APPROXIMATING TREES AS COLOURED LINEAR ORDERS AND COMPLETE AXIOMATISATIONS OF SOME CLASSES OF TREES

RUAAN KELLERMAN AND VALENTIN GORANKO

Abstract. We study the first-order theories of some natural and important classes of coloured trees, including the four classes of trees whose paths have the order type respectively of the natural numbers, the integers, the rationals, and the reals. We develop a technique for approximating a tree as a suitably coloured linear order. We then present the first-order theories of certain classes of coloured linear orders and use them, along with the approximating technique, to establish complete axiomatisations of the four classes of trees mentioned above.

§1. Introduction.

1.1. Background and motivation. The logic-based study of linear orders was comprehensively presented in the early 1980s in the still very relevant classic book [16], and there have been several important further developments since then, mentioned below. In particular, the first-order (FO) theories of various naturally arising classes of linear orders are now well-known, for example, the FO theories of each ordinal α for $\alpha < \omega^\omega$ (see [16]), as well as the FO theory of the rational numbers (implicit in [16]). Furthermore, in [1, 2] Doets studies several natural classes of *coloured* linear orders (i.e., linear orders enriched with unary predicates) and obtains complete axiomatisations for the first-order theories of: the class of coloured scattered linear orders; the class of coloured expansions of the natural numbers, from which the case of the integers follows easily; the class of coloured finite linear orders; the class of coloured complete linear orders; the class of coloured well-orders; and the class of coloured expansions of the order of the real numbers. Also closely related are the papers [13, 14], which investigate coloured finite linear orders and coloured ordinals respectively.

The study and axiomatisation of the first-order theories of naturally arising classes of trees is substantially more complicated, however, even when the first-order theory of the corresponding class of linear orders is known. A more systematic attempt was made in [8] to explore the general problem of transferring the first-order theory of a class of linear orders to the class of trees whose paths are all contained in that class of linear orders. That work left many open questions and indicated some inherent difficulties. They are mainly due to the following facts:

(i) Since paths (maximal linearly ordered chains) are special sets of nodes in a given tree, the first-order language for trees cannot, in general, impose first-order properties on all paths of the tree, but only on the *first-order definable* ones. A path P in a tree $\mathfrak{T} := (T; <)$ is called *singular* when it contains a node u such that the set $\{x \in T : u \leqslant x\}$ is a linear order within $\mathfrak{T}$. All singular paths are parametrically definable. However, trees may also contain *emerging* paths, which are paths that are not singular. In such non-definable emerging paths, behaviour in the terminal part of the path cannot generally be controlled within the expressive means of first-order logic.

(ii) The branching structure of a tree cannot be captured by the properties of its paths.

Consequently, there are very few known complete axiomatisations of first-order theories of classes of trees, in essence comprising the following classes: the class of finitely branching trees (implicit in [18]), the class of (ordinary or coloured) well-founded trees (see [2]), and the class of finite trees (see [15]). Also, [7] contains some general results on axiomatising subclasses of the class of finitely branching trees relative to the respective classes of trees with no restriction on their branching. Further, the first-order theories of the class of trees, all of whose paths contain greatest elements (leaves), and the class of trees whose paths are all isomorphic to some given ordinal α with $\alpha < \omega^{\omega}$, are investigated in [11], but without deriving complete axiomatisations of these first-order theories. Lastly, even though not directly related to the present paper, we should mention the very important works by Gurevich and Shelah [9, 10] on decidability of first-order theories of coloured trees with additional quantification over branches.

1.2. Goal and main contributions of the present work. The goal of the present paper is to study and axiomatise in first-order logic the classes of trees naturally arising from some important linear orders. More precisely, we obtain axiomatisations of the first-order theories of these classes of trees, rather than axiomatisations of those classes themselves. This amounts to the following: given a class of trees $\mathcal{K}$, we seek a recursive (i.e., decidable) set of first-order sentences Σ such that $\Sigma \subseteq \mathrm{Th}(\mathcal{K})$ and $\Sigma \models \mathrm{Th}(\mathcal{K})$. In turn, $\Sigma \models \mathrm{Th}(\mathcal{K})$ if and only if for each natural number n and each model $\mathfrak{T}$ of Σ, there exists a tree $\mathfrak{S}$ in $\mathcal{K}$ such that $\mathfrak{T}$ and $\mathfrak{S}$ satisfy the same sentences of quantifier rank at most n.

Now for any order type α, a tree whose paths are all isomorphic to α is called an α-tree. This paper addresses and solves the problems of axiomatising the first-order theories of the classes of (coloured) ω-trees, ζ-trees, η-trees, and λ-trees, where ω, ζ, η, and λ are the order types of the sets of natural numbers, integers, rational numbers, and real numbers, respectively. While the case of η-trees is easy and the case of ω-trees was essentially known from [2], the cases of ζ-trees and λ-trees turned out to be quite non-trivial. The complete axiomatisations of their first-order theories are obtained here by using a new construction for approximating a given tree by a suitably coloured linear order and then using the axiomatisations of the first-order theories of the classes of coloured expansions of ζ and λ respectively.

1.3. Structure of the paper. First, after briefly covering some basic definitions and notation in *Section* 2: *Preliminaries*, we present in *Section* 3: *First-order theories of*

coloured linear orders the axiomatisations of several important classes of coloured linear orders. All these axiomatisations are either already known in the literature or are easily derived from such known axiomatisations. However we include them here as they are necessary for the subsequent main results on trees, for the sake of completeness of the presentation and for uniformity of the underlying axioms that are used. As is to be expected, the analysis of these various classes of coloured linear orders tends to be substantially more complex than for the corresponding classes of monochromatic linear orders. We note in passing that, apart from their use in axiomatising the first-order theories of classes of trees, coloured linear orders are also useful, inter alia, as models of (possibly infinite) words over an alphabet. Then, in *Section* 4: *From trees to coloured linear orders and back* we introduce the construction used for approximating a tree as a coloured linear order, which we use in *Section* 5: *First-order theories of coloured trees* to axiomatise the first-order theories of the classes of coloured ζ-trees and coloured λ-trees. In addition, easier axiomatisations of the first-order theories of the classes of coloured ω-trees and coloured η-trees are also given. In all of these cases, the axiomatisations of the first-order theories of the classes of trees make essential use of the axiomatisations of the first-order theories of their corresponding classes of coloured linear orders. We end with a summary and some suggestions for future work in *Section* 6: *Concluding remarks*.

§2. Preliminaries.

2.1. Notation and logical preliminaries. We will be working with first-order (FO) languages with signatures that contain a binary relation $<$ (possibly, in addition to constants and unary relations). Expressions such as $x > y$, $x \leqslant y$, and $x \not< y$, when used in first-order formulas, are to be interpreted in the usual way, as abbreviations for the formulas $y < x$, $x < y \lor x = y$, and $\neg (x < y)$ respectively; similarly for other variations of the relation $<$.

Equality will always be assumed to be included in the language. The FO signature that consists only of equality plus the relation symbol $<$ will be denoted by L_0, and the signature that extends L_0 with unary relation symbols $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_k$ (regarded as colours) will be denoted by L_k. For $k \geqslant 0$, L_k' will denote the signature L_k expanded with a single constant symbol, i.e., the signature of structures of the form $(A; <, a)$ when $k = 0$ or $(A; <, a, \mathbf{c}_1, \dots, \mathbf{c}_k)$ when $k > 0$.

Consider an arbitrary fixed first-order language L. The quantifier rank of a formula φ is defined as expected and will be denoted by $\mathrm{qr}\,(\varphi)$. We denote by FO_n the set of first-order formulas (of the given signature L) of quantifier rank at most n and for any L-structures $\mathfrak{A}$ and $\mathfrak{B}$, by $\mathfrak{A} \equiv_n \mathfrak{B}$ we mean equivalence of $\mathfrak{A}$ and $\mathfrak{B}$ with respect to all sentences in FO_n, whereas $\mathfrak{A} \equiv \mathfrak{B}$ means, as usual, their elementary equivalence.

The domain of a structure $\mathfrak{A}$ will be denoted by $|\mathfrak{A}|$. The cardinality of a set X will also be denoted by $|X|$. From the context, it will always be clear which one of the two meanings applies.

Given an L-structure $\mathfrak{A}$ with domain A, the parametrically definable subsets of $\mathfrak{A}$ are those sets of the form $\{x \in A : (\mathfrak{A}; \bar{a}) \models \varphi\,(x, \bar{a})\}$, where $\varphi\,(x, \bar{y})$ is a first-order

formula with $\bar{y}$ possibly empty, and $\bar{a}$ is a tuple of elements from A of the same arity as $\bar{y}$.

Given formulas φ and $\theta = \theta(x, \bar{y})$ that have no variables in common, the relativisation of φ to θ, obtained from φ by replacing each subformula of the form $\exists u\,(\psi\,(u, \bar{w}))$ with $\exists u\,(\theta\,(u, \bar{y}) \wedge \psi\,(u, \bar{w}))$, and each subformula of the form $\forall u\,(\psi\,(u, \bar{w}))$ with the formula $\forall u\,(\theta\,(u, \bar{y}) \rightarrow \psi\,(u, \bar{w}))$, will be denoted by φ^{θ}. For a detailed account of relativisations, see, e.g., [16, p. 259]. If $\theta = \theta(x, y)$ is the formula $y \leqslant x$, with y here playing the role of a parameter, then φ^{θ} will be written simply as $\varphi^{\geqslant y}$. The formulas $\varphi^{>y}$, $\varphi^{\leqslant y}$, and $\varphi^{<y}$ are interpreted similarly. If $\theta(x, y, z)$ is the formula $y \leqslant x \wedge x < z$, with y and z now treated as parameters, then φ^{θ} will be written as $\varphi^{[y,z)}$, and similarly for the other bounded intervals. Given a structure $\mathfrak{A}$ with domain A, a (possibly empty) tuple of elements $\bar{a} \in A^k$, and a formula $\theta(x, \bar{y})$ with $\bar{y}$ being a (possibly empty) k-tuple of variables, $(\mathfrak{A}; \bar{a})^{\theta}$ will denote the set $\{v \in A : (\mathfrak{A}; \bar{a}) \models \theta\,(v, \bar{a})\}$.

For any natural number m, the linear order $(\{0, 1, \ldots, m-1\}\,;<)$, ordered in the usual way, will be denoted by $\mathbf{m}$.

We briefly recall the definitions and main result about *characteristic first-order formulas* (using the notation of [1, Section 1.6]) that will be needed.

Given a structure $\mathfrak{A}$ with domain A, natural numbers k and n, a k-tuple of elements $\bar{a} = (a_0, a_1, \ldots, a_{k-1})$ from A^k, and a k-tuple of variables $\bar{x} = (x_0, x_1, \ldots, x_{k-1})$ (where $\bar{a}$ and $\bar{x}$ are empty when $k = 0$), the *n-characteristic formula* of the structure $\mathfrak{A}$ over the tuple $\bar{a}$ is denoted by $[\![(\mathfrak{A}; \bar{a})]\!]^n(\bar{x})$ and is defined as follows:

- $[\![(\mathfrak{A}; \bar{a})]\!]^0(\bar{x}) := \bigwedge \{\varphi(\bar{x}) : \varphi \quad$ is an atomic or negated atomic formula with $\mathfrak{A} \models \varphi(\bar{a})\}$;
- $[\![(\mathfrak{A}; \bar{a})]\!]^{m+1}(\bar{x}) := \bigwedge_{a_k \in A}(\exists x_k([\![(\mathfrak{A}; \bar{a}a_k)]\!]^m(\bar{x}x_k))) \wedge \forall x_k(\bigvee_{a_k \in A}[\![(\mathfrak{A}; \bar{a}a_k)]\!]^m(\bar{x}x_k))$.

For languages with finite relational signatures it is well known (see, e.g., [5, Theorem 3.4 of Chapter 12]) that, for all natural numbers n and k there are, up to logical equivalence, only finitely many n-characteristic formulas over the class of all structures for that signature and all k-tuples in those structures. If $\bar{a}$ is the empty tuple then $[\![(\mathfrak{A}; \bar{a})]\!]^n(\bar{x})$ is written as $[\![\mathfrak{A}]\!]^n$ and is called the *n-characteristic sentence* of $\mathfrak{A}$. Here are some important facts about characteristic formulas:

1. The formula $[\![(\mathfrak{A}; \bar{a})]\!]^n(\bar{x})$ has quantifier rank n.
2. $\mathfrak{A} \models [\![(\mathfrak{A}; \bar{a})]\!]^n(\bar{a})$.
3. If $\mathfrak{B}$ is a structure in the signature of $\mathfrak{A}$ and $\bar{b}$ is a k-tuple of elements from the domain of $\mathfrak{B}$ then the following three statements are equivalent for any natural number n.
   (a) $(\mathfrak{A}; \bar{a}) \equiv_n (\mathfrak{B}; \bar{b})$.
   (b) $\mathfrak{B} \models [\![(\mathfrak{A}; \bar{a})]\!]^n(\bar{b})$.
   (c) $[\![(\mathfrak{A}; \bar{a})]\!]^n(\bar{x}) \equiv [\![(\mathfrak{B}; \bar{b})]\!]^n(\bar{x})$, where $\equiv$ denotes logical equivalence of formulas.

Given a finite relational FO signature L, the set

$$\sigma_n(L) := \{[\![\mathfrak{A}]\!]^n : \mathfrak{A} \text{ is any } L\text{-structure}\}/\equiv$$

that consists of all equivalence classes of logically equivalent n-characteristic L-sentences will be called the n-spectrum of L, and its cardinality will be denoted by $f(L,n) := |\sigma_n(L)|$. For each finite relational signature L, natural number n, and integer i with $1 \leqslant i \leqslant f(L,n)$, we fix a sentence $\tau_{L,n,i}$ such that

$$\tau_{L,n,1}, \ \tau_{L,n,2}, \ \dots, \ \tau_{L,n,f(L,n)}$$

is an enumeration of all n-characteristic L-sentences, up to logical equivalence.

Given a structure $\mathfrak{A}$, a subset A' of A that has the property that, for each $a \in A$, there exists $b \in A'$ such that $(\mathfrak{A}; a) \equiv_n (\mathfrak{A}; b)$, will be called an n-support of $\mathfrak{A}$. It follows from the properties of characteristic formulas that if $\mathfrak{A}$ has a finite relational signature then it has a finite n-support, for each n.

The reader is referred to [3] for background on Ehrenfeucht–Fraïssé games; the players Di and Sy of [3] will be called respectively Spoiler and Duplicator here and the n-round game on structures $\mathfrak{A}$ and $\mathfrak{B}$ will be denoted by $\mathrm{EF}_n(\mathfrak{A}, \mathfrak{B})$.

2.2. Linear orders. A k-coloured linear order, or simply a coloured linear order, is a structure of the type $\mathfrak{A} = (A; <, \mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_k)$ where k is a positive integer, $(A; <)$ is a linear order, and each $\mathbf{c}_i$ is a unary predicate in A, called a colour. (Thus, an element in A may have more than one colour, or none whatsoever.) The class of all k-coloured linear orders will be denoted by $\mathcal{L}_k$. Putting $\bar{\mathbf{c}} = (\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_k)$, $\mathfrak{A}$ can also be written simply as $\mathfrak{A} = (A; <, \bar{\mathbf{c}})$. The k-tuple of colours $\bar{\mathbf{c}}$ will sometimes be written as $\bar{\mathbf{c}}_k$, to emphasise that there are k colours. Given a subset B of A, we define $\bar{\mathbf{c}}\!\restriction_B := (\mathbf{c}_1\!\restriction_B, \mathbf{c}_2\!\restriction_B, \dots, \mathbf{c}_k\!\restriction_B)$ and $\mathfrak{A}^B := (B; <\!\restriction_B, \bar{\mathbf{c}}\!\restriction_B)$. To keep the notation simple, $\mathfrak{A}^B$ may sometimes also be written simply as $\mathfrak{A}^B = (B; <, \bar{\mathbf{c}})$, tacitly assuming that $<$ and $\bar{\mathbf{c}}$ are restricted in the obvious way. If $a \in A$ then $\mathfrak{A}^{\geqslant a}$ will mean $\mathfrak{A}^B$ where $B = \{x \in A : a \leqslant x\}$, and similarly for $\mathfrak{A}^{>a}$, $\mathfrak{A}^{\leqslant a}$, and $\mathfrak{A}^{<a}$. A linear order $(A; <)$ that is not enriched with any colours will be called monochromatic. For technical convenience, monochromatic linear orders may also be thought of as 0-coloured and for which the tuple of colours $\bar{\mathbf{c}}$ is empty. Given a coloured linear order $\mathfrak{A} = (A; <, \bar{\mathbf{c}})$, the monochromatic reduct $(A; <)$ will be denoted by $\mathfrak{A}^-$. Unless otherwise specified, the domains of structures $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}_i$, etc. will be denoted by A, B, and C_i, respectively. To avoid ambiguity, when several structures are under consideration, their order relations will be denoted by $<_{\mathfrak{A}}$, $<_{\mathfrak{B}}$, and $<_{\mathfrak{C}_i}$, respectively. In the interest of readability, $<_{\mathfrak{A}}$, $<_{\mathfrak{B}}$, and $<_{\mathfrak{C}_i}$ may sometimes all be denoted simply by $<$, with the understanding that the relation $<$ is then to be understood as $<_{\mathfrak{A}}$, $<_{\mathfrak{B}}$, or $<_{\mathfrak{C}_i}$, depending on the structure being worked in.

