UNIVERSITY OF OSLO
Department of Informatics

Mapping finite trees to ordinals

Master Thesis

Alfred Bratterud

November 2009

Contents

Acknowledgments

First, I have to thank my supervisor **Herman Ruge Jervell** for introducing me to logic. While there are many reasons to study logic and mathematics, Herman seems to be interested in this: discovering, rediscovering and understanding fascinating ideas. Being allowed to take part in such discovery, with such an experienced guide, has been remarkable.

I first met Herman through his textbook in logic and computability [Jer01] and [Jer04]. At a crucial point in some important proof, he would write: "you now have enough information to complete the proof". He continued this habit in a later course and a later book, some times leaving things he had not completely worked out himself as exercises. This gave me my first opportunity to discover some logic on my own, and it was no less than thrilling. As a thank you, I have included the last section about the continuum hypothesis for trees, where one problem is left as an exercise to the reader.

Second, I want to thank **Lars Kristiansen**, for a comprehensive and inspired course in mathematical logic, where the gap between logic and mathematics was bridged, and where I was guided through Gödels incompleteness theorems for the first time. I also want to thank **Dag Normann** for an excellent introduction to axiomatic set theory, where ordinals, cardinals and the levels of infinity were first explained to me. Thanks also to **Asbjørn Brændeland** for a great course on the philosophy of science where I saw logic in a wider perspective, to **Roger Antonsen** for crystal clear introduction to first order logic and to **Evgenij Thorstensen** for sparring with me. Next, I want to thank **Steinar Lund** for teaching me how to do proofs by induction. There are plenty of those in the pages to come, even some transfinite ones, but Steinar provided the base case.

A special thanks goes to my family and my friends for their support, encouragement and understanding. In logic there are axioms, in life there is family and friends, upon which everything else rests. And last but not least, thanks to **my wife Cecilie**, for all the conversations, all the meals, and for supporting me so strongly in this project - of which she understands almost nothing.

Chapter 1

Introduction

In 1931 Kurt Gödel showed that any first order theory N, with the computational strength of addition and multiplication, will contain formulas ϕ such that $N \nvdash \phi$ and $N \nvdash \neg\phi$. In particular, this formula ϕ will state exactly this fact, that it is not provable. And so the sentence must be true, because otherwise N would be inconsistent, which we know it is not[1]. In his second incompleteness theorem, Gödel showed that no such theory of arithmetic can be proven consistent from within itself, because the formula stating the consistency of N would be a Gödel sentence too. However fascinating the result, it also leaves one wanting for more information. What is it about addition and multiplication that gives a system just enough power to be incomplete? And when we know that we get "Gödel sentences" in our system, what exactly are those sentences, and are they important? In 1936 Gerhard Gentzen [Gen69] provided some more insight, by showing that the Gödel Sentence for Peano Arithmetic (PA) corresponds to transfinite induction up to ϵ_0. This puts the Gödel sentence on a scale, namely that of the ordinals, and the exact point of the scale where the Gödel sentence for arithmetic occurs, is the main subject of this thesis.

In the following I shall construct two different, but related theories, which on the surface will look nothing like classical arithmetic. They will, however, be very well suited to construct Gödel sentences, and most importantly for this thesis, to study what goes on below ϵ_0. I will approach this subject in three ways; first of all I will give a purely mathematical basis for the rest of the discourse. As a basis, I will provide the necessary definitions and axioms, and go on to proving some interesting theorems. Secondly, I will try to give the reader a guided tour through the subject matter, with emphasis on understanding ϵ_0 and its complexity. Thirdly, in aid of this effort, I will

[1] I first saw a version of this proof in [Lea00]

give some purely non-mathematical illustrations in the form of fables or short stories, hopefully giving the reader some helpful metaphors along the way. Another important goal with these parts is to give my wife some idea of what I'm doing.

1.1 Starting the tour

The first thing we need to define is the smallest infinite number, ω, and we usually do it recursively, like this:

Definition 1 (Omega)**.**

$$0 \in \omega$$
$$\alpha \in \omega \rightarrow s(\alpha) \in \omega$$

So ω is the set containing every ordinal one can build from 0, using the successor function. And since the successor function can't be broken down to something simpler, it can't really be defined, but rather it must simply be understood as meaning "adding one".

ϵ_0 is - almost unfortunately - very easy to define as well:

Definition 2 (Epsilon 0)**.**
ϵ_0 is the smallest ordinal α such that $\omega^\alpha = \alpha$

So $\omega^{\epsilon_0} = \epsilon_0$, and for every number smaller than ϵ_0, this is not the case. In other words, it is the limit of ordinal exponentiation with finite support, meaning that $\epsilon_0 = \omega^{\omega^{\omega^{\cdot^{\cdot^{\cdot}}}}}$. As long as one knows the meaning of ω and exponentiation, this makes sense, since everyone gets that the dots means "and so on", people will often say "ah, I get it", when they see the definition. This stems from a superficial understanding of the syntactical structure of the term, and without some hard thinking, one might not really understand the difference between ϵ_0 and ω^ω - other than that the latter is ... less complex. Consider for instance the following very small ordinals:

$$\omega^{\omega^{\omega^{\omega}}+\omega^{\omega^5}+\omega} \quad \text{compared to} \quad \omega^{\omega^{\omega^{\omega}}+\omega^{\omega}+\omega^{\omega 2879}}$$

Who's to say what happens in between?

In the following I shall not presume to make all of this clear and simple, because it's not. The complexity of ϵ_0 marks the very limit of what we can handle in first order arithmetic - which is a very strong mathematical theory. I shall, however, try to shine a little light on the basic structure of this number, by building it up from the bottom - or rather - planting it. Because ϵ_0 is also a tree.

1.2 Counting trees

A botanist in say, the 16th century Europe, wants to classify trees. And since this is before the time of modern taxonomy, he decides to classify them according to their branching and height, giving each type of tree a number. He has a rough idea of the plant cell, in the sense that there is a smallest piece common to every kind of wood, or at least of roughly similar size, and so he thinks this should be the unit for measuring height. The first tree in his system is one cell high, with no branches, and he writes this down as tree number one. Tree number two, of course, will be the tree that's two cells high, and still with no branching, and tree number three the one that's three cells high and with no branching, and so on. Simple. But then he soon gets into trouble, because trees can get really high - nobody knows exactly *how* high, and since microscopes are pretty rare, it's hard to say exactly how small the cells are. So the tallest tree could be ... ten thousand cells? Or a million? What if there goes a 100 cells to the millimeter, and if trees can grow as high as 100 metres? Then there will be trees as tall as 100×100.000 equals ten million cells high. That's a lot, and who knows if that's even close to the limit?

Now say he did set a limit, at a 100 million to make a seemingly generous estimate, and let the smallest tree with a branch be tree number 100 million and one. Since he don't know how short trees can be when they first branch, he assumes the worst, and lets it split right from the first cell, or height one. Now nobody really knows how long and tall one of the branches could be - for all anyone knows it could develop into the stem of a really tall tree, and so this branch could also get up to a 100 million. So he decides to make the left branch be the numbers from 100 million and one up to 200 million and one, while the right branch gets the numbers 200 million and one up to 300 million and one. But since both branches can branch again, any number of times, and all the branches can branch, there will be multiples of multiples of multiples of the number 100.000.000 and that's not at all practical.

The botanist decides that there must be a better way, and he gets his first brilliant idea; he decides to color the numbers. He lets all the aquamarine numbers denote all the trees of branching one - the ones that grow straight up from the ground with no real branches at all. And after writing that down he writes the following: "and after all the trees numbered with aquamarine, there will be the trees numbered with plain blue, of which the first will be the one with height one and branching two. And after all the trees numbered with aquamarine and blue, comes the tree with height one and branching three, which will have the beautiful color cornflower blue."

And without knowing, the botanist has invented a very small initial segment of the ordinal numbers. When we speak of the ordinals, we usually speak of infinity. This is ok, and correct, but some times it mystifies the subject matter more than it clarifies. Because instead of saying that the natural numbers go on and on towards infinity - and after infinity there is infinity plus one - one could say that aquamarine numbers comes first, then the blue and then the cornflower blue. It means almost - and we'll get back to that - the same, but it sounds a little less mysterious.

Now the botanist has made some progress, but there are still some clouds on the horizon; because he still needs to be able to combine the numbers in some way. For the first blue number he needs to decide what comes next. And how does he denote that a tree can have a long left branch with no branches - which should have the color aquamarine - but a right branch with a blue split, followed by a long stretch of aquamarine, and then another split? The botanist is now starting to uncover some of the complexities of the lower ordinals. Because by only coloring the numbers, he hasn't really gotten all that far. Say for example that he lets the trees of branching four be colored dark violet, branching five ebony, and branching six forrest green, and so on. Since the first letters of these colors are conspicuously familiar, we can use them to denote the colored numbers as follows:

$$1a, 2a, 3a, \ldots, 1b, 2b, 3b, \ldots, 1c, 2c, 3c, \ldots, 1d, \ldots, 1e, \ldots, 1f$$

Where a is aquamarine, b is blue and so on. In ordinal notation, this corresponds to

$$\omega + \omega + \omega + \omega + \omega + \omega$$

Which again is equal to a mere $\omega 6$. And so we see that we are still a very long way from ω^2, incredibly far from ω^ω, and nowhere near ϵ_0. And ϵ_0 is, take my word for it, the ordinal needed to map just the trees of branching less than three.

And so, even if the botanist were to use all the color shades discernable by the human eye, he would still only get to ωn, for some not so big n. He soon realizes the limitation of his system, and after pondering the problem for a while, he decides to make a more elaborate system, where the *combination* of numbers of different colors make up the complexity. He allready has a good basic idea; that the blue numbers are all bigger than the aquamarine ones, and so on, but inspired by the decimal numbers he decides to do one better. How about this, he suggests to a rosebush, if we let the aquamarine numbers be as they were, and let the "blue number one" be the first to follow. But now, instead of going straight to blue two, let us start all over again with aquamarine!

And so, to get from zero and all the way up to the second blue number, we first need to go through all the aquamarine numbers once, then pass "blue one", and then go through all the aquamarine numbers once more. *Then* we get to blue two. And to get to blue three, of course, we start all over again, with aquamarines. Now we're talking, the botanist thought, and started explaining to one of the budding roses how to get to the first number in cornflower blue. Because, as before, this number comes after all the blue ones. The difference is that between all the blue ones, now there are aquamarines. And so, to keep the complexity growing, there should be a whole range of blue numbers between every cornflower blue one. So, to get to "cornflower blue two", just go through all the aquamarines to blue one, go through all the aquamarines again to blue two, continue through all the aquamarines to blue three, and so on, eventually passing through all the blue ones to cornflower blue one. Repeat this whole process once more, and you're there. The rosebud looked slightly demotivated.

Again, employing the conspicuous first letters of the colors names as notation, we now have the following system:

$$1a, 2a, 3a, \ldots, 1b, 1a, 2a, 3a, \ldots, 2b, 1a, 2a, 3a, \ldots, 3b, 1a, 2a, 3a, \ldots$$

Which can be expressed as

$$\omega a + b1 + \omega a + b2 + \omega a + b3 + \omega a + \ldots$$

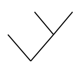
which can further be compressed to

$$\omega a \omega b \omega c \omega d \omega e \omega f \ldots$$

And finally to

$$\omega^n$$

Where n is the number of colors. And so, however more complex our system now is, it still isn't anywhere close to ϵ_0. The botanist notices this first months later. He has been busy classifying trees with his new system, and it's been working up until now; he is about to give a number to the tree of height two, that splits right away in two, with a short left branch, but a right branch that splits in two again right away. He has decided to let the branching of the trees be ordered from left to right, and he's been going on with huge trees of left branches, splitting again along the left branch which grow for a while, and then split again. But no matter how he tries, he just can't find a way to number his way further than this



We are in a sense going to follow the same path as the botanist, but in stead of coloring numbers we're going to use three flavors of constructor functions, namely $f(x)$ and $g(x, y)$, including h(0,0,0) as upper bounds. We're gonna make every combination in an orderly fashion, mapping the combinations to the ordinal numbers as we go. And as it turns out, they fit together like two sides of a zipper. We shall see that the botanists clever numbering system stranded at ω^ω, and we shall see why.

Chapter 2

Trees with fixed points

In [Jer00], Herman Ruge Jervell gives an account of how to well-order finite trees. The ordering is given by a simple formula, which for any two trees, A, B, determines whether $A = B$, $A < B$ or $B < A$. While this does gives us a total and transitive ordering, it does not give us much help in determining which ordinal a given tree corresponds to. Another feature of the Jervell ordering, is that every tree represents a unique ordinal. This means that none of the constructor functions have fixed points, since we never have $\lambda(x) = x$, for any constructor λ, but rather $\lambda(x)$ is always a tree representing a unique ordinal.

In the following I shall give an axiomatic definition of another ordering of the trees with branching < 3, where all the constructors do in deed have fixed points. I will then provide a formula which maps this ordering of trees to the ordinal numbers, and prove the correctness of the formula. Giving all the constructors fixed points will slow down the progression of the trees, since a lot of trees will be i equal to one another. Specifically, only the trees that grow from the leaves out, each subtree smaller than the next, will represent unique ordinals. But as it turns out, we will still get that $h(0, 0, 0) = \epsilon_0$, just as in the more economic Jervell-ordering.

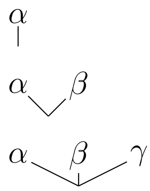
Interestingly, this means that we can have two sequences, one growing faster than the other, but still have them both reach ϵ_0 at the same time.

2.1 Definitions

Before we get to the ordering itself, we will need precise definitions of some functions and concepts.

Tree constructors: Our universe $\mathcal{T}$ will be trees, made up of functions called tree-constructors, where the arity of the function corresponds to branching of the tree at a given node;

$$f(\alpha) =$$
$$g(\alpha, \beta) =$$
$$h(\alpha, \beta, \gamma) =$$



In the following, I will only consider trees with branching ≤ 2, but $h(0, 0, 0)$ will be included as an upper limit of our trees.

