EXPRESSING CARDINALITY QUANTIFIERS IN MONADIC SECOND-ORDER LOGIC OVER CHAINS†

VINCE BÁRÁNY, ŁUKASZ KAISER, AND ALEXANDER RABINOVICH

Abstract. We study the extension of monadic second-order logic of order with cardinality quantifiers "there exists infinitely many sets" and "there exists uncountably many sets". On linear orders that require the addition of only countably many points to be complete, we show using the composition method that the second-order uncountability quantifier can be reduced to the first-order uncountability quantifier. In particular, this shows that the extension of monadic second-order logic with this quantifier has the same expressive power as monadic second-order logic over ordinals and over countable scattered linear orders. Using a Ramsey-like theorem of Shelah for dense linear orders we show how to eliminate the uncountability quantifier over the ordering of the rationals. Hence, ultimately, we give an elimination procedure that works over all countable linear orders.

§1. Introduction. The study of extensions of first-order logic with cardinality quantifiers goes back to at least Mostowski [11]. For a cardinal κ the quantifier $\exists^{\kappa} x$ asks whether there are at least κ many *elements* with a given property. For this reason these cardinality quantifiers, along with their many generalisations to be found in the literature, e.g. the Magidor-Malitz quantifiers, can rightfully be called first order. Model theoretic properties and axiomatisability of first-order logic extended with the first-order uncountability quantifier $\exists^{\aleph_1} x$ as well as the possibility of eliminating $\exists^{\aleph_1} x$ have been widely investigated [8]. See also the book [2], which presents results on decidability and other properties of first-order logic extended with cardinality quantifiers over various natural classes of structures.

In this paper we take a look at the expressive power of the extension of monadic second-order logic of order by the cardinality quantifiers $\exists^{\aleph_1} X$ and $\exists^{2^{\aleph_0}} X$ meaning "there exists uncountably many sets X" and "there exists continuum many sets X", respectively.

Monadic second-order logic of order (henceforth MLO) plays a very important role in mathematical logic and computer science. The fundamental connection between MLO and automata was discovered independently by Büchi, Elgot and Trakhtenbrot [3, 6, 14, 15] when the logic was proved to be decidable over the class of finite linear orders. Büchi proved the decidability of MLO on $(\omega, <)$ [4] and later, together with Siefkes, extended this result to the monadic theory of every ordinal up to ω_1, including ω_1 itself [5]. Shelah continued the study of MLO on ordinals and showed that the MLO theory of any ordinal $\alpha < \omega_2$ is decidable [13].

†This research was facilitated by the ESF project AutoMathA.

For a cardinal κ and a formula ψ let $\exists^{\kappa} X\ \psi$ express that "there are at least κ many *sets* X such that ψ holds". We will denote by MLO($\exists^{\kappa}$) the extension of MLO by $\exists^{\kappa}$. Furthermore we will briefly consider extensions of MLO by the predicates $\mathbf{Inf}(X) \equiv$ "the set X is infinite" and $\mathbf{Unc}(X) \equiv$ "the set X is uncountable", denoted MLO(**Inf**) and MLO(**Unc**), respectively. This is of course equivalent to allowing the use of the first-order cardinality quantifiers $\exists^{\omega} x$ and $\exists^{\aleph_1} x$, respectively, counting *elements* with a certain property. We prefer the predicate notation to avoid confusion with the second-order cardinality quantifiers.

We call a linear order *almost complete* if its completion contains only countably many new points. This class includes all ordinals, all countable scattered linear orders and of course all complete linear orders. As a kind of starting point we will prove the following.

Theorem 1. *For every* MLO *formula* $\varphi(X, \overline{Y})$ *there exists an* MLO(**Unc**) *formula* $\psi(\overline{Y})$ *that is equivalent to* $\exists^{\aleph_1} X\, \varphi(\overline{Y})$ *over the class of almost complete linear orders.*

The predicate **Unc** is easily expressible in MLO over chains of certain order types including, e.g., all ordinals. In the case of the real line little can be said in the absence of further set theoretic assumptions. Recall that in [7] Gurevich has shown, assuming the continuum hypothesis, that $\mathbf{Unc}(X)$ is expressible in MLO over a class of linear orders including the real line and relying, in fact, solely on topological properties of these orderings.

Theorem 1 immediately yields complete elimination of the uncountability quantifier over countable scattered chains. Next we prove this for chains of order type of the rationals, which enables the extension to all countable chains. Our main results are summarised in the next two theorems.

Theorem 2 (Elimination of the uncountability quantifier).

(1) For every MLO($\exists^{\aleph_1}$) *formula* $\varphi(\overline{Y})$ *there exists an* MLO *formula* $\psi(\overline{Y})$ *that is equivalent to* $\varphi(\overline{Y})$ *over the class of all ordinals.*

(2) For every MLO($\exists^{\aleph_1}$) *formula* $\varphi(\overline{Y})$ *there exists an* MLO *formula* $\psi(\overline{Y})$ *that is equivalent to* $\varphi(\overline{Y})$ *over the class of all countable linear orders.*

Furthermore, in all these cases ψ *is computable from* φ*.*

In addition to the above, the reduction will show that over countable linear orders the quantifiers $\exists^{\aleph_1} X$ and $\exists^{2^{\aleph_0}} X$ are equivalent, i.e. that the continuum hypothesis holds for MLO-definable families of sets.

Theorem 3. *For every* MLO($\exists^{\aleph_1}, \exists^{2^{\aleph_0}}$) *formula* $\varphi(X, \overline{Y})$ $\exists^{\aleph_1} X\ \varphi(X, \overline{Y})$ *is equivalent to* $\exists^{2^{\aleph_0}} X\ \varphi(X, \overline{Y})$ *over the class of all countable linear orders.*

These results generalise those of Kuske and Lohrey [9, 10] dealing solely with the ordering $(\omega, <)$.

Note that all of our results extend to cardinality quantifiers $\exists^{\aleph_0} \overline{X}$, $\exists^{\aleph_1} \overline{X}$ and $\exists^{2^{\aleph_0}} \overline{X}$ counting tuples of sets. This follows from the fact that

$$\exists^{\kappa}(X_0, X_1)\varphi \ \equiv\ \exists^{\kappa} X_0\, \exists X_1\, \varphi \ \vee\ \exists^{\kappa} X_1\, \exists X_0\, \varphi$$

for any cardinal $\kappa \geq \aleph_0$.

Organisation. After recapitulating basic definitions and notations in Section 2 and handing the infinity quantifier in Section 3, we divide the discussion into several subsections gradually working our way towards establishing the above stated theorems. In section 2.2 we recall the composition method over linear orders, the main vehicle of our arguments. In section 4 we introduce the notion of U-intervals and D-intervals and finite U-U covers underlying all our proofs and pin down their fundamental properties. Our first result is the partial reduction of cardinality quantifiers over almost-complete linear orders as stated in Theorem 1. It is presented in section 5. As corollaries we obtain complete elimination over the class of ordinals as well as over countable scattered linear orders. Section 6 sees another partial elimination result effective over all linear orders showing that the application of cardinality quantifiers can be restricted to cuts, i.e. to downward-closed subsets, i.e. to points of the completion of any given linear order. In section 7 we show that $\exists^{\aleph_1}$ and $\exists^{2^{\aleph_0}}$ are equivalent and MLO-expressible over the ordering of the rationals. In section 8 we show how to eliminate the cardinality quantifiers over linear orders uniformly decomposable into sums of linear orders when uniform elimination over the summands and elimination over the index structure are at hand. Finally, in section 9 the results of the previous sections are combined to obtain uniform and effective elimination of cardinality quantifiers over all countable linear orders.

§2. Preliminaries.