Given two coloured linear orders $\mathfrak{A} = (A; <_{\mathfrak{A}}, \bar{\mathbf{c}})$ and $\mathfrak{B} = (B; <_{\mathfrak{B}}, \bar{\mathbf{d}})$, the sum $\mathfrak{A} + \mathfrak{B}$ of $\mathfrak{A}$ and $\mathfrak{B}$ is defined to be the coloured linear order $\mathfrak{A} + \mathfrak{B} := (A \sqcup B; <, \bar{\mathbf{e}})$, where $A \sqcup B := (A \times \{0\}) \cup (B \times \{1\})$, the relation $<$ is the union of the three sets $\{((x,0),(y,0)) : x <_{\mathfrak{A}} y\}$, $\{((x,1),(y,1)) : x <_{\mathfrak{B}} y\}$, and $\{((x,0),(y,1)) : x \in A \text{ and } y \in B\}$, and $\bar{\mathbf{e}} = \bar{\mathbf{c}}\bar{\mathbf{d}}'$ where $\bar{\mathbf{d}}'$ is the tuple that is obtained from $\bar{\mathbf{d}}$ by removing all colours that are already in $\bar{\mathbf{c}}$, and for any colour $\mathbf{e}_i$ in $\bar{\mathbf{e}}$ and any $u \in A \sqcup B$, $\mathfrak{A} + \mathfrak{B} \models \mathbf{e}_i(u)$ if and only if either u has the form $u = (a,0)$ and $\mathfrak{A} \models \mathbf{e}_i(a)$, or u has the form $u = (b,1)$ and $\mathfrak{B} \models \mathbf{e}_i(b)$. To keep the notation simple, the elements $(a,0)$ and $(b,1)$ may sometimes be identified simply with the elements $a \in A$ and $b \in B$ respectively. The following Feferman–Vaught style result (see [6]) is easily proved by a straightforward application of Ehrenfeucht–Fraïssé games.

LEMMA 2.1.   *Let $\mathfrak{A}_1$, $\mathfrak{A}_2$, $\mathfrak{B}_1$, and $\mathfrak{B}_2$ be coloured linear orders, where $\mathfrak{A}_1$ and $\mathfrak{B}_1$ have the same set of colours $\bar{\mathbf{c}}$, and $\mathfrak{A}_2$ and $\mathfrak{B}_2$ have the same set of colours $\bar{\mathbf{d}}$. If $\mathfrak{A}_1 \equiv_n \mathfrak{B}_1$ and $\mathfrak{A}_2 \equiv_n \mathfrak{B}_2$ then $\mathfrak{A}_1 + \mathfrak{A}_2 \equiv_n \mathfrak{B}_1 + \mathfrak{B}_2$.*

2.3. Trees. A *forest* is a structure of type $\mathfrak{T} = (T; <)$ in which the relation $<$ is irreflexive, transitive, and left-linear (for each $a \in T$, the set $\{x : x < a\}$ is linear). A *tree* is a forest which is *left-connected*: for any $a, b \in T$ there exists $c \in T$ such that $c \leqslant a$ and $c \leqslant b$. A subset of a forest T that is maximal with respect to being connected is called a (*connected*) *component* of $\mathfrak{T}$. The elements of T are called *nodes*. The least node of $\mathfrak{T}$, if it exists, is called the *root* of $\mathfrak{T}$, and a node is called a *leaf* when it is maximal with respect to $<$. For a maximal linearly ordered subset $A \subseteq T$, the linear order $\mathsf{A} = (A; <\restriction_A)$ is called a *path* in $\mathfrak{T}$. For a *left-closed* subset B of a path A, i.e., such that for each $b \in B$, if $c < b$ then $c \in B$, too, the linear order $\mathsf{B} = (B; <\restriction_B)$ is called a *stem* (of A) in $\mathfrak{T}$. For ease of notation, the path A and the stem B will often be identified with their domains A and B. The set of all paths in $\mathfrak{T}$ will be denoted by $\mathcal{H}(\mathfrak{T})$. A set $I \subseteq T$ is called an *interval* when, for $a, b, c \in T$ with $a < b < c$, if $a, c \in I$ then $b \in I$, too.

As with linear orders, unary predicates regarded as colours may be added to the tree to obtain a *coloured tree* $(\mathfrak{T}; \bar{\mathbf{c}})$.

Given a (possibly monochromatic) tree $\mathfrak{T} = (T; <, \bar{\mathbf{c}})$ and a subset $C \subseteq T$, the structure $\mathfrak{T}^C := (C; <\restriction_C, \bar{\mathbf{c}}\restriction_C)$ is defined as with linear orders. In particular, for a node $a \in T$, the trees $\mathfrak{T}^{\geqslant a}$, $\mathfrak{T}^{>a}$, $\mathfrak{T}^{\leqslant a}$, and $\mathfrak{T}^{<a}$ are defined as the trees $\mathfrak{T}^C$ with C taken to be the set $T^{\geqslant a} := \{x \in T : a \leqslant x\}$, $T^{>a} := \{x \in T : a < x\}$, $T^{\leqslant a} := \{x \in T : x \leqslant a\}$, and $T^{<a} := \{x \in T : x < a\}$ respectively. The tree $\mathfrak{T}^{\geqslant a}$ will be called the *principal subtree generated by* a and the stem $T^{<a}$ will be called the *principal stem generated by* a.

Now, for $\mathfrak{T} = (T; <_{\mathfrak{T}}, \bar{\mathbf{c}})$ a tree and $\mathfrak{F} = (F; <_{\mathfrak{F}}, \bar{\mathbf{d}})$ a forest (where either of $\mathfrak{T}$ and $\mathfrak{F}$ may be monochromatic), and a stem B in $\mathfrak{T}$, we denote by $\mathfrak{T} +_B \mathfrak{F} := (T \sqcup F; <_{\mathfrak{T}+_B\mathfrak{F}}, \bar{\mathbf{e}})$ the tree with domain

$$|\mathfrak{T} +_B \mathfrak{F}| := T \sqcup F = (T \times \{0\}) \cup (F \times \{1\}),$$

with order relation

$$<_{\mathfrak{T}+_B\mathfrak{F}} := \Big\{ \big((x,0),(y,0)\big) : x <_{\mathfrak{T}} y \Big\} \cup \Big\{ \big((x,1),(y,1)\big) : x <_{\mathfrak{F}} y \Big\}$$
$$\cup \Big\{ \big((x,0),(y,1)\big) : x \in B \text{ and } y \in F \Big\},$$

and with colours $\bar{\mathbf{e}} = \bar{\mathbf{c}}\bar{\mathbf{d}}'$, where $\bar{\mathbf{d}}'$ is the tuple that is obtained from $\bar{\mathbf{d}}$ by removing all colours that are already in $\bar{\mathbf{c}}$, and where, for any colour $\mathbf{e}_i$ in $\bar{\mathbf{e}}$ and any $u \in T \sqcup F$, $\mathfrak{T} +_B \mathfrak{F} \models \mathbf{e}_i(u)$ if and only if either u has the form $u = (a, 0)$ and $\mathfrak{T} \models \mathbf{e}_i(a)$, or u has the form $u = (b, 1)$ and $\mathfrak{F} \models \mathbf{e}_i(b)$. In other words, $\mathfrak{T} +_B \mathfrak{F}$ is the tree obtained from $\mathfrak{T}$ by adding the forest $\mathfrak{F}$ to the end of the stem B; a depiction of this tree is given in Figure 1. As with sums of linear orders, for nodes $a \in T$ and $b \in F$, the nodes $(a, 0)$ and $(b, 1)$ in the tree $\mathfrak{T} +_B \mathfrak{F}$ will often be written simply as a and b rather than as $(a, 0)$ and $(b, 1)$, to keep the notation simpler, provided that there is no danger of confusion.

The following composition result generalises Lemma 2.1 and can also be proved by a straightforward Ehrenfeucht–Fraïssé game.
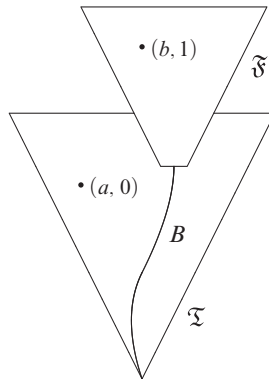
FIGURE 1. A depiction of the tree $\mathfrak{T} +_B \mathfrak{F}$.

LEMMA 2.2. *Let $\mathfrak{T}$ and $\mathfrak{S}$ be (possibly monochromatic) trees with the same set of colours $\bar{\mathbf{c}}$, A and B be stems in $\mathfrak{T}$ and $\mathfrak{S}$ respectively, and $\mathfrak{F}$ and $\mathfrak{G}$ be (possibly monochromatic) forests with the same set of colours $\bar{\mathbf{d}}$. Let $\bar{a} := (a_1, a_2, \dots, a_k) \in F^k$ and $\bar{b} := (b_1, b_2, \dots, b_k) \in G^k$ (the tuples $\bar{a}$ and $\bar{b}$ may be empty) and for each i, let $a_i' := (a_i, 1) \in |\mathfrak{T} +_A \mathfrak{F}|$ and $b_i' := (b_i, 1) \in |\mathfrak{S} +_B \mathfrak{G}|$ and let $\bar{a}' := (a_1', a_2', \dots, a_k')$ and $\bar{b}' := (b_1', b_2', \dots, b_k')$ (the tuples $\bar{a}'$ and $\bar{b}'$ will be empty when $\bar{a}$ and $\bar{b}$ are). If $(\mathfrak{T}; A) \equiv_n (\mathfrak{S}; B)$ and $(\mathfrak{F}; \bar{a}) \equiv_n (\mathfrak{G}; \bar{b})$ then $(\mathfrak{T} +_A \mathfrak{F}; \bar{a}') \equiv_n (\mathfrak{S} +_B \mathfrak{G}; \bar{b}')$.*

Given a class of (monochromatic) linear orders $\mathcal{C}$, a tree $\mathfrak{S}$ for which $\mathcal{H}(\mathfrak{S}) \subseteq \mathcal{C}$ will be called a *tree over $\mathcal{C}$*, or simply a *$\mathcal{C}$-tree*, and the class of all k-coloured $\mathcal{C}$-trees will be denoted by $\mathcal{T}_k(\mathcal{C})$ (with $k = 0$ in the case of monochromatic trees). If $\mathcal{C} = \{\alpha\}$ consists of just one linear order then $\mathcal{T}_k(\mathcal{C})$ will also be denoted by $\mathcal{T}_k(\alpha)$ and the trees in $\mathcal{T}_k(\alpha)$ will be called *α-trees*.

§3. First-order theories of coloured linear orders.

3.1. A catalogue of properties and axioms. A (possibly coloured) tree $\mathfrak{T}$, in particular, a linear order, is called:

- *forward discrete*, if for any two elements a and b, such that $a < b$, there exists an immediate successor c of a such that $a < c \leqslant b$. In the case of linear orders, forward discreteness is equivalent to stating that each non-greatest element has an immediate successor.
- *backward discrete*, when each element except for the least element, if there is one, has an immediate predecessor.
- *discrete*, when it is both forward discrete and backward discrete.
- *definably forward well-founded*, if every parametrically first-order definable non-empty set of elements contains a maximal element.[1]

---

[1]Note that, in the case of linear orders, maximal elements and greatest elements coincide, but for proper trees they need not.

- *definably backward well-founded*, if every parametrically first-order definable non-empty set of elements contains a minimal (which will also be a least) element.
- *definably bounded forward well-founded*, if every parametrically first-order definable non-empty set of elements X that is bounded above (i.e., for which there exists b such that $x \leqslant b$ for each $x \in X$) contains a maximal element.
- *definably bounded backward well-founded*, if every parametrically first-order definable non-empty set of elements X that is bounded below (i.e., there exists a such that $a \leqslant x$ for each $x \in X$) contains a minimal element.

We define the following sentences in the first-order language L_k:

1. LO expresses that the structure is a linear order, as the conjunction of the sentences expressing irreflexivity, transitivity, and totality:

$$\mathsf{LO} := \forall x\,(x \not< x) \wedge \forall x \forall y \forall z\big((x < y \wedge y < z) \to x < z\big)$$
$$\wedge\, \forall x \forall y\,(x < y \vee y < x \vee x = y)\,.$$

2. Tree expresses that the structure is a tree, as the conjunction of sentences that express irreflexivity, transitivity, left-linearity, and left-connectedness:

$$\mathsf{Tree} := \forall x\,(x \not< x) \wedge \forall x \forall y \forall z\big((x < y \wedge y < z) \to x < z\big)$$
$$\wedge\, \forall x \forall y \forall z\big((y < x \wedge z < x) \to (y < z \vee z < y \vee y = z)\big)$$
$$\wedge\, \forall x \forall y \exists z\,(z \leqslant x \wedge z \leqslant y)\,.$$

3. Le is a sentence that expresses the existence of a $<$-least element:

$$\mathsf{Le} := \exists x \forall y\,(x \leqslant y)\,.$$

4. Gr is a sentence that expresses the existence of a $<$-maximal element:

$$\mathsf{Gr} := \exists x\big(\neg \exists y\,(x < y)\big).$$

5. FD is a sentence that expresses the property of forward discreteness:

$$\mathsf{FD} := \forall x \forall y(x < y \to \exists z(x < z \leqslant y \wedge \neg \exists u(x < u < z)))\,.$$

6. BD is a sentence that expresses the property of backward discreteness:[2]

$$\mathsf{BD} := \forall x(\exists y(y < x) \to \exists y(y < x \wedge \neg \exists z(y < z < x)))\,.$$

7. De is a sentence that expresses the property of density:

$$\mathsf{De} := \forall x \forall y\,(x < y \to \exists z\,(x < z < y))\,.$$

8. FWF_k is a scheme of (infinitely many) sentences that expresses the property of definable forward well-foundedness and consists of all sentences of the form

$$\mathsf{FWF}_k(\varphi) := \forall \bar{z}(\exists x(\varphi(x, \bar{z})) \to \exists x(\varphi(x, \bar{z}) \wedge \neg \exists y(\varphi(y, \bar{z}) \wedge x < y)))\,,$$

where $\bar{z} = (z_1, z_2, \dots, z_n)$ is any (possibly empty) tuple of variables and $\varphi(x, \bar{z})$ is any formula in the language L_k.

---

[2]The reason for the different forms of FD and BD is that trees are left-linear but not right-linear.

9. BWF_k is an infinite scheme that expresses the property of definable backward well-foundedness and consists of all sentences of the form

$$\mathsf{BWF}_k(\varphi) := \forall \bar{z}(\exists x(\varphi(x,\bar{z})) \to \exists x(\varphi(x,\bar{z}) \wedge \neg \exists y(\varphi(y,\bar{z}) \wedge y < x)))$$

for any formula $\varphi(x,\bar{z})$ in the language L_k.

10. BFWF_k is an infinite scheme that expresses the property of definable bounded forward well-foundedness and consists of all sentences of the form

$$\mathsf{BFWF}_k(\varphi) := \forall \bar{z}((\exists x(\varphi(x,\bar{z})) \wedge \exists y \forall x(\varphi(x,\bar{z}) \to x \leqslant y))$$
$$\to \exists x(\varphi(x,\bar{z}) \wedge \neg \exists y(\varphi(y,\bar{z}) \wedge x < y)))$$

for any formula $\varphi(x,\bar{z})$ in the language L_k.

11. BBWF_k is an infinite scheme that expresses the property of definable bounded backward well-foundedness and consists of all sentences of the form

$$\mathsf{BBWF}_k(\varphi) := \forall \bar{z}((\exists x(\varphi(x,\bar{z})) \wedge \exists y \forall x(\varphi(x,\bar{z}) \to y \leqslant x))$$
$$\to \exists x(\varphi(x,\bar{z}) \wedge \neg \exists y(\varphi(y,\bar{z}) \wedge y < x)))$$

for any formula $\varphi(x,\bar{z})$ in the language L_k.

The following facts about the first-order theory of ω will be needed later.

Fact 3.1. The first-order theory of ω can be axiomatised by the theory

$$\{\mathsf{LO}, \mathsf{Le}, \neg\mathsf{Gr}, \mathsf{BD}, \mathsf{FD}\}$$

(see, e.g., [16, p. 254]). Moreover, the class of models of this first-order theory consists precisely of all linear orders of the form $\omega + \zeta \cdot \alpha$ where α is any linear order (see, e.g., [16, Corollary 13.12 and Proposition 13.25]).

3.2. The first-order theory of coloured finite linear orders. We define the following first-order theory:

$$\mathsf{CFLO}_k := \{\mathsf{LO}, \mathsf{Le}, \mathsf{FD}\} \cup \mathsf{FWF}_k.$$

The models of CFLO_k will be called *coloured quasi-finite linear orders*.

Lemma 3.2. *Let* $\mathfrak{A} = (A; <, \bar{\mathbf{c}})$ *be a coloured quasi-finite linear order. Then*:

1. *A has a greatest element.*
2. *$\mathfrak{A}$ is backward discrete, hence (due to FD) also discrete.*
3. *$\mathfrak{A}$ is definably backward well-founded.*
4. *For each element $a \in A$, the structures $\mathfrak{A}^{<a}$, $\mathfrak{A}^{\leqslant a}$, $\mathfrak{A}^{>a}$, and $\mathfrak{A}^{\geqslant a}$ are coloured quasi-finite linear orders too.*

Proof. (1) The set A is itself definable, hence, by FWF_k, A contains a greatest element.

(2) Given a non-least element $a \in A$, the set $A^{<a}$ is definable in $(\mathfrak{A}; a)$, hence, by FWF_k, $A^{<a}$ contains a greatest element b, which must be the immediate predecessor of a.

(3) Suppose that $\varphi(x,\bar{z})$ is a formula such that $\varphi(x,\bar{a})$ defines a non-empty set B in $(\mathfrak{A}; \bar{a})$. If B contains the least element of A then the proof is completed. Suppose otherwise. Then the set $B^< := \{x \in A : x < y \text{ for all } y \in B\}$ is non-empty.

Let $\psi(x, \bar{z}) := \forall y (\varphi(y, \bar{z}) \to x < y)$. Then $\psi(x, \bar{a})$ defines $B^<$ in $(\mathfrak{A}; \bar{a})$, hence, by FWF$_k$, $B^<$ contains a greatest element b. Since $b < y$ for all $y \in B$ and $B \neq \emptyset$, then b is not the greatest element of A, hence, by FD, b has an immediate successor b' which will be the least element of B, as required.

