Why Ordinals are Good for You

Ingo Lepper⋆ and Georg Moser⋆⋆

WWU Münster, Germany,
Institut für Mathematische Logik und Grundlagenforschung,
Einsteinstraße 62, D-48149 Münster
email: {lepper,moserg}@math.uni-muenster.de

Abstract. The course introduces the elementary notion of ordinals, following the classical (set-theoretical) foot steps. Our emphasis lies on the applicability of ordinals as 'universal scales' of wellfounded orderings.
We aim at a canonical measure of termination orderings. In particular a precise characterization of the 'strength' of well-known orderings such as 'lexicographic path orders' and 'multiset path orders' is given.
We present the surprising result that these orderings cover an impressively large segment of ordinals. Based on this result, we cover recent research into simple termination: We construct simply terminating rewrite systems that admit derivation sequences whose lengths easily capture the dimensions of the universe. This comes about by the hidden complexity of the lexicographic path ordering. Furthermore we exploit the slow-growing hierarchy to optimally characterize the derivation length of rewrite systems. In particular we study this approach in the context of rewrite system shown to be terminating by by 'lexicographic path orders' or 'multiset path orders', respectively.
The course provides an alternative approach to ordinals that directly addresses one of the most important topics of computer science: Termination orderings.

1 Introduction

The general aim of this course is to define a canonical measure of termination orders. One of the most important topics of computer science is *termination*. Whenever we are given a concrete computer program P , we are not only in need to verify that this program meets its specification, but also to verify that P will terminate on all (legal) inputs as defined in the specification.

Instead of directly studying the termination behaviour of arbitrary programs P, we will base our study on a suitable *abstract model of computation*. A suitable choice for such an abstract model would be any universal machine model, such as for example (well-known) *Turing machines*. We believe that Turing machines provide an almost perfect notion of (universal) computation. However, Turing machines are not precisely the best environment to study for example the termination behaviour of a given Java-applet. Thus we will employ below another, yet

⋆ Partially supported by DFG grant WE 2178/2–1
⋆⋆ Supported by a Marie Curie fellowship, grant number HPMF-CT-2002-015777

equivalent universal, abstract model of computation: term rewrite system (TRS). It is easy to see that any Turing machine $\mathcal{M}$ can be coded as a rewrite system. Thus the halting problem for $\mathcal{M}$ on all inputs is reducible to termination of rewrite systems.

Hence the underlying model of computation of this course are TRSs and we want to define and study a "canonical measure of termination orders" in the realm of TRSs. It should already have become clear what is meant by "termination order'. In the context of term rewriting we can make this notion precise. Suppose R is a finite rewrite system. We write $\to_R$ for the rewrite relation induced by R. Now by *termination order* we understand a well-founded *rewrite order* $(P, \prec)$, that is compatible with the set of rules R ($s \to_R t$ implies $s \succ t$). A rewrite order is a partial order that is closed under contexts and substitutions. (Alternatively we say that a rewrite order is monotone and stable.) Unfortunately it is not so easy to define the notion "canonical measure" as easy as "termination order". For one this is due to the fact that the concept of being a canonical or uniform measure cannot be described precisely in mathematical terms.

We introduces the elementary notion of *ordinals*, well-known from set theory. As a (rather naive) starting point into the realm of ordinals we say that ordinals extend the natural numbers into the transfinite. Consider as sequence of natural numbers $0, 1, 2, 3, \ldots$, then one introduces the first *infinite* ordinal, called ω, which represents the *supremum* of $0, 1, 2, 3, \ldots$. Now we allow ω to appear in the domain of well-known functions such as $+$, $\cdot$, and exponentiation. That is we introduce for example the function $\omega^{\cdot}$. Then we consider ω-*towers*:

$$\omega, \omega^{\omega}, \omega^{\omega^{\omega}}, \omega^{\omega^{\omega^{\omega}}}, \ldots \ . \tag{1}$$

Recall that ω is defined as the limit of the (infinite) sequence $0, 1, 2, 3, \ldots$. In a similar spirit one introduces ε_0 as the limit of the sequence (1). We claim that the thus generated objects are a nice tool to measure the "strength" of a given termination order.

To make this connection clear, we introduce ordinals as order types. Instead of following the usual set-theoretic footsteps, we introduce ordinals as order types. That is a *countable ordinal* is defined as a *class* of equivalent linear, and well-founded partial orders. This suits our purpose, the emphasis of the applicability of ordinals as "universal scales" of well-founded orders, nicely. Together with the introduction of the class of ordinals On we will define a linear, well-founded order $<$ on On. Below we show that the ordinal ε_0 mentioned above only comprises relatively *weak* linear, well-founded ordinals. Actually the order $<$, when restricted to ordinals $< \varepsilon_0$ is (sloppily speaking) less expressive than a prominent instance of the recursive path orders, the *multiset path order* (MPO). This triggers the question how far up in the realm of ordinals we have to travel, until we reach a class α such that the restriction of $<$ to α is of equal strength as for example the *lexicographic path order* (LPO) on some ground term algebra $\mathcal{T}(\Sigma)$.

We show the surprising result that actually this α is an impressively large ordinal, which is in proof theory referred to as *small Veblen ordinal*, and henceforth denoted as Λ. On the other hand we show that MPO on some ground term algebra $\mathcal{T}(\Sigma)$ is of equal strength as the ordinal $\varphi(\omega, 0)$, where φ denotes the binary Veblen function. To relate this to object to the possible better known ordinal ε_0 note that ε_0 can be represented as $\varphi(1, 0)$ and is thus strictly smaller than $\varphi(\omega, 0)$. (For a formal definition of Λ and the binary Veblen function φ, see Section 3.4 in Section 3 below.)

Let us collect this result in a theorem. (We have not yet give a precise definition the notion of "strength" used in the formulation below. Actually we can give such a precise (re-)formulation, and will do so below, cf. Theorem 21.)

Theorem 1 (Dershowitz and Okada 1988). *The "strength" of the class of lexicographic path orders is equivalent to the "strength" of the ordinal Λ. The "strength" of the class of multiset path orders is equivalent to the "strength" of the ordinal $\varphi(\omega, 0)$. As $\varphi(\omega, 0) \lll \Lambda$, it is safe to say that lexicographic path orders are strictly "stronger" than multiset path orders.*

Furthermore the class of all simplification orders is also connected to Λ.

Theorem 2 (Schmidt 1979). *The "strength" of the class of* all *simplification orders is equivalent to the "strength" of the ordinal Λ.*

These result are not new, neither in proof theory, nor in term rewriting theory. Dershowitz and Okada [1988] were the first to observe the strong ties between term rewriting theory and proof theory which underlie (the proof of) this result. (Note that we use a different way to denote ordinals, hence our formulation of this Theorem is different form the formulation of in Dershowitz and Okada [1988].) We prove this Theorem in Section 3. By establishing this we believe that the course provides an alternative approach to ordinals that directly addresses one of the most important topics of computer science: Termination orders.

As soon as we have accomplished the quest for a "uniform measure" of the strength of termination orders we turn to the question of *complexity* of a given program P. In our setting this amounts to a characterization of the complexity of a given TRS R. More precisely we will study upper and lower bounds for the *derivation length* of R, the longest possible sequence of rewrite steps until a normal form is reached.

We construct terminating rewrite systems R that admit derivation sequences whose lengths easily capture the dimensions of the universe. This happens in Section 5. Before we can state this result we introduce so-called *Hardy functions* $\mathrm{H}_\alpha \colon \mathbb{N} \to \mathbb{N}$. These functions form a hierarchy $\bigcup_{\alpha<\Lambda} \mathrm{H}_\alpha$ of very fast growing functions. Indeed they grow so fast that it is safe to say that for some $\alpha < \Lambda$, $\mathrm{H}_\alpha(m)$ for small m already bounds the dimensions of the universe. (The formal definition of the Hardy hierarchy is given in Section 4.2 in Section 4.)

Theorem 3 (Weiermann 1994). *For every simply terminating TRS there exists $\alpha < \Lambda$ such that the complexity of the rewrite system is dominated by H_α.*

Theorem 4 (Lepper 2003). *For every $\alpha < \Lambda$ there exists a simply (and even totally) terminating TRS whose complexity eventually dominates H_α (and thus all H_β with $\beta \leqslant \alpha$).*

(For a definition of *simple* and *total* termination see Section 2 below.)

To some extent this result is bad news, as it implies that there exist rewrite systems R, whose termination can be shown by termination orders of strength comparable to the class of lexicographic path orders, such that the complexity of R is beyond comprehension. This sheds light on the hidden complexity of the lexicographic path order.

However, there are some good news, too.

Theorem 5 (Hofbauer 1992). *Termination via MPO implies the existence of a primitive recursive bound on the complexity of the rewrite system. This result is essentially optimal.*

Theorem 6 (Weiermann 1995). *Termination via LPO implies the existence of a multiple recursive bound on the complexity of the rewrite system. This result is essentially optimal.*

Theorem 7 (Lepper 2001a). *Suppose R is a TRS terminating via the Knuth–Bendix order* (KBO)*. Then the maximal number of rewrite steps possible in R starting with s is bound by* $\mathrm{Ack}(2^{O(n)}, 0)$*, where* $\mathrm{dp}(s) \leq n$ *holds and* Ack *denotes the binary Ackermann function. (We write* $\mathrm{dp}(s)$ *to denote the term depth of s.)*

In this course material we will only establish the first theorem, using a different proof method than employed in Hofbauer [1992], see Section 6. Our approach rests on the so-called *slow growing hierarchy* $\bigcup_{\alpha<\Lambda} \mathrm{G}_\alpha$, defined in Section 4.2 below. This proof method allows a uniform treatment of all three propositions. However, we will only hint on the necessary alternations in the proof.

This course material is organized as follows. In Section 2 definitions and results which will serve as a basis for everything we will consider later are fixed. Section 3 provides the basis and the proof of the Main Theorem. Building on the notions introduced in Section 3 we introduce the reader in Section 4 to the subrecursive trade. These two sections provide the basis for the investigations into the complexity of rewrite systems in Section 5 and Section 6.

Remark 1. We assume some familiarity with basic concepts from the realms of theoretical computer science (in particular term rewriting), logic, and mathematics. No prior knowledge of the theory of ordinals is required.

2 Terms, Rewriting, and such Stuff

This section contains definitions and results which will serve as a basis for everything we will consider later. We assume familiarity with basic concepts from the realms of theoretical computer science (in particular term rewriting), logic, and mathematics. However, we fix some notations. In the following we will frequently use logical symbols such as $\wedge, \vee, \neg, \rightarrow, \forall, \exists$ in definitions and on the meta-level. As usual the binary logical operators $\wedge, \vee, \rightarrow$ are written in infix notation. We suppose that these notions (and our sloppy use of it) are familiar.

2.1 Basic definitions

We write $A \subseteq B$ if A is a subset of B, while $A \subsetneq B$ indicates that A is a proper subset of B. The set theoretic *difference* $A \setminus B$ of A and B is $\{x \in A : x \notin B\}$. By $\mathrm{card}(W)$ we denote the *cardinality* (number of elements) of a set W. Nonnegative integers are called *natural numbers* and get collected in $\mathbb{N}$. For $\{n, n+1, \ldots, m\}$ we also write $[n, m]$. The (finite ordered) *tuple* of $a_1, \ldots, a_n$ is $(a_1, \ldots, a_n)$, its *length* $|(a_1, \ldots, a_n)|$ is n. If all a_i are located in a set A, then $(a_1, \ldots, a_n)$ is a *tuple over* A. The *Cartesian product* of the sets $A_1, \ldots, A_n$ is

$$A_1 \times \cdots \times A_n := \{(a_1, \ldots, a_n) : (\forall i \in [1, n])(a_i \in A_i)\} .$$

If all the A_i coincide with A, then we write A^n for this product. Note that A^0 contains the empty tuple $(\,)$. The set $A^* := \bigcup_{n \in \mathbb{N}} A^n$ contains the tuples over A. We will extend this notion to infinite sequences in Definition 22. The *disjoint union* of $A_1, \ldots, A_n$ is

$$\biguplus_{1 \leqslant i \leqslant n} A_i := \{(i, a) : 1 \leqslant i \leqslant n \wedge a \in A_i\} .$$

We assume the notions of (n-ary) *relation* (on some set A) and *function* (from A to B) to be known. For $n = 1, 2, 3$, these are called *unary*, *binary*, and *ternary*. If $\preceq$ is a binary relation, then we usually write $a \preceq b$ instead of $\preceq (a, b)$. It is common practice to write $p \npreceq q$ instead of $\neg(p \preceq q)$.

By $f\colon X \to Y$ we indicate that f is a function from X to Y. A function f is *number-theoretic* if we have $f\colon \mathbb{N}^k \to \mathbb{N}$ for some $k \in \mathbb{N}$. The set of functions from X to Y is denoted by ${}^X Y$. If $f\colon A \to B$ and $g\colon B \to C$, then $g \circ f$ denotes the function from A to C which maps a to $g(f(a))$. For $f\colon A \to A$, the n^{th} *iteration* $f^n\colon A \to A$ of f (applied to a) is given by

$$f^0(a) := a \quad \text{and} \quad f^{n+1}(a) := f(f^n(a)) . \tag{2}$$

We will make heavy use of this notation for functions of higher arities where all but one arguments are kept fixed. In these cases a "$\cdot$" indicates the free position. So, for example,

$$g(\cdot, b)^2(c) = g(g(\cdot, b)^1(c), b) = g(g(g(\cdot, b)^0(c), b), b) = g(g(c, b), b) .$$

A binary relation $\preceq$ on a set P is

- *reflexive* if $(\forall p \in P)(p \preceq p)$,
- *irreflexive* if $(\forall p \in P)(\neg p \preceq p)$,
- *transitive* if $(\forall p, q, r \in P)((p \preceq q \wedge q \preceq r) \to p \preceq r))$,
- *symmetrical* if $(\forall p, q \in P)(p \preceq q \to q \preceq p)$,
- *antisymmetrical* if $(\forall p, q \in P)((p \neq q \wedge p \preceq q) \to \neg q \preceq p)$,
- *linear* or *total* if $(\forall p, q \in P)(p \preceq q \vee p = q \vee q \preceq p)$.

A binary relation $\preceq$ is a *equivalence relation* if $\preceq$ is reflexive, transitive and symmetric. We call an ordered pair $(P, \preccurlyeq)$ where $\preccurlyeq$ is a binary reflexive and transitive relation on P a *preorder* or *quasiorder*. The *strict part* of $(P, \preccurlyeq)$ is $(P, \prec)$ with $p \prec q \;:\Longleftrightarrow\; p \preccurlyeq q \wedge q \not\preccurlyeq p$. A *partial order* is an antisymmetrical preorder. If $(P, \preccurlyeq)$ is a partial order, then we have $(\forall p, q \in P)(p \preccurlyeq q \iff p \prec q \vee p = q)$. Hence the strict part $(P, \prec)$ is irreflexive and transitive. On the other hand, if $(P, \lhd)$ is irreflexive and transitive, then $(P, \trianglelefteq)$ with $p \trianglelefteq q \;:\Longleftrightarrow\; p \lhd q \vee p = q$ is a partial order. Therefore we will frequently introduce a partial order $(P, \preccurlyeq)$ by displaying its irreflexive and transitive strict part $(P, \prec)$, and we will not hesitate to call $(P, \prec)$ a partial order. It is common practice to write $q \succ p$ for $p \prec q$.

Let $(P, \prec)$ be a partial order and $n \geqslant 1$. We say $F \colon P^n \to P$

– is *monotone in the i^{th} argument* (with $i \in [1, n]$) if

$$F(p_1, \ldots, p_n) \succ F(q_1, \ldots, q_n)$$

holds for all $\bar{p}, \bar{q} \in P$ which satisfy $p_i \succ q_i$ and $p_j = q_j$ for $j \neq i$,
– is *monotone* if it is monotone in all arguments,
– is *weakly monotone in the i^{th} argument* (with $i \in [1, n]$) if

$$F(p_1, \ldots, p_n) \succcurlyeq F(q_1, \ldots, q_n)$$

holds for all $\bar{p}, \bar{q} \in P$ which satisfy $p_i \succcurlyeq q_i$ and $p_j = q_j$ for $j \neq i$,
– is *weakly monotone* if it is weakly monotone in all arguments, and it
– has the *(weak) subterm property* if, for all $\bar{p} \in P$ and all i we have $F(\bar{p}) \succ p_i$ (respectively $F(\bar{p}) \succcurlyeq p_i$), such that $i \in [1, n]$.

Definition 1. *A mapping $o \colon P \to P'$ of the partial order $(P, \prec)$ into the partial order $(P', \prec')$ satisfying*

$$(\forall p, q \in P)(p \prec q \implies o(p) \prec' o(q))$$

is called an embedding*. If o even satisfies*

$$(\forall p, q \in P)(p \prec q \iff o(p) \prec' o(q)) \;,$$

then it is order-preserving*. Such a mapping is an* order isomorphism *if it is bijective.*

Definition 2. *Two partial orders are* equivalent *or* order isomorphic *if there exists an order isomorphism between them.*

Consider a partial order $(P, \prec)$ and $X \subseteq P$. We call $p \in P$

– a *minimum of* X if we have $p \in X$ and $(\forall q \in X)(q \not\prec p)$,
– an *upper bound of* X if we have $(\forall q \in X)(q \preccurlyeq p)$,
– the *least element of* X if we have $p \in X$ and $(\forall q \in X)(p \preccurlyeq q)$, and
– the *supremum of* X if it is the least of the upper bounds of X.

The dual notions are *maximum*, *lower bound*, *greatest element*, and *infimum*. Existence (and uniqueness) of either maximum, minimum, supremum, or infimum of X provided, we abbreviate it by $\max X$, $\min X$, $\sup X$, and $\inf X$. We call $Y \subseteq X$ *cofinal* (in X) if $(\forall p \in X)(\exists q \in Y)(p \preccurlyeq q)$ holds.

Definition 3. *A partial order $(P, \prec)$ is* well-founded *if every nonempty subset of P contains a minimum.*

We are going to introduce multisets and the basic operations on them. A multiset is quite like a finite set, but multiple appearances of its elements are counted. For a set A, a *multiset* over A is a function $M \colon A \to \mathbb{N}$ with finite $\{a \in A : M(a) \neq 0\}$. By $\mathsf{mul}(A)$ we denote the set of multisets over A.

We sometimes use the notation $\wr \ldots \int$ for multisets, so $\wr 0,0,1,2,2,2 \int$ represents the multiset M satisfying $(\forall n \geqslant 3)(M(n) = 0)$ and $M(0) = 2$, $M(1) = 1$, $M(2) = 3$. The empty multiset is the function mapping each element of A to 0, and it is ambiguously denoted by $\emptyset$. It will always be clear from the surrounding symbols if the multiset $\emptyset$ is meant. If $a \in A$ and M is a multiset, then we use $a \in M$ for $M(a) > 0$. The *union* of the multisets M and N is denoted by $M \cup N$ and satisfies

$$(\forall a \in A)((M \cup N)(a) = M(a) + N(a)) ,$$

while the notion of subset is transferable via

$$M \subseteq N \ :\Longleftrightarrow\ (\forall a \in A)(M(a) \leqslant N(a)) .$$

By $M \setminus N$ we denote the *difference* of the multisets M and N, which is defined by $(M \setminus N)(a) := M(a) \dot{-} N(a)$.

Definition 4. *If $(\mathsf{mul}(P), \prec_{\mathsf{mul}})$ is the multiset extension of $(P, \prec)$, then we have, for all $M, N \in \mathsf{mul}(P)$,*

$$M \prec_{\mathsf{mul}} N \iff M \neq N \wedge (\forall y \in M \setminus N)(\exists x \in N \setminus M)(y \prec x) .$$

Proposition 1.

i. *The multiset extension of a (linear) partial order is a (linear) partial order.*
ii. *The multiset extension of a well-founded partial order is a well-founded partial order.*

Definition 5.

– *The* lexicographic product *of partial orders $(P_i, \prec_i)$, $i \in [1, n]$, is defined as $(P_1 \times \cdots \times P_n, \prec_{\mathsf{lex}}^{1,n})$, where $(p_1, \ldots, p_n) \prec_{\mathsf{lex}}^{1,n} (q_1, \ldots, q_n)$ holds if*

$$(\exists i \in [1, n])(p_i \prec_i q_i \wedge (\forall j \in [1, i-1])(p_j = q_j)) .$$

– *If all $(P_k, \prec_k)$ coincide with $(P, \prec)$, we write $\prec_{\mathsf{lex}}^n$ for $\prec_{\mathsf{lex}}^{1,n}$ and call the resulting $(P^n, \prec_{\mathsf{lex}}^n)$ the n-*fold *lexicographic product.*
– *The* lexicographic order *$(P^*, \prec_{\mathsf{lex}}^*)$ based on a partial order $(P, \prec)$ is defined by*

$$p \prec_{\mathsf{lex}}^* q \ :\Longleftrightarrow\ |p| < |q| \vee (|p| = |q| \wedge p \prec_{\mathsf{lex}}^{|p|} q) .$$

If two sequences of equal lengths are considered, we will often write $<_{\mathsf{lex}}$ instead of $<_{\mathsf{lex}}^n$ or $\prec_{\mathsf{lex}}^*$.

Proposition 2.

i. *The lexicographic product of (linear) partial orders is a (linear) partial order.*
ii. *The lexicographic product of well-founded partial orders is a well-founded partial order.*
iii. *The lexicographic order based on a (linear) partial order is a (linear) partial order.*
iv. *The lexicographic order based on a well-founded partial order is a well-founded partial order.*

2.2 Our sets are term-sets

A *signature* Σ is a set of function symbols, such that each function symbol $f \in \Sigma$ has a unique *arity*, denoted as $\mathrm{ar}(f)$. The set of function symbols in Σ having arity n is denoted by $\Sigma^{(n)}$. The set of terms over Σ and the countably infinite set of variables $\mathcal{V}$ is denoted as $\mathcal{T}(\Sigma, \mathcal{V})$. If no confusion can arise, the reference to the signature Σ and the set of variables $\mathcal{V}$ is dropped.

Convention: To avoid trivialities we demand that whenever we deal with a specific signature Σ, that Σ is non-empty and contains at least one *constant*, i.e. a function symbol of arity 0. This convention is kept throughout this course material.

Convention: Throughout this text, natural numbers are denoted by lowercase Latin letters ranging from a to d and i to q, possibly extended by sub- or superscripts. Sometimes we also use uppercase versions of these letters; these are either used to denote various kinds of sets, or numbers which are supposed to be fixed throughout a section, section, etc. The common names of terms are s and t, but sometimes r and u show up as well. For variables we use x, y, and z, while constants are called c, e, or k. Functions or function symbols are represented by f, g, and h. We occasionally violate these conventions, provided that it seems appropriate to do so. When we do so, it will always clear from the context what is meant.

Finite sequences of similar objects are abbreviated using a bar, for example $\bar{s}$ is a shortcut for $s_1, \dots, s_n$, and $\bar{0}$ abbreviates $0, \dots, 0$. The length of such a sequence $s_1, \dots, s_n$ should always be clear from the context, otherwise it will be denoted as $|\bar{s}|$. Empty sequences are allowed and will occur soon. If we consider sequences of the same symbol, such as k consecutive occurrences of n, then we write n^k. This will not be mixed with exponentiation.

A term t is called *ground* or *closed* if $\mathrm{var}(t) = \emptyset$, where $\mathrm{var}(t)$ denotes the set of variables in t. The set of ground terms over Σ is denoted as $\mathcal{T}(\Sigma)$. With $\mathrm{dp}(s)$ we denote the *term depth* of s, defined as follows. Set $\mathrm{dp}(s) := 0$, if $s \in \mathcal{V}$, or $c \in \Sigma^{(0)}$, and otherwise

$$\mathrm{dp}(f(s_1, \dots, s_m)) := \max\{\mathrm{dp}(s_i) \colon 1 \leq i \leq m\} + 1 \ .$$

The size of a term t is the number of symbols in t, denoted as $\mathrm{Size}(t)$. A substitution $\sigma \colon \mathcal{V} \to \mathcal{T}$ is a mapping from the set of free variables to the set of terms. The application of a substitution σ to a term t is usually written as $t\sigma$ instead of

$\sigma(t)$. Let f be an arbitrary mapping. Then the *domain* of f is denoted as $\mathrm{dom}(f)$, while the *range* of f is written as $\mathrm{rg}(f)$. Extending this notion we call the set $\{x \colon \sigma(x) \neq x\}$ the *domain*—$\mathrm{dom}(\sigma)$—of σ and the set $\{\sigma(x) \colon a \in \mathrm{dom}(\sigma)\}$ the *range* of σ, denoted as $\mathrm{rg}(\sigma)$.

The concatenation of two substitution σ and λ (such that σ is applied before λ) is written as $\sigma \circ \lambda$. An expression term e_0 is an *instance* of e if $e\sigma = e_0$ for some substitution σ. Note that we have for any term s and substitutions σ, τ $(s\sigma)\tau = s(\sigma \circ \tau)$

Let $\preceq$ be a binary relation on $\mathcal{T}(\Sigma, \mathcal{X})$. We say that $\preceq$ is

i. *closed under contexts* (*monotone*)if we have, for all $f \in \Sigma$, $\mathrm{ar}(f) = n+1$, and for all terms $s, t, \bar{s}$, $s \preceq t$ implies $f(s_1, \ldots, s, \ldots, s_n) \preceq f(s_1, \ldots, t, \ldots, s_n)$,
ii. *closed under substitutions* (*stable*) if $s \preceq t$ implies $s\sigma \preceq t\sigma$ for all terms s, t and substitutions σ.

It is obvious how the above definition is altered with respect to the ground term algebra $\mathcal{T}(\Sigma)$.

2.3 Rewrite systems

Let Σ denote a finite signature. We recall (very briefly) basic notions and concepts in term rewriting theory.

Definition 6. *A* term rewriting system *(or* rewrite system*) R over $\mathcal{T}$ is a finite set of rewrite rules (l, r). The* rewrite relation $\to_R$ *on $\mathcal{T}$ is the least binary relation on $\mathcal{T}$ containing R such that*

i. *if $s \to_R t$ and σ a substitution, then $s\sigma \to_R t\sigma$ holds, and*
ii. *if $s \to_R t$, then $f(\ldots, s, \ldots) \to_R f(\ldots, t, \ldots)$.*

When we focus on a TRS R, we will sometimes drop the subscript R. If (A) is a (named) rule from R, then $\to_{\mathrm{A}}$ indicates a rewrite step due to this rule. As usual we introduce the relation $\xrightarrow{*}$ to denote the transitive closure of $\to$. On the other hand we write $\xrightarrow{+}$ to indicate that at least one rewrite step has been performed.

Definition 7. *Let R be a TRS over Σ. A term t is* in normal form *(with respect to R) if there is no term r satisfying $t \to_R r$, and t is* a normal form *of s if $s \xrightarrow{*}_R t$ and t is in normal form.*

Definition 8. *A signature Σ which solely consists of unary symbols and exactly one constant is* monadic.

We call a TRS over a monadic signature a string rewrite system (SRS) *if the only constant of the signature does not occur in the rules. Since in this context terms may be identified with strings over the alphabet $\Sigma^{(1)}$, we will display them as strings, hence we drop the parentheses and leave out the constant resp. variable. In addition we will usually not mention the constant when introducing a monadic signature.*

Definition 9. *A rewrite system R is* terminating *if there is no infinite sequence $\langle t_i : i \in \mathbb{N} \rangle$ of terms such that*

$$t_0 \to_R t_1 \to_R \cdots \to_R t_m \to_R \cdots$$

Definition 10. *A* (nondeterministic) Turing machine (TM) *$\mathcal{M}$ is determined by*

- *a finite alphabet $\Gamma = \{a_0, \ldots, a_n\}$, whose elements are called* letters*, where a_0 is also called the* blank*,*
- *a finite set $Q = \{q_0, \ldots, q_p\}$ of* states*, and*
- *a* transition relation $\Delta \subseteq Q \times \Gamma \times Q \times \Gamma \times \{\mathrm{L}, \mathrm{R}\}$.

A configuration *consists of*

- *the contents of an (bidirectionally) infinite* tape *which contains a countable infinitude of* cells*. Each cell contains a letter (from Γ), and only finitely many cells do not contain the blank a_0.*
- *the current state and*
- *the current position (cell number) of an imagined read-write-head.*

If K and K' are configurations, then the machine $\mathcal{M}$ is able to make a transition *from K to K', abbreviated by $K \rightsquigarrow K'$, if there is $(q, a, q', a', d) \in \Delta$ such that in K the machine is in state q, reading an a, while in K' it is in state q' and the tape of K' emerged from the tape of K by replacing the a just read in by an a' and moving the head one position to the left (if $\mathrm{d} = \mathrm{L}$) or to the right (if $\mathrm{d} = \mathrm{R}$). We say that $\mathcal{M}$* halts *for the configuration K if there is no infinite $\rightsquigarrow$-descending sequence starting with K.*

It should be clear that any TRS can be simulated by a TM. But how about the opposite direction?

The *uniform halting problem* is the following problem:

Given: a TM $\mathcal{M}$.
Question: Does $\mathcal{M}$ halt for all configurations?

Theorem 8. *The uniform halting problem is undecidable.*

We want to show next that it is possible to transform any TM $\mathcal{M}$ into a corresponding SRS $RS_{\mathcal{M}}$ in such a way that $\mathcal{M}$ halts for all configurations if and only if $RS_{\mathcal{M}}$ terminates. Thus the uniform halting problem is equivalent to the question of termination of TRSs.

Theorem 9 (Huet and Lankford 1978). *Termination is undecidable for SRSs (and hence for TRSs).*

Proof. We describe the construction of the SRS $R_{\mathcal{M}}$. The first problem is to cope with the infinity of the tape. But since only finitely many cells of the tape are not blank, it mainly suffices to introduce symbols for the left and right edge of the tape and to treat anything beyond these symbols as blank.

The main problem is to prevent the TRS from doing reductions that are not intended. For this aim, we have to make more explicit which symbols are positioned left to the head and which symbols are not. This is achieved by introducing, for each letter a, symbols $\hat{a}$ and a. Our signature consists of

- symbols a and $\hat{a}$ for each $a \in \Gamma$,
- a symbol q for each state $q \in Q$, and
- the symbols $\triangleright$ and $\triangleleft$ for the left resp. right edge of the tape.

