FUNDAMENTA
MATHEMATICAE
183 (2004)

# Countable 1-transitive coloured linear orderings II

by

**G. Campero-Arena** (México, D.F.) and **J. K. Truss** (Leeds)

**Abstract.** This paper gives a structure theorem for the class of countable 1-transitive coloured linear orderings for a countably infinite colour set, concluding the work begun in [1]. There we gave a complete classification of these orders for finite colour sets, of which there are $\aleph_1$. For infinite colour sets, the details are considerably more complicated, but many features from [1] occur here too, in more marked form, principally the use (now essential it seems) of coding trees, as a means of describing the structures in our list, of which there are now $2^{\aleph_0}$.

**1. Introduction.** In a companion paper to the present one [1], we introduced the class of countable 1-transitive coloured linear orders, and gave a complete classification of these structures in the case where the colour set is finite. In addition we classified all the countable homogeneous coloured linear orders, for finite or infinite colour set. We now turn to the considerably more complex case of countably infinitely coloured 1-transitive linear orders, that is, ones in which the colour set $C$ has cardinality $\aleph_0$. For a start there are $2^{\aleph_0}$ pairwise non-isomorphic infinitely coloured orders, as opposed to the $\aleph_1$ finitely coloured ones (since for instance there are this number of pairwise non-isomorphic countable linear orders, and we can colour the points of any such by countably many distinct colours to give trivially pairwise non-isomorphic countable 1-transitive coloured orders). Now there are immediate analogues of all the examples we encountered earlier, for instance the "coloured rationals" $\mathbb{Q}_{\aleph_0}$ with infinitely many colours, and concatenations in arbitrary countable order types of disjointly coloured 1-transitive linear orders. To keep track of all of them is quite complicated, and the use of coding trees begun in [1] now seems unavoidable.

The basic definitions are as follows. A *coloured linear order* (or *coloured chain*) has the form $(X,<,F)$ where $(X,<)$ is a linearly ordered set, and $F$ is a function from $X$ into (we may assume onto) a set $C$, called the set of *colours*. By saying that $(X,<,F)$ is 1-*transitive* we mean that for any two

2000 *Mathematics Subject Classification*: Primary 06A05.
This paper is based on part of the first author's PhD thesis at the University of Leeds.

points having the same colour, there is an automorphism taking the first to the second. Our task is to describe all the countable 1-transitive coloured linear orders, which generalizes work by Morel [3] in which all monochromatic ones (that is, with $|C| = 1$) were listed. Now in the classification given in [1] of all the examples for which $C$ is finite, we used "coding trees" as a method of describing the general construction of such a coloured chain. This was not strictly necessary, as there each structure can be fairly easily described in an inductive manner. The added complications in the infinitely coloured case seem to make them indispensable here.

The definition of "coding tree" in this context is given in Section 2. For an infinite coding tree it is not even clear what should be meant by saying that it "encodes" a particular coloured linear order, and this is explained in Section 3. With this notion we are able to show that any coding tree encodes a coloured linear order, that it is countable and 1-transitive, subject to this it is unique, and that any countable 1-transitive linear order is encoded by some coding tree. More precisely, we have the following:

Theorems 3.3 and 3.5. *Any coding tree encodes some coloured linear order, and this linear order is countable and* 1*-transitive.*

Theorem 3.4. *The coloured linear order described in the above theorem is unique up to isomorphism.*

Theorem 4.12. *Any countable* 1*-transitive coloured linear order is encoded by some coding tree.*

These results therefore describe the very close connection between the method of encoding, and the structures we are aiming to classify. It is true that the coding trees which arise can themselves have extremely complicated structures, so one might object that one is classifying one class of objects in terms of another, which is just as hard to describe. Nevertheless it seems clear that the above theorems definitely provide a great deal of information, which throws considerable light on the possibilities for countable 1-transitive coloured linear orders. We give examples to illustrate non-uniqueness of an ordering encoded by a coding tree once one relaxes the countability requirement.

In view of the above remarks, we refer to our main results as providing a "structure theorem" rather than a "classification". Ideally, for a class $\mathcal{C}$ of structures to be classified, the family of classifiers should be simpler than the members of $\mathcal{C}$, and it should be possible to "read off" information about the structures and the relations between them directly, and more easily, from the classifiers alone. In many cases this is true even here, since some of the coding trees are indeed only modest extensions of finite coding trees (they may be well-founded or conversely well-founded for instance). In the general case, however, this is not true. The fact that a family of countable structures has size continuum is *not* in itself a reason for saying that it is unclassifiable, as is illustrated most clearly by Cherlin's classification of the countable homogeneous directed graphs [2] (see also [6] which gives another example), where there are continuum many structures "listed" in terms of sets of integers. Where the classifiers are themselves more complicated, it is less clear that one can regard it as a classification.

Despite this, the classifiers used, "coding trees", *do* quite directly represent the way in which the coloured linear orders are built up, in a manner directly generalizing the inductive method expected for a finite coding tree. The problem is that since coding trees are no longer necessarily well-founded or conversely well-founded, we somehow have to express what limit points in the tree represent. The solution to this is to introduce an intermediate notion, which we call "expanded coding tree". The idea is that this has a much closer connection with the encoded ordering, and we may view it as obtained from the coding tree by actually "carrying out" the instructions at each vertex, and putting in below it not just the code for what happens in the linear order, but the full ordering described at that point.

To illustrate how this works in a simple case, a vertex of the coding tree may be labelled by $\mathbb{Q}$, and it will then have a single child. The intention is that the ordering encoded at this point should represent $\mathbb{Q}$ "copies" of whatever ordering is encoded at the child. In the expanded coding tree associated with a given coding tree, there may be many vertices which correspond to this parent vertex, but each of these will now have infinitely many children indexed by $\mathbb{Q}$. So the "intention" in the original label, to take $\mathbb{Q}$ copies, has actually been "carried out" in the expanded coding tree. There are several clauses in each of the principal definitions (of "coding tree", "expanded coding tree", and "encodes"), and this key idea is followed in each case. The main difference from the case of finite coding trees is the presence of limits, both from above and below, and special but natural conditions to handle these are also imposed. The other additional label, select$_n$, is more mysterious, and further explanation of the need for this, and how it is handled, is given later.

Expanded coding trees are used in the definition of "encodes" in the following way. Since an expanded coding tree ""associated with" a coding tree tells us the intended meaning of the coding tree, we can say that a coding tree encodes a coloured linear order if it is isomorphic to the set of leaves of some associated expanded coding tree. This accords with the intuition that as we pass down the coding tree, we find out more and more detailed information about the ordering actually represented.

We remark on one technical point. The decision was taken to require that coding trees be Dedekind–MacNeille complete (a notion explained in [5]), and to require this also for expanded coding trees. This may not be strictly necessary, but in view of the proof of Theorem 4.12 it seems entirely natural, since the maximal tree of "clumps" used in the proof is "nearly" Dedekind–MacNeille complete (just requiring adjunction of ramification points for this), and the corresponding family of convex subsets coloured by members of the maximal tree actually *is* Dedekind–MacNeille complete. We briefly recall this notion here, as described in [5].

For any partially ordered set $(X,<)$, if $I \subseteq X$, we write $\bigvee I = \{x \in X : (\forall y \in I)(y \leq x)\}$ and $\bigwedge I = \{x \in X : (\forall y \in I)(x \leq y)\}$. A *Dedekind ideal* is a subset $I$ of $X$ such that $I$ and $\bigvee I$ are non-empty, and $\bigwedge \bigvee I = I$. Dedekind ideals of the form $X^{\leq x} = \{y \in X : y \leq x\}$ are called *principal* and $(X,<)$ is *Dedekind–MacNeille complete* if every Dedekind ideal is principal. For any $(X,<)$, the family of Dedekind ideals then forms a Dedekind–MacNeille complete partial order $X^{\mathrm{D}}$ under $\subset$ in which $X$ embeds via $x \mapsto X^{\leq x}$, and this is called its *Dedekind–MacNeille completion*. The Dedekind–MacNeille completion of a linear order coincides with its Dedekind completion in the usual sense, and in this paper the notion will only be required for trees (see below), for which the Dedekind–MacNeille completion is also necessarily a tree.

**2. Coding trees.** We start by defining what a coding tree will mean in this paper. This definition generalizes the one used in [1]. As with the finite coding trees, every leaf will represent a singleton colour, so this time the coding trees may be infinite, and even have dense branches. Hence we are not able to define levels on the tree, and so the construction of a linear order from a tree and the definition of "encodes" used in [1] no longer work. As mentioned above, we also require the coding trees to be Dedekind–MacNeille complete, which means that they could even be uncountable, though only countably many points contribute in a non-trivial way to the encoding process. The labelling is similar to that for finite ones, with the addition of lim and $\mathrm{select}_n$ as new possible first labels, lim given to vertices having no child, and $\mathrm{select}_n$ given to certain vertices with no parent. Here a vertex labelled lim stands for the union of orders coded at points below it, and one labelled $\mathrm{select}_n$ stands for an order encoded at one of its children. (Each such point of the tree will stand for many subsets of the finally encoded order, and this one chosen child will not be the same at each occurrence.)

For us a *tree* is a partial order in which any two vertices have an upper bound, and the points above any element are linearly ordered. A *labelled tree* is a tree together with a function $\mathcal{L}$ on its vertices. A maximal element (which must also be greatest) is called the *root*, and minimal elements are called *leaves*. If $x \prec y$ in a tree and there is no point in between, then $x$ is a *child* of $y$, and $y$ is a *parent* of $x$. Distinct children of the same parent are *siblings*.

A maximal chain containing a leaf is called a *branch*. A *ramification point* is a vertex which is the supremum of two incomparable vertices. Ramification points always exist in the Dedekind–MacNeille completion of a tree, and, in fact, for a tree to be Dedekind–MacNeille complete it is sufficient to say that all its branches are Dedekind-complete as linear orderings, and that all its ramification points lie in the tree.

If $t$ is a vertex of a tree $(\tau, \prec)$, then the relation $\sim_t$ on $\{a \in \tau : a \prec t\}$ given by $a \sim_t b$ if there is $c \in \tau$ with $a, b \preceq c \prec t$ is an equivalence relation, and the $\sim_t$-classes are called *cones* at $t$. (See [6] for instance.) The number of cones at $t$ is its *ramification order*, and it is clear that $t$ has ramification order $> 1$ if and only if it is a ramification point. We shall also require that in any coding tree, each cone of a ramification point has a greatest element. We are grateful to the referee for pointing out that this property is not an automatic consequence of Dedekind–MacNeille completeness.

A particular coloured order which features throughout is $\mathbb{Q}_n$ for $1 < n \leq \aleph_0$, which is defined to be the rationals $\mathbb{Q}$ under the usual ordering, and with a colouring function $F : \mathbb{Q} \to n$ such that between each pair of rationals, all colours appear. This exists and is unique up to isomorphism (see [4] for example). If $Y_i$ are coloured linear orders, then $\mathbb{Q}_n(Y_0, Y_1, \ldots, Y_{n-1})$ is obtained from $\mathbb{Q}_n$ by replacing every point coloured $i$ by $Y_i$. (We shall write $\mathbb{Q}_n(Y_0, Y_1, \ldots, Y_{n-1})$ etc. even though $n$ may be infinite, but in that case, this is to be interpreted as $\mathbb{Q}_n(Y_0, Y_1, \ldots)$.) If $(\gamma, <)$ is a linear order, and $\{Y_i : i \in \gamma\}$ is a family of coloured linear orders, then their *concatenation* is obtained by replacing $i$ for $i \in \gamma$ by $Y_i$ throughout (retaining the original order and colours within each $Y_i$ and ordering points in different $Y_i$s according to their position in $\gamma$). Throughout we shall write $Z$ to stand for some countable (monochromatic) non-trivial 1-transitive linear order (that is, $\mathbb{Z}^\alpha$ with $\alpha \geq 1$ or $\mathbb{Q} \cdot \mathbb{Z}^\alpha$, $\alpha \geq 0$, for some countable ordinal $\alpha$, see [1]).

We are now ready to give the definition of "coding tree".

