^{m+1} \mod G^{\deg(x)-2}$ for some nonnegative integer m.

We use (a finite) induction and assume that $x = u^{m+1} \mod G^{\deg(x)-c}$, for some positive integer $c < \deg(u)$, i.e., we assume that the equality holds for all the monomials in x and u of codegree smaller than c. To complete the proof of the theorem, we need to prove the same equality for all the monomials of codegree at most c.

By the inductive hypothesis, $x = u^{m+1} \mod G^{\deg(x)-c}$. Hence, $x = x_{c-1} + v$, where x_{c-1} is the sum of all the monomials of codegree smaller than c in x and $\deg(v) \leq \deg(x) - c$. Furthermore, x_{c-1} is precisely the sum of all the monomials of codegree smaller than c in u^{m+1}.

Let u_{c-1} be the sum of the monomials of codegree less than c in u. We set s_c to be the sum of all the monomials of codegree c in u_{c-1}^{m+1}. By construction, $u_{c-1}(x_{c-1} + s_c) = (x_{c-1} + s_c)u_{c-1} = u_{c-1}^{m+2} \mod G^{\deg(xu)-(c+1)}$, i.e., the monomials of codegree at most c are identical for the three different products.

Recall that $x = x_{c-1} + v$, where $\deg(v) \leq \deg(x) - c$. We set $x = x_{c-1} + s_c + r$, where $\deg(s_c) = \deg(x) - c$ and $\deg(r) \leq \deg(x) - c$. Let q_c be the sum of the monomials of codegree c in u. Then

$$ux = (u_{c-1} + q_c)(x_{c-1} + s_c + r) = xu = (x_{c-1} + s_c + r)(u_{c-1} + q_c) \mod G^{\deg(xu)-(c+1)}.$$

Since $u_{c-1}(x_{c-1}+s_c)=(x_{c-1}+s_c)u_{c-1} \mod G^{\deg(xu)-(c+1)}$, it follows that

$$u_{\text{top}}r+q_c x_{\text{top}}=ru_{\text{top}}+x_{\text{top}}q_c \mod G^{\deg(xu)-(c+1)},$$

where u_0 and x_0 are the top monomials in u and x in correspondence. Therefore, all these monomials are products of a top degree monomial with a codegree c monomial, and these can be broken precisely as in the codegree 1 case, assuming $\deg(x)\geq 2\deg(u)$.

We are left with the case in which $\deg(x)=\deg(u)$. In that case we write $u=u_{c-1}+q_c$ and $x=u_{c-1}+r_c \mod G^{\deg(x)-(c+1)}$, where q_c and r_c are the codegree c monomials in u and x in correspondence. Since the contributions of products of monomials of codegree smaller than c in xu and in ux are identical, we need to look only at the equation $u_0r_c+q_cx_0=x_0q_c+r_cu_0$ for the monomials of codegree c, where $x_0=u_0$ are identical monomials. By the argument that was used in the codegree 1 case (when $\deg(x)=\deg(u)$), it follows that $q_c=r_c$, and the general step of the induction is proved.

So far we may conclude that $x=u^{m+1} \mod G^{\deg(x)-\deg(u)}$. Thus, $x+u^{m+1}$ commutes with u and $\deg(x+u^{m+1})\leq\deg(x)-\deg(u)$, and the theorem follows. □

So far we assumed that the top homogeneous element of u is a monomial and that its top monomial doesn't have a proper root. We continue by allowing u to be a proper power.

Theorem 3.3. *Let $u\in FA$ be an element for which its top degree homogeneous part is a monomial, and suppose that $u=p(v)$ and the top degree monomial of v does not have a proper root. Then the centralizer of u in FA is precisely the elements in the algebra that are generated by v.*

Proof. Suppose that x is a (nontrivial) element that satisfies $xp(v)=p(v)x$. First, note that like Theorem 3.1, Theorem 3.3 can be proved easily by replacing x by $x+v^m$ for an appropriate m such that $\deg(x+v^m)<\deg(x)$ and $(x+v^m)p(v)=p(v)(x+v^m)$. However, as in the proof of Theorem 3.1 and for future purposes, we prefer to present a different proof. For that proof we assume that $\deg(v)>1$.

As in Theorem 3.1, by our analysis of homogeneous elements, the top degree homogeneous part of x must be a monomial, which is a power of the top degree monomial in v.

As in the proof of Theorem 3.1, if $\deg(x)>\deg(u)$, then the arguments that were used in the proof of Lemma 3.2, which remain valid under the assumptions of the theorem, enable us to analyze the codegree 1 monomials in x. In that case, as in the proof of Theorem 3.1, there exists an element t_1 that contains a top degree monomial and a homogeneous part of codegree 1 such that $x=ut_1$ and $x=t_1u \mod G^{\deg(x)-2}$.

Applying these arguments iteratively, as in the proof of Theorem 3.1, we get that $x=tu^m \mod G^{\deg(x)-2}$ for some t that satisfies $\deg(t)\leq\deg(u)$, which means that $tu = ut \mod G^{\deg(tu)-2}$. In particular, the top degree monomial of t must be a power of the top degree monomial of v.

In case $\deg(t) \le \deg(u) \mod G^{\deg(tu)-2}$, we apply the same argument that we used in case $\deg(t) = \deg(u)$ in the proof of Theorem 3.1. By these arguments, if $u = v^b \mod G^{\deg(u)-2}$, then $t = v^s \mod G^{\deg(t)-2}$, where s is an integer, $1 \le s \le b$. This implies that $x = v^\ell \mod G^{\deg(x)-2}$ for some positive integer ℓ.

We continue in the same way as we did in proving Theorem 3.1. We use (a finite) induction and assume that $x = v^\ell \mod G^{\deg(x)-c}$ for some positive integer $c < \deg(v)$, i.e., we assume that the equality holds for all the monomials in x and v^ℓ of codegree smaller than c. To complete the proof of the theorem, we need to prove the same equality for all the monomials of codegree at most c.

By the inductive hypothesis, $x = v^\ell \mod G^{\deg(x)-c}$. Hence, $x = x_{c-1} + h$, where x_{c-1} is the sum of all the monomials of codegree smaller than c in x and $\deg(h) \le \deg(x) - c$. Furthermore, x_{c-1} is precisely the sum of all the monomials of codegree smaller than c in v^ℓ.

Let u_{c-1} be the sum of the monomials of codegree less than c in u, and let v_{c-1} be the sum of the monomials of codegree less than c in v. We set s_c to be the sum of all the monomials of codegree c in v_{c-1}^ℓ.

We have $u = p(v)$, so we set d_c to be the sum of all the codegree c monomials in $p(v_{c-1})$. By construction: $(u_{c-1} + d_c)(x_{c-1} + s_c) = (x_{c-1} + s_c)(u_{c-1} + d_c) = (v_{c-1})^{\ell+b} \mod G^{\deg(ux)-(c+1)}$, i.e., the monomials of codegree at most c are identical for the three different products.

Recall $x = x_{c-1} + h$, where $\deg(h) \le \deg(x) - c$. We set $x = x_{c-1} + s_c + r$, where $\deg(s_c) = \deg(x) - c$ and $\deg(r) \le \deg(x) - c$. Similarly, we set $u = u_{c-1} + d_c + q$, where $\deg(q) \le \deg(u) - c$. Then

$$\begin{aligned} ux &= (u_{c-1} + d_c + q)(x_{c-1} + s_c + r) \\ &= xu = (x_{c-1} + s_c + r)(u_{c-1} + d_c + q) \mod G^{\deg(xu)-(c+1)}. \end{aligned}$$

Since $(u_{c-1} + d_c)(x_{c-1} + s_c) = (x_{c-1} + s_c)(u_{c-1} + d_c)$ modulo the same group, it follows that $u_{\text{top}} r + q x_0 = r u_0 + x_0 q \mod G^{\deg(xu)-(c+1)}$, where u_0 and x_0 are the top monomials in u and x in correspondence. Therefore, all these monomials are products of a top degree monomial with a codegree c monomial, and these can be broken precisely as in the codegree 1 case, assuming $\deg(x) > \deg(u)$.

As in the codegree 1 case, we are left with the case in which $\deg(x) \le \deg(u)$. In that case we write $u = u_{c-1} + d_c + q_c$ and $x = x_{c-1} + s_c + r_c$ as above. By the same argument that was used in that case in analyzing the codegree 1 monomials, the monomials of codegree c in x that are contained in r_c are precisely the monomials of codegree c in $v^\ell + s_c$, and the induction follows for $c \le \deg(v)$. Hence, $x = v^\ell \mod G^{\deg(x)-\deg(v)}$. Since both x and v^ℓ commute with u, the sum $x + v^\ell$ commutes with u, and $\deg(x + v^\ell) < \deg(x)$, the theorem follows. □

It is possible to use the techniques that we used in this section to analyze centralizers of general elements with monomial top homogeneous part, and centralizers of general elements, but we won't need to apply these techniques in this generality in the sequel, so we omit these generalizations.

4. Equations with a single variable

In the previous section we gave combinatorial proofs to special cases of Bergman's theorem on the structure of centralizers in free associative algebras. Such combinatorial proofs are needed in order to study the set of solutions to related systems of equations that play a central role in understanding the set of solutions to a general monomial system of equations.

In this section we study the set of solutions to monomial systems of equations with a single variable. As will be demonstrated in the next paper in this sequence, the techniques that are used in this section play an essential role in studying monomial equations with no quadratic nor free parts. Note that in the general analysis of the set of homogeneous solutions to the homogeneous system of equations that is associated with the top level of a monomial system of equations, the Makanin–Razborov diagram of such a homogeneous system of equations, as it appears in the first section of this paper (that is based on [Sela 2016]), may contain quadratic and free parts.

Recall that over free groups and semigroups equations with a single variable were analyzed in [Lorenc 1963; Appel 1968] long before the analysis of general systems of equations. The approach that we use combines the technique and results for studying equations with a single variable over a free group and semigroup with the combinatorial approach that we used in analyzing centralizers.

Lemma 4.1. *Let $u, v \in FA$ and suppose that the top homogeneous parts of u and v are monomials with no periodicity* (*i.e.*, *the top monomials in u and v contain no subword α^2 for some nontrivial word α*).

If the equation $ux = xv$ has a nontrivial solution, *then the set of solutions to the equation $ux = xv$ is a set $\{wp(v)\}$*, *where $uw = wv$ and p is an arbitrary polynomial in a single variable. Furthermore*, *the element w*, *which is the solution of minimal degree of the equation*, *is unique.*

Proof. Recall that in a free semigroup, if u and v are nontrivial and have no periodicity, then the set of solutions to the equation $ux = xv$ is $\{w_0 v^m = u^m w_0\}$, where w_0 is a fixed element, m is an arbitrary nonnegative integer, and $\text{length}(w_0) \leq \text{length}(u)$.

Also, note that since we assumed that the top homogeneous parts of u and v are monomials, then the homogeneous equation that is associated with the highest degree parts in u, x, v implies that the highest degree part of x is a monomial that satisfies the same equation in the corresponding free semigroup.

The set of solutions of the equation $ux = xv$ is a linear subspace of FA. If w_1 and w_2 are solutions to the equation $ux = xv$, and they are of the same degree, then their top homogeneous monomials are identical. Hence, $w_1 + w_2$, which is also a solution of that equation, has strictly smaller degree than w_1 and w_2. Therefore, if the equation $ux = xv$ has a solution, then it has a unique solution of minimal degree that we denote w.

If x_0 is an arbitrary solution of $ux = xv$, then there exists some nonnegative integer b such that wv^b and x_0 have the same top monomial. Since both x_0 and wv^b are solutions of the equation $ux = xv$, the sum $x_0 + wv^b$ is a solution of this equation and $\deg(x_0 + wv^b) < \deg(x_0)$. Hence, the proof of the lemma follows by induction on the degree of the solution x_0. □

Unlike the case of free groups or semigroups, the equation $ux = xv$ may have a solution, and still it can be that there are no solutions with $\deg(x) \leq \deg(u) = \deg(v)$.

Let t, μ, and ρ be arbitrary elements in the algebra FA. Let $w = t\mu t\rho t\mu t$, $v = (\rho t\mu + \mu t\rho t\mu)t$ and $u = t(\mu t\rho + \mu t\rho t\mu)$. Then $uw = wv$, and in general there is no element $y \in FA$ such that $\deg(y) \leq \deg(u) = \deg(v)$ and $uy = yv$.

To bound the degree of a minimal degree solution we need the following lemma.

Lemma 4.2. *Let FA be the free associative algebra over GF_2 that is freely generated by k elements. Let $u, v \in FA$ be as in Lemma 4.1, and suppose that the equation $ux = xv$ has a nontrivial solution. Then there exists a solution w, $uw = wv$, with* $\deg(w) \leq \deg(u) \cdot (k^{\deg(u)} + 2)$.

Proof. Suppose that $x_1 \neq 0$ satisfies $ux_1 = x_1 v$. If $\deg(x_1) \leq \deg(u) \cdot (k^{\deg(u)} + 2)$ the lemma follows. Hence, we may assume that $\deg(x_1) > \deg(u) \cdot (k^{\deg(u)} + 2)$.

We use the analysis that was applied in analyzing centralizers in the previous section. By the analysis of homogeneous elements, the top degree homogeneous part of x_1 must be a monomial. Let u_0, v_0, and x_0 be the top monomials of u, v, and x_1. Then they must satisfy $u_0 x_0 = x_0 v_0$. Therefore, there exists a monomial z_0 such that $x_0 = u_0 z_0 = z_0 v_0$.

As in analyzing centralizers, we continue the analysis of x_1 by analyzing its codegree 1 monomials. We examine the codegree 1 monomials in the products ux_1 and $x_1 v$. By the proof of Lemma 3.2 we get an element z such that $x_1 = zv = uz \bmod G^{\deg(x_1)-2}$.

We continue iteratively by analyzing products of codegree 2 in the equality $ux_1 = x_1 v$, using the equality $x_1 = uz = zv$ for the top and codegree 1 parts.

Note that monomials of codegree 2 in the equality $ux_1 = x_1 v$ that are products of monomials of codegree 0 and 1 of u, v and z, that correspond to codegree 1 monomials of u and v and codegree 1 monomials of x_1 (from the two sides of the equation), cancel in pairs. The other codegree 2 monomials in the two sides of the equation are products that are obtained as one of the following:

(1) A product of a codegree 1 monomial of x_1 with a codegree 1 monomial of v.

(2) A product of a codegree 1 monomial of u with a codegree 1 monomial of x_1. In parts (1) and (2) we need to consider only monomials that do not cancel as products of top and codegree 1 monomials of u, v and z from the two sides of the equation.

(3) A product of a codegree 2 monomial of x_1 with the top monomial of v.

(4) A product of the top monomial of u with a codegree 2 monomial of x_1.

(5) A product of a codegree 2 monomial of u with the top monomial of x_1.

(6) A product of the top monomial of x_1 with a codegree 2 monomial of v.

Note that because of the equation, for each codegree 2 monomial in the products in the two sides of the equation, either 2, 4 or all 6 options occur. However, because the top monomials of u and v have no periodicity, possibilities (5) and (6) cannot occur together, so only 2 or 4 possibilities can occur.

Suppose that (1) occurs for some codegree 2 monomial. In that case we can assume that the codegree 1 monomial of x_1 is a product of the top monomial in u with a codegree 1 monomial of z, since otherwise such a product cancels with a product of type (2) by our analysis of codegree 1 products.

If in addition to (1) only (2) occurs for that codegree 2 monomial, we add a codegree 2 monomial to z that is obtained from the given codegree 2 monomial in the product by cutting a prefix which is equal to the top monomial of u, and a suffix which is the top monomial of v.

If only (3) occurs (in addition to (1)) for the given codegree 2 monomial, we also add a codegree 2 monomial to z that is identical to the one we added in case only (1) and (2) occur. If only (4) occurs, we do not add anything. If (5) occurs we add a codegree 2 monomial to z (the same codegree 2 monomial as in the previous cases).

If (6) occurs, we do not add anything.

Suppose that (2) occurs for some codegree 2 monomial. In that case we can assume that the codegree 1 monomial of x_1 is a product of a codegree 1 monomial of z with the top monomial of v. Hence, this can be dealt with precisely as what we did in (1).

Suppose that (3) occurs. If in addition only (4) occurs, we add a codegree 2 monomial to z. If only (5) occurs, we do not add anything. If only (6) occurs, we add a codegree 2 monomial to z. Suppose that (4) occurs. This can be dealt with precisely as the case in which (3) occurs. Again, since the top monomials of u and v do not have periodicity, (5) and (6) cannot occur together. Hence, we are only left with cases in which four of the possibilities occur.

Suppose that (1) and (2) occur for some codegree 2 monomial. In this case we can assume that (1) occurs as a product of a codegree 1 monomial of z and a codegree 1 of v, and (2) as a product of a codegree 1 monomial of u and codegree 1 monomial of z (otherwise (1) and (2) cancel from our analysis of codegree 1 monomials in x_1). If in addition only (3) and (4) occur, we do not add anything. If (3) and (5) occur, we add a codegree 2 monomial to z. If (3) and (6) occur, we do not add anything. The cases in which in addition to (1) and (2), cases (4) and (5) occur or cases (4) and (6) occur, are symmetric to (3) and (5) or (3) and (6).

Suppose that (1) occurs and (2) does not. Again, we may assume that in (1), it is a product of a codegree 1 monomial of z with a codegree 1 monomial of v. If in addition (3), (4) and (5) occur, we do not add anything. If (3), (4) and (6) occur, then we add a codegree 2 monomial to z. The case in which (1) does not occur and (2) occurs is symmetric.

After possibly adding codegree 2 monomials to z, the equation that was valid for codegree 1 products is now valid for codegree 2 products, i.e., $x_1 = uz = zv \mod G^{\deg(x_1)-3}$.

We continue iteratively to construct the element z by adding higher codegree monomials, so that the constructed element z satisfies the equation $x_1 = uz = zv$ for products of higher and higher codegree. Suppose that $x_1 = uz = zv \mod G^{\deg(x_1)-d}$, i.e., that the equation holds for all the products of codegree at most $d-1$, where d is a positive integer with $d \leq \deg(u)$. We iteratively add codegree d monomials to z so that the equalities hold for all codegree d products as well.

As in analyzing codegree 2 products, products of codegree d that include monomials of codegree smaller than d of u, v and z that correspond to smaller codegree monomials of x_1 (from the two sides of the equation) cancel in pairs.

The various cases are straightforward generalizations of the cases in analyzing codegree 2 products. Suppose that a codegree d product can be presented as either

(1) an odd number of products of the top monomial of u with codegree m_i monomials of z and codegree ℓ_i monomials of v, for some subset of tuples (m_i, ℓ_i), where $m_i + \ell_i = d$ and m_i, ℓ_i are positive integers, for every index i;

(2) an odd number of products of codegree s_j monomials of u with codegree t_j monomials of z and with the top monomial of v, for some subset of tuples (s_j, t_j), where $s_j + t_j = d$ and s_j, t_j are positive integers, for every index j;

(3) a product of a codegree d monomial of x_1 with the top monomial of v;

(4) a product of the top monomial of u with a codegree d monomial of x_1;

(5) a product of a codegree d monomial of u with the top monomial of x_1;

(6) a product of the top monomial of x_1 with a codegree d monomial of v.

Note that since we assumed that the top monomials of u and v do not have periodicity, cases (5) and (6) cannot occur together unless $d = \deg(u) = \deg(v)$. In case both (5) and (6) occur for $d = \deg(u)$, it must be that both u and v contain the constant monomial 1. We have $ux_1 = x_1v$ if and only if $(u+1)x_1 = x_1(v+1)$. Hence, in case both u and v contain the constant monomial 1, we replace them by $u+1$ and $v+1$. This does not change the set of solutions, and after the change, cases (5) and (6) do not occur together for all $2 \le d \le \deg(u)$.

The analysis of codegree d products, according to the various possibilities of subsets of the six cases, is identical to the analysis that was used to analyze codegree 2 products. According to the analysis we decide what codegree d monomials to add to the element z.

After possibly adding these codegree d monomials to z, the equation that was valid for all products up to codegree $d-1$ is now valid for codegree d products, i.e., $x_1 = uz = zv \mod G^{\deg(x_1)-(d+1)}$.

Finally, we get an element z that satisfies $x_1 = uz = zv \mod G^{\deg(x_1)-\deg(u)-1}$.

After possibly changing the elements u and v so that not both of them contain the constant monomial 1, we continue the analysis of codegree d products in the two sides of the equation $ux_1 = x_1v$, for all d, $\deg(u)+1 \le d \le \deg(x_1) - \deg(u)$, precisely as we analyzed codegree d products for $2 \le d \le \deg(u)$, and iteratively add codegree d monomials to the element z. Finally, we get an element z that satisfies $x_1 = uz = zv \mod G^{\deg(u)-1}$, i.e., the equalities hold for all products up to degree $\deg(u) = \deg(v)$.

We continue by looking at the equality $uz = zv \mod G^{\deg(u)-1}$. Repeating the same argument we can find an element z_2 such that $z = uz_2 = z_2v \mod G^{\deg(u)-1}$. Continuing inductively, we get an element z_{r+1} such that

$$z_r = uz_{r+1} = z_{r+1}v \mod G^{\deg(u)-1}.$$

We are working over the free associative algebra FA_k, i.e., the algebra is over GF_2 and it is freely generated by k elements. Hence, $G^{\deg(u)-1}$ as a vector space over GF_2 has dimension bounded by $k^{\deg(u)}$. Therefore, there exist elements of distinct degrees $\{s_m \mid m = 1, \ldots, k^{\deg(u)}+1\}$ (that are the elements z_r that were constructed iteratively from a given long solution) such that $\deg(s_m) \le (1+m)\deg(u)$ and $us_m = s_mv \mod G^{\deg(u)-1}$.

By a simple pigeonhole argument, there exists a subcollection of the indices $1 \le i_1 < \cdots < i_f \le k^{\deg(u)}+1$ such that $s = s_{i_1} + \cdots + s_{i_f}$ and $us = sv$. Hence, s is a solution of the given equation, and $\deg(s) \le \deg(u) \cdot (2 + k^{\deg(u)})$. □

So far we assumed that the top degree elements of u and v are monomials that are not proper powers. First, we omit the periodicity assumption, and allow the top degree monomials of u and v to have nontrivial roots.

Lemma 4.3. *With the notation of Lemma 4.2, let $u, v \in FA$ and suppose that the top homogeneous parts of u and v are monomials. Suppose that the top degree monomial of u has nontrivial roots of degree bounded by q.*

Suppose that the equation $ux = xv$ has a nontrivial solution. Then there exists elements $w_1, \dots, w_d$, $1 \le d \le q$, such that the set of solutions to the equation $ux = xv$ is a set of the form $\{w_1 p_1(v) + \cdots + w_d p_d(v)\}$, where $uw_i = w_i v$ and $p_1, \dots, p_d$ are arbitrary polynomials in v.

Proof. Let u_0 and v_0 be the top monomials of u and v, and let x_0 be the top monomial of a solution x. Let t_0 be a primitive root of v_0. Then there exists some fixed element s_0, $\deg(s_0) < \deg(t_0)$, such that $x_0 = s_0 t_0^m$ for some nonnegative integer m. Note that the element s_0 is fixed and does not depend on the solution x, since we assumed that u_0 and v_0 are proper powers of t_0, which is primitive, and if s_0 is not fixed, t_0 must have a proper root.

Suppose that $t_0^q = v_0$. The top monomial of a solution x is of the form $x_0 = s_0 t_0^m$, so we can divide the solutions x_1 according to the residue classes of the nonpositive integers m modulo q. For each residue class for which there is a solution, we fix one of the shortest solutions in the class. We denote these shortest solutions, $w_1, \dots, w_d$, for some d, $1 \le d \le q$.

Let x be a solution. x must be in the same class as one of the fixed shortest solutions w_i. Hence, for some nonnegative integer b, both x and $w_i v^b$ are solutions and they have the same top monomial. Therefore, if $x \ne w_i$, $x + w_i v^b$ is a nontrivial solution and $\deg(x + w_i v^b) < \deg(x)$. By a finite induction, $x = w_1 p_1(v) + \cdots + w_d p_d(v)$ for some polynomials $p_1, \dots, p_d$. □

So far we analyzed the equation $ux = xv$. We use similar methods to analyze the more general equation $u_1 x u_2 = v_1 x v_2$.

Theorem 4.4. *Let $u_1, u_2, v_1, v_2 \in FA$ and suppose that the top homogeneous parts of u_i and v_i are monomials with no periodicity (i.e., the top monomials in u_i and v_i contain no subwords α^2 for some nontrivial word α), and that $\deg(u_1) > \deg(v_1)$.*

Suppose that the equation $u_1 x u_2 = v_1 x v_2$ has a solution of degree bigger than $2(\deg(u_1) + \deg(v_2))^2$.

(1) *There exist elements $s, t \in FA$ such that $u_1 = v_1 s$ and $v_2 = t u_2$.*

(2) *An element $x \in FA$ is a solution to the equation $u_1 x u_2 = v_1 x v_2$ if and only if it is a solution of the equation $sx = xt$.*

Proof. First, note that if (1) is true and x satisfies $u_1 x u_2 = v_1 x v_2$, then we have $v_1 s x u_2 = v_1 x t u_2$. Hence, $sx = xt$. Conversely, every solution of the equation $sx = xt$ satisfies $u_1 x u_2 = v_1 x v_2$, so (2) is true.

As we did in analyzing centralizers and analyzing the equation $ux = xv$, we analyze the homogeneous parts in x and in u_i and v_i going from top to bottom.

Let u_i^0 and v_i^0, $i=1,2$, be the top monomials in u_i and v_i. Let x_1 be a solution of the equation $u_1xu_2=v_1xv_2$, and suppose that $\deg(x_1)>\max(\deg(u_1),\deg(v_2))+2(\deg(u_1)-\deg(v_1))$. By our analysis of homogeneous solutions, the top homogeneous part of the solution x_1 must be a monomial as well, which we denote x_0.

Since $u_1^0x_0u_2^0=v_1^0x_0v_2^0$, there exists monomials s_0, t_0, $\deg(s_0)=\deg(t_0)$, such that $u_1^0=v_1^0s_0$, $v_2^0=t_0u_2^0$ and $x_0=f_0t_0^b=s_0^be_0$, for some positive integer b, and $\deg(f_0)=\deg(e_0)<\deg(s_0)$.

We continue by analyzing monomials of codegree 1 in u_i, v_i and x_1. By the same analysis that was used in analyzing centralizers and in Lemma 4.2, there exist elements s,t with top monomials s_0 and t_0, and an element w with top monomial w_0, $w_0t_0=s_0w_0=x_0$, such that

(i) $u_1=v_1s \mod G^{\deg(u_1)-2}$.

(ii) $v_2=tu_2 \mod G^{\deg(v_2)-2}$.

(iii) $x_1=sw=wt \mod G^{\deg(x_1)-2}$.

By iteratively applying the same construction, the above three equalities imply that $w=s^mf=et^m \mod G^{\deg(w)-2}$, for some positive integer m and elements e, f with $\deg(s)\leq\deg(e)=\deg(f)<2\deg(s)$.

We continue by analyzing products of codegree 2. First, note that as in analyzing centralizers, if we look at codegree 2 products that involve only top monomials and codegree 1 monomials from s, t, u_i, v_i and w such that the products restrict to codegree 0 or 1 monomials of x_1, u_i and v_i, then such codegree 2 products cancel in pairs from the two sides of the equation.

We further look at codegree 2 products that contain a codegree 1 monomial of u_2. If the codegree 2 product contains the top monomial of t, then such a codegree 2 product cancels with a corresponding codegree 2 product from the other side of the equation, since all the corresponding monomials of u_i, v_i and x_1 (from the two sides of the equation) are either codegree 0 or codegree 1. Hence, we look at codegree 2 products that contain codegree 1 monomials of t and u_2, and, therefore, top monomials of u_1 and w. Such a codegree 2 product, which is a product of the top monomial of u_1, a codegree 1 monomial of x_1 and a codegree 1 monomial of u_2, cancels with either

(1) a product of the top monomial of v_1, the top monomial of x_1 and a codegree 2 monomial of v_2;

(2) a product of the top monomial of v_1, a codegree 1 monomial of x_1 and a codegree 1 monomial of v_2;

(3) a product of the top monomial of v_1, a codegree 2 monomial of x_1 and the top monomial of v_2;

(4) a product of the top monomial of u_1, a codegree 2 monomial of x_1 and the top monomial of u_2;

(5) a product of the top monomial of u_1, the top monomial of x_1 and a codegree 2 monomial of u_2;

(6) a product of a codegree 1 monomial of u_1, a codegree 1 monomial of x_1 and the top monomial of u_2.

If the given codegree 2 product cancels only with a product of type (1) we don't add anything to w nor to t. Suppose that the given codegree 2 product cancels only with a product of type (2). If the codegree 1 monomial of v_2 equals the product of the top monomial of t with a codegree 1 monomial of u_2, then the codegree 2 product of type (2) cancels with a codegree 2 product from the other side of the equation that contains only codegree 0 and 1 monomials of u_1, x_1 and u_2. Hence, in case (2) we can assume that the codegree 1 monomial of v_2 is a product of a codegree 1 monomial of t with the top monomial of u_2. In that case we add a codegree 2 monomial to t and leave w unchanged.

If the given codegree 2 product cancels only with a codegree 2 product of type (3), we add a codegree 2 monomial to t and a codegree 2 monomial to w. If the given codegree 2 product cancels only with a product of type (4) we add a codegree 2 monomial to t. In case the given product cancels only with a codegree 2 product of type (5) we don't add anything (apart from the codegree 2 monomial of u_2). If it cancels only with a codegree product of type (6) we add a codegree 2 monomial to t and a codegree 2 monomial to w.

Because the top monomial of v_2 does not have periodicity, a product of type (5) cannot cancel with a product of type (3) nor (6). Hence, if five possibilities occur in addition to the given one, it must be (1)–(4) and (6). In that case, we do not add anything.

Hence, the only left possibilities are a collection of products of three different types that cancel with the given codegree 2 product. We list the various possibilities for the collections of codegree 2 products of three different types that cancel with the given codegree 2 product and indicate what we add in each possibility:

(i) Products (1)–(3) cancel. We add a codegree 2 monomial to w, apart from an existing codegree 2 monomial of v_2 (that is equal to the products of the codegree 1 monomials of t and u_2 in the given codegree 2 product).

(ii) Products (1), (2) and (4) cancel. In that case we don't add anything to w and t. A monomial of codegree 2 already appears in v_2, and is equal to the product of the given codegree 1 monomials of t and u_2.

(iii) Products (1), (2) and (5) cancel. We add a codegree 2 monomial to t, in addition to the codegree 2 monomials that already appear in u_2 and v_2.

(iv) Products (1), (2) and (6) cancel. Product (1) cancels with the given codegree 2 product. We add a codegree 2 monomial to w.

(v) Products (1), (3) and (4) cancel. We add a codegree 2 monomial to w, and the existing codegree 2 monomial to v_2.

(vi) Products (1), (3) and (6) cancel. Products (3) and (6) cancel, so this is identical to the case that only (1) occurs in addition to the given codegree 2 product.

(vii) Products (2), (3) and (4) cancel. We add a codegree 2 monomial to w and a codegree 2 monomial to t.

(viii) Products (2), (3) and (6) cancel. Like (vi), (3) and (6) cancel.

(ix) Products (1), (4) and (5) cancel. We add a codegree 2 monomial to t, and the existing codegree 2 monomials to u_2 and v_2.

(x) Products (1), (4) and (6) cancel. Product (1) cancels with the given codegree 2 product. We add a codegree 2 monomial to w.

(xi) Products (2), (4) and (5) cancel. We just add the existing codegree 2 monomial to u_2.

(xii) Products (2), (4) and (6) cancel. Like (vi), (2) and (4) cancel.

(xiii) Products (3), (4) and (6) cancel. Like (vi), (3) and (6) cancel.

So far we analyze codegree 2 products that cancel with a given codegree 2 product that is a product of the top monomial of u_1 and w and codegree 1 monomials of t and u_2. The same analysis is valid for codegree 2 products that cancel with a codegree 2 product of type (1), and the analogous cases from the left side of the equation.

We continue by analyzing case (2), i.e., those codegree 2 products that cancel with a given codegree 2 product of the top monomial of v_1, a codegree 1 monomial of x_1 that is a product of the top monomial of s and a codegree 1 monomial of w, and a codegree 1 monomial of v_2 that is equal to a product of a codegree 1 monomial of t with the top monomial of u_2. Such a given codegree 2 product can cancel with either

(1) a product of the top monomial of v_1, a codegree 2 monomial of x_1 and the top monomial of v_2;

(2) a product of the top monomial of v_1, the top monomial of x_1 and a codegree 2 monomial of v_2;

(3) a product of the top monomial of u_1, a codegree 2 monomial of x_1 and the top monomial of u_2;

(4) a product of the top monomial of u_1, the top monomial of x_1 and a codegree 2 monomial of u_2;

(5) a product of the top monomial of u_1, a codegree 1 monomial of x_1 and a codegree 1 monomial of u_2;

(6) a product of a codegree 1 monomial of u_1, which is a product of the top monomial of v_1 with a codegree 1 monomial of s, with a codegree 1 monomial of x_1, which is the product of a codegree 1 monomial of w with the top monomial of t, and with the top monomial of u_2.

If the given codegree 2 product equals only to a codegree 2 product of type (1), we add a codegree 2 monomial to w. If it equals only to a codegree 2 product of type (2), we add a codegree 2 monomial to t, apart from the existing codegree 2 monomial of v_2. If it equals only to a codegree 2 product of type (3), we do not add anything. If it equals only to a codegree 2 product of type (4) we add a codegree 2 monomial to t, apart from the existing codegree 2 monomial of u_2. We analyzed case (5) with all the codegree 2 products that it cancels with previously. If it equals only to a product of type (6) we add a codegree 2 monomial to w.

Case (5) was analyzed previously, so we can assume it does not occur. A codegree 2 product of types (1) or (6) cannot cancel with a codegree 2 product of type (4). A monomial of type (5) that cancels with a monomial of type (6) is a product of lower codegree monomials of u_i, v_i and x_1 from the two sides of the equation, so we omit this case. Hence, there are 5 cases left:

(i) Products (1), (2) and (3) cancel with the given codegree 2 product. In that case we add a codegree 2 monomial to w and a codegree 2 monomial to t, apart from the existing codegree 2 monomial of v_2.

(ii) Products (1), (2) and (6) cancel. We add a codegree 2 monomial to t.

(iii) Products (1), (3) and (6) cancel. In that case we do not add anything.

(iv) Products (2), (3) and (4) cancel. In that case we only add the already existing codegree 2 monomials of u_2 and v_2.

(v) Products (2), (3) and (6) cancel. We add a codegree 2 monomial to w and t.

Codegree 2 products that contain codegree 1 monomials of v_1 or u_1 are treated exactly in the same way. Hence, we are left with sets of codegree 2 products that cancel, and each of these codegree 2 products is a product of top monomials with codegree 2 monomials of one of the u_i, v_i or x_1. These are analyzed precisely as they are treated in the proof of Theorem 4.4 and in analyzing codegree 1 products, and in each such cancellation codegree 2 monomials may be added to either s, t or w, apart from existing codegree 2 monomials of u_i and v_i. Finally, we (possibly) added codegree 2 monomials to s, w and t, such that

(i) $u_1 = v_1 s \mod G^{\deg(u_1)-3}$.

(ii) $v_2 = t u_2 \mod G^{\deg(v_2)-3}$.

(iii) $x_1 = sw = wt \mod G^{\deg(x_1)-3}$.

We continue iteratively with products with higher codegree. Let

$$d = \min(\deg(v_1), \deg(u_2), \deg(u_1) - \deg(v_1)).$$

Let $r \leq d-1$ and suppose that we added codegree r monomials to s, w and t such that the equations above hold for all products of codegree bounded by $r-1$.

We analyze codegree r products in the same way we analyzed codegree 2 products. First, note that if a codegree r product is a product of monomials of u_i, v_i, s, t and w that correspond to products of monomials of codegree smaller than r of u_i, v_i and x_1 from the two sides of the equation, then such codegree r products cancel in pairs.

Suppose that a codegree r product is a product of the top monomials of u_1 and w, and monomials of codegree q_i of t and codegree m_i of u_2, such that $q_i + m_i = r$ and q_i, m_i are positive integers, and there are odd number of such pairs (q_i, m_i). We treat this case in the same way we treated the case of a codegree 2 product that includes a codegree 1 monomial of t and a codegree 1 monomial of u_2. This odd set of codegree r products (that are all equal) cancels with either

(1) a product of the top monomial of v_1, the top monomial of x_1 and a codegree r monomial of v_2;

(2) an odd set of codegree r products of the top monomial of v_1, codegree e_j monomial of x_1 and codegree p_j monomial of v_2, for some positive set of pairs (e_j, p_j) that satisfy $e_j + p_j = r$, and such that the codegree p_j monomial of v_2 is a product of a codegree p_j monomial of t with the top monomial of u_2;

(3) a product of the top monomial of v_1, a codegree r monomial of x_1 and the top monomial of v_2;

(4) a product of the top monomial of u_1, a codegree r monomial of x_1 and the top monomial of u_2;

(5) a product of the top monomial of u_1, the top monomial of x_1 and a codegree r monomial of u_2;

(6) an odd set of codegree r products of a codegree a_j monomial of u_1 and a codegree b_j monomial of x_1 with the top monomial of u_2, for some positive set of pairs (a_j, b_j) that satisfy $a_j + b_j = r$, and such that the codegree a_j monomial of u_1 is a product of the top monomial of v_1 and a codegree a_j monomial of s.

The treatment of the various cases is identical to what we did in analyzing codegree 2 products (cases (i)–(xiii)), just that instead of adding codegree 2 monomials to the various elements we add codegree r monomials. The other cases of codegree r

products are treated exactly as we treated codegree 2 products. Therefore, we constructed elements s, t, w for which

(i) $u_1 = v_1 s \mod G^{\deg(u_1)-d}$.

(ii) $v_2 = t u_2 \mod G^{\deg(v_2)-d}$.

(iii) $x_1 = sw = wt \mod G^{\deg(x_1)-d}$.

We divide the continuation according to the minimum between $\deg(v_1)$, $\deg(u_2)$ and $\deg(u_1) - \deg(v_1)$. First we assume that

$$d = \min\bigl(\deg(v_1), \deg(u_2), \deg(u_1) - \deg(v_1)\bigr) = \deg(u_1) - \deg(v_1).$$

In analyzing codegree d products, there are special codegree d products that we need to single out and treat separately, as they may involve cancellations between codegree d products that contain codegree d monomials of u_1 or v_1 and those that contain codegree d monomials of u_2 or v_2.

As in analyzing smaller codegree products, note that codegree d products that are products of smaller codegree monomials of the u_i, v_i, s, w and t, and correspond to smaller codegree monomials of u_i, v_i and x_1 from the two sides of the equation cancel in pairs.

We continue by analyzing codegree d products that are products of top degree monomials of u_1 and w, codegree q_i monomials of t and codegree m_i monomials of u_2, such that q_i and m_i are positive and $q_i + m_i = d$, there are odd number of such pairs (q_i, m_i), and the product of these monomials of t and u_2 is not equal to u_2^0, the top monomial of u_2.

Such codegree d products are analyzed exactly in the same way they were analyzed in codegree r products for $r < d$. Similarly, we analyze codegree d products that are obtained an odd number of times as the product of the top monomials of v_1 and s_1, a codegree e_j monomial of w and a codegree p_j monomial of v_2, such that the product of a codegree e_j monomial of w and a codegree p_j monomial of v_2 is not $w_0 u_2^0$, i.e., the product of the top monomials of w and u_2.

In a similar way we analyze codegree d products that are products of smaller codegree monomials of u_1 and w and the top monomials of t and u_2, and products of smaller codegree monomials of v_1 and s and the top monomials of w and v_2, assuming the products of these smaller degree monomials are not equal to $v_1^0 w_0$ or to v_1^0.

We continue by analyzing canceling pairs of codegree d products that are products of top monomials of v_i, u_i and x_1, with one codegree d monomial of these elements, such that this codegree d monomial of u_1 is not v_1^0, the codegree d monomial of x_1 is not w_0 and the codegree d monomial of v_2 is not u_2^0. These codegree d products are analyzed in the same way they were analyzed for smaller codegree products.

We are left with codegree d products that are either

(1) a product of a codegree d monomial of u_1 that is equal to v_1^0, with top monomials of x_1 and u_2;

(2) a product of the top monomial of v_1 with top monomials of x_1 and a codegree d monomial of v_2 that is equal to u_2^0;

(3) a product of the top monomial of u_1 with a codegree d monomial of x_1 that is equal to the top monomial of w, with the top monomial of u_2;

(4) a product of the top monomial of v_1 with a codegree d monomial of x_1 that is equal to the top monomial of w, with the top monomial of v_2;

(5) an odd set of codegree d products of the top monomial of v_1, codegree e_j monomial of x_1 and codegree p_j monomial of v_2, for some positive pairs (e_j, p_j) that satisfy $e_j + p_j = d$, such that the monomial of x_1 is a product of the top monomial of s with a codegree e_j monomial of w, and the product of each codegree e_j monomial of x_1 with a codegree p_j monomial of v_2 is equal to $w_0 v_2^0$ (the product of the top monomials of w and v_2);

(6) an odd set of codegree d products of the top monomial of u_1, codegree q_i monomials of x_1 that are products of the top monomial of w with a codegree q_i monomial of t, with codegrees m_i monomials of u_2, for some positive set of pairs (q_i, m_i) that satisfy $q_i + m_i = d$, such that the product of each codegree q_i monomial of t with a codegree m_i monomial of u_2 is equal to u_2^0;

(7) an odd set of codegree d products of codegree f_i monomials of u_1 and codegree g_i monomials of x_1 with the top monomial of u_2, for some positive pairs (f_i, g_i) that satisfy $f_i + g_i = d$, such that the monomial of x_1 is a product of a codegree g_i monomial of w with the top monomial of t, and the product of each codegree f_i monomial of u_1 with a codegree g_i monomial of x_1 is equal to $u_1^0 w_0$ (the product of the top monomials of u_1 and w);

(8) an odd set of codegree d products of codegree h_j monomials of v_1 and codegree k_j monomials of x_1 that are products of a codegree k_j monomial of s with the top monomial of w, with the top monomial of v_2, for some positive pairs (h_j, k_j) that satisfy $h_j + k_j = d$, such that the product of each codegree h_j monomial of v_1 with a codegree k_j monomial of s is equal to v_1^0.

First note that (3) exists if and only if (4) exists and they cancel each other. If (3) and (4) are the only existing possibilities, we add a codegree d monomial to w, which is the codegree d prefix or suffix of the top monomial of w. Also note that if cases (1) or (2) exist, codegree d monomials that already appear in u_1 or v_2 are added to them. Suppose that only two of the possibilities (1), (2) and (5)–(8) exist, possibly in addition to (3) and (4). We go over the various alternatives:

(i) If only (1) and (2) exist, we add the constant element 1 to s and t, and the codegree d prefix of w_0 to w, where w_0 is the top monomial of w. If (3) and (4) exists as well, we only add 1 to s and t.

(ii) If only (5) and (6) exist, we just add 1 to t. If (3) and (4) exist as well, we add the codegree d prefix of w_0 to w. The case in which only (7) and (8) exist is identical.

(iii) If only (5) and (8) exist, we add 1 to s and the codegree d prefix of w_0 to w. If (3) and (4) exist as well, we only add 1 to s. The case in which only (6) and (7) exist is treated identically.

(iv) If only (5) and (7) exist, we add the codegree d prefix of w_0 to w. If (3) and (4) exist as well, we do not add anything to any of the variables.

(v) If only (6) and (8) exist, we add 1 to s and t and the codegree d prefix of w_0 to w. If (3) and (4) exist as well, we only add 1 to s and t.

(vi) If only (1) and (8) exist, we add 1 to s and the codegree d prefix of w_0 to w. If (3) and (4) exist as well, we only add 1 to s.

(vii) If only (1) and (7) exist, we do the same as in (v), adding 1 to s and the codegree d prefix of w_0 to w. If (3) and (4) exist as well, we only add 1 to s.

(viii) If only (1) and (6) exist, we add 1 to s and t, and the codegree d prefix of w_0 to w. If (3) and (4) exist as well, we only add 1 to s and t.

(ix) If only (1) and (5) exist, we add 1 to s and the codegree d prefix of w_0 to w. If (3) and (4) exist as well, we only add 1 to s.

The cases in which only case (2) and one of the cases (5)–(8) exist are treated according to cases (vi)–(ix). Suppose that exactly four of the cases (1), (2) and (5)–(8) exist, possibly in addition to (3) and (4). We go over the alternatives:

(i) If only (1), (2), (5) and (6) exist, we add 1 to s and the codegree d prefix of w_0 to w. If (3) and (4) exist as well, we just add 1 to s. The case in which only (1), (2), (7) and (8) exist is identical.

(ii) If only (1), (2), (5) and (8) exist, we add 1 to t. If (3) and (4) exist as well, we add 1 to t and the codegree d prefix of w_0 to w. The case in which only (1), (2), (6) and (7) exist is identical.

(iii) If only (1), (2), (5) and (7) exist, we add 1 to s and t. If (3) and (4) exist as well, we add 1 to s and t and the codegree d prefix of w_0 to w.

(iv) If only (1), (2), (6) and (8) exist, we do not change any of the variables. If (3) and (4) exist as well, we add the codegree d prefix of w_0 to w.

(v) If only (5), (6), (7) and (8) exist, we add 1 to s and t. If (3) and (4) exist as well, we add 1 to s and t and the codegree d prefix of w_0 to w.

(vi) If only (1), (5), (6) and (7) exist, we add 1 to s and t. If (3) and (4) exist as well, we add 1 to s and t and the codegree d prefix of w_0 to w.

(vii) If only (1), (5), (6) and (8) exist, we add 1 to t. If (3) and (4) exist as well, we add 1 to t and the codegree d prefix of w_0 to w.

(viii) If only (1), (5), (7) and (8) exist, we add the codegree d prefix of w_0 to w. If (3) and (4) exist as well, we do not add anything to any of the variables.

(ix) If only (1), (6), (7) and (8) exist, we add 1 to t and the codegree d prefix of w_0 to w. If (3) and (4) exist as well, we just add 1 to t.

The cases in which only case (2) and three of the cases (5)–(8) exist are treated according to cases (vi)–(ix). Suppose that cases (1), (2) and (5)–(8) exist. In that case we add the codegree d prefix of w_0 to w. If cases (3) and (4) exist as well, we do not add anything to any of the variables.

This completes the analysis of codegree d products. We continue with the analysis of codegree $d+1$ products. First, as in analyzing smaller codegree products, codegree $d+1$ products that are products of smaller codegree monomials of u_i, v_i, s, t, w, that correspond to products of smaller degree monomials of u_i, v_i and x_1 from the two sides of the equation, cancel in pairs.

Lemma 4.5. *Suppose that a codegree $d+1$ product is a product of the top monomials of u_1 and w, and monomials of codegree q of t and codegree m of u_2, such that $q \geq 0$, $m > 0$ and $q+m = d+1$. Such a codegree $d+1$ product cannot be*

(1) *a product of the top monomial of v_1, a codegree e monomial of x_1 and a codegree f monomial of v_2, for $e > 0$ and $f \geq 0$, that satisfy $e+f = d+1$, and the codegree f monomial of v_2 is a product of a codegree f monomial of t with the top monomial u_2^0 of u_2;*

(2) *a product of the top monomial of u_1, a codegree $d+1$ monomial of x_1 and the top monomial of u_2.*

Proof. If such a codegree $d+1$ product can be presented as a product in the forms (1) or (2), u_2^0 has a prefix which is a suffix of t_0. Hence, v_2^0 has nontrivial periodicity that contradicts our assumptions. □

Suppose that a codegree $d+1$ product can be presented as a product of the top monomials of u_1 and w, and monomials of codegrees q_i of t and codegree m_i of u_2, such that $q_i \geq 0$, $m_i > 0$ and $q_i + m_i = d+1$, and there are an odd number of such pairs (q_i, m_i).

By Lemma 4.5, the same codegree $d+1$ product is the product of the top monomial of v_1, the top monomial of x_1 and a codegree $d+1$ monomial of v_2.

Furthermore, by the same argument that was used in the proof of Lemma 4.5, if a codegree $d+1$ product is the product of the top monomials of v_1 and x_1 and a codegree $d+1$ monomial of v_2, then it must be the product of an odd number of products of the top monomials of u_1 and w, and monomials of codegrees q_i of t and codegrees m_i of u_2, such that $q_i \geq 0$, $m_i > 0$ and $q_i + m_i = d+1$.

Lemma 4.6. *Suppose that a codegree $d+1$ product can be presented in an odd number of ways as products of the top monomial of v_1, a codegree e_j monomial of x_1, which is the product of the top monomial of s with a codegree e_j monomial of w, and a codegree f_j monomial of v_2, for positive e_j and f_j and $e_j + f_j = d+1$, and the codegree f_j monomials of v_2 are products of codegree f_j monomials of t with the top monomial u_2^0 of u_2.*

(1) *Suppose that this codegree $d+1$ product cannot be presented in an odd number of ways as products of a codegree g_j monomial of u_1, and a codegree h_j monomial of x_1 with the top monomial of u_2, where the codegree h_j monomial of x_1 is a codegree h_j monomial of w with the top monomial of t, for positive g_j and h_j and $g_j + h_j = d+1$, and the codegree g_j monomials of u_1 are products of v_1^0 with codegree g_j monomials of t.*

Then the same codegree $d+1$ product is either the product of the top monomial of u_1, a codegree $d+1$ monomial of x_1, and the top monomial of u_2, or the product of the top monomial of v_1, a codegree $d+1$ monomial of x_1 and the top monomial of v_2, and exactly one of the two occurs.

(2) *Suppose that this codegree $d+1$ product can be presented in an odd number of ways as products of a codegree g_j monomial of u_1, and a codegree h_j monomial of x_1 with the top monomial of u_2, and the codegree h_j monomial of x_1 is a codegree h_j monomial of w with the top monomial of t, for positive g_j and h_j and $g_j + h_j = d+1$, and the codegree g_j monomials of u_1 are products of v_1^0 with codegree g_j monomials of t.*

Then either the same codegree $d+1$ product is both the product of the top monomial of u_1, a codegree $d+1$ monomial of x_1, and the top monomial of u_2, and the product of the top monomial of v_1, a codegree $d+1$ monomial of x_1 and the top monomial of v_2, or none of these two possibilities occur.

Proof. Such a codegree $d+1$ product does not cancel only with the product of monomials of codegree less than d of u_i, v_i and x_1, from the two sides of the equation. By Lemma 4.5 such codegree $d+1$ products cannot be equal to the following products:

(1) top degree monomials of u_1 and w with monomials of t and u_2;

(2) top degree monomials of v_1 and x_1 and a codegree $d+1$ monomial of v_2;

(3) monomials of v_1 and s with top degree monomials of w and v_2;

(4) codegree $d+1$ monomials of u_1 with top degree monomials of x_1 and of u_2.

Therefore, such a codegree $d+1$ product must be equal to an odd number of products in the forms that are listed in the statement of the lemma. □

Lemmas 4.5 and 4.6 enable us to treat codegree $d+1$ products in a similar way to the analysis of codegree r products for $r<d$.

Suppose that a codegree $d+1$ product is obtained in an odd number of ways as the product of the top monomials of u_1 and x_1 with a codegree q_i monomial of t, and a codegree m_i monomial of u_2, such that $q_i \geq 0$, $m_i > 0$ and $q_i+m_i=d+1$. By Lemma 4.5, such a product must be equal to a product of the top monomials of v_1 and x_1 and a codegree $d+1$ monomial of v_2.

An analogous conclusion holds if a codegree $d+1$ product is obtained in an odd number of ways as the product of a codegree m_i monomial of v_1 with a codegree q_i of s with the top monomials of x_1 and v_2, such that $q_i \geq 0$, $m_i > 0$ and $q_i+m_i=d+1$.

Suppose that a codegree $d+1$ product can be presented in an odd number of ways as products of the top monomial of v_1, a codegree e_j monomial of x_1, which is the product of the top monomial of s with a codegree e_j monomial of w, and a codegree f_j monomial of v_2, for positive e_j and f_j and $e_j+p_j=d+1$, and the codegree f_j monomials of v_2 are products of codegree f_j monomials of t with the top monomial u_2^0 of u_2.

Suppose that this codegree $d+1$ product cannot be presented in an odd number of ways as products of a codegree g_j monomial of u_1, and a codegree h_j monomial of x_1, which is a codegree h_j monomial of w with the top monomial of t, for positive g_j and h_j and $g_j+h_j=d+1$, and the codegree g_j monomials of u_1 are products of v_1^0 with codegree g_j monomials of t.

If the same codegree $d+1$ product is the product of the top monomial of u_1, a codegree $d+1$ monomial of x_1, and the top monomial of u_2, we do not add anything. If it is the product of the top monomial of v_1, a codegree $d+1$ monomial of x_1 and the top monomial of v_2, we add a codegree $d+1$ monomial to w.

Suppose that this codegree $d+1$ product can be presented in an odd number of ways as products of a codegree g_j monomial of u_1, and a codegree h_j monomial of x_1, which is a codegree h_j monomial of w with the top monomial of t, for positive g_j and h_j and $g_j+h_j=d+1$, and the codegree g_j monomials of u_1 are products of v_1^0 with codegree g_j monomials of t.

If the same codegree $d+1$ product is both the product of the top monomial of u_1, a codegree $d+1$ monomial of x_1, and the top monomial of u_2, and the product of the top monomial of v_1, a codegree $d+1$ monomial of x_1 and the top monomial of v_2, then we do not add anything. If none of these two possibilities occur, we add a codegree $d+1$ monomial to w (by Lemma 4.6 either both or none occur).

Suppose that a codegree $d+1$ product can be presented only as the product of the top monomial of u_1, a codegree $d+1$ monomial of x_1 and the top monomial of u_2, and as the product of the top monomial of v_1, a codegree $d+1$ monomial of x_1, and the top monomial of v_2. In that case we add a codegree $d+1$ monomial to w.

This concludes the analysis of codegree $d+1$ products. The analysis of codegree $d+r$ products, $r<d$, is identical to the analysis of codegree $d+1$ products. Hence, we (possibly) finally add codegree $d+r$ monomials to w, and the existing codegree $d+r$ monomials to u_i and v_i for $1 \le r < d$, and do not change s and t, such that

(i) $u_1 = v_1 s \mod G^{\deg(u_1)-2d}$.

(ii) $v_2 = t u_2 \mod G^{\deg(v_2)-2d}$.

(iii) $x_1 = sw = wt \mod G^{\deg(x_1)-2d}$.

In analyzing codegree $2d$ products, as in analyzing codegree d products, there are special codegree $2d$ monomials that we need to single out and treat separately, as they may involve cancellations between codegree $2d$ products that contain codegree d or $2d$ monomials of u_1 or v_1 and those that contain codegree d or $2d$ monomials of u_2 or v_2.

As in analyzing smaller codegree products, note that codegree $2d$ products that are products of smaller codegree monomials of the u_i, v_i, s, w and t, and correspond to smaller codegree monomials of u_i, v_i and x_1 from the two sides of the equation cancel in pairs.

As in analyzing codegree d products, we continue by analyzing codegree $2d$ products that are products of top degree monomials of u_1 and w, codegree q_i monomials of t and codegree m_i monomials of u_2, such that $q_i + m_i = 2d$, there are an odd number of such pairs (q_i, m_i), and the product of these monomials of t and u_2 is not equal to a codegree d suffix of u_2^0, the top monomial of u_2 (which is a codegree $2d$ suffix of v_0^2, the top monomial of v_2). Such codegree $2d$ products must cancel with the product of the top monomials of v_1 and x_1 and a codegree $2d$ monomial of v_2. In this case we only add the already existing codegree $2d$ monomial to v_2.

Similarly, we analyze codegree $2d$ products that are obtained in an odd number of ways as the product of the top monomials of v_1 and s_1, a codegree e_j monomial of w and a codegree p_j monomial of v_2, such that the product of a codegree e_j monomial of w and a codegree p_j monomial of v_2 does not have a suffix which is the codegree d suffix of u_2^0. We analyze codegree $2d$ products that contain similar monomials of v_1, u_1 and x_1 in a similar way.

Suppose that a codegree $2d$ product is obtained in an odd number of ways as the product of the top monomials of u_1 and x_1, a codegree q_i monomial of t and a codegree m_i monomial of u_2, such that q_i and m_i are positive and $q_i + m_i = 2d$, and the product of the monomial of t and the monomial of u_2 is a codegree d monomial of u_2^0.

Because we assumed that the coefficients do not have any periodicity, such a codegree $2d$ product must cancel with either a product of the top monomials of u_1 and x_1 and a codegree $2d$ monomial of u_2, or a product of the top monomials of v_1 and x_1 and a codegree $2d$ monomial of v_2. In both of these cases we only add the already existing codegree $2d$ monomials to u_2 or v_2.

If x_1 contains a monomial that is equal to the $2d$ prefix (or suffix) of x_1^0, then the codegree $2d$ product that contains the top monomials of the u_i and this codegree $2d$ monomial of x_1 cancels with the codegree $2d$ product of the top monomials of the v_i with that codegree $2d$ monomial of x_1.

As in analyzing codegree d products, we continue by analyzing canceling pairs of codegree $2d$ products that are products of top monomials of v_i, u_i and x_1, with one codegree $2d$ monomials of these elements, such that this codegree $2d$ monomial of u_1 is not the codegree d prefix of v_1^0, the codegree $2d$ monomial of x_1 is not the codegree d prefix (or suffix) of w_0, and the codegree $2d$ monomial of v_2 is not the codegree d suffix of u_2^0. These codegree $2d$ products are analyzed in the same way they were analyzed for smaller codegree products.

We are left with codegree $2d$ products that are either

(1) a product of the top monomial of u_1 with a codegree $2d$ monomial of x_1 that is equal to a codegree d prefix (or suffix) of the top monomial of w with the top monomial of u_2;

(2) a product of the top monomial of v_1 with a codegree $2d$ monomial of x_1 that is equal to the codegree d prefix of the top monomial of w with the top monomial of v_2;

(3) an odd set of codegree $2d$ products of the top monomial of v_1, codegree e_j monomial of x_1 and codegree p_j monomial of v_2, for some positive pairs (e_j, p_j) that satisfy $e_j + p_j = 2d$, such that the codegree e_j monomial of x_1 is the product of the top monomial of s with a codegree e_j monomial of w, and the product of each codegree e_j monomial of x_1 with a codegree p_j monomial of v_2 is equal to the product of a codegree d prefix of w_0 with v_2^0;

(4) an odd set of codegree $2d$ products of codegree f_i monomials of u_1 and codegree g_i monomials of x_1 with the top monomial of u_2, for some positive pairs (f_i, g_i) that satisfy $f_i + g_i = 2d$, such that the codegree g_i monomial of x_1 is the product of a codegree g_i monomial of w with the top monomial of t, and the product of each codegree f_i monomial of u_1 with a codegree g_i monomial of x_1 is equal to the product of u_1^0 with the codegree d prefix of w_0.

Note that (1) exists if and only if (2) exists and they cancel each other. If (1) and (2) are the only existing possibilities, we add a codegree $2d$ monomial to w, which is the codegree d prefix or suffix of the top monomial of w.

If only possibilities (3) and (4) exist, we add the codegree $2d$ prefix of w_0 to w. If (1)–(4) do all exist, w remains unchanged.

This completes the analysis of codegree $2d$ products. Codegree $2d+r$ products, for $1 \le r < d$, are treated in the same way we treated codegree $d+r$ products. Codegree $3d$ products are treated in the same we treated $2d$ products, and so on. Finally, in case $d = \deg(u_1) - \deg(v_1) = \deg(v_2) - \deg(u_2)$, we obtained the conclusion of the theorem.

Suppose that $d = \min(\deg(v_1), \deg(u_2), \deg(u_1) - \deg(v_1)) = \deg(v_1)$. In that case we continue the analysis of codegree r homogeneous parts in u_i, v_i and x_1, $d \le r < \deg(u_1) - \deg(v_1)$, precisely as we analyzed the codegree r homogeneous parts for $1 \le r \le d-1$. For $r = \deg(u_1) - \deg(v_1)$, we use the same analysis that we apply for codegree d products in case $d = \deg(u_1) - \deg(v_1)$. For $r > \deg(u_1) - \deg(v_1)$, we continue the analysis of codegree r homogeneous parts according to the analysis of codegree higher than d in case $d = \deg(u_1) - \deg(v_1)$. The analysis in the case $d = \min(\deg(v_1), \deg(u_2), \deg(u_1) - \deg(v_1)) = \deg(u_2)$ is identical. □

Theorem 4.4 reduces the analysis of solutions to the equation $u_1xu_2 = v_1xv_2$ to the equation $xt = sx$, in case the equation $u_1xu_2 = v_1xv_2$ has a long enough solution, and the coefficients have no periodicity. The same techniques allow one to reduce a general equation with one variable, in case the coefficients have no periodicity.

Theorem 4.7. *Let FA be the free associative algebra over GF_2 that is freely generated by k elements. Let $u_1, \dots, u_n, v_1, \dots, v_n \in FA$ and suppose that the top homogeneous parts of u_i and v_i are monomials with no periodicity, and that for at least one index i, $1 \le i \le n$, $u_i \ne v_i$. Suppose that the equation*

$$u_1xu_2xu_3 \cdots u_{n-1}xu_n = v_1xv_2xv_3 \cdots v_{n-1}xv_n$$

has a solution x_1 of degree bigger than $2(\deg(u_1) + \cdots + \deg(u_n))^2$.

By Section 1, the top homogeneous part of the solution x_1 has to be a monomial x_1^0, and x_1^0 has to satisfy an equation in a free semigroup

$$u_1^0x_1^0u_2^0x_1^0u_3^0 \cdots u_{n-1}^0x_1^0u_n^0 = v_1^0x_1^0v_2^0x_1^0v_3^0 \cdots v_{n-1}^0x_1^0v_n^0,$$

where u_i^0 and v_i^0 are the top monomials in u_i and v_i. Every solution of this equation is semiperiodic, i.e., has to be of the form $r_0w_0^m$, where $\text{length}(r_0) < \text{length}(w_0)$ *and w_0 is primitive. We say that w_0 is the period of x_0, and we further assume that* $\text{length}(w_0) > 1$.

Suppose further that $\deg(u_i), \deg(v_i) > \text{length}(w_0)$ *for every $i = 1, \dots, n$, and that the period of the top monomial of x_1 contains no periodicity, and that in addition the top monomials from the two sides of the equation that are obtained* *from the two sides of the equation after substituting the solution x_1, contain no periodicity except the one in the top monomial of x_1 (this translates to a condition on the coefficients u_i and v_i, $1 \leq i \leq n$, in the equation).*

Then there exist some elements $s, t \in FA$, $\deg(s) = \deg(t) \leq \max \deg(u_i)$, such that

(1) *every solution of the equation $sx = xt$ is a solution of the given equation;*

(2) *every solution x_2 of the given equation that satisfies*

$$\deg(x_2) > 2(2 + k^{\deg(s)+2}) + 2(\deg(u_1) + \cdots + \deg(u_n))^2$$

is a solution of the equation $sx = xt$.

Proof. Let x_1 be a solution of the given equation that satisfies

$$\deg(x_1) > 2(\deg(u_1) + \cdots + \deg(u_n)).$$

We start by looking at the top homogeneous part of x_1, which we denote x_1^0. Clearly x_1^0 satisfies the homogeneous equation

$$u_1^0 x_1^0 u_2^0 x_1^0 u_3^0 \cdots u_{n-1}^0 x_1^0 u_n^0 = v_1^0 x_1^0 v_2^0 x_1^0 v_3^0 \cdots v_{n-1}^0 x_1^0 v_n^0,$$

where u_i^0 and v_i^0 are the top monomials in u_i and v_i.

We start the analysis of the given equation under the assumption that there exists an index i for which $\deg(u_i) \neq \deg(v_i)$. In that case there is a shift between the appearances of some of the (homogeneous) elements x_1^0 in the two sides of the (homogeneous) equation. Let i_1 be the first index i for which $\deg(u_i) \neq \deg(v_i)$. The next appearances of x_1^0 in the two sides of the equation must have a shift of $|\deg(u_{i_1}) - \deg(v_{i_1})|$. Since the top homogeneous parts of u_i and v_i are monomials, it follows that the top homogeneous part of x_1 is a monomial as well. We keep the notation x_1^0 for the top monomial of x_1.