(4) Straightforward. ⊣

More generally, observe that the following implications hold in any linear order:

$$\text{FWF}_k \Longrightarrow \text{BFWF}_k \Longrightarrow \text{BD and BWF}_k \Longrightarrow \text{BBWF}_k \Longrightarrow \text{FD}.$$

PROPOSITION 3.3. *If $\mathfrak{A} = (A; <, \bar{\mathbf{c}})$ is a coloured quasi-finite linear order then A is either finite or $\mathfrak{A}^-$ has the form $\omega + \zeta \cdot \alpha + \omega^\star$ for some linear order α.*

PROOF. Straightforward. ⊣

Proposition 3.4 below and its proof are similar to [2, Theorem 3.2] except that there, restricted definable induction instead of forward well-foundedness is used to approximate finiteness. Proposition 3.4 and its proof are also similar to the "vertical collapsing" construction used in [15] to obtain finite tree paths from infinite ones. The construction there, like Proposition 3.4, also approximates finiteness using definable forward well-foundedness. Here, however, we also give an upper bound for the size of the coloured finite linear order produced by the construction, that is n-equivalent to the given coloured quasi-finite linear order. Recall, from the discussion of characteristic sentences in Section 2.1, that $f(L_k, n)$ is the number of non-equivalent characteristic sentences of quantifier rank n in the language L_k of linear orders expanded with k colours, and that $f(L_k, n)$ is necessarily finite.

PROPOSITION 3.4. *For every coloured quasi-finite linear order $\mathfrak{A} = (A; <, \bar{\mathbf{c}}_k)$ with k colours and for each natural number n there exists a k-coloured finite linear order $\mathfrak{A}_n$ such that $\mathfrak{A} \equiv_n \mathfrak{A}_n$. Moreover, taking*

$$S := \{[\![\mathfrak{A}^{\geq u}]\!]^n : u \in A\}/\!\equiv$$

to be the set of equivalence classes of all the n-characteristic sentences of the orders $\mathfrak{A}^{\geq u}$ for $u \in A$, the cardinality of $\mathfrak{A}_n$ is given by $|A_n| = |S| \leqslant f(L_k, n)$.

PROOF. Let $\mathfrak{A}$ be a coloured quasi-finite linear order and let $m = |S|$. We start by defining three sequences of coloured linear orders: $\mathfrak{B}_1, \mathfrak{B}_2, \ldots, \mathfrak{B}_{m-1}, \mathfrak{C}_1, \mathfrak{C}_2, \ldots, \mathfrak{C}_{m-1}$, and $\mathfrak{D}_1, \mathfrak{D}_2, \ldots, \mathfrak{D}_{m-1}$, where $\mathfrak{B}_i = \mathfrak{C}_i + \mathfrak{D}_i$ and where $\mathfrak{C}_i$ and $\mathfrak{D}_i$ are substructures[3] of $\mathfrak{A}$ for each i, as described below. Informally, each $\mathfrak{C}_i$ will consist of the first i elements x in $\mathfrak{A}$ that are maximal with respect to satisfying their n-characteristic formula $[\![\mathfrak{A}^{\geq x}]\!]^n(x)$, and $\mathfrak{D}_i$ then consists of the elements in $\mathfrak{A}$ that are greater than all of the elements in $\mathfrak{C}_i$. Formally, $\mathfrak{B}_i, \mathfrak{C}_i$, and $\mathfrak{D}_i$ will have the following properties for each i:

  (i) $\mathfrak{B}_i \equiv_n \mathfrak{A}$;
 (ii) for each $u \in B_i$, $\mathfrak{B}_i^{\geq u} \equiv_n \mathfrak{A}^{\geq u}$;
(iii) $\mathfrak{C}_i$ is a coloured finite linear order with domain C_i, such that $|C_i| = i$;
(iv) $\mathfrak{D}_i$ is a coloured quasi-finite linear order of the form $\mathfrak{A}^{>u}$ for some $u \in A$;

---

[3]Elements $(x, 0)$ and $(y, 1)$ in $\mathfrak{B}_i + \mathfrak{C}_i$ will be identified with the elements $x \in B_i$ and $y \in C_i$ respectively, so that the elements of $\mathfrak{B}_i$ and $\mathfrak{C}_i$ may be treated as elements of $\mathfrak{A}$ too.

(v) for $u \in C_i$ and $v \in B_i$ with $u \neq v$, $\mathfrak{B}_i^{\geqslant u} \not\equiv_n \mathfrak{B}_i^{\geqslant v}$.

The exact construction of $\mathfrak{C}_i$ and $\mathfrak{D}_i$ will now be described. Let a be the least element of $\mathfrak{A}$. We define the formula

The set $\mathfrak{A}^{\varphi_0}$ is non-empty because $\mathfrak{A} \models \varphi_0(a)$, and it is definable. Hence, by the forward well-foundedness of $\mathfrak{A}$, $\mathfrak{A}^{\varphi_0}$ contains a greatest element a_0. Since $a, a_0 \in \mathfrak{A}^{\varphi_0}$ then $\mathfrak{A}^{\geqslant a} \equiv_n \mathfrak{A}^{\geqslant a_0}$ hence

$$\mathfrak{A} = \mathfrak{A}^{\geqslant a} \equiv_n \mathfrak{A}^{\geqslant a_0} \cong \mathfrak{A}^{[a_0, a_0]} + \mathfrak{A}^{> a_0}.$$

Let $\mathfrak{C}_1 := \mathfrak{A}^{[a_0, a_0]}$, $\mathfrak{D}_1 := \mathfrak{A}^{> a_0}$, and $\mathfrak{B}_1 := \mathfrak{C}_1 + \mathfrak{D}_1$. Observe that properties (i)–(v) are all satisfied by $\mathfrak{B}_1$, $\mathfrak{C}_1$, and $\mathfrak{D}_1$; property (iv) follows using part 4 of Lemma 3.2; property (v) follows using property (ii) along with the fact a_0 is the greatest element in $\mathfrak{A}^{\varphi_0}$; the other properties are straightforward.

Given $\mathfrak{B}_i$, $\mathfrak{C}_i$, and $\mathfrak{D}_i$ for some i with $1 \leqslant i < m - 1$, that satisfy properties (i)–(v), the coloured linear orders $\mathfrak{B}_{i+1}$, $\mathfrak{C}_{i+1}$, and $\mathfrak{D}_{i+1}$ are obtained as follows. Since $\mathfrak{D}_i$ is a coloured quasi-finite linear order, it contains a least element, say d_i. Define the formula

$$\varphi_i(x) := (\llbracket \mathfrak{D}_i \rrbracket^n)^{\geqslant x}.$$

Again, the set $\mathfrak{D}_i^{\varphi_i}$ is non-empty (because $\mathfrak{D}_i \models \varphi_i(d_i)$) and definable, hence, by the definable forward well-foundedness of $\mathfrak{D}_i$, $\mathfrak{D}_i^{\varphi_i}$ contains a greatest element e_i. Then

$$\mathfrak{D}_i = \mathfrak{D}_i^{\geqslant d_i} \equiv_n \mathfrak{D}_i^{\geqslant e_i} \cong \mathfrak{D}_i^{[e_i, e_i]} + \mathfrak{D}_i^{> e_i}. \tag{1}$$

Now take $\mathfrak{C}_{i+1} := \mathfrak{C}_i + \mathfrak{D}_i^{[e_i, e_i]}$, $\mathfrak{D}_{i+1} := \mathfrak{D}_i^{> e_i}$, and $\mathfrak{B}_{i+1} := \mathfrak{C}_{i+1} + \mathfrak{D}_{i+1}$, and again observe that properties (i)–(v) are all satisfied by $\mathfrak{B}_{i+1}$, $\mathfrak{C}_{i+1}$, and $\mathfrak{D}_{i+1}$: for property (i),

$$\mathfrak{B}_{i+1} = \mathfrak{C}_{i+1} + \mathfrak{D}_{i+1} = (\mathfrak{C}_i + \mathfrak{D}_i^{[e_i, e_i]}) + \mathfrak{D}_i^{> e_i}$$
$$\cong \mathfrak{C}_i + (\mathfrak{D}_i^{[e_i, e_i]} + \mathfrak{D}_i^{> e_i}) \equiv_n \mathfrak{C}_i + \mathfrak{D}_i = \mathfrak{B}_i \equiv_n \mathfrak{A}$$

with the first instance of $\equiv_n$ following from (1) and Lemma 2.1; properties (ii) and (iii) are straightforward; property (iv) again follows using part 4 of Lemma 3.2 using the fact that $\mathfrak{D}_i$ is a coloured quasi-finite linear order; property (v) follows using property (ii) for $\mathfrak{B}_{i+1}$, along with how $\mathfrak{C}_{i+1}$ and $\mathfrak{D}_{i+1}$ were constructed.

Finally, since $|S| = m$ there are at most m pairwise non-equivalent n-characteristic sentences that are satisfied by the structures $\mathfrak{B}_{m-1}^{\geqslant u}$ for $u \in B_{m-1}$. Using property (v), the $m - 1$ structures of the type $\mathfrak{B}_{m-1}^{\geqslant u}$, where u ranges over C_{m-1}, must all have pairwise non-equivalent n-characteristic sentences. Hence all structures of the type $\mathfrak{D}_{m-1}^{\geqslant u}$, where $u \in D_{m-1}$, must have equivalent n-characteristic sentences. In particular, the n-characteristic sentences of each of these structures $\mathfrak{D}_{m-1}^{\geqslant u}$ must be equivalent to the n-characteristic sentence of $\mathfrak{D}_{m-1}^{\geqslant b}$, where b is the greatest element of $\mathfrak{D}_{m-1}$. It follows that $|D_{m-1}| = 1$ and so $|B_{m-1}| = m$, i.e., $\mathfrak{A}_n := \mathfrak{B}_{m-1}$ may be taken to be a coloured finite linear order such that $\mathfrak{A} \equiv_n \mathfrak{A}_n$, as required. $\dashv$

LEMMA 3.5. *If $\mathfrak{A}$ and $\mathfrak{B}$ are k-coloured quasi-finite linear orders then $\mathfrak{A} + \mathfrak{B}$ is a k-coloured quasi-finite linear order, too.*

PROOF. Let n be any natural number. By Proposition 3.4 there exist k-coloured finite linear orders $\mathfrak{A}'$ and $\mathfrak{B}'$ such that $\mathfrak{A}' \equiv_n \mathfrak{A}$ and $\mathfrak{B}' \equiv_n \mathfrak{B}$. By Lemma 2.1, $\mathfrak{A} + \mathfrak{B} \equiv_n \mathfrak{A}' + \mathfrak{B}'$. Since $\mathfrak{A}' + \mathfrak{B}'$ is a k-coloured finite linear order then $\mathfrak{A}' + \mathfrak{B}' \models$ CFLO$_k$ and so $\mathfrak{A} + \mathfrak{B}$ satisfies all sentences of quantifier rank at most n in CFLO$_k$. Since n is arbitrary, it follows that $\mathfrak{A} + \mathfrak{B} \models$ CFLO$_k$, hence $\mathfrak{A} + \mathfrak{B}$ is itself a k-coloured quasi-finite linear order, as required. ⊣

3.3. The first-order theories of coloured well-orders and expansions of ω. We now define the following theories, which will be shown further to axiomatise the first-order theories of the classes of all coloured well-orders, of all coloured expansions of ω, and of all coloured expansions of $\omega^\star$ (the reverse order of ω), respectively:

$$CWO_k := \{LO\} \cup BWF_k,$$
$$C\omega_k := \{LO, \neg Gr, BD\} \cup BWF_k,$$
$$C\omega_k^\star := \{LO, \neg Le, FD\} \cup FWF_k.$$

FACT 3.6 ([2, Corollary 4.4]). *For every model $\mathfrak{A}$ of CWO$_k$ and each $n \in \mathbb{N}$, there exists a k-coloured well-order $\mathfrak{B}_n$ such that $\mathfrak{A} \equiv_n \mathfrak{B}_n$.*

Corollary 3.7 below is similar to [2, Theorem 3.1] but uses backward well-foundedness instead of definable induction for approximating the structure of coloured ω.

COROLLARY 3.7. *For every model $\mathfrak{A}$ of Cω_k and each $n \in \mathbb{N}$, there exists a k-coloured linear order $\mathfrak{B}_n$ such that $(\mathfrak{B}_n)^- \cong \omega$ and $\mathfrak{A} \equiv_n \mathfrak{B}_n$.*

PROOF. It suffices to prove the result for $n \geqslant 3$. By Fact 3.6, there exists a k-coloured well-order $\mathfrak{B}_n$ such that $\mathfrak{A} \equiv_n \mathfrak{B}_n$. Since the quantifier ranks of the sentences $\neg Gr$ and BD are at most 3, $\mathfrak{B}_n \models \{\neg Gr, BD\}$ and, since $\mathfrak{B}_n$ is a well-order, $\mathfrak{B}_n \models \{Le, FD\}$. By Fact 3.1, $(\mathfrak{B}_n)^-$ is a model of the first-order theory of ω and also a well-order, from which it follows that $(\mathfrak{B}_n)^- \cong \omega$. ⊣

Using a similar argument, it follows that the theory C$\omega_k^\star$ axiomatises the first-order theory of the class of coloured expansions of ω^*:

COROLLARY 3.8. *For every model $\mathfrak{A}$ of C$\omega_k^\star$ and each $n \in \mathbb{N}$, there exists a k-coloured linear order $\mathfrak{B}_n$ such that $(\mathfrak{B}_n)^- \cong \omega^*$ and $\mathfrak{A} \equiv_n \mathfrak{B}_n$.*

3.4. The first-order theory of coloured expansions of ζ. We now define the theory

$$C\zeta_k := \{LO, \neg Le, \neg Gr\} \cup BBWF_k \cup BFWF_k.$$

PROPOSITION 3.9. *For every model $\mathfrak{A}$ of Cζ_k and each $n \in \mathbb{N}$, there is a k-coloured linear order $\mathfrak{B}_n$, with $(\mathfrak{B}_n)^- \cong \zeta$, such that $\mathfrak{A} \equiv_n \mathfrak{B}_n$.*

PROOF. Let $\mathfrak{A} \models C\zeta_k$ and observe that $\mathfrak{A} \models \{BD, FD\}$. Pick any element $a \in A$ and let b be its immediate predecessor. Then $\mathfrak{A}^{\geqslant a}$ satisfies the theory Cω_k and $\mathfrak{A}^{\leqslant b}$ satisfies the theory C$\omega_k^\star$, hence, by Corollaries 3.7 and 3.8, there exists a k-coloured linear order $\mathfrak{C}$ such that $\mathfrak{C}^- \cong \omega$ and $\mathfrak{A}^{\geqslant a} \equiv_n \mathfrak{C}$, and there exists a k-coloured linear order $\mathfrak{D}$ such that $\mathfrak{D}^- \cong \omega^*$ and $\mathfrak{A}^{\leqslant b} \equiv_n \mathfrak{D}$. Then $\mathfrak{B}_n = \mathfrak{D} + \mathfrak{C}$ is the desired coloured linear order, since $\mathfrak{A} \equiv_n \mathfrak{B}_n$ follows from Lemma 2.1. ⊣

3.5. The first-order theory of coloured ordinals. Let α be a monochromatic ordinal such that $\omega < \alpha < \omega^\omega$. It is known (see [16, Section 13.2]) that the FO theory of α is finitely axiomatised, i.e., there exists a first-order sentence $\Phi(\alpha)$ in the language L_0 such that, for any monochromatic linear order $\mathfrak{A}$,

$$\mathfrak{A} \models \Phi(\alpha) \Longleftrightarrow \mathfrak{A} \equiv \alpha,$$

and α is the only ordinal that is a model of $\Phi(\alpha)$. The formulation of $\Phi(\alpha)$ is somewhat involved and for details we refer the reader to [16, Section 13.2]. We now define the theory

$$\mathsf{C}\alpha_k := \{\Phi(\alpha)\} \cup \mathsf{BWF}_k.$$

PROPOSITION 3.10. *For every model $\mathfrak{A}$ of $\mathsf{C}\alpha_k$ and each $n \in \mathbb{N}$ there is a k-coloured linear order $\mathfrak{B}_n$ such that $(\mathfrak{B}_n)^- \cong \alpha$ and $\mathfrak{A} \equiv_n \mathfrak{B}_n$.*

PROOF. Let $\mathfrak{A} \models \mathsf{C}\alpha_k$ and let $n \geqslant \mathrm{qr}(\Phi(\alpha))$. Since $\mathfrak{A} \models \Phi(\alpha)$ then $\mathfrak{A} \models \mathsf{LO}$. By Fact 3.6, there exists a k-coloured well-order $\mathfrak{B}_n$ such that $\mathfrak{A} \equiv_n \mathfrak{B}_n$, and since $\mathfrak{B}_n \models \Phi(\alpha)$ then $(\mathfrak{B}_n)^- \equiv \alpha$. Since $\mathfrak{B}_n$ is a well-order and α is the only ordinal that is a model of $\Phi(\alpha)$, it follows that $(\mathfrak{B}_n)^- \cong \alpha$, as required. $\dashv$

3.6. The first-order theory of coloured expansions of η. We now define the theory

$$\mathsf{C}\eta_k := \{\mathsf{LO}, \neg\mathsf{Le}, \neg\mathsf{Gr}, \mathsf{De}\}.$$

PROPOSITION 3.11. *For every model $\mathfrak{A}$ of $\mathsf{C}\eta_k$, there is a k-coloured linear order $\mathfrak{B}$, with $\mathfrak{B}^- \cong \eta$, such that $\mathfrak{A} \equiv \mathfrak{B}$.*

PROOF. By the Löwenheim–Skolem Theorem, there exists a countable structure $\mathfrak{B}$ such that $\mathfrak{B} \equiv \mathfrak{A}$. Since $\mathfrak{B} \models \mathsf{C}\eta_k$, $\mathfrak{B}$ is a countable k-coloured dense linear order without endpoints, hence $\mathfrak{B}^- \cong \eta$. $\dashv$

3.7. The first-order theory of coloured expansions of λ. An axiomatisation of the first-order theory of the class of coloured expansions of λ is obtained in [2]. For the sake of completeness of this paper we will briefly describe that axiomatisation here, as it will be needed in Section 5 to axiomatise the first-order theory of the class of λ-trees.

