Gyesik Lee

Phase Transitions in Axiomatic Thought

2005

Mathematik

Phase Transitions in Axiomatic Thought

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften im Fachbereich
Mathematik und Informatik
der Mathematisch-Naturwissenschaftlichen Fakultät
der Westfälischen Wilhelms-Universität Münster

vorgelegt von

Gyesik Lee

aus Kyunggi-Do, Südkorea

– 2005 –

Dekan:	Prof. Dr. K. Hinrichs
Erster Gutachter:	Prof. Dr. A. Weiermann
Zweiter Gutachter:	Prof. Dr. R. D. Schindler
Tag der mündlichen Prüfung:	29. 04. 2005
Tag der Promotion:	13. 07. 2005

To my family

and

all my Ourtown friends

Abstract

An aspect of the thesis is to investigate well-known ordinal notation systems for PA. It will be shown that the so-called phase transition phenomenon can be observed, i.e., there are thresholds between provability and unprovability. This investigation leads to a comparison of the ordinal notation systems.

The thesis gives also a guide how one can generally establish such phase transitions in every logic system which is strong enough in the sense of Gödel. We shall see that Friedman style miniaturizations play the central role.

Another point of the thesis is the parametrized version of the Kanamori-McAloon principle. This variants of the finite Ramsey theorem is equivalent to the Paris-Harrington principle. It will be shown that phase transitions occur with respect to the provability of the Kanamori-McAloon principle as the parameter function varies.

Contents

Chapter 1

Introduction

In the 1930s, Gödel's work showed that the proposal for the foundation of classical mathematics known as Hilbert's Program cannot be carried out. However, starting with Gentzen's consistency proof of Peano arithmetic [21], work on revised Hilbert Programs have been central to the development of proof theory: the question of finding consistency proofs is still of value, since the methods used in such proofs might provide genuine insight into the constructive content of arithmetic and stronger theories. *Ordinal analysis* belongs, for example, to the programs which have been pursued in proof theory. Gentzen's ordinal analysis of PA [21, 22, 23] and the Ackermann-Kreisel classification of the provably recursive functions of PA [1, 33] are two classic examples.

This thesis follows this spirit of ordinal analysis. We study ordinal notation systems of theories and classify the provabably recursive functions of them by defining fast growing hierarchies. On the other hand, the main results are about thresholds between provability and unprovability of some sentences in a given theory. In this sense, this thesis lies also in the line of Gödel's approach. Another aspect of this thesis is that we show how concrete mathematics can be applied in solving abstract problems of logic. And this aspect has something to do with the question if logic matters to mathematicians. Let us explain this something more.

Though Gödel's work is astonishing and remarkable, there is however a point which is not so satisfactory for mathematicians. The point is that Gödel's work talks about coding of a system into itself. And this does not so care mathematicians, since there would be still the possibility that mathematical sentences which are independent of a reasonably strong theory might be never formulated in the everyday enterprise of a mathematician.

The situation was changed when the 1977 work of Paris and Harrington [40] was published. It furnished a completely transparent theorem of finite combinatorics. The theorem deals with the so-called Paris-Harrington principle (PH) and actually a very simple finite combinatorial variation of Ramsey's Theorem. (PH) is the following Π^0_2 sentence:

> *For any natural numbers n, c, k there is an natural number ℓ such that, given $C\colon [\ell]^n \to c$, there is a set $H \subseteq \ell$ such that $C \restriction [H]^n$ is a constant function and H contains at least* $\max\{k, \min(H)\}$ *elements.*

Here $[X]^n$ is the set of all n element subsets of X. Paris and Harrington showed model-theoretically that (PH) is PA-*independent*, i.e., (PH) is not PA-provable, though it is true.

The second achievement of logicians which attracts attention of mathematicians is the *Friedman style miniaturization* of Kruskal's theorem about finite rooted trees. A finite rooted tree T is a finite partial ordering $(T, \preceq)$ such that, if T is not empty, there is a smallest element called the *root* of T and that for each $b \in T$ the set $\{a \in T\colon a \preceq b\}$ is totally ordered. Let $a \wedge b$ denote the infimum of a and b for a, $b \in T$. A finite rooted tree T_1 is called *homeomorphically embeddable* into a finite rooted tree T_2 if there is an injection $f\colon T_1 \to T_2$ such that $f(a \wedge b) = f(a) \wedge f(b)$ for all a, $b \in T_1$. Kruskal's theorem says:

> *Given a sequence of finite rooted trees $(T_k)_{k<\omega}$, there are indices $\ell < m$ such that T_ℓ is homeomorphically embeddable into T_m.*

This is a true sentence, cf. Kruskal [35]. Moreover, Friedman [49] showed that Kruskal's theorem is ATR_0-independent. The formula complexity of Kruskal's theorem is Π^1_1, i.e., it does not belong to the language of first-order arithmetic. By the Friedman style miniaturization, however, it can be transformed into a sentence of a relatively simple complexity, namely Π^0_2. This is done with the help of a certain norm function. By using it one can just talk of finite sequences of finite rooted trees instead of infinite ones:

> *For any natural number k there is a constant n so large that, for any finite sequence $T_0, \ldots, T_n$ of finite rooted trees such that the number of nodes of T_i is at most $k + f(i)$ for all $i \leq n$, there are indices $\ell < m \leq n$ such that T_ℓ is homeomorphically embeddable into T_m.*

Note that this is a Π^0_2 sentence if $f\colon \mathbb{N} \to \mathbb{N}$ is primitive recursive, i.e., it is mathematically not more complicated than its infinite prototype. Moreover, for $f(i) = i$ it is still ATR_0-independent, cf. [49]. This was later sharpened by Loebl and Matoušek [37]: if $f(i) = 4 \log_2 i$ then the miniaturization is PA-independent, and if $f(i) = \frac{1}{2} \log_2 i$ then provable in PA.

Now a question arises: is there a real number c such that the miniaturization is provable if and only if $f(i) = r \log(i)$ for some $r \leq c$? This question was one of the starting points of the pioneer work Weiermann [60]. He noticed that an analysis of the combinatorial behavior of the given norm function could really give a secret away in solving the provability problem. Through his following works it has suddenly become clear that a part of "concrete" mathematics, e.g. analytic number theory, earns respect of logicians in solving problems from "abstract" mathematics. Indeed, he found out there are phenomena which can be called *phase transitions* whose concept stems originally from physics.

1.1 Phase transitions

In physics, a phase transition is the transformation of a thermodynamic system from one phase to another. The distinguishing characteristic of a phase transition is an abrupt change in one or more physical properties, e.g. the heat capacity, with a small change in a thermodynamic variable such as the temperature. And it is one of the cornerstones of equilibrium statistical mechanics that macroscopic systems undergo phase transitions as the external parameters change. Typical examples are the transitions between the solid, liquid, and gaseous phases, i.e., evaporation, boiling, melting, freezing, sublimation, etc. For a mathematical description of phase transitions see Gibbs [24] and Lee and Yang [36, 64]

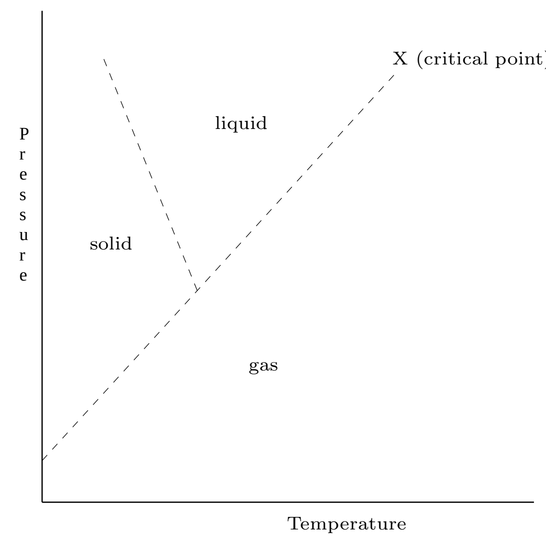

Figure 1.1: phase transitions in physics

Also in mathematics the interest in the study of phase transitions has grown, especially in random combinatorial problems. The classic combinatorial phase transition occurs in the random graph model of Erdös and Rényi [15]. There one considers a graph on n vertices with edge occupation probability α/n. As the parameter α passes through 1, the model undergoes a phase transition in the sense that the size of the largest connected component changes from order $\log n$ to order n. More recently, there has been much study of the phase transition in the random k-SAT model, both by heuristic and rigorous methods, see [8].

In k-SAT, the instances are formulas in conjunctive normal form; each formula has m clauses, and each clause has k distinct literals drawn uniformly at random from among n Boolean variables and their negations. For fixed $k \geq 2$, the model undergoes a sharp transition from solvability to insolvability as the parameter $\alpha = m/n$ passes through a particular k-dependent value, cf. Friedgut [20].

What about logic? Weiermann [60] showed that there is a real number c such that the miniaturized version of Kruskal's theorem with the parameter function $f(x) = r \log_2 x$ is provable iff $r \leq c$. Indeed, $c = \frac{1}{\log_2 \alpha}$, where $\alpha = 2.9557652\ldots$ is Otter's tree constant, cf. Otter [39].

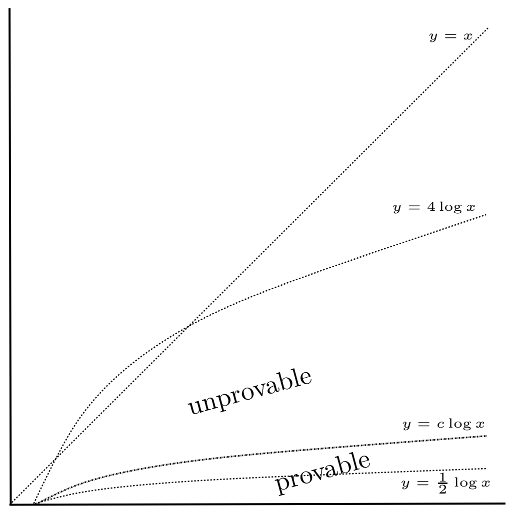

Figure 1.2: phase transitions in PA

Another example of phase transitions in logic were given in Weiermann [60] and Arai [4][1]. Using the Friedman style miniaturization of well-foundedness of ordinals below ε_0 they showed that the PA-provability of this Π^0_2 sentence with a parameter function changes abruptly as the parameter function varies. See Chapter 2.

This aspect of these two works is one of the starting points of this thesis in the sense that one demonstrates some more Friedman style miniaturizations concerning several well-known ordinal notation systems for PA which share the phase transition property. And this investigation leads to a comparison of the ordinal notation systems.

The thesis gives also a guide how one could generally construct such examples in every logic system which is strong enough in the sense of Gödel. Another point of the thesis is the investigation of the parametrized version of the Kanamori-McAloon principle which is equivalent to the Paris-Harrington principle. The two variants of the finite Ramsey theorem show also phase transitions

[1] Arai's work is based on Weiermann's work though the latter is published later.

as the growth speed of the parameter function changes. This part of the thesis is invoked by Kanamori-McAloon's work [29] and Weiermann [61] which shows phase transitions with respect to the Paris-Harrington principle.

1.2 Overview

We now give a short overview of the single parts and chapters. The thesis is composed of three parts.

Part I We introduce six well-known ordinal notation systems for PA and study them with regard to the question whether or not they are isomorphic to each other. This is closely connected with the *conceptual* problem[2] which criteria one can use to answer the question about 'natural' or 'canonical' notation systems for ordinals in proof theory. This will be done by giving certain intrinsic mathematical properties, namely phase transitions, that are independent of their possible use in proof-theoretic work.

In Chapter 2 the main results of Weiermann [60, 59] and Arai [4] are summarized. It is shown that some ordinal notation systems for PA are intrinsically different.

The main object of Chapter 3 is in contrast to show that many of the six systems are intrinsically isomorphic though at first glance they seem to have nothing to do with each other.

Part II The Kanamori-McAloon principle (KM) is a variant of the finite Ramsey theorem and equivalent to the Paris-Harrington principle (PH). Given a function $f\colon \mathbb{N} \to \mathbb{N}$ we define the parametrized version $(\mathrm{KM})_f$ and compare it with $(\mathrm{PH})_f$ both in the local level, that is, with respect to $\mathrm{I}\Sigma_n$, and in the full strength of PA. For this we construct fast growing hierarchies of Ramsey functions which are the Skolem functions of $(\mathrm{KM})_f$.

Chapter 4 contains the following: some comparisons between $(\mathrm{KM})_f$ and $(\mathrm{PH})_f$ and the upper bounds for the growth of the Skolem functions of $(\mathrm{KM})_f$ and its provability for some f.

By constructing a fast growing hierarchy we show in Chapter 5 that the parametrized Kanamori-McAloon principle undergoes phase transitions as the parameter function f varies. It is demonstrated how fast the parameter function f should grow, so that $(\mathrm{KM})_f$ is PA-independent or so that $(\mathrm{KM})_f^{n+1}$ is $\mathrm{I}\Sigma_n$-independent.

Part III In this part we present a general way of using Friedman style miniaturization to construct Π^0_2 sentences independent of a given theory beyond PA and show that they also share the phase transition property. This is directly connected with the construction of the ordinal notation system for the given theory.

[2]See Feferman [19] and Kreisel [34] for more discussions on the conceptual problems in logic.

We investigate in Chapter 6 the ordinal number system for the small Veblen ordinal $\vartheta\Omega^\omega$ which is the proof-theoretic strength of $\mathrm{ACA}_0 + \Pi^1_2$-BI. We reconstruct the ordinal notation system of $\vartheta\Omega^\omega$ given by Rathjen and Weiermann [43] as a set of closed terms and define a norm function on it. This is necessary for the construction of Π^0_2 sentences which shows phase transitions with respect to $\mathrm{ACA}_0 + \Pi^1_2$-BI.

1.3 Preliminaries

We summarize some notations and their properties which serve as basics for the whole part of the thesis.

1.3.1 Well-partial-ordering

A *quasi-ordering* is a pair $(X, \preceq)$, where X is a set and $\preceq$ is a transitive, reflexive binary relation on X. If $Y \subseteq X$ we write $(Y, \preceq)$ instead of $(Y, \preceq \restriction Y \times Y)$. A quasi-ordering $(X, \preceq)$ is called a *partial ordering* if $\preceq$ is antisymmetric, too. Modulo the equivalence relation $\cong$ on X defined by

$$x \cong y \quad \text{iff} \quad x \preceq y \quad \text{and} \quad y \preceq x$$

any quasi-ordering may be regarded as a partial ordering of the set $X/\cong$.

Given a partial ordering $(X, \preceq)$, we call $\|\cdot\| \colon X \to \mathbb{N}$ a *norm function* on X if for every $n \in \mathbb{N}$ the set $\{\alpha \in X \colon \|\alpha\| \leq n\}$ is finite. The structure $(X, \preceq, \|\cdot\|)$ is called then a *normed* partial ordering.

For any partial ordering $(X, \preceq)$ and any x, $y \in X$ we write $x \prec y$ for $x \preceq y$ and $y \not\preceq x$. A *linear ordering* is a partial ordering $(X, \preceq)$ in which any two elementary are $\preceq$-comparable.

A *well-quasi-ordering* (*wqo*) is a quasi-ordering $(X, \preceq)$ such that there is no infinite sequence $\langle x_i \rangle_{i\in\omega}$ of elements of X satisfying: $x_i \not\preceq x_j$ for all $i < j$. A *well-partial-ordering* (*wpo*) is a partial ordering which is well-quasi-ordered. $(X, \prec)$ is called *well-ordering* if $(X, \preceq)$ is a linear *wpo*. The following condition is necessary and sufficient for a partial ordering $(X, \preceq)$ to be a *wpo*:

Every extension of $\preceq$ to a linear ordering on X is a well-ordering.

Given a well-ordering $(X, \prec^+)$ and a partial ordering $(X, \preceq)$ there is a natural question concerning well-orderings and their order types: Under what condition is there a non-trivial upper bound for the order type of $(X, \prec^+)$ which depends only on $(X, \preceq)$? ("non-trivial" means "lower than the obvious upper bound obtained by considering the cardinality of X.) In [12, 1977] de Jongh and Parikh gave an answer to this question.

Definition 1.3.1. Given a *wpo* $(X, \preceq)$ we define its *maximal order type* by

$$o(X, \preceq) := \sup\{otype(\prec^+)\colon \ \prec^+ \text{ is a well-ordering on } X \text{ extending } \preceq\}.$$

$otype(\prec^+)$ denotes the order type of the well-ordering $\prec^+$.

We write $o(X)$ for $o(X, \preceq)$ if it causes no confusion.

Theorem 1.3.2 (de Jongh and Parikh [12])**.** *If $(X, \preceq)$ is a wpo, then there exists a well-ordering $\prec^+$ on X extending $\preceq$ such that $o(X, \preceq) = otype(\prec^+)$.*

1.3.2 Friedman style miniaturizations

Let T be a subsystem of second order Peano arithmetic and $\langle \mathrm{B}, \leq\rangle$ be a "reasonable" ordinal notation system of T based on a norm function $\|\cdot\|_b\colon \mathrm{B} \to \mathbb{N}$. Assume that this norm function is provably recursive in PA and that there is a uniform elementary bound on $\{\beta \in \mathrm{B}\colon \|\beta\|_b \leq n\}$ for every $n \in \mathbb{N}$.

Provably accessible elementary recursive ordinal notation systems are e.g. reasonable. More information (and proofs) about such systems can be found in Smith [51]. All of the notation systems which are usual in proof theory are "reasonable".

WO(B) is the assertion that $\langle \mathrm{B}, \leq\rangle$ is well ordered. For each $\beta \in \mathrm{B}$, WO(β) is the assertion that B is well ordered up to β, i.e. B contains no infinite descending sequence beginning with β. Though WO(B) is not provable from T we have a problem that the statement WO(B) does not belong to the domain of pure finite combinatorics, because it contains a quantifier over infinite sequence. In other words, its syntactic form is Π^1_1.

Friedman overcame this objection by replacing the Π^1_1 statement WO(B) with its so-called Π^0_2 finite miniaturization. It is a variation of the following assertion PRWO(B) that B is primitive recursively well ordered, i.e. B contains no infinite decreasing primitive recursive sequence. In an analogous way we define PRWO(β) for each $\beta \in \mathrm{B}$. Note that the assertions PRWO(B) and PRWO(β) are Π^0_2.

Definition 1.3.3 (Friedman [49], Smith [51])**.** An infinite sequence $\langle \beta_i\rangle_{i<\omega}$ from B is called *slow* if

> there is a natural number k such that $\|\beta_i\|_b \leq k + i$ for all $i \in \mathbb{N}$.

SWO$(\mathrm{B}, \leq, id)$ is the assertion that B is *slowly well ordered*, i.e. B contains no slow infinite descending sequence. By König's Lemma [32], SWO$(\mathrm{B}, \leq, id)$ is equivalent to the following Π^0_2 assertion[3]: Let $f(i) = i$.

> For any k there exists an n so large that B contains no finite descending sequence $\beta_0 > \beta_1 > \cdots > \beta_n$ such that $\|\beta_i\|_b \leq k + f(i)$ for any $i \leq n$.

This Π^0_2 assertion is also denoted by SWO$(\mathrm{B}, \leq, f)$.

[3] We use the refined version of R. Smith [51].

Let $(Q, \preceq)$ be a "*reasonable*" well-partial-ordering based on a norm function $\|\cdot\|_q \colon Q \to \mathbb{N}$. Assume that the maximal order type is the proof-theoretic ordinal of T. For more about reasonable well-partial orderings we refer to [51]. The "*slowly well partial-orderedness*" of Q, $\mathrm{SWP}(Q, \preceq, f)$, is defined as follows:

> For any k there exists an n such that for any finite sequence $\gamma_0, \ldots, \gamma_n$ from Q satisfying the condition that $\|\gamma_i\|_q \leq k + f(i)$ for any $i \leq n$ there are $\ell < m \leq n$ satisfying $\gamma_\ell \preceq \gamma_m$.

Note that $\mathrm{SWO}(\mathrm{B}, \leq, f)$ and $\mathrm{SWP}(Q, \preceq, f)$ are all true for any function $f \colon \mathbb{N} \to \mathbb{N}$ if B and Q are a well-ordering and a wpo, respectively.

Theorem 1.3.4 (Friedman [49], Smith [51])**.** *In* ACA_0 *the following assertions are pairwise equivalent*:

(i) $\mathrm{SWO}(\mathrm{B}, \leq, id)$.

(ii) $\mathrm{SWP}(Q, \preceq, id)$.

(iii) 1-*consistency of* T.

(iv) Π^0_2 *soundness of the formal system* $\mathrm{ACA}_0 + \{\mathrm{WO}(\beta) \colon \beta \in \mathrm{B}\}$.

The 1-consistency of a theory T is the assertion: if φ is a Σ^0_1 sentence provable from T, then φ is true.

Corollary 1.3.5. $\mathrm{SWO}(\mathrm{B}, \leq, id)$ *and* $\mathrm{SWP}(Q, \preceq, id)$ *are* T-*independent.*

1.3.3 Concrete mathematics

This section demonstrates a very small and simple part of the symbolic approach to combinatorial enumerations and is by no means self-contained. In his long history it has actually developed various and sophisticated systematic ways in considering combinatorial structures. However, we will concentrate on the most basic things and their properties needed for our future works. Two topics are especially useful: *generating functions* and *asymptotics*.

Over generating functions many general set-theoretic constructions can be directly translated into objects of symbolic methods. It is the most powerful way in dealing with sequences of numbers to manipulate infinite series that "generate" those sequences. We give a catalogue based on a core of important constructions which includes the operations of union, Cartesian product, sequence, powerset, and multiset. In this way, a specification language for elementary combinatorial objects is defined. The problem of enumerating a class of combinatorial structures then simply reduces to finding a proper specification, a sort of formal "grammar", for the class in terms of the basic constructions.

Another advantage of such translations becomes obvious when deriving *exact* mathematical results is not available and we, nevertheless, still would like to know something about the answer. In such cases asymptotic methods provide a convenient way to calculate good approximations to specific values for quantities of interest. The word *asymptotic* means any approximate value that gets closer and closer to the truth, when some parameter approaches a limiting value. Here of course we will be content with an introduction to the subject. We will be particularly interested in understanding the definitions of '$\sim$' and '$\mathcal{O}$' symbols.

There are many standard books the reader can refer to. Stanley [53, 54] seem to be classic and cover almost all materials on this subject. Segdewick and Flajolet [48] is somewhat compact, but deals with useful technical methods concerning generating functions and asymptotic approximations. We owe the title "concrete mathematics" to Graham, Knuth, and Patashnik [26]. For the one who wants to sniff the attractiveness and power of generating functions and asymptotics this is a very interesting book. Basic definitions and techniques are introduced through many simple, but instructive examples.

Generating functions

In the framework to be described in the following, classes of combinatorial structures are defined, either iteratively or recursively, in terms of simpler classes by means of a collection of elementary combinatorial constructions. The approach followed resembles the description of formal languages by means of context-free grammars.

Definition 1.3.6. A *class of combinatorial structures* is a pair $(\mathcal{A}, \|\cdot\|)$ where $\mathcal{A}$ is at most denumerable and the norm function $\|\cdot\|\colon \mathcal{A} \to \mathbb{N}$ is such that the inverse image of any integer is finite.

We write $\|\cdot\|_{\mathcal{A}}$ when needed. Given a class of combinatorial structures $(\mathcal{A}, \|\cdot\|)$, we consistently let

$$\mathcal{A}_n := \{\alpha \in \mathcal{A}\colon \|\alpha\| = n\}.$$

Then A_n, the number of elements of $\mathcal{A}_n$, are all finite.

Definition 1.3.7. The *generating function*[2] of a sequence $(A_n)_{n\in\omega}$ is

$$A(z) = \sum_{n\geq 0} A_n z^n.$$

The coefficient A_n of z^n in $A(z)$ is often denoted by $[z^n]A(z)$.

[2] Actually, this kind of generating function is called *ordinary* generating function. Cf. [54, 48].

