



Bachmann–Howard derivatives

Anton Freund¹ 

Received: 31 May 2021 / Accepted: 11 October 2022 / Published online: 26 October 2022
© The Author(s) 2022

Abstract

It is generally accepted that H. Friedman’s gap condition is closely related to iterated collapsing functions from ordinal analysis. But what precisely is the connection? We offer the following answer: In a previous paper we have shown that the gap condition arises from an iterative construction on transformations of partial orders. Here we show that the parallel construction for linear orders yields familiar collapsing functions. The iteration step in the linear case is an instance of a general construction that we call ‘Bachmann–Howard derivative’. In the present paper, we focus on the unary case, i.e., on the gap condition for sequences rather than trees and, correspondingly, on addition-free ordinal notation systems. This is partly for convenience, but it also allows us to clarify a phenomenon that is specific to the unary setting: As shown by van der Meeren, Rathjen and Weiermann, the gap condition on sequences admits two linearizations with rather different properties. We will see that these correspond to different recursive constructions of sequences.

Keywords Ordinal collapsing functions · Friedman’s gap condition · Bachmann–Howard ordinal · Dilators · Well partial orders

Mathematics Subject Classification 03F15 · 05C05 · 06A06 · 18A35

1 Introduction

There is clearly a parallel between Higman’s lemma [1] and Kruskal’s theorem [2] on embeddings of sequences and trees, respectively. Not least, this parallel is manifest in the fact that both results have an elegant proof by Nash-Williams’s minimal bad sequence method [3]. If one wants to make the parallel more precise, it is natural to start with the observation that both sequences and trees are recursive data types. Such

✉ Anton Freund
freund@mathematik.tu-darmstadt.de

¹ Fachbereich Mathematik, Technische Universität Darmstadt, Schlossgartenstr. 7, 64289 Darmstadt, Germany

a data type can be constructed as the initial fixed point of a suitable transformation. For example, the initial fixed point of $Z \mapsto 1 + X \times Z$ is the set $\text{Seq}(X)$ of sequences with entries in X , while the initial fixed point of $X \mapsto \text{Seq}(X)$ is the set of ordered finite trees. In [4] we have studied certain general transformations of partial orders that we call normal PO-dilators (alluding to Girard's [5] dilators on linear orders). If W is a normal PO-dilator, then its initial fixed point $\mathcal{T}W$ carries a canonical partial order. In the aforementioned examples, this order coincides with the usual embedding relation from Higman's lemma and Kruskal's theorem. Thus both these results are instances of a general fact, which we call the uniform Kruskal theorem: If the normal PO-dilator W preserves well partial orders (WPOs), then the so-called Kruskal fixed point $\mathcal{T}W$ is a WPO itself. Together with Rathjen and Weiermann, the present author has shown that the uniform Kruskal theorem is equivalent to Π_1^1 -comprehension [4] (in the setting of reverse mathematics [6]). This is particularly interesting because it means that the uniform Kruskal theorem exhausts the full strength of the minimal bad sequence method (which has been analyzed by Marcone [7]), in contrast to Kruskal's original theorem. We note that constructions similar to $W \mapsto \mathcal{T}W$ had previously been studied by Hasegawa [8, 9] and Weiermann [10].

Harvey Friedman has introduced a gap condition on embeddings of trees, which leads to a much stronger version of Kruskal's theorem (see the presentation by Simpson [11]). Schütte and Simpson [12] have studied the corresponding condition for embeddings of sequences. As observed by Hasegawa [8, 9], Weiermann [10] and van der Meeren [13], the gap condition is related to iterations of constructions such as $W \mapsto \mathcal{T}W$. One way to make this precise has been worked out by the present author [14]: Given a normal PO-dilator W and a partial order X , one can construct a relativized Kruskal fixed point $\mathcal{T}W(X)$ that comes with a bijection

$$\iota_X + \kappa_X : X + W(\mathcal{T}W(X)) \rightarrow \mathcal{T}W(X). \quad (1.1)$$

Here addition denotes disjoint union, i. e., we have functions $\iota_X : X \rightarrow \mathcal{T}W(X)$ and $\kappa_X : W(\mathcal{T}W(X)) \rightarrow \mathcal{T}W(X)$ such that $\mathcal{T}W(X)$ is the disjoint union of their images. The order on $\mathcal{T}W(X)$ is determined by certain inequalities between the values of ι_X and κ_X (see [14]). The transformation $X \mapsto \mathcal{T}W(X)$ can again be equipped with the structure of a normal PO-dilator, which we call the Kruskal derivative of W . From now on, the notation $\mathcal{T}W$ will be reserved for this PO-dilator. The single fixed point from [4] should thus be denoted by $\mathcal{T}W(0)$, where 0 stands for the empty order. The principle of Π_1^1 -comprehension is still equivalent to the statement that $\mathcal{T}W$ preserves WPOs if W does. Now that $\mathcal{T}W$ is a transformation rather than a single order, we can iterate the construction: Let \mathbb{T}_0 be the identity on partial orders, considered as a normal PO-dilator. Given \mathbb{T}_n , define \mathbb{T}_{n+1}^- as the Kruskal derivative of $\text{Seq} \circ \mathbb{T}_n$. Then put $\mathbb{T}_{n+1} := \mathbb{T}_n \circ \mathbb{T}_{n+1}^-$. In [14] it is shown that $\mathbb{T}_n(0)$ is isomorphic to the set of trees with labels in $\{0, \dots, n-1\}$, ordered according to Friedman's strong gap condition. The analogous (but much simpler) result for a certain collection of sequences is discussed below.

The equivalence between Π_1^1 -comprehension and the uniform Kruskal theorem has been derived from a previous result on the level of linear orders. Each dilator D (i. e., each suitable transformation of well orders) gives rise to a linear order ϑD , which

we call the Bachmann–Howard fixed point of D (since it relativizes the Bachmann–Howard ordinal or, more precisely, the notation system from [15]). By an earlier result of the author [16–19], Π_1^1 -comprehension is equivalent to the statement that ϑD is a well order for any dilator D . In the present paper, we will relativize the construction of ϑD to a linear order X . This results in a Bachmann–Howard fixed point $\vartheta D(X)$ over X , which comes with a bijection

$$\iota_X + \vartheta_X : X + D(\vartheta D(X)) \rightarrow \vartheta D(X),$$

analogous to (1.1). The appropriate inequalities between different values of ι_X and ϑ_X will, once again, be a crucial part of the definition. As above, the notation ϑD is now reserved for the transformation $X \mapsto \vartheta D(X)$, while the single fixed point from [19] should be denoted by $\vartheta D(0)$. We will see that the transformation ϑD can again be equipped with the structure of a dilator. This dilator ϑD will be called the Bachmann–Howard derivative of D .

Let us discuss the relation between Kruskal and Bachmann–Howard derivatives. By a linearization of a partial order P by a linear order X we shall mean an order-reflecting surjection $f : X \rightarrow P$. Here, order-reflecting means that $f(x) \leq_P f(y)$ implies $x \leq_X y$, which also ensures that f is injective. Note that this coincides with the usual notion of linearization if we identify X with its image under f . A linearization of a PO-dilator W by a dilator D is a natural family of linearizations $D(X) \rightarrow W(X)$, one for each linear order X . We will show that any linearization of W by D can be transformed into a linearization of the Kruskal derivative $\mathcal{T}W$ by the Bachmann–Howard derivative ϑD .

We now consider an application of our general constructions. In [20], van der Meeren, Rathjen and Weiermann study a certain collection of sequences with gap condition, as well as its linearization by iterated collapsing functions. We will give the following systematic reconstruction of these objects:

- (1) Let S_0^0 and T_0^0 be the identity on partial and linear orders, respectively (considered as a normal PO-dilator and a dilator in the sense of Girard).
- (2) Define S_{n+1}^0 as the Kruskal derivative of S_n^0 , and T_{n+1}^0 as the Bachmann–Howard derivative of T_n^0 .
- (3) Write 1 for the order with a single element. Then $S_n^0(1)$ coincides with the set of sequences $\overline{S}_n[0]$ from [20, Definition 12], ordered by Friedman’s strong gap condition. Furthermore, $T_n^0(1)$ coincides with the system $T_n[0]$ of collapsing functions from [20, Definitions 24 to 27].

Now the fact that $T_n[0] \cong T_n^0(1)$ is a linearization of $\overline{S}_n[0] \cong S_n^0(1)$ is immediate by the general result from the previous paragraph, which thus replaces the explicit verification in [20, Lemmas 10 and 11]. More importantly, our reconstruction clarifies two conceptual points. First, it confirms that gap condition and collapsing functions are closely related, maybe even more closely than expected: they arise by entirely parallel constructions on partial and linear orders, respectively. Secondly, the collapsing functions studied in [20] (and the variant with addition in [21]) are supposed to generalize Rathjen’s notation system for the Bachmann–Howard ordinal (see [4]). But do they

provide “the right” generalization? Our reconstruction shows that, in a certain precise sense, the answer is positive.

Let us emphasize that previous work of Hasegawa, Weiermann and van der Meeren is fundamental for the present paper. The concrete results that we derive on linearizations of the gap condition for sequences—in particular Corollaries 5.11, 5.20 and 5.21—were first shown by Rathjen, Weiermann and van der Meeren [20]. Beyond these concrete results, our approach is heavily informed by the many insights from the PhD thesis of van der Meeren [13] and his papers with Rathjen and Weiermann [22, 23] on the gap condition for trees. The achievement of the present paper is that it develops these insights in terms of dilators, which provide a framework that is both rigorous and very general.

In order to indicate that our constructions cover a larger range of applications, we sketch two possible modifications. In the first of these we put $T_0 = T_0^0$ and $T_{n+1} = T_n \circ T_{n+1}^-$, where T_{n+1}^- is defined as the Bachmann–Howard derivative of T_n (note the similarity with the clause $\mathbb{T}_{n+1} = \mathbb{T}_n \circ \mathbb{T}_{n+1}^-$ from the discussion of [14] above). It seems that $T_n(1)$ is a linearization of the order \mathbb{S}_n from [20], which is more liberal than the order $\overline{\mathbb{S}}_n[0]$. For the second modification, consider a dilator S that linearizes the PO-dilator Seq from above (e. g. take $S = \omega_2$ as below). Let D_0 be the identity on linear orders, define D_{n+1}^- as the Bachmann–Howard derivative of $S \circ D_n$, and set $D_{n+1} := D_n \circ D_{n+1}^-$. Since the construction is entirely parallel to the one from [14] (discussed above), this should yield a linearization of Friedman’s gap condition on trees. We expect that the linear orders $D_n(0)$ are closely related to the iterated collapsing functions with addition that are studied in [21]. Details of both modifications remain to be checked. We have sketched them to indicate the potential breadth of our approach.

As observed by van der Meeren, Rathjen and Weiermann [20], the linearization of $S_n^0(1) \cong \overline{\mathbb{S}}_n[0]$ by $T_n^0(1) \cong T_n[0]$ does not have maximal order type. It is expected that this phenomenon is specific to the case of sequences, i. e., that collapsing functions do exhaust the maximal order type for trees. Our approach provides some justification for this expectation, or at least a systematic explanation. To present the latter, we consider sequences on a more concrete level: The elements of $S_n^0(X)$ can be represented in the form $\langle i_1, \dots, i_k, x \rangle$ with $i_1 = 0$ and $x \in X$ (and with some further conditions, see Definition 5.1 below). In view of (1.1), the fact that S_{n+1}^0 is the Kruskal derivative of S_n^0 is witnessed, amongst others, by injections

$$\kappa_X^n : S_n^0(S_{n+1}^0(X)) \rightarrow S_{n+1}^0(X).$$

We will see that these functions are given by

$$\kappa_X^n(\langle i_1, \dots, i_k, \langle j_1, \dots, j_l, x \rangle \rangle) = \langle 0, i_1 + 1, \dots, i_k + 1, j_1, \dots, j_l, x \rangle.$$

Note that j_1 is the second entry on the right that is equal to zero, which guarantees that κ_X^n is injective. To generalize the construction from sequences to trees, think of the last entry of an element $\langle i_1, \dots, i_k, x \rangle \in S_n^0(X)$ as a leaf labelled by x . Indeed, the orders $\mathbb{T}_n(X)$ from above (studied in [14]) consist of trees with labels from $\{0, \dots, n -$

$1\} \cup X$, where labels from X are allowed at leaves only. We can now point out a crucial difference between sequences and trees: A sequence different from $\langle \rangle$ (the empty sequence) has a single leaf with known location (the last entry). This means that it suffices to record the leaf label and the rest of the sequence or, more formally, that we have a bijection

$$\{\langle \rangle\} \cup \left(X \times S_n^0(\{\star\}) \right) \cong S_n^0(X),$$

$$\langle x, \langle i_1, \dots, i_k, \star \rangle \rangle \mapsto \langle i_1, \dots, i_k, x \rangle.$$

In the case of trees, there are many possible locations for leaf labels, and a similar bijection does not appear to be available. It turns out that we get an alternative construction of the gap condition on sequences (but not on trees). To describe this construction, we recall that the set $\text{Seq}(Z)$ of finite sequences in Z is the initial Kruskal fixed point of the transformation $X \mapsto 1 + Z \times X$, i.e., that we have

$$\text{Seq}(Z) = \mathcal{T}W(0) \quad \text{with} \quad W(X) = 1 + Z \times X.$$

Here $Z \times X$ is the usual product of partial orders, where $(z, x) \leq_{Z \times X} (z', x')$ is equivalent to the conjunction of $z \leq_Z z'$ and $x \leq_X x'$. The order $1 + Z \times X$ contains a further element but no other strict inequalities. We will see that there is an isomorphism

$$S_{n+1}^0(1) \cong \text{Seq}(S_n^0(1)) \tag{1.2}$$

of partial orders, for each $n \in \mathbb{N}$. Together with (1) to (3) from above, this yields a second construction of sequences with gap condition in terms of Kruskal derivatives (or fixed points). While the two constructions coincide for partial orders, it turns out that they differ in the linear case. Given a linear order Z , we define $\omega_2(Z)$ as the initial Bachmann–Howard fixed point of the transformation $X \mapsto 1 + Z \times X$, i.e., we set

$$\omega_2(Z) = \vartheta D(0) \quad \text{with} \quad D(X) = 1 + Z \times X. \tag{1.3}$$

Note that, in the linear case, the single element of 1 lies below all elements of $Z \times X$, while $(z, x) \leq_{Z \times X} (z', x')$ holds if we have $z <_Z z'$ or $(z = z'$ and $x \leq_X x')$. More concretely, one can represent $\omega_2(Z)$ by the set of finite sequences in Z , as one checks with the help of Theorem 2.9 below. With $z_0 \frown \langle z_1, \dots, z_k \rangle := \langle z_0, \dots, z_k \rangle$, the order on $\omega_2(Z)$ is characterized by the following clauses (as before Theorem 2.2 in [24]):

- (i) we always have $\langle \rangle < z \frown \sigma$ and $\sigma < z \frown \sigma$,
- (ii) given $(z, \sigma) < (z', \sigma')$ in $Z \times \omega_2(Z)$ and $\sigma < z' \frown \sigma'$, we get $z \frown \sigma < z' \frown \sigma'$.

To turn ω_2 into a dilator (see Definition 2.1 below), it suffices to set

$$\omega_2(f)(\langle z_0, \dots, z_{k-1} \rangle) := \langle f(z_0), \dots, f(z_{k-1}) \rangle,$$

$$\text{supp}_Z(\langle z_0, \dots, z_{k-1} \rangle) := \{z_0, \dots, z_{k-1}\}.$$

For an ordinal α and a number $n \in \mathbb{N}$, the ordinal ω_n^α is explained by the recursive clauses $\omega_0^\alpha = \alpha$ and $\omega_{n+1}^\alpha = \omega^{\omega_n^\alpha}$ (towers of exponentials in the sense of ordinal arithmetic). The resulting ordinal ω_2^α is isomorphic to the order $\omega_2(1 + \alpha)$ from (1.3), as shown in [24]. We now define linear orders OT_n^0 by the recursive clauses

$$OT_0^0 = 1 \quad \text{and} \quad OT_{n+1}^0 = \omega_2(OT_n^0). \tag{1.4}$$

As we will see, OT_n^0 coincides with the order $OT_n[0]$ from [20, Section 5] (up to a typo in the cited reference, see Sect. 5 below). Important results of the cited paper can now be deduced from general facts about Bachmann–Howard fixed points: First, the parallel between (1.2) and (1.4) ensures that $OT_n[0]$ linearizes $S_n^0(1) \cong \overline{S}_n[0]$. Secondly, iterated applications of the result from [24] show that $OT_n[0]$ has order type $\omega_{2n-1} := \omega_{2n-1}^1 = \omega_{2n}^0$, for any $n > 0$. In [20] this fact was established by explicit computations that involve the addition-free Veblen functions, which may seem somewhat ad hoc. Let us recall that ω_{2n-1} is the maximal order type of the partial order $S_n^0(1) \cong \overline{S}_n[0]$, as shown in [20] (based on results from [12]).