Definition 3 (Trees of branching ≤ 2, $\mathcal{T}_2$)**.**

$$0 \in \mathcal{T}$$
$$\alpha \in \mathcal{T} \rightarrow f(\alpha) \in \mathcal{T}$$
$$\beta, \gamma \in \mathcal{T} \rightarrow g(\beta, \gamma) \in \mathcal{T}$$

Definition 4 (Universe of trees, $\mathcal{T}$)**.**

$$\mathcal{T} = \mathcal{T}_2 \cup \{h(0, 0, 0)\}$$

λ-functions: I will use λ-notation in the usual way; $\lambda x.g(g(0, x), 0)$ will be the set $\{g(g(0, x), 0) | x \in \mathcal{T}\}$. In other words, it will be a tree constructor for trees with a common root in $g(g(0, x), 0)$.

Definition 5 (Tree constructors)**.**

$$f(x)$$
$$\lambda x.g(x, \beta)$$
$$\lambda x.g(0, x)$$

for every $\beta \in \mathcal{T}_2$

These are the building blocks for all the trees. Combining these functions in the usual way will make terms, which are exactly the trees in our universe. Notice that β is any tree. This means that when we start to generate trees inductively, which is what we'll do, the set of constructors will grow as well. For each new tree we create, there's also a new constructor.

2.1.1 Coding syntax

Before we go on, let me say a few words about Gödels sentences for trees. Creating a Gödel sentence is a matter of coding the syntax of logic into the logical theory. Once you manage to do that, there will be one term within the system representing each formula in the language. Once that's done, the notion of a proof has to be coded too. But since a proof, in a deductive proof system, is really just a list of formulas, ending up with the conclusion, once you have found a way to code formulas and reason about them, the way to coding proofs is not so long. Unfortunately, doing this properly is outside the scope of this thesis, so let me just point out the following: the universal way to represent syntax is as trees. And if you think about a formula as an array, you can easily see how a proof could be looked upon as a two-dimensional array. Which again is a tree. When Gödel created coding for natural numbers, he had to employ all the strength of first order arithmetic, since the numbers is inherently linear. When we do it in trees this is not the case. We leave this matter here, since if not by this argument, it will be clear from the mapping to ordinals that the theories of trees are more than strong enough to code syntax, and thus they contain Gödel sentences.

We can now define the notion of a numeral. Numerals are functional terms, which are merely syntactical constructions. $\bar{0}$ is a numeral, frequently abbreviated to 0, since terms and trees turn out to be the same as ordinals. The numerals $\bar{n}$ are defined as follows:

Definition 6 (Numerals $\bar{n}$)**.**

$$
\begin{aligned}
f_1(0) &= f(0) \\
f_{n+1}(0) &= f(f_n(0)) \\
\bar{n} &= f_n(0)
\end{aligned}
$$

When this chapter is complete, the gap between syntactical terms of this kind and ordinals will be closed, but it's nice to go step by step. We can now define the successor function recursively. For the numerals $\bar{n}$ we will have that $s(\bar{n}) = f(\bar{n})$. But since all our constructors, including f, will have fixed points, the successor function must be defined separately:

Definition 7 (Successor)**.**

$$
\begin{aligned}
s(\bar{n}) &= f(\bar{n}) = n + 1 = 1 + n \\
s(g(\alpha, \beta)) &= g(s(\alpha), \beta)
\end{aligned}
$$

This definition gives us a hint of how the trees will grow; when we apply the successor function to a given tree, the function call will "percolate" up through the tree, giving unary branches only at the very tops.

The limit of a function λ, denoted $\lim(\lambda)$ will be defined as follows:

Definition 8 (Limit)**.** The limit function, denoted $\lim(\lambda) = \alpha$ is *represented* by the following Δ_0 formula, and thus it is computable[1]:

$$\begin{aligned} \equiv\ & \lambda(0) < \alpha \wedge \\ & \quad \forall x < \alpha . x < \lambda(x) < \alpha \end{aligned}$$

Where $\equiv$ means syntactical abbreviation.

By this definition and the axioms to come, $\lim(\lambda)$ will be the *smallest* limit of λ, when we work within in Ordering C. There are in deed other ordinals β such that $\forall x < \beta . \lambda(x) \leq \beta$.

Definition 9 (Fixed point)**.** I will let "*the* fixed point" of a function λ mean the smallest α such that $\lambda(\alpha) = \alpha$. I shall denote this $fix(\lambda)$. Further, "*a* fixed point of λ" can be any β such that $\lambda(\beta) = \beta$.

[1] A Δ_0 formula is a first order formula with only limited quantifiers; that is $\forall x < \alpha$, $\exists x < \beta$ etc. The idea here is that since we are in a well-ordering, and we are promised that something holds for some element smaller than some α or β, we can start searching below that limit, and since there is no infinitely descending sequence, the search must terminate. For a more thorough treatment, see [Lea00], chapter 4

2.2 Axioms of Ordering C

I will now give a constructive definition of the ordering, using 0, $h(0, 0, 0)$ as extremes and the fixed points as ordered constants in between. The remaining gaps will be filled inductively.

1. **Initial segment:**
   $0 < f_n(0) < g(0, 0)$

   for every $n \in \mathbb{N}$

2. **λ preserves the ordering:**
   $\alpha < \beta \leftrightarrow \lambda(\alpha) < \lambda(\beta)$

3. **Limits:**
   $\lim(f) = g(0, 0)$
   $\lim(\lambda x.g(x, \beta)) = g(0, s(\beta))$
   $\lim(\lambda x.g(0, x) = h(0, 0, 0)$

4. **Limits are fixed points:**
   $\lambda(\lim(\lambda)) = \lim(\lambda)$

5. **No mercy:**
   $(\lambda(\alpha) = \alpha \wedge \alpha < \beta) \rightarrow \lambda(\beta) = \beta$

   Where λ is any of the constructors.

2.2.1 Properties of the ordering

Now we can derive all we need to know about the ordering.

Equality is not strict: It is clear from "limits are fixed points" (axiom 4) that I use the normal equality sign = for two trees, much more often than just when they are in fact the same tree. This is fine; just read $\alpha = \beta$ as α and β represent the same ordinal. Instead of throwing away all the trees that do not represent unique ordinals, we will benefit slightly from keeping them: when get to mapping trees to ordinals, we shall in fact see that we have just inherited the equalities that hold in ordinal arithmetic, such as $1 + \omega = \omega, \quad \omega + \omega\omega = \omega\omega$ etc.

The numerals are the smallest trees: We know that $g(0, 0)$ is the smallest limit, since all the others have a successor in the second argument. We also know, from the definition of lim that all the numerals are smaller than $g(0,0)$. By the same definition we have that $f(\bar{n}) < ff(\bar{n})$, and since $0 < f(0)$ by "initial segment" (axiom 1) , all the unary trees line up at the beginning.

The limits are ordered: We have that $\lim(\lambda x.g(x,\beta)) = g(0,s(\beta))$. By the definition of lim, $\lambda(0) < \lim(\lambda)$. Thus $g(0,\beta) < g(0,s(\beta))$.

Ordering C is a well-ordering: Since I shall give a mapping between Ordering C and the ordinals, a proof of this fact will not be necessary. However, we can notice some features of the ordering, which could be developed into a real proof: First, 0 is the least element by definition, and the numerals grow from 0. There is a clear one to one mapping between the numerals and the natural numbers, which are naturally well-ordered. Secondly, the limits are ordered, and from each limit, successive calls to the constructor progress towards the next limit. Transitivity of all the constructors are integrated in the definition of lim, which in turn defines the order of the constructors. Next, we have that the constructors all preserve the ordering, making them monotonically increasing. What remains would be showing the totality of the order, which is quite clear; the universe $\mathcal{T}$ is defined in terms of the very same functions that make up the ordering.

All function limits are limit ordinals - not the other way around: All our function limits are limit ordinals, but not all limit ordinals are function limits. We have in general that:

$$fix(\lambda) = \alpha \rightarrow g(0,\lambda(\alpha))) = g(0,\alpha)$$

and since fixed points have "no mercy" (axiom), the same will be the case for any α' such that $\alpha < \alpha'$. Notice, however, that

$$fix(\lambda x.g(0,f(x))) \neq g(0,g(0,0)) \qquad \text{even if } fix(f) = g(0,0)$$

Rather, this last fact ensures that

$$g(0,f(g(0,g(0,0)))) = g(0,g(0,g(0,0)))$$

In deed, $g(0,g(0,0))$ is a limit ordinal, but this does not make it the limit of a constructor function. What we have is that $\lambda' x.g(0,\lambda(x)) = \lambda'' x.g(0,x)$ for $x > \lim(\lambda)$, meaning that λ seizes to contribute to λ' after reaching it's fixed point. So, to sum it up, any tree corresponding to a limit ordinal can be described in terms of a complex constructor function, but this function won't have it's own unique fixed point:

$g(0, g(0, 0))$ is the smallest tree greater than any tree constructed by $\lambda x.g(0, f(x))$, but to get $\lambda(x) = x$ we still need $x \geq h(0, 0, 0)$ which is the fixed point of $\lambda x.g(0, x)$.

Limits are fixed points: Since the introduced fixed points are all limits, we immediately get that they are the smallest fixed points, and so $\lim(\lambda) = fix(\lambda)$, by the definition of fix and lim. We can now derive a lot of fixed points and limits, for example:

$$\begin{aligned} fix(\lambda x.g(x, \bar{2})) &= g(0, \bar{3}) \\ fix(\lambda x.g(x, g(g(2, 2), 0))) &= g(0, s(g(g(2, 2)))) \\ &= g(0, g(s(g(2, 2)))) \\ &= g(0, g(g(3, 2))) \end{aligned}$$

Redundancy: There are some properties which can be derived both from the definition of lim and from $\lambda-$preservation; for instance that the limits are ordered. Clearly, since $\beta < s(\beta)$, $g(0, \beta) < g(0, s(\beta))$ by $\lambda-$preservation. In general, from $\alpha < \beta < \lim(\lambda)$ we can always derive that $\lambda(\alpha) < \lim(\lambda)$ and that $\lambda(\alpha) < \lim(\lambda)$. We can, however, not derive that $\lambda(\alpha) < \lambda(\beta)$, and thus not $g(0, \beta) < g(f(0), \beta)$ without the axiom.

Two more properties are important to point out; the fact that both left and right branching preserve limits. This will help us with the transfinite cases when we map the trees to ordinals.

Lemma 1 (Limit preservation)**.** *If $\alpha = \lim(\lambda)$, which means that α is the smallest tree greater than $\lambda(x)$ for any $x < \alpha$, then $g(\alpha, \beta)$ is the smallest tree greater than $g(\lambda(x), \beta)$.*

For any λ such that $\lim(\lambda) < g(0, s(\beta))$

Proof. We assume $\alpha = \lim(\lambda)$

1. Assume for contradiction that there are trees γ, δ such that $g(\lambda(x), \beta) < g(\gamma, \delta) < g(\alpha, \beta)$. Since "$\lambda$ preserves the ordering" (axiom 2) , if $\gamma \geq \alpha$ and $\delta \geq \beta$ the last inequality would not hold. Hence we must have that either

   (a) $\gamma < \alpha$ while $\delta \leq \beta$, or

   (b) $\delta < \beta$ while $\gamma \leq \alpha$

for the last inequality to hold

2. If 1b we have $\delta < \beta$ and then, by "Limits" (axiom 3), $\lim(\lambda x.g(x,\delta)) = g(0,s(\delta)) \leq g(0,\beta)$, in which case $g(\gamma,\delta) < g(\lambda(x),\beta)$ by the definition of lim, and since $\gamma \leq \alpha < \lim(\lambda x.g(0,s(\beta)))$.

3. This contradicts our assumption, and hence we must have 1a and $\gamma < \alpha$. By the definition of lim, $\lambda(x) < \lim(\lambda)$ for every $x < \lim(\lambda)$. Since we have $\gamma < \alpha = \lim(\lambda)$, we also get $\lambda(\gamma) < \alpha$. Now we have $g(\lambda(x),\beta) < g(\lambda(\gamma),\delta)$ and since $\lambda(\gamma)$ is not greater than every $\lambda(x)$, where $x < \lim(\lambda)$ (for example not $\lambda(\lambda(\gamma))$), we must have $\beta < \delta$ for the inequality to hold, by "λ preserves the ordering" (axiom 2) .

4. By "limits" (axiom 3) we have $\lim(\lambda x.g(\alpha,\beta)) = g(0,s(\beta))$ and this is smaller than or equal to $g(0,\delta)$, since $\beta < \delta$. Clearly, $g(\gamma,0) < g(\gamma,\delta)$, and by the definition of lim, $g(\alpha,\beta) < \lim(\lambda x.g(x,\beta))$, and thus $g(\alpha,\beta) < g(\gamma,\delta)$.

5. But by the assumption in step 1 $g(\gamma,\delta) < g(\alpha,\beta)$ and thus our assumption has led to a contradiction.

6. Notice that since our only requirement to x was that $x < \lim(\lambda)$, so we can also have $x = \lambda_{n-1}(y)$ by the definition of lim, and since $\lambda(\lambda_{n-1}(y)) = \lambda_n(y)$, we have also shown that $g(\lim(\lambda),\beta)$ is the smallest tree greater than $g(\lambda_n(y),\beta)$.

□

Corollary 1. *If α is the smallest tree greater than some γ, $g(\alpha,\beta)$ is the smallest tree greater than $g(\gamma,\beta)$.*

Proof. This is just a special case of the general argument in lemma 1; just let $\lambda(x) = \gamma$ and the proof is identical. □

This does not, by any means, give us that the same thing holds for right branching. Assume $\alpha = s(\beta)$ and notice the following fact:

$$\forall x < \alpha . g(0,x) < g(1,\beta)$$

In fact we have

$$\forall x < \alpha . g(0,x) < g(\gamma,\beta)$$

For any $\gamma < g(0,\alpha)$, by "limits" (axiom 3) . We do, however, have that $g(0,\lim(\lambda))$ is the smallest tree greater than every $g(\alpha,\lambda(x))$, where $\alpha < g(0,\lim(\lambda))$. These cases, according to Theorem 1, correspond directly to when we have ω^β and β is a limit ordinal. In terms of function limits, $g(0, \lim(\lambda))$ is not a limit for any constructor function, in fact it is the limit of a whole level of constructor functions.