For a number $l \in \mathbb{N}$, $l > 0$, an l-labelled linear order (or simply l-chain) is a structure $\mathcal{L} = (L, <, P_1, \ldots, P_l)$, where the P_i's are unary predicates and $(L, <)$ is a linear order.

We denote the standard ordering of natural numbers by $(\omega, <)$ or $(\mathbb{N}, <)$, the orderings of integers and rational numbers are denoted $(\mathbb{Z}, <)$ and $(\mathbb{Q}, <)$, respectively. Recall that $(\mathbb{Q}, <)$ is dense, i.e. between any two elements $x < y$ there is another element z such that $x < z < y$. A linear order $(L, <)$ is *scattered* if a non-trivial linear dense order cannot be embedded into $(L, <)$, or equivalently if $(\mathbb{Q}, <)$ cannot be embedded into it.

A subset I of a linear order $(L, <)$ is convex, if for all $x < y < z$ with $x, z \in I$ also $y \in I$. We use the word *intervals* referring to to all convex subsets, not just when they are bounded, or have endpoints in L. For intervals with endpoints $a, b \in L$, whether open or closed on any side, we will use the standard notation, such as $[a, b) = \{x \in L \mid a \leq x < b\}$, etc. Moreover, we write $L|_{[a,b)}$ for the order $L \cap [a, b)$, and for $X \subseteq L$ we use analogous notation, i.e. $X|_{[a,b]}$ for $X \cap [a, b]$. This notation is extended to k-tuples in the natural way, e.g. $\overline{X}|_I = (X_1|_I \ldots X_k|_I)$.

A linear order is *complete* if every one of its subsets has a least upper bound. In this paper a *cut* of a linearly ordered set $(L, <)$ is a downward closed set $C \subseteq L$ such that if C has a least upper bound in L then it is contained in C. A *proper cut* is a cut that has no least upper bound in L, i.e. one that has no maximal element. The *completion* of the linear order L, denoted $\overline{L}$, is defined as the linear order $(\mathcal{C}(L), \subsetneq)$ of cuts of L with the mapping $L \ni l \mapsto \{k \in L \mid k \leq l\} \in \mathcal{C}(L)$ as the canonical embedding.

2.1. Monadic logic of order.

We will work with l-chains in the relational signature $\{<, P_1, \ldots, P_l\}$ where $<$ is a binary relation symbol interpreted as a total ordering and the P_i's are unary predicates representing a labelling of the chain.

Monadic second-order logic of order, MLO for short, extends first-order logic by allowing quantification over *subsets* of the domain. MLO uses first-order variables $x, y, \dots$ interpreted as elements, and set variables $X, Y, \dots$ interpreted as subsets of the domain. Set variables will always be capitalised to distinguish them from first-order variables. The atomic formulas are $x < y$, $x \in P_i$ and $x \in X$, all other formulas are built from the atomic ones by applying boolean connectives and the universal and existential quantifiers for both kinds of variables. Concrete formulas will be given in this syntax, taking the usual liberties and short-hands, such as $X \cup Y, X \cap Y, X \subseteq Y$, guarded quantifiers and relativisations of formulas to a set.

The quantifier rank of a formula φ, denoted $\mathrm{qr}(\varphi)$, is the maximum depth of nesting of quantifiers in φ. For fixed n and l we denote by $\mathrm{Form}_{n,l}$ the set of formulas of quantifier depth $\leq n$ and with free variables among $X_1, \dots, X_l$. Let $n, l \in \mathbb{N}$ and $\mathcal{L}_1, \mathcal{L}_2$ be l-chains. We say that $\mathcal{L}_1$ and $\mathcal{L}_2$ are *n-equivalent*, denoted $\mathcal{L}_1 \equiv^n \mathcal{L}_2$, if for every $\varphi \in \mathrm{Form}_{n,l}$, $\mathcal{L}_1 \models \varphi$ iff $\mathcal{L}_2 \models \varphi$.

Clearly, $\equiv^n$ is an equivalence relation. For any $n \in \mathbb{N}$ and $l > 0$, the set $\mathrm{Form}_{n,l}$ is infinite. However, it contains only finitely many semantically distinct formulas, so there are only finitely many $\equiv^n$-classes of l-structures. In fact, we can compute representatives for these classes.

LEMMA 4 (Hintikka Lemma). *For each $n, l \in \mathbb{N}$, we can compute a* finite *set $H_{n,l} \subseteq \mathrm{Form}_{n,l}$ such that:*

- *For every l-chain $\mathfrak{A}$ there is a* unique *$\tau \in H_{n,l}$ such that $\mathfrak{T} \models \tau$.*
- *If $\tau \in H_{n,l}$ and $\varphi \in \mathrm{Form}_{n,l}$, then either $\tau \models \varphi$ or $\tau \models \neg\varphi$. Furthermore, there is an algorithm that, given such τ and φ, decides which of these two possibilities holds.*

Elements of $H_{n,l}$ are called (n, l)-Hintikka formulas.

Given an l-chain $\mathcal{L}$ we denote by $\mathrm{Tp}^n(\mathcal{L})$ the unique element of $H_{n,l}$ satisfied in $\mathcal{L}$ and call it the *n-type* of $\mathcal{L}$. Thus, $\mathrm{Tp}^n(\mathcal{L})$ determines (effectively) which formulas of quantifier-depth $\leq n$ are satisfied in $\mathcal{L}$.

We sometimes speak of the n-type of a tuple of subsets $\overline{V} = V_1, \dots, V_m$ of a given l-chain $\mathcal{L}$. This is synonymous with the n-type of the $(l+m)$-chain $(\mathcal{L}, \overline{V})$ obtained as an expansion of $\mathcal{L}$ with the predicates $P_{l+1}, \dots, P_{l+m}$ interpreted by the sets $V_1, \dots, V_m$ given. This type will be denoted by $\mathrm{Tp}^n(\mathcal{L}, \overline{V})$ and often referred to as an n-type in m variables, whereby the n-type of the $(l+m)$-structure $(\mathcal{L}, \overline{V})$ is understood. Moreover, when considering substructures, e.g. $\mathcal{L}' \subseteq \mathcal{L}$, and given sets $\overline{X} \subseteq \mathcal{L}$, we write $\mathrm{Tp}^n(\mathcal{L}', \overline{X})$ to denote $\mathrm{Tp}^n(\mathcal{L}', \overline{X} \cap \mathcal{L}')$.

2.2. The composition method. The essence of the composition method is that certain operations on structures, such as disjoint union and ordered sums of linear orders, can be projected to n-theories, i.e. there are corresponding operations mapping n-theories of constituent structures to the n-theory of the resulting structure. In other words, n-theories can be composed. The method was introduced by Shelah as an adaptation of the Feferman-Vaught Theorem to MLO [13].

DEFINITION 5 (Sums of k-chains). Given a linear order $(I,<)$ and a family of k-chains $\mathcal{L}_i = (L_i, <_i, \overline{P}_i)$ for every $i \in I$, the sum $\sum_I \mathcal{L}_i$ is defined as the k-chain over the set $\bigcup_{i\in I} L_i \times \{i\}$ such that

$$(l,i) < (l',i') \iff i < i' \text{ or } i = i' \text{ and } l <_i l',$$

with labels defined by $\overline{P}(l,i) \iff \overline{P}_i(l)$.

We write $\mathcal{L}_0 + \mathcal{L}_1$ for the sum over $(\{0,1\},<)$. For example $(\mathbb{Z},<)$ is isomorphic to $\omega^* + \omega$, where ω^* is the standard ordering of negative integers, $(\{\dots,-3,-2,-1\},<)$.

In its weakest form the composition theorem states that the n-type of a sum of chains, $\mathrm{Tp}^n\left(\sum_{i\in I}\mathcal{L}_i\right)$, is uniquely determined by the chain of n-types of the summands $\langle \mathrm{Tp}^n(\mathcal{L}_i) \mid i \in I\rangle$. A stronger statement is cited below as Theorem 24.