To summarize, the present paper introduces the general notion of Bachmann–Howard derivative. The latter allows us to give two systematic reconstructions of the gap condition on finite sequences. While the two constructions yield the same result in the case of partial orders, the versions for linear orders lead to two different systems of collapsing functions, both of which have been studied in [20]. The first construction does not realize the maximal order type (in the case of sequences) but seems to be of greater general interest, since it is readily extended from sequences to trees (cf. [14]). The second construction exploits a property that is specific to sequences, and it realizes the maximal order type.

2 Theory, part 1: Bachmann–Howard derivatives

In this section we define the notion of Bachmann–Howard derivative, by making the informal explanation from the introduction precise. We then give a proof of existence and uniqueness, including a criterion that is useful for applications.

To recall the definition of dilators, we need some terminology: Write $[X]^{<\omega}$ for the set of finite subsets of a set X . Each $f : X \rightarrow Y$ induces a function

$$[f]^{<\omega} : [X]^{<\omega} \rightarrow [Y]^{<\omega} \quad \text{with} \quad [f]^{<\omega}(a) = \{f(x) \mid x \in a\}.$$

This yields an endofunctor $[\cdot]^{<\omega}$ on the category of sets. Given $a \in [X]^{<\omega}$, we will write $\iota_a : a \hookrightarrow X$ for the inclusion map, provided that X is clear from the context. Let LO be the category of linear orders and order embeddings. We will omit the forgetful functor from orders to sets (and thus apply $[\cdot]^{<\omega}$ to orders). Conversely, a subset of an order will often be considered as a suborder. Finally, let us agree that $\text{rng}(f)$ denotes the range (in the sense of image) of f .

Definition 2.1 An LO-dilator consists of a functor $D : \text{LO} \rightarrow \text{LO}$ and a natural transformation $\text{supp} : D \Rightarrow [\cdot]^{<\omega}$ such that the so-called support condition

$$\text{supp}_Y(\sigma) \subseteq \text{rng}(f) \Rightarrow \sigma \in \text{rng}(D(f))$$

holds for any LO-morphism $f : X \rightarrow Y$ and any $\sigma \in D(Y)$. If, in addition, $D(X)$ is well founded for any well order X , then (D, supp^D) is called a WO-dilator.

Note that the converse of the implication in the definition is automatic since supp is natural. There is at most one natural transformation $\text{supp} : D \Rightarrow [\cdot]^{<\omega}$ that satisfies the support condition, since $\text{supp}_X(\sigma)$ is determined as the minimal $a \subseteq X$ with $\sigma \in \text{rng}(D(\iota_a))$. Furthermore, such a natural transformation exists if, and only if, D preserves pullbacks and direct limits, as verified in [16, Remark 2.2.2]. This means that our definition of WO-dilators coincides with Girard’s definition of dilators [5]. We have added the prefix WO for clarity, since we will later consider variants of dilators on partial orders.

As supports are uniquely determined, we often write D instead of (D, supp) . Sometimes (but not always) we then write supp^D to refer to supp . Let us write

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

if the equality holds and we have $\text{supp}_a(\sigma_0) = a$. This notation allows us to formulate a version of Girard’s normal form theorem:

Lemma 2.2 Consider an LO-dilator D and a linear order X . Each $\sigma \in D(X)$ has a unique normal form $\sigma = {}_{\text{NF}}D(\iota_a)(\sigma_0)$. The latter satisfies $a = \text{supp}_X(\sigma)$.

Proof For $\sigma = D(\iota_a)(\sigma_0)$ we have $\text{supp}_X(\sigma) = \text{supp}_a(\sigma_0)$ by naturality. Now the existence of a normal form with $a = \text{supp}_X(\sigma)$ follows from the support condition. Uniqueness holds as σ determines a and as the embedding $D(\iota_a)$ is injective. \square

As observed by Girard, the normal form theorem entails that dilators are determined (up to natural isomorphism) by their restrictions to the category of finite linear orders. Since this category is essentially small, these restrictions can be represented by sets (rather than proper classes). A formalization in second order arithmetic is available for the countable case (see e. g. [19] for details). In the present paper we do not work in a specific base theory.

To define the Bachmann–Howard derivative ϑD of an LO-dilator D we must, in particular, specify a transformation $X \mapsto \vartheta D(X)$ of linear orders. The following definition constructs the required orders from syntactic material. A more abstract characterization, which can be easier to handle in applications, will be given later. We point out that the orders $\vartheta D(X)$ are relativized versions of the Bachmann–Howard fixed points that we have constructed in [19].

Definition 2.3 Let us consider an LO-dilator D . For each linear order X , we define a set $\vartheta D(X)$ of terms and a binary relation $<_{\vartheta D(X)}$ on that set by simultaneous recursion. The set $\vartheta D(X)$ is generated by the following clauses:

- (i) For each element $x \in X$ we have a term $\bar{x} \in \vartheta D(X)$.
- (ii) Given a finite set $a \subseteq \vartheta D(X)$ that is linearly ordered by $<_{\vartheta D(X)}$, we add a term $\vartheta \langle a, \sigma \rangle$ for each element $\sigma \in D(a)$ with $\text{supp}_a^D(\sigma) = a$.

For $s, t \in \vartheta D(X)$ we stipulate that $s <_{\vartheta D(X)} t$ holds if, and only if, one of the following clauses applies:

- (i') We have $s = \bar{x}$ and $t = \bar{y}$ with $x <_X y$.
- (ii') The term s is of the form \bar{x} while t is of the form $\vartheta \langle b, \tau \rangle$.
- (iii') We have $s = \vartheta \langle a, \sigma \rangle$ and $t = \vartheta \langle b, \tau \rangle$, the restriction of $<_{\vartheta D(X)}$ to $a \cup b$ is a linear order, and one of the following holds:
 - We have $D(\iota_a)(\sigma) <_{D(a \cup b)} D(\iota_b)(\tau)$ for the inclusions $\iota_a : a \hookrightarrow a \cup b$ and $\iota_b : b \hookrightarrow a \cup b$. Furthermore, we have $s' <_{\vartheta D(X)} t$ for all $s' \in a$.
 - We have $s \leq_{\vartheta D(X)} t'$ for some $t' \in b$ (i. e., we have $s <_{\vartheta D(X)} t'$ or s and t' are the same term).

To justify the recursion in detail, one can argue as follows: In a first step, ignore the reference to $<_{\vartheta D(X)}$ in order to generate a larger set $\vartheta_0 D(X) \supseteq \vartheta D(X)$. More precisely, declare that $\vartheta_0 D(X)$ contains $\vartheta \langle a, \sigma \rangle$ whenever $a \subseteq \vartheta_0 D(X)$ is finite and $\sigma \in D(a)$ holds with respect to *some* linear order on a . Then define a length function $l : \vartheta_0 D(X) \rightarrow \mathbb{N}$ by recursion over terms, stipulating

$$l(\bar{x}) = 0 \quad \text{and} \quad l(\vartheta \langle a, \sigma \rangle) = 1 + \sum_{r \in a} 2 \cdot l(r).$$

Finally, decide $r \in \vartheta D(X)$ and $s <_{\vartheta D(X)} t$ by simultaneous recursion on $l(r)$ and $l(s) + l(t)$, respectively. For example, we can decide $r \in \vartheta D(X)$ for $r = \vartheta \langle a, \sigma \rangle$ as follows: Recursively decide $a \subseteq \vartheta D(X)$ and compute the restriction of $<_{\vartheta D(X)}$ to a . If the latter is a linear order, then check $\sigma \in D(a)$ and $\text{supp}_a^D(\sigma) = a$ with respect to it. At various places in Definition 2.3, we have required that certain restrictions of $<_{\vartheta D(X)}$ are linear. The purpose was to ensure that $D : \text{LO} \rightarrow \text{LO}$ is only applied to (morphisms of) linear orders. *Ex post*, linearity is automatic by the following proposition, which is proved as in the non-relativized case (see [19, Proposition 4.1]).

Proposition 2.4 *The relation $<_{\vartheta D(X)}$ is a linear order on $\vartheta D(X)$.*

The syntactic construction of $\vartheta D(X)$ can be hard to apply, even in simple cases. For this reason, we now develop a more abstract characterization.

Definition 2.5 Consider an LO-dilator D and a linear order X . A Bachmann–Howard fixed point of D over X consists of a linear order Z and functions $\iota : X \rightarrow Z$ and $\vartheta : D(Z) \rightarrow Z$ with the following properties:

- (i) The function $\iota : X \rightarrow Z$ is an order embedding.
- (ii) We have $\iota(x) <_Z \vartheta(\sigma)$ for all $x \in X$ and $\sigma \in D(Z)$.
- (iii) If we have $\sigma <_{D(Z)} \tau$ as well as $z <_Z \vartheta(\tau)$ for all $z \in \text{supp}_Z^D(\sigma)$, then we have $\vartheta(\sigma) <_Z \vartheta(\tau)$. Furthermore, $z <_Z \vartheta(\sigma)$ holds for any $z \in \text{supp}_Z^D(\sigma)$.

Let us extend the syntactic orders $\vartheta D(X)$ into Bachmann–Howard fixed points:

Definition 2.6 To define $\iota_X : X \rightarrow \vartheta D(X)$ and $\vartheta_X : D(\vartheta D(X)) \rightarrow \vartheta D(X)$, we stipulate $\iota_X(x) = \bar{x}$ and $\vartheta_X(\sigma) = \vartheta \langle a, \sigma_0 \rangle$ for $\sigma = \text{NF}D(\iota_a)(\sigma_0)$.

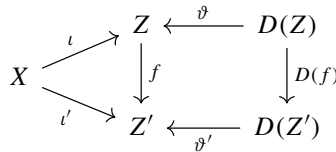
As expected, we obtain the following:

Proposition 2.7 *The tuple $(\vartheta D(X), \iota_X, \vartheta_X)$ explained by Definitions 2.3 and 2.6 is a Bachmann–Howard fixed point of D over X .*

Proof It is immediate that conditions (i) and (ii) from Definition 2.5 are satisfied. Concerning condition (iii), we point out that $s <_{\vartheta D(X)} \vartheta_X(\sigma)$ is immediate for a term of the form $s = \bar{x}$, by clause (ii') of Definition 2.3. The remaining conditions are verified as in the non-relativized case, for which we refer to [19, Theorem 4.1]. \square

In the non-relativized case, [19, Theorem 4.2] shows that $\vartheta D(0)$ can be embedded into any Bachmann–Howard fixed point of D over the empty set. Here we establish a stronger categorical property, which will be useful below (see [14, Section 3] for parallel constructions in the context of partial orders).

Definition 2.8 A Bachmann–Howard fixed point (Z, ι, ϑ) of D over X is called initial if any Bachmann–Howard fixed point (Z', ι', ϑ') of D over X admits a unique order embedding $f : Z \rightarrow Z'$ such that



is a commutative diagram.

By the usual categorical argument, initial Bachmann–Howard fixed points are unique up to isomorphism. The following yields existence and a useful criterion.

Theorem 2.9 *Consider an LO-dilator D and a linear order X . For any Bachmann–Howard fixed point (Z, ι, ϑ) of D over X , the following are equivalent:*

- (i) *We have $Z = \text{rng}(\iota) \cup \text{rng}(\vartheta)$, and there is a function $h : Z \rightarrow \mathbb{N}$ such that $z \in \text{supp}_Z^D(\sigma)$ entails $h(z) < h(\vartheta(\sigma))$, for any $\sigma \in D(Z)$.*
- (ii) *The Bachmann–Howard fixed point (Z, ι, ϑ) is initial.*

Furthermore, the Bachmann–Howard fixed point from Proposition 2.7 is initial.

Proof We first assume (i) and derive (ii). Aiming at the latter, we consider an arbitrary Bachmann–Howard fixed point (Z', ι', ϑ') of D over X . The diagram from Definition 2.8 commutes if, and only if, we have

$$\begin{array}{ll}
 f(\iota(x)) = \iota'(x) & \text{for all } x \in X, \\
 f(\vartheta(\sigma)) = \vartheta'(D(f \circ \iota_a)(\sigma_0)) & \text{for any } \sigma = \text{NF}D(\iota_a)(\sigma_0) \in D(Z).
 \end{array}$$

The idea is to read these equations as a recursive definition of f , which is possible in view of the following observations: First, ι and ϑ are injective by clauses (i) and (iii) of Definition 2.5 (see also [16, Lemma 2.1.7]). Secondly, the union $Z = \text{rng}(\iota) \cup \text{rng}(\vartheta)$ is necessarily disjoint, by clause (ii) of the same definition. Finally, the composition $f \circ \iota_a$ does only depend on values of f on arguments $z \in a = \text{supp}_Z^D(\sigma)$, for which a recursive call is justified due to $h(z) < h(\vartheta(\sigma))$. In view of these observations, there can be at most one embedding $f : Z \rightarrow Z'$ with the required properties. To show that there is one, we first define a length function $l : Z \rightarrow \mathbb{N}$ by the recursive clauses $l(\iota(x)) = 0$ and $l(\vartheta(\sigma)) = 1 + \sum_{z \in \text{supp}_Z^D(\sigma)} 2 \cdot l(z)$, which are justified as above. By simultaneous induction on $l(z)$ and $l(z_0) + l(z_1)$ one can now check that $f(z) \in Z'$ is defined and that $z_0 <_Z z_1$ entails $f(z_0) <_{Z'} f(z_1)$. The motivation for the simultaneous verification is that $f \circ \iota_a$ needs to be an embedding for $D(f \circ \iota_a)$ to be defined. The only interesting case in the induction concerns an inequality

$$z_0 = \vartheta(\sigma) <_Z \vartheta(\tau) = z_1.$$

Crucially, we must have $z <_Z \vartheta(\tau)$ for all elements $z \in \text{supp}_X^D(\sigma)$, since $\vartheta(\tau) \leq_Z z$ would entail $\vartheta(\tau) <_Z \vartheta(\sigma)$, by clause (iii) of Definition 2.5. If $\sigma <_{D(Z)} \tau$ fails, we get $\vartheta(\sigma) \leq_Z z'$ for some $z' \in \text{supp}_Z^D(\tau)$, for the same reason. Writing $\sigma = \text{NF}D(\iota_a)(\sigma_0)$ and $\tau = \text{NF}D(\iota_b)(\tau_0)$, we can use the induction hypothesis to show that one of the following must hold (write $f \circ \iota_a = f \upharpoonright (a \cup b) \circ \iota'_a$ with $\iota'_a : a \hookrightarrow a \cup b$):

- We have $D(f \circ \iota_a)(\sigma_0) <_{D(Z')} D(f \circ \iota_b)(\tau_0)$ and $z <_{Z'} f(\vartheta(\tau)) = \vartheta'(D(f \circ \iota_b)(\tau_0))$ for all $z \in [f]^{<\omega}(\text{supp}_Z^D(\sigma)) = \text{supp}_{Z'}^D(D(f \circ \iota_a)(\sigma_0))$.
- We have $\vartheta'(D(f \circ \iota_a)(\sigma_0)) \leq_{Z'} z'$ for some $z' \in \text{supp}_{Z'}^D(D(f \circ \iota_b)(\tau_0))$.

In either case, clause (iii) of Definition 2.5 yields

$$f(z_0) = \vartheta'(D(f \circ \iota_a)(\sigma_0)) <_{Z'} \vartheta'(D(f \circ \iota_b)(\tau_0)) = f(z_1).$$

Next, we use the criterion provided by (i) to show that the Bachmann–Howard fixed point $(\vartheta D(X), \iota_X, \vartheta_X)$ from Proposition 2.7 is initial. To see that we have

$$\vartheta D(X) = \text{rng}(\iota_X) \cup \text{rng}(\vartheta_X),$$

it suffices to observe that the same condition $\text{supp}_a^D(\sigma) = a$ appears in clause (ii) of Definition 2.3 and in the definition of normal forms (as given in the paragraph before Lemma 2.2). Let us now define $h : \vartheta D(X) \rightarrow \mathbb{N}$ by recursion over terms, setting $h(\bar{x}) = 0$ and $h(\vartheta(a, \sigma_0)) = 1 + \max\{h(s) \mid s \in a\}$. To see that this function has the required property, it suffices to recall that $\sigma = \text{NF}D(\iota_a)(\sigma_0) \in D(\vartheta D(X))$ entails $\text{supp}_{\vartheta D(X)}^D(\sigma) = a$, by Lemma 2.2. Finally, it is now easy to conclude that (ii) implies (i), because any initial Bachmann–Howard fixed point must be isomorphic to $(\vartheta D(X), \iota_X, \vartheta_X)$ (see the proof of [14, Theorem 3.5] for details). □

Now that we have established existence and uniqueness up to isomorphism, we will sometimes speak of ‘the’ initial Bachmann–Howard fixed point of D over X and denote ‘it’ by $\vartheta D(X)$ (i. e., this notation is no longer reserved for the specific term

systems from Definition 2.3). In [14] we have introduced a notion of Kruskal derivative for dilators of partial orders. We now define the corresponding notion in the context of linear orders.

