

A categorical construction of Bachmann–Howard fixed points

Anton Freund

ABSTRACT

Peter Aczel has given a categorical construction for fixed points of normal functors, that is, dilators which preserve initial segments. For a general dilator $X \mapsto T_X$, we cannot expect to obtain a well-founded fixed point, as the order type of T_X may always exceed the order type of X . In the present paper, we show how to construct a Bachmann–Howard fixed point of T , that is, an order $\text{BH}(T)$ with an ‘almost’ order preserving collapse $\vartheta : T_{\text{BH}(T)} \rightarrow \text{BH}(T)$. Building on previous work, we show that Π_1^1 -comprehension is equivalent to the assertion that $\text{BH}(T)$ is well-founded for any dilator T .

1. Introduction: set existence and well-foundedness

The present paper contributes to a research program known as reverse mathematics (see [19]), which aims to answer the following question: Which infinite sets are indispensable for the proof of a given mathematical theorem?

To make the previous question precise, one can consider ω -models, which are defined as collections $\mathcal{M} \subseteq \mathcal{P}(\mathbb{N})$ of subsets of the natural numbers. For such a model, we can ask whether it satisfies certain statements about finite objects (coded by natural numbers) and infinite collections of such objects (represented by subsets of \mathbb{N}). For example, the formula

$$\exists_{X \subseteq \mathbb{N}} \forall_{n \in \mathbb{N}} (n \in X \leftrightarrow \exists_{m \in \mathbb{N}} m + m = n)$$

holds in an ω -model \mathcal{M} if, and only if, \mathcal{M} contains the set of even numbers. More generally, we evaluate a formula in an ω -model \mathcal{M} by replacing all quantifiers $\exists_{X \subseteq \mathbb{N}}$ and $\forall_{X \subseteq \mathbb{N}}$ with $\exists_{X \in \mathcal{M}}$ and $\forall_{X \in \mathcal{M}}$, respectively (the quantifiers $\exists_{n \in \mathbb{N}}$ and $\forall_{n \in \mathbb{N}}$ remain unchanged).

An ω -model satisfies recursive comprehension if it contains any set that is computable relative to some of its members. Classical results of reverse mathematics show that an ω -model of recursive comprehension does always satisfy the intermediate value theorem, but not necessarily the Bolzano–Weierstraß theorem (see [19, Theorems II.6.6, III.2.2]). In the following, we assume that all ω -models satisfy recursive comprehension.

It turns out that the Bolzano–Weierstraß theorem corresponds to a stronger comprehension principle: Roughly speaking, a formula is called arithmetical if it does not involve quantification over subsets of \mathbb{N} (for example, the above formula is not arithmetical, due to the quantifier $\exists_{X \subseteq \mathbb{N}}$, but the subformula $\exists_{m \in \mathbb{N}} m + m = n$ is). An ω -model \mathcal{M} satisfies arithmetical comprehension if it contains any set that can be defined by an arithmetical formula with parameters from \mathcal{M} . Note that arithmetical comprehension implies recursive comprehension. The classical result that we have cited above does actually show the following stronger assertion: An ω -model satisfies the Bolzano–Weierstraß theorem if, and only if, it

Received 19 September 2019; revised 7 June 2019; published online 30 August 2019.

2010 *Mathematics Subject Classification* 03B30, 03D60, 03F15 (primary).

Correction added on November 12, 2020, after first online publication: Projekt Deal funding statement has been added.

© 2019 The Authors. *Bulletin of the London Mathematical Society* is copyright © London Mathematical Society. This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

satisfies arithmetical comprehension (in fact, this holds for a more general class of models, but ω -models are particularly intuitive). Hence, arithmetically defined sets are indispensable for this theorem.

Our goal is to relate set existence to the notion of well-foundedness, which plays an important role throughout mathematics. The idea is to consider transformations between well-orders, which are known as (type-one) well-ordering principles. This approach was first realized by Girard [10], who considered the transformation of an order $X = (X, \leq_X)$ into the set

$$\omega^X = \{\langle x_0, \dots, x_{n-1} \rangle \mid x_{n-1} \leq_X \dots \leq_X x_0\}$$

of finite descending sequences with entries from X , ordered lexicographically (think of Cantor normal forms). Girard has shown that an ω -model satisfies arithmetical comprehension if, and only if, it satisfies the statement that ‘ ω^X is well founded for any well-order X ’. Let us point out that the order ω^X from the previous paragraph is computable relative to X , so that $X \in \mathcal{M}$ implies $\omega^X \in \mathcal{M}$. In other words, the strength of the well-ordering principle lies in the preservation of well-foundedness, not in the existence of the set ω^X .

Many set existence principles have been characterized in terms of type-one well-ordering principles (see [7] for a list of references). However, such a characterization cannot be given for one particularly important set existence principle: An ω -model \mathcal{M} satisfies Π_1^1 -comprehension if it contains any set of the form

$$\{n \in \mathbb{N} \mid \forall X \in \mathcal{M} \varphi(n, X)\}, \tag{1.1}$$

where φ is an arithmetical formula with parameters from \mathcal{M} . The point is that any type-one well-ordering principle can be expressed by a formula $\forall X \subseteq \mathbb{N} \exists Y \subseteq \mathbb{N} \varphi$, where φ is arithmetical. It is known that Π_1^1 -comprehension cannot be equivalent to a formula of this form.

Why is Π_1^1 -comprehension interesting? Firstly, Π_1^1 -comprehension is equivalent to important mathematical theorems, such as the Cantor–Bendixson theorem or the result that any countable Abelian group can be constructed as the direct sum of a divisible and a reduced group (see [19, Theorems VI.1.6, VI.4.1]). Secondly, there is a fundamental difference between Π_1^1 -comprehension and weaker principles: An ω -model

$$\mathcal{M} = \{\mathcal{M}_0, \mathcal{M}_1, \mathcal{M}_2, \dots\}$$

of arithmetical comprehension can be built ‘from below’, by setting $\mathcal{M}_{k+1} = \{n \in \mathbb{N} \mid \varphi_k(n)\}$ for some systematic enumeration $\varphi_0, \varphi_1, \dots$ of the arithmetical formulas, where φ_k may contain the parameters $\mathcal{M}_0, \dots, \mathcal{M}_k$. An ω -model of Π_1^1 -comprehension cannot be built in the same way, since the quantifier $\forall X \in \mathcal{M}$ in (1.1) changes its range whenever we extend \mathcal{M} by a set \mathcal{M}_{k+1} , so that the construction of $\mathcal{M}_0, \dots, \mathcal{M}_k$ may be invalidated at this point. One can describe this difference by saying that arithmetical comprehension is predicative, while Π_1^1 -comprehension is impredicative (see [4] for a deeper discussion of these notions).

Rathjen [15, 16] and Montalbán [12, 13] have conjectured that Π_1^1 -comprehension can be characterized by a type-two well-ordering principle. Such a principle should take a type-one well-ordering principle $X \mapsto T_X$ as input and yield a well-order as output. In [7] (which is based on the author’s PhD thesis [6] and an earlier arXiv preprint [5]), we have made a significant step towards this conjecture. The basic idea was that the output should be a certain type of fixed-point of the input $X \mapsto T_X$.

