**Karlsruhe Reports in Informatics 2017,6**
Edited by Karlsruhe Institute of Technology,
Faculty of Informatics
ISSN 2190-4782

# A First-Order Theory of Ordinals

Peter H. Schmitt

2017

KIT – University of the State of Baden-Wuerttemberg and National
Research Center of the Helmholtz Association

# A First-Order Theory of Ordinals

Peter H. Schmitt

Karlsruhe Institute of Technology (KIT), Dept. of Informatics
Am Fasanengarten 5, 76131 Karlsruhe, Germany

**Abstract.** This technical report documents the work done leading up to a paper on a theory of ordinal submitted to the Tableaux 2017 conference.

## 1 Introduction

The theory of ordinals presented in Section 2 below is the same as in the Tableaux paper. The list of lemmas derivable in the full theory $Th_{Ord}$ in Subsection 2.3 below is considerable larger then was possible under the page restriction of that paper. Section 4 makes good on the promise in the concluding section of the Tableaux paper in giving full proofs of Theorems 2 and 3. These proof use the concept of normal forms for ordinals. References on normal forms abound in the literture. For the convenience of the reader we have however collected all the details we need in Section 3. Sections 6 and 7 on Goodstein numbers and the proof of their termination are an extension of what is covered in the Tableaux paper. We use however different names for the involved functions:

| Tableaux paper | here |
|---|---|
| $oHNF(n, m)$ | $f^{m,n}(n+1)$ |
| $oGS(n, m)$ | $o(n, m)$ |

The general three-argument function $f^{n,m}(x)$ has no counterpart in the Tableaux paper.

Furthermore this technical report contains the full JML-annotated Java program for Goodstein sequences, Figures 17 and 18. But here there is also room for the Java program that computes Goodstein sequences using BigInt, Figure 16.

Also in this report, but not in the Tableaux paper, is Section 5 containing a small example of a program termination proof using ordinals.

## 2 A Theory of Ordinals

We consulted the books [3,15] and also the books [9,10] in German on ordinal arithmetic in axiomatic set theory.

| | mathematical notation | | | Dynamic Logic | JML |
|---|---|---|---|---|---|
| **predicate** | $n < m$ | : | $(Ord, Ord)$ | $olt(n, m)$ | $\backslash ord\_less(n, m)$ |
| **functions** | $n + 1$ | : | $Ord \to Ord$ | $oadd(n, o\_1)$ | $\backslash ord\_add(n, \backslash o\_1)$ |
| | $0$ | : | $Ord$ | $o\_0$ | $\backslash o\_0$ |
| | $\omega$ | : | $Ord$ | $omega$ | $\backslash omega$ |
| **binder** | $sup_{m<b}j$ | : | $Ord \times Ord \to Ord$ | $osup\{m;\}(b, j)$ | $---$ |

**Fig. 1.** The vocabulary of the Core Theory

### 2.1 The Core Theory

We start out with a very simple core theory $Th^0_{Ord}$. This plays the same role here as Peano's theory for the arithmetic of natural number. The vocabulary of $Th^0_{Ord}$ is shown in Figure 1. In this text we use the mathematical notation throughout. The figure also gives the corresponding notation in Dynamic Logic and the corresponding JML notation that is used in annotating Java programs. In the binder symbol $m$ is the bound variable. The intended meaning of these symbols is fixed by the axioms in Figure 2.

| | | |
|---|---|---|
| 1. | $\forall x, y, z(x < y \land y < z \to x < z)$ | transitivity |
| 2. | $\forall x(\neg x < x)$ | strict order |
| 3. | $\forall x, y(x < y \lor x \doteq y \lor y < x)$ | total order |
| 4. | $\forall x(0 \leq x)$ | 0 is smallest element |
| 5. | $0 < \omega \land \neg \exists x(\omega \doteq x + 1)$ | $\omega$ is a limit ordinal |
| 6. | $\forall y(0 < y \land \forall x(x < \omega \to x + 1 < y) - > \omega \leq y)$ | $\omega$ is the least limit ordinal |
| 7. | $\forall x(x < x + 1) \land \forall x, y(x < y \to x + 1 \leq y)$ | successor function |
| 8. | $\forall z(z < \alpha \to t[z/\lambda] \leq sup_{\lambda<\alpha}t)$ | def of supremum, part 1 |
| 9. | $\forall x(\forall z(z < \alpha \to t[z/\lambda] \leq x) \to sup_{\lambda<\alpha}t \leq x)$ | def of supremum, part 2 |
| 10. | $\forall x(\forall y(y < x \to \phi(y)) \to \phi) \to \forall x \phi$ | transfinite induction scheme |

**Fig. 2.** The axioms of the Core Theory

We thus see that $<$ is to be interpreted as a strict, total, order relation. We use already here $x \leq y$ as a shorthand for $x \doteq y \lor x < y$. This predicate will be systematically included in the extended vocabulary introduced below. The constants 0 and $\omega$ are part of the core vocabulary. From the axioms we see that 0 is the least element with respect to $<$ and $\omega$ is the first infinite ordinal. Without the two axioms for $\omega$ the natural numbers with strict order and successor would be a model of the remaining axioms and nothing would have been gained over Peano's theory. We use $x + 1$ to denote the immediate successor of $x$ to avoid the introduction of an additional symbol. Furthermore $sup_{\lambda<\alpha}t$ is the strict supremum of the $\{t \mid \lambda < \alpha\}$. Here the term $t$ will typically contain the variable $\lambda$, while $\lambda$ is not allowed to occur in $\alpha$. We use the notation $t[z/\lambda]$ to denote the term that arises from $t$ by replacing $\lambda$ by $z$ everywhere. The last and most powerful axiom is the axiom of transfinite induction that is an extension the course-of-value induction scheme in finite arithmetic: Let $\phi$ be a formula that typically would contain $x$ as a free variable and let $\phi(y)$ stand for the formula obtained from $\phi$ by replacing $x$ by $y$ (assuming of course that $y$ does not occur in $\phi$, neither free nor bound). If we can prove for every $x$ the transfinite induction step $\forall y(y < x \to \phi(y)) \to \phi$ then we conclude that $\phi$ is true for all $x$.

One could ask whether the *sup* operator should be part of the core vocabulary or if it would not have been better to include it later as a definitional extension. The answer is *no!* Since we follow the usual set-up of first-order logic all functions are total. Consequently, inclusion of a function symbol in the vocabulary already implies an implicit existence axiom: the function values exist for all arguments. Adding *sup* is not a definitional extension since the associated existence claim could not be proved in the theory without *sup*.

An alternative would have been to include the following axiom scheme instead of the two parts of the definition of *sup*

$$\forall \lambda(\exists z(\forall x(x < \lambda \to t \leq z)))$$

Then it would than have been possible to show, using transfinite induction, that adding the *sup* operator is a definitional extension of this version of the core theory. The adopted approach is more straightforward.

Note, that the way we definied *sup* the formulas $sup_{\lambda<0}t \doteq 0$ and $sup_{\lambda<1}t \doteq t[0/\lambda]$ can be derived. By $t[0/\lambda]$ we denote the term that arises from $t$ by replacing everywhere the variable $\lambda$ by the constant 0.

As an alternative of the supremum *sup* defined in Figure 2 we could have defined a strict supremum $ssup_{\lambda<\alpha}t$ as the least ordinal that is strictly greater than all $t[\lambda]$. Figure 3 shows what further changes would have been entailed.

| osup | ossup |
|---|---|
| $\forall z(z < \alpha \to t[z/\lambda] \leq sup_{\lambda<\alpha}t)$ | $\forall z(z < \alpha \to t[z/\lambda] < sup_{\lambda<\alpha}t)$ |
| $\forall x(\forall z(z < \alpha \to t[z/\lambda] \leq x)$<br>$\to sup_{\lambda<\alpha}t) \leq x)$ | $\forall x(\forall z(z < \alpha \to t[z/\lambda] < x)$<br>$\to sup_{\lambda<\alpha}t) \leq x)$ |
| $sup_{\lambda<1}\ t \doteq t[0/\lambda]$ | $sup_{\lambda<1}\ t \doteq t[0/\lambda] + 1$ |
| $\forall x(lim(x) \to sup_{\lambda<x}\ \lambda \doteq x)$ | $\forall x(sup_{\lambda<x}\ \lambda \doteq x)$ |
| $\forall x(sup_{\lambda<x+1}\ t \doteq max(sup_{\lambda<x}\ t, t[x/\lambda]))$ | $\forall x(sup_{\lambda<x+1}\ t \doteq max(sup_{\lambda<x}\ t, t[x/\lambda] + 1))$ |
| $\forall x, y(lim(y) \to x * y = sup_{\lambda<y}(x * \lambda))$ | $\forall x, y(lim(y) \to x * y =$<br>if $x \doteq 0$ then 0 else $sup_{\lambda<y}(x * \lambda))$ |

**Fig. 3.** Comparison between supremum and strict supremum

Note, that the way we definied *sup* the formulas $sup_{\lambda<0}t \doteq 0$ and $sup_{\lambda<1}t \doteq t[0/\lambda] + 1$ can be derived. By $t[0/\lambda]$ we denote the term that arises from $t$ by replacing everywhere the variable $\lambda$ by the constant 0.

A real disadvantage of $ssup$ is the fact that the equation $sup_{\lambda<x} i * t = i * sup_{\lambda<x} t$ is not true. Here are counterexamples for $x$ a limit and $x$ a succsessor ordinal:

$$\begin{array}{lllll} \omega & = ssup_{n<\omega} 2 * n & \neq 2 * ssup_{n<\omega} n & = 2 * \omega \\ i * x + 1 & = ssup_{\lambda<x+1} i * \lambda & \neq i * ssup_{\lambda<x+1} \lambda & = i * (x+1) \end{array}$$

### 2.2 Extension with Auxiliary Predicates

We will describe the extensions of the core theory in several installments in the following subsections. Figure 4 shows the final vocabulary $\Sigma_{Ord}$ of the (extended) theory of ordinals $Th_{Ord}$ after all these extensions. The axioms of $Th_{Ord}$ are $Th^0_{Ord}$ plus the definitions of the new symbols in Figure 5 and Figure 7.

| | mathematical notation | Typing | Dynamic Logic | JML |
|---|---|---|---|---|
| **predicates** | $\leq$ | $(Ord, Ord)$ | *oleq* | \*ord_leq* |
| | *lim* | $(Ord)$ | *lim* | \*ord_lim* |
| **functions** | 1 | *Ord* | *o_1* | \*o_1* |
| | $\omega$ | *Ord* | *omega* | \*omega* |
| | *max* | $Ord \times Ord \to Ord$ | *omax* | \*ord_max* |
| | | | | |
| | $+$ | $Ord \times Ord \to Ord$ | *oadd* | \*ord_add* |
| | $*$ | $Ord \times Ord \to Ord$ | *otimes* | \*ord_times* |
| | ˆ | $Ord \times Ord \to Ord$ | *oexp* | \*ord_exp* |

**Fig. 4.** Extended vocabulary $\Sigma_{Ord}$

In this subsection we concentrate on the definition of the auxiliary symbols $\leq$ and $lim$ in Figure 5. In our set-up 0 is not a limit ordinal.

$$\begin{array}{lr} \forall x, y(x \leq y \leftrightarrow x \doteq y \vee x < y) & \text{(less or equal relation)} \\ \forall x(lim(x) \leftrightarrow 0 < x \wedge \neg \exists y(y + 1 \doteq x)) & \text{(limit ordinal)} \\ \forall x, y(max(x, y) \doteq \text{if } \ x \leq y \ \text{ then } \ y \ \text{ else } \ x) & \text{(maximum operator)} \end{array}$$

**Fig. 5.** Definitional extension: Axioms for auxiliary predicates

Already at this point we can derive a couple of useful lemmas from the axioms considered so far. A non-exhaustive list is shown in Figure 6.

A general comment on presentation is necessary here. In the following we group lemmas together according to the syntactical symbols involved in them.

This does not reflect the order in which the lemmas can or need to be proved in. In the KeY system the axioms and derived lemmas are formulated in a dedicated proof rule language, called the taclet language. Full explanation of the taclet language can be found in [1, Chapter 4]. The order the taclets appear in the rules files `ordRules.key` and `intOrdRules.key` is the order that is used to make sure that when the proof of a lemma is started that all the auxiliary lemmas have been proved before and that cyclic dependencies are avoided.