A coloured linear order is called *definably bounded complete* when each of its non-empty definable subsets that is bounded above has a least upper bound. For any formula $\varphi(x, \bar{z})$ (where the tuple $\bar{z}$ may be empty) we define the formula

$$\mathsf{ub}_\varphi(u, \bar{z}) := \forall x\,(\varphi(x, \bar{z}) \to x \leqslant u),$$

which expresses that u is an upper bound of the set that is defined by $\varphi(x, \bar{z})$. The property of definable bounded completeness can now be expressed by the set COM_k that consists of all sentences of the form

$$\begin{aligned}
\mathsf{COM}_k(\varphi) := \forall \bar{z}((\exists x(\varphi(x, \bar{z})) \wedge \exists u(\mathsf{ub}_\varphi(u, \bar{z}))) \\
\to \exists u(\mathsf{ub}_\varphi(u, \bar{z}) \wedge \forall v(\mathsf{ub}_\varphi(v, \bar{z}) \to u \leqslant v)))
\end{aligned}$$

for any formula $\varphi(x, \bar{z})$ in the language L_k.

Given a linear order $\mathfrak{A}$ and a partition P of A of which all members are intervals, we define the binary relation $<_\mathfrak{P}$ on P by specifying, for $X, Y \in P$, that $X <_\mathfrak{P} Y$

iff $x <_{\mathfrak{A}} y$ for some $x \in X$ and $y \in Y$. The linear order $\mathfrak{P} := (P; <_{\mathfrak{P}})$ is called a *condensation* of $\mathfrak{A}$. Given a transitive binary relation R on A, we define the relation $\sim_R$ on A by specifying, for $u, v \in A$, that $u \sim_R v$ if and only if any of the following conditions hold:

(i) $u = v$,
(ii) $u <_{\mathfrak{A}} v$ and sRt for all $s, t \in A$ such that $u \leqslant_{\mathfrak{A}} s <_{\mathfrak{A}} t \leqslant_{\mathfrak{A}} v$, or
(iii) $v <_{\mathfrak{A}} u$ and sRt for all $s, t \in A$ such that $v \leqslant_{\mathfrak{A}} s <_{\mathfrak{A}} t \leqslant_{\mathfrak{A}} u$.

It is straightforward to see that for any transitive binary relation R, the relation $\sim_R$ is an equivalence relation, hence the set $Q = {}^A/_{\sim_R}$ is a partition of A into intervals, and so the structure $\mathfrak{Q} = (Q; <_{\mathfrak{Q}})$ is a condensation of $\mathfrak{A}$. The relation R is said to *induce* the condensation $\mathfrak{Q}$. If the relation R is definable then $\mathfrak{Q}$ is called a *definable condensation*. For any formula $\varphi(x, y, \bar{z})$ (with the tuple $\bar{z}$ possibly empty) we define the formulas

$$\mathsf{bin}_\varphi(\bar{z}) := \forall x \forall y \forall v ((\varphi(x, y, \bar{z}) \wedge \varphi(y, v, \bar{z})) \to \varphi(x, v, \bar{z})),$$

$$\mathsf{part}_\varphi(u, v, \bar{z}) := u < v \wedge \forall s \forall t (u \leqslant s < t \leqslant v \to \varphi(s, t, \bar{z})),$$

$$\mathsf{eq}_\varphi(u, v, \bar{z}) := u = v \vee \mathsf{part}_\varphi(u, v, \bar{z}) \vee \mathsf{part}_\varphi(v, u, \bar{z}),$$

$$\mathsf{dense}_\varphi(\bar{z}) := \forall u \forall v \big(\big(u < v \wedge \neg \mathsf{eq}_\varphi(u, v, \bar{z})\big)$$
$$\to \exists w (u < w < v \wedge \neg \mathsf{eq}_\varphi(u, w, \bar{z}) \wedge \neg \mathsf{eq}_\varphi(w, v, \bar{z}))), \text{ and}$$

$$\mathsf{singleton}_\varphi(\bar{z}) := \forall u \forall v ((u < v \wedge \neg \mathsf{eq}_\varphi(u, v, \bar{z})) \to \exists w (u < w < v$$
$$\wedge \neg \mathsf{eq}_\varphi(u, w, \bar{z}) \wedge \neg \mathsf{eq}_\varphi(w, v, \bar{z}) \wedge \forall t (\mathsf{eq}_\varphi(w, t, \bar{z}) \to t = w))).$$

The formula $\mathsf{bin}_\varphi(\bar{z})$ expresses that $\varphi(x, y, \bar{z})$ defines a transitive binary relation R, the formula $\mathsf{part}_\varphi(u, v, \bar{z})$ expresses property (ii) in the above definition of the relation $\sim_R$, $\mathsf{eq}_\varphi(u, v, \bar{z})$ expresses that $u \sim_R v$, $\mathsf{dense}_\varphi(\bar{z})$ expresses that the condensation that is induced by R is densely ordered, and $\mathsf{singleton}_\varphi(\bar{z})$ expresses that the condensation that is induced by R contains a set of singletons that is dense in the condensation.

Following the terminology of [2], a linear order $\mathfrak{A}$ is called *definably-I* when each of its densely ordered definable condensations contains a set of singletons that is dense in the condensation. The property of a k-coloured linear order being definably-I approximates the countable chain condition of $\mathbb{R}$, and can be expressed by the set I_k that consists of all sentences of the form

$$\mathsf{I}_k(\varphi) := \forall \bar{z} \left(\big(\mathsf{bin}_\varphi(\bar{z}) \wedge \mathsf{dense}_\varphi(\bar{z}) \big) \to \mathsf{singleton}_\varphi(\bar{z}) \right)$$

for any formula $\varphi(x, y, \bar{z})$ in the language L_k and with the tuple $\bar{z}$ possibly empty.

Finally, the first-order theory of the class of k-coloured expansions of λ is axiomatised by the theory

$$C\lambda_k := \{\mathsf{LO}, \neg\mathsf{Le}, \neg\mathsf{Gr}, \mathsf{De}\} \cup \mathsf{COM}_k \cup \mathsf{I}_k,$$

as is seen from the following fact, the proof of which can be found in [2]:

FACT 3.12 [2]. *For every model* $\mathfrak{A}$ *of* $C\lambda_k$ *and each* $n \in \mathbb{N}$, *there exists a* k-coloured *linear order* $\mathfrak{B}_n$ *such that* $(\mathfrak{B}_n)^- \cong \lambda$ *and* $\mathfrak{A} \equiv_n \mathfrak{B}_n$.

§4. From trees to coloured linear orders and back.

4.1. Trees as coloured linear orders. A tree $\mathfrak{T}$ is called *branching complete* when for any two distinct paths A and B in $\mathfrak{T}$, each of the sets $A \setminus B$ and $B \setminus A$ has an infimum in $\mathfrak{T}$. Branching completeness is equivalent to the following property: for any antichain $\{a, b\} \subseteq T$, the set

$$T_{ab} := \{x \in T^{\leqslant a} : x \geqslant y \text{ for each } y \in T^{\leqslant a} \cap T^{\leqslant b}\} \tag{2}$$

has a least element. Recall the formula $\mathsf{ub}_\varphi(u, \bar{z})$ that was defined in Section 3.7 and which expresses the fact that u is an upper bound of the set that is defined by $\varphi(x, \bar{z})$. Letting $\varphi(x, z_1, z_2) := x \leqslant z_1 \wedge x \leqslant z_2$, the property of branching completeness can therefore be expressed by the sentence

$$
\begin{aligned}
\mathsf{bc} := \forall z_1 \forall z_2 (\neg(z_1 < z_2 \vee z_2 < z_1 \vee z_1 = z_2) \\
\rightarrow \exists u(u \leqslant z_1 \wedge \mathsf{ub}_\varphi(u, z_1, z_2) \wedge \forall v((v \leqslant z_1 \wedge \mathsf{ub}_\varphi(v, z_1, z_2)) \rightarrow u \leqslant v))).
\end{aligned}
$$

A sufficient condition for branching completeness is that every path in the tree is a complete linear order. Thus, in particular, every λ-tree is branching complete.

REMARK 4.1. There are other notions of completeness of trees known from the literature, such as Dedekind–MacNeille completeness and others, cf.[4] [4, 12, 17]. The notion of branching completeness introduced here is mostly ad hoc, as it is sufficient for our purposes, but it is also of some independent interest. As shown in [12], branching completeness is generally weaker than Dedekind–MacNeille completeness of trees.

For $\mathfrak{T}$ any tree, A any stem in $\mathfrak{T}$ (A will usually be a path), and $s \in A$, we define the following sets (to be explained shortly):

$$
F_l(s) := \begin{cases} \emptyset & \text{when } s \text{ has an immediate predecessor,} \\ \left(\bigcap_{t < s} T^{>t} \right) \setminus T^{\geqslant s} & \text{otherwise,} \end{cases}
$$

$$
\text{and} \quad F_u(s \backslash \mathsf{A}) := T^{>s} \setminus \left(\bigcup_{t \in \mathsf{A} \cap T^{>s}} T^{\geqslant t} \right),
$$

and, provided that these sets are non-empty, define the structures $\mathfrak{F}_l(s) := \mathfrak{T}^{F_l(s)}$ and $\mathfrak{F}_u(s \backslash \mathsf{A}) := \mathfrak{T}^{F_u(s \backslash \mathsf{A})}$. In general, either or both of the sets $F_l(s)$ and $F_u(s \backslash \mathsf{A})$ may be empty, in which case the corresponding structure is left undefined. The structures $\mathfrak{F}_l(s)$ and $\mathfrak{F}_u(s \backslash \mathsf{A})$, if defined, are forests and will be called the *lower side-forest* of s, and the *upper side-forest* of s with respect to A, respectively. Finally, the *side-forest* of s with respect to A is the forest $\mathfrak{F}(s \backslash \mathsf{A}) := (\mathfrak{T}^{F(s \backslash \mathsf{A})}; s)$ where

$$F(s \backslash \mathsf{A}) := \{s\} \cup F_l(s) \cup F_u(s \backslash \mathsf{A}).$$

The forests $\mathfrak{F}_l(s)$ and $\mathfrak{F}_u(s \backslash \mathsf{A})$ are depicted in Figure 2.

Intuitively, the upper side-forest of s with respect to A consists of those nodes that sit above s but do not sit above any nodes on $\mathsf{A} \cap T^{>s}$, and the lower side-forest of

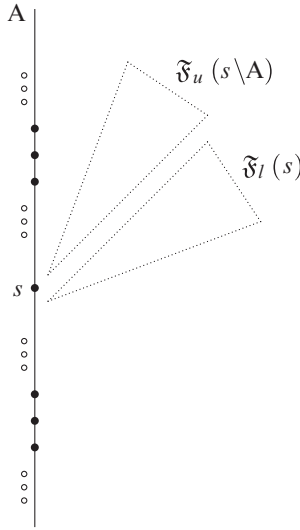---

[4]We thank the reviewer for the first two references.

FIGURE 2. The side-forests $\mathfrak{F}_l(s)$ and $\mathfrak{F}_u(s\backslash\mathsf{A})$.

s consists of those nodes that sit above $T^{<s}$ but are incomparable with s, unless s has an immediate predecessor, in which case its lower side-forest is left empty so as not to coincide with the upper side-forest of that predecessor. Observe that if A is a stem with a greatest node b then $F_u(b\backslash\mathsf{A}) = T^{>b}$.

The property of branching completeness ensures that if A is either a path in $\mathfrak{T}$, or a stem with a greatest node, then the set $\{F(s\backslash\mathsf{A})\}_{s\in\mathsf{A}}$ forms a partition of T. Indeed, each side-forest $F(s\backslash\mathsf{A})$ is non-empty since $s \in F(s\backslash\mathsf{A})$, and different side-forests $F(s\backslash\mathsf{A})$ and $F(t\backslash\mathsf{A})$ are disjoint by the way that side-forests are defined. To see that $\{F(s\backslash\mathsf{A})\}_{s\in\mathsf{A}}$ covers T, pick any node $u \in T$, and consider the following cases.

> Case 1: $u \in \mathsf{A}$. Then $u \in F(u\backslash\mathsf{A})$.
> Case 2: $u \notin \mathsf{A}$.
>> Case 2.1: $u \geqslant x$ for each $x \in \mathsf{A}$. Then A must be a stem with a greatest node b, in which case $u \in F(b\backslash\mathsf{A})$.
>> Case 2.2: There exists $v \in \mathsf{A}$ such that $\{u, v\}$ is an antichain. By the branching completeness of $\mathfrak{T}$, the set T_{vu} has a least element $w \in \mathsf{A}$.
>>> Case 2.2.1: $w < u$. Then $u \in F_u(w\backslash\mathsf{A})$.
>>> Case 2.2.2: $w \not< u$. Then $u \in F_l(w)$.

This shows that $\bigcup_{s\in\mathsf{A}} F(s\backslash\mathsf{A}) = T$.

Given $s, t \in \mathsf{A}$ such that $t > s$, the sets $F_l(s)$, $F_u(s\backslash\mathsf{A})$, and $F(s\backslash\mathsf{A})$ can be defined in $(\mathfrak{T}; s, t)$ (the expansion of $\mathfrak{T}$ obtained by adding the constants s and t) by the formulas

$$\mathsf{lsf}(x, s) := \big(\exists u\big(u < s \wedge \neg\exists v\,(u < v < s)\big) \to \neg(x = x)\big)$$
$$\wedge \big(\forall u\big(u < s \to \exists v\,(u < v < s)\big) \to \big(\forall w\,(w < s \to w < x) \wedge \neg(s \leqslant x)\big)\big),$$

$\mathsf{usf}\,(x,s,t) := s < x \land \neg \exists u\,(s < u \leqslant t \land u \leqslant x)\,,$ and

$\mathsf{sf}\,(x,s,t) := (x = s) \lor \mathsf{lsf}\,(x,s) \lor \mathsf{usf}\,(x,s,t)\,.$

(The first conjunct in $\mathsf{lsf}\,(x,s)$ covers the case where s has an immediate predecessor, and the second conjunct covers the case where s has no immediate predecessor.)

Now suppose that the tree $\mathfrak{T}$ has $k \geq 0$ colours, say $\mathfrak{T} = (T; <, \bar{\mathbf{c}})$ where $\bar{\mathbf{c}} = \bar{\mathbf{c}}_k$ (or, simply $\mathfrak{T} = (T; <)$ if $k = 0$). Recall that L_k' denotes the signature L_k of k-coloured trees, expanded with a single constant symbol, and $f\left(L_k', n\right)$ denotes the number of non-equivalent n-characteristic sentences of k-coloured trees that have an added constant symbol. Let $p = f\left(L_k', n\right)$ and define an additional set of colours, called *extended colours*, by $\bar{\mathbf{e}} := \bar{\mathbf{e}}_p = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p)$. Recall the definition of the sentence $\tau_{L,n,i}$ from the discussion on characteristic formulas in Section 2.1. Now, consider the coloured linear order

$$\mathsf{A}^{\mathfrak{T}}\,[\bar{\mathbf{e}}] := \left(\mathfrak{T}^\mathsf{A}; \bar{\mathbf{e}}\right) = (\mathsf{A}; <, \bar{\mathbf{c}}, \bar{\mathbf{e}}) \quad \text{(or, just } (\mathsf{A}; <, \bar{\mathbf{e}}) \text{ when } \mathfrak{T} \text{ is monochromatic)}$$

and where the extended colours are defined by specifying, for each $s \in \mathsf{A}$ and each i such that $1 \leqslant i \leqslant p$,

$$\mathsf{A}^{\mathfrak{T}}\,[\bar{\mathbf{e}}] \models \mathbf{e}_i\,(s) \quad \text{if and only if} \quad \mathfrak{F}\,(s \backslash \mathsf{A}) \models \tau_{L_k', n, i},$$

i.e., the extended colour of s corresponds to the n-characteristic sentence that is satisfied in $\mathfrak{F}\,(s \backslash \mathsf{A})$.

The linear order $\mathsf{A}^{\mathfrak{T}}\,[\bar{\mathbf{e}}]$ can be viewed as the k-coloured stem A of $\mathfrak{T}$ that is further enriched with p additional colours, the role of which is to capture, for each $s \in \mathsf{A}$, the n-equivalence class of the side-forest $\mathfrak{F}\,(s \backslash \mathsf{A})$ when viewed within $\mathfrak{T}$. This will allow us to approximate a tree as a coloured linear order, and thereby adapt results on coloured linear orders to trees.

The next theorem relates the first-order equivalence of two trees to the first-order equivalence of their corresponding coloured linear orders.

THEOREM 4.2. *Let $k \geq 0$ and $\mathfrak{T}_1$ and $\mathfrak{T}_2$ be k-coloured branching complete trees. Let A and B either both be paths, or both be stems with greatest elements in $\mathfrak{T}_1$ and $\mathfrak{T}_2$ respectively. Let $n \in \mathbb{N}$ and $\bar{\mathbf{e}} = \bar{\mathbf{e}}_{f\left(L_k', n\right)}$. For any $m \in \mathbb{N}$:*

1. *If $(\mathfrak{T}_1; \mathsf{A}) \equiv_{m+n+1} (\mathfrak{T}_2; \mathsf{B})$ then $\mathsf{A}^{\mathfrak{T}_1}\,[\bar{\mathbf{e}}] \equiv_m \mathsf{B}^{\mathfrak{T}_2}\,[\bar{\mathbf{e}}]$.*
2. *If $\mathsf{A}^{\mathfrak{T}_1}\,[\bar{\mathbf{e}}] \equiv_n \mathsf{B}^{\mathfrak{T}_2}\,[\bar{\mathbf{e}}]$ then $(\mathfrak{T}_1; \mathsf{A}) \equiv_n (\mathfrak{T}_2; \mathsf{B})$.*

PROOF. The result will be proved assuming that A and B are paths; the proof for the case where they are instead stems with greatest elements, is identical.