Definition 2.1. A *coding tree* is a labelled tree $(\tau, \prec \mathcal{L})$ such that

(i) $\tau$ has a root $r$,
(ii) $\tau$ has at most $\aleph_0$ leaves,
(iii) every vertex is a leaf or is above a leaf,
(iv) $\tau$ is Dedekind–MacNeille complete,
(v) all cones at ramification points of $\tau$ have greatest elements,
(vi) if $t, u, v \in \tau$ satisfy $t \prec u \prec v$, then there is $w \in \tau$ such that $t \preceq w \prec v$ having a sibling,
(vii) there are at most countably many vertices with only one child,
(viii) $\mathcal{L}$ is a labelling function $\mathcal{L} : \tau \to L$, where $L$ is a set of ordered pairs of the form $(\mathcal{F}(t), \mathcal{S}(t))$ described below.

The first label $\mathcal{F}(t)$ of a vertex $t$ is:

(a) $\mathbb{Q}_n$ for some $n$ with $1 < n \leq \aleph_0$, together with a bijection between $n$ and the children of $t$, provided that $t$ has ramification order $n$, and $t$ either has no parent, or has a parent and a sibling, or
(b) a countable linearly ordered set $(\gamma, <)$, together with a bijection between $\gamma$ and the children of $t$, provided that $t$ is a ramification point, or
(c) $\text{select}_n$ for some $n$ with $1 < n \leq \aleph_0$, together with a bijection between $n$ and the children of $t$, provided that $t$ is an infimum of points labelled by some $\mathbb{Q}_m$ or $Z$ ($m$, $Z$ not required all to be the same), or
(d) $Z$, provided that $t$ has just one child, or
(e) lim, provided that $t$ has just one cone, but no children, or
(f) 1, provided that $t$ is a leaf.

The second label $\mathcal{S}(t)$ of $t$ must satisfy:

(a) if $t$ has children, then $\mathcal{S}(t)$ is equal to the disjoint union of $\mathcal{S}(u)$ over all children $u$ of $t$,
(b) if $\mathcal{F}(t) = \lim$, then $\mathcal{S}(t) = \bigcup_{u \prec t} \mathcal{S}(u)$,
(c) if $t \neq r$ and $t$ has no parent, then $\mathcal{S}(t) = \bigcap \{\mathcal{S}(u) : t \prec u\}$,
(d) if $t$ is a leaf, then $\mathcal{S}(t)$ is a singleton $\{c\}$ for some $c \in C$, and all leaves have distinct second labels.

We remark that a non-trivial countable 1-transitive linear order may occur as $\mathcal{F}(t)$ in two possible guises, as a lexicographic product, or a concatenation, corresponding to clauses (d) and (b) respectively of the above definition. Since we can easily tell which is which by ramification order, we do not bother to signal this in the labelling. Although the bijections referred to are an integral part of the definition of the labels, we often ignore them in practice. For instance, we may talk about labels being "the same", or possibly "the same apart from the bijections", if we are just thinking of whether the label is a $\mathbb{Q}_n$ or a $\gamma$, together with the colour set (second label).

In Figures 1–5 we give some examples of coding trees, and we remark on the reasons for certain clauses in the definition.

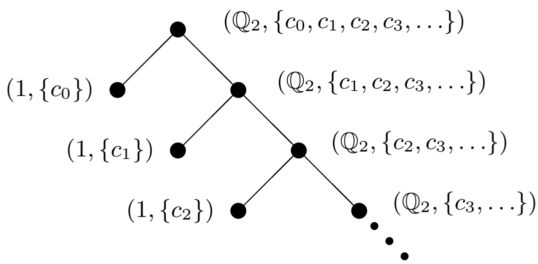


Fig. 1. In this tree every vertex is a leaf or is above a leaf

Note that every vertex in the tree in Figure 1 is a leaf or is above a leaf, as required, despite the fact that the sequence of vertices on the right has no infimum. So condition (iii) is fulfilled. We can modify this example by adding an infimum for this sequence, or indeed a whole new part of the tree at this point. For instance, the tree in Figure 2 is obtained in this way, where the infimum is labelled lim. Note that the presence of this point ensures Dedekind–MacNeille completeness. The tree in Figure 3 is another example, giving a typical occurrence of select$_n$.

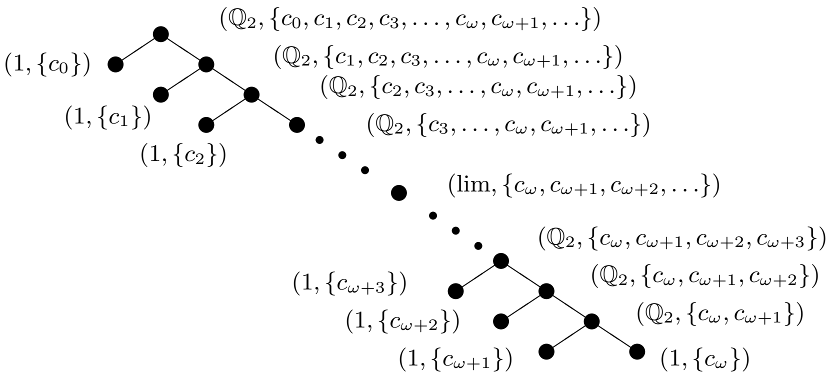


Fig. 2. The right-hand branch of this tree is Dedekind-complete

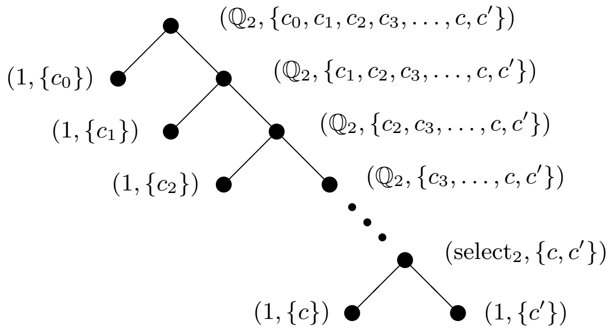


Fig. 3. A modification of the tree in Figure 1 involving select

We remark that since $\tau$ is Dedekind–MacNeille complete, it contains all its ramification points. Since any ramification point is the supremum of two leaves, of which there are at most $\aleph_0$, it follows that there are only countably many ramification points. The actual cardinality of $\tau$ may still be uncountable, though there are only countably many vertices of $\tau$ which have any "impact" on what is encoded. The others (irrational cuts in maximal chains) could be omitted, but it seems marginally simpler to include them.

Condition (vi) generalizes a similar condition imposed in the definition of coding tree in [1], and its role is to rule out having two or more consecutive lexicographic products (on the grounds that they can be performed in one go). It reads a little differently from the finite case since we also want to cover the possibility that there may be dense parts of the tree, so that "immediate successors" may not exist. The difference does mean for instance that the tree in Figure 4 is allowed, despite the fact that there are two consecutive vertices neither of which has a sibling.

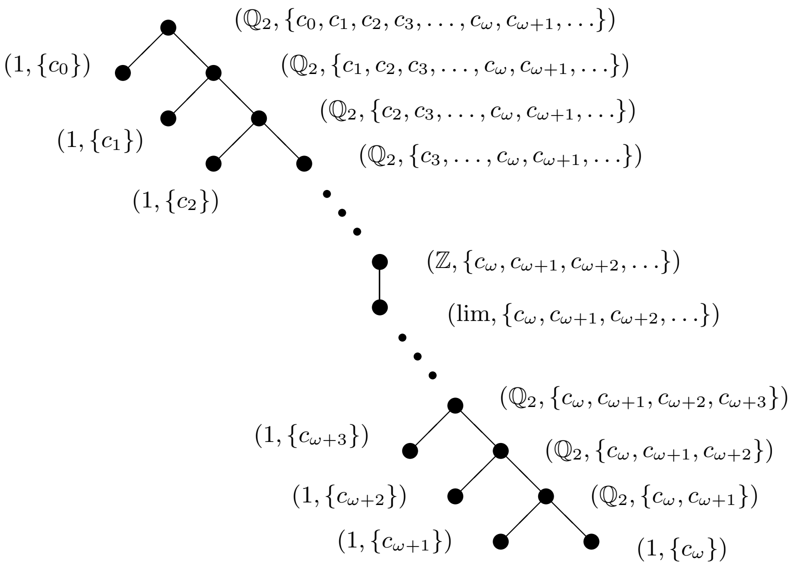


Fig. 4. This tree does not branch at every "two discrete levels"

Since every cone at a ramification point $t$ has a greatest element, $t$ must represent either a concatenation (now of a possibly infinite family), some $\mathbb{Q}_n$-combination, or a "selection" (to be explained later). For finite trees these ramification points would automatically have children; here we need to require this explicitly. For concatenations, the order in which the children are concatenated is important. In [1] the ordering was included as part of the definition of "coding tree"; here instead at this point we have said that there is a bijection given between the set of children and the order over which we are concatenating (which has essentially the same effect). Hence the greatest elements of the cones at every ramification point are linearly ordered, and this induces a linear ordering on the set of branches of the tree. See Figure 5 for an example. For vertices labelled $\mathbb{Q}_n$, the order of the children is not important, but we still need to know which is which, so we fix a bijection between $n$ and the children.

Condition (vii) ensures that there are at most countably many vertices representing lexicographic products immediately above a vertex labelled lim (without this condition, the encoded order would be uncountable). Vertices with no child are labelled lim, and will stand for the union of all the linear orders represented at the vertices below them. Since the trees are Dedekind–MacNeille complete, there may be uncountably many such points.

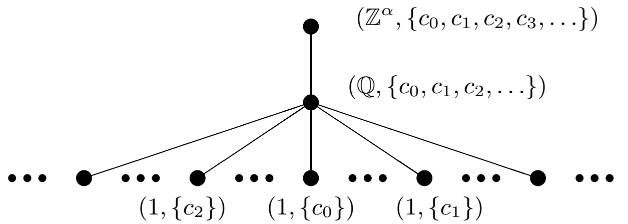


Fig. 5. The children of the vertex labelled $\mathbb{Q}$ are ordered like $\mathbb{Q}$

The labelling on infinite coding trees generalizes the one imposed on finite trees in [1]. Vertices with more than one cone may be labelled by any order type which is in 1-1 correspondence with their children, so any countable order type may appear as the first label. Note that here, in contrast to [1], we allow consecutive points labelled by order types (representing concatenations). It would be possible to rule this out, but this would result in additional, perhaps not very illuminating, technicalities. Furthermore, adopting this strategy considerably eases constructing a coding tree from a linear order (and in fact we find that concatenations over two-element sets suffice).

The conditions imposed on the second label are exactly the same as those for the finite coding trees, with natural extensions corresponding to upward or downward limit points. We remark that if $t_0$ and $t_1$ are incomparable, then $\mathcal{S}(t_0) \cap \mathcal{S}(t_1) = \emptyset$, and that if $t_0 \prec t_1$, then $\mathcal{S}(t_0) \subseteq \mathcal{S}(t_1)$.

We have only given some relatively simple examples of coding trees. More complicated ones may involve for instance branches that embed densely ordered sets, which are harder to visualize (and for which it is particularly hard to imagine what the encoded orderings look like) but which still obey the definition.

**3. How a coding tree encodes a linear order.** Coding trees are meant to help us to describe all the coloured orderings in our class. Since they need not be well-founded either upwards or downwards, no straightforward definition of what it means for such a tree to "encode" a coloured linear order is available. A further complication arises from "selection", which works rather differently from the other labels. To achieve what is wanted, we introduce the idea of an "expanded coding tree", which provides a good intermediate stage between code and encoded object. This will enable us to clarify the situation, and indeed it seems to be necessary even to *define* what we mean by saying that a coding tree encodes a coloured linear order.