Let d be the minimum positive shift between pairs of appearances of x_1^0 in the top degree equation. Then $x_1^0 = e_0(t_0)^b = (s_0)^b f_0$, for some elements t_0, s_0 in the semigroup generated by the free generators $a_1, \ldots, a_k$ of the algebra FA. Note that $\deg(s_0) = \deg(t_0) = d$, e_0 is a prefix of s_0 and suffix of t_0 and f_0 is a suffix of t_0 and prefix of s_0. Since the top monomial u_i^0 and v_i^0 have no periodicity, t_0 and s_0 have no periodicity as well.

Since we assumed that the length of x_1^0 is bigger than the sum of the lengths of the degrees $\deg(u_i)$, an appearance of x_1^0 in the product that is associated with the top monomial in the left side of the equation overlaps with the corresponding appearance of x_1^0 in the right side of the equation, and may overlap with the previous or the next appearance of x_1^0 of the right side of the equation as well. Our assumptions that $\deg(u_i), \deg(v_i) > d$ together with the assumption that the coefficients have no periodicity imply that an appearance of x_1^0 in the product that is associated with the top monomial in one side of the equation may overlap only with the corresponding appearance of x_1^0 in the other side of the equation (and not with the previous or the next one).

Let $1 \leq i_1 < \cdots < i_\ell \leq n-1$ be the indices for which there is a (nontrivial) shift between the appearances of x_1^0 in the two sides of the equation. Let $1 \leq j_1 < \cdots < j_{n-1-\ell} \leq n-1$ be the complementary indices, i.e., those indices for which there is no shift between the corresponding appearances of the monomials x_1^0 in the two sides of the equation.

We start by analyzing the codegree 1 monomials in the products that are associated with the two sides of the equation. We further assumed that the length of the period in x_1^0, i.e., $d = |\deg(u_{i_1}) - \deg(v_{i_1})| > 1$. Note that any codegree 1 monomial in the two products is a product of top monomials with a single codegree 1 monomial from one of the u_i, v_i or one of the appearances of x_1.

Suppose that $i_1 > 1$. Then x_1^0 is quasiperiodic (or rather a fractional power), its period is of length at least 2 and x_1^0 contains at least 2 periods. Hence, a codegree 1 product that contains a codegree 1 monomial of u_1 can cancel with either a codegree 1 product that contains a codegree 1 monomial in v_1 or a codegree 1 product that contains a codegree 1 monomial in the first appearance of x_1.

If the two canceling codegree 1 products contain codegree monomials of u_1 and v_1, then these two codegree 1 monomials must be equal. Otherwise, the codegree 1 product that contains a codegree 1 monomial of u_1 cancels with a codegree 1 product that contains a codegree 1 monomial of the first appearance of x_1. Now, this last codegree 1 monomial appears in the other side of the equation as well, and it can cancel only with a codegree 1 product that contains a codegree 1 monomial of v_1 that must be identical to the codegree 1 monomial of u_1 that we started with. Therefore, the codegree 1 homogeneous parts of v_1 and u_1 are equal. Continuing with the same argument iteratively, the codegree 1 homogeneous parts of the elements u_i and v_i are equal for all $i < i_1$ and $i > i_\ell$.

Let j_s be one of the indices for which there is no shift between the corresponding appearances of x_1^0 in the two sides of the equation. We look at the codegree 1 products in the two sides of the equation. Each such codegree 1 product is a product of a single codegree 1 monomial from a single appearance of x_1^0 or exactly one of the coefficients u_i or v_i, with top degree monomials. Note that the codegree 1 products that involve codegree 1 monomials of the j_s appearance of x_1 in the two sides of the equation (and top degree monomials from all the coefficients and the other appearances of x_1), are precisely the same codegree 1 products in the two sides of the equation. Hence, these do cancel. All the other codegree 1 products in the two sides of the equation contain x_1^0 in the j_s appearance of x_1. Since x_1^0 is periodic, and the length of the period is bigger than 1, a codegree 1 product that includes a codegree 1 monomial to the left of the j_s appearance of x_1 cannot be equal to a codegree 1 product that contains a codegree 1 monomial to the right of the j_s appearance of x_1.

Therefore, the left codegree 1 products (with respect to the j_s appearance of x_1) from the two sides of the equation have to cancel and the right codegree 1 products have to cancel as well. In particular, if for some index i, both the $i-1$ and the i appearances of x_1 in the two sides of the equation have no shift, then u_i and v_i have the same codegree 1 homogeneous parts.

At this point we need to examine the appearances of the variables x_1 in which there is a shift between the two sides of the equation, i.e., in places $i_1, \dots, i_\ell$, and the coefficients u_i and v_i that are connected to these appearances. To do that we break the appearances of the variables x_1, and the coefficients u_i, v_i in the two sides of the equations into regions (or intervals).

We look at the top monomial in the two sides of the equation. For each index i we add a breakpoint at the left point of the pair u_i, v_i, and to the right of that pair. We denote the variable that is associated with the region (interval) between the right point of the pair u_i, v_i and the left point of the pair u_{i+1}, v_{i+1} by w_i. The top monomial of w_i is a prefix or a suffix of the top monomial x_1^0 of x_1. We denote by q_i the variable that is associated with the region between w_{i-1} and w_i. Note that the region that is associated with q_i contains the support of u_i and v_i. If the region that is associated with q_i contains the right part of the $i-1$ appearance of x_1, we denote the variable that is associated with that right part by t_{i-1}. If it contains the left part of the i appearance of x_1, we denote the variable that is associated with that left part by s_i.

As in our previous arguments, we intend to break the solution x_1, so that $x_1 = s_i w_i = w_i t_i$, whenever the variables w_i, s_i, t_i are defined and in appropriate abelian (quotient) groups. Furthermore, each of the elements q_i can be broken according to the two sides of the equation. Hence, we intend to show that $q_i = t_{i-1} u_i s_i$, or $q_i = t_{i-1} u_i$, or $q_i = u_i s_i$, or $q_i = u_i$, and correspondingly for the elements v_i (instead of the u_i), depending on the way the elements q_i are broken in the two sides of the equation.

Because of the periodicity of x_1^0, and since we assume that the length of the period of x_1^0 is bigger than 1, a codegree 1 product that contains a codegree 1 monomial of v_i cannot cancel with a codegree 1 product that contains a codegree 1 monomial in $v_{i'}$ or $u_{i'}$ for $i \neq i'$, and likewise for the u_i.

Suppose that $q_i = v_i = t_{i-1} u_i s_i$ for the top monomials. In that case two codegree 1 products that contain codegree 1 monomials of the $i-1$ and i appearances of x_1 that are both from the v_i side, or both from the u_i side, cannot cancel. Furthermore, two codegree 1 products that cancel and belong to the two sides of the equation cannot contain codegree 1 monomials from both appearances $i-1$ and i of x_1.

Hence, in that case a pair of canceling codegree 1 products may be either

(1) codegree 1 monomials of the same appearance of x_1 from the two sides of the equation;

(2) a codegree 1 monomial of either the i or $i-1$ appearance of x_1 for one product, and a codegree 1 monomial of u_i for the second product;

(3) a codegree 1 monomial of either the i or $i-1$ appearance of x_1 for one product, and a codegree 1 monomial of v_i for the second product;

(4) a codegree 1 monomial of u_i in one product, and a codegree 1 monomial of v_i in the second product.

If case (1) occurs we add a codegree 1 monomial to w_i or w_{i-1} (depending on the appearance of x_1). In case (2) we add a codegree 1 monomial to t_{i-1} or s_i, and the existing one to u_i. In case (3) we add the existing codegree 1 monomial to v_i and a codegree 1 monomial to t_{i-1} or s_i (depending on the appearance of x_1). In case (4) we add only the existing codegree 1 monomials to v_i and u_i.

Suppose that $q_i = t_{i-1}v_i = u_i s_i$ for the top monomials. Two codegree 1 products that contain codegree 1 monomials of the $i-1$ and i appearances of x_1 from the same side of the equation cannot be equal. Furthermore, a codegree 1 product that contains a codegree 1 monomial of the $i-1$ appearance of x_1 in the u_i side cannot cancel with a codegree 1 product that contains a codegree 1 monomial in the i appearance of x_1 from the v_i side. Since we assumed that $\deg(u_i), \deg(v_i) > d$, and the coefficients have no periodicity, a codegree 1 product that contains a codegree 1 monomial of the $i-1$ appearance of x_1 in the v_i side cannot cancel with a codegree 1 product that contains a codegree 1 monomial in the i appearance of x_1 from the u_i side.

Like in the case $q_i = v_i = t_{i-1}u_i s_i$, in that case a pair of canceling codegree 1 products may be either

(1) codegree 1 monomials of the same appearance of x_1 from the two sides of the equation;

(2) a codegree 1 monomial of either the i or $i-1$ appearance of x_1 for one product, and a codegree 1 monomial of u_i for the second product;

(3) a codegree 1 monomial of either the i or $i-1$ appearance of x_1 for one product, and a codegree 1 monomial of v_i for the second product;

(4) a codegree 1 monomial of u_i in one product, and a codegree 1 monomial of v_i in the second product.

If case (1) occurs we add a codegree 1 monomial to w_i or w_{i-1} (depending on the appearance of x_1). In case (2), if a codegree 1 of the $i-1$ appearance of x_1 is part of the canceling pair, we add a codegree 1 monomial to w_{i-1} (only if the codegree 1 product is from the u_i side of the equation), a codegree 1 monomial to t_{i-1}, and the existing codegree 1 monomial to u_i. If a codegree 1 of the i appearance of x_1 is part of the cancelling pair, we add a codegree 1 monomial to s_i, and the existing codegree 1 monomial to u_i. In case (3) we do the equivalent additions for v_i, w_i, t_{i-1} and s_i. In case (4) we just add the codegree 1 existing monomials to u_i and v_i.

So far we have constructed elements w_i, t_i and s_i such that the equations $x_1 = w_i t_i$, $x_1 = s_i w_i$, $q_i = u_i s_i$ or $q_i = t_{i-1} u_i$, or $q_i = t_{i-1} u_i s_i$ or $q_i = u_i$ (and correspondingly for the v_i) hold for products of codegree at most 1. We continue by analyzing products of codegree r, $r < d$, assuming that we analyzed all the products of smaller codegree, constructed the elements w_i, s_i and t_i, and they satisfy the last equations for products of codegree at most $r-1$.

We analyze codegree r products in a similar way to their analysis in the proof of Theorem 4.4. First, note that if a codegree r product is a product of monomials of u_i, v_i, s_i, t_i and w_i that correspond to products of monomials of codegree smaller than r of u_i, v_i and all the appearances of x_1 from the two sides of the equation, then such codegree r products cancel in pairs.

Let i be an index for which $\deg(u_i) = \deg(v_i)$, and there is no shift between the $i-1$ and i appearances of x_1. By our analysis of codegree 1 monomials, the top monomials, and the codegree 1 homogeneous parts of u_i and v_i are identical. Codegree r products from one side of the equation that contain codegree r monomials of the i or $i-1$ appearances of x_1 cancel with corresponding codegree r products from the other side of the equation. Hence, a codegree r product that contains a codegree r monomial of u_i can cancel only with a codegree r product that contains a codegree r monomial of v_i. Therefore, the codegree r homogeneous part of u_i is identical to the codegree r homogeneous part of v_i. Furthermore, for the purpose of analyzing codegree r products, the given equation can be broken into finitely many equations by taking out such pairs of coefficients u_i, v_i, and the appearances of the solution x_1 that are adjacent to them.

Suppose that for some index i there is no shift between the appearances of x_1 in the two sides of the equation. In that case codegree r products that contain codegree r monomials of the i appearance of x_1 from one side of the equation cancel with codegree r products that contain codegree r monomials of that i appearance of x_1 from the other side of the equation. Hence, for the purpose of analyzing codegree r products, the given equation breaks into several equations, by taking out all the appearances of x_1 that have no shift. Therefore, for the continuation of the analysis of codegree r products, we may assume that there are no appearances of x_1 with no shift.

Since we assumed that the equation does not contain appearances of x_1 in the two sides of the equation with no shift between them, the analysis of codegree r products that contain positive codegree monomials of either u_1 or v_1, or positive codegree monomials of either u_n or v_n, is identical to the analysis of codegree r monomials in Theorem 4.4, i.e., in the equation $u_1xu_2 = v_1xv_2$. Hence, we only need to analyze codegree r products that contain positive codegree monomials from some element $q_i = u_is_i = t_{i-1}v_i$ or from an element $q_i = t_{i-1}u_is_i = v_i$.

Let $q_i = u_is_i = t_{i-1}v_i$. Since u_i and v_i have no periodicity, a codegree r product that contains a codegree r monomial of x_1 in its $i-1$ appearance cannot cancel with a codegree r from the same side of the equation that contains a codegree r monomial of x_1 in its i appearance. Furthermore, a codegree r product that contains a codegree r monomial of x_1 in its $i-1$ appearance from the u_i side of the equation, cannot cancel with a codegree r product that contains a codegree r monomial of x_1 in its i appearance from the v_i side of the equation.

Suppose that a codegree r product can be expressed as products of codegree q_j monomials of t_{i-1} and codegree m_j monomials of v_i with top monomials of the other elements in the v_i side of the equation such that $q_j \geq 0$ and m_j is positive and $q_j + m_j = r$, in an odd number of ways. Such codegree r products can cancel with either

(1) an odd number of products of codegree f_j monomials of u_i and codegree g_j monomials of s_i with top monomials of the other elements in the u_i side of the equation, such that $g_j \geq 0$ and f_j is positive and $f_j + g_j = r$;

(2) a product of a codegree r monomial of x_1 in its $i-1$ appearance with other top monomials in the v_i side of the equation;

(3) a product of a codegree r monomial of x_1 in its i appearance with other top monomials in the v_i side of the equation;

(4) a product of a codegree r monomial of x_1 in its $i-1$ appearance with other top monomials in the u_i side of the equation;

(5) a product of a codegree r monomial of x_1 in its i appearance with other top monomials in the u_i side of the equation;

(6) an odd number of products of a codegree b_j monomial of w_{i-1} with a codegree a_j monomial of t_{i-1} for positive $a_j, b_j, a_j + b_j = r$, with top monomials of the other elements from the v_i side;

(7) an odd number of products of a codegree c_j monomial of w_i with a codegree h_j monomial of s_i for positive $c_j, h_j, c_j + h_j = r$, with top monomials of the other elements from the u_i side.

If only case (1) occurs we don't add anything to any of the elements except the existing codegree r monomials of u_i and v_i. If only case (2) occurs we add a codegree r monomial to t_{i-1}. If only case (3) occurs we add a codegree r monomial to w_i and a codegree r monomial to s_i. If only case (4) occurs we add a codegree

r monomial to w_{i-1} and to t_{i-1}. If only case (5) occurs we add a codegree r monomial to s_i.

Cases (2) and (3) cannot occur together, nor cases (4) and (5), nor cases (3) and (4). If only cases (1), (2) and (4) occur, we add a codegree r monomial to w_{i-1}. If only cases (1), (2) and (5) occur, we add codegree r monomials to t_{i-1} and s_i. If only (1), (3) and (5) occur, we add a codegree r monomial to w_i.

We still need to treat cases (6) and (7). Note that the existence of these cases means that codegree r products, which were supposed to exist given the smaller codegree monomials of the various elements, may or may not exist, depending on the existence of codegree r monomials in the various appearances of the element x_1. Also, note that case (6) cannot occur with case (3), and case (7) cannot occur with case (4).

If only case (6) appears, we add a codegree r monomial to t_{i-1}. If only case (7) appears we add a codegree r monomial to s_i. If only cases (1), (2) and (6) appear, we do not add anything. If only cases (1), (2) and (7) appear, we add a codegree r monomial to t_{i-1} and to s_i. If only (1), (3) and (7) appear, we do not add anything. If only (1), (4) and (6) appear, we add a codegree r monomial to w_{i-1}. If only (1), (5) and (6) appear, we add codegree r monomials to t_{i-1} and to s_i. If only (1), (5) and (7) appear, we don't add anything.

If only (2), (4) and (6) appear, we add codegree r monomials to w_{i-1} and to t_{i-1}. If only (2), (5) and (6) appear, we add a codegree r monomial to s_i. If only (2), (5) and (7) appear, we add a codegree r monomial to t_{i-1}. If only (3), (5) and (7) appear, we add codegree r monomials to w_i and to s_i.

If only (1), (6) and (7) appear, we add codegree r monomials to t_{i-1} and to s_i. If only (2), (6) and (7) appear, we add a codegree r monomial to s_i. If only (5), (6) and (7) appear, we add a codegree r monomial to t_{i-1}. If only (1), (2), (5), (6) and (7) appear, we do not add anything.

The case in which the codegree r product is a product of case (1) is dealt with in a symmetric way. Hence, suppose that the codegree r product is not a product of case (1) and cannot be expressed in an odd number of ways as products of codegree q_j monomials of t_{i-1} and codegree m_j monomials of v_i with top monomials of the other elements in the v_i side of the equation, such that $q_j \geq 0$ and m_j is positive and $q_j + m_j = r$.

If only (2) and (4) appear, we add a codegree r monomial to w_{i-1}. If only (2) and (5) appear, we add codegree r monomials to t_{i-1} and s_i. If only (2) and (6) appear, we do not add anything. If only (2) and (7) appear, we add codegree r monomials to t_{i-1} and s_i. If only (3) and (5) appear, we add a codegree r monomial to w_i. If only (3) and (7) appear, we add a codegree r monomial to w_i. If only (4) and (6) appear, we add a codegree r monomial to w_{i-1}. If only (5) and (6) appear, we add codegree r monomials to t_{i-1} and s_i. If only (5) and (7) appear, we do not add anything. If only (6) and (7) appear, we add codegree r monomials to t_{i-1} and s_i. Finally, if (2), (5), (6) and (7) appear, we do not add anything.

As in the proof of Theorem 4.4, it can still be that a codegree r product is of type (7) and can also be presented in an odd number of ways as products of a codegree b_j monomial of w_i with a codegree a_j monomial of t_i for positive a_j, b_j, $a_j + b_j = r$, with top monomials of the other elements from the v_i side.

In that case it can either be presented only in these two forms or also in both forms (3) and (5). If it can be presented in forms (3) and (5) we do not add anything. If it cannot we add a codegree r monomial to w_i.

This concludes the construction of the elements s_i, t_i, w_i for codegree r products that involve $q_i = u_i s_i = t_{i-1} v_i$. Suppose that $q_i = v_i = t_{i-1} u_i s_i$. Since u_i and v_i have no periodicity, a codegree r product that contains a codegree r monomial of x_1 in its $i-1$ appearance cannot cancel with a codegree r product that contains a codegree r monomial of x_1 in its i appearance.

Suppose that a codegree r product can be expressed as products of codegree q_j monomials of t_{i-1}, codegree m_j monomials of u_i and codegree p_j monomials of s_i with top monomials of the other elements in the u_i side of the equation, such that $q_j, m_j, p_j \geq 0$, either $m_j > 0$ or $q_j, p_j > 0$, and $q_j + m_j + p_j = r$, in an odd number of ways. Such codegree r products can cancel with either

(1) a product of a codegree r monomial of v_i with other top monomials;

(2) a product of a codegree r monomial of x_1 in its $i-1$ appearance with other top monomials in the u_i side of the equation;

(3) a product of a codegree r monomial of x_1 in its i appearance with other top monomials in the u_i side of the equation;

(4) a product of a codegree r monomial of x_1 in its $i-1$ appearance with other top monomials in the v_i side of the equation;

(5) a product of a codegree r monomial of x_1 in its i appearance with other top monomials in the v_i side of the equation;

(6) an odd number of products of a codegree b_j monomial of w_{i-1} with a codegree a_j monomial of t_{i-1} for positive a_j, b_j, $a_j + b_j = r$, with top monomials of the other elements from the u_i side;

(7) an odd number of products of a codegree c_j monomial of w_i with a codegree h_j monomial of s_i for positive c_j, h_j, $c_j + h_j = r$, with top monomials of the other elements from the u_i side.

According to the various cases, we add monomials to the variables t_i, s_i, w_i, in a similar way to what we did in case $q_i = u_i s_i = t_{i-1} v_i$. If only case (1) occurs we don't add anything to any of the elements except the existing codegree r monomials of u_i and v_i. If only case (2) occurs we add a codegree r monomial to t_{i-1}. If only case (3) occurs we add a codegree r monomial to s_i. If only case (4) occurs we add a codegree r monomial to w_{i-1} and to t_{i-1}. If only case (5) occurs we add a codegree r monomials to w_i and to s_i.

Cases (2) and (3) cannot occur together, nor cases (4) and (5), nor cases (3) and (4), nor (2) and (5). If only cases (1), (2) and (4) occur, we add a codegree r monomial to w_{i-1}. If only (1), (3) and (5) occur, we add a codegree r monomial to w_i.

As in the case in which $q_i = u_i s_i = t_{i-1} v_i$, the existence of cases (6) and (7) means that codegree r products that were supposed to exist given the smaller codegree monomials of the various elements, may or may not exist, depending on the existence of codegree r monomials in the various appearances of the element x_1. Also, note that case (6) cannot occur with cases (3) or (5), and case (7) cannot occur with cases (2) or (4).

If only case (6) appears, we add a codegree r monomial to t_{i-1}. If only case (7) appears we add a codegree r monomial to s_i. If only cases (1), (2) and (6) appear, we do not add anything. If only (1), (3) and (7) appear, we do not add anything. If only (1), (4) and (6) appear, we add a codegree r monomial to w_{i-1}. If only (1), (5) and (7) appear, we add a codegree r monomial to w_i.

If only (2), (4) and (6) appear, we add codegree r monomials to w_{i-1} and to t_{i-1}. If only (3), (5) and (7) appear, we add codegree r monomials to w_i and to s_i.

The case in which case (1) occurs is dealt with in an analogous way. Hence, suppose that the codegree r product is not a product of case (1) and cannot be expressed in an odd number of ways as products of codegree q_j monomials of t_{i-1}, codegree m_j monomials of u_i and codegree p_j monomials of s_i with top monomials of the other elements in the u_i side of the equation, such that $q_j, m_j, p_j \geq 0$, either m_j is positive or both q_j, p_j are positive and $q_j + m_j + p_j = r$.

If only (2) and (4) appear, we add a codegree r monomial to w_{i-1}. If only (2) and (6) appear, we do not add anything. If only (3) and (5) appear, we add a codegree r monomial to w_i. If only (3) and (7) appear, we do not add anything. If only (4) and (6) appear, we add a codegree r monomial to w_{i-1}. If only (5) and (7) appear, we add a codegree r monomial to w_i.

It can still be that a codegree r product is of type (7) and can also be presented in an odd number of ways as products of a codegree b_j monomial of w_i with a codegree a_j monomial of t_i for positive $a_j, b_j, a_j + b_j = r$, with top monomials of the other elements from the v_i side. We treat this case precisely as we treated it in the case $q_i = u_i s_i = t_{i-1} v_i$.

This concludes the construction of the elements s_i, t_i, w_i for codegree r products when $r < d$. The elements w_i, t_i, and s_i that we constructed so far satisfy the equations $x_1 = w_i t_i$, $x_1 = s_i w_i$, $q_i = u_i s_i$ or $q_i = t_{i-1} u_i$, or $q_i = t_{i-1} u_i s_i$ or $q_i = u_i$ (and correspondingly for the v_i) for products of codegree smaller than d.

To continue we need to analyze products of codegree d and higher. For presentation purposes we start this analysis under the additional assumption that all the appearances of x_1 in the two sides of the equation have nontrivial shifts, i.e., the appearances of the top monomial of the solution x_1^0 in the two sides of the equality for the top monomials are shifted. This assumption enables us to analyze the higher codegree products using the arguments that were used in the proof of Theorem 4.4 and in analyzing smaller codegree products. Afterwards we drop this assumption.

As in Theorem 4.4, in analyzing codegree d products, there are special codegree d products that we need to single out and treat separately, as they may involve cancellations between codegree d products that contain codegree d monomials of u_i or v_i and those that contain codegree d monomials of u_{i+1} or v_{i+1}.

As in analyzing smaller codegree products, note that codegree d products that are products of smaller codegree monomials of the u_i, v_i, s_i, w_i and t_i, and correspond to smaller codegree monomials of u_i, v_i and x_1 from the two sides of the equation cancel in pairs.

In analyzing codegree r products for $r < d$, there is no interaction between elements in q_i and q_j for $i \neq j$. As in the proof of Theorem 4.4, in analyzing codegree d products such interaction may happen if i and j are consecutive indices. Hence, in analyzing codegree d products we need to go over the various possibilities for q_i and q_{i+1}.

Suppose that $q_i = u_i s_i = t_{i-1} v_i$. Suppose that a codegree d product can be expressed as products of codegree q_j monomials of t_{i-1} and codegree m_j monomials of v_i with top monomials of the other elements in the v_i side of the equation, such that $q_j \geq 0$ and m_j is positive and $q_j + m_j = d$, in an odd number of ways. If the q_i part of such a product is not equal to u_i^0 nor to v_i^0 (the top monomials of u_i and v_i), such codegree r products are analyzed exactly in the same way they were analyzed in codegree r products for $r < d$.

We have $u_i^0 \neq v_i^0$ because we assumed that the top monomials of the coefficients have no periodicity. If the q_i part of such a product equals v_i^0, the codegree d product may be equal to a codegree d product that contains positive codegree monomials in q_{i-1}. If the q_i part of such a product equals u_i^0, the codegree d product may be equal to a codegree d product that contains positive codegree monomials in q_{i+1}.

Suppose that the q_i part of the codegree d product equals u_i^0. Suppose further that $q_{i+1} = u_{i+1} s_{i+1} = t_i v_{i+1}$. In that case such a codegree d product can cancel with codegree d products that are either a subset of those analyzed for products of smaller codegree, or products that include positive codegree monomials of q_{i+1}:

(1) an odd number of products of codegree f_j monomials of u_i and codegree g_j monomials of s_i with top monomials of the other elements in the u_i side of the equation, such that $g_j \geq 0$ and f_j is positive and $f_j + g_j = d$;

(2) a product of a codegree d monomial of x_1 in its i appearance with other top monomials in the v_i side of the equation;

(3) a product of a codegree d monomial of x_1 in its i appearance with other top monomials in the u_i side of the equation;

(4) an odd number of products of codegree q_j monomials of t_i and codegree m_j monomials of v_{i+1} with top monomials of the other elements in the v_i side of the equation, such that $q_j \geq 0$ and m_j is positive and the product of the monomial of t_i with the monomial of v_{i+1} is v^0_{i+1};

(5) an odd number of products of codegree f_j monomials of u_{i+1} and codegree g_j monomials of s_{i+1} with top monomials of the other elements in the u_i side of the equation, such that $g_j \geq 0$ and f_j is positive and the product of the monomial of u_{i+1} with the monomial of s_{i+1} is v^0_{i+1};

(6) an odd number of products of a codegree c_j monomial of w_i with a codegree h_j monomial of s_i for positive $c_j, h_j, c_j + h_j = d$, with top monomials of the other elements from the u_i side;

(7) an odd number of products of a codegree b_j monomial of w_i with a codegree a_j monomial of t_i for positive $a_j, b_j, a_j + b_j = d$, with top monomials of the other elements from the v_i side.

Note that case (2) occurs if and only if case (3) occurs. If only one of the cases (1) or (6) occurs, we treat them as they were treated in analyzing codegree r products for $r < d$. If only case (4) or only case (5) occurs we add 1 (the identity) to s_i and t_i, and the codegree d prefix of w^0_i to w_i. If only case (7) occurs we add 1 to s_i and the codegree d prefix of w^0_i to w_i.

If only cases (1)–(3) occur, or only cases (2), (3) and (6) occur, we treat them as they were treated for codegree r products, $r < d$. If only cases (2) and (3) in addition to one of the cases (4) or (5) occur, we add 1 to s_i and t_i. If only cases (2), (3) and (7) occur, we add 1 to s_i. If only (1), (4) and (5) occur, we don't add anything. If only (1), (6) and one of (4) or (5) occur, we add 1 to t_i and the codegree d prefix of w^0_i to w_i. If only (1), (7) and one of (4) or (5) occur, we add 1 to t_i. If only (1), (6) and (7) occur, we add the codegree d prefix of w^0_i to w_i. If only (6), (7) and one of (4) or (5) occur, we add 1 to s_i and t_i and the codegree d prefix of w^0_i to w_i. If only (4), (5) and (6) occur, we add 1 to s_i. If only (4), (5) and (7) occur, we add 1 to s_i and the codegree d prefix of w^0_i to w_i.

If only (1)–(5) occur, we add the codegree d prefix of w^0_i to w_i. If only (1)–(3) and (6)–(7) occur, we do not add anything. If only (1)–(3), (6) and one of (4) or (5) occur, we add 1 to t_i. If only (1)–(3), (7) and one of (4) or (5) occur, we add 1 to t_i and the codegree d prefix of w^0_i to w_i. If only (1) and (4)–(7) occur, we add the codegree d prefix of w^0_i to w_i.

If only (2)–(6) occur, we add 1 to s_i and the prefix of codegree d of w_i^0 to w_i. If only (2)–(5) and (7) occur, we add 1 to s_i. If only (2)–(3), (6)–(7) and one of (4) or (5) occur, we add 1 to s_i and t_i and the codegree d prefix of w_i^0 to w_i. If all the possibilities (1)–(7) occur, we do not add anything.

Suppose that a codegree d product can be expressed as a product in case (1), and cannot be expressed as products of codegree q_j monomials of t_{i-1} and codegree m_j monomials of v_i with top monomials of the other elements in the v_i side of the equation, such that $q_j \geq 0$ and m_j is positive and the q_i part of the product is u_i^0 in an even number (possibly none) ways. In that case the analysis of such a product and the monomials that are added to the elements t_i, s_i and w_i are analogous to the analysis described above.

Suppose that such a codegree d product cannot be expressed as a product in case (1), but it can be expressed as a product in case (6). If only (6) and (7) occur, we add the codegree d prefix of w_i^0 to w_i. If only (6) and one of (4) or (5) occur, we add 1 to t_i and the prefix of codegree d of w_i^0 to w_i. If only (4)–(7) occur, we add the codegree d prefix of w_i^0 to w_i. If only (2)–(3) and (6)–(7) occur, we do not add anything. If only (2)–(3), (6) and one of (4) or (5) occur, we add 1 to t_i. If only (2)–(7) occur, we do not add anything.

This concludes the analysis of such codegree d products in the case that $q_{i+1} = u_{i+1}s_{i+1} = t_i v_{i+1}$. Suppose that $q_{i+1} = u_{i+1} = t_i v_{i+1} s_{i+1}$. As before, such a codegree d product can cancel with codegree d products that are either a subset of the ones that were analyzed for products of smaller codegree, or products that include positive codegree monomials of q_{i+1}:

(1) an odd number of products of codegree f_j monomials of u_i and codegree g_j monomials of s_i with top monomials of the other elements in the u_i side of the equation, such that $g_j \geq 0$ and f_j is positive and $f_j + g_j = d$;

(2) a product of a codegree d monomial of x_1 in its i appearance with other top monomials in the v_i side of the equation;

(3) a product of a codegree d monomial of x_1 in its i appearance with other top monomials in the u_i side of the equation;

(4) an odd number of products of codegree q_j monomials of t_i, codegree m_j monomials of v_{i+1} and codegree p_j monomials of s_{i+1} with top monomials of the other elements in the v_i side of the equation, such that $q_j, m_j, p_j \geq 0$, either $m_j > 0$ or $q_j, p_j > 0$, and $q_j + m_j + p_j = d$, and the product of the corresponding monomials of t_i, v_{i+1} and s_{i+1} is the codegree d suffix of u_{i+1}^0;

(5) a product of a monomial of u_{i+1}, which is the codegree d suffix of u_{i+1}^0, with the top monomials of the all the other elements from the u_i side of the equation;

(6) an odd number of products of a codegree c_j monomial of w_i with a codegree h_j monomial of s_i for positive $c_j, h_j, c_j + h_j = d$, with top monomials of the other elements from the u_i side;

(7) an odd number of products of a codegree b_j monomial of w_i with a codegree a_j monomial of t_i for positive $a_j, b_j, a_j + b_j = d$, with top monomials of the other elements from the v_i side.

Analyzing the various possibilities in this case is identical to the case $q_{i+1} = u_{i+1}s_{i+1} = t_i v_{i+1}$.

Recall that we assumed that $q_i = u_i s_i = t_{i-1} v_i$. In addition suppose that a codegree d product can be expressed as products of codegree f_j monomials of u_i and codegree g_j monomials of s_i with top monomials of the other elements in the u_i side of the equation, such that $f_j \geq 0$ and g_j is positive and $f_j + g_j = d$, in an odd number of ways, and such that the product of the monomial of u_i with the monomial of s_i is v_i^0. In that case, the codegree d product may be equal to a codegree d product that contains positive codegree monomials in q_{i-1}. Such a codegree d product can cancel with codegree d products that are either a subset of the ones that were analyzed for products of smaller codegree, or products that include positive codegree monomials of q_{i-1}:

(1) an odd number of products of codegree q_j monomials of t_{i-1} and codegree m_j monomials of v_i with top monomials of the other elements in the v_i side of the equation, such that $q_j \geq 0$ and m_j is positive and $q_j + m_j = d$;

(2) a product of a codegree d monomial of x_1 in its $i-1$ appearance with other top monomials in the u_i side of the equation;

(3) a product of a codegree d monomial of x_1 in its $i-1$ appearance with other top monomials in the v_i side of the equation;

(4) an odd number of products of codegree f_j monomials of u_{i-1} and codegree g_j monomials of s_{i-1} with top monomials of the other elements in the u_i side of the equation, such that $g_j \geq 0$ and f_j is positive and the product of the monomial of u_{i-1} with the monomial of s_{i-1} is u_{i-1}^0;

(5) an odd number of products of q_j monomials of t_{i-2} and codegree m_j monomials of v_{i-1} with top monomials of the other elements in the v_i side of the equation, such that $q_j \geq 0$ and m_j is positive and the product of the monomial of t_{i-2} with the monomial of v_{i-1} is u_{i-1}^0;

(6) an odd number of products of a codegree c_j monomial of w_{i-1} with a codegree h_j monomial of t_{i-1} for positive $c_j, h_j, c_j + h_j = d$, with top monomials of the other elements from the v_i side;

(7) an odd number of products of a codegree a_j monomial of s_{i-1} with a codegree b_j monomial of w_{i-1} for positive $a_j, b_j, a_j + b_j = d$, with top monomials of the other elements from the u_i side.

The analysis of this case is identical to the case in which the q_i part of a codegree d product is v_i^0, and there is a possible cancellation with codegree d products that contain positive codegree monomials of q_{i+1}. An identical analysis applies also when $q_{i-1} = v_{i-1} = t_{i-2}u_{i-1}s_{i-1}$.

Suppose that $q_i = u_i = t_{i-1}v_i s_i$ and $q_{i+1} = v_{i+1} = t_i u_{i+1} s_{i+1}$. Suppose that a codegree d product can be presented in an odd number of ways as products of codegree q_j monomials of t_{i-1}, codegree m_j monomials of v_i and codegree p_j monomials of s_i with top monomials of the other elements in the v_i side of the equation, such that $q_j, m_j, p_j \geq 0$, either $m_j > 0$ or $q_j, p_j > 0$, and $q_j + m_j + p_j = d$, and the product of the corresponding monomials of t_{i-1}, v_i and s_i is the codegree d prefix of u_i^0.

Such a codegree d product can cancel with codegree d products that are either a subset of the ones that were analyzed for products of smaller codegree, or products that include positive codegree monomials of q_{i+1}:

(1) a product of a monomial of u_i which is the codegree d prefix of u_i^0 with the top monomials of all the other elements from the u_i side of the equation;

(2) a product of a codegree d monomial of x_1 in its i appearance with other top monomials in the v_i side of the equation;

(3) a product of a codegree d monomial of x_1 in its i appearance with other top monomials in the u_i side of the equation;

(4) a product of a monomial of v_{i+1} which is the codegree d suffix of v_{i+1}^0 with the top monomials of the all the other elements from the v_i side of the equation;

(5) an odd number of products of codegree q_j monomials of t_i, codegree m_j monomials of u_{i+1} and codegree p_j monomials of s_{i+1} with top monomials of the other elements in the u_i side of the equation, such that $q_j, m_j, p_j \geq 0$, either $m_j > 0$ or $q_j, p_j > 0$, and $q_j + m_j + p_j = d$, and the product of the corresponding monomials of t_i, v_{i+1} and s_{i+1} is the codegree d suffix of v_{i+1}^0;

(6) an odd number of products of a codegree c_j monomial of w_i with a codegree h_j monomial of s_i for positive $c_j, h_j, c_j + h_j = d$, with top monomials of the other elements from the v_i side;

(7) an odd number of products of a codegree b_j monomial of w_i with a codegree a_j monomial of t_i for positive $a_j, b_j, a_j + b_j = d$, with top monomials of the other elements from the u_i side.

Analyzing the various possibilities in this case is identical to the case $q_i = t_{i-1}v_i = u_i s_i$. The analysis of the remaining case, in which $q_i = u_i = t_{i-1}v_i s_i$ and $q_{i-1} = v_{i-1} = t_{i-2}u_{i-1}s_{i-1}$, is identical to the previous cases as well.

This concludes the construction of the elements s_i, t_i, w_i for codegree r products when $r \leq d$, in case all the pairs of appearances of the top monomial of the solution x_1 in the two sides of the equation have nontrivial shifts. The elements w_i, t_i and s_i that we constructed so far satisfy the equations $x_1 = w_i t_i$, $x_1 = s_i w_i$, $q_i = u_i s_i$ or $q_i = t_{i-1}u_i$, or $q_i = t_{i-1}u_i s_i$ or $q_i = u_i$ (and correspondingly for the v_i) for products of codegree smaller or equal to d.

As in the proof of Theorem 4.4, we continue with the analysis of codegree $d+r$ products for $r < d$. First, as in analyzing smaller codegree products, codegree $d+r$ products that are products of smaller codegree monomials of u_i, v_i, s_i, t_i and w_i, that correspond to products of smaller codegree monomials of u_i, v_i and x_1 (in all its appearances) from the two sides of the equation, cancel in pairs. We start with two lemmas that are the analogues of Lemmas 4.5 and 4.6.

Lemma 4.8. *Suppose that a codegree $d+r$ product can be presented both as*

(1) *a product of a codegree h monomial of s_i with a codegree c monomial of w_i, for positive c, h, $c+h = d+r$, with top monomials of the other elements from the u_i side*;

(2) *a product of a codegree b monomial of w_i with a codegree a monomial of t_i for positive a, b, $a+b = d+r$, with top monomials of the other elements from the v_i side.*

Such a codegree $d+r$ product may only be presented as a product of smaller codegree monomials or (*only*) *in one of the following two products*:

(i) *a product of a codegree $d+r$ monomial of x_1 in its i appearance with other top monomials in the v_i side of the equation*;

(ii) *a product of a codegree $d+r$ monomial of x_1 in its i appearance with other top monomials in the u_i side of the equation.*

Proof. In case it can be presented as another product of a codegree $d+r$ monomial with top degree monomials, either the top monomial of s_i or the top monomial of t_i overlap with themselves with a cyclic shift. Hence they must be periodic, a contradiction to the assumption that the coefficients do not have nontrivial periodicity. □

Lemma 4.9. *With the notation of Lemma 4.8, if a codegree $d+r$ product can be presented in an odd number of ways as a product in the form* (1) *and in an even number of ways as a product of form* (2), *then such a product can be presented precisely in one of the forms* (i) *or* (ii). *If a codegree $d+r$ product can be presented precisely in one of the forms* (i) *or* (ii), *then it can be presented precisely in one of the forms* (1) *or* (2) *in an odd number of ways.*

If a codegree $d+r$ product can be presented in an odd number of ways in both forms (1) *and* (2), *then it can either be presented in both forms* (i) *and* (ii) *or in neither of them. If a codegree $d+r$ product can be presented in both forms* (i) *and* (ii) *then it can either be presented in both forms* (1) *or* (2) *in an odd number of ways*, *or in both of them in an even number of ways.*

Proof. If a codegree $d+r$ product can be presented in both forms (1) and (2) (odd or even number of times), the conclusion follows from Lemma 4.8. Suppose that it can be presented in an odd number of ways in form (1) and none in form (2). If it can also be presented as a codegree $d+r$ product that involves positive codegree monomials of u_j, v_j, s_j, t_j or x_j, for $j > i$, the top monomial of u_{i+1} must have nontrivial periodicity, a contradiction. If it can be also presented as a codegree $d+r$ product from the u_i sides of the equation that involves monomials of positive codegree monomials of u_j, s_j, t_j or x_j, for $j < i$, the top monomial of u_i must have nontrivial periodicity, a contradiction.

Suppose that the given codegree $d+r$ product can also be presented as a product of either

(1) a codegree q of t_{i-1} and a codegree m of v_i with other top monomials from the v_i side of the equation;

(2) a codegree f of u_i and a codegree g of s_i with other top monomials from the u_i side of the equation;

(3) a codegree $d+r$ product from the v_i side of the equation that involves monomials of positive codegree monomials of v_j, s_j, t_j or x_j for $j < i$.

In all these cases the suffix of length r of the top monomial of u_i is identical to the prefix of length r of the period of x. If $r \leq \deg(v_i) - d$, then v_i has nontrivial periodicity, a contradiction. Otherwise, the top monomial in the two sides of the equation contains periodicity that is not part of the periodicity of the solution x, a contradiction to our assumptions. □

Suppose that $q_i = u_i s_i = t_{i-1} v_i$, and let r be an integer, $0 < r < d$. By Lemma 4.9 if a codegree $d+r$ product can be presented in an odd number of ways in the form (1) of Lemma 4.8 then either

(1) it can be also presented in an odd number of ways as in form (2) of Lemma 4.8 and either in both forms (i) and (ii) in Lemma 4.8 or in neither of them;

(2) it can be presented in an even or no ways in form (2) of Lemma 4.8, and it can also be presented precisely in one of the forms (i) or (ii) in Lemma 4.8.

By Lemma 4.9, if a codegree $d+r$ product can be presented in form (i) of Lemma 4.8, and in even or no ways in forms (1) or (2) of that lemma, then it can also be presented in form (ii) of Lemma 4.8.

Hence, if a codegree $d+r$ product can be presented in an odd number of ways in one of the forms (1), (2), (i) or (ii), then the appearances of the codegree $d+r$ products in these forms cancel in pairs. If it appears in an odd number of ways in forms (1) and (2), and in forms (i) and (ii), we do not add anything. If it appears in an odd number of ways in forms (1) and (2) and not in the forms (i) nor (ii), we add a codegree $d+r$ monomial to w_i. If it appears in an odd number of ways in the form (1), in an even number of or no ways in the form (2), and appears in the form (i) we add a codegree $d+r$ monomial to w_i. If it appears in an odd number of ways in the form (1), in an even number of or no ways in the form (2), and in the form (ii), we do not add anything. If it appears in an even number of or no ways in the forms (1) and (2), and in both form (i) and (ii), we add a codegree $d+r$ monomial to w_i.

Therefore, if a codegree $d+r$ product can be presented in an odd number of ways as products of codegree q_j monomials of t_{i-1} and codegree m_j monomials of v_i with top monomials of the other elements in the v_i side of the equation, such that $q_j \geq 0$ and m_j is positive and $q_j + m_j = d+r$, then it must be presented in an odd number of ways as products of codegree f_j monomials of u_i and codegree g_j monomials of s_i with top monomials of the other elements in the u_i side of the equation, such that $g_j \geq 0$ and f_j is positive and $f_j + g_j = r$.

This concludes the construction of the elements s_i, t_i, w_i in case $q_i = t_{i-1}v_i = u_i s_i$ (note that the elements s_i, t_i did not change), to ensure that the equalities they are supposed to satisfy hold for products up to codegree $d+r$.

Suppose that $q_i = u_i = t_{i-1}v_i s_i$. Lemmas 4.8 and 4.9 and their proofs remain valid in this case. Hence, a codegree $d+r$ product can be expressed in an odd number of ways as products of codegree q_j monomials of t_{i-1}, codegree m_j monomials of v_i, and codegree p_j monomials of s_i with top monomials of the other elements in the v_i side of the equation, such that $q_j, m_j, p_j \geq 0$, either $m_j > 0$ or $q_j, p_j > 0$, and $q_j + m_j + p_j = d+r$, if and only if it is equal to a codegree $d+r$ monomial of u_i.

This concludes our treatment of codegree $d+r$ products for $r < d$. We continue by analyzing codegree $2d$ products. Lemmas 4.8 and 4.9 remain valid for codegree $2d$ products. Hence, the analysis of codegree $2d$ products is identical to the analysis of codegree $d+r$ products for $r < d$. The analysis of higher codegree products, for codegree up to twice the maximal degree of the elements u_i, v_i, is identical as well.

Hence, in case $\deg(u_i), \deg(v_i) > d$ and all the appearances of the elements x_1 in the two sides of the equation have nontrivial shifts, we finally constructed elements s_i, t_i, w_i that satisfy the equations

(i) $q_i = u_i s_i = t_{i-1} v_i$ or $q_i = v_i = t_{i-1} u_i s_i$ or with exchanging the appearances of u_i and v_i in the second equation;

(ii) $x_1 = s_i w_i = w_i t_i \mod G^{\deg(x_1) - 2(\deg(s_i))}$.

Therefore, s_1 and t_{n-1} are uniquely defined, and (given x_1) w_1 and w_{n-1} are uniquely defined mod $G^{\deg(w_1)-2(\deg(s_i))}$. Hence, t_1 and s_2 are uniquely defined, and w_2 is uniquely defined mod $G^{\deg(w_2)-2(\deg(s_i))}$. Continuing iteratively, all the elements s_i, t_i are uniquely defined, and the elements w_i are uniquely defined mod $G^{\deg(w_i)-2(\deg(s_i))}$.

Since $s_i w_i = w_i t_i$, it follows that $s_i x_1 = x_1 t_i$ mod $G^{\deg(s_i x_1)-2(\deg(s_i))}$. This implies that for every pair i, j, $1 \le i, j \le n$, we have $(s_i + s_j)x_1 = x_1(t_i + t_j)$ mod $G^{\deg(s_i x_1)-2(\deg(s_i))}$, so for every pair i, j either $s_i = s_j$ and $t_i = t_j$ or $s_i = s_j + 1$ and $t_i = t_j + 1$.

Since every pair (s_i, t_i) is either (s_1, t_1) or $(s_1 + 1, t_1 + 1)$, it follows that every element $\hat{x}$ that satisfies $s_1\hat{x} = \hat{x}t_1$ is a solution of the given equation. It remains to prove that every long enough solution of the given equation is a solution of the equation $s_1 x = x t_1$.

Let x_2 be a solution of the given equation that satisfies

$$\deg(x_2) > 2(2 + k^{\deg(s)+2}) + (2(\deg(u_1) + \cdots + \deg(u_n))^2.$$

By continuing the analysis of higher codegree monomials of the solution x_2, we get that there exist elements w_i such that for every index i, $1 \le i \le n$, we have $s_i w_i = w_i t_i = x_2$ mod $G^{\deg(s_1)-1}$. By the argument that was used to prove Lemma 4.2, it follows that there exists a solution $\hat{x}$ to the equation $s_1 x = x t_1$.

Note that x_2 satisfies $s_1 x_2 = x_2 t_1$ mod $G^{2\deg(s_1)-1}$. Hence, there exists an element $\hat{x}_2$ which is a solution of the equation $s_1 x = x t_1$, and $x_2 + \hat{x}_2 = r$, where $\deg(r) \le 2 + k^{\deg(s_1)+2}$.

Suppose the given equation is $v_1 x v_2 x v_3 = u_1 x u_2 x u_3$, where $\deg(v_1) < \deg(u_1)$ and $\deg(v_2) = \deg(u_2)$. In this case, $u_1 = v_1 s_1$, $t_1 u_2 = v_1 s_2$ and $v_3 = t_2 u_3$. Hence, $(\hat{x}_2 + r)v_2(\hat{x}_2 + r)t_2 = s_1(\hat{x}_2 + r)u_2(\hat{x}_2 + r)$. Since $\hat{x}_2$ is a solution to the equation $s_1 x = x t_1$, it is a solution to the given equation. Therefore

$$\hat{x}_2 v_2 r t_2 + r v_2 \hat{x}_2 t_2 = s_1 \hat{x}_2 u_2 r + s_1 r u_2 \hat{x}_2 \text{ mod } G^{\deg(r v_2 r t_2)}.$$

Hence

$$\hat{x}_2(v_2 r t_2 + t_1 u_2 r) = (r v_2 s_2 + s_1 r u_2)\hat{x}_2 \text{ mod } G^{\deg(r v_2 r t_2)}.$$

Since $s_1\hat{x}_2 = \hat{x}_2 t_1$ it follows that for any polynomial p, $p(s_1)\hat{x}_2 = \hat{x}_2 p(t_1)$. This implies $v_2 r t_2 + t_1 u_2 r = p(t_1)$ and $r v_2 s_2 + s_1 r u_2 = p(s_1)$ mod $G^{\deg(v_2 r t_2)+\deg(r)-\deg(x_2)}$ for the same polynomial p.

We have $t_1 u_2 = v_2 s_2$, so $v_2(r t_2 + s_2 r) = p(t_1)$ mod $G^{\deg(v_2 r t_2)+\deg(r)-\deg(x_2)}$. By our assumption on $\deg(x_2)$ it follows that $v_2(r t_2 + s_2 r) = p(t_1)$. Similarly, $(r t_1 + s_1 r)u_2 = p(s_1)$. Hence, $p(s_1)$ is either 0 or its leading term is of degree at least 2.

Since (s_1, t_1) equals (s_2, t_2) or $(s_2 + 1, t_2 + 1)$, we get that $v_2(r t_1 + s_1 r)u_2 = v_2 p(t_1) = p(s_1)u_2$. We look at the leading term in the two sides of the last equality.

Since we assumed that the top monomials of u_2 and v_2 are not periodic, the top monomial of u_2 must be βs_0, and the top monomial of v_2 must be $t_0\beta$, where β is a prefix of t_0 and a suffix of s_0. Hence, $t_0=\beta\alpha$ and $s_0=\alpha\beta$. But this is a contradiction, since we assumed that the periodicity in the top monomials in the two sides of the given equation is contained in the solution x_2. Therefore, $s_1r+rt_1=0$, so r is a solution of the equation $s_1x=xt_1$, which means that $x_2=\hat{x}_2+r$ is a solution to $s_1x=xt_1$ as well.

If the equation is $u_1xu_2xu_3=v_1xv_2xv_3$, and $\deg(v_1)>\deg(u_1)$, $\deg(v_3)>\deg(u_3)$, then by the same arguments we get that r (the remainder) has to satisfy the equation

$$(rt_1+s_1r)v_2s_2\hat{x}_2=\hat{x}_2t_1v_2(rt_2+s_2r).$$

That implies that if $rt_1+s_1r\neq 0$, u_2 must contain periodicity, a contradiction to our assumptions. Therefore, $rt_1+s_1r=0$, and both r and x_2 are solutions of the equations $s_1x=xt_1$.

Suppose that the length of the equation is bigger. Then x_2 is a long solution, and $x_2=\hat{x}_2+r$, where $\hat{x}_2$ is a solution of the equation $s_1x=xt_1$, and $\deg(r)\leq 2+k^{\deg(s)+2}$. In that case we get the equality

$$(\hat{x}_2+r)v_2(\hat{x}_2+r)v_3\cdots v_{n-1}(\hat{x}_2+r)t_{n-1}=s_1(\hat{x}_2+r)u_2(\hat{x}_2+r)u_3\cdots u_{n-1}(\hat{x}_2+r),$$

and since $\hat{x}_2$ is a solution of the equation $s_1x=xt_1$, we get the equality

$$\begin{aligned}rv_2\hat{x}_2v_3\cdots v_{n-1}\hat{x}_2t_{n-1}+\cdots+\hat{x}_2v_2\hat{x}_2v_3\cdots\hat{x}_2v_{n-1}rt_{n-1}\\ =s_1ru_2\hat{x}_2u_3\cdots u_{n-1}\hat{x}_2+\cdots+s_1\hat{x}_2u_2\hat{x}_2u_3\cdots\hat{x}_2u_{n-1}r \bmod G^{m_2},\end{aligned}$$

where $m_1=\deg(s_1\hat{x}_2u_2\hat{x}_2u_3\cdots\hat{x}_2u_{n-1}r)$, and $m_2=m_1-\deg(\hat{x}_2)+\deg(r)$.

That implies the equality

$$\begin{aligned}(s_1r+rt_1)u_2\hat{x}_2u_3\cdots u_{n-1}\hat{x}_2+\hat{x}_2v_2(s_2r+rt_2)u_3\hat{x}_2u_4\cdots\hat{x}_2u_n\\ +\hat{x}_2v_2\cdots\hat{x}_2v_{n-2}(s_{n-2}r+rt_{n-2})u_{n-1}\hat{x}_2u_n\\ +\hat{x}_2v_2\cdots\hat{x}_2v_{n-1}(s_{n-1}r+rt_{n-1})=0 \bmod G^{m_2}.\end{aligned}$$

Suppose that there exists an index j, $1\leq j\leq n-1$, for which $s_jr+rt_j\neq 0$. We set j_0 to be the minimal index for which s_jr+rt_j has maximal degree. We look at the top degree homogeneous part in $s_{j_0}r+rt_{j_0}$. The monomials in this homogeneous part of $s_{j_0}r+rt_{j_0}$ contribute to top degree monomials in the j_0-th product in the sum above. These top degree monomials cancel with top degree monomials from other summands that contain part of the top monomial of $\hat{x}_2$ in place of the top monomial of $s_{j_0}r+rt_{j_0}$. Hence, the top degree homogeneous part of $s_{j_0}r+rt_{j_0}$ has to be a monomial as well.

Furthermore, as for an equation of length 3, this cancellation of the top monomials implies that the top monomials of u_{j_0} and v_{j_0} contain parts of the top monomial of $\hat{x}_2$, that by our assumption is bigger than the length of the period of the top monomial of $\hat{x}_2$. Hence, as for equation of length 3, when we substitute $\hat{x}_2$ in the equation, the top monomial has periodicity that is not contained in one of the appearances of $\hat{x}_2$, a contradiction to one of the assumptions of Theorem 4.7. Therefore, for every j, $s_j r + r t_j = 0$, so r is a solution of the equation $s_1 x = x t_1$, and so is x_2.

This concludes the proof of Theorem 4.7 in case all the appearances of the top monomial of a solution x_1 in the two monomials that are the top products in the two sides of the given equation have nontrivial shifts. We still need to complete the proof in the cases in which there are appearances of the top monomial of a solution x_1 with zero shifts.

Lemma 4.10. *Let $u_1, u_2, v_1, v_2 \in FA$ satisfy $u_1 \neq v_1$, $\deg(u_i) = \deg(v_i)$, $i = 1, 2$, and suppose that the top homogeneous parts of u_i and v_i are monomials (for $i = 1, 2$) with no nontrivial periodicity. Then, if there exists a solution x_1 to the equation $u_1 x u_2 = v_1 x v_2$, and $\deg(x_1) > 2(\deg(u_1) + \deg(u_2))$, then there exist elements $s, t \in FA$ such that x is a solution of the equation $u_1 x u_2 = v_1 x v_2$ if and only if it is a solution of the equation $sx = xt$.*

Proof. The top monomials of u_1 and v_1, and of u_2 and v_2, have to be equal. We set $u_1 = v_1 + \mu_1$, $v_2 = u_2 + \mu_2$, $\deg(\mu_1) < \deg(v_1)$ and $\deg(\mu_2) < \deg(u_2)$. Hence, $(v_1 + \mu_1) x u_2 = v_1 x (u_2 + \mu_2)$, which implies $\mu_1 x u_2 = v_1 x \mu_2$. Since the top homogeneous parts of v_1 and u_2 are monomials with no periodicity, so are the top homogeneous parts of μ_1 and μ_2. Since $\deg(\mu_1) < \deg(v_1)$ and $\deg(\mu_2) < \deg(v_2)$, the conclusion of the lemma follows from Theorem 4.4. □

Proposition 4.11. *Let $u_1, u_2, u_3, v_1, v_2, v_3 \in FA$ satisfy $u_1 \neq v_1$, $u_3 \neq v_3$, $\deg(u_i) = \deg(v_i)$, $i = 1, 2, 3$, and suppose that the top homogeneous parts of u_i and v_i are monomials (for $i = 1, 2, 3$) with no nontrivial periodicity. Then, if there exists a solution x_1 to the equation $u_1 x u_2 x u_3 = v_1 x v_2 x v_3$, and the only nontrivial periodicity in the top monomials of the two sides of the equation is contained in the top monomials of the solution x_1, and $\deg(x_1) > 2(\deg(u_1) + \deg(u_2) + \deg(u_3))$, then there exist elements $s, t \in FA$ such that up to a swap between the u's and the v's:*

(1) *There exists μ_1 for which $u_1 = \mu_1(s + 1)$ and $v_1 = \mu_1 s$.*
(2) *There exists μ_2 and τ_2 for which $t\mu_2 = \tau_2 s$. Furthermore, $u_2 = \tau_2(s + 1)$ and $v_2 = (t + 1)\mu_2$.*
(3) *There exists μ_3 for which $u_3 = t\mu_3$ and $v_3 = (t + 1)\mu_3$.*

As in the conclusion of Theorem 4.7, every solution of the equation $sx = xt$ is a solution of the given equation $u_1 x u_2 x u_3 = v_1 x v_2 x v_3$. Every solution x_2 of the given equation $u_1 x u_2 x u_3 = v_1 x v_2 x v_3$ for which $\deg(x_2) > 2(2 + k^{\deg(s_1)+2} + \deg(u_1) + \deg(u_2) + \deg(u_3))$ is a solution of the equation $sx = xt$.

Proof. The top homogeneous parts of the u_i and v_i are monomials, and the equation forces these monomials to be equal. Hence, $v_i = u_i + \rho_i$, where $\deg(\rho_i) < \deg(u_i)$, $i = 1, 2, 3$. Let ρ_i^0 be the top homogeneous part in ρ_i. We start the proof by arguing that the top homogeneous part of a solution x_1 with $\deg(x_1) > 2(\deg(u_1) + \deg(u_2) + \deg(u_3))$ has to be a monomial as well.

Suppose that $\deg(\rho_1) < \max(\deg(\rho_2), \deg(\rho_3))$. In that case the top homogeneous parts have to satisfy $\rho_2^0 x_1^0 v_3^0 = u_2^0 x_1^0 \rho_3^0$. Since u_2^0 and v_3^0 are monomials, it follows that x_1^0 is a monomial, and so are ρ_2^0 and ρ_3^0. If $\deg(\rho_1^0) \geq \max(\deg(\rho_2^0), \deg(\rho_3^0))$, then ρ_1^0 has to be a monomial. This forces x_1^0 to be a monomial as well.

We look at the highest degree for which for some index i, $u_i \neq v_i$. This cannot occur for a single index i. If $u_2 = v_2$ at that highest degree, then the top monomial in u_2 (and v_2) must have periodicity, a contradiction to our assumptions. Let d be the codegree of that degree, and suppose that up to this codegree $u_3 = v_3$. In that case, the equation for codegree d products reduces to the equation $u_1 x u_2 = v_1 x v_2$. If we set $u_i = v_i + \mu_i$, $i = 1, 2, 3$, then for the codegree d products, we get the equation $\mu_1 x u_2 = v_1 x \mu_2$. This implies that the top part of μ_1 and μ_2 are monomials that are the codegree d prefix and suffix of the top monomials of v_1 and v_2 in correspondence, and that the top monomial of x_1 has a period of length d.

In that case, it must be that $u_3 = v_3$ for all the homogeneous parts of codegree less than $2d$, and hence, $\mu_1 x u_2 = v_1 x \mu_2$ for all the products up to codegree d. Therefore, there exists an element s, and an element t, such that $v_1 = u_1 = \mu_1 s \mod G^{\deg(u_1)-d}$ and $v_2 = u_2 = t\mu_2 \mod G^{\deg(u_2)-d}$.

Since $u_3 = v_3$ for all the homogeneous parts of codegree less than $2d$, and the top monomial of u_3 (and v_3) do not have nontrivial periodicity, it follows that $u_3 = v_3$. Hence, $\mu_1 x u_2 = v_1 x \mu_2$, and the conclusion follows from Theorem 4.4 in this case (note that in the statement of the proposition we assumed that $u_i \neq v_i$, $i = 1, 3$).

Suppose that for the codegree d homogeneous parts $u_i \neq v_i$ for $i = 1, 2, 3$. In that case, we get the equation

$$(v_1 + \mu_1) x u_2 x u_3 = v_1 x v_2 x (u_3 + \mu_3),$$

and $u_i = v_i$, $i = 1, 2, 3$, for all the homogeneous parts of codegree smaller than d. Hence, the top homogeneous parts of μ_1 and μ_3 are monomials, which are the codegree d prefix and suffix of the top monomials of u_1 and u_3 in correspondence. The top monomial of x_1 (the given solution to the given equation) has to be quasiperiodic (or rather fractional periodic), with a period of length d. Furthermore, $v_2 = b_2 + \mu_2$ and $u_2 = b_2 + \tau_2 \mod G^{\deg(u_2)-(d+1)}$, where the top homogeneous parts of μ_2 and τ_2 are the codegree d prefix and suffix of the top monomial of u_2 (and v_2).

We continue by looking at products of codegree $d + 1$. Every such product that contains monomials in u_i that appear also in v_i, for $i = 1, 2, 3$, cancels with a similar product from the other side of the equation. Hence, to analyze cancellations, we need to consider codegree $d+1$ products that contain monomials from μ_1 or μ_3, or monomials of codegree d and $d+1$ of u_2 and v_2 that do not appear in both.

Suppose that a codegree $d+1$ product contains a codegree $d+1$ monomial from μ_1, i.e., a codegree $d+1$ monomial in u_1 that is not in v_1. Such a codegree $d+1$ product must contain the top monomial of x_1 in its two appearances, and the top monomial of u_2 and u_3. Since the top monomial of v_1 doesn't have nontrivial periodicity, such a codegree $d+1$ product cannot cancel with a codegree $d+1$ product that contains the top monomial of v_1 (since otherwise the suffix of the top monomial of v_1 equals a shift by 1 of itself, which implies that the suffix of v_1 contains periodicity). Therefore, a codegree $d+1$ product that cancels with it must contain a codegree 1 monomial of either u_1 or v_1, or the top monomial of μ_1. Since the top monomial of v_2 contains no periodicity, if this codegree $d+1$ product contains a codegree 1 monomial of u_1 or v_1 it must contain the top monomial of μ_2. Hence, this codegree $d+1$ product has to be from the v_i side of the equation, and the codegree $d+1$ monomial of μ_1 is the codegree d prefix of a codegree 1 monomial in v_1, times the (prefix) period of the top monomial of x_1, which is the degree d suffix of v_1. If such a codegree $d+1$ product cancels with a codegree $d+1$ product that contains the top monomial of μ_1, then it must contain a codegree 1 monomial of x_1.

By the techniques that we used in the proofs of Theorem 4.4 and in the first part of Theorem 4.7, there exists an element s_1, $\deg(s_1)=d$, with a top monomial μ_1, such that $\mu_1 s_1 = u_1 = v_1 \mod G^{\deg(u_1)-2}$.

Suppose that a codegree $d+1$ product contains a codegree $d+1$ monomial of u_2 or v_2. Since the top monomial of u_2 (and v_2) contains no periodicity, such a product can cancel only with either

(1) a codegree $d+1$ product that contains a codegree 1 monomial of u_2 or v_2 and the top monomial of either μ_1 or μ_3;

(2) a codegree $d+1$ product that contains the top monomial of μ_2, and a codegree 1 monomial in the second appearance of x_1, and the top monomial of μ_1, the top monomial of u_2, and the same codegree 1 monomial in the second appearance of x_1;

(3) a codegree $d+1$ product that contains the top monomial of τ_2, and a codegree 1 monomial in the first appearance of x_1, and the top monomial of μ_3, the top monomial of v_2, and the same codegree 1 monomial in the first appearance of x_1.

Note that the two products that appear in possibilities (2) and (3) cancel each other. Hence, a codegree $d+1$ product that contains a codegree $d+1$ product that appears in u_2 or v_2, but not both, must cancel with a unique codegree $d+1$ product that is described in (1).

Suppose that a codegree $d+1$ product contains the top monomial of μ_1 and a codegree 1 monomial of u_2. Since the top monomial of u_2 (and v_2) has no periodicity, it can cancel only with a codegree $d+1$ product that contains either

(1) a codegree 1 monomial of v_2 and the top monomial of μ_3;

(2) a codegree $d+1$ monomial of u_2 or v_2;

(3) a codegree 1 monomial of the first appearance of x_1 and the top monomial of μ_2.

Similarly, suppose that a codegree $d+1$ product contains the top monomial of μ_3 and a codegree 1 monomial of v_2. It can cancel only with a codegree $d+1$ product that contains either

(1) a codegree 1 monomial of u_2 and the top monomial of μ_1;

(2) a codegree $d+1$ monomial of u_2 or v_2;

(3) a codegree 1 monomial of the second appearance of x_1 and the top monomial of τ_2.