A *configuration term* is any term t of the shape

$$\triangleright \hat{a}_{i_k} \dots \hat{a}_{i_1} q a_{j_1} \dots a_{j_h} \triangleleft$$

with $q \in Q$, $0 \leqslant h, k$ and $0 \leqslant i_1, \dots, i_k, j_1, \dots, j_h \leqslant n$. Any such t describes exactly one configuration K_t, which looks like this:

$$\begin{array}{c|c|c|c|c|c|c|c}
\hline
\dots & a_{i_k} & \dots & a_{i_1} & a_{j_1} & \dots & a_{j_h} & \dots \\
\hline
\multicolumn{4}{c}{} & \uparrow q & \multicolumn{3}{c}{}
\end{array}$$

Here the two outer dots indicate infinitely many blanks to both edges of the tape. Of course, each configuration is represented by infinitely many configuration terms.

Now each possible transition is turned into a few rules of our SRS $R_{\mathcal{M}}$, depending on the fact that we have to cope with the symbols representing the edges. For each transition $(q, a_i, q', a_j, \mathrm{R}) \in \Delta$ we add to $R_{\mathcal{M}}$ the rule $q a_i \to \hat{a}_j q'$. If $i = 0$ then we also add $q \triangleleft \to \hat{a}_j q' \triangleleft$. For each transition $(q, a_i, q', a_j, \mathrm{L}) \in \Delta$ we add $\triangleright q a_i \to \triangleright q' a_0 a_j$ and, for any $k \in \{1, \dots, n\}$, we also add $\hat{a}_k q a_i \to q' a_k a_j$ to $R_{\mathcal{M}}$. If $i = 0$ then we also have to add $\triangleright q \triangleleft \to \triangleright q' a_0 \triangleleft$ and, for any $k \in \{1, \dots, n\}$, we also add $\hat{a}_k q \triangleleft \to q' a_k a_j \triangleleft$. Thus $R_{\mathcal{M}}$ is finite.

It is obvious that for any configuration term t, $t \to t'$ implies that t' is a configuration term satisfying $K_t \rightsquigarrow K_{t'}$. On the other hand, for two configurations K and K' with $K \rightsquigarrow K'$, we find, for any configuration term t with $K_t = K$, a configuration term t' satisfying $t \to t'$ and $K_{t'} = K'$.

Thus an infinite computation of $\mathcal{M}$ induces an infinite reduction of $R_{\mathcal{M}}$. Unfortunately, the other direction is not that simple, since there may be infinite reductions starting from terms which are not configuration terms. Yet it suffices to show that any infinite reduction induces an infinite reduction starting with a configuration term. Any term t can be written as $u_1 v_1 \dots u_d v_d u_{d+1}$ such that no u_i contains a state q, the u_i may be empty and each v_i has the shape $\hat{a}_{i_k} \dots \hat{a}_{i_1} q a_{j_1} \dots a_{j_h}$ with $0 \leqslant h, k$, and it is maximal with respect to this property, that is, u_i does not end with some $\hat{a}_i$ and u_{i+1} does not start with some a_j. Now, any reduction from t requires some element of Q, thus it can not take place in an u_i. Furthermore, the u_i are left untouched (well, possibly $\triangleright$ may be replaced by $\triangleright$, but this does not count). Thus $t \to t'$ implies there is $m \in \{1, \dots, d\}$ and v'_m of the same shape as the other v_j such that $t' = u_1 v_1 \dots u_m v'_m u_{m+1} \dots u_d v_d u_{d+1}$, and additionally we have $\triangleright v_m \triangleleft \to \triangleright v'_m \triangleleft$. If there is an infinite reduction starting from t, then there is an m such that infinitely many steps of this reduction take place in $u_m v_m u_{m+1}$, thus there is an infinite reduction starting from $\triangleright v_m \triangleleft$, and this induces there is also an infinite reduction starting from a configuration term (since we may drop any superfluous $\triangleright$ and $\triangleleft$).

In the above construction the number of rules contained in R depends on the number of instructions of the $\mathcal{M}$. Dershowitz [1987b] showed that termination of

TRSs containing only two rules is already undecidable, and later Dauchet [1992] (see also Lescanne [1994]) found a way to transform TMs into equivalent TRSs containing only one rule. Thus termination of TRSs containing exactly one rule is already undecidable.

2.4 Termination proofs

A general method to prove termination of a TRS is the construction of a well-founded order on the terms which contains the rewrite relation. Another is the construction of an interpretation from the (ground) term algebra into a well-founded order. In both cases any infinite derivation corresponds to an infinite descending chain in the well-founded order, so we get termination.

Definition 11. *Let $\mathcal{X}$ be equal to $\mathcal{V}$ or $\emptyset$. A partial order $(\mathcal{T}(\Sigma, \mathcal{X}), \prec)$*

- *is a* rewrite order *if it is closed under contexts and substitutions,*
- *is a* reduction order *if it is a well-founded rewrite order,*
- *is* compatible *with a TRS R if $\to_R \subseteq \succ$, that is, if $s \to_R t$ implies $s \succ t$,*
- normalizes *a TRS R if $l\sigma \succ r\sigma$ holds for all rules $l \to r \in R$ and all ground substitutions σ, and it*
- *has the* (weak) subterm property *if $s \succ t$ (respectively $s \succcurlyeq t$) holds as soon as t is a proper subterm of s.*

Interpretations are a canonical means to prove termination. We recall (very briefly) the essential concepts and propositions.

Definition 12. *Let $(P, \prec)$ be a partial order.*

- *An* interpretation *of $\mathcal{T}(\Sigma)$ in $(P, \prec)$ is a mapping $\mathcal{I} \colon \mathcal{T}(\Sigma) \to P$.*
- *The interpretation $\mathcal{I}$ induces a binary relation $\prec_{\mathcal{I}}$ on $\mathcal{T}(\Sigma)$ via $s \prec_{\mathcal{I}} t \;:\Longleftrightarrow \mathcal{I}(s) \prec \mathcal{I}(t)$. This can be lifted to a binary relation on $\mathcal{T}(\Sigma, \mathcal{V})$ via*

$$\begin{aligned} s \prec_{\mathcal{I}} t \;&:\Longleftrightarrow \mathcal{I}(s\sigma) \prec \mathcal{I}(t\sigma) \text{ for all ground substitutions } \sigma \\ &\Longleftrightarrow s\sigma \prec_{\mathcal{I}} t\sigma \text{ for all ground substitutions } \sigma \;. \end{aligned}$$

Note that $\prec_{\mathcal{I}}$ is not total on $\mathcal{T}(\Sigma, \mathcal{V})$ because distinct variables are always incomparable. Even if $(P, \prec)$ is total, $\prec_{\mathcal{I}}$ need not be total on $\mathcal{T}(\Sigma)$ since distinct terms may be mapped to the same member of P.

Obviously $\mathcal{I}$ is an order preserving mapping from $(\mathcal{T}(\Sigma), \prec_{\mathcal{I}})$ to $(P, \prec)$, and $(\mathcal{T}(\Sigma, \mathcal{V}), \prec_{\mathcal{I}})$ is a partial order which is well-founded if $(P, \prec)$ is.

Let $\mathcal{I}$ be an interpretation of $\mathcal{T}(\Sigma)$ in the partial order $(P, \prec)$. We say that

- $\mathcal{I}$ is *monotone* if, for all $f \in \Sigma$, $\mathrm{ar}(f) = n + 1$, and all $s, t, \bar{s} \in \mathcal{T}(\Sigma)$, we have $\mathcal{I}(s) \succ \mathcal{I}(t)$ implies $\mathcal{I}(f(s_1, \ldots, s, \ldots, s_n)) \succ \mathcal{I}(f(s_1, \ldots, t, \ldots, s_n))$,
- $\mathcal{I}$ is *weakly monotone* if, for all $f, s, t, \bar{s}$, we get $\mathcal{I}(f(s_1, \ldots, s, \ldots, s_n)) \succcurlyeq \mathcal{I}(f(s_1, \ldots, t, \ldots, s_n))$ whenever $\mathcal{I}(s) \succcurlyeq \mathcal{I}(t)$ holds,
- $\mathcal{I}$ has the *subterm property* if $\mathcal{I}(f(s_1, \ldots, s_n)) \succ \mathcal{I}(s_i)$ holds for all $f \in \Sigma$, $\mathrm{ar}(f) \geq 1$, all $\bar{s} \in \mathcal{T}(\Sigma)$, and all $i \in [1, n]$, and

– $\mathcal{I}$ is a *normalization* of a TRS R if, for all rules $l \to r$ in R and for all ground substitutions σ, we have $\mathcal{I}(l\sigma) \succ \mathcal{I}(r\sigma)$.

Proposition 3. *Let $\mathcal{I}$ be an interpretation of $\mathcal{T}(\Sigma)$ in the partial order $(P, \prec)$, and let R be a TRS over Σ.*

i. *If $\mathcal{I}$ is monotone, then $(\mathcal{T}(\Sigma, \mathcal{V}), \prec_{\mathcal{I}})$ is a rewrite order. Additional well-foundedness of $(P, \prec)$ implies that $(\mathcal{T}(\Sigma, \mathcal{V}), \prec_{\mathcal{I}})$ is a reduction order.*
ii. *If $\mathcal{I}$ has the subterm property, then so does $(\mathcal{T}(\Sigma, \mathcal{V}), \prec_{\mathcal{I}})$.*
iii. *$\mathcal{I}$ normalizes R if and only if $R \subseteq \succ_{\mathcal{I}}$, i.e. if $\succ_{\mathcal{I}}$ normalizes R.*
iv. *If $\mathcal{I}$ is a monotone normalization of R, then $\mathcal{I}$ embeds $(\mathcal{T}(\Sigma), \overset{+}{\leftarrow}_R)$ into $(P, \prec)$, and R is compatible with $(\mathcal{T}(\Sigma, \mathcal{V}), \prec_{\mathcal{I}})$. Additional well-foundedness of $(P, \prec)$ implies termination of R.*

Usually, interpretations are defined in a *homomorphic* manner. For this it suffices to consider a partial order $(P, \prec)$ and, for each symbol f of the signature, an *interpreting function* $[f]$ on P of appropriate arity. Then a mapping $[\![\cdot]\!]$ of $\mathcal{T}(\Sigma)$ in $(P, \prec)$ is defined by recursion on $\mathcal{T}(\Sigma)$ via

$$[\![f(t_1, \dots, t_n)]\!] := [f]([\![t_1]\!], \dots, [\![t_n]\!]) \; . \tag{3}$$

Lemma 1. *A homomorphic mapping of* (3) *is (weakly) monotone if each of the interpreting functions is (weakly) monotone, and it has the subterm property if each of the interpreting functions has the subterm property.*

An interesting part of terminating TRSs is covered by the following concept.

Definition 13 (Ferreira 1995, 5.40). *A TRS R is* totally terminating *if there exists a well-order $(\mathcal{T}(\Sigma), \prec)$ which is compatible with R.*

Obviously, total termination implies termination.

Proposition 4. *A TRS R is totally terminating if and only if there exists an interpretation $\mathcal{I}$ of $\mathcal{T}(\Sigma)$ in a well-order $(P, \prec)$ such that $\mathcal{I}$*

– *is monotone and*
– *normalizes R.*

A major drawback of the previous characterization is the need to find everywhere (strictly) monotone interpretations, as such interpretations are usually rather hard to find, whereas weakly monotone interpretations are often much easier to construct. It turns out that weak monotonicity suffices, provided that we join it with the subterm property. This result will be of central importance in Section 5.

Theorem 10 (Touzet 1998). *A TRS R is totally terminating if there exists an interpretation $\mathcal{I}$ of $\mathcal{T}(\Sigma)$ in a well-order $(P, \prec)$ such that $\mathcal{I}$*

– *is weakly monotone,*
– *has the subterm property, and*

– *normalizes* R.

Dershowitz [1979] introduced *simplification orders*, which are rewrite orders having the subterm property. We call a TRS *simplifying* if it is compatible with a simplification order.

Theorem 11 (Kruskal's Tree Theorem). *Let Σ be a finite signature. Every simplification order $\prec$ on $\mathcal{T}$ is well-founded. Hence any simplifying TRS is terminating.*

Based on the Theorem, we coin the expression *simply terminating* to denote a TRS that is simplifying.

Remark 2. Recall that there are terminating TRSs which are not simplifying, as the SRS $ff \to fgf$ shows. It terminates because each rewrite step decreases the number of occurrences of ff, yet for any compatible simplification order $\prec$ we would get $ff \succ fgf \succ ff$.

Proposition 5 (Zantema 1994). *Total termination implies simple termination.*

In the sequel we recall some outstanding simplification orders. In particular we state the definitions of *multiset path orders* and *lexicographic path orders*.

We define a binary relation $\sim$ on $\mathcal{T}(\Sigma, \mathcal{V})$ by $s \sim t$ if $s = f(s_1, \dots, s_n)$, $t = f(t_1, \dots, t_n)$, and there is a permutation π of $[1, n]$ such that $(\forall i \in [1, n])(s_i \sim t_{\pi(i)})$. The relation is called *permutative equivalence*. Let Σ be a signature equipped with a precedence $\succ$. To simplify notation we assume $\succ$ is total and $\Sigma = \{f_1, \dots, f_N\}$, such that N denotes the cardinality of Σ. Furthermore suppose $i < j \to f_i \prec f_j$. The *multiset path order* (MPO) $\succ_{\mathsf{mpo}}$ (based on Σ and $\succ$) of $\mathcal{T}(\Sigma, \mathcal{V})$ is defined as follows.

Definition 14 (Plaisted 1978, Dershowitz 1982). $s \succ_{\mathsf{mpo}} t$ *iff*

i. $t \in var(s)$ *and* $s \neq t$, *or*
ii. $s = f_j(s_1, \dots, s_m)$, $t = f_i(t_1, \dots, t_n)$, *and*
  – *there exists* k $(1 \leq k \leq m)$ *with* $s_k \succ_{\mathsf{mpo}} t \vee s_k \sim t$, *or*
  – $j > i$ *and* $s \succ_{\mathsf{mpo}} t_l$ *for all* $l = 1, \dots, n$, *or*
  – $i = j$ *and* $\wr s_1, \dots, s_m \int \succ^{\mathsf{mul}}_{\mathsf{mpo}} \wr t_1, \dots, t_m \int$.

Let Σ be a signature equipped with a precedence $\succ$ as above. The *lexicographic path order* (LPO) $\succ_{\mathsf{lpo}}$ (based on Σ and $\succ$) of $\mathcal{T}(\Sigma, \mathcal{V})$ is defined as follows.

Definition 15. $s \succ_{\mathsf{lpo}} t$ *iff*

i. $t \in var(s)$ *and* $s \neq t$, *or*
ii. $s = f_j(s_1, \dots, s_m)$, $t = f_i(t_1, \dots, t_n)$, *and*
  – *there exists* k $(1 \leq k \leq m)$ *with* $s_k \succeq_{\mathsf{lpo}} t$, *or*
  – $j > i$ *and* $s \succ_{\mathsf{lpo}} t_l$ *for all* $l = 1, \dots, n$, *or*
  – $i = j$ *and* $s \succ_{\mathsf{lpo}} t_l$ *for all* $l = 1, \dots, n$, *and there exists an* i_0 $(1 \leq i_0 \leq m)$ *such that* $s_1 = t_1, \dots s_{i_0 - 1} = t_{i_0 - 1}$ *and* $s_{i_0} \succ_{\mathsf{lpo}} t_{i_0}$.

As an exercise we consider the (usual) binary Ackermann function, as a TRS R.

$$\begin{aligned} \mathrm{Ack}(0,m) &\to m+1 \;, \\ \mathrm{Ack}(n+1,0) &\to \mathrm{Ack}(n,1) \;, \text{ and} \\ \mathrm{Ack}(n+1,m+1) &\to \mathrm{Ack}(n,\mathrm{Ack}(n+1,m)) \;. \end{aligned}$$

Exercise 1. Define a suitable signature Σ and a (total) precedence $\succ$ on it, such that $\succ_{\mathsf{lpo}}$ (based on Σ and $\succ$) serves as a reduction order for R. Thus R is simply terminating.

Exercise 2. (+) Show that no multiset path order $\succ_{\mathsf{mpo}}$ is capable of establishing (simple) termination of R.

The following result is a folklore result in term rewriting theory.

Theorem 12 (Kamin and Lévy [1980]).

i. *If $s \succ_{\mathsf{lpo}} t$ ($s \succ_{\mathsf{mpo}} t$) then $var(t) \subseteq var(s)$.*
ii. *For any total order $\prec$ on Σ, the induced LPO $\succ_{\mathsf{lpo}}$ (MPO $\succ_{\mathsf{mpo}}$) is a simplification order on $\mathcal{T}$.*
iii. *If R is a TRS such that $\to_R$ is contained in an LPO (MPO), then R is terminating.*
iv. *Termination via LPO (MPO) implies total termination.*

Finally we want to mention Knuth–Bendix orders (KBOs), which were introduced by Knuth and Bendix [1970]. We will not define KBOs in detail, just recall the main idea. For further reading see Lankford [1979], Dershowitz [1987b], and Baader and Nipkow [1998]. The signature is equipped with a precedence and a *weight function*, which associates a nonnegative real number with each symbol. One extends this function to a weight on terms by adding the weights of the symbols. Very roughly speaking, terms are compared by first comparing their weights, then their root symbols, and finally, by recursion, their subterms. Under certain additional assumptions this results in a simplification order.

Remark 3. Note that termination via KBO is incomparable with both termination via LPO or MPO. The SRS $\{fg \to ggf\}$ (from Ferreira [1995, p. 114], see also Middeldorp and Zantema [1997, p, 148]) is not terminating via KBO as weight considerations imply $g \succ f$, whereas precedence considerations demand $f \succ g$. In contrast to this, we get termination via LPO (or MPO) using $f \succ g$. On the other hand, the SRS $\{fg \to gff\}$ terminates via KBO, but neither MPO nor LPO are able to cope with this rule.

This incomparability result shows that we cannot directly relate syntactic orders as LPO, MPO, or KBO. But still we would be interested in given a characterization of such orders that allows us the separate *strong* orders from *weak* ones. The next section will show how such a *uniform measure* can be established.

3 A Uniform Measure

As already mentioned the emphasis of this course lies on the applicability of ordinals as "universal scales" of well-founded orders and in the end we aim at a canonical measure of termination orders, such that e.g. a precise characterization of the "strength" of well-known orders such as LPO and MPO is rendered. To this avail it seems appropriate to study ordinals and ordinal notation systems to some extent. This is the purpose of this section.

3.1 Countable ordinals

In this section we introduce *countable ordinals*. Instead of following the usual set-theoretic footsteps, we introduce ordinals as order types. The central aim of this section is the introduction of the collection On of countable ordinals, together with a well-founded relation $<$ on On. Our presentation follows Pohlers [1996, pp. 41–45].

Definition 16. *Let A be a set; let $\preceq$ be a binary relation on A. We define the* field *of the relation $\preceq$, denoted as $\mathsf{Field}(\preceq)$, as follows.*

$$\mathsf{Field}(\preceq) := \{a\colon \exists x \ a \preceq x \ \textit{or} \ x \preceq a\}$$

The relation $\preceq$ is called a *partial order* if $(\mathsf{Field}(\preceq), \preceq)$ is a partially ordered set. A relation $\preceq$ is called a *total order* if $\preceq$ is a partial order which is linear. Recall that for a partial order $\preceq$ we denote its strict predicate by $\prec$. I.e. $p \prec q \iff p \preceq q \wedge p \neq q$. As above, we call $\prec$ a *strict partial order*. The relation $\preceq$ is called *well-founded* if every nonempty subset of $\mathsf{Field}(\preceq)$ has a least element. More formally

$$\forall A \, [A \subseteq \mathsf{Field}(\preceq) \wedge A \neq \emptyset \to \exists x \in A \forall y (y \preceq a \to (a = y \vee y \notin A))]$$

Definition 17. *Let $\preceq$ denote a linear partial order. Then $\preceq$ is called a* well-order *if $\preceq$ is well-founded.*

Two total orders $\leq$, $\preceq$ are *equivalent* if there is an embedding map o form $\mathsf{Field}(\leq)$ onto $\mathsf{Field}(\preceq)$. I.e. we have a surjective function o such that

$$\forall x, y \in \mathsf{Field}(\leq) \, (x \prec y \to o(x) \prec o(y)) \ ,$$

holds. We write $\leq \sim \preceq$ to denote the equivalence of $\leq$ and $\preceq$.

Exercise 3. In Definition 2, we have stated what should be understood by the equivalence of two partial orders. Prove that $\leq$ and $\preceq$ are equivalent as defined above if $(\mathsf{Field}(\leq), \leq)$ and $(\mathsf{Field}(\preceq), \preceq)$ are equivalent (in the old sense).

In term rewriting one introduces the principle of *well-founded induction* as a generalization of mathematical induction to any terminating TRS $(R, \xrightarrow{+})$: Let A

denote some property of elements of P. We formalize well-founded induction as a rule.

$$\frac{\forall x\,(\forall y \in P(x \to_R y \to A(y)) \to A(x))}{\forall x \in P\ A(x)}$$

We have only stated this rule to clarify the analogy to the principle of *transfinite induction*, defined below. Let $\preceq$ denote a well-founded relation, and assume A denotes some property of the elements of $\mathsf{Field}(\preceq)$. Consider the following rule.

Definition 18. *(Principle of Transfinite Induction.)*

$$\frac{\forall x\,(\forall y(y \prec x \to A(y)) \to A(x))}{\forall x A(x)}$$

Note that this induction principle does not have an explicite base case. This may come as a surprise. However, let p denote the $\preceq$-minimal element. Then consider the premise $\forall y(y \prec p \to A(y))$ of the induction hypothesis more closely. Obviously this premise is trivially true and hence the induction hypothesis subsumes the base case.

Proof. First note that the hypothesis entails $A(p)$ whenever $p \notin \mathsf{Field}(\preceq)$. Now we proceed indirectly. Assume the hypothesis is true and suppose the existence of an element $p \in \mathsf{Field}(\preceq)$ such that $\neg A(p)$. The set of all such elements is a subset of $\mathsf{Field}(\preceq)$. We may assume p is $\preceq$-minimal, as $\preceq$ is well-founded. Therefore for all y such that $y \prec p$ the property $A(y)$ holds. This however is the premise of the induction hypothesis. Whence $A(p)$ holds. Contradiction.

Exercise 4. The equivalence of orders is an equivalence relation.

Definition 19. *A* countable ordinal *is the equivalence class of a* countable *well-order.*

Convention: As the only ordinals we will deal here with will be countable, we will frequently drop the "countable" and simply speak of an *ordinal*.

We will use lower-case Greek letters, possibly extended by super- or subscripts, to denote ordinals. The collection of ordinals is denoted by On.

Definition 20. *A* segment *of a total order $\preceq$ is a subset M of the field of $\preceq$ such that M is an* initial part *of* $\mathsf{Field}(\preceq)$*. Put more formally*

$$\forall x \in M \forall y \in \mathsf{Field}(\preceq)(y \preceq x \to y \in M)$$

A segment M is called proper *if it is distinct from* $\mathsf{Field}(\preceq)$*.*

Let $p \in \mathsf{Field}(\preceq)$ be given. Then p induces a segment M of $\preceq$ as follows.

$$M := \{(x, y) \colon x \preceq y \wedge y \prec p\}$$

The segment M is denoted as $\preceq(p)$. Clearly this segment is a proper segment. We will use the following abbreviation. $\mathsf{Field}(\preceq(p)) := \{x \in \mathsf{Field}(\preceq) \colon x \prec p\}$. Finally we are able to define a binary relation $<$ on On. It will turn out that $<$ is a well-order itself.

Definition 21. *Let $\alpha, \beta \in \mathsf{On}$. Let the well-order $\preceq_1$ represent α and $\preceq_2$ represents β. Then $\alpha < \beta$ if there exists $z \in \mathsf{Field}(\prec)$ such that $\leq \sim \prec(z)$.*

Lemma 2. *Let the relation $<$ be defined as above, then $<$ is well-defined, irreflexive and transitive.*

Proof. Firstly we have to convince ourselves that the above given definition is indeed well-defined. Let $\leq_1, \leq_2, \preceq_1, \preceq_2$ denote (countable) well-orders such that $\leq_1 \sim \leq_2$ and $\preceq_1 \sim \preceq_2$ holds. Assume $\tau, \sigma \in \mathsf{On}$ such that τ is represented by $\leq_i$ $(i = 1, 2)$ and σ is represented by $\preceq_i$ $(i = 1, 2)$. Finally assume the existence of $z \in \mathsf{Field}(\preceq_1)$ such that $\leq_1 \sim \preceq_1(z)$. To show well-definedness it suffices to show the existence of $z \in \mathsf{Field}(\preceq_2)$ such that $\leq_2 \sim \preceq_2(z)$. This however follows easily by concatenation of the induced order-preserving functions.

Secondly, let us consider irreflexivity. The proof of irreflexivity is simplified if we first consider the following claim, which is left as an exercise.

Exercise 5. Let $\preceq$ denote a well-founded relation. Assume an embedding o from $\mathsf{Field}(\preceq)$ to $\mathsf{Field}(\preceq)$. Show that $\forall x(o(x) \succeq x)$.

Now we proceed indirectly. Assume $\sigma < \sigma$ holds for some $\sigma \in \mathsf{On}$. Let $\preceq$ denote a well-order representing σ. By definition there exists $z \in \mathsf{Field}(\preceq)$ s.t. $\preceq \sim \preceq(z)$ that is there exists an embedding o mapping $\mathsf{Field}(\preceq)$ onto $\mathsf{Field}(\preceq(z))$. Taking the claim for granted, we conclude that $\forall x(o(x) \succeq x)$ holds. In particular, if q denotes a witness for the existential quantifier in $\exists z(\preceq \sim \preceq(z))$, then $o(q) \succeq q$. However, by definition of o, $o(q) \prec q$ holds. Contradiction.

Finally, we consider transitivity. However, this case follows quite as the first part. Hence the proof is omitted and left to the reader.

By now we have defined a strict partial order $<$ on On. The following result tells us that this order $<$ is itself a well-order.

Theorem 13. *Let the relation $<$ be defined as above, then $<$ is an irreflexive well-order on On.*

Proof. First we show totality of the order $>$. Let $\preceq_1, \preceq_2$ denote representations of $\tau, \sigma \in \mathsf{On}$, respectively. Suppose $\tau \neq \sigma$.

We have to show the existence of $z \in \mathsf{Field}(\preceq_1)$ (or $z \in \mathsf{Field}(\preceq_2)$) such that $\preceq_1(z) \sim \preceq_2$ (or $\preceq_1 \sim \preceq_2(z)$). We define

$$o(x) := \min_{\preceq_2}\{z \in \mathsf{Field}(\preceq_2) \colon \forall y \prec_1 x \; o(y) \prec_2 z\}$$

Obviously o is an order-preserving map from $\mathsf{Field}(\preceq_1)$ to $\mathsf{Field}(\preceq_2)$. It is easy to see that $\mathrm{dom}(o)$ and $\mathrm{rg}(o)$ are segments of $\mathsf{Field}(\preceq_1)$, $\mathsf{Field}(\preceq_2)$, respectively. By assumption $\preceq_1 \not\sim \preceq_2$. Thus at least one of $\mathrm{dom}(o)$ or $\mathrm{rg}(o)$ has to be a proper segment. W.l.o.g. we assume the latter. Then set $z := \min\{z \in \mathsf{Field}(\preceq_2) \colon z \in \mathrm{rg}(o)\}$. Now it is easy to see that

$$\preceq_1 \sim \preceq_2(z)$$

holds. Thus linearity is established.

Finally, we have to convince ourselves that $<$ is well-founded. Assume there exists a non-empty $M \subseteq \mathsf{On}$, such that M does not admit a $<$-minimal element. In particular assume $\sigma \in M$, then there exists $\tau \in M$ and $\tau < \sigma$. Let $\preceq$ denote a representation of σ. By definition of $<$ there exists $z \in \mathsf{Field}(\preceq)$ such that $\tau \sim {\preceq}(z)$. By iterating this procedure and employing transitivity of $<$, we obtain an infinite sequence of elements of $\mathsf{Field}(\preceq)$.

$$z = z_1, z_2, \ldots, z_n, z_{n+1}, \ldots$$

A moment of reflection shows that this sequence is ordered by $\succ$ in decreasing order. Thus we have constructed an infinite $\prec$-descending sequence. Contrary to the assumption that $\preceq$ is a (countable) well-order.

Exercise 6. A binary relation $\preceq$ is well-founded if and only if there are no infinite $<$-descending sequences.

Hint: Use the ideas employed in the proof of the totality of $<$.

We are going to extend finite tuples to infinite sequences. This is achieved by identifying the tuples over A we already know with functions from finite ordinals into A.

Definition 22. *A* sequence *(over a nonempty set A) is a function whose domain is an ordinal α (and whose range is a subset of A). This α is called the* length *of the sequence. Such a sequence is usually displayed as $(a_\iota)_{\iota<\alpha}$. The sequence is* finite *if α is finite, and otherwise it is* infinite*. If β is an ordinal and A is a nonempty set, then $A^{<\beta}$ denotes the set of sequences over A having length below β.*

Note that there is an obvious isomorphism between $A^{<\omega}$ and A^*, as defined in Section 2.1. We write $\mathrm{Seq}(x)$ to state that x is a sequence, and then $|x|$ denotes its length.

Theorem 14. *If $\alpha \in \mathsf{On}$, then there exists a $\beta \in \mathsf{On}$, such that $\alpha < \beta$. We also say that the collection of countable ordinals On is* unbounded.

Proof. Let $\sigma \in \mathsf{On}$ and $\prec$ a well-order representing σ. Put

$$\begin{aligned} x \prec' y := \mathrm{Seq}(x) \land \mathrm{Seq}(y) \land |x| = |y| = 2 \land \big[&((x)_0 = 0 \land (y)_0 = 0 \land (x)_1 \prec (y)_1) \\ &\lor ((x)_0 = 0 \land (x)_1 \in \mathsf{Field}(\prec) \land (y)_0 = 1 \land (y)_1 = 1)\big] \end{aligned}$$

Hence we have extended $\prec$ by a single point $(1,1)$ at the end and therefore obtain ${\prec} \sim {\prec'}((1,1))$. Note that $(x)_0$, $(x)_1$ denote the first, resp. second projection of the pair x.

Due to the above theorem, the next definition is well-defined.

Definition 23. *Let α be an ordinal. The* successor *of α is defined as*

$$\min\{\zeta \colon \alpha < \zeta\}$$

We write $\alpha + 1$ to denote the successor of α.

The least element of On is denoted as 0. An ordinal is a *successor ordinal* if it can be written in the form $\alpha + 1$. An ordinal which is neither 0 nor an successor ordinal is called a *limit ordinal*. The collection of limit ordinals is denoted as Lim.