1. $\exists x \phi \rightarrow \exists x(\phi \wedge \forall y(y < x \rightarrow \neg\phi[y/x]))$
2. $\forall x, y, z(x \leq y \wedge y \leq z \rightarrow x \leq z)$
3. $\forall x, y, z(x \leq y \wedge y < z \rightarrow x < z)$
4. $\forall x, y, z(x < y \wedge y \leq z \rightarrow x < z)$
5. $\forall x, y(x \leq y \wedge y \leq x \rightarrow x \doteq y)$
6. $\forall x, y, z(max(x, y) < z \leftrightarrow (x < z \wedge y < z))$
7. $\forall x, y, z(z < (max(x, y) \leftrightarrow (z < x \vee z < y)))$
8. $\forall x, y, z(max(x, y) \leq z \leftrightarrow (x \leq z \wedge y \leq z))$
9. $\forall x, y, z(z \leq (max(x, y) \leftrightarrow (z \leq x \vee z \leq y)))$
10. $\forall x(max(0, x) \doteq max(x, 0) \doteq x)$
11. $sup_{\lambda<0}\ t \doteq 0$
12. $sup_{\lambda<1}\ t \doteq t[0/\lambda]$
13. $\forall x(lim(x) \rightarrow sup_{\lambda<x}\ \lambda \doteq x)$
14. $\forall x(sup_{\lambda<x+1}\ t \doteq max(sup_{\lambda<x}\ t, t[x/\lambda]))$
15. $\forall x(\forall y(y < x \rightarrow t_1[y/\lambda] \doteq t_2[y/\lambda]) \rightarrow sup_{\lambda<x}\ t_1 \doteq sup_{\lambda<x}\ t_2)$
16. $\forall \alpha_1, \alpha_2($
    $\forall x(x < \alpha_1 \rightarrow \exists y(y < \alpha_2 \wedge t_1[x/\lambda] \leq t_2[y/\lambda]))\wedge$
    $\forall y(y < \alpha_2 \rightarrow \exists x(x < \alpha_1 \wedge t_2[y/\lambda] \leq t_1[x/\lambda]))$
    $\leftrightarrow sup_{\lambda<\alpha_1}\ t_1 \doteq sup_{\lambda<\alpha_2}\ t_2)$
17. $lim(\lambda) \leftrightarrow \lambda \neq 0 \wedge \forall ov(ov < \lambda \rightarrow (ov + 1) < \lambda$
18. $\forall \lambda(t_1 \leq t_2 \rightarrow sup_{\lambda<b}\ t_1 \leq sup_{\lambda<b}\ t_2$

**Fig. 6.** Some lemmas derivable from the axioms considered so far

We now comment on the lemmas in Figure 6. Lemma 1 in Figure 6 is the least number principle, a well known equivalent to the induction axiom scheme. It is instructive to figure out why this lemma is true even if $x$ does not occur as a free variable in $\phi$. Lemmas 2 -4 are versions of transitivity involving both relations $\leq$ and $<$. Lemma 5 is an easy consequence of the definition of $\leq$ and a welcome simplification in many proofs. Lemmas 6 to 9 show how tha maximum operator behaves with respect to the order relations $<$ and $\leq$. Lemma 10 is a simplification rule that follows from the fact the 0 is the least ordinal. Equation 11 is true regardless of $t$. Lemnma 13 could be rephrased as: $x$ is the least ordinal that is greater or equal than all ordinals that are strictly less than $x$. This is only true if $x$ is a limit ordinal. In the successor case we have $sup_{\lambda<x+1}\lambda \doteq x$. Lemma 14 is usefull in proving statements involving the *sup* operator via induction. Lemma

15 helps to show that to suprema are equal especially in the case when equality between $t_1$ and $t_2$ is not obvious.

We look at a term $t$ that contains $\lambda$ as a sequence $t_\lambda$. We say sequence $t_{\lambda<\alpha_1}$ is confinal in $s_{\lambda<\alpha_2}$ if for every $x < \alpha_1$ there is $y < \alpha_2$ with $t[x/\lambda] \leq s[y/\lambda]$. If two sequences are mutually confinal in one another than they share the same supremum. This is Lemma 16 in Figure 6. Note, that we get equality of two suprema with different bounds $\alpha_1$ and $\alpha_2$. Lemma 17 gives an alternative definition of a limit ordinal.

### 2.3 Defining Ordinal Arithmetic

$$
\begin{array}{l}
\forall x(x + 0 \doteq x)\\
\forall x, y(x + (y + 1) \doteq (x + y) + 1)\\
\forall x, y(lim(y) \to x + sup_{\lambda<y}\lambda \doteq sup_{\lambda<y}(x + \lambda))\\
\forall x(x * 0 \doteq 0)\\
\forall x, y(x * (y + 1) \doteq (x * y) + x)\\
\forall x, y(lim(y) \to x * y \doteq sup_{\lambda<y}(x * \lambda))\\
\forall x(x^0 \doteq 1)\\
\forall x, y(x^{y+1}) \doteq (x^y) * x)\\
\forall x, y(lim(y) \land x \neq 0 \to x^y \doteq sup_{\lambda<y}(x^\lambda))\\
\forall y(lim(y) \to 0^y \doteq 0)
\end{array}
$$

**Fig. 7.** Definitional extension: Axioms for arithmetic operations

Figure 7 gives the usual recursive defintions of addition, multiplication and exponentiation.

To prove the equations that still hold among the three arithmetical operators we need some preparation. Figure 8 list some of the intermediate stepping stones.

1. $\forall x, y(y \neq 0 \to x < x + y)$
2. $\forall x, y(x \leq x + y)$
3. $\forall x, y(y \leq x + y)$
4. $\forall x, y, z(x < y \to z + x < z + y)$
5. $\forall x, y, z(x \leq y \to x + z \leq y + z)$
6. $\forall x, y, z(x + y < x + z \to y < z)$
7. $\forall x, y, u, w(x < y \land u < w \to x + u < y + w)$
8. $\forall x, y, u, w(x \leq y \land u \leq w \to x + u \leq y + w)$
9. $(i < \omega \land j < \omega) \to i + j < \omega$

**Fig. 8.** Lemmas involving addition and order relations

Lemma 1 in Figure 8 extends the axiom $\forall x(x < x + 1)$ from Figure 2 where we now add on the right side an arbitrary number greater or equal to 1 instead of just 1. This is, of course, proved via transfinite induction. Lemma 2 is an easy variant of Lemma 1. Addition of ordinals is not commutative, so we cannot conclude from Lemma 1 that $\forall x, y(x \neq 0 \rightarrow y < x + y)$. In fact, $y = \omega$, $x = 1$ is a counterexample. But, the version for $\leq$ instead of $<$ is provable. This is Lemma 3. Lemma 4 is also proved using transfinite induction. We remark that $\forall x, y, z(x < y \rightarrow x + z < y + z)$ is not true, as can be seen by the instantiation 0 for $x$, 1 for $y$, and $\omega$ for $z$. But, the relaxed version with $\leq$ instead of $<$ is derivable, this is Lemma 5. Lemma 6 is the reverse of Lemma 4.

1. $\forall x(0 + x \doteq x)$
2. $\forall x, y(x + y \doteq 0 \leftrightarrow x \doteq 0 \wedge y \doteq 0)$
3. $\forall x, y, z(max(z + x, z + y) \doteq z + max(x, y)$
4. $\forall x, y, z(max(x + z, y + z) \doteq max(x, y) + z$
5. $\forall x(x < \omega \rightarrow x + \omega \doteq \omega)$
6. $\forall x(lim(\lambda) \rightarrow lim(x + \lambda))$
7. $\forall x(\omega \leq x \rightarrow \exists \lambda, n(lim(\lambda) \wedge n < \omega \wedge x \doteq \lambda + n))$
8. $\forall x, y(x \leq y \rightarrow \exists z(x \doteq y + z))$
9. $\forall x, y, z(x + (y + z) \doteq (x + y) + z)$
10. $\neg \exists x(x + 1 \doteq 0)$
11. $\forall x, y(x < y \rightarrow (x + 1) < (y + 1))$
12. $\forall x, y((x + 1) \doteq (y + 1) \rightarrow x \doteq y)$
13. $\forall x, y, z((z + x) \doteq (z + y) \rightarrow x \doteq y)$
14. $sup_{\lambda<x}\ (i + t) \doteq i + sup_{\lambda<x}\ t$ if $\lambda$ does not occur in $i$ and $x > 0$
15. $i + j \doteq j$ if $\omega \leq j$ and $i < \omega$

**Fig. 9.** Lemmas on addition

Since ordinal addition is in general not commutative Lemma 1 in Figure 9 may not be immediately obvious, but it can be easily proved using ordinal induction. Lemma 5 is a fact on ordinal addtion that we have referred to already above. Lemma 6 is a useful lemma formalizing the intuition that the property of being a limit ordinal is determined by the right *end part* of the ordinal regardless of what come before. Lemma 7 gives a first general representation theorem for ordinals. In [15, Theorem 8.13] it is proved using set comprehension. This is not available in our setting. Fortunately, it turned out that there is a much simpler proof using ordinal induction. Lemma 8 required the most complex proof so far. The basic idea, however, is quite simple. As a witness for $z$ take $b$, the least ordinal such that $y \leq x + b$. It can easily be seen that such a number exists by the least number principle (Lemma 1 in Figure 6). Then a case distinction $b = 0$, $b \doteq b_0 + 1$ for some $b_0$, or $lim(b)$ leads to success. Lemma 9 is the wellknown associative law. The next three lemmas could have come earlier. Lemma 10 and 12 correspond to the Peano axioms for the natural numbers, which say that 0 is not a successor and the successor function is injective. Lemma 11 is a stepping stone in the proof of 12. Lemma 13 shows that addition on the right, with fixed left summand, is injective. Lemma 14 resisted for a while all my attempts to prove it. Since I could also not find it in [15] I was, at some point, even in doubt wether it is true at all. The inequality $sup_{\lambda<x}\ (i+t) \leq i + sup_{\lambda<x}\ t$ is simple. For the reverse inequality a proof by contradiction turned out to be the right way of attack. So assume $sup_{\lambda<x}\ (i+t) < i + sup_{\lambda<x}\ t$ and try to find a contradiction. The key to the solution, at least my solution, was the case distinction $sup_{\lambda<x}\ (i+t) < i$ or $i \leq sup_{\lambda<x}\ (i+t)$. (I later discovered that in [10] the lemma is proved by a case disctinction whether $sup_{\lambda<x}\ (i+t)$ is a limit or a successor ordinal.) In the first case we arrive at the contradiction $i \leq i + t[0/\lambda] < sup_{\lambda<x}\ (i+t) < i$. Here also the assumption $x > 0$ comes in. In the second case there is by Lemma 5 an ordinal $k$ such that $i + k \doteq sup_{\lambda<x}\ (i+t)$. By the proof-by-contradiction assumption this yiels $i+k < i+sup_{\lambda<x}\ t$ and further by Lemma 9 in Figure 8 $k < sup_{\lambda<x}\ t$. By the definition of *sup* there is $\lambda_0 < x$ with $k \leq t[\lambda_0/\lambda]$. This leads to the contradiction $i + k \leq i + t[\lambda_0/\lambda] < sup_{\lambda<x}\ (i+t)$. The *commuted* version of Lemma 14, i.e., $sup_{\lambda<x}\ (t+i) \doteq (sup_{\lambda<x}\ t) + i$ provided $\lambda$ does not occur in $i$ and $x > 0$, is - as you would have expected -not true: $\omega \doteq sup_{\lambda<\omega}\ (\lambda+1) \doteq (sup_{\lambda<\omega}\ \lambda)+1 \doteq \omega+1$ .

Lemma 15 shows a dramatic failure of commutativity for ordinal addition: A left finite ordinal summand is simply absorbed if the right summand is infinite. We found it helpful to split the proof of Lemma 15 in the cases $\omega \doteq j$ and $\omega < j$.