THEOREM 6 (Composition on linear orders I.).
Let $(I,<)$ be a linear ordering. If $(\mathcal{L}_i \mid i \in I\rangle$ and $\langle(\mathcal{L}'_i \mid i \in I\rangle$ are I-indexed sequences of chains such that $\mathrm{Tp}^n(\mathcal{L}_i) = \mathrm{Tp}^n(\mathcal{L}_i)$ *for all $i \in I$, then* $\mathrm{Tp}^n(\sum_{i\in I}\mathcal{L}_i) = \mathrm{Tp}^n(\sum_{i\in I}\mathcal{L}'_i)$.

§3. Infinity quantifier. Warming up, we observe that the second-order infinity quantifier $\exists^{\aleph_0} X$ can be eliminated uniformly over all structures with the aid of the predicate $\mathbf{Inf}(X)$ expressing that X is infinite or, equivalently, using the first-order infinity quantifier $\exists^{\aleph_0} x$.

PROPOSITION 7. *For every* $\mathrm{MLO}(\exists^{\aleph_0})$ *formula $\varphi(\overline{Y})$ there exists an* $\mathrm{MLO}(\mathbf{Inf})$ *formula $\psi(Y)$ equivalent to $\varphi(\overline{Y})$ over all structures.*

PROOF. Observe that the following are equivalent:

(1) There are only finitely many X which satisfy $\varphi(X,\overline{Y})$
(2) There is a finite set Z such that any two different sets X_1, X_2 which both satisfy $\varphi(X_i,\overline{Y})$ differ on Z, i.e.

$$\exists Z\Big(\neg\mathbf{Inf}(Z) \ \wedge\ \forall X_1X_2\ \big((\varphi(X_1,\overline{Y}) \wedge \varphi(X_2,\overline{Y}) \wedge X_1 \neq X_2) \ \rightarrow$$
$$\exists z \in Z\,(z \in X_1 \leftrightarrow z \notin X_2)\big)\Big).$$

Item (2) implies (1) as a collection of sets pairwise differing only on a finite set Z has cardinality at most $2^{|Z|}$. Conversely, if $X_1,\dots,X_k$ are all the sets that satisfy $\varphi(X_i,\overline{Y})$, then choose for every pair of distinct sets X_i, X_j an element $z_{i,j}$ in the symmetric difference of X_i and X_j and define Z as the set of the chosen elements. ⊣

As **Inf** is uniformly MLO-definable over all linear orders we have the following corollary.

COROLLARY 8. $\mathrm{MLO}(\exists^{\aleph_0})$ *collapses effectively to* MLO *over the class of linear orders.*

The converse of Proposition 7 holds as well. In fact, the predicate $\mathbf{Inf}(X)$ can be defined over all structures by the formula $\exists^\kappa Y\, Y \subseteq X$ for any $\aleph_0 \le \kappa \le 2^{\aleph_0}$. Therefore, by Proposition 7, any of the quantifiers $\exists^\kappa Y$ with $\aleph_0 < \kappa \le 2^{\aleph_0}$ can be used to define $\exists^{\aleph_0}$.

§4. U-D colouring of intervals. To eliminate a single occurrence of the uncountability quantifier from a formula $\exists^{\aleph_1} X\ \varphi(X, \overline{Y})$ we will make extensive use of the following notions for intervals.

Definition 9. Let $(L, <, X, \overline{Y})$ be a labelled chain such that $L \models \varphi(X, \overline{Y})$ and let I be an interval of $(L, <)$.

(1) I is a *U-interval* for φ, X, $\overline{Y}$ iff $L|_I \models \forall Z\, \tau(Z, \overline{Y}) \to Z = X$, where $\tau(X, \overline{Y})$ is the n-type of $(L|_I, X, \overline{Y})$ with n the quantifier rank of φ.
(2) I is a *D-interval* for φ, X, $\overline{Y}$ iff it is not a U-interval.
(3) I is an *unsplittable* D-interval for φ, X, $\overline{Y}$ iff it cannot be split into disjoint D-intervals.

Note that everything here can be formalised in MLO. For example, there is an MLO formula $DINT_\varphi(X, \overline{Y}, I)$ expressing that I is a D-interval for $\varphi, X, \overline{Y}$ and a formula $UNSP_\varphi(X, \overline{Y}, I)$ such that $L \models UNSP_\varphi(X, \overline{Y}, I)$ iff I is an unsplittable D-interval for $\varphi, X, \overline{Y}$.

Whenever $\varphi, X, \overline{Y}$ are clear from the context we will take the liberty of saying "I is an U-interval" instead of "I is U-interval for $\varphi, X, \overline{Y}$". Similarly, for D-intervals and unsplittable D-intervals.

Lemma 10. *If there is an infinite set of pairwise disjoint D-intervals for some X satisfying $\varphi(X, \overline{Y})$ then there are at least continuum many such X.*

Proof. If X is as in the assumption then the chain $(L, <, X, \overline{Y})$ can be written as a sum $\sum_{i \in I}(L_i, <_i, X \cap L_i, \overline{Y}|_{L_i})$, where I is countably infinite and for all $i \in I$ the interval L_i contains a D-interval, hence L_i is itself a D-interval for X. This means that for each $i \in I$ as above, there is a subset $X'_i \subseteq L_i$ such that

$$\mathrm{Tp}^n(L_i, <, X \cap L_i, \overline{Y}|_{L_i}) = \mathrm{Tp}^n(L_i, <, X'_i, \overline{Y}|_{L_i}).$$

We define for every subset $H \subseteq I$ the set

$$X_H = \bigcup \{X'_i \mid i \in H\} \cup \bigcup \{X \cap L_i \mid i \notin H\}.$$

It follows from the above by Theorem 6 that $\mathrm{Tp}^n(L, <, X_H, \overline{Y}) = \mathrm{Tp}^n(L, <, X, \overline{Y})$ for every $H \subseteq I$. Each of the continuum many X_H therefore satisfies $\varphi(X_H, \overline{Y})$ in L. ⊣

Definition 11 (Finite U-U cover). Let $(L, <, X, \overline{Y})$ be a labelled chain such that $L \models \varphi(X, \overline{Y})$. Let I be an interval. Intervals $I_1 \dots I_k$ constitute a *finite U-U cover* of I for $\varphi, X, \overline{Y}$ if $I = \bigcup_j I_j$ and each I_j is either a U-interval or an unsplittable D-interval for $\varphi, X, \overline{Y}$.

Again, we will most often take no mention of either φ, X, or $\overline{Y}$ when these are understood. As in the following observation.

Lemma 12. *If I has no finite U-U cover, then I can be split into two D-intervals such that one of them has no finite U-U cover.*

Proof. Because I has no finite U-U cover, it is necessarily a D-interval, but not an unsplittable D-interval. It can thus be split into D-intervals I_1, I_2 with $I_1 \cap I_2 = \emptyset$ and $I_1 \cup I_2 = I$. Now if both I_1 and I_2 had a finite U-U cover, then this would yield a finite U-U cover of I. Therefore either I_1 or I_2 has no finite U-U cover. ⊣

As a conclusion we obtain the following lemma.

LEMMA 13. *The following dichotomy holds (for each X):*

(1) either I contains infinitely many disjoint D-intervals (for X)
(2) or I has a finite U-U cover (for X).

PROOF. Assume that I has a finite U-U cover. Then it has a finite U-U cover $I_1, \ldots, I_k$ such that $I_i \cap I_j = \emptyset$ for all $1 \leq i \neq j \leq k$. As each I_j is either a U-interval or an unsplittable D-interval it cannot contain two disjoint D-intervals. This gives an upper bound $k + (k-1)$ on the size of any collection of pairwise disjoint D-intervals inside I: at most k D-intervals each contained properly in separate I_j's and at most $k-1$ D-intervals intersecting two or more of the I_j's.