Definition 24. *Assume $\succ$ is well-founded on A. We associate an* ordinal *in* On *to each element a of A.*

$$\mathrm{otype}_{\succ}(a) := \sup\{\mathrm{otype}_{\succ}(b) + 1 \colon b \in A \text{ and } a \succ b\} \quad .$$

The order type *of $\succ$, abbreviated by* $\mathrm{otype}(\succ)$*, is the supremum of* $\mathrm{otype}_{\succ}(a)+1$ *for $a \in A$.*

Sometimes we write $\mathrm{otype}(\prec)$ in place of $\mathrm{otype}(\succ)$. Note that the just given definition ought to be considered with some care. Let $a \in A$ be given. Then $\mathrm{otype}_{\succ}(a)$ is an ordinal, i.e. a class of well-orders. Note that we have not defined what 'sup' means when applied to a set of classes. Hence this definition is (strictly speaking) flawed. However, note that this is a consequence of our approach to ordinals as order-types. In the 'usual' set-theoretic setting, where ordinals are (transitive) sets this problem doesn't occur. One solution to the obstacle is to employ notions from category theory. To be more precise 'sup' should be read as a 'direct limit', cf. MacLane [1998]. We need not go into further details. The same resort to category theory can be employed in similar situations below.

Ordinals whose field is finite, are called *finite* and otherwise *infinite* Based on this we introduce the following convention. *Convention*: In the following we do not distinguish between ordinals α, $0 < \alpha < \omega$, and natural numbers. Instead we use the expressions $n \in \mathbb{N}$ and $n < \omega$ synonymously.

If $(P, \prec)$ is a well-order, then $\mathrm{otype}^{P}_{\prec}$ denotes the order isomorphism of Definition 24.

Definition 25. *Let $(P, \prec)$ be a well-order. The inverse function of* $\mathrm{otype}^{P}_{\prec}$ *is called the* enumerating function *of $(P, \prec)$ and is denoted by* $\mathrm{enum}^{P}_{\prec}$.

Definition 26. *Let $F \colon \mathsf{On} \to \mathsf{On}$ be given. We say*

- *F is* continuous *if it satisfies*

$$(\forall \lambda \in \mathsf{Lim})(F(\lambda) = \sup\{F(\alpha) : \alpha < \lambda\}) \ ,$$

- *F is* normal *if it is a continuous embedding, and*
- *α is a* fixed point *of F if $F(\alpha) = \alpha$ holds.*

Lemma 3. *A normal function has arbitrarily large fixed points.*

Proof. Let F be a normal function and pick an arbitrary β. We recursively define $G \colon \omega \to \mathsf{On}$ by $G(0) := \beta$ and $G(n+1) := F(G(n))$. Our intention is to show $\alpha := \sup\{G(n) : n < \omega\}$ is a fixed point. From Exercise 5 we infer $G(n+1) = F(G(n)) \geqslant G(n)$, and $\alpha \geqslant \beta$ follows. If there is an n satisfying

$G(n+1) = G(n)$, then we get $\alpha = G(n) = F(G(n))$, and otherwise α is a limit ordinal and we have

$$F(\alpha) = F(\sup_{n<\omega} G(n)) = \sup_{n<\omega} F(G(n)) = \sup_{n<\omega} G(n+1) = \alpha$$

since F is continuous.

3.2 The arithmetic of ordinals

Previously we introduced the principle of transfinite induction. Now we turn to *transfinite recursion.* The idea is that we wish to define a (class) function $F(\sigma)$ in terms of σ and the values of $F(\tau)$ for ordinals $\tau \prec \sigma$. Adding parameters, we arrive at the following situation. We have defined a function G, and wish to define F so that

$$F(\sigma, s_1, \ldots, s_n) = G(\sigma, (F(\tau, \overline{s}) \colon \tau < \sigma), s_1, \ldots, s_n) \tag{4}$$

The principle of transfinite recursion is provable within the framework of Set Theory, see e.g. Shoenfield [1967]. However, we will not prove it, but simply treat it as an axiom. We call (4) a definition of F by the *principle of transfinite recursion.*

A very common way of defining ordinal functions is by distinguishing between the three kinds of ordinals. This procedure is legalized by the principle of transfinite recursion.

Proposition 6. *For functions* $G \colon \mathsf{On} \to \mathsf{On}$ *and* $H \colon \mathsf{On}^2 \to \mathsf{On}$ *there exists* $F \colon \mathsf{On}^2 \to \mathsf{On}$ *satisfying*

$$F(\alpha,\beta) = \begin{cases} G(\alpha) & \text{if } \beta = 0 \ , \\ H(F(\alpha,\beta'),\alpha) & \text{if } \beta = \beta' + 1 \ , \\ \sup\{F(\alpha,\beta') : \beta' < \beta\} & \text{if } \beta \in \mathsf{Lim} \ . \end{cases}$$

Definition 27. *The (binary)* ordinal addition $\alpha + \beta$ *is defined by*

$$\alpha + \beta := \begin{cases} \alpha & \text{if } \beta = 0 \ , \\ (\alpha + \beta') + 1 & \text{if } \beta = \beta' + 1 \ , \\ \sup\{\alpha + \beta' : \beta' < \beta\} & \text{if } \beta \in \mathsf{Lim} \ . \end{cases}$$

Likewise, ordinal multiplication $\alpha \cdot \beta$ *is generated with*

$$\alpha \cdot \beta := \begin{cases} 0 & \text{if } \beta = 0 \ , \\ (\alpha \cdot \beta') + \alpha & \text{if } \beta = \beta' + 1 \ , \\ \sup\{\alpha \cdot \beta' : \beta' < \beta\} & \text{if } \beta \in \mathsf{Lim} \ . \end{cases}$$

Finally, ordinal exponentiation α^β *is given by*

$$\alpha^\beta := \begin{cases} 1 & \text{if } \beta = 0 \ , \\ \alpha^{\beta'} \cdot \alpha & \text{if } \beta = \beta' + 1 \ , \\ \sup\{\alpha^{\beta'} : 0 < \beta' < \beta\} & \text{if } \beta \in \mathsf{Lim} \ . \end{cases}$$

It is easy to see that these functions extend the usual functions on natural numbers to ordinals.

Exercise 7. Let α, β, γ be ordinals.

i. Ordinal addition is associative, but for $\alpha \geqslant \omega$ we have $1 + \alpha = \alpha < \alpha + 1$.
ii. Ordinal multiplication is associative, but $2 \cdot \omega = \omega < \omega + \omega = \omega \cdot 2$.
iii. We have $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$, but $(\omega + 1) \cdot 2 = \omega \cdot 2 + 1 < \omega \cdot 2 + 2$.
iv. We have $\alpha^\beta \cdot \alpha^\gamma = \alpha^{\beta+\gamma}$.
v. The ordinal functions $\delta \mapsto \alpha + \delta$ (with arbitrary α), $\delta \mapsto \alpha \cdot \delta$ (with $\alpha > 0$), and $\delta \mapsto \alpha^\delta$ (with $\alpha > 1$) are normal.

By the above Exercise 7.v and Lemma 3, there are arbitrarily large ordinals λ satisfying $\lambda = \omega^\lambda$. Similarly, there are arbitrarily large λ which are closed under addition, i.e. they satisfy $(\forall \alpha, \beta < \lambda)(\alpha + \beta < \lambda)$.

Definition 28.

– *Ordinals $\lambda > 0$ which are closed under addition are called* principal ordinals*. They are collected in* H.*
– *Ordinals λ satisfying $\lambda = \omega^\lambda$ are called* epsilons*. The α^{th} epsilon number is called ε_α.*
– *We introduce* ω-towers *ω_n by $\omega_0 := 1$ and $\omega_{n+1} := \omega^{\omega_n}$.*

Lemma 4.

i. *The enumerating function of H is $\alpha \mapsto \omega^\alpha$.*
ii. *We have $\varepsilon_0 = \sup \{\omega_n : n \in \omega\}$.*

Proposition 7. *For every ordinal α there are uniquely determined principal ordinals $\alpha_1 \geqslant \ldots \geqslant \alpha_n$ such that $\alpha = \alpha_1 + \cdots + \alpha_n$ holds. This is called the* additive normal form *of α, and we sometimes write $\alpha =_{\mathrm{NF}} \alpha_1 + \cdots + \alpha_n$.*

Note that the sum may be empty, yielding 0. We may combine the Proposition with Lemma 4.i.

Corollary 1. *For every ordinal α there are uniquely determined ordinals $\alpha_1 \geqslant \ldots \geqslant \alpha_n$ such that $\alpha = \omega^{\alpha_1} + \cdots + \omega^{\alpha_n}$ holds. This is called the* Cantor normal form *of α, and we sometimes write $\alpha =_{\mathrm{CNF}} \omega^{\alpha_1} + \cdots + \omega^{\alpha_n}$.*

Based on the additive normal form, it is possible to define an alternative ordinal addition which is associative and commutative.

Definition 29. *The* natural sum *$\alpha \oplus \beta$ of two ordinals $\alpha =_{\mathrm{NF}} \gamma_1 + \cdots + \gamma_n$ and $\beta =_{\mathrm{NF}} \gamma_{n+1} + \cdots + \gamma_{n+m}$ is given by $\alpha \oplus \beta := \gamma_{\pi(1)} + \cdots + \gamma_{\pi(n+m)}$, where π is any permutation of $[1, n+m]$ with $(\forall i \in [1, n+m-1])(\gamma_{\pi(i)} \geqslant \gamma_{\pi(i+1)})$.*

Exercise 8. The natural sum is commutative, associative, and monotone.

---
* H stems from the original German notion *Hauptzahlen*.

The operation $*$ *concatenates* two sequences $a = (a_\iota)_{\iota<\alpha}$ and $b = (b_\iota)_{\iota<\beta}$. By $a * b$ we denote the sequence $c = (c_\iota)_{\iota<\alpha+\beta}$ satisfying

$$c_\iota = \begin{cases} a_\iota & \text{if } \iota < \alpha \ , \\ b_\xi & \text{if } \iota = \alpha + \xi \ . \end{cases}$$

We say $a' = (a'_\iota)_{\iota<\alpha'}$ is an *extension* of $a = (a_\iota)_{\iota<\alpha}$, abbreviated by $a \leqslant_{\mathsf{ext}} a'$, if $\alpha \leqslant \alpha'$ and $(\forall \iota < \alpha)(a_\iota = a'_\iota)$.

Exercise 9. The class of sequences is partially ordered by $\leqslant_{\mathsf{ext}}$. We have $a \leqslant_{\mathsf{ext}} a'$ if and only if there is a sequence b with $a' = a * b$.

The *concatenation* of partial orders $(P_1, \prec_1), \ldots, (P_n, \prec_n)$ is $(\biguplus_{1\leqslant i\leqslant n} P_i, \prec^{1,n})$ with

$$(i,p) \prec^{1,n} (j,q) \ :\Longleftrightarrow \ i < j \vee (i = j \wedge p \prec_i q) \ .$$

Exercise 10. i. The concatenation of (linear) partial orders is a (linear) partial order.
ii. The concatenation of well-founded partial orders is a well-founded partial order, and its order type is the sum of the order types of the basic orders.
iii. The concatenation of well-orders is a well-order.

Dershowitz and Manna [1979] transmogrified the properties of principal ordinals into a well-order which does not refer to ordinals.

Definition 30. *The* multiset extension *of a given partial order* $(P, \prec)$ *is defined as* $(\mathsf{mul}(P), \prec_{\mathsf{mul}})$ *with*

$$\begin{aligned} M \prec_{\mathsf{mul}} N \ :\Longleftrightarrow \ (\exists X, Y \in \mathsf{mul}(P))(&\emptyset \neq X \subseteq N \\ &\wedge M = (N \setminus X) \cup Y \\ &\wedge (\forall y \in Y)(\exists x \in X)(y \prec x)) \ . \end{aligned}$$

Exercise 11. Previous, in Definition 4, we have given an alternative formulation of the multiset extension order. Show the equivalence of the two.

Exercise 12. The multiset extension of a well-order is a well-order.

A proof of the following Theorem can be extracted, for example, from Weiermann [1992], or from Ferreira [1995, Remark 5.22].

Theorem 15. *If* $(P, \prec)$ *is a well-founded partial order, then we have*

$$\operatorname{otype}(\mathsf{mul}(P), \prec_{\mathsf{mul}}) = \omega^{\operatorname{otype}(P,\prec)} \ .$$

Recall the definition of *lexicographic product* and *lexicographic order*, cf. Definition 5. We have the following propositions, whose proofs we leave as an exercise.

Exercise 13. i. The order type of the lexicographic product of well-founded partial orders is the reverse product of the order types of the basic orders:

$$\operatorname{otype}(P_1 \times \cdots \times P_n, \prec_{\mathsf{lex}}^{1,n}) = \operatorname{otype}(P_n, \prec_n) \cdot \cdots \cdot \operatorname{otype}(P_1, \prec_1) \ .$$

ii. The lexicographic product of well-orders is a well-order.
iii. The lexicographic order based on a well-order is a well-order.

Lemma 5. *Let* $(P, \prec)$ *be well-founded with* $\mathrm{otype}(P, \prec) \geqslant \omega$. *Then it holds that* $\mathrm{otype}(P^*, \prec^*_{\mathsf{lex}}) = \mathrm{otype}(P, \prec)^\omega$.

Proof. We put $\alpha := \mathrm{otype}(P, \prec) \geqslant \omega$. Exercise 13.i implies $\mathrm{otype}(P^n, \prec^n_{\mathsf{lex}}) = \alpha^n$. Because of Exercise 7 we have

$$\alpha^n + \alpha^{n+1} = \alpha^n \cdot 1 + \alpha^n \cdot \alpha = \alpha^n \cdot (1 + \alpha) = \alpha^n \cdot \alpha = \alpha^{n+1}$$

for all n. As $(P^*, \prec^*_{\mathsf{lex}})$ corresponds to the infinite concatenation of the $(P^n, \prec^n_{\mathsf{lex}})$, and $\mathrm{otype}(P^n, \prec^n_{\mathsf{lex}}) = \alpha^n$, we reach

$$\begin{aligned} \mathrm{otype}(P^*, \prec^*_{\mathsf{lex}}) &= \sup_{n<\omega} \big(\mathrm{otype}(P^0, \prec^0_{\mathsf{lex}}) + \cdots + \mathrm{otype}(P^n, \prec^n_{\mathsf{lex}}) \big) \\ &= \sup_{n<\omega} \big(\alpha^0 + \cdots + \alpha^n \big) = \sup_{n<\omega} \alpha^n = \alpha^\omega \ , \end{aligned}$$

using Exercise 10.ii and the definition of ordinal exponentiation.

Theorem 16. *We have* $\mathrm{otype}(\mathbb{N}^*, <^*_{\mathsf{lex}}) = \omega^\omega$.

3.3 The ordinal ε_0

In the previous section we have defined certain (normal) functions on the collection On. In particular ordinal exponentation $\omega^{\cdot}$ was defined. In Definition 28 we introduced names for the *fix-points* of this function, the *epsilons*. The main purpose of this section is the study of the first of this fix-points, called ε_0. More precisely we want to study the initial segment $<(\varepsilon_0)$ of the well-order $<$ upto ε_0. (To simplify notation below we sometimes write $<$ to denote $<(\varepsilon_0)$.) We believe that this ordinal is rather prominent. In proof theory it is well-known as the *proof theoretic ordinal* of Peano Arithmetic, and even the reader not (or not at all) familiar with proof theory may have already had contact with ordinals $< \varepsilon_0$ of the form ω_k ($k \in \mathbb{N}$), the so-called ω-towers.

As a central result of this section we will show that the class of well-orders ε_0 contains only well-orders that are rather *weak*. Especially, if they are to be compared with recursive path orders as e.g. an instance $\prec_{\mathsf{mpo}}$ of the multiset path orders. To accomplish this, we will explicitely write down a well-order $(E, \prec)$ such that $\mathrm{otype}\, E(\prec)$ equals ε_0.

This order will serve several purposes. First of all it will show that the ordinal ε_0 actually exists, i.e. the class ε_0 is non-void. Upto now, we can only be sure that the class ω^ω in non-void, cf. Theorem 16, and note that $\omega^\omega < \sup\{\omega_k \colon k \in \mathbb{N}\} = \varepsilon_0$. Secondly, the well-order $\prec$ will be defined through a *recursive* definition. This definition eases the comparisons between two ordinals $\alpha, \beta \in \mathsf{On}(\varepsilon_0)$ to some extent. Thirdly, the proof of well-foundedness of $\prec$ will show that $\mathrm{otype}\, E(\prec) \leq \mathrm{otype}\, \mathcal{T}(\Sigma)(\prec_{\mathsf{mpo}})$ (for some multiset path ordering $\prec_{\mathsf{mpo}}$ over the ground term algebra based on some suitable chosen signature Σ). This renders the relative weakness of ε_0. Finally, the well-order $(E, \prec)$ serves as a basis for further reaching well-orders $(T, \prec)$, defined in a similar way. A specific extension is defined in Section 3.4 to pin-down the precise order type of any multiset path order and any lexicographic path order.

It should not come as a surprise that our order $(E, \prec)$ will be defined as an order over notations for ordinals. Sometimes such an ordered pair $(E, \prec)$ is called an *(ordinal) notation system* for the segment of ordinals $\mathsf{On}(\varepsilon_0)$. The first part consists in the construction of an appropriate notation system for the ordinals less than ε_0 together with the order $\prec$ on it. (Here we follow the approach of Takeuti [1987].) It will become clear in the construction that $\mathrm{otype}(\prec) = \varepsilon_0$. After this is accomplished, we prove well-foundedness of the order $\prec$. This completes the argument, as the construction of this notation system is done carefully, such that irreflexivity and linearity are almost trivial by construction. We forget about the previous definition of the collection On, and the operations defined over elements of On.

Definition 31. *Recursive definition of a set E of ordinal terms and a subset P of E.*

i. $0 \in E$
ii. *If $\alpha_1, \dots, \alpha_m \in P$, then $\alpha_1 + \cdots + \alpha_m \in E$.*
iii. *If $\alpha \in E$, then $\omega^\alpha \in P$, and $\omega^\alpha \in E$.*
iv. *Only those objects are in E that have been obtained through one of the above clauses.*

The last line in the definition makes sure that all objects in E will have either the form $\alpha_1 + \cdots + \alpha_m$ for some number m and objects $\alpha_1, \dots, \alpha_m \in P$ or the form ω^α, $\alpha \in E$. It is important to note that the symbols 0, +, and ω are arbitarily chosen objects. It may be convenient to conceive these symbols as *constructors* instead of function symbols.

To simplify reading we will abbreviate the object ω^0 by 1. It follows from the definition of the set E that any object in E different from 0 can be represented in the following form.

$$\omega^{\alpha_1} + \omega^{\alpha_2} + \cdots + \omega^{\alpha_n} \ , \tag{5}$$

where each of the $\alpha_1, \alpha_2, \dots, \alpha_n$ is $\neq 0$ and has the same property. Note that for each $i = 1, \dots, n$ holds $\omega^{\alpha_i} \in P$. Now we can define the relation $=$ and $\prec$ on E.

Definition 32. *Recursive definition of the relation $=$ and $\prec$, and the binary function $+$, simultaneously.*

i. 0 *is the minimal element of $\prec$.*
ii. *If $\alpha \prec \beta$, then $\omega^\alpha \prec \omega^\beta$, and vice versa.*
iii. *Let $\alpha \in E$ contain an occurrence of* 0 *but not* 0 *itself, and let $\beta \in E$ be defined by removing this occurrence of* 0 *(as well as excessive occurrences of $+$) from α. Then $\alpha = \beta$.*
iv. *Let α, β be of the form $\omega^{\alpha_1} + \cdots + \omega^{\alpha_m}$, and $\omega^{\beta_1} + \cdots + \omega^{\beta_n}$, respectively. Then $\alpha + \beta$ is defined as*

$$\omega^{\alpha_1} + \cdots + \omega^{\alpha_m} + \omega^{\beta_1} + \cdots + \omega^{\beta_n} \ .$$

v. *Let $\alpha \in E$ be written in the form (5) and contain two consecutive terms ω^{α_i} and $\omega^{\alpha_{i+1}}$ with $\alpha_i \prec \alpha_{i+1}$. That is α has the form*

$$\cdots + \omega^{\alpha_i} + \omega^{\alpha_{i+1}} + \ldots \ .$$

Let $\beta \in E$ be obtained by removing the string "$\omega^{\alpha_i} +$ " from α, so that β is of the form

$$\cdots + \omega^{\alpha_{i+1}} + \ldots \ .$$

Then $\alpha = \beta$.

vi. *Suppose $\alpha = \omega^{\alpha_1} + \cdots + \omega^{\alpha_m}$, and $\beta = \omega^{\beta_1} + \cdots + \omega^{\beta_n}$, and suppose furthermore that $\alpha_1 \succeq \alpha_2 \succeq \cdots \succeq \alpha_m$, and $\beta_1 \succeq \beta_2 \succeq \cdots \succeq \beta_n$, holds. ($\alpha \succeq \beta$ means $(\beta \prec \alpha) \vee (\beta = \alpha)$.) Then $\alpha \prec \beta$ iff $\omega^{\alpha_i} \prec \omega^{\beta_i}$ for some i ($1 \leq i \leq m$) and for all $j = 1, \ldots, i-1$ $\omega^{\alpha_j} = \omega^{\beta_j}$ holds.*

It follows from (v) that any object $\alpha \in E$ can be uniquely represented in the form

$$\omega^{\alpha_1} + \omega^{\alpha_2} + \cdots + \omega^{\alpha_n} \ , \tag{6}$$

where $\alpha_1 \succeq \alpha_2 \succeq \cdots \succeq \alpha_m$ holds. If α is written in this way, we say that α is in *normal form*. This normal form is unique as the same holds for any α_i that is used in the construction of the normal form.

Exercise 14. The relation $\prec$ is irreflexive, linear, and transitive.

Exercise 15. In Takeuti [1987] multiplication of ordinals is introduced as follows. Suppose α is of the form (6). Let $\beta > 0$. Then $\alpha \cdot \omega^{\beta} = \omega^{\alpha_1 + \beta}$ Use Definition 27 to show that this (counter-intuitive) definition is correct. Note $\beta > 0$.

Now we are in the position to prove the well-foundedness of $(E, \prec)$. One way to establish this result would be to follow the approach by Takeuti [1987, pp. 99–101]. There the notion of *accessibility* is introduced and employed to show the well-foundedness of $(E, \prec)$. Another way would be to exploit Krukal's Tree Theorem, another standard technique in showing the well-foundedness of a given (partial) order. However, it is more to the spirit of this course to follow another (slightly more compact) approach.

Lemma 6. *$(E, \prec)$ is well-founded.*

Proof. Let $\Sigma := \{0, p, w\}$, where $\mathrm{ar}(0) = 0$, $\mathrm{ar}(p) = 2$, and $\mathrm{ar}(w) = 1$, and let $>$ denote the total precedence $w > p > 0$ on Σ. We write $\prec_{\mathsf{mpo}}$ for the multiset path order induced by $<$. We will define a mapping $o \colon (E, \prec) \to (\mathcal{T}(\Sigma), \prec_{\mathsf{mpo}})$ such that for all $\alpha, \beta \in E$, $\alpha \prec \beta \to o(\alpha) \prec_{\mathsf{mpo}} o(\beta)$ holds. This will imply that o is an embedding of $(E, \prec)$ to $(\mathcal{T}(\Sigma), \prec_{\mathsf{mpo}})$. Thus any infinite descending $\prec$-sequence in E gives rise to an infinite descending $\prec_{\mathsf{mpo}}$-sequence in $\mathcal{T}(\Sigma)$. By Theorem 12 we conclude that $(\mathcal{T}(\Sigma), \prec_{\mathsf{mpo}})$ is well-founded. This will establish the proof that $\prec$ is well-founded.

DEFINITION OF o: Let $\alpha \in E$ be given. Due to Definition 32.v, α can be written in the form (6). Thus it suffices to recursively define o on this normal form: $o(0) := 0$, $o(\omega^{\alpha}) := w(o(\alpha))$, and finally

$$o(\alpha_1 + \cdots + \alpha_n) := p(o(\alpha_1), o(\alpha_2 + \cdots + \alpha_n)) ,$$

where $n \geq 2$ and $\alpha_1 \succeq \alpha_2 \succeq \cdots \succeq \alpha_m$ holds. To prove $\alpha \prec \beta \to o(\alpha) \prec_{\mathsf{mpo}} o(\beta)$, we proceed by induction on $\mathrm{Size}(\beta)$. We abbreviate (here and henceforth) induction hypothesis by (ih).

Firstly we assume $\beta = \omega^{\beta_0}$. Let α ($\neq 0$) be written as (6). Observe that by definition of $\prec$ we have for all $i = 1, \ldots, n$ that $\alpha_i \prec \beta_0$ holds. Note that $\alpha \prec \beta$ implies $\omega^{\alpha_1} \prec \omega^{\beta_0}$ (Definition 32.vi). Which in turn implies $\alpha_1 \prec \beta_0$ (Definition 32.ii). Furthermore, by definition of α, $\alpha_1 \succeq \cdots \succeq \alpha_n$. By transitivity of $\prec$ the assertion follows. As $\mathrm{Size}(\beta_0) < \mathrm{Size}(\beta)$ (ih) is applicable to derive $o(\alpha_i) \prec_{\mathsf{mpo}} o(\beta_0)$ for all i. By definition of $\prec_{\mathsf{mpo}}$ we obtain for all $i = 1, \ldots, n$

$$o(\omega^{\alpha_i}) = w(o(\alpha_i)) \prec_{\mathsf{mpo}} w(o(\beta_0)) = o(\omega^{\beta_0}) .$$

We set $\gamma_i := \omega^{\alpha_i}$, and consider $o(\alpha)$

$$p(o(\gamma_1), p(o(\gamma_2), \cdots, p(o(\gamma_{n-1}), o(\gamma_n)) \cdots)) .$$

As $w > p$ we need to show

$$o(\gamma_1) \prec_{\mathsf{mpo}} o(\omega^{\beta_0}) \quad \text{and} \quad p(o(\gamma_2), \cdots, p(o(\gamma_{n-1}), o(\gamma_n)) \cdots) \prec_{\mathsf{mpo}} o(\omega^{\beta_0}) .$$

Obviously the first inequality follows by (ih). We iterate the argument to conclude $o(\alpha) \prec_{\mathsf{mpo}} o(\omega^{\beta_0}) = o(\beta)$.

Secondly, we consider the case where β equals $\omega^{\beta_1} + \omega^{\beta_2} + \cdots + \omega^{\beta_m}$, where $m \geq 2$ and $n > m$. We suppose α ($\neq 0$) is denoted as above. Hence, by definition there exists $i \in [1, n]$ such that $\alpha_i \prec \beta_i$. By a similar observation as before, we conclude that $\forall j \geq i (\alpha_j \prec \beta_i)$. Thus by (ih) we conclude $o(\alpha_1) = o(\beta_1), \ldots, o(\alpha_{i-1}) = o(\beta_{i-1})$, and $o(\alpha_i), \ldots, o(\alpha_n) \prec_{\mathsf{mpo}} o(\beta_i)$. We set $\delta_i := \omega^{\beta_i}$, consider $o(\beta)$

$$p(o(\delta_1), p(o(\delta_2), \cdots, p(o(\delta_{m-1}), o(\delta_m)) \cdots)) .$$

Note that $o(\gamma_j) = o(\delta_j)$ for all $j = 1, \ldots, i-1$. To show $o(\alpha) \prec_{\mathsf{mpo}} o(\beta)$ it thus suffices to show

$$\begin{aligned} &\wr o(\gamma_i), p(o(\gamma_{i+1}), \cdots, p(o(\gamma_{n-1}), o(\gamma_n)) \cdots) \int \prec_{\mathsf{mpo}}^{\mathsf{mul}} \\ &\qquad \prec_{\mathsf{mpo}}^{\mathsf{mul}} \wr o(\delta_i), p(o(\delta_{i+1}), \cdots, p(o(\delta_{m-1}), o(\delta_m)) \cdots) \int . \end{aligned}$$

This follows by the same reasoning as in the first subcase. Thus we conclude $o(\alpha) \prec_{\mathsf{mpo}} o(\beta)$.

Thirdly, we consider the case, where α, β are defined as above, but $m > n$. Then $o(\alpha) \prec_{\mathsf{mpo}} o(\beta)$ follows by one application of Subterm clause in the definition of $\prec_{\mathsf{mpo}}$. Finally, consider the case where $\alpha = 0$. Then $o(\alpha) = 0 \prec_{\mathsf{mpo}} o(\beta)$ for arbitrary $\beta \succ 0$ follows due to the setting of the precedence $w > p > 0$. This completes the argument.

Theorem 17. *The binary relation $\prec$ is an irreflexive well-order of order type ε_0.*

Proof. We take Exercise 14 for granted. Thus the irreflexivity, transitivity, and linearity of $\prec$ has already been established. Due to Lemma 6 the relation $\prec$ is well-founded. Thus $\prec$ is a well-order. It remains to show that $\prec$ is equivalent to the initial segment $<(\varepsilon_0)$. However, a close look on Definition 21, Lemma 4, Corollary 1, and Definition 32 suffices to see how an appropriate embedding from $\mathsf{Field}(\prec)$ onto $\mathsf{Field}(<)$ has to be defined.

Convention: In the following we will no longer distinguish between the order $<$ on $\mathsf{On}(\varepsilon_0)$ and the order $\prec$ on notations of ordinals E.

3.4 The small Veblen ordinal

Let K be an arbitrary natural number ≥ 2, kept fixed for the rest of this section. Based on the ordinal notation system $(E, <)$ introduced in the last section, we define a set of terms $T(K)$ (and a subset $P \subset T(K)$) together with a well-order $<$ on $T(K)$. The elements of $T(K)$ are built from 0, $+$ and the K-ary function symbol ψ.

The notation system $(E, <)$ was based on properties of the well-order $<(\varepsilon_0)$ introduced in Section 3.1. Contrary, we introduce the notation system $(T(K), <)$ as a somehow direct extension of $(E, <)$. The unary function symbol $\omega^{\cdot}$ is replaced by the K-ary function symbol ψ and the binary relation $<$ is changed to mimic this extension. As above note that the elements of $T(K)$ are *terms*, not ordinals. These ordinal terms can and will serve as representations of an initial segment of the set of ordinals On. However, we will give their *interpretation* only after we have finished the presentation of the partial ordered set $(T(K), <)$ and have established that $(T(K), <)$ is a well-order.