Definition 3.1. An *expanded coding tree* is a labelled tree $(E, \prec, \mathcal{L})$ such that:

(i) $E$ has a root $r$,
(ii) $E$ has at most $\aleph_0$ leaves,
(iii) every vertex is a leaf or is above a leaf,
(iv) $E$ is Dedekind–MacNeille complete,
(v) all cones at ramification points of $\tau$ have greatest elements,
(vi) if $t, u, v \in E$ satisfy $t \prec u \prec v$, then there is $w \in E$ labelled $\mathbb{Q}_n$, $\gamma$, or select$_n$ such that $t \prec w \preceq v$,
(vii) $\mathcal{L}$ is a labelling function $\mathcal{L} : E \to L$ for $L$ as in Definition 2.1 such that each $t \in E$ satisfies (exactly) one of the following:
   (a) $\mathcal{F}(t)$ is $\mathbb{Q}_n$ for some $n$ with $1 < n \leq \aleph_0$ together with a bijection $f$ between $\mathbb{Q}_n$ and the set of children of $t$, such that for $x, y \in \mathbb{Q}_n$, $\mathcal{S}(f(x)) = \mathcal{S}(f(y)) \Leftrightarrow x, y$ have the same colour, and then the trees below $f(x)$ and $f(y)$ are isomorphic, and the family of distinct $\mathcal{S}(f(x))$s forms a partition of $\mathcal{S}(t)$,
   (b) $\mathcal{F}(t)$ is a non-trivial countable linear order $\gamma$ together with a bijection $f$ between $\gamma$ and the set of children of $t$, and $\{\mathcal{S}(f(x)) : x \in \gamma\}$ forms a partition of $\mathcal{S}(t)$,
   (c) $t$ is an infimum of points labelled by some $\mathbb{Q}_m$ or $Z$ ($m$, $Z$ not all required to be the same), $\mathcal{F}(t)$ is select$_n$, $t$ just has one child $t^-$ say, and $\mathcal{S}(t^-) \subset \mathcal{S}(t)$,
   (d) $\mathcal{F}(t)$ is a non-trivial countable 1-transitive linear order $Z$, $t$ has children $\{t_z : z \in Z\}$, the trees below any two of the children are isomorphic, and for each $z \in Z$, $\mathcal{S}(t_z) = \mathcal{S}(t)$,
   (e) $\mathcal{F}(t) = \lim$, there is just one cone below $t$, $t$ has no children, and $\mathcal{S}(t) = \bigcup\{\mathcal{S}(u) : u \prec t\}$,
   (f) $\mathcal{F}(t) = 1$, $t$ is a leaf, and $|\mathcal{S}(t)| = 1$,

and in addition, if $t \neq r$ has no parent, then $\mathcal{S}(t) = \bigcap\{\mathcal{S}(u) : t \prec u\}$.

Any expanded coding tree has a natural "left-right" ordering on the branches (equivalently, leaves) given by $B_1 < B_2$ if at the point $t$ at which distinct branches $B_1$ and $B_2$ diverge, the child of $t$ in $B_1$ is less than the child of $t$ in $B_2$. Note that by Dedekind–MacNeille completeness, $B_1 \cap B_2$ must have a least point, and it must be a ramification point, hence labelled $\mathbb{Q}_n$, $\gamma$, or $Z$, so the ordering on the children is well-defined. This is referred to as the "branch order". It also induces an order on any antichain.

We remark that clause (vi) is slightly different from the corresponding clause in Definition 2.1. There $w$ was asserted to exist which has a sibling; here it is its *parent* $w'$ which is asserted to exist, and this will fulfil $u \prec w' \preceq v$ rather than $u \preceq w \prec v$.

Next we indicate the connection between “coding tree” and “expanded coding tree”. Roughly speaking, the latter is a “fattened” version of the former, where all the labels have been filled out in line with their intended meanings.

DEFINITION 3.2. Let $(\tau, \prec, \mathcal{L})$ be a coding tree, and $(E, \prec, \mathcal{L}')$ be an expanded coding tree. We say that $E$ is *associated with* $\tau$ if there is a function $\varphi$ from $E$ to $\tau$ which takes the root of $E$ to the root of $\tau$, each leaf of $E$ to some leaf of $\tau$, and

(i) $t_1 \prec t_2 \Rightarrow \varphi(t_1) \prec \varphi(t_2)$,
(ii) for each vertex $t$ of $E$ not labelled $\text{select}_n$, $\varphi$ maps $\{u \in E : u \preceq t\}$ onto $\{u \in \tau : u \preceq \varphi(t)\}$, and for any leaf $l$ of $E$, $\varphi$ maps $[l, r]$ onto $[\varphi(l), \varphi(r)]$,
(iii) $\mathcal{L}(\varphi(t)) = \mathcal{L}'(t)$ (as far as the second components are concerned, and also the first parts of the first components—compatibility of the bijections is ensured by the other clauses), and
(iv) if $t \in \tau$ and $\mathcal{F}(t) = \text{select}_n$, and the children of $t$ under the given bijection with $n$ are $t_0, t_1, t_2, \ldots, t_{n-1}$, then $\{s \in E : \varphi(s) = t\}$ is isomorphic to $\mathbb{Q}_n$ under the branch order, where the colouring $F : \mathbb{Q}_n \to n$ is given by $\varphi(s^-) = t_{F(s)}$ (where $s^-$ is the unique child of $s$ in $E$—see Definition 3.1(vii)(c)).

Finally, we say that the coding tree $(\tau, \prec, \mathcal{L})$ *encodes* the coloured linear order $(X, <, F)$ if there is an expanded coding tree $(E, \prec, \mathcal{L})$ associated with $\tau$ such that $X$ is (order and colour-) isomorphic to the set of leaves of $E$ under the branch order.

The tricky point in the above definition is working out how to handle points labelled $\text{select}_n$. Since we shall see later that the convex sets throughout the linear order corresponding to such a vertex $t$ are in fact ordered like $\mathbb{Q}_n$, and furthermore this holds below each vertex strictly above $t$, we have decided to include this as part of the definition, though at each occurrence in the encoded order, only one of these colours actually arises. It follows from this that there are (many) expanded coding trees which are not associated with any coding tree, but it does not seem worth strengthening the definition of expanded coding tree just for this purpose. The reason why the definition of “encodes” is so involved is that a vertex labelled $\text{select}_n$ stands for different convex sets throughout the encoded order, so to describe how it is meant to be interpreted, we somehow have to refer to the whole order. A vertex with any other label always stands for the same convex set (up to isomorphism).

We remark that it would be possible to allow points to be labelled $\text{select}_n$ even if they are not the infima of vertices labelled $\mathbb{Q}_m$ or $Z$; however this use is unnecessary, since any coloured order so encoded can also be encoded without the use of such points. For instance, the ordering encoded by a five-element coding tree $\{r, t, l_0, l_1, l_2\}$, where the root $r$ is labelled $\mathbb{Q}_2$, $t$ is labelled $\text{select}_2$, $l_0$ and $t$ are the children of $r$, and $l_1$ and $l_2$ are the children of $t$, is isomorphic to $\mathbb{Q}_3$, so can be more simply encoded.

THEOREM 3.3. *Any coding tree encodes some coloured linear order.*

*Proof.* We begin by assuming that no vertices of the given coding tree $\tau$ are labelled by $\text{select}_n$. We shall then obtain the encoding of a general coding tree as a subset of that for a suitable $\text{select}_n$-free coding tree.

The definition requires us to find an expanded coding tree associated with $\tau$. Let $B$ be a branch of $\tau$. Then $B$ is a linear order which has a least element (the leaf of $B$) and a greatest element (the root of $\tau$), but which may not be countable or well-ordered. We define functions, which we call "decoding functions" and generally write as $\sigma$, on vertices in $B$ whose first label is $\mathbb{Q}_n$ or $Z$. Although the domain of $\sigma$ is only a subset of $B$, for ease we refer to the whole of $B$ as $\operatorname{dom}\sigma$. The idea is that each $\sigma$ will correspond to one point in the final structure: the linear order encoded by $\tau$. Hence, if $l$ is the leaf of $B$, the decoding functions defined on $B$ will correspond to the points coloured $\mathcal{S}(l)$ in the linear order encoded.

A *decoding function* is a function $\sigma$ defined on a branch $B$ of $\tau$, such that for each $t \in B$ with $\mathcal{F}(t) = \mathbb{Q}_n$ or $Z$,

(i) if $\mathcal{F}(t) = \mathbb{Q}_n$, with $1 < n \leq \aleph_0$, so that $t$ has $n$ children, say $t_0, \ldots, t_{n-1}$, with $t_i$ the child of $t$ in $B$, then $\sigma(t)$ is a point of $\mathbb{Q}_n$ coloured $i$,

(ii) if $\mathcal{F}(t)$ is a countable 1-transitive linear order, then $\sigma(t) \in \mathcal{F}(t)$.

Now we cannot take *all* decoding functions, for two main reasons. First, the result would (usually) be uncountable. Second, since we want decoding functions to be linearly ordered by first difference (from the top down), we need to ensure that there always *is* a point of first difference. If we take all decoding functions, this need not be the case.

To cut down suitably, we choose suitable "default values", which the decoding functions will be required to take at all but finitely many points. For each $\mathbb{Q}_n$, and each colour $i < n$, we choose a point $a_i$ in $\mathbb{Q}_n$ coloured $i$, and for each $Z$, we choose a point $a$ of $Z$. These are called *default values.* The family $\Sigma_\tau$ of coding functions associated with $\tau$ comprises all coding functions $\sigma$ defined on a branch $B$ of $\tau$ which take the default value at all but finitely many ($\mathbb{Q}_n$ or $Z$) vertices.

We observe that $\Sigma_\tau$ may be naturally linearly ordered by first difference (from top down). For let $\sigma_1, \sigma_2$ be distinct members of $\Sigma_\tau$, having domains $B_1$ and $B_2$ respectively. By Dedekind–MacNeille completeness of $\tau$, $u =$

$\inf(B_1 \cap B_2)$ lies in $\tau$. If $\sigma_1(t) \neq \sigma_2(t)$ for some $t \succeq u$, then as $B_1$ and $B_2$ agree above $u$, and $\sigma_1, \sigma_2$ each take only finitely many non-default values, there is a greatest such point $t$, and we let $\sigma_1 < \sigma_2 \Leftrightarrow \sigma_1(t) < \sigma_2(t)$ in that copy of $\mathbb{Q}_n$ or $Z$. If $\sigma_1$ and $\sigma_2$ agree above $u$, then as $u$ is clearly a ramification point, it must be labelled by some $\gamma$, and we let $\sigma_1 < \sigma_2 \Leftrightarrow t_1 < t_2$ under the given 1-1 correspondence with $\gamma$, where $t_1$ and $t_2$ are the children of $u$ in $B_1$ and $B_2$ respectively.

Now $\Sigma_\tau$ will be the coloured linear order that is encoded by $\tau$, but to fulfil the definition, we also have to say how to form $E$, an associated expanded coding tree. We take it to consist of the set of restrictions of members of $\Sigma_\tau$ to linearly ordered subsets of $\tau$ of the form $(t, r] = \{s \in \tau : t \prec s\}$. More precisely,

$$E = \{(t, \sigma|(t, r]) : \sigma \in \Sigma_\tau,\ t \in B,\ B = \operatorname{dom}\sigma\}.$$

The labelling on $E$ is given by the first co-ordinate, $\mathcal{L}'((t, \sigma|(t, r])) = \mathcal{L}(t)$, except that we shall state precisely below how the relevant bijections in $\mathcal{L}'((t, \sigma|(t, r]))$ are obtained from those in $\mathcal{L}(t)$, and $E$ is partially ordered by extension, that is, if $t \preceq s$ in $\tau$ lie in a branch $B = \operatorname{dom}\sigma$, then $(t, \sigma|(t, r]) \preceq (s, \sigma|(s, r])$.

Under this definition it is clear that $E$ is a labelled tree, its root is $(r, \emptyset)$, and any $(t, \sigma|(t, r])$ lies above a leaf $(l, \sigma|(l, r])$, where $l$ is the leaf in $B = \operatorname{dom}\sigma$. To see that $E$ has countably many leaves, note that $\tau$ has only countably many leaves, hence only countably many branches $B$. For each $B$, there are only countably many finite sets of points labelled $\mathbb{Q}_n$ or $Z$, and for each such, there are only countably many coding functions taking non-default values at precisely the elements of this set, and hence $\Sigma_\tau$ is countable. Therefore $E$ has only countably many leaves.

Each branch of $E$ is isomorphic to a branch of $\tau$, so is Dedekind complete, and the proof that all ramification points of $E$ have been included is essentially the same argument as above, when we showed how $\Sigma_\tau$ was linearly ordered by first difference, from which it follows that $E$ is Dedekind–MacNeille complete. Note that ramification points of $\tau$ may be labelled $\mathbb{Q}_n$, $\gamma$, or in the general case, $\text{select}_n$; ramification points of $E$ may be labelled $\mathbb{Q}_n$, $\gamma$, or $Z$. The fact that all cones at ramification points have greatest elements follows from the corresponding property of $\tau$.