(1) We will reason by contraposition. Suppose that Spoiler has a winning strategy Σ in the game $\mathrm{EF}_m\left(\mathsf{A}^{\mathfrak{T}_1}\,[\bar{\mathbf{e}}], \mathsf{B}^{\mathfrak{T}_2}\,[\bar{\mathbf{e}}]\right)$. Then Spoiler can win the game $\mathrm{EF}_{m+n+1}((\mathfrak{T}_1; \mathsf{A}), (\mathfrak{T}_2; \mathsf{B}))$ as follows. For the first m rounds of the game, Spoiler plays only on the paths A and B (so, Duplicator will have no choice but to also play on these paths) by choosing nodes according to the strategy Σ in the parallel game $\mathrm{EF}_m\left(\mathsf{A}^{\mathfrak{T}_1}\,[\bar{\mathbf{e}}], \mathsf{B}^{\mathfrak{T}_2}\,[\bar{\mathbf{e}}]\right)$. If, after m rounds, the resulting nodes $a_1, a_2, \ldots, a_m \in \mathsf{A}$ and $b_1, b_2, \ldots, b_m \in \mathsf{B}$ (where a_i and b_i are the nodes chosen by the players in the i-th round) do not form a local isomorphism between the trees $(\mathfrak{T}_1; \mathsf{A})$ and $(\mathfrak{T}_2; \mathsf{B})$ then Spoiler wins the game $\mathrm{EF}_{m+n+1}((\mathfrak{T}_1; \mathsf{A}), (\mathfrak{T}_2; \mathsf{B}))$. Now, suppose the nodes do form a local isomorphism between $(\mathfrak{T}_1; \mathsf{A})$ and $(\mathfrak{T}_2; \mathsf{B})$. Then there must exist nodes

$a_q \in \mathsf{A}$ and $b_q \in \mathsf{B}$ such that a_q and b_q have different extended colours in $\mathsf{A}^{\mathfrak{T}_1}[\bar{\mathbf{e}}]$ and $\mathsf{B}^{\mathfrak{T}_2}[\bar{\mathbf{e}}]$. Spoiler can then win the game $\mathrm{EF}_{m+n+1}((\mathfrak{T}_1;\mathsf{A}),(\mathfrak{T}_2;\mathsf{B}))$ in the next $n+1$ rounds as follows.

First observe that, in rounds $m+1$ to $m+n$ of $\mathrm{EF}_{m+n+1}((\mathfrak{T}_1;\mathsf{A}),(\mathfrak{T}_2;\mathsf{B}))$, whenever Spoiler chooses a node a_i from $F(a_q \backslash \mathsf{A})$ in round i of the game, Duplicator has to choose her response from $F(b_q \backslash \mathsf{B})$, and vice versa. To see this, suppose that $a_i \in F(a_q \backslash \mathsf{A})$ but that $b_i \notin F(b_q \backslash \mathsf{B})$. Then $a_i \neq a_q$ for if $a_i = a_q$ then b_i would have to be the node b_q thus giving $b_i \in F(b_q \backslash \mathsf{B})$, a contradiction. It follows that $a_i \notin \mathsf{A}$, hence b_i cannot lie on B. So, there are two cases to consider: either $a_q < a_i$ or a_i is incomparable with a_q. If $a_q < a_i$ then it must be the case that $b_q < b_i$ too, from which, since $b_i \notin F(b_q \backslash \mathsf{B})$, there exists $b \in \mathsf{B}$ such that $b_q < b < b_i$. If Spoiler then chooses the node $b \in T_2$ in round $i+1$ of the game, he will have won, since there is no node $a \in \mathsf{A}$ for Duplicator to choose in round $i+1$ that satisfies $a_q < a < a_i$. The case where a_i is incomparable with a_q similarly leads to a win for Spoiler. Now since a_q and b_q have different extended colours in $\mathsf{A}^{\mathfrak{T}_1}[\bar{\mathbf{e}}]$ and $\mathsf{B}^{\mathfrak{T}_2}[\bar{\mathbf{e}}]$ then $\mathfrak{F}(a_q \backslash \mathsf{A}) \not\equiv_n \mathfrak{F}(b_q \backslash \mathsf{B})$. By using a winning strategy for $\mathrm{EF}_n(\mathfrak{F}(a_q \backslash \mathsf{A}),\mathfrak{F}(b_q \backslash \mathsf{B}))$, Spoiler can create a configuration of nodes in rounds $m+1$ through to n of the game $\mathrm{EF}_{m+n+1}((\mathfrak{T}_1;\mathsf{A}),(\mathfrak{T}_2;\mathsf{B}))$ that Duplicator cannot match.

(2) Suppose that Duplicator has a winning strategy in the n-round game $\mathrm{EF}_n(\mathsf{A}^{\mathfrak{T}_1}[\bar{\mathbf{e}}],\mathsf{B}^{\mathfrak{T}_2}[\bar{\mathbf{e}}])$. By decomposing the tree $(\mathfrak{T}_1;\mathsf{A})$ into the path A along with all the side-forests $\mathfrak{F}(s \backslash \mathsf{A})$, and similarly for the tree $(\mathfrak{T}_2;\mathsf{B})$, Duplicator can win the game $\mathrm{EF}_n((\mathfrak{T}_1;\mathsf{A}),(\mathfrak{T}_2;\mathsf{B}))$ by combining her winning strategy for the game $\mathrm{EF}_n(\mathsf{A}^{\mathfrak{T}_1}[\bar{\mathbf{e}}],\mathsf{B}^{\mathfrak{T}_2}[\bar{\mathbf{e}}])$ with her local strategies for the games $\mathrm{EF}_n(\mathfrak{F}(s \backslash \mathsf{A}),\mathfrak{F}(t \backslash \mathsf{B}))$ for nodes s from $\mathsf{A}^{\mathfrak{T}_1}[\bar{\mathbf{e}}]$ and t from $\mathsf{B}^{\mathfrak{T}_2}[\bar{\mathbf{e}}]$ that have the same extended colours in $\mathsf{A}^{\mathfrak{T}_1}[\bar{\mathbf{e}}]$ and $\mathsf{B}^{\mathfrak{T}_2}[\bar{\mathbf{e}}]$ respectively.   ⊣

Let $k,n \in \mathbb{N}$ and $p = f(L_k',n)$. Given a formula $\varphi(\bar{x})$ (where φ may be a sentence) in the language L_{k+p} with colours $\bar{\mathbf{c}}_k$ and $\bar{\mathbf{e}}_p$, and a variable z that does not occur in φ, let $\varphi_z^n(\bar{x})$ denote the formula that is obtained from $\varphi^{<z}$ (the relativisation of φ to the formula $\theta(u,z) := u < z$ where z is treated as a parameter) by replacing, for each i, for $1 \leqslant i \leqslant p$, every atomic formula of the form $\mathbf{e}_i(y)$ that occurs in $\varphi^{<z}$ with the formula $\tau_{L_k',n,i}^{\mathsf{sf}(u,y,z)}$ (i.e., $\tau_{L_k',n,i}$ relativised to $\mathsf{sf}(u,y,z)$ where y and z are treated as parameters).

In other words, the formula φ_z^n is the same as the relativised formula $\varphi^{<z}$, except that instead of stating that a node y has colour $\mathbf{e}_i$, φ_z^n states that the side-forest of y with respect to the stem $T^{<z}$ satisfies the characteristic sentence $\tau_{L_k',n,i}$. If φ does not contain any atomic formulas of the form $\mathbf{e}_i(y)$ then φ_z^n is simply the formula $\varphi^{<z}$.

The following result, which relates the truth of a first-order formula in a tree to the truth of that formula in the coloured linear order approximating that tree, now follows immediately from the definition of φ_z^n.

COROLLARY 4.3. *Let $k \geq 0$ and $\mathfrak{T}$ be a k-coloured tree with colours $\bar{\mathbf{c}}_k$, $n \in \mathbb{N}$, $p = f(L_k',n)$, and $\bar{\mathbf{e}} = \bar{\mathbf{e}}_p$, and let $\varphi(\bar{x})$ be a formula in the language L_{k+p} with colours $\bar{\mathbf{c}}_k$ and $\bar{\mathbf{e}}_p$. For each $a \in T$ and tuple $\bar{b}$ of nodes in $T^{<a}$ of the same arity as $\bar{x}$,*

$$(\mathfrak{T};a,\bar{b}) \models \varphi_a^n(\bar{b}) \text{ if and only if } ((T^{<a})^{\mathfrak{T}}[\bar{\mathbf{e}}];\bar{b}) \models \varphi(\bar{b}).$$

In particular, if φ is a sentence then for each $a \in T$, $(\mathfrak{T}; a) \models \varphi_a^n$ if and only if $(T^{<a})^{\mathfrak{T}}[\bar{\mathbf{e}}] \models \varphi$.

4.2. Towards axiomatising the first-order theory of ζ-trees. We will now use Theorem 4.2 and Corollary 4.3 to show how the first-order theories of certain classes of coloured $\mathcal{C}$-trees can be axiomatised. The method used in this section will have the set $\mathcal{C} = \{\zeta\}$ in mind, but can also be used on the sets $\mathcal{C} = \{\omega^\star\}$ and $\mathcal{C} = \{\zeta, \omega^\star\}$.

Say that a class of monochromatic linear orders $\mathcal{C}$ has the *fusion closure property* when the following condition holds:

> For any two linear orders $\mathfrak{A}, \mathfrak{B} \in \mathcal{C}$ and for any $a \in A$ and $b \in B$, $\mathfrak{A}^{<a} + \mathfrak{B}^{\geq b}$ is isomorphic to a linear order in $\mathcal{C}$.

Consider any class $\mathcal{C}$ of monochromatic linear orders with the following properties:

(C1) $\mathcal{C}$ does not contain the singleton linear order;
(C2) $\mathcal{C}$ has the fusion closure property;
(C3) all linear orders in $\mathcal{C}$ are discrete; and
(C4) every $\mathcal{C}$-tree is branching complete.

The property (C1) is needed to eliminate degenerate cases, (C2) is needed to ensure that paths belong to $\mathcal{C}$ when constructing models of the axiomatisation that will be given below, (C3) is to ensure that all components in the lower and upper side-forests of any $\mathcal{C}$-tree have roots, and (C4) is needed to ensure that the set of side-forests along any path form a partition of the tree.

Now let k be any natural number and let $\mathcal{S} = \mathcal{T}_k(\mathcal{C})$ be the class of k-coloured $\mathcal{C}$-trees. Suppose that the following are known:

(i) an axiomatisation $\Sigma_\uparrow^k$ of the first-order theory of the class of k-coloured trees

$$\left\{ \mathfrak{T}^{\geq a} : \mathfrak{T} \in \mathcal{S},\ a \in T \right\};$$

(ii) for each natural number n, an axiomatisation $\Sigma_\downarrow^{k,n}$ of the first-order theory of the class of $\left(k + f\left(L_k', n \right) \right)$-coloured linear orders

$$\{ (T^{<a})^{\mathfrak{T}}[\bar{\mathbf{e}}] : \mathfrak{T} \in \mathcal{S},\ a \in T \}.$$

Let

$$(\Sigma_\uparrow^k)' := \{ \forall x \left(\sigma^{\geq x} \right) : \sigma \in \Sigma_\uparrow^k \} \text{ and}$$
$$(\Sigma_\downarrow^k)' := \{ \forall z (\exists y (y < z) \rightarrow \sigma_z^n) : n \in \mathbb{N},\ \sigma \in \Sigma_\downarrow^{k,n} \}.$$

In essence, the theory $(\Sigma_\uparrow^k)'$ expresses for a k-coloured tree $\mathfrak{T}$ that each principal subtree of $\mathfrak{T}$ satisfies $\Sigma_\uparrow^k$, while, using Corollary 4.3, the theory $(\Sigma_\downarrow^k)'$ expresses of $\mathfrak{T}$ that, for each natural number n and each principal stem A in $\mathfrak{T}$, the linear order $A^{\mathfrak{T}}[\bar{\mathbf{e}}]$ satisfies $\Sigma_\downarrow^{k,n}$.

Then, the first-order theory of $\mathcal{S}$ can be axiomatised by the theory[5]

$$\mathsf{Ax}\,(\mathcal{S}) := \{\mathsf{Tree}, \mathsf{bc}\} \cup (\Sigma_\uparrow^k)' \cup (\Sigma_\downarrow^k)'$$

as is shown in the next result.

THEOREM 4.4. *Let $\mathcal{C}$ a class of linear orders that satisfies* (C1)–(C4), *and let $\mathcal{S} = \mathcal{T}_k(\mathcal{C})$ be the class of k-coloured $\mathcal{C}$-trees, for some $k \in \mathbb{N}$. For each k-coloured model $\mathfrak{T}$ of $\mathsf{Ax}\,(\mathcal{S})$ and each $n \in \mathbb{N}$, there exists a k-coloured tree $\mathfrak{S}_n \in \mathcal{S}$ such that $\mathfrak{T} \equiv_n \mathfrak{S}_n$.*

PROOF. Let $\mathfrak{T} = (T; <, \bar{\mathbf{c}})$ be a k-coloured model of $\mathsf{Ax}\,(\mathcal{S})$. Fix $n \in \mathbb{N}$; without loss of generality, we can assume that $n \geqslant \mathsf{qr}\,(\mathsf{bc})$. Now, let $p = f\left(L_k', n\right)$ and $\bar{\mathbf{e}} = \bar{\mathbf{e}}_p$. Since $\mathfrak{T} \models \mathsf{Tree}$ then $\mathfrak{T}$ is a tree. Pick any node $a \in T$ other than the root of $\mathfrak{T}$ (if there is a root) and let $A = T^{<a}$. Observe that $\mathfrak{T}$ is discrete (the property of discreteness is encoded in both of the theories $(\Sigma_\uparrow^k)'$ and $(\Sigma_\downarrow^k)'$), hence A has a greatest node, say b.

Since $\mathfrak{T} \models (\Sigma_\downarrow^k)'$ then $(\mathfrak{T}; a) \models \{\sigma_a^n : \sigma \in \Sigma_\downarrow^{k,n}\}$, hence, by Corollary 4.3, $A^{\mathfrak{T}}\,[\bar{\mathbf{e}}] \models \Sigma_\downarrow^{k,n}$. Therefore, there exists a $(k+p)$-coloured linear order $\mathfrak{B} = (B; <_{\mathfrak{B}}, \bar{\mathbf{c}}, \bar{\mathbf{e}})$ such that $\mathfrak{B} \equiv_n A^{\mathfrak{T}}\,[\bar{\mathbf{e}}]$ and for which the reduct $\mathfrak{B}' = (B; <_{\mathfrak{B}}, \bar{\mathbf{c}})$ is a coloured principal stem in some tree from $\mathcal{S}$.

Since $\mathfrak{T}$ is discrete, there are no lower side-forests in $\mathfrak{T}$. Now, for each $t \in A$, let $\{\mathfrak{F}_{t,i}\}_{i \in I_t}$ be the set of components in the upper side-forest $\mathfrak{F}_u(t \backslash A)$ of t with respect to A in the tree $\mathfrak{T}$. Since $\mathfrak{T}$ is discrete, each component $\mathfrak{F}_{t,i}$ has a root, say $r_{t,i}$. Then, since $\mathfrak{T}^{\geqslant r_{t,i}} \models \Sigma_\uparrow^k$, there exists a k-coloured tree $\mathfrak{G}_{t,i}$ such that $\mathfrak{G}_{t,i} \equiv_n \mathfrak{F}_{t,i}$ and for which $\mathfrak{G}_{t,i}$ is a principal subtree of some tree from $\mathcal{S}$. It follows that, for each node $t \in A$ there exists a forest $\mathfrak{G}_t$ such that $\mathfrak{F}_u(t \backslash A) \equiv_n \mathfrak{G}_t$ and for which each component of $\mathfrak{G}_t$ is a principal subtree of some tree from $\mathcal{S}$.

For each $s \in B$, let $t\,(s) \in A$ be a node such that s and $t\,(s)$ have the same colour from amongst the colours in $\bar{\mathbf{e}}$ in the linear orders $\mathfrak{B}$ and $A^{\mathfrak{T}}\,[\bar{\mathbf{e}}]$ respectively. (Such $t\,(s)$ exists since $\mathfrak{B} \equiv_n A^{\mathfrak{T}}\,[\bar{\mathbf{e}}]$). Define $\mathfrak{S}_n$ to be the k-coloured tree that is obtained from $\mathfrak{B}'$ by adding, to each element $s \in B$, the forest $\mathfrak{G}_{t(s)}$ as an upper side-forest of s with respect to B. It follows that for each $s \in B$, s has the same colour from amongst the colours in $\bar{\mathbf{e}}$ in both the linear orders $B^{\mathfrak{S}_n}[\bar{\mathbf{e}}]$ and $\mathfrak{B}$, hence $B^{\mathfrak{S}_n}[\bar{\mathbf{e}}] \cong \mathfrak{B}$. Since also $\mathfrak{B} \equiv_n A^{\mathfrak{T}}\,[\bar{\mathbf{e}}]$ then $B^{\mathfrak{S}_n}[\bar{\mathbf{e}}] \equiv_n A^{\mathfrak{T}}\,[\bar{\mathbf{e}}]$.

Next, we note that both $\mathfrak{S}_n$ and $\mathfrak{T}$ are branching complete. That $\mathfrak{S}_n$ is branching complete follows from the fact that every component in each of the upper side-forests $\mathfrak{F}_u(s \backslash B)$ ($\cong \mathfrak{G}_{t(s)}$) in $\mathfrak{S}_n$, being a principal subtree of a tree from $\mathcal{S}$, is branching complete. $\mathfrak{T}$ is branching complete since it is a model of the sentence bc. Now since $B^{\mathfrak{S}_n}[\bar{\mathbf{e}}] \equiv_n A^{\mathfrak{T}}\,[\bar{\mathbf{e}}]$ then, by Theorem 4.2, $\mathfrak{S}_n \equiv_n \mathfrak{T}$.