What is the correct notion of fixed point? Let us first observe that we cannot expect to obtain well-founded fixed points in the obvious sense: Consider the transformation

$$X \mapsto T_X = X \cup \{\top\}$$

that extends a given order X by a new maximal element \top . Let X be an arbitrary well-order. Then there is a unique ordinal α with $X \cong \alpha$. Clearly, we have $T_X \cong \alpha + 1$. By the uniqueness of α , we can conclude

$$T_X \not\cong X.$$

In [7], we have formulated precise conditions under which a function

$$\vartheta : T_X \rightarrow X$$

can be considered as ‘almost’ order preserving. If such a function exists, then X is called a Bachmann–Howard fixed point of the transformation T . In fact, these notions are only defined if T is a particularly uniform transformation of orders, namely a dilator in the sense of Girard [9]. Full details of all relevant definitions will be recalled in the following section. Relying on these notions, we can now state the main result of [7]:

THEOREM 1.1. *The following are equivalent over the theory $\mathbf{ATR}_0^{\text{set}}$:*

- (i) *The principle of Π_1^1 -comprehension.*
- (ii) *The statement that every dilator has a well-founded Bachmann–Howard fixed point.*

Here $\mathbf{ATR}_0^{\text{set}}$ is Simpson’s [18] set-theoretic version of arithmetical transfinite recursion (see the following section for details). The point is that $\mathbf{ATR}_0^{\text{set}}$ is a predicative theory (in the broad sense). In this crucial respect, the theory in the ‘background’ is simpler than the principle of Π_1^1 -comprehension, which is characterized by the theorem.

So what is new in the present paper? We will present an explicit construction, which transforms a given dilator T into a Bachmann–Howard fixed point $\text{BH}(T)$ of T (see Theorem 4.2). More precisely, the order $\text{BH}(T)$ can be constructed by a primitive recursive set function, as explained at the beginning of the following section. We will also show that $\text{BH}(T)$ is minimal, in the sense that it can be embedded into any other Bachmann–Howard fixed point of T (see Theorem 4.4). It follows that $\text{BH}(T)$ is well founded if, and only if, the dilator T has some well-founded Bachmann–Howard fixed point. This improves Theorem 1.1 as follows:

THEOREM 1.2. *The following are equivalent over the theory $\mathbf{ATR}_0^{\text{set}}$.*

- (i) *The principle of Π_1^1 -comprehension.*
- (ii) *The statement that the order $\text{BH}(T)$ is well founded for every dilator T .*

The crucial point is that $\text{BH}(T)$ is constructed ‘from below’, that is, in a predicative way. Hence, the impredicative principle of Π_1^1 -comprehension is reduced to a predicative construction and a statement about the preservation of well-foundedness, over a predicative background theory. Statement (ii) from Theorem 1.1 will be called the abstract Bachmann–Howard principle, since it asserts the existence of a fixed point without constructing one. Statement (ii) of Theorem 1.2 will be called the predicative Bachmann–Howard principle.

To conclude this introduction, we explain how to construct the Bachmann–Howard fixed point $\text{BH}(T)$. Our starting point is a construction due to Aczel [1, 2]: Given an endofunctor T on the category of linear orders, let X be the direct limit of the diagram

$$X_0 := \emptyset \xrightarrow{\iota_0} X_1 := T_{X_0} \xrightarrow{\iota_1 := T_{\iota_0}} X_2 := T_{X_1} \longrightarrow \dots$$

If T preserves direct limits, then we have $T_X \cong X$. Aczel has shown that X is well founded for an important class of functors, which correspond to normal functions on the ordinals (the crucial condition requires that the range of $T_f : T_X \rightarrow T_Y$ is an initial segment of the order T_Y whenever the range of $f : X \rightarrow Y$ is an initial segment of Y).

The transformation $X \mapsto T_X = X \cup \{\top\}$ that we have considered above is readily extended into a functor (set $T_f(\top) = \top$). We have seen that $T_X \cong X$ cannot hold for any well-order X . Indeed, the fixed point that results from Aczel’s construction is isomorphic to the negative integers. In order to obtain well-founded Bachmann–Howard fixed points, we will modify Aczel’s construction as follows: Given a linear order X , one can define an order $\vartheta_T(X)$ with an ‘almost’ order preserving collapse $\vartheta_X : T_X \rightarrow \vartheta_T(X)$. In fact, the situation is somewhat more complicated: To define the order relation on $\vartheta_T(X)$, we already need a function

$\iota_X : X \rightarrow \vartheta_T(X)$ between the underlying sets. After the order has been defined, we will want ι_X to be an order embedding. We will introduce a notion of (good) Bachmann–Howard system to keep track of this requirement. Relying on this notion, we will construct a diagram of the form

$$\begin{array}{ccccccc}
 T_{X_0} & \xrightarrow{T_{\iota_{X_0}}} & T_{X_1} & \longrightarrow & \cdots \\
 \downarrow \vartheta_{X_0} & & \downarrow \vartheta_{X_1} & & \\
 X_0 := \emptyset & \xrightarrow{\iota_{X_0}} & X_1 := \vartheta_T(X_0) & \xrightarrow{\iota_{X_1}} & X_2 := \vartheta_T(X_1) & \longrightarrow & \cdots
 \end{array}$$

The order $\text{BH}(T)$ will be defined as the direct limit of the orders X_n . To establish that $\text{BH}(T)$ is a Bachmann–Howard fixed point of T , we will show that the functions $\vartheta_{X_n} : T_{X_n} \rightarrow X_{n+1}$ glue to an almost order preserving function $\vartheta : T_{\text{BH}(T)} \rightarrow \text{BH}(T)$.

Finally, let us mention that it is possible to construct Bachmann–Howard fixed points in an even weaker setting: In [8], we show that $\text{BH}(T)$ can be represented by an ordinal notation system $\vartheta(T)$, which is computable relative to (a suitable representation of) T . The assertion that $\vartheta(T)$ is well founded for any dilator T will be called the computable Bachmann–Howard principle. According to [8, Theorem 4.6], this principle is still equivalent to Π_1^1 -comprehension, even over the base theory \mathbf{RCA}_0 . The construction in [8] is considerably more technical than the one in the present paper, not least because it depends on a representation of dilators in second-order arithmetic. It is often pointed out that ordinal notation systems such as $\vartheta(T)$ are difficult to understand from a purely syntactical standpoint. The present paper provides a transparent semantical construction of $\text{BH}(T)$, which is crucial for the understanding of Bachmann–Howard fixed points.

The author would like to point out that parts of this paper are based on Sections 2.2 and 2.4 of his PhD thesis [6].

2. Preliminaries: dilators and collapsing functions

In the present section, we give precise definitions of notions that have been mentioned in the introduction. We will be most interested in dilators and their Bachmann–Howard fixed points. At the end of the section, we discuss the meta theory in which the present paper is supposed to be formalized.

Dilators are particularly uniform endofunctors on the category of linear orders, with order embeddings as morphisms. In order to express the uniformity condition, we consider the finite subset functor on the category of sets, which is given by

$$\begin{aligned}
 [X]^{<\omega} &= \text{‘the set of finite subsets of } X\text{’}, \\
 [f]^{<\omega}(a) &= \{f(x) \mid x \in a\}.
 \end{aligned}$$

We will also apply $[\cdot]^{<\omega}$ to linear orders, omitting the forgetful functor to their underlying sets. Conversely, a subset of a linear order will often be considered as a suborder. The following notion is essentially due to Girard [9]:

DEFINITION 2.1. A prae-dilator consists of

- (i) an endofunctor $X \mapsto T_X$ of linear orders and
- (ii) a natural transformation $\text{supp}^T : T \Rightarrow [\cdot]^{<\omega}$ that computes supports, in the following sense: For any linear order X and any element $\sigma \in T_X$, we have $\sigma \in \text{rng}(T_{\iota_\sigma})$, where $\iota_\sigma : \text{supp}_X^T(\sigma) \hookrightarrow X$ is the inclusion.