Definition 2.10 A Bachmann–Howard derivative of an LO-dilator D consists of an LO-dilator ϑD and natural families of functions

$$\iota_X : X \rightarrow \vartheta D(X) \quad \text{and} \quad \vartheta_X : D \circ \vartheta D(X) \rightarrow \vartheta D(X)$$

such that $(\vartheta D(X), \iota_X, \vartheta_X)$ is an initial Bachmann–Howard fixed point of D over X , for each linear order X .

To avoid misunderstanding, we point out that the functions ϑ_X need not be order embeddings. Hence we do not have a natural transformation $\vartheta : D \circ \vartheta D \Rightarrow \vartheta D$ between endofunctors of linear orders (but the naturality condition is the same).

Proposition 2.11 *Assume that, for each linear order X , we are given an initial Bachmann–Howard fixed point $(\vartheta D(X), \iota_X, \vartheta_X)$ of D over X . There is a unique way to extend this data into a Bachmann–Howard derivative of D .*

Proof We first show that the given map $X \mapsto \vartheta D(X)$ can be uniquely extended into a functor. Given an embedding $f : X \rightarrow Y$ of linear orders, it is straightforward to check that $(\vartheta D(Y), \iota_Y \circ f, \vartheta_Y)$ is a Bachmann–Howard fixed point of D over X . Since $(\vartheta D(X), \iota_X, \vartheta_X)$ is initial, there is a unique embedding $\vartheta D(f) : \vartheta D(X) \rightarrow \vartheta D(Y)$ such that the following is a commutative diagram:

$$\begin{array}{ccccc}
 & & \vartheta D(X) & \xleftarrow{\vartheta_X} & D(\vartheta D(X)) \\
 & \nearrow \iota_X & \downarrow \vartheta D(f) & & \downarrow D(\vartheta D(f)) \\
 X & & \vartheta D(Y) & \xleftarrow{\vartheta_Y} & D(\vartheta D(Y)) \\
 & \searrow \iota_Y \circ f & & &
 \end{array}$$

The very same diagram (with the triangle written as a square) must commute if the functions ι_X and ϑ_X are to be natural in X . Hence there is a unique extension into a functor $\vartheta D : \text{LO} \rightarrow \text{LO}$ with the required properties. It remains to consider the extension into an LO-dilator. First observe that the required functions

$$\text{supp}_X^{\vartheta D} : \vartheta D(X) \rightarrow [X]^{<\omega}$$

are necessarily unique, as naturality and the implication from Definition 2.1 require

$$\text{supp}_X^{\vartheta D}(\sigma) = \bigcap \{a \in [X]^{<\omega} \mid \sigma \in \text{rng}(\vartheta D(\iota_a))\},$$

where $\iota_a : a \hookrightarrow X$ are the inclusions. For existence we use the characterization from part (i) of Theorem 2.9. It allows us to define supports by the recursive clauses

$$\text{supp}_X^{\vartheta D}(\iota_X(x)) = \{x\},$$

$$\text{supp}_X^{\vartheta^D}(\vartheta_X(\sigma)) = \bigcup \{ \text{supp}_X^{\vartheta^D}(s) \mid s \in \text{supp}_{\vartheta^D(X)}^D(\sigma) \}.$$

Naturality and the implication from Definition 2.1 are checked by induction. For details we refer to the analogous argument for partial orders (see [14, Theorem 4.2]). \square

Together with the last sentence of Theorem 2.9, we get existence:

Corollary 2.12 *Any LO-dilator has a Bachmann–Howard derivative.*

By an isomorphism between LO-dilators (D, supp^D) and (E, supp^E) we simply mean a natural isomorphism $\eta : D \Rightarrow E$ of functors. This is justified because the supports are automatically respected, i. e. we have $\text{supp}^E \circ \eta = \text{supp}^D$. For an isomorphism this is particularly easy to see (cf. the paragraph before [14, Theorem 4.4]). It is also true but more subtle when η is merely a natural transformation (see [5]). We can now formulate the appropriate uniqueness result:

Proposition 2.13 *If $(\vartheta^0 D, \iota^0, \vartheta^0)$ and $(\vartheta^1 D, \iota^1, \vartheta^1)$ are two Bachmann–Howard derivatives of the same LO-dilator D , then there is a unique natural isomorphism $\eta : \vartheta^0 D \Rightarrow \vartheta^1 D$ such that the diagram*

$$\begin{array}{ccccc} X & \begin{array}{l} \xrightarrow{\iota_X^0} \\ \searrow \\ \xrightarrow{\iota_X^1} \end{array} & \begin{array}{c} \vartheta^0 D(X) \\ \downarrow \eta_X \\ \vartheta^1 D(X) \end{array} & \begin{array}{c} \xleftarrow{\vartheta_X^0} D(\vartheta^0 D(X)) \\ \downarrow D(\eta_X) \\ \xleftarrow{\vartheta_X^1} D(\vartheta^1 D(X)) \end{array} \end{array}$$

commutes for every linear order X .

Proof Existence and uniqueness of isomorphisms $\eta_X : \vartheta^0 D(X) \rightarrow \vartheta^1 D(X)$ as in the diagram are due to the assumption that $(\vartheta^0 D(X), \iota_X^0, \vartheta_X^0)$ and $(\vartheta^1 D(X), \iota_X^1, \vartheta_X^1)$ are initial Bachmann–Howard fixed points of D over X . The non-trivial claim of the proposition is that these isomorphisms are natural in X . This is shown as in the corresponding result for partial orders, for which we refer to [14, Theorem 4.4]. \square

In the next section we will want to take iterated Bachmann–Howard derivatives. To see that the result is still unique, one should check that the derivatives of isomorphic LO-dilators are isomorphic. This follows from the previous result and the following observation, which is shown as in the partial case (see [14, Proposition 4.6]).

Lemma 2.14 *If $(\vartheta E, \iota, \vartheta)$ is a Bachmann–Howard derivative of E and $\eta : D \Rightarrow E$ is a natural isomorphism, then $(\vartheta E, \iota, \vartheta \bullet \eta)$ is a Bachmann–Howard derivative of D , where $\vartheta \bullet \eta : D \circ \vartheta E \Rightarrow \vartheta E$ is given by $(\vartheta \bullet \eta)_X = \vartheta_X \circ \eta_{\vartheta E(X)}$.*

Having established existence and uniqueness, we will speak of ‘the’ Bachmann–Howard derivative of D and denote ‘it’ by ϑD . To complete the basic theory of Bachmann–Howard derivatives, we should show that ϑD is a WO-dilator (i. e., preserves well foundedness) when the same holds for D . Since a particularly short proof of this fact exploits the connection with partial orders, we defer this result until Corollary 4.3 below.

3 Application, part 1: unary collapsing functions

In the introduction we have described a recursive construction of LO-dilators T_n^0 . Starting with the identity functor $T_0^0 : LO \rightarrow LO$ (which is an LO-dilator with support functions $\text{supp}_X : T_0^0(X) = X \rightarrow [X]^{<\omega}$ given by $\text{supp}_X(x) = \{x\}$), we have defined T_{n+1}^0 as the Bachmann–Howard derivative of T_n^0 . This construction is explained and justified by the results of Sect. 2. Also in the introduction, we have claimed that $T_n^0(1)$ coincides with the order $T_n[0]$ from [20] (where 1 is the order with a single element). This claim will be proved in the present section.

Due to the uniqueness results from Sect. 2 we can argue ‘the other way around’. This means that we will not, at first, consider T_n^0 as given by a recursive construction. Instead, we will give an *ad hoc* definition of T_n^0 , which extends the definition of the term systems $T_n[0] \cong T_n(1)$ from [20]. In a second step, we will define transformations $\iota^n : \text{Id} \Rightarrow T_{n+1}^0$ and $\vartheta^n : T_n^0 \circ T_{n+1}^0 \Rightarrow T_{n+1}^0$ that turn T_{n+1}^0 into a Bachmann–Howard derivative of T_n^0 . By Proposition 2.13 and Lemma 2.14, this will entail that our *ad hoc* definition and the recursive construction yield the same result after all (up to natural isomorphism).

Definition 3.1 Given a linear order X , we generate a set $T(X)$ of terms and a function $S : T(X) \rightarrow \mathbb{N} \cup \{-1\}$ by simultaneous recursion:

- (i) For each $x \in X$, include a term $\bar{x} \in T(X)$ with $S(\bar{x}) = -1$.
- (ii) Given $s \in T(X)$ and $i \geq \max\{S(s) - 1, 0\}$, add $\vartheta_i s \in T(X)$ with $S(\vartheta_i s) = i$.

For each $i \in \mathbb{N} \cup \{-1\}$, let $k_i : T(X) \rightarrow T(X)$ be given by the recursive clauses

$$k_i(\bar{x}) = \bar{x} \quad \text{and} \quad k_i(\vartheta_j s) = \begin{cases} \vartheta_j s & \text{if } j \leq i, \\ k_i(s) & \text{if } j > i. \end{cases}$$

To define a binary relation $<_{T(X)}$ on $T(X)$, we declare that $s <_{T(X)} t$ holds if, and only if, one of the following clauses applies:

- (i') We have $s = \bar{x}$ and $t = \bar{y}$ with $x <_X y$.
- (ii') The term s is of the form \bar{x} while t is of the form $\vartheta_j t'$.
- (iii') We have $s = \vartheta_i s'$ and $t = \vartheta_j t'$, and one of the following holds:
 - We have $i < j$.
 - We have $i = j$, $s' <_{T(X)} t'$ and $k_i(s') <_{T(X)} t$.
 - We have $i = j$ and $s \leq_{T(X)} k_j(t')$.

Concerning the last clause, we clarify that $s \leq_{T(X)} t$ abbreviates the disjunction of $s <_{T(X)} t$ and the statement that s and t are the same term.

To justify the definition of $<_{T(X)}$ one can employ the function $h : T(X) \rightarrow \mathbb{N}$ given by $h(\bar{x}) = 0$ and $h(\vartheta_i s) = h(s) + 1$. An easy induction shows $h(k_i(s)) \leq h(s)$. It follows that $s <_{T(X)} t$ can be decided by recursion on $h(s) + h(t)$.

Lemma 3.2 *The relation $<_{T(X)}$ on $T(X)$ is a linear order.*

Proof For a term $s = \vartheta_i s' \in T(X)$, the s -secure subterms of s are defined as follows: The term s' is s -secure. And if $\vartheta_j t$ with $j \geq i$ is s -secure, then so is t . More intuitively, t is s -secure if we have $s' = \vartheta_{j_1} \dots \vartheta_{j_n} t$ with $j_1, \dots, j_n \geq i$. If t is s -secure for $s = \vartheta_i s'$, then the following holds: First, $k_i(t)$ is s -secure. Secondly, we either have $k_i(t) = k_i(s')$, or t is $k_i(s')$ -secure with $k_i(s') = \vartheta_i s''$ for the same i . We will establish the following two statements simultaneously by induction on s . To see that the restriction to s -secure subterms is necessary, consider the counterexample that would arise from $s = \vartheta_1 \vartheta_0 \vartheta_1 \vartheta_1 0$ and $t = \vartheta_1 \vartheta_1 0 = k_1(t)$.

- (1) If t is s -secure with $s = \vartheta_i s'$, then we cannot have $s \leq_{T(X)} k_i(t)$.
- (2) We do not have $s <_{T(X)} s$.

In the induction step for (1), we use side induction on t . We may write $k_i(t) = \vartheta_i t'$, because the conclusion is trivial when $k_i(t)$ has a different form. Let us exclude all reasons for which $s \leq_{T(X)} k_i(t)$ could hold: First, note that $h(k_i(t)) \leq h(t) < h(s)$ excludes equality. Secondly, given that t is s -secure, the same holds for t' . Thus the side induction hypothesis excludes $s \leq_{T(X)} k_i(t')$. Finally, the only remaining reason would involve $k_i(s') <_{T(X)} k_i(t)$. The latter entails that $k_i(s')$ and $k_i(t)$ are different, by part (2) of the simultaneous induction hypothesis. It follows that t is $k_i(s')$ -secure, so that the main induction hypothesis excludes $k_i(s') <_{T(X)} k_i(t)$. Concerning the induction step for (2), we note that only terms of the form $s = \vartheta_i s'$ are interesting. Since the induction hypothesis excludes $s' <_{T(X)} s'$, the inequality $s <_{T(X)} s$ would require $s \leq_{T(X)} k_i(s')$. This, however, is excluded by part (1). Trichotomy and transitivity are readily established by induction on the combined term complexity (e. g., on $h(s) + h(t)$ for trichotomy between s and t). □

Together with trichotomy, statement (1) from the previous proof yields:

Corollary 3.3 *We have $k_i(s) <_{T(X)} \vartheta_i s$ whenever $\vartheta_i s \in T(X)$.*

We will be particularly interested in the following suborders of $T(X)$.

Definition 3.4 For $n \in \mathbb{N}$ and a linear order X , define $T_n(X) \subseteq T(X)$ as the suborder of terms that contain indices below n only. Equivalently, $T_n(X)$ is generated just as $T(X)$, but with the additional restriction $i < n$ in clause (ii) of Definition 3.1. Furthermore, we define $T_n^0(X) = \{s \in T_n(X) \mid S(s) \leq 0\}$ as the suborder of terms that have the form \bar{x} or outer index 0. We will also write $<_{T(X)}$ (or just $<$) for the restriction of this order to $T_n(X)$ and $T_n^0(X)$.

As mentioned before, we have the following connection:

Corollary 3.5 *The orders T_n and $T_n[0]$ from [20, Section 2.3.3] coincide with our orders $T_n(1)$ and $T_n^0(1)$, respectively (where 1 is the order with a single element).*

Proof If we write $1 = \{0\}$ and identify $\bar{0} \in T_n(1)$ with $0 \in T_n$, the definitions coincide except at one point: Definition 27 of [20] declares that $s > t$ and $\vartheta_i s \leq k_i(t)$ imply $\vartheta_i s < \vartheta_i t$. In the corresponding clause (iii') of our Definition 3.1, we have omitted the condition $s > t$, because it turns out to be superfluous: According to [20, Lemma 9] (cf. also Corollary 3.3 above) we have $k_i(t) < \vartheta_i t$. Hence it follows from transitivity that $\vartheta_i s \leq k_i(t)$ alone entails $\vartheta_i s < \vartheta_i t$. □

We want to show that T_{n+1}^0 is a Bachmann–Howard derivative of T_n^0 . Officially, this claim does only make sense once we have specified a dilator T_n^0 that extends the transformation $X \mapsto T_n^0(X)$ of linear orders. We defer this extension until later, because we want to start with the most interesting part of the construction:

Definition 3.6 For each number $n \in \mathbb{N}$ and any linear order X , we define a function $\sigma_X^n : T_n \circ T_{n+1}^0(X) \rightarrow T_{n+1}(X)$ by the recursive clauses

$$\sigma_X^n(\bar{r}) = r \text{ (with } r \in T_{n+1}^0(X) \text{) and } \sigma_X^n(\vartheta_i s) = \vartheta_{i+1} \sigma_X^n(s).$$

Now we define $\vartheta_X^n : T_n^0 \circ T_{n+1}^0(X) \rightarrow T_{n+1}^0(X)$ by setting $\vartheta_X^n(s) = \vartheta_0 \sigma_X^n(s)$. Finally, let $\iota_X^n : X \rightarrow T_{n+1}^0(X)$ be given by $\iota_X^n(x) = \bar{x}$.

Note that elements $r = \bar{x} \in T_{n+1}^0(X)$ give rise to elements $\bar{r} = \bar{\bar{x}} \in T_n \circ T_{n+1}^0(X)$. We have

$$S(\sigma_X^n(s)) = \begin{cases} S(s) - 1 & \text{if } s \text{ is of the form } \bar{\bar{x}}, \\ S(s) + 1 & \text{otherwise.} \end{cases}$$

For $i \geq 0$, it follows that the condition $i \geq S(s) - 1$ from Definition 3.1 is equivalent to $i + 1 \geq S(\sigma_X^n(s)) - 1$. This justifies the second clause in the definition of σ_X^n . To justify the definition of ϑ_X^n , it suffices to note that $s \in T_n^0 \circ T_{n+1}^0(X)$ entails $S(s) \leq 0$ and hence $0 \geq S(\sigma_X^n(s)) - 1$. To formulate the next result, we need one new piece of notation: For $s \in T(X)$, the value $k_{-1}(s) \in T(X)$ is always of the form \bar{x} with $x \in X$. We define $\underline{k} : T(X) \rightarrow X$ by setting

$$\underline{k}(s) = x \text{ for } k_{-1}(s) = \bar{x}.$$

In particular, $s \in T_n^0 \circ T_{n+1}^0(X)$ yields $\underline{k}(s) \in T_{n+1}^0(X)$. We will see that the following proposition ensures the crucial clause (iii) of Definition 2.5.