Finally, to see that $\mathfrak{S}_n$ is a $\mathcal{C}$-tree, observe that each path in $\mathfrak{S}_n$ can be written as a sum $(\mathfrak{B}^-)^{\leqslant s} + \mathsf{D}$ for some node $s \in B$ and where D is a path in the forest $\mathfrak{F}_u(s \backslash B) \cong \mathfrak{G}_{s(t)}$ in $\mathfrak{S}_n$. Since $\mathfrak{B}^-$ is a principal stem in a $\mathcal{C}$-tree and the linear orders in $\mathcal{C}$ are discrete, it follows that $(\mathfrak{B}^-)^{\leqslant s} \cong (\mathfrak{C}_1)^{<c_1}$ for some linear order $\mathfrak{C}_1 \in \mathcal{C}$ of which $\mathfrak{B}^-$ is an initial suborder and some $c_1 \in C_1$. Further, since each component of $\mathfrak{G}_{s(t)}$ is a principal subtree of some tree from $\mathcal{S}$, it follows that $\mathsf{D} \cong (\mathfrak{C}_2)^{\geqslant c_2}$ for some linear

---

[5]In fact, only the property of left-connectedness is needed from the axiom Tree; the properties of irreflexivity, transitivity, and left-linearity are all encoded in the axioms from $(\Sigma_\uparrow^k)'$ and $(\Sigma_\downarrow^k)'$.

order $\mathfrak{C}_2 \in \mathcal{C}$ of which D is a terminal suborder and some $c_2 \in C_2$. Hence, by the fusion closure property of $\mathcal{C}$, we obtain that $(\mathfrak{B}^-)^{\leqslant s} + D \cong (\mathfrak{C}_1)^{<c_1} + (\mathfrak{C}_2)^{\geqslant c_2} \in \mathcal{C}$, as required. ⊣

4.3. Applications of the results so far. Theorem 4.4 can be used to axiomatise the first-order theories of classes of trees of the type $\mathcal{S} = \mathcal{T}_k(\mathcal{C})$, where the requirements on $\mathcal{C}$ are firstly that it must satisfy the properties (C1)–(C4), and secondly, that the axiomatisations $(\Sigma_\uparrow^k)'$ and $(\Sigma_\downarrow^k)'$ are known. There are three classes $\mathcal{C}$ that fulfill these requirements, namely the sets $\{\zeta\}$, $\{\omega^\star\}$, and $\{\zeta, \omega^\star\}$.

Theorem 4.4 will hence be used in Section 5.2 to axiomatise the first-order theory of the class of k-coloured ζ-trees. The two simpler axiomatisations $(\Sigma_\uparrow^k)'$ and $(\Sigma_\uparrow^k)'$ that are needed for that purpose are those of the first-order theories of the class of k-coloured ω-trees (to be given in Section 5.1) and of the class of $(k + f(L_k', n))$-coloured expansions of the linear order $\omega^\star$ (given in Section 3.3).

In the case of the sets $\mathcal{C} = \{\omega^\star\}$ and $\mathcal{C} = \{\zeta, \omega^\star\}$, an axiomatisation of the first-order theory of the class of $(k + f(L_k', n))$-coloured expansions of $\omega^\star$ is also needed for $(\Sigma_\downarrow^k)'$. For $(\Sigma_\uparrow^k)'$, one uses in the case $\mathcal{C} = \{\omega^\star\}$ an axiomatisation of the first-order theory of the class of k-coloured trees of which all paths are finite (which can be deduced from [15]), and in the case $\mathcal{C} = \{\zeta, \omega^\star\}$ one uses an axiomatisation of the first-order theory of the class of k-coloured well-founded trees of which all paths have height at most ω (a trivial exercise). Axiomatisations for the cases $\mathcal{C} = \{\omega^\star\}$ and $\mathcal{C} = \{\zeta, \omega^\star\}$ will not be presented here however.

§5. First-order theories of coloured trees.

5.1. The first-order theory of coloured ω-trees. The following result will be needed to axiomatise the first-order theory of the class of k-coloured ω-trees.

PROPOSITION 5.1 ([2, Theorem 5.1]). *For every k-coloured model $\mathfrak{T}$ of $\{\mathsf{Tree}\} \cup \mathsf{BWF}_k$ and each $n \in \mathbb{N}$ there exists a well-founded tree $\mathfrak{S}_n$ such that $\mathfrak{S}_n \equiv_n \mathfrak{T}$.*

It now follows that the theory

$$T\omega_k := \{\mathsf{Tree}, \mathsf{Le}, \neg\mathsf{Gr}, \mathsf{BD}, \mathsf{FD}\} \cup \mathsf{BWF}_k$$

axiomatises the class of k-coloured ω-trees:

COROLLARY 5.2. *Let $\mathfrak{T}$ be any k-coloured model of $T\omega_k$, for some $k \in \mathbb{N}$. Then, for each $n \in \mathbb{N}$ there exists a k-coloured ω-tree $\mathfrak{S}_n$ such that $\mathfrak{S}_n \equiv_n \mathfrak{T}$.*

PROOF. Without loss of generality we can assume that n is at least as large as the quantifier ranks of all the sentences Tree, Le, $\neg$Gr, BD, and FD. Since $\mathfrak{T} \models \{\mathsf{Tree}\} \cup \mathsf{BWF}_k$ then, by Proposition 5.1, there exists a well-founded tree $\mathfrak{S}_n$ such that $\mathfrak{S}_n \equiv_n \mathfrak{T}$. The axioms Le, $\neg$Gr, BD, and FD ensure that each path in $\mathfrak{S}_n$ will have the properties of having a least element but no greatest element and of being discrete. Then, each path in $\mathfrak{S}_n$ will be isomorphic to ω, as required. ⊣

5.2. The first-order theory of coloured ζ-trees. Using Theorem 4.4, we will now show that the theory

$$T\zeta_k := \{\mathsf{Tree}, \neg\mathsf{Le}, \neg\mathsf{Gr}, \mathsf{BD}, \mathsf{FD}\} \cup \mathsf{BBWF}_k \cup \mathsf{BFWF}_k$$

axiomatises the first-order theory of the class of k-coloured ζ-trees.

THEOREM 5.3. *Let $\mathfrak{T}$ be a k-coloured model of $T\zeta_k$, for some $k \in \mathbb{N}$. Then, for each $n \in \mathbb{N}$ there exists a k-coloured ζ-tree $\mathfrak{S}_n$ such that $\mathfrak{S}_n \equiv_n \mathfrak{T}$.*

PROOF. Let $\mathfrak{T}$ be a k-coloured model of $T\zeta_k$. The result will be proved by showing that $\mathfrak{T} \models \mathsf{Ax}\,(\mathcal{T}_k\,(\zeta))$ and employing Theorem 4.4. It is straightforward to check that the class $\{\zeta\}$ satisfies properties (C1)–(C4). Observe that the class of principal subtrees of trees in $\mathcal{T}_k\,(\zeta)$ coincides with the class of k-coloured ω-trees, hence we take $\Sigma_\uparrow^k := T\omega_k$. For $n \in \mathbb{N}$ and $p = f\,(L_k',n)$, each linear order of the form $(S^{<a})^{\mathfrak{S}}\,[\bar{\mathbf{e}}_p]$, where $\mathfrak{S} \in \mathcal{T}_k\,(\zeta)$ and $a \in S$, is a $(k+p)$-coloured expansion of $\omega^\star$, from which it follows that we can take $\Sigma_\downarrow^{k,n} := C\omega_{k+p}^\star$ (but, see Remark 5.4 after this proof). We define $(\Sigma_\uparrow^k)'$ and $(\Sigma_\downarrow^k)'$ as in Section 4.2.

By assumption, $\mathfrak{T}$ is a model of Tree. $\mathfrak{T}$ is a model of bc since the property of branching completeness follows from the property of definable bounded backward well-foundedness. Therefore, to show that $\mathfrak{T} \models \mathsf{Ax}\,(\mathcal{T}_k\,(\zeta))$, it remains to show that $\mathfrak{T} \models (\Sigma_\uparrow^k)'$ and $\mathfrak{T} \models (\Sigma_\downarrow^k)'$.

$\mathfrak{T} \models (\Sigma_\uparrow^k)'$: Let $a \in T$. We first show that $\mathfrak{T}^{\geq a} \models T\omega_k$. It is straightforward to check that $\mathfrak{T}^{\geq a} \models \{\mathsf{Tree}, \mathsf{Le}, \neg\mathsf{Gr}, \mathsf{BD}, \mathsf{FD}\}$. To see that $\mathfrak{T}^{\geq a} \models \mathsf{BWF}_k$, let $\bar{b}$ be a tuple of nodes in $T^{\geq a}$ and suppose that the formula $\varphi(x, \bar{b})$ defines a non-empty set B in the tree $(\mathfrak{T}^{\geq a}; \bar{b})$. Then, $\varphi^{\geq a}(x, \bar{b})$ defines B in the tree $(\mathfrak{T}; a, \bar{b})$ and B is bounded below by the node a. Hence, since $\mathfrak{T}$ is a model of BBWF_k, B must contain a least element. It follows that $\mathfrak{T}^{\geq a} \models \mathsf{BWF}_k$, hence also $\mathfrak{T}^{\geq a} \models T\omega_k$. Then $(\mathfrak{T}; a) \models \{\sigma^{\geq a} : \sigma \in T\omega_k\}$, and since a was arbitrary and $\Sigma_\uparrow^k = T\omega_k$, we get $\mathfrak{T} \models (\Sigma_\uparrow^k)'$.

$\mathfrak{T} \models (\Sigma_\downarrow^k)'$: Again, let $a \in T$ and let $\mathsf{A} = (T^{<a}; <_{\mathfrak{T}})$, $n \in \mathbb{N}$, and $p = f\,(L_k', n)$. We first show that $\mathsf{A}^{\mathfrak{T}}[\bar{\mathbf{e}}_p] \models C\omega_{k+p}^\star$. Clearly, $\mathsf{A}^{\mathfrak{T}}[\bar{\mathbf{e}}_p] \models \{\mathsf{LO}, \neg\mathsf{Le}, \mathsf{FD}\}$. As for $\mathsf{A}^{\mathfrak{T}}[\bar{\mathbf{e}}_p] \models \mathsf{FWF}_k$, let $\bar{b}$ this time be a tuple of nodes in $T^{<a}$ and suppose that the formula $\varphi(x, \bar{b})$ defines a non-empty set B in the tree $(\mathsf{A}^{\mathfrak{T}}[\bar{\mathbf{e}}_p]; \bar{b})$. By Corollary 4.3, $\varphi_a^n(x, \bar{b})$ defines B in $(\mathfrak{T}; a, \bar{b})$ and B is bounded above by a, hence, using the fact that $\mathfrak{T}$ is a model of BFWF_k, B must contain a greatest element. It follows that $\mathsf{A}^{\mathfrak{T}}[\bar{\mathbf{e}}_p] \models \mathsf{FWF}_k$, hence also $\mathsf{A}^{\mathfrak{T}}[\bar{\mathbf{e}}_p] \models C\omega_{k+p}^\star$. Now, by Corollary 4.3, we have that $(\mathfrak{T}; a) \models \{\sigma_a^n : \sigma \in C\omega_{k+p}^\star\}$. Since a and n were arbitrary and $\Sigma_\downarrow^{k,n} = C\omega_{k+p}^\star$, it follows that $\mathfrak{T} \models (\Sigma_\downarrow^k)'$. $\dashv$

REMARK 5.4. In the above proof, for $n \in \mathbb{N}$ and $p = f\,(L_k', n)$, the class of linear orders of the form $(S^{<a})^{\mathfrak{S}}[\bar{\mathbf{e}}_p]$, where $\mathfrak{S} \in \mathcal{T}_k\,(\zeta)$ and $a \in S$, does not, strictly speaking, coincide with the class of $(k+p)$-coloured expansions of $\omega^\star$, because there are extended colours from $\bar{\mathbf{e}}_p$ that aren't realised in any linear orders in the first class (not all characteristic sentences $\tau_{L_k', n, i}$ are realised in side-forests), while every colour in $\bar{\mathbf{e}}_p$ is realised in some $(k+p)$-coloured expansion of $\omega^\star$. We can nonetheless take $\Sigma_\downarrow^{k,n}$ to be the theory $C\omega_{k+p}^\star$ since, for each model $\mathfrak{A}$ of $C\omega_{k+p}^\star$, every $(k+p)$-coloured expansion of $\omega^\star$ that is m-equivalent to $\mathfrak{A}$ realises exactly the same colours as $\mathfrak{A}$.

5.3. The first-order theory of coloured η-trees.

We now define the theory

$$T\eta := \{\mathsf{Tree}, \neg\mathsf{Le}, \neg\mathsf{Gr}, \mathsf{De}\}.$$

The first-order theory of the class of k-coloured η-trees is axiomatised by $T\eta$, as seen from the next theorem.

THEOREM 5.5. *For each k-coloured model $\mathfrak{T}$ of $T\eta$, there exists a k-coloured η-tree $\mathfrak{S}$ such that $\mathfrak{S} \equiv \mathfrak{T}$.*

PROOF. Given a k-coloured model $\mathfrak{T}$ of $T\eta$, by the Downward Löwenheim–Skolem Theorem there exists a countable k-coloured tree $\mathfrak{S}$ such that $\mathfrak{S} \equiv \mathfrak{T}$. Hence each path A in the tree $\mathfrak{S}$ will be a countable dense linear order without endpoints, so A $\cong \eta$, as required. ⊣

5.4. The first-order theory of coloured λ-trees. Here we will present an axiomatisation of the first-order theory of the class of k-coloured λ-trees. We will make use of Fact 3.12, along with the method of partitioning into side-forests from Section 4, to prove that our axiomatisation is complete. In essence, the axiomatisation will express the fact that the linear orders $A^{\mathfrak{T}}[\bar{\mathbf{e}}]$ obtained from a coloured λ-tree by partitioning it into side-forests are themselves coloured expansions of λ. The proof of completeness of the axiomatisation that we give here follows a similar idea as in the proof of completeness of the axiomatisation of the class of well-founded trees, given in [2], but with several essential differences.

Recall the definition of the formula φ_z^n that was given before Corollary 4.3. By applying it to each $\sigma \in C\lambda_{k+f(L_k', n)}$, for any $k \in \mathbb{N}$ we define

$$(C\lambda_k)' := \{\forall z\, (\sigma_z^n) : n \in \mathbb{N},\ \sigma \in C\lambda_{k+f(L_k', n)}\}.$$

We will show (Theorem 5.9) that the first-order theory of the class of k-coloured λ-trees is axiomatised by the set of sentences

$$T\lambda_k := \{\mathsf{Tree}, \neg\mathsf{Gr}, \mathsf{bc}\} \cup (C\lambda_k)'.$$

First, note that every k-coloured λ-tree, being branching complete, is a model of $T\lambda_k$. Next, a few technical lemmas and propositions are needed.

LEMMA 5.6. *Let $\mathfrak{T}$ be a k-coloured model of $T\lambda_k$. Then, for each $a \in T$, each component in $\mathfrak{T}^{>a}$ is a model of $T\lambda_k$, too.*

PROOF. Let $\mathfrak{F}$ be a component in $\mathfrak{T}^{>a}$. It is immediate that $\mathfrak{F}$ satisfies the sentences Tree, $\neg$Gr, and bc. To see that $\mathfrak{F}$ satisfies $(C\lambda_k)'$, let $n \in \mathbb{N}$, $p = f(L_k', n)$, and $b \in F$. It then suffices to show that $(\mathfrak{F}; b) \models \{\sigma_b^n : \sigma \in C\lambda_{k+p}\}$. Clearly $(\mathfrak{F}; b) \models (\mathsf{LO})_b^n$. It follows from the fact that $\mathfrak{T} \models (C\lambda_k)'$ that $\mathfrak{T}$ is dense, hence $(\mathfrak{F}; b) \models \{(\neg\mathsf{Le})_b^n, (\neg\mathsf{Gr})_b^n, (\mathsf{De})_b^n\}$.

Since $\mathfrak{T} \models (C\lambda_k)'$ then in particular $(\mathfrak{T}; b) \models \{\sigma_b^n : \sigma \in \mathsf{COM}_{k+p}\}$ and $(\mathfrak{T}; b) \models \{\sigma_b^n : \sigma \in \mathsf{I}_{k+p}\}$. Hence, by Corollary 4.3, we obtain $(T^{<b})^{\mathfrak{T}}[\bar{\mathbf{e}}_p] \models \mathsf{COM}_{k+p}$ and $(T^{<b})^{\mathfrak{T}}[\bar{\mathbf{e}}_p] \models \mathsf{I}_{k+p}$.

To show that $(\mathfrak{F}; b) \models \{\sigma_b^n : \sigma \in \mathsf{COM}_{k+p}\}$ it suffices, again by Corollary 4.3, to show that $(F^{<b})^{\mathfrak{F}}[\bar{\mathbf{e}}_p] \models \mathsf{COM}_{k+p}$. Let $\bar{d}$ be a tuple of nodes in $F^{<b}$ and let $\varphi(x, \bar{d})$ define the non-empty set $B \subseteq F^{<b}$ in the linear order $(F^{<b})^{\mathfrak{F}}[\bar{\mathbf{e}}_p]$. Suppose that B has an upper bound $u \in F^{<b}$. By properties of relativised formulas, $\varphi^{>a}(x, \bar{d})$ defines B

in $\left((T^{<b})^{\mathfrak{T}}[\bar{\mathbf{e}}_p]; a\right)$. Since $(T^{<b})^{\mathfrak{T}}[\bar{\mathbf{e}}_p] \models \mathsf{COM}_{k+p}$, we obtain that B has a least upper bound in $T^{<b}$, hence also in $F^{<b}$, as required.

Lastly, it will be shown that $(\mathfrak{F}; b) \models \{\sigma_b^n : \sigma \in \mathsf{I}_{k+p}\}$ by showing, instead, that $(F^{<b})^{\mathfrak{F}}[\bar{\mathbf{e}}_p] \models \mathsf{I}_{k+p}$. To this end, let $\bar{d}$ be a tuple of nodes in $F^{<b}$ and let $\varphi(x, y, \bar{d})$ define in the linear order $((F^{<b})^{\mathfrak{F}}[\bar{\mathbf{e}}_p]; \bar{d})$ a transitive binary relation on $F^{<b}$ that induces a densely ordered condensation $\mathfrak{P} = (P; <_{\mathfrak{P}})$ of $F^{<b}$. Let

$$\psi(x, y, a, \bar{d}) := (x \leqslant a \wedge x = y) \vee (a < x \wedge a < y \wedge \varphi^{>a}(x, y, \bar{d})).$$

The formula $\psi(x, y, a, \bar{d})$ defines in $\left(T^{<b}\right)^{\mathfrak{T}}[\bar{\mathbf{e}}_p]$ a transitive binary relation on $T^{<b}$ that induces a densely ordered condensation $\mathfrak{Q} = (Q; <_{\mathfrak{Q}})$ of $T^{<b}$ and $\mathfrak{P}$ is a tail of $\mathfrak{Q}$. Since $\left(T^{<b}\right)^{\mathfrak{T}}[\bar{\mathbf{e}}_p] \models \mathsf{I}_{k+p}$ then it follows that $\mathfrak{Q}$ must contain a set S of singletons that is dense in $\mathfrak{Q}$, hence $S{\upharpoonright}_P$ is a set of singletons in P that is dense in $\mathfrak{P}$, as required.   ⊣

Given a k-coloured forest $\mathfrak{F} := (F; <, \bar{\mathbf{c}})$, let F_r denote the (possibly empty) set of roots of the components of $\mathfrak{F}$ and let $F' := F \backslash F_r$. Consider the following property:

$\mathsf{Comp}_\lambda(\mathfrak{F})$: For each component G of $\mathfrak{F}^{F'}$, $\mathfrak{F}^G$ is a model of $T\lambda_k$.