1. $\forall x(1 * x \doteq x * 1 \doteq x)$
2. $\forall x(0 * x \doteq 0)$
3. $\forall x, y, z((0 < z \wedge x < y) \rightarrow z * x < z * y)$
4. $\forall x, y, z(z * x < z * y) \rightarrow (0 < z \wedge x < y))$
5. $\forall x, y, z(0 < z \wedge z * x \doteq z * y \rightarrow x \doteq y)$
6. $\forall x, y, z(x \leq y \rightarrow x * z \leq y * z)$
7. $\forall x, y(x \neq 0 \rightarrow y \leq x * y$
8. $i * j \doteq 0 \leftrightarrow i \doteq 0 \vee j \doteq 0$
9. $i * j \doteq 1 \leftrightarrow i \doteq 1 \wedge j \doteq 1$
10. $\forall x, y(x < \omega \wedge y < \omega \rightarrow x * y < \omega)$
11. $\forall x(0 < x < \omega \rightarrow x * \omega \doteq \omega)$
12. $\forall x, y, z(max(z * x, z * y) \doteq z * max(x, y))$
13. $\forall x, y, z(max(x * z, y * z) \doteq max(x, y) * z)$
14. $sup_{\lambda<x}\ (i * t) \doteq i * sup_{\lambda<x}\ t$
15. $\forall i, j, k(i * (j + k) \doteq i * j + i * k)$
16. $\forall i, j, k((i * j) * k \doteq i * (j * k)$
17. $\forall \lambda, n, i(lim(\lambda) \wedge n < \omega \rightarrow i * \lambda \doteq (i + n) * \lambda)$
18. $\forall \lambda, x((lim(\lambda) \wedge 0 < x < \omega) \rightarrow x * \lambda \doteq \lambda)$
19. $\forall i, j(1 < i \wedge 1 < j \rightarrow i + j \leq i * j)$
20. $\forall i, \lambda(0 < i \wedge lim(\lambda) \rightarrow lim(i * \lambda))$
21. $\forall i, \lambda(0 < i \wedge lim(\lambda) \rightarrow lim(\lambda * i))$

**Fig. 10.** Lemmas on multiplication

Figure 10 shows derivable properties of ordinal multiplication. Lemma 3 shows that the strict order relation is preserved by multiplication on the left provided that the left multiplyer is not 0. Multiplication on the right only preserves $\leq$, as Lemma 6 further down shows. Lemma 4 is the reverse implication from Lemma 3. Lemma 5 states that multiplication on the right, with a fixed multiplicand on the left, is an injective function. Lemmas 12 and 13 parallel Lemmas 3 and 4 from Figure 9, here for multiplication instead of addition. Lemma 14 is crucial for the proof of distributivity (Lemma 15) and multiplicative associativity (Lemma 16). After all the preparations the proof of multiplicative associativity is now straight forward. We use ordinal induction on the variable $k$. The base case is trivial. The successor induction step is proved as follows:

$$\begin{aligned} i * (j * (k + 1)) &\doteq i * (j * k + j) && \text{definition of } * \\ &\doteq i * (j * k) + i * j && \text{distributivity, Lemma 15} \\ &\doteq (i * j) * k + i * j && \text{induction hypothesis} \\ &\doteq (i * j) * (k + 1) && \text{definition of } * \end{aligned}$$

The induction step in the limit case is shown next. We us $\lambda$ instead of $k$ to signal that $k$ is a limit ordinal:

$$\begin{aligned} i * (j * \lambda) &\doteq i * sup_{x<\lambda} j * x && \text{definition of } * \\ &\doteq sup_{x<\lambda} i * (j * x) && \text{Lemma 14} \\ &\doteq sup_{x<\lambda} (i * j) * x && \text{induction hypothesis} \\ &\doteq (i * j) * \lambda && \text{definition of } * \end{aligned}$$

Lemma 18 (we are still talking about Figure 10) is a strengthening of Lemma 10: multiplicative absorbtion on the left of finite ordinals not only holds for $\omega$ but for any limit ordinal $\lambda$. Lemma 18 is an auxilliary step in the proof of 18. Lemma 19 states when addition of two ordinals is less than their product. The restrictions are necessary as can be seen by the simple examples

$$\begin{aligned} j = 0 + j &\not\leq 0 * j = 0 \\ 1 + j &\not\leq 1 * j = j \\ i = i + 0 &\not\leq i * 0 = 0 \\ i + 1 &\not\leq i * 1 = i \end{aligned}$$

Though commutativity of addition and multiplication and the second distributive law fail when all ordinals are considered they still hold true for finite ordinals. This is recorded in Figure 11.

1. $\forall x, y(x < \omega \wedge y < \omega \rightarrow x + y \doteq y + x)$
2. $\forall i, j, k((i < \omega \wedge j < \omega \wedge k > \omega) \rightarrow (i + j) * k \doteq i * k + j * k)$
3. $\forall x, y(x < \omega \wedge y < \omega \rightarrow x * y \doteq y * x)$

**Fig. 11.** Lemmas on finite ordinals

Next we turn to the investigation of the laws for exponentiation. The first three lemmas in Figure 12 cover simple equations when the base 0 or 1 or exponent is 1. Note, that for Lemma 2 the restriction on $x$ is neccessary since be definition we have $0^0 \doteq 1$. Also the restrictions in Lemma 4 are neccessary as can be seen by the following examples $2 \not< 2^0 \doteq 1$, $2 \not< 2^1 \doteq 2$, $0 \not< 0^2 \doteq 0$, and $1 \not< 1^2 \doteq 1$.

Lemma 1 of Figure 13 prepares the ground for the next lemma which is the ordinal version of division with remainder. Lemma 3 is an easy instance of Lemma 2 and a very useful representation of limit numbers.

1. $x^1 \doteq x$
2. $\forall x(0 < x \to 0^x \doteq 0)$
3. $\forall x(1^x \doteq 1)$
4. $\forall x, y(1 < x \land 1 < y \to x < x^y)$
5. $\forall x, 0 < y \to x \leq x^y)$
6. $\forall x, y(1 < x \land 0 < y \to 1 < x^y)$
7. $\forall x, y(1 < x \to 1 \leq x^y)$
8. $\forall x, y(x \neq 1 \to x * y \leq x^y)$
9. $\forall x, y(1 < x \to y \leq x^y)$
10. $\forall x, y_1, y_2(1 < x \land y_1 < y_2 \to x^{y_1} < x^{y_2})$
11. $\forall x, y_1, y_2(1 < x \land x^{y_1} < x^{y_2} \to y_1 < y_2)$
12. $\forall x_1, x_2, y(x_1 < x_2 \to x_1^y \leq x_2^y)$
13. $\forall x_1, x_2, y(x_1 < x_2 \land 0 < y \land \neg lim(y) \to x_1^y < x_2^y)$
14. $x^y \doteq 0 \to x \doteq 0 \land y \neq 0$
15. $x^y \doteq 1 \to y \doteq 0 \lor x \doteq 1$
16. $\forall x, y(x < \omega \land y < \omega \to x^y < \omega)$
17. $\forall x, y(1 < x \land x < \omega \to x^\omega \doteq \omega)$
18. $\forall x, y((0 < x \land lim(y) \to lim(y^x)$
19. $\forall x, y(1 < x \land lim(y) \to lim(x^y)$
20. $\forall x, y, z(x^{y+z} \doteq x^y * x^z)$
21. $\forall x, y, z((x^y)^z \doteq x^{y*z})$
22. $\forall b((0 < b \land \forall x(x < b \to 0 < j)) \to sup_{x<b}(i^j) = i^{sup_{x<b}(j)})$
    for all terms $i, j$ such that $x$ does not occur in $i$.

**Fig. 12.** Lemmas on exponentiation

1. $\forall x, y(y \neq 0 \to \exists z(y * z \leq x < y * (z + 1)))$
2. $\forall x, y(y \neq 0 \to \exists d \exists r(x \doteq y * d + r \land r < y))$
3. $\forall x(lim(x) \to \exists d(x \doteq \omega * d))$
4. $\forall x, y(0 < x \land 1 < y \to \exists z(y^z \leq x < y^{z+1}))$
5. $\forall x, y, u, w(x < y \land u < \omega \land w \neq 0 \to \omega^x * u + \omega^y * w \doteq \omega^y * w$
6. $\forall x, y, m, z(\ (lim(x) \land 0 < y \land 0 < m < \omega \land \exists z_1(z \doteq z_1 + 1))$
    $\to (x^y * m)^z \doteq x^{y*z} * m$
7. $\forall x, y, m, z(\ (lim(x) \land 0 < y \land 0 < m < \omega \land lim(z))$
    $\to (x^y * m)^z \doteq x^{y*z}$

**Fig. 13.** Lemmas leading to a normal form

## 3 Normal Form

For simplicity we will use in this section constants $n \in \mathbb{N}$ also for $n > 1$ as abbreviation for the $n$-fold iterated sum $1 + 1 \ldots + 1$.

A ground term is built up from constants $\{n \mid n \in \mathbb{N}\}$ and the functions symbols $+, *, exp$. The set HNFof normal forms is a subset of the set of all ground terms. It will play a special role in the following. Simultaneous with the inductive definition of normal forms we define a binary relation $\gg$ on normal forms and a rank function $rk$.

**Definition 1 (Normal Form).**

1. *A constant $n \in \mathbb{N}$ is a normal form of rank $0$, in symbols $rk(n) = 0$, and $n \gg m$ if $n$ is strictly greater than $m$ as natural numbers.*
2. *If $\alpha_i$ for $0 \leq i < k$ are normal forms such that $\alpha_0 \gg \alpha_1 \gg \ldots > \gg \alpha_{k-1} \gg 0$, $k > 0$, and $m_i \in \mathbb{N}$ for $0 \leq i < k$, and $m_i > 0$ for $0 \leq i < (k-1)$ then*

   $$t = \omega^{\alpha_0} * m_0 + \omega^{\alpha_1} * m_1 + \ldots + \omega^{\alpha_{k-2}} * m_{k-2} + m_{k-1}$$

   *is a normal form.*
   *If $k = 1$ then $rk(t) = 0$, otherwise $rk(t) = max\{rk(\alpha_i) \mid 0 \leq i < k\} + 1$.*
   *If*

   $$t_1 = \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1}$$
   $$t_2 = \omega^{\beta_0} * b_0 + \omega^{\beta_1} * b_1 + \ldots + \omega^{\beta_{s-2}} * b_{s-2} + b_{s-1}$$

   *are normal forms, with $t_1 \neq t_2$ let $t_0$ be their greatest common initial sum, i.e.*

   $$t_1 = t_0 + \omega^{\alpha_i} * a_i + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1}$$
   $$t_2 = t_0 + \omega^{\beta_j} * b_j + \ldots + \omega^{\beta_{s-2}} * b_{s-2} + b_{s-1}$$

   *Then $t_1 \gg t_2$ if $\alpha_i \gg \beta_j$ or $\alpha_i = \beta_j$ and $a_i > b_j$.*

This normal form shows some similarity to Cantor normal form. It differs in that is is hereditary, i.e. the exponents of a normal form are again normal forms and it is a strictly syntactic concept.