Two kinds of "closed forms" come up when we work with generating functions. We might have a closed form for $A(z)$, expressed in terms of z; or we might have a closed form for A_n, expressed in terms of n. For instance, binary sequences $\mathcal{S}$ and permutations $\mathcal{P}$, with the usual conventions that the size of a word is its length and the size of a permutation is the number of its elements, correspond to the counting sequences

$$S_n = 2^n \quad \text{and} \quad P_n = n! \quad \text{resp.}$$

Hence we have

$$S(z) = \sum_{n\geq 0} 2^n\, z^n = \frac{1}{1-2z} \quad \text{and} \quad P(z) = \sum_{n\beta 0} n!\, z^n.$$

The generating function $S(z)$ exists as standard analytic objects since the series converge in a neighborhood of 0, while $P(z)$ is a purely formal power series. The *radius of convergence*, *r.o.c.*, of $S(z)$ at 0 is positive and that of $P(z)$ is 0.

The sums in generating functions runs over all natural numbers, but we often find it more convenient to extend the sum over all integers. We can do this by simply putting $A_{-1} = A_{-2} = \cdots = 0$. In such cases we might still talk about the sequence $(A_n)_{n\geq 0}$, as if the A_n's didn't exist for negative n.

Admissible constructions for generating functions

We introduce some admissible operators that form the core of a specification language for combinatorial structures. $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, etc. denote classes of combinatorial structures.

Cartesian Product: Assume that $\mathcal{A}$ is the Cartesian product of $\mathcal{B}$ and $\mathcal{C}$,

$$\mathcal{A} = \mathcal{B} \times \mathcal{C},$$

the size of a pair $\alpha = (\beta, \gamma)$ being defined by $\|\alpha\|_{\mathcal{A}} = \|\beta\|_{\mathcal{B}} + \|\gamma\|_{\mathcal{C}}$. Then, the counting sequences corresponding to $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ are related by the relation

$$A_n = \sum_{k=0}^{n} B_k\, C_{n-k}.$$

Therefore, we find a product of generating functions

$$A(z) = B(z) \cdot C(z).$$

Disjoint Union: We take the sum

$$\mathcal{A} = \mathcal{B} + \mathcal{C}$$

to represent the set-theoretic disjoint union of two disjoint copies of $\mathcal{B}$ and $\mathcal{C}$. One way of formalizing this notion is to introduce two distinct "markers" ϵ_1 and ϵ_2, each of size zero, and define the disjoint union $\mathcal{A} = \mathcal{B} + \mathcal{C}$ of $\mathcal{B}$, $\mathcal{C}$ by

$$\mathcal{A} = \mathcal{B} + \mathcal{C} = (\{\epsilon_1\} \times \mathcal{B}) \cup (\{\epsilon_2\} \times \mathcal{C}).$$

Therefore, the size of $\mathcal{A}$-element coincides with that of the corresponding element in $\mathcal{B}$ or $\mathcal{C}$. We have clearly $A_n = B_n + C_n$ and

$$A(z) = B(z) + C(z).$$

Sequence: Let $\mathcal{B}$ be a class of combinatorial structures such that $\mathcal{B}$ contains no object of size 0, i.e. $[z^0]B(z) = 0$. Then the sequence class $\mathfrak{S}\{\mathcal{B}\}$ is defined as the infinite sum

$$\mathfrak{S}\{\mathcal{B}\} = \{\epsilon\} + \mathcal{B} + (\mathcal{B} \times \mathcal{B}) + (\mathcal{B} \times \mathcal{B} \times \mathcal{B}) + \cdots$$

with ϵ being a "null" structure, meaning a structure of size 0. Note that the construction $\mathcal{A} = \mathfrak{S}\{\mathcal{B}\}$ defines a proper class satisfying the finiteness condition for sizes since there are no objects of size 0 in $\mathcal{B}$. By definition of size for sums and products the size of a sequence is the sum of the sizes of its components:

$$\|\alpha\| = \|\beta_1\| + \cdots + \|\beta_\ell\|,$$

where $\alpha = (\beta_1, \ldots, \beta_\ell)$. Hence

$$A(z) = 1 + B(z) + B^2(z) + B^3(z) + \cdots = \frac{1}{1 - B(z)},$$

where the geometric sum converges in the sense of formal power series since $[z^0]B(z) = 0$.

Powerset: $\mathcal{A} = \mathfrak{P}\{\mathcal{B}\}$ is defined as the class consisting of all finite subsets of class $\mathcal{B}$ permitting no repetitions. The size of a set is the sum of the sizes of its non-repeating components:

$$\|\alpha\| = \|\beta_1\| + \cdots + \|\beta_\ell\|,$$

where $\alpha = \{\beta_1, \ldots, \beta_\ell\}$. Then

$$A(z) = \exp\left(\frac{B(z)}{1} - \frac{B(z^2)}{2} + \frac{B(z^3)}{3} - \cdots\right) = \exp\left(\sum_{k \ge 1} (-1)^{k-1} \frac{B(z^k)}{k}\right).$$

Multiset: Multisets $[\,\beta_1, \ldots, \beta_\ell\,]$ are like sets except that repetitions of elements are allowed, the notation being $\mathcal{A} = \mathfrak{M}\{\mathcal{B}\}$. Additionally, we also assume that $[z^0]B(z) = 0$. The size of a multiset is the sum of the sizes of its components:

$$\|\alpha\| = \|\beta_1\| + \cdots + \|\beta_\ell\|,$$

where $\alpha = [\beta_1, \ldots, \beta_\ell]$. Then

$$A(z) = \exp\left(\frac{B(z)}{1} + \frac{B(z^2)}{2} + \frac{B(z^3)}{3} + \cdots\right) = \exp\left(\sum_{k\geq 1} \frac{B(z^k)}{k}\right).$$

For more details about power set and multiset constructions we refer the reader to the following excellent books [53, 54, 48].

Theorem 1.3.8. *The constructions of union, Cartesian product, sequence, power set, and multiset are all admissible. The associated operators are*

$$\begin{aligned}
\text{Union:}\quad & \mathcal{A} = \mathcal{B} + \mathcal{C} \quad \textit{and} \quad A(z) = B(z) + C(z) \\
\text{Product:}\quad & \mathcal{A} = \mathcal{B} \times \mathcal{C} \quad \textit{and} \quad A(z) = B(z) \cdot C(z) \\
\text{Sequence:}\quad & \mathcal{A} = \mathfrak{S}\{\mathcal{B}\} \quad \textit{and} \quad A(z) = \tfrac{1}{1-B(z)} \\
\text{Power set:}\quad & \mathcal{A} = \mathfrak{P}\{\mathcal{B}\} \quad \textit{and} \quad A(z) = \exp\left(\textstyle\sum_{k\geq 1} (-1)^{k-1} \frac{B(z^k)}{k}\right) \\
\text{Multiset:}\quad & \mathcal{A} = \mathfrak{M}\{\mathcal{B}\} \quad \textit{and} \quad A(z) = \exp\left(\textstyle\sum_{k\geq 1} \frac{B(z^k)}{k}\right)
\end{aligned}$$

Asymptotic analysis

We say that, given two sequences $(a_n)_n$ and $(b_n)_n$ of *real* numbers, a_n is asymptotic to b_n if

$$a_n \sim b_n, \text{ i.e., } \lim_{n\to\infty} \frac{a_n}{b_n} = 1.$$

The next very helpful notational convention for asymptotic analysis was introduced by Paul Bachmann in [5], namely the $\mathcal{O}$-notation. We say that

$$a_n = \mathcal{O}(b_n)$$

when there are two constants C and n_0 such that

$$|a_n| \leq C|b_n| \quad \text{whenever } n \geq n_0,$$

where $|a|$ means the absolute value of a given real number a.

Most of the generating functions that occur in combinatorial enumerations are analytic functions. Their expansions converge in a neighborhood of the origin and suitable uses of Cauchy's integral formula make it possible to determine effective bounds for coefficients of such analytic generating functions. For the fairly common case of functions that have singularities at a finite distance the exponential growth formula relates the location of the singularities closest to the origin to the exponential order of growth of coefficients. The following shows why the singularity nearest to the origin is important.

Theorem 1.3.9 (The exponential growth formula)**.** *If $f(z)$ is analytic at* 0 *and R is the modulus of a singularity of $f(z)$ nearest to the origin, then the coefficient $f_n = [z^n]f(z)$ satisfies*

$$\limsup |f_n|^{1/n} = \frac{1}{R}.$$

In other word, for any $\epsilon > 0$:

- *$|f_n|$ exceeds $(R^{-1} - \epsilon)$ infinitely often, and*
- *$|f_n|$ is dominated by $(R^{-1} + \epsilon)$ almost everywhere.*

Definitions and properties of analytic functions, Taylor sequences, radius of convergence (*r.o.c.*), singularities, and Cauchy's integral formula can be found in any standard book on complex function theory. In particular, two theorems provide us with very important tools for the future work.

Theorem 1.3.10 (Pringsheim's lemma)**.** *If a function with a finite r.o.c. has Taylor coefficients that are nonnegative, then one of its singularities of smallest modulus is real positive.*

Proof. See Section 7.21 in [52]. □

In the following this theorem will be always applicable since the Taylor coefficients of a generating function are always nonnegative.

Theorem 1.3.11 (Weierstrass' preparation theorem)**.** *Let $F(z, w)$ be a function of two complex variables which is analytic in a neighborhood $|z-z_0| < r$, $|w-w_0| < \rho$ of the point (z_0, w_0), and suppose that*

$$F(z_0, w_0) = 0 \quad \textit{and} \quad F(z_0, w) \not\equiv 0.$$

Then there is a neighborhood $|z-z_0| < r' < r$, $|w-w_0| < \rho' < \rho$ in which $F(z, w)$ can be written as

$$F(z, w) = (A_0(z) + A_1(z) \cdot w + \cdots + A_{k-1}(z) \cdot w^{k-1} + w^k) \cdot G(z, w),$$

where k is such that

$$\frac{\partial F(z_0, w_0)}{\partial w} = \cdots = \frac{\partial^{k-1} F(z_0, w_0)}{\partial w^{k-1}} = 0, \quad \frac{\partial^k F(z_0, w_0)}{\partial w^k} \neq 0,$$

the functions $A_0(z), \ldots, A_{k-1}(z)$ are analytic if $|z - z_0| < r'$, and the function $G(z, w)$ is analytic and nonzero if $|z - z_0| < r'$, $|w - w_0| < \rho'$.

Proof. See e.g. Theorem 3.10 in [38]. □

Thus, despite the seeming generality of the equation $F(z, w) = 0$, there is a neighborhood of the point (z_0, w_0) where it is equivalent to the equation

$$A_0(z) + A_1(z) \cdot w + \cdots + A_{k-1}(z) \cdot w^{k-1} + w^k = 0,$$

which is algebraic in w.

1.3.4 Basic functions and conventions

We list some basic functions and conventions which will be used in the whole part of the thesis.

Natural numbers The small Latin letters n, $m, \ldots$ range over natural numbers.

Cardinalities Given a finite set X, $\bar{\bar{X}}$ denotes the cardinality of X.

Theories The mathematical systems we are going to talk about are PA, ACA_0, ATR_0, $\mathrm{ACA}_0 + \Pi^1_2$-BI. Their definitions and proof-theoretic properties are investigated e.g. in [50, 43].

Norm functions Given a partial ordering $(X, \preceq)$, we just write $\| \cdot \|$ generally without any subscript for any norm function on X if it causes no confusion.

Modulus Given a real number r, $|r|$ means its absolute value. There will be no confusion with the following binary length function.

Iterated binary length functions Given a nonnegative real number x, $\lfloor x \rfloor$ is the largest natural number not bigger than x and $\lceil x \rceil$ is the least natural number not less than x. Set

$$|x| := \lceil \log_2(x+1) \rceil ,$$

i.e., $|x|$ is the length of the binary representation of x. We iterate the $|\cdot|$-function:

$$|x|_0 := x \quad \text{and} \quad |x|_{m+1} := ||x|_m|$$

And we write $\log x$ for $\log_2 x$.

Inverse functions A function $f \colon \mathbb{N} \to \mathbb{N}$ is said to be *unbounded* if we have

$$\forall i, j\, [\, i \leq j \Longrightarrow h(i) \leq h(j)\,] \quad \text{and} \quad \lim_{i \to \infty} h(i) = \infty.$$

For any unbounded function f let us define its *inverse* as follows:

$$f^{-1}(i) := \min\{\ell \colon i < f(\ell)\}.$$

Then $f^{-1}(i) \leq \ell$ iff $i < f(\ell)$.

Fast growing hierarchies Given α, $\beta \in \varepsilon_0$ put

$$\alpha_0(\beta) := \beta, \quad \alpha_{n+1}(\beta) := \alpha^{\alpha_n(\beta)}, \quad \text{and} \quad \alpha_n := \alpha_n(1).$$

For any limit ordinal $\lambda < \varepsilon_0$, there is a so-called "*fundamental sequence*" of λ and defined as follows: Let $\lambda = \omega^{\lambda_1} + \cdots + \omega^{\lambda_k}$ be in Cantor normal form.

$$\lambda[n] := \begin{cases} \omega^{\lambda_1} + \cdots + \omega^{\lambda_{k-1}} + \omega^{\lambda_k - 1} \cdot (n+1) & \text{if } \lambda_k \text{ is not a limit,} \\ \omega^{\lambda_1} + \cdots + \omega^{\lambda_{k-1}} + \omega^{\lambda_k[n]} & \text{otherwise.} \end{cases}$$

Then $\lambda[n] < \lambda[n+1]$ and $\lim_{n\to\infty} = \lambda$. Given $f\colon \mathbb{N} \to \mathbb{N}$ we define

$$f^{(0)}(i) := i \quad \text{and} \quad f^{(\ell+1)}(i) := f(f^{(\ell)}(i)).$$

The Hardy-Wainer hierarchy $(H_\alpha)_{\alpha<\varepsilon_0}$ and the Schwichtenberg-Wainer hierarchy $(F_\alpha)_{\lambda<\varepsilon_0}$ are defined as follows:

$$\begin{array}{rcl} H_0(i) & = & i \\ H_{\alpha+1}(i) & = & H_\alpha(i+1) \\ H_\lambda(i) & = & H_{\lambda[n]}(i) \end{array} \quad \text{and} \quad \begin{array}{rcl} F_0(i) & = & i+1 \\ F_{\alpha+1}(i) & = & F_\alpha^{(i+1)}(i) \\ F_\lambda(i) & = & F_{\lambda[n]}(i) \end{array}$$

Further let $H_{\varepsilon_0}(i) := H_{\omega_i}(i)$ and $F_{\varepsilon_0}(i) := F_{\omega_i}(i)$. Then $F_\alpha(i) = H_{\omega^\alpha}(i)$. And it is a folklore in proof theory that H_α (resp. F_α) is provably recursive in PA iff $\alpha < \varepsilon_0$. See Fairtlough and Wainer [16] for details.

1.4 Acknowledgments

I want to express greatest thanks to Professor Andreas Weiermann for his encouragement and support in working on the subject of this thesis. He has always had an open ear during its development and given instructive motivations through numerous discussions and his works. I owe Professors Justus Diller and Wolfram Pohlers special thanks. Without their support in all aspects during the stay at the Institut für Mathematische Logik und Grundlagenforschung in Münster I would not have got through the study of logic. Their excellent lectures and the familiar atmosphere at the institute they have taken care of are the reason why I could have so much pleasure in doing mathematics. I thank Professor Ralf Dieter Schindler very much for his concern in finishing this thesis. My sincere thanks go to all my present and former colleagues as well as Martina Pfeifer at the institute. They are or were not just colleagues, but have become my good friends.

Finally, I cannot forget my family and old friends in my home. Despite the distance their mental and intellectual support has always been there when I needed it. I dedicate this thesis to them.

Part I

Ordinal Notation Systems

Chapter 2

Intrinsic differences

> The feeling is that what distinguishes such orderings are certain intrinsic mathematical properties that are independent of their possible use in proof-theoretical work.
>
> S. Feferman [19]

It is one of the so-called ‘conceptual’ problems which criteria we can use to answer the question about ‘natural’ or ‘canonical’ notation systems for ordinals in proof theory. A conceptual problem is one which cannot be stated in a precise mathematical way. See [34, 25, 19] for intensive discussions. At least since the consistency proof for Peano arithmetic by Gentzen [21], several primitive recursive ordinal notation systems have been used to give consistency proofs for formal theories. These are given by natural well-orderings of expressions in some notation systems for ordinals.

Most famous is the notation system for ordinals less than the Cantor ordinal ε_0, based on a system of expressions generated by closure under addition and exponentiation to the base ω. This ordering relation is primitive recursive, and the consistency of Peano arithmetic can be proved by transfinite induction along this ordering. Moreover, this is best possible, in the sense that for any proper initial segment of this ordering, we can prove transfinite induction applied to arbitrary arithmetical formulas up to that segment in PA.

What we shall do in the following is to point out some intrinsic common or diverse properties among a few systems of natural ordinal representation that have been of interest in proof-theoretic analysis of PA. The ordinal notation systems considered here are all historically well-known, probably except for the system extracted from the graded provability algebra. L. Beklemishev refined this system and introduced a certain combinatorial game, called the Worm principle, such that the termination of the principle cannot be proved in PA, cf. Chapter 3.

(i) **The Cantor system**

This is the closure under the function $\chi = \lambda\xi, \eta \,.\, \omega^\xi \# \eta$ of $\{0\}$, where # is the natural sum of ordinals.

(ii) **The binary trees**

A *rooted binary tree* is a set of nodes such that, if it is not empty, there is one distinguished node called the root and the remaining nodes are partitioned into two rooted binary trees. The homeomorphic embeddability relation on the set $\mathcal{B}$ of all rooted binary trees is well-founded and has the maximal order type ε_0.

(iii) **Japaridze's** GLP

The set of the so-called letterless formulas without $\top$ in the language of Japaridze's propositional polymodal logic system GLP and the consistency ordering constitute an well-founded partial ordering of height ε_0.

(iv) **Countable tree-ordinals**

A certain set of countable *tree-ordinals* resembles the Cantor system. However, the sub-tree ordering on the set builds no well-ordering, although it is well-founded and has the height ε_0.

(v) **Schütte-Simpson's ordinal notation system**

K. Schütte and G. Simpson confined the ordinal number notation system of Buchholz [10] by letting away the addition and the construction of ω^α as basic operations. This system contains a subset with a well-ordering of order type ε_0.

(vi) **Ackermann's ϵ-substitution method**

Ackermann's original paper [1] on the termination of the ϵ-substitution method used for the consistency proof of PA used a coding of ordinals up to ε_0.

2.1 The Cantor system

For any nonzero ordinal $\alpha < \varepsilon_0$ there exist a unique natural number n and uniquely determined ordinals $\alpha_1, \ldots, \alpha_n < \varepsilon_0$ such that $\alpha = \omega^{\alpha_1} + \cdots + \omega^{\alpha_n}$ and $\alpha > \alpha_1 \geq \cdots \geq \alpha_n$. It is denoted by $\alpha =_{NF} \omega^{\alpha_1} + \cdots + \omega^{\alpha_n}$ and said to be in *Cantor normal form*. Then the *norm* $N\alpha$ is the number of ω-occurrence in α: $N0 := 0$ and if $\alpha =_{NF} \omega^{\alpha_1} + \cdots + \omega^{\alpha_n}$, then

$$N\alpha := n + N\alpha_1 + \cdots + N\alpha_n.$$

This norm function gives rise to a hierarchy which can be defined by pointwise transfinite recursion[1].

Definition 2.1.1. For $\alpha < \varepsilon_0$ set

$$A_\alpha(i) := \max\{A_\beta(i) + 1 \colon \beta < \alpha \text{ and } N\beta \leq N\alpha + i\}.$$

Then the function $A_{\varepsilon_0} := \lambda i \,.\, A_{\omega_i}(i)$ grows too fast to be provably total in PA.

Theorem 2.1.2. PA $\nvdash \forall k \, \exists n (A_{\omega_k}(1) = n)$

Proof. See Arai [2] and Weiermann [57]. □

Let $f \colon \mathbb{N} \to \mathbb{N}$ be given. The Friedman style miniaturization of the well-foundedness of ε_0 gives the following true Π^0_2 assertion $\mathrm{SWO}(\varepsilon_0, <, f)$:

> For any k there exists a constant n which is so large that, for any finite ordinal sequence $\alpha_0, \ldots, \alpha_n < \varepsilon_0$ with $N\alpha_i \leq k + f(i)$ for all $i \leq n$, there exist indices $\ell < m \leq n$ satisfying $\alpha_\ell \leq \alpha_m$.

According to Friedman [49, 51] the slowly well-orderedness of ε_0, that is $\mathrm{SWO}(\varepsilon_0, <, id)$, cannot be proved in PA. Using Theorem 2.1.2, Weiermann characterized in a nearly optimal way the class of functions f such that $\mathrm{SWO}(\varepsilon_0, <, f)$ is PA-unprovable.

Theorem 2.1.3 (Weiermann [60])**.** *Let $m \in \mathbb{N}$.*

(i) $\mathrm{SWO}(\varepsilon_0, <, \lambda i \,.\, |i| \cdot \mathrm{inv}(i))$ *is* PRA*-provable.*

(ii) $\mathrm{SWO}(\varepsilon_0, <, \lambda i \,.\, |i| \cdot |i|_m)$ *is* PA*-unprovable.*

Having seen this results, Arai gave a direct connection between the slowly well-orderedness of ε_0 and the Schwichtenberg-Wainer hierarchy $(F_\alpha)_{\alpha \leq \varepsilon_0}$.

Let $f_\alpha(i) := |i| \cdot |i|_{F_\alpha^{-1}(i)}$ and $L(\cdot; F_\alpha^{-1})$ be the Skolem function of $\mathrm{SWO}(\varepsilon_0, F_\alpha^{-1})$, i.e. $L(k; F_\alpha^{-1})$ is the least n such that

> for any finite $\alpha_0, \ldots, \alpha_n < \varepsilon_0$ with $N\alpha_i \leq k + |i| \cdot |i|_{F_\alpha^{-1}(i)}$ for all $i \leq n$, there exist indices $\ell < m \leq n$ satisfying $\alpha_\ell \leq \alpha_m$.

Theorem 2.1.4 (Arai [4])**.** *Let $\alpha \leq \varepsilon_0$.*

(i) *$L(\cdot; F_\alpha^{-1})$ is primitive recursive in F_α and vice versa. Therefore, $L(\cdot; F_\alpha^{-1})$ is provably total in* PA *iff $\alpha < \varepsilon_0$.*

(ii) $\mathrm{SWO}(\varepsilon_0, f_\alpha)$ *is* PA*-unprovable iff $\alpha = \varepsilon_0$.*

[1] For more about pointwise transfinite recursion see e.g. Weiermann [57]

2.2 Binary trees

The system $(\mathcal{B}, <)$ obtained from the Feferman-Schütte notation system[2] for Γ_0 by omitting the addition terms constitutes a well-ordering of order type ε_0. Moreover, $<$ is a canonical extension of the homeomorphic embeddability relation $\trianglelefteq$ on the set of binary relation. We give here a direct definition of $(\mathcal{B}, <)$.

A *rooted binary tree* T is a set of nodes such that, if it is not empty, there is one distinguished node called the root of T and the remaining nodes are partitioned into two rooted binary trees. The following definition of the set $\mathcal{B}$ of all rooted binary trees is very convenient for our study.

Assume that a constant symbol o and a binary function symbol φ are given. Then $\mathcal{B}$ is the least set of terms defined as follows:

- $o \in \mathcal{B}$;
- if α, $\beta \in \mathcal{B}$, then $\varphi(\alpha, \beta) \in \mathcal{B}$.

$\varphi(\alpha, \beta)$ will be abbreviated by $\varphi\alpha\beta$ if it causes no confusion.

Definition 2.2.1. The homeomorphic embeddability relation $\trianglelefteq$ on $\mathcal{B}$ is the least subset of $\mathcal{B} \times \mathcal{B}$ defined as follows:

- if $\alpha = o$, then $\alpha \trianglelefteq \beta$ for all $\beta \in \mathcal{B}$;
- if $\alpha = \varphi\alpha_1\alpha_2$, $\beta = \varphi\beta_1\beta_2$, and ($\alpha \trianglelefteq \beta_1$ or $\alpha \trianglelefteq \beta_2$), then $\alpha \trianglelefteq \beta$;
- if $\alpha = \varphi\alpha_1\alpha_2$, $\beta = \varphi\beta_1\beta_2$, and ($\alpha_1 \trianglelefteq \beta_1$ and $\alpha_2 \trianglelefteq \beta_2$), then $\alpha \trianglelefteq \beta$.