Conversely, if $I_0 = I$ has no finite U-U cover then, by Lemma 12, it can be split into disjoint D-intervals I_1 and J_0, with I_1 having no finite U-U cover. Continuing in this manner we can inductively define D-intervals I_n and J_n for all $n \in \mathbb{N}$ such that each I_n is a D-interval with no finite U-U cover and $I_n \cap I_m = \emptyset$ for all $n \neq m$. ⊣

Next we refine the notion of finite U-U covers as follows.

DEFINITION 14 (Balanced cover).

(1) An unsplittable D-interval I is *left-balanced* iff $I_{<v} = \{w \in I \mid w < v\}$ is a U-interval for every $v \in I$.
(2) Similarly, an unsplittable D-interval I is *right-balanced* iff for every $v \in I$ the interval $I_{>v} = \{w \in I \mid w > v\}$ is a U-interval.
(3) An unsplittable D-interval I is *balanced* iff it is either left-balanced or right-balanced.
(4) A finite U-U cover $I_1, \ldots, I_k$ is *balanced* iff for each $1 \leq j \leq k$, I_j is either an U-interval or a balanced unsplittable interval.

LEMMA 15. *An interval I has a finite U-U cover for X iff I has a balanced U-U cover for X.*

PROOF. We show that every non-balanced unsplittable D-interval I for X can be split[1] into two intervals L and R constituting a balanced U-U cover of I for X. This immediately implies the conclusion of the lemma.

Let I be a non-balanced unsplittable D-interval for X. Because I is not right-balanced, there is a point $v \in I$ such that $I_{>v}$ is a D-interval, consequently, $I_{<v}$ must be a U-interval, since I cannot be split into two D-intervals. The following set is therefore not empty.

$$L = \{v \in I \mid I_{<v} \text{ is a U-interval for} X\}$$

By definition, L is a downward-closed subinterval of I, and it is either a U-interval or a left-balanced unsplittable D-interval.

Let $R = I \setminus L$. Because I is not left-balanced R cannot be empty. Then one of R or L is a U-interval. We have seen that if L is a D-interval then it is left-balanced. Similarly, we need to show that if R is a D-interval then it is

[1]It is important to point out that the split depends on X.

right-balanced. Notice that $I_{<v}$ is a D-interval for X for every $v \in R$, otherwise v would be in L. Therefore, since I cannot be split into disjoint D-intervals for X, $I_{>v}$ is a U-interval for X for every $v \in R$. Which means precisely that R is right-balanced. ⊣

LEMMA 16. *There is a function $N(k,l,\varphi)$ such that for every l-chain $\mathcal{L} = (L,<,\overline{Y})$ if $\{I_1,\dots,I_k\}$ is a balanced U-U cover of an interval I of $\mathcal{L}$ for each of $X_1,\dots,X_n$ then $n \le N(k,I,\varphi)$ or $X_i \cap I = X_j \cap I$ for some $i \neq j$.*

PROOF. Let K be the number of $\mathrm{qr}(\varphi)$-types in $l+1$ variables. Then, if J is a U-interval for $K+1$ sets then two of these must realise the same type on J and hence have to coincide on J. Assume now that J is left-balanced for $2K+1$ sets $X_1,\dots,X_{2K+1}$, so for each $v \in J$ the interval $J_{<v}$ is a U-interval for each of these sets X_i. If for each pair X_i, X_j with $i \neq j$ there is a point $p_{i,j} \in J$ on which these two sets differ, then all the $2K+1$ sets differ on the interval $J_{\le p}$ with $p = \max\{p_{i,j} \mid i,j \le 2K+1\}$. Therefore there are at least $K+1$ among them which are pairwise different on $J_{<p}$, which contradicts the fact that $J_{<p}$ is an U-interval for all of these. The case of right-balanced intervals is treated symmetrically.

Classify the sets X_i into 2^k classes according to which of the $I_1,\dots,I_k$ are left- or right-balanced for each X_i (considering U-intervals, say, as left-balanced). By the above, no class can contain more than $(2K)^k$ sets pairwise different on I. Therefore $N(k,l,\psi) = (4K)^k$ satisfies the claim. ⊣

Combining Lemmas 13, 15 and 16 we obtain the following criterion.

PROPOSITION 17. *Let $\mathcal{L} = (L,<,\overline{Y})$ be an l-chain and $\varphi(X,\overline{Y})$ an* MLO *formula. Then*

$$\mathcal{L} \models \neg\exists^{\aleph_1} X \ \varphi(X,\overline{Y})$$

if and only if there exists a countable subset U of the completion of L such that for every X satisfying $\varphi(X,\overline{Y})$ there is a finite balanced U-U cover of $\mathcal{L}$ the end-points of which lie in U.

Since U is a subset of the completion of $\mathcal{L}$, this condition cannot be directly expressed in MLO over $\mathcal{L}$. In certain special cases, however, we are able to use this criterion to eliminate the uncountability quantifier $\exists^{\aleph_1} X$.

§5. Almost complete linear orders. After these preparations we are ready to prove the first item of Theorem 2 concerning the collapse of MLO($\exists^{\aleph_1}$) to MLO over the class of all ordinals. This will be a corollary of the elimination step embodied in Theorem 1 and valid uniformly over all *almost complete* linear orders. Recall that a linear order L is *almost complete* if $\overline{L} \setminus L$ is countable.

THEOREM 1. For every MLO formula $\varphi(X,\overline{Y})$ there exists a MLO(**Unc**) formula $\psi(\overline{Y})$ that is equivalent to $\exists^{\aleph_1} X \ \varphi(X,\overline{Y})$ over the class of almost complete linear orders.

PROOF. We are going to show that over almost complete linear orders the condition stated in Proposition 17 can be formulated in MLO(**Unc**). First, let

U be as in Proposition 17 and let $V = U \cap L$. Note that $V \cup (\overline{L} \setminus L)$ is an over-approximation of U, which also fulfils the condition stated in Proposition 17.

Let M be a subset of L. Define an equivalence relation $\sim_M$ as follows: $x \sim_M y$ if $[x, y] \subseteq M$ or $[x, y]$ is disjoint from M. Note that for every M, the equivalence classes of $\sim_M$ are intervals in L. Moreover, the following can be formalised in MLO(**Unc**):

(i) The number of $\sim_M$ classes is finite.
(ii) Each $\sim_M$ class is a U-interval or an unsplittable D-interval.
(iii) Any extremal points of any of the $\sim_M$ classes are contained in V.

Observe that item (iii) is equivalent to the assertion that all end-points are in $V \cup (\overline{L} \setminus L)$. Note also, that if $I_0, \dots I_k$ are disjoint intervals that partition L, then there is M such that the I_i are $\sim_M$-equivalence classes. Indeed if for $i < j$ the interval I_i precedes I_j then we can take as M the set $I_0 \cup I_2 \cup \dots \cup I_{2\lfloor k/2 \rfloor}$.

The conditions of Proposition 17 can be formalised by an MLO formula expressing that there is a countable set V such that for all X satisfying $\varphi(X, \overline{Y})$ there is a set M such that the $\sim_M$ classes constitute a finite U-U cover for X and all extremal points of $\sim_M$ classes fall in V. $\dashv$

In cases where the uncountability predicate, equivalently, the first-order uncountability quantifier is MLO-definable the above technique can be used inductively to completely reduce MLO($\exists^{\aleph_1}$) to MLO. Two classes over which this is feasible are the class of countable scattered linear orders and the class of all ordinals.

Corollary 18.