**Lemma 1.** *The relation $\gg$ is total, i.e. for any two normal forms $t_1$, $t_2$*

$$\textit{either } t_1 \gg t_2 \textit{ or } t_2 \gg t_1 \textit{ or } t_1 = t_2$$

*Proof.* The proof proceeds by induction on $max(rk(t_1), rk(t_2))$
If $t_1, t_2 \in \mathbb{N}$ the claim is obvious.
In the general case, assume $t_1 \neq t_2$ and let $t_0$ be the greatest common initial sum:

$$t_1 = t_0 + \omega^{\alpha_i} * a_i + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1}$$
$$t_2 = t_0 + \omega^{\beta_j} * b_j + \ldots + \omega^{\beta_{s-2}} * b_{s-2} + b_{s-1}$$

Since $max(rk(\alpha_i), rk(\beta_j)) < max(rk(t_1), rk(t_2))$ we obtain from the induction hypothesis $\alpha_i \gg \beta_j$ or $\beta_j \gg \alpha_i$ or $\alpha_i = \beta_j$. This implies $t_1 \gg t_2$ in the first case, $t_2 \gg t_1$ in the second. In case $\alpha_i = \beta_j$ we have either $a_i > b_j$ or $b_j > a_i$, since $a_i = b_j$ would contradict the maximality of $t_0$. Thus again $t_1 \gg t_2$ in the first case, and $t_2 \gg t_1$ in the second. □

**Lemma 2.** *If $\omega^{\alpha_0} * m_0 + \omega^{\alpha_1} * m_1 + \ldots + \omega^{\alpha_{k-2}} * m_{k-2} + m_{k-1}$ is a normal form then*

$$\omega^{\alpha_0} * m_0 + \omega^{\alpha_1} * m_1 + \ldots + \omega^{\alpha_{k-2}} * m_{k-2} + m_{k-1} \quad < \quad \omega^{\alpha_0} * (m_0 + 1)$$

*is derivable in $Th_{Ord}$.*

*Proof.* The proof will proceed by induction on $k$.
In the initial case $k = 1$ the claim reduces to $m_0 < (m_0 + 1)$ which is obviously true.
For the step case assume

$$\omega^{\alpha_1} * m_1 + \omega^{\alpha_1} * m_1 + \ldots + \omega^{\alpha_{k-2}} * m_{k-2} + m_{k-1} < \omega^{\alpha_1} * (m_1 + 1) \qquad (1)$$

We will show

$$\omega^{\alpha_0} * m_0 + \omega^{\alpha_1} * m_1 + \ldots + \omega^{\alpha_{k-2}} * m_{k-2} + m_{k-1} < \omega^{\alpha_0} * (m_0 + 1) \qquad (2)$$

By Lemma 4 in Figure 8 we obtain from (1):

$$\begin{array}{l} \omega^{\alpha_0} * m_0 + \omega^{\alpha_1} * m_1 + \omega^{\alpha_1} * m_1 + \ldots + \omega^{\alpha_{k-2}} * m_{k-2} + m_{k-1} \\ < \\ \omega^{\alpha_0} * m_0 + \omega^{\alpha_1} * (m_1 + 1) \end{array} \qquad (3)$$

It thus suffices to show for $0 < a_0, a_1 < \omega$, $\alpha_0 > \alpha_1$

$$\omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 < \omega^{\alpha_0} * (a_0 + 1) \qquad (4)$$

By Lemma 8 in Figure 9 we get $k$ with $\alpha_0 = \alpha_1 + k$. Obviously, $k > 0$. This allows the following derivation

$$\begin{array}{rlr} \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 & \doteq \omega^{(\alpha_1 + k)} * a_0 + \omega^{\alpha_1} * a_1 & \text{choice of } k \\ & \doteq (\omega^{\alpha_1} * \omega^k) * a_0 + \omega^{\alpha_1} * a_1 & \text{Lem.20 in Fig.12} \\ & \doteq \omega^{\alpha_1} * (\omega^k * a_0) + \omega^{\alpha_1} * a_1 & \text{Assoc. of } * \\ & \doteq \omega^{\alpha_1} * (\omega^k * a_0 + a_1) & \text{Lem15 in Fig.10} \\ & < \omega^{\alpha_1} * (\omega^k * a_0 + \omega^k * 1) & \text{see below} \\ & \doteq \omega^{\alpha_1} * (\omega^k * (a_0 + 1)) & \text{Lem15 in Fig.10} \\ & \doteq (\omega^{\alpha_1} * \omega^k) * (a_0 + 1) & \text{Assoc. of } * \\ & \doteq \omega^{\alpha_0} * (a_0 + 1) & \text{choice of } k \end{array}$$

To fill in the gap we first observe $a_1 < \omega \leq \omega^k = \omega^k * 1$. From this $\omega^k * a_0 + a_1 < \omega^k * a_0 + \omega^k * 1$ follows via Lemma 4 in Figure 8. The last step uses Lemma 3 from Figure 10. □

**Lemma 3 (Uniqueness Lemma).**
*For any two normal forms $t_1$, $t_2$*

1. *$t_1 \gg t_2$ if and only if $Th_{Ord} \vdash t_1 > t_2$*
2. *$t_1 = t_2$ if and only if $Th_{Ord} \vdash t_1 \doteq t_2$*

*Proof.* **ad(1)** We first prove the implication from left to right. So we start from the assumption $t_1 \gg t_2$. The proof proceeds by induction on $max(rk(t_0), rk(t_1))$. If $rk(t_0) = rk(t_1) = 0$ then $t_0, t_1 \in \mathbb{N}$ and the claim is obvious.

In the inductive step let $t_0$ be the greatest common initial sum of $t_1$, $t_2$. Thus:
$t_1 = t_0 + \omega^{\alpha_i} * a_i + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1}$
$t_2 = t_0 + \omega^{\beta_j} * b_j + \ldots + \omega^{\beta_{s-2}} * b_{s-2} + b_{s-1}$
and $\alpha_i \gg \beta_j$ or $\alpha_i = \beta_j$ and $a_i > b_j$. In the first case the induction hypothesis yields $Th_{Ord} \vdash \alpha_i > \beta_j$. The following formulas are then derivable in $Th_{Ord}$:

$$\begin{array}{rlr} \omega^{\beta_j} * b_j + \ldots + \omega^{\beta_{s-2}} * b_{s-2} + b_{s-1} & < \omega^{\beta_j} * (b_j + 1) & \text{Lemma 2} \\ & < \omega^{\beta_j} * \omega & \text{since } (b_j + 1) < \omega \\ & = \omega^{(\beta_j + 1)} & \text{def. of exp} \\ & \leq \omega^{\alpha_i} & \text{since } \beta_j < \alpha_i \end{array}$$

Since $Th_{Ord} \vdash \forall x, y(y \neq 0 \to x \leq x * y)$ and $Th_{Ord} \vdash \forall x, y(x \leq x + y)$ we arrive at $Th_{Ord} \vdash t_1 > t_2$.
If $\alpha_i = \beta_j$ and $a_i > b_j$ we obtain derivability of

$$\begin{array}{rlr} \omega^{\beta_j} * b_j + \ldots + \omega^{\beta_{s-2}} * b_{s-2} + b_{s-1} & < \omega^{\beta_j} * (b_j + 1) & \text{Lemma 2} \\ & \leq \omega^{\alpha_i} * a_i & \text{case distinction} \end{array}$$

and continue from here as above. This finishes the proof of the implication $\Rightarrow$.

For the reverse implication assume $Th_{Ord} \vdash t_1 > t_2$. By Lemma 1 we have $t_1 = t_2$ or $t_2 \gg t_1$ or $t_1 \gg t_2$. The first two cases yield by rules of equality logic or by the implication just established $Th_{Ord} \vdash t_1 \doteq t_2$ respectively $Th_{Ord} \vdash t_1 > t_2$. Both contradict the strict order property of $>$. Thus we must have $t_1 \gg t_2$, as desired.

**ad(2)**
The implication $\Rightarrow$ is a trivial consequence of equational logic.
To prove the reverse implication assume $Th_{Ord} \vdash t_1 \doteq t_2$. By Lemma 1 $t_2 \gg t_1$ or $t_1 \gg t_2$ or $t_1 = t_2$. The first two choices contradict the strict order property of $>$ using part (1) of this lemma. □

We note, that Lemma 3 implies for any two normal forms $t_1$, $t_2$ the strong property $Th_{Ord} \vdash t_1 \doteq t_2$ or $Th_{Ord} \vdash t_1 > t_2$ or $Th_{Ord} \vdash t_1 < t_2$.

**Lemma 4 (Adding Normal Forms).**
*Let*
$t_1 = \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1}$
$t_2 = \omega^{\beta_0} * b_0 + \omega^{\beta_1} * b_1 + \ldots + \omega^{\beta_{s-2}} * b_{s-2} + b_{s-1}$
*be normal forms. Then*

$$t_1 + t_2 \doteq \omega^{\gamma_0} * c_0 + \omega^{\gamma_1} * c_1 + \ldots + \omega^{\gamma_{q-2}} * c_{q-2} + c_{q-1}$$

*such that*

1. *for every $0 \leq i < r - 1$ such that there is no $0 \leq j < s - 1$ with $\alpha_i = \beta_j$ there is $0 \leq k < q - 1$ with $\gamma_k = \alpha_i$ and $c_k = a_i$, and*

2. *for every $0 \leq j < s-1$ such that there is no $0 \leq i < r-1$ with $\alpha_j = \beta_i$ there is $0 \leq k < q-1$ with $\gamma_k = \beta_j$ and $c_k = b_j$, and*
3. *for every $0 \leq i < r-1$ and $0 \leq j < s-1$ with $\alpha_i = \beta_j$ there is $0 \leq k < q-1$ with $\gamma_k = \alpha_i = \beta_j$ and $c_k = a_i + b_j$,*
4. *$c_{q-1} = a_{r-1} + b_{s-1}$.*

*and vice versa for every $0 \leq k < q-1$*

1. *either there is $0 \leq i < r-1$ with $\gamma_k = \alpha_i$ and $c_k = a_i$ and there is no $0 \leq j < s-1$ with $\gamma_k = \beta_j$*
2. *or there is $0 \leq j < s-1$ with $\gamma_k = \beta_j$ and $c_k = b_j$ and there is no $0 \leq i < r-1$ with $\gamma_k = \alpha_i$*
3. *or there are $0 \leq i < r-1$ and $0 \leq j < s-1$ with $\gamma_k = \alpha_i = \beta_j$ and $c_k = a_i + b_j$.*

*Proof.* Easy computation using repeatedly associativity of multiplication and Lemma 8 from Figure 9, Lemma 20 from Figure 12, and Lemma 5 from Figure 13.
We also have to make sure that all $\gamma_i$ are normal forms. But, that is obvious since $\gamma_i = \alpha_j$ or $\gamma_i = \beta_j$ for some $j$.

**Corollary 1.** *Let*
$t = \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1}$
*be a normal form and $n < \omega$. Then*

$$t * n \doteq \omega^{\alpha_0} * (a_0 * n) + \omega^{\alpha_1} * (a_1 * n) + \ldots + \omega^{\alpha_{r-2}} * (a_{r-2} * n) + a_{r-1} * n$$

*Proof.* Follows by repeated application of the special case $t_1 = t_2$ of Lemma 4. □

**Lemma 5.** *Let*
$t = \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1}$
*be a normal form and $\delta > 0$. Then*

$$t * \omega^{\delta} \doteq \omega^{\alpha_0 + \delta}$$

*Proof.* See also [15, Thm.8.46].
By Lemma 2 and the fact that multiplication and addition on the right are non-decreasing we have

$$\omega^{\alpha_0} \leq t \leq \omega^{\alpha_0} * (m_0 + 1) \tag{5}$$

By Lemma 6 in Figure 10 we obtain

$$\omega^{\alpha_0} * \omega^{\delta} \leq t * \omega^{\delta} \leq (\omega^{\alpha_0} * (m_0 + 1)) * \omega^{\delta} \tag{6}$$

By associativity of multiplication and Lemma 11 in Figure 10 we obtain further

$$\omega^{\alpha_0} * \omega^{\delta} \leq t * \omega^{\delta} \leq \omega^{\alpha_0} * \omega^{\delta} \tag{7}$$

Actually, the strenghtening $\forall x(0 < x < \omega \to x * \omega^\alpha \doteq \omega^\alpha)$ of Lemma 11 is needed here. But, this can be easily established. From equation 7 the claim of the lemma immediately follows. □

**Lemma 6 (Multiplying Normal Forms).**
*Let*
$t_1 = \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1}$
$t_2 = \omega^{\beta_0} * b_0 + \omega^{\beta_1} * b_1 + \ldots + \omega^{\beta_{s-2}} * b_{s-2} + b_{s-1}$
*be normal forms. Then there is a normal form for* $t_1 * t_2$.

*Proof.*

$$\begin{aligned} t_1 * t_2 &= t_1 * \omega^{\beta_0} * b_0 + \ldots + t_1 * \omega^{\beta_{s-2}} * b_{s-2} + t_1 * b_{s-1} && \text{distributivity} \\ &= \omega^{\alpha_0+\beta_0} * b_0 + \ldots + \omega^{\alpha_0+\beta_{s-2}} * b_{s-2} + && \text{Lemma 5} \\ &\quad \omega^{\alpha_0} * (a_0 * b_{s-1}) + \ldots + \omega^{\alpha_{r-2}} * (a_{r-2} * b_{s-1}) + a_{r-1} * b_{s-1} && \text{Cor. 1} \end{aligned}$$

All summands are normal forms, but the constraints on the exponents might not yet bew satisifed. Lemma 4 guarantees that this can be achieved. The same lemma is needed to obtain normal forms for the exponents $\alpha_0 + \beta_j$.