Lemma 2 (The limit of limits)**.** *For any constructor function λ, and any $x < \lim(\lambda)$, the smallest tree greater than every $g(0, \lambda(x))$ is $g(0, \lim(\lambda))$*

Proof. We show that there is no β such that

$$g(0, \lambda(x)) < \beta < g(0, \lim(\lambda)) \tag{2.1}$$

for every $x < \lim(\lambda)$.

1. Assume 2.1 is true, for contradiction.

2. Since $\beta < g(0, \lim(\lambda))$, we have at most $\beta = g(\alpha, \lambda(\gamma))$ for some $\gamma < \lim(\lambda)$ by the definition of lim.

3. Now, we can assume that $\alpha < g(0, s(\lambda(\gamma)))$. If not, we get $g(\alpha, \lambda(\gamma)) = \alpha$, and we can start the analysis all over again for α, recursively, until all the fixed points are out of the tree.

4. But then

   $$\beta = g(\alpha, \lambda(\gamma)) < g(0, s(\lambda(\gamma))) \leq g(0, \lambda(\lambda(\gamma))) < g(0, \lim(\lambda))$$

   by the definition of lim, and by "λ preserves the ordering" (axiom 2) , and this contradicts the assumption.

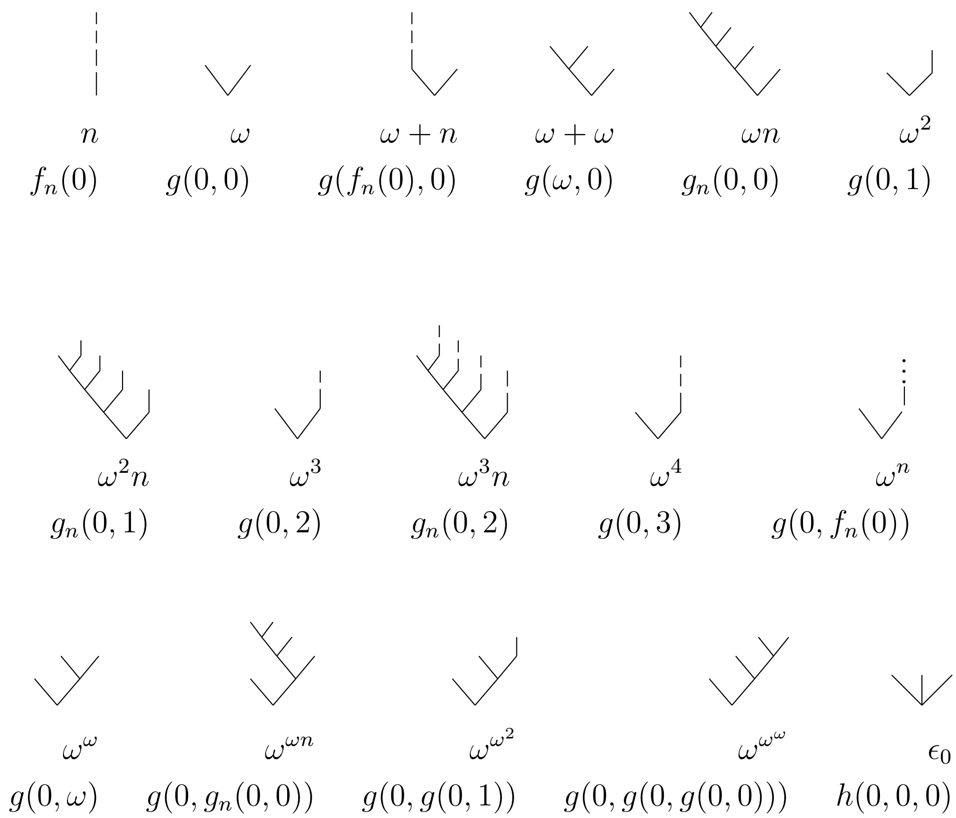
5. As noted in the previous proof, our only requirement to x was that it should be smaller than $\lim(\lambda)$, and so we can also have $x = \lambda(\lambda_{n-1}(y)) = \lambda_n(y)$.

□

For both these lemmas, notice that neither "limits are fixed points" (axiom 4) or "no mercy" (axiom 5) were employed; in fact we only needed the definition of lim and "λ preserves the ordering" (axiom 2) . Therefore we can use them again when we get to ordering A, which is *without* fixed points.

2.3 How the trees grow

All our trees grow from the leaves up, and not from the root, except the unary trees, where there is no difference. For any tree with a binary branching somewhere, however, the difference is significant. The reason for this limited form of growth is exactly the fact that we have fixed points. Graphically, the trees grow like this:



2.3.1 Why the counting stranded

Now we can see why the botanists elaborate counting system stranded. We followed it all the way up to ω^n, where n were the number of colors he used. If he were to continue to the trees beyond all of these, that is to ω^ω, he would have to have infinitely many colors. And then , to keep going further, he would have to find yet another clever trick to keep track of those infinite colors while increasing the power of the system even further.

The ordinal notation below the trees assumes that there is a mapping to the ordinals, which is something we have not yet proven. So it's about time we get around to just that.

2.4 Mapping Ordering C to the ordinals

Before we start mapping ordinals to trees, we need some facts about how they behave.I will not give a complete theory of the ordinals at this point, that would be a detour. I will rather assume that I'm working inside a standard one[2], give a brief summary of the relevant features of such a theory, and a proof sketch (no real definitions, no real proof), which is intended to provide some intuition.

2.4.1 Facts about ordinals

Notice that ω and ϵ_0 were defined in Definitions 1 and 2. Beyond this, we need to know the following:

Lemma 3 (Facts about ordinals). *We are assuming a standard theory of ordinals, in which the following holds:*

$$\begin{aligned} n + \beta &= \beta \\ n\beta &= \beta \\ \omega + \omega^2 &= \omega^2 \\ \omega\omega^\beta &= \omega^\beta \end{aligned}$$

For $n \in \mathbb{N}, \omega \leq \beta$, and in general:

$$\lim(\lambda^+ x.(\alpha + x) = \alpha\omega \tag{2.2}$$

$$\lim(\lambda^\times x.(\alpha x)) = \alpha^\omega \tag{2.3}$$

$$\lim(\lambda^{exp} x.(\beta^x)) = \epsilon_0 \tag{2.4}$$

Where $\omega \leq \beta < \epsilon_0$.

We also have that these functions preserve the ordering:

$$\alpha < \beta \leq \lim(\lambda) \rightarrow \lambda(\alpha) < \lambda(\beta) \tag{2.5}$$

proof sketch. The following is an outline of a proof, starting with 2.2.

1. For λ^+ we have:

$$\begin{aligned} \alpha + \alpha\omega &= \alpha + \alpha + \alpha\omega \\ &= \alpha + \alpha + \alpha + \dots \\ &= \alpha\omega \end{aligned}$$

[2]For a full-grown specimen of such a theory, consult [Kun80] Chapter 1 §7.

2. Similarly, for $\lambda^{\times}$:

$$\begin{aligned}\alpha\alpha^{\omega} &= \alpha\alpha\alpha^{\omega} \\ &= \alpha\alpha\alpha\ldots \\ &= \alpha^{\omega}\end{aligned}$$

3. The definition of ϵ_0 is that it is the smallest ordinal α such that $\omega^{\alpha} = \alpha$

4. We cannot not prove that the given functions respect the ordering without defining them properly, which would be a sidestep here. Suffice it to say that this must be the case if we want addition, multiplication and exponentiation to be monotonic functions - which they are.

Definition 10 (Cantor Normal Form - CNF)**.** Every ordinal can be written on the following form:

$$\omega^{\beta_1}c_1 + \omega^{\beta_2}c_2 + \ldots + \omega^{\beta_k}c_k$$

Where $b_1 > b_2 > \ldots > b_k$ are ordinals.and $c_1, c_2, \ldots, c_k$ are positive integers.

We can also let all the coefficients $c_i = 1$, and allow the exponents β_i to be equal, which gives a *simpler form*;

$$\omega^{\beta_1} + \omega^{\beta_2} + \ldots + \omega^{\beta_k}$$

Where $b_1 \geq b_2 \geq \ldots \geq b_k$ are ordinals. This form will be important in the following.

This fact is due to a theorem by Cantor, which is interesting in it's own right, but in the following we will just assume it to be true.

2.4.2 Mapping the initial segment

1. The numerals $\bar{n}$ are natural numbers. Just let $\bar{0} = 0$, and $f_n(0) = n$ for any $n \in \mathbb{N}$.

2. $g(0,0) = \omega$. Since we have that $\lim(f) = g(0,0)$ we know that $g(0,0)$ is the smallest tree such that $f_n(0) < g(0,0)$. Thus $g(0,0)$ is the smallest tree greater than all the numerals, so clearly, $g(0,0) = \omega$ by the definition of ω

Lemma 4 (Adding one)**.** $f(\alpha) = 1 + \alpha$

Proof. We have that $\omega \cup \{1\} = 1+\omega = \omega$ since $1 \in \omega$ by definition. However, $A = \{\omega\} \cup \{1\} = \omega+1$ is an entirely different set, which is bigger than ω since $A \setminus \omega = \{1\}$ whereas $\omega \setminus \omega = \emptyset$. Since $g(0,0) = \omega$ and $f(g(0,0)) = g(0,0)$ we have that $f(\omega) = 1 + \omega$. Since ω is an initial segment of every ordinal $\alpha \geq \omega$, clearly $1 + \alpha = \alpha$. And since fixed points "have no mercy" by axiom, $f(\alpha) = \alpha$ for every $\alpha \geq \omega$, and thus $f(\alpha) = 1 + \alpha$. □

2.4.3 Mapping the rest

Theorem 1 (Trees under ordering C are ordinals.)**.** *There is a mapping from Ordering C, of trees with branching* < 3*, to the ordinals* α *such that* $\omega \leq \alpha < \epsilon_0$*,given by the following formula:*

$$\begin{array}{c}\alpha \quad \beta \\ \bigvee\end{array} = \omega^{1+\beta} + \alpha$$

Equivalently:

$$g(\alpha,\beta) = \omega^{1+\beta} + \alpha$$

where α, β *are any ordinals* $< \epsilon_0$*.*

Proof. We show the following big steps, each with substeps:

Left branching is addition: $\lambda'\alpha.g(\alpha,\beta) = g(0,\beta) + \alpha$ for $\beta < h(0,0,0)$

Right branching is exponentiation: $\lambda''\beta.g(0,\beta) = \omega^{1+\beta}$

The combination of the two gives the mentioned formula: To get an ordinal from a tree θ, first use the equality λ' on θ to get out the left argument, then use the other equality, λ'', on the remaining g. This is then done recursively to get an ordinal. The process reversed gets a tree out of a given ordinal.

2.4.4 Left branching is addition:

$\lambda\alpha.g(\alpha,\beta) = g(0,\beta) + \alpha$ for $\beta < h(0,0,0)$. We have the following substeps:

1. $\alpha = s(\gamma)$. We start with α being a successor ordinal on the right side, and a successor tree on the left side of the equation. We then do induction over α.

   Base: $\alpha = 0$: Clearly $g(0,\beta) = g(0,\beta) + 0$.

Step: $g(\gamma, \beta) = g(0, \beta) + \gamma \rightarrow g(s(\gamma), \beta) = g(0, \beta) + \gamma + 1$

We shall see that both sides of the equation progress by the smallest possible step, from the antecedent to the consequent.

(a) The weakest function, by the "Limits" axiom, is f. One application of the weakest function would be the smallest step, but applying f to $g(\gamma, \beta)$ is no step, since $\lim(f) = g(0, 0)$, and fixed points have "no mercy" (axiom). We also have that $g(g(\gamma, \beta), \beta^-) = g(\gamma, \beta)$, for every $\beta^- < \beta$, by the "Limits" axiom. So the "second weakest" applicable function is $\lambda x.g(x, \beta)$. But since $g(\lambda(\bar{n}), \beta) < g(g(\bar{n}, \beta), \beta)$ for any λ such that $\lim(\lambda) < g(0, \beta)$ by the axioms "limit" and "λ-preserves the ordering", we get that applying the weakest function to the leftmost leaf is the smallest possible step. This is taken into account in the definition of the successor function, which is applied in the consequent. The successor function will, by recursion, make f applicable to any tree, by passing the function call up the left branch of the subtree until it eventually reaches a numeral; by recursively applying $s(g(\gamma, 0)) = g(s(\gamma), 0)$, we will eventually end up with $s(\bar{n}) = f(\bar{n})$ at the top of the subtree γ. So applying s to a tree is, in effect, to apply f to a numeral once. Which clearly is the smallest possible step.

(b) For the right side of the equation, clearly, adding one is the smallest next step.

(c) Since the equality holds for the antecedent by the induction hypothesis, it must hold for the succedent, since both sides have progressed by the smallest possible step.

2. Transfinite step. We need to show the following:

$$g(\lambda_n(x), \beta) = g(0, \beta) + \lambda_n(x) \rightarrow g(\lim(\lambda), \beta) = g(0, \beta) + \lim(\lambda) \quad (2.6)$$

for every $x < \lim(\lambda)$, and where $\lim(\lambda) < g(0, s(\beta))$

We are going to show that both sides of the equation progress with the smallest possible step from the antecedent to the consequent.

(a) We assume the antecedent and consider the consequent, and the left side of the equality. By Lemma 1, we have that if α is the smallest tree greater than every $\lambda_n(x)$, than $g(\alpha, \beta)$ is the smallest tree greater than $g(\lambda_n(x), \beta)$. By definition $\lim(\lambda)$ is the smallest tree greater than $\lambda_n(x)$, and so $g(\lim(\lambda), \beta)$ is the smallest tree greater than $g(\lambda_n(x), \beta)$.