Suppose $(t, \sigma|(t, r]) \prec (u, \sigma|(u, r]) \prec (v, \sigma|(v, r])$ in $E$. Then $t \prec u \prec v$ in $\tau$, so there is $s$ with $t \preceq s \prec v$ having a sibling in $\tau$. Let $w$ be the parent of $s$. Then $w$ is labelled $\mathbb{Q}_n$ or $\gamma$, as required.

Next consider the labels. If $\mathcal{F}(t) = \mathbb{Q}_n$ then $t$ has children $t_0, t_1, \dots, t_{n-1}$ under the specified 1-1 correspondence, and the possible extensions of any $\sigma|(t, r]$ to $\sigma|(t_i, r] = \sigma|[t, r]$ just depend on the value given to $\sigma(t)$, and these are given by $\mathbb{Q}_n$ with the correct colours, by the first clause of the definition of "decoding function", so this tells us which bijection to take between $\mathbb{Q}_n$ and the children of $(t, \sigma|(t,r])$ in $\mathcal{L}'((t, \sigma|(t,r]))$. If $\mathcal{F}(t) = \gamma$, then $\sigma$ is not defined on $t$, so $\sigma|(s,r] = \sigma|[t,r] = \sigma|(t,r]$ for each child $s$ of $t$. Thus the children $(s, \sigma|(s,r])$ of $(t, \sigma|(t,r])$ in $E$ are just determined by the first co-ordinate $s$, and these form a copy of $\gamma$, so this describes the bijection in this case. If $\mathcal{F}(t) = Z$, we argue as for $\mathbb{Q}_n$. This time there is just one child in $\tau$, but $Z$ children in $E$.

Finally, if $\mathcal{F}(t) = \lim$, we show that there is only one cone at $(t, \sigma|(t,r])$. Let $(t_1, \sigma_1|(t_1,r]), (t_2, \sigma_2|(t_2,r]) \prec (t, \sigma|(t,r])$, with the object of showing that they lie in the same cone below $(t, \sigma|(t,r])$. Since there is only one cone in $\tau$ below $t$, we may suppose that $t_1 = t_2$, and the definition of $\Sigma_\tau$ ensures that above some point below $t$, $\sigma_1$ and $\sigma_2$ take the same default value, so $(t_1, \sigma_1|(t_1,r])$ and $(t_2, \sigma_2|(t_2,r])$ have some common upper bound below $(t, \sigma|(t,r])$. Clearly $(t, \sigma|(t,r])$ is not a parent, since $t$ is not, so $t$ fulfils the requirements to be labelled lim.

The required properties of the second labels $\mathcal{S}$, in particular the last point (continuity on infima) follow from those of $\tau$.

It remains to remark that $E$ is associated with $\tau$, and this follows by considering the mapping $\varphi$ given by $\varphi((t, \sigma|(t,r])) = t$, which clearly preserves root, leaves, and labels (except the corresponding bijections), and $(t_1, \sigma_1|(t_1,r]) \prec (t_2, \sigma_2|(t_2,r]) \Rightarrow t_1 \prec t_2$.

Now let us move on to the general case in which $\tau$ may have points labelled by $\text{select}_n$. We consider a related coding tree $\tau'$ obtained from $\tau$ by replacing each $\text{select}_n$ label at a vertex $t$ by a concatenation; that is, with $\gamma$ equal to $n$ (finite or $\omega$), and an arbitrary bijection from $n$ to the children of the $t$. This $\tau'$ is a tree not containing any $\text{select}_n$ labels, where all possible choices for the selections are represented (in an arbitrary and ultimately unimportant order), and our task is to show how an appropriate subset of the $\Sigma_{\tau'}$ defined above will be the encoding of $\tau$. We let $E'$ be the expanded coding tree ("canonically") determined as above from $\tau'$.

From Definition 3.2 we see that we have to identify the points of $E'$ corresponding to a particular $t \in \tau$ labelled $\text{select}_n$ with $\mathbb{Q}_n$. Once we have done this, the selections that are made are then given as in Definition 3.2(iv). Now the fact that the points corresponding to $t$ are densely linearly ordered without endpoints follows easily since $t$ is the infimum of a sequence of vertices labelled by some $\mathbb{Q}_m$ or $Z$. But we recall that $\mathbb{Q}$ and $\mathbb{Q}_n$ are the same *set*—the only difference between them is that $\mathbb{Q}_n$ has a colouring, and in fact we can use an *arbitrary* indexing of the points corresponding to $t$ by $\mathbb{Q}_n$. The reason for this is that any non-empty open interval of $\mathbb{Q}_n$ is isomorphic to $\mathbb{Q}_n$, so it "cannot" make any difference which particular indexing we choose. (This is handled formally in the next two theorems.)

Now that this "colouring" has been carried out for all points of $E'$ corresponding to $\text{select}_n$ vertices of $\tau$, we can describe which subset $E$ of $E'$ we take to correspond to $\tau$. It just consists of all decoding functions $\sigma$ on branches $B$ such that for every $\text{select}_n$ vertex in $B$, the child of $t$ in $B$ is $t_i$, where $\sigma|(t,r]$ is coloured $i$.

Since $E$ is a subset of $E'$, it is a tree, and by changing the labels of points which correspond to vertices of $\tau$ labelled $\text{select}_n$ from concatenation back to $\text{select}_n$, it becomes an expanded coding tree once again. The verification of the properties of "expanded coding tree" and of the fact that it is associated with $\tau$ are straightforward. Hence $\tau$ encodes the set $\Sigma_\tau$ of leaves of $E$. ∎

Now we move on to establishing uniqueness of the order encoded (provided it is countable).

THEOREM 3.4. *Let $(\tau, \prec, \mathcal{L})$ be a coding tree. Then any two countable coloured linear orders encoded by $\tau$ are isomorphic.*

*Proof.* Suppose that $X_1$ and $X_2$ are both encoded by $\tau$, and are countable. Let $X_1$ and $X_2$ be the sets of leaves of expanded coding trees $E_1$ and $E_2$ respectively, and let $\varphi_1 : E_1 \to \tau$, $\varphi_2 : E_2 \to \tau$ be corresponding functions. The proof will be by back-and-forth, so we need to describe a suitable family of approximations to the desired isomorphism. We shall actually show that $E_1$ and $E_2$ are isomorphic, and the isomorphism between $X_1$ and $X_2$ will be the restriction.

Let $P$ be the family of isomorphisms $p$ from a finite subset of $E_1$ into $E_2$ such that

(i) the root of $E_1$ lies in $\operatorname{dom} p$, and the root of $E_2$ lies in $\operatorname{range} p$,
(ii) $\operatorname{dom} p$ and $\operatorname{range} p$ contain all their ramification points (that is, they are closed under formation of least upper bounds),
(iii) for $t \in \operatorname{dom} p$, $\varphi_1(t) = \varphi_2(p(t))$,
(iv) points of $\operatorname{dom} p$ and $\operatorname{range} p$ are either not labelled lim, or are children of points labelled $\text{select}_n$,
(v) if $t \in \operatorname{dom} p$ is a ramification point, then there is an isomorphism of the set of children of $t$ in $E_1$ to the set of children of $p(t)$ in $E_2$, such that if $u \prec t$, $u \in \operatorname{dom} p$, then the isomorphism takes the child of $t$ above $u$ to the child of $p(t)$ above $p(u)$,
(vi) if $t$ is labelled $\text{select}_n$ (so that it only has one child), then its child is also in $\operatorname{dom} p$.

We shall show that if $p \in P$ and $t \in E_1$ is not labelled lim, or is the child of a point labelled $\text{select}_n$, then there is an extension $q$ of $p$ in $P$ such that $t \in \operatorname{dom} q$ ("forth" step), and with a similar extension property for the range ("back" step). By back-and-forth it then follows from countability of the set of vertices of $E_1$, $E_2$ not labelled lim (or having a $\text{select}_n$ vertex as parent) that there is an isomorphism $f$ from the set of such points in $E_1$ to those in $E_2$, and this extends to an isomorphism $E_1 \to E_2$ by continuity.

The main thing is therefore to establish extension to the domain (extension to the range being precisely similar), so let $p$ and $t$ be given. If $t \in \operatorname{dom} p$ then we let $q = p$. Otherwise $t \notin \operatorname{dom} p$. Since $r \in \operatorname{dom} p$, there is a least vertex $v$ of $\operatorname{dom} p$ above $t$.

Case 1: There is also a vertex of dom $p$ below $t$; let $u$ be the greatest such. This is well-defined since $\operatorname{dom} p$ is finite and contains all its ramification points. Thus $p(u)$ and $p(v)$ exist, and there is a unique point $t'$ of $E_2$ such that $p(u) \leq t' \leq p(v)$ and $\varphi_2(t') = \varphi_1(t)$, so we let $q = p \cup \{(t, t')\}$. This exists since $\varphi_2$ maps $[l, r]$ onto $[\varphi_2(l), \varphi_2(r)]$ for every leaf $l \preceq t$; it is unique since if $t''$ is another then we cannot have either $t' \prec t''$ or $t'' \prec t'$ as this would give $\varphi_2(t') \prec \varphi_2(t'')$, $\varphi_2(t'') \prec \varphi_2(t')$ respectively. The argument has to be slightly modified if $t$ is labelled lim or $\text{select}_n$. In the former case, $t$ must have a parent labelled $\text{select}_n$, and in the latter, $t$ has a unique child which has to be adjoined to the domain (if not already present as $u$). So we just have to consider the possibility that the parent $t^+$ of $t$ is labelled $\text{select}_n$ (in the second case the original $t$ is renamed $t^+$) and extend to include both $t$ and $t^+$ in $\operatorname{dom} q$.

If $t$ is a ramification point, we further choose, by 1-transitivity, an isomorphism from the set of children of $t$ to those of $t'$ which takes the child of $t$ above $u$ to the child of $p(t)$ above $p(u)$.

Case 2: $v$ is minimal in $\operatorname{dom} p$. We have $\varphi_1(v) = \varphi_2(p(v))$ and $\varphi_1(t) < \varphi_1(v)$. By condition (vi), $v$ is not labelled $\text{select}_n$, and hence $\varphi_2$ maps $\{u : u \preceq p(v)\}$ onto $\{u : u \preceq \varphi_2(p(v))\}$. Therefore there is $u \preceq p(v)$ such that $\varphi_2(u) = \varphi_1(t)$, and we let $q = p \cup \{(t, u)\}$. If $t$ is labelled $\text{select}_n$ or lim we again modify the above by including a new pair of points in $\operatorname{dom} q$. In addition, if $t$ is a ramification point, we choose an isomorphism from the set of children of $t$ to those of $t'$.

Case 3: There is no vertex of $\operatorname{dom} p$ below $t$, $v$ is not minimal in $\operatorname{dom} p$, $u \in \operatorname{dom} p$, $u \prec v$ say, and the least upper bound of $u$ and $t$ is equal to $v$. Now we may suppose that $t$ is a child of $v$. For $v$, being a ramification point, has a unique child greater than or equal to $t$, and if we extend $p$ to $p'$ with this child in its domain, then this child is now minimal in $\operatorname{dom} p'$ and (if still $t \notin \operatorname{dom} p'$) we can appeal to Case 2. The fact that this extension is possible so as to fulfil the conditions follows from clause (v); in fact that tells us exactly which child of $p(v)$ the vertex $t$ should be mapped to under $q$ (and we note that $t$ will not be labelled $\text{select}_n$ or lim).

Case 4: There is no vertex of $\operatorname{dom} p$ below $t$, $v$ is not minimal in $\operatorname{dom} p$, $u \in \operatorname{dom} p$, $u \prec v$ say, but the least upper bound $w$ of $u$ and $t$ is not equal to $v$. We can extend $p$ to $p'$ in $P$ so that $w \in \operatorname{dom} p$ using Case 1, and this then reduces to Case 3, since now the least upper bound of $u$ and $t$ lies in $\operatorname{dom} p'$. ∎

We remark that without the restriction to countability of the encoded order, we do not in general get uniqueness, in contrast to [1]. For in addition to the "canonical" linear ordering $\Sigma_\tau$ as described above, comprising all those decoding functions which take the default value at all but finitely many points, we can also consider the family of all decoding functions which are allowed to take a non-default value on any (downwards) well-ordered set of points. This will give rise to an expanded coding tree, and it will still be linearly ordered by first difference from the top. However, in cases in which there is a strictly descending sequence of vertices labelled $\mathbb{Q}_n$ or $Z$ (such as those in Figures 1–4), the set of leaves is uncountable, so cannot be isomorphic to $\Sigma_\tau$. Several variations on the same idea are possible, which we omit.

