Mathematical Structures in Computer Science (2020), **30**, pp. 952–975
doi:10.1017/S0960129520000298

CAMBRIDGE UNIVERSITY PRESS

PAPER

From Kruskal's theorem to Friedman's gap condition

Anton Freund*

Fachbereich Mathematik, Technische Universität Darmstadt, Schlossgartenstr. 7, 64289 Darmstadt, Germany
*Corresponding author. Email: freund@mathematik.tu-darmstadt.de

(Received 5 March 2020; revised 20 November 2020; accepted 17 December 2020)

Abstract

Harvey Friedman's gap condition on embeddings of finite labelled trees plays an important role in combinatorics (proof of the graph minor theorem) and mathematical logic (strong independence results). In the present paper we show that the gap condition can be reconstructed from a small number of well-motivated building blocks: It arises via iterated applications of a uniform Kruskal theorem.

Keywords: Kruskal's theorem; Friedman's gap condition; labelled trees; well partial orders; dilators; Π^1_1-comprehension

1. Introduction

In this section, we explain our main result, discuss its relevance for mathematical logic and theoretical computer science and point out connections with previous work (in particular with the closely related approach of Hasegawa 1997, 2002). Our focus will be on motivation and intuitive explanation, with technical details deferred to the following sections. Nevertheless, we start by fixing a few fundamental notions: For the purpose of our paper, a tree is a finite partial order $T = (T, \leq_T)$ such that

- the order T has a unique minimal element $\langle\rangle$, called the root of T and
- for each $t \in T$, the set $\{s \in T \mid s \leq_T t\}$ is linearly ordered by $\leq_T$.

For each pair of elements $s, t \in T$, there is a $\leq_T$-maximal element $s \wedge t \in T$ with $s \wedge t \leq_T s$ and $s \wedge t \leq_T t$. An embedding of trees is given by an injective function $f : S \to T$ that satisfies

$$f(s \wedge t) = f(s) \wedge f(t)$$

for all $s, t \in S$. Since $s \leq_S t$ is equivalent to $s \wedge t = s$, this entails that f is an embedding of partial orders. Kruskal's theorem (1960) asserts the following: For any infinite sequence $T_0, T_1, \ldots$ of finite trees, there are indices $i < j$ such that T_i can be embedded into T_j.

Let us point out that Kruskal's theorem has important implications for theoretical computer science (cf. the work of Dershowitz 1982) and mathematical logic. Concerning the latter, a classical result of Schmidt (2020) and Friedman (presented by Simpson 1985) shows that Kruskal's theorem cannot be proved in $\mathbf{ATR_0}$, a relatively strong axiom system that is associated with the predicative foundations of mathematics (see the detailed explanations given by Gallier 1991 and Simpson 2009). The precise logical strength of Kruskal's theorem has been determined by Rathjen and Weiermann (1993).

By an n-tree, we mean a tree T together with a function $l : T \to \{0, \ldots, n-1\}$. An embedding between n-trees (S, l) and (T, l') is given by an embedding $f : S \to T$ of trees that satisfies the following conditions:

(i) We have $l'(f(s)) = l(s)$ for any $s \in S$.
(ii) If t is an immediate successor of $r \in S$ (i. e. if t is $\leq_S$-minimal with $r <_S t$) and we have $f(r) <_T s <_T f(t)$, then we have $l'(s) \geq l'(f(t)) = l(t)$.
(iii) If we have $s <_T f(\langle\rangle)$, then we have $l'(s) \geq l'(f(\langle\rangle)) = l(\langle\rangle)$.

Parts (ii) and (iii) constitute the famous gap condition due to Harvey Friedman (see Simpson 1985). More precisely, part (ii) on its own is known as the weak gap condition. In the present paper, we are only concerned with the strong gap condition, which is the conjunction of (ii) and (iii). The following result is known as Friedman's theorem, or also as the extended Kruskal theorem: For each number n and any infinite sequence $T_0, T_1, \ldots$ of finite n-trees, there is an embedding $T_i \to T_j$ of n-trees for some indices $i < j$.

Friedman's theorem plays a role in Robertson and Seymour's proof of their famous graph minor theorem. In fact, Friedman et al. (1987) have shown that Friedman's theorem is equivalent to the graph minor theorem for graphs of bounded treewidth, over the weak base theory $\mathbf{RCA_0}$. From the viewpoint of mathematical logic it is very significant that Friedman's theorem is unprovable in the axiom system Π^1_1-$\mathbf{CA_0}$, which is considerably stronger than $\mathbf{ATR_0}$.

The present paper shows that Friedman's gap condition results from iterated applications of a uniform Kruskal theorem. Our next aim is to explain the latter. Given a partial order X, we temporarily write $W(X)$ for the set of finite multisets with elements from X (later, the letter W will be used to denote general transformations of partial orders). Such a multiset can be written as $[x_0, \ldots, x_{n-1}]$, where the multiplicity of the entries is relevant but the order is not. To define a partial order on $W(X)$, we declare that $[x_0, \ldots, x_{m-1}] \leq_{W(X)} [y_0, \ldots, y_{n-1}]$ holds if, and only if, there is an injection $h : \{0, \ldots, m-1\} \to \{0, \ldots, n-1\}$ such that $x_i \leq_X y_{h(i)}$ holds for all $i < m$ (note the connection with Higman's lemma). Write $\mathcal{T}W$ for the set of trees, where isomorphic trees are identified. We get a bijection

$$\kappa : W(\mathcal{T}W) \to \mathcal{T}W$$

if we define $\kappa([T_0, \ldots, T_{n-1}])$ as the tree in which the root has immediate subtrees $T_0, \ldots, T_{n-1}$. Indeed, the set $\mathcal{T}W$ can be characterised as the initial fixed point of the transformation W. For $S, T \in \mathcal{T}W$ we write $S \leq_{\mathcal{T}W} T$ if there is an embedding $S \to T$ of trees. This relation can also be reconstructed in terms of the order on multisets: Writing $[X]^{<\omega}$ for the set of finite subsets of X, we define a family of functions $\operatorname{supp}^W_X : W(X) \to [X]^{<\omega}$ by setting

$$\operatorname{supp}^W_X([x_0, \ldots, x_{n-1}]) = \{x_0, \ldots, x_{n-1}\}.$$

For multisets σ and τ in $W(\mathcal{T}W)$ one can verify

$$\kappa(\sigma) \leq_{\mathcal{T}W} \kappa(\tau) \Leftrightarrow (\sigma \leq_{W(\mathcal{T}W)} \tau \text{ or } \kappa(\sigma) \leq_{\mathcal{T}W} T \text{ for some } T \in \operatorname{supp}^W_{\mathcal{T}W}(\tau)). \tag{$\star$}$$

Indeed, the first disjunct on the right corresponds to an embedding $\kappa(\sigma) \to \kappa(\tau)$ that maps the root to the root, and immediate subtrees to immediate subtrees. The second disjunct corresponds to an embedding that maps all of $\kappa(\sigma)$ into one immediate subtree of $\kappa(\tau)$.

Above, we have written $W(X)$ for the set of finite multisets with elements from X. In order to generalise the construction, we now allow W to range over a large collection of order transformations: A PO-dilator is a particularly uniform transformation W of partial orders that comes with a family of 'support' functions $\operatorname{supp}^W_X : W(X) \to [X]^{<\omega}$. In Section 2, we will recall the precise definition, as well as a normality condition for PO-dilators. For any normal PO-dilator W, one can construct a 'Kruskal fixed point' $\mathcal{T}W$ that is partially ordered according to ($\star$). Recall that a partial order X is a well partial order if any infinite sequence $x_0, x_1, \ldots$ in X admits indices $i < j$ with $x_i \leq_X x_j$. A PO-dilator W is called a WPO-dilator if $W(X)$ is a well partial order whenever the same holds for X. The uniform Kruskal theorem asserts that $\mathcal{T}W$ is a well partial order for any normal WPO-dilator W. By the previous paragraph, the usual Kruskal theorem arises as a special case. It is instructive to check that Higman's lemma is another special case (take $W(X) = 1 + Z \times X$ to generate finite lists with entries in Z). As shown by Freund et al. (2020), the uniform Kruskal theorem is equivalent to Π^1_1-comprehension (the main axiom of Π^1_1-$\mathbf{CA_0}$), over $\mathbf{RCA_0}$ together with the chain antichain principle. This result builds on a corresponding equivalence in the context of linear orders, which is due to the present author (see Freund 2019a,b, 2020).

In this paper, we show how the construction of $\mathcal{T}W$ can be relativised to a given partial order X. The result is a partial order $\mathcal{T}W(X)$ with a bijection

$$X \sqcup W(\mathcal{T}W(X)) \to \mathcal{T}W(X).$$

The point of the relativisation is that $\mathcal{T}W$ is now a transformation of partial orders (rather than a single order, as in the paper by Freund et al. 2020). We will show that $\mathcal{T}W$ can itself be equipped with the structure of a normal PO-dilator, which we call the Kruskal derivative of W. The axiom of Π^1_1-comprehension is still equivalent to the principle that $\mathcal{T}W$ is a normal WPO-dilator whenever the same holds for W. This principle will also be referred to as the uniform Kruskal theorem.

Our main aim is to reconstruct Friedman's gap condition by taking iterated Kruskal derivatives. From now on we write $M(X)$ for the set of multisets with elements from X (in the example above we have written this as $W(X)$, but W is now used for general PO-dilators). It is straightforward to equip M with the structure of a normal WPO-dilator. Our reconstruction of Friedman's gap condition proceeds via the following steps:

(1) Let $\mathbb{T}_0$ be the identity on partial orders, considered as a normal WPO-dilator.
(2) Assuming that the normal WPO-dilator $\mathbb{T}_n$ is already constructed, define the normal WPO-dilator $\mathbb{T}^-_{n+1}$ as the Kruskal derivative of $M \circ \mathbb{T}_n$.
(3) Define the normal WPO-dilator $\mathbb{T}_{n+1}$ as the composition $\mathbb{T}_n \circ \mathbb{T}^-_{n+1}$.
(4) Verify that $\mathbb{T}_n(\emptyset)$ is isomorphic to the set of n-trees, ordered according to Friedman's strong gap condition.

The construction described in steps (1)–(3) is justified by the results of Sections 2–4, which show that the recursive clauses can be satisfied by normal WPO-dilators $\mathbb{T}_n$ and determine them up to natural isomorphism. A detailed summary can be found at the end of Section 4. Step (4) is carried out in Sections 5 and 6. A summary is given at the end of Section 6.

Considering step (2) above, we see that n applications of the uniform Kruskal theorem are used to show that $\mathbb{T}_n$ is a WPO-dilator. If Π^1_2-induction is available, then one can conclude that the latter holds for all $n \in \mathbb{N}$. This accords with Friedman's result that the statement '$\mathbb{T}_n(\emptyset)$ is a well partial order for all $n \in \mathbb{N}$' can be proved in Π^1_1-$\mathbf{CA_0}$ plus Π^1_2-induction but not in Π^1_1-$\mathbf{CA_0}$ alone.

In order to explain the relevance of our result, we recall a traditional explanation of the gap condition, which is offered by Rathjen (1999): The latter writes that '[t]he gap condition imposed on the embeddings is directly related to an ordinal notation system that was used for the analysis of Π^1_1 comprehension.' In fact, he seems to suggest a clear order of precedence, by writing that 'the gap condition [...] is actually gleaned from the ordering of the terms in [one such notation system].' Furthermore, he states that '[i]t is also for that reason that criticism had been levelled against [Friedman's extended Kruskal theorem] for being too contrived or too metamathematical.'

Against this background, it may seem surprising that we can offer a completely different explanation (of a systematic rather than historic nature): The gap condition arises from iterated applications of the uniform Kruskal theorem. In the author's opinion, this explanation is decidedly mathematical, which rebuts the aforementioned criticism (note that Rathjen 1999 describes another rebuttal, which is related to the graph minor theorem). The picture becomes even more coherent if we bear in mind that the uniform Kruskal theorem is equivalent to

Π^1_1-comprehension (see Freund et al. 2020): It makes sense that we can get beyond Π^1_1-$\mathbf{CA_0}$ by suitable iterations. On the other hand, the author finds it astonishing that the iterations can take such a uniform and simple form.

The aforementioned order of precedence can also be reversed (again in a systematic rather than historic sense): Rather than explaining the gap condition in terms of ordinal notation systems, one can motivate ordinal notation systems by showing that they realise the maximal order types of certain partial orders with gap condition. Concrete results of this type have been established by van der Meeren et al. (see van der Meeren 2015 and van der Meeren et al. 2015, 2017a,b). Our result supports this approach on both a conceptual and a technical level: Conceptually, we provide an independent motivation for the gap condition. One of the main technical obstacles to more general results is notational complexity. We expect that our uniform description of the recursive construction will help to overcome this obstacle. Indeed, certain concrete cases considered by van der Meeren et al. (such as Lemma 6 of 2017a) have inspired our general approach.

Let us now discuss potential applications in computer science. The usual Kruskal theorem is an important tool in term rewriting. Specifically, it justifies termination proofs via the recursive path ordering of Dershowitz (1982). Despite the enormous success of this method, the use of Kruskal's theorem leads to a limitation (described by Tzameret 2002): Tree embeddings without the gap condition induce a subterm property (any term is bigger than its subterms) that is not compatible with all applications. The gap condition allows to transcend this limitation, since a small root label can prevent the embedding of a subtree. A concrete application of the gap condition in rewriting has been given by Ogawa (1995). As a second example, we mention that the gap condition is relevant for the priority channel systems of Haase et al. (2013). Further research is needed to assess whether our approach is fruitful for these types of application. Certainly, our construction is flexible enough to provide a variety of orders with gap condition.

Having discussed potential applications of our result to computer science, we would like to acknowledge influence in the other direction: Our approach to Kruskal fixed points owes a lot to categorical characterisations of recursive data types. The first categorical reconstruction of the gap condition by Hasegawa (1997, 2002) (cf. the detailed discussion below) has appeared in a computer science journal.

Finally, we discuss connections with the existing literature. To avoid misunderstanding, we first point out that our notion of Kruskal derivative is not related to differentiation in the sense of linear approximation (as in the work of Ehrhard and Regnier 2003). Instead, the term 'Kruskal derivative' is inspired by categorical work on derivatives of normal functions (see the paper by Freund and Rathjen 2021).