(b) For the right side, we need to show that the induction hypothesis implies $g(0, \beta) + \lim(\lambda)$. Assume for contradiction, that there is a γ such that $g(0, \beta) + \lambda_n(x) < \gamma < g(0, \beta) + \lim(\lambda)$. Since $\lim(\lambda)$ by definition is the smallest ordinal greater than every tree generated by $\lambda_n(x)$, and since + preserves the ordering by lemma 3, we must have that $\gamma < g(0, \beta) + \lim(\lambda)$. And so γ must can be at most $g(0, \beta) + \lambda_n(x) \circ \delta_1 \circ \delta_2 \ldots \circ \delta_i$ for some function $\circ$ weaker than + and $\delta_i < \lim(\lambda)$. But the only operation weaker than + is successor, and $s_n(\lambda_m) = \lambda_m + n \leq \lambda_{m+n}(x) < \lim(\lambda)$. And so $g(0, \beta) + \lim(\lambda)$ is the smallest possible step after every $g(0, \beta) + \lambda_n(x)$.

(c) Thus, by the induction hypothesis, we get that

$$g(\lim(\lambda), \beta) = g(0, \beta) + \lim(\lambda)$$

since, on both sides of the equation, we have moved to the smallest next step.

3. Fixed points: in the transfinite case of $g(\lambda_n(x), \beta)$, we assumed that x was smaller than $\lim(\lambda)$ which is equal to $fix(\lambda)$. This was not really necessary (more on that below), but treating the fixed points separately will clarify the connection between polynomials and trees. So we show that

$$fix(\lambda x.g(x, \beta)) \leq \alpha \rightarrow g(\alpha, \beta) = g(0, \beta) + \alpha$$

(a) When $g(\alpha', \beta) = \alpha'$ we also get $g(\alpha, \beta) = \alpha$ for any $\alpha' < \alpha$, since fixed points "have no mercy" (axiom). So what remains is to show that we still have $g(0, \beta) + \alpha = \alpha$.

(b) Assume the base case, where $\alpha = fix(\lambda x.g(x, \beta))$. By the definition of lim, $\alpha = g(0, s(\beta))$.

(c) Since we have shown that $\lambda x.g(x, \beta) = g(0, \beta) + x$ for $x < g(0, s(\beta))$ and also that $g(0, s(\beta))$ is the *smallest* tree bigger than $\lambda x.g(x, \beta)$, we can derive that

$$\begin{aligned} g(0, s(\beta)) &= \lim(\lambda x.g(x, \beta)) \\ &= g(g(\ldots, \beta), \beta) \\ &= g(0, \beta) + g(0, \beta) + \ldots \\ &= g(0, \beta)\omega \end{aligned}$$

(d) Thus $g(0, \beta) + g(0, s(\beta)) = g(0, \beta) + g(0, \beta)\omega$

(e) And since $\alpha + \alpha\omega = \alpha\omega$ by lemma 3, $g(0, \beta) + \alpha = \alpha$, whenever $\alpha \geq g(0, s(\beta))$.

(f) Therefore, $g(\alpha, \beta) = g(0, \beta) + \alpha$ also for $\alpha \geq fix(\lambda x.g(x, \beta))$.

In the case of the fixed points, notice that we can also get away with no proof at all, by invoking Cantor Normal Form: Since $g(0, \beta) < g(0, s(\beta))$, $g(0, \beta) + g(0, s(\beta))$ would not be on CNF. So if it is in fact an ordinal, it will be equal to some other tree, for which there is another polynomial that *is* on CNF. Since we are interested in finding ordinals, in the fixed point case $g(\alpha, \beta)$ for $\alpha \geq fix(g(x, \beta))$, we could just ignore the tree in it's current form, trim away it's root, $g(x, \beta)$ and apply the formula again to the remaining α, recursively. At some point we will get to something which is not a fixed point, and not in violation of the CNF-requirement. And that would be the ordinal also corresponding to our original tree, $g(\alpha, \beta)$.

However, when we do in fact show that the formula holds for fixed points as well, we will eventually get a mapping between all the trees and all the polynomials[3], including the ones not on CNF. This makes the formula more flexible, and shows even clearer how snugly the two structures fit together.

2.4.5 Right branching is exponentiation:

$\lambda\beta.g(0, \beta) = \omega^{1+\beta}$

1. $\beta = s(\gamma)$. We start with the successors and do induction over β

   Base: $\beta = 0$. Clearly $g(0, 0) = g(0, 0)^{1+0}$

   Step: $g(0, \beta) = \omega^{1+\beta} \rightarrow g(0, s(\beta)) = \omega^{1+\beta+1}$

   (a) Since we have that $g(0, 0) = \omega$ we have that $g(0, 0)^{1+\beta} = \omega^{\beta}\omega$.

   (b) $\lim(\lambda x.g(x, \beta)) = g(0, s(\beta))$(axiom), so $g(0, s(\beta))$ is the smallest tree τ such that $\lambda_n(x) < \lambda_n(x) < \tau$ for any $x < \tau$.

   (c) Left branching is $+$, so

$$g(g(0, \beta), \beta) = g(0, \beta) + g(0, \beta)$$

[3] To be exact, if we want a mapping *all* the polynomials, we also need support for polynomials where the first term is n, such as $n + \omega$. To get this, we could adjust our formula to this: $f_n(g(\alpha, \beta)) = n + \omega^{1+\beta} + \alpha$, but since we usually don't want formulas with that preceding n, it's omitted.

Thus $g(0, s(\beta))$ is the smallest tree τ such that

$$g(0,\beta)_1 + g(0,\beta)_2 + \ldots + g(0,\beta)_n < \tau$$

for any $n \in \mathbb{N}$. This again means that $g(0, s(\beta))$ is the limit of $g(0,\beta)n$. Since ω is the smallest ordinal greater than every n, we get that $g(0, s(\beta)) = g(0,\beta)\omega$.

(d) By the induction hypothesis, $g(0,\beta) = g(0,0)^{1+\beta}$, thus

$$\begin{aligned} g(0,\beta)\omega &= g(0,0)^{1+\beta}\omega \\ &= \omega^{1+\beta+1} \end{aligned}$$

2. Transfinite step: we show the following implication, by showing that both sides of the equality increases by the smallest possible step, from the antecedent to the succedent.

$$g(0, \lambda_n(x)) = \omega^{\lambda_n(x)} \to g(0, \lim(\lambda)) = \omega^{\lim(\lambda)}$$

for every $x < \lim(\lambda) < h(0,0,0)$

(a) The left side of the equation follows directly from Lemma 2 which gives us that $g(0, \lim(\lambda))$ is the smallest tree greater than $g(0, \lambda_n(x))$.

(b) For the right side, assume for contradiction that there is some γ such that

$$\omega^{\lambda_n(x)} < \gamma < \omega^{\lim(\lambda)}$$

(c) Since exponentiation preserves the ordering by Lemma 3, we must have that $\gamma < \omega^{\lim(\lambda)}$, and at most:

$$\begin{aligned} \gamma &= \omega^{\lambda_n(x)} \times \omega^{\lambda_n(x)} \times \ldots \times \omega^{\lambda_n(x)} + \omega^{\lambda_n(x)} + \omega^{\lambda_n(x)} + \ldots + \omega^{\lambda_n(x)} \\ &= \omega^{\lambda_n(x)m} + \omega^{\lambda_n(x)}k \end{aligned}$$

for any $\lambda_n(x) < \lim(\lambda)$

(d) This is smaller than $\omega^{\lambda_n(x)m} \times \omega^{\lambda_n(x)k} = \omega^{\lambda_n(x)m+k}$ since addition is weaker than multiplication

(e) But $\lambda_n(x)m + k = \lambda_n(x)_1 + \lambda_n(x)_2 + \ldots + \lambda_n(x)_{m+k}$

(f) and this is smaller than or equal to $\lambda_{m+k}(x)$ for any λ: For $\lambda = f_n$ and $\lambda x.g(x,\beta)$, we get equality, since this is just shown to be forms of addition. For $\lambda x.g(0,x)$, we have just shown that it is something stronger than addition, so this is ok too.

(g) So we have $\gamma < \omega^{\lambda_{m+k}} < \omega^{\lim(\lambda)}$, but this is a contradiction

(h) Thus $g(0, \beta)^{\lim(\lambda)}$ must be the smallest tree greater than $g(0, \beta)^{\lambda_n(x)}$, and the implication must hold, since both sides of the equation has progressed by the smallest possible step.

3. Fixed point: again, and for the same reasons as discussed above, we treat fixed points as a separate case.

$$\beta = h(0, 0, 0) \to g(0, \beta) = g(0, 0)^{\beta}$$

First, notice that the "1+" has disappeared on the right side of the equation. This is because every fixed point is a limit, and for any limit ordinal β, $1 + \beta = \beta$.

(a) $fix(\lambda x.g(0, x)) = h(0, 0, 0)$ by "Limits" (axiom). Therefore $g(0, \beta) = \beta$ by the definition of fix.

(b) We need that $g(0, \beta) = g(0, 0)^{\beta}$.

(c) By previous proof, for every $\gamma < h(0, 0, 0)$, $g(0, \gamma) = g(0, 0)^{\gamma} = \omega^{\gamma}$, and by definition of lim, $h(0, 0, 0)$ is the smallest ordinal such that this is the case.

(d) Thus we have that

$$\begin{aligned} h(0, 0, 0) &= g(0, g(0, g(0, \ldots))) \\ &= \left.\omega^{\omega^{\omega^{\omega}}}\right\} \omega \\ &= \epsilon_0 \end{aligned}$$

(e) By the definition of ϵ_0, it is the smallest ordinal such that $\omega^{\epsilon_0} = \epsilon_0$.

(f) Therefore, $g(0, \beta) = g(0, 0)^{\beta} = \omega^{\beta} = \beta$ for $\beta = fix(\lambda x.g(0, x))$

□

2.5 Trees on Cantor Normal Form

When we defined the successor function, the aim was to make f useful for trees greater than its fixed point. We achieve this by recursively letting the function call climb the tree, along the leftmost branch, until it reaches something smaller than its fixed point. At that point it will provide real growth to the tree. This principle can be generalized for all the constructors as follows:

Definition 11 (Left-recursive λ)**.**

$$\lambda_L(g(\alpha,\beta)) = \begin{cases} \lambda(g(\alpha,\beta)) & \text{if } \lim(\lambda) > g(0,\beta) \\ g(\lambda(\alpha),\beta) & \text{if } \lim(\lambda) \leq g(0,\beta) \end{cases}$$

For any constructor function λ.

Now we have a way of effectively avoiding the fixed points of the constructors, which gives us new constructor functions that will always provide growth. Notice that this function also includes - and makes superfluous - the definition of successor.

We can now prove that we have the following hierarchy:

Lemma 5 (Constructor strength)**.**

$$s(\alpha) < \lambda_L(\alpha) < \lambda'_L(\alpha)$$

For any $\alpha \in \mathcal{T}$*, and where* $g(0,0) < \lim(\lambda) < \lim(\lambda') < h(0,0,0)$

This looks quite a bit like the hierarchy of limits, but the difference is that while our normal constructors have fixed points, at which they stop to produce something new, these constructors do not. Their recursive structure makes sure they will always be applied to some subtree where they will make a difference to the size of the tree. Notice also that we are keeping $g(0,\alpha)$ out of this ordering by limiting the fixed points to trees below $h(0,0,0)$. This is because $g(0,\alpha)$ is allready in play, since it is the function that make all the limits (except the first one), which are, in effect, the starting points of each of the constructors. So $h(0,0,0)$ is in fact the limit for all the constructors as they increase in power.

Proof. 1. For $s(\alpha)$, the effect we get is always equivalent to going from $\bar{n}$ to $f(\bar{n})$ at the top left leaf node. The difference between the two is exactly $f(0)$.

2. For any of the other constructors $\lambda x.g(x, \beta)$, we will effectively go from α to $g(\alpha, \beta)$ somewhere in the left branch, where the recursive structure ensures that $\alpha < \lim(\lambda x.g(x, \beta))$. Thus the net difference in the left branch is $g(0, \beta)$, where $\beta > 0$. This difference is bigger than $\lim(f)$, and thus applying $s(0)$ will always yield a smaller difference than any other constructor.

3. Further, if another constructor λ' had a limit $g(0, s(\delta)) > g(0, s(\beta))$, we would have $g(0, \beta) < g(0, \delta)$ since "λ preserves the ordering" (axiom 2) . And the difference between α and $\lambda(\alpha)$ would be exactly $g(0, \delta)$ in the left branch, since λ_L ensures that $\alpha < \lim(\lambda)$.

4. Since $g(0, \beta)$ and $g(0, \delta)$ are the differences of applying λ and λ' respectively to some α, and since $g(0, \beta) < g(0, \delta)$, $\lambda(\alpha) < \lambda'(\alpha)$ whenever $\lim(\lambda) < \lim(\lambda')$

□

Now we have a convenient way of constructing all the trees which actually correspond to ordinals on Cantor Normal Form, while skipping the trees that don't.

Theorem 2. *(Trees on Cantor Normal form) The trees constructed by λ_L are on Cantor Normal form.*

Proof.
We have that

$$\omega^\alpha + \omega^\alpha \omega = \omega^\alpha \omega = \omega^{\alpha+1}$$

Which corresponds directly to:

$$\lim(g(x, \beta)) = g(0, s(\beta))$$

Now, assuming $\omega \leq \alpha$, we have $g(0, \alpha) = \omega^\alpha$ (If $\alpha < \omega$ use $\omega^{1+\alpha}$). And thus we get:

$$\begin{aligned} g(g(\alpha, \beta), \gamma) &= g(0, \gamma) + g(0, \beta) + \alpha \\ &= g(0, 0)^\gamma + g(0, 0)^\beta + \alpha \\ &= \omega^\gamma + \omega^\beta + \alpha \end{aligned}$$

Where $\alpha \leq g(0, \beta) \leq g(0, \gamma)$ because of the requirement to λ_L. Thus $\gamma \leq \beta$ and $\alpha \leq \omega^\beta \leq \omega^\gamma$, which makes sure the resulting polynomial is on CNF. □

2.6 Transfinite induction over trees

We are going to show that transfinite induction is closed under $\lambda x.g(x, \alpha)$ and under $\lambda x.g(0, x)$. This will give us a very clear picture of how the trees grow, since we have to employ the complete analysis of what happens in left and right branching.