**Theorem 1.** *For every ground term $t$ there is a normal form $t_n$ such that $t \doteq t_n$ is derivable.*

*Proof.* The proof proceeds by induction on the number of non-constant functions symbols $f(t)$ of $t$.

If $f(t) = 0$ then either $t = n$, and it is itself a normal form, or $t = \omega$, and $t \doteq \omega * 1$.

In the induction step we need to distinguish three cases:
$\mathbf{t = t_1 + t_2}$ This is covered by Lemma 4.
$\mathbf{t = t_1 * t_2}$ By the induction hypothesis we may assume that
$t_1 = \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1}$
$t_2 = \omega^{\beta_0} * b_0 + \omega^{\beta_1} * b_1 + \ldots + \omega^{\beta_{s-2}} * b_{s-2} + b_{s-1}$
are normal forms. Then there is a normal from for $t_1 * t_2$ by Lemma 6.

$\mathbf{t = t_1^{t_2}}$ By the induction hypothesis we may assume that
$t_1 = \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1}$
$t_2 = \omega^{\beta_0} * b_0 + \omega^{\beta_1} * b_1 + \ldots + \omega^{\beta_{s-2}} * b_{s-2} + b_{s-1}$
are normal forms. In a first step we rewrite $t_1^{t_2}$ using (20 ) from Figure 12 as:

$$t_1^{t_2} = t_1^{\omega^{\beta_0} * b_0} * \ldots * t_1^{\omega^{\beta_{s-2}} * b_{s-2}} * t_1^{b_{s-1}} \tag{8}$$

If we succeed to show that each summand is equivalent to a normal form then the claim follows from Lemma 6.
The last summand $t_1^{b_{s-1}}$ is equivalent to the $b_{s-1}$-fold multiplication $t_1 * \ldots * t_1$ and this is equivalent to a normal form by Lemma 6.
For the remaining summands we observe that all exponents $\omega^{\beta_i} * b_i$ for $0 \leq i \leq s-2$ are limit ordinals.

From 3 in Figure 10 and 1 in Figure 8 we obtain $\omega^{\alpha_0} \leq \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1}$ From Lemma 2 we get $\omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1} \leq \omega^{\alpha_0} * a_0 * (m_0 + 1)$. Thus

$$\omega^{\alpha_0} \leq \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1} \leq \omega^{\alpha_0} * a_0 * (m_0 + 1)$$

From 12 in Figure 12 and 7 in Figure 13 we get for any limit ordinal $\gamma$

$$\begin{aligned} \omega^{\alpha_0 * \gamma} &\leq (\omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1})^{\gamma} \\ &\leq (\omega^{\alpha_0} * a_0 * (m_0 + 1))^{\gamma} = \omega^{\alpha_0 * \gamma} \end{aligned}$$

Thus in particular for all $i$, $0 \leq i \leq s - 2$

$$t_1^{\omega^{\beta_i * b_i}} = \omega^{\alpha_0 * \omega^{\beta_i * b_i}} \tag{9}$$

Since Lemma 4 guarantees that $\alpha_0 * \omega^{\beta_i * b_i}$ is equivalent to a normal form we have finished the proof. □

**Lemma 7.**

1. *A normal form $t \notin \mathbb{N}$ is a limit ordinal exactly if the last summand is* $0$.
2. *For any limit ordinal $t \in$* HNF *there is a term $t_0$ such that*

$$Th_{Ord} \vdash t \doteq sup_{x<\omega}(t_0)$$

*and $t_0[n/x]$ is a normal form for every $n \leq \omega$*

*Proof.* **ad(1)** Easy.
**ad(2)** The proof proceeds by induction on $rk(t)$.
There are no limit ordinal normal form with rank 0. For the inductive step look at $t_1 = \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * a_{r-2}$.
In the following we will tacitly use Lemma 14 from Figure 9 and Lemma 22 from Figure 12.
Case A: $\alpha_{r-2} = \gamma + 1$
Thus $t \doteq \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\gamma} * \omega * a_{r-2}$
Case A1: $a_{r-2} = 1$ Thus $t \doteq \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\gamma} * \omega$
Now we see $t \doteq sup_{x<\omega}(\omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\gamma} * x)$
Case A2: $a_{r-2} = c + 1$ Thus
$t \doteq \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * (c + 1)$
$\doteq \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * c + \omega^{\gamma} * \omega$
Now we see $t \doteq sup_{x<\omega}(\omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * c + \omega^{\gamma} * x)$
Case B $\alpha_{r-2}$ is a limit ordinal
By induction hypothesis there is a term $\alpha$ such that $\alpha_{r-2} = sup_{x<\omega}\alpha$.
Case B1: $a_{r-2} = 1$ Thus $t \doteq \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}}$
Now we see $t \doteq sup_{x<\omega}(\omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha})$
Case B2: $a_{r-2} = c + 1$ Thus
$t \doteq \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * c + \omega^{\alpha_{r-2}}$
Now we see $t \doteq sup_{x<\omega}(\omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots \omega^{\alpha_{r-2}} * c + \omega^{\alpha})$.
It is easily checked that in all four cases $t_0[n/x]$ is a normal form for any $n \leq \omega$. □

# 4 Semantics

We follow the well established pattern for set theoretic semantics and work in a fixed informal model of set theory. In this setting all ordinals with the usual set theoretic interpretation of the constants, operations, and predicates would be a model of $Th_{Ord}$. The problem is that the collection of all ordinals is a proper class and not a set. One could now investigate how much model theory would change if classes were allowed in place of sets. We follow another line and define below a standard model, called the $\epsilon$ standard model, whose universe is an initial segment of ordinals and thus a set.

**Definition 2.** *We define the following notation:*
$\omega_0 \quad = \omega$
$\omega_{n+1} = \omega^{\omega_n}$
$\epsilon_0 \quad = sup_{n<\omega}(\omega_n)$
*It is convenient to stipulate* $\omega_{-1} = 1$.
*For example* $\omega_3 = \omega^{(\omega^\omega)}$. *We not that all* $\omega_i$ *are in* HNF.

Note that this is a semantic definition. There is no term denoting $\epsilon_0$ in $Th_{Ord}$.

**Lemma 8.**

1. *For* $n < m < \omega$ *we have* $\omega_n < \omega_m$.
2. *For all* $n, m < \omega$
   (a) $\omega_n + \omega_m < \omega_{max\{n,m\}+1}$
   (b) $\omega_n * \omega_m < \omega_{max\{n,m\}+1}$,
   (c) $\omega_n^{\omega_m} < \omega_{max\{n,m\}+2}$

*Proof.*
**ad(1)**By induction on $n$
For $n = 0 < m$ the inequality $\omega_0 = \omega < \omega^{\omega_{m-1}}$ follows from (4) in Figure 12.
In the induction step we know $\forall m(n < m \rightarrow \omega_n < \omega_m)$ and we want to prove $\forall m(n+1 < m \rightarrow \omega_{n+1} < \omega_m)$. We start with the definition $\omega_{n+1} = \omega^{\omega_n}$. From $n+1 < m$ we derive $n < m-1$ and by induction hypothesis $\omega_n < \omega_{m-1}$. Now (10) from Figure 12 implies $\omega^{\omega_n} < \omega^{\omega_{m-1}}$ which is by definition $\omega_{n+1} < \omega_m$.
**ad(2.a)**

$$\begin{array}{rll}
\omega_n + \omega_m & \leq \omega_{max\{n,m\}} + \omega_{max\{n,m\}} & \text{by (4), (5) in Fig. 8} \\
& = \omega_{max\{n,m\}} * 2 & \text{def. of } * \\
& < \omega_{max\{n,m\}} * \omega & \text{by (3) in Fig. 10} \\
& = \omega^{\omega_{max\{n,m\}-1}} * \omega & \text{def. of } \omega_{max\{n,m\}} \\
& = \omega^{\omega_{max\{n,m\}-1}+1} & \text{def. of exp.} \\
& < \omega^{\omega_{max\{n,m\}}} & \text{since } \omega_k + 1 < \omega_{k+1} \text{ and (10) in Fig. 12} \\
& = \omega_{max\{n,m\}+1} & \text{def. of } \omega_{max\{n,m\}+1}
\end{array}$$

It is easy to see that $\omega_k + 1 < \omega_{k+1}$ is true for all $k$, since by (18) in Fig. 12 all $\omega_k$ are limit ordinals and from part one $\omega_k < \omega_{k+1}$.
**ad(2.b)**

$$\begin{array}{lr}
\omega_n * \omega_m \leq \omega_{max\{n,m\}} * \omega_{max\{n,m\}} & \text{by (3) and (6 ) in Fig. 10} \\
\quad = \omega^2_{max\{n,m\}} & \text{def of exp.} \\
\quad = (\omega^{\omega_{max\{n,m\}-1}})^2 & \text{def of } \omega_{max\{n,m\}} \\
\quad = \omega^{\omega_{max\{n,m\}-1}*2} & \text{by (21) in Fig. 12} \\
\quad = \omega^{\omega_{max\{n,m\}-1}+\omega_{max\{n,m\}-1}} & \text{def. of } * \\
\quad < \omega^{\omega_{max\{n,m\}}} & \text{claim 2.a and (10) in Fig. 12} \\
\quad = \omega_{max\{n,m\}+1} & \text{def. of } \omega_{max\{n,m\}+1}
\end{array}$$

**ad(2.c)**

$$\begin{array}{lr}
\omega_n^{\omega_m} = (\omega^{\omega_{n-1}})^{\omega_m} & \text{def. of } \omega_n \\
\quad = \omega^{\omega_{n-1}*\omega_m} & \text{by (21) in Fig. 12} \\
\quad < \omega^{\omega_{max\{n-1,m\}+1}} & \text{claim 2.b and (10) in Fig. 12} \\
\quad = \omega_{max\{n-1,m\}+2} & \text{def. of } \omega_{max\{n-1,m\}+2} \\
\quad \leq \omega_{max\{n,m\}+2} & \text{part 1 of this lemma}
\end{array}$$

**Lemma 9.** *For a term $t$ in the language of $Th_{Ord}$ with the free variables $x_1, \ldots, x_n$ we denote by $f_t$ the $n$-place function that associates argument tuples $\alpha_1, \ldots, \alpha_n$ with the value $f_t(\alpha_1, \ldots, \alpha_n)$ that is obtained by evaluating term $t$ under the variable assignment $x_i \rightsquigarrow \alpha_i$.*
*For every term $t$ there is natural number $b_t < \omega$ such that*

$$\alpha_i < \omega_{m_i} \text{ for } 1 \leq i \leq n \text{ implies } f_t(\alpha_1, \ldots, \alpha_n) < \omega_{k+b_t}$$
$$\text{with } k = max\{m_i \mid 1 \leq i \leq n\}$$