If T_X is well founded for any well-order X , then (T, supp^T) is called a dilator.

We point out that our notion of prae-dilator is slightly different from Girard’s notion of pre-dilator, which involves an additional monotonicity condition. The latter is automatic for well-orders, so that the difference vanishes in the case of dilators. Also, Girard’s definition does not involve the natural transformation supp^T . Instead, it demands that T preserves direct limits and pullbacks. It is not hard to see that the two definitions are equivalent (see [6, Remark 2.2.2]). Nevertheless, it will be very useful to make the supports explicit.

To say when a function is ‘almost’ order preserving we need the following notation: Given a linear order $(X, <_X)$, we define a preorder $<^{\text{fin}}_X$ on $[X]^{<\omega}$ by stipulating

$$a <^{\text{fin}}_X b \quad :\Leftrightarrow \quad \text{‘for any } s \in a \text{ there is a } t \in b \text{ with } s <_X t\text{’}.$$

We will write $s <^{\text{fin}}_X b$ and $a <^{\text{fin}}_X t$ rather than $\{s\} <^{\text{fin}}_X b$ and $a <^{\text{fin}}_X \{t\}$ for singletons. The relation \leq^{fin}_X is defined in the same way. The following notion was introduced in [7], based on the author’s PhD thesis [6] and an earlier arXiv preprint [5]. It is inspired by the definition of the Bachmann–Howard ordinal, in particular by the variant due to Rathjen (cf. [17, Section 1]).

DEFINITION 2.2. Consider a prae-dilator (T, supp^T) and an order X . A function

$$\vartheta : T_X \rightarrow X$$

is called a Bachmann–Howard collapse if the following holds for all $\sigma, \tau \in T_X$:

- (i) If we have $\sigma <_{T_X} \tau$ and $\text{supp}^T_X(\sigma) <^{\text{fin}}_X \vartheta(\tau)$, then we have $\vartheta(\sigma) <_X \vartheta(\tau)$.
- (ii) We have $\text{supp}^T_X(\sigma) <^{\text{fin}}_X \vartheta(\sigma)$.

If such a function exists, then X is called a Bachmann–Howard fixed point of T .

We can now officially introduce the following principle:

DEFINITION 2.3. The abstract Bachmann–Howard principle is the assertion that every dilator has a well-founded Bachmann–Howard fixed point.

To give an example, we recall the functor

$$X \mapsto T_X = X \cup \{\top\}$$

that was considered in the introduction. We obtain a dilator if we set $\text{supp}^T_X(\top) = \emptyset$ and $\text{supp}^T_X(x) = \{x\}$ for $x \in X \subseteq T_X$. It is straightforward to check that the function

$$\vartheta : T_\omega \rightarrow \omega \quad \text{with} \quad \begin{cases} \vartheta(\top) = 0, \\ \vartheta(n) = n + 1, \end{cases}$$

is a Bachmann–Howard collapse. Conversely, if $\vartheta : T_X \rightarrow X$ is any Bachmann–Howard collapse, then we can define an embedding $f : \omega \rightarrow X$ by setting $f(0) = \vartheta(\top)$ and $f(n + 1) = \vartheta(f(n))$.

Let us now discuss the formalization of the previous notions: The meta theory of the present paper is primitive recursive set theory with infinity (**PRS** ω), as introduced by Rathjen [14] (see also the detailed exposition in [6, Chapter 1]). This theory has a function symbol for each primitive recursive set function in the sense of Jensen and Karp [11]. When we speak about class-sized objects of a certain kind (for example, about arbitrary endofunctors on linear orders), we need to observe two restrictions: Firstly, we will only consider class-sized objects which are primitive recursive. Secondly, we cannot quantify over all primitive recursive set functions. However, we can quantify over a primitive recursive family of class-sized functions, by quantifying over its set-sized parameters. Statements about class-sized objects should thus be read as schemata.

In our context, the restrictions from the previous paragraph are harmless: Girard [9] has shown that (prae-)dilators are essentially determined by their restrictions to the category of

natural numbers. In [8, Section 2], we deduce that any prae-dilator is naturally equivalent to one that is given by a primitive recursive set function. Indeed we show that there is a single primitive recursive family that comprises (isomorphic copies of) all prae-dilators. Thus, the abstract Bachmann–Howard principle can be expressed by a single sentence in the language of **PRS** ω . One can even represent prae-dilators in second-order arithmetic (see again [8, Section 2]), but this will not be relevant for the present paper.

Let us also point out what these methodological remarks mean for the construction of Bachmann–Howard fixed points: In the introduction, we have said that the transformation of a (prae-)dilator T into its minimal Bachmann–Howard fixed point $\text{BH}(T)$ can be achieved by a primitive recursive set function. We can now be more precise about this claim: If $(T^u)_{u \in \mathbb{V}}$ is a primitive recursive family of prae-dilators, indexed by elements of the set-theoretic universe, then the function $u \mapsto \text{BH}(T^u)$ will be primitive recursive as well.

Finally, we recall Simpson’s [18, 19] set-theoretic version $\mathbf{ATR}_0^{\text{set}}$ of arithmetical transfinite recursion: It results from **PRS** ω by adding axiom beta (which asserts that every well-founded relation can be collapsed to the \in -relation) and the axiom of countability (which asserts that every set is countable). The additional axioms of $\mathbf{ATR}_0^{\text{set}}$ are needed for Theorem 1.2, since they are used in the proof of Theorem 1.1, as given in [7].

3. Bachmann–Howard systems

In the introduction, we have mentioned linear orders $(\vartheta_T(X), <_{\vartheta_T(X)})$ that allow for an ‘almost’ order preserving collapse $\vartheta_X : T_X \rightarrow \vartheta_T(X)$. The construction of these orders proceeds in two steps. First, we must define the underlying sets:

DEFINITION 3.1. Consider a prae-dilator T . For each linear order X , we define $\vartheta_T(X)$ as the set of terms $\vartheta\sigma$ with $\sigma \in T_X$.

In view of Definition 2.2, the relation $\vartheta\sigma <_{\vartheta_T(X)} \vartheta\tau$ should depend on a comparison between $\text{supp}_X^T(\sigma)$ and $\vartheta\tau$. This is not completely straightforward, because $\text{supp}_X^T(\sigma)$ is a subset of X rather than $\vartheta_T(X)$. To resolve this problem, we introduce the following notion:

DEFINITION 3.2. Consider a linear order X together with functions $\iota_X : X \rightarrow \vartheta_T(X)$ and $L_X : X \rightarrow \omega$. Define $L_{\vartheta_T(X)} : \vartheta_T(X) \rightarrow \omega$ by

$$L_{\vartheta_T(X)}(\vartheta\sigma) := \max\{L_X(x) \mid x \in \text{supp}_X^T(\sigma)\} + 1.$$

If we have

$$L_{\vartheta_T(X)} \circ \iota_X = L_X,$$

then the tuple (X, ι_X, L_X) is called a Bachmann–Howard system (for T).

Note that one obtains $L_{\vartheta_T(X)}(\iota_X(x)) = L_X(x) < L_{\vartheta_T(X)}(\vartheta\sigma)$ for $x \in \text{supp}_X^T(\sigma)$. This allows for the following recursion:

DEFINITION 3.3. Let (X, ι_X, L_X) be a Bachmann–Howard system for the prae-dilator T . Relying on recursion over $L_{\vartheta_T(X)}(\vartheta\sigma) + L_{\vartheta_T(X)}(\vartheta\tau)$, we stipulate that $\vartheta\sigma <_{\vartheta_T(X)} \vartheta\tau$ holds precisely if one of the following clauses is satisfied.