As mentioned above, it seems that the gap condition was originally inspired by ordinal notation systems. Nevertheless, Friedman's inductive proof that $X \mapsto \mathbb{T}_n(X)$ preserves well partial orders appears to contain traces of a recursive construction (see Simpson 1985 and the discussion around Definition 6.1 below). In the author's opinion, it is non-trivial and enlightening to bring out this construction in detail. In particular, our approach separates the recursive construction and the minimal bad sequence argument, which are closely entwined in the proof of Simpson's (1985) Lemma 4.7.

The first recursive construction of the gap condition in terms of initial fixed points is due to Hasegawa (1997, 2002) (as far as the present author is aware). In the following, we explain why our construction is sufficiently different to be of interest.

Most importantly, Hasegawa's construction transforms an order transformation into a single order (Kruskal fixed point) rather than another order transformation (Kruskal derivative). For this reason, Hasegawa needs to start with a transformation in n variables in order to iterate the construction n times. In our view, this makes the construction more complicated on a conceptual and technical level:

Conceptually, our approach has the advantage that the gap condition for $n+1$ labels (i. e. the PO-dilator $\mathbb{T}_{n+1}$) is derived from the gap condition for n labels, without reference to previous stages of the construction. We hope that this will allow us to iterate the construction into the transfinite, in order to capture generalisations of Friedman's theorem by Kříž (1989) and Gordeev (1990) (but this has not been achieved yet). A more palpable consequence is that our approach yields the gap condition in its original form with vertex labels, while Hasegawa obtains a version with edge labels.

On the technical side, Hasegawa's (2002) construction involves a somewhat intricate apparatus of iterated substitutions (see the passage after Definition 2.12 of the cited paper). We would also like to point out that Hasegawa's discussion of the gap condition itself is relatively condensed (just over one page, after Corollary 2.18). The present paper provides more details, which (in the author's opinion) is of value.

Finally, we mention that Hasegawa's setting is somewhat different from ours, in ways that we have not fully analysed. One noticeable difference is that Hasegawa's functors on partial orders are liftings of analytic functors on sets. If W is such a lifting and X is a partial order, the underlying set of $W(X)$ can only depend on the underlying set of X. In the case of our PO-dilators, the underlying set of $W(X)$ can also depend on the order on X (e. g. we could put $W(X)=\{(x,y)\in X^2 \mid x\not\leq_X y\}$). Another important difference is that Hasegawa's morphisms of partial orders are order preserving while ours are order reflecting. The use of order reflecting functions is crucial for the proof of Theorem 4.5 of Freund et al. (2020). This brings us back to a main motivation for our new reconstruction: We wanted a particularly coherent picture, which integrates the equivalence with Π^1_1-comprehension (established in the paper by Freund et al. 2020) and a reconstruction of Friedman's gap condition into the same framework.

2. Relativised Kruskal Fixed Points

In this section, we recall the definition of normal PO-dilators. We then construct the relativised Kruskal fixed points $\mathcal{T}W(X)$ that were mentioned in the introduction. We will introduce these fixed points in terms of notation systems. A more semantic characterisation will follow in the next section.

Jean-Yves Girard (1981) has introduced dilators as particularly uniform transformations of linear orders. A corresponding definition for partial orders has been given by Freund et al. (2020). In order to recall the precise definition, we need some terminology: A function $f:X\to Y$ between partial orders is called a quasi embedding if $f(x)\leq_Y f(y)$ implies $x\leq_X y$. If the converse implication holds as well, then we have an embedding. The category PO consists of the partial orders as objects and the quasi embeddings as morphisms. We say that a functor $W:\mathrm{PO}\to\mathrm{PO}$ preserves embeddings if $W(f):W(X)\to W(Y)$ is an embedding whenever the same holds for $f:X\to Y$. As in the introduction, we write $[X]^{<\omega}$ for the set of finite subsets of a given set X. To turn $[\cdot]^{<\omega}$ into a functor, we define

$$[f]^{<\omega}(a)=\{f(x)\mid x\in a\}\in[Y]^{<\omega}\qquad \text{for } f:X\to Y \text{ and } a\in[X]^{<\omega}.$$

We also apply $[\cdot]^{<\omega}$ to partial orders, omitting the forgetful functor to the underlying set. Conversely, subsets of partial orders are often considered as suborders.

Definition 2.1. *A PO-dilator consists of*

(i) a functor $W:\mathrm{PO}\to\mathrm{PO}$ that preserves embeddings and
(ii) a natural transformation $\mathrm{supp}^W:W\Rightarrow[\cdot]^{<\omega}$ that satisfies the following support condition: When $f:X\to Y$ is an embedding (not just a quasi embedding) of partial orders, the arrow $[f]^{<\omega}$ in the commutative diagram

$$\begin{tikzcd} W(X) \arrow[r, "W(f)"] \arrow[d, "\mathrm{supp}^W_X"'] & W(Y) \arrow[d, "\mathrm{supp}^W_Y"] \\ {[X]^{<\omega}} \arrow[r, "{[f]^{<\omega}}"'] & {[Y]^{<\omega}} \end{tikzcd}$$

can be lifted, in the sense that any given $\sigma \in W(Y)$ lies in the range (in the sense of image) of $W(f)$ if $\mathrm{supp}^W_Y(\sigma)$ lies in the range of $[f]^{<\omega}$.

If $W(X)$ is a well partial order (wpo) for any wpo X, then W is a WPO-dilator.

The reader may have observed that the previous definition focuses on embeddings rather than quasi embeddings. The latter are important for applications to the theory of well partial orders (see Freund et al. 2020).

Concerning the support condition with respect to an embedding $f : X \to Y$, we point out that $b \in \mathrm{rng}([f]^{<\omega})$ and $b \subseteq \mathrm{rng}(f)$ are equivalent for $b \in [Y]^{<\omega}$ (where $\mathrm{rng}(\cdot)$ denotes the range). Hence, the support condition can also written as

$$\mathrm{rng}(W(f)) \supseteq \{\sigma \in W(Y) \mid \mathrm{supp}^W_Y(\sigma) \subseteq \mathrm{rng}(f)\}.$$

In fact we get an equality, since the converse inclusion $\subseteq$ holds by naturality.

When the partial order X is clear from the context, then $\iota_a : a \hookrightarrow X$ denotes the inclusion of a suborder $a \subseteq X$. For $\sigma \in W(X)$, we write

$$\sigma =_{\mathrm{NF}} W(\iota_a)(\sigma_0) \qquad \text{with } a \in [X]^{<\omega} \text{ and } \sigma_0 \in W(a)$$

if the equality holds and we have $\mathrm{supp}^W_a(\sigma_0) = a$. The latter is a uniqueness condition, which is required for the following result:

Lemma 2.2. *Consider a PO-dilator W and a partial order X. Any $\sigma \in W(X)$ has a unique normal form $\sigma =_{\mathrm{NF}} W(\iota_a)(\sigma_0)$. For the latter, we have $a = \mathrm{supp}^W_X(\sigma)$.*

Proof. Consider the values of a given $\sigma_0 \in W(a)$ under the maps in the following commutative diagram, where $\sigma = W(\iota_a)(\sigma_0)$:

$$\begin{tikzcd} \sigma_0 \in W(a) \arrow[r, "W(\iota_a)"] \arrow[d, "\mathrm{supp}^W_a"'] & W(X) \ni \sigma \arrow[d, "\mathrm{supp}^W_X"] \\ \mathrm{supp}^W_a(\sigma_0) \in [a]^{<\omega} \arrow[r, "{[\iota_a]^{<\omega}}"'] & {[X]^{<\omega}} \ni \mathrm{supp}^W_X(\sigma) \end{tikzcd}$$

Since we have $[\iota_a]^{<\omega}(a) = a$ and $[\iota_a]^{<\omega}$ is injective, we obtain

$$\sigma =_{\mathrm{NF}} W(\iota_a)(\sigma_0) \quad \Leftrightarrow \quad \mathrm{supp}^W_a(\sigma_0) = a \quad \Leftrightarrow \quad \mathrm{supp}^W_X(\sigma) = a.$$

For existence, consider $\sigma \in W(X)$ and set $a = \mathrm{supp}^W_X(\sigma)$. Due to $a \in \mathrm{rng}([\iota_a]^{<\omega})$, the support condition yields a σ_0 as in the diagram. We have $\sigma =_{\mathrm{NF}} W(\iota_a)(\sigma_0)$ by the equivalence above. For uniqueness, we assume $\sigma =_{\mathrm{NF}} W(\iota_a)(\sigma_0)$. The equivalence reveals that a is determined by σ. Just as any embedding, the function $W(\iota_a)$ is injective. Hence, σ_0 is uniquely determined as well. □

The normal forms from the previous lemma can be used to represent PO-dilators in second order arithmetic, as worked out by Freund et al. (2020). In the present paper we do not work within a particular meta theory. Given a partial order X, we define a quasi order $\leq^{\mathrm{fin}}_X$ on the set $[X]^{<\omega}$ by stipulating

$$a \leq_X^{\text{fin}} b \quad \Leftrightarrow \quad \text{for any } x \in a \text{ there is a } y \in b \text{ with } x \leq_X y.$$

We will write $a \leq_X^{\text{fin}} y$ (resp. $x \leq_X^{\text{fin}} b$) rather than $a \leq_X^{\text{fin}} \{y\}$ (resp. $\{x\} \leq_X^{\text{fin}} b$) in the case of singletons. The following normality condition turns out to be crucial:

Definition 2.3. *A PO-dilator W is normal if the function $\operatorname{supp}_X^W : W(X) \to [X]^{<\omega}$ is order preserving for any partial order X, in the sense that we have*

$$\sigma \leq_{W(X)} \tau \quad \Rightarrow \quad \operatorname{supp}_X^W(\sigma) \leq_X^{\text{fin}} \operatorname{supp}_X^W(\tau)$$

for all elements $\sigma, \tau \in W(X)$.

In many applications, the elements $\sigma, \tau \in W(X)$ are finite structures with labels in X. Then, the inequality $\operatorname{supp}_X^W(\sigma) \leq_X^{\text{fin}} \operatorname{supp}_X^W(\tau)$ corresponds to the condition that each label is mapped to a bigger one. In the paper by Freund et al. (2020), the Kruskal fixed point $\mathcal{T}W$ of a normal PO-dilator has been generated by the following inductive clause:

- If we have already generated a finite suborder $a \subseteq \mathcal{T}W$, we add a term $\circ(a, \sigma) \in \mathcal{T}W$ for each element $\sigma \in W(a)$ with $\operatorname{supp}_a^W(\sigma) = a$.

The point is that one can now define a bijection $\kappa : W(\mathcal{T}W) \to \mathcal{T}W$ by stipulating that we have $\kappa(\sigma) = \circ(a, \sigma_0)$ for $\sigma =_{\text{NF}} W(\iota_a)(\sigma_0)$. We will relativise the construction by including constant symbols $\overline{x} \in \mathcal{T}W(X)$ for elements $x \in X$ of a given partial order. At various places in the following definition, we require that $\leq_{\mathcal{T}W(X)}$ is a partial order on certain subsets of $\mathcal{T}W(X)$. We will later show that all of $\mathcal{T}W(X)$ is partially ordered by $\leq_{\mathcal{T}W(X)}$, so that these requirements become redundant. A more detailed justification of the following recursion can be found below.

Definition 2.4. *Let us consider a normal PO-dilator W. For each partial order X, we define a set $\mathcal{T}W(X)$ of terms and a binary relation $\leq_{\mathcal{T}W(X)}$ on this set by simultaneous recursion. The set $\mathcal{T}W(X)$ is generated by the following clauses:*

(i) For each element $x \in X$ we have a term $\overline{x} \in \mathcal{T}W(X)$.
(ii) Given a finite set $a \subseteq \mathcal{T}W(X)$ that is partially ordered by $\leq_{\mathcal{T}W(X)}$, we add a new term $\circ(a, \sigma) \in \mathcal{T}W(X)$ for each $\sigma \in W(a)$ with $\operatorname{supp}_a^W(\sigma) = a$.

For $s, t \in \mathcal{T}W(X)$ we stipulate that $s \leq_{\mathcal{T}W(X)} t$ holds if, and only if, one of the following clauses applies:

(i') We have $s = \overline{x}$ and $t = \overline{y}$ with $x \leq_X y$.
(ii') We have $t = \circ(b, \tau)$ and $s \leq_{\mathcal{T}W(X)} t'$ for some $t' \in b$ (where s can be of the form $\overline{x}$ or $\circ(a, \sigma)$).
(iii') We have $s = \circ(a, \sigma)$ and $t = \circ(b, \tau)$, the restriction of $\leq_{\mathcal{T}W(X)}$ to $a \cup b$ is a partial order, and we have

$$W(\iota_a)(\sigma) \leq_{W(a \cup b)} W(\iota_b)(\tau),$$

where $\iota_a : a \hookrightarrow a \cup b$ and $\iota_b : b \hookrightarrow a \cup b$ are the inclusions.

In order to justify the recursion in detail, one can argue as follows: First, generate a set $\mathcal{T}_0 W(X) \supseteq \mathcal{T}W(X)$ by including all terms $\circ(a, \sigma)$ for finite $a \subseteq \mathcal{T}_0 W(X)$, where a is not assumed to be ordered and $\sigma \in W(a)$ holds with respect to some partial order on a. Then define a length function $l_X : \mathcal{T}_0 W(X) \to \mathbb{N}$ by the recursive clauses

$$l_X(\overline{x}) = 0, \qquad l_X(\circ(a, \sigma)) = 1 + \textstyle\sum_{r \in a} 2 \cdot l_X(r).$$

Now decide $r \in \mathcal{T}W(X)$ and $s \leq_{\mathcal{T}W(X)} t$ by simultaneous recursion on $l_X(r)$ and $l_X(s) + l_X(t)$. As an example, we consider the case of $r = \circ(a, \sigma)$. For $s, t \in a$ we have $l_X(s) + l_X(t) < l_X(r)$, even when s and t are the same term (due to the factor 2 above). Recursively, we can thus determine the restriction of $\leq_{\mathcal{T}W(X)}$ to a. If the latter is a partial order, we check whether $\sigma \in W(a)$ and $\mathrm{supp}^W_a(\sigma) = a$ hold with respect to this order. When this is the case, we have $r \in \mathcal{T}W(X)$. In addition to the length functions, we need the height functions $h_X : \mathcal{T}W(X) \to \mathbb{N}$ given by

$$h_X(\overline{x}) = 0, \qquad h_X(\circ(a, \sigma)) = \max(\{0\} \cup \{h_X(r) + 1 \mid r \in a\}).$$

When there is no danger of confusion, we sometimes omit the index X. The following important observation relies on the assumption that W is normal. It confirms the intuition that $\mathcal{T}W(X)$ can be seen as a tree-like structure.