Proposition 3.7 *The function $\sigma_X^n : T_n \circ T_{n+1}^0(X) \rightarrow T_{n+1}(X)$ is an order isomorphism, for each $n \in \mathbb{N}$ and any linear order X . Furthermore, the following holds for all $s, t \in T_n^0 \circ T_{n+1}^0(X)$:*

- If we have $s < t$ and $\underline{k}(s) < \vartheta_X^n(t)$, then we have $\vartheta_X^n(s) < \vartheta_X^n(t)$.
- We have $\underline{k}(s) < \vartheta_X^n(s)$.

Proof An easy induction over an arbitrary term $r \in T_{n+1}(X)$ shows that it lies in the range of σ_X^n . As preparation for the rest of the proof, one inductively shows

$$k_{i+1}(\sigma_X^n(s)) = \sigma_X^n(k_i(s))$$

for $i \in \mathbb{N} \cup \{-1\}$. In the crucial case of a term $s = \bar{r}$, this follows from $k_i(s) = s$ and $\sigma_X^n(s) = r = k_{i+1}(r)$, where the latter relies on $S(r) \leq i + 1$ due to $r \in T_{n+1}^0(X)$. Now an easy induction on $h(s) + h(t)$ shows that $s < t$ implies $\sigma_X^n(s) < \sigma_X^n(t)$.

This implication does automatically upgrade to an equivalence, as we are concerned with linear orders. In particular σ_X^n is injective, and indeed an order isomorphism. Concerning the claims about ϑ_X^n , we first note that $\underline{k}(s) < \vartheta_X^n(t)$ entails

$$k_0(\sigma_X^n(s)) = \sigma_X^n(k_{-1}(s)) = \underline{k}(s) < \vartheta_0\sigma_X^n(t).$$

Since we already know that $s < t$ entails $\sigma_X^n(s) < \sigma_X^n(t)$, we get

$$\vartheta_X^n(s) = \vartheta_0\sigma_X^n(s) < \vartheta_0\sigma_X^n(t) = \vartheta_X^n(t)$$

under the given assumptions. Arguing as before, we see that the remaining claim amounts to $k_0(\sigma_X^n(s)) < \vartheta_0\sigma_X^n(s)$, which holds by Corollary 3.3. □

As promised, we now extend T_n^0 (and in the process also T_n) into an LO-dilator.

Definition 3.8 For each embedding $f : X \rightarrow Y$ of linear orders, we define a function $T(f) : T(X) \rightarrow T(Y)$ by the recursive clauses

$$T(f)(\bar{x}) = \overline{f(x)} \quad \text{and} \quad T(f)(\vartheta_i s) = \vartheta_i T(f)(s).$$

Furthermore, we define functions $\text{supp}_X^T : T(X) \rightarrow [X]^{<\omega}$ by setting

$$\text{supp}_X^T(s) = \{\underline{k}(s)\}.$$

We will also write supp_X^T for the restrictions of this function to $T_n(X)$ and to $T_n^0(X)$. By $T_n(f) : T_n(X) \rightarrow T_n(Y)$ and $T_n^0(f) : T_n^0(X) \rightarrow T_n^0(Y)$ we denote the restrictions of $T(f)$ with (co-)domains as given.

It is immediate that we have $S(T(f)(s)) = S(s)$, which confirms that the functions $T(f)$, $T_n(f)$ and $T_n^0(f)$ are well-defined with the indicated codomains. We have given recursive definitions because they easily generalize from sequences to more complicated data types (cf. the treatment of trees in [14]). In the present case, it may simplify matters if we observe

$$T(f)(\vartheta_{j_1} \dots \vartheta_{j_n} \bar{x}) = \vartheta_{j_1} \dots \vartheta_{j_n} \overline{f(x)} \quad \text{and} \quad \text{supp}_X^T(\vartheta_{j_1} \dots \vartheta_{j_n} \bar{x}) = \{x\}.$$

Let us now verify the following:

Proposition 3.9 *The previous definition extends T_n and T_n^0 into LO-dilators.*

Proof A straightforward induction over $s \in T(X)$ shows $k_i(T(f)(s)) = T(f)(k_i(s))$. Given an order embedding $f : X \rightarrow Y$, one can now check

$$s <_{T(X)} t \quad \Rightarrow \quad T(f)(s) <_{T(Y)} T(f)(t)$$

by induction over $h(s) + h(t)$ (for $h : T(X) \rightarrow \mathbb{N}$ as given after Definition 3.1). Two more easy inductions show that T respects identity morphisms and compositions. It

follows that T_n and T_n^0 are endofunctors of linear orders. By the definition of \underline{k} and the first line of this proof, we have $\underline{k}(T(f)(s)) = f(\underline{k}(s))$. This yields

$$\text{supp}_Y^\top \circ T(f)(s) = \{\underline{k}(T(f)(s))\} = [f]^{<\omega}(\{\underline{k}(s)\}) = [f]^{<\omega} \circ \text{supp}_X^\top(s),$$

so that supp^\top is natural. To conclude that T_n is a dilator, we show

$$\text{supp}_Y^\top(r) \subseteq \text{rng}(f) \Rightarrow r \in \text{rng}(T_n(f))$$

by induction over $r \in T_n(Y)$, still for $f : X \rightarrow Y$. In the base case of a term $r = \bar{y}$ we observe $\text{supp}_Y^\top(r) = \{\underline{k}(\bar{y})\} = \{y\}$. By the antecedent of our implication we may write $y = f(x)$, which yields $r = T_n(f)(\bar{x})$ as desired. In the step for $r = \vartheta_i r'$ we note that $k_{-1}(r) = k_{-1}(r')$ entails $\text{supp}_Y^\top(r) = \text{supp}_Y^\top(r')$. Given the antecedent of our implication, we can thus invoke the induction hypothesis to get $r' = T_n(f)(r'_0)$ for some $r'_0 \in T_n(X)$. In view of $S(r'_0) = S(T(f)(r'_0)) = S(r')$ we may form the term $\vartheta_i r'_0 \in T_n(X)$ to get $r = T_n(f)(\vartheta_i r'_0) \in \text{rng}(T_n(f))$. In order to deduce the analogous implication for T_n^0 , we need only observe that $T_n(f)(r_0) = r \in T_n^0(Y)$ entails $S(r_0) = S(r) \leq 0$ and hence $r_0 \in T_n^0(X) \subseteq T_n(X)$. \square

The following theorem is the main result of this section. We write ι^n and ϑ^n for the families of functions $\iota_X^n : X \rightarrow T_{n+1}^0(X)$ and $\vartheta_X^n : T_n^0 \circ T_{n+1}^0(X) \rightarrow T_{n+1}^0(X)$, which are indexed by the linear order X (cf. Definition 3.6).

Theorem 3.10 *The Bachmann–Howard derivative of T_n^0 is given by $(T_{n+1}^0, \iota^n, \vartheta^n)$, for each number $n \in \mathbb{N}$.*

Proof A straightforward induction over terms shows that the functions σ_X^n from Definition 3.6 are natural in X . One can conclude that the same holds for ι_X^n and ϑ_X^n . In view of Definition 2.10, it remains to show that $(T_{n+1}^0(X), \iota_X^n, \vartheta_X^n)$ is an initial Bachmann–Howard fixed point of T_n^0 over X , whenever X is a linear order. Clauses (i) and (ii) of Definition 2.5 are immediate by our constructions. Clause (iii) holds by Proposition 3.7 and the definition of supp^\top . To complete the proof, we verify the criteria from part (i) of Theorem 2.9. The first criterion demands

$$T_{n+1}^0(X) = \text{rng}(\iota_X^n) \cup \text{rng}(\vartheta_X^n).$$

To see that this holds, consider a term $\vartheta_0 s \in T_{n+1}^0(X)$. We note that $s \in T_{n+1}(X)$ must satisfy $S(s) \leq 1$, due to Definition 3.1. From Proposition 3.7 we know that σ_X^n is surjective, which yields $s = \sigma_X^n(s')$ for some $s' \in T_n \circ T_{n+1}^0(X)$. By the paragraph after Definition 3.6 we get $S(s') = -1$ or $S(s') = S(s) - 1 \leq 0$, which means that we even have $s' \in T_n^0 \circ T_{n+1}^0(X)$. We now see

$$\vartheta_0 s = \vartheta_0 \sigma_X^n(s') = \vartheta_X^n(s') \in \text{rng}(\vartheta_X^n) \quad \text{with} \quad \vartheta_X^n : T_n^0 \circ T_{n+1}^0(X) \rightarrow T_{n+1}^0(X),$$

as required. The second criterion requires a function $h : T_{n+1}^0(X) \rightarrow \mathbb{N}$ with

$$h(r) < h(\vartheta_X^n(s)) \quad \text{for any } s \in T_n^0 \circ T_{n+1}^0(X) \text{ and } r \in \text{supp}_{T_{n+1}^0(X)}^\top(s).$$

We show that this holds for (the restriction of) the function $h : T(X) \rightarrow \mathbb{N}$ from the paragraph after Definition 3.1. By our definition of supports, the only possibility is $r = \underline{k}(s)$. By induction over $s \in T_n \circ T_{n+1}^0(X)$ one can verify $h(\underline{k}(s)) \leq h(\sigma_X^n(s))$. In view of $h(\vartheta_X^n(s)) = h(\vartheta_0\sigma_X^n(s)) = h(\sigma_X^n(s)) + 1$ this yields the claim. \square

The LO-dilator T_n^0 has been defined in two different ways: First, we have constructed T_n^0 in terms of iterated Bachmann–Howard derivatives, according to steps (1) and (2) from the introduction. Secondly, we have given an *ad hoc* construction of T_n^0 in Definitions 3.1, 3.4 and 3.8. The results of the two constructions coincide by Theorem 3.10, as explained in the first two paragraphs of this section. Let us point out that there is, nevertheless, an interesting difference: For the *ad hoc* definition of T_n^0 , we needed to define T_n as an auxiliary construct (which also appears in many inductive verifications). In contrast, the recursive construction via Bachmann–Howard derivatives yields T_n^0 directly.

4 Theory, part 2: connecting with Kruskal derivatives

In Sect. 2 we have introduced the Bachmann–Howard derivative ϑD of an LO-dilator D . A parallel construction on the level of partial orders was previously studied in [14]: For each suitable dilator W on partial orders, it yields the so-called Kruskal derivative $\mathcal{T}W$. In the present section we establish fundamental connections between Bachmann–Howard and Kruskal derivatives, i. e., between the linear and the partial case.

Important notions from [14] will be recalled informally, but the reader may need to consult the cited reference for precise definitions. A function $f : X \rightarrow Y$ between partial orders is a quasi embedding if it is order reflecting, i. e., if $f(x) \leq_Y f(x')$ implies $x \leq_X x'$. We consider the category PO of partial orders with the quasi embeddings as morphisms. A PO-dilator is a functor $W : PO \rightarrow PO$ that satisfies certain conditions, in particular a support condition as in Definition 2.1 above (see [14, Definition 2.1] for details). We call W a WPO-dilator if, in addition, $W(X)$ is a well partial order whenever the same holds for X . By $W \upharpoonright LO : LO \rightarrow PO$ we denote the restriction of a PO-dilator W to the category of linear orders. Also, we sometimes consider an LO-dilator D as a functor $D : LO \rightarrow PO$, i. e., we implicitly compose it with the inclusion $LO \hookrightarrow PO$. We can then consider ν as in the following:

Definition 4.1 By a quasi embedding of an LO-dilator D into a PO-dilator W we mean a natural transformation $\nu : D \Rightarrow W \upharpoonright LO$.

Note that ν consists of a quasi embedding $\nu_X : D(X) \rightarrow W(X)$ for each linear order X . If a PO-dilator W satisfies a certain normality condition, then it has an essentially unique Kruskal derivative $\mathcal{T}W$ (see Definition 2.3 and Section 4 of [14]). The latter comes with natural families of functions

$$\iota_X^W : X \rightarrow \mathcal{T}W(X) \quad \text{and} \quad \kappa_X^W : W \circ \mathcal{T}W(X) \rightarrow \mathcal{T}W(X),$$

indexed by the partial order X . From Sect. 2 above we recall that the Bachmann–Howard derivative ϑD of an LO-dilator D comes with functions

$$\iota_X^D : X \rightarrow \vartheta D(X) \quad \text{and} \quad \vartheta_X^D : D \circ \vartheta D(X) \rightarrow \vartheta D(X),$$

where X does now range over linear orders. Together with Theorem 4.8 below, the following result provides the connection between Kruskal and Bachmann–Howard derivatives. The theorem extends [4, Theorem 4.5], in which ν_X^+ is only constructed for the empty order $X = 0$. Even though the main idea remains the same, we provide full details, as the setting and notation in [4] are somewhat different.

Theorem 4.2 *Let $\nu : D \Rightarrow W \upharpoonright LO$ be a quasi embedding of an LO-dilator D into a normal PO-dilator W . There is a unique quasi embedding $\nu^+ : \vartheta D \Rightarrow TW \upharpoonright LO$ such that the diagram*

$$\begin{array}{ccc}
 & \vartheta D(X) & \xleftarrow{\vartheta_X^D} D \circ \vartheta D(X) \\
 \iota_X^D \nearrow & \downarrow \nu_X^+ & \downarrow W(\nu_X^+) \circ \nu_{\vartheta D(X)} \\
 X & & \\
 \iota_X^W \searrow & TW(X) & \xleftarrow{\kappa_X^W} W \circ TW(X)
 \end{array}$$

commutes for each linear order X .

Proof By Lemma 2.2, each $\sigma \in D \circ \vartheta D(X)$ has a normal form $\sigma = {}_{\text{NF}}D(\iota_a)(\sigma_0)$ with $a \subseteq \vartheta D(X)$ and $\sigma_0 \in D(a)$. Note that $\nu_{\vartheta D(X)} \circ D(\iota_a) = W(\iota_a) \circ \nu_a$ holds since ν is natural. Hence the diagram in the theorem commutes if, and only if, we have

$$\begin{aligned}
 \nu_X^+(\iota_X^D(x)) &= \iota_X^W(x) \quad \text{for } x \in X, \\
 \nu_X^+(\vartheta_X^D(\sigma)) &= \kappa_X^W \circ W(\nu_X^+ \circ \iota_a) \circ \nu_a(\sigma_0) \quad \text{for } \sigma = {}_{\text{NF}}D(\iota_a)(\sigma_0).
 \end{aligned}$$

The idea is to read these equations as recursive clauses, which is justified as follows: According to Definition 2.10, the tuple $(\vartheta D(X), \iota_X^D, \vartheta_X^D)$ is an initial Bachmann–Howard fixed point of D over X . By (the proof of) Theorem 2.9, it follows that the functions ι_X^D and ϑ_X^D are injective, and that $\vartheta D(X)$ is the disjoint union of their ranges. Furthermore, the same theorem yields a function $h : \vartheta D(X) \rightarrow \mathbb{N}$ such that

$$s \in \text{supp}_{\vartheta D(X)}^D(\sigma) \Rightarrow h(s) < h(\vartheta_X^D(\sigma))$$

holds for any element $\sigma \in D \circ \vartheta D(X)$. Here supp^D is the support that comes with the LO-dilator D (see Definition 2.1 and the discussion that follows it). Now recall that $\sigma = {}_{\text{NF}}D(\iota_a)(\sigma_0)$ entails $\text{supp}_{\vartheta D(X)}^D(\sigma) = a$, by Lemma 2.2. This means that the clauses above define $\nu_X^+(s)$ by recursion over $h(s)$. More precisely, a straightforward induction on $h(s)$ shows that the value $\nu_X^+(s)$ is uniquely determined, i.e., that there is at most one quasi embedding ν_X^+ such that the diagram in the theorem commutes. The proof of existence is somewhat more subtle, since we must simultaneously show

that v_X^+ is a quasi embedding, in order to ensure that $W(v_X^+ \circ \iota_a)$ is defined. Let us define $l : \vartheta D(X) \rightarrow \mathbb{N}$ by stipulating $l(\iota_X^D(x)) = 0$ and

$$l(\vartheta_X^D(\sigma)) = 1 + \sum_{s \in a} 2 \cdot l(s) \quad \text{for } \sigma = \text{NF}D(\iota_a)(\sigma_0),$$

which is itself a recursion based on h . By simultaneous induction on the values $l(r)$ and $l(s) + l(t)$ one can now show that $v_X^+(r) \in \mathcal{T}W(X)$ is defined and that we have

$$v_X^+(s) \leq_{\mathcal{T}W(X)} v_X^+(t) \quad \Rightarrow \quad s \leq_{\vartheta D(X)} t.$$