*Proof.* The proof proceeds by structural induction on $t$. The claim is trivial if $t$ is just a variable or if $t$ is a constant.
If $t = t_1 + t_2$ there are by induction hypthesis bounds $b_{t_1}$, $b_{t_2}$ such that for all tuples $\alpha_1, \ldots, \alpha_n$ with $\alpha_i < \omega_{m_i}$ for all $1 \leq i \leq n$ and $k = max\{m_i \mid 1 \leq i \leq n\}$ we have $f_{t_1}(\alpha_1, \ldots, \alpha_n) < \omega_{k+b_{t_1}}$ and $f_{t_2}(\alpha_1, \ldots, \alpha_n) < \omega_{k+b_{t_2}}$. By (2.a) of Lemma 8 we get $f_t(\alpha_1, \ldots, \alpha_n) < \omega_{k+b+1}$ with $b = max\{b_{t_1}, b_{t_2}\}$.
The cases $t = t_1 * t_2$ and $t = t_1^{t_2}$ are handled analogously.
It remains to consider $t = sup_{x_0 < t_1}(t_2)$.
By induction hypothesis there are bounds $b_{t_1}$, $b_{t_2}$ such that for all arguments $\alpha_0, \alpha_1, \ldots, \alpha_n$ with $\alpha_i < \omega_{m_i}$ we know $f_{t_1}(\alpha_1, \ldots, \alpha_n) < \omega_{max\{m_1,\ldots,m_n\}+b_{t_1}}$ and $f_{t_2}(\alpha_0, \alpha_1, \ldots, \alpha_n) < \omega_{max\{m_0,m_1,\ldots,m_n\}+b_{t_2}}$. Observe, that the variable $x_0$, to which $\alpha_0$ is assigned, is not allowed to occur in $t_1$. For fixed $\alpha_1, \ldots, \alpha_n$ with $\alpha_i < \omega_{m_i}$ $\omega_{max\{m_1,\ldots,m_n\}+b_{t_1}}$ is an upper bound for the assignments to $x_0$. We thus get for all instantiations $\alpha_0$ for $x_0$ that $f_{t_2}(\alpha_0, \alpha_1, \ldots, \alpha_n) < \omega_{max\{m_1,\ldots,m_n\}+b_{t_1}+b_{t_2}}$. This is an uppper bound also for the supremum, i.e. $f_t(\alpha_1, \ldots, \alpha_n) < \omega_{max\{m_1,\ldots,m_n\}+b_{t_1}+b_{t_2}}$. □

**Definition 3 ($\epsilon$-Standard Modell).**
*The $\epsilon$ standard model $\mathcal{S} = (U, <, 0, \omega, +, *, exp, sup)$ has as universe $U$ the set of all ordinals strictly less than $\epsilon_0$: $U = \{\alpha \mid \alpha < \epsilon_0\}$. The two constants, the ordering, ordinal addition, multiplication, exponentiation and the supremum operator sup are determined by the usual set theoretic definitions.*
*Lemma 8 guarantees that the results of the arithmetic operations never exceed $\epsilon_0$*

*and Lemma 9 guarantees that also the result of the sup operator stays below $\epsilon_0$ for any choice of the terms $t_1$, $t_2$.*

**Theorem 2.** *$\mathcal{S}$ is a model for $Th_{Ord}$.*

*Proof.* Obvious.

Note, that as a conseqeunce of Theorem 2 the existence of $\epsilon_0$ cannot be proved in $Th_{Ord}$.

**Lemma 10.** *The structure* $(\mathrm{HNF}, \gg)$ *is a wellordering.*

*Proof.* Assume for the sake of a contradiction that there is an infinite decreasing chain $t_1 \gg t_2 \gg \ldots \gg t_n \gg \ldots$ of constant terms $t_n \in \mathrm{HNF}$.
By Lemma 3 we have $Th_{Ord} \vdash t_n > t_{n+1}$ for every $n \in \mathbb{N}$. Thus by Theorem 2 also for $n \in \mathbb{N}$ that $t_n^{\mathcal{S}} > t_{n+1}^{\mathcal{S}}$. This contradicts the well-foundedness of the $\epsilon$ standard model $\mathcal{S}$. □

Since $Th_{Ord}$ is a first-order theory there will be $\omega$-nonstandard models, i.e. models $\mathcal{M}$ such that there are elements $o$ with $o <^{\mathcal{M}} \omega^{\mathcal{M}}$ and $n^{\mathcal{M}} <^{\mathcal{M}} o$ for every finite ordinal $n$.

**Definition 4** **($\omega$-Standard Modell).**
*A model $\mathcal{M}$ of $Th_{Ord}$ is called a $\omega$-standard model if for every $o <^{\mathcal{M}} \omega^{\mathcal{M}}$ there is a finite ordinal $n$ such that $o <^{\mathcal{M}} n^{\mathcal{M}}$.*

**Theorem 3.** *Let $\mathcal{M}$ be an $\omega$-standard model of $Th_{Ord}$. Then $\mathcal{M}$ contains an initial segment that is isomorphic to the $\epsilon$-standard moodel.*

*Proof.* We start out by proving

$$\begin{array}{l}\text{For every } t_0 \in \mathrm{HNF} \text{ and every } o \in M \text{ with } o <^{\mathcal{M}} t_0^{\mathcal{M}} \\ \text{there is } t \in \mathrm{HNF} \text{ such that } o = t^{\mathcal{M}}.\end{array} \tag{10}$$

The proof of this claim proceeds by induction on the wellfounded ordering $(\mathrm{HNF}, \gg)$ (see Lemma 10). The initial case and the successor step are simple. So let us assume that claim (10) is true for all $t_1 \in \mathrm{HNF}$ with $t_0 \gg t_1$ and aim to show that it is true for $t_0$.

Using Lemma 7(2)) we find a term $t'_0$ such that $t_0^{\mathcal{M}} = sup_{x<\omega}^{\mathcal{M}}(t'_0)$. By the defining axioms of the supremum operator there is $o' <^{\mathcal{M}} \omega^{\mathcal{M}}$ with $o <^{\mathcal{M}} (t')_0^{\mathcal{M}}[o'/x]$. Since $\mathcal{M}$ was assumed to be a $\omega$ standard model we must have $o' = n^{\mathcal{M}}$ for some constant $n$. Thus $o <^{\mathcal{M}} t_1^{\mathcal{M}}$ for $t_1 = t'_0[n/x]$. Using Lemmata 1 and 3 we must have $t_0 \gg t_1$ since the only other options, $t_1 \gg t_2$, $t_0 = t_1$, would contradict $t_0^{\mathcal{M}} = sup_{x<\omega}^{\mathcal{M}}(t'_0)$. By our inductive assumption there must be a constant term $t \in \mathrm{HNF}$ that names $o$ in $\mathcal{M}$.

Thus the set $E = \{o \in M \mid o <^{\mathcal{M}} t^{\mathcal{M}} \text{ for some } t \in \mathrm{HNF}\}$ where $M$ is - as usual - the universe of $\mathcal{M}$ equals $\{t^{\mathcal{M}} \mid t \text{ a ground term}\}$.

We close this section with some observations in the informal model of set theory-

**Lemma 11.** *1.* $\omega^{\epsilon_0} = \epsilon_0$
*2.* $\delta < \omega^{\delta}$ *for all* $0 < \delta < \epsilon_0$

*Proof.* **add 1**

$$\begin{aligned}
\omega^{\epsilon_0} &= \omega^{sup_{n<\omega}\omega_n} && \text{Definition 2}\\
&= sup_{n<\omega}\omega^{\omega_n} && \text{Lemma 22 in Figure 12}\\
&= sup_{n<\omega}\omega_{n+1} && \text{Definition 2}\\
&= sup_{n<\omega}\omega_n && \text{Lemma 16 in Figure 6}\\
&= \epsilon_0 && \text{Definition 2}
\end{aligned}$$

**add 2** We procced by induction on $\delta$. We assume that $\delta' < \omega^{\delta'}$ for all $0 < \delta' < \delta$. By (3) in the proof of Theorem 3 we know that $\delta$ can be represented by a term in HNF

$$\delta = \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1}$$

Thus

$$\begin{aligned}
\omega^{\delta} &= \omega^{\omega^{\alpha_0}*a_0+\omega^{\alpha_1}*a_1+\ldots+\omega^{\alpha_{r-2}}*a_{r-2}+a_{r-1}} && \text{previous equation}\\
&= \omega^{\omega^{\alpha_0}*a_0} * \omega^{\omega^{\alpha_1}*a_1} * \ldots * \omega^{\omega^{\alpha_{r-2}}*a_{r-2}} * \omega^{a_{r-1}} && \text{Lemma 20 in Figure 12}\\
&= \omega^{\omega^{\alpha_0}*a_0} + \omega^{\omega^{\alpha_1}*a_1} + \ldots + \omega^{\omega^{\alpha_{r-2}}*a_{r-2}} + \omega^{a_{r-1}} && \text{Lemma 19 in Figure 10}\\
&> \omega^{\alpha_0} * a_0 + \omega^{\alpha_1} * a_1 + \ldots + \omega^{\alpha_{r-2}} * a_{r-2} + a_{r-1} && \text{induction hypothesis}\\
&= \delta
\end{aligned}$$

We add that because of Lemmas 1 and 3 in Figure 8 we know $\omega^{\alpha_i} * a_i < \delta$ for all $0 \leq i < r$. This guarantees that the induction hypothesis is applicable in the above reasoning. □

## 5 Application Scenario

We will consider applications in the area of program verification. Ordinals will never occur in the programs but only in their specifications. We thus need a way to link program data to ordinals. Figure 14 shows an axiomatisation of the function $onat : Int \rightarrow Ord$ that maps the positive natural numbers into corresponding ordinals less than $\omega$. For negative arguments *onat* is undefined via underspecification. Figure 14 also show usful derived lemmas. We use in this figure and also later on overloaded syntax. Thus, whether 0 denotes an interger or an ordinal, wether + is ordinal addition or addition of natural numbers can be found out by looking at the type information.

As a possible application area we propose program termination proofs. This was already suggested by Alan Turing in [2] (see also the corrected and commented account [5]). As a simple example we want to prove termination of the program in Figure 15. As in all pratical examples that we know of program termination can be proved within Peano Arithmetic the advantage of using transfinite ordinals is seen in simplifying specification and verification.

Definitional Extension

1. $onat(0) \doteq 0$
2. $\forall n(0 \leq n \rightarrow onat(n+1) \doteq onat(n)+1)$

Derived Lemmas

3. $onat(1) \doteq 1$
4. $\forall n, m(0 \leq n \wedge 0 \leq m \rightarrow onat(n+m) \doteq onat(n) + onat(m)$
5. $\forall n, m((0 \leq n \wedge 0 \leq m) \rightarrow (onat(n) \doteq onat(m) \rightarrow n \doteq m))$
6. $\forall n, m((0 \leq n \wedge 0 \leq m) \rightarrow (onat(n) < onat(m) \leftrightarrow n < m))$
7. $\forall n(0 \leq n \rightarrow onat(n) < \omega)$
8. $\forall i_1, i_2, j_1, j_2 \; ((0 \leq i_1 \wedge 0 \leq i_2 \wedge 0 \leq j_1 \wedge 0 \leq j_2) \rightarrow$
   $\omega * onat(i_1) + onat(j_1) < \omega * onat(i_2) + onat(j_2)$
   $\leftrightarrow i_1 < i_2 \vee (i_1 \doteq i_2 \wedge j_1 < j_2))$

**Fig. 14.** Positive integers as ordinals

Figure 15 shows a possible pattern. In a while loop two positive integer variables $x$ and $y$ are reassigned. Either $y$ is decremented by 1, if not already 0, and $x$ is left alone or $x$ is decremented by 1, if it is nor already 0, and $y$ is assigned an arbitrary natural number. The point is that an upper bound on the new $y$ may not be known or too complicated to estimate. In the while loop in lines 7 to 14 both actions happen. It can easily be proved that $\omega * onat(x) + onat(y)$ is variant for this loop. Here *onat* is a function that injects Java integers into ordinals. In line 8/9 of course the JML syntax for the variant has to be used.

One should, however, not overestimate the usefulness of ordinals in termination proofs. For the program shown in Figure 15 the KeY system would allow to specify $(x, y)$ in the decreases clause, which would be interpreted as the lexicographical ordering of pairs of positive integers. The proof completes automatically.

## 6 Goodstein Sequences

Goodstein sequences were first introduce in the paper [14]. This section is mainly based on the paper [8].

The hereditary base-$n$ notation for a natural number $m$ is obtained from its ordinary base-$n$ notation

$$m = m_k \cdot n^k + m_{k-1} \cdot n^{k-1} + \ldots m_1 \cdot n + m_0, \quad 0 \leq m_i < n, m_k \neq 0$$

by also writing the exponents $k$, $k-1$, …, $n+1$ in base-$n$ notation and again the thus arising exponents, and so on.

The following formal recursive definition is taken from [8].