Definition 12 (Progressive property)**.** A first order formula F with one variable, x, free in F is *progressive* when the following is the case:

$$PROG(F) \equiv \forall x < \alpha.F(x) \to F(\alpha)$$

This is the standard transfinite case in transfinite induction.

Definition 13 (Transfinite induction up to some ordinal)**.** Transfinite induction up to some ordinal α for some property F is defined as follows:

$$TI(\alpha, F) \equiv PROG(F) \to \forall \beta < \alpha.F(\beta)$$

Theorem 3 (Transfinite induction is closed under left branching)**.**

$$TI(g(0, \beta), F) \to TI(g(\alpha, \beta), F)$$

for any $\alpha, \beta \in \mathcal{T}_2$, given some assumptions (†) about F and TI.

Proof.

1. Assume $PROG(F)$ and $TI(g(0, \beta), F)$. We immediately get $F(g(0, \beta))$.
2. Assume
$$\forall x < \gamma.F(g(x, \beta)) \qquad (\dagger)$$
3. Clearly γ is the smallest tree greater than every $x < \gamma$. By lemma 1 $g(\gamma, \beta)$ is the smallest tree greater than every $g(x, \beta)$ for every $x < \gamma$.
4. But then, by our assumption, we have $\forall x < g(\gamma, \beta).F(x)$. And since F is progressive, we get $F(g(\gamma, \beta))$.
5. Thus we have shown that $\forall x < \gamma.F(x, \beta) \to F(g(\gamma, \beta))$, which means that $PROG(\lambda x.F(g(x, \beta)))$
6. Now assume
$$TI(\alpha, \lambda x.F(g(x, \beta))) \qquad (\dagger)$$
We immediately get $\forall y < \alpha.F(g(y, \beta))$ by modus ponens.

7. Since we allready have $\forall x < g(0, \beta).F(x)$ we can conclude
8. $\forall x < g(\alpha, \beta).F(g(\alpha, \beta))$
9. Which means that we have $TI(g(\alpha, \beta), F)$.

□

About the assumptions (†)

We assume two things in our proof, apart from the antecedent in the theorem:

1. $\forall x < \gamma.F(g(x, \beta))$ for some γ.
2. $TI(\alpha, \lambda x.F(g(x, \beta)))$

These might look like strong assumptions, but what is important is that we are not assuming what we are trying to prove, and we are not assuming anything about the function $\lambda x.g(x, \beta)$. We are merely assuming that F holds for certain ordinals expressed in terms of this function.

Certainly, the assumptions might not hold, in which case our proof would not continue. This, however, would not be due to the function in question, but rather due to some limitation in the predicate F, or because we were trying to prove something stronger than what we can get from the progressive property. And this is exactly the limitation we want "Closed under transfinite induction" to entail.

Theorem 4.

$$TI(g(0, 0), F) \to TI(g(0, \alpha), F)$$

Proof.

1. We start by assuming $PROG(F)$ and $TI(g(0, 0), F)$, and get $F(g(0, 0))$.
2. Since transfinite induction is closed under left branching, we have $TI(g(\alpha, 0), F)$ given some acceptable assumptions (†), and for any α.
3. This means that we have $\forall x < g(\alpha, 0).F(x)$ for every α. Since we only need every α that contributes to $g(x, 0)$, we can assume $\alpha < g(0, 1)$.
4. By "limits" (axiom 3) the smallest tree larger than this is $g(0, 1)$, which means we have $\forall x < g(0, 1).F(x)$.
5. But since F is progressive, we get $F(g(0, 1))$.

6. Now assume we have $F(0, \beta)$ for some β. By the same argument we get $F(g(\alpha, \beta))$ for every $\alpha < g(0, s(\beta))$.

7. Again, by "limits" (axiom 3) , we have $\forall x < g(0, s(\beta)).F(x)$.

8. And this means that we have $TI(g(0, s(\beta)), F)$ which completes the proof for successor trees.

Now we need to show that we can get from a tree with a successor subtree on the right branch, to tree with a limit in the right branch.

1. Assume we have $TI(g(0, \lambda(x)), F)$ for every $x < \lim(\lambda)$, where the case of $\lambda = s$ is the base step. Now we have that the smallest next tree is $F(g(0, \lim(\lambda)))$, by Lemma 2.

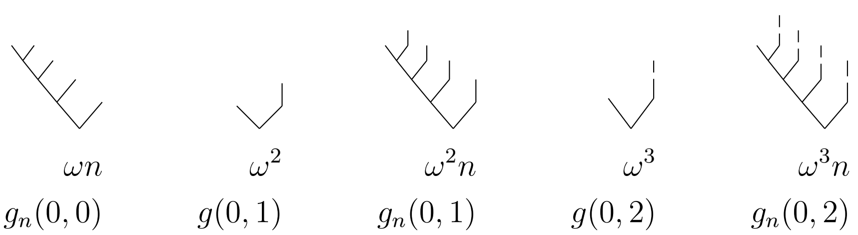
2. And then we have $\forall x < g(0, \lim(\lambda)).F(x)$

3. which means that we have $TI(g(0, \lim(\lambda)), F)$

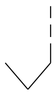
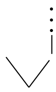
$\square$

The abstract property

In the previous proof we make a very big last step when we assume that we have $TI(g(0, \lambda(x)), F)$. I mention that $\lambda = s$ is a base step in this process, and I shall now try to give some further insight as to how we get from $\lambda = s$ to the more powerful constructors. Let's first take a look at where we can get with only $\lambda = s$ and "initial segment" (axiom 1) . Clearly, we have the following development:



Now, pay attention to what happens in the immediate right branch. We are using the fact that the smallest tree greater than $g(\lambda(x), \bar{n})$ is $g(0, s(\bar{n}))$, and this gives the unary growth:

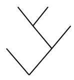
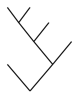
  *And a small leap...* 

ω^4 $\quad$ ω^n $\quad$ ω^ω

$g(0,3)$ $\quad$ $g(0, f_n(0))$ $\quad$ $g(0, g(0,0))$

When we prove transfinite induction for the right branch, for each successor step in the right branch, we employ the fact that we allready have exhausted the left branch. So since we have shown that we get $F(g(\alpha, 0))$ for any α, this means that we have $\forall x < g(0, s(\beta)).F(x)$ by the definition of lim. By the same definition, and since F is progressive, we can go on and get $F(g(0, s(\beta)))$. This process continues, and by lemma 2 we now have $\forall y < g(0, g(0,0)).F(y)$, and again, since F is progressive, we get $F(g(0, g(0,0)))$.

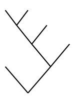
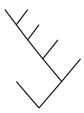
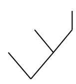
And now for the abstract property. Clearly, by applying this principle over and over, recursively, we get

ω^ω $\quad$ $\omega^{\omega 2}$ $\quad$ $\omega^{\omega 3}$

$g(0, g(0,0))$ $\quad$ $g(0, g(g(0,0),0))$ $\quad$ $g(0, g(g(g(0,0),0),0))$

And so on, since this is just applying the successor function, and then lemma 2 to get to the next successor limit. Now we need to generalize this whole process, and we do so implicitly; because the process we have just seen, actually amounts to the more complex process $g(0, \lambda x.(g(x,0)))$. And by the powerful lemma 2, we now have $\forall x < g(0, g(0,1)).F(x)$, and since F is progressive, $F(g(0, g(0,1)))$:

  *Abstract leap...* 

$\omega^{\omega 3}$ $\quad$ $\omega^{\omega 4}$ $\quad$ ω^{ω^2}

Chapter 3

Intermezzo: The towers of the emperor

An emperor wants to build a line of great towers to commemorate himself for all eternity. The first tower will be in the center of the capital, right outside the royal palace. It will be made from bricks of gold and it will be the highest tower ever built. And from the first tower, there will be a vast succession of towers, equally high, stretching all the way through the great empire.

In the farthest outskirts of the land, at the edge of the endless steppes, lies a small patch of land belonging to an old monastery, where a lonely monk resides. The monastery is inconveniently placed exactly where the emperor wants to place his last, magnificent tower, and the mere idea of being remembered as slightly less magnificent makes the emperors face red with offense. But as the emperor cannot find any publicly acceptable reason to order the demolition of the monastery, he sends a messenger to the monk, generously offering him a new piece of land, and stone and labour to build a new and better place of meditation. A few weeks later the monk comes walking into the city of the emperor, dusty and worn from the long journey. The king hears word of his arrival and has his servants arrange for food and shelter which the monk accepts. After a night of rest, the monk is granted audience to the king, and he presents his reply.

-People of my order have never been the kinds to stand in the way of kings, he says, and I shall not be the exception. But I must humbly ask for something in return

- I am sure that can be arranged, said the emperor, feeling hopeful.

- Very well, said the monk, then I must ask for your permission to take one stone, every day, from the smallest and farthest of your towers, to use in the building of our new monastery. However, I have been taught that one shall

not only take, said the monk, so I will certainly also give back in plenty, and I will commit to this in writing.
Naturally, the king was skeptical about letting someone take away from his magnificent tower, even from the least magnificent one, but he was also curious about the payment he would receive in return.
- So tell me, said the emperor, what would you suggest be the nature of this contract?
And thus the monk brought forth a modest scroll with the following text:

1. I will take one stone every day, from the least and most distant of the emperors towers, with the one exception specified below. And when I have passed away, someone else will come in my place and take one stone, and after him another, and so on.

2. For every stone I take, I will grant the emperor a contract guaranteeing the right to build any given number of new towers, into the barren land which I oversee. And until the specified number of new towers are completed, I will refrain from taking another stone. The one condition will be that the new towers will not be higher than the one from which I removed a stone, and that there will not be added to their height after their completion. This way it is certain that no tower in the barren land will be higher or more glorious than the one in the royal city.

3. I shall also ask the emperor to refrain from adding to the height of a tower from which I have taken a stone. In stead, he will be granted the right to build as many new towers as he can imagine, of the same height, as specified above.

4. To make sure I give back more stones than I take away, when ever I take one stone from a new tower - which was granted the emperor as repayment - I shall give the stone back to the emperor, on the condition that it be added to one of the original towers. It may be added to any one of these, as long as no tower farther from the center is higher than one closer. This way, what is taken away from a new tower, will equally be added to increase the height and glory of the towers the emperor first built.

The emperor was baffled by this great willingness to pay, but also a little apprehensive about the peculiar nature of the contract. And so, to clarify, the emperor asked the monk the following questions:

- You have said that when you take one stone from my tower, i get to build as many new towers as I like, of height equal to the tower from

which you removed a stone. If I only chose to build one more tower, from only two stones, is it not true then, that you must repay me these two stones before you can take another?
- It is indeed so, the monk replied.

- This means that by adding one tower of only two stones, I get back in two-fold what you have taken away. Is it then, not also so, that since I can choose to build as many new towers as I like, I can chose to get back a hundred-fold and a thousand-fold, and much much more, for which ever stone you take another away from the last tower?
- This is true, the monk replied.

- And further, said the emperor, let us assume that you take away one stone from my glorious tower, and that I chose to build a hundred new towers, each from a hundred stones. Is it not so, that if you remove one stone from the most distant of these new towers, I get to build as many towers I like, from 99 stones each? And that if you remove one stone from the most distant of these, I get to build as many new towers as I like from 98 stones each? - Yes, said the monk. This is true.

- Finally, said the emperor, when you are busy collecting stones from the new towers and return them to me, I can make sure they are added successively to my original towers. And so, since the size of your repayment can be of my choosing, I can chose to have the original towers increased to whichever height I want. Is it not so then, that should you completely take down one of the original towers, the next tower in line may have increased to such a vast height that it has never before been conceived? And so, when you have taken one stone from this next tower I can build a whole new succession of towers, each as high as this tower now is, so that the new line of towers may be longer and higher than ever before imaginable?
- Your reasoning is sound, said the monk. What you say is true.

So, to sum it all up, said the emperor, for each stone you take away, I can chose the rate of interest, and the rate of interest upon interest and so on, and all of this will be added to my original towers. These in turn will determine the height of the next new line of towers, which will be higher than all the new towers built before them. So, by signing this contract, I am guaranteed that my line of towers can be longer and greater than any man has ever dreamed of. Surely, this is a contract worthy of a king.

But before I sign, said the emperor after a moment of reflection, indulge me: why would you grant me such a bargain?

- In the long run, this will be to the benefit of the empire. Said the monk.
- I see. Said the emperor. So you too see the true value of a great symbol of the emperors greatness.

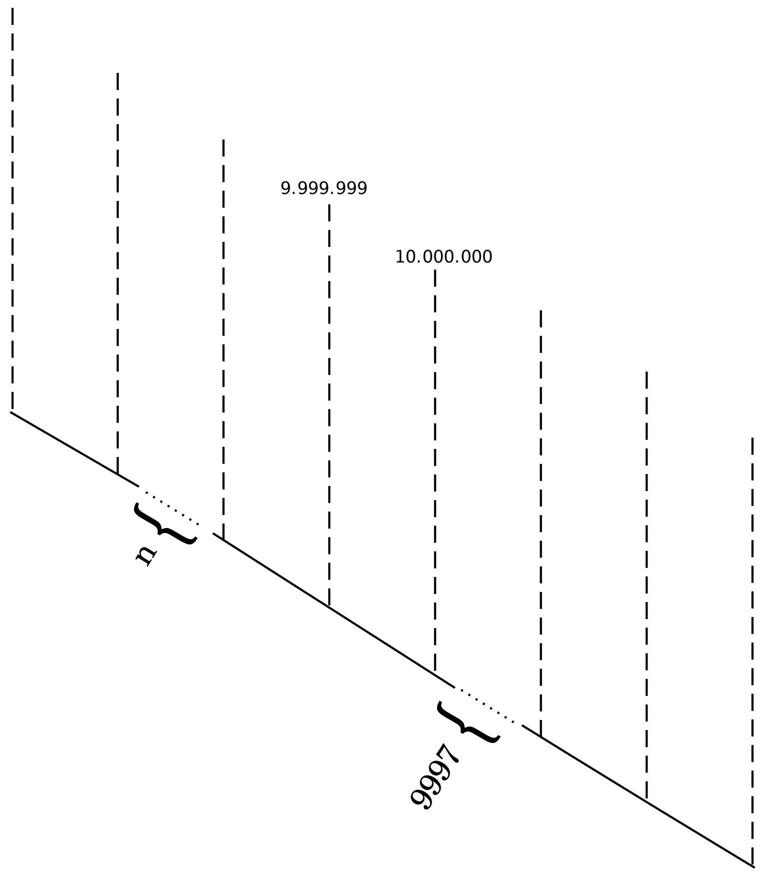
And so the emperor and the monk both signed the scroll, and a carefully written and verified copy. The emperor with his glorious golden seal, and the monk with his name.