Furthermore, a codegree $d+1$ product that contains the top monomial of μ_2 cannot cancel with a codegree $d+1$ product that contains the top monomial of τ_2.

Hence, we can look at the collection of codegree $d+1$ products that contain the top monomial of μ_1 and the entire collection of codegree 1 monomials of u_2. Each such product cancels with precisely one product that contains either a codegree $d+1$ monomial of u_2 or v_2, or a codegree 1 monomial of the first appearance of x_1 and the top monomial of μ_2, or a codegree 1 monomial of v_2 and the top monomial of μ_3. A similar statement holds for codegree $d+1$ products that contain a codegree 1 monomial of v_2 and the top monomial of μ_3.

Therefore, there exist elements $t_1, s_2, b, w_1, w_2, \tau_2, \mu_2$ such that

(1) $t_1\mu_2 = v_2$ and $\tau_2 s_2 = v_2 \mod G^{\deg(u_2)-2}$, $\deg(s_2) = \deg(t_1) = d$.

(2) $b + \tau_2 = u_2$ and $b + \mu_2 = v_2 \mod G^{\deg(u_2)-(d+2)}$.

(3) $x_1 = s_1 w_1 = w_1 t_1 = s_2 w_2 = w_2 t_2 \mod G^{\deg(x_1)-2}$.

We continue by induction for $1 \le r \le d$, and assume that for $r < d$ there exist elements $t_1, s_2, b, w_1, w_2, \tau_2, \mu_2$ such that the equalities that were true for the top 2 homogeneous parts and codegree d and codegree $d+1$ monomials hold for the top r monomials, and for codegree $d+r-1$ monomials,

(1) $t_1\mu_2 = v_2$ and $\tau_2 s_2 = v_2 \mod G^{\deg(u_2)-r}$, $\deg(s_2) = \deg(t_1) = d$.

(2) $b + \tau_2 = u_2$ and $b + \mu_2 = v_2 \mod G^{\deg(u_2)-(d+r)}$.

(3) $x_1 = s_1 w_1 = w_1 t_1 = s_2 w_2 = w_2 t_2 \mod G^{\deg(x_1)-r}$.

We continue by studying codegree $d+r$ products. All such products that involve only monomials of codegree less than d of the u_i, v_i, $1 \le i \le 3$, cancel in pairs. All such products that involve only monomials of codegree less than $d+r$ of the u_i, v_i, $1 \le i \le 3$, and codegree less than r of x_1 (in its two appearances from both sides of the equation) cancel in pairs by the induction hypothesis.

Hence, to analyze the structure of u_1 and v_1 (and hence, of μ_1 and s_1) we only need to consider codegree $d+r$ products that contain

(i) a codegree $d+r$ monomial of u_1 that does not appear in v_1 and vice versa;

(ii) a codegree r monomial of v_1 and the top monomial of μ_2;

(iii) a codegree $d+q_j$ monomial of μ_1, $q_j < r$, and a codegree $r-q_j$ monomial of the first appearance of x_1;

(iv) a codegree p_j monomial of v_1, $p_j < r$, and a codegree $r-p_j$ monomial of the first appearance of x_1 and the top monomial of μ_2.

A product of type (iv) that cancels with products of type (i) or (ii) must cancel with a corresponding product of type (iii) by our induction hypothesis. A product of type (iii) that cancels with a product of type (i) or (ii) and in which q_j is positive, and the codegree $r-q_j$ monomial of the first appearance of x_1 is obtained as a product of a codegree $r-m_j$ monomial of s_1 with a codegree m_j-q_j monomial of w_1, for $q_j < m_j < r$, cancels with a product of type (iv).

Therefore, to analyze the structure of u_1, v_1, s_1 and w_1, we consider only those codegree $d+r$ products that can be presented either in form (i) or (ii), that we denote (1) and (2) in the sequel, or in the form

(3) a product of the top monomial of μ_1, and a codegree r monomial of the first appearance of x_1.

A codegree $d+r$ product that can be presented in one of the forms (1)–(3) can cancel with either

(4) an odd number of products of a codegree $d+q_j$ monomial of μ_1, $0 < q_j < r$, and a codegree $r-q_j$ monomial of the first appearance of x_1;

(5) an odd number of products of a codegree $d+q_j$ monomial of μ_1, $0 < q_j < r$, and a product of a codegree $r-q_j$ monomial of s_1 with the top monomial of w_1;

(6) an odd number of products of a codegree $d+q_j$ monomial of μ_1, $q_j < r$, and a product of a codegree $r-m_j$ monomial of s_1 with a codegree m_j-q_j monomial of w_1, where $q_j < m_j < r$;

(7) an odd number of products of a codegree p_j monomial of v_1, $0 < p_j < r$, and a codegree $r-p_j$ monomial of the first appearance of x_1 and the top monomial of μ_2;

(8) a product of the top monomial of v_1, a codegree r monomial of the first appearance of x_1 and the top monomial of μ_2;

(9) an odd number of products of a codegree p_j monomial of v_1, $0 < p_j < r$, and a codegree m_j monomial of the first appearance of x_1, $0 < m_j$, $p_j + m_j < r$ and a codegree $d + r - p_j - m_j$ monomial of μ_2;

(10) an odd number of products of a codegree $d + q_j$ monomial of μ_1, a codegree m_j monomial of the first appearance of x_1, $0 < m_j$, $q_j + m_j < r$, and a codegree $r - m_j - q_j$ monomial of u_2.

If (1) or (2) occur, (8) cannot occur, and (6) occurs if and only if (7) occurs as well. If (1) occurs, (3) cannot occur. Suppose that (1) occurs. If in addition only (2) occurs, we add a codegree $d + r$ monomial to μ_1. If in addition to (1) only (4) and (5) occur, we also add a codegree $d + r$ monomial to μ_1. If in addition to (1) only (5), (6) and (7) occur, we add a codegree $d + r$ monomial to μ_1. If (1) occurs, (9) and (10) cannot occur.

Suppose that (2) occurs. If in addition only (3) occurs (and in addition possibly (4), (6) and (7)) we add a codegree r monomial to s_1. If in addition to (2) only (4) and (5) occur, we do not add anything. If in addition to (2) only (5), (6) and (7) occur, we do the same. If (2) occurs, (8)–(10) cannot occur.

Suppose that (3) occurs. The codegree r monomial of x_1 cannot be presented both as a product of the top monomial of s_1 with a codegree r monomial of w_1, and as a codegree r monomial of s_1 with the top monomial of w_1. We look at all the possible ways to present the codegree r monomial of x_1 as a product of a codegree q_j monomial of s_1 with a codegree $r - q_j$ monomial of w_1, for $0 < q_j < r$. If the number of such products is odd we don't add anything. If the number is even, we either add a codegree r monomial to s_1 or a codegree r monomial to w_1 (but not both). The validity of this addition of a codegree r monomial to either s_1 or w_1 can be verified by going over the possible cancellation of the given codegree $d + r$ product with all the other possible forms of such a product.

This concludes the adaptation of s_1, μ_1 and w_1 to include codegree r monomials. The same adaptation works for t_2, μ_3 and w_2. It is still left to analyze u_2 and v_2 in order to add codegree r monomials to μ_2 and τ_2 such that the equalities that by induction hold for the top codegree $r - 1$ parts of these elements will hold for the top codegree r part.

To analyze the structure of u_2 and v_2 (and hence, of μ_2, τ_2, t_1 and s_2) we start by observing the following:

(i) The codegree $d + r$ products that contain either a positive codegree monomial of u_1 or a positive codegree monomial of the first appearance of x_1, a codegree $d + q_j$ monomial of τ_2, $q_j < r$, a monomial of the second appearance of x_1, and a monomial of u_3, cancel with codegree $d + r$ products that contain either a positive codegree monomial of v_1 or a positive codegree monomial of the first appearance of x_1, a codegree p_j monomial of v_2, $p_j < r$, a monomial of the second appearance of x_1, and a monomial of μ_3.

(ii) The codegree $d+r$ products that contain a monomial of v_1, a monomial of the first appearance of x_1, a codegree $d+q_j$ monomial of μ_2, $q_j < r$, and either a positive codegree monomial of the second appearance of x_1, or a positive codegree monomial of v_3, cancel with codegree $d+r$ products that contain a monomial of μ_1, a monomial of the first appearance of x_1, a codegree p_j monomial of u_2, $p_j < r$, either a positive codegree monomial of the second appearance of x_1, or a positive codegree monomial of u_3.

Hence, to analyze the structure of u_2 and v_2 we only need to consider codegree $d+r$ products that contain

(i) a codegree $d+r$ monomial of u_2 or of v_2;

(ii) the top monomial of μ_1 and a codegree r monomial of u_2 or a codegree r monomial of v_2 and the top monomial of μ_3;

(iii) a codegree $d+q_j$ monomial of τ_2, $q_j < r$, and a codegree $r-q_j$ monomial of the second appearance of x_1 or a codegree $r-q_j$ of the first appearance of x_1 and a codegree $d+q_j$ monomial of μ_2, $q_j < r$;

(iv) the top monomial of μ_1, a codegree $r-p_j$ monomial of the first appearance of x_1 and a codegree p_j monomial of u_2, $p_j < r$, or a codegree p_j monomial of v_2, $p_j < r$ and a codegree $r-p_j$ monomial of the second appearance of x_1 and the top monomial of μ_3.

If there are two products of codegree $d+r$ of type (i), they cancel each other, and we can ignore them in analyzing codegree $d+r$ products. Therefore, to analyze the structure of u_2, v_2, s_2, t_1, μ_2 and τ_2, we consider only those codegree $d+r$ products that can be presented either in form (i) or (ii), that we denote (1) and (2) in the sequel, or in codegree $d+r$ products in the form

(3) a product of the top monomial of τ_2, and a codegree r monomial of the second appearance of x_1;

(4) a codegree r monomial of the first appearance of x_1, and the top monomial of μ_2;

(5) an odd number of products of a codegree $d+q_j$ monomial of τ_2, $0 < q_j < r$, and a codegree $r-q_j$ monomial of the second appearance of x_1;

(6) an odd number of products of a codegree $d+q_j$ monomial of τ_2, $0 < q_j < r$, and a product of a codegree $r-q_j$ monomial of s_2 with the top monomial of w_2;

(7) an odd number of products of a codegree $d+q_j$ monomial of τ_2, $q_j < r$, and a product of a codegree $r-m_j$ monomial of s_2 with a codegree $m_j - q_j$ monomial of w_2, where $q_j < m_j < r$;

(8) an odd number of products of a codegree p_j monomial of v_2, $0 < p_j < r$, and a codegree $r - p_j$ monomial of the second appearance of x_1 and the top monomial of μ_3;

(9) a product of the top monomial of v_2, a codegree r monomial of the second appearance of x_1 and the top monomial of μ_3.

And similarly, from the other sides of the equation,

(10) an odd number of products of a codegree $r-q_j$ monomial of the first appearance of x_1, and a codegree $d+q_j$ monomial of μ_2, $0 < q_j < r$;

(11) an odd number of products of the top monomial of w_1, a codegree $r - q_j$ monomial of t_1, and a codegree $d+q_j$ monomial of μ_2, $0 < q_j < r$;

(12) an odd number of products of a codegree $m_j - q_j$ monomial of w_1, a codegree $r-m_j$ monomial of t_1, a codegree $d+q_j$ monomial of μ_2, $q_j < r$, $q_j < m_j < r$;

(13) an odd number of products of the top monomial of μ_1, a codegree $r - p_j$ monomial of the first appearance of x_1, and a codegree p_j monomial of u_2, $0 < p_j < r$;

(14) a product of the top monomial of μ_1, a codegree r monomial of the first appearance of x_1, and the top monomial of u_2.

Suppose that (1) occurs. If only one of the possibilities in (2) occurs, we add a codegree $d+r$ monomial to μ_2 or τ_2, depending which of the two possibilities in (2) occurs. If (1) occurs, (3) and (4) cannot occur. If in addition to (1) only (5) occurs, then (6) or (7) must occur and not both. If only (5) and (6) occur, we add a codegree $d+r$ monomial to τ_2. If in addition to (1), (5) and (7) occur, then (8) must occur as well, and hence at least an additional possibility must occur. If in addition to (1), (8) occurs, then (5) and (7) must occur as well, so an additional possibility must occur. If (1) occurs, (9) cannot occur. The possibilities (10)–(14) are parallel to (5)–(9) and are dealt with accordingly.

Suppose that (1) and the two possibilities in (2) occur. If in addition only (5) and (6) occur, we add a codegree $d+r$ monomial only to μ_2, and if only (10) and (11) occur, we add a codegree $d+r$ monomial to τ_2. Suppose that (1) and only one of the products in the form (2) occur, without loss of generality the product from the v_i side, i.e., the one that contains μ_3. If in addition (5), (6), (10) and (11) occur, we add a codegree $d+r$ monomial to μ_2.

Suppose that one of the possibilities in (2) occurs, without loss of generality the one from the v_i side. If the only additional product that cancels with it is also a product in form (2) from the u_i side of the equation, we add a codegree $d+r$ monomial to both τ_2 and μ_2. If in addition to the form (2) only possibility (3) occurs, we add a codegree r monomial to s_2. Form (4) cannot occur. If only (5) and (6) occur, we do not add anything. If (5) and (7) occur, (8) must occur as well. Form (9) cannot occur. If in addition (10) and (11) occur, we add a codegree $d+r$ monomial to both τ_2 and μ_2. If only (3), (5), (6), (10) and (11) occur, we add a codegree r monomial to s_2, and a codegree $d+r$ monomial to both μ_2 and τ_2.

Suppose that the two possibilities in part (2) occur. In that case (3) cannot occur. If in addition (5), (6), (11) and (12) occur, we do not add anything. Suppose that (3) occurs. In that case (4) cannot occur. If in addition only (5) and (6) occur, we add a codegree r monomial to s_2. If in addition to (3) only (9) occurs, we add a codegree r monomial to w_2. If (3) occurs, then (10)–(14) cannot occur. If (4) occurs the analysis is analogous to the case in which (3) occurs.

Suppose that (5) and (6) occur. In that case (9) cannot occur. If (10) and (11) occur as well, we add a codegree $d+r$ monomial to both τ_2 and μ_2.

This concludes our treatment of codegree $d+r$ products for $r<d$. So far we proved that

(1) $\mu_1 s_1 = u_1 = v_1 \bmod G^{\deg(u_1)-d}$, $\deg(s_1)=d$, $u_1 = v_1+\mu_1 \bmod G^{\deg(u_1)-2d}$.

(2) $t_1\mu_2 = v_2$ and $\tau_2 s_2 = v_2 \bmod G^{\deg(u_2)-d}$, $\deg(s_2)=\deg(t_1)=d$, $\deg(\mu_2) = \deg(\tau_2)=\deg(u_2)-d$.

(3) $b_2+\tau_2 = u_2$ and $b_2+\mu_2 = v_2 \bmod G^{\deg(u_2)-(2d)}$.

(4) $x_1 = s_1 w_1 = w_1 t_1 = s_2 w_2 = w_2 t_2 \bmod G^{\deg(x_1)-d}$.

We continue by analyzing codegree $2d$ products. The analysis of codegree $2d$ products is similar to the analysis of codegree $d+r$ products for $r<d$. In their analysis we use the following observations:

(i) All the codegree $2d$ products that contain monomials of codegree smaller than d from the elements u_i, v_i and x in its two appearances cancel in pairs.

(ii) All the codegree $2d$ products that contain a monomial of codegree bigger than d, from b_1, b_2 or b_3, cancel in pairs.

Hence, we need to analyze only those codegree $2d$ products that contain monomials from either $\mu_1, \mu_2, \tau_2, \mu_3$, or monomials of codegree d from b_1, b_2, b_3. To analyze the elements $u_1, v_1, b_1, \mu_1, s_1$ and w_1, we need to analyze codegree $2d$ products that contain one of the following:

(i) a codegree $2d$ monomial of u_1 that does not appear in v_1 and vice versa;

(ii) a codegree d monomial of v_1 and the top monomial of μ_2;

(iii) a codegree $d+q_j$ monomial of μ_1, $q_j<d$, and a codegree $d-q_j$ monomial of the first appearance of x_1;

(iv) a codegree p_j monomial of v_1, $p_j < d$, and a codegree $d - p_j$ monomial of the first appearance of x_1 and the top monomial of μ_2.

(v) Note that the codegree $2d$ product that contains the top monomials of μ_1 and τ_2 cancels with the product that contains the top monomials of μ_2 and μ_3. Also the codegree $2d$ products that contain a codegree d monomial of u_1 which is from b_1 (i.e., also a monomial of v_1), and the top monomial of τ_2, cancel with the products that contain the same codegree d monomial from v_1, and the top monomial of μ_3.

Because of (v), the analysis of codegree $2d$ monomials of u_1 and v_1 is identical to the analysis of codegree $d+r$ monomials of these elements. This concludes the construction of the element s_1, and adds codegree $2d$ monomials to μ_1, and codegree d monomials to w_1. The analysis of the elements u_3, v_3, b_3, μ_3 and w_2 is identical.

We continue by analyzing the codegree $2d$ monomials in u_2, v_2, τ_2 and μ_2. The observations (i) and (ii) that we used in analyzing the codegree $d+r$ monomials of these elements for $r < d$ remain valid for codegree $2d$ monomials. In addition by part (v) in the analysis of codegree $2d$ monomials of u_1 and v_1, it follows that the codegree $2d$ product that contains the top monomials of μ_1 and τ_2 cancels with the product that contains the top monomials of μ_2 and μ_3. Hence, the rest of the analysis of codegree $2d$ monomials of u_2 and v_2 is identical to the analysis of codegree $d+r$ monomials of these elements for $r < d$.

We continue by analyzing higher codegree products and monomials. We assume inductively for $r > 0$ that

(1) $\mu_1(s_1+1) = u_1$ and $\mu_1 s_1 = v_1 \mod G^{\deg(u_1)-(d+r)}$, $\deg(s_1) = d$, $u_1 = v_1 + \mu_1 \mod G^{\deg(u_1)-(2d+r)}$.

(2) $(t_1+1)\mu_2 = v_2$, $\tau_2(s_2+1) = u_2$ and $t_1\mu_2 = \tau_2 s_2 \mod G^{\deg(u_2)-(d+r)}$, $\deg(s_2) = \deg(t_1) = d$, $\deg(\mu_2) = \deg(\tau_2) = \deg(u_2) - d$.

(3) $b_2 + \tau_2 = u_2$ and $b_2 + \mu_2 = v_2 \mod G^{\deg(u_2)-(2d+r)}$.

(4) $x_1 = s_1 w_1 = w_1 t_1 = s_2 w_2 = w_2 t_2 \mod G^{\deg(x_1)-(d+r)}$.

And we continue by analyzing codegree $2d+r$ products. The analysis is similar to the analysis of codegree $d+r$ and codegree $2d$ products. We use the following observations:

(i) All the codegree $2d+r$ products that contain monomials of u_i, v_i, $i = 1, 2, 3$, that are all of codegree smaller than d cancel in pairs. In particular, all the codegree $2d+r$ products that contain a monomial of x of codegree bigger than $d+r$, in one of its two appearances, cancel in pairs.

(ii) All the codegree $2d+r$ products that contain monomials from all b_1, b_2 and b_3 cancel in pairs.

(iii) A codegree $2d+r$ product that contains a monomial from μ_1 of codegree more than d, and a monomial from the first appearance of x, such that the sum of their codegrees is less than $2d+r$, an element from b_2 and an element from b_3, cancels with a product that contains an element from b_1, an element from the first appearance of x, an element from μ_2 and the same element from b_3. The same holds for products that contain monomials from b_1, b_2 and μ_3 with parallel restrictions.

(iv) A codegree $2d+r$ product that contains a monomial from b_1, a monomial from τ_2 of codegree bigger than d, and a monomial from b_3, such that the sum of the codegrees of the monomial from τ_2 and the monomial from the second appearance of x is smaller than $2d+r$, cancels with a product that contains the same monomials of b_1 and the first appearance of x, a monomial from b_2 and a monomial from μ_3. The same holds for products that contain monomials from b_1, μ_2 and b_3 with parallel restrictions.

(v) A codegree $2d+r$ product that contains a monomial from μ_1 of codegree bigger than d, a monomial from τ_2, and a monomial from b_3, cancels with a product that contains a monomial from b_1, a monomial from μ_2, and a monomial from μ_3. The same holds for products that contain monomials from b_1, μ_2 and μ_3 with parallel restrictions.

Hence, like in the analysis of codegree $2d$ products, to analyze the elements u_1, v_1, b_1, μ_1 and w_1, we need to analyze codegree $2d+r$ products that contain one of the following:

(i) a codegree $2d+r$ monomial of u_1 that does not appear in v_1 and vice versa;

(ii) a codegree $d+r$ monomial of v_1 (which is a monomial of b_1) and the top monomial of μ_2;

(iii) a codegree $d+q_j$ monomial of μ_1, $q_j < d+r$, and a codegree $d+r-q_j$ monomial of the first appearance of x_1;

(iv) a codegree p_j monomial of v_1 (which is a monomial of b_1), $p_j < d+r$, and a codegree $d+r-p_j$ monomial of the first appearance of x_1 and the top monomial of μ_2.

Hence, the analysis of codegree $2d+r$ monomials of u_1 and v_1 is identical to the analysis of codegree $d+r$ and $2d$ monomials of these elements. Note that in analyzing products of codegree greater than $2d+r$, the element s_1 is already fixed, and we only add codegree $2d+r$ monomials to μ_1 and b_1, and codegree $d+r$ monomials to w_1. The analysis of the elements u_3, v_3, b_3, μ_3 and w_2 is identical.

We continue by analyzing the codegree $2d$ monomials in u_2, v_2, τ_2 and μ_2. The observations (i)–(v) that we used in analyzing the codegree $2d+r$ monomials of b_1 and μ_1 imply that analyzing codegree $2d+r$ monomials of b_2, τ_2 and μ_2 is similar to the analysis of the codegree $d+r$ monomials of these elements. Hence, we can finally deduce that

(1) $\mu_1(s_1+1)=u_1$ and $\mu_1 s_1=v_1$, $\deg(s_1)=d$ and $u_1=v_1+\mu_1$.

(2) $(t_1+1)\mu_2=v_2$, $\tau_2(s_2+1)=u_2$ and $t_1\mu_2=\tau_2 s_2$, $\deg(s_2)=\deg(t_1)=d$ and $\deg(\mu_2)=\deg(\tau_2)=\deg(u_2)-d$.

(3) $b_2+\tau_2=u_2$ and $b_2+\mu_2=v_2$.

(4) $x_1=s_1w_1=w_1t_1=s_2w_2=w_2t_2 \mod G^{\deg(x_1)-\deg(u_1u_2u_3)}$.

This proves the structure of the coefficients in the statement of Proposition 4.11. Suppose that there exists a solution x_2 to the given equation, and

$$\deg(x_2)>2(2+k^{\deg(s)+2}+\deg(u_1)+\deg(u_2)+\deg(u_3)).$$

As in the analysis of the same equation in case there are shifts between the appearances of the element x_2, we can continue the analysis of higher codegree monomials of the solution x_2, and get that there exist elements w_i, $i=1,2$, that satisfy $s_iw_i=w_it_i=x_1 \mod G^{\deg(s_1)-1}$. By the argument that was used to prove Lemma 4.2, it follows that there exists a solution $\hat{x}$ to the equations $s_ix=xt_i$, $i=1,2$.

The element x_2 satisfies $s_1x_2=x_2t_1 \mod G^{2\deg(s_1)-1}$. Hence, there exists an element $\hat{x}_2$, which is a solution of the equation $s_1x=xt_1$, and $x_2+\hat{x}_2=r$, where $\deg(r)\leq 2+k^{\deg(s_1)+2}$.

Also, x_2 is a solution to the equation $v_1xv_2xv_3=u_1xu_2xu_3$, where $v_1=\tau_1s_1$, $u_1=\tau_1(s_1+1)$, $v_2=(t_1+1)\mu_2$, $u_2=\tau_2(s_2+1)$, $v_3=(t_2+1)\mu_3$, $u_3=t_2\mu_3$, and $\tau_2s_2=t_1\mu_2$. Hence

$$\tau_1(s_1+1)(\hat{x}_2+r)\tau_2(s_2+1)(\hat{x}_2+r)t_2\mu_3=\tau_1s_1(\hat{x}_2+r)(t_1+1)\mu_2(\hat{x}_2+r)(t_2+1)\mu_3.$$

Therefore

$$\begin{aligned}(s_1+1)r\tau_2(s_2+1)\hat{x}_2t_2+(s_1+1)\hat{x}_2\tau_2(s_2+1)rt_2\\ =s_1r(t_1+1)\mu_2\hat{x}_2(t_2+1)+s_1\hat{x}_2(t_1+1)\mu_2r(t_2+1) \mod G^{\deg(s_1r\tau_2s_2rt_2)}.\end{aligned}$$

Since $s_1\hat{x}_2=\hat{x}_2t_1$, this implies

$$\begin{aligned}((s_1+1)r\tau_2(s_2+1)s_2+s_1r(t_1+1)\mu_2(s_2+1))\hat{x}_2\\ =\hat{x}_2((t_1+1)\tau_2(s_2+1)rt_2+t_1(t_1+1)\mu_2r(t_2+1)) \mod G^{\deg(s_1r\tau_2s_2rt_2)}.\end{aligned}$$

Therefore,

$$(s_1+1)r\tau_2(s_2+1)s_2+s_1r(t_1+1)\mu_2(s_2+1)=p(s_1),$$
$$(t_1+1)\tau_2(s_2+1)rt_2+t_1(t_1+1)\mu_2r(t_2+1)=p(t_1)$$

for some polynomial p.

This implies that $r\tau_2 s_2 + s_1 r\mu_2$ is a polynomial in s_1, and $\tau_2 r t_2 + t_1\mu_2 r$ is a polynomial in t_1. Hence, $(rt_1 + s_1 r)\mu_2$ is a polynomial in s_1, and $\tau_2(rt_2 + s_2 r)$ is a polynomial in t_1.

Since we assumed that the top monomials of the coefficient do not contain periodicity, it cannot be that the top monomials of τ_2 and μ_2 are equal, and equal to the top monomials of t_1 and s_1. Hence, $rt_1 + s_1 r = rt_2 + s_2 r \neq 1$.

If $\deg(\tau_2) = \deg(s_1)$, then the top monomials of s_1 and t_1 are equal, and the top monomials of u_2 and v_2 have periodicity, a contradiction. The top monomial of τ_2 has no periodicity, so $\deg(\tau_2) < 2\deg(s_1)$. If $\deg(\tau_2) > \deg(s_1)$, then necessarily the top monomials of u_2 and v_2 contain periodicity, a contradiction.

Suppose that $\deg(\tau_2) < \deg(s_1)$. If the top monomial of τ_2 is the same as the top monomial of μ_2, then the top monomials of the two sides of the equation contain periodicity, a contradiction. If the top monomials of μ_2 and τ_2 are distinct, then the top monomials of u_2 and v_2 contain periodicity, a contradiction.

Therefore, $rt_1 + s_1 r = 0$, so r is a solution of the equation $s_1 x = x t_1$ and so is $x_2 = \hat{x}_2 + r$, and the conclusion of Proposition 4.11 follows. □

Proposition 4.11 and its proof enable us to prove Theorem 4.7 in case there are no shifts, i.e., in case the degrees of the elements u_i, v_i satisfy $\deg(u_i) = \deg(v_i)$ for all indices i.

Proposition 4.12. *Let $u_1, \dots, u_n, v_1, \dots, v_n \in FA$, where FA is the free associative algebra over GF_2 that is generated by k elements, and suppose that the equation*

$$u_1 x u_2 x u_3 \cdots u_{n-1} x u_n = v_1 x v_2 x v_3 \cdots v_{n-1} x v_n$$

has a solution x_1 of degree bigger than $2(\deg(u_1)+\cdots+\deg(u_n))^2$. Suppose further that

1. *For every index i, $1 \le i \le n$, $\deg(u_i) = \deg(v_i)$.*
2. *The top homogeneous parts of u_i and v_i are monomials with no periodicity.*
3. *For some index i, $u_i \neq v_i$.*
4. *All the periodicity in the top monomials that are associated with the top monomials of the two sides of the equation after substituting the solution x_1 is contained in the periodicity of the top monomial of the solution x_1.*

Then there exist some elements $s, t \in FA$, $\deg(s) = \deg(t) < \min\deg(u_i)$, such that

1. *Every solution of the equation $sx = xt$ is a solution of the given equation.*
2. *Every solution x_2 of the given equation that satisfies*

$$\deg(x_2) > 2(2 + k^{\deg(s)+2} + \deg(u_1) + \cdots + \deg(u_n))$$

   is also a solution of the equation $sx = xt$.

(3) *For every index* i, $1 \le i \le r$, *for which* $u_i \neq v_i$, *there exist elements* τ_i, μ_i *such that the elements* u_i, v_i *are either* $\tau_i(s_i+1)$ *or* $(t_{i-1}+1)\mu_i$ *or* $\tau_i s_i$ *or* $t_{i-1}\mu_i$, *where the elements* s_i *are either* s *or* $s+1$, *and the elements* t_i *are either* t *or* $t+1$, *and* $t_{i-1}\mu_i = \tau_i s_i$.

Proof. The proof of the structure of the coefficients is similar to the proof of Proposition 4.11. Given the structure of the coefficients, it is clear that every solution of the equation $sx = xt$ is a solution of the given equation. It is left to prove that every long enough solution of the given equation is a solution of the equation $sx = xt$.

Suppose that x_2 is a solution of the given equation that satisfies

$$\deg(x_2) > 2(2 + k^{\deg(s)+2} + \deg(u_1) + \cdots + \deg(u_n)).$$

By the argument that we used in Proposition 4.11, it follows that the equation $sx = xt$ has a solution, and that $x_2 = \hat{x}_2 + r$, where $\hat{x}_2$ is a solution to the equation $sx = xt$, and $\deg(r) \le 2 + k^{\deg(s)+2}$.

In that case we get the equality

$$\begin{aligned}\tau_1(s_1+1)(\hat{x}_2+r)\tau_2(s_2+1)(\hat{x}_2+r)\cdots\tau_{n-1}(s_{n-1}+1)(\hat{x}_2+r)t_{n-1}\mu_n \\ = \tau_1 s_1(\hat{x}_2+r)(t_1+1)\mu_2(\hat{x}_2+r)\cdots(t_{n-2}+1)\mu_{n-1}(\hat{x}_2+r)(t_{n-1}+1)\mu_n,\end{aligned}$$

and since $\hat{x}_2$ is a solution of the equation $s_1 x = x t_1$, we get the equality

$$\begin{aligned}&(s_1+1)r\tau_2(s_2+1)\hat{x}_2\cdots\tau_{n-1}(s_{n-1}+1)\hat{x}_2 t_{n-1} \\ &\qquad + \cdots + (s_1+1)\hat{x}_2\tau_2(s_2+1)\hat{x}_2\cdots\tau_{n-1}(s_{n-1}+1)r t_{n-1} \\ &\quad = s_1 r(t_1+1)\mu_2\hat{x}_2\cdots(t_{n-2}+1)\mu_{n-1}\hat{x}_2(t_{n-1}+1) \\ &\qquad + \cdots + s_1\hat{x}_2(t_1+1)\mu_2\hat{x}_2\cdots(t_{n-2}+1)\mu_{n-1}r(t_{n-1}+1) \mod G^{m_2},\end{aligned}$$

where

$$m_1 = \deg((s_1+1)r\tau_2(s_2+1)\hat{x}_2\cdots\tau_{n-1}(s_{n-1}+1)\hat{x}_2 t_{n-1})$$

and $m_2 = m_1 - \deg(\hat{x}_2) + \deg(r)$.

By the same argument that we used in the proof of Proposition 4.11, since the top monomials of the coefficients u_i, v_i, $i = 1, \ldots, n$, do not have periodicity, and since the top monomial in the two sides of the equation after substituting the solution x_2 has no periodicity, except the one that is contained in the appearances of the top monomial of x_2, it follows that for some i, $1 \le i \le n-1$, $s_i r = r t_i$. Hence, r is a solution to the equation, $sx = xt$, and so is $x_2 = \hat{x}_2 + r$, since both $\hat{x}_2$ and r are solutions to this equation. □

At this point we need to consider equations in which some of the appearances of the elements x are shifted, and some are not.

Lemma 4.13. *Let* $u_1, u_2, u_3, v_1, v_2, v_3 \in FA$ *satisfy* $u_1 \neq v_1$, $\deg(u_1) = \deg(v_1)$, $\deg(u_2) > \deg(v_2)$, $\deg(v_3) > \deg(u_3)$, *where FA is the free associative algebra over* GF_2 *that is generated by k elements.*

Suppose that the top homogeneous parts of u_i *and* v_i *are monomials* (*for* $i = 1, 2, 3$) *with no nontrivial periodicity. If there exists a solution* x_1 *to the equation* $u_1 x u_2 x u_3 = v_1 x v_2 x v_3$, *and the only nontrivial periodicity in the top monomials of the two sides of the equation after substituting* x_1 *is contained in the top monomial of the solution* x_1 (*this translates to a condition on the top monomials of the coefficients*), *and* $\deg(x_1) > 2(\deg(u_1) + \deg(u_2) + \deg(u_3))^2$, *then there exist elements* $s, t \in FA$, *such that either*

(1) *There exists* μ_1 *for which* $u_1 = \mu_1(s+1)$ *and* $v_1 = \mu_1 s$.

(2) *There exist* μ_2 *and* s_2, t_2 *for which* $(t+1)\mu_2 = v_2$ *and* $t\mu_2 s_2 = u_2$. *Furthermore,* $v_3 = t_2 u_3$ *and the pair* (s_2, t_2) *is either* (s, t) *or* $(s+1, t+1)$.

or

(1) *There exists* μ_1 *for which* $u_1 = \mu_1 s$ *and* $v_1 = \mu_1(s+1)$.

(2) *There exist* μ_2 *and* s_2, t_2 *for which* $(t+1)\mu_2 = u_2$ *and* $v_2 s_2 = t\mu_2$. *Furthermore,* $v_3 = t_2 u_3$ *and the pair* (s_2, t_2) *is either* (s, t) *or* $(s+1, t+1)$.

As in the conclusion of Theorem 4.7, every solution of the equation $sx = xt$ *is a solution of the given equation* $u_1 x u_2 x u_3 = v_1 x v_2 x v_3$. *Every solution* x_2 *of the given equation* $u_1 x u_2 x u_3 = v_1 x v_2 x v_3$ *that satisfies*

$$\deg(x_2) > 2(2 + k^{\deg(s_1)+2} + \deg(u_1) + \deg(u_2) + \deg(u_3))$$

is also a solution of the equation $sx = xt$.

Proof. The proof is similar to the proof of Proposition 4.11. □

At this point we can complete the proof of Theorem 4.7. We already analyzed the case in which there are nontrivial shifts between (the top monomials of) pairs of appearances of the variable x in the two sides of the equation. Propositions 4.11 and 4.12 analyze the case in which there are no shifts between pairs of appearances of the variable x in the two sides of the equation, and Lemma 4.13 analyzes the case $n = 3$ in which there is a pair with no shift and a pair with a shift.

By the techniques that were used in proving Proposition 4.11 and in analyzing the case in which there are nontrivial shifts between pairs of appearances of the variable x, if there is a pair of coefficients, u_i, v_i such that $u_i = v_i$ and the $i-1$ (hence, also the i-th) pair of appearances of the variable x has no shift, then the equation breaks into two equations, the first contains the coefficients $u_1, \dots, u_{i-1}, v_1, \dots, v_{i-1}$, and the second contains the coefficients $u_{i+1}, \dots, u_n$, $v_{i+1}, \dots, v_n$. Therefore, in the sequel we may assume that there is no such pair of coefficients u_i, v_i.

Then there exist some elements $s, t \in FA$, $\deg(s) = \deg(t) < \min \deg(u_i)$, and elements $s_1, \dots, s_{n-1}, t_1, \dots, t_{n-1}$, such that

(1) For every index i, the pair (s_i, t_i) is either (s, t) or $(s+1, t+1)$.

(2) For every pair of coefficients, u_i, v_i for which the two pairs of appearances of the variable x from the two sides of the pair of coefficients have no nontrivial shift, either $u_i = v_i$ or there exist elements τ_i and μ_i such that either $u_i = \tau_i s_i$ and $v_i = \tau_i(s_i + 1)$ (or vice versa), or $u_i = t_{i-1}\mu_i$ and $v_i = (t_{i-1} + 1)\mu_i$ (or vice versa), or $u_i = (t_{i-1} + 1)\mu_i$ and $v_i = \tau_i(s_i + 1)$ (or vice versa).

(3) If $\deg(u_1) = \deg(v_1)$, either $u_1 = v_1$ or there exists τ_1 such that $u_1 = \tau_1 s_1$ and $v_1 = \tau_1(s_1 + 1)$ (or vice versa). If $\deg(u_n) = \deg(v_n)$, either $u_n = v_n$ or there exists an element μ_n such that $u_n = t_{n-1}\mu_n$ and $v_n = (t_{n-1} + 1)\mu_n$ (or vice versa).

(4) For every pair u_i, v_i for which the two pairs of appearances of the variable x from the two sides of the pair of coefficients have nontrivial shifts, $u_i s_i = t_{i-1} v_i$ (or vice versa), or $u_i = t_{i-1} v_i s_i$ (or vice versa).

(5) If $\deg(u_1) \neq \deg(v_1)$, then $u_1 = v_1 s_1$ or vice versa. If $\deg(u_n) \neq \deg(v_n)$, then $u_n = t_{n-1} v_n$ or vice versa.

(6) Suppose that $\deg(u_i) \neq \deg(v_i)$, $1 < i < n$, there is no shift between the $i-1$ appearances of the variable x, and there is a nontrivial shift between the i-th appearances of the variable x from the two sides of the equation. Then either $u_i s_i = v_i$ or vice versa, in which case the original equation can be broken into two equations, the first contains the first $i-1$ pairs of coefficients, and the second contains the last $n+1-i$ pairs of coefficients, or $v_i = (t_{i-1} + 1)\mu_i$ and $u_i s_i = t_{i-1}\mu_i$ (or vice versa), or $u_i = (t_{i-1} + 1)\mu_i$ and $v_i = t_{i-1}\mu_i s_i$ (or vice versa).

(7) Suppose that $\deg(u_i) \neq \deg(v_i)$, $1 < i < n$, there is no shift between the i-th appearances of the variable x, and there is a nontrivial shift between the $i-1$ appearances of the variable x from the two sides of the equation. Then either $t_{i-1} u_i = v_i$ or vice versa, in which case the original equation can be broken into two equations, the first contains the first i pairs of coefficients, and the second contains the last $n-i$ pairs of coefficients, or $v_i = \tau_i(s_i + 1)$ and $t_{i-1} u_i = \tau_i s_i$ (or vice versa), or $u_i = \tau_i(s_i + 1)$ and $v_i = t_{i-1}\tau_i s_i$ (or vice versa).

This description of the coefficients in a general equation with one variable, in which the coefficients have no periodicity, and the top homogeneous parts of the coefficients are monomials, finally implies:

(1) Every solution of the equation $sx = xt$ is a solution of the given equation.

(2) Every solution x_2 of the given equation that satisfies

$$\deg(x_2) > 2(2 + k^{\deg(s)+2} + \deg(u_1) + \cdots + \deg(u_n))$$

is also a solution of the equation $sx = xt$.

The proof of (1) follows from the structure of the coefficients, and the proof of (2) follows by the argument that was used to prove (2) for the case in which there are no shifts between the various appearances of the top monomial of the solution x_2 in the two sides of the given equation in Proposition 4.12.

This concludes the proof of Theorem 4.7. □

Acknowledgments

I would like to thank Eliyahu Rips, who inspired me years ago to try and study varieties over noncommutative algebras, and David Kazhdan for his continuing interest. I am grateful to the referee, who went carefully over the text that is not easy to follow and verify, and suggested quite a few changes that improved the paper.

References

[Amitsur 1966] S. A. Amitsur, "Rational identities and applications to algebra and geometry", *J. Algebra* **3** (1966), 304–359. MR Zbl

[Appel 1968] K. I. Appel, "One-variable equations in free groups", *Proc. Amer. Math. Soc.* **19** (1968), 912–918. MR Zbl

[Bergman 1969] G. M. Bergman, "Centralizers in free associative algebras", *Trans. Amer. Math. Soc.* **137** (1969), 327–344. MR Zbl

[Bestvina and Feighn 1995] M. Bestvina and M. Feighn, "Stable actions of groups on real trees", *Invent. Math.* **121**:2 (1995), 287–321. MR Zbl

[Casals-Ruiz and Kazachkov 2011] M. Casals-Ruiz and I. V. Kazachkov, "On systems of equations over free products of groups", *J. Algebra* **333** (2011), 368–426. MR Zbl

[Cohn 1971] P. M. Cohn, *Free rings and their relations*, London Math. Soc. Monographs **2**, Academic Press, London, 1971. MR Zbl

[Guba 1986] V. S. Guba, "Equivalence of infinite systems of equations in free groups and semigroups to finite subsystems", *Mat. Zametki* **40**:3 (1986), 321–324. In Russian; translated in *Math. Notes* **40**:3 (1986), 688–690. MR Zbl

[Lorenc 1963] A. A. Lorenc, "Solution of systems of equations in one unknown in free groups", *Dokl. Akad. Nauk SSSR* **148** (1963), 1253–1256. In Russian; translated in *Soviet Math. Dokl.* **4** (1963), 262–265. MR Zbl

[Makanin 1977] G. S. Makanin, "The problem of the solvability of equations in a free semigroup", *Mat. Sb.* (*N.S.*) **103**:2 (1977), 147–236. In Russian; translated in *Math. USSR-Sb.* **32**:2 (1977), 129–198. MR Zbl

[Malcev 1948] A. I. Malcev, "On the embedding of group algebras in division algebras", *Doklady Akad. Nauk SSSR* (*N.S.*) **60** (1948), 1499–1501. MR Zbl

[Neumann 1949] B. H. Neumann, "On ordered division rings", *Trans. Amer. Math. Soc.* **66** (1949), 202–252. MR Zbl

[Sela 2001] Z. Sela, "Diophantine geometry over groups, I: Makanin–Razborov diagrams", *Publ. Math. Inst. Hautes Études Sci.* 93 (2001), 31–105. MR Zbl

[Sela 2016] Z. Sela, "Word equations, I: Pairs and their Makanin–Razborov diagrams", preprint, 2016. arXiv 1607.05431

[Shelah 1990] S. Shelah, *Classification theory and the number of nonisomorphic models*, 2nd ed., Studies Logic Found. Math. **92**, North-Holland Publishing Co., Amsterdam, 1990. MR Zbl

Received 20 Mar 2022. Revised 28 Mar 2024.

ZLIL SELA:

zlil.sela@mail.huji.ac.il
Mathematics Department, Hebrew University, Jerusalem, Israel

Model Theory
Vol. 3, No. 3, 2024
https://doi.org/10.2140/mt.2024.3.801

When does $\aleph_1$-categoricity imply ω-stability?

John T. Baldwin, Michael C. Laskowski and Saharon Shelah

For an $\aleph_1$-categorical atomic class, we clarify the space of types over the unique model of size $\aleph_1$. Using these results, we prove that if such a class has a model of size $\beth_1^+$ then it is ω-stable.

1. Introduction

Our principal result is:

Theorem 1.1. *If an atomic class* At *is* $\aleph_1$*-categorical and has a model of size* $(2^{\aleph_0})^+$, *then* At *is* ω*-stable.*

This result springs from several related problems in the study of $L_{\omega_1,\omega}$: the role of $\beth_{\omega_1}$, the possible necessity of the weak continuum hypothesis, the absoluteness of $\aleph_1$-categoricity.

For first order logic, Morley [1965] proved, en route to his categoricity theorem, that an $\aleph_1$-categorical first order theory is ω-stable (né totally transcendental). The existence of a saturated Ehrenfeucht–Mostowski model of cardinality $\aleph_1$ that is generated by a well-ordered set of indiscernibles is crucial to the proof. The construction of such indiscernibles via the Erdős–Rado theorem and Ehrenfeucht–Mostowski models is tied closely to the existence of "large" (i.e., of size $\beth_{\omega_1}$) models of the theory.

The compactness of first order logic yields the full upward Löwenheim–Skolem–Tarski (LST) theory for $L_{\omega,\omega}$: if ψ has an infinite model it has arbitrarily large models. But for $L_{\omega_1,\omega}$, the LST-theorem replaces "an infinite model" by "a model of size $\beth_{\omega_1}$." The proof proceeds by using iterations of the Erdős–Rado theorem to find infinite sets of indiscernibles and to transfer size via Ehrenfeucht–Mostowski models.

All authors are grateful to an individual who prefers to remain anonymous for providing typing services that were used during the work on the paper. Baldwin was partially supported by Simons travel grant G3535. Laskowski was partially supported by NSF grant DMS-2154101, and is grateful to Danielle Ulrich for helpful discussions. Shelah would like to thank the NSF and BSF for partially supporting this research through grant NSF 2051825/BSF 3013005232 with M. Malliaris. This is paper 1244 in the Shelah Archive.
MSC2020: primary 03C45; secondary 03C48, 03C75.
Keywords: atomic class, categoricity, ω-stable.

By an atomic class we mean the *atomic* models (i.e., each finite sequence in each model realizes a principal type over the empty set) of a complete theory in a countable first order language. Every complete sentence in $L_{\omega_1,\omega}$ defines such a class because Chang's theorem translates the sentence to a first order theory omitting types and the language can be expanded to make all realized types atomic [Baldwin 2009, Chapter 6].

Shelah calls an atomic class excellent if it satisfies an n-amalgamation property for all n and structures of arbitrary cardinality. Shelah [1983a; 1983b] proved in ZFC: If an atomic class K is excellent and has an uncountable model then

(1) it has models of arbitrarily large cardinality;

(2) if it is categorical in one uncountable power it is categorical in all uncountable powers.

He also obtained a partial converse; under the very weak generalized continuum hypothesis ($2^{\aleph_n} < 2^{\aleph_{n+1}}$ for $n < \omega$): an atomic class $\boldsymbol{K}$ that has at least one uncountable model and is categorical in $\aleph_n$ for each $n < \omega$ is excellent. Thus under VWGCH the "Hanf number" for existence and for categorical atomic classes is reduced from $\beth_{\omega_1}$ to $\aleph_\omega$.

This raises the question. Does an $\aleph_1$-categorical atomic class have arbitrarily large models? Shelah [1975] showed it has a model in $\aleph_2$.

For the authors, work on this problem began by searching for sentences of $L_{\omega_1,\omega}$ for which $\aleph_1$-categoricity can be altered by forcing.[1] The third author proposed an example, but the first author objected to the proof and the second author proved in ZFC that the putative example was not $\aleph_1$-categorical.

In a series of papers the authors show that $\aleph_1$categorical atomic classes (or even simply $< 2^{\aleph_1}$ atomic models in $\aleph_1$) exhibit some "superstable-like" behavior. In [Baldwin et al. 2016] we introduced the appropriate notion of an algebraic type for atomic classes, *pseudoalgebraic* (Definition 3.2.2) and proved there that for an atomic class with $< 2^{\aleph_1}$ models in $\aleph_1$ the pseudoalgebraic types were dense. This is analogous to every nonalgebraic formula being extendible to a weakly minimal formula in a superstable theory. In [Laskowski and Shelah 2019] it is shown that an atomic class with few models in $\aleph_1$ is "pcl-small", i.e., there are few types over the pseudoclosure of any finite set (which is a weakening of ω-stability) and here we show that $\aleph_1$ categoricity and the existence of an atomic model of size $\beth_1^+$ implies ω-stability.

The search for weakened conditions for ω-stability is partially motivated by asking whether the absoluteness of $\aleph_1$-categoricity for first order logic (given by

[1]For sentences of $L_{\omega_1,\omega}(Q)$, such sentences exist; see [Shelah 1987, §6], expounded as [Baldwin 2009, §17]. A non-ω-stable sentence with no models above the continuum is given, where $\aleph_1$-categoricity fails under CH but holds under Martin's axiom.

the equivalence to ω-stable and no two-cardinal model) extends to atomic classes. Baldwin [2012] proved that either arbitrarily large models ($\beth_{\omega_1}$) or ω-stability suffices for such an absolute characterization. Our main theorem reduces the $\beth_{\omega_1}$ to $\beth_1^+$.

In Section 2 we investigate *constrained* types over models and investigate their relation to $\aleph_1$-categoricity and ω-stability. The notion of a constrained type is just a renaming; a type $p \in S(M)$ is constrained just if it does not split over a finite subset. Such a type is definable in the standard use in model theory — the existence of a schema such that for all $\boldsymbol{m} \in M$, $\phi(\boldsymbol{x}, \boldsymbol{m}) \in p \leftrightarrow d_\phi(\boldsymbol{x}, \boldsymbol{m})$. In Sections 2.2 and 2.3.1 we introduce "constrained" and limit types (over models) and investigate them under the assumption of $\aleph_1$-categoricity. From this, we prove the main theorem. However, our results in Section 2.2 depend on a major hypothesis, the existence of an uncountable model in which every limit type is constrained. In Section 3 we pay back our debt. By proving Theorem 2.3.2, we show the existence of a model of size $\aleph_1$ in which every limit type is constrained, using only the existence of an uncountable model. Although the proof there uses forcing, by appealing to the absoluteness given by Keisler's model existence theorem for sentences of $L_{\omega_1,\omega}(Q)$, the result is really a theorem of ZFC.

2. Constrained types, $\aleph_1$-categoricity and ω-stability

Throughout this article, T denotes a complete theory in a countable language for which there is an uncountable atomic model. At denotes the class of atomic models of T. In everything that follows, we only consider atomic sets, i.e., sets for which every finite tuple is isolated by a complete formula. Throughout, M, N denote atomic models and A, B atomic sets. We write $\boldsymbol{a}$, $\boldsymbol{b}$ for finite atomic tuples, and $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ denote finite tuples of variables.

We repeatedly use the fact that the countable atomic model M is unique up to isomorphism. Vaught [1961] showed the existence of an uncountable atomic model is equivalent to the countable atomic model having a proper elementary extension. The only types we consider are either over an atomic model or are over a finite subset of a model. In either case, we only consider types realized in atomic sets.

For general background see [Baldwin 2009] and more specifically [Baldwin et al. 2016].

2.1. *Constrained types and filtrations.*

Definition 2.1.1. Fix a countable complete theory T with monster model $\mathcal{M}$. $\mathrm{At} = \mathrm{At}_T$ denotes the collection of atomic models of T.

(1) For $M \in At$, $S_{at}(M)$ is the collection of $p(\boldsymbol{x}) \in S(M)$ such that if $\boldsymbol{a} \in \mathcal{M}$ realizes p, $M\boldsymbol{a}$ is an atomic set.

(2) At is *ω-stable* if for every/some countable $M \in \mathrm{At}$, $S_{at}(M)$ is countable.

The reader is cautioned that the definition of ω-stability is not equivalent to the classical notion (i.e., $S(M)$ countable) but within the context of atomic sets, this revised notion of ω-stability plays an analogous role. The spaces $S_{at}(M)$ are typically not compact. However, if M is countable, then $S_{at}(M)$ is a G_δ subset of the full Stone space $S(M)$, and thus is a Polish space. In particular, if At is not ω-stable, then $S_{at}(M)$ contains a perfect set.

Definition 2.1.2. (1) A type $p \in S_{at}(M)$ *splits over* $F \subseteq M$ if there exist tuples $\boldsymbol{b}, \boldsymbol{b}' \subseteq M$ and a formula $\phi(\boldsymbol{x}, \boldsymbol{y})$ such that $\text{tp}(\boldsymbol{b}/F) = \text{tp}(\boldsymbol{b}'/F)$, but $\phi(\boldsymbol{x}, \boldsymbol{b}) \wedge \neg\phi(\boldsymbol{x}, \boldsymbol{b}') \in p$.

(2) We call $p \in S_{at}(M)$ *constrained* if p does not split over some finite $F \subseteq M$ and *unconstrained* if p splits over every finite subset of M.

(3) For any atomic model M, let $C_M := \{p \in S_{at}(M) : p \text{ is constrained}\}$. We say At has *only constrained types* if $S_{at}(N) = C_N$ for every atomic model N.

We use the term constrained in place of "does not split over a finite subset" for its brevity, which is useful in subsequent definitions.

Remark 2.1.3. The concepts in clauses (2) and (3) above give a method of proving that an atomic class is ω-stable. We show At is ω-stable holds if and only if both

(a) C_M is countable for some/every countable atomic M and

(b) At has only constrained types.

Right to left is well-known: ω-stability immediately implies (a) and the deduction of (b) is standard [Baldwin 2009, Lemma 20.8]. Under the assumption of $\aleph_1$-categoricity, Theorem 2.2.1 gives (a) and Theorem 2.4.4 gives three equivalents of (b). However, the short proof of Theorem 2.4.4 makes crucial use of Theorem 2.3.2, whose lengthy proof is relegated to Section 3.

The constrained types $p \in C_M$ are those that have a defining scheme over a uniform finite set of parameters, i.e., if $p \in S_{at}(M)$ does not split over $\boldsymbol{a}$, then for every parameter-free $\phi(\boldsymbol{x}, \boldsymbol{y})$, there is an $\boldsymbol{a}$-definable formula $d_p\boldsymbol{x}\phi(\boldsymbol{x}, \boldsymbol{y})$ such that for any $\boldsymbol{b} \in M^{|\boldsymbol{y}|}$, $\phi(\boldsymbol{x}, \boldsymbol{b}) \in p$ if and only if $M \models d_p\boldsymbol{x}\phi(\boldsymbol{x}, \boldsymbol{b})$. We record three easy facts about extensions and restrictions of types.

Lemma 2.1.4. (1) *If M is a countable atomic model and $p \in S_{at}(M)$ then p is realized in an atomic extension of M.*

(2) *For any atomic models $M \preceq N$ and $A \subseteq M$ is finite, then for any $q \in S_{at}(N)$ that does not split over A, the restriction $q\restriction_M$ does not split over A; and any $p \in S_{at}(M)$ that does not split over A has a unique nonsplitting extension $q \in S_{at}(N)$.*

(3) *If some atomic N has an unconstrained $p \in S_{at}(N)$, then for every countable $A \subseteq N$, there is a countable $M \preceq N$ with $A \subseteq M$ for which the restriction $p\restriction_M$ is unconstrained.*

(4) At *has only constrained types if and only if $S_{at}(M) = C_M$ for every/some countable atomic model M.*

Proof. (1) Suppose $\boldsymbol{a}$ realizes p in the monster model $\mathcal{M} \succeq M$. $\mathcal{M}$ need not be atomic, but $M \cup \{\boldsymbol{a}\}$ is a countable atomic subset. Since every atomic model N is ω-homogeneous, a "forth construction" shows that for every countable atomic $S \subseteq \mathcal{M}$, there is an $(\mathcal{M}, N)$-elementary map $f : S \to N$. Thus there is an atomic $M' \succeq M$ containing $\boldsymbol{a}$.

(2) The first statement is immediate. For the second, given $p(\boldsymbol{x}) \in S_{at}(M)$ non-splitting over A, put

$$q(\boldsymbol{x}) := \left\{\phi(\boldsymbol{x},\boldsymbol{b}) : \boldsymbol{b} \in N^{|\boldsymbol{y}|}, \phi(\boldsymbol{x},\boldsymbol{b}') \in p \text{ for some } \boldsymbol{b}' \in M \text{ with } \mathrm{tp}(\boldsymbol{b}'/A) = \mathrm{tp}(\boldsymbol{b}/A)\right\}$$

(3) We construct $M \preceq N$ as the union of an increasing elementary ω-chain $M_n \preceq N$ of countable, elementary substructures of N with $A \subseteq M_0$ and, for each $n \in \omega$, $p\restriction_{M_{n+1}}$ splits over every finite $F \subseteq M_n$. It follows that $M^* := \bigcup\{M_n : n \in \omega\}$ is as required.

(4) Left to right is immediate. For the converse, assume there is some atomic N with an unconstrained type $p \in S_{at}(N)$. By (2) there is a countable $M \preceq N$ with $p\restriction_M$ unconstrained. $\square$

Much of the paper concerns analyzing atomic models N of size $\aleph_1$. It is useful to consider any such N as a direct limit of a family of countable, atomic submodels.

Definition 2.1.5. For a model N of size $\aleph_1$, a *filtration of* N is a continuous, increasing sequence $(M_\alpha : \alpha \in \omega_1)$ of countable, elementary substructures with $N = \bigcup_{\alpha\in\omega_1} M_\alpha$.

When N is atomic, then in any filtration $(M_\alpha : \alpha \in \omega_1)$ of N, each of the countable models are isomorphic. As well, any two filtrations $(M_\alpha : \alpha \in \omega_1)$ and $(M'_\alpha : \alpha \in \omega_1)$ agree on a club. Thus, for any given countable $M \preceq N$, we have that $\{\alpha \in \omega_1 : M \preceq M_\alpha \text{ and } M_\alpha = M'_\alpha\}$ is club as well.

2.2. *$\aleph_1$-categoricity implies C_M is countable.* Throughout this subsection, At is an atomic class that admits an uncountable model and M *denotes a fixed copy of the countable atomic model.* We aim to count the set

$$C_M = \{p \in S_{at}(M) : p \text{ is constrained}\}.$$

Theorem 2.2.5 yields the main result of the subsection:

Theorem 2.2.1. *If* At *is $\aleph_1$-categorical, then C_M is countable for every/some countable atomic model M.*

As M is countable, the natural action of $\mathrm{Aut}(M)$ on the set M induces an action of $\mathrm{Aut}(M)$ on $S_{at}(M)$. When M is atomic, a useful characterization of $p \in C_M$ is: C_M consists of those elements of $S_{at}(M)$ whose orbits are countable. However, for the results in this section we only require the easy half of this statement.

Lemma 2.2.2. *Suppose $p \in C_M$ and M' is any countable, atomic model. Then*:

(1) *$\{\pi(p) : \pi : M \to M'$ an isomorphism$\}$ is a countable set of constrained types in $S_{at}(M')$.*

(2) *There is a countable atomic $M^* \succ M'$ realizing $\pi(p)$ for every isomorphism $\pi : M \to M'$.*

Proof. (1) Choose a finite $A \subseteq M$ over which p does not split. As M' is countable, A has only countably many images under isomorphisms $\pi : M \to M'$, and it follows immediately from nonsplitting that if $\pi_1, \pi_2 : M \to M'$ are isomorphisms satisfying $\pi_1(a) = \pi_2(a)$ for each $a \in A$, then $\pi_1(p) = \pi_2(p)$.

(2) Using (1), let $\{q_i : i < \gamma \leq \omega\} \subseteq S_{at}(M')$ be the set of all images of p under isomorphisms $\pi : M \to M'$. We recursively construct an increasing sequence of countable models $\{M_i : i < \gamma\}$ with $M_0 = M'$ and, for each $i < \gamma$, M_i contains a realization of q_j for every $j < i$. Supposing $i < \gamma$ and M_i has been defined, let $q_i^* \in S_{at}(M_i)$ be the unique ([Baldwin 2009, Theorem 19.9]) nonsplitting extension of $q_i \in S_{at}(M')$. Then letting d_i realize q_i^*, let $M_{i+1} \in At$ be an elementary extension of M_i containing $M_i \cup \{d_i\}$. Then $\bigcup_{i<\omega} M_i$ works. □

Definition 2.2.3. Suppose $(M_\beta : \beta < \omega_1)$ is a filtration of some $N \in \mathrm{At}$ of size $\aleph_1$. For each $\beta < \omega_1$, let

$$R_N^\beta := \{p \in C_M : \pi(p) \text{ is realized in } N \text{ for every isomorphism } \pi : M \to M_\beta\}$$

and let $R_N := \{p \in C_M : p \in R_N^\beta$ for a stationary set of $\beta \in \omega_1\}$.

As any two filtrations of N agree on a club, it follows that R_N is independent of the choice of filtration of N. Similarly, R_N is an isomorphism invariant, i.e., if $N \cong N'$ are each atomic models of size $\aleph_1$, then $R_N = R_{N'}$. We record two facts about R_N.

Lemma 2.2.4. (1) *For any $N \in \mathrm{At}$ of size $\aleph_1$, $|R_N| \leq \aleph_1$.*

(2) *For any $p \in C_M$ there is some $N \in \mathrm{At}$ of size $\aleph_1$ such that $p \in R_N$.*

Proof. (1) Choose any sequence $\langle p_i : i \in \omega_2\rangle$ from R_N and we will show that $p_i = p_j$ for some distinct i, j. Fix a filtration (M_α) of N. We shrink the sequence in two stages. First, for each $i < \omega_2$, let $\alpha(i) \in \omega_1$ be least such that $p_i \in R_N^{\alpha(i)}$. By pigeonhole and reindexing we may assume $\alpha(i) = \alpha^*$ for all i, i.e., each $p_i \in R_N^{\alpha^*}$. Now fix any isomorphism $\pi : M \to M_{\alpha^*}$. By definition of $R_N^{\alpha^*}$, $\pi(p_i)$ is realized in N for every p_i. But, as $|N| = \aleph_1$, there is $c^* \in N$ realizing both $\pi(p_i)$ and $\pi(p_j)$ for some distinct i, j. Thus, $\pi(p_i) = \pi(p_j)$, hence $p_i = p_j$.

(2) Fix $p \in C_M$. Using Lemma 2.2.2(2) at each level, construct a continuous, increasing elementary sequence M_α of countable atomic models such that, for every $\alpha < \omega_1$, $\pi(p)$ is realized in $M_{\alpha+1}$ for every isomorphism $\pi : M \to M_\alpha$. Put $N := \bigcup_{\alpha<\omega_1} M_\alpha$. Then (M_α) is a filtration of N and $p \in R^\alpha_N$ for every $\alpha < \omega_1$. Thus, $p \in R_N$. □

We are now able to prove the theorem below, which clearly implies Theorem 2.2.1.

Theorem 2.2.5. *If C_M is uncountable, then $I(\mathrm{At}, \aleph_1) = 2^{\aleph_1}$.*

Proof. It is easily verified that C_M is an F_σ subset of the Polish space $S_{at}(M)$, so on general grounds, C_M is either countable or else it contains a perfect set.

Our proof is nonuniform, depending on the relative sizes of $2^{\aleph_0}$ and $2^{\aleph_1}$. First, under weak CH, i.e., $2^{\aleph_0} < 2^{\aleph_1}$ then combining arguments of Keisler [1970] and Shelah [Baldwin 2009, Theorem 18.16] shows if $I(\mathrm{At}, \aleph_1) \neq 2^{\aleph_1}$, then At is ω-stable, so $S_{at}(M)$ is countable. As $C_M \subseteq S_{at}(M)$, C_M is countable as well.

On the other hand, assume $2^{\aleph_0} = 2^{\aleph_1}$, so in particular WCH fails. Under this assumption, we will prove that if C_M is uncountable, then $I(\mathrm{At}, \aleph_1) = 2^{\aleph_0}$, which equals $2^{\aleph_1}$ under our cardinal hypotheses for this case. Indeed, choose representatives $\{N_i : i \in \kappa\}$ for the isomorphism classes of atomic models of size $\aleph_1$. If C_M is uncountable, then as noted the first sentence of the proof, C_M contains a perfect set and so $|C_M| = 2^{\aleph_0}$. But by Lemma 2.2.4, $C_M \subseteq \bigcup\{R_{N_i} : i \in \kappa\}$ and $|R_{N_i}| \leq \aleph_1$ for each $i \in \kappa$. As we are assuming $2^{\aleph_0} > \aleph_1$, we conclude $\kappa \geq 2^{\aleph_0}$, as required. □

2.3. *Limit types and $\aleph_1$-categoricity.*

Definition 2.3.1. A type $p \in S_{at}(N)$ is a *limit type* if the restriction $p\restriction_M$ is realized in N for every countable $M \preceq N$.

Trivially, for every N, every type in $S_{at}(N)$ realized in N is a limit type. Since we allow $M = N$ in the definition of a limit type, if M is countable, then the only limit types in $S_{at}(M)$ are those realized in M.

Also, if $(M_\alpha : \alpha \in \omega_1)$ is a filtration of N, then a type $p \in S_{at}(N)$ is a limit type if and only if N realizes $p\restriction_{M_\alpha}$ for cofinally many α.

The long proof of the following crucial theorem is relegated to Section 3. Note that there are no additional assumptions on At, other than the existence of an uncountable, atomic model.

Theorem 2.3.2. *If* At *admits an uncountable, atomic model, then there is some $N \in \mathrm{At}$ with $|N| = \aleph_1$ for which every limit type in $S_{at}(N)$ is constrained.*

Here, we sharpen this result under the additional assumption of $\aleph_1$-categoricity.