**Definition 5.** *For $n, m \in \mathbb{N}$ with $n > 1$ and a new constant $x$ define by recursion on $m$ a term $f^{m,n}(x)$. To this end the $n$- addic expansion of $m$ is computed:*

$$m = n^k \cdot a_k + n^{k-1} \cdot a_{k-1} + \ldots n \cdot a_1 + a_0$$

```
class Aclass {
  /*@ normal_behaviour
    @   requires 0 <= x &&  0 <= y;
    @*/
  void  method(int x, int y) {
    /*@ loop_invariant
      @  x>=0 && y>=0;
      @  decreases  \ord_add(\ord_times(\omega,\onat(x)),
                             \onat(y))
      @*/
     while (x>0 || y>0)
            { if (x>0) {x = x-1 ; y = g(y);}
              if (y>0) {y=y-1;}
            }
    }
    /*@ normal_behaviour
      @    ensures \result > 0;
      @*/
    int /*strictly_pure*/ g(int p){
}
```

**Fig. 15.** An Example Programm

*Thus we have* $a_i < n$ *for all* $0 \leq i \leq k$ *and* $a_k > 0$. *Now set*

$$\begin{array}{ll} f^{m,n}(x) & = \Sigma_{i=0}^{k} a_i x^{f^{i,n}(x)} \\ f^{0,n}(x) & = 0 \end{array}$$

*From the term* $f^{m,n}(x)$ *the hereditary base-*$n$ *expansion of* $m$ *is obtained by replacing* $x$ *by* $n$.

$f^{m,n}(x)$ is far from being additive in the first argument, i.e., in general $f^{m_1+m_2,n}(x) \neq f^{m_1,n}(x) + f^{m_2,n}(x)$. But, in the following very special case additivity pervails.

**Lemma 12.** *Let* $m_1 = n^k \cdot a_k + n^{k-1} \cdot a_{k-1} \ldots n^r \cdot a_r$ *and* $n^r > m_2$ *then*

$$f^{m_1+m_2,n}(x) = f^{m_1,n}(x) + f^{m_2,n}(x)$$

*Proof.* Ovious. □

**Lemma 13.** *In this lemma* $f^{m,n}(n)$ *is not the term but the value obtain by evaluating it.*

1. $f^{i,n}(x) = i$ *for all* $i$ *with* $0 \leq i < n$
2. $f^{m,n}(n) = m$

*Proof* To prove claim (1) we start with the $n$-addic expansion of $i$ which under the present assumption is $i = i \cdot n^0$. Thus $f^{i,n}(x) = i \cdot x^{f^{0,n}} = i \cdot x^0 = i \cdot 1 = i$.

Claim (2) is proved by induction on $m$. The case $m = 0$ follows directly form the definition. We assume the claim for $m$ and set out to prove it for $m + 1$. Let

$$m = a_k \cdot n^k + a_{k-1} \cdot n^{k-1} + \ldots + a_1 \cdot n + a_0$$

the $n$-addic expansion of $m$. From this we derive the $n$-addic expansion of $m + 1$

$$m + 1 = a_k \cdot n^k + \ldots + (a_j + 1) \cdot n^j$$

where $0 \leq j \leq k + 1$ with $a_j < n - 1$ and $a_r = n - 1$ for all $0 \leq r < j$. Here we also stipulate $a_{k+1} = 0$
In the extreme cases we get

$$m + 1 = a_k \cdot n^k + a_{k-1} \cdot n^{k-1} + \ldots + a_1 \cdot n + (a_0 + 1)$$

in case $j = 0$ and

$$m + 1 = n^{k+1}$$

in case $j = k + 1$. In this last case we obtain by definition $f^{m+1,n}(n) = n^{f^{k+1,n}(n)}$. Since $k + 1 \leq m$ we obtain from the induction hypothesis $f^{m+1,n}(n) = n^{k+1}$. Since the assumptions for this case are $a_r = n - 1$ for all $0 \leq r \leq k$ we have $n^{k+1} = (a_j \cdot n^k + a_{k-1} \cdot n^{k-1} + \ldots + a_1 \cdot n + a_0) + 1$ Thus $f^{m+1,n}(n) = m + 1$.

For the cases $j \leq k$ we obtain

$$\begin{aligned} f^{m+1,n}(n) &= a_k \cdot n^{f^{k,n}(n)} + \ldots + (a_j + 1) \cdot n^{fj,n(n)} && \text{definition of} f \\ &= a_k \cdot n^k + \ldots + (a_j + 1) \cdot n^j && \text{induction hypothesis} \\ &= a_k \cdot n^k + a_{k-1} \cdot n^{k-1} + \ldots + a_1 \cdot n + (a_0 + 1) && \text{case assumptions on } j \\ &= m + 1 \end{aligned}$$

□

| *Example 1.* | | |
|---|---|---|
| base-2 | 35 | $= 2^5 + 2^1 + 2^0$ |
| hereditary base-2 | 35 | $= 2^{2^2+1} + 2 + 1$ |
| base-3 | 100 | $= 3^4 + 2 \cdot 3^2 + 3^0$ |
| hereditary base-3 | 100 | $= 3^{3+1} + 2 \cdot 3^2 + 1.$ |

**Definition 6 ( Next numner function).**
*The function $G_n(m)$ is defined by*

$$\begin{aligned} G_n(0) &= 0 \\ G_n(m) &= f^{m,n}(n + 1) - 1 \end{aligned}$$

**Definition 7 (Goodstein sequence for $m$).**

$$\begin{aligned} m_0 &= m \\ m_{i+1} &= G_{i+2}(m_i) \end{aligned}$$

*Thus*

$$m_0 = m \quad m_1 = G_2(m_0) \quad m_2 = G_3(m_1) \quad m_3 = G_4(m_2) \quad \ldots$$

Another way to notate the Goodstein sequence for $m$ is:

$$m, G_2(m), G_3(G_2(m)), \ldots, G_{i+1}(G_i(\ldots G_2(m)\ldots))\ldots$$

Using the $f$ notation that leads to a slanting tower of exponents is less suitable as e.g.,

$$m_2 = f^{f^{m,2}(3)-1,3}(4) - 1$$

shows.

*Example 2.* The Goodstein sequence for $m = 3$

| | | | |
|---|---|---|---|
| $m_0$ | by definition | | 3 |
| $m_1$ | write 3 in her. base 2 notation | $2^1 + 1$ | |
| | replace 2 by 3 minus 1 | $3^1 + 1 - 1$ | 3 |
| $m_2$ | write 3 in her. base 3 notation | $3^1$ | |
| | replace 3 by 4 minus 1 | $4^1 - 1$ | 3 |
| $m_3$ | write 3 in her. base 4 notation | 3 | |
| | replace 4 by 5 minus 1 | $3 - 1$ | 2 |
| $m_4$ | write 2 in her. base 5 notation | 2 | |
| | replace 5 by 6 minus 1 | $2 - 1$ | 1 |
| $m_5$ | write 1 in her. base 6 notation | 1 | |
| | replace 6 by 7 minus 1 | $1 - 1$ | 0 |

*Example 3.* Initial part of the Goodstein sequence for $m = 4$

| | | | |
|---|---|---|---|
| | | 4 | |
| $2^{2^1}$ | $3^{3^1} - 1$ | 26 | $\omega^\omega$ |
| $3^2 * 2 + 3^1 * 2 + 2$ | $4^2 * 2 + 4^1 * 2 + 2 - 1$ | 41 | $\omega^2 * 2 + \omega * 2 + 2$ |
| $4^2 * 2 + 4^1 * 2 + 1$ | $5^2 * 2 + 5^1 * 2 + 1 - 1$ | 60 | $\omega^2 * 2 + \omega * 2 + 1$ |
| $5^2 * 2 + 5^1 * 2$ | $6^2 * 2 + 6^1 * 2 - 1$ | 83 | $\omega^2 * 2 + \omega * 2$ |
| $6^2 * 2 + 6^1 * 1 + 5$ | $7^2 * 2 + 7^1 * 1 + 5 - 1$ | 109 | $\omega^2 * 2 + \omega + 5$ |
| $7^2 * 2 + 7^1 * 1 + 4$ | $8^2 * 2 + 8^1 + 1 + 4 - 1$ | 139 | $\omega^2 * 2 + \omega + 4$ |
| $8^2 * 2 + 8^1 * 1 + 3$ | $9^2 * 2 + 9^1 * 1 + 3 - 1$ | 173 | $\omega^2 * 2 + \omega + 3$ |
| $9^2 * 2 + 9^1 * 1 + 2$ | $10^2 * 2 + 10^1 * 1 + 2 - 1$ | 211 | $\omega^2 * 2 + \omega + 2$ |
| $10^2 * 2 + 10^1 * 1 + 1$ | $11^2 * 2 + 11^1 * 1 + 1 - 1$ | 253 | $\omega^2 * 2 + \omega + 1$ |
| $11^2 * 2 + 11^1 * 1$ | $12^2 * 2 + 12^1 * 1 - 1$ | 299 | $\omega^2 * 2 + \omega$ |
| $12^2 * 2 + 11$ | $13^2 * 2 + 10$ | 348 | $\omega^2 * 2 + 11$ |
| | | | |
| | | 1058 | |
| $23^2 * 2$ | $24^2 * 2 - 1$ | 1151 | $\omega^2 * 2$ |
| $24^2 + 24 * 23 + 23$ | $25^2 + 25 * 23 + 23 - 1$ | 1222 | $\omega^2 + \omega * 23 + 23$ |

Figure 16 shows a runnable Java program that computes the Goodstein sequence for $m$, where you enter $m$ as a command line parameter. Note, that BigIntergers have to be used, since numbers get large very quickly. For $m \geq 4$

the program would certainly stop, but not during your lifetime. The Goodstein sequence for $m = 4$ e.g., terminates after $10^{121210700}$ steps in order of magnitude. The program ist thus incremental and prompts the user to specify how many more numbers in the sequence should be computed next. Entering 0 will terminate the program.

Figure 17 shows a Java program for Goodstein sequences without big integers and without input output. Since in the default setting the KeY system treats Java integers as mathematical integers this is the program we need to verify. Since exponention is not part of the Java language the method `intPow` from Figure 18 is needed.

```
import java.io.*;
import java.math.BigInteger;

public class Goodstein{
  static BigInteger m;

public static BigInteger changeBase(BigInteger m, BigInteger
oldBase){
 BigInteger acc = BigInteger.ZERO;
 BigInteger exp = BigInteger.ZERO;
  while (m.compareTo(BigInteger.ZERO)>0) {
  BigInteger[]  divArray = m.divideAndRemainder(oldBase);
  acc = acc.add((divArray[1]).multiply(intPow(oldBase.add(BigInteger.ONE),
           changeBase(exp,oldBase))));
  m = divArray[0];
  exp = exp.add(BigInteger.ONE);
   }
    return acc;}

  public static BigInteger intPow(BigInteger base, BigInteger exp){
    BigInteger r = BigInteger.ONE;
    while (!exp.equals(BigInteger.ZERO)) {
      r = r.multiply(base);
      exp = exp.subtract(BigInteger.ONE);
  }
    return r;}

 public static void main(String[] args){
  int answer;
  int bunch = 0;
   m = new BigInteger(args[0]);
   System.out.println("Goodstein␣sequence␣for␣" + m);
   System.out.println("");
   BigInteger  base = new BigInteger("2");
   while (m.compareTo(BigInteger.ZERO)!=0) {
     if (bunch == 0) {
     System.out.println("Continue␣n␣steps");
     answer  = LiesInt();
     if (answer == 0) {break;}
    else { bunch = answer;}}
  System.out.print("Next␣Goodstein␣number␣" + m + "␣␣␣");
  System.out.println("Base␣" + base);
  m = changeBase(m,base).subtract(BigInteger.ONE);
  base = base.add(BigInteger.ONE);
  bunch = bunch -1;
   } }
 static int LiesInt() {
 DataInput StdEingabe = new DataInputStream(System.in);
 int ergebnis = 0;
  try{ ergebnis = Integer.parseInt(StdEingabe.readLine()); }
  catch (IOException io) {}
  return ergebnis;}
}
```

**Fig. 16.** Runnable program for Goodstein sequences

```
public class Goodstein{
    static int start;

    /*@ requires m>0;
      @ diverges false;
      @*/

    public static void goodstein(int m){
        int base = 2;
        while (m>0) {
        m = changeBase(m,base)-1;
        base++;
        }}

    public static int changeBase(int m, int oldBase){
        int acc = 0;
        int exp = 0;
        while (m>0) {
            acc = acc + (m%oldBase)*intPow(oldBase+1,changeBase(exp,oldBase));
            m = m/oldBase;
            exp++;
        }
        return acc;
    }}
```

**Fig. 17.** Goodstein program for verification

```
  public static int intPow(int base, int exp){
        int r = 1;
        while (exp != 0) {
            r = r*base;
        exp--;}
        return r;}
```

**Fig. 18.** Auxiliary method for the Goodstein progrm

## 7 Termination

Goodstein considered in his paper [14] more general sequences involving a non-decreasing function $f : \mathbb{N} \to \mathbb{N}$ as a parameter. The Goodstein sequences considered here and in the Kirby and Paris paper [8] are obtained by the choice $f(i) = i + 2$. A short proof of the termination of general Goodstein sequences is given in [13, Theorem 2.5] while [8] presents a much more involved argument. For the convience of the reader we review the special case here giving full proofs,

The most important function that in the end *measures* the steps to termination is given by the next definition making use of the term $f^{m,n}$ as defined in Definition 5:

**Definition 8.** *For $n > 1$, $m \geq 0$ the function $o(n,m) : \mathbb{N} \times \mathbb{N} \to Ord$ is defined by*

$$o(n,m) = f^{m,n}(\omega)$$

*Example 4.* $12 = 2^3 \cdot 1 + 2^2 \cdot 1$ is the 2-addic expansion of 12. Thus $o(2,12) = \omega^{o(2,3)} \cdot 1 + \omega^{o(2,2)} \cdot 1$. Since $o(2,3) = \omega + 1$ and $o(2,2) = \omega$ we get as final result

$$o(2,12) = \omega^{\omega+1} + \omega^{\omega}$$

The next lemma is Lemma 2.3(iii) in [13].