(1) For every MLO($\exists^{\aleph_1}$) *formula* $\varphi(\overline{Y})$ *there exists an* MLO *formula* $\psi(\overline{Y})$ *that is equivalent to* $\varphi(\overline{Y})$ *over the class of countable scattered linear orders.*

(2) For every MLO($\exists^{\aleph_1}$) *formula* $\varphi(\overline{Y})$ *there exists an* MLO *formula* $\psi(\overline{Y})$ *that is equivalent to* $\varphi(\overline{Y})$ *over the class of all ordinals.*

(3) (Under CH) For every MLO($\exists^{\aleph_1}$) *formula* $\varphi(\overline{Y})$ *there exists an* MLO *formula* $\psi(\overline{Y})$ *that is equivalent to* $\varphi(\overline{Y})$ *over the reals.*

Proof. In all three cases one eliminates successively all uncountability quantifiers from ψ from the inside out by an application of Theorem 1 followed by the elimination of the predicate **Unc**(X).

Over countable linear orders the predicate **Unc** is vacuously always false, hence the first claim.

Gurevich [7] proved (assuming the continuous hypothesis) that the predicate "the set X is uncountable" is expressible in MLO over the reals. This proves the third claim.

It is well-known that "the set X is uncountable" is expressible in MLO over the class of all ordinals. Recall that a subset X of an ordinal α is uncountable iff the order type of $(X, <)$ is greater than or equal to ω_1 iff there is a subset $Y \subseteq X$ such that the cofinality of the order-type of $(Y, <)$ is strictly greater than ω. This is the case precisely if every subset $Z \subseteq Y$ such that the order-type of $(Z, <)$ is ω is bounded in Y. Formally, "the set X is uncountable" is equivalent over ordinals to

$$\exists Y \subseteq X \ \forall Z \subseteq Y \ \omega(Z) \rightarrow \exists y \in Y \forall z \in Z \ z < y$$

where $\omega(Z)$ expresses that the order type of $(Z,<)$ is ω, for instance by saying that Z is infinite and $[0,z)\cap Z$ is finite for every $z\in Z$. Finiteness can be expressed e.g. as shown in Proposition 7. $\dashv$

Observation of the proof reveals that $\exists^{\aleph_1}$ and $\exists^{2^{\aleph_0}}$ are equivalent over countable scattered linear orders, but of course not over all ordinals. In [9, 10] Kuske and Lohrey obtained similar results over $(\omega,<)$.

§6. Reduction to counting cuts. In this section we show that the existence of uncountably many sets satisfying an MLO formula can be reduced to the existence of uncountably many cuts (downward closed sets) satisfying some MLO formula effectively obtainable from the prior one. Cuts are of course just representations of points of the completion of the underlying linear order. Hence we show that $\exists^{\aleph_1}$ over a linear order L reduces to **Unc** over its completion $\overline{L}$. In the case of almost complete linear orders it was of course sufficient to consider **Unc** over L as we did in Theorem 1. We say that a point of $\overline{L}$ is an *extremal point* of a given finite U-U cover if it is the supremum or the infimum of any of its constituent intervals. When convenient we may blur the distinction between a cut C of a linear order L and the corresponding point $\sup C$ of the completion $\overline{L}$.

DEFINITION 19. A cut C is an *essential cut for* X, equivalently, $\sup C\in\overline{L}$ is an *essential point for* X, if every interval I such that I intersects both C and its complement (i.e. $\inf I\le\sup C<\sup I$ in $\overline{L}$) is a D-interval for X.

LEMMA 20.

(i) *If* $\sup C$ *is an essential point for* X *then it is an extremal point of every finite U-U cover for* X.
(ii) *If there is a finite balanced U-U cover for some* X *then there is also one whose non-essential extremal points belong to* L.

PROOF.

(i) Assume indirectly that there is an interval I of some balanced U-U cover for X and points $v,w\in I$ such that $v\in C$ and $w\notin C$. Then (v,w) must be a U-interval for X because either $\{x\in I\mid x<w\}$ or $\{x\in I\mid v<x\}$ is a U-interval and (v,w) is contained in both.
(ii) Consider wlog. a finite balanced U-U cover consisting of disjoint intervals bounded by consecutive elements $\sigma_1<\sigma_2<\ldots<\sigma_t$ of the completion $\overline{L}$. If some σ_j does not fall in L and neither is it an essential cut-point for X then there are points $v<\sigma_j<w$ in L such that (v,w) is a U-interval for X. Wlog. $\sigma_{j-1}<v<w<\sigma_{j+1}$. Hence the points $\sigma_1<\ldots<\sigma_{j-1}<v<w<\sigma_{j+1}<\ldots<\sigma_t$ give rise to a new balanced U-U cover for X with fewer non-essential cut-points in $\overline{L}\setminus L$. Continuing this way in a finite number of iterations we arrive at a finite balanced U-U cover for X consisting of disjoint intervals the extremal points of which are either essential cut-points for X or fall inside L.

$\dashv$

Combining Proposition 17 and Lemma 20 (ii) we obtain the following generalisation of Theorem 1.

PROPOSITION 21. *To every* MLO *formula* $\varphi(X,\overline{Y})$ *one can effectively associate an* MLO($\exists^{\aleph_1}$) *formula* $\psi(\overline{Y})$ *equivalent to* $\exists^{\aleph_1}X\ \varphi(X,\overline{Y})$ *over all linear orders and such that in* ψ *the uncountability quantifier only occurs in the restricted form* $\exists^{\aleph_1}C\ (\mathrm{cut}(C)\wedge\vartheta)$.

PROOF. Proposition 17 tells us that $\neg\exists^{\aleph_1}X\ \varphi(X,\overline{Y})$ is equivalent to the existence of a countable subset $U\in\overline{L}$ containing all extremal points of some finite U-U cover for each X satisfying $\varphi(X,\overline{Y})$. According to Lemma 20 (ii) if there is one such countable set $U\in\overline{L}$ then there is also one containing only those points of $\overline{L}\setminus L$, which are essential for some set X satisfying $\varphi(X,\overline{Y})$. Thus we see that $\neg\exists^{\aleph_1}X\ \varphi(X,\overline{Y})$ holds over any chain if and only if

- there is a countable set $U\subseteq L$ such that every X satisfying $\varphi(X,\overline{Y})$ has a finite balanced U-U cover with non-essential cut-points in U,
- and there are altogether only countably many cuts essential for some set X satisfying $\varphi(X,\overline{Y})$.

Note that a set $U\subseteq L$ is countable iff $\neg\exists^{\aleph_1}C\big(\mathrm{cut}(C)\wedge\exists x\in U\ \forall y(y\in C\leftrightarrow y\leq x)\big)$. It is also straightforward to give an MLO formula $ECUT_{\varphi}(C,X,\overline{Y})$ expressing that C is an essential cut for X. The second condition can be thus formalised as $\neg\exists^{\aleph_1}C\big(\mathrm{cut}(C)\wedge\exists X\,\varphi(X,\overline{Y})\wedge ECUT_{\varphi}(C,X,\overline{Y})\big)$. In both of these formulas the quantifier $\exists^{\aleph_1}$ is only used in the restricted form $\exists^{\aleph_1}C\ \mathrm{cut}(C)\wedge\vartheta$, which proves the lemma. ⊣

§7. Rationals (counting Dedekind cuts). In this section we show how the uncountability quantifier can be eliminated from monadic second-order logic of order over the rationals. It will also be apparent that over the rationals $\exists^{\aleph_1}X\varphi$ and $\exists^{2^{\aleph_0}}X\varphi$ are equivalent for any MLO formula φ. Recall that a *proper cut* is a cut having no supremum in the underlying linear order. A cut is non-trivial if neither itself nor its complement is empty.