The next day, the monk began his long journey back to the monastery, while the emperor gathered all his advisers, eager to get started with the towers. It was decided that they would start with a great line of ten thousand and one towers, each a hundred stones wide, a hundred stones deep and a thousand stones high. This way every tower would consist of ten million stones, which would give quite a potential for the first new line of towers, once the monk removed the first stone.

When all the plans were drawn, the work of securing and clearing all the property began. This was in itself a tedious process, and it took a year before the horde of workers finally came close to the the area where the monastery was. Once again the monk left his place of meditation, and went out to meet the emperors captain, who was supervising the work. The captain saw him coming, and asked with a loud voice: Hey monk, have you changed your mind? Remember, we have a written contract!

- Yes, said the monk, when he finally reached the captain. I shall honor my end of the deal. There is just one thing I would like to ask the emperor before I move, if you don't mind.
- Fair enough, said the captain, what is it?
- Tell the emperor that I am surprised. If he wants to build something to make the world rememberhim for all eternity, why would he build something that will inevitably be gone in time?
- What do you mean? Said the captain. Surely, you have made it so that these towers may continue to grow as high and far as the emperor could ever want?
- Yes, said the monk. As high as he wants. And when that height is finally reached, they will dwindle away, slowly but steadily, until there is not another stone left. I have written down the explanation of this fact, he continued, holding up a scroll, but surely, the emperor and all his advisers must know this allready.

The monk handed over the scroll to the captain, and returned to his monastery, confident that it would not be taken away.



The emperors towers
- after the first stone

3.1 What the Monk knew

The monk, of course, knew about Ordering C. Notice from the illustration, that the towers are all really a tree, and since they are ordered in height, and the tree have maximum branching of two, it must be a tree in ordering C.

Chapter 4

Trees without fixed points

I shall now give an account of the growth for trees without fixed points. Here we have Jervells ordering as a guideline.

Definition 14. (Left and right subtrees)

$$l(g(\alpha, \beta)) = \alpha$$

$$r(g(\alpha, \beta) = \beta$$

Definition 15 (Pivot point)**.**
The pivot point of a tree $\alpha = g(\gamma, \delta)$, denoted $\mathcal{P}(\alpha)$, is the subtree of α with the greatest right subtree. If there are more than one, we chose the lowest one. We have a Δ_0-representation of this function, which makes it decidable:

$$\exists \beta \preceq \alpha . \forall \gamma \preceq \alpha . (r(\gamma) < r(\beta) \vee \\ r(\gamma) = r(\beta) \wedge \gamma \preceq \beta)$$

Where $\alpha \prec \beta$ means that α is a subtree of β.

We can actually define the whole Ordering C in terms of pivot points:

Definition 16. (Ordering C ordinals)
$\mathbb{C}$ is the set of numerals, and every tree x such that x is a pivot point, and all subtrees of x are pivot points. :

$$\mathbb{C} = \{\alpha \mid \forall x \preceq \alpha . \mathcal{P}(x) = x\} \cup \{\bar{n} \mid \bar{n} = f_n(0) \text{ for any } n \in \mathbb{N}\}$$

Definition 17 (A minimal tree)**.**
A tree is *minimal* if it is on the form $g(\alpha, \beta)$, and it is the smallest tree equal to itself.

$$\forall x \prec g(\alpha, \beta) . x \neq g(\alpha, \beta)$$

Theorem 5 (Ordering C are the trees in $\mathbb{C}$)**.**
The minimal trees of Ordering C is the set $\mathbb{C}$

Proof. Under ordering C, for any tree (or subtree) $\gamma = g(\alpha, \beta)$, if it is *minimal*, the right subtree β, is smaller than the right subtree of α. This must be the case, since otherwise we would have $g(\alpha, \beta) = \alpha$, by "limits" (axiom 3) , in which case γ would not be minimal. □

Definition 18 (Pivot point application)**.** We define two general function schema's, meant to apply any constructor function to the left or right subtree of the pivot point of a tree.

$$\lambda_L(g(\alpha,\beta)) = \begin{cases} g(\lambda(\alpha),\beta) & \text{if } g(\alpha,\beta) = \mathcal{P}(g(\alpha\beta)) \\ g(\lambda_L(\alpha),\beta) & \text{otherwise} \end{cases}$$

$$\lambda_R(g(\alpha,\beta)) = \begin{cases} g(\alpha,\lambda(\beta)) & \text{if } g(\alpha,\beta) = \mathcal{P}(g(\alpha\beta)) \\ g(\lambda_R(\alpha),\beta) & \text{otherwise} \end{cases}$$

Definition 19 (Multiple applications of g)**.** When we write g_m to denote multiple applications of g, we will always mean the following:

$$g_{n+1}(\alpha,\beta) = g_n(g(\alpha,\beta)\beta)$$

The point here is that $g(\alpha, \beta)$ is a constructor function where only the left branch varies. This notation is useful when we define the axioms of ordering A, in the following.

4.1 Axioms of Ordering A

I now give a constructive definition of the ordering A. The constructors are the same as before.

1. **Initial segment:**
   $0 < f_n(0) < g(0,0)$

   for every $n \in \mathbb{N}$

2. **λ preserves the ordering:**
   $\alpha < \beta \leftrightarrow \lambda(\alpha) < \lambda(\beta)$

3. **Constructor order:**
   $\alpha < f_n(\alpha) < g_m(\alpha, \beta^-) < g_i(\alpha, \beta) < \lambda_{Lj}(\alpha) < \lambda'_R(\alpha)$

   Assuming that:

   (a) $\mathcal{P}(g(\alpha, \beta)) = \mathcal{P}(\alpha)$

   (b) $r(\mathcal{P}(\lambda'_L(\alpha))) < r(\mathcal{P}(\lambda''_R(\alpha)))$

   (c) λ and λ' are any constructors except $\lambda x.g(0, x)$.

   (d) $\beta^- < \beta$ and n, m, i, j are any natural numbers greater than 0.

4.1.1 Properties of the ordering

The axioms are intended to describe the following growth of the trees, from any starting tree α:

1. First, f is applied any number of times to α

2. Then $\lambda x.g(x, 0)$ is applied, any number of times to α

3. Then $\lambda x.g(x, 1)$ is applied, any number of times to α

4. Then $\lambda x.g(x, \beta)$ is applied, any number of times, for any β, in increasing order, that does not change the pivot point of α

5. When these possibilities are emptied, the process starts over at the head of the pivot point of α.

6. Finally, when all the constructors are applied to the left subtree of the pivot point, that does not increase the right subtree of the pivot point, the process starts over at the right subtree of the pivot point.

4.1.2 Ordering A is the Jervell Ordering

In [Jer00] on page 3, we have what we shall call the Jervell-ordering. It is defined there as follows:

A<B: The ordering to be defined

A$\leq$ ⟨B⟩: $\exists B' \in \langle B \rangle . A \leq B'$

⟨A⟩ $\leq$B: $\forall A' \in \langle A \rangle . A' < B$

⟨A⟩ < ⟨B⟩: The lexicographic ordering of the two sequences.

Definition 20 (The Jervell Ordering, $\mathcal{J}$)**.**

$$A < B \Leftrightarrow A \leq \langle B \rangle . \vee . \langle A \rangle < \langle B \rangle \wedge \langle A \rangle < B$$

Theorem 6 (Ordering A is the Jervell Ordering)**.**

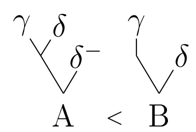
$$\alpha <_A \beta \Leftrightarrow \alpha <_J \beta$$

Proof. The proof is in two steps, one for each direction in the equivalence.

1. $\alpha <_A \beta \Rightarrow \alpha <_J \beta$; all the inequalities derivable in $\mathcal{J}$, can be derived in Ordering A. To show this implication, I will derive all the axioms of Ordering A from $\mathcal{J}$, beginning with the crucial Constructor order (axiom 3). This axiom consists of four inequalities, where the first one is derivable from axiom 2. Hence the two first, and most important steps in the proof, are the two last inequalities. I will abbreviate $g_i(\alpha, \beta)$ to $\lambda_n(\alpha)$, for readability.

   (a) $\lambda_n(\alpha) < f_L(\alpha)$ where $\mathcal{P}(\lambda(\alpha)) = \mathcal{P}(\alpha)$: the limit of successive applications of any function[1] to some α, is adding one to the head of its pivot point. This is the case as long as the function itself does not change the pivot point of α. The proof is by induction over n:

   For a base case, let $n = 1$. We need to show that



[1]The axiom says $\lambda_L(\alpha)$, not $f_L(\alpha)$, but f is the weakest function and hence the critical case to consider.

For any $\delta^- < \delta$

We make sure one of the two requirements of $\mathcal{J}$ordering are fulfilled:

i. $A \leq \langle B \rangle$ is clearly not the case: B has two subtrees. The right one, δ is a subtree of A, so it must be smaller. The left one is $f(\gamma)$, but $g(\gamma, \delta)$ is a subtree of α, and comparing them lexicographically gives us that the latter is bigger. Since this requirement does not hold, we need the two next ones to hold to fulfill the disjunction.

ii. $\langle A \rangle < B$. The two subtrees of A must be smaller than B. We have $\delta^- < \delta$ by definition, so $\delta^- < g(x, \delta)$ since $\delta < g(x, \delta)$ regardless of x. For the next one, we compare $g(\gamma, \delta)$ to $g(f(\gamma), \delta)$ lexicographically. The first one is smaller, since $\gamma < f(\gamma)$ by "Initial segment" (axiom 1).

iii. For the last requirement, clearly $\langle A \rangle < \langle B \rangle$. They differ at the right subtree, and since $\delta^- < \delta$ by definition, A is the smaller tree.

Now for the inductive step. We assume

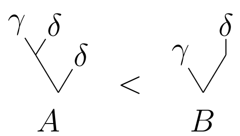
$$A' = \lambda_n(g(\gamma, \delta)) < g(f(\gamma), \delta) = B$$

We are still also assuming that $\mathcal{P}(\lambda(\alpha)) = \mathcal{P}(\alpha)$, and so applying λ once more would give $A = g(\lambda_n(g(\gamma, \delta)), \delta^-)$ for some $\delta^- < \delta$. Now the two last requirements in $\mathcal{J}$must hold:

i. $\lambda_n(g(\gamma, \delta)) < B$ by the induction hypothesis and $\delta^- < \delta$ by definition. Thus both subtrees of A are smaller than B.

ii. We now compare the trees lexicographically, and they differ at the right subtree; since $\delta^- < \delta$ we can conclude that $A < B$ and we are done.

(b) $\lambda_L n(\alpha) < f_R(\alpha)$: the limit of applying any function successively to the head of the pivot point, is adding one to its tail. This, of course, under the condition that the function itself does not increase the head of the pivot point. Again, the proof is by induction over n.
For a base case, let $n = 1$. We need to show that



The two last requirements in $\mathcal{J}$holds:

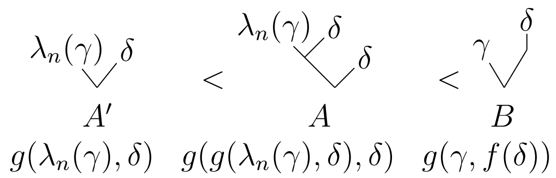
i. $\langle A \rangle < B$: We start with the right subtree of A, and δ is also a subtree of B, so that's ok. For the left subtree, $g(\gamma, \delta) < g(f(\gamma), \delta)$ since $\gamma < f(\gamma)$ by "Initial segment" (axiom 1).

ii. $\langle A \rangle < \langle B \rangle$. The trees differ in the right subtree, and since $\delta < f(\delta)$ we get $A < B$ and we are done.

For the inductive step, assume

$$A' = g(\lambda_n(\gamma), \delta) < g(\gamma, f(\delta)) = B$$

We are also still assuming that $tl(\mathcal{P}(\lambda(\alpha))) = tl(\mathcal{P}(\alpha))$, meaning that the constructor does not change the right subtree of the pivot point. Thus we know that $\lambda(\alpha) \leq g(\alpha, \delta)$. So, assume the worst for λ, and one more application gives the following to be shown:



The first inequality is obvious, since the left operand is a subtree of the right. The last one is the important one.

i. $\langle A \rangle < B$. For the right subtree, $\delta < f(\delta)$ by "initial segment" (axiom 1). For the left subtree, this is the induction hypothesis.

ii. $\langle A \rangle < \langle B \rangle$. The trees differ in the right subtree, and again, since $\delta < f(\delta)$ we get $A < B$ and we are done.

(c) The remaining axioms are straight forward:

i. "Initial segment" (axiom 1) is taken for granted in $\mathcal{J}$

ii. "λ preserves the ordering": it suffices to notice that whenever we apply the same function λ to some α, β where $\alpha < \beta$, α and β remain subtrees. The requirement $\langle \alpha \rangle < \langle \beta \rangle$ makes sure the difference in subtrees still count.