Corollary 2.3.3. *If* At *is* $\aleph_1$*-categorical and* $N \in$ At *has size* $\aleph_1$, *then for every* $p \in S_{at}(N)$, $\in C_N \leftrightarrow p \in \{$*limit types in* $S_{at}(N)\}$.

Proof. The hard direction of the equality is Theorem 2.3.2. For the converse, by the assumption of $\aleph_1$-categoricity it suffices to construct *some* $N \in$ At of size $\aleph_1$ for which every $p \in C_N$ is a limit type. For this, first note that for every countable atomic M, since C_M is countable by Theorem 2.2.1, iterating Lemma 2.1.4(1) ω times yields a countable atomic $M' \succeq M$ that realizes every $p \in C_M$. Using this, construct a strictly increasing, continuous elementary chain $(M_\alpha : \alpha \in \omega_1)$ of countable, atomic models such that for each $\alpha \in \omega_1$, $M_{\alpha+1}$ realizes every $p \in C_{M_\alpha}$. Put $N := \bigcup_{\alpha \in \omega_1} M_\alpha$. We claim that every $p \in C_N$ is a limit type. So fix $p \in C_N$ and choose any countable $M \preceq N$. Choose a finite $A \subseteq N$ for which p does not split over A and choose $\alpha \in \omega_1$ so that $M \cup A \subseteq M_\alpha$. By Lemma 2.1.4(2), $p{\restriction}_{M_\alpha}$ is constrained, hence it is realized in $M_{\alpha+1} \subseteq N$. As any such realization in N realizes $p{\restriction}_M$, p is a limit type. □

2.4. *Characterizing ω-stability.* In this subsection, we first derive Lemma 2.4.3 that gives three consequences of ω-stability in terms of the behavior of constrained types. Then, taking Theorem 2.3.2 as a black box (proved in Section 3), Theorem 2.4.4 shows that each of these conditions is equivalent to ω-stability under the assumption of $\aleph_1$-categoricity. Finally, Theorem 2.4.5 asserts that the existence of a model in $\beth_1^+$ and $\aleph_1$-categoricity implies condition (1) of Theorem 2.4.4 and thus ω-stability.

Definition 2.4.1.
- A *proper constrained pair* is a pair $N \precneqq N'$ of atomic models such that $\mathrm{tp}(\boldsymbol{c}/N)$ is constrained for every tuple $\boldsymbol{c} \in N'$.
- A *proper relatively* $\aleph_1$*-saturated pair* is a proper pair $N \precneqq N'$ such that, for every countable $M \preceq N$, every type $p \in S(M)$ realized in N' is realized in N.

Note that in (2), both models must be uncountable, whereas (1) makes sense for countable models as well. Of course, in (2) it would be equivalent to say that "every type over every countable set $A \subseteq N$ that is realized in N' is realized in N," but we choose the definition above to conform with our convention about only looking at types over models.

Lemma 2.4.2. *Let* At *be any atomic class.*

(1) *If both* (M, M') *and* (M', M'') *are constrained pairs*, *then* (M, M'') *is a constrained pair as well.*

(2) *If* (M, M') *is a constrained pair of countable atomic models*, *then there is an uncountable* N *with a filtration* $(M_\alpha : \alpha \in \omega_1)$ *such that* (M_α, N) *is a constrained pair for every* $\alpha \in \omega_1$.

Proof. (1) Choose any $\boldsymbol{c} \in M''$. As (M', M'') is a constrained pair, choose $\boldsymbol{b} \in M'$ such that $\mathrm{tp}(\boldsymbol{c}/M')$ does not split over $\boldsymbol{b}$. As (M, M') is a constrained pair, choose $\boldsymbol{a} \in M$ such that $\mathrm{tp}(\boldsymbol{b}/M)$ does not split over $\boldsymbol{a}$. We claim that $\mathrm{tp}(\boldsymbol{cb}/M)$ does not split over $\boldsymbol{a}$, which clearly suffices. To see this, choose any $\boldsymbol{m}_1, \boldsymbol{m}_2$ from M such that $\mathrm{tp}(\boldsymbol{m}_1\boldsymbol{a}) = \mathrm{tp}(\boldsymbol{m}_2\boldsymbol{a})$. By nonsplitting, this implies $\mathrm{tp}(\boldsymbol{m}_1\boldsymbol{ab}) = \mathrm{tp}(\boldsymbol{m}_2\boldsymbol{ab})$. Now both $\boldsymbol{m}_1\boldsymbol{a}$ and $\boldsymbol{m}_2\boldsymbol{a}$ are from M', hence $\mathrm{tp}(\boldsymbol{m}_1\boldsymbol{abc}) = \mathrm{tp}(\boldsymbol{m}_2\boldsymbol{abc})$ as $\mathrm{tp}(\boldsymbol{c}/M')$ does not split over $\boldsymbol{b}$.

(2) As M is a countable atomic model that is the lower part of a constrained pair, so is any other countable, atomic model. Thus, we can form a continuous, increasing chain $(M_\alpha : \alpha \in \omega_1)$ of countable atomic models with $(M_\alpha, M_{\alpha+1})$ a constrained pair for each α. This chain is a filtration of the atomic $N := \bigcup\{M_\alpha : \alpha \in \omega_1\}$. That each (M_α, N) is a constrained pair follows from (1). □

We record the following consequences of ω-stability in atomic classes. It is noteworthy that $\aleph_1$-categoricity plays no role in Lemma 2.4.3, and without additional assumptions, none of these imply ω-stability. However, following this, with Theorem 2.4.4 we see that when coupled with $\aleph_1$-categoricity, each of these conditions implies ω-stability.

Lemma 2.4.3. *Suppose* At *is an ω-stable atomic class that admits an uncountable atomic model. Then*

(1) At *has only constrained types*;

(2) At *has a proper constrained pair*; *and*

(3) At *has a proper*, *relatively $\aleph_1$-saturated pair.*

Proof. (1) For an ω-stable atomic class, one can define ([Baldwin 2009, Definition 19.1]) a splitting rank on types $p \in S_{at}(N)$ for any model N such that ([Baldwin 2009, Theorem 19.8]): for any atomic model N and any $p \in S_{at}(N)$, then choosing $\phi(x, \boldsymbol{a}) \in p$ to be a complete formula of smallest rank, p does not split over $\boldsymbol{a}$. That is, p is constrained.

(2) Choose any countable, atomic model M. Since At admits an uncountable atomic model, there is a countable, proper, atomic elementary extension $M' \succ M$. By (1), $\mathrm{tp}(c/M)$ is constrained for every $c \in M'$, hence (M, M') is a proper constrained pair.

(3) We first argue that there is an *atomically saturated* model N of size $\aleph_1$. That is, for every countable $M \preceq N$, N realizes every $p \in S_{at}(M)$. The existence of an uncountable, atomically saturated N is easy. Using Lemma 2.4.3(1) all types for At are constrained. Then, using Lemma 2.1.4(1) and (2) as in the proof of Corollary 2.3.3, build a union of a continuous elementary chain $(M_\alpha : \alpha \in \omega_1)$ of countable atomic models with the property that for each $\alpha < \omega_1$, $M_{\alpha+1}$ realizes every $p \in S_{at}(M_\alpha)$. The existence of such an $M_{\alpha+1}$ is immediate since $S_{at}(M_\alpha)$ is countable and every $p \in S_{at}(M_\alpha)$ can be realized in some countable, atomic elementary extension.

Now, given an atomically saturated model N of size $\aleph_1$, recall that if At is ω-stable, then every model of size $\aleph_1$ has a proper atomic extension N'; see, e.g., the proof of 19.26 of [Baldwin 2009]. But then (N, N') is a proper, relatively $\aleph_1$-saturated pair. □

Given Theorem 2.2.1 and Corollary 2.3.3 (the latter depending on the promised Theorem 2.3.2), we give short proofs of our main results.

Theorem 2.4.4. *The following are equivalent for an* $\aleph_1$*-categorical atomic class* At:

(1) At *has a proper, relatively* $\aleph_1$*-saturated pair.*

(2) At *has a proper constrained pair.*

(3) At *has only constrained types.*

(4) At *is* ω*-stable.*

Proof. We will show (1) ⇒ (2) ⇒ (3) ⇒ (4), which in light of Lemma 2.4.3 suffices.

(1) ⇒ (2): Suppose (M^*, M^{**}) is a proper, relatively $\aleph_1$-saturated pair of atomic models, and by way of contradiction suppose that (M^*, M^{**}) is not a proper constrained pair. Choose $c \in M^{**}$ such that $p := \operatorname{tp}(c/M^*)$ is unconstrained. Then, by iterating Lemma 2.1.4(3), we construct a continuous, elementary chain $(M_\alpha : \alpha \in \omega_1)$ of countable, elementary substructures of M^* such that, for every $\alpha \in \omega_1$, $p\restriction_{M_\alpha}$ is unconstrained, but is realized in $M_{\alpha+1}$. To accomplish this, by Lemma 2.1.4(3), choose a countable $M_0 \preceq M^*$ such that $p\restriction_{M_0}$ is unconstrained. At countable limits, take unions. Finally, given a countable $M_\alpha \preceq M^*$, by relative $\aleph_1$-saturation choose $c_\alpha \in M^*$ realizing $p\restriction_{M_\alpha}$ and then apply Lemma 2.1.4(3) to the set $M_\alpha \cup \{c_\alpha\}$ to get $M_{\alpha+1} \preceq M^*$ with $p\restriction_{M_{\alpha+1}}$ unconstrained. Let $N := \bigcup\{M_\alpha : \alpha \in \omega_1\}$. Then N has size $\aleph_1$ and the type $p\restriction_N$ is an unconstrained limit type, contradicting $\aleph_1$-categoricity by Corollary 2.3.3.

(2) ⇒ (3): Assume that (N^*, N^{**}) is a proper constrained pair (of any cardinality). By an easy Löwenheim–Skolem argument (in the pair language) there is a proper constrained pair (M, M') of countable atomic models. By Lemma 2.4.2(2), there is an atomic model N of size $\aleph_1$ with a filtration $(M_\alpha : \alpha \in \omega_1)$ such that (M_α, N) is a constrained pair for every $\alpha \in \omega_1$.

Now, by way of contradiction, assume (3) fails. By Lemma 2.1.4(4), $S_{at}(M)$ contains an unconstrained type for every countable atomic model M. Thus, for any such M, there is a countable atomic $M' \succ M$ containing a realization of an unconstrained type. By iterating this ω_1 times, we construct a continuous, elementary chain $(M'_\alpha : \alpha \in \omega_1)$ for which $M'_{\alpha+1}$ contains a realization of an unconstrained type in $S_{at}(M'_\alpha)$. Let $N' := \bigcup\{M'_\alpha : \alpha \in \omega_1\}$. Note that (M'_α, N') is never a constrained pair. But this contradicts $\aleph_1$-categoricity: If $f : N \to N'$ were an isomorphism, then there would be (club many) $\alpha \in \omega_1$ such that $f\restriction_{M_\alpha}$

maps M_α onto M'_α, hence maps the pair (M_α, N) onto (M'_α, N'). As the former is a constrained pair, while the latter is not, we obtain a contradiction.

(3) $\Rightarrow$ (4): Assume At has only constrained types and let M be any countable, atomic model. This means that $S_{at}(M) = C_M$. However, as At is $\aleph_1$-categorical, C_M is countable by Theorem 2.2.1. Thus, $S_{at}(M)$ is countable, which is the definition of At being ω-stable. □

With this result in hand, it is easy to deduce the main theorem. This is the *only* use of the existence of a model in $\beth_1^+$. We imitate the classical proof that for every $\kappa \geq |L|$, every L-theory with an infinite model has a κ^+-saturated model of size 2^κ, to prove clause (1) of Theorem 2.4.4 and thus deduce ω-stability.

Theorem 2.4.5. *If an atomic class* At *is* $\aleph_1$*-categorical and has a model of size* $(2^{\aleph_0})^+$, *then* At *is* ω*-stable.*

Proof. Let M^{**} be an atomic model of size $(2^{\aleph_0})^+$. We construct a relatively $\aleph_1$-saturated elementary substructure $M^* \preceq M^{**}$ of size $2^{\aleph_0}$ as the union of a continuous chain $(N_\alpha : \alpha \in \omega_1)$ of elementary substructures of M^{**}, each of size $2^{\aleph_0}$, where, for each $\alpha < \omega_1$ and each of the $2^{\aleph_0}$ countable $M \preceq N_\alpha$, $N_{\alpha+1}$ realizes each of the at most $2^{\aleph_0}$ $p \in S(M)$ that is realized in M^{**}. ω-stability is immediate from (1) $\Rightarrow$ (4) in Theorem 2.4.4. □

3. Paying our debt

The whole of this section is aimed at proving Theorem 2.3.2: If a countable theory T has an uncountable atomic model, then it has one in which every limit type is constrained.[2] The proof relies heavily on Keisler's completeness theorem that implies "model existence" of sentences of $L_{\omega_1,\omega}(Q)$ is absolute between forcing extensions. In the first subsection, we explicitly give an $L_{\omega_1,\omega}(Q)$ sentence Ψ^* in a countable language extending the language of T such that in any set-theoretic universe, Ψ^* has a model of size $\aleph_1$ if and only if there is an atomic model of size $\aleph_1$ with every limit type constrained.

The second subsection describes family of striated formulas [Baldwin et al. 2016]. Such formulas are used to describe a c.c.c. forcing notion $(\mathbb{P}, \leq)$ in the third subsection. There, we prove that $(\mathbb{P}, \leq)$ forces the existence of an atomic model of size $\aleph_1$ with every limit type constrained. Thus, we conclude that Ψ^* has a model of size $\aleph_1$ in a c.c.c. forcing extension, so by the absoluteness described above, $\mathbb{V}$ has a model of Ψ^* of size $\aleph_1$, yielding our requested model.

[2] By the correspondence described in the introduction, it follows immediately that any complete $L_{\omega_1,\omega}$-sentence with an uncountable model has an uncountable model with every limit type constrained.

3.1. *Finding a requisite sentence Ψ^* of $L_{\omega_1,\omega}(Q)$.* This subsection is devoted to proving the following proposition.

Proposition 3.1.1. *Let T be a first order L-theory for a countable language with an uncountable model in* At, *the class of atomic models. There is a sentence $\Psi^* \in (L^*)_{\omega_1,\omega}(Q)$ in an expanded (but still countable) language $L^* \supseteq L$ for which the following are equivalent:*

(1) *There is a model $N^* \models \Psi^*$.*

(2) *There is an atomic model $N \models T$ of size $\aleph_1$ such that every limit type of N is constrained.*

Whereas the L-reduct of any $N^* \models \Psi^*$ will satisfy (2), it is noteworthy that in proving (2) $\Rightarrow$ (1), the model $N^* \models \Psi^*$ we produce is not necessarily an expansion of a given N witnessing (2).

The relevant Ψ^* is defined in Definition 3.1.6. As we will be interested in arbitrary models of a sentence and because "is a well ordering" is not expressible in $L_{\omega,\omega}(Q)$, we need to generalize the notion of a filtration.

Definition 3.1.2. A linear order $(I, \leq)$ is ω_1-*like* if it has cardinality $\aleph_1$, but, letting $\mathrm{pred}(i)$ denote $\{j \in I : j < i\}$, for every $i \in I$, $|\mathrm{pred}(i)| \leqslant \aleph_0$.

If N is any set and $(I, \leq)$ is ω_1-like, then an $(I, \leq)$-*scale* is a surjective function $f : N \to I$ such that $f^{-1}(i)$ is countable for every $i \in I$.

If $f : N \to I$ is a scale, put $A_i := f^{-1}(\mathrm{pred}(i))$ for every $i \in I$, and note that each A_i is countable.

Observe that "being an ω_1-like linear order" is expressible by a sentence of $L_{\omega_1,\omega}(Q)$—the point is that any uncountable linear order $(I, \leq)$ for which $\mathrm{pred}(i)$ is countable for every $i \in I$ has both size and cofinality $\aleph_1$. Similarly, if an uncountable set N has an $(I, \leq)$-scale, then N must have size $\aleph_1$.

We consider the sets $(A_i : i \in I)$ to be a surrogate for a filtration of N; A_i replaces M_α. We now define a tree order on types over certain countable subsets of a model with cardinality $\aleph_1$ of T.

Definition 3.1.3. Fix T, N as in Proposition 3.1.1. Suppose $(I, \leq)$ is an ω_1-like linear order and $f : N \to I$ is a scale.

(1) Define an equivalence relation E_f on $(N \times I)$ as $(a, i) E_f (b, j)$ if and only if $i = j$ and $\mathrm{tp}(a/A_i) = \mathrm{tp}(b/A_i)$. Thus each equivalence class corresponds to a type.

(2) Define a strict partial order $\prec_f$ on $(N \times I)/E_f$ as: $[(a, i)] \prec_f [(b, j)]$ if and only if $i <_I j$; $\mathrm{tp}(a/A_i) = \mathrm{tp}(b/A_i)$; and $\mathrm{tp}(b/A_j)$ splits over every finite $F \subseteq A_i$.

(3) A $\prec_f$-*chain* is a sequence of types linearly ordered by $\prec_f$ (hence splitting).

It is evident that $((N \times I)/E_f, \prec_f)$ is *tree-like* in that the $\prec_f$-predecessors of every E_f-class are linearly ordered by $\prec_f$. Moreover, since $(I, \leq)$ is ω_1-like, every E_f-class has only countably many $\prec_f$-predecessors.

Lemma 3.1.4. *Let N be any atomic model of size $\aleph_1$, $(I, \leq)$ be ω_1-like, $f : N \to I$ be any scale and I, E_f, A_i, and $\prec_f$ be as in Definition 3.1.3. The following are equivalent*:

(1) *There exists an f such that $\mathcal{T}_f = ((N \times I)/E_f, \prec_f)$ has an uncountable $\prec_f$-chain.*

(2) *Some limit type in $S_{at}(N)$ is unconstrained.*

(3) *For every f, $\mathcal{T}_f = ((N \times I)/E_f, \prec_f)$ has an uncountable $\prec_f$-chain.*

Proof. (3) $\Rightarrow$ (1) is immediate. For (1) $\Rightarrow$ (2), suppose for some f, $C \subseteq \mathcal{T}_f$ is an uncountable $\prec_f$-chain. As $[(a, i)] \prec_f [(b, j)]$ implies $i < j$ and since $(I, \leq)$ is ω_1-like, $\pi_2(C) := \{i \in I : \exists a \in N[(a, i)] \in C\}$ is cofinal in I. Therefore $\bigcup\{A_i : i \in \pi_2(C)\} = N$. Also, as $[(a, i)] \prec_f [(b, j)]$ implies $\operatorname{tp}(a/A_i) = \operatorname{tp}(b/A_i)$, there is a unique $p \in S_{at}(N)$ defined as $p := \bigcup\{\operatorname{tp}(a/A_i) : (a, i) \in C\}$. Furthermore, as $\operatorname{tp}(b/A_j)$ splits over every finite $F \subseteq A_i$, it follows that p is unconstrained. Recalling Definition 2.3.1(2), it remains to show that p is a limit type. Choose a filtration $\overline{M} = (M_\alpha)$ of N and argue that $p\restriction_{M_\alpha}$ is realized in N for every $\alpha \in \omega_1$. Given $\alpha \in \omega_1$, choose $i \in \pi_2(C)$ such that $M_\alpha \subseteq A_i$. Then each $a \in N$ for which $(a, i) \in C$ realizes $p\restriction_{A_i}$ and hence realizes $p\restriction_{M_\alpha}$. So p is a limit type.

(2) $\Rightarrow$ (3). Suppose N has an unconstrained limit type $p \in S_{at}(N)$ and fix a scale f. Also choose a filtration $(M_\alpha : \alpha \in \omega_1)$ of N. To construct an uncountable chain $\mathcal{T}_f$ we repeatedly use the following claim.

Claim 3.1.5. *For every countable $B \subseteq N$ there is $i \in I$ such that*

- *$B \subseteq A_i$;*
- *$p\restriction_{A_i}$ is realized; and*
- *$p\restriction_{A_i}$ splits over every finite $F \subseteq B$.*

Proof. Given a countable $B \subseteq N$, since $p \in S_{at}(N)$ splits over every finite $F \subseteq N$, there is a countable $B^* \supseteq B$ such that $p\restriction_{B^*}$ splits over every finite $F \subseteq B$. Now choose $i \in I$ such that $B^* \subseteq A_i$ and then choose $\alpha \in \omega_1$ such that $A_i \subseteq M_\alpha$. Since p is a limit type, choose $c \in N$ realizing $p\restriction_{M_\alpha}$ and hence $p\restriction_{A_i}$. □

Iterating Claim 3.1.5 ω_1 times yields a strictly increasing sequence $(i_\alpha : \alpha \in \omega_1)$ from $(I, \leq)$ and $(c_\alpha : \alpha \in \omega_1)$ from N, where at each stage α, we take $B = \bigcup\{A_{i_\beta} : \beta < \alpha\}$. It follows directly from the definition of $\prec_f$ that $(c_\beta, i_\beta) \prec_f (c_\alpha, i_\alpha)$ whenever $\beta < \alpha$, so $((N \times I)/E, \prec_f)$ has an uncountable chain. □

With Lemma 3.1.4 in hand, we now define the sentence Ψ^* described in Proposition 3.1.1.

Definition 3.1.6. Let $L^* := L \cup \{I, \leq_I, f, E, \prec_f\} \cup \{\mathbb{Q}, \leq_{\mathbb{Q}}, H\}$ and let Ψ^* be a set of $L_{\omega_1,\omega}(Q)$-axioms ensuring that, for any $N^* \models \Psi^*$,

(1) the L-reduct N of N^* is an atomic model of T (as well, N denotes the universe of N^*);

(2) N is uncountable;

(3) $I \subseteq N$ and $(I, \leq_I)$ is an ω_1-like linear order;

(4) $f : N \to I$ is a scale; (recall: $A_i := f^{-1}(\mathrm{pred}(i))$);

(5) $E \subseteq N \times I$ satisfies $(a, i)E(b, j)$ if and only if $i = j$ and $\mathrm{tp}_L(a/A_i) = \mathrm{tp}_L(b/A_i)$;

(6) for all $[(a, i)], [(b, j)] \in (N \times I)/E$, $[(a_i)] \prec_f [(b, j)]$ if and only if $i < j$, $\mathrm{tp}_L(a/A_i) = \mathrm{tp}_L(b/A_i)$, and $\mathrm{tp}_L(b/A_j)$ splits over every finite $F \subseteq A_i$;

(7) $\mathbb{Q} \subseteq N$ and $(\mathbb{Q}, \leq_{\mathbb{Q}})$ is a countable model of DLO;

(8) $H : N \times I \to \mathbb{Q}$ satisfies: For all (a, i), (b, j),
   (a) if $(a, i)E(b, j)$ then $H(a, i) = H(b, j)$; and
   (b) if $[(a, i)] \prec_f [(b, j)]$, then $H(a, i) <_{\mathbb{Q}} H(b, j)$.

We verify that this sentence Ψ^* works for Proposition 3.1.1.

Proof of Proposition 3.1.1. For (1) $\Rightarrow$ (2) assume $N^* \models \Psi^*$ and let N be the L-reduct of N^*. Then $(I, \leq)$ is an ω_1-ordering and $f : N \to I$ is a scale, so $|N| = \aleph_1$. Moreover, as the ordering on $(\mathbb{Q}, \leq)$ forbids a strictly increasing ω_1 sequence, the existence of the function H forbids $T = ((N \times I)/E, \prec_f)$ having an uncountable $\prec_f$-chain. Thus, by Lemma 3.1.4, every limit type in $S_{at}(N)$ is constrained.

The converse is more involved. Assume we are given $N \in \mathrm{At}$ of size $\aleph_1$ with every limit type in $S_{at}(N)$ constrained. Under this assumption, with the help of Lemma 3.1.7 we will show that a model $N^* \models \Psi^*$ can be found in some generic extension $\mathbb{V}[G]$ of $\mathbb{V}$ by a c.c.c. forcing extension. Once we have that, it follows by the absoluteness gleaned from Keisler's model existence theorem for sentences of $L_{\omega_1,\omega}(Q)$ that a model of Ψ^* exists in $\mathbb{V}$, giving Proposition 3.1.1(1).

So, given N as in Proposition 3.1.1(2), choose arbitrary subsets $I, \mathbb{Q} \subseteq N$ of cardinality $\aleph_1, \aleph_0$, respectively and choose orderings $\leq_I$ and $\leq_{\mathbb{Q}}$ as required by Ψ^*. Fix an arbitrary scale $f : N \to I$ and interpret E and $\prec_f$ as required. Since N has every limit type constrained, it follows from Lemma 3.1.4 that $\prec_f$ has no uncountable chains.

It only remains to find a function $H : N \times I \to \mathbb{Q}$ as requested by Ψ^*. For this, we turn to forcing, and invoke the following general lemma,[3] taking X to be $(N \times I/E)$ and $\prec$ to be $\prec_f$.

[3]The statement of Lemma 3.1.7 is reminiscent of how one specializes Aronszajn trees by forcing, and the ideas of the proof can be found in Section 2 of [Baumgartner 1970].

Lemma 3.1.7. *Suppose* $(X, \prec)$ *is any strict partial order satisfying*

(1) $|X| = \aleph_1$;

(2) *for every* $a \in X$, *the induced suborder* $(\mathrm{pred}(a), \prec)$ *is a countable linear order*; *and*

(3) $(X, \prec)$ *has no uncountable chain.*

Then there is a c.c.c. forcing $(\mathbb{P}, \leq)$ *such that in any generic* $\mathbb{V}[G]$ *there is a function* $H : X \to \mathbb{Q}$ *such that if* $a \prec b$, *then* $H(a) <_{\mathbb{Q}} H(b)$.

Proof. The partial order $(\mathbb{P}, \leq)$ is simply the set of all finite approximations of such an H. That is, $\mathbb{P}$ is the set of all functions $h : X_0 \to \mathbb{Q}$ with $X_0 \subseteq X$ finite such that for all $a, b \in X_0$, if $a \prec b$, then $h(a) <_{\mathbb{Q}} h(b)$, ordered by inclusion, i.e., ($\sharp$) $h \leq h'$ if and only if $h \subseteq h'$. It is easily checked that this forcing will produce (in $\mathbb{V}[G]$) a total function $H : X \to \mathbb{Q}$ as desired. The nontrivial part is showing that $(\mathbb{P}, \leq)$ has the c.c.c. For this, choose any uncountable set $Y = \{h_\alpha : \alpha \in \omega_1\} \subseteq \mathbb{P}$ and assume, by way of contradiction, that $h_\alpha \cup h_\beta \notin \mathbb{P}$ for distinct $\alpha, \beta \in \omega_1$. By passing to a subset of Y, we may assume $|\mathrm{dom}(h_\alpha)| = n$ for some fixed $n \in \omega$ and we argue by contradiction. If $n = 1$, i.e., $\mathrm{dom}(h_\alpha) = \{a^\alpha\}$, then by passing to a further subset, there is a single $m^* \in \mathbb{Q}$ such that $h_\alpha(a^\alpha) = m^*$ for every α. The only way we could have $h_\alpha \cup h_\beta \notin \mathbb{P}$ would be if a^α, a^β were distinct, but $\prec$-comparable. But then $C = \{a^\alpha : \alpha \in \omega_1\}$ would be an uncountable chain in $(X, \prec)$, contradicting our assumption.

So, assume $n > 1$ and we have proved (c.c.c.) for all $n' < n$. To ease notation, enumerate the universe X with order type ω_1. For each α, write $\mathrm{dom}(h_\alpha) = (a_1^\alpha, \ldots, a_n^\alpha)$ in increasing order, subject to this enumeration. By the Δ-system lemma, there is an uncountable subset and a root r such that $\mathrm{dom}(h_\alpha) \cap \mathrm{dom}(h_\beta) = r$ for all distinct pairs α, β. If $r \neq \varnothing$, we can apply our inductive hypothesis to the family of sets $\{\mathrm{dom}(h_\alpha) \backslash r : \alpha \in \omega_1\}$, so we may assume $r = \varnothing$, i.e., the domains $\{\mathrm{dom}(h_\alpha) : \alpha \in \omega_1\}$ are pairwise disjoint. Again, passing to a subsequence, we may assume that with respect to the global enumeration of X $a_n^\alpha < a_1^\beta$ for all $\alpha < \beta$. Additionally, we may assume there are $\{m_1, \ldots, m_n\} \subseteq \mathbb{Q}$ such that $h_\alpha(a_i^\alpha) = m_i$ for all $\alpha \in \omega_1$ and $i \in \{1, \ldots, n\}$.

Now fix $\alpha < \beta$. In order for $h_\alpha \cup h_\beta$ to not be in $\mathbb{P}$, there must be some $p(\alpha, \beta), q(\alpha, \beta) \in \{1, \ldots, n\}$ such that $a^\alpha_{p(\alpha,\beta)}$ and $a^\beta_{q(\alpha,\beta)}$ are $\prec$-comparable. As a bookkeeping device, fix a uniform[4] ultrafilter $\mathcal{U}$ on ω_1.

Thus, for any $\alpha \in \omega_1$, there is some $S_\alpha \in \mathcal{U}$, some $p(\alpha), q(\alpha) \in \{1, \ldots, n\}$ such that, by ($\sharp$), for every $\beta \in S_\alpha$, $a^\alpha_{p(\alpha)}$ and $a^\beta_{q(\alpha)}$ are $\prec$-comparable. However, since $\mathrm{pred}(a^\alpha_{p(\alpha)})$ was assumed to be countable, there is $S^*_\alpha \subseteq S_\alpha$, $S^*_\alpha \in \mathcal{U}$ such that $a^\alpha_{p(\alpha)} \prec a^\beta_{q(\alpha)}$ for all $\beta \in S^*_\alpha$.

[4]That is, every $Y \in \mathcal{U}$ has cardinality $\aleph_1$. Equivalently, $\mathcal{U}$ contains all of the cocountable subsets of ω_1.

Similarly, there is some $S \in \mathcal{U}$ and some $p^*, q^* \in \{1, \ldots, n\}$ such that for all $\alpha \in S$ and for all $\beta \in S^*_\alpha$ we have $a^\alpha_{p^*} \prec a^\beta_{q^*}$. We obtain our contradiction by showing that

$$C = \{a^\alpha_{p^*} : \alpha \in S\}$$

is an uncountable chain in $(X, \prec)$. Since $\mathcal{U}$ is uniform, C is uncountable. To get comparability, choose any $\alpha, \gamma \in S$. As $S^*_\alpha, S^*_\gamma \in \mathcal{U}$, there is $\beta \in S^*_\alpha \cap S^*_\gamma$. It follows that $a^\alpha_{p^*} \prec a^\beta_{q^*}$ and $a^\gamma_{p^*} \prec a^\beta_{q^*}$. From our assumptions on $(X, \prec)$, $(\mathrm{pred}(a^\beta_{q^*}), \prec)$ is a linear order, so $a^\alpha_{p^*}$ and $a^\gamma_{p^*}$ are $\prec$-comparable. □

3.2. *Extendible and striated formulas.* Throughout this section, we work with the atomic models of a complete, first order theory T in a countable language that has an uncountable atomic model. We expound model theoretic properties needed in the forcing construction of Section 3.3.

Remark 3.2.1. In this section we work with complete formulas $\theta(\boldsymbol{w})$, usually with a prescribed partition of the free variables. Regardless of the partition, for any subsequence $\boldsymbol{v} \subseteq \boldsymbol{w}$, we use the notation $\theta\restriction_{\boldsymbol{v}}$ to denote the complete formula in the variables $\boldsymbol{v}$ that is equivalent to $\exists \boldsymbol{u}\theta(\boldsymbol{v}, \boldsymbol{u})$, where $\boldsymbol{u} = (\boldsymbol{w} \setminus \boldsymbol{v})$.

Definition 3.2.2. (1) A complete formula $\phi(x, \boldsymbol{a})$ is *pseudoalgebraic*[5] if for some/any countable M with $\boldsymbol{a} \in M$ and any $N \supsetneq M$, $\phi(N, \boldsymbol{a}) = \phi(M, \boldsymbol{a})$.

(2) $b \in \mathrm{pcl}(\boldsymbol{a}, M)$, written $b \in \mathrm{pcl}(\boldsymbol{a})$, if and only if $b \in N$ for every $N \preceq M$ with $\boldsymbol{a} \subseteq N$.

(3) A complete formula $\theta(\boldsymbol{z}; \boldsymbol{x})$ is *extendible* if there is a pair $M \preceq N$ of countable, atomic models and $\boldsymbol{b} \subseteq M$, $\boldsymbol{a} \subseteq N \setminus M$ such that $N \models \theta(\boldsymbol{b}, \boldsymbol{a})$.

Note that an atomic class has an uncountable model if and only if it has a non-pseudoalgebraic type.

The definition of an extendible formula depends on the partition of its free variables. As we require extendible formulas to be complete, they are not preserved under adjunction of dummy variables. If $\lg(\boldsymbol{x}) = 1$, then $\theta(\boldsymbol{z}, x)$ being extendible is equivalent to it being complete, with $\theta(\boldsymbol{z}, x)$ not pseudoalgebraic. Much of the utility of the notion is given by the following fact.

Fact 3.2.3. (1) *If $\theta(\boldsymbol{z}; \boldsymbol{x})$ is extendible, then for any countable, atomic M and any $\boldsymbol{b} \in M^{\lg(\boldsymbol{z})}$ and $\boldsymbol{a} \in M^{\lg(\boldsymbol{x})}$ such that $M \models \theta(\boldsymbol{b}, \boldsymbol{a})$, there is $M_0 \preceq M$ such that $\boldsymbol{b} \subseteq M_0$ and $\boldsymbol{a} \subseteq M \setminus M_0$.*

(2) *If $\theta(\boldsymbol{z}; \boldsymbol{x})$ is extendible and $\boldsymbol{z}' \subseteq \boldsymbol{z}$ and $\boldsymbol{x}' \subseteq \boldsymbol{x}$, then the restriction $\theta\restriction_{\boldsymbol{z}';\boldsymbol{x}'}$ is extendible as well.*

(3) *Any complete formula $\theta(\boldsymbol{z}; \boldsymbol{x})$ is extendible if and only if $\theta\restriction_{\boldsymbol{z},x_i}$ is not pseudo-algebraic for every $x_i \in \boldsymbol{x}$.*

[5]The careful distinctions of pseudoalgebraicity "in a model" of [Baldwin et al. 2016] are avoided because we have assumed there is an uncountable atomic model.

Proof. (1) As $\theta(z; \boldsymbol{x})$ is extendible, choose countable atomic models $M' \preceq N'$, $\boldsymbol{b}' \subseteq M$ and $\boldsymbol{a}' \subseteq N' \setminus M'$ such that $N' \models \theta(\boldsymbol{b}', \boldsymbol{a}')$. As $\theta(z; \boldsymbol{x})$ is complete, there is an isomorphism $f : N' \to M$ with $f(\boldsymbol{b}') = \boldsymbol{b}$ and $f(\boldsymbol{a}') = \boldsymbol{a}$. Then $M_0 := f(M')$ is as desired.

(2) This follows easily from the proof of (1).

(3) Left to right follows easily from (2). We prove the converse by induction on $\lg(\boldsymbol{x})$. For $\lg(\boldsymbol{x}) = 1$ this is immediate, so assume this holds when $\lg(\boldsymbol{x}) = n$. Choose a complete $\theta(z; \boldsymbol{x}, x_n)$ such that $\lg(\boldsymbol{x}) = n$ and $\theta\restriction_{z,x_i}$ is non-pseudoalgebraic for each $i \le n$. Choose any countable, atomic N and $\boldsymbol{b}, \boldsymbol{a}, a_n$ from N so that $N \models \theta(\boldsymbol{b}, \boldsymbol{a}, a_n)$. By (1), it suffices to find some $M_0 \preceq N$ with $\boldsymbol{b} \subseteq M_0$ and $\boldsymbol{a}a_n \subseteq N \setminus M_0$. To obtain this, since $\exists x_n \theta(z; \boldsymbol{x}, x_n)$ is extendible by (2), (1) implies there is $M \preceq N$ with $\boldsymbol{b} \subseteq M$ and $\boldsymbol{a} \subseteq N \setminus M$. Thus, if $a_n \in N \setminus M$, we can take $M_0 := M$ and we are done. If not, then as $\boldsymbol{b}a_n \subseteq M$ we can apply (1) to M and the extendible $\exists \boldsymbol{x} \theta(z; \boldsymbol{x}, x_n)$ to get $M_0 \preceq M$ with $\boldsymbol{b} \subseteq M_0$ and $a_n \in M \setminus M_0$. □

Next, we consider the "transitive closure" of extendibility.

Definition 3.2.4. An *n-striated formula* is a complete formula $\theta(\boldsymbol{y}_0, \dots, \boldsymbol{y}_{n-1})$ whose free variables are partitioned into n pieces such that, for every $i < n$, letting $z = (\boldsymbol{y}_0, \dots, \boldsymbol{y}_i)$ and $\boldsymbol{x} = (\boldsymbol{y}_i, \dots, \boldsymbol{y}_{n-1})$, we have $\theta(z, \boldsymbol{x})$ extendible.

A *striated formula* is an n-striated formula for some n.

A *realization* of an n-striated formula $\theta(\boldsymbol{y}_0, \dots, \boldsymbol{y}_{n-1})$ is an n-chain $M_0 \preceq M_1 \preceq M_{n-1}$ of countable, atomic models, together with tuples $\boldsymbol{a}_0, \dots, \boldsymbol{a}_{n-1}$ with $\boldsymbol{a}_0 \subseteq M_0$ and $\boldsymbol{a}_i \subseteq M_i \setminus M_{i-1}$ for every $0 < i < n$ such that $M_{n-1} \models \theta(\boldsymbol{a}_0, \dots, \boldsymbol{a}_{n-1})$.

Iterating Fact 3.2.3, we see that a partitioned complete formula $\theta(\boldsymbol{y}_0, \dots, \boldsymbol{y}_{n-1})$ is n-striated if and only if for some countable atomic M and some $(\boldsymbol{a}_0, \dots, \boldsymbol{a}_{n-1})$ from M with $M \models \theta(\boldsymbol{a}_0, \dots, \boldsymbol{a}_{n-1})$, there are $M_0 \preceq M_1 \preceq \cdots \preceq M_{n-2} \preceq M$ with $\boldsymbol{a}_0 \subseteq M_0$, $\boldsymbol{a}_i \subseteq M_i \setminus M_{i-1}$ for $0 < i < n-2$ and $\boldsymbol{a}_{n-1} \cap M_{n-2} = \varnothing$.

Using this characterization, if $\theta(\boldsymbol{y}_0, \dots, \boldsymbol{y}_{n-1})$ is n-striated and we modify the partition of θ by fusing together two adjacent tuples, then the resulting partition yields an $(n-1)$-striated formula. Going forward, we have the following amalgamation property for striated formulas.

Lemma 3.2.5. *Suppose $\alpha(z, \boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ and $\beta(z, \boldsymbol{y}_1, \dots, \boldsymbol{y}_m)$ are striated and $\alpha\restriction_z$ is equivalent to $\beta\restriction_z$. Then there is a striated $\psi(z, \boldsymbol{x}_1, \dots, \boldsymbol{x}_n, \boldsymbol{y}_1, \dots, \boldsymbol{y}_m)$ extending $\alpha(z, \boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \wedge \beta(z, \boldsymbol{y}_1, \dots, \boldsymbol{y}_m)$.*

Proof. Choose an $(n+1)$-chain $M_0 \preceq M_1 \preceq \cdots \preceq M_n$ and $\boldsymbol{b}, \boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ realizing α (so $\boldsymbol{b} \subseteq M_0$ and $\boldsymbol{a}_i \subseteq M_i \setminus M_{i-1}$ for each i) and choose similarly an $(m+1)$-chain $N_0 \preceq N_1 \preceq \cdots \preceq N_m$ and $\boldsymbol{c}, \boldsymbol{d}_1, \dots, \boldsymbol{d}_m$ realizing β. As $\alpha\restriction_z$ is equivalent to $\beta\restriction_z$, there is an isomorphism $f : N_0 \to M_n$ with $f(\boldsymbol{c}) = \boldsymbol{b}$. Choose $M_{n+m} \succeq M_n$ for which there is an isomorphism $f^* : N_m \to M_{n+m}$ extending f. Now, for $i \le m$ put

$M_{n+i} := f^*(N_i)$. (Note this is compatible with our previous placements.) Also, for each $1 \le i \le m$, put $\boldsymbol{a}_{n+i} := f^*(\boldsymbol{d}_i)$. Finally, put $\psi(z, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_m) := \operatorname{tp}(\boldsymbol{b}, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_{n+m})$. Then the $(n+m+1)$-chain $M_0 \preceq \cdots \preceq M_{n+m}$, together with $\boldsymbol{b}, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_{n+m}$ witness that ψ is striated. □

3.3. *The forcing.* We continue our assumption that we have a fixed complete theory T in a countable language with an uncountable atomic model. We fix an ω_1-like dense linear order $(I, \le)$ with least element 0 and fix a continuous, increasing (necessarily cofinal) sequence $\langle J_\alpha : \alpha \in \omega_1 \rangle$ of initial segments of I. Also, fix a set $X = \{x_{t,m} : t \in I, m \in \omega\}$ of distinct variable symbols and, for each $\alpha \in \omega_1$, let $X_\alpha = \{x_{t,m} : t \in J_\alpha, m \in \omega\}$. Our forcing below will describe a complete diagram in the variables X corresponding to an atomic model N of size $\aleph_1$ and the countable substructures N_α corresponding to the variables X_α will be a filtration of N.

Definition 3.3.1. The forcing $(\mathbb{P}, \le)$ consists of all conditions

$$p = \big(u_p, \ell(p), \{k_{p,i} : i < \ell(p)\}, \theta_p(\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{\ell(p)-1})\big)$$

satisfying the following properties:

(1) u_p is a finite subset $\{s_0, \ldots, s_{\ell(p)-1}\} \subseteq I$. We always write the elements of u_p in ascending order.

(2) $\ell(p) = |u_p|$.

(3) If $u_p \neq \varnothing$, then $0 \in u_p$.

(4) Each $k_{p,i} \in \omega$ and denotes $\lg(\boldsymbol{y}_i)$ in θ_p.

(5) $\theta_p(\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{\ell(p)-1})$ is an $\ell(p)$-striated formula, where each $\boldsymbol{y}_i = (\boldsymbol{x}_{s_i, j} : j < k_{p,i})$ is the initial segment of the s_i-th row of X of length $k_{p,i}$.

The ordering on $\mathbb{P}$ is natural, i.e., $p \le_{\mathbb{P}} q$ if and only if $u_p \subseteq u_q$, the free variables of θ_p are contained in the free variables of θ_q and $\theta_q \vdash \theta_p$.

We remark that the effect of requiring $0 \in u_p$ whenever u_p is nonempty is to ensure that if θ_p entails "$x_{\alpha_i, j} \in \operatorname{pcl}(\varnothing)$", then $\alpha_i = 0$. That is, in the generic model we construct, all pseudoalgebraic complete types of singletons will be contained in M_0.

It is easily verified that $(\mathbb{P}, \le)$ is c.c.c. (See [Baldwin et al. 2016, Claim 4.3.7] for a verification of this in an extremely similar setting.) As well, $(\mathbb{P}, \le)$ is highly homogeneous. In particular, we record the following facts, with (1) following from $(I, \le)$ being dense and ω_1-like.

Fact 3.3.2. (1) *For all $\alpha < \omega_1$ and for all finite $u_1, u_2 \subseteq I \setminus J_\alpha$ with $|u_1| = |u_2|$ and min$(I \setminus J_\alpha)$ (if it exists) $\notin u_1 \cup u_2$, then there is an order isomorphism $\sigma \in \operatorname{Aut}(I, \le)$ with $\sigma(u_1) = u_2$ and $\sigma\restriction_{J_\alpha} = \mathrm{id}$.*

(2) *Any order isomorphism $\sigma \in \operatorname{Aut}(I, \le)$ induces both a permutation $\sigma' \in \operatorname{Sym}(X)$ given by $\sigma'(x_{t,m}) = x_{\sigma(t),m}$ and an automorphism $\sigma^* \in \operatorname{Aut}(\mathbb{P}, \le)$ given by*

$$\sigma^*(p) = \big(\sigma(u_p), \ell_p, \{k_{p,i} : i < \ell(p)\}, \theta_p(\sigma'(\boldsymbol{y}_0), \ldots, \sigma'(\boldsymbol{y}_{\ell(p)-1}))\big).$$

We record three additional density conditions about $(\mathbb{P}, \leq)$ whose verifications depend on the following fact.

Lemma 3.3.3. *Suppose $\delta(x)$ is a non-pseudoalgebraic 1-type. Then for every countable atomic N and every $\boldsymbol{e} \subseteq N$, there are $M \preceq N$ and $c \in N \setminus M$ such that $\boldsymbol{e} \subseteq M$ and $N \models \delta(c)$.*

Proof. From the definition of (non)-pseudoalgebraicity, fix countable atomic $M^* \preceq N^*$ and $c^* \in N^* \setminus M^*$ with $N^* \models \delta(c^*)$. Choose any isomorphism $f : N \to M^*$ and put $\boldsymbol{e}^* := f(\boldsymbol{e})$. Now, choose an isomorphism $g : N^* \to N$ with $g(\boldsymbol{e}^*) = \boldsymbol{e}$. Put $M := g(M^*)$ and $c := g(c^*)$. Then $\boldsymbol{e} \subseteq M$, $c \in N \setminus M$, and $N \models \delta(c)$. □

The forcing is surjective in the sense that for every condition p and every variable there is an extension of p that includes the variable.

Lemma 3.3.4 (surjective). *For every $p \in \mathbb{P}$ and $x_{t,m} \in X$, there is $q \in \mathbb{P}$, $q \geq p$ with $\boldsymbol{x}_q = \boldsymbol{x}_p \cup \{x_{t,m}\}$.*

Proof. We may assume that $p \neq 0$ and that $x_{t,m} \notin \boldsymbol{x}_p$. Choose $M_0 \preceq M_1 \preceq \cdots \preceq M_{n-1}$ and $\boldsymbol{e}_0 \ldots \boldsymbol{e}_{n-1}$ realizing θ_p (so $\boldsymbol{e}_0 \subseteq M_0$, $\boldsymbol{e}_i \subseteq M_i \setminus M_{i-1}$ for $0 < i < n$ and $M_{n-1} \models \theta(\boldsymbol{e}_0, \ldots, \boldsymbol{e}_{n-1})$).

We first handle the case where $m = 0$. In this case, it must be that $t \notin u_p$. Choose j maximal such that $s_j < t$. Apply Lemma 3.3.3 to M_j and $\boldsymbol{e}_0 \ldots \boldsymbol{e}_j$ to get $M_j^* \preceq M_j$ and $c \in M_j \setminus M_j^*$ with $M_j \models \delta(c)$ and $\boldsymbol{e}_0 \ldots, \boldsymbol{e}_j \subseteq M_j^*$. Now let $f : M_j \to M_j^*$ be an isomorphism fixing $\boldsymbol{e}_0 \ldots, \boldsymbol{e}_j$ pointwise. Then the type $\mathrm{tp}(\boldsymbol{e}_0, \ldots, \boldsymbol{e}_j, c, \boldsymbol{e}_{j+1}, \ldots, \boldsymbol{e}_{n-1})$ and the $(n+1)$-chain $f(M_0) \preceq \ldots f(M_j) \preceq M_j \preceq M_{j+1} \preceq \ldots M_{n-1}$ describes an $(n+1)$-striated formula θ. Let $q \in \mathbb{P}$ be the element with $\boldsymbol{x}_q = \boldsymbol{x}_p \cup \{x_{t,0}\}$ with $\theta_q(\boldsymbol{x}_q)$ being the complete formula generating this type.

If $m > 0$, then we apply the previous case to ensure that $x_{t,0} \in \boldsymbol{x}_p$. Say $t = s_j$, the j-th element of u_p. But then, given any $\boldsymbol{e}_0, \ldots, \boldsymbol{e}_{n-1}$ and $M_0 \preceq \cdots \preceq M_{n-1}$ realizing θ_p, extend $\boldsymbol{x}_{p,t}$ to include $x_{t,m}$ by making each "new" element of $\boldsymbol{e}_j$ equal to the element $e_{j,0} \in M_j$. □

The notational issue in what follows is the placement of free variables. For $p \in \mathbb{P}$, there is an explicit ordering to the variables $\boldsymbol{x}_p$ occurring in $\theta(\boldsymbol{x}_p)$, but when we consider extensions $\phi(\boldsymbol{v}, \boldsymbol{x}_p)$, we do not want to specify where the v_i's fit in the sequence. Recall Definition 3.3.1(5).

Lemma 3.3.5 (Henkin). *Suppose $p \in \mathbb{P}$ and $\theta_p(\boldsymbol{x}_p) \vdash \exists \boldsymbol{v} \phi(\boldsymbol{v}, \boldsymbol{x}_p)$. Then there is $q \in \mathbb{P}$, $q \geq p$ for which the variables in $(\boldsymbol{x}_q \setminus \boldsymbol{x}_p)$ consist of a realization of $\phi(\boldsymbol{v}, \boldsymbol{x}_p)$ (in some order). Moreover, if $p \neq 0$, then can be chosen with $u_q = u_p$.*

Proof. Arguing by induction, we may assume $\boldsymbol{v} = \{v\}$ is a singleton, and we may further assume that $\phi(v, \boldsymbol{x}_p)$ describes a complete type. Let $\boldsymbol{e}_0, \ldots, \boldsymbol{e}_{n-1}$ and $M_0 \preceq \cdots \preceq M_{n-1}$ witness the truth and striation of θ_p and choose any $b \in M_{n-1}$

such that $M_{n-1} \models \phi(b, \boldsymbol{e}_p)$. Let $j \leq n-1$ be least such that $b \in M_j$. (Note that if $\phi(v, \boldsymbol{x}_p) \vdash$ '$v \in \mathrm{pcl}(\varnothing)$', then we must have $j = 0$.) Let $\boldsymbol{x}_q = \boldsymbol{x}_p \cup \{x_{s_j, k_{p,j}}\}$. Then, letting $\boldsymbol{e}_j^* = \boldsymbol{e}_j b$, we have a striation $\boldsymbol{e}_0, \ldots, \boldsymbol{e}_{j-1}, \boldsymbol{e}_j^*, \boldsymbol{e}_{j+1}, \ldots, \boldsymbol{e}_{n-1}$ using the same n-chain of models $M_0 \preceq \ldots M_{n-1}$. Put

$$\theta_q(\boldsymbol{x}_q) := \mathrm{tp}(\boldsymbol{e}_0, \ldots, \boldsymbol{e}_{j-1}, \boldsymbol{e}_j^*, \boldsymbol{e}_{j+1}, \ldots, \boldsymbol{e}_{n-1}).$$

Then $q \in \mathbb{P}$ and $q \geq p$. □

Lemma 3.3.6. *Suppose $p, q, r \in \mathbb{P}$ with $p \leq q$, $p \leq r$, $\boldsymbol{x}_q \cap \boldsymbol{x}_r = \boldsymbol{x}_p$, and for some $t \in I$, $u_q \subseteq I_{<t}$ and $(u_r \setminus u_p) \subseteq I_{>t}$. Suppose further that there are $M \preceq N$ and $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ with $\boldsymbol{b} \cap \boldsymbol{c} = \boldsymbol{a}$, $\boldsymbol{b} \subseteq M$, and $(\boldsymbol{c} \setminus \boldsymbol{a}) \subseteq N \setminus M$ with $N \models \theta_p(\boldsymbol{a}) \wedge \theta_q(\boldsymbol{b}) \wedge \theta_r(\boldsymbol{c})$. Then there is $r^* \in \mathbb{P}$, $r^* \geq q$, $r^* \geq r$ with $\boldsymbol{x}_{r^*} = \boldsymbol{x}_q \cup \boldsymbol{x}_r$ and $\theta_{r^*} = \mathrm{tp}(\boldsymbol{b}, (\boldsymbol{c} \setminus \boldsymbol{a}))$.*

Proof. Arguing by induction, we may additionally assume that $u_r = u_p \cup \{s^*\}$ for some single $s^* > t$. That is, $\boldsymbol{x}_q \setminus \boldsymbol{x}_p$ lies on a single level of X. Since $q \in \mathbb{P}$, there is a striation of $\boldsymbol{b} = \boldsymbol{b}_0, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_{n-1}$ induced by the rows of $\boldsymbol{x}_q$. As $\boldsymbol{b} \subseteq M$, we can find an n-chain $M_0 \preceq M_{n-1}$ of models with $M_{n-1} = M$, $\boldsymbol{b}_0 \subseteq M_0$ and $\boldsymbol{b}_i \subseteq M_i \subseteq M_{i-1}$ for all $0 < i < n$. As $(\boldsymbol{c} \setminus \boldsymbol{a}) \subseteq N \setminus M$ and as $(\boldsymbol{x}_q \setminus \boldsymbol{x}_p)$ consists of a single row (and since $s^* > t$) it follows that the $(n+1)$-tuple $\boldsymbol{b}_0, \ldots, \boldsymbol{b}_{n-1}, (\boldsymbol{c} \setminus \boldsymbol{a})$ is realized in the $(n+1)$-chain $M_0 \preceq \cdots \preceq M \preceq N$. Choose $\boldsymbol{x}_{r^*} = \boldsymbol{x}_q \cup \boldsymbol{x}_r$ and put $\theta_{r^*} = \mathrm{tp}(\boldsymbol{b}_0, \ldots, \boldsymbol{b}_{n-1}, (\boldsymbol{c} \setminus \boldsymbol{a}))$. Then $r^* \in \mathbb{P}$ and both $r^* \geq q$, $r^* \geq r$ hold. □

Armed with these lemmas, we can now prove the main fact about the forcing $(\mathbb{P}, \leq)$ and the generic model $\underset{\sim}{N}$ of T. For general forcing notation see [Kunen 1980]. However, note that contrary to Kunen, we use the convention that $p \leqslant q$ means q is a stronger condition, carrying more information.

Notation 3.3.7. In what follows, when dealing with L-formulas, we will use the letters $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$, possibly with decorations to denote free variables. By contrast, tuples denoted by $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ denote finite tuples from X. Thus, for example, $\eta(\boldsymbol{v}, \boldsymbol{z})$ has free variables $\boldsymbol{v}$, and $\boldsymbol{z}$ is a fixed tuple from X.

We first establish that $(\mathbb{P}, \leq)$ forces an uncountable atomic model of T. This initial lemma only uses the surjective and Henkin density conditions (Lemmas 3.3.4 and 3.3.5). More details of this initial construction can be found in [Baldwin et al. 2016, §4.4].

Lemma 3.3.8. *There are $\mathbb{P}$-names $\underset{\sim}{N}$ and $\underset{\sim}{N}_\alpha$ for each $\alpha \in \omega_1$ such that*

$$(\mathbb{P}, \leq) \Vdash \text{“}\underset{\sim}{N} \in \mathrm{At},\ |\underset{\sim}{N}| = \aleph_1, \textit{ and } (\underset{\sim}{N}_\alpha : \alpha \in \omega_1) \textit{ is a filtration of } \underset{\sim}{N}\text{”}.$$

Proof. For every n-ary atomic L-formula $\phi(\boldsymbol{u})$, choose a $\mathbb{P}$-name $\underset{\sim}{\phi}$ such that, for every generic subset $G \subseteq \mathbb{P}$ (recalling Remark 3.2.1),

$$\underset{\sim}{\phi}[G] = \{\boldsymbol{x} \in X^n : T + \theta_p\restriction_{\boldsymbol{x}} \vdash \phi(\boldsymbol{x}) \text{ for some } p \in G\}.$$

In particular, for the atomic formula of equality, we have a $\mathbb{P}$-name $\underset{\sim}{E}$ representing "equality" on X^2. As each θ_p is consistent with T, it follows that $(\mathbb{P}, \leq)$ forces that $\underset{\sim}{E}$ is an L-congruence. Choose a $\mathbb{P}$-name $\underset{\sim}{N}$ representing L-structure whose universe is the quotient $X/\underset{\sim}{E}$ and whose atomic formulas are interpreted as $\underset{\sim}{\phi}$, and choose $\mathbb{P}$-names $\underset{\sim}{N}_\alpha$ for the substructure with universe $X_\alpha/\underset{\sim}{E}$.

Continuing, for every L-formula $\psi(\boldsymbol{u})$ (with quantifiers) choose a $\mathbb{P}$-name $\underset{\sim}{\psi}$ analogously to $\underset{\sim}{\phi}$. Using the Henkin density conditions, a straightforward induction on the quantifier complexity of ψ shows that for every $\boldsymbol{x} \in X^n$ and generic $G \subseteq \mathbb{P}$,

$$\mathbb{V}[G] \models \text{“}\underset{\sim}{N} \models \psi(\boldsymbol{x}/\underset{\sim}{E})\text{”} \iff \boldsymbol{x} \in \underset{\sim}{\psi}[G]$$

and similarly for each $\underset{\sim}{N}_\alpha$. From this, it follows that $(\mathbb{P}, \leq)$ forces that each $\underset{\sim}{N}_\alpha \preceq \underset{\sim}{N}$. As each θ_p is a (consistent) complete formula with respect to T, $(\mathbb{P}, \leq)$ also forces that $\underset{\sim}{N}$ is an atomic model of T. Finally, since each θ_p is a striated formula, we see that $(\mathbb{P}, \leq)$ also forces $\underset{\sim}{N}_{\alpha+1}$ properly extends $\underset{\sim}{N}_\alpha$, hence forces $|\underset{\sim}{N}| = \aleph_1$. □

It remains to show that $(\mathbb{P}, \leq)$ forces that $\underset{\sim}{N}$ has every limit type constrained. For this, we note a consequence of splitting inside an atomic model.

Remark 3.3.9. Suppose $M \preceq N$ are atomic, $\boldsymbol{a} \in M$, $\boldsymbol{b} \in N$, but $\mathrm{tp}(\boldsymbol{b}/M)$ splits over $\boldsymbol{a}$. Then, letting $\theta(\boldsymbol{u})$ isolate the complete type of $\boldsymbol{a}$ and $\theta'(\boldsymbol{w}, \boldsymbol{u})$ isolate the complete type of $\boldsymbol{ba}$, there must be a complete formula $\eta(\boldsymbol{v}, \boldsymbol{u}) \vdash \theta(\boldsymbol{u})$ and two contradictory complete formulas $\delta_1(\boldsymbol{w}, \boldsymbol{v}, \boldsymbol{u})$ and $\delta_2(\boldsymbol{w}, \boldsymbol{v}, \boldsymbol{u})$, each extending the (incomplete) formula $\eta(\boldsymbol{v}, \boldsymbol{u}) \wedge \theta'(\boldsymbol{w}, \boldsymbol{u})$.

Proposition 3.3.10. $(\mathbb{P}, \leq)$ *forces every limit type in* $S_{at}(\underset{\sim}{N})$ *to be constrained.*

Proof. To ease notation, in what follows write ψ in place of the more cumbersome $\underset{\sim}{\psi}$ throughout the argument. Call a function $b : \omega_1 \to \underset{\sim}{N}$ a *limit sequence* if, for all $\alpha \leq \beta$, $\mathrm{tp}(b(\alpha)/\underset{\sim}{N}_\alpha) = \mathrm{tp}(b(\beta)/\underset{\sim}{N}_\alpha)$. Now, if $(\mathbb{P}, \leq)$ does not force that every limit type is constrained, then there is some $p^* \in \mathbb{P}$ and some $\mathbb{P}$-name $\underset{\sim}{\boldsymbol{b}}$ and some club $C \subseteq \omega_1$ such that

$$p^* \Vdash \underset{\sim}{\boldsymbol{b}} \text{ is a limit sequence with } \mathrm{tp}(\underset{\sim}{\boldsymbol{b}}(\alpha)/\underset{\sim}{N}_\alpha) \text{ unconstrained for every } \alpha \in C.$$

(Since $(\mathbb{P}, \leq)$ is c.c.c. we can find such a club in $\mathbb{V}$.)

For each $\alpha \in C$, choose $p_\alpha \in \mathbb{P}$, $p_\alpha \geq p^*$ and $x^*_\alpha \in X$ such that

$$p_\alpha \Vdash \underset{\sim}{\boldsymbol{b}}(\alpha) = x^*_\alpha.$$

We will eventually reach a contradiction by finding some $q^* \geq p^*$ and some $\alpha < \beta$ from C such that

$$q^* \Vdash \mathrm{tp}(x^*_\alpha/N_\alpha) \neq \mathrm{tp}(x^*_\beta/N_\alpha)$$

contradicting that $p^* \Vdash \underset{\sim}{b}$ is a limit sequence. By a routine Δ-system argument, find a "root" $p_0 \in \mathbb{P}$, some $\gamma^* \in \omega_1$, and a stationary set $S \subseteq C$ satisfying

- $p_0 \leq p_\alpha$ for all $\alpha \in S$;
- $u_{p_0} \subseteq J_{\gamma^*}$ (first paragraph of Section 3.3); and
- for all $\alpha < \beta$ in S,
  - $\boldsymbol{x}_{p_\alpha} \cap X_{\gamma^*} = \boldsymbol{x}_{p_0}$;
  - $\max(u_{p_\alpha}) < \min(u_{p_\beta} \setminus u_{p_0})$;
  - $\lg(p_\alpha) = \lg(p_\beta)$ and $k_{p_\alpha} = k_{p_\beta}$; and
  - the formulas θ_{p_α} and θ_{p_β} have the same syntactic shape (one formula can be obtained from the other by substituting the free variables).

Note that we do not require $p_0 \geq p^*$. As notation, we write $\boldsymbol{z}$ for $\boldsymbol{x}_{p_0}$ and note that $\boldsymbol{z} \subseteq X_{\gamma^*}$. Now fix, for the remainder of the argument, some $\alpha < \beta$ from S. To obtain our desired contradiction, we first concentrate on p_α. Write $\theta_{p_\alpha}(\boldsymbol{y}, \boldsymbol{z})$ and note that $\boldsymbol{y}$ is disjoint from X_{γ^*}. We apply Remark 3.3.9, noting that $p_\alpha \Vdash \operatorname{tp}(x^*_\alpha / \underset{\sim}{N}_\alpha)$ splits over $\boldsymbol{z}$. Choose a complete formula $\eta(\boldsymbol{v}, \boldsymbol{z})$ implying $\theta_{p_0}(\boldsymbol{z})$ and contradictory complete formulas $\delta_1(x^*_\alpha, \boldsymbol{v}, \boldsymbol{z})$ and $\delta_2(x^*_\alpha, \boldsymbol{v}, \boldsymbol{z})$, each extending $\eta(\boldsymbol{v}, \boldsymbol{z}) \wedge \theta^*_{p_\alpha}(x^*_\alpha, \boldsymbol{z})$, where $\theta^*_{p_\alpha}$ is the restriction of the compete formula $\theta_{p_\alpha}(\boldsymbol{y}, \boldsymbol{z})$.

By Henkin, choose $q_0 \in \mathbb{P}$, $q_0 \geq p_0$ with $u_{q_0} \subseteq J_\alpha$ and $\theta_{q_0}(\boldsymbol{z}', \boldsymbol{z}) := \eta(\boldsymbol{z}', \boldsymbol{z})$. Next, we use Lemma 3.3.6 twice. In both cases we start with $p_0 \leq q_0$ and $p_0 \leq p_\alpha$. Our first application gives $r^1_\alpha \in \mathbb{P}$ extending both q_0 and p_α with $\theta_{r^1_\alpha}(\boldsymbol{y}, \boldsymbol{z}', \boldsymbol{z}) \vdash \delta_1(x^*_\alpha, \boldsymbol{z}', \boldsymbol{z})$. The second application gives $r^2_\alpha \in \mathbb{P}$, also extending both q_0 and p_α with $\theta_{r^2_\alpha}(\boldsymbol{y}, \boldsymbol{z}', \boldsymbol{z}) \vdash \delta_2(x^*_\alpha, \boldsymbol{z}', \boldsymbol{z})$.