LEMMA 22. *Let* $\psi(C,Y_1,\dots,Y_M)$ *be an* MLO *formula and* $V_1,\dots,V_M\subseteq\mathbb{Q}$. *Then there are uncountably many — and in fact continuum many — Dedekind cuts* C *of* $\mathbb{Q}$ *satisfying* $(\mathbb{Q},<)\models\psi(C,\overline{V})$ *if and only if there is a subset* $D\subseteq\mathbb{Q}$ *such that* $(D,<)$ *is dense and for every non-trivial proper cut* C *of* $(D,<)$ *the Dedekind cut* $C'=\{q\in\mathbb{Q}\mid\exists p\in C:\ q<p\}$ *of rationals satisfies* $\psi(C',\overline{V})$.

PROOF. Only the necessity of the above condition requires consideration. To that end assume that there are uncountably many cuts satisfying ψ and say that two rationals q and q' are *close*, $q\asymp q'$, if $[\min(q,q'),\max(q,q')]$ contains only countably many cuts satisfying ψ; and *far* otherwise. This defines an equivalence relation each equivalence class of which constitutes an interval of the rationals. These intervals are naturally linearly ordered and form a dense ordering. Indeed, by assumption there are at least two classes and by definition no two classes can form adjacent intervals, for otherwise their union would have to be part of a single class. In other words between any two points far apart there must be a third, which is far form both of these.

Assign to every pair $[q]_{\asymp}<[q']_{\asymp}$ of $\asymp$-classes as its colour the n-theory of the interval $L_{[q,q')}=\bigcup\{[p]_{\asymp}\mid[q]_{\asymp}\leq[p]_{\asymp}<[q']_{\asymp}\}$, where n is the quantifier rank of ψ.

$$\nu([q]_{\asymp},[q']_{\asymp}) = \mathrm{Tp}^n(L_{[q,q')},<,Y_1\cap L_{[q,q')},\dots,Y_M\cap L_{[q,q')})$$

By composition, ν defines an additive binary colouring on $\mathbb{Q}/_{\asymp}$. At this point we invoke a Ramsey-like theorem for dense linear orders due to Shelah [13, Theorem 1.3] asserting that there is an open interval $\mathcal{I}$ of $\mathbb{Q}/_{\asymp}$ and a subset $\mathcal{O}\subset\mathcal{I}$, which is dense in $\mathcal{I}$ and is ν-homogeneous. In other words there exists an n-theory τ such that $\nu([q]_{\asymp},[q']_{\asymp}) = \tau$ for all $[q]_{\asymp} < [q']_{\asymp}$ in $\mathcal{O}$. Let D be an arbitrary complete set of representatives of the $\asymp$-classes in $\mathcal{I}$. In particular $(D,<)$ is dense, countable and without endpoints. Let $D_0 = \{q\in D \mid [q]_{\asymp}\in\mathcal{O}\}$ and let $I = \bigcup\mathcal{I}$.

Consider now a non-trivial proper cut C of D and let $C' = \{q\in\mathbb{Q}\mid \exists p\in C:\ q<p\}$ as in the statement of this lemma. Because D_0 is dense in D there exist $\mathbb{Z}$-chains $\ldots<p_{-2}<p_{-1}<p_0<p_1<\ldots$ in $D_0\cap C$ and $\ldots<q_{-2}<q_{-1}<q_0<q_1<\ldots$ in $D_0\setminus C$ such that $C=\{d\in D\mid \exists z\ p_z\le d<p_{z+1}\}$ and, similarly, $D\setminus C=\{d\in D\mid \exists z\ q_z\le d<q_{z+1}\}$. In particular, there is no $d\in D$ such that $p_z<d<q_z$ for all $z\in\mathbb{Z}$, which also means that there is in fact no $q\in\mathbb{Q}$ such that $p_z<q<q_z$ for all $z\in\mathbb{Z}$. Therefore we have $I\cap C'=\{q\in\mathbb{Q}\mid \exists z\ p_z\le q<p_{z+1}\}$ and $I\setminus C=\{q\in\mathbb{Q}\mid\exists z\ q_z\le q<q_{z+1}\}$. By composition and homogeneity of $\mathcal{O}$, the n-theories $\mathrm{Tp}^n(I\cap C',<,Y_1\cap I\cap C',\dots,Y_M\cap I\cap C')$ and $\mathrm{Tp}^n(I\setminus C',<,Y_1\cap I\setminus C',\dots,Y_M\cap I\setminus C')$ are obtained as the $\mathbb{Z}$-fold product of τ with itself and as such are independent of the choice of C. By the composition theorem again it follows that either every C' as above satisfies $\psi(C',Y_1,\dots,Y_M)$ or none does. The latter possibility can be immediately ruled out on the grounds that any two points of D are by definition far apart meaning that there must be uncountably many Dedekind cuts between them satisfying ψ of which at most countably many do not induce, equivalently, are not induced by a proper cut of D. ⊣

The condition in the above statement is clearly MLO expressible. Combined with Proposition 21 this yields full and effective elimination of $\exists^{\aleph_1}$ over $(\mathbb{Q},<)$.

Proposition 23 (Elimination of $\exists^{\aleph_1}$ over the rationals). *For every* MLO*-formula* $\varphi(X,\overline{Y})$ *one can compute an* MLO*-formula* $\psi(\overline{Y})$ *equivalent to both* $\exists^{\aleph_1}X\ \varphi(X,\overline{Y})$ *and* $\exists^{2^{\aleph_0}}X\ \varphi(X,\overline{Y})$ *over the standard ordering of the rationals.*

Proof. The proof is by induction on the structure of the formula. To eliminate an inner-most occurrence of the uncountability quantifier one applies first Proposition 21 followed by a number of applications of Lemma 22. ⊣

§8. Sums of linear orderings. In the following we will make use of a more informative statement on composition of types on sums of linear orderings as formulated by Shelah.

Theorem 24 (Composition on linear orders II. [13]). *For every* MLO*-formula* $\varphi(\overline{X})$ *in the signature of l-chains having m free variables and quantifier rank n, and given the enumeration* $\tau_1(\overline{X}),\dots,\tau_k(\overline{X})$ *of* $H_{n,l+m}$*, there exists an* MLO*-formula* $\theta(Q_1,\dots,Q_k)$ *computable from the above and such that for every linear ordering* $\mathfrak{I}=(I,<^I)$ *and family* $\{\mathcal{L}_i\mid i\in I\}$ *of l-chains and subsets* $V_1,\dots,V_m$

of $\sum_{i\in I}\mathcal{L}_i$,

$$\sum_{i\in I}\mathcal{L}_i \models \varphi(\overline{V}) \iff \mathfrak{I} \models \theta(Q_1,\dots,Q_k)$$

where the predicates $\overline{Q}$ *form a partition of* I *induced by* $\overline{V}$ *as follows: for each* $1\le r\le k$,

$$Q_r = Q_r^{I;\overline{V}} = \{i\in I \mid \mathrm{Tp}^n(\mathcal{L}_i,\overline{V}) = \tau_r\}.$$

Using this theorem we can formulate some general conditions allowing to reduce the problem of eliminating the uncountability quantifier $\exists^{\aleph_1}$ over sums of linear orderings to eliminating $\exists^{\aleph_1}$ over the index structure as well as eliminating it uniformly over the summands. This, of course, assuming that the sum is already given.