Definition 33. *Let $K \geq 2$ be given. Recursive definition of a set $T(K)$ of ordinal terms, a subset $P \subset T(K)$, and a binary relation $>$ on $T(K)$.*

i. $0 \in T(K)$.
ii. *If $\alpha_1, \dots, \alpha_m \in P$ and $\alpha_1 \geq \dots \geq \alpha_m$, then $\alpha_1 + \dots + \alpha_m \in T(K)$.*
iii. *If $\alpha_1, \dots, \alpha_K \in T(K)$, then $\psi(\alpha_1, \dots, \alpha_K) \in P$ and $\psi(\alpha_1, \dots, \alpha_K) \in T(K)$.*
iv. *$\alpha \neq 0$ implies $\alpha > 0$.*
v. *$\alpha > \beta_1, \dots, \beta_m$ and $\alpha \in P$ implies $\alpha > \beta_1 + \dots + \beta_m$.*
vi. *Let $\alpha = \alpha_1 + \dots + \alpha_m$, $\beta = \beta_1 + \dots + \beta_n$. Then $\alpha > \beta$ iff*
  - *$m > n$, and for all i ($i \in \{1, \dots, n\}$) $\alpha_i = \beta_i$, or*
  - *there exists i ($i \in \{1, \dots, m\}$) such that $\alpha_1 = \beta_1, \dots, \alpha_{i-1} = \beta_{i-1}$, and $\alpha_i > \beta_i$.*

vii. *Let $\alpha = \psi(\alpha_1, \dots, \alpha_K)$, $\beta = \psi(\beta_1, \dots, \beta_K)$. Then $\alpha > \beta$ iff*
  - *there exists k ($1 \leq k \leq K$) with $\alpha_i \geq \beta$, or*
  - *$\alpha > \beta_l$ for all $l = 1, \dots, K$ and there exists an i_0 ($1 \leq i_0 \leq K$) such that $\alpha_1 = \beta_1, \dots \alpha_{i_0-1} = \beta_{i_0-1}$ and $\alpha_{i_0} > \beta_{i_0}$.*

Remark 4. – Note that, as in Section 3.3, the set $P \in T(K)$ corresponds to the set of additive principal numbers H in On. We will elaborate on the interpretation of the K-ary functions ψ below.
– Note the close correspondence between the definition of $<$ on terms of the form $\psi(\overline{\alpha})$ and $\psi(\overline{\beta})$ and the way terms are compared in lexicographic path orders. In particular compare Definition 15.ii in Section 2.4 and Definition 33.vii above.

As for the elements of On, we (ambiguously) use lower-case Greek letters to denote the elements of $T(K)$. Furthermore we formally define $\alpha + 0 = 0 + \alpha = \alpha$ for all $\alpha \in T(K)$. Note that $0 \notin P$.

Definition 34. *To relate the elements of $T(K)$ to more expressive ordinal notations, we define $1 := \psi(\overline{0})$, $\omega := \psi(\overline{0}, 1)$, $\varepsilon_0 := \psi(\overline{0}, 1, 0)$, and $\Gamma_0 := \psi(\overline{0}, 1, 0, 0)$.*

Let Lim be the set of elements in T which are neither 0 nor of the form $\alpha + 1$. Elements of Lim are called *limit* ordinal terms. Note that we use the same notation for the set of limit ordinals in On and a subset of T. However, no confusion will arise from this.

Exercise 16. Let $(T(K), <)$ be defined as above. Show that $<$ is a strict total order on T.

Hint: Let Size(α) denote the number of symbols in the ordinal term α. Exploiting induction on Size(α) one easily verifies that the order $<$ is well-defined.

Theorem 18. *Let $(T(K), <)$ be defined as above. Then $<$ is an irreflexive (countable) well-order on T.*

Proof. We take the result of the exercise for granted. Hence it remains to establish that $<$ is well-founded. (The fact that $<$ is a countable order is trivial.) We sketch the definition of an embedding o mapping $(T(K), <)$ onto $(\mathcal{T}(\Sigma), \prec_{\mathsf{lpo}})$. In the first step, we consider terms in P. That is we consider $\alpha = \psi(\alpha_1, \ldots, \alpha_m)$, and $\beta = \psi(\beta_1, \ldots, \beta_n)$. Then it is obvious how to define an embedding for α, β.

In the second step, we consider terms $\alpha = \alpha_1 + \cdots + \alpha_m$ and $\beta = \beta_1 + \cdots + \beta_n$. We assume the embedding o has already be defined for P. Then we set

$$o(\alpha) = o(\alpha_1 + \cdots + \alpha_m) := p(o(\alpha_1), o(\alpha_2 + \cdots + \alpha_m))$$

and similar for β. It is easy to check that $\alpha < \beta$ implies $o(\alpha) \prec_{\mathsf{lpo}} o(\beta)$.

Finally, a close look to the definition of $(T(K), <)$ reveals that a comparison between any terms $\alpha, \beta \in T(K)$ can always be reduced to a comparison of the above kinds. This completes the construction of the embedding o. Hence the well-foundedness of $<$ follows from the well-foundedness of $\prec_{\mathsf{lpo}}$.

In the following proposition we want to relate the *order type* of the well-order $(T(K), <)$ and the wellfounded partial order $(\mathcal{T}(\Sigma, \mathcal{V}), \prec_{\mathsf{lpo}})$. Concerning the latter it is best to momentarily restrict our attention to the well-order $(\mathcal{T}(\Sigma), \prec_{\mathsf{lpo}})$. We write $\mathcal{A}(k)$ to denote the set of signatures Σ such that the maximal arity of $f \in \Sigma$ is $\leq k$.

Theorem 19. i. *For any number $k \geq 1$, there exists an embedding from the well-order $(\mathcal{T}(\Sigma), \prec_{lpo})$ into $(T(k+1), <)$, where $\Sigma \in \mathcal{A}(k)$.*
ii. *For any number $k > 2$, there exists an embedding from $(T(K), <)$ into $(\mathcal{T}(\Sigma), \prec_{lpo})$, where $\Sigma \in \mathcal{A}(k)$.*
iii. *Finally, we have*

$$\sup_{2 \leq k \in \mathbb{N}} (\mathrm{otype}(T(k), <)) = \sup_{k \in \mathbb{N}} \{\mathrm{otype}\,(\prec_{lpo}) \colon \prec_{lpo} \text{ is a LPO over } \Sigma \in \mathcal{A}(k)\} \ .$$

Proof. The first two assertions are a consequence of the well-order proof of $(T, <)$. We only comment on the stated lower bound in the second one. The statement fails for $(T(2), <)$ and $(\mathcal{T}(\Sigma(2)), \prec_{\mathsf{lpo}})$ if only one binary function symbol is present in Σ. The presence of the binary function symbol $+$ in $T(2)$ can make the order $<$ more expressive than $\prec_{\mathsf{lpo}}$. This difference vanishes for $k \geq 3$. The third assertion is a direct consequence of the first two.

We denote the supremum of the order types of the notation systems $T(K)$ as Λ.

Definition 35.

$$\Lambda := \sup_{2 \leqslant k} (\mathrm{otype}(T(k), <)) = \sup_{k \in \mathbb{N}} \{\mathrm{otype}\,(\prec_{lpo}) : \prec_{lpo} \text{ is a LPO over } \Sigma \in \mathcal{A}(k)\} \ .$$

Remark 5. It follows from work by Schmidt [1979] that the ordinal Λ is equal to the so-called *small Veblen ordinal*. For the convenience of readers familiar with (ordinal) proof theory we mention that the small Veblen ordinal is sometimes denoted as $\vartheta(\Omega^\omega)$, where Ω stands either for the first uncountable ordinal, of the first not recursively representable ordinal ω_1^{CK}. We will not go into any details here.

Now we turn to multiset path orders. We write $\mathcal{C}(k)$ to denote the set of signatures Σ such that the cardinality of Σ is $\leq k$.

Theorem 20. i. *For any number k, there exists an embedding from the well-order $(\mathcal{T}(\Sigma), \prec_{\mathsf{mpo}})$ into $(T(2), <(\psi(k+1, 0)))$, where $\Sigma \in \mathcal{C}(k)$.*
ii. *For any number k, there exists an embedding from $(T(2), <(\psi(k, 0)))$ into $(\mathcal{T}(\Sigma), \prec_{\mathsf{mpo}})$, where $\Sigma \in \mathcal{C}(k+2)$.*
iii. *Finally, we obtain*

$$\sup_{k \in \mathbb{N}} (\mathrm{otype}(T(2), <(\psi(k, 0)))) = \sup_{k \in \mathbb{N}} \{\mathrm{otype}\,(\prec_{\mathsf{mpo}}) \colon \prec_{\mathsf{mpo}} \text{ is a MPO over } \Sigma \in \mathcal{C}(k)\} \ .$$

Proof. For the first part we recursively define a mapping $o \colon (\mathcal{T}(\Sigma), \prec_{\mathsf{mpo}}) \to (T(2), <)$ as follows. If $s = f_j(s_1, \ldots, s_m)$, then set $o(s) := \psi(j, o(s_1) + \cdots + o(s_m))$. Then it is easy to see that $o(s) < \psi(k+1, 0)$ for all $s \in \mathcal{T}(\Sigma)$. Furthermore it is easy to see that for all terms $s, t \in \mathcal{T}(\Sigma)$, $s \prec_{\mathsf{mpo}} t \to o(s) < o(t)$. We leave this as an exercise. Hence o is indeed an embedding and therfore $\mathrm{otype}(\mathcal{T}(\Sigma), \prec_{\mathsf{mpo}}) \leq \psi(k+1, 0)$, holds.

Now consider the second part. We suppose Σ contains a constant symbol 0, a binary function symbol p, and k unary function symbols $f_0, \ldots, f_{k-1}$. We recursively define an embedding $o\colon (T(2), <(\psi(k,0))) \to (\mathcal{T}(\Sigma), \prec_{\mathsf{mpo}})$ by

$$o(0) := 0 \quad \text{and} \quad o(\alpha_1 + \cdots + \alpha_m) := p(o(\alpha_1), o(\alpha_2 + \cdots + \alpha_n)) \ ,$$

where $\alpha_i \in P$ for all $i = 1, \ldots, m$. Finally

$$o(\psi(i,\alpha)) := f_i(o(\alpha)) \quad \text{for all } i = 0, \ldots, k-1 \ .$$

This completes the definition of o. (Again the proof that o is indeed an embedding is left as an exercise.)

The third assertion follows from the first two.

3.5 Interpretation

Now we turn to the *interpretation* of the notation system $(T(K), <)$. The only tricky part in the interpretation of the objects occurring in $T(K)$ is to give an interpretation for the K-ary function symbol ψ. To this end, we are going to introduce a part of the *Veblen [1908] function* φ and its fixed point free variant ψ. This part suffices to reach Λ.

For $\alpha_1, \ldots, \alpha_k \in \mathsf{On}$ with $k > 0$ we intend to recursively define the branch $\varphi_{\bar{\alpha}}\colon \mathsf{On} \to \mathsf{On}$ of the *Veblen function*. It is advisable to interchangeably use $\varphi_{\bar{\alpha}}(\beta)$ and $\varphi(\bar{\alpha}, \beta)$, thus regarding φ as a function from the ordinal sequences of lengths larger than 1 into the ordinals. The principal ordinals H are enumerated by $\varphi_{\bar{0}}$. If $\alpha_k > 0$, then $\varphi_{\bar{\alpha}}$ is the enumerating function of

$$\{\beta : (\forall \gamma < \alpha_k)(\varphi(\alpha_1, \ldots, \alpha_{k-1}, \gamma, \beta) = \beta)\} \ ,$$

and otherwise we have $(\alpha_1, \ldots, \alpha_k) = (\alpha_1, \ldots, \alpha_i, \bar{0}, 0)$ with $\alpha_i > 0$. Here we let $\varphi_{\bar{\alpha}}$ be the enumerating function of

$$\{\beta : (\forall \gamma < \alpha_i)(\varphi(\alpha_1, \ldots, \alpha_{i-1}, \gamma, \beta, \bar{0}, 0) = \beta)\} \ .$$

Obviously $\varphi_{\bar{0},\bar{\alpha}} = \varphi_{\bar{\alpha}}$ holds (Exercise). By definition $\varphi_{\bar{0},1}$ enumerates the epsilons. The φ function lacks the subterm property since it admits fixed points. Therefore we concentrate on ψ, the *fixed point free* version of φ.

Definition 36. *We let $\psi(\alpha_1, \ldots, \alpha_k, \beta)$ be $\varphi(\bar{\alpha}, \beta + 1)$ if $\beta = \beta_0 + n$ for some $n \in \mathbb{N}$ and $\beta_0 \in \mathsf{Lim} \cup \{0\}$ with $\varphi(\bar{\alpha}, \beta_0) \in \{\alpha_1, \ldots, \alpha_k, \beta_0\}$, and otherwise $\psi(\bar{\alpha}, \beta)$ is just $\varphi(\bar{\alpha}, \beta)$.*

Now we are in the position to restate the central result of this course, compare Theorem 19 and Theorem 20.

Theorem 21 (Dershowitz and Okada 1988). *Let Σ be a finite signature. Let $\prec_{\mathsf{lpo}}$ ($\prec_{\mathsf{mpo}}$) denote a lexicographic (multiset) path order based on Σ. Then*

i. $\sup_{k \in \mathbb{N}} \{\text{otype}(\prec_{lpo}) : \prec_{lpo}$ *is a* LPO *over* $\Sigma \in \mathcal{A}(k)\} = \Lambda$*, our denotation of the* small Veblen ordinal *and*

ii. $\sup_{k\in\mathbb{N}} \{\text{otype}\,(\prec_{\mathsf{mpo}}) : \prec_{\mathsf{mpo}}$ *is a* MPO *over* $\Sigma \in \mathcal{C}(k)\} = \varphi(\omega, 0)$.

We denote the first infinite ordinal closed under $+$ and the k-ary ψ ($k \geq 2$) as Δ_k:

$$\Delta_k := \varphi(1, \underbrace{0, \ldots, 0}_{k \text{ times}}) = \psi(1, \underbrace{0, \ldots, 0}_{k \text{ times}}) \,. \tag{7}$$

Thus Δ_2 coincides with the ordinal Γ_0 celebrated by Gallier [1991]. (In other contexts the ordinal Γ_0 is sometimes called the *ordinal of predicativity*.) Note that $\Delta_k = \text{otype}((T(k), <))$ (for $k \geq 2$) follows directly from the definitions.

Another denotation for the small Veblen number Λ is $\vartheta(\Omega^\omega)$.The connection between the Veblen function φ and ϑ was illuminated by Schmidt [1979]. The equality $\Lambda = \vartheta(\Omega^\omega)$ comes about as follows.

$$\Lambda = \sup\{\text{otype}(T(k), <) : k \in \mathbb{N} \wedge k \geq 2\} = \sup\{\Delta_k : k \in \mathbb{N} \wedge k \geq 2\} = \vartheta(\Omega^\omega) \,.$$

The following theorem relates the strength of LPO to arbitrary simplification orders.

Theorem 22 (Schmidt 1979).

$$\sup\{\text{otype}(\prec) : \prec \textit{ is a simplification order}\} = \Lambda$$

By Proposition 7, for every ordinal $\alpha > 0$ there are uniquely determined principal ordinals $\alpha_0 \geqslant \ldots \geqslant \alpha_n$ such that $\alpha = \alpha_0 + \cdots + \alpha_n$ holds. In addition, for every principal $\alpha < \Delta_k$ there are uniquely determined $\alpha_1, \ldots, \alpha_{k+1}$ below α satisfying $\alpha = \psi(\bar{\alpha})$, cf. Buchholz [1993]. So every $\alpha < \Delta_k$ can be associated with a unique representation solely built up from 0, $+$ and the K-ary ψ. We call this the *k-normal form* of α. The next Lemma lists some of the basic properties of ψ. Recall from Definition 5 that the lexicographic order of ordinal tuples having the same length is denoted by $<_{\mathsf{lex}}$.

Lemma 7. *Let* $\alpha_1, \ldots, \alpha_{k+1}$ *and* $\gamma_1, \ldots, \gamma_{k+1}$ *be given.*

i. *Each* $\psi(\bar{\alpha})$ *is a principal ordinal and, except for* $\psi(\bar{0}) = 1$*, a limit ordinal.*
ii. *The function* ψ *has the subterm property and is monotone.*
iii. $\psi(\bar{\alpha}) > \psi(\bar{\gamma})$ *is equivalent to* $[\bar{\alpha} >_{\mathsf{lex}} \bar{\gamma} \wedge \psi(\bar{\alpha}) > \bar{\gamma}] \vee \psi(\bar{\gamma}) \leqslant \bar{\alpha}$.

Exercise 17. It is an easy but interesting exercise to relate the above Lemma to the definition of the notation system $(T(k+1), <)$. In particular note the correspondence between Lemma 7.iii and Definition 33.vii.

Recall from (2) the notion $f(\bar{x}, \cdot, \bar{y})^n(z)$ of the n^{th} iteration of the unary function $v \mapsto f(\bar{x}, v, \bar{y})$ on z. The following Lemma contains all properties of ψ we will rely on later.

Lemma 8. *Let* $i \in [1, k]$*;* $m, n \in \mathbb{N}$*, and ordinals* $\alpha_1, \ldots, \alpha_{k+1}$*,* $\gamma_1, \ldots, \gamma_{k+1}$*,* δ, δ' *be given.*

i. *If* $n > m$*, then* $\psi(\bar{\alpha}) \cdot n > \psi(\bar{\alpha}) \cdot m$.

ii. *If* $\psi(\bar\alpha) > \psi(\bar\gamma)$, *then* $\psi(\bar\alpha) > \psi(\bar\gamma) \cdot n + m$.
iii. *If* $n \geqslant m$; $\delta \geqslant \delta'$, *and at least one of the inequalities is proper, then*

$$\psi(\alpha_1, \dots, \alpha_i, \cdot, \alpha_{i+1}, \dots, \alpha_k)^n(\delta) > \psi(\alpha_1, \dots, \alpha_i, \cdot, \alpha_{i+1}, \dots, \alpha_k)^m(\delta') \ .$$

iv. *If* $(\gamma_1, \dots, \gamma_i) >_{\mathsf{lex}} (\alpha_1, \dots, \alpha_i)$ *and* $\psi(\bar\gamma) > \alpha_1, \dots, \alpha_k, \delta$, *then*

$$\psi(\bar\gamma) > \psi(\alpha_1, \dots, \alpha_i, \cdot, \alpha_{i+1}, \dots, \alpha_k)^n(\delta) \ .$$

Proof. The first point follows from $\psi(\bar\alpha) \geq 1$. Under the conditions of (ii), $\psi(\bar\alpha)$ is a principal limit ordinal. For (iii) we utilize the subterm property and the monotonicity of ψ, respectively. Finally, (iv) is established by induction on n using Lemma 7.iii.

Exercise 18. Show the above Lemma without reference to the collection On, i.e. purely on basis of the definition of $<$ as introduced in Definition 33.

Remark 6. Note that in the remaining part of this course material, we need not make use of the above defined *interpretation* of the notation system $(T(K), <)$, but can rely directly on the properties of $(T(K), <)$ instead. Thus we can make sure that the objects we are dealing with are always concretely.

4 Slow-growing and Fast-growing functions

The purpose of this section is to introduce hierarchies of subrecursive functions, defined through transfinite recursion along the terms in the ordinal notation system $T(K)$, where K is an arbitrary natural number ≥ 2, cf. Definition 33. Note that we can work purely within the ordinal notation system T based upon the K-ary ψ function symbol. In particular no reference is made to the class On.

Convention: Throughout this section K is supposed to be fixed. To simplify denotation we will drop the reference to K henceforth.

4.1 Fundamental sequences

To each ordinal term $\alpha \in T$ we assign a canonical sequence of ordinal terms $\langle \alpha[x] \colon x \in \mathbb{N} \rangle$, the *fundamental sequence*. We have to wade through some technical definitions. We define the set $\mathrm{IS}_{\overline{\alpha}}(\gamma)$, the set of *interesting subterms* of γ (relative to $\overline{\alpha}$) by induction on γ. We set $\mathrm{IS}_{\overline{\alpha}}(0) := \emptyset$, $\mathrm{IS}_{\overline{\alpha}}(\gamma_1 + \cdots + \gamma_m) := \bigcup_{i=1}^m \mathrm{IS}_{\overline{\alpha}}(\gamma_i)$, and finally

$$\mathrm{IS}_{\overline{\alpha}}(\psi(\gamma_1, \dots, \gamma_K)) := \begin{cases} \{\psi(\overline{\gamma})\} & \text{if } (\gamma_1, \dots, \gamma_{K-1}) \geq_{\mathsf{lex}} (\alpha_1, \dots, \alpha_{K-1}) \\ \bigcup_{i=1}^K \mathrm{IS}_{\overline{\alpha}}(\gamma_i) & \text{otherwise.} \end{cases}$$

The (relative to $\overline{\alpha}$) *maximal interesting subterm* $\mathrm{MS}_{\overline{\alpha}}(\gamma_1, \dots, \gamma_n)$ of a non-empty sequence $(\gamma_1, \dots, \gamma_n)$ is defined as the maximum of the terms occurring in $\mathrm{IS}_{\overline{\alpha}}(\gamma_i)$. Let $>_{\mathsf{lex}}$ denote the lexicographic order on sequences of ordinal terms induced by $>$. Let $\overline{\alpha} = \alpha_1, \ldots, \alpha_N \in T$ and $\beta \in T$. Then set

$$\mathsf{Fix}(\overline{\alpha}) := \{\psi(\overline{\gamma}, \delta) \colon \overline{\gamma} >_{\mathsf{lex}} \overline{\alpha} \text{ and } \psi(\overline{\gamma}, \delta) > \alpha_i \text{ for all } i = 1, \ldots, K-1\} . \quad (8)$$

Within this section λ (possibly extended by a subscript) will denote a limit ordinal term.

Definition 37. *Recursive definition of $\alpha[x]$ for $x < \omega$.*

$$\begin{aligned}
0[x] &:= 0 \\
(\alpha_1 + \cdots + \alpha_m)[x] &:= \alpha_1 + \cdots + \alpha_m[x] \qquad m > 1; \alpha_1 \geq \cdots \geq \alpha_m \\
\psi(\overline{0})[x] &:= 0 \\
\psi(\overline{0}, \beta + 1)[x] &:= \psi(\overline{0}, \beta) \cdot (x+1) \\
\psi(\overline{0}, \lambda)[x] &:= \psi(\overline{0}, \lambda[x]) \qquad \lambda \notin \mathsf{Fix}(\overline{0}) \\
\psi(\overline{0}, \lambda)[x] &:= \lambda \cdot (x+1) \qquad \lambda \in \mathsf{Fix}(\overline{0}) \\
\psi(\alpha_1, \ldots, \alpha_i + 1, \overline{0}, 0)[x] &:= \psi(\alpha_1, \ldots, \alpha_i, \cdot, \overline{0})^{x+1}(0) \\
\psi(\alpha_1, \ldots, \alpha_i + 1, \overline{0}, \beta + 1)[x] &:= \psi(\alpha_1, \ldots, \alpha_i, \cdot, \overline{0})^{x+1}(\psi(\alpha_1, \ldots, \alpha_i + 1, \overline{0}, \beta)) \\
\psi(\alpha_1, \ldots, \alpha_i + 1, \overline{0}, \lambda)[x] &:= \psi(\alpha_1, \ldots, \alpha_i + 1, \overline{0}, \lambda[x]) \qquad \lambda \notin \mathsf{Fix}(\overline{\alpha}, \overline{0}) \\
\psi(\alpha_1, \ldots, \alpha_i + 1, \overline{0}, \lambda)[x] &:= \psi(\alpha_1, \ldots, \alpha_i, \cdot, \overline{0})^{x+1}(\lambda) \qquad \lambda \in \mathsf{Fix}(\overline{\alpha}, \overline{0}) \\
\psi(\alpha_1, \ldots, \lambda_i, \overline{0}, 0)[x] &:= \psi(\alpha_1, \ldots, \lambda_i[x], \overline{0}, \mathrm{MS}_{\overline{\alpha}, \lambda_i, \overline{0}}(\overline{\alpha}, \lambda_i)) \\
\psi(\alpha_1, \ldots, \lambda_i, \overline{0}, \beta + 1)[x] &:= \psi(\alpha_1, \ldots, \lambda_i[x], \overline{0}, \psi(\alpha_1, \ldots, \lambda_i, \overline{0}, \beta)) \\
\psi(\alpha_1, \ldots, \lambda_i, \overline{0}, \lambda)[x] &:= \psi(\alpha_1, \ldots, \lambda_i, \overline{0}, \lambda[x]) \qquad \lambda \notin \mathsf{Fix}(\overline{\alpha}, \overline{0}) \\
\psi(\alpha_1, \ldots, \lambda_i, \overline{0}, \lambda)[x] &:= \psi(\alpha_1, \ldots, \lambda_i[x], \overline{0}, \lambda) \qquad \lambda \in \mathsf{Fix}(\overline{\alpha}, \overline{0})
\end{aligned}$$

The above definition is given in such a way as to simplify the comparison between the fundamental sequences for T and the fundamental sequences for the set of ordinal terms $T(2)$ (built from 0, +, and a 2-ary function symbol ψ) as presented in Weiermann [2001]. Note that our definition is equivalent to the more compact one presented in Lepper [2003]. The following proposition is stated without proof. A proof (for a slightly different assignment of fundamental sequences) can be found in Buchholz [2003].

Theorem 23. *Let $\alpha \in T$ be given; assume $x < \omega$. If $\alpha > 0$, then $\alpha > \alpha[x]$. For $\alpha > 1$ we get $\alpha[x] > 0$, and if $\alpha \in \mathsf{Lim}$, then $\alpha[x+1] > \alpha[x]$. Finally, for $\alpha \in \mathsf{Lim}$, $\sup_{x<\omega} \alpha[x] = \alpha$.*

In the definition of $\psi(\alpha_1, \ldots, \lambda_i, \overline{0}, 0)[x]$ we introduce at the last position of ψ the term $\mathrm{MS}_{\overline{a}, \overline{0}}(\overline{\alpha})$. We cannot simply dispense of this term. To see this, we momentarily consider only 3-ary ψ-functions, set $\varGamma_0 := \psi(1, 0, 0)$ and calculate $\psi(0, \varGamma_0[x], 0)$:

$$\begin{aligned}
\psi(0, \varGamma_0[x], 0) &= \psi(0, \psi(1, 0, 0)[x], 0) \\
&= \psi(0, \psi(0, \cdot, 0)^{x+1}(0), 0) = \psi(0, \cdot, 0)^{x+2}(0) < \psi(1, 0, 0) .
\end{aligned}$$

Hence for every $x < \omega$ we have $\psi(0, \varGamma_0[x], 0) < \varGamma_0 < \psi(0, \varGamma_0, 0)$.

Definition 38 (Bachmann 1967). *We say that a ordinal notation system T together with an assignment of fundamental sequences is a Bachmann system $(T, \cdot[\cdot])$ if*

$$\alpha[x] < \beta < \alpha \to \alpha[x] \leq \beta[0]$$

holds for all $\alpha, \beta \in T$.

As a side-remark we want to mention that the given assignment of fundamental sequences even fulfills the *Bachmann* property, i.e. $(T, \cdot[\cdot])$ is a Bachmann system.

4.2 Subrecursive hierarchies

Utilizing Definition 37 we are now in the position to define subrecursive hierarchies of ordinal functions. We start with a technical definition.

Definition 39. *A function $f\colon \mathbb{N} \to \mathbb{N}$* (eventually) dominates *$g\colon \mathbb{N}^n \to \mathbb{N}$ if*

$$g(m_1, \ldots, m_n) < f(\max\{m_1, \ldots, m_n\}) \tag{9}$$

holds for (almost) all $m_1, \ldots, m_n$, where "almost all" means "all but finitely many".

We abbreviate eventual domination by $g <_{\mathrm{ed}} f$. The canonical variant using $\leqslant$ in (9) is called $\leqslant_{\mathrm{ed}}$. We extend the notion to sets X and Y of number-theoretic functions, where $X \leqslant_{\mathrm{ed}} Y$ means $(\forall g \in X)(\exists f \in Y)(g \leqslant_{\mathrm{ed}} f)$. If additionally $Y \leqslant_{\mathrm{ed}} X$ holds, then we mark this by $X \approx Y$.

We assume the reader is familiar with the usual definitions of the set Elem of *elementary functions* and the set Prec of *primitive recursive functions*. We will briefly recall a nice classification of the set Prec through the (binary) Ackermann function.

Definition 40 (Ritchie 1965). *The (binary)* Ackermann function *is given by* $\mathrm{Ack}(n, m) := \mathrm{Ack}_n(m)$*, with its branches* $\mathrm{Ack}_n\colon \mathbb{N} \to \mathbb{N}$ *generated via*

$$\mathrm{Ack}_0(m) := m + 1 \quad \textit{and} \quad \mathrm{Ack}_{n+1}(m) := \mathrm{Ack}_n^{m+1}(1)\ .$$

Usually the (binary) Ackermann function is defined differently, as follows, cf. Péter [1935].

$$\begin{aligned} \mathrm{Ack}(0, m) &= m + 1\ , \\ \mathrm{Ack}(n + 1, 0) &= \mathrm{Ack}(n, 1)\ , \text{ and} \\ \mathrm{Ack}(n + 1, m + 1) &= \mathrm{Ack}(n, \mathrm{Ack}(n + 1, m))\ . \end{aligned}$$

Exercise 19. Show that both formulations of the Ackermann function are equivalent. Show that Ack is monotone in both its arguments. Show that the Ack_n are primitive recursive.

Theorem 24.

i. *For any primitive recursive function* $f\colon \mathbb{N}^k \to \mathbb{N}$ *there is an* n *such that*

$$f(m_1,\dots,m_k) < \mathrm{Ack}_n(\max\{m_1,\dots,m_k\})$$

holds for all $\bar{m} \in \mathbb{N}$*. Hence* f *is dominated by* Ack_n*.*

ii. *We have* $\textsc{Prec} \approx \{\mathrm{Ack}_n : n \in \mathbb{N}\}$*.*

iii. Ack *is not primitive recursive.*

iv. *For all* n *we have* $\mathrm{Ack}(n,n) \leqslant \mathrm{Ack}(n+2,0)$*.*

We extend the binary function Ack to higher arities.