Lemma 2.5. *Consider a normal PO-dilator W and a partial order X. For any $s, t \in \mathcal{T}W(X)$, the inequality $s \leq_{\mathcal{T}W(X)} t$ implies $h_X(s) \leq h_X(t)$.*

Proof. One argues by induction on $l(s) + l(t)$. For $s = \overline{x}$ and $t = \overline{y}$, the claim is immediate. The other cases are checked as for the non-relativised order $\mathcal{T}W$. Details can be found in the proof of Lemma 3.5 of Freund et al. (2020). In order to explain the importance of normality, we briefly show where it is needed: Assume that $s = \circ(a, \sigma) \leq_{\mathcal{T}W(X)} \circ(b, \tau) = t$ holds because of $W(\iota_a)(\sigma) \leq_{W(a \cup b)} W(\iota_b)(\tau)$. Using normality, one deduces $a \leq^{\mathrm{fin}}_{\mathcal{T}W(X)} b$. This means that any $s' \in a$ is bounded by some $t' \in b$, so that $h(s') \leq h(t') < h(t)$ holds by induction hypothesis. Since the element $s' \in a$ was arbitrary, we get $h(s) \leq h(t)$. □

Using the previous lemma, one can prove what was promised above:

Proposition 2.6. *The relation $\leq_{\mathcal{T}W(X)}$ is a partial order on the set $\mathcal{T}W(X)$, for any normal PO-dilator W and any partial order X.*

Proof. The argument is very similar to the non-relativised case, for which we refer to Proposition 3.6 of Freund et al. (2020): One establishes reflexivity, antisymmetry and transitivity by simultaneous induction on the added length of the terms that are involved. The new cases are straightforward. Consider, for example, $r \leq_{\mathcal{T}W(X)} s \leq_{\mathcal{T}W(X)} t$ with $t = \overline{z}$ for some $z \in X$. Then we must have $s = \overline{y}$ and $r = \overline{x}$ with $x \leq_X y \leq_X z$, which yields $r \leq_{\mathcal{T}W(X)} t$. To show where the previous lemma is needed, we consider one case in the proof of antisymmetry: Assume that $s \leq_{\mathcal{T}W(X)} \circ(b, \tau) = t$ holds because we have $s \leq_{\mathcal{T}W(X)} t'$ for some $t' \in b$. Lemma 2.5 yields $h(s) \leq h(t') < h(t)$, making $t \leq_{\mathcal{T}W(X)} s$ impossible. Normality is needed to prove Lemma 2.5, and also for one case in the proof of transitivity (cf. Proposition 3.6 of Freund et al. 2020). □

Concerning preservation of well partial orders, we have the following result:

Proposition 2.7. *Let W be a normal WPO-dilator. If X is a well partial order, then so is $\mathcal{T}W(X)$.*

Proof. The argument is very similar to the non-relativised case, which is proved as Theorem 3.10 of Freund et al. (2020): Aiming at a contradiction, consider a well partial order X and an infinite sequence $t_0, t_1, \dots \subseteq \mathcal{T}W(X)$ such that $t_i \leq_{\mathcal{T}W(X)} t_j$ fails for all $i < j$. As in the cited proof, Nash-Williams' (1963) minimal bad sequence argument can be used to reduce to the following situation: There is a well partial order $Z \subseteq \mathcal{T}W(X)$ with $a_i \subseteq Z$ for all t_i of the form $t_i = \circ(a_i, \sigma_i)$. Once this is accomplished, note that only finitely many t_i can have the form $t_i = \overline{x_i}$ with $x_i \in X$. After deleting these, we can conclude as in the proof of Theorem 3.10 of Freund et al. (2020). □

For a suitable formalisation of PO-dilators in the language of second-order arithmetic, the result of Proposition 2.7 is equivalent to Π^1_1-comprehension, over $\mathbf{RCA_0}$ together with the chain antichain principle. Indeed, the main result of Freund et al. (2020) shows that Π^1_1-comprehension follows from Proposition 2.7 restricted to $X=\emptyset$, over the same base theory. Conversely, the axiom of Π^1_1-comprehension justifies the minimal bad sequence argument, which is the crucial ingredient for the proof of Proposition 2.7.

3. A Categorical Characterisation

The term systems $\mathcal{T}W(X)$ from the previous section can be hard to handle, both in general arguments and in concrete examples. To resolve this issue, the present section provides a more semantic approach. We begin with a general notion:

Definition 3.1. *Consider a normal PO-dilator W and a partial order X. A Kruskal fixed point of W over X consists of a partial order Z and functions $\iota : X \to Z$ and $\kappa : W(Z) \to Z$ that satisfy*

$$\begin{aligned}
&\iota(x) \leq_Z \iota(y) \quad \Rightarrow \quad x \leq_X y \quad \text{(for } x, y \in X\text{)},\\
&\iota(x) \leq_Z \kappa(\tau) \quad \Leftrightarrow \quad \iota(x) \leq_Z^{\text{fin}} \operatorname{supp}_Z^W(\tau) \quad \text{(for } x \in X \text{ and } \tau \in W(Z)\text{)},\\
&\kappa(\sigma) \not\leq_Z \iota(y) \quad \text{holds for all } \sigma \in W(Z) \text{ and } y \in X,\\
&\kappa(\sigma) \leq_Z \kappa(\tau) \quad \Leftrightarrow \quad \sigma \leq_{W(Z)} \tau \text{ or } \kappa(\sigma) \leq_Z^{\text{fin}} \operatorname{supp}_Z^W(\tau) \quad \text{(for } \sigma, \tau \in W(Z)\text{)}.
\end{aligned}$$

Note that the third condition entails $\operatorname{rng}(\iota) \cap \operatorname{rng}(\kappa) = \emptyset$. Also note that we do not demand that $x \leq_X y$ implies $\iota(x) \leq_Z \iota(y)$. This will become important in the proof of Theorem 4.2. The last equivalence in Definition 3.1 corresponds to the definition of Kruskal fixed points over $X=\emptyset$ in Definition 3.7 of Freund et al. (2020). The following is justified by Lemma 2.2.

Definition 3.2. *Consider a normal PO-dilator W. For each partial order X, we define functions $\iota_X : X \to \mathcal{T}W(X)$ and $\kappa_X : W(\mathcal{T}W(X)) \to \mathcal{T}W(X)$ by stipulating*

$$\begin{aligned}
\iota_X(x) &= \overline{x},\\
\kappa_X(\sigma) &= \circ(a, \sigma_0) \quad \text{for } \sigma =_{\text{NF}} W(\iota_a)(\sigma_0).
\end{aligned}$$

Let us verify that $\mathcal{T}W(X)$ has the desired structure:

Theorem 3.3. *We consider a normal PO-dilator W and a partial order X. The order $\mathcal{T}W(X)$ and the functions ι_X and κ_X form a Kruskal fixed point of W over X.*

Proof. In view of Definition 2.4, it is immediate that $\iota_X(x) = \overline{x} \leq_{\mathcal{T}W(X)} \overline{y} = \iota_X(y)$ implies (and is in fact equivalent to) $x \leq_X y$ and that $\kappa_X(\sigma) = \circ(a, \sigma_0) \leq_{\mathcal{T}W(X)} \overline{y} = \iota_X(y)$ is always false. For an element $\tau =_{\text{NF}} W(\iota_b)(\tau_0)$, we also get

$$\iota_X(x) = \overline{x} \leq_{\mathcal{T}W(X)} \circ(b, \tau_0) = \kappa_X(\tau) \quad \Leftrightarrow \quad \iota_X(x) \leq^{\text{fin}}_{\mathcal{T}W(X)} b = \operatorname{supp}^W_{\mathcal{T}W(X)}(\tau).$$

The remaining equivalence is verified as in the non-relativised case, for which we refer to Theorem 3.8 of Freund et al. (2020). □

To obtain a unique characterisation, we use the following categorical description, which extends Definition 3.7 of Freund et al. (2020):

Definition 3.4. *Consider a normal PO-dilator W and a partial order X. A Kruskal fixed point (Z, ι, κ) of W over X is called initial if any Kruskal fixed point (Z', ι', κ') of W over X admits a unique quasi embedding $f : Z \to Z'$ such that*

$$\begin{tikzcd} & Z \arrow[d, "f"] & W(Z) \arrow[l, "\kappa"'] \arrow[d, "W(f)"] \\ X \arrow[ur, "\iota"] \arrow[dr, "\iota'"'] & & \\ & Z' & W(Z) \arrow[l, "\kappa'"] \end{tikzcd}$$

is a commutative diagram.

Just like other initial objects, initial Kruskal fixed points are unique up to isomorphism. The following criterion will be very useful.

Theorem 3.5. *For a Kruskal fixed point (Z, ι, κ) of a normal PO-dilator W over a partial order X, the following are equivalent:*

(i) We have $\mathrm{rng}(\iota) \cup \mathrm{rng}(\kappa) = Z$, and $x \leq_X y$ implies $\iota(x) \leq_Z \iota(y)$ for $x, y \in X$. Furthermore, there is a function $h : Z \to \mathbb{N}$ such that

$$s \in \mathrm{supp}^W_Z(\sigma) \quad \Rightarrow \quad h(s) < h(\kappa(\sigma))$$

holds for any $s \in Z$ and any $\sigma \in W(Z)$.

(ii) The Kruskal fixed point (Z, ι, κ) is initial.

Proof. Let us first show that condition (i) implies (ii). For $s \in Z$ we define $l(s) \in \mathbb{N}$ by recursion on $h(s)$, setting

$$l(\iota(x)) = 0 \qquad \text{and} \qquad l(\kappa(\sigma)) = 1 + \textstyle\sum_{s \in \mathrm{supp}^W_Z(\sigma)} 2 \cdot l(s).$$

Note that each element of Z is covered by exactly one clause, since Definition 3.1 and part (i) of the present theorem provide $\mathrm{rng}(\iota) \cap \mathrm{rng}(\kappa) = \emptyset$ and $\mathrm{rng}(\iota) \cup \mathrm{rng}(\kappa) = Z$. Now consider another Kruskal fixed point (Z', ι', κ'). We first show that there is at most one quasi embedding $f : Z \to Z'$ such that the diagram from Definition 3.4 commutes. This condition amounts to the equations

$$\begin{aligned} f(\iota(x)) &= \iota'(x) && \text{for } x \in X, \\ f(\kappa(\sigma)) &= \kappa'(W(f)(\sigma)) = \kappa'(W(f{\restriction}a)(\sigma_0)) && \text{for } \sigma =_{\mathrm{NF}} W(\iota_a)(\sigma_0) \in W(Z), \end{aligned}$$

where $f{\restriction}a = f \circ \iota_a : a \to Z'$ is the restriction of f. Once again, each argument of f is covered by exactly one of these clauses. From Lemma 2.2, we know that $\sigma =_{\mathrm{NF}} W(\iota_a)(\sigma_0)$ implies $\mathrm{supp}^W_Z(\sigma) = a$. Now a straightforward induction on $l(s)$ shows that $f(s)$ is uniquely determined. To establish existence we read the above as recursive clauses. We verify

$$\begin{aligned} r \in Z \quad &\Rightarrow \quad f(r) \in Z', \\ f(s) \leq_{Z'} f(t) \quad &\Rightarrow \quad s \leq_Z t \end{aligned}$$

by simultaneous induction on $l(r)$ and $l(s) + l(t)$. Let us verify the first claim for $r = \kappa(\sigma)$ with $\sigma =_{\mathrm{NF}} W(\iota_a)(\sigma_0)$. For $s, t \in a$ we have $l(s) + l(t) < l(r)$. Hence, the simultaneous induction hypothesis ensures that $f{\restriction}a$ is a quasi embedding. We may thus form $W(f{\restriction}a)$, as needed for the clause that defines the value $f(r) \in Z'$. Let us now show that f is a quasi embedding. For $s = \iota(x)$ and $t = \iota(y)$ we see that

$$f(s) = f(\iota(x)) = \iota'(x) \leq_{Z'} \iota'(y) = f(\iota(y)) = f(t)$$

implies $x \leq_X y$. By (i) this implies $s = \iota(x) \leq_Z \iota(y) = t$, as required. For $s = \iota(x)$ and $t = \kappa(\tau)$ with $\tau =_{\mathrm{NF}} W(\iota_b)(\tau_0)$, a glance at Definition 3.1 reveals that $f(s) = \iota'(x) \leq_{Z'} \kappa'(W(f{\restriction}b)(\tau_0)) = f(t)$ implies

$$f(s) \leq^{\text{fin}}_{Z'} \operatorname{supp}^W_{Z'}(W(f\restriction b)(\tau_0)) = [f\restriction b]^{<\omega}(\operatorname{supp}^W_b(\tau_0)) = [f]^{<\omega}(b),$$

where the equalities rely on naturality of supports and on $f\restriction b = f \circ \iota_b$. Since $t' \in b = \operatorname{supp}^W_Z(\tau)$ implies $l(t') < l(\kappa(\tau)) = l(t)$, the induction hypothesis yields $s \leq^{\text{fin}}_Z \operatorname{supp}^W_Z(\tau)$, which implies $s = \iota(x) \leq_Z \kappa(\tau) = t$. For $s = \kappa(\sigma)$ and $t = \iota(y)$ it suffices to observe that $f(s) \leq_{Z'} f(t)$ cannot hold. Finally, we consider the case of elements $s = \kappa(\sigma)$ and $t = \kappa(\tau)$ with $\sigma =_{\text{NF}} W(\iota_a)(\sigma_0)$ and $\tau =_{\text{NF}} W(\iota_b)(\tau_0)$. If

$$f(s) = \kappa'(W(f\restriction a)(\sigma_0)) \leq_{Z'} \kappa'(W(f\restriction b)(\tau_0)) = f(t)$$

holds because of $f(s) \leq^{\text{fin}}_{Z'} \operatorname{supp}^W_{Z'}(W(f\restriction b)(\tau_0))$, then one argues as above. Now assume that we have

$$W(f\restriction a)(\sigma_0) \leq_{W(Z')} W(f\restriction b)(\tau_0).$$

The induction hypothesis ensures that $f\restriction(a \cup b) : a \cup b \to Z'$ is a quasi embedding. Here, it is crucial that we argue by induction on $l(s) + l(t)$, not on $h(s) + h(t)$. Let us factor $f\restriction a = f\restriction(a \cup b) \circ \iota'_a$ and $f\restriction b = f\restriction(a \cup b) \circ \iota'_b$, where $\iota'_a : a \hookrightarrow a \cup b$ and $\iota'_b : b \hookrightarrow a \cup b$ are the inclusions. Then the last inequality amounts to

$$W(f\restriction(a \cup b)) \circ W(\iota'_a)(\sigma_0) \leq_{W(Z')} W(f\restriction(a \cup b)) \circ W(\iota'_b)(\tau_0).$$

Since $W(f\restriction(a \cup b))$ is a quasi embedding, we get $W(\iota'_a)(\sigma_0) \leq_{W(a\cup b)} W(\iota'_b)(\tau_0)$. Now compose both sides with the embedding $W(\iota)$, where $\iota : a \cup b \hookrightarrow Z$ is the inclusion. This yields

$$\sigma = W(\iota_a)(\sigma_0) = W(\iota \circ \iota'_a)(\sigma_0) \leq_{W(Z)} W(\iota \circ \iota'_b)(\tau_0) = W(\iota_b)(\tau_0) = \tau.$$

The latter implies $s = \kappa(\sigma) \leq_Z \kappa(\tau) = t$, which completes the proof that (i) implies (ii).