Corollary 25. *Let* $\varphi(X,\overline{Y})$ *be an* MLO *formula of quantifier rank* N, *and let* $\mathcal{L}=\sum_{i\in I}\mathcal{L}_i$ *be an ordered sum of chains, and* $\overline{V}$ *subsets of* $\mathcal{L}$. *Let the enumeration of the* N*-types be given by* $\tau_1(X,\overline{Y}),\dots,\tau_K(X,\overline{Y})$ *and let* $\theta(\overline{T})$ *be the formula delivered by Theorem 24. Then there are uncountably many* $U\subseteq\mathcal{L}$ *satisfying* $\mathcal{L}\models\varphi(U,\overline{V})$ *iff*

(a) there is one such U *having infinitely many disjoint* D*-intervals; or*

(b) there is one such U *and an index* $i\in I$ *so that* $\mathcal{L}_i\models\exists^{\aleph_1}Z\ \tau_r(Z,\overline{V}|_{\mathcal{L}_i})$ *where* τ_r *is the* N*-type of* $(U|_{\mathcal{L}_i},\overline{V}|_{\mathcal{L}_i})$ *on* $\mathcal{L}_i$*; or*

(c) The set of those partitions $\overline{P}$ *of* I *that are induced by* $\overline{V}$ *and some* U *satisfying* $\psi(U,\overline{V})$ *is uncountable. This can be expressed by an* MLO($\exists^{\aleph_1}$)*-formula over the index structure:*

$$(I,<)\models\exists^{\aleph_1}\overline{P} : Part(\overline{P})\wedge\bigwedge_{r=1}^{K} P_r\subseteq Q_r\wedge\theta(\overline{P})$$

where $Part(\overline{P})$ *states that* $\overline{P}$ *partition* I *and for each* $r=1\dots K$ *the set* $Q_r=\{i\in I\mid \mathcal{L}_i\models\exists X'\,\tau_r(X',\overline{V}|_{\mathcal{L}_i})\}$.

Proof. Each of the three conditions is sufficient to yield uncountably many sets U satisfying φ. For (a) this was proved in Lemma 10 by the weaker form of the composition theorem. Similarly, for (b) this also follows directly already from the weaker composition theorem. Finally, for condition (c) this follows from the fact that, for every one of the uncountably many tuples $\overline{P}$ accounted for, there is a distinct set U inducing the type-predicates $\overline{P}$ and fulfilling $\varphi(U,\overline{V})$.

Conversely, if condition (c) fails then there are only countably many colourings of I with type predicates $\overline{P}$ induced by some U satisfying $\varphi(U,\overline{V})$. By failure of (a) for each of these type predicates we have for all but finitely many indices i that $i\in P_r$ implies that τ_r uniquely defines $U\cap L_i$ from $\overline{V}|_{L_i}$. Finally, if condition (b) fails too, then on each of the finitely many remaining intervals L_i there are also only countably many choices for $U\cap L_i$.

To see that the formalisation of condition (c) provided above is sound note that by Theorem 24 every U satisfying $\varphi(U,\overline{V})$ induces (together with $\overline{V}$) a partition $\overline{P}$ satisfying the given formula. Conversely, each tuple $\overline{P}$ satisfying it fulfils all the following: It forms a partition, it is induced by some set U together with

$\overline{V}$ as ensured by $\bigwedge_{r=1}^{K} P_r \subseteq Q_r$, and every U inducing it must satisfy $\varphi(U, \overline{V})$ thanks to $\theta(\overline{P})$. ⊣

Furthermore, if $\exists^{\aleph_1}$ is equivalent to $\exists^{2^{\aleph_0}}$ both over the index structure and on each of the summands then, by the proof of the previous Lemma, these two quantifiers are also equivalent over the sum.

A crucial point as to the applicability of the above claim is that it assumes a given factorisation of a linear ordering as a sum. We introduce the notion of definable splitting to facilitate the use of the above technique on classes of linear orderings over which an appropriate factorisation is uniformly definable.

Definition 26 (Splitting). Let $\mathcal{L} = (L, <, \dots)$ be a chain and $\theta(x, y)$ a formula with x and y first-order variables.

(1) We call θ a *splitting* of $\mathcal{L}$ if $\{(a, b) \in L^2 \mid \mathcal{L} \models \theta(a, b)\}$ is an equivalence relation whose every class is an interval.
(2) For a splitting θ of $\mathcal{L}$ let $\sim_\theta^{\mathcal{L}}$ denote the equivalence relation defined by θ in $\mathcal{L}$, let $I_{\mathcal{L}/\theta}$ be the set of $\sim_\theta^{\mathcal{L}}$-classes and $Ind_{\mathcal{L}/\theta} := (I_{\mathcal{L}/\theta}, <)$ the natural ordering of $I_{\mathcal{L}/\theta}$ according to representatives. Call $Ind_{\mathcal{L}/\theta}$ the *indexing order* of $\mathcal{L}$ and $\mathcal{S}_{\mathcal{L}/\theta} = \{\mathcal{L}|_I \mid I \in I_{\mathcal{L}/\theta}\}$ the *summand structures* of $\mathcal{L}$ w.r.t. θ.
(3) Let $\mathcal{C}$ be a class of labelled chains. We call θ a *splitting* of $\mathcal{C}$ iff θ splits every $\mathcal{L} \in \mathcal{C}$. Let $Ind_{\mathcal{C}/\theta} := \{Ind_{\mathcal{L}/\theta} \mid \mathcal{L} \in \mathcal{C}\}$ and $\mathcal{S}_{\mathcal{C}/\theta} := \bigcup_{\mathcal{L} \in \mathcal{C}} \mathcal{S}_{\mathcal{L}/\theta}$. Call $\mathrm{Ind}_{\mathcal{C}/\theta}$ and $\mathcal{S}_{\mathcal{C}/\theta}$ the *class of indexing chains* of $\mathcal{C}$ and the *class of summand structures* of $\mathcal{C}$ w.r.t. θ, respectively.

Theorem 27. *Let $\mathcal{C}$ be a class of labelled chains and θ a* splitting *of $\mathcal{C}$. If*

(1) MLO $+ \exists^{\aleph_1}$ *collapses effectively to* MLO *over the* class of indexing chains *of $\mathcal{C}$ w.r.t. θ, and*
(2) MLO$+\exists^{\aleph_1}$ *collapses effectively to* MLO *over the* class of summand structures *of $\mathcal{C}$ w.r.t. θ,*

then MLO $+ \exists^{\aleph_1}$ *collapses effectively to* MLO *over $\mathcal{C}$.*

Proof. Consider a formula $\varphi(X, \overline{Y})$ of MLO. We give a formula $\alpha \vee \beta \vee \gamma$ expressing in MLO the disjunction of the three conditions of Corollary 25 equivalent to $\exists^{\aleph_1} X \varphi(X, \overline{Y})$ uniformly over each $\mathcal{L} \in \mathcal{C}$ with the factorisation as defined by θ. Let $\tau_1, \dots, \tau_K$ be en enumeration of $\mathrm{Tp}(N, 1+M)$ where N is the quantifier rank of φ and $M = |\overline{Y}|$.

Condition (a) can be expressed in MLO uniformly over all chains of a given signature. For instance by requiring the existence of an X satisfying $\varphi(X, \overline{Y})$ and an infinite set D such that every interval containing at least two points of D is a D-interval for X:

$$\alpha = \exists X\ \exists D\ \mathbf{Inf}(D) \wedge \forall \text{ interval } I\ (\exists d \neq d' \in D \cap I) \rightarrow DINT_\varphi(X, \overline{Y}, I).$$

The use of $\mathbf{Inf}(D)$ above is, of course, just a shorthand, it can be eliminated as in Proposition 7.

Condition (b) can be easily expressed in MLO relying on the elimination procedure for $\mathcal{S}_{\mathcal{C}/\theta}$. By the latter, one obtains for each N-type $\tau_r(X, \overline{Y})$ an MLO formula $\nu_r(\overline{Y})$ equivalent to $\exists^{\aleph_1} Z\ \tau_r(Z, \overline{Y})$ over $\mathcal{S}_{\mathcal{C}/\theta}$. Using these, condition (b) can be written as

$$\beta = \exists X\, \varphi(X, \overline{Y}) \wedge \exists I\, \exists x\; CLASS_\theta(x, I) \wedge \bigvee_r (\tau_r^I(X, \overline{Y}) \wedge \nu_r^I(\overline{Y}))$$

where $CLASS_\theta(x, I) = \forall y(y \in I \leftrightarrow \theta(x, y))$ defines I as the equivalence class of x with respect to θ; and where for a formula ψ, we denote by ψ^I the relativisation of ψ to I.