- (i) We have $\sigma <_{T_X} \tau$ and $[\iota_X]^{<\omega}(\text{supp}_X^T(\sigma)) <_{\vartheta_T(X)}^{\text{fin}} \vartheta\tau$.
- (ii) We have $\tau <_{T_X} \sigma$ and $\vartheta\sigma \leq_{\vartheta_T(X)}^{\text{fin}} [\iota_X]^{<\omega}(\text{supp}_X^T(\tau))$.

Let us establish the following basic property:

LEMMA 3.4. *If (X, ι_X, L_X) is a Bachmann–Howard system, then $(\vartheta_T(X), <_{\vartheta_T(X)})$ is a linear order.*

Proof. The antisymmetry of $<_{\vartheta_T(X)}$ follows easily from the antisymmetry of $<_{T_X}$. Trichotomy for $\vartheta\sigma$ and $\vartheta\tau$ is established by induction on $L_{\vartheta_T(X)}(\vartheta\sigma) + L_{\vartheta_T(X)}(\vartheta\tau)$: By symmetry, we may assume $\sigma <_{T_X} \tau$. For an arbitrary $x \in \text{supp}_X^T(\sigma)$, the induction hypothesis provides $\iota_X(x) <_{\vartheta_T(X)} \vartheta\tau$ or $\vartheta\tau \leq_{\vartheta_T(X)} \iota_X(x)$. If the former holds for all $x \in \text{supp}_X^T(\sigma)$, we have $\vartheta\sigma < \vartheta\tau$ by clause (i) of the previous definition. If we have $\vartheta\tau \leq_{\vartheta_T(X)} \iota_X(x)$ for some $x \in \text{supp}_X^T(\sigma)$, we get $\vartheta\tau <_{\vartheta_T(X)} \vartheta\sigma$ by clause (ii). Finally, we argue by induction on $L_{\vartheta_T(X)}(\vartheta\rho) + L_{\vartheta_T(X)}(\vartheta\sigma) + L_{\vartheta_T(X)}(\vartheta\tau)$ to show that $\vartheta\rho <_{\vartheta_T(X)} \vartheta\sigma <_{\vartheta_T(X)} \vartheta\tau$ implies $\vartheta\rho <_{\vartheta_T(X)} \vartheta\tau$. If $\vartheta\rho <_{\vartheta_T(X)} \vartheta\sigma$ and $\vartheta\sigma <_{\vartheta_T(X)} \vartheta\tau$ hold by the same clause of the previous definition, then it is easy to conclude by induction hypothesis. Now assume that $\vartheta\rho <_{\vartheta_T(X)} \vartheta\sigma$ holds by clause (i) while $\vartheta\sigma <_{\vartheta_T(X)} \vartheta\tau$ holds by clause (ii). This means that we have

$$\begin{aligned} \rho <_{T_X} \sigma \quad \text{and} \quad [\iota_X]^{<\omega}(\text{supp}_X^T(\rho)) <_{\vartheta_T(X)}^{\text{fin}} \vartheta\sigma, \\ \tau <_{T_X} \sigma \quad \text{and} \quad \vartheta\sigma \leq_{\vartheta_T(X)}^{\text{fin}} [\iota_X]^{<\omega}(\text{supp}_X^T(\tau)). \end{aligned}$$

If we have $\rho <_{T_X} \tau$ or $\tau <_{T_X} \rho$, then we can conclude by induction hypothesis. It remains to exclude the case $\rho = \tau$: By the assumption $\vartheta\sigma \leq_{\vartheta_T(X)}^{\text{fin}} [\iota_X]^{<\omega}(\text{supp}_X^T(\tau))$, pick an element $x \in \text{supp}_X^T(\tau) = \text{supp}_X^T(\rho)$ with $\vartheta\sigma \leq_{\vartheta_T(X)} \iota_X(x)$. In view of $[\iota_X]^{<\omega}(\text{supp}_X^T(\rho)) <_{\vartheta_T(X)}^{\text{fin}} \vartheta\sigma$, we also have $\iota_X(x) <_{\vartheta_T(X)} \vartheta\sigma$. The induction hypothesis allows us to conclude $\iota_X(x) <_{\vartheta_T(X)} \iota_X(x)$ by transitivity. This contradicts antisymmetry, as desired. Finally, assume that $\vartheta\rho <_{\vartheta_T(X)} \vartheta\sigma$ holds by clause (ii) while $\vartheta\sigma <_{\vartheta_T(X)} \vartheta\tau$ holds by clause (i). This means that we have

$$\begin{aligned} \sigma <_{T_X} \rho \quad \text{and} \quad \vartheta\rho \leq_{\vartheta_T(X)}^{\text{fin}} [\iota_X]^{<\omega}(\text{supp}_X^T(\sigma)), \\ \sigma <_{T_X} \tau \quad \text{and} \quad [\iota_X]^{<\omega}(\text{supp}_X^T(\sigma)) <_{\vartheta_T(X)}^{\text{fin}} \vartheta\tau. \end{aligned}$$

The assumption $\vartheta\rho \leq_{\vartheta_T(X)}^{\text{fin}} [\iota_X]^{<\omega}(\text{supp}_X^T(\sigma))$ yields an $x \in \text{supp}_X^T(\sigma)$ with $\vartheta\rho \leq_{\vartheta_T(X)} \iota_X(x)$. By $[\iota_X]^{<\omega}(\text{supp}_X^T(\sigma)) <_{\vartheta_T(X)}^{\text{fin}} \vartheta\tau$, we also have $\iota_X(x) <_{\vartheta_T(X)} \vartheta\tau$. Using the induction hypothesis, we can conclude $\vartheta\rho <_{\vartheta_T(X)} \vartheta\tau$, as required. \square

We can now define the collapsing functions mentioned in the introduction:

DEFINITION 3.5. Let (X, ι_X, L_X) be a Bachmann–Howard system for the prae-dilator T . We define a function $\vartheta_X : T_X \rightarrow \vartheta_T(X)$ by setting $\vartheta_X(\sigma) = \vartheta\sigma$.

Let us recover the conditions from Definition 2.2:

PROPOSITION 3.6. *Assume that (X, ι_X, L_X) is a Bachmann–Howard system for the prae-dilator T . Then the following holds for all $\sigma, \tau \in T_X$:*

- (i) *if we have $\sigma <_{T_X} \tau$ and $[\iota_X]^{<\omega}(\text{supp}_X^T(\sigma)) <_{\vartheta_T(X)}^{\text{fin}} \vartheta_X(\tau)$, then $\vartheta_X(\sigma) <_{\vartheta_T(X)} \vartheta_X(\tau)$ holds;*
- (ii) *we have $[\iota_X]^{<\omega}(\text{supp}_X^T(\sigma)) <_{\vartheta_T(X)}^{\text{fin}} \vartheta_X(\sigma)$.*