To show that (ii) implies (i), we first establish (i) for the Kruskal fixed point $(\mathcal{T}W(X), \iota_X, \kappa_X)$ from Theorem 3.3. Any element $\circ(a, \sigma_0) \in \mathcal{T}W(X)$ arises as $\kappa_X(\sigma)$ for $\sigma = W(\iota_a)(\sigma_0)$. Here the condition $\operatorname{supp}^W_a(\sigma_0) = a$ from Definition 2.4 ensures that σ is in normal form. This shows $\operatorname{rng}(\iota_X) \cup \operatorname{rng}(\kappa_X) = \mathcal{T}W(X)$. The requirement that $x \leq_X y$ implies $\iota_X(x) = \overline{x} \leq_{\mathcal{T}W(X)} \overline{y} = \iota_X(y)$ is immediate by Definition 2.4. A height function $h_X : \mathcal{T}W(X) \to \mathbb{N}$ has been defined before Lemma 2.5. For arbitrary elements $\sigma =_{\text{NF}} W(\iota_a)(\sigma_0) \in W(\mathcal{T}W(X))$ and $s \in \operatorname{supp}^W_{\mathcal{T}W(X)}(\sigma) = a$, the construction of h_X entails

$$h_X(s) < h_X(\circ(a, \sigma_0)) = h_X(\kappa_X(\sigma)),$$

just as needed. Since we have already shown that (i) implies (ii), we can conclude that $(\mathcal{T}W(X), \iota_X, \kappa_X)$ is an initial Kruskal fixed point of W over X. If (Z, ι, κ) is any initial Kruskal fixed point as in (ii), we get an isomorphism $f : Z \to \mathcal{T}W(X)$ with $f \circ \iota = \iota_X$ and $f \circ \kappa = \kappa_X \circ W(f)$. As $W(f)$ is an isomorphism, the ranges of κ_X and $\kappa_X \circ W(f)$ coincide. Hence we can conclude $\mathcal{T}W(X) = \operatorname{rng}(f \circ \iota) \cup \operatorname{rng}(f \circ \kappa)$ and then $Z = \operatorname{rng}(\iota) \cup \operatorname{rng}(\kappa)$. We also learn that $x \leq_X y$ implies

$$f \circ \iota(x) = \iota_X(x) \leq_{\mathcal{T}W(X)} \iota_X(y) = f \circ \iota(y)$$

and then $\iota(x) \leq_Z \iota(y)$. Finally, we define $h : Z \to \mathbb{N}$ by setting $h(s) = h_X(f(s))$. For $s \in \operatorname{supp}^W_Z(\sigma)$ we have

$$f(s) \in [f]^{<\omega}(\operatorname{supp}^W_Z(\sigma)) = \operatorname{supp}^W_{\mathcal{T}W(X)}(W(f)(\sigma)).$$

We can conclude

$$h(s) = h_X(f(s)) < h_X(\kappa_X \circ W(f)(\sigma)) = h_X(f \circ \kappa(\sigma)) = h(\kappa(\sigma)),$$

as required for (i). □

The following result, which extends Theorem 3.8 of Freund et al. (2020), was shown as part of the previous proof. It is important since it establishes the existence of initial Kruskal fixed points.

Corollary 3.6. *For each normal PO-dilator W and each partial order X, the Kruskal fixed point $(\mathcal{T}W(X), \iota_X, \kappa_X)$ is initial.*

For later use we also record the following result.

Lemma 3.7. *Let (Z, ι, κ) be an initial Kruskal fixed point of a normal PO-dilator W over a partial order X. Consider another Kruskal fixed point (Z', ι', κ') and the unique quasi embedding $f : Z \to Z'$ with $f \circ \iota = \iota'$ and $f \circ \kappa = \kappa' \circ W(f)$. If $x \leq_X y$ implies $\iota'(x) \leq_{Z'} \iota'(y)$ for all $x, y \in X$, then f is an embedding.*

Proof. Define $l : Z \to \mathbb{N}$ as in the proof of Theorem 3.5. In the latter, we have used induction on $l(s) + l(t)$ to show that $f(s) \leq_{Z'} f(t)$ implies $s \leq_Z t$. Assuming that $x \leq_X y$ implies $\iota'(x) \leq_{Z'} \iota'(y)$, one can read the given argument in reverse, to show that $s \leq_Z t$ does also imply $f(s) \leq_{Z'} f(t)$. $\square$

So far, the notation $\mathcal{T}W(X)$ has been reserved for the term systems constructed in Definition 2.4. In the following we will also use $\mathcal{T}W(X)$ for other initial Kruskal fixed points of W over X. This is harmless, as all these fixed points are equivalent.

4. Kruskal Derivatives

Consider a normal PO-dilator W. As shown in the previous section, each partial order X gives rise to an initial Kruskal fixed point $(\mathcal{T}W(X), \iota_X, \kappa_X)$. In the present section, we show that the transformation $X \mapsto \mathcal{T}W(X)$ of partial orders can be extended into a normal PO-dilator $\mathcal{T}W$. More precisely, we will show that there is an essentially unique extension in the sense of the following definition.

Definition 4.1. *A Kruskal derivative of a normal PO-dilator W is a tuple $(\mathcal{T}W, \iota, \kappa)$ that consists of a normal PO-dilator $\mathcal{T}W$ and two families of functions*

$$\iota_X : X \to \mathcal{T}W(X) \quad \text{and} \quad \kappa_X : W(\mathcal{T}W(X)) \to \mathcal{T}W(X)$$

indexed by the partial order X, such that the following properties are satisfied:

(i) *The tuple $(\mathcal{T}W(X), \iota_X, \kappa_X)$ is an initial Kruskal fixed point of W over X, for each partial order X.*
(ii) *The diagram*

$$\begin{tikzcd}
X \arrow[r, "\iota_X"] \arrow[d, "f"'] & \mathcal{T}W(X) \arrow[d, "\mathcal{T}W(f)"] & W(\mathcal{T}W(X)) \arrow[l, "\kappa_X"'] \arrow[d, "W(\mathcal{T}W(f))"] \\
Y \arrow[r, "\iota_Y"'] & \mathcal{T}W(Y) & W(\mathcal{T}W(Y)) \arrow[l, "\kappa_Y"]
\end{tikzcd}$$

commutes for any quasi embedding $f : X \to Y$ between partial orders.

Let us begin by proving existence:

Theorem 4.2. *Each normal PO-dilator has a Kruskal derivative.*

Proof. Consider a normal PO-dilator W. For each partial order X, Corollary 3.6 provides an initial Kruskal fixed point $(\mathcal{T}W(X), \iota_X, \kappa_X)$ of W over X. Given a quasi embedding $f : X \to Y$, it is easy to see that $(\mathcal{T}W(Y), \iota_Y \circ f, \kappa_Y)$ is a Kruskal fixed point of W over X as well. Since $(\mathcal{T}W(X), \iota_X, \kappa_X)$ is initial, Definition 3.4 entails that there is a unique quasi embedding $\mathcal{T}W(f):\mathcal{T}W(X)\to\mathcal{T}W(Y)$ such that

$$\begin{tikzcd} & \mathcal{T}W(X) \arrow[d, "\mathcal{T}W(f)"] & W(\mathcal{T}W(X)) \arrow[l, "\kappa_X"'] \arrow[d, "W(\mathcal{T}W(f))"] \\ X \arrow[ur, "\iota_X"] \arrow[dr, "\iota_Y\circ f"'] & & \\ & \mathcal{T}W(Y) & W(\mathcal{T}W(Y)) \arrow[l, "\kappa_Y"] \end{tikzcd}$$

commutes, as required by Definition 4.1. To obtain a Kruskal derivative $(\mathcal{T}W,\iota,\kappa)$, it suffices to turn $\mathcal{T}W$ into a normal PO-dilator. In order to show that $\mathcal{T}W$ is a functor, one checks that $\mathcal{T}W(g)\circ\mathcal{T}W(f)$ satisfies the condition that characterises $\mathcal{T}W(g\circ f)$. If $f:X\to Y$ is an embedding, then $x\leq_X y$ implies $\iota_Y\circ f(x)\leq_{\mathcal{T}W(Y)}\iota_Y\circ f(y)$, since ι_Y must satisfy the condition from part (i) of Theorem 3.5. Hence Lemma 3.7 ensures that $\mathcal{T}W(f)$ is again an embedding, as required in part (i) of Definition 2.1. It remains to exhibit suitable support functions

$$\operatorname{supp}^{\mathcal{T}W}_X:\mathcal{T}W(X)\to[X]^{<\omega}.$$

In view of Theorem 3.5, we can recursively define

$$\begin{aligned}\operatorname{supp}^{\mathcal{T}W}_X(\iota_X(x))&=\{x\},\\ \operatorname{supp}^{\mathcal{T}W}_X(\kappa_X(\sigma))&=\bigcup\{\operatorname{supp}^{\mathcal{T}W}_X(s)\mid s\in\operatorname{supp}^W_{\mathcal{T}W(X)}(\sigma)\}.\end{aligned}$$

To show naturality, one verifies

$$\operatorname{supp}^{\mathcal{T}W}_Y(\mathcal{T}W(f)(s))=[f]^{<\omega}(\operatorname{supp}^{\mathcal{T}W}_X(s))$$

by induction on $h_X(s)$, where $h_X:\mathcal{T}W(X)\to\mathbb{N}$ is as in part (i) of Theorem 3.5. To verify the support condition from part (ii) of Definition 2.1, we need to establish

$$\operatorname{supp}^{\mathcal{T}W}_Y(s)\subseteq\operatorname{rng}(f)\quad\Rightarrow\quad s\in\operatorname{rng}(\mathcal{T}W(f))$$

for an embedding $f:X\to Y$ (recall that the converse implication is automatic). We use induction on $h_Y(s)$. For $s=\iota_Y(y)$, we see that $\{y\}=\operatorname{supp}^{\mathcal{T}W}_Y(s)\subseteq\operatorname{rng}(f)$ yields $y=f(x)$ for some $x\in X$. This entails

$$s=\iota_Y\circ f(x)=\mathcal{T}W(f)\circ\iota_X(x)\in\operatorname{rng}(\mathcal{T}W(f)).$$

Now consider $s=\kappa_Y(\sigma)$. For any $s'\in\operatorname{supp}^W_{\mathcal{T}W(Y)}(\sigma)$ we have $h_Y(s')<h_Y(s)$ and

$$\operatorname{supp}^{\mathcal{T}W}_Y(s')\subseteq\operatorname{supp}^{\mathcal{T}W}_Y(s)\subseteq\operatorname{rng}(f),$$

so that the induction hypothesis yields $s'\in\operatorname{rng}(\mathcal{T}W(f))$. Thus we get

$$\operatorname{supp}^W_{\mathcal{T}W(Y)}(\sigma)\subseteq\operatorname{rng}(\mathcal{T}W(f)).$$

Now the support condition for the PO-dilator W ensures that we have $\sigma=W(\mathcal{T}W(f))(\sigma_0)$ for some $\sigma_0\in W(\mathcal{T}W(X))$. We then obtain

$$s=\kappa_Y\circ W(\mathcal{T}W(f))(\sigma_0)=\mathcal{T}W(f)\circ\kappa_X(\sigma_0)\in\operatorname{rng}(\mathcal{T}W(f)),$$

as required. It remains to show that the PO-dilator $\mathcal{T}W$ is normal. We verify

$$s\leq_{\mathcal{T}W(X)}t\quad\Rightarrow\quad\operatorname{supp}^{\mathcal{T}W}_X(s)\leq^{\mathrm{fin}}_X\operatorname{supp}^{\mathcal{T}W}_X(t)$$

by induction on $h_X(s)+h_X(t)$. If we have $s=\iota_X(x)\leq_{\mathcal{T}W(X)}\iota_X(y)=t$, then we must have $x\leq_X y$ and hence

$$\operatorname{supp}^{\mathcal{T}W}_X(s)=\{x\}\leq^{\mathrm{fin}}_X\{y\}=\operatorname{supp}^{\mathcal{T}W}_X(t).$$

Now consider the case of an inequality $s \leq_{\mathcal{T}W(X)} \kappa_X(\tau) = t$ that holds because $s \leq_{\mathcal{T}W(X)} t'$ is true for some $t' \in \operatorname{supp}^W_{\mathcal{T}W(X)}(\tau)$. In view of $h_X(t') < h_X(t)$, the induction hypothesis yields

$$\operatorname{supp}^{\mathcal{T}W}_X(s) \leq^{\mathrm{fin}}_X \operatorname{supp}^{\mathcal{T}W}_X(t') \subseteq \operatorname{supp}^{\mathcal{T}W}_X(t).$$