Next, we use the fact that the forcing $(\mathbb{P}, \leq)$ is highly homogeneous. Due to the similarity of p_α and p_β found by the Δ-system argument and described in the third bullet point just above, Fact 3.3.2 gives an automorphism σ of $(\mathbb{P}, \leq)$ sending p_α to p_β, fixing q_0. Put $r^2_\beta := \sigma(r^2_\alpha)$. We now apply Lemma 3.2.5 to $q_0 \leq r^1_\alpha$ and $q_0 \leq r^2_\beta$ to get $q^* \in \mathbb{P}$ with $q^* \geq r^1_\alpha$ and $q^* \geq r^2_\beta$. However, this is impossible, as

$$q^* \Vdash \delta_1(x^*_\alpha, \boldsymbol{z}', \boldsymbol{z}) \wedge \delta_2(x^*_\beta, \boldsymbol{z}', \boldsymbol{z}),$$

contradicting $p^* \Vdash \operatorname{tp}(x^*_\alpha / \underset{\sim}{N}_\alpha) = \operatorname{tp}(x^*_\beta / \underset{\sim}{N}_\alpha)$ since δ_1 and δ_2 were chosen to be contradictory. □

Proof of Theorem 2.3.2. Theorem 2.3.2 follows easily from Propositions 3.1.1 and 3.3.10 and Keisler's model existence result for $L_{\omega_1,\omega}(Q)$. In particular, in some c.c.c. forcing extension $\mathbb{V}[G]$, by Proposition 3.3.10, there is an uncountable atomic model of T with every limit type constrained. Hence, by (2) $\Rightarrow$ (1) of Proposition 3.1.1, there is a model of Ψ^* in $\mathbb{V}[G]$. By the absoluteness of model existence from Keisler's theorem, there is also a model of Ψ^* in $\mathbb{V}$. Hence, by (1) $\Rightarrow$ (2) of Proposition 3.1.1, we obtain the existence of an atomic model of T in $\mathbb{V}$ such that all limit types are constrained. □

References

[Baldwin 2009] J. T. Baldwin, *Categoricity*, University Lecture Series **50**, American Mathematical Society, Providence, RI, 2009. MR Zbl

[Baldwin 2012] J. T. Baldwin, "Amalgamation, absoluteness, and categoricity", pp. 22–50 in *Proceedings of the* 11*th Asian Logic Conference*, edited by T. Arai et al., World Scientific, 2012. MR Zbl

[Baldwin et al. 2016] J. T. Baldwin, M. C. Laskowski, and S. Shelah, "Constructing many atomic models in $\aleph_1$", *J. Symb. Log.* **81**:3 (2016), 1142–1162. MR Zbl

[Baumgartner 1970] J. E. Baumgartner, *Results and independence proofs in combinatorial set theory*, Ph.D. thesis, University of California, Berkeley, 1970. MR Zbl

[Keisler 1970] H. J. Keisler, "Logic with the quantifier "there exist uncountably many"", *Ann. Math. Logic* **1** (1970), 1–93. MR Zbl

[Kunen 1980] K. Kunen, *Set theory: an introduction to independence proofs*, Studies in Logic and the Foundations of Mathematics **102**, North-Holland, 1980. MR Zbl

[Laskowski and Shelah 2019] M. C. Laskowski and S. Shelah, "A strong failure of $\aleph_0$-stability for atomic classes", *Arch. Math. Logic* **58**:1-2 (2019), 99–118. MR Zbl

[Morley 1965] M. Morley, "Categoricity in power", *Trans. Amer. Math. Soc.* **114** (1965), 514–538. MR Zbl

[Shelah 1975] S. Shelah, "Categoricity in $\aleph_1$ of sentences in $L_{\omega_1,\omega}(Q)$", *Israel J. Math.* **20**:2 (1975), 127–148. Shelah index 48. MR Zbl

[Shelah 1983a] S. Shelah, "Classification theory for nonelementary classes, I: The number of uncountable models of $\psi \in L_{\omega_1,\omega}$, Part A", *Israel J. Math.* **46**:3 (1983), 212–240. Shelah index 87a. MR Zbl

[Shelah 1983b] S. Shelah, "Classification theory for nonelementary classes, I: The number of uncountable models of $\psi \in L_{\omega_1,\omega}$, Part B", *Israel J. Math.* **46**:4 (1983), 241–273. Shelah index 87b. Zbl

[Shelah 1987] S. Shelah, "Classification of nonelementary classes, II: Abstract elementary classes", pp. 419–497 in *Classification theory* (Chicago, IL, 1985), edited by J. T. Baldwin, Lecture Notes in Math. **1292**, Springer, 1987. MR Zbl

[Vaught 1961] R. L. Vaught, "Denumerable models of complete theories", pp. 303–321 in *Infinitistic Methods, Proceedings of the Symposium on the Foundations of Mathematics* (Warsaw, 1959), Państwowe Wydawnictwo Naukowe, Warsaw, 1961. Zbl

Received 26 Aug 2023. Revised 13 Feb 2024.

JOHN T. BALDWIN:
jbaldwin@uic.edu
University of Illinois at Chicago, Chicago, IL, United States

MICHAEL C. LASKOWSKI:
laskow@umd.edu
University of Maryland, College Park, MD, United States

SAHARON SHELAH:
shelah@math.huji.ac.il
Hebrew University of Jerusalem, Jerusalem, Israel

Model Theory
Vol. 3, No. 3, 2024
https://doi.org/10.2140/mt.2024.3.825

A New Kim's Lemma

Alex Kruckman and Nicholas Ramsey

Kim's Lemma is a key ingredient in the theory of forking independence in simple theories. It asserts that if a formula divides, then it divides along every Morley sequence in type of the parameters. Variants of Kim's Lemma have formed the core of the theories of independence in two orthogonal generalizations of simplicity — namely, the classes of NTP_2 and NSOP_1 theories. We introduce a new variant of Kim's Lemma that simultaneously generalizes the NTP_2 and NSOP_1 variants. We explore examples and nonexamples in which this lemma holds, discuss implications with syntactic properties of theories, and ask several questions.

1. Introduction

The simple theories are a class of first-order theories which admit a structure theory built upon a good notion of independence. Nonforking independence was introduced by Shelah [1990] in the context of classification theory for stable theories, but was later shown to be meaningful in a broad class of unstable theories. Shelah's characterization [1980] of simple theories in terms of their saturation spectra, together with Hrushovski's work [2002] on bounded PAC structures and structures of finite S_1-rank, and the work of Cherlin and Hrushovski [2003] on quasi-finite theories, all made use of a circle of ideas concerning independence and amalgamation. These ideas were subsequently distilled and consolidated into the core results of simplicity theory by Kim [1998] and Kim and Pillay [1997], organized around

Ramsey was supported by NSF grant DMS-2246992 and an AMS-Simons Travel Grant.
MSC2020: 03C45.
Keywords: classification theory, forking, dividing, Kim-dividing.

the good behavior of nonforking independence in this setting. A key ingredient in this theory is a result known as *Kim's Lemma*, which establishes that, in a simple theory, a formula $\varphi(x; b)$ divides over a set A if and only if $\varphi(x; b)$ divides along every Morley sequence over A in $\mathrm{tp}(b/A)$. Kim's lemma says that dividing is always witnessed by "generic" indiscernible sequences and changes the existential quantifier in the definition of dividing ("*there is* an A-indiscernible sequence such that...") into a universal one ("*for every* Morley sequence over A..."). Kim [2001] later showed that Kim's lemma characterizes the simple theories.

More recent developments have highlighted the centrality of Kim's lemma to the theory of nonforking independence and its generalizations. In particular, the theories of independence in NTP_2 and NSOP_1 theories are based on two orthogonal generalizations of Kim's lemma.[1] For NTP_2 theories, the equivalence between dividing and dividing along all generic sequences is preserved, but this equivalence requires a stronger notion of genericity. More specifically, Chernikov and Kaplan [2012] showed that, in an NTP_2 theory, a formula $\varphi(x; b)$ divides over a model M if and only if $\varphi(x; b)$ divides along Morley sequences for every strictly M-invariant type extending $\mathrm{tp}(b/M)$. This variant of Kim's lemma was shown to characterize NTP_2 theories in [Chernikov 2014].

On the other hand, in NSOP_1 theories, the equivalence between dividing and dividing along generic sequences no longer holds in general. Nonetheless, at the generic scale, there is an analogue of Kim's lemma: a formula $\varphi(x; b)$ divides along *some* generic sequence in $\mathrm{tp}(b/M)$ over a model M if and only if it divides along *every* such sequence. More precisely, Kaplan and the second-named author introduced *Kim-dividing*, which is defined so that a formula $\varphi(x; b)$ Kim-divides over a model M if $\varphi(x; b)$ divides along some Morley sequence for a global M-invariant type extending $\mathrm{tp}(b/M)$. It was shown in [Kaplan and Ramsey 2020] that, in an NSOP_1 theory, $\varphi(x; b)$ Kim-divides over M if and only if it divides along Morley sequences for every global M-invariant type extending $\mathrm{tp}(b/M)$ and that, moreover, this variant of Kim's lemma characterizes NSOP_1 theories.

We introduce a "New Kim's Lemma" that simultaneously generalizes the Kim's Lemmas for NTP_2 and NSOP_1 theories. The starting point is an observation about the broom lemma of Chernikov and Kaplan [2012]. This lemma is the key step in showing that, in NTP_2 theories, types over models always have global strict invariant extensions, which generate the generic sequences needed to get a Kim's lemma for NTP_2 theories. However, an inspection of the proof shows that this fact really bundles together two separate statements. The first is that in NTP_2

[1] As a consequence of Mutchnik's work [2022b], we now know that the properties NSOP_1, NSOP_2, and NTP_1 are equivalent at the level of theories. In this paper, we primarily refer to NSOP_1 theories (rather than to NSOP_2 or NTP_1 theories), since the notion of Kim-independence was originally developed in [Kaplan and Ramsey 2020] under this hypothesis.

theories, Kim-dividing and forking independence coincide over models. The second is that, in any theory whatsoever, types over models extend to global *Kim-strict* invariant types, where Kim-strictness relaxes the nonforking independence condition required for strictness to one that only requires non-Kim-forking; see Theorem 2.26 below.

The statement of New Kim's Lemma, then, suggests itself (see Definition 3.7 below): a formula $\varphi(x; b)$ Kim-divides over a model M if and only if it divides along Morley sequences for every Kim-strictly M-invariant type extending $\mathrm{tp}(b/M)$. This variant of Kim's Lemma coincides with the Chernikov–Kaplan Kim's Lemma in NTP_2 theories (since there, Kim-forking agrees with forking over models, and hence Kim-strict invariant types are strict invariant), and coincides with the Kaplan–Ramsey Kim's Lemma in NSOP_1 theories (since there, Kim-independence is symmetric, so invariant types are automatically Kim-strict).

In Section 3, we survey the Kim's lemmas of the past and introduce our New Kim's Lemma. We also observe that New Kim's Lemma implies that Kim-forking equals Kim-dividing at the level of formulas. In Section 4, we show that our variant of Kim's Lemma holds in some examples of interest, including parametrized dense linear orders and the two-sorted theory of an infinite dimensional vector space over a real closed field with a bilinear form which is alternating and nondegenerate or symmetric and positive-definite. Our choice of examples is motivated by the idea that structures obtained by "generically putting together" NTP_2 and NSOP_1 behavior should satisfy New Kim's Lemma. We show, however, that New Kim's Lemma does *not* hold in the generic triangle-free graph, suggesting that it could serve as a meaningful dividing line among theories.

In Section 5, we try to relate New Kim's Lemma to syntactic properties of formulas. Our approach here reverses the usual order of explanation in neostability theory, which typically begins with a syntactic property (e.g., the tree property, SOP_1, TP_2) and then tries to establish a structure theory for theories without this property. In contrast, we are starting with a structural feature and trying to find a way of characterizing it syntactically. We introduce a new combinatorial configuration, which we provisionally call the bizarre tree property (BTP). The class of NBTP theories (those without BTP) contains both NTP_2 and NSOP_1, and all NBTP theories satisfy New Kim's Lemma. However, we do not obtain an exact characterization.

The antichain tree property (ATP), which was introduced in [Ahn and Kim 2024] and developed in [Ahn et al. 2023], is another combinatorial configuration generalizing TP_2 and SOP_1. We observe that NBTP implies NATP. But it is not clear whether there is an implication in either direction between NATP and New Kim's Lemma, or whether NBTP and NATP are equivalent. Figure 1 summarizes the state of affairs.

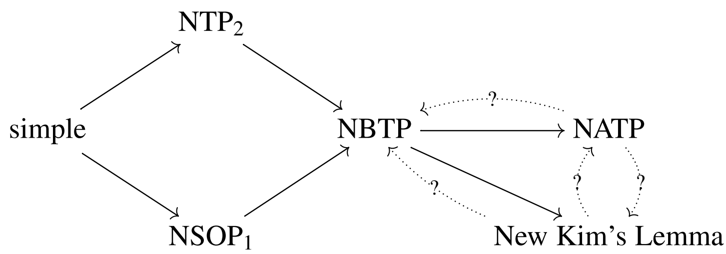


Figure 1. The current state of known implications.

While this paper was in preparation, two closely related preprints appeared:

- Kim and Lee [2023] establish a different variant of Kim's Lemma for NATP theories. Similarly to our work here, they do not prove that this Kim's Lemma characterizes NATP. In the context of NATP, they also study dividing along coheir sequences which are Kim-strict in the sense of this paper.
- Hanson [2023] studies a number of variants of Kim's Lemma which are related to ours. In particular, he succeeds in characterizing the class of NCTP theories by means of a variant of Kim's Lemma. Here CTP is the *comb tree property* (which was introduced by Mutchnik [2022b] under the name ω-DCTP$_2$). The class of NCTP theories contains the NBTP theories and its contained in the NATP theories.

At the moment, the NATP theories are the class beyond NSOP$_1$ and NTP$_2$ with the most developed syntactic theory; it would be very satisfying if these three approaches coincide. We conclude in Section 6 with several questions on where the theory might go from here.

2. Preliminaries

Throughout, T is a complete L-theory and $\mathbb{M} \models T$ is a monster model. As usual, all tuples come from $\mathbb{M}$, all sets are small subsets of $\mathbb{M}$, and all models are small elementary submodels of $\mathbb{M}$.

When α is an ordinal, we view the set $\alpha^{<\omega}$ of all finite sequences from α as a tree, with the tree partial order denoted by $\trianglelefteq$. The root of the tree is the empty sequence $\langle\,\rangle$. For $\rho \in \alpha^\omega$ and $i < \omega$, $\rho \mid i \in \alpha^{<\omega}$ is the restriction of ρ to i. We write $\eta^\frown \nu$ for concatenation of sequences. We write $\eta \perp \nu$ when η and ν are incomparable in the tree order. An *antichain* is a set of pairwise incomparable elements.

2A. ***Tree properties.*** We will begin by recalling the definitions of a number of tree properties and the known implications between them. The following three tree properties were introduced by Shelah [1990] under different names as part of his analysis of forking in stable theories.[2] He introduced the "tree property" terminology in [Shelah 1980] and Kim [2001] subsequently dubbed the latter two as TP_1 and TP_2.

Definition 2.1. Let $\varphi(x; y)$ be a formula:

(1) We say $\varphi(x; y)$ has the *tree property* (TP) if there is $k < \omega$ and a tree of tuples $(a_\eta)_{\eta \in \omega^{<\omega}}$ satisfying the following conditions:
   (a) For all $\rho \in \omega^\omega$, $\{\varphi(x; a_{\rho \mid i}) : i < \omega\}$ is consistent.
   (b) For all $\eta \in \omega^{<\omega}$, $\{\varphi(x; a_{\eta^\frown \langle j \rangle}) : j < \omega\}$ is k-inconsistent.

(2) We say $\varphi(x; y)$ has the *tree property of the first kind* (TP_1) if there is a tree of tuples $(a_\eta)_{\eta \in \omega^{<\omega}}$ satisfying the following conditions:
   (a) For all $\rho \in \omega^\omega$, $\{\varphi(x; a_{\rho \mid i}) : i < \omega\}$ is consistent.
   (b) For all $\eta, \nu \in \omega^{<\omega}$, if $\eta \perp \nu$, then $\{\varphi(x; a_\eta), \varphi(x; a_\nu)\}$ is inconsistent.

(3) We say $\varphi(x; y)$ has the *tree property of the second kind* (TP_2) if there is $k < \omega$ and an array $(a_{i,j})_{i,j<\omega}$ satisfying the following conditions:
   (a) For all $f : \omega \to \omega$, $\{\varphi(x; a_{i, f(i)}) : i < \omega\}$ is consistent.
   (b) For all $i < \omega$, $\{\varphi(x; a_{i,j}) : j < \omega\}$ is k-inconsistent.

(4) We say T is NTP (NTP_1, NTP_2) if no formula has TP (TP_1, TP_2, respectively) modulo T. An NTP theory is also called a *simple theory*.

The next property was introduced by Džamonja and Shelah [2004].

Definition 2.2. [Džamonja and Shelah 2004, Definition 2.2] We say $\varphi(x; y)$ has the 1-*strong order property* (SOP_1) if there is a tree of tuples $(a_\eta)_{\eta \in 2^{<\omega}}$ satisfying the following conditions:

- For all $\rho \in 2^\omega$, the set of formulas $\{\varphi(x; a_{\rho \mid i}) : i < \omega\}$ is consistent.
- For all $\nu, \eta \in 2^{<\omega}$, if $\nu^\frown\langle 0 \rangle \trianglelefteq \eta$ then $\{\varphi(x; a_\eta), \varphi(x; a_{\nu^\frown\langle 1 \rangle})\}$ is inconsistent.

T is $NSOP_1$ if no formula has SOP_1 modulo T.

Our last property was introduced much more recently by Ahn and Kim [2024].

Definition 2.3. [Ahn and Kim 2024, Definition 4.1] We say $\varphi(x; y)$ has the *antichain tree property* (ATP) if there is a tree of tuples $(a_\eta)_{\eta \in 2^{<\omega}}$ satisfying the following conditions:

(1) If $X \subseteq 2^{<\omega}$ is an antichain, then $\{\varphi(x; a_\eta) : \eta \in X\}$ is consistent.

(2) If $\eta \trianglelefteq \nu \in 2^{<\omega}$, then $\{\varphi(x; a_\eta), \varphi(x; a_\nu)\}$ is inconsistent.

T is NATP if no formula has ATP modulo T.

---

[2] TP, TP_1, and TP_2 were first introduced under the rather cumbersome labels $\kappa_{\mathrm{cdt}}(T) = \infty$, $\kappa_{\mathrm{sct}}(T) = \infty$, and $\kappa_{\mathrm{inp}}(T) = \infty$, respectively.

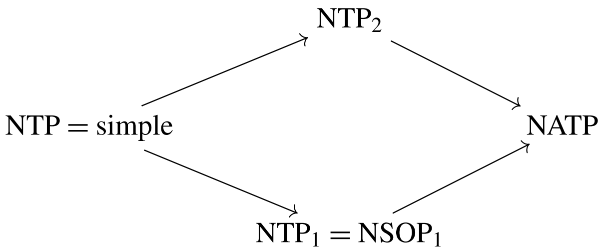


Figure 2. A summary of known implications.

Fact 2.4. Here is a summary of the known implications, which are depicted in Figure 2:

(1) The simple theories are the intersection of the NTP$_1$ and NTP$_2$ theories, i.e., T is simple if no formula has TP$_1$ or TP$_2$ modulo T [Shelah 1990, Theorem III.7.11].

(2) A theory T is NSOP$_1$ if and only if it is NTP$_1$ [Mutchnik 2022b, Theorem 1.6].[3]

(3) The NATP theories (properly) contain both the NTP$_1$ and NTP$_2$ theories [Ahn and Kim 2024, Propositions 4.4 and 4.6].

2B. *Forking and dividing.* In this section, we introduce a number of refinements of Shelah's notions of forking and dividing, based on the idea that, when a formula divides, it can be useful to study which indiscernible sequences it divides along.

Definition 2.5. Suppose $\varphi(x; b)$ is a formula, C is a set, and $I = (b_i)_{i<\omega}$ is a C-indiscernible sequence in $\text{tp}(b/C)$ (meaning that b_i realizes $\text{tp}(b/C)$ for all $i < \omega$). We say that φ *divides along* I (over C) if $\{\varphi(x; b_i) : i < \omega\}$ is inconsistent.

Definition 2.6. Suppose $\varphi(x; b)$ is a formula and C is a set:

(1) We say $\varphi(x; b)$ *divides* over C if it divides along some C-indiscernible sequence in $\text{tp}(b/C)$.

(2) We say $\varphi(x; b)$ *forks* over C if there are formulas $(\psi_i(x; c_i))_{i<n}$ with $n < \omega$ such that $\varphi(x; b) \models \bigvee_{i<n} \psi_i(x; c_i)$ and each $\psi_i(x; c_i)$ divides over C.

(3) The notation $a \mathop{\smile\hspace{-0.9em}\mid}\limits^{d}_{C} b$ means that $\text{tp}(a/Cb)$ contains no formula that divides over C and, similarly, $a \mathop{\smile\hspace{-0.9em}\mid}\limits^{f}_{A} b$ means that $\text{tp}(a/Cb)$ contains no formula that forks over A.

[3]The theorem as stated in [Mutchnik 2022b] says that every NSOP$_2$ theory is NSOP$_1$. Prior to the appearance of this result, it was well-known that NSOP$_1$ implies NSOP$_2$ and NSOP$_2$ is equivalent to NTP$_1$, see, e.g., [Kim and Kim 2011].

We will be primarily concerned with extremely "generic" sequences, i.e., Morley sequences for global invariant types.

Definition 2.7. A *global partial type* $\pi(x)$ is a consistent set of formulas over $\mathbb{M}$. A *global type* is a global partial type which is complete. For a set C, we say the global partial type $\pi(x)$ is C*-invariant* if, for all formulas $\varphi(x; y)$, if $b \equiv_C b'$, then $\varphi(x; b) \in \pi$ if and only if $\varphi(x; b') \in \pi$.

An important class of examples of global C-invariant types are the types that are finitely satisfiable in C. In any theory T, if $M \models T$, every type over M has a global extension which is finitely satisfiable in M (and therefore M-invariant). See Remark 2.9 below.

Over a general set C, there may be no global C-invariant types whatsoever. For this reason, when we want to work with invariant types (such as in the definition of Kim-dividing below), we usually work over a model.

Definition 2.8. Suppose $M \models T$:

(1) We write $a \mathop{\smile\!\!\!\!\!\!|}^{i}_{M} b$ if $\mathrm{tp}(a/Mb)$ extends to a global M-invariant type.

(2) We write $a \mathop{\smile\!\!\!\!\!\!|}^{u}_{M} b$ if $\mathrm{tp}(a/Mb)$ extends to a global type finitely satisfiable in M.

Remark 2.9. The u superscript comes from "ultrafilter", since global M-finitely satisfiable types all arise from the following construction: if $p(x) \in S_x(M)$, then $\{\varphi(M) : \varphi(x) \in p\} \subseteq \mathcal{P}(M^x)$ generates a filter on M^x. If $\mathcal{D}$ is an ultrafilter on M^x extending this filter, then

$$\mathrm{Av}(\mathcal{D}, \mathbb{M}) = \{\varphi(x) \in L(\mathbb{M}) : \varphi(M) \in \mathcal{D}\}$$

is a global type extending p which is finitely satisfiable in M. We write $\mathrm{Av}(\mathcal{D}, B)$ for $\mathrm{Av}(\mathcal{D}, \mathbb{M})$ restricted to parameters coming from B.

Definition 2.10. If q is a global C-invariant type, then a *Morley sequence* over C for q is a sequence $(a_i)_{i<\omega}$ such that $a_i \models q|_{Ca_{<i}}$ for all $i < \omega$.

Fact 2.11. By invariance, every Morley sequence over C for q is C-indiscernible. Furthermore, for a fixed global C-invariant type q extending $\mathrm{tp}(b/C)$, if $\varphi(x; b)$ divides along some Morley sequence over C for q, then it divides along every Morley sequence over C for q.

Definition 2.12. Suppose $\varphi(x; b)$ is a formula and M is a model:

(1) We say $\varphi(x; b)$ *Kim-divides* over M if it divides along a Morley sequence over M for some global M-invariant type extending $\mathrm{tp}(b/M)$.

(2) We say $\varphi(x; b)$ *Kim-forks* over M if there are formulas $(\psi_i(x; c_i))_{i<n}$ with $n < \omega$ such that $\varphi(x; b) \models \bigvee_{i<n} \psi_i(x; c_i)$ and each $\psi_i(x; c_i)$ Kim-divides over M.

(3) The notation $a \mathop{\smile\hskip-0.9em^{|}\,}^{Kd}_M b$ means that $\operatorname{tp}(a/Mb)$ contains no formula that Kim-divides over M and, similarly, $a \mathop{\smile\hskip-0.9em^{|}\,}^{K}_M b$ means that $\operatorname{tp}(a/Mb)$ contains no formula that Kim-forks over M.

Kim-independence was introduced by Kaplan and Ramsey [2020], in the context of NSOP_1 theories. They showed that if T is NSOP_1, then Kim-forking is equivalent to Kim-dividing, and $\mathop{\smile\hskip-0.9em^{|}\,}^{K}$ satisfies many of the good properties of $\mathop{\smile\hskip-0.9em^{|}\,}^{f}$ in simple theories. The definition of Kim-dividing was inspired by an earlier suggestion of Kim [2009] for studying independence in NTP_1 theories.

Remark 2.13. In a general theory, Kim-dividing as we have defined it is not always preserved under adding dummy parameters. That is, suppose $\varphi(x; y)$ is a formula, and write $\widehat{\varphi}(x; y, z)$ for the same formula consider in a larger variable context by appending unused variables z. It is possible that there are tuples b and c such that $\varphi(x; b)$ Kim-divides over M but $\widehat{\varphi}(x; b, c)$ does not Kim-divide over M. The reason is that $\mathop{\smile\hskip-0.9em^{|}\,}^{i}$ does not satisfy left-extension in general. More explicitly, if $q(y)$ is a global M-invariant type extending $\operatorname{tp}(b/M)$ (and witnessing the Kim-dividing of $\varphi(x; b)$), there may be no global M-invariant type $r(y, z)$ extending both $q(y)$ and $\operatorname{tp}(bc/M)$. Hanson [2023, Appendix C] has produced an explicit example of this behavior.

As a result, we have to be careful about parameters when working with Kim-dividing in arbitrary theories. For example, if $\varphi(x; b)$ Kim-forks, we cannot assume in general that the witnessing Kim-dividing formulas $(\psi_i(x; c_i))_{i<n}$ have the same tuple of parameters. This will cause us some trouble in Section 2C below.

All this suggests to us that our definition of Kim-dividing may not be the "right" one outside of the context of NSOP_1 theories. If T is NSOP_1, then a formula Kim-divides over a model M if and only if it Kim-divides along a coheir sequence over M (a Morley sequence for a global type finitely satisfiable in M). And if Kim-dividing were defined as dividing along a coheir sequence, then the issue with dummy parameters would not arise, since $\mathop{\smile\hskip-0.9em^{|}\,}^{u}$ always satisfies left-extension. However, focusing only on coheir sequences seems potentially too restrictive, and the definition of Kim-dividing in terms of invariant Morley sequences is well-established, so we retain it for this paper.

The diagram below depicts the implications between the notions of independence defined in this section:

$$\begin{array}{ccccccc}
a \mathop{\smile\hskip-0.9em^{|}\,}^{u}_M b & \longrightarrow & a \mathop{\smile\hskip-0.9em^{|}\,}^{i}_M b & \longrightarrow & a \mathop{\smile\hskip-0.9em^{|}\,}^{f}_M b & \longrightarrow & a \mathop{\smile\hskip-0.9em^{|}\,}^{K}_M b \\
 & & & & \downarrow & & \downarrow \\
 & & & & a \mathop{\smile\hskip-0.9em^{|}\,}^{d}_M b & \longrightarrow & a \mathop{\smile\hskip-0.9em^{|}\,}^{Kd}_M b
\end{array}$$

Fact 2.14 [Chernikov and Kaplan 2012; Adler 2014]. In NTP_2 theories, a formula $\varphi(x;b)$ divides over a model M if and only if it Kim-divides over M. Further, forking and dividing coincide over models. So when T is NTP_2, $\mathop{\smile\!\!\!\!\!\!|}^f_M = \mathop{\smile\!\!\!\!\!\!|}^d_M = \mathop{\smile\!\!\!\!\!\!|}^K_M = \mathop{\smile\!\!\!\!\!\!|}^{Kd}_M$.

It is a fact that simple theories are characterized by symmetry of $\mathop{\smile\!\!\!\!\!\!|}^f$ [Kim 2001, Theorem 2.4]. So in a simple theory, if p is a global M-invariant type and a realizes $p|_{MB}$, then $B \mathop{\smile\!\!\!\!\!\!|}^f_M a$ (since $a \mathop{\smile\!\!\!\!\!\!|}^i_M B$ implies $a \mathop{\smile\!\!\!\!\!\!|}^f_M B$ and $\mathop{\smile\!\!\!\!\!\!|}^f$ is symmetric). Outside of the simple context, it can be useful to consider invariant types which always satisfy this instance of symmetry. These "strict" invariant types play an important role in Chernikov and Kaplan's analysis [2012] of forking in NTP_2 theories.

Similarly, NSOP_1 theories are characterized by symmetry of $\mathop{\smile\!\!\!\!\!\!|}^K$, so it makes sense in our context to consider "Kim-strict" invariant types, which are the analogue of strict invariant types for Kim-forking.

Definition 2.15. Suppose $p \in S(\mathbb{M})$ is a global M-invariant type:

(1) We say p is a *strict invariant type* over M when, for any set B, if $a \models p|_{MB}$, then $B \mathop{\smile\!\!\!\!\!\!|}^f_M a$.

(2) We say p is a *Kim-strict invariant type* over M when, for any set B, if $a \models p|_{MB}$, then $B \mathop{\smile\!\!\!\!\!\!|}^K_M a$.

(3) A formula $\varphi(x;b)$ *strictly divides* over M if it divides along a Morley sequence for some global strictly M-invariant type extending $\mathrm{tp}(b/M)$.

(4) A formula $\varphi(x;b)$ *Kim-strictly divides* over M if it divides along a Morley sequence for some global Kim-strictly M-invariant type extending $\mathrm{tp}(b/M)$.

Finally, for each of the variants of dividing defined above, we can also consider changing the quantifier from dividing along *some* to dividing along *every* indiscernible sequence of the appropriate kind.

Definition 2.16. We say a formula $\varphi(x;b)$ *universally Kim-divides* over M if it divides along Morley sequences for *every* global M-invariant type extending $\mathrm{tp}(b/M)$.[4] Similarly, we say $\varphi(x;b)$ *universally strictly divides* over M if it divides along Morley sequences for *every* global strict M-invariant type extending $\mathrm{tp}(b/M)$, and we say $\varphi(x;b)$ *universally Kim-strictly divides* over M if it divides along Morley sequences for *every* global Kim-strict M-invariant type extending $\mathrm{tp}(b/M)$.

Remark 2.17. For completeness, we could say a formula $\varphi(x;b)$ *universally divides* over C if it divides along *every* C-indiscernible sequence in $\mathrm{tp}(b/C)$. Note,

[4] Universal Kim-dividing is called "strong Kim-dividing" in [Kaplan et al. 2019] and "Conant-dividing" in [Mutchnik 2022a].

however, that since the constant sequence with $b_i = b$ for all i is C-indiscernible, a universally dividing formula is inconsistent.

2C. *The broom lemma.* It is clear that universal Kim-dividing implies Kim-dividing, since every type over a model M extends to a global M-invariant type (see Remark 2.9). However, it is not so clear that universal (Kim-)strict dividing implies (Kim-)strict dividing.

Chernikov and Kaplan [2012] proved that in an NTP_2 theory, every type over a model M extends to a global strictly M-invariant type, using a device they called the broom lemma. It turns out that their argument applies to all theories, if we replace strict invariance with Kim-strict invariance.

A key step in the Chernikov–Kaplan argument is that forking implies quasi-dividing in the sense of the following definition.

Definition 2.18. A formula $\varphi(x; b)$ *quasi-divides* over M if the conjunction of finitely many conjugates of $\varphi(x; b)$ over M is inconsistent. That is, if there exist $(b_i)_{i<k}$ with $k < \omega$ and $b_i \equiv_M b$ for all $i < k$ such that $\bigwedge_{i<k} \varphi(x; b_i)$ is inconsistent.

Remark 2.19. We could say that $\varphi(x; b)$ *quasi-forks* over M if there are formulas $(\psi_i(x; c_i))_{i<n}$ with $n < \omega$ such that $\varphi(x; b) \models \bigvee_{i<n} \psi_i(x; c_i)$ and each $\psi_i(x; c_i)$ quasi-divides over M. It is worth noting that $a \mathop{\smile\!\!\!\!\!\!\;|}^{i}_{M} b$ if and only if $\mathrm{tp}(a/Mb)$ contains no formula which quasi-forks. But we will not make use of this fact.

The original broom lemma argument from [Chernikov and Kaplan 2012] does not appear to generalize directly to our context. But Adler [2014] used a variant of the broom lemma, which he called the vacuum cleaner lemma, to give a simplified proof of some of the Chernikov–Kaplan results on NTP_2 theories. Adler's proof [2014, Lemma 3] goes through verbatim to prove the following result, in the context of an arbitrary theory T.

Lemma 2.20 (vacuum cleaner for Kim-dividing). *Let $\pi(x)$ be an M-invariant partial type and suppose*

$$\pi(x) \models \psi(x; b) \vee \bigvee_{i<n} \varphi_i(x; c),$$

where $b \mathop{\smile\!\!\!\!\!\!\;|}^{i}_{M} c$ and each $\varphi_i(x; c)$ Kim-divides over M. Then $\pi(x) \models \psi(x; b)$.[5]

Corollary 2.21. *Suppose $\theta(x; b) \models \bigvee_{i<n} \varphi_i(x; c)$, where each $\varphi_i(x; c)$ Kim-divides over M. Then $\theta(x; b)$ quasi-divides over M.*

Proof. Let $\pi(x) = \{\theta(x; b') : b' \equiv_M b\}$ and let ψ be $\perp$. By Lemma 2.20, $\pi(x)$ is inconsistent, so, by compactness, $\theta(x; b)$ quasi-divides. □

[5] A similar modified broom lemma played a key role in Mutchnik's proof [2022b] of the equivalence of NSOP_1 and NSOP_2.

Corollary 2.21 seems to say that Kim-forking formulas quasi-divide. But, as noted in Remark 2.13 above, we cannot assume in general that in the finite disjunction $\bigvee_{i<n} \varphi_i(x; c_i)$ witnessing Kim-forking, all of the Kim-dividing formulas have the same tuple of parameters c. Unfortunately, this assumption seems crucial in Adler's proof of the vacuum cleaner lemma. As in Remark 2.13, this would not be an issue if we defined Kim-dividing in terms of dividing along coheir sequences.

Nevertheless, it is true in general that Kim-forking formulas quasi-divide. We present an alternative proof, based on an idea due to Hanson.

Lemma 2.22. *Let $\varphi(x; b)$ be a formula. Suppose that the conjunction of finitely many conjugates of $\varphi(x; b)$ over M entails a formula which quasi-divides over M. Then $\varphi(x; b)$ quasi-divides over M.*

Proof. By hypothesis, there exist $(b_i)_{i<k}$ with $b_i \equiv_M b$ for all $i < k$ such that $\bigwedge_{i<k} \varphi(x; b_i) \models \psi(x; c)$, and $\psi(x; c)$ quasi-divides over M. Then there exist $(c_j)_{j<n}$ with $c_j \equiv_M c$ for all $j < n$ such that $\bigwedge_{j<n} \psi(x; c_j)$ is inconsistent.

For each $j < n$, pick $(b_{i,j})_{i<k}$ such that $b_{0,j} \cdots b_{(k-1),j} c_j \equiv_M b_0 \cdots b_{k-1} c$. Then

$$\bigwedge_{j<n} \bigwedge_{i<k} \varphi(x; b_{i,j}) \models \bigwedge_{j<n} \psi(x; c_j).$$

For all $i < k$ and $j < n$, $b_{i,j} \equiv_M b_i \equiv_M b$, so this is a finite conjunction of conjugates of $\varphi(x; b)$ over M which is inconsistent. □

Lemma 2.23. *Suppose $\varphi(x; b)$ Kim-divides over M. Then for any $(b_i)_{i<\ell}$ such that $b_i \equiv_M b$ for all $i < \ell$, $\bigvee_{i<\ell} \varphi(x; b_i)$ quasi-divides over M.*

Proof. Write $\Phi(x; \bar{b})$ for the formula $\bigvee_{i<\ell} \varphi(x; b_i)$. Our goal is to show that $\Phi(x; \bar{b})$ quasi-divides. Let $q(y)$ be a global M-invariant type extending $\mathrm{tp}(b/M)$ and witnessing that $\varphi(x; b)$ Kim-divides over M. Let k be such that, if $(b'_i)_{i<\omega}$ is a Morley sequence for q over M, $\{\varphi(x; b'_i) : i < k\}$ is inconsistent.

Write $\ell_*^{\leq m}$ for the set of functions $\eta\colon n \to \ell$ with $0 < n \leq m$, that is, $\ell_*^{\leq m} = \ell^{\leq m} \setminus \{\langle\,\rangle\}$. We will prove by induction that for all $m \leq k$, we can find $(b_\eta)_{\eta \in \ell_*^{\leq m}}$ such that:

(1) For each $\rho \in \ell^m$, $(b_\rho, b_{\rho|_{m-1}}, \ldots, b_{\rho|_1})$ begins a Morley sequence in q over M.

(2) For each $\eta \in \ell^{<m}$, writing $\bar{b}'_\eta$ for the tuple $(b_{\eta^\frown\langle i\rangle})_{i<\ell}$, we have $\bar{b}'_\eta \equiv_M \bar{b}$.

In the base case, when $m = 0$, $\ell_*^{\leq m}$ is empty, and the conditions are satisfied vacuously.

For the inductive step, suppose we are given $F_0 = (b_\eta)_{\eta \in \ell_*^{\leq m}}$ satisfying the conditions, with $m < k$. Let b''_0 realize $q|_{MF_0}$. By condition (1), we now have that for each $\rho \in \ell^m$, $(b_\rho, b_{\rho|_{m-1}}, \ldots, b_{\rho|_1}, b''_0)$ begins a Morley sequence in q over M.

Since $b''_0 \equiv_M b_0$, we can pick $(b''_i)_{0<i<\ell}$ so that $(b''_i)_{i<\ell} \equiv_M \bar{b}$. Now, for each $0 < i < \ell$, pick F_i so that $F_i b''_i \equiv_M F_0 b''_0$. Reindex so that we have a forest indexed by $\ell_*^{\leq(m+1)}$, with $(b_i'')_{i<\ell}$ as the "bottom layer" $\bar{b}'_{\langle\,\rangle}$. This completes the inductive construction.

Now we have $(b_\eta)_{\eta\in\ell_*^{\leq k}}$ satisfying (1) and (2). Observe that

$$\bigwedge_{\eta\in\ell^{<k}}\bigvee_{i<\ell}\varphi(x;b_{\eta\frown\langle i\rangle})\models\bigvee_{\rho\in\ell^k}\bigwedge_{1\leq i\leq k}\varphi(x;b_{\rho|_i}).$$

By (1), for each $\rho\in\ell^k$, $\bigwedge_{1\leq i\leq k}\varphi(x;b_{\rho|_i})$ is inconsistent. Thus the left-hand side, which is $\bigwedge_{\eta\in\ell^{<k}}\Phi(x;\bar{b}'_\eta)$, is inconsistent. By (2), this shows that $\Phi(x;\bar{b})$ quasi-divides over M. □

Lemma 2.24. *Suppose $(\varphi_i(x;b_i))_{i<n}$ are formulas, each of which Kim-divides over M. For each $i<n$, let $\theta_i(x;c_i)$ be a disjunction of finitely many conjugates of $\varphi_i(x;b_i)$. Then $\bigvee_{i<n}\theta_i(x;c_i)$ quasi-divides over M.*

Proof. By induction on n. When $n=0$, the disjunction is $\perp$, which quasi-divides over M. For the inductive step, we consider $\bigvee_{i<n+1}\theta_i(x;c_i)$. Now $\theta_n(x;c_n)$ is a disjunction of finitely many conjugates of $\varphi_n(x;b_n)$. By Lemma 2.23, $\theta_n(x;c_n)$ quasi-divides over M, so there are $(c_{nj})_{j<k}$ with $c_{nj}\equiv_M c_n$ for all $j<k$ such that $\bigwedge_{j<k}\theta_n(x;c_{nj})$ is inconsistent.

For each $j<k$, pick $(c_{ij})_{i<n}$ such that $c_{0j}\cdots c_{nj}\equiv_M c_0\cdots c_n$. Consider the conjunction

$$\bigwedge_{j<k}\bigvee_{i<n+1}\theta_i(x;c_{ij}).$$

Whenever this formula is true, there must be some $j<k$ such that some disjunct $\theta_i(x;c_{ij})$ with $i\neq n$ is true, since $\bigwedge_{j<k}\theta_n(x;c_{nj})$ is inconsistent. Thus

$$\bigwedge_{j<k}\bigvee_{i<n+1}\theta_i(x;c_{ij})\models\bigvee_{i<n}\bigvee_{j<k}\theta_i(x;c_{ij}).$$

Since each formula $\bigvee_{j<k}\theta_i(x;c_{ij})$ is a disjunction of finitely many conjugates of $\varphi_i(x;b_i)$, by induction $\bigvee_{i<n}\bigvee_{j<k}\theta_i(x;c_{ij})$ quasi-divides over M. By Lemma 2.22, $\bigvee_{i<n+1}\theta_i(x;c_{ij})$ quasi-divides over M. □

Corollary 2.25. *Every formula which Kim-forks over M quasi-divides over M.*

Proof. Suppose $\varphi(x;b)$ Kim-forks over M. Then $\varphi(x;b)\models\bigvee_{i<n}\psi_i(x;c_i)$ such that each $\psi_i(x;c_i)$ Kim-divides over M. By Lemma 2.24 (taking each θ_i to be $\psi_i(x;c_i)$), $\bigvee_{i<n}\psi(x;c_i)$ quasi-divides over M, and hence so does $\varphi(x;b)$ by Lemma 2.22. □

Theorem 2.26. *Every type over $M\models T$ has a Kim-strict M-invariant global extension.*

Proof. Given $p(x) = \text{tp}(a/M)$, consider the following collection of formulas:

$$p(x) \cup \{\psi(x; c) \leftrightarrow \psi(x; c') : c \equiv_M c'\} \cup \{\neg\varphi(x; b) : \varphi(a; y) \text{ Kim-forks over } M\}.$$

We must show that this is a consistent partial type. Suppose not; then, by compactness,

$$p(x) \cup \{\psi(x; c) \leftrightarrow \psi(x; c') : c \equiv_M c'\} \models \varphi(x; b),$$

for some formula $\varphi(x; y)$ such that $\varphi(a; y)$ Kim-forks over M.

By Corollary 2.25, there are $(a_i)_{i<m}$ such that $a_i \equiv_M a$ for all $i < m$ and $\{\varphi(a_i, y) : i < m\}$ is inconsistent. Let $r(x_0, \ldots, x_{m-1})$ be a global M-invariant type extending $\text{tp}(a_0, \ldots, a_{m-1}/M)$, and for $j < m$, let $r(x_j)$ be the restriction of r to formulas with free variables from x_j. Then each $r(x_j)$ is a global M-invariant type extending $p(x_j)$, so $r(x_j) \models \varphi(x_j, b)$. Thus,

$$r(x_0, \ldots, x_{m-1}) \models \bigwedge_{j<m} \varphi(x_j; b),$$

and therefore $\exists y \bigwedge_{j<m} \varphi(x_j, y) \in r$. This contradicts the fact that r extends $\text{tp}(a_0, \ldots, a_{m-1}/M)$. □

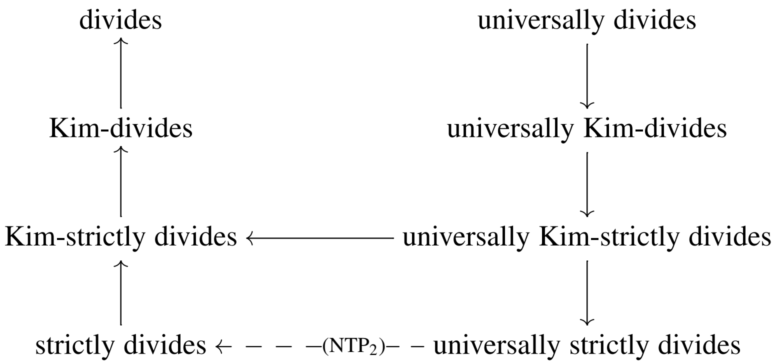
Corollary 2.27. *If $\varphi(x; b)$ universally Kim-strictly divides over M, then it Kim-strictly divides over M.*

Proof. By Theorem 2.26, $\text{tp}(b/M)$ has a Kim-strict M-invariant global extension $q(y)$. Since $\varphi(x; b)$ universally Kim-strictly divides over M, it divides along Morley sequences for q. Thus, it Kim-strictly divides over M. □

Note that the only properties of Kim-forking used in the proof of Theorem 2.26 are (a) that the Kim-forking formulas form an ideal (i.e., they are closed under finite disjunctions), and (b) that every Kim-forking formula quasi-divides. In unpublished work, Hanson has shown that there is a largest M-invariant ideal which contains only quasi-dividing formulas, called the "fracturing" ideal. The proof of Theorem 2.26 works just as well to show that $\text{tp}(a/M)$ extends to a global M-invariant extension containing no formula $\varphi(x; b)$ such that $\varphi(a; y)$ fractures.

Remark 2.28. Chernikov and Kaplan [2012, Subsection 5.1] present an example, due to Martin Ziegler, of a theory T in which there is a model $M \models T$ and a type over M with no global extension that is strict invariant over M. This shows that, in general, Theorem 2.26 cannot be improved to establish the existence of global strict invariant types over models in arbitrary theories.

We conclude this section with a diagram showing the implications between the various notions of dividing (over models) introduced in Section 2B. All implications hold in an arbitrary theory, except for the implication from universally strictly divides to strictly divides, which requires NTP_2:



3. A diversity of Kim's lemmas

In this section, we survey the characterizations of simplicity, NSOP_1, and NTP_2 by Kim's Lemmas, and we introduce our new Kim's Lemma. We begin with the original Kim's Lemma in the context of simple theories.

Theorem 3.1 [Kim 1998, Proposition 2.1; 2001, Theorem 2.4]. *The following are equivalent*:

(1) *T is simple.*

(2) *For all sets C, if a formula $\varphi(x; b)$ divides over C, then it divides along every $\mathop{\smile\!\!\!\!\!\!\mid}^{f}$-Morley sequence over C.*

In this paper, we are primarily interested in Morley sequences for global invariant types over models (rather than $\mathop{\smile\!\!\!\!\!\!\mid}^{f}$-Morley sequences over arbitrary sets), so we are led to consider the following variant of (2):

(3) For all models M, if a formula $\varphi(x; b)$ divides over M, then it universally Kim-divides over M.

Note that (3) is a weakening of (2), since it restricts to the special case of models, and since every Morley sequence for a global M-invariant type is a $\mathop{\smile\!\!\!\!\!\!\mid}^{f}$-Morley sequence over M. But (3) is still strong enough to characterize simplicity.

The equivalence of (1) and (3) has not (to our knowledge) appeared explicitly in the literature, but it does follow directly from facts in the literature. We have already observed that (1) implies (2) and (2) implies (3). Conversely, (3) implies, in particular, that Kim-dividing implies universal Kim-dividing, so T is NSOP_1 (by the Kim's Lemma for NSOP_1 theories, Theorem 3.3 below). Thus T is an NSOP_1 theory in which dividing and Kim-dividing coincide over models, so T is simple by [Kaplan and Ramsey 2020, Proposition 8.4].

For the reader's convenience, and to give an indication of the typical flavor of arguments relating variants of Kim's Lemma to combinatorial configurations like the tree property, we will also give a self-contained proof of the equivalence of (1) and (3).

Proof. (1)$\Longrightarrow$(3). Suppose (3) fails, so there is a model $M \models T$, a formula $\varphi(x; b)$ that divides over M, and a global M-invariant type $q \supseteq \mathrm{tp}(b/M)$ such that $\varphi(x; b)$ does not divide along Morley sequences over M for q. Let $(b_i)_{i<\omega}$ be an M-indiscernible sequence in $\mathrm{tp}(b/M)$ such that $\{\varphi(x; b_i) : i < \omega\}$ is inconsistent (and hence k-inconsistent for some k). By induction, we will build for each $n < \omega$, a tree $(c_\eta)_{\eta\in\omega^{\leq n}}$ satisfying the following:

- For all $\eta \in \omega^{<n}$, $(c_{\eta\frown\langle i\rangle})_{i<\omega} \equiv_M (b_i)_{i<\omega}$.
- For all $\nu \in \omega^n$, $(c_\nu, c_{\nu\,|\,(n-1)}, \ldots, c_{\nu\,|\,0})$ begins a Morley sequence in q over M.

For $n = 0$, we define $c_{\langle\,\rangle} = b$. The conditions are trivially satisfied.

For the inductive step, we are given a tree $(c_{\eta,0})_{\eta\in\omega^{\leq n}}$. Since $c_{\langle\,\rangle,0} \models q|_M$, we have $c_{\langle\,\rangle,0} \equiv_M b$, and there is a sequence $(c_{\langle\,\rangle,i})_{i<\omega}$ beginning with $c_{\langle\,\rangle,0}$ such that $(c_{\langle\,\rangle,i})_{i<\omega} \equiv_M (b_i)_{i<\omega}$. For each i, $c_{\langle\,\rangle,i} \equiv_M c_{\langle\,\rangle,0}$, so we can choose a tree $(c_{\eta,i})_{\eta\in\omega^{\leq n}}$ with root $c_{\langle\,\rangle,i}$ such that $(c_{\eta,0})_{\eta\in\omega^{\leq n}} \equiv_M (c_{\eta,i})_{\eta\in\omega^{\leq n}}$. Let $c_{\langle\,\rangle}$ be a realization of $q|_{M\{c_{\eta,i}:\eta\in\omega^{\leq n},i<\omega\}}$. Then we reindex to define a tree $(c_\eta)_{\eta\in\omega^{\leq n+1}}$ by setting $c_{\langle i\rangle\frown\eta} = c_{\eta,i}$ for all $i < \omega$ and $\eta \in \omega^{\leq n}$.

Note that, for each n, the tree $(c_\eta)_{\eta\in\omega^{\leq n}}$ that we constructed has the following properties. First, for each $\eta \in \omega^{<n}$, $\{\varphi(x; c_{\eta\frown\langle i\rangle}) : i < \omega\}$ is k-inconsistent, by the first bullet point above. Secondly, for all $\nu \in \omega^n$, $\{\varphi(x; c_{\nu\,|\,\ell}) : \ell \leq n\}$ is consistent, by the second bullet point and our assumption on q. By compactness, $\varphi(x; y)$ has TP, and T is not simple.

(3)$\Longrightarrow$(1) Suppose T has TP witnessed by $\varphi(x; y)$, $k < \omega$, and $(a_\eta)_{\eta\in\omega^{<\omega}}$. Fix a Skolemization T^{Sk} of T. The same data shows that $\varphi(x; y)$ has TP modulo T^{Sk}.

By compactness, we can obtain a tree $(a_\eta)_{\eta\in\kappa^{<\omega}}$, where $\kappa > 2^{|T|}$, and which satisfies the obvious extensions of the defining conditions of the tree property.

We build an array $(b_{i,j})_{i,j<\omega}$ and $\rho \in \kappa^\omega$ with the following properties (in T^{Sk}):

- $b_{i,0} = a_{\rho\,|\,(i+1)}$ for all $i < \omega$ (and therefore $\{\varphi(x; b_{i,0}) : i < \omega\}$ is consistent).
- For all $i < \omega$, $\{\varphi(x; b_{i,j}) : j < \omega\}$ is k-inconsistent.
- For all $i < \omega$, $(b_{i,j})_{j<\omega}$ is indiscernible over $(b_{\ell,0})_{\ell<i}$.

We proceed by recursion on i. Given $\rho\,|\,n$ and $(b_{i,j})_{i<n,j<\omega}$, let $\eta = \rho\,|\,n$, and consider the sequence $(a_{\eta\frown\langle\alpha\rangle})_{\alpha<\kappa}$. By the conditions on κ, we can find a subsequence $I = (a_{\eta\frown\langle\alpha_j\rangle})_{j<\omega}$ such that each $a_{\eta\frown\langle\alpha_j\rangle}$ satisfies the same complete type $p(y)$ over $(b_{i,0})_{i<n}$. Let $(b_{n,j})_{j<\omega}$ be a sequence which is indiscernible and locally based on I over $(b_{i,0})_{i<n}$ (i.e., realizes the Ehrenfeucht–Mostowski type

of I over $(b_{i,0})_{i<n}$). It follows that each $b_{n,j}$ satisfies $p(y)$, so we can assume that $b_{n,0} = a_{\eta\frown\langle\alpha_0\rangle}$ and let $\rho(n) = \alpha_0$. It also follows that $\{\varphi(x; b_{n,j}) : j < \omega\}$ is k-inconsistent. This completes the construction.

For each $i < \omega$, let $\bar{b}_i = (b_{i,j})_{j<\omega}$, and let $J = (\bar{b}_i)_{i<\omega}$. Let $J' = (\bar{b}'_i)_{i<\omega+\omega}$ be a sequence which is indiscernible and locally based on J (over $\varnothing$). Writing each $\bar{b}'_i$ as $(b'_{i,j})_{j<\omega}$, we retain consistency of $\{\varphi(x; b'_{i,0}) : i < \omega + \omega\}$, k-inconsistency of $\{\varphi(x; b'_{i,j}) : j < \omega\}$ for all $i < \omega+\omega$, and indiscernibility of $(b'_{i,j})_{j<\omega}$ over $(b'_{\ell,0})_{\ell<i}$ for all $i < \omega + \omega$.

Let M be the Skolem hull of $(b'_{i,0})_{i<\omega}$. By indiscernibility, $\text{tp}(b'_{\omega,0}/M(b'_{i,0})_{i>\omega})$ is finitely satisfiable in M and therefore extends to a global M-finitely satisfiable (and therefore M-invariant) type q. Moreover, by indiscernibility, $b'_{\omega+i,0} \models q|_{M(b'_{n,0})_{n>\omega+i}}$ for all i, which shows that for all n, $(b'_{\omega+n,0}, \dots, b'_{\omega,0})$ begins a Morley sequence for q over M. By construction, $\{\varphi(x; b'_{\omega+i,0}) : i < \omega\}$ is consistent, so $\varphi(x; b'_{\omega,0})$ does not divide along Morley sequences for q over M. However, $\varphi(x; b'_{\omega,0})$ does divide along the M-indiscernible sequence $\bar{b}'_\omega$.

Taking the reduct back to T, the restriction $q|_L$ of q to L-formulas is still finitely satisfiable in M, Morley sequences in q are also Morley sequences in $q|_L$, and the M-indiscernible sequence $\bar{b}'_\omega$ remains M-indiscernible in the reduct. Thus $\varphi(x; b'_{\omega,0})$ divides but does not universally Kim-divide with respect to T, and (3) fails. □

Example 3.2. Let T_E be the theory of an equivalence relation E with infinitely many classes, each of which is infinite. T_E is a simple theory (in fact, it is ω-stable). Let $M \models T_E$, and let b be an element of $\mathbb{M}$ in an equivalence class which is not represented in M. There are three types of M-indiscernible sequence $(b_i)_{i<\omega}$ in $\text{tp}(b/M)$: (a) constant sequences, in which $b_i = b_j$ for all $i, j < \omega$, (b) sequences contained in one equivalence class, in which $b_i \neq b_j$ but $b_i E b_j$ for all $i \neq j$, and (c) sequences that move across equivalence classes, in which $\neg b_i E b_j$ for all $i \neq j$.

The formula xEb divides along sequences of type (c), but not along sequences of type (a) or (b). Is there a general explanation for this behavior? Kim's Lemma gives the answer: the dividing formula xEb universally Kim-divides, and every Morley sequence for a global M-invariant type extending $\text{tp}(b/M)$ has type (c).

Indeed, if $q(y)$ is a global M-invariant type extending $\text{tp}(b/M)$, we will show that q cannot contain the formula yEc for any $c \in \mathbb{M}$. If cEm for some $m \in M$, then since $\neg yEm \in \text{tp}(b/M)$, $\neg yEc \in q$. And if the equivalence class of c is not represented in M, then letting c' be another element inequivalent to c whose equivalence class is not represented in M, q cannot contain both yEc and yEc', but $\text{tp}(c/M) = \text{tp}(c'/M)$, so by invariance q does not contain yEc. It follows that a Morley sequence for q has type (c).

Next, we turn to the Kim's Lemma characterization of NSOP_1 theories.

Theorem 3.3 [Kaplan and Ramsey 2020, Theorem 3.16]. *The following are equivalent*:

(1) T *is* NSOP_1.

(2) *For all models* M, *if a formula* $\varphi(x;b)$ *Kim-divides over* M, *then it universally Kim-divides over* M.

Example 3.4. T^*_{feq}, the generic theory of parametrized equivalence relations, is NSOP_1 and has TP_2. It is the complete theory of the Fraïssé limit of the Fraïssé class $\mathcal{K}_{\mathrm{feq}}$. The language has two sorts, O and P, and one ternary relation yE_xz, where the subscript x has type P and y and z have type O. A finite structure A is in $\mathcal{K}_{\mathrm{feq}}$ if for all $a \in P(A)$, E_a defines an equivalence relation on $O(A)$.

Let $M \models T^*_{\mathrm{feq}}$, let $c \in P(\mathbb{M}) \setminus P(M)$, and let $b \in O(\mathbb{M})$ such that the E_c-class of b is not represented in $O(M)$. The formula xE_cb divides over M, along any M-indiscernible sequence $(b_i, c_i)_{i<\omega}$ such that $c_i = c$ for all i and $\neg b_iE_cb_j$ for all $i \neq j$. But if $p(y,z)$ is a global M-invariant type extending $\mathrm{tp}(bc/M)$ and $I = (b_i, c_i)_{i<\omega}$ is a Morley sequence for p, then $c_i \neq c_j$ for all $i \neq j$, and xE_cb does not divide along I. Indeed, by compactness and the genericity properties of the Fraïssé limit, if $(c_i)_{i<\omega}$ is any sequence of pairwise distinct elements of $P(\mathbb{M})$, and C_i is an E_{c_i} class for each $i < \omega$, then we can find $a \in O(\mathbb{M})$ such that $a \in C_i$ for all $i \in \omega$. It follows that xE_cb does not Kim-divide, and hence does not universally Kim-divide, so the Kim's Lemma for simple theories fails in T^*_{feq}.

Now let $m \in P(M)$, and let $b' \in O(\mathbb{M})$ such that the E_m-class of b' is not represented in $O(M)$. Then the formula xE_mb' Kim-divides over M, and, as predicted by the Kim's Lemma for NSOP_1 theories, it universally Kim-divides over M. Indeed, if $p(y,z)$ is any global M-invariant type extending $\mathrm{tp}(b'm/M)$, and $I = (b_i, m_i)_{i<\omega}$ is a Morley sequence for p, then $m_i = m$ for all $i \in \omega$ and $(b_i)_{i<\omega}$ is an indiscernible sequence of type (c) for E_m, according to the terminology in Example 3.2. Thus xE_mb' divides along I.

Finally, we turn to the Kim's Lemma characterization of NTP_2 theories.

Theorem 3.5 [Chernikov and Kaplan 2012, Lemma 3.14; Chernikov 2014, Theorem 4.9]. *The following are equivalent*:

(1) T *is* NTP_2.

(2) *For all models* M, *if a formula* $\varphi(x;b)$ *divides over* M, *then it universally Kim-strictly divides over* M.

Note that the notion of Kim-strict dividing does not appear in [Chernikov and Kaplan 2012] or [Chernikov 2014]. Instead, Chernikov and Kaplan prove that (1) is equivalent to (3):

(3) For all models M, if a formula $\varphi(x;b)$ divides over M, then it universally strictly divides over M.

But since Kim-strict invariant types coincide with strict invariant types in NTP$_2$ theories (by Fact 2.14), and universal Kim-strict dividing implies universal strict dividing in arbitrary theories, it follows immediately that (1), (2), and (3) are all equivalent. We have chosen to focus on Kim-strict dividing because it behaves better outside of the NTP$_2$ context (by Theorem 2.26 and Remark 2.28).

Example 3.6. DLO, the theory of dense linear orders without endpoints, is NTP$_2$ (in fact, it is NIP) and has SOP$_1$ (in fact, it has SOP). Let $M \models$ DLO, and $b < c$ be two elements in $\mathbb{M} \setminus M$ living in the same cut in M (so there is no $m \in M$ with $b < m < c$). Now $q(y, z) = \text{tp}(bc/M)$ has three global M-invariant extensions. By quantifier elimination, each is determined by the order relations between y and z and the elements $d \in \mathbb{M}$ living in the same cut in M as b and c:

(1) Let q_1 be the global type containing $d < y < z$ for all such d. A Morley sequence $(b_i, c_i)_{i<\omega}$ for q_1 has

$$b_0 < c_0 < b_1 < c_1 < b_2 < c_2 < \cdots .$$

(2) Let q_2 be the global type containing $y < z < d$ for all such d. A Morley sequence $(b_i, c_i)_{i<\omega}$ for q_2 has

$$\cdots < b_2 < c_2 < b_1 < c_1 < b_0 < c_0.$$

(3) Let q_3 be the global type containing $y < d < z$ for all such d. A Morley sequence $(b_i, c_i)_{i<\omega}$ for q_3 has

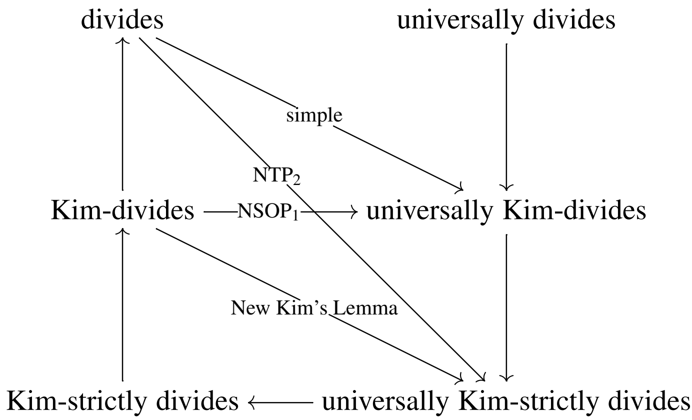
$$\cdots < b_2 < b_1 < b_0 < c_0 < c_1 < c_2 < \cdots .$$

The formula $b < x < c$ divides along Morley sequences for q_1 and q_2, but not along Morley sequences for q_3. This shows that the Kim's Lemma for NSOP$_1$ theories fails in DLO: Kim-dividing does not imply universal Kim-dividing. But the Kim's Lemma for NTP$_2$ theories explains which Morley sequences we should expect a dividing formula to divide along. Indeed, the dividing formula $b < x < c$ universally Kim-strictly divides, and we will show that q_1 and q_2 are Kim-strict, while q_3 is not.

Suppose $A \subseteq \mathbb{M}$, and suppose $b'c' \models q_i|_{MA}$ for some $i \in \{1, 2, 3\}$. If $i = 1$ or 2, then there is no $a \in A$ such that $b < a < c$, and it follows that $A \mathop{\smile\!\!\!\!\!\!|}^K_M b'c'$. So q_1 and q_2 are Kim-strict.

On the other hand, if $i = 3$, and if A contains an element a living in the same cut in M and b and c, then $b' < a < c'$. Thus $\text{tp}(A/Mb'c')$ contains the Kim-dividing formula $b' < x < c'$, and $A \not\mathop{\smile\!\!\!\!\!\!|}^K_M b'c'$. So q_3 is not Kim-strict.

We can now fill in the diagram from the end of Section 2 with the implications coming from the variants of Kim's Lemma which hold in various contexts, as well as our New Kim's Lemma:



Definition 3.7. T satisfies *New Kim's Lemma* if for all models M, if a formula $\varphi(x; b)$ Kim-divides over M, then it universally Kim-strictly divides over M.

We will give some examples and nonexamples of New Kim's Lemma in the next section. For now, let us observe a simple consequence. Variants of Kim's Lemma allow us to prove that the relevant notions of forking and dividing coincide, and the usual proof works here as well.

Proposition 3.8. *Suppose T satisfies New Kim's Lemma and $M \models T$. Then a formula $\varphi(x; b)$ Kim-forks over M if and only if it Kim-divides over M.*

Proof. Kim-dividing implies Kim-forking by definition. So suppose $\varphi(x; b)$ Kim-forks over M. Then $\varphi(x; b) \models \bigvee_{j<n} \psi_j(x; c_j)$, where each $\psi_j(x; c_j)$ Kim-divides over M.

By Theorem 2.26, let $q(y, z_1, \ldots, z_n)$ be a global Kim-strict invariant type extending $\mathrm{tp}(bc_0 \cdots c_{n-1}/M)$, and let $I = (b^i, c_0^i, \ldots, c_{n-1}^i)_{i<\omega}$ be a Morley sequence for q over M.