In other words, property $\mathsf{Comp}_\lambda(\mathfrak{F})$ states that for each component X in $\mathfrak{F}$, if $\mathfrak{F}^X$ has no root then $\mathfrak{S}^X$ is a model of $T\lambda_k$, while if $\mathfrak{F}^X$ has a root r then $\mathfrak{F}^Y$ is a model of $T\lambda_k$ for each component Y in $\left(\mathfrak{F}^X\right)^{>r}$.

LEMMA 5.7. *Let $\mathfrak{T} := (T; <_{\mathfrak{T}}, \bar{\mathbf{c}})$ be a k-coloured model of $T\lambda_k$. Then, for each $a \in T$ and $n \in \mathbb{N}$ there exists a k-coloured tree $\mathfrak{S} := (S; <_{\mathfrak{S}}, \bar{\mathbf{c}})$ and a node $b \in S$ such that*

(i) $(\mathfrak{T}; a) \equiv_n (\mathfrak{S}; b)$,
(ii) $(\mathfrak{S}^{<b})^- \cong \lambda$, *and*
(iii) $\mathsf{Comp}_\lambda\left(\mathfrak{S}^J\right)$ *holds, where $J := S \backslash S^{\leqslant b}$.*

PROOF. Let $\mathsf{A} := T^{\leqslant a}$, $n' = n + 1$, $p = f\left(L'_k, n'\right)$, and $\bar{\mathbf{e}} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p)$. Let $\mathfrak{F}_1, \mathfrak{F}_2, \ldots, \mathfrak{F}_q$ be all, up to n'-equivalence, side-forests of the form $\mathfrak{F}(t\backslash\mathsf{A})$ with $t \in \mathsf{A}$. Without loss of generality, we can assume that $\mathfrak{F}_i \models \tau_{L'_k, n', i}$ for each i. Each $\mathfrak{F}_i$ has the form $\mathfrak{F}_i := \left(F_i; <_{\mathfrak{F}_i}, \bar{\mathbf{c}}, d_i\right)$ where $d_i \in \mathsf{A}$.

Since $(\mathfrak{T}; a) \models \{\sigma_a^{n'} : \sigma \in C\lambda_{k+p}\}$ then by Corollary 4.3, $(T^{<a})^{\mathfrak{T}}[\bar{\mathbf{e}}] \models C\lambda_{k+p}$, hence, by Fact 3.12, there exists a $(k+p)$-coloured linear order $\mathfrak{B}_1$ such that $(T^{<a})^{\mathfrak{T}}[\bar{\mathbf{e}}] \equiv_{n'} \mathfrak{B}_1$ and $\mathfrak{B}_1^- \cong \lambda$. Let $\mathfrak{B}_2 := \left(\mathsf{A}^{\mathfrak{T}}[\bar{\mathbf{e}}]\right)^{\{a\}}$ be the restriction of $\mathsf{A}^{\mathfrak{T}}[\bar{\mathbf{e}}]$ to the set $\{a\}$ and define the $(k+p)$-coloured linear order $\mathfrak{B} := \mathfrak{B}_1 + \mathfrak{B}_2 = (B; <_{\mathfrak{B}}, \bar{\mathbf{c}}, \bar{\mathbf{e}})$. By Lemma 2.1, $\mathfrak{B} \equiv_{n'} \mathsf{A}^{\mathfrak{T}}[\bar{\mathbf{e}}]$.

For each $x \in B$, let $\chi(x)$ denote the value of i for which $\mathfrak{B} \models \mathbf{e}_i(x)$. We define

$$S := \bigcup_{x \in B} \left(F_{\chi(x)} \times \{x\}\right),$$

and we also define the binary relation $<_{\mathfrak{S}}$ on S by specifying, for all pairs $(u, x), (v, y) \in S$, that

$$(u, x) <_{\mathfrak{S}} (v, y) \Longleftrightarrow [x <_{\mathfrak{B}} y \text{ and } u = d_{\chi(x)}) \text{ or } (x = y \text{ and } u <_{\mathfrak{F}_{\chi(x)}} v)].$$

Now, we define $\mathfrak{S} = (S; <_{\mathfrak{S}}, \bar{\mathbf{c}})$ by specifying that, for each $s = (u, x)$ in S, $\mathfrak{S} \models \mathbf{c}_i(s)$ if and only if $\mathfrak{F}_{\chi(x)} \models \mathbf{c}_i(u)$. Let $D := \{(d_{\chi(x)}, x) : x \in B\}$ and $\mathsf{D} := (D; <_{\mathfrak{S}})$. $\mathfrak{S}$ is simply the tree that is obtained from $\mathfrak{B}$ by replacing each $x \in B$ with the side-forest $\mathfrak{F}_{\chi(x)}$, and D is the path in $\mathfrak{S}$ that consists of all the d_i's in these side-forests.

We are going to apply Theorem 4.2. For that, we will need to know that $\mathfrak{T}$ and $\mathfrak{S}$ are branching complete. $\mathfrak{T}$ is branching complete since it satisfies bc. It follows that every lower side-forest in $\mathfrak{T}$—in particular, each of the lower side-forests in $\mathfrak{F}_1, \mathfrak{F}_2, \ldots, \mathfrak{F}_q$—must contain a root. For $\mathfrak{S}$, let $r, w \in S$ be incomparable and let S_{rw} be defined as in (2) at the beginning of Section 4.1. We consider the following two cases, each with three subcases, for the location of r and w:

Case 1: r and w belong to the same set $F_{\chi(x)} \times \{x\}$. We consider three sub-cases:

1.1 If r and w belong to the same component from $F_{\chi(x)} \times \{x\}$ then S_{rw} contains a least element, by the fact that $\mathfrak{F}_{\chi(x)}$ is branching complete.

1.2 If r and w belong to different components from $F_{\chi(x)} \times \{x\}$, but with r lying in the upper side-forest from $F_{\chi(x)} \times \{x\}$, then $(d_{\chi(x)}, x)$ is the least element of S_{rw}.

1.3 If r and w belong to different components from $F_{\chi(x)} \times \{x\}$, but with r lying in the lower side-forest from $F_{\chi(x)} \times \{x\}$, then the root of that lower side-forest that contains r is the least element of S_{rw}.

Case 2: r belongs to the set $F_{\chi(x)} \times \{x\}$ and w belongs to a different set $F_{\chi(y)} \times \{y\}$. Then either $x <_{\mathfrak{B}} y$ or $y <_{\mathfrak{B}} x$.

2.1 If $x <_{\mathfrak{B}} y$ and r lies in the upper side-forest from $F_{\chi(x)} \times \{x\}$ then $(d_{\chi(x)}, x)$ is the least element of S_{rw}.

2.2 If $x <_{\mathfrak{B}} y$ and r lies in the lower side-forest from $F_{\chi(x)} \times \{x\}$ then the root of that lower side-forest that contains r is the least element of S_{rw}.

2.3 If $y <_{\mathfrak{B}} x$ then $(d_{\chi(y)}, y)$ is the least element of S_{rw}.

Hence, $\mathfrak{S}$ is branching complete. Since $\mathsf{D}^{\mathfrak{S}}[\bar{\mathbf{e}}] \cong \mathfrak{B} \equiv_{n'} \mathsf{A}^{\mathfrak{T}}[\bar{\mathbf{e}}]$ it follows, using Theorem 4.2, that $(\mathfrak{T}; \mathsf{A}) \equiv_{n'} (\mathfrak{S}; \mathsf{D})$. Letting b be the greatest element of D, it then follows that $(\mathfrak{T}; a) \equiv_n (\mathfrak{S}; b)$ and clearly $(\mathfrak{S}^{<b})^- \cong \mathfrak{B}_1^- \cong \lambda$. That the property $\mathsf{Comp}_\lambda(\mathfrak{S}^J)$ holds follows from the fact that, for each component F in $\mathfrak{S}^J$, the tree $\mathfrak{S}^F$ is either a copy of some lower side-forest of $\mathfrak{T}$ (which will necessarily have a root), or a copy of a component in some upper side-forest of $\mathfrak{T}$ (which will not have a root since $\mathfrak{T}$ satisfies the density property that is encoded in $(C\lambda_k)'$); the claim then follows by Lemma 5.6. ⊣

LEMMA 5.8. *Let* $\mathfrak{T} := (T; <_{\mathfrak{T}}, \bar{\mathbf{c}})$ *be a* k-*coloured model of the theory* $T\lambda_k$, *let* $\{a_1, a_2, \ldots, a_m\}$ *be an antichain in* $\mathfrak{T}$, *and let* $\bar{a} := (a_1, a_2, \ldots, a_m)$. *Then, for each* $n \in \mathbb{N}$ *there exists a* k-*coloured tree* $\mathfrak{S} := (S; <_{\mathfrak{S}}, \bar{\mathbf{c}})$ *and a tuple of nodes* $\bar{b} := (b_1, b_2, \ldots, b_m) \in S^m$ *such that*

(i) $(\mathfrak{T}; \bar{a}) \equiv_n (\mathfrak{S}; \bar{b})$,

(ii) $(\mathfrak{S}^{<b_i})^- \cong \lambda$ *for each* i, *and*

(iii) $\mathsf{Comp}_\lambda(\mathfrak{S}^J)$ *holds, where* $J := S \setminus \bigcup_{i=1}^m S^{\leqslant b_i}$.

PROOF. The result will be proved by induction on m. For $m = 1$, the result is precisely Lemma 5.7. Now let $m \geqslant 2$ be an integer such that the claim holds for each positive integer less than m and let $C := \{a_1, a_2, \ldots, a_m\}$ be an antichain in $\mathfrak{T}$. Consider the stem $H := \bigcap_{i=1}^m T^{<a_i}$ and let $H' := \{x \in T : x \geqslant y \text{ for each } y \in H\}$.

Now, let $\{F_i\}_{i \in I}$ be the set of all components in $\mathfrak{T}^{H'}$ and define the tree $\mathfrak{F}_i := \mathfrak{T}^{F_i}$ for each $i \in I$. Each component F_i must contain fewer than m (possibly none) elements from C, for, if $C \subseteq F_i$ then $H \cap F_i \neq \emptyset$—a contradiction. Assume, without loss of generality, that the components F_i that contain nodes from C are precisely $F_1, F_2, \ldots, F_r$. Let $C \cap F_i = \{a_{i,1}, a_{i,2} \ldots, a_{i,s_i}\}$ and $\bar{a}_i := (a_{i,1}, a_{i,2} \ldots, a_{i,s_i})$ for $1 \leqslant i \leqslant r$. Two cases will be considered, depending on whether H contains a greatest element or not.

Case 1: H contains a greatest element a. By Lemma 5.7, there exists a k-coloured tree $\mathfrak{S}_0 = (S_0; <_{\mathfrak{S}_0}, \bar{\mathbf{c}})$ and a node $b \in S_0$ such that $(\mathfrak{T}; a) \equiv_n (\mathfrak{S}_0; b)$ and $(\mathfrak{S}_0^{<b})^- \cong \lambda$ and $\mathsf{Comp}_\lambda(\mathfrak{S}_0^{J_0})$ holds for $J_0 := S_0 \backslash S_0^{\leqslant b}$.

By Lemma 5.6, each of the trees $\mathfrak{F}_i$ is a model of $T\lambda_k$. By the induction hypothesis, there exists, for each i with $1 \leqslant i \leqslant r$, a k-coloured tree $\mathfrak{G}_i = (G_i; <_{\mathfrak{G}_i}, \bar{\mathbf{c}})$ and a tuple $\bar{b}_i := (b_{i,1}, b_{i,2}, \ldots, b_{i,s_i})$ of nodes in G_i such that

(i′) $(\mathfrak{F}_i; \bar{a}_i) \equiv_n (\mathfrak{G}_i; \bar{b}_i)$,
(ii′) $(\mathfrak{G}_i^{<b_{i,j}})^- \cong \lambda$ for each j with $1 \leqslant j \leqslant s_i$, and
(iii′) $\mathsf{Comp}_\lambda(\mathfrak{G}_i^{J_i})$ holds, where $J_i := G_i \backslash \bigcup_{j=1}^{s_i} G_i^{\leqslant b_{i,j}}$.

Now, we define the tree $\mathfrak{S}$ as follows. Let $\mathfrak{G}$ be the disjoint union of all the trees $(\mathfrak{G}_i; \bar{b}_i)$ for $1 \leqslant i \leqslant r$, along with all the trees $\mathfrak{F}_i$ for $i \in I \backslash \{1, 2, \ldots, r\}$. Let $S_0' := S_0 \backslash (S_0)^{>b}$, $\mathfrak{S}_0' := \mathfrak{S}_0^{S_0'}$ be the restriction of $\mathfrak{S}_0$ to S_0', $\mathsf{B} := (S_0')^{\leqslant b}$, and $\mathfrak{S} := \mathfrak{S}_0' +_\mathsf{B} \mathfrak{G}$.

To see that $(\mathfrak{T}; \bar{a}) \equiv_n (\mathfrak{S}; \bar{b})$, let $T' := T \backslash T^{>a}$, $\mathfrak{T}' := \mathfrak{T}^{T'}$, $\mathsf{A} := T^{\leqslant a}$, and $\mathfrak{F} := \mathfrak{T}^{>a}$. Since $(\mathfrak{T}; a) \equiv_n (\mathfrak{S}_0; b)$ then $(\mathfrak{T}'; \mathsf{A}) \equiv_n (\mathfrak{S}_0'; \mathsf{B})$, and it is clear that $(\mathfrak{F}; \bar{a}_1 \bar{a}_2 \cdots \bar{a}_r) \equiv_n (\mathfrak{G}; \bar{b}_1 \bar{b}_2 \cdots \bar{b}_r)$. By Lemma 2.2, $(\mathfrak{T}' +_\mathsf{A} \mathfrak{F}; \bar{a}_1 \bar{a}_2 \cdots \bar{a}_r) \equiv_n (\mathfrak{S}_0' +_\mathsf{B} \mathfrak{G}; \bar{b}_1 \bar{b}_2 \cdots \bar{b}_r)$, from which it follows that $(\mathfrak{T}; \bar{a}) \equiv_n (\mathfrak{S}; \bar{b})$.

Next, we will show that $(\mathfrak{S}^{<b_i})^- \cong \lambda$ for each i. Suppose that $b_i = b_{j,l} \in G_j$. Since $(\mathfrak{S}^{<b})^- \cong (\mathfrak{S}_0^{<b})^- \cong \lambda$ and $(\mathfrak{S}^{(b,b_i)})^- \cong (\mathfrak{G}_j^{<b_{j,l}})^- \cong \lambda$ (where $\mathfrak{S}^{(b,b_i)}$ denotes the restriction of $\mathfrak{S}$ to the interval (b, b_i)), it follows that $(\mathfrak{S}^{<b_i})^- \cong (\mathfrak{S}^{<b})^- + 1 + (\mathfrak{S}^{(b,b_i)})^- \cong \lambda + 1 + \lambda \cong \lambda$.

Finally, the claim that $\mathsf{Comp}_\lambda(\mathfrak{S}^J)$ holds for $J = S \backslash \bigcup_{i=1}^m S^{\leqslant b_i}$ follows by the corresponding properties that hold in $\mathfrak{S}_0$ and in each of the trees $\mathfrak{G}_i$, along with using Lemma 5.6 when working in the trees $\mathfrak{F}_i$ for $i \in I \backslash \{1, 2, \ldots, r\}$.

Case 2: H does not contain a greatest element. It follows by the branching completeness of $\mathfrak{T}$ that each $\mathfrak{F}_i$ must contain a least element. Fix $i_0 \in I$ and let a be the root of $\mathfrak{F}_{i_0}$ and $\mathsf{A} := T^{<a}$. Let $T' = T \backslash \bigcup_{i \in I} F_i$ and $\mathfrak{T}' = \mathfrak{T}^{T'}$. Observe that, as A does not have a greatest element, $\mathsf{A}^{\mathfrak{T}}[\bar{\mathbf{e}}] = \mathsf{A}^{\mathfrak{T}'}[\bar{\mathbf{e}}]$.

Let $p := f(L_k', n)$ and $\bar{\mathbf{e}} := \bar{\mathbf{e}}_p$. Since $(\mathfrak{T}; a) \models \{\sigma_a^n : \sigma \in C\lambda_{n+p}\}$ then $\mathsf{A}^{\mathfrak{T}}[\bar{\mathbf{e}}] \models C\lambda_{n+p}$. Hence, by Fact 3.12, there exists an $(n+p)$-coloured linear order $\mathfrak{B} = (B; <_{\mathfrak{B}}, \bar{\mathbf{c}}, \bar{\mathbf{e}})$ such that $\mathfrak{B} \equiv_n \mathsf{A}^{\mathfrak{T}}[\bar{\mathbf{e}}]$ and $\mathfrak{B}^- \cong \lambda$. As with the construction of the tree $\mathfrak{S}$ in the proof of Lemma 5.7, we can add side-forests to the elements of B to obtain a k-coloured tree $\mathfrak{S}_0 := (S_0; <_{\mathfrak{S}_0}, \bar{\mathbf{c}})$, and a path D in $\mathfrak{S}_0$, such that $\mathsf{D}^{\mathfrak{S}_0}[\bar{\mathbf{e}}] \cong \mathfrak{B}$, from which

$$\mathsf{A}^{\mathfrak{T}'}[\bar{\mathbf{e}}] = \mathsf{A}^{\mathfrak{T}}[\bar{\mathbf{e}}] \equiv_n \mathfrak{B} \cong \mathsf{D}^{\mathfrak{S}_0}[\bar{\mathbf{e}}]. \tag{3}$$

$\mathfrak{T}'$ is branching complete, due to the fact that $\mathfrak{T}$ satisfies bc, and it can be shown that $\mathfrak{S}_0$ is branching complete, similarly to how it was shown that the tree $\mathfrak{S}$ in the proof of Lemma 5.7 was branching complete. It then follows, using (3) and Theorem 4.2, that $(\mathfrak{T}'; A) \equiv_n (\mathfrak{S}_0; D)$.