Theorem 2.2.2 (Higman [28])**.** $(\mathcal{B}, \trianglelefteq)$ *is a wpo.*

The following theorem is an unpublished result by de Jongh.

Theorem 2.2.3 (de Jongh)**.** $o(\mathcal{B}, \trianglelefteq) = \varepsilon_0$.

In case of $(\mathcal{B}, \trianglelefteq)$ we easily find a well-ordering $<$ on $\mathcal{B}$ with the order type ε_0.

Definition 2.2.4. $<$ is the least binary relation on $\mathcal{B}$ defined as follows:

- if $\alpha = o$ and $\beta \neq o$, then $\alpha < \beta$;
- if $\alpha = \varphi\alpha_1\alpha_2$ and $\beta = \varphi\beta_1\beta_2$, then $\alpha < \beta$ if one of the following hold:
  - $\alpha_1 < \beta_1$ and $\alpha_2 < \beta$,
  - $\alpha_1 = \beta_1$ and $\alpha_2 < \beta_2$,
  - $\alpha_1 > \beta_1$ and $\alpha \leq \beta_2$.

[2] See e.g. [17, 18, 45, 46]

Lemma 2.2.5. *$<$ is a well-ordering on $\mathcal{B}$ such that $\trianglelefteq \subseteq \leq$ and otype$(<) = \varepsilon_0$.*

Since ACA_0 does not prove the well-foundedness of ε_0, it cannot prove that $(\mathcal{B}, \trianglelefteq)$ and $(\mathcal{B}, \leq)$ are *wpo*s. And the claim that $(\mathcal{B}, \trianglelefteq)$ is a *wpo* is Π^1_1, and the same for $(\mathcal{B}, \leq)$. One can translate them into Π^0_2-formulae in the language of first order Peano arithmetic by Friedman style miniaturization. A norm function is needed: $\| \cdot \| \colon \mathcal{B} \to \omega$ is defined as follows:

- $\|o\| := 0$,
- $\|\varphi\alpha\beta\| := \|\alpha\| + \|\beta\| + 1$.

That is, $\|\alpha\|$ is the number of occurrences of φ in $\alpha \in \mathcal{B}$.

Using a standard coding, it is evident that $\mathcal{B}$, $\trianglelefteq$, $<$, and $\| \cdot \|$ are all primitive recursively definable in PA. Given $f \colon \mathbb{N} \to \mathbb{N}$ we call $(\mathcal{B}, \trianglelefteq)$ *slowly well-partial-ordered* by f if the following holds:

> For any k there exists a constant n which is so large that, for any finite sequence $\alpha_0, \ldots, \alpha_n$ of finite trees with $|\alpha_i| \leq k + f(i)$ for all $i \leq n$, there exist indices $\ell < m \leq n$ satisfying $\alpha_\ell \trianglelefteq \alpha_m$.

This is denoted by $\mathrm{SWP}(\mathcal{B}, \trianglelefteq, f)$. The *slowly well-orderedness* $\mathrm{SWO}(\mathcal{B}, <, f)$ is defined similarly.

Given a primitively recursive real number r put $f_r(i) := r|i|$.

Theorem 2.2.6 (Weiermann [59])**.**

(i) *If $r > \frac{1}{2}$, then $\mathrm{SWP}(\mathcal{B}, \trianglelefteq, f_r)$ and $\mathrm{SWO}(\mathcal{B}, <, f_r)$ are PA-unprovable.*

(ii) *If $r \leq \frac{1}{2}$, then $\mathrm{SWP}(\mathcal{B}, \trianglelefteq, f_r)$ and $\mathrm{SWO}(\mathcal{B}, <, f_r)$ are PRA-provable.*

Remark: Intrinsic differences

Theorem 2.1.3 and Theorem 2.2.6 point out that the Cantor system for ε_0 differs intrinsically from $(\mathcal{B}, <)$:

- $\mathrm{PA} \vdash \mathrm{SWO}(\varepsilon_0, <, \lambda i \,.\, r \cdot |i|)$ for any r;
- $\mathrm{PA} \nvdash \mathrm{SWO}(\mathcal{B}, <, \lambda i \,.\, r \cdot |i|)$ if $r > \frac{1}{2}$.

Chapter 3

Intrinsic isomorphisms

One of the popular criteria for the problem on *canonical* ordinal notation systems of a given theory is formulated as the question whether there is a natural ordinal classification of all provably total functions of the theory. In a significant number of well-known theories in mathematical logic there do exist plausible solutions, such as the classifications of provably recursive functions through Hardy-Wainer or Schwichtenberg-Wainer hierarchies. However, there still remains many questions about ordinal notation systems of a theory: How can it be explained that proof-theoretic ordinals are sensitive to the choice of particular proof systems? What are the intrinsic properties that distinguishe some ordering from other? Are they independent of their possible use in proof-theoretic work?

One of the latest approaches to these questions is done by Beklemishev [7]. He was concerned with the question of recovering a ordinal notation system *from* a given theory. He posed the question what would make a formal theory possible to rigorously specify its canonical ordinal notation system? He pointed out that an algebraic view point of proof theory, e.g., a well-behaved notion of *graded provability algebra*, could give a positive answer. In case of Peano arithmetic, such a view point is indeed helpful in clarifying the question where a canonical ordinal notation system comes from and how the whole process can be specified.

This chapter studies some more notation systems for ε_0 and demonstrates that the systems, including Beklemishev's system, share some common intrinsic properties. This will be done by showing the following:

- They build more or less the same *structured* systems. A structured system of countable ordinals is a system in which an arbitrary, but fixed "fundamental sequence" has been assigned to each limit.

- Some Friedman style independence results will be achieved such that the results are essentially regardless of the systems.

3.1 The graded provability algebra

Let T be an elementarily represented, sound fragment of PA such that $\mathrm{I}\Sigma_1 \subseteq \mathrm{T}$. The *Lindenbaum boolean algebra* $\mathcal{L}_{\mathrm{T}}$ is the set of all sentences modulo provable equivalence in T.

Let n-$\mathsf{Con}(\mathrm{T})$ denote a natural formula expressing that the theory $\mathrm{T}+Th_{\Pi_n}(\mathbb{N})$ is consistent, where $Th_{\Pi_n}(\mathbb{N})$ is the set of all true arithmetical Π_n sentences.

The *graded provability algebra* of T, $\mathcal{M}_{\mathrm{T}}$, is the structure of Lindenbaum boolean algebra $\mathcal{L}_{\mathrm{T}}$ with the *n-consistency operator* $\langle n\rangle_{\mathrm{T}}$, $n \in \mathbb{N}$, defined by $\langle n\rangle_{\mathrm{T}}\varphi := n\text{-}\mathsf{Con}(\mathrm{T}+\varphi)$.

The subscript T will be suppressed if the underlying theory is known from the context. The *n-provability operator* $[n]$ is defined by $[n]\varphi := \neg\langle n\rangle\neg\varphi$. $\langle 0\rangle\varphi$ is usually written by $\diamond\varphi$ and $[0]\varphi$ by $\Box\varphi$. Terms of the graded provability algebra correspond to propositional polymodal formulas.

GLP based on the identities of $\mathcal{M}_{\mathrm{T}}$ is an extension of the Gödel-Löb system GL[1]: Let m, $n \in \mathbb{N}$.

- **Axioms**
  - Boolean tautologies
  - $\langle n\rangle(\varphi \vee \psi) \to (\langle n\rangle\varphi \vee \langle n\rangle\psi)$
  - $\neg\langle n\rangle\neg\top$
  - $\langle n\rangle\varphi \to \langle n\rangle(\varphi \wedge \neg\langle n\rangle\varphi)$
  - $\langle n\rangle\varphi \to \langle m\rangle\varphi$ for $m \leq n$
  - $\langle m\rangle\varphi \to [n]\langle m\rangle\varphi$ for $m < n$
- **Rules**
  - modus ponens
  - $\varphi \to \psi \vdash \langle n\rangle\varphi \to \langle n\rangle\psi$

Then we have

$$\mathrm{GLP} \vdash \varphi(\vec{x}) \quad \text{iff} \quad \mathcal{M}_{\mathrm{T}} \models \forall\vec{x}(\varphi(\vec{x}) = \top).$$

Let S be the set of all finite words in the alphabet $\mathbb{N}$, including the empty word Λ. S_n is the restriction of S to the alphabet $\{n, n+1, \dots\}$. We identify each element $\alpha = n_1 \cdots n_k$ of S with its *modal interpretation* $\langle n_1\rangle \cdots \langle n_k\rangle\top$.

We write $\alpha \sim \beta$ if $\mathrm{GLP} \vdash \alpha \leftrightarrow \beta$. And $\alpha = \beta$ means the graphical identity. The orderings $<_n$ are defined on S by:

$$\alpha <_n \beta \quad \text{iff} \quad \mathrm{GLP} \vdash \beta \to \langle n\rangle\alpha.$$

[1]Cf. Boolos [9] for more about provability logic

Note that $<_n$ are transitive and irreflexive. Below we summarize Beklemishev's results, referring the reader to [7].

Given $\alpha \in S$ let α^k denote the k times iterated concatenation of α. The function $o\colon S \to \varepsilon_0$ is given as follows:

- $o(0^k) = k$;
- if $\alpha = \alpha_0 0 \cdots 0 \alpha_n$, where all $\alpha_i \in S_1$ and not all of them empty, then

$$o(\alpha) = \omega^{o(\alpha_n^-)} + \cdots + \omega^{o(\alpha_0^-)}.$$

Here γ^- is obtained from $\gamma \in S_1$ by replacing every letter $m+1$ with m.

Note that some of the elements of S are pairwise equivalent. However, there is a set of elements which represent each equivalence class, namely the set NF of *normal forms*. We define $\alpha \in NF$ by recursive induction on the *width* $w(\alpha)$, i.e. the number of different letters occurring in α.

- if $w(\alpha) \leq 1$, then $\alpha \in NF$;
- assume $w(\alpha) > 2$ and let n be the smallest letter in α such that graphically $\alpha = \alpha_0 n \cdots n \alpha_k$, where all $\alpha_i \in S_{n+1}$. Then $\alpha \in NF$ if all $\alpha_i \in NF$ and $o(\alpha_{i+1}) \not<_{n+1} (\alpha_i)$ for any $i < k$.

Theorem 3.1.1 (Beklemishev [7]). *Let α, $\beta \in S$.*

(i) *$(S, <_0)$ is a well-partial ordering of height ε_0.*

(ii) *Every word $\alpha \in S$ has an uniquely defined equivalent normal form.*

(iii) *If $\alpha \sim \beta$ then $o(\alpha) = o(\beta)$.*

(iv) *If $\alpha <_0 \beta$ then $o(\alpha) < o(\beta)$.*

(v) *$o \restriction NF\colon NF \to \varepsilon_0$ is an order-preserving isomorphism.*

It is also possible to assign fundamental sequences to each element of S. For $\alpha \in S$ and any $k \in \mathbb{N}$ we define $\alpha[k] \in S$ as follows:

- if $\alpha = \langle 0 \rangle \beta$ then $\alpha[k] = \beta$;
- if $\alpha = \langle n+1 \rangle \gamma \langle m \rangle \beta$, where $\gamma \in S_{n+1}$ and $m \leq n$, then $\alpha[k] = (n\gamma)^{k+1} m\beta$.

Theorem 3.1.2 (Beklemishev [7]). *Let $\alpha = \langle n+1 \rangle \beta$ and $k \in \mathbb{N}$.*

(i) *If $\alpha \in NF$, then $\alpha[k] \in NF$.*

(ii) $\alpha[k] <_0 \alpha[k+1] <_0 \alpha$.

(iii) *For every $\beta \in S$ there is a natural number ℓ such that $\beta <_0 \alpha[\ell]$.*

3.2 Worms, Hydras, and tree-ordinals

The *Hydra battle* introduced by L. Kirby and J. Paris [30] has an isomorphic formulation in terms of ordinals, namely fundamental sequences for ordinals below ε_0. Let $\cdot[\cdot]$ denote the standard assignments of fundamental sequences and let $\alpha(0) = \alpha$ and $\alpha(i+1) = \alpha(i)[i]$. Then the fact that chopping of the rightmost head is a winning strategy for Hercules is formalized by:

for any n there exists an i such that $\omega_n(i) = 0$

This is a true Π^0_2 sentence which is PA-unprovable since

$$H_\alpha(0) \leq \min\{i \colon \alpha(i) = 0\}.$$

Another combinatorial game similar to the Hydra battle is introduced by Beklemishev in [6] as an application of proof-theoretic analysis to his ordinal notation system S based on the concept of graded provability algebra. This principle deals with objects called *worms* and is hence called the *Worm principle*. We shall see that, modulo some isomorphism, they are mutually translatable.

A *worm* is just a finite function with natural numbers as values. We identify the worm $f \colon [\,0, n\,] \to \mathbb{N}$ with the list $f(0) \cdots f(n)$ or $\langle f(0), \ldots, f(n) \rangle$. We call $f(n)$ the *head* of the worm. $\varnothing$ denotes the empty function. Let W be the set of all worms and W_n the subset of W whose elements have values at least n.

A worm game begins with a worm and at each step we chop off its head. In response the worm grows in length according to some rules. Formally, we specify a function $\mathit{next} \colon W \times \mathbb{N} \to W$. Let α range over worms.

- $\mathit{next}(\varnothing, k) := \varnothing$.
- Let $\alpha = a_0 \cdots a_n$.
  - If $a_n = 0$, then $\mathit{next}(\alpha, k) := a_0 \cdots a_{n-1}$.
  - If $a_n > 0$, let $m := \max\{i < n \colon a_i < a_n\}$. We define

$$\mathit{next}(\alpha, k) := r * \underbrace{s * s * \cdots * s}_{k+1 \text{ times}}.$$

where $r = \langle a_0, \ldots, a_m \rangle$, $s = \langle a_{m+1}, \ldots, a_{n-1}, a_n - 1 \rangle$.

Here $*$ means the concatenation function of worms. Now let $\alpha(0) := \alpha$ and $\alpha(n+1) := \mathit{next}(\alpha(n), n+1)$. Then the *Worm principle* says that **E**very **W**orm **D**ies:

EWD := for any worm α there exists an n such that $\alpha(n) = \varnothing$

Note that **EWD** is a Π^0_2 sentence since $\alpha(n)$ is defined primitive recursively and that the size of maximal element of worms cannot increase. Hence $\alpha(n) = \beta$ can be written out as a Δ_0 formula in three variables.

Theorem 3.2.1 (Beklemishev [6]).

(i) **EWD** *is true, but* PA-*unprovable.*

(ii) **EWD** *is* PA-*equivalent to* 1-Con(PA).

In order to emphasize the relevance to S we use below another notation $\alpha[\![k]\!]$ instead of $next(\alpha, k)$.

Definition 3.2.2. Let α, β, $\gamma \in W$.

$$\alpha[\![k]\!] := next(\alpha, k) = \begin{cases} \varnothing & \text{if } \alpha = \varnothing, \\ \beta & \text{if } \alpha = \beta 0, \\ \beta m(\gamma n)^{k+1} & \text{if } \alpha = \beta m \gamma \langle n+1 \rangle,\ \gamma \in W_{n+1},\ m \leq n. \end{cases}$$

The next question is what is responsible for the PA-unprovability of **EWD**. Note that the Skolem function of **EWD** should grow too fast to be provably total in PA. Below we characterize the growth rate conditions which make the function grow fast. Given $f \colon \mathbb{N} \to \mathbb{N}$ and $\alpha \in W$ set

$$\alpha(f, 0) := \alpha, \quad \alpha(f, n+1) := \alpha(f, n)[\![f(n+1)]\!],$$

and define

$$\mathbf{EWD}(f) :\equiv \forall \alpha\, \exists n\, \alpha(f, n) = \varnothing.$$

Then $\mathbf{EWD} = \mathbf{EWD}(id)$. And $\mathbf{EWD}(f)$ remains Π^0_2 if f is primitive recursive.

Now we analyse growth rates of the Skolem functions of $\mathbf{EWD}(f)$ in terms of fast growing hierarchies. Notice that the correspondence o between S and the Cantor system for ε_0 defined above cannot be used in its original form since the correspondence is not one-to-one. This is the reason why we should turn our attention to tree-ordinals instead of (normal) ordinals below ε_0.

Definition 3.2.3. The set Ω of countable *tree-ordinals* is generated inductively as follows:

- $0 \in \Omega$;
- if $\alpha \in \Omega$, then $\alpha + 1 := \alpha \cup \{\alpha\} \in \Omega$;
- if $\alpha_n \in \Omega$ for all $n \in \mathbb{N}$, then $\alpha := \langle \alpha_n \rangle_{n \in \mathbb{N}} \in \Omega$.

λ will always denote a *limit* $\lambda = \langle \lambda_n \rangle_n := \langle \lambda_n \rangle_{n\in\mathbb{N}}$. Addition, multiplication, and exponentiation are defined as usual:

- **Addition:** $\alpha + 0 := \alpha$; $\alpha + (\beta + 1) := (\alpha + \beta) + 1$; $\alpha + \lambda := \langle \alpha + \lambda_n \rangle_n$
- **Multiplication:** $\alpha \cdot 0 := 0$; $\alpha \cdot (\beta + 1) := (\alpha \cdot \beta) + \alpha$; $\alpha \cdot \lambda := \langle \alpha \cdot \lambda_n \rangle_n$
- **Exponentiation:** $\alpha^0 := 1$; $\alpha^{(\beta+1)} := \alpha^\beta \cdot \alpha$; $\alpha^\lambda := \langle \alpha^{\lambda_n} \rangle_n$

We define also a set $\mathbb{T} \subseteq \Omega$ of tree-ordinals which correspond to (normal) ordinals up to ε_0. Set

$$n := 0 + \underbrace{1 + \cdots + 1}_{n\,\text{times}} \quad \text{and} \quad \omega := \langle 1 + n \rangle_n .$$

Definition 3.2.4. $\mathbb{T}$ is defined inductively as follows:

- $0 \in \mathbb{T}$;
- if $\alpha_0, \ldots, \alpha_n \in \mathbb{T}$, then also $\omega^{\alpha_0} + \cdots + \omega^{\alpha_n} \in \mathbb{T}$.

Note that each tree-ordinal in $\mathbb{T}$ represents a unique determined (ordered) tree figure other than the ordinals in the Cantor system. And though there is a canonical fundamental sequence for each limit tree-ordinal, we shall make some modifications for technical reasons. This modifications will have no significant effect on the fast growing hierarchy we consider. We write $\alpha \cdot m$ for $\underbrace{\alpha + \cdots + \alpha}_{m\text{ times}}$.

Definition 3.2.5 (Fundamental sequences for tree-ordinals)**.** Let $\alpha \in \mathbb{T}$.

- If $\alpha = 0$, then $\alpha[k] = 0$.
- If $\alpha = \beta + 1$, then $\alpha[k] = \beta$.
- If $\alpha = n + \omega$ for some $n \in \mathbb{N}$, then $\alpha[k] = n + k + 1$.
- If $\alpha = \beta + \omega$ and $\beta \neq n$ for any $n \in \mathbb{N}$, then $\alpha[k] = \beta + k + 2$.
- If $\alpha = \beta + \omega^{\gamma+1}$ and $\gamma \neq 0$, then $\alpha[k] = \beta + \omega^\gamma \cdot (k + 1) + 1$.
- If $\alpha = \beta + \omega^\lambda$ and λ a limit, then $\alpha[k] = \beta + \omega^{\lambda[k]}$.

Definition 3.2.6. The *sub-tree ordering* $\prec$ is the transitive closure of the rule:

$$\alpha[m] \prec \alpha \text{ for all } \alpha \in \mathbb{T} \setminus \{0\} \text{ and } m \in \mathbb{N}.$$

Theorem 3.2.7. *The set $\{\beta \,|\, \beta \prec \alpha\}$ is well-ordered by $\prec$ and of order type less than ε_0.*

Proof. Cf. Fairtlough and Wainer [16]. □

Now we are going to establish an one-to-one and onto correspondence between worms and tree-ordinals from $\mathbb{T}$. We use the same notion o as in the case of the correspondence between S and ε_0 since the former is indeed a kind of extension of the latter.

Definition 3.2.8. $o\colon W \to \mathbb{T}$ is defined recursively as follows:

- $o(0^k) := k$;
- if $\alpha = \alpha_0 0 \cdots 0 \alpha_n$, where all $\alpha_i \in W_1$ and not all of them empty, then

$$o(\alpha_0 0 \alpha_1 0 \cdots 0 \alpha_n) := \omega^{o(\alpha_0^-)} + \cdots + \omega^{o(\alpha_n^-)}.$$

The function $o\colon W \to \mathbb{T}$ is one-to-one and onto, since $g\colon \mathbb{T} \to W$ defined by

- $g(k) := 0^k$;
- if $\alpha = \omega^{\alpha_0} + \cdots + \omega^{\alpha_n}$ and $\alpha_i \neq 0$ for some $i \leq n$, then

$$g(\alpha) = g(\alpha_0)^+ 0 \cdots 0 g(\alpha_n)^+,$$

is obviously the inverse function of o, where β^+ is obtained from $\beta \in W$ by replacing every letter m with $m+1$.

Lemma 3.2.9. *Let α, $\beta \in W$. Then*

$$o(\alpha 0 \beta) = \begin{cases} o(\alpha) + 1 + o(\beta) & \textit{if } 0^m \in \{o(\alpha), o(\beta)\} \textit{ for some } m \in \mathbb{N}, \\ o(\alpha) + o(\beta) & \textit{otherwise.} \end{cases}$$

Proof. (i) Let $\alpha = 0^m$ and $\beta = 0^n$ for some m, $n \in \omega$. Then

$$o(\alpha 0 \beta) = o(0^{m+1+n}) = m + 1 + n = o(\alpha) + 1 + o(\beta).$$

(ii) Let $\alpha = 0^m$ for some $m \in \omega$ and $\beta = \beta_0 0 \cdots 0 \beta_n$, where all $\beta_j \in W_1$ and not all of them empty. Then

$$\begin{aligned} o(\alpha 0 \beta) &= o(0^{m+1} \beta_0 0 \cdots 0 \beta_n) \\ &= m + 1 + \omega^{o(\beta_0^-)} + \cdots + \omega^{o(\beta_n^-)} \\ &= o(\alpha) + 1 + o(\beta). \end{aligned}$$

(iii) Similar for the case that $\beta = 0^n$ for some $n \in \omega$ and $\alpha = \alpha_0 0 \cdots 0 \alpha_m$, where all $\alpha_i \in W_1$ and not all of them empty.

(iv) Let $\alpha = \alpha_0 0 \cdots 0 \alpha_m$ and $\beta = \beta_0 0 \cdots 0 \beta_n$, where all α_i, $\beta_j \in W_1$ and there are some $\alpha_i \neq \varnothing$ and $\beta_j \neq \varnothing$. Then

$$\begin{aligned} o(\alpha 0 \beta) &= o(\alpha_0 0 \cdots 0 \alpha_m 0 \beta_0 0 \cdots 0 \beta_n) \\ &= \omega^{o(\alpha_0^-)} + \cdots + \omega^{o(\alpha_m^-)} + \omega^{o(\beta_0^-)} + \cdots + \omega^{o(\beta_n^-)} \\ &= o(\alpha) + o(\beta). \end{aligned}$$

The proof is now complete. □

If needed, this lemma will be used tacitly. Let β^{-n} be obtained from $\beta \in W_n$ by replacing every letter m with $m - n$. β^{+n} is similarly defined for $\beta \in W$. Below we write $o(\beta) < \omega$ for $o(\beta) = k$ for some $k \in \mathbb{N}$ and $o(\beta) \geq \omega$ otherwise.

Theorem 3.2.10. *$o(\alpha[\![k]\!]) = o(\alpha)[k]$ for all $\alpha \in W$.*

Proof. There is nothing to prove if $\alpha = \varnothing$. If $\alpha = \beta 0$, then $\alpha[\![k]\!] = \beta$ and

$$(o(\alpha))[k] = (o(\beta) + 1)[k] = o(\beta) = o(\alpha[\![k]\!]).$$

Now let $\alpha = \beta m \gamma \langle n+1 \rangle$, where $m \leq n$ and $\gamma \in W_{n+1}$.