2. $\alpha <_J \beta \Rightarrow \alpha <_A \beta$: We still need to show the other direction of the equivalence. We assume $A <_A B$ and show that this implies that $\mathcal{J}$holds:

(a) $A \leq \langle B \rangle$. Clearly, if A is smaller than some immediate subtree of B it will be smaller than B. This means $B = \lambda(\beta)$ and $A < \beta$, for some λ, β, and we get directly from "λ preserves the ordering" (axiom 2) that this implies $\lambda(A) < \lambda(\beta)$. Since $A < \lambda(A)$ by "Constructor order" (axiom 3), we get $A < \lambda(A) < \lambda(\beta) = B$

(b) $\langle A \rangle < B$. This follows from the previous, since the negation of the statement would mean $B \leq \langle A \rangle$, in which case we would have $B < A$.

(c) $\langle A \rangle < \langle B \rangle$. When we have gotten to this requirement we can assume that the first one did not hold, or that we are allready in the subtree of a B' with the greatest chance of surviving this test. That's the pivot point of B'. We can also assume that we have the second requirement, making all the subtrees of A smaller than B, but we might just as well do them both at the same time. This is done by comparing pivot points; we take the subtree of both A and B with the greatest chance of "surviving" this test. Now "Constructor order" (axiom 3) gives us that the right subtree takes precedence over anything in the left subtree, and "initial segment" (axiom 1) gives us that binary branching takes precedence over unary branching. If both left and right subtrees are equal, it means that either $A \prec B$ or there is some place closer to the root, where the test would fail, according to $\mathcal{J}$. This is then covered by the first inequality in "Constructor order" (axiom 3).

□

4.2 Mapping Ordering A to the ordinals

Theorem 7 (Trees under Ordering A are ordinals). *For any tree τ under ordering A, there is an ordinal on CNF given by the following formula for $\mathcal{P}(\tau)$:*

$$\begin{array}{c}\beta \quad \gamma \\ \diagdown\!\diagup \\ \alpha\end{array} = \omega^{\omega^{\gamma}}(1+\beta)+\alpha$$

Where α is a tree where with upper left leaf node $\lambda(0)$, and where τ has α as root, but $\lambda(g(\beta,\gamma))$ as upper left leaf node.

Proof.

1. $\lambda(g(\beta,\gamma)) = g(\beta,\gamma) + \lambda(0)$: applying any function to the pivot point of a tree is addition, so long as applying the function changes the pivot

point of the tree. The last requirement is allready assumed in the formula, since the function has allready been applied, so to speak, and the formula describes the pivot point of τ after this application. As a base case, I will show that the proposition holds for f and $\lambda x.g(x,0)$, and do induction from there.

(a) First let's establish that $f(\alpha) = s(\alpha)$: For the initial segment, this is explicit in the axiom. Further, we have that "λ preserves the ordering" (axiom 2), and thus any other function will be stronger. Thus, applying f will always be the smallest possible step, which is in essence the successor.

(b) Further, we get that $f_n(\alpha) = \alpha + n$, by simple induction over n.

(c) Now, since $\lambda x.g(x,0)$ is the second weakest constructor by "Initial segment" (axiom 1), $g(g(\beta,\gamma),0)$ must be next after $f_n(g(\alpha,\beta))$, and since $g(0,0)$ is the smallest tree greater than every $f_n(0)$, $g(\beta,\gamma) + g(0,0)$ is next after $g(\beta,\gamma) + f_n(0)$.

(d) And then our proposition holds up to and including $\lambda x.g(x,0)$, since in both sides of the equation, we have progressed by the smallest possible steps.

(e) Now for any constructor where the right subtree is a successor tree: assume $g(g(\beta',\gamma'),\delta) = g(\beta',\gamma') + g(0,\delta)$.

(f) We get $g_n(g(\beta,\gamma),\delta) = g(\beta,\gamma) + g_n(0,\delta)$ directly, since β' was arbitrary and could be $g_{n-1}(g(\beta'',\gamma),\delta)$.

(g) But now, the smallest possible next step for both sides of the equation is to increase δ to $f(\delta)$ since $f(\delta) = s(\delta)$, which is always the smallest next step..

(h) And so $g(g(\beta,\gamma),f(\delta)) = g(\beta,\gamma) + g(0,f(\delta))$.

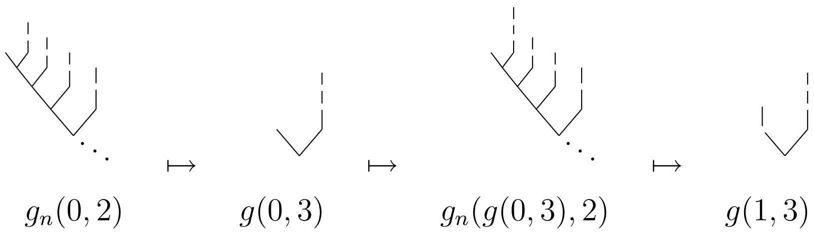
(i) For constructors with limit trees on the right, this follows from lemma 2, the same way as in ordering C. The proof of the lemma uses only the definition of lim and "λ preserves the ordering", so it is the same for ordering A.

2. $g(\beta,\gamma) = g(0,\gamma) \times (1+\beta)$: The left subtree of the pivot point is a multiplier.

(a) The case where $\beta = 0$ or $\gamma = 0$ are trivial. Since they are not so informative, so we also do the case $\beta = \gamma = 1$: so we need to show that $g(1,1) = g(0,1) \times 2$. This follows directly from "Constructor order" (axiom 3). First, notice that $g(0,1)$ is the limit of $g_n(0,0)$

by the last inequality. But now we can apply the exact same functions again, to $g(0,1)$ and get $g_n(g(0,1),0)$. But this time the pivot point does not change, and since the requirement that $\mathcal{P}(g(0,1)) = \mathcal{P}(0)$ does not hold here, the next step is not to apply the next constructor, but to apply the smallest constructor f to the head of the pivot point. Thus the limit of this new process is $g(1,1)$, and so, since the process $g_n(x,0)$ for every $n \in \mathbb{N}$, first produces $g(0,1)$ and then produces $g(1,1)$ when applied once more, $g(1,1) = g(0,1) \times 2$.

(b) For a general case of successor trees, we use the generalization of the same principle: For any $\gamma = s(\gamma^-)$, we always have that $g(0,\gamma)$ is the limit of successive applications of $\lambda x.g(0,\gamma^-)$ to 0, and that when we repeat this process, applying it to the result of the first run, it does not change the pivot point any more, since that was just increased by one:



Where the right branch is arbitrary, but chosen to be 2 and 3 for practical reasons.

(c) So assume $g(\beta,\gamma) = g(0,\gamma) \times (1+\beta)$ for some successor tree $\beta = s(\beta^-)$, and we need to show $g(f(\beta),\gamma) = g(0,\gamma) \times (1 + f(\beta))$

(d) By "Constructor order" (axiom 3) $g(f(\beta),\gamma)$ is the limit of $\lambda_n(g(\beta,\gamma))$ for any λ such that $\mathcal{P}(\lambda(\alpha)) = \mathcal{P}(\alpha)$.

(e) But by the same axiom, $g(\beta,\gamma)$ is the limit of $\lambda_n(\beta^-,\gamma)$. And thus, adding one to β happens when we apply λ_n once more, to the result of the previous run. And thus, given the induction hypothesis, $g(f(\beta),\gamma) = g(0,\gamma) \times (1 + f(\beta))$.

(f) For the transfinite case, where β is a limit ordinal, as always, this is directly by Lemma 2 and the facts about ordinals employed in the proof for Ordering C. If β is the smallest tree greater than every $\beta^- < \beta$, then $g(\beta,\gamma)$ is the smallest tree greater than every $g(\beta^-,\gamma)$. And in ordinal arithmetic, if β is the smallest ordinal greater than every $\beta^- < \beta$, then $\alpha \times \beta$ is the smallest ordinal greater than $\alpha \times \beta^-$. And thus, by the induction hypothesis,

$g(\beta, \gamma) = g(0, \gamma) \times \beta$ since both sides of the equality have increased by the smallest possible step.

(g) Note: the "1+" part of $1 + \beta$ is there to make sure $g(0, \alpha)$ does not turn out to be $g(0, \alpha) \times 0 = 0$. This fact comes out clearly in the first part of the proof, and when β is a limit tree it can be omitted, since $1 + \beta = \beta$ for any $\beta \geq \omega$.

3. $g(0, \gamma) = \omega^{\omega^\gamma}$: Right branching is a strong form of $\omega-$exponentiation.

(a) The base case of $\gamma = 0$ is clear: $g(0, 0) = \omega$, and so $g(0, 0) = \omega^{\omega^0} = \omega^1 = \omega$. In order to make the principle clearer, let's also do $\gamma = 1$. We have shown that left branching is a multiplier, and since $g(0, 0) = \omega$, then $g_n(0, 0) = \omega_1 \times \omega_2 \times \ldots \times \omega_n = \omega^n$. Since $\lim(\lambda x.g(x, 0) = g(0, 1))$ by "Constructor order" (axiom 3) we get that $g(0, 1) = \omega^\omega = \omega^{\omega^1}$.

(b) For any tree γ we have the same principle; since left branching is a multiplier and $\lim(\lambda x.g(x, \gamma) = g(0, f(\gamma))$ by "Constructor order" (axiom 3), and assuming the induction hypothesis, we get

$$\begin{aligned} g(\beta, f(\gamma)) &= g(0, \gamma)_1 \times g(0, \gamma)_2 \times g(0, \gamma) \times \ldots \\ &= \omega^{\omega^\gamma} \times \omega^{\omega^\gamma} \times \omega^{\omega^\gamma} \times \ldots \text{by ind.hyp.} \\ &= \omega^{\omega^\gamma + \omega^\gamma + \omega^\gamma + \ldots} \\ &= \omega^{\omega^\gamma \omega} \\ &= \omega^{\omega^{\gamma+1}} \end{aligned}$$

and thus the equality holds for any successor tree $f(\gamma)$.

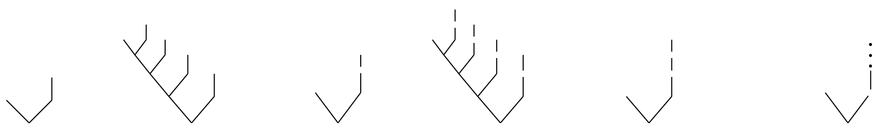
(c) For the transfinite case, this is as always by Lemma 2. If γ is the smallest tree greater than every γ^-, then $g(0, \gamma)$ is the smallest tree greater than every tree $g(0, \gamma^-)$. And in ordinal arithmetic, ω^γ is the smallest ordinal greater than ω^{γ^-} for every $\gamma^- < \gamma$, as shown in the proof for Ordering C. And thus, by the induction hypothesis, $g(0, \gamma) = g(0, 0)^{g(0,0)^\gamma} = \omega^{\omega^\gamma}$, since on both sides of the equality we have moved to the smallest possible next step.

□

4.3 How the trees grow

I shall now give some examples and try to explain how the trees grow.

n	ω	$\omega + n$	ωn	ω^2	ω^n
$f_n(0)$	$g(0,0)$	$f_n(g(0,0))$	$g(f_n(0),0)$	$g(\omega,0)$	$g_n(0,0)$



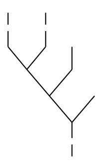
ω^ω	$\omega^{\omega n}$	ω^{ω^2}	$\omega^{\omega^2 n}$	ω^{ω^3}	ω^{ω^n}
$g(0,1)$	$g_n(0,1)$	$g(0,2)$	$g_n(0,2)$	$g(0,3)$	$g(0,f_n(0))$

Since we have that

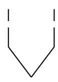
$$\omega^\omega \omega^\omega \dots \omega^\omega = \omega^{\omega n}$$

ω^{ω^ω}	$\omega^{\omega^{\omega^\omega}}$	$\dots$	ϵ_0
$g(0,\omega)$	$g(0,\omega^\omega)$	$g(0,\omega^{\omega^\omega})$	$h(0,0,0)$

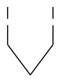
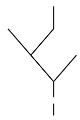
Now, let's analyze some ugly tree α in ordering A:

$\alpha =$ 

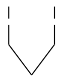
We start by identifying the pivot point, $\mathcal{P}(\alpha)$, which in this case is the highest binary split.

$\mathcal{P}(\alpha) =$ 

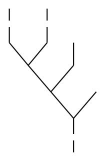
Now we can detach the pivot point and identify the first term of the polynomial:

$\alpha =$  $+$ 

and we keep identifying pivot points further down to get all the terms in the polynomial:

$\alpha =$  $+$  $+$  $+$

And finally, we apply the formula in theorem 7 and get the ordinal.

 $= \omega^{\omega^2}3 + \omega^{\omega} + \omega + 2$

Chapter 5

Discussion

5.1 Ordering B

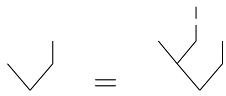
I have given two orderings, C, and A, but there is also an ordering B, or actually a whole class of orderings which can be defined as follows:

- B_1 is the ordering where f does not have a fixed point, but the other constructors do.
- B_2 is the ordering where $\lambda x.g(x, \beta)$ do not have fixed points, but f does.
- B_β for $\beta > 2$ is any ordering where some of the constructors have fixed points and some don't.

The reason for the naming of the orderings is that they were discovered in that order; It started with the Jervell Ordering A, which I wanted to map to the Ordinals. The formula describing this ordering is highly recursive, and it soon gets very hard to check by hand which is the bigger of two trees. So I used hypothetical-deductive method and set up different hypothesis about how the trees grow, and for pairs of trees I made predictions about which would be smaller. I checked my predictions against the Jervell-formula, and found out if I was correct. One of the hypothesis seemed promising, and I found that it mapped nicely to the ordinals like so:

$$\underset{n}{\overset{\alpha \quad 1+\beta}{\bigvee}} = \omega^{\omega+\beta} + \omega\alpha + n$$

Which is only defined from trees greater than $g(0,1)$, and ordinals greater than ω^ω, since trees the ordering below this point seemed pretty clear. It soon turned out that by this ordering, we would get that



Since the first one translates to ω^ω and the second to $\omega^\omega + \omega^{\omega+1}$, and

$$\begin{aligned}\omega^\omega + \omega^{\omega+1} &= \omega^\omega + \omega^\omega\omega \\ &= \omega^\omega + \omega^\omega + \omega^\omega + \omega^\omega + \dots \\ &= \omega^\omega\omega \\ &= \omega^{\omega+1}\end{aligned}$$

This, of course, is an example of a fixed point in the constructors, and in this case $g(0,\beta)$ for any β all have fixed points, which makes this ordering B_1. Since these "hybrid orderings" are so many, it seemed natural to concentrate on the ordering I define as ordering C, as a counterpoint to the Jervell Ordering. They do exist, however, and they could be interesting to develop further, especially, maybe, in light of the last section which deals with infinite trees.