Next we establish 1-transitivity of the encoded order. Because of the presence of select$_n$ vertices, it was easier to begin by demonstrating uniqueness.

THEOREM 3.5. *If* $(\tau, \prec, \mathcal{L})$ *is a coding tree*, *the coloured linear order* $\Sigma_\tau$ *it encodes described above is countable and* 1*-transitive.*

*Proof.* Let $E$ be associated with $\tau$ as given above, so that $\Sigma_\tau$ is the set of leaves of $E$. We saw above that $\Sigma_\tau$ is countable.

To prove that $(\Sigma_\tau, <_\tau, F_\tau)$ is 1-transitive, suppose that $\sigma_0$ and $\sigma_1$ in $\Sigma_\tau$ have the same colour. Then they have the same domain, the branch $B = [l, r]$, say. We adapt the proof of Theorem 3.4, using $E_1 = E_2 = E$, the expanded coding tree canonically associated with $\tau$ (of which $\Sigma_\tau$ is the set of leaves). Start with the map $p$ fixing $r$ and taking $\sigma_0$ to $\sigma_1$ (and if $r$ is a ramification point of $E$, an automorphism of its set of children taking the child in $\sigma_0$ to the child in $\sigma_1$). Then $p \in P$, so by the proof of Theorem 3.4 there is an automorphism of $E$ extending $p$, and this restricts to an automorphism of $\Sigma_\tau$ taking $\sigma_0$ to $\sigma_1$. This establishes 1-transitivity. ∎

**4. Finding a coding tree for a linear order.** In the previous section we found a unique linear order encoded by a given coding tree, which was countable and 1-transitive. Now we turn to the converse, and show that any countable 1-transitive coloured linear order $(X, <, F)$ is encoded by some coding tree (which however need not be unique). Following the ideas of [1], where an inductive method of building the tree was used which involved "bunching" suitable colours together, we introduce the idea of a "clump", which is a subset of $C$ colouring some convex subset of $X$. The coding tree is formed from a maximal tree of clumps on addition of a few additional vertices (the ones to be labelled as lexicographic products or selections) and adjoining labels. A straightforward induction is impossible here, since the tree will be infinite and not necessarily well-founded. Instead it has to be defined explicitly, and its Dedekind–MacNeille completeness used in place of induction. At the same time we shall build an associated expanded coding tree, and this will demonstrate that the coding tree formed does encode the original linear order.

DEFINITION 4.1. Let $(X,<,F)$ be a linear order with colour set $C$. A *clump* is a non-empty subset $C_0$ of $C$ such that $C_0 = \{F(z) : z \in I\}$ for some convex subset $I$ of $X$.

LEMMA 4.2. *Let $(X,<,F)$ be a 1-transitive coloured linear order. Then for any clump $C_0$ and non-empty convex subset $I$ of $X$ such that $F(I) \subseteq C_0$ there is a unique maximal convex subset of $X$ containing $I$ coloured only by members of $C_0$, and this is 1-transitive.*

*Proof.* Let $P = \{J \subseteq X : I \subseteq J,\ J \text{ is convex, and } F(J) \subseteq C_0\}$. Then $\bigcup P$ is convex, since all the members of $P$ contain $I$. Hence $\bigcup P$ is the desired maximal convex set coloured by $C_0$. To see that $F(\bigcup P) = C_0$ (rather than just $F(\bigcup P) \subseteq C_0$), note that as $C_0$ is a clump, $F(J) = C_0$ for some convex $J$. By 1-transitivity, we may suppose that $J \cap I \neq \emptyset$. Hence $J \cup I \in P$, so $F(\bigcup P) = C_0$.

For 1-transitivity of $\bigcup P$, let $x, y \in \bigcup P$. As $X$ is 1-transitive, there is an automorphism $g$ taking $x$ to $y$. Since $F(g(\bigcup P)) \subseteq C_0$ and $\bigcup P \cap g(\bigcup P) \neq \emptyset$, it follows that $g(\bigcup P) \in P$, so $g(\bigcup P) \subseteq \bigcup P$. Similarly $g^{-1}(\bigcup P) \subseteq \bigcup P$, so $g(\bigcup P) = \bigcup P$. ∎

LEMMA 4.3. *If $(X,<,F)$ is a countable 1-transitive coloured linear order, then the union of two clumps that intersect is a clump, and the union of any chain of clumps is a clump.*

*Proof.* Let $C_1$ and $C_2$ be clumps having non-trivial intersection, and choose convex sets $I_1$ and $I_2$ whose points are coloured by $C_1$, $C_2$ respectively. By 1-transitivity, we may suppose that $I_1 \cap I_2 \neq \emptyset$. Then $I_1 \cup I_2$ is convex, and is coloured by $C_1 \cup C_2$, so $C_1 \cup C_2$ is a clump.

Now let $\mathcal{C}$ be a chain of clumps. As $C$ is countable, so is $\bigcup \mathcal{C}$, so we may let $\bigcup \mathcal{C} = \{c_n : n \in \omega\}$. (If $\bigcup \mathcal{C}$ is finite, the result is immediate.) Choose $C_n \in \mathcal{C}$ by induction. Let $C_0 \in \mathcal{C}$ be arbitrary. If $C_n$ has been chosen, let $C_{n+1} \supseteq C_n$ be a member of $\mathcal{C}$ containing $c_n$. This gives a chain $C_0 \subseteq C_1 \subseteq C_2 \subseteq \cdots$ in $\mathcal{C}$ with union equal to $\bigcup \mathcal{C}$. Choose inductively a convex set $I_n$ such that $C_n = \{F(z) : z \in I_n\}$. Let $I_0$ be convex so that $C_0 = \{F(z) : z \in I_0\}$. Given $I_n$ coloured by $C_n$, by Lemma 4.2 there is a convex set $I_{n+1} \supseteq I_n$ such that $C_{n+1} = \{F(z) : z \in I_{n+1}\}$. Let $I = \bigcup_{n\in\omega} I_n$. It is easy to check that $\bigcup \mathcal{C} = \bigcup_{n\in\omega} C_n = \{F(x) : x \in I\}$, so it follows that $\bigcup \mathcal{C}$ is a clump. ∎

We remark in passing that the intersection of two clumps $C_1$ and $C_2$ of $X$, even when it is non-empty, is not always a clump. For instance if $(X,<)$ is $\mathbb{Z}$, coloured periodically by the four colours $c_0, c_1, c_2, c_3$, then $\{c_0, c_1, c_2\}$ and $\{c_0, c_2, c_3\}$ are clumps, but their intersection $\{c_0, c_2\}$ is not. We shall also see below that the intersection of a chain of clumps need not be a clump.

DEFINITION 4.4. Let $(X,<,F)$ be a coloured linear order with colour set $C$. A *tree of clumps* for $X$ is a set of clumps $\tau$ of $X$ which is a tree when ordered by proper inclusion and which contains $C$ and each singleton.

LEMMA 4.5. *Any countable* 1*-transitive coloured linear order has a maximal tree of clumps.*

*Proof.* This is immediate by Zorn's Lemma, since being a tree is preserved under taking unions of chains. ∎

LEMMA 4.6. *Any maximal tree of clumps* $(\tau,\subset)$ *for a countable* 1*-transitive coloured linear order is closed under unions of chains.*

*Proof.* Let $\mathcal{C}$ be a chain of clumps in $\tau$. Then by Lemma 4.3, $\bigcup\mathcal{C}$ is a clump. We see that $\tau\cup\{\bigcup\mathcal{C}\}$ is a tree of clumps, from which, by maximality, $\bigcup\mathcal{C}\in\tau$ will follow. Let $C'\in\tau$, and let $C'\subseteq D_1, D_2\in\tau\cup\{\bigcup\mathcal{C}\}$. If $D_1, D_2\neq\bigcup\mathcal{C}$, then as $\tau$ is a tree, $D_1\subseteq D_2$ or $D_2\subseteq D_1$. Suppose therefore that $D_1=\bigcup\mathcal{C}$ and $D_2\neq\bigcup\mathcal{C}$. Since $C'\subseteq D_1, D_2$, there is $c\in D_1\cap D_2$. By definition of "tree of clumps", $\{c\}\in\tau$, so $\{c\}\subseteq D_1, D_2$. Since $D_1=\bigcup\mathcal{C}$, $c\in C''$ for some $C''\in\mathcal{C}$, and hence $c\in C'''$ for all $C'''\supseteq C''$ in $\mathcal{C}$. As $\{c\}\subseteq C''', D_2$, both of which lie in $\tau$, and as $\tau$ is a tree, $D_2\subseteq C'''$ or $C'''\subseteq D_2$. If the former holds for some $C'''$ then $D_2\subseteq\bigcup\mathcal{C}=D_1$. Otherwise the latter holds for all $C'''\supseteq C''$ in $\mathcal{C}$, so $D_1=\bigcup\mathcal{C}\subseteq D_2$. Thus $\tau\cup\{\bigcup\mathcal{C}\}$ is a tree, as required. ∎

What we would really like is for the maximal tree of clumps to be Dedekind–MacNeille complete. Unfortunately, because of the possible presence of "selections", this need not be true. If however we form the Dedekind–MacNeille completion of $\tau$, then at least, by the general theory of these completions (or directly) and by Lemma 4.6, one sees that this is obtained from $\tau$ just by adjoining ramification points, which we write as $\tau^+$. Thus its description in terms of $\tau$ is fairly explicit. In fact we may take the points of $\tau^+$ to be subsets of $C$, since any point of $\tau^+-\tau$ is the infimum of a descending chain of members of $\tau$, and we just represent it by the intersection of this chain. The fact that $\tau^+-\tau$ may be non-empty explains why the intersection of a chain of clumps is not necessarily a clump, since the intersection of the family of clumps greater than $t\in\tau^+-\tau$ is equal to $t$ itself.

LEMMA 4.7. *Let $(\tau, \subset)$ be a tree of clumps for a coloured linear order $(X, <, F)$. Then for all $x \in \tau^+$ such that $x$ is not a leaf*, $x = \bigcup\{y \in \tau^+ : y \subset x\}$.

*Proof.* Clearly $\bigcup\{y \in \tau^+ : y \subset x\} \subseteq x$. Now, let $c \in x$. By definition, $\{c\} \in \tau$. As $x$ is not a leaf, $\{c\} \subset x$, so $c \in \bigcup\{y \in \tau^+ : y \subset x\}$, and $x \subseteq \bigcup\{y \in \tau^+ : y \subset x\}$. ∎

LEMMA 4.8. *Let $(\tau, \subset)$ be a maximal tree of clumps for a coloured linear order $(X, <, F)$. If $x$ is a vertex in $\tau^+$ having no parent and $x \neq r$, then* $x = \bigcap\{y \in \tau^+ : x \subset y\}$.

*Proof.* That $x \subseteq \bigcap\{y \in \tau^+ : x \subset y\}$ is immediate. Now, let $c \in \bigcap\{y \in \tau^+ : x \subset y\}$. If $c \notin x$, then $\{c\} \not\subset x$. On the other hand, $\{c\} \subset y$ for all $y \supset x$. Since $\tau^+$ is Dedekind–MacNeille complete, $\sup\{\{c\}, x\} \in \tau^+$. Clearly $x \subset \sup\{\{c\}, x\}$. As $x$ has no parent, $x \subset y \subset \sup\{\{c\}, x\}$ for some $y$. But then $x \subset y$ and $\{c\} \subset y$, and so $\sup\{\{c\}, x\} \subseteq y$, contradiction. Hence $\bigcap\{y \in \tau^+ : x \subset y\} \subseteq x$. ∎

LEMMA 4.9. *Let $(\tau, \subset)$ be a maximal tree of clumps for a coloured linear order $(X, <, F)$. Then for all $x \in \tau^+$ such that $x$ is not a leaf, the union of any maximal antichain $A$ of the subtree $\{z \in \tau^+ : z \subset x\}$ is equal to $x$.*

*Proof.* Clearly $\bigcup A \subseteq x$. Now, let $c \in x$. Then, since $\tau$ is a tree of clumps, and $x$ is not a leaf, $\{c\} \subset x$. As $A$ is a maximal antichain of $\{z \in \tau^+ : z \subset x\}$, there must be $y \in A$ comparable with $\{c\}$. Since $\{c\}$ is minimal, $c \in y$, and so $c \in \bigcup A$. ∎

LEMMA 4.10. *Let $(\tau, \subset)$ be a maximal tree of clumps for a* 1*-transitive coloured linear order $(X, <, F)$, and let $t \in \tau^+ - \tau$. Then for some $n$, $t$ has children $t_0, t_1, \ldots, t_{n-1}$, $1 < n \leq \aleph_0$, which all lie in $\tau$, and every element below $t$ lies below some $t_i$. Furthermore, for any $u \succ t$ in $\tau$, and any maximal convex $u$-coloured subset $Y$ of $X$, the family of maximal $t_i$-coloured subsets of $Y$ for $i < n$ is order-isomorphic to $\mathbb{Q}_n$.*

*Proof.* By Lemma 4.3, the union of each cone of elements of $\tau$ below $t$ is a clump, and so this establishes the first part. Since $t \notin \tau$, $n \geq 2$.