1. case: $\gamma = \varnothing$ and $\beta m = \varnothing$, i.e. $\alpha = \langle n+1 \rangle$ and $\alpha[\![k]\!] = n^{k+1}$.

$$o(\alpha)[k] = \omega_{n+1}[k] = \omega_n(k+1) = o(n^{k+1}) = o(\alpha[\![k]\!]).$$

2. case: $\gamma = \varnothing$ and $\beta m \neq \varnothing$. We use an induction on n.

(2.a) m is the minimum of the occurrences in βm.

$$\begin{aligned} o(\alpha) &= o(\beta m \langle n+1 \rangle) \\ &= \omega_m(o(\beta^{-m} 0 \langle n+1-m \rangle)) \\ &= \begin{cases} \omega_m(o(\beta^{-m}) + \omega_{n+1-m}) & , o(\beta^{-m}) \geq \omega, \\ \omega_m(o(\beta^{-m}) + 1 + \omega_{n+1-m}) & , o(\beta^{-m}) < \omega. \end{cases} \end{aligned}$$

And

$$o(\alpha)[k] = \begin{cases} \omega_m(o(\beta^{-m}) + \omega_{n-m}(k+1)) & , o(\beta^{-m}) \geq \omega,\ n > m, \\ \omega_m(o(\beta^{-m}) + k + 2) & , o(\beta^{-m}) \geq \omega,\ n = m, \\ \omega_m(o(\beta^{-m}) + 1 + \omega_{n-m}(k+1)) & , o(\beta^{-m}) < \omega. \end{cases}$$

On the other hand,

$$\begin{aligned}
o(\alpha[\![k]\!]) &= o(\beta m n^{k+1}) \\
&= \omega_m(o(\beta^{-m}0\langle n-m\rangle^{k+1})) \\
&= \begin{cases} \omega_m(o(\beta^{-m}) + o(\langle n-m\rangle^{k+1})) & , o(\beta^{-m}) \geq \omega,\ n > m, \\ \omega_m(o(\beta^{-m}) + 1 + k + 1) & , o(\beta^{-m}) \geq \omega,\ n = m, \\ \omega_m(o(\beta^{-m}) + 1 + o(\langle n-m\rangle^{k+1})) & , o(\beta^{-m}) < \omega, \end{cases} \\
&= \begin{cases} \omega_m(o(\beta^{-m}) + \omega_{n-m}(k+1)) & , o(\beta^{-m}) \geq \omega,\ n > m, \\ \omega_m(o(\beta^{-m}) + k + 2) & , o(\beta^{-m}) \geq \omega,\ n = m, \\ \omega_m(o(\beta^{-m}) + 1 + \omega_{n-m}(k+1)) & , o(\beta^{-m}) < \omega. \end{cases}
\end{aligned}$$

Note that the case $n = 0$ is also proved since m should then be 0.

(2.b) $n > 0$, $m > p$ and $\beta = \beta_0 p \cdots p \beta_{t+1}$, where all $\beta_i \in W_{p+1}$.

$$\begin{aligned}
o(\alpha) &= o(\beta_0 p \cdots p \beta_{t+1} m \langle n+1\rangle) \\
&= \omega_p(o(\beta_0^{-p}0 \cdots 0 \beta_{t+1}^{-p})\langle m-p\rangle\langle n+1-p\rangle) \\
&= \omega_p(\omega^{o(\beta_0^{-p-1})} + \cdots + \omega^{o(\beta_t^{-p-1})} + \omega^{o(\beta_{t+1}^{-p-1}\langle m-p-1\rangle\langle n-p\rangle)})
\end{aligned}$$

Since $o(\beta_{t+1}^{-p-1}\langle m-p-1\rangle\langle n-p\rangle)$ is a limit, we have

$$\begin{aligned}
o(\alpha)[k] &= \omega_p(\omega^{o(\beta_0^{-p-1})} + \cdots + \omega^{o(\beta_t^{-p-1})} + \omega^{o(\beta_{t+1}^{-p-1}\langle m-p-1\rangle\langle n-p\rangle)[k]}) \\
&\overset{i.h.}{=} \omega_p(\omega^{o(\beta_0^{-p-1})} + \cdots + \omega^{o(\beta_t^{-p-1})} + \omega^{o((\beta_{t+1}^{-p-1}\langle m-p-1\rangle\langle n-p\rangle)[\![k]\!])}) \\
&= \omega_p(\omega^{o(\beta_0^{-p-1})} + \cdots + \omega^{o(\beta_t^{-p-1})} + \omega^{o(\beta_{t+1}^{-p-1}\langle m-p-1\rangle\langle n-p-1\rangle^{k+1})})
\end{aligned}$$

On the other hand,

$$\begin{aligned}
o(\alpha[\![k]\!]) &= o(\beta_0 p \cdots p \beta_{t+1} m n^{k+1}) \\
&= \omega_p(o(\beta_0^{-p}0 \cdots 0 \beta_{t+1}^{-p})\langle m-p\rangle\langle n-p\rangle^{k+1}) \\
&= \omega_p(\omega^{o(\beta_0^{-p-1})} + \cdots + \omega^{o(\beta_t^{-p-1})} + \omega^{o(\beta_{t+1}^{-p-1}\langle m-p-1\rangle\langle n-p-1\rangle^{k+1})})
\end{aligned}$$

3. case: $\gamma \neq \varnothing$ and $\beta m = \varnothing$, i.e. $\alpha = \gamma\langle n+1\rangle$, $\gamma \in W_{n+1}$.

$$\begin{aligned}
o(\alpha) &= \omega_{n+1}(o(\gamma^{-n-1}0)) \\
&= \omega_{n+1}(o(\gamma^{-n-1}) + 1)
\end{aligned}$$

Since $o(\gamma^{-n-1}) > 0$, we have

$$o(\alpha)[k] = \omega_n(\omega^{o(\gamma^{-n-1})} \cdot (k+1) + 1).$$

On the other hand,

$$\begin{aligned}
o(\alpha[\![k]\!]) &= o((\gamma n)^{k+1}) \\
&= \omega_n(o((\gamma^{-n}0)^{k+1})) \\
&= \omega_n(\omega^{o(\gamma^{-n-1})} \cdot (k+1) + 1).
\end{aligned}$$

4. case: $\gamma \neq \varnothing$ and $\beta m \neq \varnothing$. The claim will be shown by induction on n.

(4.a) m is the minimum of the occurrences in βm.

$$\begin{aligned}
o(\alpha) &= o(\beta m\gamma\langle n+1\rangle)\\
&= \omega_m(o(\beta^{-m}0\gamma^{-m}\langle n+1-m\rangle))\\
&= \begin{cases}\omega_m(o(\beta^{-m})+\omega_{n+1-m}(o(\gamma^{-n-1}0))) & ,o(\beta^{-m})\geq\omega,\\ \omega_m(o(\beta^{-m})+1+\omega_{n+1-m}(o(\gamma^{-n-1}0))) & ,o(\beta^{-m})<\omega.\end{cases}\\
&= \begin{cases}\omega_m(o(\beta^{-m})+\omega_{n+1-m}(o(\gamma^{-n-1})+1)) & ,o(\beta^{-m})\geq\omega,\\ \omega_m(o(\beta^{-m})+1+\omega_{n+1-m}(o(\gamma^{-n-1})+1)) & ,o(\beta^{-m})<\omega.\end{cases}
\end{aligned}$$

Since $o(\gamma^{-n-1})>0$, we have

$$o(\alpha)[k] = \begin{cases}\omega_m(o(\beta^{-m})+\omega_{n-m}(\omega^{o(\gamma^{-n-1})}\cdot(k+1)+1)) & ,o(\beta^{-m})\geq\omega,\\ \omega_m(o(\beta^{-m})+1+\omega_{n-m}(\omega^{o(\gamma^{-n-1})}\cdot(k+1)+1)) & ,o(\beta^{-m})<\omega.\end{cases}$$

On the other hand,

$$\begin{aligned}
o(\alpha[\![k]\!]) &= o(\beta m(\gamma n)^{k+1})\\
&= \omega_m(o(\beta^m0(\langle n-m\rangle\gamma^{-m})^{k+1}))\\
&= \begin{cases}\omega_m(o(\beta^{-m})+o((\gamma^{-m})^{k+1}\langle n-m\rangle)) & ,\ o(\beta^{-m})\geq\omega,\\ \omega_m(o(\beta^{-m})+1+o((\gamma^{-m})^{k+1}\langle n-m\rangle)) & ,\ o(\beta^{-m})<\omega,\end{cases}\\
&= \begin{cases}\omega_m(o(\beta^{-m})+\omega_{n-m}(o((\gamma^{-n})^{k+1}0))) & ,\ o(\beta^{-m})\geq\omega,\\ \omega_m(o(\beta^{-m})+1+\omega_{n-m}(o((\gamma^{-n})^{k+1}0))) & ,\ o(\beta^{-m})<\omega,\end{cases}\\
&= \begin{cases}\omega_m(o(\beta^{-m})+\omega_{n-m}(\omega^{o(\gamma^{-n-1})}\cdot(k+1)+1)) & ,\ o(\beta^{-m})\geq\omega,\\ \omega_m(o(\beta^{-m})+1+\omega_{n-m}(\omega^{o(\gamma^{-n-1})}\cdot(k+1)+1)) & ,\ o(\beta^{-m})<\omega.\end{cases}
\end{aligned}$$

Note that the case $n=0$ is also proved since m should then be 0.

(4.b) $n>0$, $m>p$ and $\beta=\beta_0p\cdots p\beta_{t+1}$, where all $\beta_i\in W_{p+1}$.

$$\begin{aligned}
o(\alpha) &= o(\beta_0p\cdots p\beta_{t+1}m\gamma\langle n+1\rangle)\\
&= \omega_p(o(\beta_0^{-p}0\cdots 0\beta_{t+1}^{-p}\langle m-p\rangle\gamma^{-p}\langle n+1-p\rangle))\\
&= \omega_p(\omega^{o(\beta_0^{-p-1})}+\cdots+\omega^{o(\beta_t^{-p-1})}+\omega^{o(\beta_{t+1}^{-p-1}\langle m-p-1\rangle\gamma^{-p-1}\langle n-p\rangle)})
\end{aligned}$$

Since $o(\beta_{t+1}^{-p-1}\langle m-p-1\rangle\gamma^{-p-1}\langle n-p\rangle)$ is a limit, we have

$$\begin{aligned}
o(\alpha)[k] &= \omega_p(\omega^{o(\beta_0^{-p-1})}+\cdots+\omega^{o(\beta_t^{-p-1})}+\omega^{o(\beta_{t+1}^{-p-1}\langle m-p-1\rangle\gamma^{-p-1}\langle n-p\rangle)[k]})\\
&\overset{i.h.}{=} \omega_p(\omega^{o(\beta_0^{-p-1})}+\cdots+\omega^{o(\beta_t^{-p-1})}+\omega^{o((\beta_{t+1}^{-p-1}\langle m-p-1\rangle\gamma^{-p-1}\langle n-p\rangle)[\![k]\!])})\\
&= \omega_p(\omega^{o(\beta_0^{-p-1})}+\cdots+\omega^{o(\beta_t^{-p-1})}+\omega^{o(\beta_{t+1}^{-p-1}\langle m-p-1\rangle(\gamma^{-p-1}\langle n-p-1\rangle)^{k+1})})
\end{aligned}$$

On the other hand,

$$\begin{aligned} o(\alpha[\![k]\!]) &= o(\beta_0 p \cdots p \beta_{t+1} m (\gamma n)^{k+1}) \\ &= \omega_p(o(\beta_0^{-p} 0 \cdots 0 \beta_{t+1}^{-p} \langle m-p \rangle (\gamma^{-p} \langle n-p \rangle)^{k+1})) \\ &= \omega_p(\omega^{o(\beta_0^{-p-1})} + \cdots + \omega^{o(\beta_t^{-p-1})} + \omega^{o(\beta_{t+1}^{-p-1} \langle m-p-1 \rangle (\gamma^{-p-1} \langle n-p-1 \rangle)^{k+1})}) \end{aligned}$$

This completes the proof. □

Remember that an ordinal $\alpha \neq 0$ below ε_0 is said to be in Cantor normal form if $\alpha = \omega^{\alpha_0} + \cdots + \omega^{\alpha_n}$ and $\alpha > \alpha_0 \geq \cdots \geq \alpha_n$. We can demand the same property from every α_i and so on. Furthermore, if we make no difference between an ordinal below ε_0 in Cantor normal form and an tree-ordinal which has the same tree figure, we can specify a set $\mathbb{B}$ of all tree-ordinals in so-called Cantor normal form. This implies in turn that $\mathbb{B}$ corresponds isomorphically to the set $NF \subseteq S$ of all words in normal forms.

Let $NF(W) \subseteq W$ be the set of all worms which are converses of a word in NF. The worms in $NF(W)$ are also said to be in Cantor normal form and the set $NF(W)$ is isomorphic to ε_0.

Lemma 3.2.11. *$NF(W)$ can be characterized inductively as follows:*

- *$\varnothing$ and any worm of length 1 belong to $NF(W)$;*
- *assume that the length of the worm α is larger than 1 and $\alpha = \alpha_0 0 \cdots 0 \alpha_n$, where all $\alpha_i \in W_1$. Then $\alpha \in NF(W)$ iff all $\alpha_i^- \in NF(W)$ and $o(\alpha_{j+1}^-) \leq o(\alpha_j^-)$ for all $j < n$.*

Note that $o(\alpha) \in \mathbb{B}$ for every $\alpha \in NF(W)$, so we might talk about the linear order $<$ of ordinals. It is also obvious that $\alpha[\![k]\!] \in NF(W)$ for all $\alpha \in NF(W)$ and $k \in \mathbb{N}$. Let $\prec_0$ be the well-order on $NF(W)$ induced by the isomorphism o.

Lemma 3.2.12. *$o \restriction NF(W) \colon NF(W) \to \mathbb{B}$ is an order-preserving isomorphism.*

Having established an correspondence between W and $\mathbb{T}$ (resp. between $NF(W)$ and $\mathbb{B}$) it is now obvious that the Worm principle is the counterpart of the Hydra battle game on the tree-ordinals in $\mathbb{T}$ (resp. on the ordinals up to ε_0). On the other hand, the Hydra battle game has a direct connection to the Hardy-Wainer hierarchy. It is a folklore that the Hardy-Wainer hierarchy up to ε_0 features exactly the provably recursive functions in PA. Fairtlough and Wainer [16] showed that an similar characterization of provably recursive functions in PA is possible using the tree-ordinals from $\mathbb{T}$. Furthermore, Weiermann [58] made an refinement in such a way that how fast heads of a hydra should be multiplied at cutting off the right-most head, so that the Hydra battle game on the ordinals up to ε_0 remains unprovable in PA. Using the same idea we show that an analogous process is possible using the tree-ordinals from $\mathbb{T}$.

First we recall some well-known definitions and lemmata from subrecursive hierarchy theory based on the fundamental sequences defined in Definition 3.2.5. Let f, g range over unary arithmetical functions, k, n, x over $\mathbb{N}$, and α, β, λ, etc. over $\mathbb{T}$.

Definition 3.2.13. Let $\lambda \in Lim$.

(i) $P^f_x 0 := 0$, $P^f_x(\alpha+1) := \alpha$ and $P^f_x\lambda := P^f_x(\lambda[f(x)])$.

(ii) $Q^f_x 0 := 0$, $Q^f_x(\alpha+1) := \alpha$ and $Q^f_x\lambda := \lambda[f(x)]$.

(iii) Let $R \in \{P, Q\}$.

- $R_x\alpha := R^{id}_x\alpha$.
- $\alpha \succ^{R,n}_f \beta$ if $\beta = R^f_n \cdots R^f_1\alpha$.
- $\alpha \succ^R_f \beta$ if $\beta = R^f_n \cdots R^f_1\alpha$ for some positive n.
- $\alpha \succ^R_k \beta$ if $\alpha \succ^R_f \beta$, where $f \equiv k$.
- $\alpha \succeq^R_k \beta$ if $\alpha \succ^R_k \beta$ or $\alpha = \beta$.

(iv) $G_x(0) := 0$, $G_x(\alpha+1) = G_x(\alpha)+1$ and $G_x(\lambda) := G_x(\lambda[x])$.

(v) $H^f_0(x) := x$, $H^f_{\alpha+1}(x) := H^f_\alpha(x+1)$ and $H^f_\lambda(x) := H^f_{\lambda[f(x)]}(x)$.

(vi) $H_\alpha := H^{id}_\alpha$.

(vii) $\mathrm{mc}(m) := m$ and $\mathrm{mc}(\alpha) := \max\{m_1, \ldots, m_n, \mathrm{mc}(\alpha_1), \ldots, \mathrm{mc}(\alpha_n)\}$, where $\alpha = \omega^{\alpha_1} \cdot m_1 + \cdots + \omega^{\alpha_n} \cdot m_n$ such that $\alpha_i > \alpha_{i+1}$ for each $i < n$.

(viii) $\alpha_0(\beta) := \beta$, $\alpha_{n+1}(\beta) := \alpha^{\alpha_n(\beta)}$ and $\alpha_n := \alpha_n(1)$.

(ix) $\varepsilon_0 := \langle\omega_{n+1}\rangle_n$ and $\varepsilon_0[k] := \omega_{k+1}$.

(x) $H_{\varepsilon_0}(x) := H_{\varepsilon_0[x]}(x)$.

Note that $G_1(\omega_{n+1}) \geq 2_n(2)$. In fact, $G_1(\omega_2) > 2^2$. It is another difference compared to the hierarchy with ordinals up to ε_0.

Theorem 3.2.14. *Let* α, $\beta \in \mathbb{T}$.

(i) G_α *is increasing (strictly if* α *infinite), and if* $\beta \prec \alpha[n]$*, then* $G_\beta(n) < G_\alpha(n)$ *for all* n *and* G_α *eventually dominates* G_β.

(ii) H_α *is strictly increasing, and if* $\beta \prec \alpha[n]$*, then* $H_\beta(n) < H_\alpha(n)$ *for all* n *and* H_α *eventually dominates* H_β.

(iii) H_α *is provably recursive in* PA.

(iv) *Every provably recursive function in* PA *is dominated by* H_α *for some* α.

(v) H_{ε_0} *is not provably recursive in* PA.

Proof. See [16]. □

Given $R \in \{P, Q\}$ set $R_x^{(0)}\alpha := \alpha$ and $R_x^{(i+1)}\alpha := R_x R_x^{(i)}\alpha$. For $\alpha \in \mathbb{T}$ and $n \in \mathbb{N}$ let $\alpha[\omega := n]$ be the natural number obtained by replacing every occurrence of ω in α with n.

Lemma 3.2.15. *Let* $R \in \{P, Q\}$.

(i) $R_x(\alpha + \beta) = \alpha + R_x\beta$ *for* $\beta \neq 0$.

(ii) *If* $\alpha \succ_x^P \beta$ *then* $\alpha \succ_x^Q \beta$.

(iii) *If* $\alpha \succ_x^R \beta$ *then* $\gamma + \alpha \succ_x^R \gamma + \beta$.

(iv) *If* $\alpha \succ_x^R \beta$ *then* $\omega^\alpha \succ_x^R \omega^\beta$.

(v) *If* λ *is a limit then* $\lambda[x+1] \succ_0^Q \lambda[x]$.

(vi) *If* λ *is a limit then* $\lambda[x+1] \succ_{x+1}^Q \lambda[x] + 1$.

(vii) *If* $x > 0$ *then* $\omega^{\alpha+1} \succ_x^Q \omega^\alpha + \omega^\alpha$.

(viii) *If* $x \geq 0$ *then* $\omega^{\alpha+1} \succ_x^Q \omega^\alpha + 1$.

(ix) *If* $x > 0$ *then* $\omega_{n+1}(\alpha + 1) \succ_x^Q \omega_{n+1}(\alpha) + \omega_{n+1}$.

(x) *If* $\alpha > 0$ *then* $\alpha \succeq_{x+1}^Q P_x\alpha + 1$.

(xi) $\alpha \succ_x^Q P_x\alpha$.

(xii) *If* f, g *are increasing, where* $g(i) \leq f(i)$ *for all* i, *and* $\alpha \succ_g^{R,m} \beta$, *then* $\alpha \succ_f^{R,n} \beta$ *for some* $n \geq m$.

(xiii) *If* $\alpha \succeq_x^Q \beta \succ_x^{P,m} \gamma$ *then* $\alpha \succ_x^{P,n} \gamma$ *for some* $n \geq m$.

(xiv) *There are at most* $G_{x+1}(\alpha)$ *elements in* $\{\beta \prec \alpha \colon \mathrm{mc}(\beta) \leq x+1\}$.

(xv) $\alpha[\omega := x+1] \leq G_x(\alpha) \leq \alpha[\omega := x+2]$.

(xvi) $G_x(\alpha) = \min\{i \colon P_x^{(i)}\alpha = 0\}$.

(xvii) $H_\alpha(x) = \min\{i \colon P_{i+x-1} \cdots P_x\alpha = 0\} + x$, *so*

$$H_\alpha(x) = \min\{i \geq x \colon P_i \cdots P_x\alpha = 0\} + 1.$$

Proof. (i) ~ (xvi) are more or less obvious. (xvii) is proved in [16]. □

Given $n \in \mathbb{N}$ set $g_n(i) := |i|_n$ and given $\alpha \in \mathbb{T} \cup \{\varepsilon_0\}$ set $f_\alpha(i) := |i|_{H_\alpha^{-1}(i)}$.