Finally, assume that $s = \kappa_X(\sigma) \leq_{\mathcal{T}W(X)} \kappa_X(\tau) = t$ holds due to $\sigma \leq_{W(\mathcal{T}W(X))} \tau$. Since W is normal, we get $\operatorname{supp}^W_{\mathcal{T}W(X)}(\sigma) \leq^{\mathrm{fin}}_{\mathcal{T}W(X)} \operatorname{supp}^W_{\mathcal{T}W(X)}(\tau)$. Given an arbitrary $s' \in \operatorname{supp}^W_{\mathcal{T}W(X)}(\sigma)$, we may thus pick a $t' \in \operatorname{supp}^W_{\mathcal{T}W(X)}(\tau)$ with $s' \leq_{\mathcal{T}W(X)} t'$. By induction hypothesis we get

$$\operatorname{supp}^{\mathcal{T}W}_X(s') \leq^{\mathrm{fin}}_X \operatorname{supp}^{\mathcal{T}W}_X(t') \subseteq \operatorname{supp}^{\mathcal{T}W}_X(t).$$

Since $s' \in \operatorname{supp}^W_{\mathcal{T}W(X)}(\sigma)$ was arbitrary, this establishes

$$\operatorname{supp}^{\mathcal{T}W}_X(s) = \bigcup \{\operatorname{supp}^{\mathcal{T}W}_X(s') \mid s' \in \operatorname{supp}^W_{\mathcal{T}W(X)}(\sigma)\} \leq^{\mathrm{fin}}_X \operatorname{supp}^{\mathcal{T}W}_X(t),$$

as required. □

Let us highlight some of the information that is implicit in the previous proof:

Remark 4.3. In order to construct a Kruskal derivative of a specific PO-dilator, one can follow the proof of Theorem 4.2. The latter shows that we only need to find a family of initial Kruskal fixed points. The extension into a Kruskal derivative is then automatic. In particular, the fact that one obtains a normal PO-dilator does not need to be verified in each specific case. Also observe that the functor $\mathcal{T}W$ was uniquely determined by the initial Kruskal fixed points $\mathcal{T}W(X)$. The choice of support functions is also unique (as for any PO-dilator), since $\operatorname{supp}^{\mathcal{T}W}_X(s)$ must be the smallest set $a \subseteq X$ with $s \in \operatorname{rng}(\mathcal{T}W(\iota_a))$, where $\iota_a : a \hookrightarrow \mathcal{T}W(X)$ is the inclusion: In one direction, the support condition from part (ii) of Definition 2.1 ensures $s \in \operatorname{rng}(\mathcal{T}W(\iota_a))$ for $a = \operatorname{supp}^{\mathcal{T}W}_X(s)$. In the other direction, naturality entails that $s = \mathcal{T}W(\iota_a)(s_0)$ yields

$$a \subseteq [\iota_a]^{<\omega}(\operatorname{supp}^{\mathcal{T}W}_a(s_0)) = \operatorname{supp}^{\mathcal{T}W}_X(\mathcal{T}W(\iota_a)(s_0)) = \operatorname{supp}^{\mathcal{T}W}_X(s).$$

We have not included this information in the statement of Theorem 4.2, because a more general uniqueness result will be shown below.

To prepare our uniqueness result, we recall that two PO-dilators $(V, \operatorname{supp}^V)$ and $(W, \operatorname{supp}^W)$ are equivalent if there is a natural isomorphism $\eta : V \Rightarrow W$ of functors. It may also seem reasonable to demand

$$\operatorname{supp}^W_X \circ \eta_X = \operatorname{supp}^V_X$$

for any partial order X. However, the latter turns out to be automatic. Girard (1981) has shown that this is the case for any natural transformation between dilators of linear orders (cf. also Lemma 2.17 of Freund and Rathjen 2021, which is closer to our notation). One can check that the proof remains valid for partial orders. In the case of an isomorphism, the argument is particularly simple: Given $\sigma \in V(X)$, we invoke Lemma 2.2 to write $\sigma = V(\iota_a)(\sigma_0)$ with $a = \operatorname{supp}^V_X(\sigma)$. We then get

$$\begin{aligned}\operatorname{supp}^W_X \circ \eta_X(\sigma) = \operatorname{supp}^W_X(\eta_X \circ V(\iota_a)(\sigma_0)) = \operatorname{supp}^W_X(W(\iota_a) \circ \eta_a(\sigma_0)) = \\ = [\iota_a]^{<\omega}(\operatorname{supp}^W_a(\eta_a(\sigma_0))) \subseteq \operatorname{rng}(\iota_a) = a = \operatorname{supp}^V_X(\sigma).\end{aligned}$$

By applying the same argument to the inverse of η, we also get

$$\operatorname{supp}^V_X(\sigma) = \operatorname{supp}^V_X \circ \eta^{-1}_X(\eta_X(\sigma)) \subseteq \operatorname{supp}^W_X(\eta_X(\sigma)) = \operatorname{supp}^W_X \circ \eta_X(\sigma).$$

The following result shows that Kruskal derivatives are essentially unique.

Theorem 4.4. *For any two Kruskal derivatives $(\mathcal{T}^0W, \iota^0, \kappa^0)$ and $(\mathcal{T}^1W, \iota^1, \kappa^1)$ of a normal PO-dilator W, there is a natural isomorphism $\eta : \mathcal{T}^0W \Rightarrow \mathcal{T}^1W$ such that the following diagram commutes for any partial order X:*

$$\begin{tikzcd}
 & \mathcal{T}^0W(X) \arrow[dd, "\eta_X"] & W(\mathcal{T}^0W(X)) \arrow[l, "\kappa^0_X"'] \arrow[dd, "W(\eta_X)"] \\
X \arrow[ur, "\iota^0_X"] \arrow[dr, "\iota^1_X"'] & & \\
 & \mathcal{T}^1W(X) & W(\mathcal{T}^1W(X)) \arrow[l, "\kappa^1_X"]
\end{tikzcd}$$

Proof. For each order X, the fact that $(\mathcal{T}^0W(X), \iota^0_X, \kappa^0_X)$ and $(\mathcal{T}^1W(X), \iota^1_X, \kappa^1_X)$ are initial Kruskal fixed points of W over X implies that there is an isomorphism $\eta_X : \mathcal{T}^0W(X) \to \mathcal{T}^1W(X)$ with $\eta_X \circ \iota^0_X = \iota^1_X$ and $\eta_X \circ \kappa^0_X = \kappa^1_X \circ W(\eta_X)$. It remains to show that the resulting family η is natural. Given a quasi embedding $f : X \to Y$ between partial orders, we show

$$\eta_Y \circ \mathcal{T}^0W(f)(s) = \mathcal{T}^1W(f) \circ \eta_X(s)$$

by induction on $h^0_X(s)$, where $h^0_X : \mathcal{T}^0W(X) \to \mathbb{N}$ is as in part (i) of Theorem 3.5. For elements of the form $s = \iota^0_X(x)$, it suffices to observe that the marked subdiagrams in the following commute (note that the outer square is our current target):

$$\begin{tikzcd}
\mathcal{T}^0W(X) \arrow[rrrrr, "\mathcal{T}^0W(f)"] \arrow[dddd, "\eta_X"'] & & & & & \mathcal{T}^0W(Y) \arrow[dddd, "\eta_Y"] \\
 & & {} \arrow[r, phantom, "\#"] & {} & & \\
 & \# & X \arrow[uull, "\iota^0_X"'] \arrow[ddll, "\iota^1_X"] \arrow[r, "f"] & Y \arrow[uurr, "\iota^0_Y"'] \arrow[ddrr, "\iota^1_Y"] & \# & \\
 & & {} \arrow[r, phantom, "\#"] & {} & & \\
\mathcal{T}^1W(X) \arrow[rrrrr, "\mathcal{T}^1W(f)"'] & & & & & \mathcal{T}^1W(Y)
\end{tikzcd}$$

Given an element of the form $s = \kappa^0_X(\sigma)$, we invoke Lemma 2.2 to write $\sigma =_{\mathrm{NF}} W(\iota_a)(\sigma_0)$, where $\iota_a : a \hookrightarrow \mathcal{T}^0W(X)$ is the inclusion. For any $s' \in a = \operatorname{supp}^W_{\mathcal{T}^0W(X)}(\sigma)$ we have $h^0_X(s') < h^0_X(s)$, so that the induction hypothesis yields

$$\eta_Y \circ \mathcal{T}^0W(f) \circ \iota_a = \mathcal{T}^1W(f) \circ \eta_X \circ \iota_a.$$

Hence, the inner square in the following commutes after composition with $W(\iota_a)$:

$$\begin{tikzcd}
\mathcal{T}^0W(X) \arrow[rrrr, "\mathcal{T}^0W(f)"] \arrow[dddd, "\eta_X"'] & & {} \arrow[d, phantom, "\#"] & & \mathcal{T}^0W(Y) \arrow[dddd, "\eta_Y"] \\
 & W(\mathcal{T}^0W(X)) \arrow[ul, "\kappa^0_X"'] \arrow[rr, "W(\mathcal{T}^0W(f))"'] \arrow[dd, "W(\eta_X)"'] & {} & W(\mathcal{T}^0W(Y)) \arrow[ur, "\kappa^0_Y"] \arrow[dd, "W(\eta_Y)"] & \\
{} \arrow[r, phantom, "\#"] & {} & W(a) \arrow[ul, "W(\iota_a)"'] & {} & {} \arrow[l, phantom, "\#"] \\
 & W(\mathcal{T}^1W(X)) \arrow[dl, "\kappa^1_X"'] \arrow[rr, "W(\mathcal{T}^1W(f))"] & {} & W(\mathcal{T}^1W(Y)) \arrow[dr, "\kappa^1_Y"] & \\
\mathcal{T}^1W(X) \arrow[rrrr, "\mathcal{T}^1W(f)"'] & & {} \arrow[u, phantom, "\#"] & & \mathcal{T}^1W(Y)
\end{tikzcd}$$

We can conclude that the outer square of the diagram commutes with respect to the given element $s = \kappa^0_X \circ W(\iota_a)(\sigma_0) \in \mathcal{T}^0 W(X)$, which completes the induction step. □

Given a normal PO-dilator W, we will write $\mathcal{T}W$ for 'its' Kruskal derivative, even though the latter is only determined up to isomorphism. The following is an immediate consequence of Proposition 2.7.

Corollary 4.5. *If W is a normal WPO-dilator, then so is its Kruskal derivative $\mathcal{T}W$.*

In the following we will consider iterated Kruskal derivatives. To ensure that the iterations are essentially unique, we now show that equivalent PO-dilators have equivalent Kruskal derivatives.

Proposition 4.6. *Consider a natural isomorphism $\eta : V \Rightarrow W$ between normal PO-dilators. If $(\mathcal{T}W, \iota, \kappa)$ is a Kruskal derivative of W, then $(\mathcal{T}W, \iota, \kappa \circ \eta)$ is a Kruskal derivative of V, where $\kappa \circ \eta$ is defined by $(\kappa \circ \eta)_X = \kappa_X \circ \eta_{\mathcal{T}W(X)}$.*

Proof. It is straightforward to verify that $(\mathcal{T}W(X), \iota_X, (\kappa \circ \eta)_X)$ is an initial Kruskal fixed point of V over X, for any partial order X. To provide a representative part of the verification, we show that $(\kappa \circ \eta)_X(\sigma) \leq_{\mathcal{T}W(X)} (\kappa \circ \eta)_X(\tau)$ is equivalent to

$$\sigma \leq_{V(\mathcal{T}W(X))} \tau \quad \text{or} \quad (\kappa \circ \eta)_X(\sigma) \leq^{\mathrm{fin}}_{\mathcal{T}W(X)} \mathrm{supp}^V_{\mathcal{T}W(X)}(\tau),$$

as required by Definition 3.1. Since $(\mathcal{T}W(X), \iota_X, \kappa_X)$ is a Kruskal fixed point of W over X, the same definition entails that $(\kappa \circ \eta)_X(\sigma) \leq_{\mathcal{T}W(X)} (\kappa \circ \eta)_X(\tau)$ is equivalent to the disjunction of $\eta_{\mathcal{T}W(X)}(\sigma) \leq_{W(\mathcal{T}W(X))} \eta_{\mathcal{T}W(X)}(\tau)$ and

$$(\kappa \circ \eta)_X(\sigma) \leq^{\mathrm{fin}}_{\mathcal{T}W(X)} \mathrm{supp}^W_{\mathcal{T}W(X)}(\eta_{\mathcal{T}W(X)}(\tau)).$$

The first disjunct is equivalent to $\sigma \leq_{V(\mathcal{T}W(X))} \tau$. To relate the second disjuncts, it suffices to recall that we have

$$\mathrm{supp}^W_{\mathcal{T}W(X)}(\eta_{\mathcal{T}W(X)}(\tau)) = \mathrm{supp}^V_{\mathcal{T}W(X)}(\tau).$$

To conclude that $(\mathcal{T}W, \iota, \kappa \circ \eta)$ is a Kruskal derivative of V over X, it suffices to observe that the diagram

$$\begin{tikzcd}
X \arrow[r, "\iota_X"] \arrow[d, "f"'] & \mathcal{T}W(X) \arrow[d, "\mathcal{T}W(f)"] & W(\mathcal{T}W(X)) \arrow[l, "\kappa_X"'] \arrow[d, "W(\mathcal{T}W(f))"] & V(\mathcal{T}W(X)) \arrow[l, "\eta_{\mathcal{T}W(X)}"'] \arrow[d, "V(\mathcal{T}W(f))"] \\
Y \arrow[r, "\iota_Y"'] & \mathcal{T}W(Y) & W(\mathcal{T}W(Y)) \arrow[l, "\kappa_Y"] & V(\mathcal{T}W(Y)) \arrow[l, "\eta_{\mathcal{T}W(Y)}"]
\end{tikzcd}$$

commutes for any quasi embedding $f : X \to Y$ (cf. Definition 4.1). □

In the introduction, we have sketched a construction of normal WPO-dilators $\mathbb{T}_n$ by recursion over n. We can now make this rigorous, by showing that the recursive clauses (steps (1)–(3) from the introduction) can be satisfied and determine the result of the recursion up to natural isomorphism.