Finally, to express condition (c) of Corollary 25 one needs to

- choose a set I of representatives of all the equivalence classes defined by θ,
- relativise to I the MLO formula $\rho(\overline{Q})$ equivalent to $\exists^{\aleph_1} \overline{P} \ldots$ of condition (c) over $Ind_{\mathcal{C}/\theta}$ as delivered by the elimination procedure for this class,
- and substitute into ρ^I the sets $\overline{Q}$ defined as

$$\Omega(\overline{Q}) = \bigwedge_{r=1}^{K} \forall x(x \in Q_r \leftrightarrow \exists X', L : CLASS_\theta(x, L) \wedge X' \subseteq L \wedge \tau_r^L(X', \overline{Y})).$$

With the customary shorthand "$\exists!\, y$" meaning "there is a unique y such that" this formula takes the form

$$\gamma = \exists I\; (\forall x\; \exists!\, y \in I : \theta(x, y)) \wedge \exists Q_1, \ldots, Q_K\; \Omega(\overline{Q}) \wedge \rho^I(\overline{Q}).$$

$\dashv$

§9. All countable linear orders. At last we are in a position to conclude that the quantifiers $\exists^{\aleph_1}$ and $\exists^{2^{\aleph_0}}$ are equivalent and can be effectively eliminated from MLO($\exists^{\aleph_1}, \exists^{2^{\aleph_0}}$) uniformly over all countable chains. Indeed, it is well known that every countable linear order arises as a dense sum of scattered linear orders, i.e. in the form $\sum_{q \in D} L_q$ where each L_q is a countable scattered linear order and D is either a singleton or is isomorphic to the standard ordering of the rationals with or without additional left and/or right extremal elements.

Let $\theta(x, y)$ be the MLO-formula expressing that for no subset A of $L_{[x,y]}$ is $(A, <)$ a dense ordering, i.e. that $L_{[x,y]}$ is a scattered linear order. Over any countable chain θ defines an equivalence relation partitioning it into intervals coinciding with the summands L_q as above. Thus, θ is a splitting of the class of all countable linear orders.

Taking advantage of Theorem 27 and using the previously proven collapse results over the class of countable scattered linear orders (Corollary 18) and over the rationals (Proposition 23), — which trivially extends to the rationals with either one or both endpoints added — we obtain uniform effective elimination of $\exists^{\aleph_1}$ over the class of all countable chains. This completes the proof of Theorem 2(2) and Theorem 3 follows similarly.

§10. Further results. Observe that combining Theorems 2 and 27 further elimination results can be obtained for a hierarchy of classes of linear orders. Starting from the classes of ordinals (and their reverses) and the class of countable linear orders, effective elimination of $\exists^{\aleph_1}$ can be derived for e.g. ordinal sums of countable linear orders, or for countable sums of ordinals and reverse ordinals, and so on for any finite number of iterations of summation. However, this transformation of formulas is not uniform over the union of these classes.

The same technique employed here can be adapted to obtain similar results over trees. The more complex structure of trees does pose, however, some interesting additional challenges. In [1] we have reported on a preliminary treatment of the tree case.

A simple (labelled) tree is a structure endowed with a partial order $<$ and any number of unary predicates such that the graph of $<$ consists of all transitive edges of a directed tree in the graph theoretic sense with its root as minimal element and such that the distance of every node from the root is finite and every node of the tree has finite out-degree. In [1] we have shown the following theorem generalising an earlier result of Niwinski [12].

THEOREM 28. *Over the class of simple trees every* $\mathrm{MLO}(\exists^{\aleph_0}, \exists^{\aleph_1}, \exists^{2^{\aleph_0}})$ *formula is effectively equivalent to an* MLO *formula. Moreover,* $\exists^{\aleph_1} X\, \varphi(X, \overline{Y})$ *and* $\exists^{2^{\aleph_0}} X\, \varphi(X, \overline{Y})$ *are equivalent for every* MLO *formula* φ *with parameters* $\overline{Y}$ *over every simple tree.*

REFERENCES

[1] V. Bárány, Ł. Kaiser, and A. Rabinovich, *Cardinality Quantifiers in MLO over Trees*, ***Computer Science Logic, CSL'09, Proceedings***, Lecture Notes in Computer Science, vol. 5771, Springer, 2009, pp. 117–131.

[2] A. Baudisch, D. Seese, H.-P. Tuschik, and M. Weese, ***Decidability and generalized quantifiers***, Akademie-Verlag Berlin, 1980.

[3] Julius R. Büchi, *Weak second-order arithmetic and finite automata*, ***Zeitschrift für mathematische Logik und Grundlagen der Mathematik***, vol. 6 (1960), pp. 66–92.

[4] ———, *On decision method in restricted second order arithmetic*, ***Proceedings of the international congress on logic, methodology and philosophy of science***, 1962, pp. 1–11.

[5] Julius R. Büchi and Dirk Siefkes, ***The monadic second order theory of all countable ordinals***, Lecture Notes in Mathematics, vol. 328, Springer, 1973.

[6] Calvin C. Elgot, *Decision problems of finite automata design and related arithmetics*, ***Transactions of the American Mathematical Society***, vol. 98 (1961), pp. 21–51.

[7] Y. Gurevich, *Monadic theory of order and topology, I*, ***Israel Journal of Mathematics***, vol. 27 (1977), pp. 299–319.

[8] M. Kaufmann, *The Quantifier "There Exists Uncountably Many" and Some of Its Relatives*, ***Model Theoretic Logics*** (J. Barwise and S. Feferman, editors), Springer-Verlag, 1985, pp. 123–176.

[9] Dietrich Kuske and Markus Lohrey, *First-order and counting theories of ω-automatic structures*, ***Proceedings of fossacs'06***, 2006, pp. 322–336.

[10] ———, *First-order and counting theories of omega-automatic structures*, ***The Journal of Symbolic Logic***, vol. 73 (2008), pp. 129–150.

[11] A. Mostowski, *On a generalization of quantifiers*, ***Fundamenta Mathematicae***, vol. 44 (1957), pp. 12–36.

[12] Damian Niwiński, *On the cardinality of sets of infinite trees recognizable by finite automata*, ***Mathematical foundations of computer science, mfcs'91, proceedings***, LNCS, vol. 520, Springer, 1991, pp. 367–376.

[13] Saharon Shelah, *The monadic theory of order*, ***Annals of Mathematics***, vol. 102 (1975), no. 3, pp. 379–419.

[14] Boris A. Trakhtenbrot, *Finite automata and the logic of monadic predicates*, ***Dokl. Akad. Nauk SSSR***, vol. 140 (1961), pp. 326–329.

[15] ———, *Finite automata and the logic of one-place predicates*, ***Sibirian Mathematical Journal***, vol. 13 (1962), pp. 103–131, (in Russian).

OXFORD UNIVERSITY COMPUTING LABORATORY
WOLFSON BUILDING, PARKS ROAD
OXFORD OX1 3QD, UNITED KINGDOM
E-mail: vbarany@comlab.ox.ac.uk

MATHEMATICAL FOUNDATIONS OF COMPUTER SCIENCE
RWTH AACHEN D-52056 AACHEN, GERMANY
E-mail: kaiser@logic.rwth-aachen.de

TEL AVIV UNIVERSITY SCHOOL OF COMPUTER SCIENCE
RAMAT AVIV, TEL AVIV 69978, ISRAEL
E-mail: rabinoa@post.tau.ac.il