Lemma 3.2.16. *Given a limit $\lambda \in \mathbb{T}$ let $\beta := \omega_{n+1}(\lambda) + \omega_{n+1}$. Then there exists an $i \geq H_\lambda(1)$ such that $\beta \succ_{g_n}^{P,i} \omega_{n+1}(0)$.*

Proof. Let $L := H_\lambda(1) - 1 = \min\{i \colon P_i \cdots P_1\lambda = 0\}$. By definition we have $g_n(i) \geq 1$ for all i. Further we obtain

$$\begin{array}{lll} \beta & = & \omega_{n+1}(\lambda) + \omega_{n+1} \\ & \succ_1^P & \omega_{n+1}(\lambda) + P_1\omega_{n+1} \\ & \succ_1^P & \omega_{n+1}(\lambda) + P_1P_1\omega_{n+1} \\ & \succ_1^P & \cdots \end{array}$$

Hence, there exists $i_0 \geq 2_n(2)$ such that $\beta \succ_1^{P,i_0} \omega_{n+1}(\lambda)$ since

$$\min\{n \colon P_1^{(n)}\omega_{n+1} = 0\} = G_1\omega_{n+1} \geq 2_n(2).$$

And for $i \geq i_0$ we have $2 \leq g_n(i)$. In addition, we have

$$\begin{array}{lll} \omega_{n+1}(\lambda) & \succ_2^P & \omega_{n+1}(P_1\lambda + 1) \\ & \succ_2^Q & \omega_{n+1}(P_1\lambda) + \omega_{n+1} \\ & \succ_2^P & \omega_{n+1}(P_1\lambda) + P_2\omega_{n+1} \\ & \succ_2^P & \omega_{n+1}(P_1\lambda) + P_2P_2\omega_{n+1} \\ & \succ_2^P & \cdots \end{array}$$

Therefore, there exists $i_0 \geq 2_n(2)$ such that $\beta \succ_1^{P,i_0} \omega_{n+1}(\lambda)$ since

$$\min\{k \colon P_2^{(k)}\omega_{n+1} = 0\} = G_2\omega_{n+1} \geq 3_n(3).$$

This process shows that given $k \leq L$ there is a sequence $\langle i_\ell \rangle_{\ell \leq k}$ such that

$$i_\ell \geq (\ell + 2)_n(\ell + 2)$$

for all $\ell \leq k$ and

$$\beta \succ_{g_n}^{P,i_0+i_1+\cdots+i_k} \omega_{n+1}(P_k \cdots P_1\lambda).$$

The assertion follows now from the fact that $i_0 + \cdots + i_L \geq L + 1 = H_\lambda(1)$. ☐

Lemma 3.2.17. *Given $n \geq 2$ and $\alpha := \omega_{n+1}(\omega_n) + \omega_{n+1}$ there is $\delta \geq \omega_{n+1}(0)$ such that $\alpha \succ_{f_{\varepsilon_0}}^{P,i} \delta$ for some $i \geq H_{\omega_n}(1)$.*

Proof. If $k \leq H_{\omega_n}(1) =: i_0$, then $H_{\varepsilon_0}^{-1}(k) \leq H_{\varepsilon_0}^{-1}(i_0) \leq n$. Hence

$$f_{\varepsilon_0}(k) = |k|_{H_{\varepsilon_0}^{-1}(k)} \geq |k|_n = g_n(k).$$

By Lemma 3.2.16 there exists $\delta > \omega_{n+1}(0)$ such that $\alpha \succ_{g_n}^{P,i_0} \delta$. This implies that there is $i \geq i_0$ such that $\alpha \succ_{f_{\varepsilon_0}}^{P,i} \delta$. ☐

Theorem 3.2.18. *Let n be a natural number.*

(i) $\mathrm{PA} \nvdash \forall k\, \exists m\, Q_m^{f_{\varepsilon_0}} \cdots Q_1^{f_{\varepsilon_0}} \omega_k = 0.$

(ii) $\mathrm{PA} \nvdash \forall k\, \exists m\, Q_m^{g_n} \cdots Q_1^{g_n} \omega_k = 0.$

Proof. It follows from the fact that the function $\lambda i \,.\, H_{\omega_i}(1)$ is not provably recursive in PA. Cf. [16]. □

Theorem 3.2.19. *Let $\alpha \in \mathbb{T}$.*

$$\mathrm{PRA} \vdash \forall k\, \exists m\, Q_m^{f_\alpha} \cdots Q_1^{f_\alpha} \omega_k = 0.$$

Proof. Assume that k is large enough. How large k should be will be obvious from the context. We claim that

$$Q_m^{f_\alpha} \cdots Q_1^{f_\alpha} \omega_k = 0,$$

where $m := 2_{H_{\omega^\alpha \cdot 2}(k)}$. Assume otherwise. Since

$$\mathrm{mc}(\omega_k[f_\alpha(1)] \cdots [f_\alpha(i)]) \leq f_\alpha(i) + 2 \leq f_\alpha(m) + 2$$

for every $i \leq m$ we have by Lemma 3.2.15.(xiv)

$$m \leq G_{2+f_\alpha(m)}(\omega_k).$$

By Lemma 3.2.15.(xv)

$$\begin{aligned} m &\leq (4 + |2_{H_{\omega^\alpha \cdot 2}(k)}|_{H_\alpha^{-1}(2_{H_{\omega^\alpha \cdot 2}(k)})})_k(1) \\ &\leq (4 + |2_{H_{\omega^\alpha \cdot 2}(k)}|_{H_{\omega^\alpha}(k)})_k(1) \\ &= (5 + 2_{H_{\omega^\alpha \cdot 2}(k) - H_{\omega^\alpha}(k)})_k(1) \\ &< 2_{H_{\omega^\alpha \cdot 2}(k)} \; = \; m \end{aligned}$$

for sufficiently large k. Contradiction! □

Note that, for any α, $\beta \in \mathbb{T}$, there is $k \in \mathbb{N}$ such that

$$\min\{m \colon Q_m^{f_\alpha} \cdots Q_1^{f_\alpha} \beta = 0\} \leq \min\{m \colon Q_m^{f_\alpha} \cdots Q_1^{f_\alpha} \omega_k = 0\}.$$

The existence of such k is provable in PA.

Theorem 3.2.20. *Let $n \in \mathbb{N}$ and $\alpha \in \mathbb{T} \cup \{\varepsilon_0\}$.*

(i) $\mathrm{PA} \nvdash \mathbf{EWD}(|\cdot|_n)$.

(ii) $\mathrm{PA} \nvdash \mathbf{EWD}(|\cdot|_{H_\alpha^{-1}(\cdot)})$ *iff* $\alpha = \varepsilon_0$.

Proof. Obvious by Theorem 3.2.10, Theorem 3.2.18, and Theorem 3.2.19. □

Remark 3.2.21. Of course, we can talk of the same independence results concerning the Worm principles with the worms from $NF(W)$ only.

3.3 Schütte-Simpson's ordinal notation system

Schütte and Simpson [47] introduced another interesting ordinal notation system $\pi_0(\omega)$ for ε_0. It is a segment of $\pi(\omega)$ defined by letting out the addition and the function $\lambda\alpha\,.\,\omega^\alpha$ in the construction of the ordinal notation system developed by Buchholz [10].

At first glance, the new defined ordinal terms seem so artificial that it would make no sense to say about their meaning. Hence it is all the more meaningful to see a canonical correspondence between them and worms which in turn can be interpreted in terms of some graded provability algebras.

In the following, we will proceed as in [47]. However with different proofs in the sense that we don't refer to the original collapsing functions. This seems to be somewhat more technical, but has the advantage that one can easily see the correspondence between Schütte-Simpson's system and Beklemishev's one.

The small Greek letters α, β, $\gamma, \ldots$ range over ordinals. We set $\Omega_0 := 0$ and, for $i > 0$, Ω_i the i-th regular ordinal and $\Omega_\omega := \sup\{\Omega_i \colon i < \omega\}$.

Definition 3.3.1. We define $B_i^m(\alpha)$, $B_i(\alpha)$ and $\pi_i(\alpha)$ (by the main induction on α and the subsidiary induction on m):

(B1) if $\gamma = 0$ or $\gamma < \Omega_i$, then $\gamma \in B_i^m(\alpha)$;

(B2) if $i \le j$, $\beta < \alpha$, $\beta \in B_j(\beta)$, and $\beta \in B_i^m(\alpha)$, then $\pi_j\beta \in B_i^{m+1}(\alpha)$;

(B3) $B_i(\alpha) := \cup\{B_i^m(\alpha) \colon m < \omega\}$;

(B4) $\pi_i\alpha := \min\{\eta \colon \eta \notin B_i(\alpha)\}$.

Lemma 3.3.2.

(i) *If $k < m$, then $B_i^k(\alpha) \subseteq B_i^m(\alpha)$.*

(ii) *If $i \le j$ and $\alpha \le \beta$, then $B_i(\alpha) \subseteq B_j(\beta)$, $\pi_i\alpha \le \pi_j\beta$.*

(iii) $\Omega_i \le \pi_i\alpha < \Omega_{i+1}$.

(iv) *If $\gamma \in B_i(\alpha)$ and $\gamma < \Omega_{i+1}$, then $\gamma < \pi_i\alpha$.*

(v) *If $\alpha \in B_i(\alpha)$ and $\alpha < \beta$, then $\pi_i\alpha < \pi_i\beta$.*

(vi) *If $\alpha \in B_i(\alpha)$, $\beta \in B_i(\beta)$, and $\pi_i\alpha = \pi_i\beta$, then $\alpha = \beta$.*

Proof. Standard. Cf. [47]. □

Definition 3.3.3. $\pi(\omega)$ is defined as follows:

- $0 \in \pi(\omega)$;
- if $\alpha \in \pi(\omega)$ and $\alpha \in B_i(\alpha)$ then $\pi_i\alpha \in \pi(\omega)$.

To see that $\pi(\omega)$ may be considered as a primitive recursive ordinal notation system we must be able to decide the relation $\alpha \in B_i(\alpha)$ for $\alpha \in \pi(\omega)$. For this purpose we introduce an auxiliary concept of coefficients sets. The idea stems from Rathjen and Weiermann [43].

Definition 3.3.4. Inductive definition of a set of ordinals $K_i\alpha$ for $\alpha \in \pi(\omega)$.

(i) $K_i(0) := \varnothing$;

(ii) $K_i(\pi_j\alpha) := \begin{cases} \{\alpha\} \cup K_i(\alpha) & \text{if } i \leq j, \\ \varnothing & \text{otherwise.} \end{cases}$

Lemma 3.3.5. *Let $\alpha \in \pi(\omega)$.*

$$K_i(\alpha) < \beta \Longleftrightarrow \alpha \in B_i(\beta)$$

Proof. By induction on α. □

Theorem 3.3.6. *Let α, $\beta \in \pi(\omega)$.*

(i) *The set $\pi(\omega)$ is primitive recursive and can be characterized as follows:*

- $0 \in \pi(\omega)$*;*
- *if $\alpha \in \pi(\omega)$ and $K_i(\alpha) < \alpha$, then $\pi_i\alpha \in \pi(\omega)$.*

(ii) $\pi(\omega) = B_0(\Omega_\omega)$.

(iii) *$\alpha < \beta$ if one of the following three cases holds:*

- *$\alpha = 0$ and $\beta \neq 0$;*
- *$\alpha = \pi_i\delta$, $\beta = \pi_j\gamma$, and $i < j$;*
- *$\alpha = \pi_i\delta$, $\beta = \pi_i\gamma$, and $\delta < \gamma$.*

Proof. (i) and (ii) are obvious. (iii) follows from Lemma 3.3.2. □

Remark 3.3.7. By Theorem 3.3.6.(iii) it can be decided whether $\alpha < \beta$, $\alpha = \beta$, or $\alpha > \beta$ for any α, $\beta \in \pi(\omega)$. In other words, $\alpha < \beta$ can be read as a Δ_0-formula in PA with two variables.

We define $\pi_0(\omega)$ as the set of ordinals from $\pi(\omega)$ which are less than Ω_1. Then

$$\pi_0(\omega) = \{\alpha \in \pi(\omega) \mid \alpha = 0 \text{ or } \alpha = \pi_0\beta \text{ for some } \beta \in \pi(\omega)\} = \pi_0\Omega_\omega.$$

We consider every element of $\pi(\omega)$ as a term defined according to the induction and call it an *ordinal term*. Below α, $\beta, \gamma, \ldots$ denote ordinal terms. If we use abbreviations

$$i_1 \cdots i_k 0 := \pi_{i_1} \cdots \pi_{i_k} 0,$$

then every $\alpha \in \pi_0(\omega)$ is of the form

$$\alpha = 0\alpha_1 0 \cdots 0\alpha_n 0^m 0$$

for some n, $m \in \mathbb{N}$, where $\alpha_i \in W_1$. Note that, if $n = 0$, then it is of the form $0^m 0$, so 0 if $m = 0$, too.

The following lemma reveals something about the relationship between the elements of $\pi_0(\omega)$ and $NF(W)$.

Lemma 3.3.8. *Let $\alpha = 0\alpha_1 0 \cdots 0\alpha_n 0^m 0$ be in $\pi_0(\omega)$ with $n > 1$. If $\alpha_i = \varnothing$ for some i, $1 \leq i < n$, then $\alpha_{i+1} = \varnothing$, too.*

Proof. Assume $\alpha_i = \varnothing$ and $\alpha_{i+1} \neq \varnothing$ for some $i < n$. Then α has the form $\pi_0 \cdots \pi_0 \pi_0 \pi_l \cdots 0$ for some $\ell > 0$. However, this cannot be in $\pi_0(\omega)$, since

$$K_0(\pi_0 \pi_\ell \cdots 0) = \{\pi_\ell \cdots, \ldots\} \not< \pi_0 \pi_\ell \cdots 0.$$

Hence $\alpha_{i+1} = \varnothing$. □

Remark 3.3.9. Hence we may assume for every $\alpha \in \pi_0(\omega)$ that α is of the form

$$\alpha = 0\alpha_1 0 \cdots 0\alpha_n 0^m 0,$$

where $\alpha_i \in W_1 \setminus \{\varnothing\}$ if $n \geq 1$.

By a *functional* we mean a finite sequence γ of natural numbers such that $\gamma 0 \in \pi(\omega)$. For $\alpha = i_1 \cdots i_n 0 \in \pi(\omega)$, $n \geq 1$, define a functional $\bar{\alpha}$ by

$$\bar{\alpha} := \langle i_1 + 1 \rangle \cdots \langle i_n + 1 \rangle.$$

Lemma 3.3.10. *Let α, $\beta > 0$ and $\gamma, \delta \in \pi_0(\omega)$.*

(i) *$\bar{\alpha}\gamma \in \pi(\omega) \setminus \pi_0(\omega)$.*

(ii) *$\bar{\alpha}\gamma < \bar{\beta}\delta$ iff $\alpha < \beta$, or $\alpha = \beta$ and $\gamma < \delta$.*

(iii) *$K_{i+1}(\bar{\alpha}\gamma) < \bar{\beta}\gamma$ iff $K_i(\alpha) < \beta$.*

Proof. By induction on α. □

Lemma 3.3.11. *For every $\gamma \in \pi(\omega) \setminus \pi_0(\omega)$ there are unique $\alpha > 0$ and $\delta \in \pi_0(\omega)$ such that $\gamma = \bar{\alpha}\delta$. In fact, if $\gamma = \gamma_1 0 \gamma_2$ with $\gamma_1 \in W_1$, then $\gamma = \bar{\alpha} 0 \gamma_2$, where $\alpha := \gamma_1^- 0$.*

Proof. The uniqueness of α and δ follows from Lemma 3.3.10.(ii). It remains to show that $\alpha := \gamma_1^- 0 \in \pi(\omega)$. By induction on the length of γ_1 we show the claim.

- $\gamma_1 = \langle i + 1 \rangle$. Then $\alpha = \pi_i 0$ is obviously in $\pi_0(\omega)$.

- $\gamma_1 = \langle i+1\rangle\eta$, where $\eta 0\gamma_2 \in \pi(\omega) \setminus \pi_0(\omega)$ and $K_{i+1}(\eta) < \eta$. Then by I.H. $\beta := \eta^- 0 \in \pi(\omega)$ with $\eta = \bar{\beta}0\gamma_2$, and $K_i(\beta) < \beta$ follows from $K_{i+1}(\eta) < \eta$. So $\alpha = i\beta \in \pi(\omega)$. □

Lemma 3.3.12. *Let $\alpha, \beta > 0$ and $\delta \in \pi_0(\omega)$. If $K_0(\bar{\alpha}\delta) < \bar{\beta}\delta$, then $K_0(\alpha) < \beta$.*

Proof. By induction on α.

- If $\alpha = \pi_i 0$, then it is obvious since $\beta > 0$.
- Let $\alpha = \pi_i\eta$ and $\eta > 0$. Then $\bar{\alpha}\delta = \pi_{i+1}\bar{\eta}\delta$. Since $K_0(\bar{\alpha}\delta) < \bar{\beta}\delta$ we have $\bar{\eta}\delta < \bar{\beta}\delta$ and $K_0(\bar{\eta}\delta) < \bar{\beta}\delta$. By I.H. $K_0(\eta) < \beta$, hence $K_0(\pi_i\eta) < \beta$. □

Lemma 3.3.13. *Let $\alpha, \beta > 0$ such that $K_0(\alpha) < \beta$ and $\delta \in \pi_0(\omega)$.*

$$K_0(\bar{\alpha}\delta) < \bar{\beta}\delta \quad \textit{iff} \quad \delta < \pi_0\bar{\beta}\delta$$

Proof. If $K_0(\bar{\alpha}\delta) < \bar{\beta}\delta$, then $K_0(\delta) < \bar{\beta}\delta$. Hence $\delta \in B_0(\bar{\beta}\delta)$. By Theorem 3.3.2.(iv) we have $\delta < \pi_0\bar{\beta}\delta$. Now assume $\delta < \pi_0\bar{\beta}\delta$, i.e., $K_0(\delta) < \bar{\beta}\delta$. There are two cases.

- If $\alpha = \pi_0 0$, then $\bar{\alpha}\delta = \pi_{i+1}\delta$. Hence $K_0(\bar{\alpha}\delta) < \bar{\beta}\delta$.
- Let $\alpha = \pi_i\eta$, $\eta > 0$, and $\bar{\alpha}\delta = \pi_{i+1}\bar{\eta}\delta$. Since $K_0(\alpha) < \beta$ then $\eta < \beta$ and $K_0(\eta) < \beta$. Hence $\bar{\eta}\delta < \bar{\beta}\delta$ and $K_0(\bar{\eta}\delta) < \bar{\beta}\delta$ by I.H.. The claim follows. □

Definition 3.3.14. Define $[\alpha]$ for $\alpha \in \pi_0(\omega)$ as follows:

- $[0] := 0$;
- if $\alpha = 0^{m+1}0$ for some m, then $[\alpha] := 0\bar{\alpha} = 01^{m+1}$;
- if $\alpha = 0\beta$ with $\beta \in \pi(\omega) \setminus \pi_0(\omega)$, then $[\alpha] := 0\bar{\beta}$.

At first glance this definition seems to be somewhat different from the original one. But it isn't because of Lemma 3.3.8.

Lemma 3.3.15. *If $\alpha, \delta \in \pi_0(\omega)$, then $[\alpha]\delta \in \pi_0(\omega)$ iff $\delta < [\alpha]\delta$.*

Proof. By induction on α.

- If $\alpha = 0$, then $[\alpha]\delta = \pi_0\delta$. Hence $[\alpha]\delta \in \pi_0(\omega)$ iff $K_0(\delta) < \delta$. This is exactly the case if $\delta < \pi_0\delta = [\alpha]\delta$ because of Lemma 3.3.8.
- Let $\alpha = \pi_0\beta$, $K_0(\beta) < \beta$, and $[\alpha]\delta = \pi_0\bar{\gamma}\delta$, where $\alpha = \gamma = 0^{m+1}0$ for some m by Lemma 3.3.8 if $\beta \in \pi_0(\omega)$, and $\gamma = \beta$ otherwise. In the first case, $\beta < \alpha$, hence $K_0(\beta) < \alpha$ and $K_0(\alpha) < \alpha$. Therefore, we have $K_0(\gamma) < \gamma$ in both cases. By Lemma 3.3.10.(i) $\bar{\gamma}\delta \in \pi(\omega)$. Moreover, by Lemma 3.3.13 $K_0(\bar{\gamma}\delta) < \bar{\gamma}\delta$ iff $\delta < \pi_0\bar{\gamma}\delta = [\alpha]\delta$. This was to show. □

Corollary 3.3.16. *$[\alpha]$ is a functional for every $\pi_0(\omega)$.*

The following characterization is obvious.

Lemma 3.3.17. *Let α, β, γ, δ, $[\alpha]\gamma$ and $[\beta]\delta$ be from $\pi_0(\omega)$. Then $[\alpha]\gamma < [\beta]\delta$ in exactly one of the following two cases:*

- *$\alpha < \beta$:*
- *$\alpha = \beta$ and $\gamma < \delta$.*

Note that if $\alpha = 0\alpha_1 0 \cdots 0\alpha_n 0^m 0 \in \pi_0(\omega)$ and $\alpha_i = 1\eta$, then $\eta = 1^k$ for some k. If not, we would have $K_0(\alpha_i 0 \cdots 0\alpha_n 0^m 0) \not< \alpha_i 0 \cdots 0\alpha_n 0^m 0$ which is not allowed. Hence the following lemma makes sense.

Lemma 3.3.18. *For every $\gamma \in \pi_0(\omega) \setminus \{0\}$ there are unique $\alpha, \eta \in \pi_0(\omega)$ such that $\gamma = [\alpha]\eta$. In fact, if $\gamma = 0\beta 0\delta$ and $\beta \in W_1$, then*

$$\gamma = \begin{cases} [0]0\delta & \textit{if } \beta = \varnothing, \\ [\beta']0\delta & \textit{otherwise}, \end{cases}$$

where $\beta' := \begin{cases} \beta^- 0 & \textit{if } \beta = 1^k \textit{ for some } k, \\ 0\beta^- 0 & \textit{if } \beta = j\eta \textit{ for some } \eta \textit{ and } j \geq 2. \end{cases}$

Proof. The uniqueness follows from Lemma 3.3.17. Let

$$\gamma' := \begin{cases} 0 & \text{if } \beta = \varnothing, \\ \beta' & \text{otherwise.} \end{cases}$$

We claim $\gamma' \in \pi_0(\omega)$ and $\gamma = [\gamma']0\delta$. In case of $\beta = \varnothing$ it is obvious. Let $\beta \neq \varnothing$. Since $K_0(\beta 0\delta) < \beta 0\delta$

$$K_0(0\delta) < \beta 0\delta \quad \text{and} \quad 0\delta < 0\beta 0\delta = \gamma$$

and by Lemma 3.3.11 and Lemma 3.3.12

$$\beta^- 0 \in \pi(\omega) \quad \text{and} \quad K_0(\beta^- 0) < \beta^- 0.$$

Hence $0\beta^- 0 \in \pi_0(\omega)$. If $\beta = 1\eta$ for some η, then $\beta^- 0 = 0^m$ for some m. So $[\beta'] = [\beta^- 0] = 0\beta^{-+} = 0\beta$. If $\beta = j\eta$ for some η and $j \geq 2$, then $[\beta'] = [0\beta' 0] = 0\beta^{-+} = 0\beta$. ◻

Lemma 3.3.19. *Let $n > 0$ and $\alpha = 0\alpha_1 0 \cdots 0\alpha_n 0^m 0$ from $\pi_0(\omega)$ with a non-empty α_n.*

(i) *$\alpha'_1 \geq \cdots \geq \alpha'_n$.*

(ii) *$\alpha = [\alpha'_1] \cdots [\alpha'_n] 0^m 0$.*

Proof. (ii) is true by Lemma 3.3.18. For (i) note that for every $i < n$

$$[\alpha'_{i+1}] \cdots [\alpha'_n]0^m 0 < [\alpha'_i] \cdots [\alpha'_n]0^m 0$$

by Lemma 3.3.15. The claim follows now by Lemma 3.3.17. □

Definition 3.3.20. Define $\breve{o}\colon \pi_0(\omega) \to \varepsilon_0$ by

$$\breve{o}(0\alpha_1 0 \cdots 0\alpha_n 0^m 0) := \omega^{\breve{o}(\alpha'_1)} + \cdots + \omega^{\breve{o}(\alpha'_n)} + m,$$

where α'_i is defined as above.

Theorem 3.3.21. *$\breve{o}\colon \pi_0(\omega) \to \varepsilon_0$ is an order-preserving isomorphism.*

Proof. Define $\breve{g} : \varepsilon_0 \to \pi_0(\omega)$ by

$$\breve{g}(\omega^{\alpha_1} + \cdots + \omega^{\alpha_n} + m) := 0\breve{g}(\alpha_1)''0 \cdots 0\breve{g}(\alpha_n)''0^m 0,$$

where $\alpha_1 \geq \cdots \geq \alpha_n > 0$ and

$$\beta'' := \begin{cases} \langle j+1\rangle\gamma^+ & \text{if } \beta = 0j\gamma 0 \text{ and } j \geq 1, \\ 1^k & \text{if } \beta = 0^k 0 \text{ for some } k. \end{cases}$$

Then we have obviously $\breve{g} \circ \breve{o} = \breve{o} \circ \breve{g} = id$. Note only that $(\alpha'')' = \alpha$ for every $\alpha \in \pi_0(\omega)$ and $(\beta')'' = \beta$ for every $\beta \in W_1$. That $\breve{o}$ and $\breve{g}$ are order-preserving follows from Lemma 3.3.19. □

Theorem 3.3.21 is interesting in the sense that it together with the definition of o gives simple and canonical order-preserving isomorphisms among $\pi_0(\omega)$, $NF(W)$, and $NF \subseteq S$. Indeed, $\iota_1\colon \pi_0(\omega) \to NF(W)$ and $\iota_2\colon \pi_0(\omega) \to NF$ are order-preserving isomorphisms:

$$\iota_1(0\alpha_1 0 \cdots 0\alpha_n 0^m 0) := \alpha_1 0 \cdots 0\alpha_n 0^m$$

and

$$\iota_2(0\alpha_1 0 \cdots 0\alpha_n 0^m 0) := 0^m \alpha_n^* 0 \cdots 0\alpha_1^*,$$

where β^* is the converse of β.

Remark 3.3.22. The question is why it is so. Why is the correspondence between Schütte-Simpson's system and Beklemishev's one so obvious, at least judging by its appearance? A naive answer is that the former and its order type represent really the proof-theoretic strength of PA as it is so in case of the graded provability algebra of PA.

3.4 Phase transitions

In this section we devote ourselves to some phase transitions demonstrating the intrinsic relations among the Cantor system, Beklemishev's system, and Schütte-Simpson's system. As in case of the Cantor system we need some norm functions. Let $lh(\alpha)$ be the length of the word α and $ht(\alpha)$ a maximal component of α^+.