To establish this implication, we distinguish cases according to the forms of s and t . First assume $s = \iota_X^D(x)$ and $t = \iota_X^D(y)$. Then the antecedent of our implication amounts to $\iota_X^W(x) \leq_{\mathcal{T}W(X)} \iota_X^W(y)$. According to [14, Definition 3.1] we get $x \leq_X y$. Now $s \leq_{\vartheta D(X)} t$ follows by clause (i) of Definition 2.5 above. By the same definition, we always have $s \leq_{\vartheta D(X)} t$ for terms of the form $s = \iota_X^D(x)$ and $t = \vartheta_X^D(\tau)$. For terms $s = \vartheta_X^D(\sigma)$ and $t = \iota_X^D(y)$, say with $\sigma = \text{NF}D(\iota_a)(\sigma_0)$, we have

$$v_X^+(s) = \kappa_X^W \circ W(v_X^+ \circ \iota_a) \circ v_a(\sigma_0) \not\leq_{\mathcal{T}W(X)} \iota_X^W(x) = v_X^+(t),$$

again by [14, Definition 3.1]. It remains to compare terms $s = \vartheta_X^D(\sigma)$ and $t = \vartheta_X^D(\tau)$. We write $\sigma = \text{NF}D(\iota_a)(\sigma_0)$ and $\tau = \text{NF}D(\iota_b)(\tau_0)$ and assume

$$\kappa_X^W \circ W(v_X^+ \circ \iota_a) \circ v_a(\sigma_0) = v_X^+(s) \leq_{\mathcal{T}W(X)} v_X^+(t) = \kappa_X^W \circ W(v_X^+ \circ \iota_b) \circ v_b(\tau_0).$$

According to [14, Definition 3.1], this inequality can hold for two different reasons. In the first case we have $v_X^+(s) \leq_{\mathcal{T}W(X)} t'$ for some element

$$t' \in \text{supp}_{\mathcal{T}W(X)}^W (W(v_X^+ \circ \iota_b) \circ v_b(\tau_0)) = [v_X^+ \circ \iota_b]^{<\omega} \left(\text{supp}_b^W (v_b(\tau_0)) \right).$$

By [4, Lemmas 4.2 and 4.4] any quasi embedding of an LO-dilator into a PO-dilator respects supports. This means that we have $\text{supp}_b^W (v_b(\tau_0)) = \text{supp}_b^D(\tau_0) = b$, where the last equality comes from the normal form condition (see the paragraph before Definition 2.2). We can thus write $t' = v_X^+(t'_0)$ for some $t'_0 \in b$. The latter entails that we have $l(t'_0) < l(t)$, so that we get $s \leq_{\vartheta D(X)} t'_0$ by induction hypothesis. By Lemma 2.2 we have $b = \text{supp}_{\vartheta D(X)}^D(\tau)$, so that clause (iii) of Definition 2.5 allows us to conclude $t'_0 <_{\vartheta D(X)} \vartheta_X^D(\tau) = t$. Now transitivity yields $s \leq_{\vartheta D(X)} t$, as desired (in fact the inequality is strict in this case). In the remaining case, the above inequality $v_X^+(s) \leq_{\mathcal{T}W(X)} v_X^+(t)$ holds because we have

$$W(v_X^+ \circ \iota_a) \circ v_a(\sigma_0) \leq_{W \circ \mathcal{T}W(X)} W(v_X^+ \circ \iota_b) \circ v_b(\tau_0).$$

Let us factor $\iota_a = \iota_{a \cup b} \circ \iota'_a$ and $\iota_b = \iota_{a \cup b} \circ \iota'_b$ with $\iota_{a \cup b} : a \cup b \hookrightarrow \vartheta D(X)$. The induction hypothesis ensures that $v_X^+ \circ \iota_{a \cup b}$ is a quasi embedding, which allows us to

form the quasi embedding $W(v_X^+ \circ \iota_{a \cup b})$. The previous inequality thus entails

$$W(\iota'_a) \circ v_a(\sigma_0) \leq_{W(a \cup b)} W(\iota'_b) \circ v_b(\tau_0). \tag{4.1}$$

Due to the naturality of v , we get $v_{a \cup b} \circ D(\iota'_a)(\sigma_0) \leq_{W(a \cup b)} v_{a \cup b} \circ D(\iota'_b)(\tau_0)$. Since $v_{a \cup b}$ reflects the order while $D(\iota_{a \cup b})$ preserves it, this entails

$$\sigma = D(\iota_a)(\sigma_0) = D(\iota_{a \cup b}) \circ D(\iota'_a)(\sigma_0) \leq_{D \circ \vartheta D(\sigma)} D(\iota_{a \cup b}) \circ D(\iota'_b)(\tau_0) = \tau.$$

In order to conclude $s = \vartheta_X^D(\sigma) \leq_{\vartheta D(X)} \vartheta_X^D(\tau) = t$ by clause (iii) of Definition 2.5, it remains to show that we have

$$r <_{\vartheta D(X)} \vartheta_X^D(\tau) \text{ for all } r \in \text{supp}_{\vartheta D(X)}^D(\sigma) = a.$$

Analogous to the above, we get

$$a = \text{supp}_a^D(\sigma_0) = [\iota'_a]^{<\omega} \left(\text{supp}_a^W(v_a(\sigma_0)) \right) = \text{supp}_{a \cup b}^W(W(\iota'_a) \circ v_a(\sigma_0)),$$

as well as $b = \text{supp}_{a \cup b}^W(W(\iota'_b) \circ v_b(\tau_0))$. Due to inequality (4.1) and the assumption that W is normal (cf. [14, Definition 2.3]), it follows that any $r \in a$ admits an $r' \in b$ with $r \leq_{\vartheta D(X)} r'$ (note that the inequality holds in $\vartheta D(X)$ because $a \cup b$ is a suborder of the latter). Once again, clause (iii) of Definition 2.5 yields $r' <_{\vartheta D(X)} \vartheta_X^D(\tau)$. Now transitivity allows us to conclude $r <_{\vartheta D(X)} \vartheta_X^D(\tau)$, as needed. This completes the simultaneous proof that v_X^+ is well defined and a quasi embedding. It remains to show naturality. Given a quasi embedding $f : X \rightarrow Y$, we prove

$$v_Y^+ \circ \vartheta D(f)(s) = TW(f) \circ v_X^+(s)$$

by induction over $h(s)$, for $h : \vartheta D(X) \rightarrow \mathbb{N}$ as above. The crucial case concerns a term $s = \vartheta_X^D(\sigma)$, say with $\sigma = \text{NF}D(\iota_a)(\sigma_0)$. Using the naturality of ϑ^D , we get

$$\vartheta D(f)(s) = \vartheta D(f) \circ \vartheta_X^D \circ D(\iota_a)(\sigma_0) = \vartheta_Y^D \circ D(\vartheta D(f) \circ \iota_a)(\sigma_0).$$

In order to apply v_Y^+ to the expression on the right side, we need to determine the normal form of $D(\vartheta D(f) \circ \iota_a)(\sigma_0)$. Consider the restriction $\vartheta D(f) \upharpoonright a : a \rightarrow b$ with codomain $b := [\vartheta D(f)]^{<\omega}(a)$. In view of $\vartheta D(f) \circ \iota_a = \iota_b \circ (\vartheta D(f) \upharpoonright a)$ we get

$$D(\vartheta D(f) \circ \iota_a)(\sigma_0) = D(\iota_b)(\sigma_1) \text{ for } \sigma_1 := D(\vartheta D(f) \upharpoonright a)(\sigma_0) \in D(b).$$

This expression is in normal form, since the naturality of supports yields

$$\text{supp}_b^D(\sigma_1) = [\vartheta D(f) \upharpoonright a]^{<\omega} \left(\text{supp}_a^D(\sigma_0) \right) = [\vartheta D(f) \upharpoonright a]^{<\omega}(a) = b,$$

where $\text{supp}_a^D(\sigma_0) = a$ comes from the normal form condition for $\sigma = \text{NF}D(\iota_a)(\sigma_0)$. By the recursive definition of v_Y^+ , we now obtain

$$v_Y^+ \circ \vartheta D(f)(s) = \kappa_Y^W \circ W(v_Y^+ \circ \iota_b) \circ v_b(\sigma_1).$$

The naturality of v yields $v_b(\sigma_1) = W(\vartheta D(f) \upharpoonright a) \circ v_a(\sigma_0)$ and hence

$$W(\iota_b) \circ v_b(\sigma_1) = W(\iota_b \circ (\vartheta D(f) \upharpoonright a)) \circ v_a(\sigma_0) = W(\vartheta D(f) \circ \iota_a) \circ v_a(\sigma_0).$$

Since $r \in a$ entails $h(r) < h(s)$, we have $v_Y^+ \circ \vartheta D(f) \circ \iota_a = TW(f) \circ v_X^+ \circ \iota_a$ by induction hypothesis. Putting things together, we can finally conclude

$$\begin{aligned} v_Y^+ \circ \vartheta D(f)(s) &= \kappa_Y^W \circ W(TW(f) \circ v_X^+ \circ \iota_a) \circ v_a(\sigma_0) \\ &= TW(f) \circ \kappa_X^W \circ W(v_X^+ \circ \iota_a) \circ v_a(\sigma_0) = TW(f) \circ v_X^+(s), \end{aligned}$$

as needed for the inductive proof that v^+ is natural. □

We now deduce the result that was promised at the end of Sect. 2. Our proof is somewhat indirect but nevertheless instructive, as it connects several fundamental facts. Even though we do not formalize the present paper in a specific meta theory, we point out that the following argument uses Π_1^1 -comprehension, in the form of the minimal bad sequence lemma. This is unavoidable by the results of [17, 19].

Corollary 4.3 *If D is a WO-dilator (i. e., preserves well foundedness), then so is its Bachmann–Howard derivative ϑD .*

Proof That ϑD is an LO-dilator is guaranteed by Definition 2.10 (but see also Proposition 2.11). It remains to show that $\vartheta D(X)$ is a well order when the same holds for X . According to [4, Section 5] there is a quasi embedding $v : D \Rightarrow W_D \upharpoonright \text{LO}$ into a normal WPO-dilator W_D . By the previous theorem we get a quasi embedding

$$v_X^+ : \vartheta D(X) \rightarrow TW_D(X)$$

for each linear order X . Assume that X is a well order and hence a well partial order. Then $TW_D(X)$ is also a well partial order, by [14, Proposition 2.7] (which is proved by the minimal bad sequence method of Nash-Williams [3]). Given an infinite sequence s_0, s_1, \dots in $\vartheta D(X)$, we get a sequence $v_X^+(s_0), v_X^+(s_1), \dots$ in $TW_D(X)$. Since the latter is a well partial order, there are $i < j$ such that we have $v_X^+(s_i) \leq_{TW_D(X)} v_X^+(s_j)$. We can infer $s_i \leq_{\vartheta D(X)} s_j$, as v_X^+ is a quasi embedding and hence order reflecting. This shows that $\vartheta D(X)$ is a well order. □

As an immediate consequence of our general approach, we re-obtain the following known result about the orders $T_n[0]$ from [20, Section 2.3.3].

Corollary 4.4 *The order $T_n[0]$ is well founded for each $n \in \mathbb{N}$.*

Proof In the previous section we have studied LO-dilators T_n^0 with $T_n[0] \cong T_n^0(1)$. It suffices to show that these are WO-dilators. We argue by induction on n . In the base we need only observe $T_n^0(X) \cong X$. The step is covered by the previous corollary, as T_{n+1}^0 is the Bachmann–Howard derivative of T_n^0 , by Theorem 3.10. \square

Let $f : X \rightarrow Y$ be a quasi embedding between partial orders. If X is linear and f is surjective, then we call f a linearization (of Y by X). This coincides with the usual notion if we identify X with its image under f . Linearizations are particularly important, as they are related to maximal order types and independence results (see e.g. [11, 25–27]). We introduce the corresponding functorial notion:

Definition 4.5 Consider an LO-dilator D and a PO-dilator W . A quasi embedding $\nu : D \Rightarrow W \upharpoonright \text{LO}$ is called a linearization (of W by D) if $\nu_X : D(X) \rightarrow W(X)$ is surjective for each linear order X .

Our next goal is to identify a condition which ensures that the quasi embedding $\nu^+ : \vartheta D \Rightarrow TW \upharpoonright \text{LO}$ from Theorem 4.2 is a linearization. As the following example shows, the assumption that $\nu : D \Rightarrow W \upharpoonright \text{LO}$ is a linearization does not suffice.

Example 4.6 We define a transformation $X \mapsto W(X)$ of partial orders by

$$W(X) = \{(x, x') \in X^2 \mid x' \not\leq_X x\},$$

$$(x, x') \leq_{W(X)} (y, y') \iff x \leq_X y \text{ and } x' \leq_X y'.$$

If $f : X \rightarrow Y$ is a quasi embedding, then $x' \not\leq_X x$ implies $f(x') \not\leq_Y f(x)$. We thus get a function $W(f) : W(X) \rightarrow W(Y)$ by setting $W(f)(x, x') = (f(x), f(x'))$. Let us also define functions $\text{supp}_X^W : W(X) \rightarrow [X]^{<\omega}$ by putting $\text{supp}_X^W(x, x') = \{x, x'\}$. It is straightforward to verify that this data constitutes a normal PO-dilator. Here it is crucial that the support condition

$$\text{supp}_Y^W(\sigma) \subseteq \text{rng}(f) \implies \sigma \in \text{rng}(W(f))$$

is only required when $f : X \rightarrow Y$ is an embedding, rather than just a quasi embedding (see [14, Definition 2.1]). In the present case, the support condition does in fact imply that f is order preserving. For each linear order X , we put

$$D(X) = \{(x, x') \in X^2 \mid x <_X x'\},$$

$$(x, x') \leq_{D(X)} (y, y') \iff x <_X y \text{ or } (x = y \text{ and } x' \leq_X y').$$

Given that X is linear, the underlying sets of $D(X)$ and $W(X)$ are equal. We may thus declare that $D(f) : D(X) \rightarrow D(Y)$ and $\text{supp}_X^D : D(X) \rightarrow [X]^{<\omega}$ coincide with $W(f)$ and supp_X^W , respectively, for any embedding $f : X \rightarrow Y$ of linear orders. It is easy to see that this yields an LO-dilator. The identity maps $\nu_X : D(X) \rightarrow W(X)$ constitute a linearization $\nu : D \Rightarrow W \upharpoonright \text{LO}$. Theorem 4.2 gives rise to a quasi embedding $\nu^+ : \vartheta D \Rightarrow TW \upharpoonright \text{LO}$, which we shall now describe concretely. For each

partial order X , we consider a collection $\mathcal{B}(X)$ of structured binary trees with leaf labels from X , which is recursively generated as follows:

- For each $x \in X$ we have an element $\bar{x} \in \mathcal{B}(X)$ (a single node with label x).
- Given $s, s' \in \mathcal{B}(X)$ we add an element $\circ(s, s') \in \mathcal{B}(X)$ (the tree in which the root has immediate subtrees s and s').

To extend \mathcal{B} into a functor (currently on sets but later on orders), we stipulate

$$\mathcal{B}(f)(\bar{x}) = \overline{f(x)} \quad \text{and} \quad \mathcal{B}(f)(\circ(s, s')) = \circ(\mathcal{B}(f)(s), \mathcal{B}(f)(s')).$$

We also define support functions $\text{supp}_X^{\mathcal{B}} : \mathcal{B}(X) \rightarrow [X]^{<\omega}$ by setting

$$\text{supp}_X^{\mathcal{B}}(\bar{x}) = \{x\} \quad \text{and} \quad \text{supp}_X^{\mathcal{B}}(\circ(s, s')) = \text{supp}_X^{\mathcal{B}}(s) \cup \text{supp}_X^{\mathcal{B}}(s').$$

To describe the Kruskal derivative $\mathcal{T}W$ of W , we first define a partial order $\leq_{\mathcal{B}(X)}^W$ on $\mathcal{B}(X)$. The latter coincides with the usual embeddability relation for labelled trees, which can be given by the following recursive clauses:

- (i) If we have $x \leq_X y$, then we have $\bar{x} \leq_{\mathcal{B}(X)}^W \bar{y}$.
- (ii) We have $s \leq_{\mathcal{B}(X)}^W \circ(s, s')$ and $s' \leq_{\mathcal{B}(X)}^W \circ(s, s')$.
- (iii) If we have $s \leq_{\mathcal{B}(X)}^W t$ and $s' \leq_{\mathcal{B}(X)}^W t'$, then we have $\circ(s, s') \leq_{\mathcal{B}(X)}^W \circ(t, t')$.