Definition 41. *For* $k > 2$ *the* k-ary Ackermann function *is defined by*

$$\begin{aligned}
\mathrm{Ack}(\bar{0},m) &:= m+1\ ,\\
\mathrm{Ack}(\bar{l},n+1,0) &:= \mathrm{Ack}(\bar{l},n,1)\ ,\\
\mathrm{Ack}(\bar{l},n+1,m+1) &:= \mathrm{Ack}(\bar{l},n,\mathrm{Ack}(\bar{l},n+1,m))\ ,\ \textit{and}\\
\mathrm{Ack}(l_1,\dots,l_{i-1},l_i+1,0,\bar{0},m) &:= \mathrm{Ack}(\bar{l},m,\bar{0},m)\ .
\end{aligned}$$

We slightly simplify the usual definition of the set of multiple recursive functions. See Péter 1936 for the original definition.

Definition 42 (Péter 1936). *The set* $\textsc{Mrec}$ *of* multiple recursive functions *is the smallest set of number-theoretic functions which contains zero functions of arbitrary arities, the successor, and the projections and is closed under substitution and all* k*-ary Ackermann functions.*

Definition 43. *Based (and depending) on our Bachmann system* $(T,\cdot[\cdot])$*, we recursively define three hierarchies of number-theoretic functions. Consider* $\alpha,\lambda \in T$ *with* $\lambda \in \mathsf{Lim}$*.*

The fast growing functions $(\mathrm{F}_\gamma)_{\gamma\in T}$ *are based on iterations*

$$\mathrm{F}_0(m) := m+1, \quad \mathrm{F}_{\alpha+1}(m) := \mathrm{F}_\alpha^{m+1}(m), \quad \textit{and} \quad \mathrm{F}_\lambda(m) := \mathrm{F}_{\lambda[m]}(m)\ .$$

Of slightly slower growth (below ε_0*) are the* Hardy functions $(\mathrm{H}_\gamma)_{\gamma\in T}$

$$\mathrm{H}_0(m) := m, \quad \mathrm{H}_{\alpha+1}(m) := \mathrm{H}_\alpha(m+1), \quad \textit{and} \quad \mathrm{H}_\lambda(m) := \mathrm{H}_{\lambda[m]}(m)\ .$$

Finally, the slow growing functions $(\mathrm{G}_\gamma)_{\gamma\in T}$

$$\mathrm{G}_0(m) := 0, \quad \mathrm{G}_{\alpha+1}(m) := \mathrm{G}_\alpha(m)+1, \quad \textit{and} \quad \mathrm{G}_\lambda(m) := \mathrm{G}_{\lambda[m]}(m)\ .$$

Hardy [1904] used the H_γ to construct a set of real numbers of cardinality $\aleph_1$, and their first appearance in the field of subrecursive hierarchies is Wainer [1972]. Robbin [1965] investigated the fast growing functions up to ω^ω as the canonical extension of (variants of) the Grzegorczyk [1953] functions to the transfinite. Later this approach was extended to ε_0, see Löb and Wainer [1970], Schwichtenberg [1971], and Wainer [1970]. Weiermann [1997] observed that the choice of the underlying assignment of fundamental sequences is vital to the slow growing hierarchy. In contrast to this, the other two hierarchies are not that much affected by changes of the assignment. However, the slow growing functions possesses some nice properties. Any primitive recursive function f over $\mathbb{N}$

$$\begin{aligned} f(a,0) &:= h(a) \\ f(a,b+1) &:= g(a,b,f(a,b)) \end{aligned}$$

can be easily extended to T by extending the definition of f with a third clause for the 'limit' case. Assume ρ, χ express the extensions of h, g, respectively. Set

$$\begin{aligned} \phi(\alpha,0) &:= \rho(\alpha) \\ \phi(\alpha,\beta+1) &:= \chi(\alpha,\beta,\phi(\alpha,\beta)) \\ \phi(\alpha,\lambda) &:= \sup_{x\in\mathbb{N}} \phi(\alpha,\lambda[x]) \ . \end{aligned}$$

With the use of the slow-growing hierarchy a clear connection between the number-theoretic function f and its lifting $\phi\colon T \to T$ can be given. To make the argument easily presentable we swap the argument in the definition of the G function. We write $\mathrm{G}_n(\alpha)$ instead of $\mathrm{G}_\alpha(n)$. I.e. for each fixed $n \in \mathbb{N}$ there is a function $\mathrm{G}_n\colon T \to \mathbb{N}$. We prove that

$$\mathrm{G}_n(\phi(\alpha,\beta)) = f(\mathrm{G}_n(\alpha), \mathrm{G}_n(\beta))$$

assuming we have already for fixed n, and all α, β, γ.

$$\begin{aligned} \mathrm{G}_n(\rho(\alpha)) &= h(\mathrm{G}_n(\alpha)) \\ \mathrm{G}_n(\chi(\alpha,\beta,\gamma)) &= g(\mathrm{G}_n(\alpha), \mathrm{G}_n(\beta), \mathrm{G}_n(\gamma)) \ . \end{aligned}$$

The result follows from a simple induction on β. Hence G_n constitutes a homomorphism, collapsing the arithmetic on T down to $\mathbb{N}$. Using the recipe above, the following Lemma is easy.

Lemma 9. *Let $n \in \mathbb{N}$ be fixed. Then*

$$\begin{aligned} \mathrm{G}_n(\alpha+\beta) &= \mathrm{G}_n(\alpha) + \mathrm{G}_n(\beta) \\ \mathrm{G}_n(\alpha\cdot\beta) &= \mathrm{G}_n(\alpha) \cdot \mathrm{G}_n(\beta) \\ \mathrm{G}_n(\alpha^\beta) &= \mathrm{G}_n(\alpha)^{\mathrm{G}_n(\beta)} \ . \end{aligned}$$

As properties of the fast-growing hierarchies note the following facts.

Exercise 20.

$$\mathrm{H}_{\alpha+\beta} = \mathrm{H}_\alpha \circ \mathrm{H}_\beta \quad \text{and} \quad \mathrm{H}_{\omega^\alpha} = \mathrm{F}_\alpha \ .$$

Definition 44. *Let $<_{(x)}$ denote the* transitive closure *of $.[x]$. I.e.*

$$\alpha <_{(x)} \beta \ :\Longleftrightarrow \ \alpha <_{(x)} \beta[x] \vee \alpha = \beta[x] \ .$$

Lemma 10. *For all $\alpha \in T$.*

i. G_α *is increasing (strictly if α is infinite), and if $\beta <_{(n)} \alpha$, then* $\mathrm{G}_\beta(m) < \mathrm{G}_\alpha(m)$ *for all $m > n$.*

ii. H_α *and* F_α *are strictly increasing, and if* $\beta <_{(n)} \alpha$*, then* $\mathrm{H}_\beta(m) < \mathrm{H}_\alpha(m)$*, and* $\mathrm{F}_\beta(m) < \mathrm{F}_\alpha(m)$ *for all* $m > n$*.*

Proof. We only prove that (a) G_α is increasing and (b) that for all $n < m$, $\beta <_{(n)} \alpha \to \mathrm{G}_\beta(m) < \mathrm{G}_\alpha(m)$ by simultaneous induction on α. To prove the second assertion of the proposition, follow the pattern of the given proof. We abbreviate induction hypothesis as (ih).

Let $\alpha = 0$. Then $\mathrm{G}_0(m) = 0 = \mathrm{G}_0(m+1)$. Now consider a successor ordinal $\alpha + 1$. Then

$$\mathrm{G}_{\alpha+1}(m) = \mathrm{G}_\alpha(m) + 1 \leq \mathrm{G}_\alpha(m+1) + 1 = \mathrm{G}_{\alpha+1}(m+1) \ .$$

Finally consider the case of a limit ordinal λ. By definition and (ih)

$$\mathrm{G}_\lambda(m) = \mathrm{G}_{\lambda[m]}(m) \leq \mathrm{G}_{\lambda[m]}(m+1) \ .$$

To proceed we have to prove the following claim.

Claim. For any m, n, $n < m$, and $\lambda \in \mathsf{Lim}$, $\lambda[n] <_{(0)} \lambda[m]$ holds.

Assume the claim has already been established. Then (ih) yields

$$\mathrm{G}_{\lambda[m]}(m+1) \leq \mathrm{G}_{\lambda[m+1]}(m+1) = \mathrm{G}_\lambda(m+1) \ ,$$

yielding property (a). Now we prove the claim. Due to Theorem 23 we obtain $\lambda[n] < \lambda[m]$. Now we apply the Bachmann property $(\lambda[n] < \lambda[m] < \lambda) \to (\lambda[n] \leq \lambda[m][0])$ repeatedly to derive $\lambda[n] = \lambda[m][0] \ldots [0]$. Thus we obtain $\lambda[n] <_{(0)} \lambda[m]$.

We turn to property (b). The case $\alpha = 0$ is trivial. Assume the case of a successor ordinal, and suppose $\beta <_{(n)} \alpha + 1$. Hence either $\beta = \alpha$ or $\beta <_{(n)} \alpha$. Assume the former, then

$$\mathrm{G}_\beta(m) < \mathrm{G}_\beta(m) + 1 = \mathrm{G}_\alpha(m) + 1 = \mathrm{G}_{\alpha+1}(m) \ .$$

on the other hand, we have by (ih)

$$\mathrm{G}_\beta(m) < \mathrm{G}_\alpha(m) < \mathrm{G}_\alpha(m) + 1 = \mathrm{G}_{\alpha+1}(m) \ .$$

Finally consider a limit ordinal λ. Then assume $\beta <_{(n)} \lambda$, and assume $\beta = \lambda[n]$. By the claim, we obtain $\lambda[n] <_{(0)} \lambda[m]$ and thus

$$\mathrm{G}_\beta(m) = \mathrm{G}_{\lambda[n]}(m) < \mathrm{G}_{\lambda[m]}(m) < \mathrm{G}_\lambda(m) \ .$$

Theorem 25. *Let* $\alpha, \beta \in T$*. Assume* $\beta < \alpha$*. Then* G_β *(*H_β*,* F_β*) is eventually dominated by* G_α *(*H_α*,* F_α*).*

Proof. This time, we argue with respect to the Hardy hierarchy. We employ induction on α. If $\beta < \alpha$, then there exists n such that $\beta < \alpha[n]$ by Theorem 23. Hence, by (ih), H_β is eventually dominated by $\mathrm{H}_{\alpha[n]}$. If $m > n$, then $\mathrm{H}_{\alpha[n]}(m) < \mathrm{H}_\alpha(m)$, by the Proposition. Thus H_α eventually dominates H_β.

In the literature another variant of the Hardy hierarchy occurs, whose definition is slightly different.

Definition 45. *Consider* $\alpha, \lambda \in T$ *with* $\lambda \in \mathsf{Lim}$. *We define a slightly faster growing variant of the* Hardy functions $(\mathrm{H}'_\gamma)_{\gamma \in T}$

$$\mathrm{H}'_0(m) := m, \quad \mathrm{H}'_{\alpha+1}(m) := \mathrm{H}'_\alpha(m+1), \quad \textit{and} \quad \mathrm{H}'_\lambda(m) := \mathrm{H}'_{\lambda[m]}(m+1)\ ,$$

and the related counting functions $(\mathrm{L}_\gamma)_{\gamma \in T}$

$$\mathrm{L}_0(m) := 0, \quad \mathrm{L}_{\alpha+1}(m) := \mathrm{L}_\alpha(m+1)+1, \quad \textit{and} \quad \mathrm{L}_\lambda(m) := \mathrm{L}_{\lambda[m]}(m)+1\ .$$

This *variant* of the Hardy hierarchy allows a more direct assessment of the longest possible Hydra battle, see Section 5. The following lemma relates the functions H'_α and L_α.

Lemma 11. *For any* $\alpha < \mathit{\Lambda}$ *and any* $m \in \mathbb{N}$ *we have* $\mathrm{H}'_\alpha(m) = \mathrm{H}'_{\mathrm{L}_\alpha(m)}(m) = m + \mathrm{L}_\alpha(m)$.

Proof. The second equation is easily established by noting that for finite n we have $\mathrm{H}'_n(m) = n + m$. We prove the first equation by induction on α. The case $\alpha = 0$ follows from $\mathrm{L}_0(n) = 0$. For $\alpha \neq 0$ we put $\gamma := \alpha[m]$ and gladly see

$$\mathrm{H}'_\alpha(m) = \mathrm{H}'_\gamma(m+1) = \mathrm{H}'_{\mathrm{L}_\gamma(m+1)}(m+1) = \mathrm{H}'_{\mathrm{L}_\gamma(m+1)+1}(m) = \mathrm{H}'_{\mathrm{L}_\alpha(m)}(m)$$

suffices.

Exercise 21. Let $\alpha \in T$, $m \in \mathbb{N}$ be given. Show that $\mathrm{H}_\alpha(m) \leq \mathrm{H}'_\alpha(m)$. Furthermore show that the second assertion of Exercise 20 fails when $\mathrm{H}_{\omega^\alpha}$ is replaced by $\mathrm{H}'_{\omega^\alpha}$.

Taking the exercise for granted, we see that the newly defined variant of the Hardy hierarchy grows (slightly) faster than the original one. Furthermore the two hierarchies are different. However, we have the following result, showing that the growth-rate of both functions coincide.

Lemma 12. $\displaystyle\bigcup_{\alpha \in T} \mathrm{H}_\alpha \approx \bigcup_{\alpha \in T} \mathrm{H}'_\alpha$.

Proof. The inclusion form left to right follows from the Exercise. To show the other direction it suffices to show

$$\mathrm{H}'_\alpha(m) \leq \mathrm{H}_{\alpha+1}(m) \quad \text{for any } \alpha \in T, m \in \mathbb{N}. \tag{10}$$

First we consider the following claim.

Claim. Let $\lambda \in \mathsf{Lim}$. Then $\lambda[x] + 1 \leq_{(0)} \lambda[x+1]$.

Assume the claim has already established. We proceed by induction on α. The case $\alpha = 0$ is trivial. Consider a successor ordinal $\alpha + 1$. Then

$$\mathrm{H}'_{\alpha+1}(m) = \mathrm{H}'_\alpha(m+1) \leq \mathrm{H}_{\alpha+1}(m+1) = \mathrm{H}_{(\alpha+1)+1}(m)\ .$$

Finally consider a limit ordinal λ. Then we obtain by (ih) and an application of Lemma 10

$$\begin{aligned} \mathrm{H}'_\lambda(m) &= \mathrm{H}'_{\lambda[m]}(m+1) \\ &\leq \mathrm{H}_{\lambda[m]+1}(m+1) \\ &\leq \mathrm{H}_{\lambda[m+1]}(m+1) = \mathrm{H}_\lambda(m+1) = \mathrm{H}_{\lambda+1}(m) \ . \end{aligned}$$

It remains to prove the claim. Let $\lambda \in \mathsf{Lim}$. Due to Theorem 23 we obtain $\lambda[x]+1 \leq \lambda[x+1] < \lambda$. Hence following the pattern of the proof of Lemma 10 we obtain either $\lambda[x]+1 <_{(0)} \lambda[x+1]$ or $\lambda[x]+1 = \lambda[x+1]$. Hence the claim follows.

To see that the name of the hierarchy $\bigcup_{\alpha \in T} \mathrm{G}_\alpha$ is appropriate, we calculate an example. Take e.g. G_ω:

$$\mathrm{G}_\omega(x) = \mathrm{G}_{\psi(\bar{0})\cdot(x+1)}(x) = \mathrm{G}_{x+1}(x) = \mathrm{G}_x(x)+1 = x+1 \ .$$

Furthermore note the following facts.

Exercise 22. $\mathrm{G}_{\varepsilon_0}$ majorizes the elementary functions Elem. F_ω majorizes the primitive recursive functions Prec. That is its growth rate is comparable to the (binary) Ackermann function.

Furthermore the class Mrec of multiple recursive functions can be characterized by

Theorem 26 (Péter 1967, Robbin 1965).

$$\textsc{Mrec} \approx \bigcup_{\gamma<\omega^\omega} \mathrm{F}_\gamma \approx \bigcup_{\gamma<\omega^{\omega^\omega}} \mathrm{H}_\gamma$$

The following theorem states a (surprising) connection between the slow- and fast-growing hierarchy. This theorem shows the difference in the growth-rate between the slow-growing and the fast-growing functions. (Note that due to the second result of Exercise 20 $\mathrm{H}_{\varepsilon_0}$ and $\mathrm{F}_{\varepsilon_0}$ have the same growth rate.) See e.g. Girard [1981], Cichon and Wainer [1983], Weiermann [2001] for further reading on the Hierarchy Comparison Theorem.

Theorem 27 (The Hierarchy Comparison Theorem).

$$\bigcup_{\alpha \in T(K)} \mathrm{G}_\alpha \approx \bigcup_{\gamma<\omega^{K+1}} \mathrm{F}_\gamma \ .$$

Proof. We do not give a detailed proof, but only state the main idea. In Weiermann [2001] the hierarchy comparison theorem has been established for the set of ordinal terms $T(2)$ (built from 0, +, and the binary function symbol ψ). To extend the result to T it suffices to follow the pattern of the proof in Weiermann [2001].

The difficult direction is to show that every function in the hierarchy $\{\mathrm{F}_\gamma \colon \gamma < \omega^{K+1}\}$ is majorized by some G_α. To show this one in particular needs to extend the proofs of Lemma 5 and Theorem 1 in Weiermann [2001] adequately. The reversed direction follows by standard techniques, cf. Cichon and Wainer [1983].

Theorem 28.

$$\bigcup_{\alpha<\psi(\omega,0)} \mathrm{G}_\alpha \approx \bigcup_{\gamma<\omega} \mathrm{F}_\gamma \approx \textsc{Prec} .$$

Proof. Again the difficult direction is to show that every F_γ with $\gamma < \omega$ is majorized by some G_α. This follows directly from Theorem 1 in Weiermann [2001]. To show the other direction one follows the proof of Theorem 27.

5 Simple Termination is Really Really Complex

If termination of a TRS R is established by showing R is compatible with a simplification order, then this can be used to impose an upper bound on the worst case behaviour of R. The worst-case behavior of a terminating TRS can be measured by the lengths of its derivations.

First we define $\mathrm{dl}_R \colon \mathcal{T}(\Sigma) \to \mathbb{N}$ in such a way that $\mathrm{dl}_R(s)$ is the maximal length of a derivation starting from s:

$$\mathrm{dl}_R(s) := \max\{\mathrm{dl}_R(t) + 1 : s \to_R t\},$$

using the convention $\max\emptyset = 0$. This is well-defined by an invocation of König's Lemma since the finiteness of R implies $\to_R$ is finitely branching. Now we introduce the *derivation length function* or *complexity* $\mathrm{Dl}_R \colon \mathbb{N} \to \mathbb{N}$ by

$$\mathrm{Dl}_R(n) := \max\{\mathrm{dl}_R(s) : s \in \mathcal{T}(\Sigma) \wedge \mathrm{dp}(s) \leqslant n\}.$$

The condition $\mathrm{dp}(s) \leqslant n$ guarantees weak monotonicity. It suffices to consider only closed terms because any derivation may be transformed, under preservation of depths and sizes, to a derivation of equal length containing only closed terms: simply apply a substitution which maps all variables to constants. One glance at Definition 24 shows dl_R corresponds to a certain order type.

Lemma 13. *For any terminating TRS R we have* $\mathrm{otype}(\mathcal{T}(\Sigma), \overset{+}{\leftarrow}_R) \leqslant \omega$ *since* $\mathrm{dl}_R(s)$ *is just the order type of s in the partial order* $(\mathcal{T}(\Sigma), \overset{+}{\leftarrow}_R)$.

We know from Theorems 5 and 6 that

- a termination proof via MPO implies there is a primitive recursive bound on the derivation length function and
- a termination proof via LPO implies there is a multiple recursive bound on the derivation length function.

Both results are essentially optimal. In Chapter 6 we will encounter proofs of these Theorems which relate the order types of the involved orders to the slow growing hierarchy of Definition 43. In the general case, slow growing functions have to be replaced with Hardy functions.

Theorem 29 (Weiermann 1994). *For every simply terminating TRS R there exists $\alpha < \Lambda$ such that the complexity of R is dominated by* H_α.

We recall from Theorem 22 that Λ is the supremum of the order types of simplification orders. One may wonder if there is a closer connection between the complexity of the TRS and the order type of the compatible order. And indeed, there is a rather close connection.

Theorem 30 (Buchholz et al. 1994, Theorem 1). *If a TRS R is compatible with a simplification order of order type α then the complexity of R is dominated by some H_β with $\beta < \alpha + \omega^\omega$.*

The aim of this section is to show that these gigantic bounds (recall from Theorem 26 that all multiple recursive functions are eventually dominated by H_α for $\alpha = \omega^{\omega^\omega}$) are essentially optimal. For $k \geqslant 1$ we construct TRSs R_k which are able to simulate, for all $\alpha < \Delta_{k+1}$, computational processes which are closely related to computing $\mathrm{H}_\alpha(n)$. More precisely, we will see below that the Hardy function is related to an iterated application of fundamental sequences in the following sense:

$$\mathrm{H}_\alpha(n) = n + \min\{m : \alpha[n][n+1]\dots[n+m-1] = 0\}$$

For reasons which will be elucidated later this iterated application is called the *Hydra battle*. The battles will be simulated by the R_k, yielding a complexity Dl_{R_k} related to $\mathrm{H}_{\Delta_{k+1}}$.

The structure of this section is as follows. In Section 5.2 we have to give a slight reformulation of the notation system for Λ of Definition 37 since in a term rewriting setting we have to stick to a $+$ with fixed arity. Consequently, we have to adapt the definition of fundamental sequences to this new setting. For a fixed $k \geqslant 1$, the TRS R_k mentioned above is then introduced in Section 5.2, and it is shown that its complexity eventually dominates all H_α with $\alpha < \Delta_{k+1}$. Total termination of R_k is established in Section 5.3 using Touzet's technically smooth characterization of total termination from Theorem 10. The R_k are given in a uniform manner, and for $k > l$ the TRS R_k can be regarded as a proper extension of R_l. Thus the R_k constitute a hierarchy of totally terminating TRSs, and the complexity of any simply terminating TRS is eventually dominated by the complexities of almost all R_k. This stepwise approach from below is inevitable as, due to Theorem 29, it is not possible to define a simply terminating TRS which is able to simulate Hydra battles for all ordinals below Λ.

We are now going to formally introduce Hydrae and Hydra battles and to show the connection between these battles and the Hardy hierarchy.

Definition 46. *For $m \in \mathbb{N}$ we introduce*

$$\alpha[n,m] := \begin{cases} \alpha & \textit{if } n > m\ , \\ (\alpha[n,m-1])[m] & \textit{otherwise.} \end{cases}$$

Thus $\alpha[n,m] = \alpha[n][n+1]\dots[m]$. The main results concerning subrecursive hierarchies in Chapter 4 are related to the ordinal notation system $T(k)$, cf. Definition 33.

The *battle of Hercules and the Hydra* from Kirby and Paris [1982] is closely connected to the Hardy functions. This is emphasized by the fact that *Hardy* is an anagram for *Hydra*. The name of the battle refers to the famous fight of Hercules against the Hydra, a multi-headed monster. In the battle, Hercules tries to chop off the heads of the Hydra with his sword, but each time he chops off one head a couple of new heads grow out of the fresh wound. However, Hercules wins the fight as soon as he realizes that the heads stop to grow when he burns fresh wounds with a torch.

The connection to the theoretical battle defined below is as follows. We take the place of Hercules and fight against a Hydra configuration (α, n) with $\alpha < \Delta_{k+1}$ and $n \in \mathbb{N}$. At each configuration (β, m) of the battle we manage to chop off something from the Hydra, reaching the new configuration $(\beta[m], m+1)$. Thus many new heads (in the term) grow as long as β is a limit ordinal, whereas no new heads appear only if β is a successor or zero. Of course nothing new happens when we arrive at 0.

Definition 47. *A* Hydra *is an ordinal below Δ_{k+1}. For each Hydra α the ordered pair $c := (\alpha, n)$ is called a* configuration*. The* next configuration *c^+ for c is $(\alpha[n], n+1)$, and the* Hydra battle *for the configuration c is the sequence $(c_m)_{m<\omega}$ of configurations with $c_0 = c$ and $c_{m+1} = c_m^+$. The minimal m such that the Hydra in c_m is 0 is called the* length *of the battle.*

An immediate consequence of Theorem 23 is $(\forall\alpha > 0)(\forall n)(\alpha > \alpha[n])$, thus the length of the battle is well-defined for each configuration. The following Lemma gathers a few basic facts about Hydra battles, and it links the length of a battle to the counting functions L_α of Definition 45.

Lemma 14. *Let $c := (\alpha, n)$ be a configuration and consider its battle $(c_m)_{m\in\mathbb{N}}$.*

i. *For all m we have $c_m = (\alpha[n, n+m-1], n+m)$.*
ii. *The length of the battle is the minimal m such that $\alpha[n, n+m-1] = 0$.*
iii. *The length of the battle is $\mathrm{L}_\alpha(n)$.*

Proof. We can show (i) by an induction on m, and (ii) is an easy consequence of this. An induction on α establishes (iii).

The counting functions are tailored exactly for the Hydra battle, and hence the Hardy functions are also closely related to the battle.

Proposition 8. *The function which maps n to the length of the battle for the configuration (Δ_{k+1}, n) eventually dominates all H_α with $\alpha < \Delta_{k+1}$.*

Proof. We combine Lemma 11 and Lemma 14.iii to see that the length of the battle for (Δ_{k+1}, n) is $\mathrm{H}'_{\Delta_{k+1}}(n) - n$. Since Δ_{k+1} is a limit ordinal, Theorem 25 and (10) establish the claim.

5.1 Encoding all Hydrae

Since we work in a rewriting setting, we have to encode all Hydrae below Δ_{k+1} by terms using a $+$ of finite arity. For these terms the fundamental sequences of Definition 37 have to be adapted.

Definition 48. *The signature Σ_0 consists of*

- *the constant* $\mathsf{0}$,
- *the unary (successor)* S,
- *the binary* $+$, *and*
- *the* $k+1$*-ary* P, *which represents* ψ.

Each $s \in \mathcal{T}(\Sigma_0)$ *has a* value $\mathrm{val}(s) < \Delta_{k+1}$, *which is calculated by interpreting* $\mathsf{0}$, S, $+$, P *with* 0, *the successor function, the ordinal sum, and* ψ, *respectively.*

The symbol S is not really needed for encoding Hydrae since the terms $\mathsf{S}s$ and $+(s, \mathsf{P}(\bar{\mathsf{0}}))$ have the same value. We use S as a syntactic indicator of successor ordinals, and additionally its presence will simplify various calculations.

Obviously, each ordinal below Δ_{k+1} can be denoted by terms of $\mathcal{T}(\Sigma_0)$. To mimic fundamental sequences on terms we introduce a set of *standard terms*. Because our $+$ has fixed arity and because we do not want to bother about distinguishing between $+(+(s,t),u)$ and $+(s,+(t,u))$, there will usually be distinct standard terms denoting the same ordinal. Furthermore, for $\gamma < \Delta_{k+1}$ denoted by the standard term s and for $n > 0$ we intend to denote $\gamma + n$ by $\mathsf{S}^n s$, thus using $+$ only for certain additions of (standard terms for) limit ordinals. This will be useful later when we treat derivations of standard terms.

Recall from Definition 29 that $\oplus$ denotes the natural sum of two ordinals. A pair (λ, μ) of limit ordinals is called *compatible* if $\lambda + \mu = \lambda \oplus \mu$ holds.

Definition 49. *The set* $\mathcal{D} \subseteq \mathcal{T}(\Sigma_0)$ *of* standard terms *is the smallest superset of* $\{\mathsf{0}\}$ *which is closed under* S *and these rules:*

- $\bar{s} \in \mathcal{D}$ *and* $\bar{s} \neq \bar{\mathsf{0}} \implies \mathsf{P}(\bar{s}) \in \mathcal{D}$,
- $s, t \in \mathcal{D}$ *and* $(\mathrm{val}(s), \mathrm{val}(t))$ *compatible* $\implies +(s,t) \in \mathcal{D}$.

By $\mathcal{D}(\alpha)$ *we denote the collection of standard terms with value* α.

The following Lemma should be no surprise.

Lemma 15.

i. *If* s' *is a proper subterm of* $s \in \mathcal{D}(\alpha)$, *then there is* $\beta < \alpha$ *with* $s' \in \mathcal{D}(\beta)$.
ii. *For all* $\alpha < \Delta_{k+1}$ *we have* $\mathcal{D}(\alpha) \neq \emptyset$.
iii. $\mathcal{D}$ *is the union of the* $\mathcal{D}(\alpha)$ *with* $\alpha < \Delta_{k+1}$.

Definition 50. *A formal* multiplication *for a term* s *and* $n > 0$ *is defined by*

$$s \times n := +(\cdot, s)^{n-1}(s) \ .$$

It will be important that the recursion occurs in the first argument.

Our aim is now to mimic Definition 37, the definition of $\alpha[n]$, for the members of $\mathcal{D}$. We encounter some difficulties on the way. Our rather granulated definition of standard terms implies that a formal equivalent to fundamental sequences for standard terms will not always produce standard terms. For example, if we defined $\mathsf{P}(\bar{\mathsf{0}}, \mathsf{S}s)[n]$ to be $\mathsf{P}(\bar{\mathsf{0}}, s) \times (n+1)$, this would result in occurrences of the nonstandard term $\mathsf{P}(\bar{\mathsf{0}})$ for $s = \mathsf{0}$. We overcome this obstacle for $d \in \mathcal{D}(\alpha)$ by simultaneously defining $d\langle n\rangle$ and $d[n]$, where $d\langle n\rangle$ is a formal equivalent to $\alpha[n]$ which need not be a standard term but has a uniform definition, while $d[n]$ is a refinement of $d\langle n\rangle$ and an element of $\mathcal{D}(\alpha[n])$. Later we will work in a TRS which is able to reduce $d\langle n\rangle$ to $d[n]$.