For all $j < n$, $I_j = (c_j^i)_{i<\omega}$ is also a Morley sequence over M for a global Kim-strict invariant type, namely the restriction of q to formulas in the single variable z_j. By New Kim's Lemma, $\psi_j(x; c_j)$ divides along I_j.

Suppose for contradiction that $\varphi(x; b)$ does not divide along $I_* = (b^i)_{i<\omega}$. Then there exists a satisfying $\{\varphi(x; b^i) : i < \omega\}$. For each $i < \omega$, since $b^i c_0^i \cdots c_{n-1}^i \equiv_M bc_0 \cdots c_{n-1}$, there exists $j < n$ such that $\models \psi_j(a; c_j^i)$. By the pigeonhole principle, there is some $j < n$ such that for infinitely many $i < \omega$, a satisfies $\psi_j(x; c_j^i)$. This contradicts the fact that $\psi_j(x; c_j)$ divides along I_j. Thus $\varphi(x; b)$ divides along I_*. Since I_* is a Morley sequence over M for the restriction of q to formulas in the single variable y, $\varphi(x; b)$ Kim-divides over M. □

4. Examples

4A. *Parametrized linear orders.* In this section, we introduce the theory DLO_p of parametrized dense linear orders without endpoints, and we show that it satisfies New Kim's Lemma. The choice of this example is motivated by the examples in Section 3: DLO_p is to DLO (Example 3.6) as T^*_{feq} (Example 3.4) is to T_E (Example 3.2).

The language L has two sorts, O and P, and one ternary relation $y <_x z$, where the subscript x has type P and y and z have type O. For an L-structure A, we write A_P and A_O for the two sorts. Let $L_<$ be the language $\{<\}$, where $<$ is a binary relation. Given $a \in A_P$, we write A_a for the $L_<$-structure $(A_O, <_a)$.

Let $\mathcal{K}$ be the class of all finite structures A such that for all $a \in A_P$, $<_a$ is a linear order on A_O. This is a special case of the parametrization construction introduced in [Chernikov and Ramsey 2016, Section 6.3], applied to the class of finite linear orders. By [loc. cit., Lemma 6.3], $\mathcal{K}$ is a Fraïssé class with disjoint amalgamation. Let DLO_p be the theory of its Fraïssé limit. By disjoint amalgamation, DLO_p has trivial acl. By [loc. cit., Lemma 6.4], if $M \models \mathrm{DLO}_p$, then for all $m \in M_P$, $M_m \models \mathrm{DLO}$.

If $C \subseteq \mathbb{M}_O$ and $\varphi(x)$ is an $L_<$-formula with parameters in C, then, for each $m \in \mathbb{M}_P$, we write $\varphi_m(x)$ for the L-formula obtained by replacing each instance of $<$ with $<_m$. Likewise, if $q(x)$ is a partial $L_<$-type over C, we write $q_m(x)$ for $\{\varphi_m(x) : \varphi(x) \in q\}$. Note that $q_m(x)$ is a partial L-type over Cm.

Fact 4.1 [Chernikov and Ramsey 2016, Lemma 6.5]. Suppose $C \subseteq \mathbb{M}_O$, $(b_i)_{i\in I}$ is a family of distinct elements of $\mathbb{M}_P$, and for each $i \in I$, $p^i(x)$ is a consistent nonalgebraic $L_<$-type over C in $\mathbb{M}_{b_i}$. Then $\bigcup_{i\in I} p^i_{b_i}(x)$ is a consistent partial L-type over $C(b_i)_{i\in I}$.

Recall that a *coheir sequence* over A is a Morley sequence for a global type finitely satisfiable in A. The following lemma is a general fact that is easy and well-known.

Lemma 4.2. *Suppose $M \models T$ and $I = (a_i)_{i<\omega}$ is a coheir sequence over M. Then given any b, there exists $(b_i)_{i<\omega}$ such that $(a_i, b_i)_{i<\omega}$ is a coheir sequence over M and* $\mathrm{tp}(a_i b_i/M) = \mathrm{tp}(a_0 b/M)$ *for all $i < \omega$.*

Proof. Suppose $(a_i)_{i<\omega}$ is a coheir sequence for the global M-finitely satisfiable type $p(x)$. Let N be an $|M|^+$-saturated model containing M. Let a^* realize $p \mid N$, so $a^* \mathop{\smile\hspace{-0.9em}\mid}^u_M N$. By left extension for $\mathop{\smile\hspace{-0.9em}\mid}^u$, we can find b^* such that $\mathrm{tp}(a^*b^*/M) = \mathrm{tp}(a_0 b/M)$ and such that $a^*b^* \mathop{\smile\hspace{-0.9em}\mid}^u_M N$. By saturation of N, $\mathrm{tp}(a^*b^*/N)$ has a unique global M-invariant extension $q(x, y)$, which is finitely satisfiable in M. Likewise, $p(x) \subseteq q(x, y)$, since the restriction of q to formulas in context x is the unique global M-invariant extension of $\mathrm{tp}(a^*/N) = p \mid N$. Let $(a^*_i b^*_i)_{i<\omega}$ be a Morley sequence for q over M. Since $(a^*_i)_{i<\omega}$ and $(a_i)_{i<\omega}$ are both Morley sequences for

p over M, there is an automorphism σ of $\mathbb{M}$ over M such that $\sigma(a_i^*) = a_i$ for all i. Let $b_i = \sigma(b_i^*)$. □

Lemma 4.3. *Let $M \models \mathrm{DLO}_p$. Then $A \mathop{\smile\!\!\!\!\!\vert}^{Kd}_M B$ if and only if*

(1) *$A \cap B \subseteq M$, and*

(2) *for every $m \in M_P$, and for all $b <_m a <_m b'$ with $a \in A_O \setminus M_O$ and $b, b' \in B_O \setminus M_O$, there exists $m' \in M_O$ such that $b <_m m' <_m b'$ (i.e., b and b' live in different $<_m$-cuts in M_O).*

Condition (2) in the statement of Lemma 4.3 can be more succinctly stated as: for every $m \in M_P$, $A_O \mathop{\smile\!\!\!\!\!\vert}^{f}_{M_O} B_O$ in the $L_<$-structure $\mathbb{M}_m$. Nevertheless, we will prove and use the more concrete characterization.

Proof. Suppose $A \mathop{\smile\!\!\!\!\!\vert}^{Kd}_M B$. In any theory, $A \mathop{\smile\!\!\!\!\!\vert}^{Kd}_M B$ implies $A \cap B \subseteq M$, so we have (1). For (2), assume for contradiction that $b <_m a <_m b'$, with $m \in M_P$, $a \in A_O \setminus M_O$, and $b, b' \in B_O \setminus M_O$, and b and b' live in the same $<_m$-cut in M_O, i.e., there is no $m' \in M_O$ such that $b <_m m' <_m b'$.

We will find a global type $q(y, y')$ extending $\mathrm{tp}(bb'/M)$ and finitely satisfiable in M, such that the formula $\varphi(x; b, b')$: $b <_m x <_m b'$ divides along Morley sequences for q over M. We may assume that the set $C = \{c \in M_O \mid c <_m b\}$ is nonempty and has no greatest element. The other case, when the set $D = \{d \in M_O \mid b <_m d\}$ is nonempty and has no least element, is symmetrical.

Consider the filter on $M_O^{yy'}$ generated by

$$\{\psi(M) : \psi(y, y') \in \mathrm{tp}(bb'/M)\} \cup \{(e, e') \mid e' \in C\}.$$

By quantifier elimination, a set Y in this filter contains the intersection of:

(1) $\{(e, e') \mid e' \in C\}$,

(2) A set $\{(e, e') \mid c <_m e <_m e' <_m d\}$ for some $c \in C$ and $d \notin C$, or $\{(e, e') \mid c <_m e <_m e'\}$ for some $c \in C$, and

(3) finitely many nonempty sets in $M_O^{yy'}$, each defined in terms of an order $<_{m'}$ for $m' \neq m$ in M_P.

Since C has no greatest element, we can pick some $c' \in C$ with $c <_m c'$. Then replacing (1) and (2) in the intersection with $\{(e, e') \mid c <_m e <_m e' <_m c'\}$, the intersection of these sets is nonempty, by the extension axioms for the Fraïssé limit, and contained in Y. Thus the filter is proper and extends to an ultrafilter $\mathcal{D}$.

Let $q = \mathrm{Av}(\mathcal{D}, \mathbb{M})$. Suppose $I = (b_i, b_i')_{i<\omega}$ is a Morley sequence for q over M. Since each b_i realizes $\mathrm{tp}(b/M)$,

$$\{(e, e') \in M_O^{yy'} \mid e' <_m b_i\} = \{(e, e') \in M_O^{yy'} \mid e' \in C\} \in \mathcal{D}.$$

So $b_{i+1} <_m b_{i+1}' <_m b_i$ for all $i < \omega$. Thus the $<_m$-intervals (b_i, b_i') are pairwise disjoint, and $b <_m x <_m b'$ divides along I.

Let b^* be a tuple enumerating $B \setminus \{b, b'\}$. By Lemma 4.2, there exists $(b^*_i)_{i<\omega}$ such that $J = (b_i, b'_i, b^*_i)_{i<\omega}$ is a coheir sequence over M and $\text{tp}(b_i, b'_i, b^*_i) = \text{tp}(b, b', b_*) = \text{tp}(B/M)$ for all $i < \omega$. The formula $b <_m x <_m b'$ is contained in $\text{tp}(A/MB)$ and divides along J, which contradicts $A \mathop{\smile\!\!\!\!\!\!|}^{Kd}_M B$.

Conversely, suppose conditions (1) and (2) hold. We may assume that A is disjoint from M (and hence also from B, by (1)), since $(A \setminus M) \mathop{\smile\!\!\!\!\!\!|}^{Kd}_M B$ implies $A \mathop{\smile\!\!\!\!\!\!|}^{Kd}_M B$.

Let $p(x, x', y) = \text{tp}(AB/M)$, where x enumerates A_O, x' enumerates A_P, and y enumerates B. Let $(B_i)_{i<\omega}$ be a Morley sequence for a global M-invariant type extending $\text{tp}(B/M)$. Let $C = M \cup \bigcup_{i<\omega} B_i$. It suffices to show that $q(x, x') = \bigcup_{i<\omega} p(x, x', B_i)$ is a consistent partial type over C.

Let $q_O(x)$ be the subset of q that only mentions the variables x (those of type O). For each $c \in C_P$, let $q_c(x)$ be set of atomic and negated atomic formulas in $q_O(x)$ involving the relation $<_c$. Then there is a partial $L_<$-type $q^c(x)$ over C_O in $\mathbb{M}_c$ such that $(q^c)_c$ is equivalent to q_c. We will show that each $q^c(x)$ is consistent.

If $c \notin M$, then since the B_i are pairwise disjoint over M, there is a unique $i < \omega$ such that $c \in (B_i)_P$. Then $q_c(x)$ is contained in $p(x, x', B_i)$, which is consistent, and hence $q^c(x)$ is consistent as well.

Suppose $c \in M$, and assume for contradiction that $q^c(x)$ is inconsistent. By compactness and density of $\mathbb{M}_c$, there is some variable z from x and some $b_i \in B_i$ and $b'_j \in B_j$ for $i, j < \omega$ such that $b_i \leq b'_j$ in $\mathbb{M}_c$, but $q^c(x)$ entails $b'_j < z < b_i$. Let b and b' be the elements of B corresponding to b_i and b'_j, respectively, and let a be the element of A corresponding to the variable z. Then $b' < a < b$, so by (2) there is some $m' \in M_O$ such that $b' < m' < b$ in $\mathbb{M}_c$. But since $B_i \equiv_M B_j \equiv_M B$, $b'_j < m' < b_i$, contradicting $b_i \leq b'_j$.

Since A is disjoint from B, each type $q^c(x)$ is nonalgebraic, so by Fact 4.1, $\bigcup_{c \in C_P} q_c(x)$ is consistent. By quantifier elimination, $q_O(x)$ is consistent.

Let A' realize $q_O(x)$. It remains to show that $q(A', x')$ is consistent. Each variable in x' is of type P. Since each atomic formula contains at most one variable of type P, it suffices to show that for each variable z in x', the set $r(z)$ of all atomic and negated atomic formulas from $q(A', x')$ involving the relation $<_z$ is consistent.

The type $r(z)$ specifies a linear order on $A'M_O$, which extends to a linear order on $A'M_O(B_i)_O$ for all $i < \omega$. Using the amalgamation property for linear orders, we can find a linear order on $A'C_O$ extending each of the given linear orders. By compactness and the extension axioms for the Fraïssé limit, we can find $c \in \mathbb{M}_P$ such that $<_c$ induces this linear order on $A'C_O$. This completes the proof. □

Theorem 4.4. DLO_p *satisfies the New Kim's Lemma.*

Proof. Suppose $\varphi(x; b)$ Kim-divides over $M \models T$. To show that $\varphi(x; b)$ universally Kim-strictly divides, let $p(y)$ be a global Kim-strict M-invariant type extending $\text{tp}(b/M)$, and let $I = (b_n)_{n<\omega}$ be a Morley sequence for p. Suppose for contradiction that $\{\varphi(x; b_n) : n < \omega\}$ is consistent, realized by a. By Ramsey's theorem, compactness, and an automorphism, we may assume that $(b_n)_{n<\omega}$ is Ma-indiscernible. Now it suffices to show that $a \mathop{\smile\!\!\!\!\!\mid}^{Kd}_M b_0$, since this will contradict the fact that the Kim-dividing formula $\varphi(x; b_0)$ is in $\text{tp}(a/Mb_0)$.

Let A be the set enumerated by a, and let B_n be the set enumerated by b_n for all n. Since $B_1 \mathop{\smile\!\!\!\!\!\mid}^{i}_M B_0$, $B_1 \cap B_0 \subseteq M$. If $c \in A \cap B_0$, then since $AB_0 \equiv_M AB_1$, also $c \in B_1$, so $c \in M$. Thus $A \cap B_0 \subseteq M$.

Now suppose $m \in M_P$ and $d_0 <_m c <_m d_0'$, with $c \in A_O \setminus M_O$ and $d_0, d_0' \in (B_0)_O \setminus M_O$. Suppose for contradiction that there is no $m' \in M_O$ such that $d_0 <_m m' <_m d_0'$. Let d_1 and d_1' be the elements of $(B_1)_O$ corresponding to d_0 and d_0' in $(B_0)_O$. Since $B_1 \equiv_{MA} B_0$, $d_1 <_m c <_m d_1'$, and there is no $m' \in M_O$ such that $d_1 <_m m' <_m d_1'$.

Since p is Kim-strict, $B_0 \mathop{\smile\!\!\!\!\!\mid}^{Kd}_M B_1$ and $B_1 \mathop{\smile\!\!\!\!\!\mid}^{Kd}_M B_0$. By Lemma 4.3, d_0, d_0', d_1, and d_1' are distinct, neither d_0 nor d_0' are in the $<_m$-interval (d_1, d_1'), and neither d_1 nor d_1' are in the $<_m$-interval (d_0, d_0'). It follows that the $<_m$-intervals (d_0, d_0') and (d_1, d_1') are disjoint. This contradicts the fact that c is in both of them.

So there is $m' \in M_O$ such that $d_0 <_m m' <_m d_0'$. By Lemma 4.3, $A \mathop{\smile\!\!\!\!\!\mid}^{Kd}_M B_0$. □

4B. *Bilinear forms over real closed fields.* Let T_∞^{RCF} be the two-sorted theory of an infinite-dimensional vector space over a real closed field with a bilinear form, which is assumed to be either alternating and nondegenerate, or symmetric and positive-definite. This is really two theories, one for each type of bilinear form, but our arguments are identical in both cases so we will not notationally distinguish them. The language has a sort V for the vector space, equipped with the language of abelian groups, a sort R for the real closed field of scalars, equipped with the language of ordered rings, a function symbol $\cdot\colon R \times V \to V$ for scalar multiplication, and a function symbol $[-,-]\colon V \times V \to R$ for the bilinear form.

By [Granger 1999], T_∞^{RCF} is the model companion of the theory of a vector space over a real closed field with an alternating (or symmetric and positive-definite) bilinear form. By [Dobrowolski 2023], this theory additionally has quantifier-elimination in an expanded language, containing, for each n, a predicate I_n on V^n, such that $I_n(v_1, \ldots, v_n)$ holds if and only if $v_1, \ldots, v_n$ are linearly independent, as well as $(n+1)$-ary "coordinate functions" $F_{n,i} : V^{n+1} \to R$ for each $1 \le i \le n$. These functions are which are interpreted so that, if $v_1, \ldots, v_n$ are linearly independent and $w = \sum_{i=1}^n \alpha_i v_i$, then $F_{n,i}(\bar{v}, w) = \alpha_i$, and $F_{n,i}(\bar{v}, w) = 0$ otherwise.

When A is a subset of $\mathbb{M} \models T_\infty^{\text{RCF}}$, we write A_R for the elements of the field sort and A_V for the elements of the vector space sort.

Remark 4.5. As a consequence of quantifier elimination and elementary linear algebra, the field sort R is stably embedded. More precisely, suppose C is a substructure of $\mathbb{M}$. If $\varphi(x)$ is a formula with parameters from C such that every variable is in the field sort R, then $\varphi(x)$ is equivalent to a formula $\psi(x)$ in the language of ordered rings with parameters from C_R. Consequently, for any tuple a from $\mathbb{M}_R$ and any substructure C, $\mathrm{tp}_{\mathrm{RCF}}(a/C_R)$ entails $\mathrm{tp}(a/C)$.

If W is a set of vectors, we write $\langle W\rangle$ for the linear span of W with scalars from the field $\mathbb{M}_R$ (so $\langle W\rangle$ is a large set). By $\dim(W)$, we mean the dimension of $\langle W\rangle$ as a vector space over $\mathbb{M}_R$.

Suppose A, B, and C are substructures of $\mathbb{M}$. We write $A \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{RCF}}_C B$ to mean that A_R and B_R are forking-independent over C_R in the reduct of $\mathbb{M}_R$ to a model of RCF. We write $A \mathop{\smile\!\!\!\!\!\!|}^{V}_C B$ to mean $\langle A_V\rangle \cap \langle B_V\rangle \subseteq \langle C_V\rangle$. Our goal is to show that T^{RCF}_∞ satisfies New Kim's Lemma, which will involve characterizing $\mathop{\smile\!\!\!\!\!\!|}^{Kd}$ in this theory in terms of $\mathop{\smile\!\!\!\!\!\!|}^{\mathrm{RCF}}$ and $\mathop{\smile\!\!\!\!\!\!|}^{V}$. The argument is the analogue of [Kaplan and Ramsey 2020, Proposition 9.37] (incorporating the corrections of [Dobrowolski 2023, Proposition 8.12]). A similar characterization of Kim-independence in the theory of a bilinear form on a vector space over an NSOP_1 field occurs in [Bossut 2023].

We begin with another general lemma which, in conjunction with Lemma 4.2, will allow us to upgrade a coheir sequence in $\mathrm{tp}_{\mathrm{RCF}}(B_R/M_R)$ in RCF to a coheir sequence in $\mathrm{tp}(B/M)$ in T^{RCF}_∞.

Lemma 4.6. *Suppose $L \subseteq L'$ are languages, T' is an L'-theory and $T = T' \restriction L$. If $A \subseteq B$ and $I = (c_i)_{i<\omega}$ is a coheir sequence over A in T, then there is $I' \models \mathrm{tp}_L(I/A)$ which is a coheir sequence in T' over B.*

Proof. If $(c_i)_{i<\omega}$ is a coheir sequence over A in T, there is some ultrafilter $\mathcal{D}$ on A^n, where n is the length of c_0, such that I is a Morley sequence over A in the global A-finitely satisfiable type $\mathrm{Av}_L(\mathcal{D}, \mathbb{M})$. To see this, stretch I to $(c_i)_{i<\omega+1}$ and observe that the family of sets $\{\varphi(A; c_{<\omega}) : \varphi(x; c_{<\omega}) \in \mathrm{tp}(c_\omega/Ac_{<\omega})\} \subseteq \mathcal{P}(A^n)$ generates a filter and hence extends to an ultrafilter $\mathcal{D}$. It is easily checked that this $\mathcal{D}$ works. Let $\mathcal{E}$ be the ultrafilter on B^n induced by $\mathcal{D}$, i.e., a subset $X \subseteq B^n$ satisfies $X \in \mathcal{E}$ if and only if $X \cap A^n \in \mathcal{D}$. Then we can take I' to be Morley over B in the global B-finitely satisfiable type $\mathrm{Av}_{L'}(\mathcal{E}, \mathbb{M})$. □

Lemma 4.7. *If $M \models T^{\mathrm{RCF}}_\infty$ and $A \mathop{\smile\!\!\!\!\!\!|}^{Kd}_M B$, then $A \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{RCF}}_M B$.*

Proof. Because RCF is an NTP_2 theory, any dividing formula divides along some coheir sequence by [Chernikov and Kaplan 2012, Lemma 3.12]. So if $A \not\mathop{\smile\!\!\!\!\!\!|}^{\mathrm{RCF}}_M B$, then there is a formula $\varphi(x; b)$ in $\mathrm{tp}_{\mathrm{RCF}}(A_R/M_R B_R)$ and a coheir sequence $I = (B_i)$ over M_R in $\mathrm{tp}_{\mathrm{RCF}}(B_R/M_R)$ such that $\varphi(x; b)$ divides along I. By Lemmas 4.6 and 4.2, there is a coheir sequence $I' = (B'_i)_{i<\omega}$ over M in $\mathrm{tp}(B/M)$ such that $((B'_i)_R)_{i<\omega} \equiv^{\mathrm{RCF}}_{M_R} I$. Then $\varphi(x; b)$ divides along I', and $A \not\mathop{\smile\!\!\!\!\!\!|}^{Kd}_M B$. □

Lemma 4.8. *Suppose $M \models T^{\mathrm{RCF}}_\infty$:*

(1) *If $A \mathop{\smile\!\!\!\!\!\!|}^{u}_M B$, then $A \mathop{\smile\!\!\!\!\!\!|}^{V}_M B$.*

(2) *If* $(B_i)_{i<\omega}$ *is a* $\mathop{\smile\!\!\!\!\!\mid}^{V}_{M}$*-independent sequence (i.e.,* $B_i \mathop{\smile\!\!\!\!\!\mid}^{V}_{M} B_0 \cdots B_{i-1}$ *for all* $i < \omega$*), and there exists* A' *such that* $A'B_i \equiv_M AB$ *for all* $i < \omega$*, then* $A \mathop{\smile\!\!\!\!\!\mid}^{V}_{M} B$*.*

(3) *If* $A \mathop{\smile\!\!\!\!\!\mid}^{Kd}_{M} B$*, then* $A \mathop{\smile\!\!\!\!\!\mid}^{V}_{M} B$*.*

Proof. Suppose that $A \not\mathop{\smile\!\!\!\!\!\mid}^{V}_{M} B$. Then $\langle A_V \rangle \cap \langle B_V \rangle \not\subseteq \langle M_V \rangle$, so there exists a vector v, a finite linearly independent tuple a from A_V, and a finite linearly independent tuple b from B_V such that $v \in \langle a \rangle \cap \langle b \rangle$ and $v \notin \langle M_V \rangle$. Let $C = \langle b \rangle \cap \langle M_V \rangle$, and note that C is a subspace of the finite-dimensional space $\langle b \rangle$. Let c be a finite basis for C. Note that the formula $\varphi(x; b, c)$:

$$\exists w (I_{|a|}(x) \wedge \neg I_{|a|+1}(x, w) \wedge \neg I_{|b|+1}(b, w) \wedge I_{|c|+1}(c, w)),$$

which asserts that x is linearly independent and $\langle x \rangle \cap \langle b \rangle \not\subseteq \langle c \rangle$, is in $\operatorname{tp}(a/Mb)$. With the above notation set, we now prove (1) and (2).

For (1), assume for contradiction that $A \mathop{\smile\!\!\!\!\!\mid}^{u}_{M} B$. Since $\operatorname{tp}(a/Mb)$ is finitely satisfiable in M, there is some $a' \in M_V$ satisfying $\varphi(a'; b, c)$. Let w' be the witness to the existential quantifier. Then $w' \in \langle a' \rangle \subseteq \langle M_V \rangle$ and $w' \in \langle b \rangle$, so $w' \in \langle b \rangle \cap \langle M_V \rangle = C$. But $w' \notin \langle c \rangle$, contradiction.

For (2), assume for contradiction that there exists a $\mathop{\smile\!\!\!\!\!\mid}^{V}_{M}$-independent sequence $(B_i)_{i<\omega}$ and A' such that $A'B_i \equiv_M AB$ for all $i < \omega$. Let $(b_i)_{i<\omega}$ be the restriction of this sequence to the tuples b_i from B_i corresponding to the tuple b in B, and let a' be the tuple from A' corresponding to the tuple a in A. Let $k = \dim(\langle a' \rangle) = |a'|$, and let $v_0, \ldots, v_k$ be such that $v_i \in \langle a' \rangle \cap \langle b_i \rangle \setminus \langle c \rangle$ for all $i < k+1$. Since these $k+1$ vectors are all in $\langle a' \rangle$, they are not linearly independent, and we can write one of them, say v_j, as a linear combination of $v_0, \ldots, v_{j-1}$. Then $v_j \in \langle b_j \rangle \cap \langle b_0, \ldots, b_{j-1} \rangle \setminus \langle c \rangle$. But since $b_j \mathop{\smile\!\!\!\!\!\mid}^{V}_{M} b_0 \cdots b_{j-1}$, $\langle b_j \rangle \cap \langle b_0, \ldots, b_{j-1} \rangle \subseteq \langle b_j \rangle \cap \langle M_V \rangle = \langle c \rangle$, contradiction.

For (3), let $(B_i)_{i<\omega}$ be a coheir sequence in $\operatorname{tp}(B/M)$. Since $A \mathop{\smile\!\!\!\!\!\mid}^{Kd}_{M} B$, by compactness there exists A' such that $A'B_i \equiv_M AB$ for all $i < \omega$. By (1), $(B_i)_{i<\omega}$ is a $\mathop{\smile\!\!\!\!\!\mid}^{V}_{M}$-independent sequence, and by (2), $A \mathop{\smile\!\!\!\!\!\mid}^{V}_{M} B$. □

Theorem 4.9. *If* $M \models T^{\mathrm{RCF}}_{\infty}$, $A = \operatorname{acl}(AM)$, $B = \operatorname{acl}(BM)$, *then* $A \mathop{\smile\!\!\!\!\!\mid}^{Kd}_{M} B$ *if and only if* $A \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{RCF}}_{M} B$ *and* $A \mathop{\smile\!\!\!\!\!\mid}^{V}_{M} B$*.*

Proof. One direction is Lemmas 4.7 and 4.8(3).

In the other direction, suppose that $A \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{RCF}}_{M} B$ and $A \mathop{\smile\!\!\!\!\!\mid}^{V}_{M} B$. Let $(B_i)_{i<\omega}$ be a Morley sequence over M for a global M-invariant type extending $\operatorname{tp}(B/M)$. Since $A \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{RCF}}_{M} B$, we can find A'_R such that $A'_R(B_i)_R \equiv^{\mathrm{RCF}}_{M_R} A_R B_R$ for all $i < \omega$. By Remark 4.5, $A'_R B_i \equiv_M A_R B$ for all $i < \omega$. Let $\tilde{R}$ be the field $(\operatorname{acl}(A'_R(B_i)_{i<\omega}))_R$.

Let $\bar{m} = (m_i)_{i<\alpha}$ be a tuple from M_V which is a basis of $\langle M_V \rangle$. Choose $\bar{a} = (a_i)_{i<\beta}$ from A_V such that $\bar{a}\bar{m}$ is a basis of $\langle A_V \rangle$ and choose $\bar{b}_i = (b_{i,j})_{j<\gamma}$ from $(B_i)_V$ such that $\bar{m}\bar{b}_i$ is a basis of $\langle (B_i)_V \rangle$. Since $(B_i)_{i<\omega}$ is a $\mathop{\smile\!\!\!\!\!\mid}^{i}_{M}$-independent sequence, by Lemma 4.8(3) it is also a $\mathop{\smile\!\!\!\!\!\mid}^{V}_{M}$-independent sequence. Thus $\bar{m}$ and

$(\bar{b}_i)_{i<\omega}$ are linearly independent. Let $\tilde{V} = \langle \bar{m}(\bar{b}_i)_{i<\omega}\rangle_{\tilde{R}}$, the vector space over $\tilde{R}$ spanned by this basis. Note that, unlike $\langle \bar{m}(\bar{b}_i)_{i<\omega}\rangle$, this is a small set, and it contains $(B_i)_{i<\omega}$, since $\tilde{R}$ contains the values of the coordinate functions $F_{n,i}$ on tuples from $(B_i)_{i<\omega}$. Let $\tilde{N}$ be the substructure of $\mathbb{M}$ with $\tilde{N}_R = \tilde{R}$ and $\tilde{N}_V = \tilde{V}$. Note that if we give the symbols θ_n and $F_{n,i}$ their intended interpretations in $\tilde{N}$, they agree with the interpretations of these symbols in $\mathbb{M}$.

Let $\bar{a}' = (a_i')_{i<\beta}$ be a tuple of new vectors (not in $\mathbb{M}_V$) of the same length as $\bar{a}$. Let W be the $\tilde{R}$-vector space extending $\tilde{V}$ with basis $\bar{a}'$, $\bar{m}$, and $(\bar{b}_i)_{i<\omega}$. We build a structure N extending $\tilde{N}$ with $N_R = \tilde{N}_R = \tilde{R}$ and $N_V = W$. The field structure and vector space structure have been determined, so it remains to define the bilinear form $[-,-]^N$. To do this, it suffices to define the form on every pair of basis vectors for W such that at least one comes from $\bar{a}'$, and extend linearly.

For all $i < \alpha$, $i' < \beta$, $j' < \beta$, $j < \omega$, and $k < \gamma$, set

$$[a'_{i'}, a'_{j'}]^N = [a_{i'}, a_{j'}]^{\mathbb{M}}, \quad [a'_{i'}, m_i]^N = [a_{i'}, m_i]^{\mathbb{M}}, \quad [a'_{i'}, b_{j,k}]^N = [a_i, b_{0,k}]^{\mathbb{M}}.$$

These conditions uniquely determine a bilinear form on all pairs of vectors from W, which is alternating or symmetric and positive-definite, as required by T_∞^{RCF}. We can extend the language to include the θ_n and $F_{n,i}$ in the natural way, and the interpretations of these symbols agree with those on $\tilde{N}$, since $\tilde{N}_R = N_R$.

Now we can embed N into $\mathbb{M}$ over $\tilde{N}$. Let A'_V be the image under this embedding of the subset of N corresponding to A_V, and let $A' = (A'_R, A'_V)$. It follows by construction and quantifier elimination that $A'B_i \equiv_M AB$ for all $i < \omega$. Thus $A \mathop{\smile\!\!\!\!\!\!|}\nolimits_M^{Kd} B$. $\square$

Theorem 4.10. *The theory T_∞^{RCF} satisfies New Kim's Lemma.*

Proof. Let $M \models T_\infty^{\mathrm{RCF}}$ and suppose $\varphi(x; b)$ Kim-divides over M. Let $I = (b_i)_{i<\omega}$ be a Morley sequence over M for a global Kim-strict M-invariant type $q(y) \supseteq \mathrm{tp}(b/M)$. We would like to show that $\varphi(x; b_i)$ divides along I. Assume, towards contradiction, that there exists a realizing $\{\varphi(x; b_i) : i < \omega\}$. By Ramsey's theorem, compactness, and an automorphism, we may assume that $(b_i)_{i<\omega}$ is indiscernible over $A = \mathrm{acl}(Ma)$. For each $i < \omega$, let $B_i = \mathrm{acl}(Mb_i)$, with each B_i enumerated in such a way that $(B_i)_{i<\omega}$ remains indiscernible over A.

Since $(b_i)_{i<\omega}$ is a $\mathop{\smile\!\!\!\!\!\!|}\nolimits_M^{i}$-independent sequence, it is a $\mathop{\smile\!\!\!\!\!\!|}\nolimits_M^{K}$-independent sequence, and thus $(B_i)_{i<\omega}$ is a $\mathop{\smile\!\!\!\!\!\!|}\nolimits_M^{K}$-independent sequence. By Lemma 4.8(3), $(B_i)_{i<\omega}$ is a $\mathop{\smile\!\!\!\!\!\!|}\nolimits_M^{V}$-independent sequence, and since $AB_i \equiv_M AB_0$ for all $i < \omega$, $A \mathop{\smile\!\!\!\!\!\!|}\nolimits_M^{V} B_0$ by Lemma 4.8(2).

We now claim that $((B_i)_R)_{i<\omega}$ is a (Kim-)strict Morley sequence over M_R in RCF. Let N be an $|M|^+$-saturated model containing M and $(B_i)_{i<\omega}$. Let b_ω realize $q|_N$, and let $B_\omega = \mathrm{acl}(Mb_\omega)$. Since $(b_i)_{i\leq\omega}$ is a Morley sequence over M, and hence M-indiscernible, and $(B_i)_{i<\omega}$ is M-indiscernible, we can enumerate B_ω in such a way that $(B_i)_{i\leq\omega}$ remains M-indiscernible.

Since q is Kim-strict, $b_\omega \mathop{\perp\!\!\!\perp}^K_M N$ and $N \mathop{\perp\!\!\!\perp}^K_M b_\omega$, so $B_\omega \mathop{\perp\!\!\!\perp}^K_M N$ and $N \mathop{\perp\!\!\!\perp}^K_M B_\omega$, and hence $B_\omega \mathop{\perp\!\!\!\perp}^{\mathrm{RCF}}_M N$ and $N \mathop{\perp\!\!\!\perp}^{\mathrm{RCF}}_M B_\omega$, by Theorem 4.9. Since RCF is an NIP theory, $\mathop{\perp\!\!\!\perp}^i_M = \mathop{\perp\!\!\!\perp}^f_M$ in RCF; see [Simon 2015, Corollary 5.22]. Thus $\mathrm{tp}_{\mathrm{RCF}}((B_\omega)_R/N_R)$ extends to a global M_R-invariant type q_* which is strict over M_R in RCF. Indeed, suppose for contradiction that $C_R \subseteq \mathbb{M}_R$, $B'_R \models q_*|_{N_R C_R}$, and $C_R \not\mathop{\perp\!\!\!\perp}^f_{M_R} B'_R$ in RCF. Then $c \not\mathop{\perp\!\!\!\perp}^f_{M_R} B'_R$ for some finite tuple c from C_R, whose type over M_R is realized by $c' \in N_R$. Then $c' \not\mathop{\perp\!\!\!\perp}^f_{M_R} B'_R$ in RCF by invariance of q_*, contradicting $N \mathop{\perp\!\!\!\perp}^{\mathrm{RCF}}_M B_\omega$.

By M-indiscernibility of $(B_i)_{i\leq\omega}$, $(B_i)_R \models q_*|_{M_R(B_{<i})_R}$ for all i, so $((B_i)_R)_{i<\omega}$ is a strict Morley sequence over M_R in RCF. Since $((B_i)_R)_{i<\omega}$ is A_R-indiscernible, it follows that $A \mathop{\perp\!\!\!\perp}^{\mathrm{RCF}}_M B_0$ by the NTP_2 Kim's Lemma (Theorem 3.5).

Since $A \mathop{\perp\!\!\!\perp}^{\mathrm{RCF}}_M B_0$ and $A \mathop{\perp\!\!\!\perp}^V_M B_0$, by Theorem 4.9, $A \mathop{\perp\!\!\!\perp}^{Kd}_M B_0$. This contradicts the fact that $\mathrm{tp}(A/MB_0)$ contains the formula $\varphi(x; b_0)$, which Kim-divides over M, since $b_0 \equiv_M b$. □

4C. *Nonexample: the Henson graph.* The Henson graph, or generic triangle-free graph, is the Fraïssé limit of the class of finite triangle-free graphs. Its complete theory $T_\triangle$ is SOP_3 and NSOP_4. Conant [2017] analyzed forking and dividing in $T_\triangle$ in detail. We will use the following characterization of $\mathop{\perp\!\!\!\perp}^f$.

Fact 4.11 [Conant 2017, Theorem 5.3]. Suppose that A and B are sets in $\mathbb{M} \models T_\triangle$ and $M \models T_\triangle$. Then $A \mathop{\perp\!\!\!\perp}^f_M B$ if and only if $A \cap B \subseteq M$ and for all $a \in A$ and $b \neq c \in B \setminus M$, if aRb and aRc, then there exists $m \in M$ such that mRb and mRc.

We will show that a very weak variant of Kim's Lemma fails in $T_\triangle$: strict dividing does not imply universal strict dividing. Since strict dividing implies Kim-dividing and universal Kim-strict dividing implies universal strict dividing, it follows that $T_\triangle$ fails to satisfy New Kim's Lemma.

Theorem 4.12. *Modulo $T_\triangle$, there is a formula which strictly divides but does not universally strictly divide. Thus $T_\triangle$ does not satisfy New Kim's Lemma.*

Proof. Let $M \models T$. Let b and c be elements of $\mathbb{M} \setminus M$ with $\neg bRc$, such that b has a single neighbor in M, call it m, and c has no neighbors in M.[6] Consider the formula $\varphi(x; b, c)$: $xRb \wedge xRc$. It suffices to find two strict global M-invariant types $p(y, z)$ and $q(y, z)$ extending $\mathrm{tp}(b, c/M)$ such that $\varphi(x; b, c)$ divides along Morley sequences for p but does not divide along Morley sequences for q.

Let $p(y, z)$ extend $\mathrm{tp}(bc/M)$ by including, for each $d \in \mathbb{M} \setminus M$, $y \neq d$, $z \neq d$ and $\neg yRd$. Additionally, include zRd if $d \models \mathrm{tp}(b/M)$ and $\neg zRd$ otherwise. We claim this defines a consistent partial type. Any inconsistency would come from a

[6] Really, all we will use is that the set of neighbors of b in M is nonempty and disjoint from the set of neighbors of c in M.

triangle involving the variables and elements of $\mathbb{M}$. Such a triangle cannot contain y, since $\neg yRz$ and y has an edge to exactly one element of $\mathbb{M}$, namely m. Since z only has edges to realizations of $\mathrm{tp}(b/M)$, any triangle containing z contains two realizations of $\mathrm{tp}(b/M)$. But no two realizations of $\mathrm{tp}(b/M)$ are adjacent, since they are both adjacent to m.

By quantifier elimination, this partial type determines a complete M-invariant type over $\mathbb{M}$. Letting $I = (b_i, c_i)_{i<\omega}$ be a Morley sequence for p over M, $\varphi(x; b, c)$ divides along I, since $\{\varphi(x; b_1, c_1), \varphi(x; b_2, c_2)\}$ entails $\{xRb_1, xRc_2\}$, and b_1Rc_2.

Now let $q(y, z)$ extend $\mathrm{tp}(bc/M)$ by including, for each $d \in \mathbb{M} \setminus M$, $y \neq d$, $z \neq d$, $\neg yRd$, and $\neg zRd$. This defines a consistent partial type, since the only edge from a variable to an element of $\mathbb{M}$ is the single edge from y to m. Again, by quantifier elimination, this determines a complete M-invariant type over $\mathbb{M}$. And if $J = (b_i, c_i)_{i<\omega}$ is a Morley sequence for q over M, then $\varphi(x; b, c)$ does not divide along J. Indeed, since there are no edges among the vertices $\{b_i, c_i : i < \omega\}$, $\{\varphi(x; b_i, c_i) : i < \omega\}$ does not induce any triangles.

It remains to show that both p and q are strict. Let $A \subseteq \mathbb{M}$, $b_0, c_0 \models p|_{MA}$, and $b_1, c_1 \models q|_{MA}$. We would like to show that for $i \in \{0, 1\}$, $A \mathop{\smile\!\!\!\!\!\mid}^f_M b_i c_i$. In each case, $A \cap \{b_i, c_i\} = \varnothing \subseteq M$, and there is no $a \in A$ such that aRb_i and aRc_i (since b_i is not adjacent to any element of $A \setminus M$, and c_i is not adjacent to any element of M). By Fact 4.11, $A \mathop{\smile\!\!\!\!\!\mid}^f_M b_i c_i$. □

5. Syntax

In this section, we isolate a tree property, provisionally called BTP, which generalizes TP_2 and SOP_1, and we show that NBTP theories satisfy New Kim's Lemma. We also show that NBTP theories are NATP. We have not succeeded in proving that New Kim's Lemma characterizes NBTP theories.

For ordinals $\alpha, \beta \leq \omega$, write $\alpha_*^{<\beta}$ for the forest obtained by removing the root from $\alpha^{<\beta}$:

- A *left-leaning path* in $\alpha_*^{<\beta}$ is a sequence (λ_n) such that if $\lambda_n = \eta^\frown\langle i\rangle$, then $\eta^\frown\langle j\rangle \lhd \lambda_{n+1}$ for some $j \leq i$.
- A *right-veering path* in $\alpha_*^{<\beta}$ is a sequence (ρ_n) such that if $\rho_n = \eta^\frown\langle i\rangle$, then $\eta^\frown\langle j\rangle \trianglelefteq \rho_{n+1}$ for some $j > i$.

Note that to get to the next element in a left-leaning path, one *optionally* moves leftward to a sibling and then moves *strictly* upward to a descendent, while in a right-veering path, one moves *strictly* rightward to a sibling, and then *optionally* moves upward to a descendent.

Definition 5.1. A formula $\varphi(x; y)$ has k-BTP (k*-bizarre tree property*) with $k < \omega$ if there exists a forest of tuples $(a_\eta)_{\eta \in \omega_*^{<\omega}}$ satisfying the following conditions:

- For every left-leaning path $(\lambda_n)_{n<\omega}$, $\{\varphi(x; a_{\lambda_n}) : n < \omega\}$ is consistent.
- For every right-veering path $(\rho_n)_{n\in\omega}$, $\{\varphi(x; a_{\rho_n}) : n < \omega\}$ is k-inconsistent.

A theory T has BTP if there is some formula $\varphi(x; y)$ and some $k < \omega$ such that φ has k-BTP. Otherwise, T is NBTP.

Theorem 5.2. *Suppose T is* NBTP. *Then T satisfies New Kim's Lemma.*

Proof. We prove the contrapositive. If New Kim's Lemma fails, then we have a formula $\varphi(x; b)$, a model $M \models T$, and global M-invariant types $p(y)$ and $q(y)$ extending $\mathrm{tp}(b/M)$ such that $p(y)$ is Kim-strict and $\varphi(x; b)$ divides along Morley sequences for q but not along Morley sequences for p. Fix $k < \omega$ such that if $(b_i)_{i<\omega}$ is a Morley sequence for q, then $\{\varphi(x; b_i) : i < \omega\}$ is k-inconsistent.

For arbitrary m and n in ω, we will build a finite forest $(a_\eta)_{\eta\in m_*^{<n}}$ such that:

- For every left-leaning path $(\lambda_i)_{1\leq i\leq \ell}$ in $m_*^{<n}$, $(a_{\lambda_\ell}, \dots, a_{\lambda_1})$ starts a Morley sequence for p over M, and hence $\{\varphi(x; a_{\lambda_i}) : 1 \leq i \leq \ell\}$ is consistent.
- For every right-veering path $(\rho_i)_{1\leq i\leq \ell}$ in $m_*^{<n}$, $(a_{\rho_\ell}, \dots, a_{\rho_1})$ starts a Morley sequence for q over M, and hence $\{\varphi(x; a_{\lambda_i}) : 1 \leq i \leq \ell\}$ is k-inconsistent.

By compactness, this will suffice to show that $\varphi(x; y)$ has k-BTP.

Fix $m < \omega$ with $m > 0$, and proceed by induction on n. The base cases $n = 0$ and $n = 1$ are trivial, since $m_*^{<n}$ is empty.

Suppose we are given $F_0 = (a_\eta)_{\eta\in m_*^{<n}}$ satisfying the induction hypothesis. Let b_0 realize $p|_{MF_0}$. Since p is Kim-strict, $F_0 \mathop{\smile\!\!\!\!\!\mid}^K_M b_0$. Let $r(z, y) = \mathrm{tp}(F_0 b_0/M)$.

By induction on $1 \leq \ell \leq m$, we now find $(b_i, F_i)_{i<\ell}$ such that:

(1) $F_i \equiv_M F_0$ for all $i < \ell$.

(2) b_i realizes $p|_{MF_j}$ if $i \leq j$.

(3) $(b_i, b_{i+1}, \dots, b_{\ell-1})$ starts a Morley sequence in q over MF_j if $i > j$.

In the base case $\ell = 1$, b_0 and F_0 satisfy the conditions.

Given $(b_i, F_i)_{i<\ell}$ satisfying (1)–(3) for $\ell < m$, let b_ℓ realize $q|_{M(b_i,F_i)_{i<\ell}}$. Then (3) is satisfied for $\ell+1$. Since $r(z, b_0) = \mathrm{tp}(F_0/Mb_0)$ does not Kim-divide over M and $(b_i)_{i<\ell+1}$ starts a Morley sequence for a global M-invariant type, $\bigcup_{i<\ell+1} r(z, b_i)$ is consistent. Let F_ℓ realize this type. Then (1) is satisfied for $\ell + 1$. Now since $r(F_0, z) = \mathrm{tp}(b_0/MF_0) = p|_{MF_0}$, p is M-invariant, and $F_\ell \equiv_M F_0$, we have, for all $i < \ell + 1$, $\mathrm{tp}(b_i/MF_\ell) = r(F_\ell, z) = p|_{MF_\ell}$, and thus (2) is satisfied for $\ell + 1$.

Having constructed $(b_i, F_i)_{i<m}$, we reindex to define the forest $(a'_\eta)_{\eta\in m_*^{<n+1}}$. By (1), we can write $F_i = (a^i_\eta)_{\eta\in m_*^{<n}}$, and each F_i satisfies the induction hypothesis. Set $a'_{\langle i\rangle} = b_{m-i-1}$, and $a'_{\langle i\rangle^\frown\eta} = a^{m-i-1}_\eta$. Note that the reindexing by $(m - i - 1)$ means that our sequence $(b_i, F_i)_{i<m}$ proceeds leftward in the new forest.

A left-leaning path in the new forest begins with at most one element b_i at the bottom level and is followed by some left-leaning path in F_j with $i \leq j$. By (2) and induction, the reverse sequence starts a Morley sequence for p over M. A right-veering path in the new forest may begin with elements $b_{i_1}, \ldots, b_{i_\ell}$ at the bottom level, with $i_1 > \cdots > i_\ell$, and is followed by a right-veering path in some F_j with $i_\ell > j$. By (3) and induction, the reverse sequence starts a Morley sequence for q over M. □

We now situate NBTP relative to the other tree properties.

Proposition 5.3. *If T is* NTP_2, *then T is* NBTP.

Proof. Assume $\varphi(x; y)$ has k-BTP, witnessed by $(a_\eta)_{\eta \in \omega_*^{<\omega}}$. Consider the array $(b_{i,j})_{i,j<\omega}$ with $b_{i,j} = a_{(0^i)\frown\langle j\rangle}$, where 0^i denotes the string of length i consisting of all 0's.

For all $f : \omega \to \omega$, the sequence $(\lambda_i)_{i<\omega}$ with $\lambda_i = (0^i)\frown\langle f(i)\rangle$ is a left-leaning path. So $\{\varphi(x; b_{i,f(i)}) : i < \omega\} = \{\varphi(x; a_{\lambda_i}) : i < \omega\}$ is consistent.

For all $i < \omega$, the sequence $(\rho_j)_{j<\omega}$ with $\rho_j = (0^i)\frown\langle j\rangle$ is a right-veering path. So $\{\varphi(x; b_{i,j}) : j < \omega\} = \{\varphi(x; a_{\rho_j}) : j < \omega\}$ is k-inconsistent.

Thus $\varphi(x; y)$ has TP_2. □

When $k > 2$, a witness to k-BTP does not directly contain a witness to SOP_1, but rather a variant of SOP_1 with k-inconsistency instead of 2-inconsistency. So for the implication from NSOP_1 to NBTP, we will use the following alternative characterization of SOP_1 from [Kaplan and Ramsey 2020].

Fact 5.4 [Kaplan and Ramsey 2020, Proposition 2.4]. *T has SOP_1 if and only if there exists $k < \omega$ and an array $(c_{i,j})_{i<\omega, j<2}$ such that:*

- $c_{n,0} \equiv_{(c_{i,j})_{i<n,j<2}} c_{n,1}$ *for all $n < \omega$.*
- $\{\varphi(x; c_{i,0}) : i < \omega\}$ *is consistent.*
- $\{\varphi(x; c_{i,1}) : i < \omega\}$ *is k-inconsistent.*

Proposition 5.5. *If T is* NSOP_1, *then T is* NBTP.

Proof. Assume $\varphi(x; y)$ has k-BTP, witnessed by $(a_\eta)_{\eta \in \omega_*^{<\omega}}$. Consider the binary subtree $(b_\eta)_{\eta \in 2^{<\omega}}$ with $b_\eta = a_{\langle 0\rangle\frown\eta}$. This tree does not witness SOP_1, but it does have the following properties, which will be sufficient to obtain SOP_1:

- For any $\rho \in 2^\omega$, $\{\varphi(x; b_{\rho|n}) : n < \omega\}$ is consistent (since the corresponding sequence in our original forest is a left-leaning path).
- For any $\mu_1, \ldots, \mu_k \in 2^{<\omega}$ such that $\mu_i\frown\langle 1\rangle \trianglelefteq \mu_{i+1}$ for all $1 \le i < k$, $\{\varphi(x; b_{\mu_i\frown\langle 0\rangle}) : 1 \le i \le k\}$ is inconsistent (since the corresponding sequence in our original forest is a right-veering path of length k).

By compactness, we can obtain a tree $(b_\eta)_{\eta \in 2^{<\kappa}}$, where $\kappa > |S_y(T)|$, which satisfies the obvious extensions of the two properties above.

Following the proof of [Chernikov and Ramsey 2016, Proposition 5.2], we define $(\eta_i, \nu_i)_{i<\omega}$ in $2^{<\kappa}$ by recursion. Given $(\eta_i, \nu_i)_{i<n}$ (and setting $\eta_{-1} = \langle\,\rangle$ when $n = 0$), let $\mu_\alpha = \eta_{n-1}{}^\frown(1^\alpha)^\frown\langle 0\rangle$ for all $\alpha < \kappa$. Since $\kappa > |S_y(T)|$, there are $\alpha < \beta < \kappa$ such that b_{μ_α} and b_{μ_β} have the same type over $(b_{\eta_i}, b_{\nu_i})_{i<n}$. Let $\nu_n = \mu_\alpha$ and $\eta_n = \mu_\beta$. Directly from the construction, we have the following properties:

(1) $b_{\eta_n} \equiv_{(b_{\eta_i}, b_{\nu_i})_{i<n}} b_{\nu_n}$ for all n.

(2) If $i < j$, then $\eta_i \lhd \eta_j, \nu_j$.

(3) For all i, $(\eta_i \wedge \nu_i)^\frown\langle 1\rangle \trianglelefteq \eta_i$ and $(\eta_i \wedge \nu_i)^\frown\langle 0\rangle = \nu_i$.

Now, in the statement of Fact 5.4, set $c_{i,0} = b_{\eta_i}$ and $c_{i,1} = b_{\nu_i}$ for all $i < \omega$. We have $c_{n,0} \equiv_{(c_{i,j})_{i<n,j<2}} c_{n,1}$ by (1). Since $(\eta_i)_{i<\omega}$ is a chain in $2^{<\kappa}$ by (2), $\{\varphi(x; c_{i,0}) : i < \omega\} = \{\varphi(x; b_{\eta_i}) : i < \omega\}$ is consistent. And setting $\mu_i = (\eta_i \wedge \nu_i)$ for all i, note that by (2) and (3), $\nu_i = \mu_i^\frown\langle 0\rangle$, and $\mu_i^\frown\langle 1\rangle \trianglelefteq \eta_i \trianglelefteq (\eta_j \wedge \nu_j) = \mu_j$ when $i < j$. So $\{\varphi(x; c_{i,1}) : i < \omega\} = \{\varphi(x; b_{\mu_i^\frown\langle 0\rangle}) : i < \omega\}$ is k-inconsistent. Thus T has SOP_1. □

Proposition 5.6. *If T is* NBTP, *then T is* NATP.

Proof. Assume $\varphi(x; y)$ has ATP, witnessed by $(a_\eta)_{\eta\in 2^{<\omega}}$.

Define a map $e\colon \omega^{<\omega} \to 2^{<\omega}$ by recursion on the length of the input sequence:

$$e(\langle\,\rangle) = \langle\,\rangle,$$
$$e(\eta^\frown\langle i\rangle) = e(\eta)^\frown\langle 0\rangle^\frown(1^{2i}).$$

Note that if $\eta \trianglelefteq \nu$, then $e(\eta) \trianglelefteq e(\nu)$.

Now define $f : \omega^{<\omega} \to 2^{<\omega}$ by $f(\eta) = e(\eta)^\frown\langle 1\rangle$, and consider the tree $(b_\eta)_{\eta\in\omega_*^{<\omega}}$ with $b_\eta = a_{f(\eta)}$.

If $(\lambda_n)_{n<\omega}$ is a left-leaning path, we claim that $\{f(\lambda_n) : n < \omega\}$ is an antichain in $2^{<\omega}$, and hence $\{\varphi(x; b_{\lambda_n}) : n < \omega\} = \{\varphi(x; a_{f(\lambda_n)}) : n < \omega\}$ is consistent.

So fix $n < m$ in ω. Writing $\lambda_n = \eta^\frown\langle i\rangle$, we have $\eta^\frown\langle j\rangle \lhd \lambda_{n+1}$ for some $j \le i$. Now if $\eta^\frown\langle j\rangle \lhd \nu$, then also $\eta^\frown\langle j\rangle \lhd \nu'$ whenever ν' is a descendent of ν or a descendent of a leftward sibling of ν. Since $(n+1) \le m$, it follows that $\eta^\frown\langle j\rangle \lhd \lambda_m$. Let $j' < \omega$ be such that $\eta^\frown\langle j\rangle^\frown\langle j'\rangle \trianglelefteq \lambda_m$.

Now $f(\lambda_n) = e(\eta^\frown\langle i\rangle)^\frown\langle 1\rangle = e(\eta)^\frown\langle 0\rangle^\frown(1^{2i+1})$. On the other hand, $f(\lambda_m) = e(\lambda_m)^\frown\langle 1\rangle$ has as an initial segment $e(\eta^\frown\langle j\rangle^\frown\langle j'\rangle) = e(\eta)^\frown\langle 0\rangle^\frown(1^{2j})^\frown\langle 0\rangle^\frown(1^{2j'})$. Since $2i+1 \ne 2j$, $f(\lambda_n) \perp f(\lambda_m)$, as desired.

If $(\rho_n)_{n<\omega}$ is a right-veering path, we claim that $f(\rho_n) \trianglelefteq f(\rho_{n+1})$ for all $n < \omega$. From this, it follows that the values $\{f(\rho_n) : n < \omega\}$ are pairwise comparable, and hence $\{\varphi(x; b_{\rho_n}) : n < \omega\} = \{\varphi(x; a_{f(\rho_n)}) : n < \omega\}$ is 2-inconsistent.

So fix $n < \omega$. Writing $\rho_n = \eta^\frown\langle i\rangle$, we have $\eta^\frown\langle j\rangle \trianglelefteq \rho_{n+1}$ for some $j > i$. Now $f(\rho_n) = e(\eta^\frown\langle i\rangle)^\frown\langle 1\rangle = e(\eta)^\frown\langle 0\rangle^\frown(1^{2i+1})$. On the other hand, $f(\rho_{n+1}) = e(\rho_{n+1})^\frown\langle 1\rangle$ has as an initial segment $e(\eta^\frown\langle j\rangle) = e(\eta)^\frown\langle 0\rangle^\frown(1^{2j})$. Since $2i+1 < 2j$, $f(\rho_n) \trianglelefteq f(\rho_{n+1})$, as desired.

Thus $\varphi(x; y)$ has 2-BTP. □

6. Questions

We have left open several natural directions for future work. In our view, the main problem is to find a syntactic characterization of the theories satisfying New Kim's Lemma. We have shown that NBTP implies New Kim's Lemma, but it is open whether this implication reverses. No implication in either direction is known between New Kim's Lemma and NATP. In light of Hanson's preprint [Hanson 2023], we are also interested in the relationship between New Kim's Lemma and the property NCTP explored there.

Question 6.1. Is New Kim's Lemma equivalent to one or more of the syntactic properties NATP, NBTP, or NCTP?

However, it is conceivable that there simply is no syntactic property that characterizes New Kim's Lemma. One way of making this precise is to recall the following very general definition, due to Shelah.

Definition 6.2 [Shelah 2000, Definition 5.17].

- For $n < \omega$, an *n-code* (for a partial type) is a pair $A = (A_+, A_-)$ of disjoint subsets of $[n] = \{0, \dots, n-1\}$. Given a formula $\varphi(x; y)$ and tuples $a_0, \dots, a_{n-1} \in \mathbb{M}^y$, the partial type coded by $A = (A_+, A_-)$ is

$$q_A(x) = \{\varphi(x; a_i) : i \in A_+\} \cup \{\neg\varphi(x; a_i) : i \in A_-\}.$$

- For $n < \omega$, an *n-pattern* (of consistency and inconsistency) is a pair (C, I) of disjoint sets of n-codes. A *finite pattern* is an n-pattern for some $n < \omega$. We say that a formula $\varphi(x; y)$ *exhibits* the n-pattern (C, I) if there are tuples $a_0, \dots, a_{n-1} \in \mathbb{M}^y$ such that for every code $A \in C$, $q_A(x)$ is consistent, and for every code $A \in I$, $q_A(x)$ is inconsistent.
- A property of formulas P is *definable by patterns* if there is a set $\mathcal{F}$ of finite patterns such that $\varphi(x; y)$ has property P if and only if $\varphi(x; y)$ exhibits every pattern in $\mathcal{F}$.
- A property Q of theories is *definable by patterns* if there is a property P of formulas which is defined by patterns, and T has property Q if and only if there is some formula $\varphi(x; y)$ which has property P.[7]

[7] Shelah calls a property of theories which is definable by patterns "weakly simply high straight". This is a special case of a related notion that Shelah calls "straightly defined".

Each of the properties TP, TP_1, TP_2, SOP_1, ATP, and BTP considered in this paper are definable by patterns: let $\mathcal{F}$ consist of one pattern for each finite subset of the infinite pattern of consistency and inconsistency defining the property, and apply compactness.

Question 6.3. Is the class of theories in which New Kim's Lemma fails definable by patterns?

It would be nice to have a larger stock of examples of theories satisfying New Kim's Lemma. To this end, we would like it to be easier to check that New Kim's Lemma holds, and to have more constructions for producing theories satisfying New Kim's Lemma.

Question 6.4. Does it suffice to show that New Kim's Lemma holds for formulas in a single free variable to establish that it holds for all formulas?

The analogous fact is known for each of the properties NTP, NTP_1, NTP_2, $NSOP_1$, and NATP: to prove that a theory has one of these properties, it suffices to check that no formula $\varphi(x; y)$ has the corresponding property, where x is a single variable. These arguments typically push against the syntactic definition of the property, so it is hard to envision what a solution to this question might look like without first resolving Question 6.3. In light of this, it makes sense to ask Question 6.4 with New Kim's Lemma replaced by NBTP.

The theory DLO_p examined in Section 4A is a special case of a general construction, developed in [Chernikov and Ramsey 2016], for "parametrizing" arbitrary Fraïssé limits with disjoint amalgamation. As shown in [loc. cit., Corollary 6.3], the parametrization of a Fraïssé limit with a simple theory is always $NSOP_1$. It seems likely that the arguments in Section 4A generalize to provide a positive answer to the following question.

Question 6.5. Suppose $\mathcal{K}$ is a Fraïssé class with disjoint amalgamation, and let $\mathcal{K}_{\text{pfc}}$ be the parametrized version of $\mathcal{K}$, as defined in [loc. cit., Section 6.3]. Let T and T_{pfc} be the theories of the Fraïssé limits of $\mathcal{K}$ and $\mathcal{K}_{\text{pfc}}$, respectively. If T satisfies New Kim's Lemma (or if T is NTP_2), does T_{pfc} satisfy New Kim's Lemma?

There is a theme in the literature that "generic constructions" (i.e., those involving taking a model companion) often produce properly $NSOP_1$ theories. For example, interpolative fusion, introduced in [Kruckman et al. 2021], is a general method for "generically putting together" multiple theories over a common reduct. Tran, Walsberg, and Kruckman [Kruckman et al. 2022] showed that the interpolative fusion of stable theories over a stable base theory is always $NSOP_1$ (and, under mild hypotheses, the interpolative fusion of $NSOP_1$ theories over a stable base theory is always $NSOP_1$).

If theories satisfying New Kim's Lemma are to generalize NSOP_1 theories in an analogous way to how NTP_2 theories generalize simple theories, and how NIP theories generalize stable theories, then the following seems like a reasonable conjecture.

Question 6.6. Does the interpolative fusion of NIP theories over a stable base theory always satisfy New Kim's Lemma?

Questions 6.5 and 6.6 are also meaningful with New Kim's Lemma replaced by NBTP.

Finally, since the Kim's Lemma surveyed in Section 3 form the cornerstones of the theories of independence in simple, NSOP_1, and NTP_2 theories, one might hope that a satisfying theory of Kim-independence, generalizing the theory of $\mathop{\smile\!\!\!\!\!\!|}^f$ in NTP_2 theories and of $\mathop{\smile\!\!\!\!\!\!|}^K$ in NSOP_1 theories, could be developed on the basis of New Kim's Lemma. A natural first step would be the chain condition.

Definition 6.7. We say $\mathop{\smile\!\!\!\!\!\!|}^K$ satisfies the *chain condition over models* if whenever $M \models T$, $a \mathop{\smile\!\!\!\!\!\!|}^K_M b$, and $I = (b_i)_{i<\omega}$ is a Morley sequence for a global M-invariant type extending $\mathrm{tp}(b/M)$, there exists a' such that $a'b_i \equiv_M ab$ for all $i < \omega$, I is Ma'-indiscernible, and $a' \mathop{\smile\!\!\!\!\!\!|}^K_M I$.

Question 6.8. If T satisfies New Kim's Lemma, does $\mathop{\smile\!\!\!\!\!\!|}^K$ satisfy the chain condition over models?

One motivation for this question is that $\mathop{\smile\!\!\!\!\!\!|}^f$ satisfies the chain condition over models in NTP_2 theories, see [Ben Yaacov and Chernikov 2014, Theorem 2.9] (and the chain condition is the key step in the proof of the variant of the independence theorem for NTP_2 theories in that paper). The proof of the chain condition in [loc. cit.] uses both the Kim's Lemma for NTP_2 theories and the syntactic definition of NTP_2. So here again, if Question 6.8 has a positive answer, it may be necessary to first resolve Question 6.3.

References

[Adler 2014] H. Adler, "Kim's lemma for NTP_2 theories: a simpler proof of a result by Chernikov and Kaplan", *Rend. Semin. Mat. Univ. Politec. Torino* **72**:3-4 (2014), 121–126. MR Zbl

[Ahn and Kim 2024] J. Ahn and J. Kim, "SOP_1, SOP_2, and antichain tree property", *Ann. Pure Appl. Logic* **175**:3 (2024), Paper No. 103402, 26. MR Zbl

[Ahn et al. 2023] J. Ahn, J. Kim, and J. Lee, "On the antichain tree property", *J. Math. Log.* **23**:2 (2023), art. it. 2250021. MR Zbl

[Ben Yaacov and Chernikov 2014] I. Ben Yaacov and A. Chernikov, "An independence theorem for NTP_2 theories", *J. Symb. Log.* **79**:1 (2014), 135–153. MR Zbl

[Bossut 2023] Y. Bossut, "A note on some example of NSOP1 theories", preprint, 2023. arXiv 2304.05404

[Cherlin and Hrushovski 2003] G. Cherlin and E. Hrushovski, *Finite structures with few types*, Annals of Mathematics Studies **152**, Princeton University Press, 2003. MR Zbl

[Chernikov 2014] A. Chernikov, "Theories without the tree property of the second kind", *Ann. Pure Appl. Logic* **165**:2 (2014), 695–723. MR Zbl

[Chernikov and Kaplan 2012] A. Chernikov and I. Kaplan, "Forking and dividing in NTP_2 theories", *J. Symbolic Logic* **77**:1 (2012), 1–20. MR Zbl

[Chernikov and Ramsey 2016] A. Chernikov and N. Ramsey, "On model-theoretic tree properties", *J. Math. Log.* **16**:2 (2016), art. id. 1650009. MR Zbl

[Conant 2017] G. Conant, "Forking and dividing in Henson graphs", *Notre Dame J. Form. Log.* **58**:4 (2017), 555–566. MR Zbl

[Dobrowolski 2023] J. Dobrowolski, "Sets, groups, and fields definable in vector spaces with a bilinear form", *Ann. Inst. Fourier* (*Grenoble*) **73**:5 (2023), 1795–1841. MR Zbl

[Džamonja and Shelah 2004] M. Džamonja and S. Shelah, "On $\triangleleft^*$-maximality", *Ann. Pure Appl. Logic* **125**:1-3 (2004), 119–158. MR Zbl

[Granger 1999] N. Granger, *Stability*, *simplicity*, *and the model theory of bilinear forms*, Ph.D. thesis, University of Manchester, 1999.