One can check that this turns \mathcal{B} into a normal PO-dilator. Indeed, \mathcal{B} is the Kruskal derivative of $X \mapsto X^2$, where the order on X^2 extends the one on $W(X)$ in the obvious way. In view of $W(X) \subsetneq X^2$, we now define $\mathcal{T}W(X) \subseteq \mathcal{B}(X)$ as the smallest suborder that contains all elements \bar{x} and contains $\circ(s, s')$ for any $s, s' \in \mathcal{T}W(X)$ with $s' \not\leq_{\mathcal{B}(X)}^W s$. We can turn $\mathcal{T}W$ into a normal PO-dilator by restricting the constructions from above. Let us define

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

by setting $\iota_X^W(x) = \bar{x}$ and $\kappa_X^W(s, s') = \circ(s, s')$. As one readily verifies, these functions witness that $\mathcal{T}W$ is the Kruskal derivative of W (so that the notation $\mathcal{T}W$ is indeed justified). To describe the Bachmann–Howard derivative ϑD of D , we first refine $\leq_{\mathcal{B}(X)}^W$ into a linear order $\leq_{\mathcal{B}(X)}^D$ on $\mathcal{B}(X)$. The latter is characterized by the following clauses together with (i-iii) above (with $\leq_{\mathcal{B}(X)}^D$ at the place of $\leq_{\mathcal{B}(X)}^W$):

- (iv) We have $\bar{x} <_{\mathcal{B}(X)}^D \circ(t, t')$ for any terms of the given forms.
- (v) If we have $s <_{\mathcal{B}(X)}^D t$ and $s' <_{\mathcal{B}(X)}^D \circ(t, t')$, we have $\circ(s, s') <_{\mathcal{B}(X)}^D \circ(t, t')$.

Analogous to the above, let $\vartheta D(X) \subseteq \mathcal{B}(X)$ be the smallest (linear) suborder that contains all elements \bar{x} and contains $\circ(s, s')$ for any $s, s' \in \vartheta D(X)$ with $s <_{\mathcal{B}(X)}^D s'$. To turn ϑD into an LO-dilator, it suffices to restrict the functions $\mathcal{B}(f)$ and $\text{supp}_X^{\mathcal{B}}$ from above. We define

$$\iota_X^D : X \rightarrow \vartheta D(X) \quad \text{and} \quad \vartheta_X^D : D \circ \vartheta D(X) \rightarrow \vartheta D(X)$$

as the restrictions of ι_X^W and κ_X^W , i.e. by $\iota_X^D(x) = \bar{x}$ and $\vartheta_X^D(s, s') = \circ(s, s')$. These functions witness that ϑD is the Bachmann–Howard derivative of D . Since $\leq_{\mathcal{B}(X)}^D$ refines $\leq_{\mathcal{B}(X)}^W$, we have $\vartheta D(X) \subseteq \mathcal{T}W(X)$. One can verify that the inclusion maps

$$\nu_X^+ : \vartheta D(X) \hookrightarrow \mathcal{T}W(X)$$

satisfy the conditions from Theorem 4.2, which characterize ν^+ uniquely. We now consider $X = \{0, 1, 2\}$ with the usual linear order. The element $\circ(\circ(\bar{0}, \bar{1}), \bar{2}) \in \mathcal{B}(X)$ is contained in $\mathcal{T}W(X)$ but not in $\vartheta D(X)$, so that ν_X^+ is not surjective. This means that $\nu^+ : \vartheta D \Rightarrow \mathcal{T}W \upharpoonright \text{LO}$ is no linearization, even though $\nu : D \Rightarrow W \upharpoonright \text{LO}$ is one.

The previous example suggests the following notion:

Definition 4.7 A PO-dilator W is called flat if the support condition

$$\text{supp}_Y^W(\sigma) \subseteq \text{rng}(f) \Rightarrow \sigma \in \text{rng}(W(f))$$

holds for any quasi embedding $f : X \rightarrow Y$ between partial orders (recall that [14, Definition 2.1] does only require this condition for embeddings).

As expected, we get the following (cf. Theorem 4.2):

Theorem 4.8 Consider a linearization $\nu : D \Rightarrow W \upharpoonright \text{LO}$ of a normal PO-dilator W by an LO-dilator D . If W is flat, then $\nu^+ : \vartheta D \Rightarrow \mathcal{T}W \upharpoonright \text{LO}$ is again a linearization.

Proof Given a linear order X , we need to show that $\nu_X^+ : \vartheta D(X) \rightarrow \mathcal{T}W(X)$ is surjective. According to [14, Definition 4.1], the order $\mathcal{T}W(X)$ comes with functions $\iota_X^W : X \rightarrow \mathcal{T}W(X)$ and $\kappa_X^W : W \circ \mathcal{T}W(X) \rightarrow \mathcal{T}W(X)$ that turn it into an initial Kruskal fixed point. By [14, Theorem 3.5] we get a function $h : \mathcal{T}W(X) \rightarrow \mathbb{N}$ with

$$r \in \text{supp}_{\mathcal{T}W(X)}^W(\sigma) \Rightarrow h(r) < h(\kappa_X^W(\sigma)).$$

For $s \in \mathcal{T}W(X)$ we shall now show $s \in \text{rng}(\nu_X^+)$ by induction on $h(s)$. Also by [14, Theorem 3.5], it suffices to consider the following two cases: First, let us look at an element $s = \iota_X^W(x)$. Here we obtain $s = \nu_X^+(\iota_X^D(x)) \in \text{rng}(\nu_X^+)$ by the diagram in Theorem 4.2. Secondly, consider $s = \kappa_X^W(\sigma)$. The induction hypothesis yields

$$\text{supp}_{\mathcal{T}W(X)}^W(\sigma) \subseteq \text{rng}(\nu_X^+).$$

Let us recall that ν_X^+ is a quasi embedding but no embedding (unless $\mathcal{T}W(X)$ is linear). Hence the support condition from [14, Definition 2.1] does not apply. But since W is flat we get $\sigma \in \text{rng}(W(\nu_X^+))$ anyway. We also know that $\nu_{\vartheta D(X)}$ is surjective, by the assumption that ν is a linearization. Thus we may write

$$\sigma = W(\nu_X^+) \circ \nu_{\vartheta D(X)}(\sigma_0) \text{ for some } \sigma_0 \in D \circ \vartheta D(X).$$

Due to the diagram in Theorem 4.2, we then obtain

$$s = \kappa_X^W(\sigma) = \nu_X^+ \circ \vartheta_X^D(\sigma_0) \in \text{rng}(\nu_X^+),$$

as needed to complete the inductive proof. \square

In the following section we will construct sequences with gap condition via iterated Kruskal derivatives. The following result will ensure that all PO-dilators in this construction are flat.

Proposition 4.9 *We consider a normal PO-dilator W . If W is flat, then so is its Kruskal derivative $\mathcal{T}W$.*

Proof Given a quasi embedding $f : X \rightarrow Y$, we need to prove the support condition

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

for arbitrary $s \in \mathcal{T}W(Y)$. Similarly to the previous proof, this can be achieved by induction on $h(s)$, where $h : \mathcal{T}W(Y) \rightarrow \mathbb{N}$ with

$$r \in \text{supp}_{\mathcal{T}W(Y)}^W(\sigma) \Rightarrow h(r) < h(\kappa_Y^W(\sigma))$$

is provided by [14, Theorem 3.5]. An element $s = \iota_Y^W(y)$ has support $\{y\}$, by the proof of [14, Theorem 4.2]. Given that the premise of the support condition holds, we thus get $y = f(x)$ for some $x \in X$. The diagram in [14, Definition 4.1] yields

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

as required. It remains to consider an element $s = \kappa_Y^W(\sigma)$. Here we have

$$\text{supp}_Y^{\mathcal{T}W}(s) = \bigcup \{ \text{supp}_X^{\mathcal{T}W}(r) \mid r \in \text{supp}_{\mathcal{T}W(X)}^W(\sigma) \},$$

again by the proof of [14, Theorem 4.2]. Assuming the premise of the support condition for s , we can thus invoke the induction hypothesis to get

$$\text{supp}_{\mathcal{T}W(X)}^W(\sigma) \subseteq \text{rng}(\mathcal{T}W(f)).$$

Using the assumption that W is flat, we obtain

$$\sigma = W \circ \mathcal{T}W(f)(\sigma_0) \quad \text{for some } \sigma_0 \in W \circ \mathcal{T}W(X).$$

Now the diagram in [14, Definition 4.1] yields

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

as needed for the inductive proof that $\mathcal{T}W$ is flat. \square

5 Application, part 2: linearizing the gap condition for sequences

As shown in Sect. 3, one can reconstruct certain collapsing functions by taking iterated Bachmann–Howard derivatives (of dilators on linear orders). In the present section we show that sequences with Friedman’s gap condition arise from a completely parallel construction in terms of Kruskal derivatives (of dilators on partial orders). This confirms that there is an extremely tight connection between collapsing functions and the gap condition. Also in this section, we give a second iterative construction that yields the same gap condition but different collapsing functions. This explains a phenomenon from [20], as discussed in the introduction of the present paper.

Definition 5.1 For each partial order X , let $S(X)$ and $S : S(X) \rightarrow \mathbb{N} \cup \{-1\}$ be generated by clauses (i) and (ii) from Definition 3.1 (with S at the place of T , so that $S(X) = T(X)$ when X is linear). Also, let $k_i : S(X) \rightarrow S(X)$ for $i \in \mathbb{N} \cup \{-1\}$ be given as in the cited definition. To define a binary relation $\leq_{S(X)}$ on $S(X)$, we declare that $s \leq_{S(X)} t$ holds if, and only if, one of the following clauses is satisfied:

- (i') We have $s = \bar{x}$ and $t = \bar{y}$ with $x \leq_X y$.
- (ii') We have $t = \vartheta_j t'$ with $s \leq_{S(X)} k_j(t')$ (where s can be of the form \bar{x} or $\vartheta_i s'$).
- (iii') We have $s = \vartheta_i s'$ and $t = \vartheta_i t'$ (for the same i) with $s' \leq_{S(X)} t'$.

Furthermore, let $S_n(X) \subseteq S(X)$ consist of the terms with indices $i < n$ only, and put $S_n^0(X) = \{s \in S_n(X) \mid S(s) \leq 0\}$ (cf. Definition 3.4). We will always consider these subsets with (the restrictions of) the relation $\leq_{S(X)}$ (sometimes written as \leq).

Let $h : S(X) \rightarrow \mathbb{N}$ be given as in the paragraph after Definition 3.1. Once again we have $h(k_i(s)) \leq h(s)$, so that $s \leq_{S(X)} t$ can be decided by recursion on $h(s) + h(t)$. Let us begin with a preparatory result:

Lemma 5.2 *If we have $s \leq_{S(X)} t$, then we get $k_i(s) \leq_{S(X)} k_i(t)$ for any $i \in \mathbb{N}$.*

Proof We argue by induction on $h(s) + h(t)$. First, assume that $s \leq \vartheta_j t' = t$ holds because of $s \leq k_j(t')$. We then get $k_i(s) \leq k_i(k_j(t'))$ by induction hypothesis. For $j \leq i$ we have $k_i(k_j(t')) = k_j(t')$, so that we obtain $k_i(s) \leq k_j(t')$ and then

$$k_i(s) \leq \vartheta_j t' = k_i(\vartheta_j t') = k_i(t).$$

For $j > i$ we get $k_i(k_j(t')) = k_i(t')$ and thus $k_i(s) \leq k_i(t') = k_i(t)$. Secondly, assume that $s = \vartheta_j s' \leq \vartheta_j t' = t$ holds because of $s' \leq t'$. For $i < j$ we have $k_i(s) = k_i(s')$ and $k_i(t) = k_i(t')$, so that it suffices to invoke the induction hypothesis. For $i \geq j$ we have $k_i(s) = s$ and $k_i(t) = t$, which means that the claim is trivial. □

The following is known for $X = 1 = \{0\}$ (see e.g. [20]), but the author has not found a reference for the (straightforward) generalization.

Lemma 5.3 *The relation $\leq_{S(X)}$ on $S(X)$ is a partial order.*

Proof Reflexivity is immediate by induction over terms. To establish antisymmetry one first shows that $s \leq t$ entails $h(s) \leq h(t)$, by induction over $h(s) + h(t)$. The

point is that $t \leq s$ becomes impossible if $s \leq t = \vartheta_j t'$ holds by clause (ii') above, since we then have

$$h(s) \leq h(k_j(t')) \leq h(t') < h(t).$$

Antisymmetry between s and t is now straightforward by induction on $h(s) + h(t)$. To show that $r \leq s$ and $s \leq t$ imply $r \leq t$, one uses induction on $h(r) + h(s) + h(t)$. In the most interesting case we have $r \leq s = \vartheta_i s'$ because of $r \leq k_i(s')$, while $s \leq t = \vartheta_i t'$ is due to $s' \leq t'$. The previous lemma ensures $k_i(s') \leq k_i(t')$. Using the induction hypothesis, we get $r \leq k_i(t')$ and then $r \leq \vartheta_i t' = t$. □

For $X = 1 = \{0\}$ we get $\bar{0} \leq_{S(1)} s$ by a straightforward induction over $s \in S(X)$. This means that our order $(S_n(1), \leq_{S(1)})$ coincides with the order (T_n, \leq) defined in [20, Section 2.3.3]. According to [20, Lemma 10], we thus have the following (see the cited reference for a precise definition of the gap condition):

Corollary 5.4 *If we identify $\vartheta_{i_1} \dots \vartheta_{i_n} 0 \in S_n(1)$ with the sequence (i_1, \dots, i_n) , then $\leq_{S(1)}$ coincides with the strong gap embeddability relation due to H. Friedman.*

Analogous to Sect. 3, we will show that S_{n+1}^0 is the Kruskal derivative of S_n^0 . Let us first turn S_n and S_n^0 into PO-dilators.

Definition 5.5 For each quasi embedding $f : X \rightarrow Y$ between partial orders, let $S(f) : S(X) \rightarrow S(Y)$ be defined by the clauses from Definition 3.8 (with S at the place of T). Let $\underline{k} : S(X) \rightarrow X$ be given as in the paragraph before Proposition 3.7. For a partial order X , we now define $\text{supp}_X^S : S(X) \rightarrow [X]^{<\omega}$ by $\text{supp}_X^S(s) = \{\underline{k}(s)\}$. We will also write supp_X^S for the restrictions of this function to $S_n(X)$ and to $S_n^0(X)$. Furthermore, we write $S_n(f) : S_n(X) \rightarrow S_n(Y)$ and $S_n^0(f) : S_n^0(X) \rightarrow S_n^0(Y)$ for the restrictions of $S(f)$ with the indicated (co-)domains.

As expected, we have the following:

Proposition 5.6 *Definition 5.5 extends S_n and S_n^0 into normal PO-dilators.*

Proof Most verifications are completely parallel to the proof of Proposition 3.9, but two additional observations are needed: First, the map $f \mapsto S(f)$ on morphisms preserves not only embeddings but also quasi embeddings, as required by clause (i) of [14, Definition 2.1]. Secondly, normality amounts to the implication

$$s \leq_{S(X)} t \implies \underline{k}(s) \leq_X \underline{k}(t),$$

which holds by Lemma 5.2 (since we have $\underline{k}(r) = x$ for $k_{-1}(r) = \bar{x}$). □

Let us now describe the extension into a Kruskal derivative:

Definition 5.7 For each number $n \in \mathbb{N}$ and any partial order X , let

$$\sigma_X^n : S_n \circ S_{n+1}^0(X) \rightarrow S_{n+1}(X)$$

be given by the clauses from Definition 3.6. Now define functions

$$\kappa_X^n : S_n^0 \circ S_{n+1}^0(X) \rightarrow S_{n+1}^0(X) \quad \text{and} \quad \iota_X^n : X \rightarrow S_{n+1}^0(X)$$

by setting $\kappa_X^n(s) = \vartheta_0 \sigma_X^n(s)$ and $\iota_X^n(x) = \bar{x}$.

Arguing as in the proof of Proposition 3.7, one can show the following:

Proposition 5.8 *The function $\sigma_X^n : S_n \circ S_{n+1}^0(X) \rightarrow S_{n+1}(X)$ is an order isomorphism, for each number $n \in \mathbb{N}$ and any partial order X .*

Proof To show that σ_X^n is surjective and order preserving, one argues as in the proof of Proposition 3.7. It remains to show that σ_X^n reflects the order, i. e., that

$$\sigma_X^n(s) \leq \sigma_X^n(t) \quad \Rightarrow \quad s \leq t$$

holds for $s, t \in S_n \circ S_{n+1}^0(X)$. We argue by induction on $h(s) + h(t)$ and discuss the two most interesting cases: First consider terms of the form $s = \bar{r}$ and $t = \vartheta_j t'$. In view of $r \in T_{n+1}^0(X)$ we have $r = \bar{x}$ or $\vartheta_0 r'$, which means that

$$\sigma_X^n(s) = r \leq \vartheta_{j+1} \sigma_X^n(t') = \sigma_X^n(t)$$

can only be due to $\sigma_X^n(s) \leq k_{j+1}(\sigma_X^n(t')) = \sigma_X^n(k_j(t'))$, where the equality comes from the proof of Proposition 3.7. By induction hypothesis we get $s \leq k_j(t')$ and then $s \leq \vartheta_j t' = t$, as desired. Let us also consider $s = \vartheta_i s'$ and $t = \bar{r}$. For terms of these forms, the inequality $s \leq t$ is always false. To show that $\sigma_X^n(s) \leq \sigma_X^n(t)$ is false as well, we observe

$$S(\sigma_X^n(s)) = S(\vartheta_{i+1} \sigma_X^n(s')) = i + 1 > 0 \geq S(r) = S(\sigma_X^n(t)),$$

where $S(r) \leq 0$ is due to $r \in T_n^0(X)$. On the other hand, a straightforward induction on $h(s_0) + h(t_0)$ shows that $s_0 \leq t_0$ implies $S(s_0) \leq S(t_0)$. □

In the following, ι^n and κ^n denote the families of functions $\iota_X^n : X \rightarrow S_{n+1}^0(X)$ and $\kappa_X^n : S_n^0 \circ S_{n+1}^0(X) \rightarrow S_{n+1}^0(X)$, respectively, indexed by the partial order X . We recall from [14, Section 4] that Kruskal derivatives are essentially unique. Hence the following means that the recursive construction from the introduction is uniquely realized by the normal PO-dilators S_n^0 that we have defined in the present section.