Definition 3.4.1. $\check{N} : \pi_0(\omega) \to \mathbb{N}$ and $\hat{N} : NF(W) \to \mathbb{N}$ are defined as follows:

$$\begin{aligned}\check{N}(0\alpha_1 0 \cdots 0\alpha_n 0^m 0) &:= \hat{N}(\alpha_1 0 \cdots 0\alpha_n 0^m) \\ &:= m + n \dot{-} 1 + \sum_{i=1}^{n} lh(\alpha_i) + \sum_{i=1}^{n}\sum_{k=0}^{n_i} a_{ik}\,,\end{aligned}$$

where $\alpha_i = a_{i0} \cdots a_{in_i} \in W_1$.

Roughly speaking, $\check{N}\alpha$ and $\hat{N}\alpha$ are the addition of the length of α and all of its components. Given $X \in \{\varepsilon_0,\, NF,\, NF(W),\, \pi_0(\omega)\}$ and $f\colon \mathbb{N} \to \mathbb{N}$ define $\mathrm{SWO}(X, \sqsubseteq, f)$ by:

> for any k there exists a constant n which is so large that, for any finite sequence $\alpha_0, \ldots, \alpha_n$ from X with $\tilde{N}\alpha_i \le k + f(i)$ for all $i \le n$, there exist indices $\ell < m \le n$ satisfying $\alpha_\ell \sqsubseteq \alpha_m$.

Here $\tilde{N} \in \{N,\, \check{N},\, \hat{N}\}$ and $\sqsubseteq \in \{\le,\, \le_0,\, \preceq\}$ depending on X.

Lemma 3.4.2. *Let* $\alpha \in \{\pi_0(\omega),\, NF\}$.

(i) $N(\check{o}(\alpha)) \le \check{N}\alpha$ *and* $\check{N}(\alpha^{+p}) \le (ht(\alpha) + p) \cdot N(\check{o}(\alpha))$ *for every* $\alpha \in \pi_0(\omega)$.

(ii) $N(o(\alpha)) \le \hat{N}\alpha$ *and* $\hat{N}(\alpha^{+p}) \le (ht(\alpha) + p) \cdot N(o(\alpha))$ *for every* $\alpha \in NF$.

Proof. It suffices to show (ii) since (i) follows from it. We write just N for $\hat{N}$. There will be no confusions. Let $\alpha = 0^m\alpha_1 0 \cdots 0\alpha_n \in NF$. We show the claim by induction on the maximal component in α. Note that $o(\alpha) = \omega^{o(\alpha_n^-)} + \cdots + \omega^{o(\alpha_1^-)} + m$.

If $n = 0$ it is obvious. Now assume $n > 0$.

$$\begin{aligned}N(o(\alpha)) &= m + n + \sum_{i=1}^{n} N(o(\alpha_i^-)) \\ &\le m + n + \sum_{i=1}^{n} N\alpha_i^- \quad \text{(by I.H.)} \\ &\le m + n - 1 + \sum_{i=1}^{n} N\alpha_i = N\alpha\end{aligned}$$

Further, we have

$$\begin{aligned} N\alpha^{+p} &= N(\alpha_1^{+p}) + \cdots + N(\alpha_n^{+p}) + p(m+n-1) + (m+n-1) \\ &= N((\alpha_1^-)^{+(p+1)}) + \cdots + N((\alpha_n^-)^{+(p+1)}) + (p+1)(m+n-1) \\ &\leq (ht(\alpha_1^-) + p + 1) \cdot N(o(\alpha_1^-)) + \cdots + (ht(\alpha_n^-) + p + 1) \cdot N(o(\alpha_n^-)) \\ &\quad + (p+1)(m+n-1) \text{ by I.H.} \\ &\leq (ht(\alpha) + p)(N(o(\alpha_1^-)) + \cdots + N(o(\alpha_n^-)) + m + n - 1) \\ &= (ht(\alpha) + p) \cdot N(o(\alpha)) \end{aligned}$$

This completes the proof. □

Lemma 3.4.2 implies that the norm condition does not causes any essential difference in transformations between every two systems from ε_0, NF, $NF(W)$, $\pi_0(\omega)$. Hence the following theorem is a direct consequence of Theorem 2.1.3 and Theorem 2.1.4.

Let X be one of the systems ε_0, NF, $NF(W)$, $\pi_0(\omega)$.

Theorem 3.4.3. *Let n be a natural number and $\alpha \leq \varepsilon_0$.*

(i) $\mathrm{PRA} \vdash \mathrm{SWO}(X, \sqsubseteq, \lambda i \,.\, |i| \cdot \mathrm{inv}(i))$.

(ii) $\mathrm{PA} \nvdash \mathrm{SWO}(X, \sqsubseteq, \lambda i \,.\, |i| \cdot |i|_n)$.

(iii) $\mathrm{PA} \nvdash \mathrm{SWO}(X, \sqsubseteq, \lambda i \,.\, |i| \cdot |i|_{H_\alpha^{-1}(i)})$ *iff* $\alpha = \varepsilon_0$.

Remark: Ackermann's consistency proof for PA

T. Arai made an observation in [3] that Ackermann's consistency proof of PA in [1] based on the ϵ-substitution method uses a coding of ordinals up to ε_0 which is similar to Schütte-Simpson's ordinal notation system $\pi_0(\omega)$. Indeed, it is much more obvious to see there is a connection with Beklemishev's system $NF(W)$.

Part II

Ramsey Functions

Chapter 4

Finite Ramsey theorems

The Paris-Harrington principle is the first PA-independent sentence that mathematicians could encounter in their customary enterprise after Gödel introduced his classical work. It is a slight variation of the finite Ramsey theorem.

Another relevant principle is introduced by Kanamori and McAloon [29] who showed that the two principles are equivalent. In this part we investigate a generalization of the Kanamori-McAloon principle by some parameter functions and show phase transitions both in PA and in its fragments $\mathrm{I}\Sigma_n$.

We recall the original finite Ramsey theorem. Let the small Latin letters range over natural numbers and X over sets of natural numbers. Put

$$[X]^n := \{Y \subseteq X : \bar{\bar{Y}} = n\}.$$

If C is a function with domain $[X]^n$, we write $C(x_1, \ldots, x_n)$ for $C(\{x_1, \ldots, x_n\})$, where $x_1 < x_2 < \cdots < x_n$. Also, we identify each natural numbers ℓ with the set of its predecessors, i.e. $\ell = \{0, \ldots, \ell - 1\}$. Let

$$X \rightarrow (k)^n_c$$

denote the following:

for any $C\colon [X]^n \rightarrow c$, there is H s.t. C is constant on $[H]^n$ and $\bar{\bar{H}} \geq k$.

We call H *homogeneous* or *monochromatic* for f. One might try to notice the notation $X \rightarrow (k)^n_c$ as follows:

> given any coloring C with the n-dimensional domain $[X]^n$ and the range of c colors, there is a subset H of X such that C is $k$onstant on $[H]^n$ and $\bar{\bar{H}} \geq k$.

Ramsey [42] proved that for any n, c, k there is ℓ such that $\ell \to (k)^n_c$. Erdös and Rado [14] gave an upper bound for such ℓ depending super-exponentially on n, c and k.

Given $f\colon \mathbb{N} \to \mathbb{N}$ call $C\colon [X]^n \to \mathbb{N}$ *f-regressive* if $C(s) < f(\min(s))$ for any $s \in [X]^n$ such that $f(\min(s)) > 0$. If f is the identity function it is just called *regressive*. A set H is *min-homogeneous* for C if $C(s) = C(t)$ for any s, $t \in [H]^n$ such that $\min(s) = \min(t)$; that is, $f \restriction [H]^n$ depends only on the minimum element. Given n, k

$$X \to (k)^n_{f\text{-}reg}$$

denotes that, whenever $C\colon [X]^n \to \mathbb{N}$ is f-regressive, there is a subset H of X such that H is min-homogeneous for C and $\bar{\bar{H}} \geq k$. Then given $f\colon \mathbb{N} \to \mathbb{N}$ set

$$(\mathrm{KM})_f :\equiv \text{for any } n,\, k \text{ there is } \ell \text{ such that } \ell \to (k)^n_{f\text{-}reg}.$$

Kanamori and McAloon [29] proved that $(\mathrm{KM})_f$ is true for any $f\colon \mathbb{N} \to \mathbb{N}$. They even showed that $(\mathrm{KM}) :\equiv (\mathrm{KM})_{id}$ is PA-provably equivalent to the Paris-Harrington theorem.

The Paris-Harrington theorem is a variant of the finite Ramsey theorem. Given $H \subseteq \mathbb{N}$ and $f\colon \mathbb{N} \to \mathbb{N}$, we say that H is *f-large* if $\bar{\bar{H}} \geq f(\min(H))$. For $f = id$ we just say *large*. The notation

$$X \to^*_f (k)^n_c$$

denotes that for any $C\colon [X]^n \to c$ there is an f-large monochromatic set H for C such that $\bar{\bar{H}} \geq \max\{k, \min(H)\}$. Paris and Harrington [40] showed that the proposition

$$(\mathrm{PH}) :\equiv \text{for any } n,\, c,\, k \text{ there is } \ell \text{ such that } \ell \to^*_{id} (k)^n_c$$

is PA-independent. Weiermann [61] generalized this by considering f-largeness:

$$(\mathrm{PH})_f :\equiv \text{for any } n,\, c,\, k \text{ there is } \ell \text{ such that } \ell \to^*_f (k)^n_c$$

Here the largeness is replaced with f-largeness, i.e., $\bar{\bar{H}} \geq \max\{k,\, f(\min(H))\}$. He characterized via the fast growing hierarchy for which function class of f $(\mathrm{PH})_f$ remains unprovable in PA. We shall give an analogous classification of functions f such that $(\mathrm{KM})_f$ remains PA-independent and show that some refinements are also possible with respect to $\mathrm{I}\Sigma_n$. For this we work with

$$(\mathrm{KM})^n_f :\equiv \text{for any } k \text{ there is } \ell \text{ such that } \ell \to (k)^n_{f\text{-}reg}.$$

Weiermann worked with $(\mathrm{PH})^n_f$.

$$(\mathrm{PH})^n_f :\equiv \text{for any } c,\, k \text{ there is } \ell \text{ such that } \ell \to^*_f (k)^n_c.$$

4.1 Paris-Harrington vs. Kanamori-McAloon

We give some combinatorial connections between the Paris-Harrington principles and the Kanamori-McAloon principles. These results can be formalized in $I\Sigma_1$ or in *Primitive Recursive Arithmetic*, PRA.

Lemma 4.1.1. *Let* $f,\, g\colon \mathbb{N} \to \mathbb{N}$ *be increasing such that*

$$\forall i\, \exists m\, \forall \ell \geq m\, (\, f(\ell) \geq g(\ell) + i\,)\,.$$

Then $(\mathrm{PH})_f$ *implies* $(\mathrm{KM})_g$.

Proof. Assume

$$\forall n, x, c, k\, \exists y\, ([x, y] \rightarrow^*_f (k)^n_c\,)\,.$$

To show is

$$\forall n, x, k\, \exists y\, ([x, y] \rightarrow (k)^n_{g\text{-}reg}\,)\,.$$

Given n, x, k choose m and y such that

$$\forall \ell \geq m\, (\, f(\ell) \geq g(\ell) + n\,) \quad \text{and} \quad [x+m, y] \rightarrow^*_f (n+k)^{n+1}_3.$$

We claim

$$[x, y] \rightarrow (k)^n_{g\text{-}reg}.$$

Let $C\colon [x, y]^n \to \mathbb{N}$ be g-regressive. Define $D_0\colon [x+m, y]^{n+1} \to 3$ as follows:

$$D_0(x_0, \ldots, x_n) = \begin{cases} 0 & \text{if } C(x_0 - m, x_1 - m, \ldots, x_{n-1} - m) \\ & \qquad = C(x_0 - m, x_2 - m, \ldots, x_n - m), \\ 1 & \text{if } C(x_0 - m, x_1 - m, \ldots, x_{n-1} - m) \\ & \qquad < C(x_0 - m, x_2 - m, \ldots, x_n - m), \\ 2 & \text{if } C(x_0 - m, x_1 - m, \ldots, x_{n-1} - m) \\ & \qquad > C(x_0 - m, x_2 - m, \ldots, x_n - m). \end{cases}$$

By assumption there is $Y_0 \subseteq [x+m, y]$ homogeneous for D_0 such that

$$\bar{\bar{Y}}_0 \geq \max\{f(\min(Y_0)),\, n+k\}.$$

Put $Y := \{i - m\colon i \in Y_0\}$. Then Y is homogeneous for $D\colon [x, y]^{n+1} \to 3$ defined by:

$$D(x_0, \ldots, x_n) = \begin{cases} 0 & \text{if } C(x_0, x_1, \ldots, x_{n-1}) = C(x_0, x_2, \ldots, x_n), \\ 1 & \text{if } C(x_0, x_1, \ldots, x_{n-1}) < C(x_0, x_2, \ldots, x_n), \\ 2 & \text{if } C(x_0, x_1, \ldots, x_{n-1}) > C(x_0, x_2, \ldots, x_n). \end{cases}$$

Further, we have

$$\bar{\bar{\bar{Y}}} = \bar{\bar{\bar{Y}}}_0 \geq f(\min(Y_0)) = f(\min(Y) + m) \geq g(\min(Y)) + n.$$

Let $x_0 < x_1 < \cdots < x_{g(x_0)+n-1}$ be the first $g(x_0) + n$ elements of Y, where $x_0 = \min(Y)$. Then

$$D(x_0, x_1, \ldots, x_n) = D(x_0, x_2, \ldots, x_{n+1}) = \cdots = D(x_0, x_{g(x_0)}, \ldots, x_{g(x_0)+n-1}).$$

We claim $D \restriction [H]^{n+1} \equiv 0$. Assume otherwise, then

$$C(x_0, x_1, \ldots, x_{n-1}) < C(x_0, x_2, \ldots, x_n) < \cdots < C(x_0, x_{g(x_0)+1}, \ldots, x_{g(x_0)+n-1})$$

or

$$C(x_0, x_1, \ldots, x_{n-1}) > C(x_0, x_2, \ldots, x_n) > \cdots > C(x_0, x_{g(x_0)+1}, \ldots, x_{g(x_0)+n-1}).$$

Hence $C(x_0, x_i, \ldots, x_{i+n-1}) \geq g(x_0)$ for some i. This contradicts the fact that C is g-regressive.

Now let H be the set of the first k elements of Y and $z_1 < z_2 < \cdots < z_n$ the last n elements of Y. We claim H is min-homogeneous for C. Let $x_0 < x_1 < \cdots < x_{n-1}$ and $x_0 < y_1 < \cdots < y_{n-1}$ be from H. Then

$$\begin{aligned} C(x_0, x_1, \ldots, x_{n-1}) &= C(x_0, x_2, \ldots, x_{n-1}, z_1) \\ &= C(x_0, x_3, \ldots, x_{n-1}, z_1, z_2) \\ &\vdots \\ &= C(x_0, z_1, \ldots, z_{n-1}). \end{aligned}$$

The same holds for $C(x_0, y_1, \ldots, y_{n-1})$. We showed the claim. □

Lemma 4.1.2. *Let $d \in \mathbb{N}$.*

(i) $\forall i \, \exists m \, \forall \ell \geq m \, (|\ell|_d \geq |\ell|_{d+1} + i\,)$.

(ii) *Given $\alpha < \varepsilon_0$ there is $\beta < \varepsilon_0$ such that $\alpha < \beta$ and*

$$\forall i \, \exists m \, \forall \ell \geq m \, (|\ell|_{H_\beta^{-1}(\ell)} \geq |\ell|_{H_\alpha^{-1}(\ell)} + i\,)\,.$$

Proof. The first claim is obvious. We show the second one.

Given α set $\beta := \alpha + \omega$. Let $i \in \mathbb{N}$ be given. Putting $m := H_{\alpha+i+3}(i+1)$ we claim

$$\forall \ell \geq m \, (|\ell|_{H_\beta^{-1}(\ell)} \geq |\ell|_{H_\alpha^{-1}(\ell)} + i\,)\,.$$

Let $\ell \geq m$ be given. Then there is $p > i$ such that

$$H_{\alpha+i+3}(p) \leq \ell < H_{\alpha+i+3}(p+1).$$

Since $H_{\alpha+i+3}(p) = H_\alpha(p+i+3)$, we have $H_\alpha^{-1}(\ell) \geq p+i+3$. Hence

$$|\ell|_{H_\alpha^{-1}(\ell)} + i \leq |\ell|_{p+i+3} + i.$$

On the other hand

$$|\ell|_{H_\beta^{-1}(\ell)} = |\ell|_{H_{\alpha+\omega}^{-1}(\ell)} \geq |\ell|_{H_{\alpha+i+3}^{-1}(\ell)} \geq |\ell|_{p+1} \geq |\ell|_{p+i+3} + i$$

since $\ell \geq m \geq 2^{2^i}$. □

Corollary 4.1.3. *Let $d \in \mathbb{N}$ and $\alpha < \varepsilon_0$.*

(i) $(\mathrm{PH})_{|\cdot|_d}$ *implies* $(\mathrm{KM})_{|\cdot|_{d+1}}$.

(ii) $(\mathrm{PH})_{|\cdot|_{H_{\alpha+\omega}^{-1}(\cdot)}}$ *implies* $(\mathrm{KM})_{|\cdot|_{H_\alpha^{-1}(\cdot)}}$.

In the next chapter the following theorem will be proved .

Theorem 4.1.4. $(\mathrm{KM})^n_{|\cdot|_{n \dot{-} 3}}$ *implies* $(\mathrm{PH})^n$ *for any* $n > 1$.

The relation is much more compact with respect to (PH) and (KM).

Theorem 4.1.5 ([40, 29])**.** *The following are equivalent in* $\mathrm{I}\Sigma_1$*:*

(i) (PH)*,*

(ii) (KM)*,*

(iii) 1-Con(PA)*.*

Theorem 4.1.6 ([40, 29])**.** *If* $n > 0$*, the following are equivalent in* $\mathrm{I}\Sigma_1$*:*

(i) $(\mathrm{PH})^{n+1}$*,*

(ii) $(\mathrm{KM})^{n+1}$*,*

(iii) 1-$\mathsf{Con}(\mathrm{I}\Sigma_n)$*.*

Theorem 4.1.7 ([40, 29])**.** *Let* $n > 0$*.*

(i) $\mathrm{I}\Sigma_n \vdash (\mathrm{PH})^n \wedge (\mathrm{KM})^n$*.*

(ii) $\mathrm{I}\Sigma_1 \vdash (\mathrm{PH})^{n+1} \leftrightarrow (\mathrm{KM})^{n+1}$*.*

(iii) $\mathrm{I}\Sigma_n \nvdash (\mathrm{PH})^{n+1}$*.*

(iv) $\mathrm{I}\Sigma_n \nvdash (\mathrm{KM})^{n+1}$*.*

We are now going to introduce a more general concept. Let n, c, k, ℓ, s be natural numbers such that $s \leq n$, $1 \leq n \leq k$ and $1 \leq c$. U, V, W, etc denote finite sets of natural numbers.

Definition 4.1.8. Let $C\colon [\ell]^n \to c$ be a coloring. Call a set H *s-homogeneous* for C if for any s-element set $U \subseteq H$ and for any $(n-s)$-element sets $V, W \subseteq H$ such that $\max U < \min\{\min V, \min W\}$, we have

$$C(U \cup V) = C(U \cup W).$$

$(n-1)$-homogeneous sets are called *end-homogeneous*.

Note that 0-homogeneous sets are homogeneous and 1-homogeneous sets are min-homogeneous. Let

$$X \to_s \langle k \rangle_c^n$$

denote that given any coloring $C\colon [X]^n \to c$, there is H s-homogeneous for C such that $\overline{\overline{H}} \geq k$. The following lemma shows a connection between s-homogeneity and homogeneity.

Lemma 4.1.9. *Let $s \leq n$ and assume*

(i) $\ell \to_s \langle k \rangle_c^n$,

(ii) $k - n + s \to (m - n + s)_c^s$.

Then we have

$$\ell \to (m)_c^n.$$

Proof. Let $C\colon [\ell]^n \to c$ be given. Then (i) implies that there is $H \subseteq \ell$ such that $|H| = k$ and H is s-homogeneous for C. Let $z_1 < \cdots < z_{n-s}$ be the last $n-s$ elements of H. Set $H_0 := H \setminus \{z_1, \ldots, z_{n-s}\}$. Then $\overline{\overline{H_0}} = k - n + s$. Define $D\colon [H_0]^s \to c$ by

$$D(x_1, \ldots, x_s) := C(x_1, \ldots, x_s, z_1, \ldots, z_{n-s}).$$

By (ii) there is Y_0 such that $Y_0 \subseteq H_0$, $\overline{\overline{Y_0}} = m - n + s$, and homogeneous for D. Hence $D \restriction [Y_0]^s = e$ for some $e < c$. Set $Y := Y_0 \cup \{z_1, \ldots, z_{n-s}\}$. Then $\overline{\overline{Y}} = m$ and Y is homogeneous for C. Indeed, we have for any sequence $x_1 < \cdots < x_n$ from Y

$$C(x_1, \ldots, x_n) = C(x_1, \ldots, x_s, z_1, \ldots, z_{n-s}) = D(x_1, \ldots, x_s) = e.$$

The proof is complete. □

4.2 Ramsey functions and provability

Given n, s such that $s \leq n$ define $R^s_\mu(n, \cdot, \cdot)\colon \mathbb{N}^2 \to \mathbb{N}$ by

$$R^s_\mu(n, c, k) := \min\{\ell \colon \ell \to_s \langle k \rangle^n_c\}.$$

Then

- $R^0_\mu(1, c, k - n + 1) = c \cdot (k - n) + 1$,
- $R^n_\mu(n, c, k) = R^s_\mu(n, 1, k) = k$,
- $R^s_\mu(n, c, n) = n$,
- $R^s_\mu(n, c, k) \leq R^{s-1}_\mu(n, c, k)$ for any $s > 0$.

R^s_μ are called *Ramsey functions.* Set

$$R(n, c, k) := R^0_\mu(n, c, k) \quad \text{and} \quad R_\mu(n, c, k) := R^1_\mu(n, c, k).$$

Define a binary operation $*$ by putting, for positive natural numbers x and y,

$$x * y := x^y.$$

Further, we put for $p \geq 3$

$$x_1 * x_2 * \cdots * x_p := x_1 * (x_2 * (\cdots * (x_{p-1} * x_p) \cdots))$$

Erdös and Rado [14] gave an upper bound for $R(n, c, k)$: Given n, c, k such that $c \geq 2$ and $k \geq n \geq 2$, we have

$$R(n, c, k) \leq c * (c^{n-1}) * (c^{n-2}) * \cdots * (c^2) * (c \cdot (k - n) + 1).$$

This estimate turned out to be very useful in Weiermann [61]. However, we shall need somewhat more sharp estimate to deal with regressive functions and min-homogeneous sets.

Theorem 4.2.1 ($\mathrm{I}\Sigma_1$)**.** *Let $2 \leq n \leq k$, $0 < s \leq n$, and $2 \leq c$.*

$$R^s_\mu(n, c, k) \leq c * (c^{n-1}) * (c^{n-2}) * \cdots * (c^{s+1}) * (k - n + s) * s.$$

In particular, $R_\mu(2, c, k) \leq c^{k-1}$.

Proof. The proof construction below is motivated by Erdös and Rado [14]. We shall work with s-homogeneity instead of homogeneity.

Let X be a finite set. In the following construction we assume that $\bar{\bar{X}}$ is large enough. How large it should be will be determined after the construction has been defined. Throughout this proof the letter Y denotes subsets of X such that $\bar{\bar{Y}} = n-2$.