Definition 51. *By $\mathcal{D}(\mathsf{Lim})$ we denote the set of standard terms whose values are limit ordinals, while the analogies to $\mathsf{Fix}(\bar{\alpha})$ and $\mathrm{MS}_{\bar{\alpha}}(\bar{\beta})$ on $\mathcal{D}$ are called $\mathsf{Fix}(\bar{s})$ and $\mathrm{MS}_{\bar{s}}(\bar{t})$.*

Definition 52. *For $d \in \mathcal{D}$ and $n \in \mathbb{N}$ we simultaneously define $d\langle n\rangle$ and $d[n]$, both members of $\mathcal{T}(\Sigma_0)$, by recursion on $\mathcal{D}$:*

$$\mathsf{0}\langle n\rangle := \mathsf{0} \tag{11a}$$
$$\mathsf{S}s\langle n\rangle := s \tag{11b}$$
$$+(s,t)\langle n\rangle := +(s,t[n]) \tag{11c}$$
$$\mathsf{P}(\bar{s},t)\langle n\rangle := \mathsf{P}(\bar{s},t[n]) \qquad \text{if } t \in \mathcal{D}(\mathsf{Lim}) \setminus \mathsf{Fix}(\bar{s}) \tag{11d}$$
$$\mathsf{P}(\bar{\mathsf{0}},\mathsf{S}t)\langle n\rangle := \mathsf{P}(\bar{\mathsf{0}},t) \times (n+1) \tag{11e}$$
$$\mathsf{P}(\bar{\mathsf{0}},t)\langle n\rangle := t \times (n+1) \tag{11f}$$
$$\mathsf{P}(s_1,\ldots,\mathsf{S}s_i,\bar{\mathsf{0}},\mathsf{S}t)\langle n\rangle := \mathsf{P}(\bar{s},\cdot,\bar{\mathsf{0}})^{n+1}(\mathsf{P}(s_1,\ldots,\mathsf{S}s_i,\bar{\mathsf{0}},t)) \tag{11g}$$
$$\mathsf{P}(s_1,\ldots,\mathsf{S}s_i,\bar{\mathsf{0}},t)\langle n\rangle := \mathsf{P}(\bar{s},\cdot,\bar{\mathsf{0}})^{n+1}(t) \tag{11h}$$
$$\mathsf{P}(s_1,\ldots,s_i,\bar{\mathsf{0}},\mathsf{0})\langle n\rangle := \mathsf{P}(s_1,\ldots,s_i[n],\bar{\mathsf{0}},\mathrm{MS}_{\bar{s},\bar{\mathsf{0}}}(\bar{s})) \tag{11i}$$
$$\mathsf{P}(s_1,\ldots,s_i,\bar{\mathsf{0}},\mathsf{S}t)\langle n\rangle := \mathsf{P}(s_1,\ldots,s_i[n],\bar{\mathsf{0}},\mathsf{P}(s_1,\ldots,s_i,\bar{\mathsf{0}},t)) \tag{11j}$$
$$\mathsf{P}(s_1,\ldots,s_i,\bar{\mathsf{0}},t)\langle n\rangle := \mathsf{P}(s_1,\ldots,s_i[n],\bar{\mathsf{0}},t)\ . \tag{11k}$$

Similar to Definition 37, $s_i \neq \mathsf{0}$ is required for (11i)–(11k).

If $d = +(s,t)$ and $t[n] = \mathsf{S}^i t'$ where i is as large as possible, we put

$$+(s,t)[n] := \begin{cases} \mathsf{S}^i s & \text{if } t' = \mathsf{0}\ , \\ \mathsf{S}^i{+}(s,t') & \text{otherwise.} \end{cases}$$

Moreover, we demand $\mathsf{P}(\bar{\mathsf{0}},\mathsf{S0})[n] := \mathsf{S}^{n+1}\mathsf{0}$ as well as

$$\mathsf{P}(\mathsf{0},\ldots,\mathsf{S0},\bar{\mathsf{0}},\mathsf{0})[n] := \mathsf{P}(\mathsf{0},\ldots,\mathsf{0},\cdot,\bar{\mathsf{0}})^n(\mathsf{S0})\ ,$$

and in all remaining cases we put $d[n] := d\langle n\rangle$. For $m \in \mathbb{N}$ we further introduce $d[n,m]$ in analogy with the $\alpha[n,m]$ of Definition 46.

We now show that this definition is correct and meets our requirements.

Lemma 16. *Let $\alpha < \Delta_{k+1}$ and $n \in \mathbb{N}$. For $d \in \mathcal{D}(\alpha)$ we have $\mathrm{val}(d\langle n\rangle) = \alpha[n]$ and $d[n] \in \mathcal{D}(\alpha[n])$.*

Proof (Proof by induction on $\mathcal{D}$). As the definition of $d\langle n\rangle$ just copies Definition 37, the first statement is immediate from the induction hypothesis. Because in the above definition recursion is only used for standard terms s and t denoting limit ordinals, we have $s[n] \neq \mathsf{0}$ and $t[n] \neq \mathsf{0}$ according to the induction hypothesis and Theorem 23. Since $\mathrm{MS}_{\bar{s},\bar{\mathsf{0}}}(\bar{s})$, being a subterm of some $s_j \in \mathcal{D}$, cannot be $\mathsf{P}(\bar{\mathsf{0}})$, the only possible occurrences of $\mathsf{P}(\bar{\mathsf{0}})$ in $d\langle n\rangle$ are the ones we gave special treatment in the definition of $d[n]$. In both cases it is obvious that $d\langle n\rangle$ and $d[n]$ have the same value and that $d[n]$ is standard.

Now let $d = +(s,t)$ and let i, t' be as in the definition of $d[n]$. Since d is standard, we know the pair (τ,μ) with $\tau := \mathrm{val}(s)$ and $\mu := \mathrm{val}(t)$ is compatible. The statement obviously holds if $t' = \mathsf{0}$. So let t' denote a limit ordinal, say μ'. By Definition 49 we have to show (τ,μ') is compatible. This is done by proving $\mu' \leqslant \mu$. The induction hypothesis yields $t[n]$ is a standard term with $\mathrm{val}(t[n]) = \mu[n]$. Because μ is a limit, Theorem 23 implies $\mu > \mu[n]$ holds, thus $\mu' \leqslant \mu[n] < \mu$. So $d[n]$ is standard and has the correct value. In the remaining cases the statement easily follows from the induction hypothesis.

5.2 Simulating all Hydra battles

We are now prepared to gradually define the TRS R, which is intended to simulate all Hydra battles below $\varDelta_{k+1}$. Therefore $\varSigma_0$ has to be enlarged by new symbols whose meaning will be elucidated in the following definitions.

Definition 53. *The signature $\varSigma$ consists of $\varSigma_0$ enriched by*

- *the unary* $\bullet$, $\circ$, *and* $[]$,
- *the* $k+1$*-ary* M,
- *the* $i+1$*-ary* J_i, *for* $1 \leqslant i \leqslant k$,
- *the* $i+1$*-ary* Q_{ij}, *for* $1 \leqslant j \leqslant i \leqslant k$, *and*
- *the* $i+2$*-ary* R_i, *for* $1 \leqslant i \leqslant k$.

It will sometimes be necessary to reduce a term to one of its subterms. Since R is intended to be totally and thus simply terminating, we may introduce, for all symbols $f \in \varSigma^{(n)}$ with $n > 0$ and for $1 \leqslant i \leqslant n$, embedding rules

$$(\mathrm{S}_i f) \quad f(x_1,\ldots,x_n) \to x_i \; .$$

Because $\mathsf{0}$ is the only constant, the following result is blatantly trivial, even though we did not yet define the whole of R.

Lemma 17. *Each $s \in \mathcal{T}(\varSigma)$ reduces in less than $\mathrm{dp}(s)+1$ steps to its unique normal form $\mathsf{0}$. If s' is a subterm of s, then $s \xrightarrow{*} s'$.*

The promised rules which enable us to reduce $d\langle n\rangle$ to $d[n]$ are

$$(\mathrm{F1}) \quad \mathsf{P}(\bar{\mathsf{0}}) \to \mathsf{S0}\;, \qquad (\mathrm{F2}) \quad +(x,\mathsf{S}y) \to \mathsf{S}{+}(x,y)\;.$$

Lemma 18. *For $d \in \mathcal{D}$ we have $d\langle n\rangle \xrightarrow{*} d[n]$.*

Proof. The difference between $d\langle n\rangle$ and $d[n]$ for $d = \mathsf{P}(\bar{\mathsf{0}},\mathsf{S0},\bar{\mathsf{0}},\mathsf{0})$ consists of one single $\mathsf{P}(\bar{\mathsf{0}})$ which is replaced with $\mathsf{S0}$. This can be handled by (F1). For $d = \mathsf{P}(\bar{\mathsf{0}},\mathsf{S0})$ and $n > 0$ we have to show $\mathsf{P}(\bar{\mathsf{0}}) \times n \xrightarrow{+} \mathsf{S}^n\mathsf{0}$, which is done by induction on n. The case $n = 1$ is again established by (F1), while the induction hypothesis yields $\mathsf{P}(\bar{0}) \times (n+1) \xrightarrow{+} +(\mathsf{S}^n\mathsf{0}, \mathsf{P}(\bar{0}))$. Now we get

$$+(\mathsf{S}^n\mathsf{0}, \mathsf{P}(\bar{0})) \to_{\mathrm{F1}} +(\mathsf{S}^n\mathsf{0}, \mathsf{S}\mathsf{0}) \to_{\mathrm{F2}} \mathsf{S}{+}(\mathsf{S}^n\mathsf{0}, \mathsf{0}) \to_{\mathrm{S}_1+} \mathsf{S}^{n+1}\mathsf{0} \ .$$

It suffices for the remaining case $d = +(s, t)$ to note that

$$+(s, \mathsf{S}^i t') \xrightarrow{*}_{\mathrm{F2}} \mathsf{S}^i{+}(s, t') \to_{\mathrm{S}_1+} \mathsf{S}^i s$$

is possible for arbitrary i and t'.

Following Touzet [1998], R is to regard $\bullet[]^{n+1}d$ with $d \in \mathcal{D}$ as a term which encodes the battle configuration $(\mathrm{val}(d), n)$. Since we want to simulate Hydra battles at full length we intend, for $d \neq \mathsf{0}$, to make possible derivations $\bullet[]^{n+1}d \xrightarrow{+} \bullet[]^{n+2}d[n]$, which can then be iterated until $\bullet[]^{n+m}\mathsf{0}$ is reached.

For some calculations it will be necessary to facilitate $\bullet[]^{n+1}d \xrightarrow{+} []^{n+1}\bullet^{n+1}d$, so that $\bullet^{n+1}$ may be moved to the top of subterms of d as material which can be deleted in subderivations. When we reach a point where d can safely be modified into something close to $d\langle n\rangle$, we do so and put a $\circ$ on top of the new subterm. This $\circ$ will enable us to create $\bullet[]$ in front of $[]^{n+1}d[n]$, furthermore, recursions like the one needed for (11c) can be simulated. The required rules are variations on rules of Touzet [1998]:

$$\begin{array}{llll} (\mathrm{N1}) & \bullet[]x \to []\bullet\bullet x \ , & (\mathrm{N2}) & []\circ x \to \circ[][]x \ , \\ (\mathrm{N3}) & \circ x \to []x \ , & (\mathrm{N4}) & []x \to \bullet x \ . \end{array}$$

Lemma 19. *For $n > 0$ and $s \in \mathcal{T}(\Sigma)$ we have*

i. $[]^n s \xrightarrow{+} \bullet^n s \xrightarrow{+} s$
ii. $\bullet[]^n s \xrightarrow{+} []^n\bullet^n s$
iii. $[]^n\circ s \xrightarrow{+} \bullet[]^{n+1}s$.

Proof. For (i) we rely on (N4) and $(\mathrm{S}_1\bullet)$, while (ii) follows from (i) and $\bullet[]^n s \xrightarrow{+} []^n\bullet^{2^n}s$, which is shown by induction on n using

$$\bullet^m[]s \xrightarrow{*}_{\mathrm{N1}} []\bullet^{2m}s \ , \tag{12}$$

which in turn is shown by induction on $m \geqslant 0$ using (N1). We get (iii) from

$$[]^n\circ s \xrightarrow{+}_{\mathrm{N2}} \circ[]^{2n}s \xrightarrow{*} \circ[]^{n+1}s \to_{\mathrm{N3}} [][]^{n+1}s \to_{\mathrm{N4}} \bullet[]^{n+1}s \ .$$

Its first step is won like (12), and the second one relies on $2n \geqslant n+1$ and (i).

To simulate cases like (11c) we have to import $\bullet$ and $[]$ into standard terms. For $f \in \{\mathsf{S}, +, \mathsf{P}\}$ with arity n and for $1 \leqslant i \leqslant n$ we thus introduce the rule

$$(\mathrm{D}_i f) \quad \bullet f(\bar{x}) \to f(x_1, \ldots, []x_i, \ldots, x_n) \ .$$

Lemma 20. *For $s, t, \bar{s} \in \mathcal{T}(\Sigma)$; $n \geqslant 0$, and $1 \leqslant i \leqslant k+1$ we have*

i. $\bullet^{n+1}{+}(s,t) \xrightarrow{+} {+}(s, \bullet^{n+1} t)$
ii. $\bullet^{n+1}\mathsf{P}(\bar{s}) \xrightarrow{*} \bullet\mathsf{P}(s_1, \ldots, \bullet^n s_i, \ldots, s_{k+1}) \xrightarrow{+} \mathsf{P}(s_1, \ldots, \bullet^{n+1} s_i, \ldots, s_{k+1})$
iii. $\bullet^{n+1}\mathsf{P}(\bar{s}) \xrightarrow{+} \mathsf{P}(s_1, \ldots, [\!]^{n+1} s_i, \ldots, s_{k+1})$
iv. $\bullet^{n+1}\mathsf{P}(s_1, \ldots, \mathsf{S}s_i, \ldots, s_{k+1}) \xrightarrow{+} \mathsf{P}(s_1, \ldots, \mathsf{S}[\!]^{n+1} s_i, \ldots, s_{k+1})$.

Proof. To settle (i), ($\mathrm{D}_2{+}$) is applied $n+1$ times, and afterwards we rely on Lemma 19.i. With little changes, using ($\mathrm{D}_i\mathsf{P}$) and ($\mathrm{D}_1\mathsf{S}$) instead of ($\mathrm{D}_2{+}$), the remaining points follow.

As mentioned earlier, importing $\bullet^n$ or $[\!]^n$ shall enable us to locally reduce until it is safe to create a $\circ$ on top of the subterm we treated. Sometimes such a $\circ$ has to be exported. This is achieved by these rules:

$$
\begin{array}{lll}
(\mathrm{E}_2{+}) & {+}(x, \circ y) \to \circ{+}(x,y) & \\
(\mathrm{E}_i\mathsf{P}) & \mathsf{P}(x_1, \ldots, \circ x_i, \ldots, x_{k+1}) \to \circ\mathsf{P}(\bar{x}) & \text{for } 1 \leqslant i \leqslant k+1\ .
\end{array}
$$

In order to simulate (11e)–(11h) we need rewrite rules for a special kind of multiplication and for iterations of P. Since multiplication amounts to iterating $+$, the rules are very similar:

$$
\begin{array}{lll}
(\mathrm{R}\mathsf{M}) & \mathsf{M}(\bar{x}, [\!]y) \to {+}(\mathsf{M}(\bar{x},y), \mathsf{P}(\bar{x},y)) & \\
(\mathrm{R}\mathsf{J}_i) & \mathsf{J}_i(x_1, \ldots, [\!]x_i, y) \to \mathsf{P}(\bar{x}, \mathsf{J}_i(\bar{x},y), \bar{0}) & \text{for } 1 \leqslant i \leqslant k\ .
\end{array}
$$

Lemma 21. *For $\bar{s}, t \in \mathcal{T}(\Sigma)$; $n > 0$, and $1 \leqslant i \leqslant k$ we have*

i. $\mathsf{M}(\bar{s}, [\!]^n t) \xrightarrow{+} \mathsf{P}(\bar{s}, t) \times n$
ii. $\mathsf{J}_i(s_1, \ldots, [\!]^n s_i, t) \xrightarrow{+} \mathsf{P}(s_1, \ldots, s_i, \cdot, \bar{0})^n(t)$.

Proof. As both statements are treated similarly by induction on n, we only prove (i) in detail. For the start we have

$$\mathsf{M}(\bar{s}, [\!]t) \to_{\mathrm{RM}} {+}(\mathsf{M}(\bar{s},t), \mathsf{P}(\bar{s},t)) \to_{\mathrm{S}_2+} \mathsf{P}(\bar{s},t)\ ,$$

and the induction step is

$$
\begin{aligned}
\mathsf{M}(\bar{s}, [\!]^{n+1} t) &\to_{\mathrm{RM}} {+}(\mathsf{M}(\bar{s}, [\!]^n t), \mathsf{P}(\bar{s}, [\!]^n t)) \\
&\xrightarrow{+} \quad {+}(\mathsf{M}(\bar{s}, [\!]^n t), \mathsf{P}(\bar{s}, t)) \\
&\xrightarrow{+} \quad {+}(\mathsf{P}(\bar{s}, t) \times n, \mathsf{P}(\bar{s}, t))\ ,
\end{aligned}
$$

where we used Lemma 19.i and the induction hypothesis for the last two steps. Statement (ii) relies on ($\mathrm{R}\mathsf{J}_i$) and ($\mathrm{S}_{i+1}\mathsf{J}_i$) instead of (RM) and ($\mathrm{S}_2{+}$).

We now present the rules intended to carry out the transformations prescribed by Definition 52. Because it is not easy to distinguish between cases like (11d) and (11k), Lemma 20 and the rules ($\mathrm{E}_i\mathsf{P}$) facilitate both possible transformations. Hence R is also able to simulate wrong battles, i.e. battles based on assignments of fundamental sequences which differ from our assignment. In the wrong cases the ordinals denoted are smaller, thus this shall pose no problem to our intended normalization. Likewise, transformations as the one needed in (11h) are made possible for arbitrary terms t, and the t in (11f) is supposed to be an element of $\mathsf{Fix}(\bar{0})$ whereas the associated rule (H3) below treats arbitrary terms beginning with P. The final rules are:

$$
\begin{array}{lr@{\;}l}
\text{(H1)} & \bullet\mathsf{S}x \to & \circ x \\
\text{(H2)} & \mathsf{P}(\bar{0},\mathsf{S}y) \to & \circ\mathsf{M}(\bar{0},y) \\
\text{(H3)} & \mathsf{P}(\bar{0},\mathsf{P}(\bar{x},y)) \to & \circ\mathsf{M}(\bar{x},y) \\
(\mathrm{H}_i4) & \mathsf{P}(x_1,\dots,\mathsf{S}x_i,\bar{0},y) \to & \circ\mathsf{J}_i(\bar{x},y) \\
(\mathrm{H}_i5) & \mathsf{P}(x_1,\dots,\mathsf{S}x_i,\bar{0},\mathsf{S}y) \to & \circ\mathsf{J}_i(\bar{x},\mathsf{P}(x_1,\dots,\mathsf{S}x_i,\bar{0},y)) \\
(\mathrm{H}_{ij}6) & \bullet\mathsf{P}(x_1,\dots,x_i,\bar{0},0) \to & \mathsf{Q}_{ij}(x_1,\dots,\bullet x_i,x_j) \\
(\mathrm{R}\mathsf{Q}_{ij}) & \mathsf{Q}_{ij}(x_1,\dots,\circ x_i,y) \to & \circ\mathsf{P}(\bar{x},\bar{0},y) \\
(\mathrm{H}_i7) & \bullet\mathsf{P}(x_1,\dots,x_i,\bar{0},\mathsf{S}y) \to & \mathsf{R}_i(x_1,\dots,\bullet x_i,x_i,y) \\
(\mathrm{R}\mathsf{R}_i) & \mathsf{R}_i(x_1,\dots,\circ x_i,y,z) \to & \circ\mathsf{P}(\bar{x},\bar{0},\mathsf{P}(x_1,\dots,x_{i-1},y,\bar{0},z))
\end{array}
$$

for i and j with $1 \leqslant j \leqslant i \leqslant k$.

Proposition 9. *For $d \in \mathcal{D}$ with $d \neq 0$ and for $n \geqslant 0$ we have*

$$\bullet\rrbracket^{n+1}d \xrightarrow{+} \rrbracket^{n+1}\bullet^{n+1}d \xrightarrow{+} \rrbracket^{n+1}\circ d[n] \xrightarrow{+} \begin{cases}\circ d[n] \\ \bullet\rrbracket^{n+2}d[n]\ .\end{cases}$$

Before we prove this, we would like to point out its main implication. By Lemma 16, R is able to simulate Hydra battles for all configurations (α,n) with $0 < \alpha < \Delta_{k+1}$ at full length:

$$\bullet\rrbracket^{n+1}d \xrightarrow{+} \bullet\rrbracket^{n+2}d[n] \xrightarrow{+} \bullet\rrbracket^{n+3}d[n,n+1] \xrightarrow{+} \dots \xrightarrow{+} \bullet\rrbracket^{n+l}0 \xrightarrow{+} 0\ .$$

It is obviously not wise to strive after a similar result for $d = 0$.

Proof (Proof by induction on $\mathcal{D}$). As mentioned before, a close look at Definition 52 shows that whenever $u[n]$ (with u being a subterm of d) is used to define $d[n]$ we have $u \neq 0$. This observation enables us to rely on the induction hypothesis when required.

A quick glance at Lemmata 18 and 19 assures us that it suffices to show

$$\bullet^{n+1}d \xrightarrow{+} \circ d\langle n\rangle\ .$$

For (11b) we employ Lemma 19.i to get $\bullet^{n+1}\mathsf{S}s \xrightarrow{*} \bullet\mathsf{S}s \to_{\mathrm{H1}} \circ s = \circ d\langle n\rangle$, while (11c) is handled by Lemma 20 and the induction hypothesis:

$$\bullet^{n+1}+(s,t) \xrightarrow{+} +(s,\bullet^{n+1}t) \xrightarrow{+} +(s,\circ t[n]) \to_{\mathrm{E}_2+} \circ+(s,t[n])\ .$$

The treatment of (11d) and (11k) is very close to this, relying on $(\mathrm{E}_i\mathsf{P})$ instead of (E_2+). For (11f) we note that $d = \mathsf{P}(\bar{0},u)$ with $u \in \mathsf{Fix}(\bar{0})$ holds. According to (8), u is some $\mathsf{P}(\bar{s},t)$. Applying Lemmata 20.ii,iii and 21.i,

$$\begin{aligned} \bullet^{n+1}\mathsf{P}(\bar{0},\mathsf{P}(\bar{s},t)) &\xrightarrow{+} \mathsf{P}(\bar{0},\mathsf{P}(\bar{s},[\!]^{n+1}t)) \\ &\to_{\mathrm{H3}} \circ\mathsf{M}(\bar{s},[\!]^{n+1}t) \\ &\xrightarrow{+} \circ(\mathsf{P}(\bar{s},t)\times(n+1)) \end{aligned}$$

follows. The proof of (11e) is very similar (using (H2)) and therefore left out here. For (11h) we need Lemmata 20.iv and 21.ii:

$$\begin{aligned} \bullet^{n+1}\mathsf{P}(s_1,\ldots,\mathsf{S}s_i,\bar{0},t) &\xrightarrow{+} \mathsf{P}(s_1,\ldots,\mathsf{S}[\!]^{n+1}s_i,\bar{0},t) \\ &\to_{\mathrm{H}_i4} \circ\mathsf{J}_i(s_1,\ldots,[\!]^{n+1}s_i,t) \\ &\xrightarrow{+} \circ\mathsf{P}(s_1,\ldots,s_i,\cdot,\bar{0})^{n+1}(t)\ , \end{aligned}$$

while we care for (11g) in much the same way, replacing (H_i4) by (H_i5). The treatment of $\mathrm{MS}_{\bar{s},\bar{0}}(\bar{s})$ in (11i) requires a new idea, since R does not know which of the s_j has $\mathrm{MS}_{\bar{s},\bar{0}}(\bar{s})$ as a subterm. By virtue of Lemma 20.ii and the induction hypothesis, for each j with $1 \leqslant j \leqslant i$ and for $s_i \neq \mathsf{0}$, we can show

$$\begin{aligned} \bullet^{n+1}\mathsf{P}(s_1,\ldots,s_i,\bar{0},\mathsf{0}) &\xrightarrow{*} \bullet\mathsf{P}(s_1,\ldots,\bullet^{n}s_i,\bar{0},\mathsf{0}) && (13a) \\ &\xrightarrow{+} \mathsf{Q}_{ij}(s_1,\ldots,\bullet^{n+1}s_i,s_j) && (13b) \\ &\xrightarrow{+} \mathsf{Q}_{ij}(s_1,\ldots,\circ s_i[n],s_j) \\ &\to_{\mathsf{RQ}_{ij}} \circ\mathsf{P}(s_1,\ldots,s_i[n],\bar{0},s_j)\ . && (13c) \end{aligned}$$

To get from (13a) to (13b), we make use of $(\mathrm{H}_{ij}6)$ and, in case of $j = i$, Lemma 19.i. Now we may incorporate Lemma 17 to reduce the second s_j in (13c) to any of its subterms. Since $\mathrm{MS}_{\bar{s},\bar{0}}(\bar{s})$ is a subterm of some s_j, we reach our goal. When we use (H_i7) and (RR_i) instead of $(\mathrm{H}_{ij}6)$ and (RQ_{ij}), the result for (11j) is easily obtained.

Corollary 2. *If R terminates, then its complexity Dl_R eventually dominates all H_α with $\alpha < \Delta_{k+1}$.*

Proof. As mentioned in (7), $\Delta_{k+1} = \psi(1,\bar{0})$ holds where $\bar{0}$ has length $k+1$. Proposition 8 tells us the function which maps n to the length of the Hydra battle for $c_n := (\Delta_{k+1}[n], n+1)$ eventually dominates all H_α with $\alpha < \Delta_{k+1}$. We take a short digression from our fixed k to $k+1$, recall $\psi_{0,\bar{\alpha}} = \psi_{\bar{\alpha}}$, and see

$$\Delta_{k+1}[n] = \psi(0,\cdot,\bar{0})^{n+1}(0) = \psi(\cdot,\bar{0})^{n+1}(0)$$

holds where the first ψ is $k+2$-ary and the second one is, as usual, $k+1$-ary. Since $s_n := \bullet[\!]^{n+2}\mathsf{P}(\cdot,\bar{0})^n(\mathsf{S0})$ encodes c_n, the length of the battle for c_n is majorized by $\mathrm{dl}_R(s_n)$. Because of $\mathrm{dp}(s_{n+1}) = \mathrm{dp}(s_n) + 2$, Dl_R grows much faster than any H_α with $\alpha < \Delta_{k+1}$.

Since the complexity of a TRS terminating via either MPO, LPO, or KBO is bounded by a multiple recursive function (see Theorems 5, 6, and 7), termination of R cannot be established by one of these orders.

We want to demonstrate where LPO fails. R owes much of its strength to the interplay between $\bullet$, $[\!]$, and $\circ$ on the one hand and $+$ and P on the other hand. No LPO is able to prove termination of a TRS containing the rules (N3), (N4), ($\mathrm{D}_1\mathsf{P}$), and ($\mathrm{E}_1\mathsf{P}$), since the first three rules require $\circ \succ [\!] \succ \bullet \succ \mathsf{P}$ while the fourth rule implies $\mathsf{P} \succ \circ$.

5.3 Proof of total termination

We order $\mathcal{P} := (\Delta_{k+1} \setminus \{0\}) \times \omega \times \omega$ by $\prec$, which is the lexicographic product of the usual $<$ on these sets of ordinals. By Exercise 13.i, $(\mathcal{P}, \prec)$ is a well-order with (using Exercise 7 and the fact that Δ_{k+1} is an epsilon)

$$\mathrm{otype}(\mathcal{P}, \prec) = \omega \cdot \omega \cdot \Delta_{k+1} = \omega^2 \cdot \omega^{\Delta_{k+1}} = \omega^{2+\Delta_{k+1}} = \omega^{\Delta_{k+1}} = \Delta_{k+1}\ .$$

We identify $(\alpha, 0, 0) \in \mathcal{P}$ and α to avoid lengthy notations. Thus $\alpha > \beta$ implies $\alpha \succ (\beta, m, n) \succcurlyeq \beta$. We now define interpreting functions for all symbols of our rewrite system. The homomorphic mapping (cf. (3)) based on these functions will establish total termination for R via Theorem 10. Note that the homomorphic mapping is not the real interpretation of R in a well-order in the sense of Proposition 4, it is rather used to define such an interpretation via Theorem 10. In spite of this, we call $[\![s]\!]$ the interpretation of s.

Definition 54. *For arbitrary elements $p = (\alpha, m, n)$, $p_l = (\alpha_l, m_l, n_l)$ (with $1 \leqslant l \leqslant k$), $q = (\beta, m', n')$ and $r = (\gamma, m'', n'')$ of $\mathcal{P}$ we put*

$$\begin{aligned}
[\mathsf{0}] &:= 1\\
[\mathsf{S}](p) &:= \alpha + 1\\
[+](p, q) &:= \alpha \oplus \beta \oplus \beta\\
[\mathsf{P}](\bar{p}, q) &:= \psi(\bar{\alpha}, \beta)\\
[\bullet](p) &:= (\alpha, m, n+1)\\
[[\!]](p) &:= (\alpha, m+1, 0)\\
[\circ](p) &:= \alpha + 1\\
[\mathsf{M}](\bar{p}, q) &:= \psi(\bar{\alpha}, \beta) \cdot (3m' + 1)\\
[\mathsf{J}_i](p_1, \dots, p_i, q) &:= \psi(\alpha_1, \dots, \alpha_i, \cdot, \bar{1})^{2m_i+1}(\beta)\\
[\mathsf{Q}_{ij}](p_1, \dots, p_i, q) &:= \psi(\alpha_1, \dots, \max\{\alpha_j, \beta\}, \dots, \alpha_i, \bar{1})\\
[\mathsf{R}_i](p_1, \dots, p_i, q, r) &:= \psi(\alpha_1, \dots, \max\{\alpha_i, \beta\}, \bar{1}, \gamma + 1)\ .
\end{aligned}$$

Note that some components of q, r and p_l are never used. Furthermore the $[f]$ with $f \in \Sigma_0$ intentionally forget the two trailing components of p, q and p_l. Hence, for $s, t, \bar{s} \in \mathcal{T}(\Sigma)$, we have

$$\begin{aligned}
&[\![\mathsf{S}[\!]s]\!] = [\![\mathsf{S}s]\!] = [\![\circ s]\!] = [\![\circ[\!]s]\!],\\
&[\![+([\!]s, t)]\!] = [\![+(s, [\!]t)]\!] = [\![+(s, t)]\!],\\
&[\![\mathsf{P}(s_1, \dots, [\!]s_i, \dots, s_{k+1})]\!] = [\![\mathsf{P}(\bar{s})]\!].
\end{aligned} \tag{14}$$

This remains true when $[]$ is replaced with $\bullet$.