Theorem 5.9 *The Kruskal derivative of S_n^0 is given by $(S_{n+1}^0, \iota^n, \kappa^n)$.*

Proof Part (i) of [14, Definition 4.1] requires that $(S_{n+1}^0(X), \iota_X^n, \kappa_X^n)$ is an initial Kruskal fixed point of S_n^0 over X , for each partial order X . The most interesting conditions from the definition of Kruskal fixed point (see [14, Definition 3.1]) demand

$$\begin{aligned} \bar{x} \leq \kappa_X^n(t) &\Leftrightarrow \bar{x} \leq \underline{k}(t), \\ \kappa_X^n(s) \leq \kappa_X^n(t) &\Leftrightarrow s \leq t \text{ or } \kappa_X^n(s) \leq \underline{k}(t) \end{aligned}$$

for $x \in X$ and $s, t \in S_n^0 \circ S_{n+1}^0(X)$. In view of $\kappa_X^n(t) = \vartheta_0 \sigma_X^n(t)$, the left side of the first equivalence is indeed equivalent to

$$\bar{x} \leq k_0(\sigma_X^n(t)) = \sigma_X^n(k_{-1}(t)) = \underline{k}(t),$$

as in the proof of Proposition 3.7. Similarly, the left side of the second equivalence is equivalent to the disjunction of $\sigma_X^n(s) \leq \sigma_X^n(t)$ and $\kappa_X^n(s) \leq \underline{k}(t)$. This is equivalent to the right side, as σ_X^n is an order embedding. To show that our fixed point is initial we use the criterion from [14, Theorem 3.5], which demands that we have $S_{n+1}^0(X) = \text{rng}(l_X^n) \cup \text{rng}(\kappa_X^n)$ and $h(\underline{k}(s)) < h(\kappa_X^n(s))$ for some $h : S_{n+1}^0(X) \rightarrow \mathbb{N}$. These requirements can be verified as in the proof of Theorem 3.10. It remains to show that part (ii) of [14, Definition 4.1] is satisfied, i. e., that the functions l_X^n and κ_X^n are natural in X . This is readily reduced to the naturality of σ_X^n , which can be established by induction over terms, as in the proof of Theorem 3.10. \square

We can now use our general approach to re-derive two results that were previously shown by explicit computations. In view of Corollaries 3.5 and 5.4, we focus on the case of $X = 1$. The following is a special case of H. Friedman’s result on tree embeddings with the gap condition (see [11]). For the case of sequences, the result was analyzed by Schütte and Simpson [12] (see also [20, Section 2.2]).

Corollary 5.10 *The order $S_n^0(1)$ (sequences with gap condition, cf. Corollary 5.4) is a well partial order for each $n \in \mathbb{N}$.*

Proof We argue by induction on n to show that $X \mapsto S_n^0(X)$ preserves well partial orders. The base case is immediate in view of $S_n^0(X) \cong X$. The induction step is covered by [14, Corollary 4.5], given that S_{n+1}^0 is the Kruskal derivative of S_n^0 . \square

As in the previous section, a surjective quasi embedding $f : X \rightarrow Y$ between a linear order X and a partial order Y is called a linearization. The following result was first established by a concrete verification in [20, Lemma 11], which provides additional information: it shows that one can take v_1^n to be the identity on the underlying set $T_n^0(1) = S_n^0(1)$. This information could also be tracked through our general constructions, but we will not do so: the forte of our approach is precisely that fewer concrete computations are necessary.

Corollary 5.11 *For each $n \in \mathbb{N}$ we have a linearization $v_1^n : T_n^0(1) \rightarrow S_n^0(1)$ of the partial order $S_n^0(1)$ (sequences with gap condition) by the linear order $T_n^0(1)$ (collapsing functions, cf. Sect. 3).*

Proof As preparation, we use induction on n to show that S_n^0 is flat in the sense of Definition 4.7. The induction step is covered by Proposition 4.9. For the base case we recall $S_0^0(Y) = \{\bar{y} \mid y \in Y\}$. Given $f : X \rightarrow Y$ with

$$\text{supp}_Y^S(\bar{y}) = \{\underline{k}(\bar{y})\} = \{y\} \subseteq \text{rng}(f),$$

we write $y = f(x)$ to get $\bar{y} = \overline{f(x)} = S_0^0(f)(\bar{x}) \in \text{rng}(S_0^0(f))$. The point is that this works for any quasi embedding f , not just for embeddings. By recursion on n ,

we now construct linearizations $\nu^n : \mathbb{T}_n^0 \Rightarrow \mathbb{S}_n^0 \upharpoonright \text{LO}$ in the sense of Definition 4.5. For $n = 0$ we can define ν_X^0 as the identity map on $\mathbb{T}_0^0(X) = \{\bar{x} \mid x \in X\} = \mathbb{S}_0^0(X)$. In the step we define ν^{n+1} as the quasi embedding $(\nu^n)^+$ from Theorem 4.2. This is justified because \mathbb{T}_{n+1}^0 and \mathbb{S}_{n+1}^0 are the Bachmann–Howard and Kruskal derivative of \mathbb{T}_n^0 and \mathbb{S}_n^0 , respectively, by Theorems 3.10 and 5.9. From Theorem 4.8 we learn that ν^{n+1} is a linearization, given that the same holds for ν^n and that \mathbb{S}_n^0 is flat. \square

The linearization of $\mathbb{S}_n^0(1)$ by $\mathbb{T}_n^0(1)$ does not realize the maximal order type, as shown in [20] (see also the introduction of the present paper). We will now present a different construction of $\mathbb{S}_n^0(1)$ in terms of iterated Kruskal fixed points. Interestingly, the parallel construction in terms of Bachmann–Howard fixed points yields the order $OT_n[0]$ from [20, Section 5], which differs from $\mathbb{T}_n^0(1)$ and does realize the maximal order type of $\mathbb{S}_n^0(1)$. The idea is to consider the following transformations W_n , which are readily shown to be normal PO-dilators.

Definition 5.12 For each $n \in \mathbb{N}$ and any partial order X we put

$$W_n(X) = 1 + \mathbb{S}_n^0(1) \times X = \{0\} \cup \{(s, x) \mid s \in \mathbb{S}_n^0(1) \text{ and } x \in X\}.$$

To define a partial order on $W_n(X)$, we declare that 0 is incomparable to all other elements, and that we have

$$(s, x) \leq_{W_n(X)} (s', x') \iff s \leq_{\mathbb{S}_n^0(1)} s' \text{ and } x \leq_X x'.$$

Given a quasi embedding $f : X \rightarrow Y$, we define $W_n(f) : W_n(X) \rightarrow W_n(Y)$ by $W_n(f)(0) = 0$ and $W_n(f)(s, x) = (s, f(x))$. Also, let $\text{supp}_X : W_n(X) \rightarrow [X]^{<\omega}$ be given by $\text{supp}_X(0) = \emptyset$ and $\text{supp}_X(s, x) = \{x\}$.

We will show that $\mathbb{S}_{n+1}^0(1)$ is the initial Kruskal fixed point of W_n over $0 = \emptyset$. This gives rise to a recursive construction, since W_n does only depend on $\mathbb{S}_n^0(1)$, and since initial objects are unique up to isomorphism.

Definition 5.13 To construct a function $\pi_n : \mathbb{S}_n(1) \times \mathbb{S}_{n+1}^0(1) \rightarrow \mathbb{S}_{n+1}(1)$, we define $\pi_n(s, t)$ by recursion over the term s , setting

$$\pi_n(\bar{0}, t) = t \text{ and } \pi_n(\vartheta_i s, t) = \vartheta_{i+1} \pi_n(s, t).$$

We then define $\kappa_n : W_n(\mathbb{S}_{n+1}^0(1)) \rightarrow \mathbb{S}_{n+1}^0(1)$ by $\kappa_n(0) = \bar{0}$ and $\kappa_n(s, t) = \vartheta_0 \pi_n(s, t)$. Let us also agree to write $\iota_n : 0 \rightarrow \mathbb{S}_{n+1}^0(1)$ for the empty function.

To see that the definition of π_n is justified, one should observe that we have

$$S(\pi_n(s, t)) = \begin{cases} S(t) \leq 0 & \text{if } s = 0, \\ S(s) + 1 & \text{otherwise.} \end{cases}$$

This also shows that $s \in \mathbb{S}_n^0(1)$ entails $S(\pi_n(s, t)) \leq 1$, as needed to justify the definition of κ_n . As promised, we have the following:

Theorem 5.14 *The tuple $(S_{n+1}^0(1), \iota_n, \kappa_n)$ is the initial Kruskal fixed point of W_n over 0 , for each number $n \in \mathbb{N}$.*

Proof Let us abbreviate $Z = S_{n+1}^0(1)$. To show that we have a Kruskal fixed point, we need to prove that the equivalence

$$\kappa_n(\sigma) \leq_Z \kappa_n(\tau) \iff \sigma \leq_{W_n(Z)} \tau \text{ or } \kappa_n(\sigma) \leq_Z t \text{ for some } t \in \text{supp}_Z(\tau)$$

is satisfied for all $\sigma, \tau \in W_n(Z)$ (cf. [14, Definition 3.1]). For $\sigma = 0$, both sides of the equivalence hold, since we have $\bar{0} \leq_Z t$ for any $t \in Z$ (see the paragraph before Corollary 5.4). For $\sigma \neq 0$ and $\tau = 0$, both sides of the equivalence fail. Now consider $\sigma = (s, t)$ and $\tau = (s', t')$. Then our equivalence amounts to

$$\vartheta_0 \pi_n(s, t) \leq_Z \vartheta_0 \pi_n(s', t') \iff (s \leq_{S(1)} s' \text{ and } t \leq_Z t') \text{ or } \vartheta_0 \pi_n(s, t) \leq_Z t'.$$

Given that $t' \in S_{n+1}^0(1)$ entails $k_{i+1}(t') = t'$, an auxiliary induction over s' yields

$$k_{i+1}(\pi_n(s', t')) = \pi_n(k_i(s'), t').$$

For $i = -1$ we get $k_{-1}(s') = \bar{0}$ and hence $k_0(\pi_n(s', t')) = \pi_n(\bar{0}, t') = t'$. In view of Definition 5.1, this entails that the last equivalence reduces to

$$\pi_n(s, t) \leq_{S(1)} \pi_n(s', t') \iff s \leq_{S(1)} s' \text{ and } t \leq_Z t'.$$

The latter can be shown by induction over $h(s) + h(s')$, where we admit s, s' from the bigger set $S_n(1) \supset S_n^0(1)$ to make the induction go through. For $s = \bar{0} = s'$, both sides of our equivalence amount to $t \leq_Z t'$. Now consider $s = \bar{0}$ and $s' = \vartheta_j r'$. In view of $S(\pi_n(s, t)) = S(t) \leq 0 < j + 1$ and $\pi_n(s', t') = \vartheta_{j+1} \pi_n(r', t')$, the left side of the desired equivalence is equivalent to

$$\pi_n(s, t) \leq_{S(1)} k_{j+1}(\pi_n(r', t')) = \pi_n(k_j(r'), t').$$

Inductively, this is equivalent to the conjunction of $s \leq_{S(1)} k_j(r')$ and $t \leq_Z t'$. We can conclude since both $s \leq_{S(1)} s'$ and $s \leq_{S(1)} k_j(r')$ are automatic for $s = \bar{0}$. If we have $s = \vartheta_i r$ and $s' = \bar{0}$, then both sides of the desired equivalence fail. To see this, note that we have

$$S(\pi_n(s, t)) = S(\vartheta_{i+1} \pi_n(r, t)) = i + 1 > 0 \geq S(t') = S(\pi_n(s', t'))$$

and $S(s) = i > -1 = S(s')$. At the same time, a straightforward induction shows that $s_0 \leq_{S(1)} s_1$ entails $S(s_0) \leq S(s_1)$. Finally, consider $s = \vartheta_i r$ and $s' = \vartheta_j r'$. The left side of the desired equivalence can hold for two reasons: First, assume that we have $\pi_n(s, t) \leq_{S(1)} k_{j+1}(\pi_n(r', t'))$. As before, we get $s \leq_{S(1)} k_j(r')$ and $t \leq_Z t'$. The former entails $s \leq_{S(1)} s'$, as needed for the right side. Now assume that the left side holds because we have $i = j$ and $\pi_n(r, t) \leq_{S(1)} \pi_n(r', t')$. Inductively we get $r \leq_{S(1)} r'$ and $t \leq_Z t'$, and the former entails $s \leq_{S(1)} s'$. By reading the given argument

backwards, one obtains the implication from right to left. To show that our Kruskal fixed point is initial, we apply the criterion from [14, Theorem 3.5]. As preparation, one shows that any $r \in S_{n+1}(1)$ lies in the range of π_n , using induction over r . To conclude that κ_n is surjective, it suffices to observe that $\vartheta_0 r \in S_{n+1}^0(1)$ requires $S(r) \leq 1$, so that $r = \pi_n(s, t)$ forces $S(s) \leq 0$ and hence

$$(s, t) \in S_n^0(1) \times S_{n+1}^0(1) \subseteq W_n(S_{n+1}^0(1)).$$

Finally, we need to verify that

$$r \in \text{supp}_Z(\sigma) \Rightarrow h(r) < h(\kappa_n(\sigma))$$

holds for all $\sigma \in W_n(Z) = W_n(S_{n+1}^0(1))$. For $\sigma = 0$ we have $\text{supp}_Z(\sigma) = \emptyset$, so that the condition is void. For $\sigma = (s, t)$ we have $\text{supp}_Z(\sigma) = \{t\}$, which means that the claim amounts to $h(t) < h(\kappa_n(s, t))$. This reduces to $h(t) \leq h(\pi_n(s, t))$, which is readily established by induction over s . □

We have just seen a reconstruction of $S_n^0(1)$ via iterated Kruskal fixed points. In the following, we show that the parallel construction for Bachmann–Howard fixed points yields the orders $\text{OT}_n[0]$ from [20, Section 5]. Let us first recall the latter.

Definition 5.15 For each $n \in \mathbb{N}$ we construct a set OT_n of terms. Simultaneously, each term $s \in \text{OT}_n$ is associated with a number $S(s) \in \{-1, \dots, n - 1\}$ and finite subsets $K_i(s) \subseteq \text{OT}_n$ for $i \geq -1$, according to the following clauses:

- (i) We have a term $\bar{0} \in \text{OT}_n$ with $S(\bar{0}) = -1$ and $K_i(\bar{0}) = \emptyset$.
- (ii) Given $s, t \in \text{OT}_n$ and $i < n$ with $i \geq \max\{S(s) - 1, S(t), 0\}$ and $K_i(s) = \emptyset$, we add a term $\theta_i st \in \text{OT}_n$ with $S(\theta_i st) = i$ and

$$K_j(\theta_i st) = \begin{cases} \{\theta_i st\} & \text{if } i \leq j, \\ K_j(s) \cup K_j(t) & \text{otherwise.} \end{cases}$$

To define a binary relation $<_{\text{OT}}$ on OT_n , we declare that $r <_{\text{OT}} r'$ holds if, and only if, one of the following clauses applies:

- (i') We have $r = \bar{0}$ and $r' \neq \bar{0}$.
- (ii') We have $r = \theta_i st$ and $r' = \theta_j s't'$ with $i < j$.
- (iii') We have $r = \theta_i st$ and $r' = \theta_i s't'$, and one of the following holds:
 - We have $s <_{\text{OT}} s'$ and $t <_{\text{OT}} t' = \theta_i s't'$.
 - We have $s = s'$ and $t <_{\text{OT}} t'$.
 - We have $\theta_i st = r \leq_{\text{OT}} t'$ (i. e. $r <_{\text{OT}} t'$ or $r = t'$ as terms).

Finally, we set $\text{OT}_n^0 = \{s \in \text{OT}_n \mid S(s) \leq 0\}$. We will also write $<_{\text{OT}}$ for the restriction of the given relation to this set.