Let $C\colon [X]^n \to c$ be given and $x_1 < \ldots < x_{n-1}$ the first $n-1$ elements of X. Given $x \in X \setminus \{x_1, \ldots, x_{n-1}\}$ put

$$C_{n-1}(x) := C(x_1, \ldots, x_{n-1}, x).$$

Then $\mathrm{Im}(C_{n-1}) \subseteq c$, and there is $X_n \subseteq X \setminus \{x_1, \ldots, x_{n-1}\}$ such that C_{n-1} is constant on X_n and

$$\bar{\bar{X}}_n \geq c^{-1} \cdot (\bar{\bar{X}} - n + 1).$$

Let $x_n := \min X_n$ and given $x \in X_n - \{x_n\}$ put

$$C_n(x) := \prod \{C(Y \cup \{x_n, x\})\colon Y \subseteq \{x_1, \ldots, x_{n-1}\}\}.$$

Then $\mathrm{Im}(C_n) \subseteq c * \binom{n-1}{n-2}$, and there is $X_{n+1} \subseteq X_n - \{x_n\}$ such that C_n is constant on X_{n+1} and

$$\bar{\bar{X}}_{n+1} \geq c^{-\binom{n-1}{n-2}} \cdot (\bar{\bar{X}}_n - 1)\,.$$

Generally, let $p \geq n$, and suppose that $x_1, \ldots, x_{p-1}$ and $X_n, X_{n+1}, \ldots, X_p$ have been defined, and that $X_p \neq \varnothing$. Then let $x_p := \min X_p$ and for $x \in X_p - \{x_p\}$ put

$$C_p(x) := \prod \{C(Y \cup \{x_p, x\})\colon Y \subseteq \{x_1, \ldots, x_{p-1}\}\}.$$

Then $\mathrm{Im}(C_p) \subseteq c * \binom{p-1}{n-2}$, and there is $X_{p+1} \subseteq X_p - \{x_p\}$ such that C_p is constant on X_{p+1} and

$$\bar{\bar{X}}_{p+1} \geq c^{-\binom{p-1}{n-2}} \cdot (\bar{\bar{X}}_p - 1)\,.$$

Now put

$$\ell := 1 + R^s_\mu(n-1, c, k-1).$$

Then $\ell \geq k > n$. If $\bar{\bar{X}}$ is sufficiently large, then $X_p \neq \varnothing$, $n \leq p \leq \ell$, so that $x_1, \ldots, x_\ell$ exist. Note also that $x_1 < \cdots < x_\ell$. For $1 \leq \rho_1 < \cdots < \rho_{n-1} < \ell$ put

$$D(\rho_1, \ldots, \rho_{n-1}) := C(x_{\rho_1}, \ldots, x_{\rho_{n-1}}, x_\ell).$$

By definition of ℓ there is $Z \subseteq \{1, \ldots, \ell-1\}$ such that Z is s-homogeneous for D and $\bar{\bar{Z}} = k-1$. Finally, we put

$$X' := \{x_\rho\colon \rho \in Z\} \cup \{x_\ell\}.$$

We claim that X' is min-homogeneous for C. Let

$$H := \{x_{\rho_1}, \dots, x_{\rho_n}\} \quad \text{and} \quad H' = \{x_{\eta_1}, \dots, x_{\eta_n}\}$$

be two subsets of X' such that $\rho_1 = \eta_1, \dots, \rho_s = \eta_s$ and

$$1 \leq \rho_1 < \cdots < \rho_n \leq \ell, \quad 1 \leq \eta_1 < \cdots < \eta_n \leq \ell.$$

Since x_{ρ_n}, $x_\ell \in X_{\rho_n}$, we have $C_{\rho_{n-1}}(x_{\rho_n}) = C_{\rho_{n-1}}(x_\ell)$ and hence

$$C(x_{\rho_1}, \dots, x_{\rho_{n-1}}, x_{\rho_n}) = C(x_{\rho_1}, \dots, x_{\rho_{n-1}}, x_\ell).$$

Similarly, we show that

$$C(x_{\eta_1}, \dots, x_{\eta_{n-1}}, x_{\eta_n}) = C(x_{\eta_1}, \dots, x_{\eta_{n-1}}, x_\ell).$$

In addition, since $\{x_{\rho_1}, \dots, x_{\rho_{n-1}}\} \cup \{x_{\eta_1}, \dots, x_{\eta_{n-1}}\} \subseteq X'$, we have

$$D(\rho_1, \dots, \rho_{n-1}) = D(\eta_1, \dots, \eta_{n-1}),$$

i.e.,

$$C(x_{\rho_1}, \dots, x_{\rho_{n-1}}, x_\ell) = C(x_{\eta_1}, \dots, x_{\eta_{n-1}}, x_\ell).$$

This means that $C(H) = C(H')$. So X' is s-homogeneous for C.

We now return to the question how large $\bar{\bar{X}}$ should be in order to ensure that the construction above can be carried through.

Set

$$\begin{aligned} t_n &:= c^{-1} \cdot (\bar{\bar{X}} - n + 1), \\ t_{p+1} &:= c^{-\binom{p-1}{n-2}} \cdot (t_p - 1) \qquad (n \leq p < \ell). \end{aligned}$$

Then we require that $t_\ell > 0$, where

$$\begin{aligned} t_\ell &= c^{-\binom{\ell-2}{n-2}} \cdot \Big(c^{-\binom{\ell-3}{n-2}} \cdot \Big(\cdots \Big(c^{-\binom{n-1}{n-2}} \cdot \big(t_n - 1\big)\Big) \cdots\Big) - 1\Big) \\ &= c^{-\binom{\ell-2}{n-2} - \cdots - \binom{n-1}{n-2}} \cdot t_n - c^{-\binom{\ell-2}{n-2} - \cdots - \binom{n-1}{n-2}} - \cdots - c^{-\binom{\ell-2}{n-2} - \binom{\ell-3}{n-2}} - c^{-\binom{\ell-2}{n-2}}. \end{aligned}$$

Since $c = c^{\binom{n-2}{n-2}}$ a sufficient condition on $\bar{\bar{X}}$ is then

$$\bar{\bar{X}} - n + 1 > c^{\binom{\ell-3}{n-2} + \cdots + \binom{n-2}{n-2}} + c^{\binom{\ell-4}{n-2} + \cdots + \binom{n-2}{n-2}} + \cdots + c^{\binom{n-2}{n-2}}.$$

A possible value is

$$\bar{\bar{X}} = n + \sum_{p=n-1}^{\ell-2} c^{\binom{p}{n-1}},$$

so that

$$\begin{aligned}
R^s_\mu(n,c,k) &\leq n + \sum_{p=n-1}^{\ell-2} c^{\binom{p}{n-1}} \quad \leq \quad n + \sum_{p=n-1}^{\ell-2} c^{p^{n-1}} \\
&\leq n + \sum_{p=n-1}^{\ell-2} \left(c^{(p+1)^{n-1}} - c^{p^{n-1}}\right) \\
&= n + c^{(\ell-1)^{n-1}} - c^{(n-1)^{n-1}} \\
&\leq c^{(\ell-1)^{n-1}} \\
&= c^{R_\mu(n-1,c,k-1)^{n-1}}.
\end{aligned}$$

Hence

$$R^s_\mu(n,c,k) * n \quad \leq \quad (c^n) * R^s_\mu(n-1,c,k-1) * (n-1).$$

After $(n-s)$ times iterated applications of the inequality we get

$$\begin{aligned}
R^s_\mu(n,c,k) * n &\leq (c^n) * (c^{n-1}) * \cdots * (c^{s+1}) * R^s_\mu(s,c,k-n+s) * s \\
&= (c^n) * (c^{n-1}) * \cdots * (c^{s+1}) * (k-n+s) * s.
\end{aligned}$$

This completes the proof. □

Remark 4.2.2. Lemma 26.4 in [13] gives a slight sharper estimate for $s = n-1$:

$$R^{n-1}_\mu(n,c,k) \leq n + \sum_{i=n-1}^{k-2} c^{\binom{i}{n-1}}$$

Corollary 4.2.3. *Let $2 \leq n \leq k$ and $2 \leq c$.*

$$R_\mu(n,c,k) \leq c * (c^{n-1}) * (c^{n-2}) * \cdots * (c^2) * (k-n+1).$$

Now we come back to f-regressiveness.

Definition 4.2.4. Given $f\colon \mathbb{N} \to \mathbb{N}$ set

$$R_f(n,x,k) := \min\{\ell\colon [x,\ell] \to (k)^n_{f\text{-}reg}\}.$$

Lemma 4.2.5. *Given $n \geq 1$ and $\alpha \leq \varepsilon_0$ set $f^n_\alpha(i) := \sqrt[H_\alpha^{-1}(i)]{|\, i \,|_{n-1}}$. Then*

$$R_{f^n_\alpha}(n+1,x,k) \leq 2_{n-1}(H_\alpha(x+q)^{k+p})$$

for some p, $q \in \mathbb{N}$ depending (primitive-recursively) on n, x, k.

Proof. Given n, x, k note first that there are two natural numbers p and q such that $n < p < q$ and

$$\ell := 2_{n-1}(H_\alpha(x+q)^{k+n}+1) + H_\alpha(x+q) \leq 2_{n-1}(H_\alpha(x+q)^{k+p}) =: m.$$

Let $C\colon [x,m]^{n+1} \to \mathbb{N}$ be any f^n_α-regressive function and

$$D\colon [H_\alpha(x+q), \ell]^{n+1} \to \mathbb{N}$$

be defined from C by restriction. Then for any $y \in [H_\alpha(x+q), \ell]$, we have

$$\begin{aligned} \sqrt[H_\alpha^{-1}(y)]{|y|_{n-1}} &\leq \sqrt[H_\alpha^{-1}(H_\alpha(x+q))]{|2_{n-1}(H_\alpha(x+q)^{k+p})|_{n-1}} \\ &= \sqrt[x+q]{H_\alpha(x+q)^{k+p}+1}. \end{aligned}$$

Hence

$$\mathrm{Im}(D) \subseteq \lfloor H_\alpha(x+q)^{(k+p)/(x+q)} \rfloor + 1.$$

Put now $c := \lfloor H_\alpha(x+q)^{(k+p)/(x+q)} \rfloor + 1$. Then

$$(c) * (c^n) * \cdots * (c^2) * (k-n) < 2_{n-1}(H_\alpha(x+q)^{k+n}+1)$$

if q is sufficiently larger than p. By Theorem 4.2.1 there is H min-homogeneous for D, hence for C, such that $\bar{\bar{H}} \geq k$. □

Theorem 4.2.6. *Let* $n \geq 1$.

(i) $(\mathrm{KM})_{\log^*}$ *is provable in* $\mathrm{I}\Sigma_1$.

(ii) $(\mathrm{KM})^{n+1}_{|\cdot|_n}$ *is provable in* $\mathrm{I}\Sigma_1$.

(iii) $(\mathrm{KM})^{n+1}_{\sqrt[H_\alpha^{-1}(\cdot)]{|\cdot|_{n-1}}}$ *is provable in* $\mathrm{I}\Sigma_n$ *if* $\alpha < \omega_{n+1}$.

Proof. Given n, $k \geq 1$ we claim that

$$\max\{R_{\log^*}(n,x,k),\ R_{|\cdot|_n}(n+1,x,k)\} \leq x + 2_n(x+k) =: \ell$$

if x is sufficiently large.

(i) Given n, x, k let $C\colon [x,\ell]^n \to \mathbb{N}$ be $\log^*$-regressive. Put $c := x+k$. Then

$$\log^* \ell \leq c$$

and

$$c * (c^{n-1}) * (c^{n-2}) * \cdots * (c^2) * (k \dot{-} n + 1) < 2_n(x+k)$$

if x is sufficiently larger than n and k. By Theorem 4.2.1 we can find H min-homogeneous for C such that $\bar{\bar{H}} \geq k$.

(ii) Given n, x, k let $C\colon [x,\ell]^{n+1} \to \mathbb{N}$ be $|\cdot|_n$-regressive. Put $c := 2x+k$. Then

$$|\ell|_n \le c$$

and

$$c * (c^n) * (c^{n-1}) * \cdots * (c^2) * (k \dot{-} n) < 2_n(x+k)$$

if x is sufficiently larger than n and k. By Theorem 4.2.1 we can find $H \subseteq [x,\ell]$ min-homogeneous for C such that $\bar{\bar{H}} \ge k$.

(iii) H_α is provably recursive in $\mathrm{I}\Sigma_n$ for $\alpha < \omega_{n+1}$. See e.g. [55, 56] for a proof. Then the assertion follows from 4.2.5 □

Theorem 4.2.7. *Given* $\alpha \le \varepsilon_0$ *set* $f_\alpha := |\cdot|_{H_\alpha^{-1}(\cdot)}$. *Then*

$$R_{f_\alpha}(n,x,k) \le 2_{n+1}(H_\alpha(x+k))$$

if x *is large enough.*

Proof. Given n, x, k set $\ell := 2_{n+1}(H_\alpha(x+k))$. Then

$$m := 2_n(H_\alpha(x+k)) + H_\alpha(x+k) \le \ell.$$

Let $C\colon [x,N]^n \to \mathbb{N}$ be any f_α-regressive function. Define from C

$$D\colon [H_\alpha(x+k), m]^n \to \mathbb{N}$$

by restriction. Note that, for any $y \in [H_\alpha(x+k), m]$, we have

$$\begin{aligned} |y|_{H_\alpha^{-1}(y)} &\le |2_{n+1}(H_\alpha(x+k))|_{H_\alpha^{-1}(H_\alpha(x+k))} \\ &< H_\alpha(x+k), \end{aligned}$$

if $x+k > n+1$. Hence,

$$\mathrm{Im}(D) \subseteq H_\alpha(x+z).$$

In addition, we have for $c := H_\alpha(x+k)$

$$(c) * (c^{n-1}) * \cdots * (c^2) * (k-n+1) < 2_n(H_\alpha(x+k)),$$

if x is large enough. By Theorem 4.2.1 there is Y min-homogeneous for D, hence for C, such that $\bar{\bar{Y}} \ge k$. □

Corollary 4.2.8. *Let* $n > 0$.

(i) $(\mathrm{KM})^{n+1}_{|\cdot|_{H_\alpha^{-1}(\cdot)}}$ *is provable in* $\mathrm{I}\Sigma_n$ *for any* $\alpha < \omega_{n+1}$.

(ii) $(\mathrm{KM})_{|\cdot|_{H_\alpha^{-1}(\cdot)}}$ *is provable in* PA *for any* $\alpha < \varepsilon_0$.

Proof. H_α is provably recursive in $\mathrm{I}\Sigma_n$ for any $\alpha < \omega_{n+1}$. See e.g. [55, 56]. The assertion follows now from Theorem 4.2.7. □

Chapter 5

Fast growing functions

In this chapter we classify function classes of f such that R_f build fast growing hierarchies. We owe a great deal of this chapter to [14], [29] and [61]. We want to express special thanks to J. Paris and A. Kanamori who sent a personal handwritten note of J. Paris. It contains a purely combinatorial proof that $(\mathrm{KM})^n$ implies $(\mathrm{PH})^n$ for $n \geq 2$ and is one of the starting points of this thesis.

We begin with a foretaste, that is, we study the growth rate of $R_f(2, \cdot, \cdot)$ using the functions F_α, $\alpha \leq \omega$.

5.1 Ackermannian Ramsey functions

Kojman and Shelah [31] gave a very short and elementary, but a little technical proof that $R_{id}(2, \cdot, \cdot)$ is Ackermannian. The first proof of this fact had already been given in Kanamori and McAloon [29] but with model-theoretic methods. Here we are going to give a complete classification of the parameter functions such that the resulting Ramsey functions are Ackermannian.

Given $c \geq 1$ set

$$R_\mu(c, k) := R_\mu(2, c, k).$$

Then $R_\mu(1, k) = k$ and $R_\mu(c, 2) = 2$.

Lemma 5.1.1. *$R_\mu(c, k) \leq 2 \cdot c^{k-2}$ for c, $k \geq 2$.*

Proof. See Lemma 26.4 in [13]. □

Note that Lemma 26.4 in [13] talks about *end-homogeneous* sets. However, if we confine ourselves to the 2-dimensional case it is just about min-homogeneous sets.

Throughout this section m denotes a fixed positive natural number. Set

$$g_A(i) := \lfloor \sqrt[A^{-1}(i)]{i} \rfloor\,, \quad g_m(i) := \lfloor \sqrt[F_m^{-1}(i)]{i} \rfloor \quad \text{and} \quad h_m(i) := \lfloor \sqrt[m]{i} \rfloor\,.$$

where $A := F_\omega$. Given $f \in \{g_A, g_m, h_m\}$ set

$$R_f(k) := R_f(2, 0, k) = \min\{\ell \colon \ell \to (k)^2_{f\text{-}reg}\} - 1.$$

Theorem 5.1.2. *R_{g_m} is primitive recursive.*

Proof. Given $k \geq 2$ set $p := F_m(k^2)^{k+1}$. We claim that

$$R_{g_m}(k) \leq p$$

Set $\ell := F_m(k^2)^k + F_m(k^2) < p$ and let $C \colon [p]^2 \to \mathbb{N}$ be a g_m-regressive function. Consider the function $D \colon [F_m(k^2), \ell]^2 \to \mathbb{N}$ defined from C by restriction. For any $y \in [F_m(k^2), \ell]$ we have

$$\sqrt[F_m^{-1}(y)]{y} \leq \sqrt[F_m^{-1}(F_m(k^2))]{F_m(k^2)^{k+1}} = F_m(k^2)^{(k+1)\cdot k^{-2}}\,,$$

hence $\mathrm{Im}(D) \subseteq F_m(k^2)^{(k+1)\cdot k^{-2}} + 1$. On the other hand,

$$2 \cdot (F_m(k^2)^{(k+1)\cdot k^{-2}} + 1)^{k-2} < (F_m(k^2)^{(k+1)\cdot k^{-2}+1})^{k-1} < F_m(k^2)^k + 1.$$

By Lemma 5.1.1 we find a set H min-homogeneous for D, hence for C, such that $|H| \geq k$. This implies that R_{g_m} is primitive recursive. Note that class of primitive recursive functions is closed under the bounded μ-operator, see [41]. □

We show further that some refinements of the ideas elaborated by Kanamori and McAloon in [29] give us the counterpart of Theorem 5.1.2. Define a sequence of strictly increasing functions $f_{m,n}$ for as follows:

$$f_{m,n}(i) := \begin{cases} i+1 & \text{if } n = 0, \\ f_{m,n-1}^{(\lfloor \sqrt[m]{i} \rfloor)}(i) & \text{otherwise.} \end{cases}$$

Note that $f_{m,n}$ are strictly increasing.

Lemma 5.1.3. *$R_{h_m}(R_\mu(c, i+3)) \geq f_{m,c}(i)$ for all c and i.*

Proof. Let $k := R_\mu(c, i+3)$ and define a function $C_m \colon [R_{h_m}(k)]^2 \to \mathbb{N}$ as follows:

$$C_m(x, y) := \begin{cases} 0 & \text{if } f_{m,c}(x) \leq y, \\ \ell & \text{otherwise,} \end{cases}$$

where the number ℓ is defined by

$$f^{(\ell)}_{m,p}(x) \leq y < f^{(\ell+1)}_{m,p}(x)$$

where $p < c$ is the maximum such that $f_{m,p}(x) \le y$. Note that C_m is h_m-regressive since $f_{m,p}^{(\lfloor \sqrt[m]{x} \rfloor)}(x) = f_{m,p+1}(x)$. Let H be a k-element subset of $R_{h_m}(k)$ which is min-homogeneous for C_m. Define a c-coloring $D_m \colon [H]^2 \to c$ by

$$D_m(x,y) := \begin{cases} 0 & \text{if } f_{m,c}(x) \le y, \\ p & \text{otherwise,} \end{cases}$$

where p is as above. Then there is a $(i+3)$-element set $X \subseteq H$ homogeneous for D_m. Let $x < y < z$ be the last three elements of X. Then $i \le x$. Hence, it suffices to show that $f_{m,c}(x) \le y$ since $f_{m,c}$ is an increasing function.

Assume $f_{m,c}(x) > y$. Then $f_{m,c}(y) \ge f_{m,c}(x) > z$ by the min-homogeneity. Let $C_m(x,y) = C_m(x,z) = \ell$ and $D_m(x,y) = D_m(x,z) = D_m(y,z) = p$. Then

$$f_{m,p}^{(\ell)}(x) \le y < z < f_{m,p}^{(\ell+1)}(x).$$

By applying $f_{m,p}$ we get the contradiction that $z < f_{m,p}^{(\ell+1)}(x) \le f_{m,p}(y) \le z$. □

It remains to show that R_{h_m} is not primitive recursive. We are going to show this by comparing the functions $f_{m,n}$ with the Ackermann function.

Lemma 5.1.4. *Let $i \ge 4^m$ and $\ell \ge 0$.*

(i) $(2i+2)^m < f_{m,\ell+2m^2}(i)$ *and* $f_{m,\ell+2m^2}((2i+2)^m) < f_{m,\ell+2m^2}^{(2)}(i)$.

(ii) $F_n(i) < f_{m,n+2m^2}^{(2)}(i)$.

Proof. (1) By induction on k it is easy to show that $f_{m,k}(i) > (\lfloor \sqrt[m]{i} \rfloor)^k$ for any $i > 0$. Hence for $i \ge 4^m$

$$f_{m,2m^2}(i) > (\lfloor \sqrt[m]{i} \rfloor)^{2m^2} \ge (\lfloor \sqrt[m]{i} \rfloor)^{m^2} \cdot 2^{m^2+m} \ge (\sqrt[m]{i+1})^{m^2} \cdot 2^m = (2i+2)^m$$

since $2 \cdot \lfloor \sqrt[m]{i} \rfloor \ge \sqrt[m]{i+1}$. The second claim follows from the first one.

(2) By induction on n we show the claim. If $n = 0$ it is obvious. Suppose the claim is true for n. Let $i \ge 4^m$ be given. Then by induction hypothesis we have $F_n(i) \le f_{m,n+2m^2}^{(2)}(i)$. Hence

$$F_{n+1}(i) = F_n^{(i+1)}(i) \le f_{m,n+2m^2}^{(2i+2)}(i) \le f_{m,n+2m^2+1}((2i+2)^m) < f_{m,n+2m^2+1}^{(2)}(i).$$

The induction is now complete. □

Corollary 5.1.5. *$F_n(i) \le f_{m,n+2m^2+1}(i)$ for any $i \ge 4^m$.*

Theorem 5.1.6. *R_{h_m} and R_{g_A} are not primitive recursive.*

Proof. Lemma 5.1.3 and Corollary 5.1.5 imply that R_{h_m} is not primitive recursive. For the second assertion we claim that

$$p := R_{g_A}(R_\mu(i + 2i^2 + 1, 4^i + 3)) > A(i)$$

for all i. Assume to the contrary that $p \leq A(i)$ for some i. Then for any $\ell \leq p$ we have $A^{-1}(\ell) \leq i$, hence $\sqrt[i]{\ell} \leq \sqrt[A^{-1}(\ell)]{\ell}$. Hence

$$\begin{aligned} R_{g_A}(R_\mu(i + 2i^2 + 1, 4^i + 3)) &\geq R_{h_i}(R_\mu(i + 2i^2 + 1, 4^i + 3)) \\ &\geq f_{i,i+2i^2+1}(4^i) \\ &> A(i) \end{aligned}$$

by Lemma 5.1.3 and Corollary 5.1.5. Contradiction! □

5.2 Fast growing Ramsey functions

The emphasis of Section 4.2 lies in combinatorial results which give upper bounds for Ramsey functions in primitive recursive way. We shall discuss now how one can make use of the results from Kanamori and McAloon [29] to get some refinement results of the parametrized Kanamori-McAloon principle. For this we base ourselves on the independence results proved in Paris and Harrington [40] and Weiermann [61]. Let's start with some technical lemmata from [29]. The proofs will be repeated because they help to understand the following ideas.

Lemma 5.2.1 (Kanamori and McAloon [29])**.** *Let $I \subseteq \mathbb{N}$. If $C\colon [I]^n \to \mathbb{N}$ is regressive, then $H \subseteq I$ is min-homogeneous for C iff every $Y \subseteq H$ of cardinality $n+1$ is min-homogeneous for C.*

Proof. If $H \subseteq I$ is not min-homogeneous, let $x_0 < \cdots < x_{n-1}$ be the lexicographically least sequence drawn from H such that there are $x_0 < y_1 < \cdots < y_{n-1}$ all from H with $C(x_0, x_1, \ldots, x_{n-1}) \neq C(x_0, y_1, \ldots, y_{n-1})$, where we can take $\langle y_1, \ldots, y_{n-1}\rangle$ to be the lexicographically least with this property. If i is the least such that $x_i \neq y_i$, then $x_i < y_i$. Hence we have

$$C(x_0, \ldots, x_{n-1}) = C(x_0, \ldots, x_i, y_i, \ldots, y_{n-2})$$

and thus $Y = \{x_0, \ldots, x_i, y_i, \ldots, y_{n-1}\}$ is not min-homogeneous for C. □

The following lemma is a slight modification of Lemma 3.3 in [29].