Both the use of m' in $[\mathsf{M}]$ and the use of m_i in $[\mathsf{J}_i]$ are based on $[[]]$, as m' and m_i count the appearances of $[]$ at the positions important for (RM) and (RJ_i), respectively. The definitions of $[\mathsf{Q}_{ij}]$ and $[\mathsf{R}_i]$ reflect the duplication of x_j and x_i in (H$_{ij}$6) and (H$_i$7), respectively. Here we profit from Theorem 10, since taking the maximum violates monotonicity. Note that only $[\bullet]$ is monotone, as the other functions ignore the third component of their first arguments.

Lemma 22. *The mapping $[\![\cdot]\!]$ is weakly monotone and has the subterm property.*

Proof. Due to Lemma 1 it suffices to show that the functions interpreting the symbols of Σ are weakly monotone and have the subterm property. We start with the latter one. The proof for $[+]$ uses the fact that the first components of elements of $\mathcal{P}$ are larger than 0, and all interpretations involving ψ rely on its subterm property and its monotonicity, hence on Lemma 7.ii..

Weak monotonicity is obvious for all interpreting functions from $[\mathsf{0}]$ to $[\circ]$, and a moment's reflection establishes it for $[\mathsf{Q}_{ij}]$ and $[\mathsf{R}_i]$. The result for $[\mathsf{M}]$ and $[\mathsf{J}_i]$ relies on Lemma 8.

Proposition 10. *R is totally terminating.*

Proof. Because of Theorem 10 and Lemma 22 it remains to prove that $[\![\cdot]\!]$ normalizes R. Let a ground substitution σ be given. We denote the values of σ for x, x_i (with $1 \leq i \leq k+1$), y and z by s, s_i, t and u with interpretations (α, m, n), (α_i, m_i, n_i), (β, m', n') and (γ, m'', n''). Our goal will be established by showing

$$\forall (l,r) \in R \quad p := [\![l\sigma]\!] \succ [\![r\sigma]\!] =: p' \ .$$

For all rules (S$_i f$) we can fall back upon the subterm property of the interpreting functions under consideration. The subterm properties of $[\bullet]$ and $[\circ]$, sometimes in combination with (14), settle (N2), (N3), (D$_i f$), (H1), (H$_{ij}$6), and (H$_i$7). For example, we can treat (N3) by $p = [\![\circ s]\!] = [\![\circ [] s]\!] \succ [\![[] s]\!] = p'$, and

$$\begin{aligned} p &= [\![\bullet \mathsf{P}(s_1, \ldots, s_i, \bar{0}, \mathsf{S}t)]\!] \succ [\![\mathsf{P}(s_1, \ldots, s_i, \bar{0}, \mathsf{S}t)]\!] \\ &= [\![\mathsf{P}(s_1, \ldots, \bullet s_i, \bar{0}, \mathsf{S}t)]\!] = [\![\mathsf{R}_i(s_1, \ldots, \bullet s_i, s_i, t)]\!] = p' \end{aligned}$$

suffices for (H$_i$7). In the same way we build on $[\![[] s]\!] = [\![[] \bullet s]\!]$ to treat (N1) and (N4), while (F2) and (E$_2$+) both rely on $\alpha \oplus (\beta+1) \oplus (\beta+1) > \alpha \oplus \beta \oplus \beta + 1$. The monotonicity of ψ and Lemma 8.ii yield the result we are after for (F1), (E$_i$P), and (H2), while (i) and (iii) of Lemma 8 settle (RM) and (RJ_i). The subterm property and the monotonicity of ψ, joined with (ii) and (iv) of Lemma 8, provide us with everything we need for (H3), (H$_i$4) and (H$_i$5). It remains to take care of (RQ_{ij}) and (RR_i). For the former we have $p \succcurlyeq \psi(\alpha_1, \ldots, \alpha_i + 1, \bar{1})$ by weak monotonicity, and $p \succ \alpha_1, \ldots, \alpha_i, 1, \beta$ by the subterm property of ψ. Under these conditions Lemma 7.iii implies $p \succ p'$. Via similar reasoning we can handle (RR_i).

Combining Corollary 2, Proposition 10, and the aforementioned Theorem 29 of Weiermann we come to a satisfying conclusion.

Theorem 31. *For every $\alpha < \Lambda$ there exists a simply (and even totally) terminating TRS whose complexity eventually dominates H_α (and thus all H_β with $\beta \leqslant \alpha$.*

On the other hand, for every simply terminating TRS there exists $\alpha < \Lambda$ such that the complexity of the rewrite system is dominated by H_α.

In conclusion we learn that simply terminating TRSs are blessed with an enormous computational strength. This strength is by no means exhausted by the common classes of simplification orders like MPO, LPO, or KBO. The question arises whether there is a natural class of simplification orders whose corresponding TRSs may attain complexities beyond multiple recursion.

5.4 Some notes

A somewhat more abstract version of this section can be found in Lepper [2001b].

We should stress that most of the proofs presented in this section are inspired by Touzet [1998]. Recall from Definition 28 that $\omega_0 = 1$ and $\omega_{n+1} = \omega^{\omega_n}$. Touzet showed the following:

Theorem 32 (Touzet 1998). *For any $n \in \mathbb{N}$ there is a totally terminating TRS whose complexity eventually dominates all H_α with $\alpha < \omega_n$.*

The TRS used for $n+1$ can be identified with a proper extension (concerning both symbols and rules) of the one used for n. As TRSs have to be finite, it was not possible to reach $\mathrm{H}_{\varepsilon_0}$ with this construction.

There are related results, which also rely on applications of subrecursive hierarchies to term rewriting. For example, the construction behind the following result is based on a generalized Hardy hierarchy below ω_3.

Theorem 33 (Cichon and Tahhan Bittar 1998). *For any simply terminating SRS R there is $\alpha < \omega_3$ such that the complexity of R is dominated by H_α.*

Touzet established the optimality of this upper complexity bound.

Theorem 34 (Touzet 1999). *For any $\alpha < \omega_3$ there is a totally terminating SRS whose complexity eventually dominates H_α.*

So for strings it is all just the same as for terms: these results are directly connected to the supremum of the order types of simplification orders on strings.

Theorem 35 (de Jongh and Parikh 1977). *If $(\mathcal{T}(\Sigma), \prec)$ is a simplification order on strings then*

$$\mathrm{otype}(\mathcal{T}(\Sigma), \prec) < \omega_3 \ .$$

The frequent appearance of the Hardy functions in this section give strong evidence to the following quote of Touzet [1999]:

"the Hardy hierarchy is the right tool for connecting derivation length and order type."

Though this is true for the set of *all* simplification orders (no matter if we consider terms or strings), it may yield far too large bounds for smaller and more uniform sets of simplification orders.

An outstanding example for the validity of Touzet's claim is KBO with a maximal (tiny) order type of ω^ω and complexities equivalent to $\mathrm{Ack}(2^{O(n)}, 0)$, see Lepper [2001a]. The Hardy function $\mathrm{H}_{\omega^\omega}$ is a version of the (binary) Ackermann function Ack (cf. Exercises 22 and 20). Far less convincing examples are MPO and LPO, which are not able to make full use of their huge order types. For the complexities occurring within termination via MPO, the H_α with $\alpha < \omega^\omega$ suffice, while those of termination via LPO are controlled by the H_α with $\alpha < \omega_3$. A tight approach to these orders is presented in the next section.

6 Exploiting the Slow-growing Hierarchy

We start this section by presenting a general outline. Let terms $s = t_0, t_1, \ldots, t_n$ be given, such that $s \to_R t_1 \to_R \cdots \to_R t_n$ holds, where t_n is in normal form and term-depth of s is $\leq m$. Assume $\to_R$ is contained in a termination order $\succ$. Hence

$$s \succ t_1 \succ \cdots \succ t_n \ .$$

Assume further the sequence $\langle s, t_1, \ldots, t_n \rangle$ is chosen so that n is maximal. Then in the realm of classifications of derivation lengths one usually defines an *interpretation* $\mathcal{I} \colon \mathcal{T}(\Sigma, \mathcal{V}) \to \mathbb{N}$ such that

$$\mathcal{I}(s) > \mathcal{I}(t_1) > \cdots > \mathcal{I}(t_n) \ ,$$

holds. The existence of such an interpretation then directly yields a bound on the derivation length.

The problem is to guess the right interpretation from the beginning. More often than not this is not at all obvious. Therefore we want to generate the interpretation function directly from the termination order in an intrinsic way. To this avail we proceed as follows. We separate $\mathcal{I}$ into an *ordinal interpretation* $\pi \colon \mathcal{T}(\Sigma) \to T$ and an ordinal theoretic function $g \colon T(K) \to \mathbb{N}$, where T denotes a suitable chosen set of terms representing an initial segment of the ordinals, cf. Definition 33.

Firstly, we can employ the connection between the termination order $\succ$ and the order on the notation system T, as elaborated on in Section 3. Secondly, it turns out that g can be defined in terms of the slow-growing function $\mathrm{G}_x \colon T \to \mathbb{N}$; $x \in \mathbb{N}$. (Note that we have swapped the usual denotation of arguments, see Definition 43.) This approach work smoothly when the TRS R is compatible with (i) a MPO $\succ_{\mathsf{mpo}}$, (ii) a LPO $\succ_{\mathsf{lpo}}$, or (iii) a KBO $\succ_{\mathsf{kbo}}$, respectively. To simplify the presentation of the proof we restrict our attention to a rewrite system R whose termination can be shown by a *multiset path order* $\succ_{\mathsf{mpo}}$. We will comment on the necessary extension for the other two cases in Section 6.5. Let terms $s = t_0, t_1, \ldots, t_n$ be given, such that $s \to_R t_1 \to_R \cdots \to_R t_n$ holds, where t_n is in normal form and $\mathrm{dp}(s) \leq m$. By our choice of R this implies

$$s \succ_{\mathsf{mpo}} t_1 \succ_{\mathsf{mpo}} \cdots \succ_{\mathsf{mpo}} t_n \ . \tag{15}$$

We define a ground substitution ρ: $\rho(x) = c$, for all $x \in \mathcal{V}$. Let $>$ denote the well-order on the ordinal notation system T as stated in Definition 33. Let $l, r \in \mathcal{T}(\Sigma, \mathcal{V})$. Depending on m and properties of R, we show the existence of a natural number h such that $l \succ_{\mathsf{lpo}} r$ implies $\pi(l\rho) > \pi(r\rho)$ and $\mathrm{G}_h(\pi(l\rho)) > \mathrm{G}_h(\pi(r\rho))$, respectively. Employing this form of an *Interpretation Theorem* we conclude from (15) for some $\alpha \in T$, so that $\alpha < \psi(\omega, 0)$

$$\alpha > \pi(s\rho) > \pi(t_1\rho) > \cdots > \pi(t_n\rho) \ .$$

and consequently

$$\mathrm{G}_h(\alpha) > \mathrm{G}_h(\pi(s\rho)) > \mathrm{G}_h(\pi(t_1\rho)) > \cdots > \mathrm{G}_h(\pi(t_n\rho)) \ .$$

Thus $\mathrm{G}_h(\alpha)$ calculates an upper bound for n. Therefore the *complexity* of R can be measured in terms of the *slow-growing hierarchy* below $\psi(\omega, 0)$. Hence, due to Theorem 28 we obtain a primitive recursive upper bound for n.

6.1 The underlying rewrite system R

We fix some notations. Let $\Sigma = \{f_1, \ldots, f_N\}$ denote a finite signature. The cardinality N is assumed to be fixed in the sequel. We set $K := \max\{\mathrm{ar}(f) \colon f \in \Sigma\}$. In Section 2 we introduced the notion of *term depth*. To simplify the presentation below, we will alter this definition to our purpose.

Definition 55. *We define $\tau_{\mathrm{MPO}}(s) := K$, if $s \in \mathcal{V}$ or $s \in \Sigma^{(0)}$ and otherwise*

$$\tau_{\mathrm{MPO}}(f(s_1, \ldots, s_m)) := \max\{\tau_{\mathrm{MPO}}(s_i) \colon 1 \leq i \leq m\} + 3 \ .$$

Let R denote an arbitrary but fixed finite rewrite system. By assumption the rewrite relation $\to_R$ is compatible with in the multiset path order $\succ_{\mathsf{mpo}}$, hence terminating.

6.2 The interpretation theorem

Assume T equals $T(2)$ the notation system based on the binary function symbol ψ as introduced in Definition 33.

Definition 56. *Recursive definition of the interpretation function $\pi \colon \mathcal{T}(\Sigma) \to T$. If $s = f_j(s_1, \ldots, s_m)$, then set*

$$\pi(s) := \psi(j, \pi(s_1) \oplus \cdots \oplus \pi(s_m) + 1) \ .$$

In the sequel of this section we show that π defines an *interpretation* for R on $(T, <)$; i.e. we establish the following theorem.

Theorem 36. *For all $s, t \in \mathcal{T}(\Sigma)$ we have $s \to_R t$ implies $\pi(s) > \pi(t)$.*

Unfortunately this is not strong enough. The problem being that $\alpha > \beta$ implies that G_α eventually dominates G_β, only, compare Theorem 25. Whereas to proceed with our general program we need an interpretation theorem for a binary relation $\succ$ on T, such that $\alpha \succ \beta \Rightarrow \mathrm{G}_\alpha(x) > \mathrm{G}_\beta(x)$ holds for all x.

We introduce a notion of a generalized system of fundamental sequences for $(T, <(\psi(\omega,0)))$. Based on this generalized notion, it is then possible to define a suitable order $\succ$.

Definition 57 (Generalized system of fundamental sequences). *Recursive definition of* $(\alpha)^x$ *for* $x < \omega$.

i. $(0)^x := \emptyset$
ii. *Assume* $\alpha = \alpha_1 + \cdots + \alpha_m$; $m > 1$. *Then* $\beta \in (\alpha)^x$ *if either*
  – $\beta = \alpha_1 \oplus \cdots \alpha_i^* \cdots \oplus \alpha_m$ *and* $\alpha_i^* \in (\alpha_i)^x$ *holds, or*
  – $\beta = \alpha_i$.
iii. *Assume* $\alpha = \psi(\overline{\alpha})$. *Then* $\beta \in (\alpha)^x$ *if*
  – $\beta = \psi(\alpha_1^*, \alpha_2)$ *or* $\beta = \psi(\alpha_1, \alpha_2^*)$, *and* $\alpha_i^* \in (\alpha_i)^x$ *for* $i \in \{1,2\}$ *or*
  – $\beta = \alpha_i + x$, *where* $\alpha_i > 0$, *or*
  – $\beta = \psi(\overline{\alpha})[x]$.

By recursion we define the *transitive closure* of the ownership $(\alpha)^x \ni \beta$. (We ambiguously denote this binary relation by the same symbol as the binary relation introduced in Definition 44. No confusion will arise from this.)

$$(\alpha >_{(x)} \beta) \Leftrightarrow (\exists \gamma \in (\alpha)^x (\gamma >_{(x)} \beta \vee \gamma = \beta)) \; .$$

Let $\alpha, \beta \in T$. It is easy to verify that $\alpha >_{(x)} \beta$ (for some $x < \omega$) implies $\alpha > \beta$. If no confusion can arise we write α^x instead of $(\alpha)^x$. Instead of $\alpha >_{(x)} \beta$ we sometimes write $\beta <_{(x)} \alpha$.

Lemma 23. *(Subterm Property) Let* $x < \omega$ *be arbitrary.*

i. $\alpha <_{(x)} \gamma_1 \oplus \cdots \alpha \cdots \oplus \gamma_m$.
ii. $\alpha, \beta <_{(x)} \psi(\alpha, \beta)$.

Proof. The first assertion is trivial. The second assertion follows by the definition of $<_{(x)}$ and assertion i).

Lemma 24. *(Monotonicity Property) Let* $x < \omega$ *be arbitrary.*

i. *If* $\alpha >_{(x)} \beta$, *then* $\gamma_1 \oplus \cdots \alpha \cdots \oplus \gamma_m >_{(x)} \gamma_1 \oplus \cdots \beta \cdots \oplus \gamma_m$.
ii. *If* $\alpha >_{(x)} \beta$, *then* $\psi(\alpha, \gamma) >_{(x)} \psi(\beta, \gamma)$ *and* $\psi(\gamma, \alpha) >_{(x)} \psi(\gamma, \beta)$, *respectively.*

Proof. We employ induction on α to prove i). We may assume that $\alpha > 0$. By definition of $\alpha >_{(x)} \beta$ we either have (a) that there exist $\delta \in \alpha^x$ and $\delta >_{(x)} \beta$ or (b) $\beta \in \alpha^x$. Firstly, one considers the latter case. Then $(\gamma_1 \oplus \cdots \beta \cdots \oplus \gamma_m) \in (\gamma_1 \oplus \cdots \alpha \cdots \oplus \gamma_m)^x$ holds by Definition 57. Therefore $(\gamma_1 \oplus \cdots \alpha \cdots \oplus \gamma_m) >_{(x)} (\gamma_1 \oplus \cdots \beta \cdots \oplus \gamma_m)$ follows. Now, we consider the case (a). By assumption $\delta >_{(x)} \beta$ holds, by (ih) this implies $(\gamma_1 \oplus \cdots \delta \cdots \oplus \gamma_m) >_{(x)} (\gamma_1 \oplus \cdots \beta \cdots \oplus \gamma_m)$.

Now $(\gamma_1 \oplus \cdots \alpha \cdots \oplus \gamma_m) >_{(x)} (\gamma_1 \oplus \cdots \delta \cdots \oplus \gamma_m)$ follows by definition of $>_{(x)}$, if we replace β by δ in the proof of the second case. This completely proves i).

To prove ii) we proceed by induction on α. By definition of $\alpha >_{(x)} \beta$ we have either (i) $\delta \in \alpha^x$ and $\delta >_{(x)} \beta$ or (ii) $\beta \in \alpha^x$. It is sufficient to consider the latter case, the first case follows from the second as above. By Definition 57, $\beta \in \alpha^x$ implies $\psi(\beta,\gamma) \in (\psi(\alpha,\gamma))^x$ and $\psi(\gamma,\beta) \in (\psi(\gamma,\alpha))^x$, respectively. Thus $\psi(\alpha,\gamma) >_{(x)} \psi(\beta,\gamma)$ or $\psi(\gamma,\alpha) >_{(x)} \psi(\gamma,\beta)$ follows.

In the sequel we show the existence of a natural number d, such that for all $s,t \in \mathcal{T}$, and any ground substitution ρ, $s \to_R t$ implies $\pi(s\rho) >_{(d)} \pi(t\rho)$. Theorem 36 follows then as a corollary. The proof is involved, and makes use of a sequence of lemmas.

We start with the following lemma that will be applied in various places below. It generalizes the fact $\alpha >_{(x+1)} \alpha[x]$, i.e. sequences of descents along the $(x+1)$-branch lead us to a (single) decent along the x-branch of the generalized system of fundamental sequences for $(T,<)$, cf. Definition 57.

Lemma 25. *Assume $\alpha,\beta \in \mathsf{Lim}$; $x \geq 1$. If $\alpha >_{(x)} \beta$, then $\alpha >_{(x+1)} \beta+1$ holds.*

To prove the lemma we exploit the following auxiliary lemma.

Lemma 26. *We assume the assumptions and notation of Lemma 25; assume Lemma 25 holds for all $\gamma,\delta \in \mathsf{Lim}$ with $\gamma,\delta < \alpha$. Then $\alpha >_{(x+1)} \alpha[x+1] \geq_{(x+1)} \alpha[x]+1$.*

Proof. The lemma follows by induction on the form of α by analyzing all cases of Definition 37.

Proof. (of Lemma 25) The proof proceeds by induction on the form of α. We consider only the case where $\alpha = \psi(\alpha_1,\alpha_2)$. The case where $\alpha = \alpha_1 + \cdots + \alpha_m$ is similar but simpler.

By definition of $\alpha >_{(x)} \beta$ we have either (i) $\gamma \in \alpha^x$ and $\gamma >_{(x)} \beta$ or (ii) $\beta \in \alpha^x$. Assume for $\gamma \in \alpha^x$ we have already shown that $\gamma+1 <_{(x+1)} \alpha$. Then for $\beta <_{(x)} \gamma$, we conclude by (ih) and the Subterm Property $\beta+1 <_{(x+1)} \gamma <_{(x+1)} \gamma+1 <_{(x+1)} \alpha$. Hence, it suffices to consider the second case. We proceed by case distinction on the form of β.

Case $\beta = \psi(\alpha_1^*,\alpha_2)$ where $\alpha_1^* \in (\alpha_1)^x$. Note that $\alpha_1 < \alpha$, hence (ih) is applicable to establish $\alpha_1^*+1 <_{(x+1)} \alpha_1$. Furthermore by the Subterm Property follows $\alpha_1^* <_{(x+1)} \alpha_1^*+1$ and therefore $\psi(\alpha_1^*,\alpha_2) <_{(x+1)} \psi(\alpha_1^*+1,\alpha_2)$ holds with Monotonicity. Applying (ih) wrt. $\psi(\alpha_1^*+1,\alpha_2)$ we obtain

$$\psi(\alpha_1^*,\alpha_2)+1 <_{(x+1)} \psi(\alpha_1^*+1,\alpha_2) <_{(x+1)} \psi(\alpha_1,\alpha_2) = \alpha \ .$$

The last inequality follows again by an application of the Monotonicity Property. The case $\beta = \psi(\alpha_1,\alpha_2^*)$ is similar.

Case $\beta = \alpha_i + x$: Then $(\alpha_i + x)+1 = \alpha_i + (x+1) <_{(x+1)} \alpha$.

Case $\beta = \psi(\overline{\alpha})[x]$. Clearly $\beta \in \mathsf{Lim}$. Then the auxiliary lemma becomes applicable. Thus $\psi(\overline{\alpha})[x]+1 \leq_{(x+1)} \alpha[x+1] <_{(x+1)} \alpha$.

The following lemma provides a relation between the generalized system of fundamental sequences for $(T, <)$, introduced in this section, and the "usual" notion of fundamental sequence $\cdot[\cdot]\colon T \times \omega \to T$. Note that the restriction to $\alpha \in T$ such that α is in the range of π is essential.

Lemma 27. *Let $t \in \mathcal{T}(\Sigma)$, let $d \geq 3$, and set $\alpha := \pi(t)$. Then $\beta \in \alpha^{d-3}$ implies $\beta \leq_{(d)} \alpha[d-3]$.*

Proof. The proof proceeds by induction on α. (We abbreviate induction hypothesis as (ih).) Assume α has the form $\psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m + 1)$ for $j, m < \omega$. We proceed by case-distinction on β.

Assume $\beta = \psi(j^*, \alpha_1 \oplus \cdots \alpha_m + 1)$, where $j^* \in (j)^{d-3}$, i.e. $j^* \leq_{(d)} j - 1 <_{(d)} j$. By definition we obtain

$$\alpha[d-3] = \psi(j-1, \cdot)^{d-2}(\psi(j, \alpha_1 \oplus \cdots \alpha_m)) \geq_{(d)} \psi(j-1, \psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m)) \ .$$

Furthermore, by definition we obtain $\psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m) >_{(d)} \alpha_1 \oplus \cdots \oplus \alpha_m + 1$. This together with Monotonicity yields $\psi(j-1, \psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m)) \geq_{(d)} \psi(j-1, \alpha_1 \oplus \cdots \oplus \alpha_m + 1)$. Putting everything together renders

$$\alpha[d-3] >_{(d)} \psi(j-1, \alpha_1 \oplus \cdots \oplus \alpha_m + 1) >_{(d)} \beta \ .$$

Now assume $\beta = \psi(j, \gamma)$, where $\gamma \in (\alpha_1 \oplus \cdots \oplus \alpha_m + 1)^{d-3}$. Assume w.l.o.g. $\gamma = \alpha_1 \oplus \cdots \oplus \alpha_i^* \oplus \alpha_m + 1$, so that $\alpha_i^* \in \alpha_i^{d-3}$. Employing (ih) and Lemma 26 we conclude $\alpha_i^* \leq_{(d)} \alpha_i[d-3] \leq_{(d)} \alpha_i[d-3] + 1 \leq_{(d)} \alpha_i$ Thus we have $(\alpha_1 \oplus \cdots \oplus \alpha_i \oplus \cdots \oplus \alpha_m) \geq_{(d)} (\alpha_1 \oplus \cdots \oplus \alpha_i[d-3] \oplus 1 \cdots \oplus \alpha_m)$ and furthermore $(\alpha_1 \oplus \cdots \oplus \alpha_i[d-3] \oplus 1 \cdots \oplus \alpha_m) \geq_{(d)} (\alpha_1 \oplus \cdots \oplus \alpha_i[d-3] \oplus \cdots \oplus \alpha_m \oplus 1) \geq_{(d)} \gamma$. Putting this together, we obtain

$$\alpha[d-3] \geq_{(d)} \psi(j-1, \psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m)) >_{(d)} \psi(j, \alpha_1 \oplus \cdots \alpha_m) \geq_{(d)} \beta \ .$$

Assume $\beta = j + (d-3)$. Then clearly

$$\alpha[d-3] >_{(d)} \psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m) >_{(d)} j + d >_{(d)} \beta$$

On the other hand assume $\beta = (\alpha_1 \oplus \cdots \oplus \alpha_m + 1) + (d-3)$. Then we have

$$\begin{aligned} \alpha[d-3] &>_{(d)} \psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m) \\ &>_{(d)} (\alpha_1 \oplus \cdots \oplus \alpha_m) + d) \\ &>_{(d)} (\alpha_1 \oplus \cdots \oplus \alpha_m) + (d-3) + 1 \\ &= \beta \end{aligned}$$

Finally assume $\beta = \alpha[d-3]$. Then the lemma holds almost trivially via Lemma 26.

In the following we feel free to frequently employ the Subterm and the Monotonicity Property without further notice. The following lemma is purely technical. Recall that K denotes the maximal arity of a function symbol in Σ.

Lemma 28. *Let $t \in \mathcal{T}(\Sigma)$ and assume $\alpha := \pi(t)$. Furthermore assume $d \geq K + 3$. Then $\alpha[d-1] >_{(d)} \alpha[d-3] \cdot K + 1$.*

Proof. Let α be of the form $\psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m + 1)$, where $j, m < \omega$. Note that by definition $d \geq K$ holds. Then

$$\begin{aligned}
\alpha[d-1] &= \psi(j-1, \cdot)^d(\psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m)) \\
&= \psi(j-1, \cdot)^{2+(d-2)}(\psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m)) \\
&= \psi(j-1, \psi(j-1, \psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m + 1)[d-3])) \\
&>_{(d)} \psi(j-1, \alpha[d-3] + 1) \\
&\geq_{(d)} \psi(0, \alpha[d-3] + 1) \\
&>_{(d)} \psi(0, \alpha[d]) \cdot (d+1) \\
&>_{(d)} (\alpha[d] \cdot K) + \alpha[d-3] \\
&>_{(d)} (\alpha[d] \cdot K) + 1
\end{aligned}$$

It is a crucial observation that $0 <_{(x)} \alpha$ holds for any $x < \omega$, $\alpha \in T$. (This follows by a simple induction on α.) Hence the last inequality follows as by definition $\alpha > 0$.

The following three lemmas provide the ground for the Interpretation theorem. These are the crucial arguments in the proof.

Lemma 29. *Let $t \in \mathcal{T}(\Sigma)$ be given. Assume $\tau_{\mathrm{MPO}}(t) \leq d$, and $f_j \in \Sigma$. If $f_j \succ_{\mathsf{lpo}} t$, then $\pi(f_j) >_{(d)} \pi(t)$.*

Proof. We proceed by induction on $\mathrm{dp}(t)$. Set $\alpha := \pi(f_j)$, and $\beta := \pi(t)$.

Case $\mathrm{dp}(t) = 0$: Then by assumption $t = f_i \in \Sigma$, $i < j$. Hence $i <_{(d)} j$ holds and we conclude $\pi(t) = \psi(i, 1) <_{(d)} \psi(j, 1) = \pi(f_j)$.

Case $\mathrm{dp}(t) > 0$: Let $t = f_i(t_1, \ldots, t_n)$. Set $\beta_l := \pi(t_l)$ for all $l = 1, \ldots, n$. By (ih) one obtains $\beta_l <_{(\tau_{\mathrm{MPO}}(t_l))} \alpha$ for all $l = 1, \ldots, n$, i.e. $\beta_l <_{(d-3)} \alpha$ holds for all l. (Apply Lemma 25 if necessary.)

For all l, we need only consider the case where $\beta_l \in \alpha^{d-3}$. By Lemma 27 this implies $\beta_l \leq_{(d)} \alpha[d-3]$. We consider $\alpha[d] = \pi(f_j)[d]$ and apply the following sequence of descents via $>_{(d)}$. Apart from the definition of the generalized system of fundamental sequences, we employ (ih) and Lemma 28.

$$\begin{aligned}
\alpha[d] &= \psi(j-1, \cdot)^{d+1}(\psi(j, 0)) \\
&= \psi(j-1, \psi(j-1, \cdot)^d(\psi(j, 0))) \\
&= \psi(j-1, \psi(j, 1)[d-1]) \\
&>_{(d)} \psi(j-1, \psi(j, 1)[d-3] \cdot K + 1) \\
&\geq_{(d)} \psi(j-1, \alpha[d-3] \oplus \cdots \oplus \cdot\alpha[d-3] + 1) \\
&\geq_{(d)} \psi(j-1, \beta_1 \oplus \cdots \oplus \beta_n + 1) \\
&\geq_{(d)} \psi(i, \beta_1 \oplus \cdots \oplus \beta_n + 1) = \beta \quad .
\end{aligned}$$

This proves the lemma completely.

Lemma 30. *Let $f_i(t_1,\dots,t_n), f_j(s_1,\dots,s_m) \in \mathcal{T}(\Sigma)$ be given; let $d \geq K+3$.*

i. *Assume $i < j$ and $\pi(f_j(\overline{s})) >_{(d-3)} \pi(t_l)$ for all $l = 1,\dots,n$. Then we obtain $\pi(f_j(\overline{s})) >_{(d)} \pi(f_i(\overline{t}))$ holds.*
ii. *Assume $i = j$ and assume the existence of sets X, Y with*

$$\{\pi(s_1),\dots,\pi(s_m)\} - X \cup Y = \{\pi(t_1),\dots,\pi(t_n)\}$$

such that for all $\beta \in Y$, exists $\alpha \in X$ and $\alpha >_{(d-3)} \beta$. Then $\pi(f_j(\overline{s})) >_{(d)} \pi(f_i(\overline{t}))$ holds.

Proof. Set $\alpha := \pi(f_j(\overline{s}))$; $\beta := \pi(f_i(\overline{t}))$; finally set $\alpha_i := \pi(s_i)$ for all $i = 1,\dots,m$, and $\beta_i := \pi(t_i)$ for all $i = 1,\dots,n$.