According to step (1) from the introduction, $\mathbb{T}_0$ is the identity functor on the category of partial orders, viewed as a normal WPO-dilator. Let us show that there is, indeed, a unique extension of the identity functor into such a dilator: To establish existence, we define support functions $\mathrm{supp}^{\mathbb{T}_0}_X : \mathbb{T}_0(X) = X \to [X]^{<\omega}$ by setting $\mathrm{supp}^{\mathbb{T}_0}_X(x) = \{x\}$. One readily verifies that the relevant conditions (see Definitions 2.1 and 2.3) are satisfied. Uniqueness follows from the more general observation in Remark 4.3.

To prepare the recursion step, we give precise definitions of the finite multiset construction and of the composition of PO-dilators. Let us write $X^{<\omega}$ for the set of finite sequences $\langle x_0, \dots, x_{n-1}\rangle$ with entries $x_i \in X$. We say that two sequences $\langle x_0, \dots, x_{m-1}\rangle$ and $\langle y_0, \dots, y_{n-1}\rangle$ in $X^{<\omega}$ are equivalent if, and only if, there is a bijective function $h : \{0, \dots, m-1\} \to \{0, \dots, n-1\}$ such that we have $x_i = y_{h(i)}$ for all $i < m = n$. Let us write $[x_0, \dots, x_{n-1}]$ for the equivalence class of $\langle x_0, \dots, x_{n-1}\rangle$ with respect to this equivalence relation. From $[x_0, \dots, x_{n-1}]$ one can recover the multiplicity but not the order of the entries. The quotient set

$$M(X) = \{[x_0, \dots, x_{n-1}] \mid \langle x_0, \dots, x_{n-1}\rangle \in X^{<\omega}\}$$

is called the set of finite multisets with elements from X. We declare that

$$[x_0, \dots, x_{m-1}] \leq_{M(X)} [y_0, \dots, y_{n-1}]$$

holds if, and only if, there is an injection $g : \{0, \dots, m-1\} \to \{0, \dots, n-1\}$ such that we have $x_i \leq_X y_{g(i)}$ for all $i < m$. One can check that this is well defined and yields a partial order on $M(X)$ (for antisymmetry, use induction on the number of elements). Higman's lemma entails that $M(X)$ is a well partial order if the same holds for X. Given a (quasi) embedding $f : X \to Y$, one can define a (quasi) embedding $M(f) : M(X) \to M(Y)$ by setting

$$M(f)([x_0, \dots, x_{n-1}]) = [f(x_0), \dots, f(x_{n-1})].$$

A family of functions $\mathrm{supp}^M_X : M(X) \to [X]^{<\omega}$ can be given by

$$\mathrm{supp}^M_X([x_0, \dots, x_{n-1}]) = \{x_0, \dots, x_{n-1}\}.$$

It is straightforward to check that this turns M into a normal WPO-dilator in the sense of Definitions 2.1 and 2.3. In order to compose PO-dilators V and W one first takes their composition as functors. To get a PO-dilator, one defines functions $\mathrm{supp}^{V\circ W}_X : V \circ W(X) \to [X]^{<\omega}$ by setting

$$\mathrm{supp}^{V\circ W}_X(\sigma) = \bigcup\{\mathrm{supp}^W_X(s) \mid s \in \mathrm{supp}^V_{W(X)}(\sigma)\}.$$

The composition $V \circ W$ is normal (is a WPO-dilator, respectively) if V and W have the same property. If there are natural isomorphisms between V and V' and between W and W', then there is one between $V \circ W$ and $V' \circ W'$.

We can now come to the recursion step. Inductively, we may assume that the construction has produced a normal WPO-dilator $\mathbb{T}_n$ that is essentially unique (i. e. determined up to natural isomorphism). As we have just seen, it follows that the composition $M \circ \mathbb{T}_n$ is a normal WPO-dilator and essentially unique as well. Following step (2) from the introduction, we want to define $\mathbb{T}^-_{n+1}$ as the Kruskal derivative of $M \circ \mathbb{T}_n$. In order to justify this, we need to invoke several of our results: First, Theorem 4.2 tells us that $M \circ \mathbb{T}_n$ has a Kruskal derivative $(\mathbb{T}^-_{n+1}, \iota, \kappa)$. Second, Theorem 4.4 and Proposition 4.6 ensure that the first component $\mathbb{T}^-_{n+1}$ of this derivative is essentially unique (note that Proposition 4.6 is needed since $\mathbb{T}_n$ was only determined up to isomorphism). Finally, Corollary 4.5 shows that $\mathbb{T}^-_{n+1}$ is a normal WPO-dilator. The rest of the recursion step is straightforward: Following step (3) from the introduction, we define $\mathbb{T}_{n+1}$ as the composition $\mathbb{T}_n \circ \mathbb{T}^-_{n+1}$. The latter is still a normal WPO-dilator and essentially unique.

5. The Gap Orders as PO-Dilators

At the end of the previous section, we have given a rigorous version of steps (1)–(3) from the introduction, which describe a recursive construction of normal WPO-dilators $\mathbb{T}_n$. In the remainder of this paper, we will show that the resulting orders $\mathbb{T}_n(\emptyset)$ are isomorphic to the set of n-trees with Friedman's strong gap condition, as claimed in step (4) from the introduction. Our approach is as follows: In the present section, we give an ad hoc definition of PO-dilators $\mathbb{T}_n$ for which the relation with the gap condition is clear. It is a slight abuse of notation that we write $\mathbb{T}_n$ in both the recursive construction and our ad hoc definition. However, there is no actual danger of confusion, since the present section is exclusively concerned with our ad hoc definition. In Section 6, we will show that the PO-dilators $\mathbb{T}_n$ from our ad hoc definition satisfy the clauses of the recursive construction, so that both versions of $\mathbb{T}_n$ coincide after all (up to natural isomorphism).

In order to define PO-dilators $\mathbb{T}_n$, we first specify transformations $X \mapsto \mathbb{T}_n(X)$ of partial orders. Intuitively, the underlying set of the partial order $\mathbb{T}_n(X)$ consists of the finite trees with labels in $\{0, \ldots, n-1\} \cup X$, where labels from X may only occur at the leaves. More formally, this set admits the following recursive description:

Definition 5.1. *Given a number $n \in \mathbb{N}$ and a partial order X, we generate a set $\mathbb{T}_n(X)$ by the following recursive clauses:*

(i) For each $x \in X$ we have an element $\overline{x} \in \mathbb{T}_n(X)$.
(ii) Whenever we have constructed an element $\sigma = [t_0, \ldots, t_{m-1}] \in M(\mathbb{T}_n(X))$, we add an element $i \star \sigma \in \mathbb{T}_n(X)$ for each natural number $i < n$.

Let us also define

$$\mathbb{T}_n^-(X) = \{\overline{x} \mid x \in X\} \cup \{0 \star \sigma \mid \sigma \in M(\mathbb{T}_n(X))\} \subseteq \mathbb{T}_n(X),$$

provided that we have $n > 0$.

We define height functions $h_X^n : \mathbb{T}_n(X) \to \mathbb{N}$ by the recursive clauses

$$h_X^n(\overline{x}) = 0, \qquad h_X^n(i \star [t_0, \ldots, t_{m-1}]) = \max(\{0\} \cup \{h_X^n(t_k) + 1 \mid k < m\}).$$

The following definition decides $s \leq_{\mathbb{T}_n(X)} t$ by recursion on $h_X^n(s) + h_X^n(t)$.

Definition 5.2. *To define a binary relation $\leq_{\mathbb{T}_n(X)}$ on the set $\mathbb{T}_n(X)$ we stipulate*

$$\overline{x} \leq_{\mathbb{T}_n(X)} t \quad\Leftrightarrow\quad \begin{cases} \text{either } t = \overline{y} \text{ with } x \leq_X y, \\ \text{or } t = j \star [t_0, \ldots, t_{m-1}] \text{ and } \overline{x} \leq_{\mathbb{T}_n(X)} t_l \text{ for some } l < m, \end{cases}$$

$$i \star \sigma \leq_{\mathbb{T}_n(X)} t \quad\Leftrightarrow\quad \begin{cases} t = i \star \tau \text{ with } \sigma \leq_{M(\mathbb{T}_n(X))} \tau, \text{ or } t = j \star [t_0, \ldots, t_{m-1}] \\ \text{with } j \geq i \text{ and } i \star \sigma \leq_{\mathbb{T}_n(X)} t_l \text{ for some } l < m. \end{cases}$$

In the case of $n > 0$, we define $\leq_{\mathbb{T}_n^-(X)}$ as the restriction of $\leq_{\mathbb{T}_n(X)}$ to $\mathbb{T}_n^-(X)$.

A straightforward induction shows

$$s \leq_{\mathbb{T}_n(X)} t \quad\Rightarrow\quad h_X^n(s) \leq h_X^n(t).$$

Similarly to the proof of Proposition 2.6, one can deduce that $\leq_{\mathbb{T}_n(X)}$ is a partial order on $\mathbb{T}_n(X)$. In the introduction we have given the usual definition of Friedman's gap condition for embeddings of n-trees. As promised, the connection with Definition 5.2 is rather straightforward:

Proposition 5.3. *The partial order $\mathbb{T}_n(\emptyset)$ is isomorphic to the set of n-trees, ordered according to Friedman's strong gap condition.*

Proof. For $s = i \star [s(0), \ldots, s(k-1)] \in \mathbb{T}_n(\emptyset)$, we recursively define T_s as the n-tree with root label i and immediate subtrees $T_{s(0)}, \ldots, T_{s(k-1)}$. Clearly this yields a bijection. By induction on $h_X^n(s) + h_X^n(t)$ one can show that $s \leq_{\mathbb{T}_n(\emptyset)} t$ holds if, and only if, there is an embedding $f : T_s \to T_t$ that satisfies Friedman's gap condition. An inequality

$$s = i \star \sigma = i \star [s(0), \dots, s(k-1)] \leq_{\mathbb{T}_n(\emptyset)} i \star [t(0), \dots, t(m-1)] = i \star \tau = t$$

that holds because of $\sigma \leq_{M(\mathbb{T}_n(X))} \tau$ corresponds to an embedding $f : T_s \to T_t$ that maps the root to the root. Indeed, the inequalities $s(j) \leq_{\mathbb{T}_n(\emptyset)} t(l_j)$ that witness $\sigma \leq_{M(\mathbb{T}_n(X))} \tau$ correspond to the restrictions

$$f_j = f \restriction T_{s(j)} : T_{s(j)} \to T_{t(l_j)} \subseteq T_t.$$

At this point, it is crucial that we consider the strong gap condition: Writing root (T) for the root of T, the gap below $f_j(\text{root}\,(T_{s(j)}))$ in $T_{t(l_j)}$ corresponds to the gap between $f(\text{root}\,(T_s))$ and $f(\text{root}\,(T_{s(j)}))$ in T_t. An inequality

$$s = i \star \sigma \leq_{\mathbb{T}_n(\emptyset)} j \star [t_0, \dots, t_{m-1}] = t$$

that holds because of $j \geq i$ and $s \leq_{\mathbb{T}_n(X)} t_l$ with $l < m$ corresponds to an embedding $f : T_s \to T_t$ with range contained in $T_{t(l)} \subseteq T_t$. The condition $j \geq i$ accounts for the fact that root (T_t) lies in the gap below $f(\text{root}\,(T_s))$ in T_t but not in $T_{t(l)}$. $\square$

Our next goal is to extend $\mathbb{T}_n$ and $\mathbb{T}^-_{n+1}$ into PO-dilators.

Definition 5.4. *Given a quasi embedding $f : X \to Y$ between partial orders, we define a function $\mathbb{T}_n(f) : \mathbb{T}_n(X) \to \mathbb{T}_n(Y)$ by the recursive clauses*

$$\mathbb{T}_n(f)(\overline{x}) = \overline{f(x)}, \qquad \mathbb{T}_n(f)(i \star [t_0, \dots, t_{m-1}]) = i \star [\mathbb{T}_n(f)(t_0), \dots, \mathbb{T}_n(f)(t_{m-1})].$$

For $n > 0$ we observe that $\mathbb{T}_n(f)$ restricts to $\mathbb{T}^-_n(f) : \mathbb{T}^-_n(X) \to \mathbb{T}^-_n(Y)$. We also define a family of functions $\text{supp}^{\mathbb{T}_n}_X : \mathbb{T}_n(X) \to [X]^{<\omega}$ by stipulating

$$\text{supp}^{\mathbb{T}_n}_X(\overline{x}) = \{x\}, \qquad \text{supp}^{\mathbb{T}_n}_X(i \star [t_0, \dots, t_{m-1}]) = \bigcup \{\text{supp}^{\mathbb{T}_n}_X(t_l) \mid l < m\}.$$

We will write $\text{supp}^{\mathbb{T}^-_n}_X$ for the restriction of $\text{supp}^{\mathbb{T}_n}_X$ to $\mathbb{T}^-_n(X)$.

Let us verify that we obtain the desired structure:

Proposition 5.5. *The previous definition yields normal PO-dilators $\mathbb{T}_n$ and $\mathbb{T}^-_{n+1}$.*

Proof. Given a quasi embedding f, an easy induction on $h^n_X(s) + h^n_X(t)$ shows

$$\mathbb{T}_n(f)(s) \leq_{\mathbb{T}_n(Y)} \mathbb{T}_n(f)(t) \quad \Rightarrow \quad s \leq_{\mathbb{T}_n(X)} t.$$

If f is an embedding, then the converse implication holds as well. Also by induction, one readily checks that $\mathbb{T}_n$ is a functor and that $\text{supp}^{\mathbb{T}_n}$ is a natural transformation. In order to conclude that $\mathbb{T}_n$ is a PO-dilator, one needs to establish the support condition from part (ii) of Definition 2.1. By induction on s, one can indeed show

$$\text{supp}^{\mathbb{T}_n}_Y(s) \subseteq \text{rng}\,(f) \quad \Rightarrow \quad s \in \text{rng}\,(\mathbb{T}_n(f))$$

for $s \in \mathbb{T}_n(Y)$, where $f : X \to Y$ is an embedding (recall that the converse implication is automatic). To see that $\mathbb{T}^-_{n+1}$ does also satisfy the support condition, one observes that $\mathbb{T}_{n+1}(f)(s) \in \mathbb{T}^-_{n+1}(Y)$ implies $s \in \mathbb{T}^-_{n+1}(X)$. To establish the normality condition from Definition 2.3, one verifies

$$s \leq_{\mathbb{T}_n(X)} t \quad \Rightarrow \quad \text{supp}^{\mathbb{T}_n}_X(s) \leq^{\text{fin}}_X \text{supp}^{\mathbb{T}_n}_X(t)$$

by induction on $h^n_X(s) + h^n_X(t)$. $\square$

Corollary 6.6 below establishes the stronger result that $\mathbb{T}_n$ and $\mathbb{T}^-_{n+1}$ are normal WPO-dilators. In view of Proposition 5.3 this implies that the trees with Friedman's gap condition form a well partial order. To prove this fact, one needs a combination of Π^1_1-comprehension and Π^1_2-induction (cf. Simpson 1985).