[Hanson 2023] J. E. Hanson, "Bi-invariant types, reliably invariant types, and the comb tree property", preprint, 2023. arXiv 2306.08239

[Hrushovski 2002] E. Hrushovski, "Pseudo-finite fields and related structures", pp. 151–212 in *Model theory and applications*, edited by L. Bélair et al., Quad. Mat. **11**, Aracne, Rome, 2002. MR Zbl

[Kaplan and Ramsey 2020] I. Kaplan and N. Ramsey, "On Kim-independence", *J. Eur. Math. Soc.* (*JEMS*) **22**:5 (2020), 1423–1474. MR Zbl

[Kaplan et al. 2019] I. Kaplan, N. Ramsey, and S. Shelah, "Local character of Kim-independence", *Proc. Amer. Math. Soc.* **147**:4 (2019), 1719–1732. MR Zbl

[Kim 1998] B. Kim, "Forking in simple unstable theories", *J. London Math. Soc.* (2) **57**:2 (1998), 257–267. MR Zbl

[Kim 2001] B. Kim, "Simplicity, and stability in there", *J. Symbolic Logic* **66**:2 (2001), 822–836. MR Zbl

[Kim 2009] B. Kim, "NTP_1 theories", Slides, Stability Theoretic Methods in Unstable Theories, BIRS, 2009, available at https://stats.birs.ca/workshops/2009/09w5113/files/Kim.pdf.

[Kim and Kim 2011] B. Kim and H.-J. Kim, "Notions around tree property 1", *Ann. Pure Appl. Logic* **162**:9 (2011), 698–709. MR Zbl

[Kim and Lee 2023] J. Kim and H. Lee, "Some remarks on Kim-dividing in NATP theories", preprint, 2023. arXiv 2211.04213

[Kim and Pillay 1997] B. Kim and A. Pillay, "Simple theories", *Ann. Pure Appl. Logic* **88**:2-3 (1997), 149–164. MR Zbl

[Kruckman et al. 2021] A. Kruckman, C.-M. Tran, and E. Walsberg, "Interpolative fusions", *J. Math. Log.* **21**:2 (2021), art. id. 2150010. MR Zbl

[Kruckman et al. 2022] A. Kruckman, M. C. Tran, and E. Walsberg, "Interpolative fusions II: Preservation results", preprint, 2022. arXiv 2201.03534

[Mutchnik 2022a] S. Mutchnik, "Conant-independence and generalized free amalgamation", preprint, 2022. arXiv 2210.07527

[Mutchnik 2022b] S. Mutchnik, "On NSOP_2 theories", preprint, 2022. arXiv 2206.08512

[Shelah 1980] S. Shelah, "Simple unstable theories", *Ann. Math. Logic* **19**:3 (1980), 177–203. MR Zbl

[Shelah 1990] S. Shelah, *Classification theory and the number of nonisomorphic models*, 2nd ed., Studies in Logic and the Foundations of Mathematics **92**, North-Holland, Amsterdam, 1990. MR Zbl

[Shelah 2000] S. Shelah, "On what I do not understand (and have something to say), model theory", *Math. Japon.* **51**:2 (2000), 329–377. MR Zbl

[Simon 2015] P. Simon, *A guide to NIP theories*, Lecture Notes in Logic **44**, Association for Symbolic Logic, Chicago, IL, 2015. MR Zbl

Received 16 Sep 2023. Revised 2 Apr 2024.

ALEX KRUCKMAN:

akruckman@wesleyan.edu

Department of Mathematics and Computer Science, Wesleyan University, Middletown, CT, United States

NICHOLAS RAMSEY:

sramsey5@nd.edu

Department of Mathematics, University of Notre Dame, Notre Dame, IN, United States

Model Theory
Vol. 3, No. 3, 2024
https://doi.org/10.2140/mt.2024.3.861

Sparse graphs and the fixed points on type spaces property

Rob Sullivan

We examine the topological dynamics of the automorphism groups of ω-categorical sparse graphs resulting from Hrushovski constructions. Specifically, we consider the fixed points on type spaces property, which a structure M has if, for all $n \in \mathbb{N}$, every $\mathrm{Aut}(M)$-subflow of the space $S_n(M)$ of n-types has a fixed point. Extending a result of Evans, Hubička and Nešetřil, we show that there exists an ω-categorical structure M, resulting from a Hrushovski construction, such that no ω-categorical expansion of M has the fixed points on type spaces property.

1. Introduction

The paper [Evans et al. 2019] is concerned with the topological dynamics of the automorphism groups of sparse graphs, in the context of the Kechris–Pestov–Todorčević correspondence [Kechris et al. 2005]. One of the key results of [Evans et al. 2019] is the following:

Theorem [Evans et al. 2019, Theorem 1.2]. *There exists an ω-categorical structure M such that no ω-categorical expansion has an extremely amenable automorphism group.*

We recall that, for a Hausdorff topological group G, a *G-flow* is a continuous action of G on a nonempty compact Hausdorff space X, and we say that G is *extremely amenable* if every G-flow has a G-fixed point.

In this paper, we show that the above result holds even in the context of a more restricted class of flows: subflows of type spaces. Let M be a relational structure. Following [Meir and Sullivan 2023], we say that M has the *fixed points on type spaces property* (FPT), if, for each $n \in \mathbb{N}_+$, every subflow of $S_n(M)$ has an $\mathrm{Aut}(M)$-fixed point, where $S_n(M)$ denotes the Stone space of n-types with parameters from M and the action is given by translation of parameters in formulae. This property is studied in depth in [loc. cit.], and may be thought of as a restriction of extreme amenability to a subclass of flows which occur naturally in a model-theoretic context.

This project formed the first part of the PhD of the author at Imperial College London, under the supervision of professor David Evans.
MSC2020: 03C15, 05C55, 05D10, 20B27, 37B05.
Keywords: sparse graphs, Hrushovski constructions, omega-categorical, type spaces, orientations.

The main result of this paper is as follows.

Theorem 5.1. *There is an ω-categorical structure M such that no ω-categorical expansion has FPT, the fixed points on type spaces property.*

The structure M appearing in both these results is a particular type of ω-categorical sparse graph known as an ω-categorical Hrushovski construction (first introduced in [Hrushovski 1988] — a clear introductory exposition may be found in [Evans 2013]). A graph A is *k-sparse* if for all finite $B \subseteq A$, the number of edges of B is at most k times the number of vertices of B.

The proof strategy for Theorem 5.1 is as follows. A central fact in the analysis of sparse graphs is that a graph is k-sparse if and only if it is k-orientable: its edges may be directed so that each vertex has at most k out-edges. This fact is well known to graph theorists [Nash-Williams 1964], and the proof is by Hall's marriage theorem (see Proposition 2.6).

For any k-sparse graph M, the space $\mathrm{Or}(M)$ of k-orientations of M (with the subspace topology from 2^{M^2}) gives an $\mathrm{Aut}(M)$-flow (Lemma 2.8). As in [Evans et al. 2019], we specialize to the case $k = 2$ (results generalize straightforwardly to any k). Theorem 1.2 of [loc. cit.], the result of Evans, Hubička and Nešetřil mentioned above, then immediately results from the following, using the Ryll–Nardzewski theorem:

Proposition 5.2 (adapted from [Evans et al. 2019, Theorem 3.7]). *Let M be an infinite 2-sparse graph in which all vertices have infinite degree. Let $G = \mathrm{Aut}(M)$.*

Consider the G-flow $G \curvearrowright \mathrm{Or}(M)$. If $H \leq G$ fixes a 2-orientation of M, then H has infinitely many orbits on M^2.

To prove Theorem 5.1, we also use the above result. Letting M be the ω-categorical Hrushovski construction detailed in Section 3, we define a notion of when a 1-type *encodes* an orientation of M. We then define an $\mathrm{Aut}(M)$-flow morphism $u : S_1(M) \to 2^{M^2}$ which sends each orientation-encoding 1-type to the orientation it encodes. Let M' be an expansion of M with FPT, and let $H = \mathrm{Aut}(M')$. Then H must fix a point in the subflow of orientation-encoding 1-types, so fixes an orientation. We then use Proposition 5.2 to see that H has infinitely many orbits on M^2, so H is not oligmorphic, and therefore by the Ryll–Nardzewski theorem we see that M' is not ω-categorical. Thus M has no ω-categorical expansion with FPT.

2. Background

In this section, we present the sufficient background material on topological dynamics, sparse graphs and Fraïssé classes with distinguished substructures ("strong Fraïssé classes") in order to be able to construct the ω-categorical examples of sparse graphs (ω-categorical Hrushovski constructions) given in Section 3.

We assume that the reader is familiar with the classical Fraïssé theory, the pointwise convergence topology on automorphism groups of first-order structures and the Ryll–Nardzewski theorem. (The background for these three topics can be found in [Hodges 1993, Chapter 7] and [Evans 2013, Sections 1–2].)

The background material in this section has been mostly adapted from [Evans et al. 2019] and [Evans 2013].

All first-order languages considered in this article will be countable and relational.

2A. ***Topological dynamics.*** A central object of study in topological dynamics is the following (see [Auslander 1988] for a thorough background):

Definition 2.1. A *G-flow* is a continuous action $G \curvearrowright X$ of a Hausdorff topological group G on a nonempty compact Hausdorff topological space X.

We will often simply write X to refer to the G-flow $G \curvearrowright X$ when this is clear from context. Given a G-flow on X, $\overline{G \cdot x}$, the orbit closure of a point $x \in X$, is a G-invariant compact subset of X. In general, a nonempty compact G-invariant subset $Y \subseteq X$ defines a *subflow* by restricting the G-action to Y.

Let X, Y be G-flows. A *G-flow morphism* $X \to Y$ is a continuous map $\alpha : X \to Y$ such that $\alpha(g \cdot x) = g \cdot \alpha(x)$ (this property is called *G-equivariance*). A surjective G-flow morphism $X \to Y$ is called a *factor* of X, and we will also say that Y is a factor of X when the morphism is contextually implied. Bijective G-flow morphisms are isomorphisms, as they are between compact Hausdorff spaces.

2B. ***Graphs.*** We work with graphs in first-order logic as follows. Let $\mathcal{L}$ be a first-order language consisting of a single binary relation symbol E. A *graph* consists of an $\mathcal{L}$-structure (A, E^A) where the binary relation $E^A \subseteq A^2$ is symmetric and irreflexive. We call A the *vertex set*, and write E_A for the set of unordered pairs $\{a, b\}$ such that $(a, b) \in E^A$. We call E_A the *edge set*, and this will usually be the relevant set we work with in this paper, following the usual graph-theoretic definition of a graph — rather than the symmetric set E^A of ordered pairs, which we only introduce for the sake of first-order structure formalism. We will usually just write A to denote the graph (A, E^A) when this is clear from context. We will often write $\sim$ instead of E in formulae to indicate adjacency.

By the above definition, here we only work with *simple* graphs: graphs having no loops on a single vertex or multiple edges between two vertices.

Definition 2.2. Let (A, E^A) be a graph. A set $\rho^A \subseteq A^2$ is an *orientation* of (A, E^A) if

- $\rho^A \subseteq E^A$;
- for each $(x, y) \in E^A$, exactly one of $(x, y), (y, x)$ is in ρ^A.

We may visualize the above definition as follows: an orientation of a graph consists of a direction for each edge.

Note that the above definition implies that ρ^A contains no directed loops or directed 2-cycles. We will refer to (A, E^A, ρ^A) as an *oriented graph*.

Definition 2.3. Let (A, E^A, ρ^A) be an oriented graph.

If $(x, y) \in \rho^A$, we refer to (x, y) as an *out-edge* of x and as an *in-edge* of y. We call y an *out-vertex* of x, and x an *in-vertex* of y.

The *out-neighborhood* $\mathrm{N}_+(x)$ of x consists of the out-vertices of x. The *in-neighborhood* $\mathrm{N}_-(x)$ of x consists of the in-vertices of x. The *out-degree* $\mathrm{d}_+(x)$ of x is defined to be $\mathrm{d}_+(x) = |\mathrm{N}_+(x)|$, and the *in-degree* $\mathrm{d}_-(x)$ of x is defined to be $\mathrm{d}_-(x) = |\mathrm{N}_-(x)|$.

When we refer to a subgraph of a graph, or an oriented subgraph of an oriented graph, we mean a substructure in the model-theoretic sense. For graph theorists, these substructures would usually be referred to as *induced* subgraphs.

(We use the full notation for structures in this section for clarity, but henceforth we will usually denote graphs (A, E^A) by A, and oriented graphs (A, E^A, ρ^A) by A or (A, ρ^A).)

2C. *Sparse graphs.*

Definition 2.4. Let $k \in \mathbb{N}_+$. A graph A is *k-sparse* if for all $B \subseteq_{\text{fin.}} A$, we have $|E_B| \leq k|B|$.

Definition 2.5. Let (A, ρ^A) be an oriented graph. Let $k \in \mathbb{N}_+$. We call ρ^A a *k-orientation* if for $x \in A$, we have $\mathrm{d}_+(x) \leq k$. We refer to (A, ρ^A) as a *k-oriented graph*.

If an undirected graph A has a k-orientation, we say it is *k-orientable*.

The following proposition is well-known to graph theorists [Nash-Williams 1964], and will be a key tool here. We present the proof as it is relatively brief.

Proposition 2.6 [Evans et al. 2019, Theorem 3.4]. *Let A be a countable graph. Then A is k-orientable if and only if it is k-sparse.*

Proof. $\Rightarrow$: straightforward.

$\Leftarrow$: We prove the statement for finite A, and then the statement for countably infinite A follows by a straightforward Kőnig's lemma argument. We wish to produce a k-orientation of A, and to do this we must direct each edge. We will use Hall's marriage theorem [Bollobás 1998, III.3], which for the convenience of the reader we briefly restate: for a finite bipartite graph B with left set X and right set Y, there is an X-saturated matching if and only if $|W| \leq |\mathrm{N}_B(W)|$ for $W \subseteq X$. (Here $\mathrm{N}_B(W)$ denotes the neighborhood of W in B.)

Form a bipartite graph B with left set E_A and right set $A \times [k]$, and place an edge between $e \in E_A$ and $(x, i) \in A \times [k]$ if $x \in e$. Given a left-saturated matching, if e is matched to (x, i), we orient e outwards from x, and this gives a k-orientation of A.

To see that a left-saturated matching exists, take $W \subseteq E_A$. Let V be the set of vertices of the edges which lie in W. Then $|\mathrm{N}_B(W)| = k|V|$, and as A is k-sparse, we have that $k|V| \geq |\mathrm{E}_A(V)|$, where $\mathrm{E}_A(V)$ is the set of edges in A whose vertices lie in V. As $|\mathrm{E}_A(V)| \geq |W|$, by Hall's marriage theorem there exists a left-saturated matching of the bipartite graph B. □

For presentational simplicity, we will work with $k = 2$. Our results generalize straightforwardly for $k > 2$.

Note. In this paper, we may occasionally say "oriented graph" to in fact mean "2-oriented graph". We will try to avoid this in general, but when this does occur the meaning will be clear from context.

Definition 2.7. Let M be a 2-sparse graph. We let $\mathrm{Or}(M) \subseteq 2^{M^2}$ denote the topological space of 2-orientations of M, where the topology is given by the subspace topology from the Cantor space 2^{M^2}.

Lemma 2.8. *Let M be a 2-sparse graph. Then $\mathrm{Or}(M)$ is an $\mathrm{Aut}(M)$-flow with the natural action*

$$g \cdot \rho = \{(gx, gy) : (x, y) \in \rho\}.$$

Proof. By Proposition 2.6, we see that $\mathrm{Or}(M)$ is nonempty, and it is immediate that $\mathrm{Or}(M)$ is $\mathrm{Aut}(M)$-invariant. It therefore remains to show that $\mathrm{Or}(M)$ is closed in 2^{M^2}: if $\sigma \in 2^{M^2}$ is not a 2-orientation, then this is witnessed on a finite set, so $2^{M^2} \setminus \mathrm{Or}(M)$ is open. □

2D. *Graph predimension.* One way to characterize 2-sparsity is in terms of a particular notion of *graph predimension*.

Definition 2.9. Let A be a finite graph. We define the *predimension* $\delta(A)$ of A to be $\delta(A) = 2|A| - |E_A|$.

For $B \subseteq A$, we define the *relative predimension of A over B* to be $\delta(A/B) = \delta(A) - \delta(B)$.

We immediately see that, for A a finite graph, A is 2-sparse if and only if for all $B \subseteq A$ we have $\delta(B) \geq 0$.

2E. *Strong classes.* For the ω-categorical Hrushovski constructions in Section 3, we will need to take a class of sparse graphs where we only consider particular distinguished embeddings between structures in the class, and for this we require the definition below. In the subsequent section, we will construct Fraïssé classes where we only permit these distinguished embeddings between finite structures.

Definition 2.10. Let $\mathcal{K}$ be a class of finite $\mathcal{L}$-structures closed under isomorphisms. Let $\mathcal{S} \subseteq \mathrm{Emb}(\mathcal{K})$ be a class of embeddings between structures in $\mathcal{K}$ satisfying that

(S1) $\mathcal{S}$ contains all isomorphisms;

(S2) $\mathcal{S}$ is closed under composition;

(S3) if $f : A \to C$ is in $\mathcal{S}$ and $f(A) \subseteq B \subseteq C$ with $B \in \mathcal{K}$, then $f : A \to B$ is in $\mathcal{S}$.

Then we call $(\mathcal{K}, \mathcal{S})$ a *strong class*, and call the elements of $\mathcal{S}$ *strong embeddings*. (This is originally due to Hrushovski [1988]. An accessible exposition of strong classes is in [Evans 2013, Section 3].)

If $A, B \in \mathcal{K}$, $A \subseteq B$ and the inclusion map $\iota : A \hookrightarrow B$ is in $\mathcal{S}$, then we write $A \leq B$ and say A is a *strong substructure* of B. We then have that

(L1) $\leq$ is reflexive;

(L2) $\leq$ is transitive;

(L3) if $A \leq C$ and $A \subseteq B \subseteq C$ with $B \in \mathcal{K}$, then $A \leq B$.

We will often write $(\mathcal{K}, \leq)$ instead of $(\mathcal{K}, \mathcal{S})$, and we will refer to the elements of $\mathcal{S}$ as $\leq$*-embeddings*.

If $(\mathcal{K}, \leq)$ is a strong class (i.e., $\mathcal{S}$ satisfies (S1), (S2), (S3)), then we have that for $f : A \to B$ in $\mathcal{S}$, if $X \leq A$, then $f(X) \leq B$.

Definition 2.11. Suppose $(\mathcal{K}, \leq)$ is a strong class. Let $A_0 \leq A_1 \leq \cdots$ be an increasing $\leq$-chain of structures in $\mathcal{K}$, and let $M = \bigcup_{i \in \mathbb{N}} A_i$. Let $A \subseteq_{\text{fin.}} M$.

Then we write $A \leq M$, and say that A is a *strong substructure* of M, or that A is $\leq$*-closed* in M, to mean that there is some A_i $(i \in \mathbb{N})$ with $A \leq A_i$.

Given $A \in \mathcal{K}$ and an embedding $f : A \to M$, we will likewise say that f is a $\leq$-embedding if $f(A) \leq M$.

The above definition is independent of the choice of $\leq$-chain. To see this, suppose M is also the union of the elements of an increasing $\leq$-chain $B_0 \leq B_1 \leq \cdots$ of $\mathcal{K}$-structures. Take any A_i $(i \in \mathbb{N})$. Then $A_i \subseteq B_j$ for some $j \in \mathbb{N}$, and $B_j \subseteq A_k$ for some $k \geq i$. As $A_i \leq A_k$, by (L3) we have $A_i \leq B_j$.

Let $g \in \mathrm{Aut}(M)$. Take a pair $A_i \leq A_j$ $(i < j)$. Then $g|_{A_j} : A_j \to gA_j$ is an isomorphism, so $g|_{A_j} \in \mathcal{S}$, and so $gA_i \leq gA_j$. Thus M is also the union of the increasing $\leq$-chain $gA_0 \leq gA_1 \leq \cdots$. So if $A \leq M$, then $gA \leq M$: that is, all $g \in \mathrm{Aut}(M)$ preserve $\leq$.

2F. *Fraïssé theory for strong classes.* We now develop an analogue of the classical Fraïssé theory for strong classes. We omit the proofs and state the relevant material as a series of definitions and lemmas. (For the classical Fraïssé theory, originally developed in [Fraïssé 1954], see [Hodges 1993, Chapter 7], and for a more complete treatment of Fraïssé theory for strong classes, see [Evans 2013, Section 3].)

Definition 2.12. Let $(\mathcal{K}, \leq)$ be a strong class of $\mathcal{L}$-structures.

- $(\mathcal{K}, \leq)$ has the *joint embedding property* (JEP) if for $A_0, A_1 \in \mathcal{K}$, there is $B \in \mathcal{K}$ with $\leq$-embeddings $f_0 : A_0 \to B$, $f_1 : A_1 \to B$.
- $(\mathcal{K}, \leq)$ has the *amalgamation property* (AP) if, for any pair of $\leq$-embeddings $B_0 \xleftarrow{f_0} A \xrightarrow{f_1} B_1$, there exists $C \in \mathcal{K}$ and a pair of $\leq$-embeddings $B_0 \xrightarrow{g_0} C \xleftarrow{g_1} B_1$ such that $g_0 \circ f_0 = g_1 \circ f_1$.
- For $A, B_0, B_1 \in \mathcal{K}$ with $A \leq B_0, B_1$, the *free amalgam* C of B_0, B_1 over A is the $\mathcal{L}$-structure C whose domain is the disjoint union of B_0, B_1 over A and whose relations R^C are exactly the unions $R^{B_0} \cup R^{B_1}$ of the relations R^{B_0}, R^{B_1} on B_0, B_1 (for R a relation symbol in $\mathcal{L}$). If for all $\mathcal{L}$-structures $A, B_0, B_1 \in \mathcal{K}$ with $A \leq B_0, B_1$ we have that the free amalgam C of B_0, B_1 over A is in $\mathcal{K}$ with $B_0, B_1 \leq C$, then we say that $(\mathcal{K}, \leq)$ is a *free amalgamation class*.

We will usually not mention the distinguished class of embeddings in our terminology, as it will be clear from context and the fact that we are working with strong classes. For instance, we say that $(\mathcal{K}, \leq)$ has the amalgamation property, even though perhaps more strictly we should say that $(\mathcal{K}, \leq)$ has the $\leq$-amalgamation property.

In the following definitions and lemmas, let $(\mathcal{K}, \leq)$ be a strong class, and let M be the union of an increasing $\leq$-chain $A_1 \leq A_2 \leq \cdots$ of finite structures in $(\mathcal{K}, \leq)$.

Definition 2.13. The $\leq$-*age* of M, written $\mathrm{Age}_{\leq}(M)$, is the class of $A \in \mathcal{K}$ such that there is a $\leq$-embedding $A \to M$.

The class $(\mathrm{Age}_{\leq}(M), \leq)$ is a $\leq$-hereditary strong subclass of $(\mathcal{K}, \leq)$, and it has the $\leq$-joint embedding property.

Definition 2.14. M has the $\leq$-*extension property* if for all $A, B \in \mathrm{Age}_{\leq}(M)$ and $\leq$-embeddings $f : A \to M$, $g : A \to B$, there exists a $\leq$-embedding $h : B \to M$ with $h \circ g = f$.

M is $\leq$-*ultrahomogeneous* if each isomorphism $f : A \to A'$ between strong substructures A, A' of M extends to an automorphism of M.

(Again, when it is clear from context, we will often omit the $\leq$- prefix and just say that M has the extension property or is ultrahomogeneous.)

Lemma 2.15. *Let M' also be a union of an increasing $\leq$-chain in $\mathcal{K}$. Suppose M, M' have the same $\leq$-age and both have the $\leq$-extension property. Then M, M' are isomorphic.*

Lemma 2.16. *M is $\leq$-ultrahomogeneous if and only if M has the $\leq$-extension property.*

Lemma 2.17. *Suppose M is $\leq$-ultrahomogeneous. Then the class* $(\mathrm{Age}_{\leq}(M), \leq)$ *has the amalgamation property.*

Definition 2.18. Let $(\mathcal{K}, \leq)$ be a strong class. We say that $(\mathcal{K}, \leq)$ is an *amalgamation class* or *Fraïssé class* if $(\mathcal{K} \leq)$ contains countably many isomorphism types, contains structures of arbitrarily large finite size, and has the joint embedding and amalgamation properties.

Theorem 2.19 (Fraïssé–Hrushovski). *Let $(\mathcal{K}, \leq)$ be an amalgamation class. Then there is a structure M which is a union of an increasing $\leq$-chain in $\mathcal{K}$ such that M is $\leq$-ultrahomogeneous and $\mathrm{Age}_{\leq}(M) = \mathcal{K}$, and M is unique up to isomorphism amongst structures with these properties.*

*We call this structure the **Fraïssé limit** or **generic structure** of $\mathcal{K}$.*

3. ω-categorical sparse graphs

The material in this section is based on [Evans et al. 2019] and the unpublished notes [Evans 2013], with some minor modifications, and constitutes further background required for Section 5.

We now construct an amalgamation class of sparse graphs whose Fraïssé limit is ω-categorical. Specifically, this will be a version of the ω-categorical Hrushovski construction M_F, first seen in [Hrushovski 1988]. We will do this by defining a notion of closure (i.e., a particular notion of strong substructure), d-closure, which will be uniformly bounded. The relevance of this can be seen in the lemma below.

Lemma 3.1 [Evans et al. 2019, Remark 2.8]. *Let $(\mathcal{K}, \leq)$ be an amalgamation class such that for each $n \in \mathbb{N}$, $(\mathcal{K}, \leq)$ has only finitely many isomorphism classes of structures of size n. Suppose there is a function $h : \mathbb{N} \to \mathbb{N}$ such that for $B \in \mathcal{K}$ and $A \subseteq B$ with $|A| \leq n$, there exists $A \subseteq C \leq B$ with $|C| \leq h(n)$.*

Then the Fraïssé limit M of $(\mathcal{K}, \leq)$ is ω-categorical.

(The function h will be a uniform bound on the size of $\leq$-closures.)

Proof. By the Ryll–Nardzewski theorem, it suffices to show that, for $n \geq 1$, $\mathrm{Aut}(M)$ has finitely many orbits on M^n. Take $n \geq 1$. As there are only finitely many isomorphism types of structures of size $\leq h(n)$ in $\mathcal{K}$ and M is $\leq$-ultrahomogeneous, we have that $\mathrm{Aut}(M)$ has finitely many orbits on $\{\bar{c} \in M^{h(n)} : \bar{c} \leq M\}$. We can extend any $\bar{a} \in M^n$ to an element of this set (note that in ordered tuples, we can have repeats of elements). If $\bar{a}, \bar{a}'$ are not in the same orbit, then nor will their extensions be, so we are done. □

Definition 3.2. Let $\mathcal{C}_{>0}$ be the class of finite graphs A such that for nonempty $B \subseteq A$, we have $\delta(B) > 0$.

We note that for $A \in \mathcal{C}_{>0}$, if $A' \subseteq A$ then $A' \in \mathcal{C}_{>0}$.

Definition 3.3. Take $A, B \in \mathcal{C}_{>0}$ with $A \subseteq B$. We say that A is d-closed in B, written $A \leq_d B$, if for all $A \subsetneq C \subseteq B$, we have $\delta(A) < \delta(C)$.

Lemma 3.4 (submodularity [Evans 2013, Lemma 3.7]). *Let A be a finite graph, and let $B, C \subseteq A$. Then we have that*

$$\delta(B \cup C) \leq \delta(B) + \delta(C) - \delta(B \cap C).$$

We have equality if and only if $E_{B \cup C} = E_B \cup E_C$, i.e., B, C are freely amalgamated over $B \cap C$ in A.

The proof of the above lemma is straightforward. We now prove some basic properties of $\leq_d$.

Lemma 3.5 [Evans 2013, Lemma 3.10]. *Let $B \in \mathcal{C}_{>0}$. Then*:

(1) $A \leq_d B$, $X \subseteq B \Rightarrow A \cap X \leq_d X$.

(2) $A \leq_d C \leq_d B \Rightarrow A \leq_d B$.

(3) $A_1, A_2 \leq_d B \Rightarrow A_1 \cap A_2 \leq_d B$.

Proof. (1) Take $A \cap X \subsetneq Y \subseteq X$. Note that $A \cap Y = A \cap X$. By submodularity,

$$\delta(A \cup Y) \leq \delta(A) + \delta(Y) - \delta(A \cap Y) = \delta(A) + \delta(Y) - \delta(A \cap X),$$

so $\delta(Y) - \delta(A \cap X) \geq \delta(A \cup Y) - \delta(A) > 0$, using the fact that $A \subsetneq A \cup Y \subseteq B$.

(2) We may assume $A \neq C$. Take $A \subsetneq X \subseteq B$. By (1) applied to $C \leq_d B$ and $X \subseteq B$, we have $C \cap X \leq_d X$. Also we have $A \subseteq C \cap X \subseteq C$. So, as $A \leq_d C$, we have $\delta(A) < \delta(X)$.

(3) By (1), $A_1 \cap A_2 \leq_d A_1$. Then use (2). □

For $B \in \mathcal{C}_{>0}$, by part (3) of the previous lemma we see that for $A \subseteq B$ we have that $\bigcap_{A \subseteq A' \leq_d B} A' \leq_d B$, so we can define the *d-closure* of A in B as this intersection, written $\mathrm{cl}^d_B(A)$.

Lemma 3.6 [Evans 2013, Lemma 3.12]. *Let $B \in \mathcal{C}_{>0}$ and let $A \subseteq B$. Then $\delta(A) \geq \delta(\mathrm{cl}^d_B(A))$.*

Proof. Amongst all $A \subseteq X \subseteq B$, consider those for which $\delta(X)$ is smallest, and then out of these choose a C of greatest size. By the first stage of selection, we have $\delta(C) \leq \delta(A)$, and by the second stage, if $C \subsetneq D \subseteq B$ then $\delta(C) < \delta(D)$, so $C \leq_d B$. So $\mathrm{cl}^d_B(A) \subseteq C \subseteq B$, and as $\mathrm{cl}^d_B(A) \leq_d B$, we have $\delta(\mathrm{cl}^d_B(A)) \leq \delta(C)$. □

Lemma 3.7 [Evans 2013, Lemma 3.15]. *$(\mathcal{C}_{>0}, \leq_d)$ is a free amalgamation class.*

Proof. It only remains to check the free amalgamation property (which implies JEP). We prove a stronger claim. Given $A, B_1, B_2 \in \mathcal{C}_{>0}$ such that $A \leq_d B_1$ and $A \subseteq B_2$, with $B_1, B_2 \subseteq E$, where E is the free amalgam of B_1, B_2 over A, we claim that $B_2 \leq_d E$. Once we have the claim, note that $\varnothing \leq_d B_2 \leq_d E$ implies that $E \in \mathcal{C}_{>0}$.

Take $B_2 \subsetneq X \subseteq E$. Then letting $Y = X \cap B_1$, we have $Y \supsetneq A$ and $X = B_2 \cup Y$, and X is the free amalgam of B_2, Y over A. So

$$\delta(X) = \delta(B_2 \cup Y) = \delta(B_2) + \delta(Y) - \delta(A),$$

and so, as $A \leq_d B_1$,

$$\delta(X) - \delta(B_2) = \delta(Y) - \delta(A) > 0. \qquad \square$$

The Fraïssé limit $M_{>0}$ of $(\mathcal{C}_{>0}, \leq_d)$ is not ω-categorical, as for $A \subseteq_{\text{fin.}} M_{>0}$, there is no uniform bound on $|\text{cl}^d(A)|$ in terms of $|A|$.

To construct ω-categorical examples, as mentioned at the start of this section, we consider subclasses of $\mathcal{C}_{>0}$ in which d-closure is uniformly bounded.

Definition 3.8. Let $F : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be a continuous, strictly increasing function with $F(0) = 0$ and $F(x) \to \infty$ as $x \to \infty$. We define

$$\mathcal{C}_F := \{B \in \mathcal{C}_{>0} : \delta(A) \geq F(|A|) \text{ for all } A \subseteq B\}.$$

Note that if $B \in \mathcal{C}_F$ and $C \subseteq B$, then $C \in \mathcal{C}_F$.

Lemma 3.9 [Evans 2013, Theorem 3.19; Evans et al. 2019, Theorem 4.14].

(1) *For $B \in \mathcal{C}_F$, $A \subseteq B$, we have $|\text{cl}^d_B(A)| \leq F^{-1}(2|A|)$.*

(2) *If $(\mathcal{C}_F, \leq_d)$ is an amalgamation class, then its Fraïssé limit M_F is ω-categorical.*

Proof. (1) From Lemma 3.6 and the fact that $\text{cl}^d_B(A) \in \mathcal{C}_F$, we have $F(|\text{cl}^d_B(A)|) \leq \delta(\text{cl}^d_B(A)) \leq \delta(A) \leq 2|A|$.

(2) This follows from Lemma 3.1. $\square$

Definition 3.10. Suppose that $(\mathcal{C}_F, \leq_d)$ is an amalgamation class, and write M_F for its Fraïssé limit.

For $A \subseteq M_F$ with A infinite, we say that $A \leq_d M_F$ if $A \cap X \leq_d X$ for all finite $X \subseteq M_F$.

(Note that if A is finite, then $A \leq_d M_F$ if and only if $A \cap X \leq_d X$ for all finite $X \subseteq M_F$, by part (1) of Lemma 3.5, so this definition is consistent.)

Similarly we define $\text{cl}^d_{M_F}(A)$ as the smallest $\leq_d$-closed subset of M_F containing A. (This is well-defined: intersections of $\leq_d$-closed subsets of M_F are $\leq_d$-closed, by part (2) of Lemma 3.5.)

Let A be a graph, possibly infinite, which is embeddable in M_F. We say that an embedding $f : A \to M_F$ is a *$\leq_d$-embedding* if $f(A) \leq_d M_F$.

We now describe a method for constructing the control function F to ensure that $(\mathcal{C}_F, \leq_d)$ is a free amalgamation class.

Lemma 3.11 (adapted from [Evans 2013, Example 3.20; Evans et al. 2019, Example 4.15]). *Let $n \in \mathbb{N}$. Let F be as in Definition 3.8, and assume additionally that*

- *F is piecewise smooth*;
- *its right derivative F' is decreasing*;
- *$F'(x) \leq 1/x$ for $x > n$*;
- *for A, B_1, $B_2 \in \mathcal{C}_F$ with $A \leq_d B_1, B_2$ and $|B_1| < n$, $|B_2| < n$, the free amalgam of B_1, B_2 over A lies in $\mathcal{C}_F$.*

Then $(\mathcal{C}_F, \leq_d)$ is a free amalgamation class.

Proof. Let $A, B_1, B_2 \in \mathcal{C}_F$, with $A \leq_d B_1, B_2$. We may assume $|B_1| \geq n$ and $|B_1| \geq |B_2|$. Let E be the free amalgam of B_1, B_2 over A. By Lemma 3.7, $E \in \mathcal{C}_{>0}$ and $B_1, B_2 \leq_d E$. We need to show that $E \in \mathcal{C}_F$. Assuming $E \neq B_1, B_2$, we have $A \neq B_1, B_2$. Suppose $X \subseteq E$: we need to show that $\delta(X) \geq F(|X|)$. As X is the free amalgam of $B_1 \cap X$, $B_2 \cap X$ over $A \cap X$ and as $A \cap X \leq_d B_i \cap X$, it suffices to check just for $X = E$.

We have that

$$\delta(E) = \delta(B_1) + \delta(B_2) - \delta(A) = \delta(B_1) + (|B_2| - |A|)\frac{\delta(B_2) - \delta(A)}{|B_2| - |A|}.$$

As $|B_1| \geq |B_2|$ and as $A \leq_d B_1$ with $A \neq B_1$, we have

$$\frac{\delta(B_2) - \delta(A)}{|B_2| - |A|} \geq \frac{1}{|B_1|}.$$

So

$$\delta(E) \geq \delta(B_1) + \frac{|B_2| - |A|}{|B_1|} \geq F(|B_1|) + \frac{|B_2| - |A|}{|B_1|},$$

and as the conditions on F ensure that $F(x + y) \leq F(x) + y/x$ for $x \geq n$, we have

$$\delta(E) \geq F(|B_1| + |B_2| - |A|) = F(|E|). \qquad \square$$

4. The fixed points on type spaces property (FPT)

The following is folklore:

Lemma 4.1. *Let M be an $\mathcal{L}$-structure, and let $G = \mathrm{Aut}(M)$ with the pointwise convergence topology. Then, for each $n \geq 1$, G acts continuously on the Stone space $S_n(M)$ of n-types with parameters in M, with the action given by*

$$g \cdot p(\bar{x}) = \{\phi(\bar{x}, g\bar{m}) : \phi(\bar{x}, \bar{m}) \in p(\bar{x})\}.$$

That is, $G \curvearrowright S_n(M)$ with the action defined above is a G-flow.

See [Meir and Sullivan 2023, Lemma 4.1] for a proof. (The proof is relatively straightforward and follows via a compactness argument.)

Note that we define the action of G on $\mathcal{L}(M)$-formulae as

$$g \cdot \phi(\bar{x}, \bar{m}) = \phi(\bar{x}, g\bar{m}).$$

Definition 4.2 [Meir and Sullivan 2023, Definition 4.2]. Let M be an $\mathcal{L}$-structure and let $G = \mathrm{Aut}(M)$. We say that M has the *fixed points on type spaces property* (FPT) if every subflow of $G \curvearrowright S_n(M)$, $n \geq 1$, has a fixed point.

Note that FPT is equivalent to every orbit closure $\overline{G \cdot p(\bar{x})}$ in $S_n(M)$ having a fixed point.

The below lemma will play a key role in the proof of Theorem 5.1. (See Section 2A for the definition of a factor.)

Lemma 4.3. *Let M be an $\mathcal{L}$-structure and let $G = \mathrm{Aut}(M)$. Suppose that M has FPT. Then every subflow of each factor of $G \curvearrowright S_n(M)$, $n \geq 1$, has a fixed point.*

The proof is straightforward.

5. An ω-categorical structure such that no ω-categorical expansion has FPT

We will now discuss the main result of this paper, which is new.

Theorem 5.1. *There is an ω-categorical structure M such that no ω-categorical expansion has FPT, the fixed points on type spaces property.*

The ω-categorical structure M in the above theorem will be a particular case of the 2-sparse graph M_F, the ω-categorical Hrushovski construction from Section 3.

The proof will depend on the following key result from [Evans et al. 2019]:

Proposition 5.2 [Evans et al. 2019, Theorem 3.7]. *Let M be an infinite 2-sparse graph in which all vertices have infinite degree. Let $G = \mathrm{Aut}(M)$, and let $H \leq G$.*

Consider the G-flow $G \curvearrowright \mathrm{Or}(M)$. If H fixes a 2-orientation of M, then H has infinitely many orbits on M^2.

Before giving the details of the proof of Theorem 5.1, we first give an informal general outline.

The informal overview of the proof is as follows. Let $G = \mathrm{Aut}(M)$. For each orientation $\tau \in \mathrm{Or}(M)$, we will define a notion of when a 1-type $p(x) \in S_1(M)$ *encodes* τ (see Figure 1), and we will have an associated "decoding" G-morphism $u : S_1(M) \to 2^{M^2}$ sending each orientation-encoding 1-type back to the orientation it encodes. We will show that each orientation has a 1-type encoding it, and thus u contains $\mathrm{Or}(M)$ in its image. Now, let M' be an expansion of M with FPT, and let H denote its automorphism group. There is an H-factor map $w : S_1(M') \to S_1(M)$ given by the restriction map, and so composing with u we see that $\mathrm{Or}(M)$ is an H-subflow of a factor of $S_1(M')$. Thus, by Lemma 4.3, as M' has FPT, H fixes an orientation of M. Therefore, by Proposition 5.2, H will have infinitely many orbits on M^2, and thus by the Ryll–Nardzewski theorem M' cannot be ω-categorical.

We now start the formal details of the proof of Theorem 5.1, which proceeds in three parts.

Part 1: specify the control function. We begin with a description of the control function F and properties of the class $\mathcal{C}_F$ used to produce the structure M for Theorem 5.1. It will become clearer in later steps why we take control functions satisfying the below conditions.

Lemma 5.3. *Let F be a control function for the class $\mathcal{C}_F$ satisfying the conditions of Definition 3.8, and additionally assume that*

- *F is piecewise smooth, and its right derivative $F'(x)$ is decreasing*;
- $F(1) = 2$, $F(2) = 3$;
- $F'(x) \leq 2/(8x+1)$ *for $x \geq 2$, where F' denotes the right derivative.*

Then

(1) *$\mathcal{C}_F$ contains a point and an edge, and points and edges are d-closed in elements of $\mathcal{C}_F$*;

(2) *$(\mathcal{C}_F, \leq_d)$ is a free amalgamation class*;

(3) *each vertex of M_F has infinite degree (where M_F is the Fraïssé limit of $(\mathcal{C}_F, \leq_d)$)*;

(4) *if $a_0a_1 \cdots a_{n-1}$ is a path, then $a_0a_1 \cdots a_{n-1} \in \mathcal{C}_F$*;

(5) $F(4) < 4$, $F(5) < 4$, $F(6) < 4$;

(6) *if $abcd$ is a 4-cycle, then $abcd \in \mathcal{C}_F$.*

Proof. (1) As $F(1) = 2$, if x is a point then $\delta(\{x\}) = 2 = F(|\{x\}|)$, so $\{x\} \in \mathcal{C}_F$. If ab is an edge, then $\delta(ab) = 3 = F(2)$, so $ab \in \mathcal{C}_F$. As F is strictly increasing, points and edges are d-closed in elements of $\mathcal{C}_F$.

(2) Take $A, B_1, B_2 \in \mathcal{C}_F$ with $A \leq_d B_1, B_2$. Then as $F'(x) \leq 2/(8x+1) < 1/x$ for $x \geq 2$, by Lemma 3.11 we need only check the case $|B_1|, |B_2| \leq 1$, and the only nontrivial case is where $A = \varnothing$: if b_1, b_2 are nonadjacent points then $\delta(\{b_1, b_2\}) = 4 > F(2)$. So $(\mathcal{C}_F, \leq_d)$ is a free amalgamation class.

(3) Let $k \geq 1$. Let $ax \in \mathcal{C}_F$ be an edge. The point a is d-closed in ax, and so by taking the free amalgamation of k copies $ax_1, \ldots, ax_k$ of ax over a, we have that the star graph S_k is in $\mathcal{C}_F$ (where S_k is the complete bipartite graph $K_{1,k}$). Using the $\leq_d$-extension property of M_F, this implies that each vertex of M_F has infinite degree.

(4) Proceed by induction, and obtain $a_0 \cdots a_{n-1} \in \mathcal{C}_F$ by the free amalgamation of $a_0 \cdots a_{n-2}$ and $a_{n-2}a_{n-1}$ over a_{n-2}.

(5) F is strictly increasing, and so it suffices to show $F(6) < 4$. $F(6) \leq F(2) + \int_2^6 2/(8x+1)\,\mathrm{d}x = 3 + \frac{1}{4}\log(49) - \frac{1}{4}\log(17) < 4$.

(6) Let $abcd \subseteq M_F$ be a 4-cycle. Then $\delta(abcd) = 4 > F(4)$. For $C \subsetneq abcd$, C either consists of a path of length 2, an edge, two nonadjacent points or a single point. All of these lie in $\mathcal{C}_F$. □

Throughout the rest of the proof of Theorem 5.1 in Section 5, we will assume F is a control function satisfying the conditions of Lemma 5.3, and we write $M = M_F$. The first three conditions of the above lemma are relatively standard; the fourth condition is the one that is particularly specific to our example, constituting a mild additional restriction on F.

Note that control functions satisfying the conditions of the above lemma do indeed exist: for example, take F piecewise linear with $F(0) = 0$, $F(1) = 2$, $F(2) = 3$, and then for $x \geq 2$ define $F(x) = \frac{1}{4}\log(8x+1) + 3 - \frac{1}{4}\log(17)$.

Part 2: types encoding orientations, the encoding lemma, and its use in proving the main theorem. Given an orientation $\tau \in \mathrm{Or}(M)$, we will define a particular notion of when a 1-type *encodes* τ.

We write $\mathcal{L}$ for the language of graphs (so M is an $\mathcal{L}$-structure).

Definition 5.4. For $a, b \in M$, we define the *label formula* $f(x, a, b)$ in the language $\mathcal{L}(M)$ (with constants from M and free variable x) to be:

$$f(x,a,b) \equiv (x \neq a \wedge x \neq b \wedge a \neq b \wedge a \sim b) \wedge$$
$$(\exists l_1, l_2, l_3, l_4)\Bigg(\bigg(\bigwedge_{i<j} l_i \neq l_j\bigg) \wedge \bigg(\bigwedge_i l_i \neq x \wedge l_i \neq a \wedge l_i \neq b\bigg) \wedge$$
$$(x \sim l_1 \wedge l_1 \sim l_2 \wedge l_2 \sim l_3 \wedge l_3 \sim l_4 \wedge l_4 \sim l_1 \wedge l_2 \sim a \wedge l_4 \sim a \wedge l_3 \sim b)\Bigg).$$

(See Figure 1.)

Let $\tau \in \mathrm{Or}(M)$, and let $p(x) \in S_1(M)$. We say that $p(x)$ *encodes* τ if $p(x)$ contains the following set of formulae:

$$\{f(x,a,b) : (a,b) \in \tau\} \cup \{\neg f(x,a,b) : (a,b) \in M^2 \setminus \tau\}.$$

Informally, $p(x)$ encodes τ if, for every pair $(a, b) \in M^2$, we have that $(a, b) \in \tau$ if and only if a, b are adjacent and (a, b) has a "label structure" $L_{(a,b)} = \{x, l_1^{(a,b)}, l_2^{(a,b)}, l_3^{(a,b)}, l_4^{(a,b)}\}$ attached to it, where all label structures intersect exactly in the "head vertex" x. See Figure 1 for an example of this (where $p(x)$ has been realized as a point c).

We define the *decoding map* $u : S_1(M) \to 2^{M^2}$ by

$$u(p(x)) = \{(a,b) \in M^2 : f(x,a,b) \in p(x)\}.$$

(Note that we will often use subset notation when formally we in fact mean the characteristic function of that subset within M^2.)

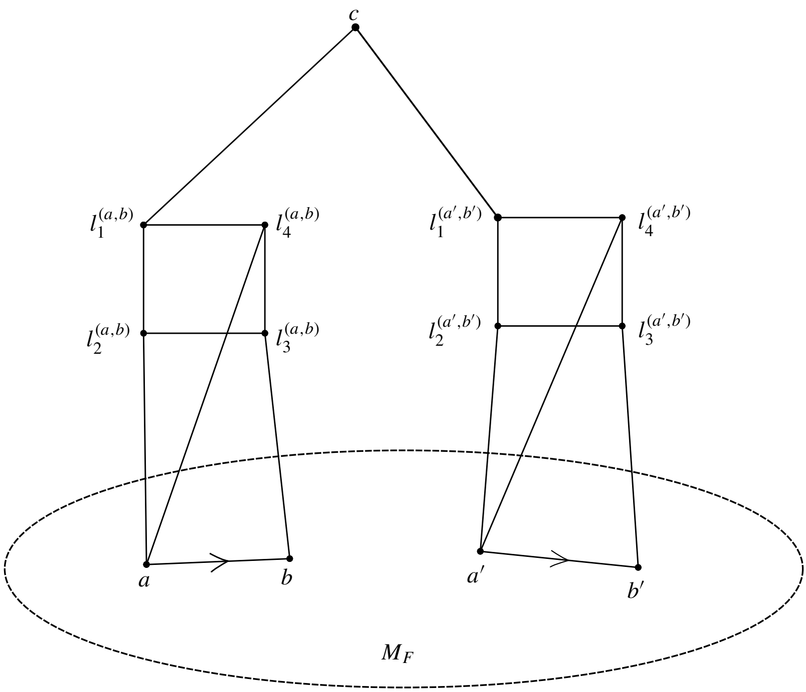


Figure 1. Encoding an orientation using label structures.

It is immediate that if $p(x)$ encodes τ, then $u(p(x)) = \tau$.

The proof of the following is straightforward.

Lemma 5.5. *The map u is a G-flow morphism.*

Now we turn to the key result used in the proof of Theorem 5.1, which we call the *encoding lemma*:

Lemma 5.6. *For each orientation $\tau \in \mathrm{Or}(M)$, there exists a type $p(x) \in S_1(M)$ encoding τ. Thus $\mathrm{Or}(M)$ is a subflow of the image of the decoding map u.*

Before proving Lemma 5.6, whose proof involves a significant amount of technical work, we show how to use it to prove Theorem 5.1.

Proof of Theorem 5.1 given Lemma 5.6. Let M' be an expansion of M with FPT. Let $H = \mathrm{Aut}(M)$.

We have a surjective H-flow morphism $w : S_1(M') \to S_1(M)$ given by restriction, i.e.,

$$w(p(x)) = \{\varphi(x) \in p(x) : \varphi(x) \text{ is a formula in the language } \mathcal{L}(M)\}.$$

We have that $u : S_1(M) \to 2^{M^2}$ is a G-flow morphism, and $\mathrm{Or}(M)$ is a G-subflow of 2^{M^2} contained in the image of u. By considering u as an H-flow morphism to its image, we have that $u \circ w$ is an H-factor of $S_1(M')$ with $\mathrm{Or}(M)$ as a subflow of its image. As M' has FPT, by Lemma 4.3 we see that H fixes an orientation on M. By Proposition 5.2, H has infinitely many orbits on M^2, and so is not oligomorphic. Therefore M' is not ω-categorical, by the Ryll–Nardzewski theorem. □

Part 3: the proof of the encoding lemma. We now prove Lemma 5.6. This forms the bulk of the technical work in this paper.

Let $\tau \in \mathrm{Or}(M)$. To show that there exists a type encoding τ, it suffices to show that the set of formulae

$$\Lambda(x) = \{f(x, a, b) : (a, b) \in \tau\} \cup \{\neg f(x, a, b) : (a, b) \in M^2 \setminus \tau\}$$

is finitely satisfiable in M itself: this implies via compactness that there exists a type $p(x) \in S_1(M)$ containing this set of formulae.

Again, before beginning the proof of Lemma 5.6 we provide a brief informal overview. In order to show the finite satisfiability of $\Lambda(x)$, we will take a finite d-closed substructure A of M and show that the set $\Lambda_A(x)$ is satisfiable in M, where $\Lambda_A(x)$ consists of the formulae in $\Lambda(x)$ with parameters only from A. We will construct a graph B with head vertex c of label structures over A (as in Figure 1) such that $A \leq_d B$ and $B \in \mathcal{C}_F$. Therefore in fact we may assume $A \leq_d B \leq_d M$, using the $\leq_d$-extension property. We will then show that the only label structures in M over any pair of elements of A can be found in B, using the fact that B is d-closed in M, and thus we have that $M \models \Lambda_A(c)$.

We now begin the formal proof. Let $A \leq_d M$ be finite. We define a graph B as

- B includes A as a substructure;
- add a new vertex c to B, with $c \notin A$;
- for $(a, b) \in \tau|_A$ (i.e., the edge ab is oriented from a to b in the orientation τ and ab is an edge of A), add to B four new vertices $l_1^{(a,b)}, l_2^{(a,b)}, l_3^{(a,b)}, l_4^{(a,b)}$ and new edges

$$cl_1^{(a,b)}, \quad l_1^{(a,b)}l_2^{(a,b)}, \quad l_2^{(a,b)}l_3^{(a,b)}, \quad l_3^{(a,b)}l_4^{(a,b)}, \quad l_4^{(a,b)}l_1^{(a,b)}$$

  and add two edges $l_2^{(a,b)}a$, $l_4^{(a,b)}a$ (to the "start vertex" a) and one edge $l_3^{(a,b)}b$ (to the "end vertex" b).

For $(a, b) \in \tau|_A$, let $L_{(a,b)} = \{c, l_1^{(a,b)}, l_2^{(a,b)}, l_3^{(a,b)}, l_4^{(a,b)}\}$. Informally, each $(a, b) \in \tau|_A$ has its orientation labeled by $L_{(a,b)}$.

We have that

$$B = A \cup \bigcup_{(a,b)\in\tau|_A} L_{(a,b)}$$

and the $L_{(a,b)}$ intersect only in c. We will show that $A \leq_d B$ and $B \in \mathcal{C}_F$.

It is recommended that the reader consult Figure 1 during the technical lemmas in this part of the proof.

Lemma 5.7. *We have* $A \leq_d B$.

Proof. For $A \subsetneq C \subseteq B$, we need to show $\delta(C) > \delta(A)$.

First consider the case where A consists of a single edge ab, with $(a, b) \in \tau$ (recall that we chose the control function F so that edges are always d-closed). Then, suppressing subscripts for notational convenience, we have $B = \{c, l_1, l_2, l_3, l_4, a, b\}$. We calculate the relative predimension of some $A \subsetneq C \subseteq B$ in the table below.

$C \setminus A$	$\delta(C/A)$
l_2	1
l_3	1
l_4	1
l_1, l_2	2
l_1, l_4	2
l_2, l_3	1
l_3, l_4	1
c, l_1	3
l_1, l_2, l_3	2
l_1, l_2, l_4	2
l_1, l_3, l_4	2
l_2, l_3, l_4	1
l_1, l_2, l_3, l_4	1
c, l_1, l_2, l_3, l_4	2

The remaining cases result from free amalgamations over A, and so also have positive predimension (as if two graphs X, Y are freely amalgamated over Z, then $\delta((X \cup Y)/Z) = \delta(X/Z) + \delta(Y/Z)$). The remaining cases are where $C \setminus A$ is equal to $\{l_1\}$, $\{c\}$, $\{l_1, l_3\}$, $\{l_2, l_4\}$ or $\{c\} \cup X$, where $X \subseteq \{l_2, l_3, l_4\}$.

Now consider the general case of finite $A \leq_d M$. Given $A \subsetneq C \subseteq B$, the vertices of C consist of A together with subsets $J_{(a,b)}$ of $L_{(a,b)}$ for each $(a, b) \in \tau|_A$. For $(a, b) \in \tau|_A$, let $J'_{(a,b)} = J_{(a,b)} \cup A$.

If $c \notin C$, then the $J'_{(a,b)}$ are freely amalgamated over A, and so from the single-edge case we see that $\delta(C/A) > 0$.

We now consider the case where $c \in C$. If $l_1^{(a,b)} \notin J_{(a,b)}$ for all $(a, b) \in \tau|_A$, then C consists of a vertex c with no neighbors together with a free amalgamation over A of each of the $J'_{(a,b)} \setminus \{c\}$, for $(a, b) \in \tau|_A$. So, from the single-edge case and the fact that $\delta(\{c\}) = 2$, we have that $\delta(C/A) > 0$.

If $c \in C$ and there exists $(a', b') \in \tau|_A$ with $l_1^{(a',b')} \in J_{(a',b')}$, then C is the free amalgam over A of each of the $J'_{(a,b)} \setminus \{c\}$ for which $l_1^{(a,b)} \notin J_{(a,b)}$ (where $(a, b) \in \tau|_A$), together with

$$\bigcup\{J'_{(a,b)} : (a,b) \in \tau|_A,\, l_1^{(a,b)} \in J_{(a,b)}\}.$$

Therefore we need only consider the case where $l_1^{(a,b)} \in J_{(a,b)}$ for all $(a,b) \in \tau|_A$. The single-edge calculation shows that $\delta(J_{(a,b)} \setminus \{c\}/A) \geq 1$ for each $J_{(a,b)}$, and these $J'_{(a,b)} \setminus \{c\}$ are freely amalgamated over A. Each addition of an edge $cl_1^{(a,b)}$ reduces the predimension by one, but the single addition of the vertex c adds two to the predimension, so in total $\delta(C/A) > 0$. □

Lemma 5.8. *For $(a,b) \in \tau|_A$, we have that the substructures of B given by $\{a, b, l_1^{(a,b)}, l_2^{(a,b)}, l_3^{(a,b)}, l_4^{(a,b)}\}$ and $L_{(a,b)}$ lie in $\mathcal{C}_F$.*

Proof. We write l_1, l_2, l_3, l_4, suppressing superscripts.

To show that $\{a, b, l_1, l_2, l_3, l_4\} \in \mathcal{C}_F$, we consider each subset $C \subseteq \{a, b, l_1, l_2, l_3, l_4\}$ and show that $\delta(C) \geq F(|C|)$. To speed up the process of checking each subset C, in the below table we show that certain subsets $C \subseteq \{a, b, l_1, l_2, l_3, l_4\}$ lie in $\mathcal{C}_F$, and therefore every $C' \subseteq C$ must satisfy $\delta(C') \geq F(|C'|)$.

C	proof that $C \in \mathcal{C}_F$
$l_1l_2l_3l_4$, l_2l_3ab, l_3l_4ab, $l_1l_2l_4a$, $l_2l_3l_4a$	C is a 4-cycle
$l_1l_2l_3ab$	free amalgam of l_2l_3ab, l_1l_2 over l_2
$l_1l_3l_4ab$	free amalgam of l_3l_4ab, l_1l_4 over l_4
$l_1l_2l_4ab$	free amalgam of $l_1l_2l_4a$, ab over a
$l_1l_2l_3l_4b$	free amalgam of $l_1l_2l_3l_4$, l_3b over l_3

We now check the remaining subsets $C \subseteq \{a, b, l_1, l_2, l_3, l_4\}$ by directly calculating the predimension:

C	$\delta(C)$	$F(\lvert C\rvert)$
$l_2l_3l_4ab$	4	$F(5) < 4$
$l_1l_2l_3l_4a$	4	$F(5) < 4$
$l_1l_2l_3l_4ab$	4	$F(6) < 4$

We have now shown that $\{a, b, l_1, l_2, l_3, l_4\} \in \mathcal{C}_F$. For the second part of the lemma, we obtain $L_{(a,b)} \in \mathcal{C}_F$ via the free amalgam of $L_{(a,b)}$ and cl_1 over l_1 (recalling that we have defined our control function F so that points are always d-closed). □

Lemma 5.9. *We have that $B \in \mathcal{C}_F$.*

Proof. We have to show that $\delta(C) \geq F(|C|)$ for $C \subseteq B$. The vertices of C consist of $C \cap A$ together with subsets $J_{(a,b)}$ of $L_{(a,b)}$ for each $(a,b) \in \tau|_A$ (some of these $J_{(a,b)}$ may be empty). For $(a,b) \in \tau|_A$, let $J'_{(a,b)} = J_{(a,b)} \cup (C \cap A)$.

First we consider the case where $c \notin C$. C is then the free amalgam of the $J'_{(a,b)}$ (where $(a,b) \in \tau|_A$) over $C \cap A$. Given that $\mathcal{C}_F$ is a free amalgamation class and $C \cap A \leq_d C$, it therefore suffices to show that $J'_{(a,b)} \in \mathcal{C}_F$ for $(a,b) \in \tau|_A$. Fix $(a,b) \in \tau|_A$. To show that $J'_{(a,b)} \in \mathcal{C}_F$, as $J'_{(a,b)}$ is a free amalgam of $J_{(a,b)} \cup (\{a,b\} \cap C)$ and $C \cap A \in \mathcal{C}_F$ over $\{a,b\} \cap C \in \mathcal{C}_F$, it suffices to show that $J_{(a,b)} \cup (\{a,b\} \cap C)$ lies in $\mathcal{C}_F$, and we have already checked this in Lemma 5.8.

Now we consider the case where $c \in C$. If $l_1^{(a,b)} \notin J_{(a,b)}$ for each $(a,b) \in \tau|_A$, then C consists of a vertex c with no neighbors together with the free amalgam over $C \cap A$ of each $J'_{(a,b)} \setminus \{c\}$, and so we are done by the first case in the previous paragraph. Otherwise, C is the free amalgam over $C \cap A$ of

$$\bigcup \{J'_{(a,b)} : l_1^{(a,b)} \in J_{(a,b)}, (a,b) \in \tau|_A\}$$

with each $J'_{(a,b)} \setminus \{c\}$ for which $l_1^{(a,b)} \notin J_{(a,b)}$, and so using the first case considered above we may reduce to the case where each nonempty $J_{(a,b)}$ contains $l_1^{(a,b)}$.

Similarly, we may exclude the case where C contains sets $J_{(a,b)}$ for which $J_{(a,b)} = \{c, l_1^{(a,b)}, l_3^{(a,b)}\}$, as C is the free amalgam over $C \cap A$ of

$$\bigcup \{J'_{(a,b)} : (a,b) \in \tau|_A, J_{(a,b)} \neq \{c, l_1^{(a,b)}, l_3^{(a,b)}\}\} \cup$$
$$\bigcup \{\{c, l_1^{(a,b)}\} \cup (C \cap A) : J_{(a,b)} = \{c, l_1^{(a,b)}, l_3^{(a,b)}\}\}$$

with each $\{l_3^{(a,b)}\} \cup (C \cap A)$ (which lies in $\mathcal{C}_F$ by Lemma 5.8) for which $J_{(a,b)} = \{c, l_1^{(a,b)}, l_3^{(a,b)}\}$. We may likewise freely amalgamate over c to exclude the cases where C contains sets $J_{(a,b)}$ for which $J_{(a,b)} = \{c, l_1^{(a,b)}\}$, or for which $J_{(a,b)}$ is any subset of $L_{(a,b)}$ but we have $a, b \notin C \cap A$.

So, the case remaining is where C consists of $C \cap A$ together with sets $J_{(a,b)}$ containing $c, l_1^{(a,b)}$ and at least one of $l_2^{(a,b)}, l_4^{(a,b)}$, where each $J_{(a,b)}$ has some edge to $C \cap A$. We need to show that $\delta(C) \geq F(|C|)$.

We now calculate the relative predimension over $A \cup \{c\}$ of each remaining possible $J_{(a,b)} \cup X$, $X \subseteq \{a,b\}$, in the following table, where we label each structure as Y_i, $1 \leq i \leq 11$:

$J_{(a,b)} \cup X$	label	$\delta(J_{(a,b)} \cup X / A \cup \{c\})$
cl_1l_2a	Y_1	1
cl_1l_4a	Y_2	1
$cl_1l_2l_3a$	Y_3	2
$cl_1l_2l_3b$	Y_4	2
$cl_1l_2l_3ab$	Y_5	1
$cl_1l_3l_4a$	Y_6	2
$cl_1l_3l_4b$	Y_7	2
$cl_1l_3l_4ab$	Y_8	1
$cl_1l_2l_3l_4a$	Y_9	1
$cl_1l_2l_3l_4b$	Y_{10}	2
$cl_1l_2l_3l_4ab$	Y_{11}	0

We write k_i for how many times Y_i occurs in C. We also write $\delta_i = \delta(Y_i / A \cup \{c\})$. Let $\lambda_i = |\{l_1, l_2, l_3, l_4\} \cap Y_i|$.

Then, recalling that the vertex c also adds 2 to the predimension, we have that

$$\delta(C) = \sum_{1 \le i \le 11} \delta_i k_i + 2 + \delta(C \cap A).$$

Now,

$$\begin{aligned} F(|C|) &= F\left(1 + |C \cap A| + \sum_{1 \le i \le 11} \lambda_i k_i\right) \\ &\le F\left(1 + |C \cap A| + 4 \sum_{1 \le i \le 11} k_i\right) \\ &= F\left(1 + |C \cap A| + 4(k_4 + k_7 + k_{10}) + 4 \sum_{i \le 11, i \notin \{4,7,10\}} k_i\right). \end{aligned}$$

As $\tau|_{C \cap A}$ is a 2-orientation, we have that each $a \in C \cap A$ can have at most two label structures with a as the starting vertex (i.e., with edges to a from l_2, l_4), and so

$$\sum_{i \le 11, i \notin \{4,7,10\}} k_i \le 2|C \cap A|.$$

So

$$F(|C|) \le F(8|C \cap A| + 4(k_4 + k_7 + k_{10}) + 1 + |C \cap A|).$$

As $F(u+v) \le F(u) + vF'(u)$ and $F'(x) \le \frac{2}{8x+1}$ for $x \ge 2$, we have that if $|C \cap A| \ge 2$, then

$$\begin{aligned} F(|C|) &\le F(|C \cap A|) + \frac{2}{8|C \cap A| + 1}(8|C \cap A| + 4(k_4 + k_7 + k_{10}) + 1) \\ &< F(|C \cap A|) + 2 + k_4 + k_7 + k_{10} \\ &\le \delta(C). \end{aligned}$$

If $|C \cap A| = 1$, then

$$\begin{aligned} F(|C|) &\le F(1 + |C \cap A|) + (8|C \cap A| + 4(k_4 + k_7 + k_{10}))F'(1 + |C \cap A|) \\ &= 3 + \tfrac{2}{8 \cdot 2 + 1}(8|C \cap A| + 4(k_4 + k_7 + k_{10})) \\ &< 4 + \tfrac{8}{17}(k_4 + k_7 + k_{10}) \\ &\le \delta(C) \end{aligned}$$

(as $\delta(C \cap A) = 2$). □

Lemma 5.10. *Let $\tau \in \mathrm{Or}(M)$. Then the set of formulae*

$$\Lambda(x) = \{f(x,a,b) : (a,b) \in \tau\} \cup \{\neg f(x,a,b) : (a,b) \in M^2 \setminus \tau\}$$

is finitely satisfiable in M.