Now let $\mathcal{Z}$ be the family of maximal convex $t_i$-coloured subsets of $Y$ as $i$ varies, with the corresponding colours. Let $i, j < n$ be two colours. Then $t$ is the least upper bound in $\tau^+$ of $t_i$ and $t_j$. First let us see that it cannot be the case that all members of $\mathcal{Z}$ coloured $i$ precede all members of $\mathcal{Z}$ coloured $j$. Suppose otherwise for a contradiction, and let $Z_i$, $Z_j$ be members of $\mathcal{Z}$ coloured $i$, $j$ respectively. Thus $Z_i < Z_j$. Let $Z$ be the convex hull of $Z_i \cup Z_j$ and let $v = F(Z)$. Then $v$ is a clump. We show that it is the least upper bound of $t_i$ and $t_j$ in $\tau$, which will contradict the fact that this least upper bound is actually $t$, which does not lie in $\tau$.

First note that $\tau \cup \{v\}$ is a tree. For suppose $u \subseteq v, v'$ where $u, v' \in \tau$. We have to show that $v \subseteq v'$ or $v' \subseteq v$. Suppose that $v' \not\subseteq v$. Then a maximal $v'$-coloured convex set $W$ that intersects $Z$ must extend to the left of $Z_i$ or to the right of $Z_j$, or both. Suppose the former. As $W \cap Z \neq \emptyset$, also $W \cap Z_i \neq \emptyset$, so $t_i \cap v' \neq \emptyset$. As $t_i, v' \in \tau$, and $\tau$ is a tree containing all singletons, $v' \subseteq t_i$ or $t_i \subset v'$. The former contradicts $v' \not\subseteq v$, so $t_i \subset v'$. Since $t_i$ is a maximal clump in $\tau$ disjoint from $t_j$, $v' \cap t_j \neq \emptyset$, and repeating the argument, $t_j \subset v'$. We can now show that $v \subseteq v'$. Let $c \in v$. If $c \in t_i \cup t_j$ then by what we have just shown we know that $c \in v'$. Otherwise $c = F(x)$ for some $x$ with $Z_i < x < Z_j$. Now as $W$ is maximal convex $v'$-coloured, and $t_i, t_j \subseteq v'$, $W$ has maximal convex $t_i$, $t_j$-coloured subsets $Z_i'$, $Z_j'$ respectively. By Lemma 4.2, $Y$ is 1-transitive, and this enables us to assume that $Z_i' = Z_i$. If $x < Z_j'$ then it follows that $x \in W$ so $c = F(x) \in v'$ as required. If not, then we must have $Z_j' < x < Z_j$. By 1-transitivity of $Y$ again, there is an automorphism $f$ taking $Z_j$ to $Z_j'$. This must fix $Y$, and since all members of $\mathcal{Z}$ coloured $i$ precede all members of $\mathcal{Z}$ coloured $j$, $Z_i < f(Z_j')$. Therefore $Z_i < f(Z_j') < f(x) < Z_j'$, and it follows that $f(x) \in W$, so $c = F(f(x)) \in v'$.

This concludes the proof that $\tau \cup \{v\}$ is a tree. As $\tau$ is a maximal tree of clumps, $v \in \tau$. To see that $v$ is the supremum of $t_i$ and $t_j$ in $\tau$, since $t_i, t_j \subseteq v$ is already known, we just need to suppose that $t_i, t_j \subseteq v' \in \tau$ and show that $v \subseteq v'$. However, the argument of the previous paragraph applies again, giving the result.

We are now able to deduce that for each $i < n$, $\mathcal{Z}$ has no least member coloured $i$, and it will follow in particular that it has no left endpoint (and by a similar argument that it has no right endpoint). Suppose on the contrary that $Z_i$ is such. Since $n \geq 2$, there is another colour $j$. Since not every member of $\mathcal{Z}$ coloured $i$ precedes every member of $\mathcal{Z}$ coloured $j$, there is a member $Z_j$ of $\mathcal{Z}$ coloured $j$ with $Z_j < Z_i$. Similarly there are $Z_i' < Z_j'$ coloured $i, j$ respectively the other way round. By 1-transitivity we may assume that $Z_j = Z_j'$, so that $Z_i' < Z_i$, contrary to $Z_i$ being the least member of $\mathcal{Z}$ coloured $i$.

Now we have to demonstrate density. Let $Z_i < Z_j$ in $\mathcal{Z}$ be coloured $i$ and $j$ (which may now be equal), and let $k < n$ be a given colour. First suppose that $k \neq i$. There must be some point between $Z_i$ and $Z_j$, as otherwise $i \neq j$ by maximality of $Z_i$ convex $t_i$-coloured, and so $t_i \cup t_j$ would be a clump, and therefore equal the least upper bound of $t_i$ and $t_j$ in $\tau$, contradiction. Now consider $i$ and $k$. Note that $Z_i$ is in fact a maximal convex subset of $Y$ which is coloured by a set disjoint from $t_k$, as otherwise there would be an element of the cone at $t$ containing $t_i$ strictly greater than $t_i$. Since by the first part of the proof we know that there are points both to the left and the right of $Z_i$ coloured $k$, this tells us that there are such points arbitrarily close to $Z_i$ on left and right, and in particular, to the right of $Z_i$ and the left of $Z_j$. This gives density if $k \neq i$. Now suppose that $k = i$, and let $l < n$ be different from $i$. Then there is $Z_l$ between $Z_i$ and $Z_j$, and, repeating, another $Z'_l$ coloured $l$ between $Z_i$ and $Z_l$; hence applying density to $Z'_l < Z_l$, there is $Z'_i$ coloured $i$ between these, and this also lies between $Z_i$ and $Z_j$. ∎

We shall work with some maximal tree of clumps $(\tau, \subset)$ for a countable 1-transitive coloured linear order $(X, <, F)$, and extend $\tau^+$ to a coding tree for $X$ by adding some extra vertices to form $\tau^*$ (the ones representing lexicographic products, that is, vertices having only one child), and assigning labels. For any clump $t$, we write $X_t$ for a maximal convex subset of $X$ coloured by $t$, as given by Lemma 4.2. Since $(X, <, F)$ is 1-transitive, by that lemma, $X_t$ is 1-transitive. To avoid confusion between vertices of the tree and sets of colours, we write $C_t$ for the set of colours that the clump $t \in \tau$ represents (even though $C_t = t$). We recall that in [1], the notation $\pi'$ for a partition $\pi$ of the colour set was used, and it stood for the least refinement of $\pi$ into clumps, using the present terminology. One or two results from [1] will also be needed in what follows.

We define $(\tau^*, \prec, \mathcal{L})$ from $(\tau^+, \subset)$ as follows. Let $v \in \tau^+$.

If $v$ is a leaf, then it represents a singleton clump $\{c\}$, so $X_v$ is a (monochromatic) countable 1-transitive linear order $Z$. If $Z \neq 1$, put a parent $v^*$ above $v$ (parent to no other vertex) in $\tau^*$, and let $\mathcal{L}(v^*) = (Z, \{c\})$, and $\mathcal{L}(v) = (1, \{c\})$. If $Z = 1$, we just let $\mathcal{L}(v) = (1, \{c\})$.

If $v$ has only one cone, then it lies in $\tau$, and has no child (because if it had any, then it could only be one, and this child would have to be a proper subset of $v$, but, by Lemma 4.9, the union of any maximal antichain below $v$ equals $v$). So we may choose a chain $v_0 \subset v_1 \subset \cdots \subset v_n \subset \cdots$ of clumps whose supremum is $v$. Choose a maximal convex subset $X_{v_n}$ of $X$ coloured by $v_n$ by induction. Let $X_{v_0} \subset X_v$ be arbitrary (since $v_0 \subset v$, there is some $X_{v_0}$ contained in $X_v$). Given $X_{v_n}$, by Lemma 4.2 there is a unique maximal convex subset $X_{v_{n+1}}$ of $X$ coloured by $v_{n+1}$ containing $X_{v_n}$. By Lemma 4.2 again, $\bigcup_{n\in\omega} X_{v_n} \subseteq X_v$.

If $X_v \neq \bigcup_{n\in\omega} X_{v_n}$, let $x \in X_v - \bigcup_{n\in\omega} X_{v_n}$. Then there is $x' \in \bigcup_{n\in\omega} X_{v_n}$ such that $F(x') = F(x)$. As $X_v$ is 1-transitive, there is an automorphism $g$ of $X_v$ taking $x'$ to $x$. Then the image $g \bigcup_{n\in\omega} X_{v_n}$ of $\bigcup_{n\in\omega} X_{v_n}$ under $g$ is isomorphic to $\bigcup_{n\in\omega} X_{v_n}$. If $\bigcup_{n\in\omega} X_{v_n} \cap g \bigcup_{n\in\omega} X_{v_n} \neq \emptyset$, then $X_{v_m} \cap gX_{v_m} \neq \emptyset$ for some $m \in \omega$ such that $x' \in X_{v_m}$. Since $X_{v_m}$ is a maximal convex subset coloured by $C_{v_m}$, $X_{v_m} = gX_{v_m}$, contradicting $x \notin \bigcup_{n\in\omega} X_{v_n}$. Hence, if $X_v \neq \bigcup_{n\in\omega} X_{v_n}$, then $X_v$ is the union of pairwise disjoint copies of $\bigcup_{n\in\omega} X_{v_n}$. As $X_v$ is 1-transitive, $X_v \cong Z \cdot \bigcup_{n\in\omega} X_{v_n}$ for some countable 1-transitive linear order $Z$ (which is 1 if $X_v = \bigcup_{n\in\omega} X_{v_n}$). If $Z$ is non-trivial, put a parent $v^*$ above $v$ (parent to no other vertex) in $\tau^*$, and let $\mathcal{L}(v^*) = (Z, C_v)$ and $\mathcal{L}(v) = (\lim, C_v)$. If $X_v = \bigcup_{n\in\omega} X_{v_n}$, we just let $\mathcal{L}(v) = (\lim, C_v)$.

If $v \in \tau^+ - \tau$ then we label $v$ by $\mathrm{select}_n$, where $n$ is its ramification order. Otherwise, suppose that $v \in \tau$ is a ramification point with $n$ cones (where $1 < n \leq \aleph_0$). We show that these cones have greatest elements. Let $A$ and $B$ be distinct cones at $v$. Since $\tau$ is a tree, every element of $A$ is disjoint from every element of $B$. By Lemma 4.6, $\bigcup A, \bigcup B \in \tau$, and as they are disjoint, they are both proper subsets of $v$. Hence $\bigcup A$ lies in $A$, so $A$ has a greatest element, and similarly for all other cones at $v$. Let $v_0, v_1, \ldots, v_{n-1}$ be the greatest elements of the cones at $v$. Then these lie in $\tau$, since any member of $\tau^+ - \tau$ is the infimum of an infinite descending sequence. By Lemma 4.2, $X_v$ is 1-transitive. If $\pi_v = \{C_{v_0}, C_{v_1}, \ldots, C_{v_{n-1}}\}$, then, by Lemma 4.9, $\pi_v$ is a partition of $C_v$. Since each $C_{v_i}$ is a clump, $X_{v_i}$ is defined for each $i$. By 1-transitivity of $X$, there is $X'_{v_i} \cong X_{v_i}$ with $X'_{v_i} \subset X_v$. Clearly $X'_{v_i}$ is convex and coloured exactly by $C_{v_i}$. Hence, $\pi_v = \pi'_v$, in the notation of [1]. We look at $X_v \,/\!\sim_{\pi'_v}$, which, by [1, Lemma 1.3], is 1-transitive and coloured by $\pi'_v$.