6. Reconstructing the Gap Condition

In the present section, we show that the normal PO-dilators $\mathbb{T}_n$ that were defined in Section 5 satisfy the recursive clauses from the introduction. At the end of the section, we will explain how this completes our reconstruction of the gap condition.

To avoid misunderstanding, we stress that the following considerations refer to the PO-dilators $\mathbb{T}_n$ from Section 5. As explained in the first paragraph of that section, this involves a slight abuse of notation. The following maps π^n_X are closely related to the transformation $T \mapsto T^*$ from Section 4 of Simpson (1985).

Definition 6.1. *For each partial order X, we define* $\pi^n_X : \mathbb{T}_n \circ \mathbb{T}^-_{n+1}(X) \to \mathbb{T}_{n+1}(X)$ *by the recursive clauses*

$$\pi^n_X(\overline{t}) = t, \qquad \pi^n_X(i \star [s_0, \ldots, s_{m-1}]) = (i+1) \star [\pi^n_X(s_0), \ldots, \pi^n_X(s_{m-1})],$$

where the first clause relies on the inclusion $\mathbb{T}^-_{n+1}(X) \subseteq \mathbb{T}_{n+1}(X)$.

Intuitively speaking, an element of $\mathbb{T}_n \circ \mathbb{T}^-_{n+1}(X)$ is a finite tree with labels from the set $\{0, \ldots, n-1\} \cup \mathbb{T}^-_{n+1}(X)$, where the labels from $\mathbb{T}^-_{n+1}(X)$ can only occur at leaves. The function π^n_X increases the labels from $\{0, \ldots, n-1\}$ and "unravels" the leaf labels. Hence, the leaves of $s \in \mathbb{T}_n \circ \mathbb{T}^-_{n+1}(X)$ correspond to the minimal nodes of $\pi^n_X(s) \in \mathbb{T}_{n+1}(X)$ that have a label in $\{0\} \cup X$. The following can be seen as a strengthening of Lemma 4.5 of Simpson (1985), which essentially proves that the components π^n_X are order preserving.

Proposition 6.2. *The family* $\pi^n : \mathbb{T}_n \circ \mathbb{T}^-_{n+1} \Rightarrow \mathbb{T}_{n+1}$ *is a natural isomorphism.*

Proof. To show that π^n_X is surjective, we verify $t \in \operatorname{rng}(\pi^n_X)$ by induction over the recursive construction of $t \in \mathbb{T}_{n+1}(X)$ according to Definition 5.1 (alternatively by induction over $h^{n+1}_X(t) \in \mathbb{N}$). If t is of the form $\overline{x}$ or $0 \star \sigma$, then we have $t \in \mathbb{T}^-_{n+1}(X)$, so that we obtain $\overline{t} \in \mathbb{T}_n \circ \mathbb{T}^-_{n+1}(X)$ as well as $t = \pi^n_X(\overline{t}) \in \operatorname{rng}(\pi^n_X)$. Let us now consider an element of the form $t = (i+1) \star [t_0, \ldots, t_{m-1}]$, with $i+1 < n+1$ and $t_l \in \mathbb{T}_{n+1}(X)$ for $l < m$. Inductively we get $t_l = \pi^n_X(s_l)$, which yields $i \star [s_0, \ldots, s_{m-1}] \in \mathbb{T}_n \circ \mathbb{T}^-_{n+1}(X)$ and

$$t = (i+1) \star [t_0, \ldots, t_{m-1}] = \pi^n_X(i \star [s_0, \ldots, s_{m-1}]) \in \operatorname{rng}(\pi^n_X).$$

To conclude that π^n_X is an isomorphism, we show

$$s \leq_{\mathbb{T}_n \circ \mathbb{T}^-_{n+1}(X)} t \quad \Leftrightarrow \quad \pi^n_X(s) \leq_{\mathbb{T}_{n+1}(X)} \pi^n_X(t)$$

by induction on $h^n_{\mathbb{T}^-_{n+1}(X)}(s) + h^n_{\mathbb{T}^-_{n+1}(X)}(t)$. For $s = \overline{s'}$ and $t = \overline{t'}$ it suffices to invoke Definition 5.2. Now consider $s = \overline{s'}$ and $t = j \star [t_0, \ldots, t_{m-1}]$. Inductively we get

$$s \leq_{\mathbb{T}_n \circ \mathbb{T}^-_{n+1}(X)} t \quad \Leftrightarrow \quad \pi^n_X(s) \leq_{\mathbb{T}_{n+1}(X)} \pi^n_X(t_l) \text{ for some } l < m.$$

Note that $\pi^n_X(s) = s' \in \mathbb{T}^-_{n+1}(X)$ must be of the form $\overline{x}$ or $0 \star \sigma$. In view of $j+1 \neq 0$ and $j+1 \geq 0$, the right side of the previous equivalence is thus equivalent to

$$\pi^n_X(s) \leq_{\mathbb{T}_{n+1}(X)} (j+1) \star [\pi^n_X(t_0), \ldots, \pi^n_X(t_{m-1})] = \pi^n_X(t).$$

For $s = i \star [s_0, \ldots, s_{k-1}]$ and $t = \overline{t'}$, we cannot have $s \leq_{\mathbb{T}_n \circ \mathbb{T}^-_{n+1}(X)} t$. We also see

$$\pi_X^n(s) = (i+1) \star [\pi_X^n(s_0), \ldots, \pi_X^n(s_{k-1})] \not\leq_{\mathbb{T}_{n+1}(X)} t' = \pi_X^n(t),$$

since an inequality would require $t' = j \star [t_0, \ldots, t_{m-1}]$ with $j \geq i+1$, in contrast to $t' \in \mathbb{T}_{n+1}^-(X)$. For $s = i \star [s_0, \ldots, s_{k-1}]$ and $t = j \star [t_0, \ldots, t_{m-1}]$ the claim is readily deduced from the induction hypothesis (due to $i \geq j \Leftrightarrow i+1 \geq j+1$). To complete the proof, we verify the naturality property

$$\pi_Y^n \circ (\mathbb{T}_n \circ \mathbb{T}_{n+1}^-)(f)(t) = \mathbb{T}_{n+1}(f) \circ \pi_X^n(t),$$

arguing by induction over (the height of) the element $t \in \mathbb{T}_n \circ \mathbb{T}_{n+1}^-(X)$. For $t = \bar{s}$ we compute

$$\pi_Y^n \circ (\mathbb{T}_n \circ \mathbb{T}_{n+1}^-)(f)(t) = \pi_Y^n \circ \mathbb{T}_n(\mathbb{T}_{n+1}^-(f))(\bar{s}) = \pi_Y^n(\overline{\mathbb{T}_{n+1}^-(f)(s)}) =$$
$$= \mathbb{T}_{n+1}^-(f)(s) = \mathbb{T}_{n+1}(f)(s) = \mathbb{T}_{n+1}(f) \circ \pi_X^n(t).$$

The induction step for $t = j \star [t_0, \ldots, t_{m-1}]$ is straightforward. □

The following lemma will be needed below. Intuitively, the equivalence says that a tree with root label 0 can be embedded into another tree if, and only if, it can be embedded into a subtree with root label 0. This is true because the gap condition below a node with label 0 is automatic.

Lemma 6.3. *We have*

$$s \leq_{\mathbb{T}_{n+1}(X)} \pi_X^n(t) \quad \Leftrightarrow \quad s \leq^{\mathrm{fin}}_{\mathbb{T}_{n+1}^-(X)} \operatorname{supp}^{\mathbb{T}_n}_{\mathbb{T}_{n+1}^-(X)}(t)$$

for all $s \in \mathbb{T}_{n+1}^-(X)$ and all $t \in \mathbb{T}_n \circ \mathbb{T}_{n+1}^-(X)$.

Proof. We establish the claim by induction over (the height of) t. For $t = \overline{t'}$ it suffices to observe that we have $\pi_X^n(t) = t'$ and $\operatorname{supp}^{\mathbb{T}_n}_{\mathbb{T}_{n+1}^-(X)}(t) = \{t'\}$. To prove the claim for $t = j \star [t_0, \ldots, t_{m-1}]$, we recall a step from the previous proof: For $s \in \mathbb{T}_{n+1}^-(X)$ we have observed

$$s \leq_{\mathbb{T}_{n+1}(X)} \pi_X^n(t) \quad \Leftrightarrow \quad s \leq_{\mathbb{T}_{n+1}(X)} \pi_X^n(t_l) \text{ for some } l < m.$$

Together with

$$\operatorname{supp}^{\mathbb{T}_n}_{\mathbb{T}_{n+1}^-(X)}(t) = \bigcup \{\operatorname{supp}^{\mathbb{T}_n}_{\mathbb{T}_{n+1}^-(X)}(t_l) \mid l < m\},$$

this reduces the claim to the induction hypothesis. □

The central result of our reconstruction states that $\mathbb{T}_{n+1}^-$ is a Kruskal derivative of $M \circ \mathbb{T}_n$, where M is the finite multiset dilator (cf. the end of Section 4). In view of Definition 4.1, we need the following additional structure:

Definition 6.4. *For any partial order X we define a function $\iota_X^n : X \to \mathbb{T}_{n+1}^-(X)$ by $\iota_X^n(x) = \bar{x}$. To define $\kappa_X^n : M \circ \mathbb{T}_n \circ \mathbb{T}_{n+1}^-(X) \to \mathbb{T}_{n+1}^-(X)$ we stipulate*

$$\kappa_X^n([s_0, \ldots, s_{m-1}]) = 0 \star [\pi_X^n(s_0), \ldots, \pi_X^n(s_{m-1})],$$

for $s_0, \ldots, s_{m-1} \in \mathbb{T}_n \circ \mathbb{T}_{n+1}^-(X)$. We will write ι^n and κ^n for the families of functions ι_X^n and κ_X^n that are indexed by the partial order X.

Let us now prove the promised result:

Theorem 6.5. *For any number $n \in \mathbb{N}$, the tuple $(\mathbb{T}_{n+1}^-, \iota^n, \kappa^n)$ is a Kruskal derivative of the normal PO-dilator $M \circ \mathbb{T}_n$.*

Proof. From Proposition 5.5 we know that $\mathbb{T}_{n+1}^-$ is a normal PO-dilator. It remains to verify conditions (i) and (ii) from Definition 4.1. Let us begin by showing that the tuple $(\mathbb{T}_{n+1}^-(X), \iota_X^n, \kappa_X^n)$ is a Kruskal fixed point of $M \circ \mathbb{T}_n$ over X, for each partial order X. In order to establish the first and third condition from Definition 3.1, we invoke Definition 5.2 to get

$$\iota_X^n(x) = \overline{x} \leq_{\mathbb{T}_{n+1}^-(X)} \overline{y} = \iota_X^n(y) \quad \Leftrightarrow \quad x \leq_X y,$$
$$\kappa_X^n([s_0, \ldots, s_{k-1}]) = 0 \star [\pi_X^n(s_0), \ldots, \pi_X^n(s_{k-1})] \not\leq_{\mathbb{T}_{n+1}^-(X)} \overline{y} = \iota_X^n(y).$$

To verify the second and fourth condition, we recall the definition of composition from the end of Section 4. It reveals that an element $\tau = [t_0, \ldots, t_{m-1}] \in M \circ \mathbb{T}_n \circ \mathbb{T}_{n+1}^-(X)$ has support

$$\operatorname{supp}_{\mathbb{T}_{n+1}^-(X)}^{M \circ \mathbb{T}_n}(\tau) = \bigcup \{\operatorname{supp}_{\mathbb{T}_{n+1}^-(X)}^{\mathbb{T}_n}(t) \mid t \in \operatorname{supp}_{\mathbb{T}_n \circ \mathbb{T}_{n+1}^-(X)}^M(\tau)\} = \bigcup \{\operatorname{supp}_{\mathbb{T}_{n+1}^-(X)}^{\mathbb{T}_n}(t_l) \mid l < m\}.$$

For $s \in \mathbb{T}_{n+1}^-(X)$, we can thus invoke Lemma 6.3 to get

$$s \leq_{\mathbb{T}_{n+1}^-(X)}^{\text{fin}} \operatorname{supp}_{\mathbb{T}_{n+1}^-(X)}^{M \circ \mathbb{T}_n}(\tau) \quad \Leftrightarrow \quad s \leq_{\mathbb{T}_{n+1}(X)} \pi_X^n(t_l) \text{ for some } l < m.$$

For $s = \overline{x}$, we now see that the second condition from Definition 3.1 requires that

$$\iota_X^n(x) = \overline{x} \leq_{\mathbb{T}_{n+1}^-(X)} 0 \star [\pi_X^n(t_0), \ldots, \pi_X^n(t_{m-1})] = \kappa_X^n(\tau)$$

holds if, and only if, we have $\overline{x} \leq_{\mathbb{T}_{n+1}^-(X)} \pi_X^n(t_l)$ for some $l < m$. This is true by Definition 5.2. For $\sigma = [s_0, \ldots, s_{k-1}]$, the fourth and final condition from Definition 3.1 requires that

$$\kappa_X^n(\sigma) = 0 \star [\pi_X^n(s_0), \ldots, \pi_X^n(s_{k-1})] \leq_{\mathbb{T}_{n+1}^-(X)} 0 \star [\pi_X^n(t_0), \ldots, \pi_X^n(t_{m-1})] = \kappa_X^n(\tau)$$

is equivalent to the disjunction

$$\sigma \leq_{M \circ \mathbb{T}_n \circ \mathbb{T}_{n+1}^-(X)} \tau \quad \text{or} \quad \kappa_X^n(\sigma) \leq_{\mathbb{T}_{n+1}(X)} \pi_X^n(t_l) \text{ for some } l < m.$$