Proof. Claim (i) is immediate by the definitions. To establish claim (ii), we consider the auxiliary function $E_X : \vartheta_T(X) \rightarrow [\vartheta_T(X)]^{<\omega}$ with

$$E_X(\vartheta\sigma) := \{\vartheta\sigma\} \cup \bigcup \{E_X(\iota_X(x)) \mid x \in \text{supp}_X^T(\sigma)\},$$

which is defined by recursion on $L_{\vartheta_T(X)}(\vartheta\sigma) > L_{\vartheta_T(X)}(\iota_X(x))$. One may think of $\vartheta\rho \in E_X(\vartheta\sigma)$ as a subterm of $\vartheta\sigma$: A straightforward induction on $L_{\vartheta_T(X)}(\vartheta\sigma)$ shows that $\vartheta\rho \in E_X(\vartheta\sigma)$

implies $E_X(\vartheta\rho) \subseteq E_X(\vartheta\sigma)$ and $L_{\vartheta_T(X)}(\vartheta\rho) \leq L_{\vartheta_T(X)}(\vartheta\sigma)$. The crucial step towards claim (ii) is the implication

$$\vartheta\rho \in E_X(\vartheta\sigma) \quad \Rightarrow \quad \vartheta\rho \leq_{\vartheta_T(X)} \vartheta\sigma,$$

which we prove by induction on $L_{\vartheta_T(X)}(\vartheta\rho) + L_{\vartheta_T(X)}(\vartheta\sigma)$. Let us distinguish three cases: If we have $\rho = \sigma$, then the claim is immediate. Now assume $\rho <_{T_X} \sigma$. To infer $\vartheta\rho <_{\vartheta_T(X)} \vartheta\sigma$, we must establish $[\iota_X]^{<\omega}(\text{supp}_X^T(\rho)) <_{\vartheta_T(X)}^{\text{fin}} \vartheta\sigma$. For an arbitrary $x \in \text{supp}_X^T(\rho)$, we have

$$L_{\vartheta_T(X)}(\iota_X(x)) < L_{\vartheta_T(X)}(\vartheta\rho) \leq L_{\vartheta_T(X)}(\vartheta\sigma),$$

as well as

$$\iota_X(x) \in E_X(\iota_X(x)) \subseteq E_X(\vartheta\rho) \subseteq E_X(\vartheta\sigma).$$

So the induction hypothesis yields $\iota_X(x) \leq_{\vartheta_T(X)} \vartheta\sigma$. Also note that $\iota_X(x)$ and $\vartheta\sigma$ cannot be the same term, since we have $L_{\vartheta_T(X)}(\iota_X(x)) < L_{\vartheta_T(X)}(\vartheta\sigma)$. Thus, we get $\iota_X(x) <_{\vartheta_T(X)} \vartheta\sigma$, as required. Finally, we consider the case $\sigma <_{T_X} \rho$. By the definition of $E_X(\vartheta\sigma)$, we may pick a $y \in \text{supp}_X^T(\sigma)$ with $\vartheta\rho \in E_X(\iota_X(y))$. The induction hypothesis provides $\vartheta\rho \leq_{\vartheta_T(X)} \iota_X(y)$ and thus

$$\vartheta\rho \leq_{\vartheta_T(X)}^{\text{fin}} [\iota_X]^{<\omega}(\text{supp}_X^T(\sigma)).$$

Then we can conclude $\vartheta\rho <_{\vartheta_T(X)} \vartheta\sigma$ by definition. To deduce claim (ii) of the proposition, we observe that $x \in \text{supp}_X^T(\sigma)$ yields

$$\iota_X(x) \in E_X(\iota_X(x)) \subseteq E_X(\vartheta\sigma).$$

We have just shown that this implies $\iota_X(x) \leq_{\vartheta_T(X)} \vartheta\sigma$. Due to $L_{\vartheta_T(X)}(\iota_X(x)) < L_{\vartheta_T(X)}(\vartheta\sigma)$, the terms $\iota_X(x)$ and $\vartheta\sigma$ must be different. So indeed we get $\iota_X(x) <_{\vartheta_T(X)} \vartheta\sigma = \vartheta_X(\sigma)$. \square

In particular, we have shown that the condition $\tau <_{T_X} \sigma$ in clause (ii) of Definition 3.3 becomes redundant: The implication

$$\vartheta\sigma \leq_{\vartheta_T(X)}^{\text{fin}} [\iota_X]^{<\omega}(\text{supp}_X^T(\tau)) \quad \Rightarrow \quad \vartheta\sigma <_{\vartheta_T(X)} \vartheta\tau$$

follows from $[\iota_X]^{<\omega}(\text{supp}_X^T(\tau)) <_{\vartheta_T(X)}^{\text{fin}} \vartheta_X(\tau)$ and transitivity. To define the linear order $<_{\vartheta_T(X)}$, we have relied on a function $\iota_X : X \rightarrow \vartheta_T(X)$ which respects the length assignments L_X and $L_{\vartheta_T(X)}$. Now that we have an order on $\vartheta_T(X)$ we want ι_X to respect it as well:

DEFINITION 3.7. A Bachmann–Howard system (X, ι_X, L_X) is called good if the function $\iota_X : X \rightarrow \vartheta_T(X)$ is an order embedding.

Note that this justifies the arrow $T_{\iota_{X_0}}$ in the second diagram from the introduction: If ι_{X_0} is an embedding of X_0 into $X_1 = \vartheta_T(X_0)$, then $T_{\iota_{X_0}}$ is an embedding of T_{X_0} into T_{X_1} . Based on this arrow, we can also construct the arrow $\iota_{X_1} : X_1 \rightarrow X_2$:

DEFINITION 3.8. Let (X, ι_X, L_X) be a good Bachmann–Howard system. We define a function $\iota_{\vartheta_T(X)} : \vartheta_T(X) \rightarrow \vartheta_T(\vartheta_T(X))$ by setting $\iota_{\vartheta_T(X)}(\vartheta\sigma) := \vartheta T_{\iota_X}(\sigma)$.

Note that the empty order, together with the unique functions $\iota_\emptyset : \emptyset \rightarrow \vartheta_T(\emptyset)$ and $L_\emptyset : \emptyset \rightarrow \omega$, is a good Bachmann–Howard system for any prae-dilator. Once we have a starting point, we can use the following result to construct iterations:

THEOREM 3.9. Consider a prae-dilator T . If (X, ι_X, L_X) is a good Bachmann–Howard system for T , then so is $(\vartheta_T(X), \iota_{\vartheta_T(X)}, L_{\vartheta_T(X)})$.

Proof. Abbreviate $\vartheta_T^2(X) := \vartheta_T(\vartheta_T(X))$ and consider $L_{\vartheta_T^2(X)} : \vartheta_T^2(X) \rightarrow \omega$ as constructed in Definition 3.2. Since supp^T is a natural transformation and X is a Bachmann–Howard system, we can compute

$$\begin{aligned} L_{\vartheta_T^2(X)} \circ \iota_{\vartheta_T(X)}(\vartheta\sigma) &= L_{\vartheta_T^2(X)}(\vartheta T_{\iota_X}(\sigma)) &&= \\ &= \max\{L_{\vartheta_T(X)}(s) \mid s \in \text{supp}_{\vartheta_T(X)}^T(T_{\iota_X}(\sigma))\} + 1 = \\ &= \max\{L_{\vartheta_T(X)}(s) \mid s \in [\iota_X]^{<\omega}(\text{supp}_X^T(\sigma))\} + 1 = \\ &= \max\{L_{\vartheta_T(X)}(\iota_X(x)) \mid x \in \text{supp}_X^T(\sigma)\} + 1 &&= \\ &= \max\{L_X(x) \mid x \in \text{supp}_X^T(\sigma)\} + 1 &&= L_{\vartheta_T(X)}(\vartheta\sigma). \end{aligned}$$