5.2 The smallest possible step

When I map the orderings to the ordinals, I rely heavily on the notion of "The smallest possible step" and that if we, on both sides of an equality, increase the operand with the smallest possible step, the equality still holds. First I use it to show that if $f(\bar{\alpha})$ is the smallest possible tree that can be constructed from $\bar{\alpha}$, and if $s(\alpha)$ is the smallest possible ordinal that can be constructed from α, then $f(\bar{\alpha})$ maps to $s(\alpha)$. Granted, of course, the induction hypothesis that $\bar{\alpha}$ maps to α. Secondly, I use it in the transfinite cases; if $\bar{\alpha}$ is the smallest tree greater than every tree $\lambda(x)$, and if α is the smallest ordinal greater than every α^-, then $\bar{\alpha}$ maps to α. Again, granted the induction hypothesis, which is that $\lambda(x)$ maps to α^- for any $\lambda(x) < \bar{\alpha}$ and every $\alpha^- < \alpha$.. The assumption is intuitively ok, but if it were to be employed in a formal proof calculus, it would have to be axiomatized. I have chosen not to do this, for the same reason that I have chosen not to formally define the language I use; I wanted to focus on understanding how the trees grow, precisely describing this growth axiomatically, and proving that there is a mapping to the ordinals. The theory is created as we go, and in a "Worst case scenario", should I end up employing all the strength of axiomatic set theory, that would have been fine. Formalizing the theory further would surely be interesting, and doing this might be a candidate for future projects.

5.3 Nice properties of Ordering C

As we saw in "The Towers of the Emperor", the whole process involved could be described directly in terms of a tree of ordering C. This, in turn, can be translated directly to a polynomial on CNF by the provided formula. When we look at the process of the towers, directly as a tree, proving that the growth must terminate, is simply a matter of removing subtrees. The exact same would hold, if we increased the complexity of the sequence to involve the whole process involved in building ϵ_0. In this case, the sequence would not be a long left branch with all unary right branches, but a long left branch, where all the right branches could be arbitrary high branches with arbitrary many splits. One can imagine this process as a next generation tower challenge, a millennium later, in space. Now one could build towers at an angle, out from the sides of other towers, as long as they remained ordered along the left branch, decreasing in size from the root and out.

This would correspond directly to a goodstein-sequence[1], and again, proving termination would be a matter of manipulating trees. And since the trees clearly exists, this might make the proof easier to swallow for those skeptical of infinite ordinals.

5.4 Ending the tour: The effect of fixed points

Since we now have two formulas describing the complexity of trees in terms of ordinals, we can measure exactly how much it matters that we "waste trees" by having fixed points for constructors. Let's compare the formulas: we have

$$\omega^{1+\gamma} + \beta \quad \text{versus} \quad \omega^{\omega^{\gamma}}(1+\beta) + \alpha$$

for Ordering C, and Ordering A, respectively, where the relevant difference is ω^{γ} compared to $\omega^{\omega^{\gamma}}$. Clearly the second formula describes a much more complex process, but compared to ϵ_0, which is an infinitely high tower of $\omega-$exponents, a difference in height of one $\omega-$exponent between two trees, which is essentially what we have, is insignificant. It is, however, worth

[1]See [Kir82] for a definition, proof of termination, and proof that termination cannot be shown in PA

thinking about the fact that we have two different processes, one provably slower than the other, that both reaches ϵ_0 at the same time. And this is due to the comparison we just did; because as the tower of exponents grow toward infinity, a lead of one gets reduced to nothing and vanishes, just like for $m + n$, $\omega + \omega n$ and $\omega \omega^n$ where the first term vanishes when n reaches ω.

But let us stop for a moment right there, because what we just found out puts ϵ_0 in perspective. Remember in "The Towers of the Emperor", and in the anecdote about the botanist, how much complexity there is just below ω^ω. Compare all that complexity to ω, which is the process of just adding one, and this difference is exactly one ω exponent.

Let's go through that again. Think about the process of just adding one, forever. Jump from there, to the whole process involved in the story of the monk: he removes one stone from a tower, just to see an arbitrary amount of new towers emerge. He keeps removing stones from those new towers, just to see arbitrary many new towers, just one stone lower, emerge, and so on. Eventually all the stones from all the new towers are removed, and added back to the original towers, which then grow to arbitrary new heights. All this, just to have the monk remove one more stone, and another new range of towers emerge, this time arbitrary much higher than the first time. This process has a complexity that can be described by the clever version of colored numbers in the anecdote of the botanist: infinitely many the aquamarine numbers, stuffed in between each of infinitely many blue numbers, stuffed in between each of infinitely many cornflower blue numbers, and, and so on, for *any natural number of colors.* Both of these processes can again be described by trees smaller than ω^ω. So going from just increasing a row of dots by one, which is ω, to all of the mentioned, which is ω^ω, is an almost incomprehensible jump in complexity. And this is exactly what we *gain* when we remove fixed points from the ordering, and, what *becomes completely insignificant* as the complexity keeps growing towards ϵ_0.

Just for fun: The continuum hypothesis for trees

Let's define a constructor for "Dense trees", which will be the trees that have equal left and right subtrees:

Definition 21 (Dense trees)**.**

$$\tilde{g}_1(\alpha, \alpha) = g(\alpha, \alpha)$$
$$\tilde{g}_{n+1}(\alpha, \alpha) = g(\tilde{g}_n(\alpha, \alpha), \tilde{g}_n(\alpha, \alpha))$$

Now assume that we added an axiom to our theory of trees:

Definition 22 (Infinite binary tree)**.**

$$\tilde{g}_n(0, 0) \prec \tau_\omega$$

For every $n \in \mathbb{N}$.

Clearly τ_ω has height ω, and width 2^ω, where the exponentiation *does not* have finite support. So at the top of the tree, the leaf nodes are just as densely packed as the real numbers. But we also know that any smaller dense tree will have height n and width 2^n by the definition of ω.

A tree with ω number of leaf nodes might be constructed if we could find a way to let the sequence $g(f(0), f(0))$, $g(f(g(f(0), f(0))), f(g(f(0), f(0))))$, continue and then making sure, for each step that the height doubles as well as the width. But that's a little tricky.

There is a subtle and fascinating connection here, between ω and 2^ω, that can be summed up in one tree. We leave finding that tree as an exercise to the reader.

Exercise 1. *Working within the theories discussed in this paper, and allowing recursive constructs like Definition 21 and 22, construct a tree with between ω and 2^ω number of leaf nodes.*

Chapter 6

Summary and future work

To some extent I have been writing about things that are allready well established concepts and truths within the field of mathematical logic. The goal with this, however, has been to put my own work into a wider perspective, and to motivate it properly. The axiomatic ordering that I call Ordering C is new, as far as I know, and so are the theorems I prove about it. Ordering A is, as I prove in Theorem 6, the same ordering as given in [Jer00], but until now there has only been estimates about which trees correspond to which ordinals, for a very few key ordinals. I have given a new axiomatization of the ordering, with emphasis on how the trees grow. I have proven that this axiomatization does indeed describe the same ordering as in [Jer00] for the trees of branching < 3, and given a simple formula that for any tree, gives its respective ordinal. I have also provided such a formula for Ordering C, and this gives us an exact measure of how fixed points inhibit the complexity in growth of trees.

Herman Ruge Jervell introduced me to the idea of using trees to construct ordinals, and this is a central part of his work. He gave the original ordering that I have built on, proved that it was a well-ordering and showed how the initial segment maps to the ordinals[1]. He pointed out the interesting fact that his constructor functions do not have fixed points, which makes every tree represent a unique ordinal. Trying to make sense of the highly recursive and compact formula that defines the Jervell ordering is what made me toy around with different growth functions, including ones with fixed points. This resulted in ordering B, which is exactly like ordering C, except that f does not have a fixed point. At some point I decided to find out what happened if *all* the constructors had fixed points, and a shortly there after I found the mapping formula and realized that this too would grow towards

[1] Again, I will refer to [Jer00], but he also has several other papers and projects where he deals with this in many ways on his webpage: `http://folk.uio.no/herman`.

ϵ_0. Since this gives us a counterpart to the Jervell ordering, with respect to fixed points, it seemed like a more interesting ordering to focus on than the hybrid ordering B.

I think it's important at this point to emphasize the fact that the Jervell-ordering is much more general, since it deals with trees of *any* branching, whereas my theories deal with trees of branching less than three. However, since ϵ_0 marks the limit of ordinal exponentiation for exponentiation with finite support, a formula mapping trees of higher branching to the ordinals above ϵ_0 would, if it exists at all, have to involve something more than ω and exponentiation with finite support. And so, such a formula would not look like the ones given in this paper, but rather it might involve ϵ_0 as a constant as well as ω. In any case, expanding ordering A and ordering C to involve constructors of higher arity seems to be a project worthwhile; it might shed some more light on the processes involved in the higher ordinals, for example the debated [2] ordinal Γ_0.

It would also be interesting to look closer at what happens when we include infinite trees. When we get the cardinality $\aleph_1$ just by letting a dense tree grow to height ω, we are in fact deriving the notion of the real line from Euclid's 10th proposition, stating that "Any straight line can be bisected"[Euc56]. Because this is what happens in the infinite dense tree, it is bisection *ad infinitum*. This is also exactly what happens in Conways definition of the numbers [Con76] (which is where I got the idea), where the real numbers are "born on day ω". He attributes this construction to Dedekind and Eudoxus, which makes this a very old notion, and so it strikes me as a "about time" that a theory for the reals based on trees should be developed. One could argue that the process employed in letting the dense tree grow infinitely high is just like a power set for subtrees, but we have not introduced anything like the power set axiom as is done in ZFC[3], we have just used the same principle employed in defining ω, which is simple induction.

Last, but not least: we have that Ordering C can be mapped directly to Goodstein-sequences, as mentioned in 5.3. We also have that Kriby-Paris Theorem[Kir82], which states that the termination of Goodstein-sequences cannot be shown in PA, leads to leads to Gentzens Theorem [Gen69]. In turn Gentzens Theorem, as mentioned in the introduction, identifies the Gödel sentence for PA as transfinite induction up to ϵ_0. Finally, since coding syntax comes very natural in theories for trees, we can make Gödel sentences much

[2]This is the Feferman-Shütte ordinal, which for long has been referred to as the limit of "predicative reasoning". This claim has recently received strong criticism, in [Wea09] among other places.

[3]Zermelo Fraenkel with the axiom of Choice. See [Kun80] for a comprehensive introduction

easier in Ordering C than in any classical theory of arithmetic. These facts leads me to believe that Ordering C might be a theory in which one could further clarify the connection between Gödel sentences and ϵ_0.

Bibliography

[Con76] J.H. Conway. *On Numbers and Games.* Academic Press, London, 1976.

[Euc56] Euclid. *The Thirteen books of The Elements*, volume 1. Dover, New York, 1956.

[Gen69] Gerhard Gentzen. The consistency of elementary number theory. In *The Collected Papers of Gerhard Gentzen*, pages 132–214. North-Holland Publishing Co., Amsterdam, 1969.

[Jer00] Herman Ruge Jervell. How to Well Order Finite Trees. *Institut Mittag-Leffler Preprint series: Mathematical Logic*, 2000.

[Jer01] Herman Ruge Jervell. *Logikk og beregnbarhet.* Unipub forlag, Oslo, Norway, 2001.

[Jer04] Herman Ruge Jervell. *Formal Logic, a guided tour.* Unipub forlag - Oslo Academic Press, Oslo, Norway, 2004.

[Kir82] Jeff Kirby, Nick Paris. Accessible independence results for peano arithmetic. `http://reference.kfupm.edu.sa/content/a/c/accessible_independence_results_for_pean_59864.pdf`, 1982.

[Kun80] Kenneth Kunen. *Set Theory*, volume 102 of *Studies in logic and the foundation of mathematics.* Elsevier B.V, Amsterdam, 1980.

[Lea00] Christopher C. Leary. *A Friendly Introduction to Mathematical Logic.* Prentice Hall, Upper Saddle River, New Jersey, 2000.

[Wea09] Nik Weaver. Predicativity beyond Gamma_0. arXiv:math/0509244v3 [math.LO], 2009.

Revisions made after deadline

During the official presentation of this thesis, it was pointed out that some important definitions were not stated explicitly in the text, or worse, left out completely. This did not ultimately affect the grade the paper was given, but it felt important to put them in there before publication. On the other hand, this is an electronic publication of a Master Thesis, produced during a strictly limited period of time, and thus it should not be changed into something essentially different. After a fruitful dialogue with the Department of Informatics, it was agreed that the following changes do not alter the content of the work in any essential way, except making it more accessible to others, and were thus considered acceptable:

1. Definition 12 of the progressive property $PROG$ was added on page 30
2. Definition of the subtree operator, $\prec$, was included in Definition 15, on page 39
3. The formula defining a Pivot point, in definition 15 on page 39, was changed to reflect the text (which was right all along). In the original version the formula made no sense.
4. Definition of the Jervell Ordering was included on the top of page 42
5. References to the mentioned definitions have been updated where necessary
6. The headings of the two sections regarding the axioms of orderings A and C have been changed from "Ordering A" and "Ordering C" to "The axioms of ordering A" and "The axioms of ordering C".
7. Some typos might have been corrected

Should there for any reason be a need for the original manuscript, please contact the author, the Department of Informatics, or the library at the Faculty of Mathematics and Natural Sciences (MatNat), which all have copies.