**Lemma 14.** *For $m_1 > m_2$ and $n > 1$*

$$f^{m_1,n}(\omega) > f^{m_2,n}(\omega)$$

*Thus also $o(n,m_1) > o(n,m_2)$*

*Proof.* We prove by induction on $b$:

$$f^{i,n}(\omega) > f^{j,n}(\omega) \text{ for all } i,j \text{ with } b \geq i > j \text{ and all } n > 1$$

In the initial case $b = i = 1$, $j = 0$ we have $f^{1,n}(\omega) = 1 > 0 = f^{0,n}(\omega)$.

In the inductive step from $b$ to $b + 1$ is suffices by transitivity of the ordinal ordering $>$ to show $f^{b+1,n}(\omega) > f^{b,n}(\omega)$.

Let

$$b = n^k \cdot a_k + n^{k-1} \cdot a_{k-1} + \ldots + n \cdot a_1 + a_0 \tag{11}$$

be the $n$-addic expansion of $b$ with $0 \leq a_i < n$ and $a_k \neq 0$ and thus

$$f^{b,n}(\omega) = \omega^{f^{k,n}(\omega)} \cdot a_k + \omega^{f^{k-1,n}(\omega)} \cdot a_{k-1} + \ldots + \omega \cdot a_1 + a_0 \tag{12}$$

Let $j$, $0 \leq j \leq k+1$ be such that $a_j < n - 1$ and $a_i = n - 1$ for all $0 \leq i < j$. Then

$$b + 1 = \begin{cases} n^{k+1} & \text{if } j = k+1 \\ n^k \cdot a_k + \ldots + n^j \cdot (a_j + 1) & \text{if } j \leq k \end{cases} \tag{13}$$

is the $n$-addic expansion of $b+1$ which yields

$$f^{b+1,n}(\omega) = \begin{cases} \omega^{f^{k+1,n}(\omega)} & \text{if } j = k+1 \\ \omega^{f^{k,n}(\omega)} \cdot a_k + \ldots + \omega^{f j,n(\omega)} \cdot (a_j + 1) & \text{if } j \leq k \end{cases} \tag{14}$$

The case $k = 0$ and thus $b = a_0$ is easy. We may thus assume $k > 0$. Then $k+1 \leq b$ and we obtain from the induction hypothesis

$$f^{k+1,n}(\omega) > f^{k,n}(\omega) > \ldots > f^{j,n}(\omega)$$

Comparing (12) and (14) we conclude by the rules of ordinal arithmetic $f^{b+1,n}(\omega) > f^{b,n}(\omega)$, as desired. □

The following lemma will play a crucial role 5 in the upcomming proof.

**Lemma 15.**

1. *For all* $n, m \in \mathbb{N}$ *with* $n > 1 \quad f^{m,n}(x) = f^{f^{m,n}(n+1),n+1}(x)$
2. $o(n,m) = o(n+1, f^{m,n}(n+1))$

*Proof.* It suffices to prove (1). (2) is only a notational variation.
The proof proceeds by induction on $m$. For $m < n$ $f^{m,n}(x) = m = f^{m,n+1}(x)$. Thus also $f^{m,n}(n+1) = m$. In total $f^{f^{m,n}(n+1),n+1}(x) = f^{m,n+1}(x) = m = f^{m,n}(x)$.

Assume now that the claim $f^{k,n}(x) = f^{f^{k,n}(n+1),n+1}(x)$ is true for all $k < m$ and we set out to prove it for $m$. To this end consider the n-adic expansion of $m$:

$$m = n^r a_r + n^{r-1} a_{r-1} + \ldots + n a_1 + a_0$$

By definition

$$f^{m,n}(x) = x^{f^{r,n}(x)} a_r + x^{f^{r-1,n}(x)} a_{r-1} + \ldots + x a_1 + a_0 \tag{15}$$

$$f^{m,n}(n+1) = (n+1)^{f^{r,n}(n+1)} a_r + \ldots + (n+1) a_1 + a_0 \tag{16}$$

Since 16 is the (n+1)-addic expansion of $f^{m,n}(n+1)$ we obtain again from the definition

$$f^{f^{m,n}(n+1),n+1} = x^{f^{f^{r,n}(n+1),n+1}(x)} a_r + x^{f^{f^{r-1,n}(n+1)}(x)} a_{r-1} + \ldots + x a_1 + a_0 \tag{17}$$

Since $r < m$ the induction hypothesis applies and yield the desired result. □

**Theorem 4.** *The Goodstein sequence*

$$m_0 = m \quad m_1 = G_2(m_0) \quad m_2 = G_3(m_1) \quad m_3 = G_4(m_2) \quad \ldots$$

*terminates for every* $m \geq 0$.

*Proof.* We will show that $o(n{+}2, m_n)$ decreases, i.e. $o(n{+}2, m_n) > o(n{+}3, m_{n+1})$:

$$\begin{aligned}
o(n+2, m_n) &= o(n+3, f^{n+2,m_n}(n+3)) && \textit{Lemma } 15(ii)\\
&> o(n+3, f^{n+2,m_n}(n+3) - 1) && \textit{Lemma } 14\\
&= o(n+3, G_{n+2}(m_n)) && \textit{Def. } 6\\
&= o(n+3, m_{n+1}) && \textit{Def. } 7
\end{aligned}$$

□

Definition 8 serves theoretical purposes very well. The following recursive definition of $o(n, m)$ is more amenable for machine assisted reasoning.

**Definition 9.** *The function $o(n, m) : Int \times Int \to Ord$ is defined by*

1. *$o(n, m) = m$ for all $n \geq 2$ and $0 \leq m < n$*
2. *For $2 \leq n$ let $m = n^k a_k + c$ with $1 \leq k$, $0 < a_k < n$ and $c < n^k$ then*

$$o(n, m) = \omega^{o(n,k)} a_k + o(n, c)$$

This is a valid recursive definition since $k < m$ and $c < m$.
For $m \neq 0$ there is a unique $k$ with $n^k \leq m < n^{k+1}$. Furthermore, there are unique $a, c$ with $c < n^k$ such that $m = n^k a + c$. By choice of $k$ we nust have $0 < a < n$. If $k = 0$ then $m < n$. Thus, if $m = 0$ or $k = 0$ then $o(n, m)$ is given by clause 1 in Definition 9 otherwise clause 2 applies. Thus $o(n, m)$ is unambiguously fixed for $m, n \in \mathbb{N}$ and $n \geq 2$. Function values for negative $m$ and smaller $n$ are left open.

## 8 Related Work

The papers [4,7] deal with ordinals in the context of the Isabelle proof assistent. The second paper is in fact a contribution to The Archive for Formal Proofs (AFP) http://afp.sourceforge.net/ and is only accessible to readers familiar with the Isabelle\HOL syntax and semantics. [4] copes with the problem that Isabelle\HOL is a strongly typed logic while ZF is an untyped axiomatization of set theory. Ordinals are replaced in the HOL approach by well-orders. Note, there no concept of an ordinal as an equivalence class (class as opposed to set) of order-isomorphic well-orders, nor is there a unique representation of such an equivalence class.

The research report [6] describes constructions of ordinals in Coq.

In the context of ACL2 the paper [12] and its predecessor [11] present an algorithm for solving problems in ordinal arithmetic working on a kind of nested Cantor Normal Form representation.

In [16] a finite rewrite system is presented whose termination encodes the termination of Goodstein sequences and termination of the rewrite system is proved. This is the first automatic termination proof for Goodstein sequences.

## References

1. W. Ahrendt, B. Beckert, R. Bubel, R. Hähnle, P. H. Schmitt, and M. Ulbrich, editors. *Deductive Software Verification - The KeY Book - From Theory to Practice*, volume 10001 of *Lecture Notes in Computer Science*. Springer, 2016.
2. A.M.Turing. Chceking a large routine. In *Conference on High Speed Automatics Calculating Machines*, pages 67–69. University Mathematical Laboratory, Cambridge, 1949.
3. H. Bachmann. *Transfinite Zahlen*, volume 1 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer Verlag, 2 edition, 1967.
4. J. C. Blanchette, A. Popescu, and D. Traytel. Cardinals in Isabelle/HOL. In G. Klein and R. Gamboa, editors, *Interactive Theorem Proving - 5th International Conference, ITP 2014, Held as Part of the Vienna Summer of Logic, VSL 2014, Vienna, Austria, July 14-17, 2014. Proceedings*, volume 8558 of *Lecture Notes in Computer Science*, pages 111–127. Springer, 2014.
5. F.L.Morris and C.B.Jones. An early program proof by Alan Turing. *Annals of the History of Computation*, 6(2):139 – 14, 1964.
6. J. Grimm. Implementation of three types of ordinals in Coq. Research Report RR-8407, CRISAM - Inria Sophia Antipolis, 2013.
7. B. Huffman. Countable ordinals. *Archive of Formal Proofs*, Nov. 2005. http://afp.sf.net/entries/Ordinal.shtml, Formal proof development.
8. L. Kirby and J. Paris. Accessible independence results for Peano Arithmetic. *Bulletin of the London Mathematical Society*, 14(4), 1982.
9. D. Klaua. *Kardinal- und Ordinalzahlen, Teil 1*. Wissenschaftliche Taschenbücher: Mathematik,Physik. Vieweg Braunschweig, 1974.
10. D. Klaua. *Kardinal- und Ordinalzahlen, Teil 2*. Wissenschaftliche Taschenbücher: Mathematik,Physik. Vieweg Braunschweig, 1974.
11. P. Manolios and D. Vroon. Algorithms for ordinal arithmetic. In F. Baader, editor, *Automated Deduction - CADE-19, 19th International Conference on Automated Deduction Miami Beach, FL, USA, July 28 - August 2, 2003, Proceedings*, volume 2741 of *Lecture Notes in Computer Science*, pages 243–257. Springer, 2003.
12. P. Manolios and D. Vroon. Ordinal arithmetic: Algorithms and mechanization. *J. Autom. Reasoning*, 34(4):387–423, 2005.
13. M. Rathjen. Goodstein revisited. *ArXiv e-prints*, May 2014.
14. R.L.Goodstein. On the restricted ordinal theorem. *Journal of Symbolic Logic*, 9:33–41, 1944.
15. G. Takeuti and W. M.Zaring. *Introduction to Axiomatic Set Theory*, volume 1 of *Graduate Texts in Mathematics*. Springer Verlag, 1971.
16. S. Winkler, H. Zankl, and A. Middeldorp. Beyond Peano Arithmetic – Automatically Proving Termination of the Goodstein Sequence. In F. van Raamsdonk, editor, *24th International Conference on Rewriting Techniques and Applications (RTA 2013)*, volume 21 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 335–351, Dagstuhl, Germany, 2013. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.