Firstly assume $i < j$ and $\alpha >_{(d-3)} \beta_l$ for all $l = 1,\dots,n$. As above, we consider only the case where $\beta_l \in (\alpha)^{d-3}$. By Lemma 27 this implies $\beta_l \leq_{(d)} \alpha[d-3]$. The other case follows easily.

We consider $\pi(f_j(\overline{s}))[d] = \alpha[d]$ and apply the following sequence of descents via $>_{(d)}$. Apart from the definition of the generalized system of fundamental sequences, we employ the assumption of the lemma and Lemma 28.

$$\begin{aligned}
\alpha[d] &= \psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m + 1)[d] \\
&= \psi(j-1,\cdot)^{d+1}(\psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m)) \\
&= \psi(j-1, \psi(j-1,\cdot)^{d}(\psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m))) \\
&\geq_{(d)} \psi(j-1, \alpha[d-3] \cdot K + 1) \\
&\geq_{(d)} \psi(j-1, \alpha[d-1] \oplus \cdots \oplus \cdot\alpha[d-1] + 1) \\
&\geq_{(d)} \psi(j-1, \beta_1 \oplus \cdots \oplus \beta_n + 1) \\
&\geq_{(d)} \psi(i, \beta_1 \oplus \cdots \oplus \beta_n + 1) = \beta \quad .
\end{aligned}$$

This proves the first case of the lemma. Now assume the existence of sets X, Y with $\{\alpha_1,\dots,\alpha_m\} - X \cup Y = \{\beta_1,\dots,\beta_n\}$ such that for all $\beta_j \in Y$, exists $\alpha_i \in X$ and $\alpha_i >_{(d-3)} \beta_j$.

We fix i, j; by Lemma 27 we conclude $\alpha[d-3] \geq_{(d)} \beta$. Furthermore

$$\alpha_i >_{(d)} \alpha_i[d-1] >_{(d)} \alpha_i[d-3] \cdot K + 1 >_{(d)} \beta_j \cdot K \geq_{(d)} \beta_j \cdot \mathrm{card}(Y)$$

holds by Lemma 25 and Lemma 28.

As α, β were arbitrarily chosen we conclude $\alpha_1 \oplus \cdots \oplus \alpha_m >_{(d)} \beta_1 \oplus \cdots \oplus \beta_n$. In conclusion we obtain

$$\pi(s\rho) = \psi(j, \alpha_1 \oplus \cdots \oplus \alpha_m + 1) >_{(d)} \psi(i, \beta_1 \oplus \cdots \oplus \beta_n + 1) = \pi(t\rho) \ .$$

This proves part two.

Lemma 31. *Let $s, t \in \mathcal{T}$ be given. Assume $s = f_j(s_1,\dots,s_m)$, ρ is a ground substitution, $\tau_{\text{MPO}}(t) \leq d$. Assume further $s_k \succ_{\mathsf{mpo}} u$ and $\tau_{\text{MPO}}(u) \leq d$ implies $\pi(s_k\rho) >_{(d)} \pi(u\rho)$ for all $u \in \mathcal{T}$. Then $s \succ_{\mathsf{mpo}} t$ implies $\pi(s\rho) >_{(d)} \pi(t\rho)$.*

Proof. The proof is by induction on d. (Note that $d \geq K$ by definition.)

Case $d = K$: This implies $\mathrm{dp}(t) = 0$; therefore $t \in \mathcal{V}$ or $t = f_i \in \Sigma$. Consider $t \in \mathcal{V}$. Then t is a subterm of s. Hence there exists k $(1 \leq k \leq m)$ s.t. t is subterm of s_k. Hence $s_k \succeq_{\mathsf{mpo}} t$, and by assumption this implies $\pi(s_k\rho) >_{(d)} \pi(t\rho)$, and therefore $\pi(s\rho) >_{(d)} \pi(t\rho)$ by the Subterm Property.

Now assume $t = f_i \in \Sigma$. As $s \succ_{\mathsf{mpo}} t$ by assumption either $i < j$ or $s_k \succeq_{\mathsf{mpo}} t$ holds. In the latter case, the assumptions render $\pi(s_k\rho) \geq_{(d)} \pi(t\rho)$; hence $\pi(s\rho) >_{(d)} \pi(t\rho)$. Otherwise, $\pi(s\rho) = \psi(j, \pi(s_1\rho) \oplus \cdots \oplus \pi(s_m\rho) + 1)$, while $\pi(t\rho) = \pi(t) = \psi(i, 1)$. As $\pi(s_k\rho) >_{(x)} 0$ holds for arbitrary $x < \omega$, we conclude $\pi(s\rho) >_{(d)} \pi(t\rho)$.

Case $d > K$: Assume $\mathrm{dp}(t) > 0$. (Otherwise, the proof follows the pattern of the case $d = 0$.) Let $t = f_i(t_1, \ldots, t_n)$, and clearly $\tau_{\mathrm{MPO}}(t_l) \leq (d-3)$ for all $l = 1, \ldots, n$. We start with the following observation: Assume there exists i_0 s.t. $s \succ_{\mathsf{mpo}} t_l$ holds for all $l = i_0 + 1, \ldots, n$. Then by (ih) we have $\pi(s\rho) >_{(d-3)} \pi(t_l\rho)$.

We proceed by case-distinction on $s \succ_{\mathsf{mpo}} t$. Assume firstly there exists k $(1 \leq k \leq m)$ s.t. $s_k \succeq_{\mathsf{mpo}} t$. Utilizing the assumptions of the lemma, we conclude $\pi(s\rho) >_{(d)} \pi(t\rho)$. Now assume $i < j$ and $s \succ_{\mathsf{mpo}} t_l$ for all $l = 1, \ldots, n$. By the observation $\pi(s\rho) >_{(d-3)} \pi(t_l\rho)$ holds. Hence Lemma 30i becomes applicable and therefore $\pi(s\rho) >_{(Kd)} \pi(t\rho)$ holds true.

Finally assume $i = j$; and $(s_1, \ldots, s_m) \succ^{\mathsf{mul}}_{\mathsf{mpo}} (t_1, \ldots, t_n)$. By assumption of the lemma for each $k = 1, \ldots, m$ and $j = 1, \ldots, n$ $s_k \succ_{\mathsf{mpo}} t_j$ implies $\pi(s_k\rho) >_{(d-3)} \pi(t_j\rho)$. Set $\alpha_i := \pi(s_i\rho)$ for all $i = 1, \ldots, m$ and $\beta_i := \pi(t_i\rho)$ for all $i = 1, \ldots, n$. By assumption $(s_1, \ldots, s_m) \succ^{\mathsf{mul}}_{\mathsf{mpo}} (t_1, \ldots, t_n)$ implies the existence of sets X, Y with $\{\alpha_1, \ldots, \alpha_m\} - X \cup Y = \{\beta_1, \ldots, \beta_n\}$ such that for all $\beta_j \in Y$, exists $\alpha_i \in X$ and $\alpha_i >_{(d-3)} \beta_j$. Hence Lemma 30ii becomes applicable and therefore $\pi(s\rho) >_{(d)} \pi(t\rho)$ holds true.

The following lemma tells us that we can find an ordinal $\alpha < \psi(\omega, 0)$ that bounds effectively all ordinal interpretations of terms in the ground term algebra.

Lemma 32. *Let $t \in \mathcal{T}(\Sigma)$ be given, assume $\tau_{\mathrm{MPO}}(t) \leq d$. Then $\psi(N+1, 1) >_{(d)} \pi(t)$.*

Proof. The proof is by induction on $\mathrm{dp}(t)$ and follows the proof of Lemma 29.

Finally, we can collect the pieces.

Theorem 37. *Let $l, r \in \mathcal{T}$ be given. Assume ρ is a ground substitution and $\tau_{\mathrm{MPO}}(t) \leq d$. Then $l \succ_{\mathsf{mpo}} r$ implies $\pi(l\rho) >_{(d)} \pi(r\rho)$.*

Proof. We proceed by induction on $\mathrm{dp}(s)$.

Case $\mathrm{dp}(s) = 0$: Then s can either be a constant or a variable. As $s \succ_{\mathsf{mpo}} t$ holds, we can exclude the latter case. Hence assume $s = f_j$. As $f_j \succ_{\mathsf{lpo}} t$, t is closed. Hence the assumptions of the theorem imply the assumptions of Lemma 29 and we conclude $\pi(s\rho) = \pi(s) >_{(d)} \pi(t) = \pi(t\rho)$.

Case $\mathrm{dp}(s) > 0$: Then s can be written as $f_j(s_1, \ldots, s_m)$. By (ih) $s_k \succ_{\mathsf{mpo}} u$ and $\mathrm{dp}(u) \leq d$ imply $\pi(s_k\rho) >_{(d)} \pi(t\rho)$. Therefore the present assumptions contain the assumptions of Lemma 31 and hence $\pi(s\rho) >_{(d)} \pi(t\rho)$ follows.

Theorem 38 (The Interpretation Theorem.). *Let R denote a finite rewrite system whose induced rewrite relation is contained in $\succ_{\mathsf{mpo}}$. Then there exists $d < \omega$, such that for all $l, r \in \mathcal{T}$, and any ground substitution ρ $l \to_R r$ implies $\pi(l\rho) >_{(d)} \pi(r\rho)$.*

The number d is given effectively by setting d equal to $\max\{\tau_{\mathrm{MPO}}(r) \colon \exists l\ (l, r) \in R\}$.

Proof. The theorem follows as a direct corollary to Theorem 37.

6.3 Collapsing theorem

We define a variant of the slow-growing hierarchy, cf. Definition 43, suitable for our purposes.

Definition 58. *Recursive definition of the function $\widetilde{\mathrm{G}}_\alpha \colon \omega \to \omega$ for $\alpha \in T$.*

$$\begin{aligned} \widetilde{\mathrm{G}}_0(x) &:= 0 \\ \widetilde{\mathrm{G}}_\alpha(x) &:= \max\{\widetilde{\mathrm{G}}_\beta(x) \colon \beta \in (\alpha)^x\} + 1 \ . \end{aligned}$$

Lemma 33. *Let $\alpha \in T$, $\alpha > 0$ be given. Assume $x < \omega$ is arbitrary.*

i. *$\widetilde{\mathrm{G}}_\alpha$ is increasing. (Even strictly if $\alpha > \omega$.)*
ii. *If $\alpha >_{(x)} \beta$, then $\widetilde{\mathrm{G}}_\alpha(x) > \widetilde{\mathrm{G}}_\beta(x)$.*

Proof. Both assertions follow by induction over $<$ on α.

We need to know that this variant of the slow-growing hierarchy is indeed slow-growing. We show this by verifying that the hierarchies $\{\widetilde{\mathrm{G}}_\alpha \colon \alpha < \psi(\omega, 0)\}$ and $\{\mathrm{G}_\alpha \colon \alpha < \psi(\omega, 0)\}$ coincide with respect to growth-rate. It is a triviality to verify that there exists $\beta \in T$ such that $\widetilde{\mathrm{G}}_\beta$ majorizes G_α. (Simply set $\beta = \alpha$.) The other direction is less trivial. One first proves that for any $\alpha < \psi(\omega, 0)$ there exists $\gamma < \omega$ such that $\widetilde{\mathrm{G}}_\alpha(x) \leq \mathrm{F}_\gamma(x)$ for almost all x. Then one employs Theorem 28 to establish the existence of $\beta \in T$ such that $\widetilde{\mathrm{G}}_\alpha(x) \leq \mathrm{G}_\beta(x)$ holds for almost all x.

Theorem 39. $\displaystyle\bigcup_{\alpha<\psi(\omega,0)} \mathrm{G}_\alpha \approx \bigcup_{\alpha<\psi(\omega,0)} \widetilde{\mathrm{G}}_\alpha \approx \bigcup_{\gamma<\omega} \mathrm{F}_\gamma \approx \textsc{Prec}\ .$

6.4 Complexity bounds

The *complexity* of a terminating finite rewrite system R is measured by the *derivation length* function.

Definition 59. *The derivation length function $\mathrm{Dl}_R \colon \omega \to \omega$. Let $m < \omega$ be fixed. Then we define*

$$\mathrm{Dl}_R(m) := \max\{n \colon \exists t_1, \dots, t_n \in \mathcal{T}\ \ ((t_1 \to_R \cdots \to_R t_n) \wedge (\mathrm{dp}(t_1) \leq m))\} \ .$$

By assumption R is a finite rewrite system over $\mathcal{T}$ such that $\rightarrow_R$ is contained in a multiset path order. Now assume that there exist $s = t_0, t_1, \ldots, t_n \in \mathcal{T}$ with $\tau_{\text{MPO}}(s) \leq m$ such that

$$s \rightarrow_R t_1 \rightarrow_R \cdots \rightarrow_R t_n$$

holds. By our choice of R this implies $s \succ_{\mathsf{mpo}} t_1 \succ_{\mathsf{mpo}} \cdots \succ_{\mathsf{mpo}} t_n$. By assumption on Σ there exists $c \in \Sigma$, with $\text{ar}(c) = 0$.

We define a ground substitution ρ: $\rho(x) = c$, for all $x \in \mathcal{V}$. Let $d < \omega$ be defined as $\max\{\tau_{\text{MPO}}(r) \colon \exists l\ (l, r) \in R\}$. Recall that N denotes the cardinality of Σ. We conclude from the Interpretation Theorem and Lemma 32, $\pi(s\rho) >_{(d)} \pi(t_1\rho) >_{(d)} \cdots >_{(d)} \pi(t_n\rho)$ and $\psi(N+1, 0) >_{(m)} \pi(s\rho)$.

Setting $h := \max\{d, m\}$ and utilizing Lemma 25, we obtain $\psi(N+1, 0) >_{(h)} \pi(s\rho) >_{(h)} \cdots >_{(h)} \pi(t_n\rho)$. An application of Lemma 33.ii yields

$$\widetilde{\mathrm{G}}_{\psi(N+1,0)}(h) > \widetilde{\mathrm{G}}_{\pi(s\rho)}(h) > \cdots > \widetilde{\mathrm{G}}_{\pi(t_n\rho)}(h) \ .$$

Thus we have established a primitive-recursive upper bound for the derivation length of R if $\rightarrow_R$ is contained in a multiset path order. Furthermore, this bound is essentially optimal, cf. Hofbauer [1992]. Contrary to the original proof inHofbauer [1992], we can circumvent technical calculations with functions on the natural numbers and can shed light on the way the slow-growing hierarchy relates the order type of the termination order $\succ$ to the bound on the length of reduction sequences along $\rightarrow_R$.

6.5 Extensions

As already mentioned the above presented argument is general applicable. Using the same approach we obtain (essentially) optimal bounds on the derivation length of rewrite systems R where the induced rewrite relation $\rightarrow_R$ is

i. contained in a *lexicographic path order* $\succ_{\mathsf{lpo}}$, or
ii. contained in a *Knuth–Bendix order* $\succ_{\mathsf{kbo}}$.

We briefly indicate how the above proof needs to be adapted in the former case. Let R denote a rewrite system whose termination can be shown via a *lexicographic path order* $\succ_{\mathsf{lpo}}$. Firstly the definition of the interpretation function π has to be changed as follows.

Definition 60. *Recursive definition of the interpretation function* $\pi \colon \mathcal{T}(\Sigma) \rightarrow T(K+1)$*, where* K *denotes the maximal arity of a function symbol in* Σ*. If* $s = f_j \in \Sigma$*, then set* $\pi(s) := \psi(j, \overline{0})$*. Otherwise, let* $s = f_j(s_1, \ldots, s_m)$ *and set*

$$\pi(s) := \psi(j, \pi(s_1), \ldots, \pi(s_m) + 1, \overline{0}) \ .$$

Secondly it is necessary to adapt the specific notion of *term depth* employed above. Instead of Definition 55 we employ the following definition.

Definition 61. *We define* $\tau_{\text{LPO}}(s) := 0$*, if* $s \in \mathcal{V}$ *or* $s \in \Sigma^{(0)}$ *and otherwise*

$$\tau_{\text{LPO}}(f(s_1, \ldots, s_m)) := \max\{\tau_{\text{LPO}}(s_i) \colon 1 \leq i \leq m\} + 2 \ .$$

Then the presented proof needs only partial changes. It suffices to reprove the crucial lemmas in the above proofs. I.e. Lemma 29, 30, 31, and 32, respectively need to be reformulated and reproved.

As we have packed the technical details into the altered definition of term depth, Definition 55, the reformulation of these lemmas can almost be dropped. However it is important to note that in (the altered form of) Lemma 32 we need the full strength of the ordinal notation system T to find an upper bound. I.e. if $t \in \mathcal{T}(\Sigma)$ be given, assume $\tau_{\text{LPO}}(t) \leq d$. Then $\psi(K+1, \overline{0}) >_{(d)} \pi(t)$.

It remains to relate the hierarchy $\bigcup_{\alpha \in T} \mathrm{G}_\alpha$ to the multiple recursive functions. This is obtained through an application of the Hierarchy Comparison Theorem, cf. Chapter 4. Thus we can establish a multiple recursive upper bound for the derivation length of R if $\to_R$ is contained in a lexicographic path order. Furthermore, this bound is essentially optimal, cf. Weiermann [1995]. A complete presentation of the argument is given in Moser and Weiermann [2003].

Now let us consider the case where the given finite rewrite system induces a rewrite relation $\to_R$ that is contained in a KBO. Unfortunately the definition of the ordinal interpretation π is not as simple as in Definition 56 and 60. Technically this is due to the more involved definition of KBOs that seemingly does not fit the chosen approach as easily as a lexicographic or multiset path order does. In the next section we will briefly discuss some not so technical reasons for this difference.

As the order $\succ_{\mathsf{kbo}}$ itself is distinctively different form the above considered orders, the proof needs more considerable changes. Due to these difficulties we will not longer dwell on the extension of the presented method to the Knuth Bendix order. However, the reader should bear in mind that the same method is applicable in this context as well. Of course, rendering optimal bounds.

6.6 Bibliographic notes

In this section we have exploited a very nice feature of lexicographic path orders and multiset path orders, namely the fact that these orders abide to the following principle. This principle—henceforth referred to as (CP)—claims that the complexity (or derivation length function) of a rewrite system for which termination is provable using a termination order of order type α is eventually dominated by a function from the *slow-growing hierarchy* along α.

A. Cichon (implicitly) stated this principle in Cichon [1992] as a valid principle for all rewrite systems. In Cichon [1992] the attempt was made to prove the correctness of (CP) for the (i) *multiset path order* ($\succ_{\mathsf{mpo}}$) and the (ii) *lexicographic path order* ($\succ_{\mathsf{lpo}}$). Unfortunately, the proof was bugged, cf. Buchholz [1995].

However, Hofbauer [1992] proved that $\succ_{\mathsf{mpo}}$ as termination order implies primitive recursive derivation length, while Weiermann [1995] showed that $\succ_{\mathsf{lpo}}$ as termination order implies multiply-recursive derivation length. If one regards the order types of $\succ_{\mathsf{mpo}}$ and $\succ_{\mathsf{lpo}}$, respectively, then these results imply the correctness of (CP) for (i) and (ii). Buchholz [1995] has given an alternative proof of (CP) for (i) and (ii). His proof avoids the (sometimes lengthy) calculations with functions from subrecursive hierarchies in Hofbauer [1992], Weiermann [1995]. Instead a clever application of proof-theoretic results is used.

The mentioned proofs Hofbauer [1992], Weiermann [1995], Buchholz [1995] of (CP)—with respect to (i) and (ii)—are *indirect*. I.e. without direct reference to the slow-growing hierarchy. On the other hand, the above proof of the fact that termination proofs by multiset path orders imply primitive derivation lengths (together with its mentioned extension to lexicographic path orders) is a *direct* proof of (CP) for (i) and (ii).

By now, we know from the work of Touzet [1998] and Lepper [2001a, 2003] that (CP) fails to hold in general. In Section 5 we have shown that for any $\alpha < \Lambda$ there exist simply terminating TRSs such that the respective derivation length function eventually dominates H_α. Hence, employing Theorem 22, (CP) cannot be valid for these rewrite systems. Another interesting counter-example are TRSs terminating via KBO. It was shown in Lepper [2001a] that the corresponding derivation length function lifes within $\mathrm{Ack}(2^{O(n)}, 0)$. Furthermore it was shown that the maximal order type of a KBO is ω^ω. Thus, employing Theorem 27 together with Exercise 19, (CP) fails to hold for rewrite systems whose termination is shown via a KBO.

However, as already mentioned, the proof method presented in this chapter is applicable for $\succ_{\mathsf{kbo}}$, too, yielding the (essentially) optimal bound. Hence we can employ the slow growing hierarchy usefully, even if (CP) fails to hold. This leads to the interesting question what the genuine feature of a termination order $\succ$ ought to look like, such that (CP) holds with respect to a rewrite system whose termination has been shown by $\succ$. A question we are not yet able to answer.

Bibliography

Franz Baader and Tobias Nipkow. *Term Rewriting and All That.* Cambridge University Press, Cambridge, 1998.

H. Bachmann. *Transfinite Zahlen.* Springer-Verlag, Berlin, 1967.

W. Buchholz. Ordinal notations and fundamental sequences. Unpublished manuscript; availabe at `www.mathematik.uni-muenchen.de/~buchholz/`, 2003.

Wilfried Buchholz. Kruskals Theorem impliziert $WF(\vartheta(\Omega^\omega))$. Notes in German, 1993.

Wilfried Buchholz. Proof-theoretic analysis of termination proofs. *Annals of Pure and Applied Logic*, 75:57–65, 1995.

Wilfried Buchholz, E. Adam Cichon, and Andreas Weiermann. A uniform approach to fundamental sequences and hierarchies. *Mathematical Logic Quarterly*, 40:273–286, 1994.

E. A. Cichon and E. Tahhan Bittar. Ordinal recursive bounds for Higman's theorem. *Theoretical Computer Science*, 201(1–2):63–84, 1998.

E. A. Cichon and S. S. Wainer. The slow-growing and the Grzegorczyk hierarchies. *Journal of Symbolic Logic*, 48(2):399–408, 1983.

E. Adam Cichon. Termination orderings and complexity characterisations. In Peter Aczel, Harold Simmons, and Stanley S. Wainer, editors, *Proof Theory*, pages 171–193, Cambridge, 1992. Cambridge University Press.

Max Dauchet. Simulation of Turing machines by a left-linear rewrite rule. In N. Dershowitz, editor, *Rewriting Techniques and Applications*, volume 355 of *Lecture Notes in Computer Science*, pages 109–120, Berlin, April 1989. Springer-Verlag.

Max Dauchet. Simulation of Turing machines by a regular rewrite rule. *Theoretical Computer Science*, 103(2):409–420, 1992. A preliminary version appeared as Dauchet [1989].

D. H. J. de Jongh and R. J. Parikh. Well-partial orderings and hierarchies. *Indagationes Mathematicae*, 39(3):195–206, 1977.

N. Dershowitz. Corrigendum to termination of rewriting. *Journal of Symbolic Computation*, 4:409–410, 1987a.

Nachum Dershowitz. A note on simplification orderings. *Information Processing Letters*, 9(5):212–215, 1979.

Nachum Dershowitz. Orderings for term-rewriting systems. *Theoretical Computer Science*, 17:279–301, 1982.

Nachum Dershowitz. Termination of rewriting. *Journal of Symbolic Computation*, 3(1–2):69–115, 1987b. ISSN 0747-7171. See also Dershowitz [1987a].

Nachum Dershowitz and Zohar Manna. Proving termination with multiset orderings. *Communications of the ACM*, 22(8):465–476, 1979.

Nachum Dershowitz and Mitsuhiro Okada. Proof-theoretic techniques for term rewriting theory. In *Third Annual Symposium on Logic in Computer Science*, pages 104–111. IEEE Computer Society, 1988.

Maria da Conceição Fernández Ferreira. *Termination of term rewriting: Well-foundedness, totality and transformations.* PhD thesis, Universiteit Utrecht, 1995. URL http://www.library.uu.nl/digiarchief/dip/diss/01847318/inhoud.htm.

Jean H. Gallier. What's so special about Kruskal's theorem and the ordinal Γ_0? A survey of some results in proof theory. *Annals of Pure and Applied Logic*, 53(3):199–260, 19 September 1991.

Jean-Yves Girard. Π^1_2-logic I. Dilators. *Annals of Mathematical Logic*, 21:75–219, 1981.

Andrzej Grzegorczyk. Some classes of recursive functions. *Rozprawy Matematyczne*, 4, 1953.

G. H. Hardy. A theorem concerning the infinite cardinal numbers. *Quarterly Journal of Mathematics*, 35:87–94, 1904.

D. Hofbauer. Termination proofs by multiset path orderings imply primitive recursive derivation lengths. *Theoretical Computer Science*, 105(1):129–140, 1992.

G. Huet and D. S. Lankford. On the uniform halting problem for term rewriting systems. Report 283, Institut de Recherche en Informatique et en Automatique, Le Chesnay, France, 1978.

Samuel Kamin and Jean-Jacques Lévy. Two generalizations of the recursive path ordering. Unpublished manuscript, 1980.

Laurie Kirby and Jeff Paris. Accessible independence results for Peano arithmetic. *Bulletin of the London Mathematical Society*, 14:285–293, 1982.

D. E. Knuth and P. B. Bendix. Simple word problems in universal algebras. In J. Leech, editor, *Computational Problems in Abstract Algebra*, pages 263–297, Oxford, 1970. Pergamon Press.

Joseph B. Kruskal. Well-quasi-orderings, the tree theorem, and Vázsonyi's conjecture. *Transactions of the American Mathematical Society*, 95:210–225, 1960.

D. S. Lankford. On proving term rewriting systems are Noetherian. Memo MTP-3, Mathematics Department, Louisiana Tech. University, Ruston, LA, May 1979. Revised October 1979.

I. Lepper. Simply terminating rewrite systems with long derivations. *Archive for Mathematical Logic*, 2003. To appear. www.math.uni-muenster.de/logik/publ/pre/3.html.

Ingo Lepper. Derivation lengths and order types of Knuth–Bendix orders. *Theoretical Computer Science*, 269(1–2), 2001a.

Ingo Lepper. *Simplification Orders in Term Rewriting. Derivation Lengths, Order Types, and Computability.* PhD thesis, WWU Münster, 2001b.

Pierre Lescanne. On termination of one rule rewrite systems. *Theoretical Computer Science*, 132(1–2):395–401, 1994.

M. H. Löb and S. S. Wainer. Hierarchies of number-theoretic functions I. *Archiv für Mathematische Logik*, 13:39–51, 1970.

Saunders MacLane. *Categories for the Working Mathematician.* Graduate Texts in Mathematics 5. Springer, 1998.

Aart Middeldorp and Hans Zantema. Simple termination of rewrite systems. *Theoretical Computer Science*, 175(1):127–158, 1997.

Georg Moser and Andreas Weiermann. Relating derivation lengths with the slow-growing hierarchy directly. In *Rewriting Techniques and Applications*, 2003. to appear.

David A. Plaisted. A recursively defined ordering for proving termination of term rewriting systems. Report R-78-943, Department of Computer Science, University of Illinois, Urbana, IL, September 1978.

Wolfram Pohlers. Computability theory of hyperarithmetical sets, 1996. Lecture notes Summer Term 1996.

Rósza Péter. Konstruktion nichtrekursiver Funktionen. *Mathematische Annalen*, 111:42–60, 1935.

Rósza Péter. Über die mehrfache Rekursion. *Mathematische Annalen*, 113: 489–527, 1936.

Rósza Péter. *Recursive Functions.* Academic Press, New York, 1967.

R. W. Ritchie. Classes of recursive functions based on Ackermann's function. *Pacific Journal of Mathematics*, 15:1027–1044, 1965.

Joel W. Robbin. *Subrecursive Hierarchies.* PhD thesis, Princeton University, Princeton, 1965.

Diana Schmidt. *Well-Partial Orderings and their Maximal Order Types.* Habilitationsschrift, Fakultät für Mathematik der Ruprecht-Karl-Universität Heidelberg, 1979.

Helmut Schwichtenberg. Eine Klassifikation der ε_0-rekursiven Funktionen. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 17:61–74, 1971.

J. R. Shoenfield. *Mathematical Logic.* Addison-Wesley, 1967.

Gaisi Takeuti. *Proof Theory.* Number 81 in Studies in Logic and the Foundations of Mathematics. North-Holland Publishing Company, Amsterdam, 2 edition, 1987.

Hélène Touzet. Encoding the Hydra battle as a rewrite system. In L. Brim, J. Gruska, and J. Zlatuška, editors, *Mathematical Foundations of Computer Science*, volume 1450 of *Lecture Notes in Computer Science*, pages 267–276, Berlin, 1998. Springer-Verlag.

Hélène Touzet. A characterization of multiply recursive functions with Higman's lemma. In Paliath Narendran and Michael Rusinowitch, editors, *Rewriting Techniques and Applications*, volume 1631 of *Lecture Notes in Computer Science*, pages 163–174, Berlin, 1999. Springer-Verlag.

O. Veblen. Continuous increasing functions of finite and transfinite ordinals. *Transactions of the American Mathematical Society*, 9:280–292, 1908.

S. S. Wainer. A classification of the ordinal recursive functions. *Archiv für Mathematische Logik*, 13:136–153, 1970.

S. S. Wainer. Ordinal recursion, and a refinement of the extended Grzegorczyk hierarchy. *Journal of Symbolic Logic*, 37(2):281–292, 1972.

A. Weiermann. Proving termination for term rewriting systems. In E. Börger, G. Jäger, H. Kleine Büning, and M. M. Richter, editors, *Computer Science Logic. 5th Workshop, CSL'91*, volume 626 of *Lecture Notes in Computer Science*, pages 419–428, Berlin, 1992. Springer-Verlag.

A. Weiermann. Termination proofs for term rewriting systems by lexicographic path orderings imply multiply recursive derivation lengths. *Theoretical Computer Science*, 139(1–2):355–362, March 1995.

A. Weiermann. Some interesting connections between the slow growing hierarchy and the Ackermann function. *Journal of Symbolic Logic*, pages 609–628, 2001.

Andreas Weiermann. Complexity bounds for some finite forms of Kruskal's theorem. *Journal of Symbolic Computation*, 18(5):463–488, November 1994.

Andreas Weiermann. Sometimes slow growing is fast growing. *Annals of Pure and Applied Logic*, 90(1–3):91–99, 15 December 1997.

H. Zantema. Termination of term rewriting: Interpretation and type elimination. *Journal of Symbolic Computation*, 17(1):23–50, 1994.