Proof. Let $A \leq_d M$ be finite, and let

$$\Phi_A(x) = \{f(x,a,b) : (a,b) \in \tau|_A\},\ \Psi_A(x) = \{\neg f(x,a,b) : (a,b) \in A^2 \setminus \tau\}.$$

Let $\Lambda_A(x) = \Phi_A(x) \cup \Psi_A(x)$. We will show that $\Lambda_A(x)$ is satisfiable in M.

Let $B \supseteq A$ be as constructed previously, with distinguished head vertex c. As $A \leq_d B$ (Lemma 5.7) and $B \in \mathcal{C}_F$ (Lemma 5.9), we may use the $\leq_d$-extension property of M to assume that $A \leq_d B \leq_d M$.

It is immediate from the construction of B that $B \models \Phi_A(c)$ and hence $M \models \Phi_A(c)$, as for each $(a,b) \in \tau|_A$, there is a label structure $L_{(a,b)}$ attached.

We now show that $M \models \Psi_A(c)$. It suffices to show that for $(a,b) \in A^2$, if $M \models f(c,a,b)$ then the l_i, $1 \leq i \leq 4$, that $f(c,a,b)$ specifies must lie in $\mathrm{cl}^d_M(\{a,b,c\})$ and therefore in B, as $B \leq_d M$. We show that $\{l_1,l_2,l_3,l_4\} \subseteq \mathrm{cl}^d_M(\{a,b,c\})$ in the table below.

X/Y	$\delta(X/Y)$
$l_1,l_2,l_3,l_4/a,b,c$	0
$l_1,l_2,l_3/l_4,a,b,c$	-1
$l_1,l_2,l_4/l_3,a,b,c$	-1
$l_1,l_3,l_4/l_2,a,b,c$	-1
$l_2,l_3,l_4/l_1,a,b,c$	-1
$l_1,l_2/l_3,l_4,a,b,c$	-1
$l_1,l_3/l_2,l_4,a,b,c$	-2
$l_1,l_4/l_2,l_3,a,b,c$	-1
$l_2,l_3/l_1,l_4,a,b,c$	-1
$l_2,l_4/l_1,l_3,a,b,c$	-2
$l_3,l_4/l_1,l_2,a,b,c$	-1
$l_1/l_2,l_3,l_4,a,b,c$	-1
$l_2/l_1,l_3,l_4,a,b,c$	-1
$l_3/l_1,l_2,l_4,a,b,c$	-1
$l_4/l_1,l_2,l_3,a,b,c$	-1

This completes the proof of Lemma 5.10. □

The above lemma implies, via compactness, that there exists a type $p(x)$ containing the set $\{f(x,a,b) : (a,b) \in \tau\} \cup \{\neg f(x,a,b) : (a,b) \in M^2 \setminus \tau\}$, and thus $p(x)$ encodes τ. This completes the proof of the encoding lemma (Lemma 5.6), and therefore the proof of Theorem 5.1.

Acknowledgements

The author would like to thank David Evans for his supervision during this project, which formed the first half of the author's PhD thesis. The author would also like to thank the anonymous referee for their helpful comments, including a simplification of the proof of the main theorem.

References

[Auslander 1988] J. Auslander, *Minimal flows and their extensions*, North-Holland Mathematics Studies **153**, North-Holland, Amsterdam, 1988. MR Zbl

[Bollobás 1998] B. Bollobás, *Modern graph theory*, Graduate Texts in Mathematics **184**, Springer, 1998. MR Zbl

[Evans 2013] D. M. Evans, "Homogeneous structures, omega-categoricity and amalgamation constructions", unpublished notes from talks given at the Hausdorff Institute for Mathematics, Bonn, 2013, http://wwwf.imperial.ac.uk/~dmevans/Bonn2013_DE.pdf.

[Evans et al. 2019] D. M. Evans, J. Hubička, and J. Nešetřil, "Automorphism groups and Ramsey properties of sparse graphs", *Proc. Lond. Math. Soc.* (3) **119**:2 (2019), 515–546. MR Zbl

[Fraïssé 1954] R. Fraïssé, "Sur l'extension aux relations de quelques propriétés des ordres", *Ann. Sci. École Norm. Sup.* (3) **71** (1954), 363–388. MR Zbl

[Hodges 1993] W. Hodges, *Model theory*, Encyclopedia of Mathematics and its Applications **42**, Cambridge University Press, 1993. MR Zbl

[Hrushovski 1988] E. Hrushovski, "A stable $\aleph_0$-categorical pseudoplane", unpublished notes, 1988, http://people.maths.ox.ac.uk/hrushovski/Some%20older%20texts/stable-alephzero-categorical.pdf.

[Kechris et al. 2005] A. S. Kechris, V. G. Pestov, and S. Todorcevic, "Fraïssé limits, Ramsey theory, and topological dynamics of automorphism groups", *Geom. Funct. Anal.* **15**:1 (2005), 106–189. MR Zbl

[Meir and Sullivan 2023] N. Meir and R. Sullivan, "The externally definable Ramsey property and fixed points on type spaces", preprint, 2023. Zbl arXiv 2309.06522

[Nash-Williams 1964] C. S. J. A. Nash-Williams, "Decomposition of finite graphs into forests", *J. London Math. Soc.* **39** (1964), 12. MR Zbl

Received 18 Oct 2023. Revised 15 May 2024.

ROB SULLIVAN:

robertsullivan1990@gmail.com
Institut für Mathematische Logik und Grundlagenforschung, University of Münster, Münster, Germany

Model Theory
Vol. 3, No. 3, 2024
https://doi.org/10.2140/mt.2024.3.883

Correction to the article Kim-independence in positive logic

Jan Dobrowolski and Mark Kamsma

Volume 1:1 (2022), 55–113

The proof of the independence theorem for Kim-independence in positive thick NSOP_1 theories by Dobrowolski and Kamsma (*Model Theory* **1** (2022), 55–113) contains a gap. The theorem is still true, and in this corrigendum we give a different proof.

1. Introduction

The proof of the independence theorem for Kim-independence in thick NSOP_1 theories [Dobrowolski and Kamsma 2022, Theorem 7.7] contains a gap. Everything in that proof is fine up to the point where it is argued how the theorem follows from what is called "Claim 2" (at the bottom of page 88). By compactness, an M-indiscernible sequence $(g_i h_i g'_i h'_i g''_i h''_i)_{i \in \mathbb{Z}}$ is extracted from the data from Claim 2. However, it may be that the properties $(h''_i g''_{i+1})_{i \in \mathbb{Z}} \models (q'|_{z_0, y})^{\otimes \mathbb{Z}}|_M$ and $h_i g_{i+1} \equiv_{M h_{>i} g_{>i+1} h''_{>i} g''_{i+1}} h''_i g''_{i+1}$ are not carried over.

The theorem, as stated, is still true, and in this corrigendum we give a different proof. We assume familiarity with [Dobrowolski and Kamsma 2022].

2. Technical tools

We reformulate the chain condition in a form that will be useful to us.

Lemma 2.1 (chain condition). *Let T be a thick NSOP_1 theory. Suppose that $a \mathop{\smile\!\!\!\!\!\!\mid\,}^K_M b$ and that $(b_i)_{i<\omega}$ is a Morley sequence in some global M-Ls-invariant type with $b_0 = b$. Then, writing $p(x, b) = \mathrm{tp}(a/Mb)$, we have that*

$$\bigcup_{i<\omega} p(x, b_i)$$

does not Kim-divide over M.

Kamsma is supported by the EPSRC grant EP/X018997/1 .
MSC2020: primary 03C45; secondary 03C10.
Keywords: Kim-independence, Kim-dividing, positive logic, NSOP1 theory, independence theorem.

Proof. Let $q(x)$ be the global M-Ls-invariant type in which $(b_i)_{i<\omega}$ is a Morley sequence. As $a \mathop{\smile\!\!\!\!\!\mid}^K_M b$ we have by [Dobrowolski and Kamsma 2022, Proposition 4.2] that there is an Ma-indiscernible $(b'_i)_{i<\omega} \models q^{\otimes\omega}|_M$ with $b'_0 = b$. So we have $(b'_i)_{i<\omega} \equiv_M (b_i)_{i<\omega}$ and we let a^* be such that $a(b'_i)_{i<\omega} \equiv_M a^*(b_i)_{i<\omega}$. Then $(b_i)_{i<\omega}$ is Ma^*-indiscernible, and so $a^* \mathop{\smile\!\!\!\!\!\mid}^K_M (b_i)_{i<\omega}$ by [loc. cit., Lemma 6.1]. We conclude by noting that $a^*b_i \equiv_M a^*b_0 \equiv_M ab'_0 = ab$ for all $i < \omega$. □

Proposition 2.2 (being Ls-invariant is type-definable). *Let T be a thick theory. Let C be some parameter set and let $N \supseteq C$ be $(2^{|C|+\lambda_T})^+$-saturated (possibly N is the monster). Define $\Sigma(x)$ to be the following partial type over N*

$$\bigcup\{\mathrm{d}_C(xb, xb') \leq 2 : b, b' \in N \text{ are finite tuples such that } \mathrm{d}_C(b, b') \leq 1\}.$$

Then a type $q(x)$ over N is C-Ls-invariant if and only if $\Sigma(\alpha)$ for $\alpha \models q$.

Proof. Let $q(x)$ be a C-Ls-invariant type over N and let $\alpha \models q$. Let $b, b' \in N$ be finite tuples such that $\mathrm{d}_C(b, b') \leq 1$. Then there is a C-indiscernible sequence $(b_i)_{i<\omega}$ with $b_0b_1 = bb'$, which we may assume to be in N by saturation. Using saturation again, we find a λ_T-saturated $C \subseteq M \subseteq N$ such that $(b_i)_{i<\omega}$ is M-indiscernible. In particular this means that $bM \equiv^{\mathrm{Ls}}_C b'M$ and so $\alpha bM \equiv^{\mathrm{Ls}}_C \alpha b'M$. It follows that $\alpha b \equiv^{\mathrm{Ls}}_M \alpha b'$ and thus by our choice of M we get $\mathrm{d}_C(\alpha b, \alpha b') \leq 2$. As b, b' were arbitrary, we conclude that $\models \Sigma(\alpha)$.

For the other direction we let $q(x)$ be a type over N such that for $\alpha \models q$ we have $\models \Sigma(\alpha)$. Now let $d, d' \in N$ be (potentially infinite tuples) such that $d \equiv^{\mathrm{Ls}}_C d'$. Let $n < \omega$ be such that $\mathrm{d}_C(d, d') \leq n$, we claim that $\mathrm{d}_C(\alpha d, \alpha d') \leq 2n$, which implies the required $\alpha d \equiv^{\mathrm{Ls}}_C \alpha d'$. By thickness we have that the condition $\mathrm{d}_C(\alpha d, \alpha d') \leq 2n$ is given by

$$\bigcup\{\mathrm{d}_C(\alpha b, \alpha b') \leq 2n : b \subseteq d \text{ and } b' \subseteq d' \text{ are finite matching tuples}\}.$$

So we have reduced the problem to the case where d and d' are finite. By saturation then there are $d = d_0, d_1, \dots, d_n = d'$ in N such that $\mathrm{d}_C(d_i, d_{i+1}) \leq 1$ for all $0 \leq i < n$. By assumption we thus have that $\mathrm{d}_C(\alpha d_i, \alpha d_{i+1}) \leq 2$ for all $0 \leq i < n$. We conclude that $\mathrm{d}_C(\alpha d, \alpha d') \leq 2n$, as required. □

Proposition 2.3 (extending Ls-invariant types). *Let T be a thick theory. Let $N \supseteq C$ be $(2^{|C|+\lambda_T})^+$-saturated. Suppose that $p(x) = \mathrm{tp}(a/N)$ is a C-Ls-invariant type, then $p(x)$ extends to a unique global C-Ls-invariant type $q(x)$.*

Proof. Let $\Sigma(x)$ be the global partial type from Proposition 2.2 expressing C-Ls-invariance. We will show that $p(x) \cup \Sigma(x)$ is finitely satisfiable. So let $\varphi(x, e) \in p(x)$, where e is a tuple of parameters from N, and let $\Sigma_0(x) \subseteq \Sigma(x)$ be finite. Let $b_1, \dots, b_n$ and $b'_1, \dots, b'_n$ be the finite tuples that occur in $\Sigma_0(x)$, so $\mathrm{d}_C(b_i, b'_i) \leq 1$ for all $1 \leq i \leq n$. By saturation of N we find $d_1, \dots, d_n, d'_1, \dots, d'_n \in N$ such that

$d_1 \cdots d_n d'_1 \cdots d'_n \equiv_{Ce} b_1 \cdots b_n b'_1 \cdots b'_n$. So for all $1 \leq i \leq n$ we have $\mathrm{d}_C(d_i, d'_i) \leq 1$, and hence $\mathrm{d}_C(ad_i, ad'_i) \leq 2$ by Proposition 2.2 applied to $p(x)$. Now let a^* be such that $ad_1 \cdots d_n d'_1 \cdots d'_n \equiv_{Ce} a^* b_1 \cdots b_n b'_1 \cdots b'_n$. Then by construction we have that $\models \varphi(a^*, e)$ and $\models \Sigma_0(a^*)$, which proves finite satisfiability of $p(x) \cup \Sigma(x)$. By compactness we then find a realisation α of $p(x) \cup \Sigma(x)$, so that $q(x) = \mathrm{tp}(\alpha/\mathfrak{M})$ is our desired C-Ls-invariant type. The uniqueness claim follows from [Dobrowolski and Kamsma 2022, Fact 7.6]. □

We recall from [loc. cit., Definition 3.12] that $a \mathop{\smile\!\!\!\!\!\!\;|}\nolimits_C^{iLs} b$ means that $\mathrm{tp}(a/Cb)$ extends to a global C-Ls-invariant type.

Proposition 2.4. *Let T be a thick theory. If $(a_i)_{i<\omega}$ is a C-indiscernible sequence such that $a_i \mathop{\smile\!\!\!\!\!\!\;|}\nolimits_C^{iLs} a_{<i}$ for all $i < \omega$ then $(a_i)_{i<\omega}$ is a Morley sequence in some global C-Ls-invariant type.*

Proof. By compactness we find a_ω such that $(a_i)_{i \leq \omega}$ is C-indiscernible. Set $p(x) = \mathrm{tp}(a_\omega / Ca_{<\omega})$ and let $\Sigma(x)$ be the global partial type from Proposition 2.2. We claim that $p(x) \cup \Sigma(x)$ is consistent. Indeed, for any finite $p'(x) \subseteq p(x)$ there is some $i < \omega$ so that $p'(x)$ only contains parameters in $Ca_{<i}$, and so $\models p'(a_i)$ by C-indiscernibility. As $a_i \mathop{\smile\!\!\!\!\!\!\;|}\nolimits_C^{iLs} a_{<i}$ we then have that $p'(x)$ extends to a global C-Ls-invariant type $q'(x)$, and any realisation of $q'(x)$ will then be a realisation of $p'(x) \cup \Sigma(x)$. So $p(x) \cup \Sigma(x)$ is finitely satisfiable and hence consistent.

Let α^* be a realisation of $p(x) \cup \Sigma(x)$ and set $q^*(x) = \mathrm{tp}(\alpha^*/\mathfrak{M})$, so $q^*(x)$ is global C-Ls-invariant. Let $a^* \equiv^{\mathrm{Ls}}_{Ca_{<\omega}} \alpha^*$, then there is $f \in \mathrm{Aut}(\mathfrak{M}/Ca_{<\omega})$ such that $f(a^*) = a_\omega$. Set $q = f(q^*)$, so $q(x)$ is global C-Ls-invariant by [loc. cit., Lemma 3.8(i)] with $p(x) \subseteq q(x)$ and, letting α be a realisation of q, we have $\alpha \equiv^{\mathrm{Ls}}_{Ca_{<\omega}} a_\omega$.

For any $i < \omega$ we thus have $a_i \equiv^{\mathrm{Ls}}_{Ca_{<i}} a_\omega \equiv^{\mathrm{Ls}}_{Ca_{<i}} \alpha$. We therefore have $a_{<i} \models q^{\otimes i}|_C$ for all $i < \omega$ and so $(a_i)_{i<\omega} \models q^{\otimes \omega}|_C$. So $(a_i)_{i<\omega}$ is the automorphic image over C of a Morley sequence over C, hence it is itself a Morley sequence in a (potentially different) global C-Ls-invariant type. □

3. Spread out trees

We recall various definitions concerning trees and trees of parameters (which we will from now on also simply call trees) from [Kaplan and Ramsey 2020]. In particular, we will work with the ill-founded trees $\mathcal{T}_\alpha$ from [loc. cit., Definition 5.1] and we use the same notation, so we assume familiarity with those definitions. We refer to [Kamsma 2024] for the definitions and terminology involving s-indiscernibility, str-indiscernibility and generalised EM-types. We slightly adjust [Kaplan and Ramsey 2020, Definition 5.7] to fit our situation.

Definition 3.1. Let $(a_\eta)_{\eta \in \mathcal{T}_\alpha}$ be a tree and let M be an e.c. model:

(i) We call $(a_\eta)_{\eta \in \mathcal{T}_\alpha}$ *spread out over M* if for all $\eta \in \mathcal{T}_\alpha$ with $\mathrm{dom}(\eta) = [\beta+1, \alpha)$ for some $\beta < \alpha$, there is a global M-Ls-invariant type $q_\eta \supseteq \mathrm{tp}(a_{\trianglerighteq \eta^\frown\langle 0\rangle}/M)$ such that $(a_{\trianglerighteq \eta^\frown\langle i\rangle})_{i<\omega}$ is a Morley sequence in q_η over M.

(ii) A *Morley tree over M* is an str-indiscernible and spread out tree over M.

(iii) A *tree Morley sequence over M* is a branch in an infinite height Morley tree over M.

Lemma 3.2. *Suppose that $(a_i)_{i<\omega}$ is a tree Morley sequence over M*:

(i) *If $b_i \subseteq a_i$ for each $i < \omega$, of matching length and position, then $(b_i)_{i<\omega}$ is a tree Morley sequence over M.*

(ii) *Fix $1 \leq n < \omega$ and define $d_i = (a_{ni}, \dots, a_{ni+n-1})$ for all $i < \omega$. Then $(d_i)_{i<\omega}$ is a tree Morley sequence over M.*

Proof. This is essentially [Kaplan and Ramsey 2020, Lemma 5.9], but we work with slightly different definitions, so we go through the proof here. Part (i) is clear, because being a Morley tree is preserved under taking subtuples. For (ii) we let $(b_\eta)_{\eta \in \mathcal{T}_\omega}$ be a Morley tree such that $(a_i)_{i<\omega}$ is a branch in $(b_\eta)_{\eta \in \mathcal{T}_\omega}$. We may assume that $(a_i)_{i<\omega}$ is the branch indexed by the constant zero functions. We define $j : \mathcal{T}_\omega \to \mathcal{T}_\omega$ so that for $\eta \in \mathcal{T}_\omega$ with $\mathrm{dom}(\eta) = [k, \omega)$ we have $\mathrm{dom}(j(\eta)) = [nk+n-1, \omega)$ and

$$j(\eta)(m) = \begin{cases} \eta((m-(n-1))/n) & \text{if } n \mid (m-(n-1)), \\ 0 & \text{otherwise,} \end{cases}$$

for all $m \in [nk+n-1, \omega)$. We define $(c_\eta)_{\eta \in \mathcal{T}_\omega}$ by $c_\eta = (b_{j(\eta)}, \dots, b_{j(\eta)^\frown\langle 0\rangle^{n-1}})$. This corresponds to the n-fold elongation of $(b_\eta)_{\eta \in \mathcal{T}_\omega}$ from [Chernikov and Ramsey 2016]. One then straightforwardly verifies that $(c_\eta)_{\eta \in \mathcal{T}_\omega}$ is a Morley tree over M, so $(c_{\zeta_i})_{i<\omega}$ is a tree Morley sequence over M. For $i < \omega$ we have

$$c_{\zeta_i} = (b_{\zeta_{ni+n-1}}, \dots, b_{\zeta_{ni}}) = (a_{ni+n-1}, \dots, a_{ni}),$$

so by reversing the order of the tuples we see that $(d_i)_{i<\omega}$ is a tree Morley sequence over M. □

Lemma 3.3 (Kim's lemma for tree Morley sequences). *Let T be a thick NSOP$_1$ theory. Let M be an e.c. model and let $\Sigma(x, b)$ be a partial type over M. Then the following are equivalent*:

(i) *$\Sigma(x, b)$ Kim-divides over M.*

(ii) *For some tree Morley sequence $(b_i)_{i<\omega}$ over M with $b_0 = b$ we have that $\bigcup_{i<\omega} \Sigma(x, b_i)$ is inconsistent.*

(iii) *For every tree Morley sequence* $(b_i)_{i<\omega}$ *over* M *with* $b_0 = b$ *we have that* $\bigcup_{i<\omega} \Sigma(x, b_i)$ *is inconsistent.*

Proof. This is [Kaplan and Ramsey 2020, Corollary 5.14], whose proof is really found in [loc. cit., Proposition 5.13]. Our setting requires some minor extra verifications, which we will do below, but the proof is essentially the same.

Given the existence of tree Morley sequences starting with b (Lemma 3.10), the equivalence of these three statements reduces to proving that for any tree Morley sequence $(b_i)_{i<\omega}$ over M with $b_0 = b$ we have that $\Sigma(x, b)$ Kim-divides if and only if $\bigcup_{i<\omega} \Sigma(x, b_i)$ is inconsistent.

Let $(c_\eta)_{\eta\in\mathcal{T}_\omega}$ be a Morley tree over M such that $(b_i)_{i<\omega}$ is a branch in that tree, which we may assume to be the constant zero branch. For $i < \omega$ define $\eta_i \in \mathcal{T}_\omega$ to be the function with domain $[i, \omega)$ such that

$$\eta_i(j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

By str-indiscernibility, the sequences $(c_{\zeta_i})_{i<\omega}$ and $(c_{\eta_i})_{i<\omega}$ are M-indiscernible. We claim that $(c_{\eta_i})_{i<\omega}$ is a Morley sequence over M in a global M-Ls-invariant type. Indeed, because $(c_\eta)_{\eta\in\mathcal{T}_\omega}$ is spread out over M we have that $c_{\eta_i} \mathop{\smile\!\!\!\!\!\!\;|}\nolimits_M^{iLs} (c_{\eta_j})_{j<i}$ for all $i < \omega$. So the claim follows from Proposition 2.4. By str-indiscernibility we also have for all $i < \omega$ that c_{ζ_i}, c_{η_i} starts an $M(c_{\zeta_j}, c_{\eta_j})_{j>i}$-indiscernible sequence. So since T is NSOP$_1$ we can apply [Dobrowolski and Kamsma 2022, Lemma 5.10] to conclude that $\bigcup_{i<\omega} \Sigma(x, c_{\zeta_i})$ is inconsistent if and only if $\bigcup_{i<\omega} \Sigma(x, c_{\eta_i})$ is inconsistent. The former is just $\bigcup_{i<\omega} \Sigma(x, b_i)$, and the latter is inconsistent if and only if $\Sigma(x, b)$ Kim-divides by Kim's lemma for NSOP$_1$ theories [loc. cit., Proposition 4.4], which concludes the proof. □

Fact 3.4 (tree modelling theorems). *Let T be a thick theory*:

(i) *Let* $(a_\eta)_{\eta\in\mathcal{T}_\alpha}$ *be a tree of tuples and let* C *be any set of parameters, then there is a tree* $(b_\eta)_{\eta\in\mathcal{T}_\alpha}$ *that is s-indiscernible over* C *and* EM$_s$*-based on* $(a_\eta)_{\eta\in\mathcal{T}_\alpha}$ *over* C.

(ii) *Let* C *be any parameter set,* κ *any cardinal, and let* $\lambda = \beth_{(2^{|T|+|C|+\kappa})^+}$. *Given any tree* $(a_\eta)_{\eta\in\mathcal{T}_\lambda}$ *of* κ*-tuples that is s-indiscernible over* C, *there is a tree* $(b_\eta)_{\eta\in\mathcal{T}_\omega}$ *that is str-indiscernible over* C *str-based on* $(a_\eta)_{\eta\in\mathcal{T}_\lambda}$ *over* C. *The latter means that for any finite tuple* $\bar\eta \in \mathcal{T}_\omega$ *there is* $\bar\nu \in \mathcal{T}_\lambda$ *such that* $\bar\eta$ *and* $\bar\nu$ *have the same str-quantifier-free type and* $b_{\bar\eta} \equiv_C a_{\bar\nu}$.

Proof. Part (i) is [Kamsma 2024, Theorem 4.6], which is essentially just compactness applied to [Dobrowolski and Kamsma 2022, Proposition 5.8]. Part (ii) is [Kamsma 2024, Theorem 4.8], which is technically stated for well-founded trees, but its proof applies to the ill-founded trees we are interested in here. □

Lemma 3.5. *Let T be a thick theory. Suppose that $(a_\eta)_{\eta\in\mathcal{T}_\alpha}$ is s-indiscernible and spread out over M and that $(b_\eta)_{\eta\in\mathcal{T}_\omega}$ is str-based on $(a_\eta)_{\eta\in\mathcal{T}_\alpha}$ over M, then $(b_\eta)_{\eta\in\mathcal{T}_\omega}$ is spread out over M.*

Proof. Let $\eta \in \mathcal{T}_\omega$, we have to show that $(b_{\trianglerighteq \eta^\frown\langle i\rangle})_{i<\omega}$ is a Morley sequence in some global M-Ls-invariant type. We claim that $b_{\trianglerighteq \eta^\frown\langle i\rangle} \mathop{\smile\!\!\!\!\!\!\,|}\nolimits^{iLs}_M (b_{\trianglerighteq \eta^\frown\langle j\rangle})_{j<i}$ for all $i < \omega$. This is indeed enough, because $(b_{\trianglerighteq \eta^\frown\langle i\rangle})_{i<\omega}$ is M-indiscernible by str-indiscernibility over M, and so the result follows by Proposition 2.4.

We prove the claim by showing that for all $i < \omega$ and all finite $b \subseteq b_{\trianglerighteq \eta^\frown\langle i\rangle}$ and $b' \subseteq (b_{\trianglerighteq \eta^\frown\langle j\rangle})_{j<i}$ we have $b \mathop{\smile\!\!\!\!\!\!\,|}\nolimits^{iLs}_M b'$, which is enough by Proposition 2.2. Let $\bar{\nu}_{i_1}, \dots, \bar{\nu}_{i_n}$ be finite tuples in $\mathcal{T}_\omega$ such that $i_1 < \cdots < i_n < \omega$ and $\bigwedge \bar{\nu}_{i_k} \trianglerighteq \eta^\frown\langle i_k\rangle$ for all $1 \le k \le n$. By str-basing there are $\gamma, \bar{\mu}_{i_1}, \dots, \bar{\mu}_{i_n}$ in $\mathcal{T}_\alpha$ such that $\gamma\bar{\mu}_{i_1}\cdots\bar{\mu}_{i_n}$ has the same str-quantifier-free type as $\eta\bar{\nu}_{i_1}\cdots\bar{\nu}_{i_n}$ and $b_\eta b_{\bar{\nu}_{i_1}}\cdots b_{\bar{\nu}_{i_n}} \equiv_M a_\gamma a_{\bar{\mu}_{i_1}}\cdots a_{\bar{\mu}_{i_n}}$. We now have reduced the problem to showing that $a_{\bar{\mu}_{i_n}} \mathop{\smile\!\!\!\!\!\!\,|}\nolimits^{iLs}_M a_{\bar{\mu}_{i_1}}\cdots a_{\bar{\mu}_{i_{n-1}}}$. As $\gamma \triangleleft \bigwedge \bar{\mu}_{i_n}$, there must be some $m < \omega$ such that $\bigwedge \bar{\mu}_{i_n} \trianglerighteq \gamma^\frown\langle m\rangle$. Furthermore, we have for every $1 \le k < n$ that $\gamma \triangleleft \bigwedge \bar{\mu}_{i_k}$ and $\bigwedge \bar{\mu}_{i_k} <_{\text{lex}} \bigwedge \bar{\mu}_{i_n}$, and so $\bigwedge \bar{\mu}_{i_k} \trianglerighteq \gamma^\frown\langle j\rangle$ for some $j < m$. Because $(a_\eta)_{\eta\in\mathcal{T}_\alpha}$ is spread out over M we have $a_{\trianglerighteq \gamma^\frown\langle m\rangle} \mathop{\smile\!\!\!\!\!\!\,|}\nolimits^{iLs}_M (a_{\trianglerighteq \gamma^\frown\langle j\rangle})_{j<m}$, and so $a_{\bar{\mu}_{i_n}} \mathop{\smile\!\!\!\!\!\!\,|}\nolimits^{iLs}_M a_{\bar{\mu}_{i_1}}\cdots a_{\bar{\mu}_{i_{n-1}}}$, as required. □

Corollary 3.6. *Let T be a thick theory, and let C be some parameter set and κ some cardinal. Set $\lambda = \beth_{(2^{\kappa+2^{\lambda_T+|C|}})^+}$. Given a tree $(a_\eta)_{\eta\in\mathcal{T}_\lambda}$ of κ-tupes that is s-indiscernible and spread out over C, there is a Morley tree $(b_\eta)_{\eta\in\mathcal{T}_\omega}$ over C that is str-Ls-based on $(a_\eta)_{\eta\in\mathcal{T}_\lambda}$ over C. The latter means that for any finite tuple $\bar{\eta} \in \mathcal{T}_\omega$ there is $\bar{\nu} \in \mathcal{T}_\lambda$ such that $\bar{\eta}$ and $\bar{\nu}$ have the same str-quantifier-free type and $b_{\bar{\eta}} \equiv^{\text{Ls}}_C a_{\bar{\nu}}$.*

Proof. By [Dobrowolski and Kamsma 2022, Fact 2.12] there is λ_T-saturated $M \supseteq C$ with $|M| \le 2^{\lambda_T+|C|}$. As $\kappa + |T| + |M| \le \kappa + |T| + 2^{\lambda_T+|C|} = \kappa + 2^{\lambda_T+|C|}$, we can use Fact 3.4(ii) to find a tree $(b_\eta)_{\eta\in\mathcal{T}_\omega}$ that is str-indiscernible over M and str-based on $(a_\eta)_{\eta\in\mathcal{T}_\lambda}$ over M. In particular $(b_\eta)_{\eta\in\mathcal{T}_\omega}$ is str-based on $(a_\eta)_{\eta\in\mathcal{T}_\lambda}$ over C, so it is spread out over C by Lemma 3.5 and hence it is a Morley tree over C. Finally, by str-basing, we have that for any finite tuple $\bar{\eta} \in \mathcal{T}_\omega$ there is $\bar{\nu} \in \mathcal{T}_\lambda$ such that $\bar{\eta}$ and $\bar{\nu}$ have the same str-quantifier-free type and $b_{\bar{\eta}} \equiv_M a_{\bar{\nu}}$. By our choice of M this implies $b_{\bar{\eta}} \equiv^{\text{Ls}}_C a_{\bar{\nu}}$, as required. □

The following key lemma in constructing spread out trees is due to N. Ramsey, for which we take terminology from [Chernikov et al. 2023, Definition 1.14].

Definition 3.7. We call a sequence of trees $((a^i_\eta)_{\eta\in\mathcal{T}_\alpha})_{i<\omega}$ *mutually s-indiscernible over C* if $(a^i_\eta)_{\eta\in\mathcal{T}_\alpha}$ is s-indiscernible over $C((a^j_\eta)_{\eta\in\mathcal{T}_\alpha})_{j\neq i, j<\omega}$ for all $i < \omega$.

Lemma 3.8. *Let T be a thick theory and let $(a_\eta)_{\eta\in\mathcal{T}_\alpha}$ be a tree that is s-indiscernible over M. Then there is a Morley sequence $((a^i_\eta)_{\eta\in\mathcal{T}_\alpha})_{i<\omega}$ in some global M-Ls-invariant type with $(a^0_\eta)_{\eta\in\mathcal{T}_\alpha} = (a_\eta)_{\eta\in\mathcal{T}_\alpha}$ that is mutually s-indiscernible over M.*

Proof. Let $q((x_\eta)_{\eta\in\mathcal{T}_\alpha}) \supseteq \mathrm{tp}((a_\eta)_{\eta\in\mathcal{T}_\alpha}/M)$ be a global M-Ls-invariant type. Let $N \supseteq M$ be $(2^{|M|+\lambda_T})^+$-saturated, and let $(a'_\eta)_{\eta\in\mathcal{T}_\alpha} \models q|_N$. Apply the s-modelling theorem (Fact 3.4(i)) to find a tree $(a''_\eta)_{\eta\in\mathcal{T}_\alpha}$ that is s-indiscernible over N and EM_s-based on $(a'_\eta)_{\eta\in\mathcal{T}_\alpha}$ over N.

Claim 3.8.1. *The type* $\mathrm{tp}((a''_\eta)_{\eta\in\mathcal{T}_\alpha}/N)$ *is M-Ls-invariant.*

Proof of claim. By Proposition 2.2 it is enough to show that for any finite $b, b' \in N$ with $\mathrm{d}_M(b, b') \leq 1$ we have $\mathrm{d}_M((x_\eta)_{\eta\in\mathcal{T}_\alpha}b, (x_\eta)_{\eta\in\mathcal{T}_\alpha}b') \leq 2 \subseteq \mathrm{tp}((a''_\eta)_{\eta\in\mathcal{T}_\alpha}/N)$. By thickness we have that $\mathrm{d}_M((x_\eta)_{\eta\in\mathcal{T}_\alpha}b, (x_\eta)_{\eta\in\mathcal{T}_\alpha}b') \leq 2$ is given by

$$\bigcup\{\mathrm{d}_M(x_{\bar\eta}b, x_{\bar\eta}b') \leq 2 : \bar\eta \text{ is a finite tuple in } \mathcal{T}_\alpha\}.$$

Let $\bar\eta$ be any finite tuple in $\mathcal{T}_\alpha$. For any $\bar\nu$ that has the same s-quantifier-free type as $\bar\eta$ we have that $\mathrm{d}_M(x_{\bar\nu}b, x_{\bar\nu}b') \leq 2 \subseteq \mathrm{tp}((a'_\eta)_{\eta\in\mathcal{T}_\alpha}/N)$ by Proposition 2.2, because $\mathrm{tp}((a'_\eta)_{\eta\in\mathcal{T}_\alpha}/N) = q|_N$ is M-Ls-invariant. We thus see that $\mathrm{d}_M(x_{\bar\eta}b, x_{\bar\eta}b') \leq 2 \subseteq \mathrm{EM}_s((a'_\eta)_{\eta\in\mathcal{T}_\alpha}/N) \subseteq \mathrm{tp}((a''_\eta)_{\eta\in\mathcal{T}_\alpha}/N)$, which concludes the proof of the claim. □

By Claim 3.8.1, Proposition 2.3 and our choice of N there is a unique global M-Ls-invariant type $q''((x_\eta)_{\eta\in\mathcal{T}_\alpha}) \supseteq \mathrm{tp}((a''_\eta)_{\eta\in\mathcal{T}_\alpha}/N)$. Let $((b^i_\eta)_{\eta\in\mathcal{T}_\alpha})_{i<\omega}$ be a Morley sequence in q'' over N.

Claim 3.8.2. *The sequence* $(b^i_\eta)_{\eta\in\mathcal{T}_\alpha}$ *is mutually s-indiscernible over N.*

Proof of claim. Fix $i < \omega$. We prove by induction on $k \geq i$ that $(b^i_\eta)_{\eta\in\mathcal{T}_\alpha}$ is s-indiscernible over $N((b^j_\eta)_{\eta\in\mathcal{T}_\alpha})_{j\neq i, j<k}$.

For the base case $k = i$ we need to prove that $(b^i_\eta)_{\eta\in\mathcal{T}_\alpha}$ is s-indiscernible over $N((b^j_\eta)_{\eta\in\mathcal{T}_\alpha})_{j<i}$. Let $\bar\eta, \bar\nu \in \mathcal{T}_\alpha$ be finite tuples with the same s-quantifier-free type. As $(b^i_\eta)_{\eta\in\mathcal{T}_\alpha} \equiv_N (a''_\eta)_{\eta\in\mathcal{T}_\alpha}$, we have that it is s-indiscernible over N. So there is a single type (after renaming variables) $p(y) = \mathrm{tp}(b^i_{\bar\eta}/N) = \mathrm{tp}(b^i_{\bar\nu}/N)$, which is M-Ls-invariant by Claim 3.8.1. Since $q''(x_{\bar\eta})$ and $q''(x_{\bar\nu})$ are both global M-Ls-invariant extensions of $p(y)$ we have that $q''(x_{\bar\eta}) = q''(x_{\bar\nu})$, after renaming variables. By construction $b^i_{\bar\eta} \models q''(x_{\bar\eta})|_{N((b^j_\eta)_{\eta\in\mathcal{T}_\alpha})_{j<i}}$ and $b^i_{\bar\nu} \models q''(x_{\bar\nu})|_{N((b^j_\eta)_{\eta\in\mathcal{T}_\alpha})_{j<i}}$, so $b^i_{\bar\eta} \equiv_{N((b^j_\eta)_{\eta\in\mathcal{T}_\alpha})_{j<i}} b^i_{\bar\nu}$ follows, as required.

For the successor step we have $k > i$ and we assume that $(b^i_\eta)_{\eta\in\mathcal{T}_\alpha}$ is s-indiscernible over $N((b^j_\eta)_{\eta\in\mathcal{T}_\alpha})_{j\neq i, j<k}$. Let $\bar\eta, \bar\nu \in \mathcal{T}_\alpha$ be finite tuples with the same s-quantifier-free type. By the induction hypothesis we have

$$b^i_{\bar\eta} \equiv^{\mathrm{Ls}}_{N((b^j_\eta)_{\eta\in\mathcal{T}_\alpha})_{j\neq i, j<k}} b^i_{\bar\nu},$$

where we get equivalence of Lascar-strong types instead of just normal types from s-indiscernibility; see e.g., [Kamsma 2024, Proposition 4.5]. As $(b^k_\eta)_{\eta\in\mathcal{T}_\alpha}$ realises an M-Ls-invariant type over $N((b^j_\eta)_{\eta\in\mathcal{T}_\alpha})_{j<k}$ and $N \supseteq M$ we get

$$(b^k_\eta)_{\eta\in\mathcal{T}_\alpha} b^i_{\bar\eta} \equiv^{\mathrm{Ls}}_{N((b^j_\eta)_{\eta\in\mathcal{T}_\alpha})_{j\neq i, j<k}} (b^k_\eta)_{\eta\in\mathcal{T}_\alpha} b^i_{\bar\nu},$$

which completes the induction step and thus the proof of the claim. □

We have $(b^0_\eta)_{\eta\in\mathcal{T}_\alpha} \equiv_M (a''_\eta)_{\eta\in\mathcal{T}_\alpha} \equiv_M (a'_\eta)_{\eta\in\mathcal{T}_\alpha} \equiv_M (a_\eta)_{\eta\in\mathcal{T}_\alpha}$, where the middle equality of types follows because $(a'_\eta)_{\eta\in\mathcal{T}_\alpha}$ is s-indiscernible over M and so its EM_s-type over M is maximal (i.e., is the same as its type over M) and $(a''_\eta)_{\eta\in\mathcal{T}_\alpha}$ is in particular EM_s-based on $(a'_\eta)_{\eta\in\mathcal{T}_\alpha}$ over M. So by an automorphism we find $((a^i_\eta)_{\eta\in\mathcal{T}_\alpha})_{i<\omega} \equiv_M ((b^i_\eta)_{\eta\in\mathcal{T}_\alpha})_{i<\omega}$, with $(a^0_\eta)_{\eta\in\mathcal{T}_\alpha} = (a_\eta)_{\eta\in\mathcal{T}_\alpha}$, which is then as required by construction of $((b^i_\eta)_{\eta\in\mathcal{T}_\alpha})_{i<\omega}$ and Claim 3.8.2. □

Remark 3.9. Lemma 3.8 is in fact a missing ingredient in [Kaplan and Ramsey 2020], in particular in the inductive steps in their Lemmas 5.11 and 6.4. There they replace some spread out tree A by an s-indiscernible tree B locally based on A (in our terminology: EM_s-based). However, this process might not preserve the property of being spread out. By replacing the inductive step by Lemma 3.8, the argument can be fixed.

In existing work on Kim-independence over arbitrary sets there is the same issue, as discussed in [Chernikov et al. 2023, page 7]. This can be fixed in a similar manner: [loc. cit., Lemma 1.15] is a variant of Lemma 3.8 over arbitrary sets (in full first-order logic), and can then be used in the inductive steps in the same way.

We also remark that this is not an issue in [Dobrowolski and Kamsma 2022], because the proofs there make use of a different notion called "q-spread-out". The point of this notion is that it is type-definable, so it can be captured by the EM_s-type. The gap in the proof of the Independence Theorem that this corrigendum addresses is of a different nature.

The following lemma illustrates the use of Lemma 3.8 and completes the proof of Lemma 3.3.

Lemma 3.10. *Let T be a thick theory. For any a and M there is a tree Morley sequence $(a_i)_{i<\omega}$ over M with $a_0 = a$.*

Proof. Let λ be the cardinal from Corollary 3.6, where M and $|a|$ take the respective roles of C and κ there. By induction on $\alpha \leq \lambda$ we will construct trees $(a^\alpha_\eta)_{\eta\in\mathcal{T}_\alpha}$, such that:

(1) For all $\eta \in \mathcal{T}_\alpha$ we have $a^\alpha_\eta \equiv_M a$.

(2) The tree $(a^\alpha_\eta)_{\eta\in\mathcal{T}_\alpha}$ is spread out and s-indiscernible over M.

(3) For all $\beta < \alpha$ we have $a^\alpha_{\iota_{\beta\alpha}(\eta)} = a^\beta_\eta$ for all $\eta \in \mathcal{T}_\beta$.

We start by setting $a^0_\varnothing = a$. For a limit stage ℓ, we set $a^\ell_{\iota_{\beta\ell}(\eta)} = a^\beta_\eta$, where β ranges over all ordinals $< \ell$ and η ranges over all elements in $\mathcal{T}_\beta$. This is well-defined by property (3), and properties (1) and (2) follow immediately from the induction hypothesis.

For the successor step we suppose $(a^\alpha_\eta)_{\eta\in\mathcal{T}_\alpha}$ has been constructed. By Lemma 3.8 we find a Morley sequence $((a^\alpha_{\eta,i})_{\eta\in\mathcal{T}_\alpha})_{i<\omega}$ in some global M-Ls-invariant type with $(a^\alpha_{\eta,0})_{\eta\in\mathcal{T}_\alpha} = (a^\alpha_\eta)_{\eta\in\mathcal{T}_\alpha}$ that is mutually s-indiscernible over M. Define a tree $(b_\eta)_{\eta\in\mathcal{T}_{\alpha+1}}$ by setting $b_\varnothing = a$ and $b_{\langle i\rangle^\frown\eta} = a^\alpha_{\eta,i}$ for all $\eta \in \mathcal{T}_\alpha$ and $i < \omega$. The EM_s-type of $(b_\eta)_{\eta\in\mathcal{T}_{\alpha+1}}$ over M satisfies the following properties:

(i) It contains $\text{tp}((b_\eta)_{\eta\in\mathcal{T}_{\alpha+1}\setminus\{\varnothing\}}/M)$. This is because $(b_{\trianglerighteq\langle i\rangle})_{i<\omega}$ forms an M-indiscernible sequence of trees that is mutually s-indiscernible over M.

(ii) The EM_s-type specifies that the type of the root is $\text{tp}(a/M)$.

We apply Fact 3.4(i) to find an s-indiscernible tree $(a^{\alpha+1}_\eta)_{\eta\in\mathcal{T}_{\alpha+1}}$ over M that is EM_s-based over M on $(b_\eta)_{\eta\in\mathcal{T}_{\alpha+1}}$. By an automorphism and (i) we may assume that $a^{\alpha+1}_{\langle i\rangle^\frown\eta} = b_{\langle i\rangle^\frown\eta} = a^\alpha_{\eta,i}$ for all $\eta \in \mathcal{T}_\alpha$ and $i < \omega$, and so (3) is satisfied. This then also implies that (2) is satisfied and (1) is satisfied by (ii), completing the inductive construction.

We thus have constructed a tree $(a^\lambda_\eta)_{\eta\in\mathcal{T}_\lambda}$ that is spread out and s-indiscernible over M with $a^\lambda_\eta \equiv_M a$ for all $\eta \in \mathcal{T}_\lambda$. We can now apply Corollary 3.6 to find a Morley tree $(a_\eta)_{\eta\in\mathcal{T}_\omega}$ that is str-Ls-based on $(a^\lambda_\eta)_{\eta\in\mathcal{T}_\lambda}$ over M. In particular $a_\eta \equiv_M a$ for all $\eta \in \mathcal{T}_\omega$, and so by an automorphism we may assume $a_{\zeta_0} = a$. Then setting $a_i = a_{\zeta_i}$ for all $i < \omega$ we obtain the required tree Morley sequence $(a_i)_{i<\omega}$. □

4. The independence theorem

We now give a new proof of the independence theorem [Dobrowolski and Kamsma 2022, Theorem 7.7]. The statement remains exactly the same. The proof is essentially that of [Kaplan and Ramsey 2020, Theorem 6.5], with Lemma 3.8 mixed in.

Lemma 4.1. *Let T be a thick NSOP$_1$ theory. Suppose that $a \mathop{\smile\!\!\!\!\!\mid}^K_M b$ and fix some cardinal κ. Suppose that $q(x, y) = \text{tp}(N/\mathfrak{M})$ is a global M-Ls-invariant type such that $q|_x$ extends* $\text{Lstp}(b/M)$, *where $N \supseteq M$ is $\beth_\omega(\lambda_T + |Mab| + |\mathcal{T}_\kappa|)$-saturated and the x variable matches b. If $(b_\eta)_{\eta\in\mathcal{T}_\alpha}$, with $\alpha \leq \kappa$, is a tree that is spread out over M, such that for all $\eta \in \mathcal{T}_\alpha$ we have $b_\eta \equiv^{\text{Ls}}_M b$ and $b_\eta \models (q|_x)|_{Mb_{\triangleright\eta}}$, then, writing $p(x, b) = \text{tp}(a/Mb)$,*

$$\bigcup_{\eta\in\mathcal{T}_\alpha} p(x, b_\eta)$$

does not Kim-divide over M.

Proof. We follow the proof of [Kaplan and Ramsey 2020, Lemma 6.2], replacing their use of [loc. cit., Proposition 6.1] by [Dobrowolski and Kamsma 2022, Proposition 7.5]. The proof is by induction on α. For $\alpha = 0$ there is nothing to do, and limit stages follow from the induction hypothesis by finite character. Now suppose that $(b_\eta)_{\eta\in\mathcal{T}_{\alpha+1}}$ is as in the statement. By the induction hypothesis we have that

$$\bigcup_{\eta\trianglerighteq\langle 0\rangle} p(x, b_\eta)$$

does not Kim-divide over M. Because $(b_\eta)_{\eta\in\mathcal{T}_{\alpha+1}}$ is spread out we have that $(b_{\trianglerighteq\langle i\rangle})_{i<\omega}$ is a Morley sequence in some global M-Ls-invariant type. By the chain condition Lemma 2.1 we then have that

$$\bigcup_{i<\omega}\bigcup_{\eta\trianglerighteq\langle i\rangle} p(x, b_\eta)$$

does not Kim-divide over M. At the same time we have $b_\varnothing \models (q|_x)|_{Mb_{\triangleright\varnothing}}$ and so by our assumptions on q we have $b_\varnothing \mathop{\smile\!\!\!\!\!\mid}^{*}_{M} b_{\triangleright\varnothing}$; see [Dobrowolski and Kamsma 2022, Definition 7.3]. Using that $p(x, b_\varnothing)$ does not Kim-divide (because $a \mathop{\smile\!\!\!\!\!\mid}^{K}_{M} b$), we can apply the weak independence theorem [loc. cit., Proposition 7.5] to see that

$$p(x, b_\varnothing)\cup\bigcup_{i<\omega}\bigcup_{\eta\trianglerighteq\langle i\rangle} p(x, b_\eta)$$

does not Kim-divide (here we implicitly used the assumption that $b_\eta \equiv^{\mathrm{Ls}}_M b$ for all $\eta\in\mathcal{T}_{\alpha+1}$). Unfolding definitions, this is exactly saying that

$$\bigcup_{\eta\in\mathcal{T}_{\alpha+1}} p(x, b_\eta),$$

does not Kim-divide, completing the induction step and thereby the proof. □

Lemma 4.2 (zig-zag lemma). *Let T be a thick NSOP$_1$ theory. Suppose that $b \mathop{\smile\!\!\!\!\!\mid}^{K}_{M} c$. Then there is a global M-Ls-invariant type $q(x, y) = \mathrm{tp}(N/\mathfrak{M})$, where $N \supseteq M$ is some $\beth_\omega(\lambda_T + |Mbc|)$-saturated model and $q|_x$ extends $\mathrm{tp}(b/M)$, and a tree Morley sequence $(b_i, c_i)_{i<\omega}$ over M such that*:

(i) *If $i \le j$ then $b_i c_j \equiv_M bc$.*

(ii) *If $i > j$ then $b_i \models (q|_x)|_{Mc_j}$.*

Proof. We basically verify that the proof of [Kaplan and Ramsey 2020, Lemma 6.4] goes through, while fixing a gap by mixing in a use of Lemma 3.8 (see also Remark 3.9).

Let λ be the cardinal from Corollary 3.6, where the C and κ are M and $|bc|$ respectively. Let $N \supseteq Mb$ be $\beth_\omega(|\mathcal{T}_\lambda|)$-saturated (note that $|\mathcal{T}_\lambda| \ge \lambda_T + |Mbc|$). Let $q(x, y)$ be a global M-Ls-invariant extension of $\mathrm{Lstp}(N/M)$, where the x variable matches b. In particular, for $\beta \models q|_x$ we have $\beta \equiv^{\mathrm{Ls}}_M b$. We write $p(z, b) = \mathrm{tp}(c/Mb)$. By induction on $\alpha \le \lambda$ we will construct trees $(b^\alpha_\eta, c^\alpha_\eta)_{\eta\in\mathcal{T}_\alpha}$, such that:

(1) For all $\eta\in\mathcal{T}_\alpha$ we have $b^\alpha_\eta \models (q|_x)|_{Mb^\alpha_{\triangleright\eta}c^\alpha_{\triangleright\eta}}$ and $b_\eta \equiv^{\mathrm{Ls}}_M b$.

(2) For all $\eta\in\mathcal{T}_\alpha$ we have $c^\alpha_\eta \models \bigcup_{\nu\trianglerighteq\eta} p(z, b^\alpha_\nu)$.

(3) The tree $(b^\alpha_\eta, c^\alpha_\eta)_{\eta\in\mathcal{T}_\alpha}$ is spread out and s-indiscernible over M.

(4) For all $\beta < \alpha$ we have $b^\alpha_{\iota_{\beta\alpha}(\eta)} c^\alpha_{\iota_{\beta\alpha}(\eta)} = b^\beta_\eta c^\beta_\eta$ for all $\eta\in\mathcal{T}_\beta$.

We start by setting $b^0_\varnothing c^0_\varnothing = bc$. For a limit stage ℓ, we set $b^\ell_{\iota_{\beta\ell}(\eta)} c^\ell_{\iota_{\beta\ell}(\eta)} = b^\beta_\eta c^\beta_\eta$, where β ranges over all ordinals $< \ell$ and η ranges over all elements in $\mathcal{T}_\beta$. This is well-defined by property (4), and properties (1)–(3) then follow immediately from the induction hypothesis.

For the successor step we suppose $(b^\alpha_\eta, c^\alpha_\eta)_{\eta \in \mathcal{T}_\alpha}$ has been constructed. Using Lemma 3.8 we find a Morley sequence $((b^\alpha_{\eta,i}, c^\alpha_{\eta,i})_{\eta \in \mathcal{T}_\alpha})_{i<\omega}$ in some global M-Ls-invariant type with $(b^\alpha_{\eta,0}, c^\alpha_{\eta,0})_{\eta \in \mathcal{T}_\alpha} = (b^\alpha_\eta, c^\alpha_\eta)_{\eta \in \mathcal{T}_\alpha}$ that is mutually s-indiscernible over M. Define a tree $(d_\eta, e_\eta)_{\eta \in \mathcal{T}_{\alpha+1}}$ by setting $d_{\langle i \rangle ^\frown \eta} e_{\langle i \rangle ^\frown \eta} = b^\alpha_{\eta,i} c^\alpha_{\eta,i}$ for all $\eta \in \mathcal{T}_\alpha$ and $i < \omega$. This leaves us to define $d_\varnothing$ and $e_\varnothing$. Let $\beta \models q|_x$ and pick $d_\varnothing$ such that

$$d_\varnothing \equiv^{\mathrm{Ls}}_{M d_{\rhd \varnothing} e_{\rhd \varnothing}} \beta.$$

We can then apply Lemma 4.1 to the tree $(d_\eta)_{\eta \in \mathcal{T}_{\alpha+1}}$ to see that

$$\bigcup_{\eta \in \mathcal{T}_{\alpha+1}} p(z, d_\eta)$$

does not Kim-divide over M. In particular, this set is consistent and so we can let $e_\varnothing$ be a realisation of this set. The EM_s-type of $(d_\eta, e_\eta)_{\eta \in \mathcal{T}_{\alpha+1}}$ over M satisfies the following properties:

(i) It contains $\mathrm{tp}((d_\eta, e_\eta)_{\eta \in \mathcal{T}_{\alpha+1} \setminus \{\varnothing\}}/M)$. This is because $(d_{\trianglerighteq \langle i \rangle}, e_{\trianglerighteq \langle i \rangle})_{i<\omega}$ forms an M-indiscernible sequence sequence of trees that is mutually s-indiscernible over M.

(ii) It contains the type $r(x_\varnothing, (x_\eta)_{\eta \rhd \varnothing}, (z_\eta)_{\eta \rhd \varnothing}) = \mathrm{tp}(d_\varnothing, d_{\rhd \varnothing}, e_{\rhd \varnothing}/M)$, and note that by construction $r(x, d_{\rhd \varnothing}, e_{\rhd \varnothing}/M) = (q|_x)|_{M d_{\rhd \varnothing} e_{\rhd \varnothing}}$. Indeed, let $\bar{\eta}$ and $\bar{\nu}$ be two finite tuples in $\mathcal{T}_{\alpha+1}$ with the same s-quantifier free type that do not contain the root. Then we have $d_{\bar{\eta}} e_{\bar{\eta}} \equiv^{\mathrm{Ls}}_M d_{\bar{\nu}} e_{\bar{\nu}}$, see (i) for the justification. The claim then follows from M-Ls-invariance of q.

(iii) It captures that $d_\varnothing \equiv^{\mathrm{Ls}}_M d_{\langle i \rangle}$ for all $i < \omega$. By construction we have $\mathrm{d}_M(d_\varnothing, d_{\langle 0 \rangle}) \leq n$ for some $n < \omega$, so $\mathrm{d}_M(d_\varnothing, d_{\langle i \rangle}) \leq n+1$ for all $i < \omega$. By thickness $\mathrm{d}_M(x_\varnothing, x_{\langle i \rangle}) \leq n+1$ is type-definable over M, and this partial type is thus contained in the EM_s-type.

(iv) It captures that $e_\varnothing \models \bigcup_{\nu \trianglerighteq \varnothing} p(z, d_\nu)$.

We apply Fact 3.4(i) to find an s-indiscernible tree $(b^{\alpha+1}_\eta, c^{\alpha+1}_\eta)_{\eta \in \mathcal{T}_{\alpha+1}}$ over M that is EM_s-based over M on $(d_\eta, e_\eta)_{\eta \in \mathcal{T}_{\alpha+1}}$. By an automorphism and (i) we may assume that $b^{\alpha+1}_{\langle i \rangle ^\frown \eta} c^{\alpha+1}_{\langle i \rangle ^\frown \eta} = d_{\langle i \rangle ^\frown \eta} e_{\langle i \rangle ^\frown \eta} = b^\alpha_{\eta,i} c^\alpha_{\eta,i}$ for all $\eta \in \mathcal{T}_\alpha$ and $i < \omega$, and so (4) is satisfied. This then also implies that (3) is satisfied. Finally, (1) is satisfied because of (ii) and (iii) and (2) is satisfied because of (iv), in both cases combined with the induction hypothesis. This completes the inductive construction.

We thus have constructed a tree $(b^\lambda_\eta, c^\lambda_\eta)_{\eta \in \mathcal{T}_\lambda}$ satisfying (1)–(3). We now apply Corollary 3.6 to find a Morley tree $(b_\eta, c_\eta)_{\eta \in \mathcal{T}_\omega}$ over M str-Ls-based on $(b^\lambda_\eta, c^\lambda_\eta)_{\eta \in \mathcal{T}_\lambda}$

over M. Property (2) is clearly preserved under str-Ls-basing. To see that property (1) is preserved under str-Ls-basing we show that, for any $\eta \in \mathcal{T}_\omega$ and finite tuple $\bar{\nu}$ in $\mathcal{T}_\omega$, we have $b_\eta \models (q|_x)|_{Mb_{\bar{\nu}}c_{\bar{\nu}}}$. Indeed, by str-Ls-basing we find $\gamma, \bar{\mu} \in \mathcal{T}_\omega$ such that $\gamma\bar{\mu}$ has the same str-quantifier-free type as $\eta\bar{\nu}$ and $b_\eta b_{\bar{\nu}} c_{\bar{\nu}} \equiv^{\mathrm{Ls}}_M b^\lambda_\gamma b^\lambda_{\bar{\mu}} c^\lambda_{\bar{\mu}}$. Let $\beta \models q|_x$, then we have by M-Ls-invariance of $q|_x$ that

$$b_\eta b_{\bar{\nu}} c_{\bar{\nu}} \equiv^{\mathrm{Ls}}_M b^\lambda_\gamma b^\lambda_{\bar{\mu}} c^\lambda_{\bar{\mu}} \equiv \beta b^\lambda_{\bar{\mu}} c^\lambda_{\bar{\mu}} \equiv^{\mathrm{Ls}}_M \beta b_{\bar{\nu}} c_{\bar{\nu}},$$

as required. So setting $(b_i, c_i) = (b_{\zeta_i}, c_{\zeta_i})$ for all $i < \omega$ we find our desired tree Morley sequence. □

Theorem 4.3 (independence theorem). *Let T be a thick NSOP$_1$ theory. Suppose that $a \equiv^{\mathrm{Ls}}_M a'$, $a \mathop{\smile\!\!\!\!\!\mid}^K_M b$, $a' \mathop{\smile\!\!\!\!\!\mid}^K_M c$ and $b \mathop{\smile\!\!\!\!\!\mid}^K_M c$. Then there is a'' with $a'' \equiv^{\mathrm{Ls}}_{Mb} a$, $a'' \equiv^{\mathrm{Ls}}_{Mc} a'$ and $a'' \mathop{\smile\!\!\!\!\!\mid}^K_M bc$.*

Proof. We now have all the tools in place to follow the proof of [Kaplan and Ramsey 2020, Theorem 6.5]. To get our conclusion about Lascar strong types, we apply the same trick as at the start of [Dobrowolski and Kamsma 2022, Theorem 7.7]: as described there we may assume b and c to enumerate λ_T-saturated e.c. models containing M. So we have reduced our goal to proving that, for $p_0(x, b) = \mathrm{tp}(a/Mb)$ and $p_1(x, c) = \mathrm{tp}(a'/Mc)$, the partial type $p_0(x, b) \cup p_1(x, c)$ does not Kim-divide over M.

Let $(b_i, c_i)_{i<\omega}$ and $q(x, y)$ be as in Lemma 4.2, and we may assume $b_1 c_1 = bc$. Let a'' be such that $a'' c_0 \equiv^{\mathrm{Ls}}_M a'c$, which can be done because $c = c_1 \equiv^{\mathrm{Ls}}_M c_0$. We then have $a \equiv^{\mathrm{Ls}}_M a''$ as well as $a \mathop{\smile\!\!\!\!\!\mid}^K_M b_1$, $a'' \mathop{\smile\!\!\!\!\!\mid}^K_M c_0$ and $b_1 \mathop{\smile\!\!\!\!\!\mid}^*_M c_0$, because $b_1 \models (q|_x)|_{Mc_0}$, so by [Dobrowolski and Kamsma 2022, Proposition 7.5] we have that $p_0(x, b_1) \cup p_1(x, c_0)$ does not Kim-divide over M. Since $(b_i, c_i)_{i<\omega}$ is a tree Morley sequence over M, we can apply both parts of Lemma 3.2 to see that $(b_{2i+1}, c_{2i})_{i<\omega}$ is a tree Morley sequence over M. Hence by Kim's lemma for tree Morley sequences (Lemma 3.3) we have that

$$\bigcup_{i<\omega} p_0(x, b_{2i+1}) \cup p_1(x, c_{2i})$$

is consistent. Thus

$$\bigcup_{i<\omega} p_0(x, b_{2i+1}) \cup p_1(x, c_{2i+2})$$

is consistent, as this is contained in the above set. Again, by Lemma 3.2, we have that $(b_{2i+1}, c_{2i+2})_{i<\omega}$ is a tree Morley sequence over M. Since $b_1 c_2 \equiv_M bc$ we thus have by Kim's lemma for tree Morley sequences (Lemma 3.3) again that $p_0(x, b) \cup p_1(x, c)$ does not Kim-divide over M, which finishes the proof. □

Acknowledgements

We would like to thank W. Kho [2023], who found the gap while working on his Master's thesis. We would also like to thank I. Kaplan and N. Ramsey for discussions leading to this proof.

References

[Chernikov and Ramsey 2016] A. Chernikov and N. Ramsey, "On model-theoretic tree properties", *J. Math. Log.* **16**:2 (2016), 1650009, 41. MR Zbl

[Chernikov et al. 2023] A. Chernikov, B. Kim, and N. Ramsey, "Transitivity, lowness, and ranks in NSOP_1 theories", *J. Symb. Log.* **88**:3 (2023), 919–946. MR Zbl

[Dobrowolski and Kamsma 2022] J. Dobrowolski and M. Kamsma, "Kim-independence in positive logic", *Model Theory* **1** (2022), 55–113. MR Zbl

[Kamsma 2024] M. Kamsma, "Positive indiscernibles", *Arch. Math. Logic* (2024).

[Kaplan and Ramsey 2020] I. Kaplan and N. Ramsey, "On Kim-independence", *J. Eur. Math. Soc. (JEMS)* **22**:5 (2020), 1423–1474. MR Zbl

[Kho 2023] W. Kho, *A primer of* NSOP_1 *theories*, master's thesis, University of Oxford, 2023. Zbl

Received 15 Jan 2024. Revised 27 Mar 2024.

JAN DOBROWOLSKI:

dobrowol@math.uni.wroc.pl
Department of Mathematics, University of Manchester, Manchester, United Kingdom

MARK KAMSMA:

mark@markkamsma.nl
School of Mathematical Sciences, Queen Mary University of London, London, UK

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the submission page.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles are usually in English or French, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not refer to bibliography keys. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and a Mathematics Subject Classification for the article, and, for each author, affiliation (if appropriate) and email address.

Format. Authors are encouraged to use LaTeX and the standard amsart class, but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should normally be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages — Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc. — allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with as many details as you can about how your graphics were generated.

Bundle your figure files into a single archive (using zip, tar, rar or other format of your choice) and upload on the link you been provided at acceptance time. Each figure should be captioned and numbered so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript with all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables.

White Space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Model Theory

no. 3 vol. 3 2024