To reduce this to Definition 5.2, it suffices to note that we have

$$\sigma \leq_{M \circ \mathbb{T}_n \circ \mathbb{T}_{n+1}^-(X)} \tau \;\Leftrightarrow\; [\pi_X^n(s_0), \ldots, \pi_X^n(s_{k-1})] \leq_{M \circ \mathbb{T}_{n+1}(X)} [\pi_X^n(t_0), \ldots, \pi_X^n(t_{m-1})],$$

as π_X^n is an embedding. To show that each Kruskal fixed point $(\mathbb{T}_{n+1}^-(X), \iota_X^n, \kappa_X^n)$ is initial, we use the criterion from Theorem 3.5. Since $\pi_X^n : \mathbb{T}_n \circ \mathbb{T}_{n+1}^-(X) \to \mathbb{T}_{n+1}(X)$ is surjective, we indeed get $\operatorname{rng}(\iota_X^n) \cup \operatorname{rng}(\kappa_X^n) = \mathbb{T}_{n+1}^-(X)$. Now recall the function $h_X^{n+1} : \mathbb{T}_{n+1}(X) \to \mathbb{N}$ that was specified before the statement of Defintion 5.2 above. We will also write h_X^{n+1} for the restriction of this function to $\mathbb{T}_{n+1}^-(X) \subseteq \mathbb{T}_{n+1}(X)$. In order to apply Theorem 3.5, we need to establish

$$s \in \operatorname{supp}_{\mathbb{T}_{n+1}^-(X)}^{M \circ \mathbb{T}_n}(\tau) \quad \Rightarrow \quad h_X^{n+1}(s) < h_X^{n+1}(\kappa_X^n(\tau))$$

for $s \in \mathbb{T}_{n+1}^-(X)$ and $\tau \in M \circ \mathbb{T}_n \circ \mathbb{T}_{n+1}^-(X)$. So assume $s \in \operatorname{supp}_{\mathbb{T}_{n+1}^-(X)}^{M \circ \mathbb{T}_n}(\tau)$ with $\tau = [t_0, \ldots, t_{m-1}]$. By the above we get $s \in \operatorname{supp}_{\mathbb{T}_{n+1}^-(X)}^{\mathbb{T}_n}(t_l)$ for some $l < m$. Then Lemma 6.3 yields $s \leq_{\mathbb{T}_{n+1}(X)} \pi_X^n(t_l)$. As observed after Definition 5.2, this implies $h_X^{n+1}(s) \leq h_X^{n+1}(\pi_X^n(t_l))$ and hence

$$h_X^{n+1}(s) < h_X^{n+1}(0 \star [\pi_X^n(t_0), \ldots, \pi_X^n(t_{m-1})]) = h_X^{n+1}(\kappa_X^n(\tau)).$$

We have now verified all conditions from Definition 3.1 and Theorem 3.5, which shows that $(\mathbb{T}_{n+1}^-(X), \iota_X^n, \kappa_X^n)$ is an initial Kruskal fixed point of $M \circ \mathbb{T}_n$ over X. To conclude that $(\mathbb{T}_{n+1}^-, \iota^n, \kappa^n)$ is a Kruskal derivative of $M \circ \mathbb{T}_n$, it remains to establish condition (ii) from Definition 4.1. Given a quasi embedding $f : X \to Y$, we first compute

$$\iota_Y^n \circ f(x) = \overline{f(x)} = \mathbb{T}_{n+1}(f)(\overline{x}) = \mathbb{T}_{n+1}^-(f) \circ \iota_X^n(x).$$

For $\tau = [t_0, \ldots, t_{m-1}] \in M \circ \mathbb{T}_n \circ \mathbb{T}_{n+1}^-(X)$ we also get

$$\begin{aligned}\kappa^n_Y\circ(M\circ\mathbb{T}_n)(\mathbb{T}^-_{n+1}(f))(\tau)&=\kappa^n_Y([(\mathbb{T}_n\circ\mathbb{T}^-_{n+1})(f)(t_0),\ldots,(\mathbb{T}_n\circ\mathbb{T}^-_{n+1})(f)(t_{m-1})])\\&=0\star[\pi^n_Y\circ(\mathbb{T}_n\circ\mathbb{T}^-_{n+1})(f)(t_0),\ldots,\pi^n_Y\circ(\mathbb{T}_n\circ\mathbb{T}^-_{n+1})(f)(t_{m-1})]\\&=0\star[\mathbb{T}_{n+1}(f)\circ\pi^n_X(t_0),\ldots,\mathbb{T}_{n+1}(f)\circ\pi^n_X(t_{m-1})]\\&=\mathbb{T}_{n+1}(f)(0\star[\pi^n_X(t_0),\ldots,\pi^n_X(t_{m-1})])=\mathbb{T}^-_{n+1}(f)\circ\kappa^n_X(\tau),\end{aligned}$$

just as required by Definition 4.1. □

To conclude this paper, we explain how the previous results complete our reconstruction of the gap condition due to Harvey Friedman: In the introduction, we have described a construction of normal WPO-dilators $\mathbb{T}_n$ by recursion over n. A rigorous version of this construction has been given at the end of Section 4. What remains to be shown is that the order $\mathbb{T}_n(\emptyset)$ that arises from this construction is isomorphic to the set of n-trees, ordered according to Friedman's strong gap condition. It may appear that this is the result of Proposition 5.3, but this impression results from an abuse of notation: The cited proposition does not refer to our recursive construction, but rather to an ad hoc definition of PO-dilators $\mathbb{T}_n$ in Section 5. In order to obtain the desired result, we must show that our abuse of notation is permissible, or in other words: that the PO-dilators $\mathbb{T}_n$ from the recursive construction coincide with those defined in Section 5, up to natural isomorphism. In Section 4 we have seen that the result of the recursive construction is essentially unique. For this reason, it suffices to show the following:

> *The normal PO-dilators $\mathbb{T}_n$ defined in Section 5 satisfy the recursive clauses given by steps (1)–(3) from the introduction, at least up to natural isomorphism.*

Step (1) demands that $\mathbb{T}_0$ is the identity functor on the category of partial orders, at least up to natural isomorphism. Since clause (ii) of Definition 5.1 does not apply for $n=0$, the required isomorphism has components $X\ni x\mapsto\overline{x}\in\mathbb{T}_0(X)$. Step (2) demands that $\mathbb{T}^-_{n+1}$ is the Kruskal derivative of $M\circ\mathbb{T}_n$, which is true by Theorem 6.5. The isomorphism $\mathbb{T}_{n+1}\cong\mathbb{T}_n\circ\mathbb{T}^-_{n+1}$ required for step (3) is provided by Proposition 6.2. Now that we have justified our abuse of notation, we can indeed invoke Proposition 5.3 to complete step (4) from the introduction: the order $\mathbb{T}_n(\emptyset)$ that arises from our recursive construction is isomorphic to the set of n-trees, ordered according to Friedman's strong gap condition.

Finally, we deduce a result about the preservation of well partial orders, which was promised at the end of Section 5. Except for the formulation in terms of PO-dilators, it is due to Friedman (see Simpson 1985).

Corollary 6.6. *For each $n\in\mathbb{N}$, the normal PO-dilators $\mathbb{T}_n$ and $\mathbb{T}^-_{n+1}$ preserve well partial orders (which means that they are normal WPO-dilators). In particular, Friedman's gap condition yields a well partial order on the set of n-trees.*

Proof. By the discussion above, we may identify the PO-dilators $\mathbb{T}_n$ from Section 5 with those from our recursive construction. With this in mind, the result about $\mathbb{T}_n$ follows from the discussion at the end of Section 4 (i. e. essentially by iterated applications of Corollary 4.5). The result about $\mathbb{T}^-_{n+1}$ was implicit in that discussion. If a more explicit argument is demanded, one can simply invoke the fact that $\mathbb{T}^-_{n+1}(X)\subseteq\mathbb{T}_{n+1}(X)$ holds for any partial order X. The result about Friedman's (strong and a fortiori weak) gap condition follows by Proposition 5.3. □

Acknowledgements

I am very grateful to Jeroen van der Meeren, Michael Rathjen and Andreas Weiermann. I owe them many of the ideas that were fundamental for the present paper.

References

Dershowitz, N. (1982). Orderings for term-rewriting systems. *Theoretical Computer Science* **17** 279–301.

Ehrhard, T. and Regnier, L. (2003). The differential lambda-calculus. *Theoretical Computer Science* **309** 1–41.

Freund, A. (2019a). Π^1_1-comprehension as a well-ordering principle. *Advances in Mathematics* **355**. Article no. 106767, 65 pp.

Freund, A. (2019b). A categorical construction of Bachmann-Howard fixed points. *Bulletin of the London Mathematical Society* **51** (5) 801–814.

Freund, A. (2020). Computable aspects of the Bachmann-Howard principle. *Journal of Mathematical Logic* **20** (2). Article no. 2050006, 26 pp.

Freund, A. and Rathjen, M. (2021). Derivatives of normal functions in reverse mathematics. *Annals of Pure and Applied Logic* **172** (2). Article no. 102890, 49 pp.

Freund, A., Rathjen, M. and Weiermann, A. (2020). Minimal bad sequences are necessary for a uniform Kruskal theorem. Preprint available as arXiv:2001.06380.

Friedman, H., Robertson, N. and Seymour, P. (1987). Metamathematics of the graph minor theorem. In: Simpson, S. (ed.) *Logic and Combinatorics*, Contemporary Mathematics, vol. 65, American Mathematical Society, 229–261.

Gallier, J. (1991). What's so special about Kruskal's theorem and the ordinal Γ_0? A survey of some results in proof theory. *Annals of Pure and Applied Logic* **53** 199–260.

Girard, J.-Y. (1981). Π^1_2-logic, part 1: Dilators. *Annals of Pure and Applied Logic* **21** 75–219.

Gordeev, L. (1990). Generalizations of the Kruskal-Friedman theorems. *The Journal of Symbolic Logic* **55** (1) 157–181.

Haase, C., Schmitz, S. and Schnoebelen, P. (2013). The power of priority channel systems. In: D'Argenio, P. and Melgratti, H. (eds.) *CONCUR 2013 — Concurrency Theory*, Lecture Notes in Computer Science, vol. 8052, Springer, 319–333.

Hasegawa, R. (1997). An analysis of divisibility orderings and recursive path orderings. In: Shyamasundar, R. and Ueda, K. (eds.) *Advances in Computing Science — ASIAN'97*, Lecture Notes in Computer Science, vol. 1345, 283–296.

Hasegawa, R. (2002). Two applications of analytic functors. *Theoretical Computer Science* **272** 113–175.

Kříž, I. (1989). Well-quasiordering finite trees with gap-condition. Proof of Harvey Friedman's conjecture. *Annals of Mathematics* **130** (1) 215–226.

Kruskal, J. (1960). Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture. *Transactions of the American Mathematical Society* **95** (2) 210–225.

Nash-Williams, C. S. (1963). On well-quasi-ordering finite trees. *Proceedings of the Cambridge Philosophical Society* **59** 833–835.

Ogawa, M. (1995). Simple gap termination on term graph rewriting systems. Research report, RIMS, Kyoto University. http://www.kurims.kyoto-u.ac.jp/ kyodo/kokyuroku/contents/pdf/0918-10.pdf (accessed 23.09.2020).

Rathjen, M. (1999). The realm of ordinal analysis. In: Cooper, S. and Truss, J. (eds.) *Sets and Proofs*, London Mathematical Society Lecture Note Series, Cambridge University Press, 219–280.

Rathjen, M. and Weiermann, A. (1993). Proof-theoretic investigations on Kruskal's theorem. *Annals of Pure and Applied Logic* **60** 49–88.

Schmidt, D. (2020). Well-partial orderings and their maximal order types. In: Schuster, P., Seisenberger, M. and Weiermann, A. (eds.) *Well-Quasi Orders in Computation, Logic, Language and Reasoning*, Trends in Logic (Studia Logica Library), vol. 53, Cham, Springer, 351–391. Originally Habilitationsschrift, Universität Heidelberg, 1979.

Simpson, S. (1985). Nonprovability of certain combinatorial properties of finite trees. In: Harrington, L. A., Morley, M. D., Sčědrov, A., and Simpson, S. G. (eds.) *Harvey Friedman's Research on the Foundations of Mathematics*, Studies in Logic and the Foundations of Mathematics, vol. 117, North-Holland, 87–117.

Simpson, S. (2009). *Subsystems of Second Order Arithmetic*, Perspectives in Logic, Cambridge University Press, Cambridge.

Tzameret, I. (2002). *Kruskal-Friedman Gap Embedding Theorems over Well-Quasi-Orderings*. MSc thesis, Tel-Aviv University. www.cs.tau.ac.il/thesis/thesis/Tzameret-Iddo-MSc-Thesis.pdf (accessed 23.09.2020).

van der Meeren, J. (2015). *Connecting the Two Worlds: Well-Partial-Orders and Ordinal Notation Systems*. PhD thesis, Ghent University. https://biblio.ugent.be/publication/8094697/file/8094746.pdf (accessed 25.09.2020).

van der Meeren, J., Rathjen, M. and Weiermann, A. (2015). Well-partial-orderings and the big Veblen number. *Archive for Mathematical Logic* **54** (1–2) 193–230.

van der Meeren, J., Rathjen, M. and Weiermann, A. (2017a). An order-theoretic characterization of the Howard-Bachmann-hierarchy. *Archive for Mathematical Logic* **56** (1–2) 79–118.

van der Meeren, J., Rathjen, M. and Weiermann, A. (2017b). Ordinal notation systems corresponding to Friedman's linearized well-partial-orders with gap-condition. *Archive for Mathematical Logic* **56** 607–638.

Cite this article: Freund A (2020). From Kruskal's theorem to Friedman's gap condition. *Mathematical Structures in Computer Science* **30**, 952–975. https://doi.org/10.1017/S0960129520000298