This shows that $(\vartheta_T(X), \iota_{\vartheta_T(X)}, L_{\vartheta_T(X)})$ is a Bachmann–Howard system. We can now invoke Definition 3.3 and Lemma 3.4 to equip $\vartheta_T^2(X)$ with a linear order. To show that $\vartheta_T(X)$ is good, we establish the implication

$$s <_{\vartheta_T(X)} t \quad \Rightarrow \quad \iota_{\vartheta_T(X)}(s) <_{\vartheta_T^2(X)} \iota_{\vartheta_T(X)}(t),$$

by induction on $L_{\vartheta_T(X)}(s) + L_{\vartheta_T(X)}(t)$. First assume that $s = \vartheta\sigma <_{\vartheta_T(X)} \vartheta\tau = t$ holds by clause (i) of Definition 3.3. This means that we have

$$\sigma <_{T_X} \tau \quad \text{and} \quad [\iota_X]^{<\omega}(\text{supp}_X^T(\sigma)) <_{\vartheta_T(X)}^{\text{fin}} t.$$

Clearly, we have $T_{\iota_X}(\sigma) <_{T_{\vartheta_T(X)}} T_{\iota_X}(\tau)$. To conclude

$$\iota_{\vartheta_T(X)}(s) = \vartheta T_{\iota_X}(\sigma) <_{\vartheta_T^2(X)} \vartheta T_{\iota_X}(\tau) = \iota_{\vartheta_T(X)}(t),$$

we need to establish

$$[\iota_{\vartheta_T(X)}]^{<\omega}(\text{supp}_{\vartheta_T(X)}^T(T_{\iota_X}(\sigma))) <_{\vartheta_T^2(X)}^{\text{fin}} \iota_{\vartheta_T(X)}(t).$$

For any $r \in \text{supp}_{\vartheta_T(X)}^T(T_{\iota_X}(\sigma)) = [\iota_X]^{<\omega}(\text{supp}_X^T(\sigma))$, we have $r <_{\vartheta_T(X)} t$ by assumption. The induction hypothesis yields $\iota_{\vartheta_T(X)}(r) <_{\vartheta_T^2(X)} \iota_{\vartheta_T(X)}(t)$, as required. A similar argument applies if $s <_{\vartheta_T(X)} t$ holds by clause (ii) of Definition 3.3. \square

By Definition 3.5, we obtain a collapse $\vartheta_{\vartheta_T(X)} : T_{\vartheta_T(X)} \rightarrow \vartheta_T(\vartheta_T(X))$. The following shows that the diagram from the introduction commutes:

PROPOSITION 3.10. *Assume that X and thus $\vartheta_T(X)$ is a good Bachmann–Howard system for a prae-dilator T . Then we have $\iota_{\vartheta_T(X)} \circ \vartheta_X = \vartheta_{\vartheta_T(X)} \circ T_{\iota_X}$.*

Proof. Unravelling definitions, we compute

$$\vartheta_{\vartheta_T(X)}(T_{\iota_X}(\sigma)) = \vartheta T_{\iota_X}(\sigma) = \iota_{\vartheta_T(X)}(\vartheta\sigma) = \iota_{\vartheta_T(X)}(\vartheta_X(\sigma)),$$

as promised. \square

4. The minimal Bachmann–Howard fixed point

In the previous section, we have given a detailed construction of the diagram from the introduction. The goal of this section is to investigate its direct limit. We have already observed that the empty order \emptyset , together with the unique functions $\iota_\emptyset : \emptyset \rightarrow \vartheta_T(\emptyset)$ and $L_\emptyset : \emptyset \rightarrow \omega$, is a good Bachmann–Howard system for any prae-dilator. Together with Theorem 3.9, we can construct the following objects:

DEFINITION 4.1. Consider a prae-dilator T . We build a sequence of good Bachmann–Howard systems by setting

$$(X_0, \iota_{X_0}, L_{X_0}) := (\emptyset, \iota_\emptyset, L_\emptyset),$$

$$(X_{n+1}, \iota_{X_{n+1}}, L_{X_{n+1}}) := (\vartheta_T(X_n), \iota_{\vartheta_T(X_n)}, L_{\vartheta_T(X_n)}).$$

Define the order $\text{BH}(T)$ as the direct limit of the system $(X_n, \iota_{X_n} : X_n \rightarrow X_{n+1})_{n \in \omega}$. It comes with embeddings $j_{X_n} : X_n \rightarrow \text{BH}(T)$ that satisfy $j_{X_{n+1}} \circ \iota_{X_n} = j_{X_n}$.

As explained in the introduction, the present paper is supposed to be formalized in primitive recursive set theory ($\mathbf{PRS}\omega$). Let us briefly discuss the formalization of the above constructions (more details can be found in [6, Sections 1.1, 1.2, 2.2]): Given a primitive recursive family $(T^u)_{u \in \mathbb{V}}$ of prae-dilators, it is straightforward to see that the transformation $(u, X) \mapsto \vartheta_{T^u}(X)$ is a primitive recursive set function, and that the properties from the previous section can be established in $\mathbf{PRS}\omega$. Write X_n^u for the Bachmann–Howard systems from the above definition, constructed with respect to T^u . Invoking primitive recursion along the ordinals we see that $(u, n) \mapsto X_n^u$ is a primitive recursive set function. It follows that the transformation of u into the underlying set of the direct limit $\text{BH}(T^u)$ is primitive recursive as well, since the latter can be explicitly represented by

$$\text{BH}(T^u) = \{(n, s) \mid s \in X_{n+1}^u \wedge s \notin \text{rng}(\iota_{X_n^u})\}.$$

Similarly, one checks that the universal property is witnessed by a primitive recursive transformation (see [6, Lemma 2.2.17]). In particular, we can use the universal property (in the category of sets) to construct the limit order on $\text{BH}(T^u)$. Thus, we finally learn that $u \mapsto (\text{BH}(T^u), <_{\text{BH}(T^u)})$ is a primitive recursive set function. Let us now come to the first of our main results:

THEOREM 4.2. For each prae-dilator T , the order $\text{BH}(T)$ is a Bachmann–Howard fixed point of T .

Proof. In order to construct a Bachmann–Howard collapse $\vartheta : T_{\text{BH}(T)} \rightarrow \text{BH}(T)$, we will exploit the fact that $T_{\text{BH}(T)}$ is a direct limit of the system

$$(T_{X_n}, T_{\iota_{X_n}} : T_{X_n} \rightarrow T_{X_{n+1}})_{n \in \omega}.$$

Indeed, Girard’s original definition explicitly demands that (prae-)dilators preserve direct limits. Since we have worked with a different formulation of the definition, we shall give a short proof of this fact: Consider an arbitrary $\sigma \in T_{\text{BH}(T)}$. Since the support $\text{supp}_{\text{BH}(T)}^T(\sigma)$ is a finite subset of $\text{BH}(T)$, it is contained in the range of some embedding j_{X_n} . Using clause (ii) of Definition 2.1, we can infer that σ lies in the range of $T_{j_{X_n}}$. Thus, we have established

$$T_{\text{BH}(T)} = \bigcup_{n \in \omega} \text{rng}(T_{j_{X_n}}),$$

which ensures that $T_{\text{BH}(T)}$, together with the functions $T_{j_{X_n}} : T_{X_n} \rightarrow T_{\text{BH}(T)}$, is the desired direct limit (both in the category of linear orders and in the category of sets). Relying on Definition 3.5, let us now consider the functions

$$j_{X_{n+1}} \circ \vartheta_{X_n} : T_{X_n} \rightarrow \text{BH}(T).$$

We can use Proposition 3.10 to compute

$$\begin{aligned} (j_{X_{n+2}} \circ \vartheta_{X_{n+1}}) \circ T_{\iota_{X_n}} &= j_{X_{n+2}} \circ (\vartheta_{\vartheta_T(X_n)} \circ T_{\iota_{X_n}}) \\ &= j_{X_{n+2}} \circ (\iota_{\vartheta_T(X_n)} \circ \vartheta_{X_n}) = (j_{X_{n+2}} \circ \iota_{X_{n+1}}) \circ \vartheta_{X_n} = j_{X_{n+1}} \circ \vartheta_{X_n}. \end{aligned}$$