Similarly to Case 1, for each i, where $1 \leqslant i \leqslant r$, there exists a k-coloured tree $\mathfrak{G}_i$ and a tuple $\bar{b}_i := (b_{i,1}, b_{i,2}, \ldots, b_{i,s_i})$ of nodes in G_i, such that properties (i$'$) and (iii$'$) hold, in addition to property (ii$'$) replaced by:

(iv$'$) $(\mathfrak{G}_i^{<b_{i,j}})^- \cong 1 + \lambda$ for each j with $1 \leqslant j \leqslant s_i$.

Now, we define $\mathfrak{G}$ as in Case 1 and let $\mathfrak{S} := \mathfrak{S}_0 +_D \mathfrak{G}$. By similar arguments to those used in Case 1 and Lemma 5.7, it can again be seen that properties (i)–(iii) hold. ⊣

THEOREM 5.9. *Let $\mathfrak{T}$ be a k-coloured model of $T\lambda_k$. Then, for each $n \in \mathbb{N}$ there exists a k-coloured λ-tree $\mathfrak{S}$ such that $\mathfrak{T} \equiv_n \mathfrak{S}$.*

PROOF. Without loss of generality, we can assume that $n \geqslant 5$. We define the k-coloured trees $\mathfrak{T}_0, \mathfrak{T}_1, \mathfrak{T}_2, \ldots$ and sets $H_0, H_1, H_2, \ldots$ with each $H_i \subseteq T_i$, as follows. First, take $\mathfrak{T}_0 := \mathfrak{T}$ and $H_0 = \emptyset$. Now, given $\mathfrak{T}_i$ and H_i, let $T_i' := T_i \backslash H_i$ and $\mathfrak{T}_i' := \mathfrak{T}_i^{T_i'}$, and suppose that these have the following properties:

(a) H_i is a left-closed subset of T_i and each path in $\mathfrak{H}_i := \mathfrak{T}_i^{H_i}$ is isomorphic to $\lambda + 1$;
(b) $\mathfrak{T}_i$ is dense and branching complete;
(c) $\mathsf{Comp}_\lambda(\mathfrak{T}_i')$ holds.

Note that property (a) is satisfied vacuously by H_0, property (b) is satisfied by $\mathfrak{T}_0$ (with the density property being encoded in the theory $(C\lambda_k)'$ that is satisfied by $\mathfrak{T}$), and property (c) is trivially satisfied by $\mathfrak{T}_0'$. The construction of $\mathfrak{T}_{i+1}$ and H_{i+1} is described below.

Let F be any component in $\mathfrak{T}_i'$ and put $\mathfrak{F} := \mathfrak{T}_i^F$. If $i = 0$ then $\mathfrak{F} = \mathfrak{T}$ while if $i > 0$ then $\mathfrak{F}$ can be represented in one of the forms (i) and (ii) below as follows. For each path X in $\mathfrak{H}_i$ (which can be viewed as a stem in $\mathfrak{T}_i$), since X has a greatest node and $\mathfrak{T}_i$ is branching complete, the set of side-forests $\{F(x \backslash X)\}_{x \in X}$ form a partition of T_i. There must hence exist a path A in $\mathfrak{H}_i$ and a node $d \in A$ such that F is contained in the side-forest of d with respect to A. In particular, $(\mathfrak{T}_i^{<d})^- \cong \lambda$ and either $F \subseteq F_u(d \backslash A)$ or $F \subseteq F_l(d)$. Now, if $F \subseteq F_u(d \backslash A)$ then, using the fact that $\mathfrak{T}_i$ is dense, $\mathfrak{F}$ has the following form:

(i) $\mathfrak{F}$ has no root and $F \subseteq T_i^{>d}$ with $(\mathfrak{T}_i^{\leqslant d})^- \cong \lambda + 1$,

while, if $F \subseteq F_l(d)$ then, using the fact that $\mathfrak{T}_i$ is branching complete and dense, $\mathfrak{F}$ has the following form:

(ii) $\mathfrak{F}$ has a root r and $F = T_i^{\geqslant r}$ with $(\mathfrak{T}_i^{<r})^- \cong \lambda$.

In the construction that follows, we hence take $\mathfrak{F} = \mathfrak{T}$ when $i = 0$ and if $i > 0$ then we can restrict our attention to components that have the form of case (i), because in case (ii) each component F' in $\mathfrak{T}_i^{>r}$ will have the form of case (i) (but, with r fulfilling the role of d) and the subsequent construction is then performed using this F' rather than F.

Let $\{a_1, a_2, \ldots, a_p\}$ be an antichain in $\mathfrak{F}$ such that $\bigcup_{i=1}^p F^{\leqslant a_i}$ contains an n-support for $\mathfrak{F}$ (recall the definition of 'n-support' at the end of Section 2.1) and

let $\bar{a} := (a_1, a_2, \ldots, a_p)$. By property (c) and Lemma 5.8, there exists a k-coloured tree $\mathfrak{G}$ and a tuple of nodes $\bar{b} := (b_1, b_2, \ldots, b_p) \in G^p$ such that: $(\mathfrak{F}; \bar{a}) \equiv_n (\mathfrak{G}; \bar{b})$, $(\mathfrak{G}^{<b_j})^- \cong \lambda$ for each j, and $\mathsf{Comp}_\lambda(\mathfrak{G}^J)$ holds for $J := G \setminus \left(\bigcup_{j=1}^p G^{\leqslant b_j} \right)$. It then follows that

(†): for each $a \in F$ there exists $b \in \bigcup_{i=1}^p G^{\leqslant b_i}$ such that $(\mathfrak{F}; a) \equiv_{n-1} (\mathfrak{G}; b)$.

The tree $\mathfrak{T}_{i+1}$ is now obtained from $\mathfrak{T}_i$ by replacing $\mathfrak{F}$ with $\mathfrak{G}$ and by performing a similar replacement for every other component of $\mathfrak{T}_i'$, while H_{i+1} is the union over all these components $\mathfrak{G}$ of the sets $\bigcup_{j=1}^p T_{i+1}^{\leqslant b_j}$.

Note that since $\mathfrak{H}_{i+1}$ is obtained from $\mathfrak{H}_i$ by adding only finitely many new paths corresponding to each component F in $\mathfrak{T}_i'$, none of the paths in $\mathfrak{H}_{i+1}$ are emerging paths, and hence there is no possibility of producing 'unwanted' paths that are isomorphic to λ by diagonalisation.

Each path in $\mathfrak{H}_{i+1}$ will be of the form

$$(\mathfrak{T}_i^{\leqslant d})^- + (\mathfrak{S}^{\leqslant b_j})^- \cong (\lambda + 1) + (\lambda + 1) \cong \lambda + 1$$

(or simply of the form $(\mathfrak{S}^{\leqslant b_j})^- \cong \lambda + 1$ when $i = 0$) so property (a) holds for H_{i+1}. To see that property (b) holds for $\mathfrak{T}_{i+1}$, first note that each of the trees $\mathfrak{G}$ that was used to replace the tree $\mathfrak{F}$ in the above construction is itself dense and branching complete, since $\mathfrak{F} \equiv_n \mathfrak{G}$ and the properties of denseness and branching completeness can be expressed by sentences of quantifier rank at most 5. We can now use an argument similar to the one used in Lemma 5.6 to prove that the new tree that was constructed there was branching complete. The claim that property (c) holds for $\mathfrak{T}_{i+1}'$ follows from the fact that each of the trees $\mathfrak{G}$ that was used to replace the tree $\mathfrak{F}$ in the above construction satisfies $\mathsf{Comp}_\lambda(\mathfrak{G}^J)$ for $J = G \setminus \left(\bigcup_{j=1}^p G^{\leqslant b_j} \right)$.

Observe, moreover, that $\mathfrak{T}_{i+1}$ and H_{i+1} also have the following properties:

(d) $\mathfrak{H}_{i+1}$ is an end-extension of $\mathfrak{H}_i$,
(e) $(\mathfrak{T}_i, t)_{t \in H_i} \equiv_n (\mathfrak{T}_{i+1}, t)_{t \in H_i}$,
(f) for each $a \in T_i'$, there exists $b \in H_{i+1}$ such that

$$(\mathfrak{T}_i; t, a)_{t \in H_i} \equiv_{n-1} (\mathfrak{T}_{i+1}; t, b)_{t \in H_i}.$$

Property (e) follows from the fact that in the construction of $\mathfrak{T}_{i+1}$, $\mathfrak{H}_i$ was left intact and the components in $\mathfrak{T}_i'$ were all replaced by n-equivalent trees. Property (f) follows from (†).

Finally, let $\mathfrak{S} := \bigcup_{i=0}^\infty \mathfrak{H}_i$.
(It follows from property (d) that this union is defined.)

Note that even though every path in each $\mathfrak{H}_i$ is singular, the (infinite) union $\mathfrak{S}$ of all these $\mathfrak{H}_i$'s may contain emerging paths.

$\mathfrak{S}$ is a λ-tree since, clearly, the monochromatic reduct of each of its paths is isomorphic to $(\lambda + 1) \cdot \omega \cong \lambda$. To show that $\mathfrak{T} \equiv_n \mathfrak{S}$, we will describe a winning strategy for Player II for the game $\mathrm{EF}_n(\mathfrak{T}, \mathfrak{S})$. Let $\bar{t} := (t_1, t_2, \ldots, t_m) \in T^m$ and $\bar{s} := (s_1, s_2, \ldots, s_m) \in S^m$ represent the nodes that have been chosen by the two players after m moves (with $\bar{t} = \varepsilon = \bar{s}$ when $m = 0$), and such that the following property holds:

$[\star]_{i,\bar{t},\bar{s}} \colon \bar{s} \in (H_i)^m$ and $(\mathfrak{T};\bar{t}) \equiv_{n-m} (\mathfrak{T}_i;\bar{s})$.

Observe that, by property (e), $[\star]_{i,\bar{t},\bar{s}}$ holds trivially when $m = 0$, and if $i < j$ and $[\star]_{i,\bar{t},\bar{s}}$ holds then $[\star]_{j,\bar{t},\bar{s}}$ holds, too. We will consider two cases for the choice of Spoiler in the $(m+1)$-th round of the game:

- Suppose that Spoiler chooses the node $t_{m+1} \in T$. By $[\star]_{i,\bar{t},\bar{s}}$ there exists $u_{m+1} \in T_i$ such that $(\mathfrak{T};\bar{t}t_{m+1}) \equiv_{n-m-1} (\mathfrak{T}_i;\bar{s}u)$. If $u \in H_i$, then take $s_{m+1} := u$ as the response of Duplicator for round $m+1$ of the game and observe that $[\star]_{i,\bar{t}t_{m+1},\bar{s}s_{m+1}}$ holds. Otherwise, by property (f) there exists $v \in H_{i+1}$ such that $(\mathfrak{T}_i;\bar{s}u) \equiv_{n-1} (\mathfrak{T}_{i+1};\bar{s}v)$, hence also $(\mathfrak{T};\bar{t}t_{m+1}) \equiv_{n-m-1} (\mathfrak{T}_{i+1};\bar{s}v)$. Now, take $s_{m+1} := v$ as the response of Duplicator for round $m+1$ of the game and observe that $[\star]_{i+1,\bar{t}t_{m+1},\bar{s}s_{m+1}}$ holds.
- Suppose that Spoiler chooses $s_{m+1} \in S$, say with $s_{m+1} \in H_j$. If $j \leqslant i$ then $s_{m+1} \in H_i$. Hence, by $[\star]_{i,\bar{t},\bar{s}}$, there exists $u \in T$ such that $(\mathfrak{T};\bar{t}u) \equiv_{n-m-1} (\mathfrak{T}_i;\bar{s}s_{m+1})$. Take $t_{m+1} := u$ as the response of Duplicator and observe that $[\star]_{i,\bar{t}t_{m+1},\bar{s}s_{m+1}}$ holds. On the other hand, if $i < j$, then $[\star]_{j,\bar{t},\bar{s}}$ holds, too. Duplicator will again have a response $t_{m+1} \in T$ such that $[\star]_{j,\bar{t}t_{m+1},\bar{s}s_{m+1}}$ holds.

Clearly, the tuples $(t_1, t_2, \ldots, t_n) \in T^n$ and $(s_1, s_2, \ldots, s_n) \in S^n$ thus defined form a local isomorphism between $\mathfrak{T}$ and $\mathfrak{S}$, as required. $\dashv$

§6. Concluding remarks. The study of classes of trees and their first-order theories is still in an early stage, and still much less explored than the study of first-order theories of classes of linear orders. In particular, the complete first-order axiomatisations of several important classes of trees have so far been open. The work presented in this paper was intended towards advancing that study and, in particular, answering some of these open questions. Our main contributions can be summarised as follows.

1. The main new technique developed here is a method for approximating any given (possibly, coloured) tree as a coloured linear order, so that the first-order theory of the resulting coloured linear order can be used to determine the first-order theory of the given tree. Using this technique, we have obtained a complete axiomatisation of the first-order theory of the class of coloured ζ-trees.

2. Another major new result in this work is the complete axiomatisation of the first-order theory of the class of coloured λ-trees. Our completeness proof of that axiomatisation mimics, modulo several essential complications, the completeness proof of the axiomatisation of the class of well-founded trees given in [2]. In fact, in the introduction of Section 5 of [2], Doets says that the completeness proof that is presented there "can be considered as a paradigm for a method applicable in a variety of situations, where the models considered belong to certain types of partial orderings (trees being the simplest example) and the Π_1^1-property involved can be either well-foundedness, converse well-foundedness, or, more generally, some kind of completeness."

3. We have also obtained and presented a few easier cases of complete axiomatisations of the first-order theories of other tree classes of natural interest, including ω-trees and η-trees. In addition, the paper summarised

the axiomatisations of the first-order theories of important classes of linear orders needed for axiomatising the first-order theories of the classes of trees mentioned above, though most of the former axiomatisations were already known from the literature.

Some directions for future work include the following.

- Use the methods developed here to axiomatise the first-order theories of other classes of trees of natural interest, such as the coloured trees in which all paths are complete linear orders.
- Generalise the method of [2] to a result of the kind of Theorem 4.4, presented here. Moreover, the only class to which Theorem 4.4 was applied here was $\{\zeta\}$, but it can also be applied to the classes $\{\omega^\star\}$ and $\{\zeta, \omega^\star\}$.
- The methods developed here can possibly be modified to apply to other, more general ordered structures, such as partially ordered sets and lattices.

REFERENCES

[1] K. DOETS, *Completeness and definability: Applications of the Ehrenfeucht game in second-order and intensional logic*, Ph.D. thesis, University of Amsterdam, 1987.

[2] ———, *Monadic Π_1^1-theories of Π_1^1-properties*. **Notre Dame Journal of Formal Logic**, vol. 30 (1989), no. 2, pp. 224–240.

[3] ———, **Basic Model Theory**, CSLI Publications & The European Association for Logic Language and Information, Stanford, 1996.

[4] M. DROSTE, **Structure of Partially Ordered Sets with Transitive Automorphism Groups,**, Memoirs of the American Mathematical Society, vol. 57, no. 334, American Mathematical Society, Providence, 1985.

[5] H.D. EBBINGHAUS, J. FLUM, and W. THOMAS, **Mathematical Logic**, Springer, New York, 1996.

[6] S. FEFERMAN and R.L. VAUGHT, The first order properties of products of algebraic systems. **Fundamenta Mathematicae**, vol. 47 (1959), pp. 57–103.

[7] V. GORANKO, *Trees and finite branching.* **Proceedings of the 2nd Panhellenic Logic Symposium**, 1999, pp. 91–101.

[8] V. GORANKO and R. KELLERMAN, Classes and theories of trees associated with a class of linear orders. **Logic Journal of the IGPL**, vol. 19 (2010), pp. 217–232.

[9] Y. GUREVICH and S. SHELAH, *The decision problem for branching time logic*, this JOURNAL, vol. 50 (1985), pp. 668–681.

[10] ———, *To the decision problem for branching time logic*, **Foundations of Logic and Linguistics: Problems and Their Solutions** (G. Dorn and P. Weingartner, editors), Plenum, New York, 1985, pp. 181–198.

[11] R. KELLERMAN, **First-order theories of bounded trees**, 2019, under submission.

[12] R. KELLERMAN, V. GORANKO, and A. ZANARDO, *Structural theory of trees. II. Completeness and completions of trees*, 2021, under submission.

[13] F. MWESIGYE and J.K. TRUSS, *Classification of finite coloured linear orderings*. **Order**, vol. 28 (2011), pp. 387–397.

[14] ———, Ehrenfeucht–Fraïssé games on ordinals. **Annals of Pure and Applied Logic**, vol. 169 (2018), no. 7, pp. 616–636.

[15] J. ROGERS, R. BACKOFEN, and K. VIJAY-SHANKAR, A first-order axiomatization of the theory of finite trees. **Journal of Logic, Language and Information**, vol. 4 (1995), pp. 5–39.

[16] J.G. ROSENSTEIN, **Linear Orderings**, Academic Press, New York, 1982.

[17] M. RUBIN, **The Reconstruction of Trees from Their Automorphism Groups**, Contemporary Mathematics, vol. 151, American Mathematical Society, Providence, 1993.

[18] J. SCHMERL, *On $\aleph_0$-categoricity and the theory of trees*. **Fundamenta Mathematicae**, vol. 94 (1977), pp. 121–128.

DEPARTMENT OF MATHEMATICS AND APPLIED MATHEMATICS
UNIVERSITY OF PRETORIA
PRETORIA, SOUTH AFRICA
E-mail: ruaan.kellerman@up.ac.za

DEPARTMENT OF PHILOSOPHY
STOCKHOLM UNIVERSITY
STOCKHOLM, SWEDEN
and
SCHOOL OF MATHEMATICS (VISITING PROFESSORSHIP)
UNIVERSITY OF THE WITWATERSRAND
JOHANNESBURG, SOUTH AFRICA
E-mail: valentin.goranko@philosophy.su.se