The reader may have noticed that our definition of $<_{\text{OT}}$ looks somewhat different from the one in [20, Definition 37]. This has the following context: First, the cited

reference also defines $<_{OT}$ on a larger set that contains terms $\theta_i st$ with $K_i(s) \neq \emptyset$. This leads to additional conditions that are void in our case. Secondly, we have declared that $\theta_i st \leq_{OT} t'$ entails $\theta_i st <_{OT} \theta_i s't'$. In [20], this implication has $s' <_{OT} s$ as an additional assumption. At the same time, [20, Lemma 23] shows that $t' < \theta_i s't'$ does always hold. As it is transitive, the order from [20] will thus satisfy our stronger implication, so that the two definitions coincide. On a different note, [20, Definition 40] declares that we have

$$OT_n[0] \stackrel{?}{=} \{s \in OT_n \mid s <_{OT} \theta_0 \overline{00}\}.$$

It is easy to see that this yields $OT_n[0] = \{\overline{0}\}$, which contradicts [20, Corollary 6]. Presumably, this is a typo that should be corrected into

$$OT_n[0] = \{s \in OT_n \mid s <_{OT} \theta_1 \overline{00}\}.$$

Strictly speaking, this is only explained when we have $n > 1$, so that $\theta_1 \overline{00}$ is available. For the corrected definition one can show $OT_n[0] = OT_n^0$ by a straightforward induction over terms. The following was left implicit in [20]:

Lemma 5.16 *The relation $<_{OT}$ is a linear order on OT_n , for each $n \in \mathbb{N}$.*

Proof Define $h : OT_n \rightarrow \mathbb{N}$ recursively by $h(\overline{0}) = 0$ and $h(\theta_i st) = h(s) + h(t) + 1$. A straightforward induction over $h(r) + h(s) + h(t)$ shows that the conjunction of $r <_{OT} s$ and $s <_{OT} t$ implies $r <_{OT} t$. We can now establish $r \not<_{OT} r$ by induction over $h(r)$. This is readily reduced to $r = \theta_i st \not<_{OT} t$. In view of $\theta_i st \in OT_n$ we must have $S(t) \leq i$, which leaves a term of the form $t = \theta_i s't'$ as the only interesting case (with the same i as in $r = \theta_i st$). We trivially have $\theta_i s't' \leq_{OT} t$, which provides the inequality $t = \theta_i s't' <_{OT} \theta_i st = r$. Now if $r \leq_{OT} t$ was true, then transitivity would yield $t <_{OT} t$, which contradicts the induction hypothesis. Finally, trichotomy between s and t follows by a straightforward induction over $h(s) + h(t)$. \square

Parallel to the above, our aim is to show that OT_{n+1}^0 is the initial Bachmann–Howard fixed point of the following LO-dilators D_n over the empty order 0.

Definition 5.17 For each $n \in \mathbb{N}$ and any linear order X we put

$$D_n(X) = 1 + OT_n^0 \times X = \{0\} \cup \{(s, x) \mid s \in OT_n^0 \text{ and } x \in X\}.$$

To define a linear order on $D_n(X)$, we declare that 0 is the smallest element, and that we have

$$(s, x) <_{D_n(X)} (s', x') \iff s <_{OT} s' \text{ or } (s = s' \text{ and } x <_X x').$$

For any embedding $f : X \rightarrow Y$, we define a function $D_n(f) : D_n(X) \rightarrow D_n(Y)$ by $D_n(f)(0) = 0$ and $D_n(f)(s, x) = (s, f(x))$. Also, let $\text{supp}_X : D_n(X) \rightarrow [X]^{<\omega}$ be given by $\text{supp}_X(0) = \emptyset$ and $\text{supp}_X(s, x) = \{x\}$.

Let us construct the functions that Definition 2.5 requires:

Definition 5.18 Let the map $\text{OT}_n \ni s \mapsto s^+ \in \text{OT}_{n+1}$ be given by the recursive clauses $\bar{0}^+ = \bar{0}$ and $(\theta_i s t)^+ = \theta_{i+1} s^+ t^+$. Now define $\vartheta_n : D_n(\text{OT}_{n+1}^0) \rightarrow \text{OT}_{n+1}^0$ by

$$\vartheta_n(0) = \bar{0} \quad \text{and} \quad \vartheta_n(s, t) = \theta_0 s^+ t.$$

Let us also agree to write $\iota_n : 0 \rightarrow \text{OT}_{n+1}^0$ for the empty function.

To justify the definition of $s^+ \in \text{OT}_{n+1}$, one should simultaneously check

$$K_{i+1}(s^+) = \{t^+ \mid t \in K_i(s)\} \quad \text{and} \quad S(s^+) = \begin{cases} S(s) - 1 & \text{if } s = \bar{0}, \\ S(s) + 1 & \text{otherwise.} \end{cases}$$

In particular we get $K_0(s^+) = \emptyset$, as an easy induction over s yields $K_{-1}(s) = \emptyset$. For $s \in \text{OT}_n^0$ we also obtain $S(s^+) \leq 1$, so that we indeed have $\vartheta_n(s, t) = \theta_0 s^+ t \in \text{OT}_{n+1}^0$ for $(s, t) \in D_n(\text{OT}_{n+1}^0)$. We can now establish the promised result:

Theorem 5.19 *The tuple $(\text{OT}_{n+1}^0, \iota_n, \vartheta_n)$ is an initial Bachmann–Howard fixed point of D_n over 0 , for each number $n \in \mathbb{N}$.*

Proof As preparation one checks that $s <_{\text{OT}} t$ implies $s^+ <_{\text{OT}} t^+$, by a straightforward induction over $h(s) + h(t)$ (cf. [20, Lemma 24]). Abbreviating $Z = \text{OT}_{n+1}^0$, we now verify the two conditions from clause (iii) of Definition 2.5. First, we need to show that $z \in \text{supp}_Z(\sigma)$ implies $z <_Z \vartheta_n(\sigma)$, for any $\sigma \in D_n(Z)$. In the case of $\sigma = 0$ we have $\text{supp}_Z(\sigma) = \emptyset$, which means that the condition is void. For $\sigma = (s, t)$ we have $\text{supp}_Z(s, t) = \{t\}$, so that the claim amounts to

$$t <_Z \vartheta_n(s, t) = \theta_0 s^+ t.$$

In the proof of Lemma 5.16 we have shown $\theta_0 s^+ t \not<_{\text{OT}} t$, which suffices due to trichotomy (in view of $S(t) \leq 0$ it is easy to give a direct argument as well). The second condition from clause (iii) of Definition 2.5 does essentially amount to

$$(s_0, t_0) <_{D_n(Z)} (s_1, t_1) \text{ and } t_0 <_Z \theta_0 s_1^+ t_1 \implies \theta_0 s_0^+ t_0 <_Z \theta_0 s_1^+ t_1.$$

Assuming that the left side holds, we have $s_0^+ = s_1^+$ and $t_0 <_{\text{OT}} t_1$, or $s_0^+ <_{\text{OT}} s_1^+$ and $t_0 <_Z \theta_0 s_1^+ t_1$, using the fact that we have shown as preparation. In either case, the right side follows by clause (iii') of Definition 5.15. To show that our fixed point is initial, we use the criterion from Theorem 2.9. Let us first observe

$$h(t) < h(s^+) + h(t) + 1 = h(\theta_0 s^+ t) = h(\vartheta_n(s, t)),$$

for $h : \text{OT}_{n+1}^0 \rightarrow \mathbb{N}$ as in the proof of Lemma 5.16. It remains to show that the function $\vartheta_n : D_n(\text{OT}_{n+1}^0) \rightarrow \text{OT}_{n+1}^0$ is surjective. An easy induction shows that any element $s \in \text{OT}_{n+1}$ with $K_0(s) = \emptyset$ can be written as $s = s_0^+$ for some $s_0 \in \text{OT}_n$. For $S(s) \leq 1$ we get $s_0 \in \text{OT}_n^0$, and it is straightforward to conclude. \square

Finally, we use our general approach to re-derive two known results. The following was first shown as a consequence of [20, Lemma 25].

Corollary 5.20 *For each $n \in \mathbb{N}$ we have a linearization $f_n : OT_n^0 \rightarrow S_n^0(1)$ of the partial order $S_n^0(1)$ (sequences with gap condition) by the linear order OT_n^0 (binary collapsing functions, cf. [20, Section 5]).*

Proof We construct f_n by recursion over $n \in \mathbb{N}$. In the base case, let f_0 be the identity on $OT_0^0 = \{0\} = S_0^0(1)$. Given f_n , we get a linearization $\nu : D_n \Rightarrow W_n$ by setting $\nu_X(0) = 0$ and $\nu_X(s, x) = (f_n(s), x)$. It is easy to see that W_n is flat (cf. Definition 4.7). Theorems 4.2 and 4.8 yield a linearization $\nu^+ : \vartheta D_n \Rightarrow TW_n$. From Theorems 5.14 and 5.19 we know that $TW_n(0)$ and $\vartheta D_n(0)$ coincide with $S_{n+1}^0(1)$ and OT_{n+1}^0 , respectively (up to isomorphism). Hence setting

$$f_{n+1} = \nu_0^+ : OT_{n+1}^0 = \vartheta D_n(0) \rightarrow TW_n(0) = S_{n+1}^0(1)$$

completes the recursive construction. □

The following result was first shown in [20, Corollaries 6 and 7]. The proof in the cited paper goes via an addition-free variant of the Veblen functions. In the author’s opinion, our general approach leads to a more transparent argument:

Corollary 5.21 *For each $n \geq 1$, the linear order OT_n^0 has order type ω_{2n-1} , which means that it realizes the maximal order type of the partial order $S_n^0(1)$.*

Proof Recall that we have $\omega_0^\alpha = \alpha$ and $\omega_{n+1}^\alpha = \omega_n^{\omega^\alpha}$. We will show

$$OT_n^0 \cong 1 + \omega_{2n}^0$$

by induction over $n \in \mathbb{N}$. This entails the first claim of the corollary, as $n \geq 1$ yields

$$1 + \omega_{2n}^0 = \omega_{2n}^0 = \omega_{2n-1}^1 = \omega_{2n-1}.$$

In the base case of our induction, it suffices to recall $OT_0^0 = \{\bar{0}\}$ and $\omega_0^0 = 0$. The induction step relies on [24, Theorem 2.2]. To state the latter, consider an arbitrary ordinal α . Analogous to Definition 5.17, we get an LO-dilator D^α with

$$D^\alpha(X) = 1 + (1 + \alpha) \times X.$$

According to the cited theorem, we have $\vartheta D^\alpha(0) \cong \omega_2^\alpha$. From Theorem 5.19 we know $OT_{n+1}^0 \cong \vartheta D_n(0)$ with $D_n(X) = 1 + OT_n^0 \times X$. Now the induction hypothesis provides a natural isomorphism $D_n \cong D^\alpha$ for $\alpha = \omega_{2n}^0$. In view of Lemma 2.14 (and the uniqueness of initial fixed points), we obtain

$$OT_{n+1}^0 \cong \vartheta D_n(0) \cong \vartheta D^{\omega_{2n}^0}(0) \cong \omega_2^{\omega_{2n}^0} = \omega_{2(n+1)}^0 = 1 + \omega_{2(n+1)}^0,$$

as needed to complete the induction step. Finally, let us give references for the fact that $S_n^0(1)$ has maximal order type ω_{2n-1} , for $n \geq 1$. In the paragraph after Lemma 5.3 we have observed that our order $S_n^0(1)$ coincides with the order $(T_n[0], \preceq)$ from [20, Section 2.2.3]. The cited reference shows that this order is isomorphic to a certain collection $\bar{S}_n[0]$ of sequences with the strong gap condition. According to [20, Lemma 3 and Corollary 2] (based on work by Schütte and Simpson [12]), that collection has maximal order type ω_{2n-1} , for $n \geq 1$. \square

Funding Open Access funding enabled and organized by Projekt DEAL. The work of Anton Freund has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project number 460597863.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Higman, G.: Ordering by divisibility in abstract algebras. *Proc. Lond. Math. Soc.* **3**(2), 326–336 (1952)
2. Kruskal, J.: Well-quasi-ordering, the tree theorem, and Vazsonyi’s conjecture. *Trans. Am. Math. Soc.* **95**(2), 210–225 (1960)
3. Nash-Williams, C.S.J.A.: On well-quasi-ordering finite trees. *Proc. Camb. Philos. Soc.* **59**, 833–835 (1963)
4. Freund, A., Rathjen, M., Weiermann, A.: Minimal bad sequences are necessary for a uniform Kruskal theorem. *Adv. Math.* **400**, article no. 108265, 44 pages (2022)
5. Girard, J.-Y.: Π_1^1 -logic, part 1: dilators. *Ann. Pure Appl. Logic* **21**, 75–219 (1981)
6. Simpson, S.G.: Subsystems of second order arithmetic. In: *Perspectives in Logic*. Cambridge University Press (2009)
7. Marcone, A.: On the logical strength of Nash–Williams’ theorem on transfinite sequences. In: Hodges, W., Hyland, M., Steinhorn, C., Truss, J. (eds.) *Logic: From Foundations to Applications* (Oxford). Oxford University Press, pp. 327–351 (1996)
8. Hasegawa, R.: Well-ordering of algebras and Kruskal’s theorem. In: Jones, N.D., Hagiya, M., Sato, M. (eds.) *Logic, Language and Computation, Lecture Notes in Computer Science*, vol. 792 (1994)
9. Hasegawa, R.: An analysis of divisibility orderings and recursive path orderings. In: Shyamasundar, R.K., Ueda, K. (eds.) *Advances in Computing Science—ASIAN’97, Lecture Notes in Computer Science*, vol. 1345 (1997)
10. Weiermann, A.: A computation of the maximal order type of the term ordering on finite multisets. In: Ambos-Spies, K., Löwe, B., Merkle, W. (eds.) *Mathematical Theory and Computational Practice. CiE 2009, Lecture Notes in Computer Science*, vol. 5635 (2009)
11. Simpson, S.G.: Nonprovability of certain combinatorial properties of finite trees. In: Harrington, L.A., Morley, M.D., Sčědřov, A., Simpson, S.G. (eds.) *Harvey Friedman’s Research on the Foundations of Mathematics, Studies in Logic and the Foundations of Mathematics*, vol. 117, pp. 87–117. North-Holland (1985)
12. Schütte, K., Simpson, S.G.: Ein in der reinen Zahlentheorie unbeweisbarer Satz über endliche Folgen von natürlichen Zahlen. *Archiv für Mathematische Logik und Grundlagenforschung* **25**, 75–89 (1985)
13. van der Meeren, J.: Connecting the two worlds: well-partial-orders and ordinal notation systems, PhD thesis, Ghent University (2015)
14. Freund, A.: From Kruskal’s theorem to Friedman’s gap condition. *Math. Struct. Comput. Sci.* **8**(30), 952–975 (2020)

15. Rathjen, M., Weiermann, A.: Proof-theoretic investigations on Kruskal's theorem. *Ann. Pure Appl. Logic* **60**, 49–88 (1993)
16. Freund, A.: Type-Two Well-Ordering Principles, Admissible Sets, and Π_1^1 -Comprehension, PhD thesis, University of Leeds. <http://etheses.whiterose.ac.uk/20929/> (2018)
17. Freund, A.: Π_1^1 -comprehension as a well-ordering principle. *Adv. Math.* **355**, article no. 106767, 65 pages (2019)
18. Freund, A.: A categorical construction of Bachmann–Howard fixed points. *Bull. Lond. Math. Soc.* **51**(5), 801–814 (2019)
19. Freund, A.: Computable aspects of the Bachmann–Howard principle. *J. Math. Logic* **20**(2), article no. 2050006, 26 pages (2020)
20. van der Meeren, J., Rathjen, M., Weiermann, A.: Ordinal notation systems corresponding to Friedman's linearized well-partial-orders with gap-condition. *Arch. Math. Logic* **56**, 607–638 (2017)
21. Weiermann, A., Wilken, G.: Ordinal arithmetic with simultaneously defined theta-functions. *Math. Logic Q.* **57**(2), 116–132 (2011)
22. van der Meeren, J., Rathjen, M., Weiermann, A.: Well-partial-orderings and the big Veblen number. *Arch. Math. Logic* **54**(1–2), 193–230 (2015)
23. van der Meeren, J., Rathjen, M., Weiermann, A.: An order-theoretic characterization of the Howard–Bachmann–hierarchy. *Arch. Math. Logic* **56**(1–2), 79–118 (2017)
24. Freund, A.: Predicative collapsing principles. *J. Symb. Logic* **85**(1), 511–530 (2020)
25. de Jongh, D., Parikh, R.: Well-partial orderings and hierarchies. *Indagationes Mathematicae* **80**(3), 195–207 (1977)
26. Aschenbrenner, M., Pong, W.Y.: Orderings of monomial ideals. *Fundamenta Mathematicae* **181**, 27–74 (2004)
27. Knight, J.F., Lange, K.: Lengths of developments in $K((G))$. *Selecta Mathematica* **25**, article no. 14, 36 pages (2019)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.