Now the universal property of $T_{\text{BH}(T)}$ yields a function

$$\vartheta : T_{\text{BH}(T)} \rightarrow \text{BH}(T) \quad \text{with} \quad \vartheta \circ T_{j_{X_n}} = j_{X_{n+1}} \circ \vartheta_{X_n}.$$

We have to verify the conditions from Definition 2.2: Aiming at condition (i), consider elements $\sigma, \tau \in T_{\text{BH}(T)}$ with

$$\sigma <_{T_{\text{BH}(T)}} \tau \quad \text{and} \quad \text{supp}_{\text{BH}(T)}^T(\sigma) <_{\text{BH}(T)}^{\text{fin}} \vartheta(\tau).$$

Pick n large enough to write $\sigma = T_{j_{X_n}}(\sigma_0)$ and $\tau = T_{j_{X_n}}(\tau_0)$ with $\sigma_0, \tau_0 \in T_{X_n}$. Then we have $\sigma_0 <_{T_{X_n}} \tau_0$, as well as

$$\begin{aligned} [j_{X_{n+1}}]^{<\omega} \circ [\iota_{X_n}]^{<\omega}(\text{supp}_{X_n}^T(\sigma_0)) &= [j_{X_n}]^{<\omega}(\text{supp}_{X_n}^T(\sigma_0)) \\ &= \text{supp}_{\text{BH}(T)}^T(T_{j_{X_n}}(\sigma_0)) <_{\text{BH}(T)}^{\text{fin}} \vartheta(\tau) = \vartheta \circ T_{j_{X_n}}(\tau_0) = j_{X_{n+1}} \circ \vartheta_{X_n}(\tau_0). \end{aligned}$$

Since $j_{X_{n+1}}$ is an order embedding we obtain $[\iota_{X_n}]^{<\omega}(\text{supp}_{X_n}^T(\sigma_0)) <_{X_{n+1}}^{\text{fin}} \vartheta_{X_n}(\tau_0)$. Now Proposition 3.6 yields $\vartheta_{X_n}(\sigma_0) <_{X_{n+1}} \vartheta_{X_n}(\tau_0)$ and then

$$\vartheta(\sigma) = j_{X_{n+1}} \circ \vartheta_{X_n}(\sigma_0) <_{\text{BH}(T)} j_{X_{n+1}} \circ \vartheta_{X_n}(\tau_0) = \vartheta(\tau).$$

To establish condition (ii) of Definition 2.2, we again write $\sigma = T_{j_{X_n}}(\sigma_0)$. By Proposition 3.6, we have $[\iota_{X_n}]^{<\omega}(\text{supp}_{X_n}^T(\sigma_0)) <_{X_{n+1}}^{\text{fin}} \vartheta_{X_n}(\sigma_0)$. This implies

$$\begin{aligned} \text{supp}_{\text{BH}(T)}^T(\sigma) &= \text{supp}_{\text{BH}(T)}^T(T_{j_{X_n}}(\sigma_0)) = [j_{X_n}]^{<\omega}(\text{supp}_{X_n}^T(\sigma_0)) \\ &= [j_{X_{n+1}}]^{<\omega} \circ [\iota_{X_n}]^{<\omega}(\text{supp}_{X_n}^T(\sigma_0)) <_{\text{BH}(T)}^{\text{fin}} j_{X_{n+1}} \circ \vartheta_{X_n}(\sigma_0) = \vartheta \circ T_{j_{X_n}}(\sigma_0) = \vartheta(\sigma), \end{aligned}$$

just as required. □

The previous results were formulated for arbitrary prae-dilators, whether they preserve well-foundedness or not. Restricting our attention to dilators, we obtain a more explicit version of the Bachmann–Howard principle:

DEFINITION 4.3. The predicative Bachmann–Howard principle is the assertion that $\text{BH}(T)$ is well founded for any dilator T .

The nomenclature alludes to the view that the construction of $\text{BH}(T)$ is predicatively acceptable, since it is realized by a primitive recursive set function (cf. [3]). To avoid misunderstanding, we point out that the well-foundedness of $\text{BH}(T)$ cannot be established by predicative means: Indeed, we will see that the predicative Bachmann–Howard principle is equivalent to Π_1^1 -comprehension. This equivalence also ensures that the predicative Bachmann–Howard principle is sound, which is not trivial at all (in general, well-foundedness is not preserved under direct limits). Theorem 4.2 shows that the predicative Bachmann–Howard principle implies its abstract counterpart. The converse implication follows from the fact that $\text{BH}(T)$ is the minimal Bachmann–Howard fixed point:

THEOREM 4.4. *Consider a prae-dilator T . The order $\text{BH}(T)$ can be embedded into any Bachmann–Howard fixed point of T .*

Proof. Let Y be a Bachmann–Howard fixed point of T , witnessed by a Bachmann–Howard collapse $\vartheta_Y : T_Y \rightarrow Y$. Given a good Bachmann–Howard system (X, ι_X, L_X) and an embedding $h_X : X \rightarrow Y$, we can define a function $h_X^\vartheta : \vartheta_T(X) \rightarrow Y$ by

$$h_X^\vartheta(\vartheta\sigma) = \vartheta_Y(T_{h_X}(\sigma)).$$

If we have

$$h_X^\vartheta \circ \iota_X = h_X,$$

then h_X is called a Bachmann–Howard interpretation of X . The main step of the present proof is to establish the following claim: If h_X is a Bachmann–Howard interpretation of X , then h_X^ϑ is a Bachmann–Howard interpretation of $\vartheta_T(X)$ (in particular it is an order embedding). Based on this claim, we can conclude as follows: Clearly, the empty map $h_{X_0} : X_0 = \emptyset \rightarrow Y$ is a Bachmann–Howard interpretation of X_0 . Iteratively we can then construct Bachmann–Howard interpretations

$$h_{X_{n+1}} := h_{X_n}^\vartheta : X_{n+1} \rightarrow Y$$

of the orders from Definition 4.1. The definition of Bachmann–Howard interpretation ensures $h_{X_{n+1}} \circ \iota_{X_n} = h_{X_n}$. Thus, the universal property of the limit $\text{BH}(T)$ allows us to glue the embeddings $h_{X_n} : X_n \rightarrow Y$ to the desired embedding of $\text{BH}(T)$ into Y . In order to establish the open claim, we consider a Bachmann–Howard interpretation $h_X : X \rightarrow Y$. The implication

$$s <_{\vartheta_T(X)} t \rightarrow h_X^\vartheta(s) <_Y h_X^\vartheta(t)$$

can be established by induction on $L_{\vartheta_T(X)}(s) + L_{\vartheta_T(X)}(t)$. Let us first consider the case where $s = \vartheta\sigma <_{\vartheta_T(X)} \vartheta\tau = t$ holds because of

$$\sigma <_{T_X} \tau \quad \text{and} \quad [\iota_X]^{<\omega}(\text{supp}_X^T(\sigma)) <_{\vartheta_T(X)}^{\text{fin}} t.$$