If $n = 2$ (that is, $v$ has only two cones), then $X_v \,/\!\sim_{\pi'_v}$ is a 2-coloured 1-transitive linear order. Hence, by [1, Theorem 2.1], $X_v \,/\!\sim_{\pi'_v}$ is isomorphic to $\mathbb{Q}_2$ or to $Z \cdot 2$, where $Z$ is a countable 1-transitive linear order.

If $X_v \,/\!\sim_{\pi'_v} \cong \mathbb{Q}_2$, let $\mathcal{L}(v) = (\mathbb{Q}_2, C_v)$, and choose an arbitrary bijection from 2 to the children of $v$.

If $X_v \,/\!\sim_{\pi'_v} \cong Z \cdot 2$, let the bijection take 0 to the left child of $v$, and 1 to its right child.

If $Z \neq 1$, then $Z$ is a non-trivial countable 1-transitive linear order. Put a parent $v^*$ above $v$ (parent to no other vertex) in $\tau^*$, and let $\mathcal{L}(v^*) = (Z, C_v)$, and $\mathcal{L}(v) = (2, C_v)$.

If $Z = 1$, we just let $\mathcal{L}(v) = (2, C_v)$.

If $n > 2$ (that is, $v$ has more than two cones), we prove that $X_v \,/\!\sim_{\pi'_v} \cong \mathbb{Q}_n$, where the $n$ colours are the $C_{v_i}$.

First to see that $X_v \,/\!\sim_{\pi'_v}$ is dense, suppose not for a contradiction. Then there are consecutive $x < y \in X_v \,/\!\sim_{\pi'_v}$. Now, $x$ and $y$ cannot be coloured by the same $C_{v_i}$, because if they were, then $x \sim_{\pi'_v} y$, and so they would have been identified in the quotient. Let $C_{v_i}$ be the colour of $x$, and $C_{v_j}$ $(\neq C_{v_i})$ be the colour of $y$. Let $I_x$ be the convex set (the $\sim_{\pi'_v}$-class) in $X_v$ represented by $x$, and let $I_y$ be the corresponding convex set for $y$. Since $x$ and $y$ are adjacent, $I_x \cup I_y$ is convex and is coloured exactly by $C_{v_i} \cup C_{v_j}$. Hence, $\tau' = \tau \cup \{v_i \cup v_j\}$ is a tree of clumps for $X$, contrary to the maximality of $\tau$. Thus, $X_v \,/\!\sim_{\pi'_v}$ is dense.

Secondly, let $C_{v_j}$ be a colour of $X_v \,/\!\sim_{\pi'_v}$, and let $x < y \in X_v \,/\!\sim_{\pi'_v}$. Let $u, w \in X_v \,/\!\sim_{\pi'_v}$ with $x < u < w < y$, and suppose that for each $z \in [u, w] \subset X_v \,/\!\sim_{\pi'_v}$, $z$ is not coloured by $C_{v_j}$. Let $C_{[u,w]}$ be the subset of $\pi'_v$ colouring the members of $[u, w]$. Now, there are $C_{v_l}, C_{v_m} \in C_{[u,w]}$ such that $C_{v_l} \neq C_{v_m}$ (as otherwise $u \sim_{\pi'_v} w$ in $X_v$, and $u$ and $w$ would have been identified in the quotient), and also $C_{v_l} \neq C_{v_j} \neq C_{v_m}$. But by definition

$C_{[u,w]}$ is a clump in $X_v/{\sim_{\pi'_v}}$. Hence $\{\bigcup C_{v_i} : C_{v_i} \in C_{[u,w]}\}$ is a clump in $X_v$, and so $\tau' = \tau \cup \{\bigcup C_{v_i} : C_{v_i} \in C_{[u,w]}\}$ is a tree of clumps, contrary to the maximality of $\tau$. It follows that there is $z \in [u, w]$ coloured by $C_{v_j}$.

Since $X_v/{\sim_{\pi'_v}}$ is dense and 1-transitive, it has no endpoints, so $X_v/{\sim_{\pi'_v}} \cong \mathbb{Q}_n$, and we let $\mathcal{L}(v) = (\mathbb{Q}_n, C_v)$, with an arbitrary bijection from $n$ to $\{v_0, \ldots, v_{n-1}\}$.

This completes the definition of $(\tau^*, \prec, \mathcal{L})$. It remains to show that it is a coding tree, and that it encodes the original coloured linear order $X$. We remark that in $\tau^*$, at most two elements are concatenated at a time, since we took a maximal tree of clumps in which all possible concatenations (of colours which could be grouped together as clumps) had been performed. This means that the coding tree may not be as economical as it could be. It would be possible to coalesce consecutive concatenations, but this would add technical details which we prefer to avoid.

Lemma 4.11. *Let $(X, <, F)$ be a countable 1-transitive linear order coloured by $C$. Then $(\tau^*, \prec, \mathcal{L})$ (as defined above) is a coding tree.*

*Proof.* We go through the clauses in Definition 2.1.

(i) There is a root $r$ in $\tau$, namely $C$, which is also the root of $\tau^+$. This is also the root of $\tau^*$, unless we added a vertex above the root of $\tau^+$ in forming $\tau^*$, in which case this added vertex is the root.

(ii) There are at most $\aleph_0$ leaves in $\tau$, since $X$, and hence $C$, is countable, and any leaf of $\tau^*$ is also a leaf of $\tau$.

(iii) Every vertex $v$ in $\tau$ is a non-empty subset of $C$, so there is $c \in C$ with $\{c\} \subseteq v$. Since in $\tau^*$ the vertices that may have been added are all above vertices in $\tau$, every vertex in $\tau^*$ is above a leaf.

(iv) We know that $\tau^+$ is Dedekind–MacNeille complete. Note that the extra vertices which may have been added to form $\tau^*$ are all immediately above vertices in $\tau^+$. Hence, all the ramification points of $\tau^*$ are in $\tau^+$. Also, there are no new points which could form an upper bounded sequence with no supremum. This is because if new points occur unboundedly in this sequence, their supremum will equal the supremum of their set of children. Hence $\tau^*$ is Dedekind–MacNeille complete.

(v) We remarked in the above proof that in each of the cases $v \in \tau^+ - \tau$, $v \in \tau$, any cone at a ramification point $v$ has a greatest element.

(vi) Take $t, u, v \in \tau^*$ such that $t \prec u \prec v$. We show that there is $w \in \tau^*$ with $t \preceq w \prec v$ such that $w$ has a sibling.

Let $t^-$ be the unique child of $t$, if it exists, and $t^- = t$ otherwise; similarly for $u^-$, $v^-$. First we show that there are $t', v'$ such that $t' \in \tau$, $v' \in \tau^+$, and $t^- \preceq t' \prec v' \preceq v$. If $t \in \tau^* - \tau^+$ we let $t' = t^-$. If $t \in \tau^+ - \tau$ then $t$ is the infimum of a descending sequence of vertices in $\tau$, so we may let $t'$ be one of these which is strictly below $u^-$. And if $t \in \tau$, we let $t' = t$. If $u \in \tau^+$ we let

$v' = u$. Otherwise $u^- \prec u$. If $t' \prec u^-$ we let $v' = u^-$. If not, then $t = t' = u^-$ and $u \prec v^-$, so we may let $v' = v^-$.

Now with $t', v'$ chosen, $t' \subset v'$ (as sets of colours). Pick $c \in v' - t'$, and let $w$ be a maximal clump contained in $v'$ and containing $t'$ but not $c$ (which exists by Lemma 4.6), and let $w'$ be a maximal clump contained in $v'$ and containing $c$ but disjoint from $w$ (using Lemma 4.6 again). Then $w, w' \subset v'$, $w \cap w' = \emptyset$, and $\sup(w, w')$ is a ramification point above $t$ and below $v$. Furthermore, each of $w$, $w'$ is the supremum of a cone below $\sup(w, w')$, so $w$ has $w'$ as a sibling.

Finally, observe that if $w = t^-$, then as $w$ has a sibling, we cannot have $t^- \prec t$, which would mean that $w$ was $t$'s only child. Hence $t \preceq w$.

(vii) Suppose for a contradiction that there are uncountably many vertices in $\tau^*$ with only one child. These vertices represent uncountably many subsets of $X$ isomorphic to a lexicographic product of a 1-transitive countable linear order with a coloured linear order. We shall show that we can choose distinct elements of uncountably many of these, contrary to countability of $X$.

Now any vertex of $\tau^*$ lies above a leaf, and there are only countably many leaves. Hence, there is a leaf $l = \{c\}$ say, such that uncountably many points with only one child are above $l$. Since $\tau^*$ has only countably many ramification points, there is an uncountable set $U$ of vertices above $l$ with just one child labelled lim. Choose a point $x$ coloured $c$. Then by Lemma 4.2, for each $t \in \tau$ above $l$ not in $U$ or labelled lim there is a unique maximal convex $t$-coloured set $X_t$ with $x \in X_t$. We may extend this choice to points $t$ labelled lim by letting $X_t = \bigcup\{X_u : u \prec t,\ X_u \text{ defined}\}$, and to points $t \in \tau^+$ by letting $X_t = \bigcap\{X_u : t \prec u,\ u \in \tau\}$. Clearly $t_1 \prec t_2 \Leftrightarrow X_{t_1} \subset X_{t_2}$. Each point of $U$ is the parent $t^+$ of a unique $t$ for which $X_t$ has been chosen, and by choice of the labels, the unique maximal convex $t$-coloured set $X_{t^+}$ with $x \in X_t$ is a lexicographic product of the form $Z \cdot X_t$, where $Z$ is a non-trivial 1-transitive linear order (equal to $\mathcal{F}(t^+)$). Choose an element $x_t$ of $X_{t^+} - X_t$. Then all these choices are distinct, so $\{x_t : t \in U\}$ is an uncountable subset of $X$, contradiction.

(viii) We have defined a labelling function $\mathcal{L}$ which assigns to every vertex $v \in \tau^*$ a label $(\mathcal{F}(v), \mathcal{S}(v))$, and bijections in the relevant cases. We first check that $\tau^*$ obeys the conditions for $\mathcal{F}(t)$.

If $t$ ramifies into $n$ cones, then $t \in \tau^+$. If $t \in \tau^+ - \tau$ then $t$ is labelled select$_n$, and $t$ is the infimum of a descending sequence of members of $\tau$. We need to see that we can take the members of this sequence to be labelled just by $\mathbb{Q}_m$ or $Z$. If not, there is $u \succ t$ such that no member of $(t, u]$ is so labelled. Choose a maximal $u$-coloured set $Y$. By Lemma 4.10 there are maximal $t_0$-, $t_1$-coloured subsets $X_0 < X_1$ of $Y$, where $t_0$ and $t_1$ are distinct children of $t$. Let $X$ be the convex hull of $X_0 \cup X_1$. By Dedekind–MacNeille completeness of $\tau^+$, there is a minimal member $v$ of $\tau^+$ containing $F(X)$. If $v$ is labelled select$_m$ then some child of $v$ contains $F(X)$, contrary to minimality. Hence $v$ is labelled lim or $\gamma$.

If $v$ is labelled lim, let $V \supseteq X$ be maximal convex $v$-coloured ($\subseteq Y$). Suppose that $v_0 \subset v_1 \subset v_2 \subset \cdots$ in $\tau$ with $v = \bigcup_{i\in\omega} v_i$, and choose maximal convex $v_i$-coloured $V_0 \subset V_1 \subset V_2 \subset \cdots \subset V$. Since $v$ is labelled lim, $V = \bigcup_{i\in\omega} V_i$. Hence for some $i$, $V_i \cap X_0 \neq \emptyset$, and hence also $v_i \cap t_0 \neq \emptyset$. Since $\tau$ is a tree, $v_i \subseteq t_0$ or $t_0 \subseteq v_i$. If $v_i \subseteq t_0$ for every $i$ then $v \subseteq t_0$, contradiction. Hence $t_0 \subseteq v_i$. As $V_i \cap X_0 \neq \emptyset$, $X_0 \subseteq V_i$. Similarly $X_1 \subseteq V_j$ for some $j$. As $V_{\max(i,j)}$ is convex, $X \subseteq V_{\max(i,j)}$. By minimality of $v$, $v \subseteq v_{\max(i,j)}$—contradiction.