Then we get $T_{h_X}(\sigma) <_{T_Y} T_{h_X}(\tau)$. Also recall that $L_{\vartheta_T(X)}(\iota_X(x)) < L_{\vartheta_T(X)}(s)$ holds for any $x \in \text{supp}_X^T(\sigma)$. Thus, the definition of Bachmann–Howard interpretation and the induction hypothesis yield

$$\begin{aligned} \text{supp}_Y^T(T_{h_X}(\sigma)) &= [h_X]^{<\omega}(\text{supp}_X^T(\sigma)) \\ &= [h_X^\vartheta]^{<\omega} \circ [\iota_X]^{<\omega}(\text{supp}_X^T(\sigma)) <_Y^{\text{fin}} = h_X^\vartheta(t) = \vartheta_Y(T_{h_X}(\tau)). \end{aligned}$$

In view of Definition 2.2, we obtain the desired inequality

$$h_X^\vartheta(s) = \vartheta_Y(T_{h_X}(\sigma)) <_Y \vartheta_Y(T_{h_X}(\tau)) = h_X^\vartheta(t).$$

Next, assume that $s = \vartheta\sigma <_{\vartheta_T(X)} \vartheta\tau = t$ holds because of $\tau <_{T_X} \sigma$ (which is in fact redundant) and $s \leq_{\vartheta_T(X)}^{\text{fin}} [\iota_X]^{<\omega}(\text{supp}_X^T(\tau))$. Parallel to the above, we obtain

$$h_X^\vartheta(s) \leq_Y^{\text{fin}} \text{supp}_Y^T(T_{h_X}(\tau)).$$

Since Definition 2.2 provides the inequality

$$\text{supp}_Y^T(T_{h_X}(\tau)) <_Y^{\text{fin}} \vartheta_Y(T_{h_X}(\tau)) = h_X^\vartheta(t),$$

we can infer $h_X^\vartheta(s) <_Y h_X^\vartheta(t)$ by transitivity. So far we have established that $h_X^\vartheta : \vartheta_T(X) \rightarrow Y$ is an embedding. To conclude that it is a Bachmann–Howard interpretation, we consider the function $h_X^{\vartheta^2} := (h_X^\vartheta)^\vartheta : \vartheta_T(\vartheta_T(X)) \rightarrow Y$ and compute

$$h_X^{\vartheta^2} \circ \iota_{\vartheta_T(X)}(\vartheta\sigma) = h_X^{\vartheta^2}(\vartheta T_{\iota_X}(\sigma)) = \vartheta_Y(T_{h_X^\vartheta} \circ T_{\iota_X}(\sigma)) = \vartheta_Y(T_{h_X}(\sigma)) = h_X^\vartheta(\vartheta\sigma),$$

using the assumption that h_X is a Bachmann–Howard interpretation. \square

We can now complete the proof of Theorem 1.2, which was stated in the introduction:

Proof. In view of Theorem 1.1 (which was established in [7], based on similar results in [5, 6]), it suffices to show that the abstract and the predicative Bachmann–Howard principle are equivalent. To show that the former implies the latter, we assume that Y is a well-founded Bachmann–Howard fixed point of a given dilator T . By the previous theorem, there is an order embedding of $\text{BH}(T)$ into Y . This ensures that $\text{BH}(T)$ is well founded as well, as demanded by the predicative Bachmann–Howard principle. For the other direction, we consider a dilator T and assume that $\text{BH}(T)$ is well founded. From Theorem 4.2, we know that $\text{BH}(T)$ is a Bachmann–Howard fixed point of T . Thus, $\text{BH}(T)$ itself serves as a witness for the abstract Bachmann–Howard principle. \square

As explained in the introduction, the point of the predicative Bachmann–Howard principle is that it separates the construction of a Bachmann–Howard fixed point from the question of well-foundedness. Thus, it splits the impredicative principle of Π_1^1 -comprehension into a predicative construction and a statement about the preservation of well-foundedness.

Acknowledgements. Open access funding enabled and organized by Projekt DEAL.

References

1. P. ACZEL, ‘Mathematical problems in logic’, PhD Thesis, University of Oxford, 1966.
2. P. ACZEL, ‘Normal functors on linear orderings’, *J. Symb. Log.* 32 (1967) 430. (Abstract to a paper presented at the annual meeting of the Association for Symbolic Logic, Houston, Texas, 1967.)
3. S. FEFERMAN, ‘Predicatively reducible systems of set theory’, *Axiomatic set theory* Proceedings of Symposia in Pure Mathematics, vol. XIII, part II (ed. T. J. Jech; American Mathematical Society, Providence, RI, 1974) 11–32.
4. S. FEFERMAN, ‘Predicativity’, *The Oxford handbook of philosophy of mathematics and logic* (ed. S. Shapiro; Oxford University Press, Oxford, 2005) 590–624.
5. A. FREUND, ‘A higher Bachmann-Howard principle’, 2017, arXiv:1704.01662.
6. A. FREUND, ‘Type-two well-ordering principles, admissible sets, and Π_1^1 -comprehension’, PhD Thesis, University of Leeds, 2018.
7. A. FREUND, ‘ Π_1^1 -comprehension as a well-ordering principle’, *Adv. Math.* 355 (2019), <https://doi.org/10.1016/j.aim.2019.106767>.
8. A. FREUND, ‘Computable aspects of the Bachmann-Howard principle’, *J. Math. Log.*, to appear.
9. J.-Y. GIRARD, ‘ Π_2^1 -logic, part 1: dilators’, *Ann. Pure Appl. Logic* 21 (1981) 75–219.
10. J.-Y. GIRARD, *Proof theory and logical complexity*, vol. 1, Studies in Proof Theory (Bibliopolis, Napoli, 1987).
11. R. B. JENSEN and C. KARP, ‘Primitive recursive set functions’, *Axiomatic set theory* Proceedings of Symposia in Pure Mathematics, vol. XIII, part I (ed. Dana S. Scott; American Mathematical Society, 1971) 143–176.
12. A. MONTALBÁN, ‘Ordinal functors and $\Pi_1^1 - \mathbf{CA}_0$ ’, December 2009, draft.
13. A. MONTALBÁN, ‘Open questions in reverse mathematics’, *Bull. Symb. Log.* 17 (2011) 431–454.
14. M. RATHJEN, ‘A proof-theoretic characterization of the primitive recursive set functions’, *J. Symb. Log.* 57 (1992) 954–969.
15. M. RATHJEN, *Omega models and well-ordering principles, beta models and functors*, Talk at the Reverse Mathematics Workshop, University of Chicago, September, 2011.
16. M. RATHJEN, ‘ ω -models and well-ordering principles’, *Foundational adventures: essays in honor of Harvey M. Friedman* (ed. N. Tennant; College Publications, London, 2014) 179–212.
17. M. RATHJEN and A. WEIERMANN, ‘Proof-theoretic investigations on Kruskal’s theorem’, *Ann. Pure Appl. Logic* 60 (1993) 49–88.

18. S. G. SIMPSON, 'Set theoretic aspects of \mathbf{ATR}_0 ', *Logic Colloquium '80 Studies in Logic and the Foundations of Mathematics* (eds D. van Dalen, D. Lascar and T. J. Smiley; North Holland, Amsterdam, 1982) 255–271.
19. S. G. SIMPSON, *Subsystems of second order arithmetic*, Perspectives in Logic (Cambridge University Press, Cambridge, 2009).

Anton Freund
Fachbereich Mathematik
Technische Universität Darmstadt
Schlossgartenstr. 7
64289 Darmstadt
Germany
freund@mathematik.tu-darmstadt.de