The remaining possibility is that $v$ is labelled $\gamma$, which must actually be 2 by the above proof. Suppose that $V = V_0 \cup V_1$, where $V_0$, $V_1$ are nonempty, disjoint, and $V_0 < V_1$. By minimality of $v$, $F(V_0), F(V_1) \not\supseteq F(X)$, so $V_0, V_1 \not\supseteq X$. Therefore $\sup V_0 < \sup X_1$. By Lemma 4.10 there are intervals contained in $Y$ coloured $t_0$ and $t_1$ to the right of $X_1$, and hence contained in $V_1$. Similarly there are such intervals contained in $V_0$. This gives $F(V_0) \cap F(V_1) \neq \emptyset$, contrary to $V$ being the concatenation of $V_0$ and $V_1$ (which requires the concatenated orders to be disjointly coloured).

Otherwise, $t \in \tau$. If $n = 2$, then, by the construction of $\tau$, $\mathcal{F}(v)$ is $\mathbb{Q}_2$ or 2 (and 2 is clearly the order type of the children of $t$), and if $n > 2$, then $\mathcal{F}(t) = \mathbb{Q}_n$. Also, if $t$ has a parent and no sibling, then this parent is in $\tau^*$, but not in $\tau$, and so $\mathcal{F}(t)$ has to be 2.

If there is only one cone at $t$, and $t$ is not a leaf, then

- if $t$ has a child (it can only have one), then $t \in \tau^* - \tau$, and all such were labelled by a 1-transitive countable linear order,
- if $t$ does not have a child, then, by construction of $\tau^*$, $\mathcal{F}(t) = \lim$.

If $t$ is a leaf, then, by construction of $\tau^*$, $\mathcal{F}(t) = 1$.

Finally, we see that $\tau^*$ obeys the conditions for $\mathcal{S}(t)$. If $t \in \tau^+$ and $t$ is not a leaf, then, by Lemma 4.7, $t = \bigcup\{u \in \tau : u \subset t\}$. Now, in $\tau^*$, if $x \in \tau$, $\mathcal{S}(x)$ is the clump $x$ stands for, and if $x \notin \tau$, $\mathcal{S}(x)$ is the clump its child stands for. Hence, if $t$ is not a leaf, $\mathcal{S}(t)$ is the union of the second labels of vertices under $t$. Also, leaves in $\tau^*$ are the same as leaves in $\tau$, and so all leaves in $\tau^*$ have distinct singleton colours as their second labels.

Thus, $(\tau^*, \prec, \mathcal{L})$ is a coding tree. ■

Now we need to prove that this coding tree actually encodes the linear order we started with.

THEOREM 4.12. *Let $(X, <, F)$ be a 1-transitive countable coloured linear order. Then the coding tree $(\tau^*, \prec, \mathcal{L})$ as defined above encodes $(X, <, F)$.*

*Proof.* We form an expanded coding tree $E^*$ associated with $\tau^*$. To begin with we take as vertices all subsets of $X$ which are maximal convex subject to being coloured by some fixed member of $\tau$. To see that this family $E$ is a tree, note that certainly $X \in E$ is its root. Otherwise, if $Y_1 \subseteq Y_2, Y_3$ in $E$ are maximal convex sets for $t_1, t_2, t_3$ respectively, then $t_1 \subseteq t_2, t_3$, and since $\tau$ is a tree, $t_2 \subseteq t_3$ or $t_3 \subseteq t_2$, from which $Y_2 \subseteq Y_3$ or $Y_3 \subseteq Y_2$ follows by Lemma 4.2.

Next we extend to $E^+$ by adding in points corresponding to elements in $\tau^+ - \tau$. Each of these will actually equal one of its children as far as the corresponding subset of $X$ is concerned, but it will be distinguished by means of its labels. To be more precise, for each point $t$ of $\tau^+ - \tau$ with children $t_0, t_1, \ldots, t_{n-1}$ and subset $Y_i$ corresponding to some $t_i$, we add another vertex to $E$ above $Y_i$.

Finally, we extend to $E^*$ and assign labels as follows.

If $t \in \tau^+ - \tau$ has ramification order $n$, we label each corresponding point $Y$ in $E^+ - E$ by select$_n$, and let $\mathcal{S}(Y) = \mathcal{S}(t)$.

If $Y$ is a singleton, then we label it by $(1, \{c\})$, where $\{c\} = F(Y)$. If $Y$ is monochromatic but not a singleton, it is a non-trivial countable 1-transitive linear order, and we replace it by an upper vertex labelled $(Y, \{c\})$, and for each $y \in Y$ a lower point $y$, labelled $(1, \{c\})$.

If $Y$ has just one cone and corresponds to a member of $\tau$, then we choose an increasing chain $Y_0 \subset Y_1 \subset Y_2 \subset \cdots$ in $E$ such that $F(Y) = \bigcup_{n\in\omega} F(Y_n)$. Then either $Y = \bigcup_{n\in\omega} Y_n$, in which case we label $Y$ by $(\lim, F(Y))$, or $Y \cong Z \cdot \bigcup_{n\in\omega} Y_n$ for some non-trivial countable 1-transitive linear order $Z$, and we replace $Y$ by an upper point $Y$, labelled $(Z, F(Y))$, and for each $z \in Z$ a lower point being the corresponding copy of $\bigcup_{n\in\omega} Y_n$, labelled $(\lim, F(Y))$.

If $Y$ has ramification order greater than 1, then as before, $Y$ is either a lexicographic product of the form $Z \cdot (Y_0{}^\wedge Y_1)$ $(n = 2)$, where $Z$ is countable and 1-transitive, or is a $\mathbb{Q}_n$-combination of its children. In the former case we label $Y$ by $(2, F(Y))$ if $Z = 1$, and replace $Y$ by an upper vertex, being $Y$ labelled $(Z, F(Y))$, and for each $z \in Z$ a lower vertex which is the corresponding copy of $Y_0{}^\wedge Y_1$ labelled $(2, F(Y))$ if $Z \neq 1$. In the latter case we label $Y$ by $(\mathbb{Q}_n, F(Y))$.

Thus in all cases the elements of $E^*$ chosen are subsets of $X$, and the partial ordering is compatible with inclusion in the sense that $Y_0 \preceq Y_1 \Leftrightarrow Y_0 \subseteq Y_1$. Furthermore, except for the case of vertices labelled select$_n$, $Y_0 \prec Y_1 \Leftrightarrow Y_0 \subset Y_1$.

To conclude the verification that $E^*$ is an expanded coding tree, we just have to show that clauses (iv) and (vi) hold (since (v) is an immediate consequence of the corresponding property of $\tau$). We first see that $E$ is Dedekind–MacNeille complete, from which Dedekind–MacNeille completeness of $E^*$ is immediate. Let $\mathcal{C}$ be a maximal chain in $E$. Then $\{F(Y) : Y \in \mathcal{C}\}$ is a maximal chain in $\tau$, so this is Dedekind complete, and isomorphic to $\mathcal{C}$, so $\mathcal{C}$

is also Dedekind complete. Next consider $X_0, X_1 \in E$. Let $X$ be the convex hull of $X_0 \cup X_1$, and as above find a minimal member $v$ of $\tau^+$ containing $F(X)$. Minimality implies that $v$ is not labelled select$_n$, so $v \in \tau$. Let $Y \supseteq X$ be maximal convex $v$-coloured. We see that $Y$ is the least upper bound of $X_0$ and $X_1$ in $E$. Suppose that $Z \supseteq X_0, X_1$, and let $t = F(Z)$. Since $Z$ is convex, $Z \supseteq X$. As $E$ is a tree, $Y \subseteq Z$ or $Z \subset Y$. The latter would give $t \subset v$, contrary to minimality of $v$. Hence $Y \subseteq Z$ as required. This shows that $E$ and hence also $E^*$ is Dedekind–MacNeille complete.

For (v), suppose that $X \prec Y \prec Z$ in $E^*$. If $Y$ or $Z$ is labelled select$_n$, we may take $W$ to equal one of them so labelled. Otherwise, $X \subset Y \subset Z$. If either $Y$ or $Z$ is labelled lim, then we may replace them by smaller sets in $E$ not labelled lim. The result now follows, since we cannot have consecutive lexicographic products.

Now that $E^*$ has been defined, we just need to show that it is associated with $\tau^*$, and that the set of leaves of $E^*$ is isomorphic to $X$. The latter is immediate, since all members of $X$ feature as singletons of $E^*$, and the ordering and colours are correct.

To associate $\tau^*$ with $E^*$ we give the map $\varphi$. Since $E^*$ was defined directly from $\tau^*$, it is clear how this should be defined. In the most straightforward case, $t \in \tau$ gives rise to just one vertex in $t^*$ (that is, where the set corresponding to the clump is not a non-trivial lexicographic product), in which case every maximal $t$-coloured subset of $X$ in $E$ is mapped to $t$ by $\varphi$. For elements $t$ of $\tau^+ - \tau$ we introduced points of $E^*$ corresponding to exactly the same subset of $X$ as the child, but with a different label, and each such is mapped to $t$. Finally, points $t^+$ of $\tau^* - \tau^+$ correspond to lexicographic products. In $\tau^*$ points come in pairs, $t$ and $t^+$ such that $t$ is the only child of $t^+$, and the corresponding members of $E^*$ also come in pairs such that the upper one is the lexicographic product by some $Z$ of the lower one. We map each such upper point to $t^+$, and the corresponding lower point to $t$.

It remains to verify the clauses of Definition 3.2 for this $\varphi$. The first (order-preserving) is immediate, as is (iii) (preservation of labels).

For (ii), consider a vertex $t$ of $E$ not labelled select$_n$, and let $u \prec \varphi(t)$ in $\tau^*$. We treat various cases.

First suppose that $u, \varphi(t)$ are clumps replaced by a single point in $\tau^*$. Then $t$ is maximal $\varphi(t)$-coloured. Picking $c \in u$, there is $x \in t$ coloured $c$, so by Lemma 4.2 there is a maximal $u$-coloured set containing $x$, and this is contained in $t$ and maps to $u$ under $\varphi$. This argument is easily modified if $u$ and/or $\varphi(t)$ are in a pair of points of $\tau^*$ corresponding to a single clump. By assumption $\varphi(t) \notin \tau^+ - \tau$. If $u \in \tau^+ - \tau$ then a preimage of $u$ under $\varphi$ was specifically included in $E^*$ below $t$.

Now let $l$ be a leaf of $E$, and let $\varphi(l) \preceq u$. Thus $u$ is a convex subset of $X$, and if $t = F(u)$, then $t$ lies in $\tau$ with $l \subseteq t$. If $u$ is modified (by addition of a parent or child) in passing to $\tau^*$, then we can similarly modify $t$ so that $\varphi(t) = u$, and it follows that $\varphi$ maps $[l, r]$ onto $[\varphi(l), \varphi(r)]$.

Finally, property (iv) follows by Lemma 4.10. ∎

**References**


[1] G. Campero-Arena and J. K. Truss, *1-transitive, coloured, countable linear orderings I*, J. Combin. Theory Ser. A 105 (2004), 1–13.

[2] G. L. Cherlin, *The classification of countable homogeneous directed graphs and countable homogeneous n-tournaments*, Mem. Amer. Math. Soc. 131 (1998), no. 621.

[3] A. C. Morel, *A class of relation types isomorphic to the ordinals*, Michigan Math. J. 12 (1965), 317–321.

[4] J. K. Truss, *Infinite permutation groups; subgroups of small index*, J. Algebra 120 (1989), 494–515.

[5] —, *Betweenness relations and cycle-free partial orders*, Math. Proc. Cambridge Philos. Soc. 119 (1996), 631–643.

[6] R. Warren, *The structure of k-CS-transitive cycle-free partial orders*, Mem. Amer. Math. Soc. 129 (1997), no. 614.



Departamento de Matemáticas
Facultad de Ciencias
Universidad Nacional Autónoma de México
Ciudad Universitaria
México, D.F. 04510, Mexico
E-mail: gabriela@lya.fciencias.unam.mx

Department of Pure Mathematics
University of Leeds
Leeds LS2 9JT, England
E-mail: j.k.truss@leeds.ac.uk



*Received 17 July 2003;*
*in revised form 16 June 2004 and 7 January 2005*