Lemma 5.2.2. *If $C\colon [x, y]^n \to y$ is f-regressive and $f(i) \leq i$, then there is an C' such that*

- *$C'\colon [x, y]^{n+1} \to y$ is $2|f| + 1$-regressive, and*
- *if H' is min-homogeneous for C', then $H = H' - (7 \cup \{\max(H')\})$ is min-homogeneous for C.*

Proof. Given $x \in \mathbb{N}$ note that any $y < x$ can be represented as $(y_0, \ldots, y_{d-1})_2$ in binary notation, where $d = |x|$. Write $C(s) = (C_0(s), \ldots, C_{d-1}(s))_2$, where $s \in [x, y]^n$ and $d = |f(\min(s))|$. Define C' on $[x, y]^{n+1}$ by:

- $C'(x_0, \ldots, x_n) = 0$ if either $x_0 < 7$, or $\{x_0, \ldots, x_n\}$ is min-homogeneous for C;
- $C'(x_0, \ldots, x_n) = 2i + C_i(x_0, \ldots, x_{n-1}) + 1$, otherwise, where $i < |f(x_0)|$ is the least such that $\{x_0, \ldots, x_n\}$ is not min-homogeneous for C_i.

Then C' is $2|f|+1$-regressive. (Notice that $2\log(x+1)+1 \leq x$ for every $x \geq 7$.) Suppose that H' is min-homogeneous for C' and H is as described. If $C' \restriction [H]^{n+1} = \{0\}$, then we are done by the previous lemma. Suppose on the contrary that there were $x_0 < \cdots < x_n$ all in H such that $C'(x_0, \ldots, x_n) = 2i + C_i(x_0, \ldots, x_{n-1}) + 1$. Given any s, $t \in [\{x_0, \ldots, x_n\}]^n$ with $\min(s) = \min(t) = x_0$, note that

$$C'(s \cup \{\max(H')\}) = C'(x_0, \ldots, x_n) = C'(t \cup \{\max(H')\})$$

by min-homogeneity. But then, $C_i(s) = C_i(t)$, so that $\{x_0, \ldots, x_n\}$ were min-homogeneous for C_i after all, which contradicts the assumption. □

Define a sequence of functions as follows:

$$b_0(i) := i \quad \text{and} \quad b_{m+1}(i) := 2 \cdot |b_m(i)| + 1$$

Lemma 5.2.3. *Let n, $m \geq 1$.*

$$\mathrm{I}\Sigma_1 \vdash (\mathrm{KM})^{n+m}_{b_m} \rightarrow (\mathrm{KM})^n$$

Proof. Assume $(\mathrm{KM})^{n+m}_{b_m}$. Let x and k be given. Note that given any ℓ there is t_ℓ such that $7 \leq b_\ell(i)$ for any $i \geq t_\ell$. We may assume $x \geq t_m$. By assumption we can find y such that

$$[x, y] \rightarrow (k+m)^{n+m}_{b_m\text{-}reg}.$$

We claim $[x, y] \rightarrow (k)^n$. To see this let $C\colon [x, y]^n \rightarrow y$ be regressive. Applying the previous lemma m times we get $C'\colon [x, y]^{n+m} \rightarrow y$, b_m-regressive. Therefore, there is H' min-homogeneous for C' such that $\bar{\bar{H}} \geq k+m$, so

$$H = H' - \{\text{the last } m \text{ elements of } H'\}$$

is min-homogeneous for C and $\bar{\bar{H}} \geq k$. □

Corollary 5.2.4. *Let $m > 0$.*

(i) $\mathrm{I}\Sigma_1 \vdash (\mathrm{KM})_{|\cdot|_m} \longrightarrow (\mathrm{KM})$.

(ii) $\mathrm{PA} \nvdash (\mathrm{KM})_{|\cdot|_m}$, *i.e.* $R_{|\cdot|_m}$ *is not provably recursive in* PA.

Remark 5.2.5. Later in Theorem 5.2.13 we will see that $R_{|\cdot|_m}$ build a fast growing hierarchy.

Lemma 5.2.6 ([29])**.** *There are three regressive functions* $\eta_1, \eta_2, \eta_3 : [\mathbb{N}]^2 \to \mathbb{N}$ *such that whenever* H' *is min-homogeneous for all of them, then*

$$H = H' \setminus \{\textit{the last three elements of } H'\}$$

has the property that $x < y$ *both in* H *implies* $x^x \leq y$*.*

Proof. Define $\eta_1, \eta_2, \eta_3 : [\mathbb{N}]^2 \to \mathbb{N}$ by:

$$\eta_1(x,y) \;=\; \begin{cases} 0 & \text{if } 2x \leq y, \\ y - x & \text{otherwise,} \end{cases}$$

$$\eta_2(x,y) \;=\; \begin{cases} 0 & \text{if } x^2 \leq y, \\ u & \text{otherwise, where } u \cdot x \leq y < (u+1) \cdot x, \end{cases}$$

$$\eta_3(x,y) \;=\; \begin{cases} 0 & \text{if } x^x \leq y, \\ v & \text{otherwise, where } x^v \leq y < x^{v+1}. \end{cases}$$

Suppose that H' is as hypothesized, and let $z_1 < z_2 < z_3$ be the last three elements of H'. If $x < y$ are both in $H' \setminus \{z_3\}$, then since $\eta_1(x,y) = \eta_1(x, z_3)$, clearly we must have $\eta_1(x,y) = 0$. Hence, η_1 on $[H' \setminus \{z_3\}]^2$ is constantly 0.

Next, assume that $x < y$ are both in $H' \setminus \{z_2, z_3\}$ and $\eta_2(x,y) = u > 0$. then

$$u \cdot x \leq y < z_2 < u \cdot x + x$$

by min-homogeneity, and so we have

$$u \cdot x + x \leq y + x \leq y + y \leq z_2$$

by the previous paragraph. But this leads to the contradiction $z_2 < z_2$. Hence, η_2 on $[H' \setminus \{z_2, z_3\}]^2$ is constantly 0.

Finally, we can iterate the argument to show that η_3 on $[H' \setminus \{z_1, z_2, z_3\}]^2$ is constantly 0, and so the proof is complete. □

The following theorem stems from the brilliant, purely combinatorial idea sketch by J. Paris. He gave a combinatorial proof of the propositions

$$(\mathrm{KM})^n \to (\mathrm{KM})^n_{2^{2^{\cdot}}} \quad \text{and} \quad (\mathrm{KM})^n_{2^{2^{\cdot}}} \to (\mathrm{PH})^n$$

for $n \geq 2$. During a trial to give a direct proof of the proposition

$$(\mathrm{KM})^n \to (\mathrm{PH})^n,$$

it turned out that the idea could be optimalized. The basic idea is to get very large min-homogeneous sets such that some fine thinning can be chosen whose every two elements lie far away enough from each other. This is also one of the basic ideas of Paris' original proof. We shall demand somewhat more, and this will be achieved by fitting Lemma 5.2.6 into the construction of such sets.

Theorem 5.2.7 ($I\Sigma_1$)**.** *Let $n \geq 2$.*

(i) $(\mathrm{KM})^n_{|\cdot|_{n \dot{-} 3}} \to (\mathrm{PH})^n$.

(ii) $(\mathrm{KM})^n \to (\mathrm{PH})^n$.

Proof. The second assertion follows from the first. For a better readability we give here a proof of the first assertion for $n = 3$. Though the following construction is general enough, a general proof for arbitrary $n \geq 2$ will be given later.

Given x, z, N let y be such that

$$[x, y] \longrightarrow (m)^3_{reg},$$

where m comes from

$$m \longrightarrow (\ell + 7)^3_4,$$

where $\ell = 21z + 22$. Such an m exists by the finite Ramsey theorem (provable in $I\Sigma_1$). We may assume that x is very much larger than $\max\{7, z, N, M\}$, where M is so large that for all $i \geq M$

$$2|i| + 1 < i, \quad 2^{2^{|i|_3}} < i, \quad \text{and} \quad |i|_3 > 2.$$

These conditions will be used to ensure that the following function g is regressive. Claim that

$$[x, y] \longrightarrow^* (N)^3_z.$$

For this let $f : [x, y]^3 \to z$ and define a regressive function $g : [x, y]^3 \to y$ as follows:

Let $\alpha < \beta < \gamma$ be from $[x, y]$. First define a function Q_γ by:

- $Q_\gamma(0) := x$, $Q_\gamma(1) := x + 1$, and if $Q_\gamma \restriction i$ defined then
- $Q_\gamma(i) :=$ the smallest t such that $t > Q_\gamma(i-1)$, $t < \gamma$, and

$$\forall j,\, k < i\, [\, j < k \longrightarrow f(Q_\gamma(j), Q_\gamma(k), t) = f(Q_\gamma(j), Q_\gamma(k), \gamma)\,].$$

- If such a t does not exist, set $Q_\gamma := Q_\gamma \restriction i$.

Now put for $j < k \in \alpha \cap \mathrm{dom}(Q_\gamma)$

$$R_{\gamma\alpha}(j, k) := f(Q_\gamma(j), Q_\gamma(k), \gamma).$$

Notice that $Q_\gamma \restriction \alpha$ can be regained from f and $R_{\gamma\alpha}$, i.e. γ is not necessary.

For $\beta \in \mathrm{dom}(Q_\gamma)$ define a function $P_{\gamma\beta}$ by

- $P_{\gamma\beta}(0) := x$, and if $P_{\gamma\beta} \restriction i$ defined then
- $P_{\gamma\beta}(i) :=$ the smallest t such that $t > P_{\gamma\beta}(i-1)$, $t \in \mathrm{Im}(Q_\gamma \restriction \beta)$, and

$$\forall j < i (f(P_{\gamma\beta}(j), t, \gamma) = f(P_{\gamma\beta}(j), Q_\gamma(\beta), \gamma)).$$

- If such a t does not exist, set $P_{\gamma\beta} := P_{\gamma\beta} \restriction i$.

Define for $j \in \alpha \cap \mathrm{dom}(P_{\gamma\beta})$

$$S_{\gamma\beta\alpha}(j) := f(P_{\gamma\beta}(j), Q_\gamma(\beta), \gamma)\,.$$

Notice that $P_{\gamma\beta} \restriction \alpha$ can be regained from f, $S_{\gamma\beta\alpha}$ and γ.

Before we define g we need one more assistant function which guarantees some distances between numbers. Remember the three regressive functions η_j, $j = 1, 2, 3$, from Lemma 5.2.6. Applying Lemma 5.2.2 define $\bar\eta_j : [x,y]^3 \to 2|y| + 1$, $j = 1, 2, 3$, such that, if $\bar H$ min-homogeneous for all $\bar\eta_j$, then $H := \bar H - \{\text{the last element of } \bar H\}$ is min-homogeneous for all η_j from Lemma 5.2.6. Let

$$h(\alpha,\beta,\gamma) := \begin{cases} 0 & \text{if } \bar\eta_j(\alpha,\beta,\gamma) = 0 \text{ for each } j \in \{1,2,3\}, \\ j & \text{otherwise, where } j \text{ is the least s.t. } \bar\eta_j(\alpha,\beta,\gamma) \neq 0. \end{cases}$$

Then define g on $[x,y]^3$ as follows:

- If $h(\alpha,\beta,\gamma) = j > 0$, then

$$g(\alpha,\beta,\gamma) := \bar\eta_j(\alpha,\beta,\gamma);$$

- Assume $h(\alpha,\beta,\gamma) = 0$.
  - ➤ Unless $x \le |\alpha|_3 < |\beta|_3 < |\gamma|_3$, then

  $$g(\alpha,\beta,\gamma) := 0.$$

  Assume now additionally $x \le |\alpha|_3 < |\beta|_3 < |\gamma|_3$.
  - ➤ If $|\beta|_3 \notin \mathrm{dom}(Q_{|\gamma|_3})$, then

  $$g(\alpha,\beta,\gamma) := \langle R_{|\gamma|_3|\alpha|_3}, |\gamma|_3 \pmod{2z|\alpha|_3}\rangle.$$

  - ➤ If $|\beta|_3 \in \mathrm{dom}(Q_{|\gamma|_3})$, then

  $$g(\alpha,\beta,\gamma) := \langle R_{|\gamma|_3|\alpha|_3}, S_{|\gamma|_3|\beta|_3|\alpha|_3}, |\gamma|_3 \pmod{2z|\alpha|_3}\rangle.$$

Here $\langle \cdot, \cdots, \cdot \rangle$ are suitable coding functions s.t. for all $\alpha < \beta < \gamma$

$$\langle R_{\gamma\alpha}, S_{\gamma\beta\alpha}, \gamma \pmod{2z\alpha} \rangle \leq 2^{2^{\alpha}}, \tag{$*$}$$

if α is large enough. This is possible since $\mathrm{dom}(R_{\gamma\alpha}) \cup \mathrm{dom}(S_{\gamma\beta\alpha}) \subseteq \alpha$ and $\mathrm{Im}(R_{\gamma\alpha}) \cup \mathrm{Im}(S_{\gamma\beta\alpha}) \subseteq z$. The finite functions R and S are of course to be understood as their codes. (In fact, it is not so important which coding function should be used.) From now on assume that $(*)$ is always satisfied for any $\alpha \geq x$. Then

$$g(\alpha, \beta, \gamma) \leq \max\{\bar{\eta}_j(\alpha, \beta, \gamma), 2^{2^{|\alpha|_3}}\} < \alpha,$$

since $\alpha > M$, so g is regressive.

Let X_0 be min-homogeneous for g and homogeneous for h with $|X_0| \geq \ell + 7$, and set

$$\begin{aligned} X_1 &:= X_0 - \{\text{the last four elements of } X_0\}, \\ X &:= X_1 - \{\text{the first three elements of } X_1\}, \end{aligned}$$

hence $|X| \geq \ell$. Let also Y' be the set of every third element of X and Y the set of every second element of Y', i.e. Y is the set of every 6th element of X, so $|Y| \geq \ell/7 > 3z + 3$.

Claim 1: $h \restriction [X_1]^3$ is the constant function 0.

Proof of Claim 1: Let $a < b < c < d$ be the last four elements of X_0 and assume $h \restriction [X_1]^3 = 1$. Then

$$h \restriction [X_0]^3 = 1 \quad \text{and} \quad g \restriction [X_0]^3 = \bar{\eta}_1 \restriction [X_0]^3.$$

It follows that X_0 is min-homogeneous for $\bar{\eta}_1$, so $X_0 \setminus \{d\}$ is min-homogeneous for η_1. By the proof of Lemma 5.2.6 $\eta_1 \restriction [X_0 \setminus \{c, d\}]^2 = 0$. Hence $\bar{\eta}_1 \restriction [X_0 \setminus \{c, d\}]^3 = 0$ contradicting $h \restriction [X_0]^3 = 1$. Therefore, $h \restriction [X_0]^3 \neq 1$ and $\bar{\eta}_1 \restriction [X_0]^3 = 0$ since it is a constant function. In particular, X_0 is min-homogeneous for $\bar{\eta}_1$, and so $\eta_1 \restriction [X_0 \setminus \{c, d\}]^2 = 0$.

Finally, we can iterate the same argument to show that $\eta_2 \restriction [X_0 \setminus \{b, c, d\}]^2 = 0$ and $\eta_3 \restriction [X_0 \setminus \{a, b, c, d\}]^2 = 0$. It follows that $h \restriction [X_1]^3 \notin \{1, 2, 3\}$, and so we should have $h \restriction [X_1]^3 = 0$. q.e.d.

Then by Lemma 5.2.2 and Lemma 5.2.6 it follows for all $\alpha < \beta \in X_1$ that $2^{\alpha} < \beta$, hence $|\alpha|_3 < |\beta|_3$ if $|\alpha|_3 > 2$. And since there are three elements from X_1 which are smaller than $\min(X)$ we also have $x \leq |\alpha|_3$ for all $\alpha \in X$. Therefore, $g \restriction [X]^3 > 0$.

Furthermore, we need to show the following claims.

(1) Let $\alpha < \beta < \delta < \gamma$ from Y'.

Claim 2: $z|\alpha|_3 < |\delta|_3 - |\beta|_3$.

Proof of Claim 2: Let $\tau < \rho \in X$ such that $\alpha < \tau < \rho < \beta < \delta$. Since

$$g(\alpha, \tau, \rho) = g(\alpha, \tau, \beta) = g(\alpha, \tau, \delta)$$

we have

$$|\rho|_3 \pmod{2z|\alpha|_3} = |\beta|_3 \pmod{2z|\alpha|_3} = |\delta|_3 \pmod{2z|\alpha|_3}.$$

Then for some $n_1 < n_2 \in \mathbb{N}$

$$|\beta|_3 = |\rho|_3 + 2z|\alpha|_3 \cdot n_1 \quad \text{and} \quad |\delta|_3 = |\rho|_3 + 2z|\alpha|_3 \cdot n_2\,,$$

hence $|\delta|_3 - |\beta|_3 > z|\alpha|_3$. q.e.d.

Claim 3: $|\beta|_3 \in \mathrm{dom}(Q_{|\gamma|_3})$, $Q_{|\gamma|_3}(|\beta|_3) < |\delta|_3$, $Q_{|\gamma|_3} \restriction |\beta|_3 = Q_{|\delta|_3} \restriction |\beta|_3$, and $\mathrm{dom}(R_{|\gamma|_3|\alpha|_3}) = |\alpha|_3$.

Proof of Claim 3: Let τ, $\rho \in X$ such that $\beta < \tau < \rho < \delta$. Since

$$g(\beta, \tau, \rho) = g(\beta, \tau, \gamma)$$

we have $R_{|\rho|_3|\beta|_3} = R_{|\gamma|_3|\beta|_3}$ and $Q_{|\rho|_3} \restriction |\beta|_3 = Q_{|\gamma|_3} \restriction |\beta|_3$. Hence, for each $j < k \in |\beta|_3 \cap \mathrm{dom}(Q_{|\rho|_3}) = |\beta|_3 \cap \mathrm{dom}(Q_{|\gamma|_3})$

$$f(Q_{|\rho|_3}(j), Q_{|\rho|_3}(k), |\rho|_3) = f(Q_{|\gamma|_3}(j), Q_{|\gamma|_3}(k), |\gamma|_3).$$

Let $\mu := |\beta|_3 \cap \mathrm{dom}(Q_{|\rho|_3})$. Then $\mu = |\beta|_3$, since otherwise it would follow that $Q_{|\gamma|_3}(\mu)$ were defined. Note that $|\rho|_3 < |\gamma|_3$. This contradicts the definition of μ. Here we used the fact that $Q_{|\rho|_3} \restriction \mu = Q_{|\gamma|_3} \restriction \mu$. In the same way we show that

$$|\beta|_3 \in \mathrm{dom}(Q_{|\gamma|_3}) \quad \text{and} \quad Q_{|\gamma|_3}(|\beta|_3) \leq |\rho|_3 < |\delta|_3.$$

Replacing $|\gamma|_3$ by $|\delta|_3$ above one can see that

$$Q_{|\delta|_3} \restriction |\beta|_3 = Q_{|\gamma|_3} \restriction |\beta|_3.$$

Therefore, $\mathrm{dom}(R_{|\gamma|_3|\alpha|_3}) = |\alpha|_3$. q.e.d.

(2) Let $\alpha < \beta < \delta < \gamma < \eta$ from Y.

Claim 4: $\mathrm{dom}(S_{|\gamma|_3|\beta|_3|\alpha|_3}) = |\alpha|_3$ and $P_{|\gamma|_3|\beta|_3}(|\alpha|_3) < |\beta|_3$.

Proof of Claim 4: Let $\tau \in Y'$ such that $\alpha < \tau < \beta$, then $g(\alpha, \tau, \gamma) = g(\alpha, \beta, \gamma)$. By the same arguments as above we may talk about $S_{|\gamma|_3|\tau|_3|\alpha|_3} = S_{|\gamma|_3|\beta|_3|\alpha|_3}$, i.e.

$$\mu := |\alpha|_3 \cap \mathrm{dom}(P_{|\gamma|_3|\tau|_3}) = |\alpha|_3 \cap \mathrm{dom}(P_{|\gamma|_3|\beta|_3})$$

and for each $j < \mu$

$$f(P_{|\gamma|_3|\tau|_3}(j), Q_{|\gamma|_3}(|\tau|_3), |\gamma|_3) = f(P_{|\gamma|_3|\beta|_3}(j), Q_{|\gamma|_3}(|\beta|_3), |\gamma|_3).$$

(Remember that $S_{|\gamma|_3|\beta|_3|\alpha|_3}$ is a code of a finite function.) As above we can show that $\mu = |\alpha|_3$ and $P_{|\gamma|_3|\beta|_3}(|\alpha|_3) \leq Q_{|\gamma|_3}(|\tau|_3) < |\beta|_3$. q.e.d.

Claim 5: $P_{|\gamma|_3|\delta|_3}(|\alpha|_3) < |\beta|_3$ and $P_{|\delta|_3|\beta|_3} \restriction |\alpha|_3 = P_{|\eta|_3|\gamma|_3} \restriction |\alpha|_3$.

Proof of Claim 5: Let τ be as above. Replacing β by δ above one sees that

$$P_{|\gamma|_3|\delta|_3}(|\alpha|_3) \leq Q_{|\gamma|_3}(|\tau|_3) < |\beta|_3.$$

$P_{|\delta|_3|\beta|_3} \restriction |\alpha|_3 = P_{|\eta|_3|\gamma|_3} \restriction |\alpha|_3$ follows directly from $g(\alpha,\beta,\delta) = g(\alpha,\gamma,\eta)$. q.e.d.

Now let $Y = \{\alpha_i \mid i \leq k\}$ in ascending order and pick sets

$$Z_i \subseteq \mathrm{Im}(P_{|\alpha_k|_3|\alpha_{k-1}|_3} \restriction [\, |\alpha_{3i}|_3, |\alpha_{3i+1}|_3)),$$

where $i < [k/3] - 1$, such that

$$|Z_i| \geq \frac{|\alpha_{3i+1}|_3 - |\alpha_{3i}|_3}{z}$$

and the function $\lambda t.f(t, Q_{|\alpha_k|_3}(|\alpha_{k-1}|_3), |\alpha_k|_3)$ is constant on Z_i, with say constant value $c_i < z$. This is possible since $P_{|\alpha_k|_3|\alpha_{k-1}|_3}$ is strictly increasing. Since $[k/3] - 1 > z$ we have $c_{i_0} = c_{i_1}$for some $i_0 < i_1 < [k/3] - 1$. We claim that $Z_{i_0} \cup Z_{i_1}$ is a large homogeneous set for f.

- f is constant on $[Z_{i_0} \cup Z_{i_1}]^3$:

  Let $u < v < w$ from $Z_{i_0} \cup Z_{i_1}$. Then

$$\begin{aligned} f(u,v,w) &= f(u,v,|\alpha_k|_3) && (\text{by } \{u,v,w\} \subseteq \mathrm{Im}(Q_{|\alpha_k|_3})) \\ &= f(u, Q_{|\alpha_k|_3}(|\alpha_{k-1}|_3), |\alpha_k|_3) && (\text{by def. of } P_{|\alpha_k|_3|\alpha_{k-1}|_3}) \end{aligned}$$

  On the other hand, we have

$$f(u, Q_{|\alpha_k|_3}(|\alpha_{k-1}|_3), |\alpha_k|_3) = f(u', Q_{|\alpha_k|_3}(|\alpha_{k-1}|_3), |\alpha_k|_3)$$

  for all $u' \in Z_{i_0} \cup Z_{i_1}$. So f is homogeneous on $Z_{i_0} \cup Z_{i_1}$.

- $Z_{i_0} \cup Z_{i_1}$ is relatively large:

$$\begin{aligned} \min(Z_{i_0} \cup Z_{i_1}) &\leq \min(Z_{i_0}) < |\alpha_{3i_0+2}|_3 \\ &< \frac{|\alpha_{3i_1+1}|_3 - |\alpha_{3i_1}|_3}{z} \\ &\leq |Z_{i_1}| < |Z_{i_0} \cup Z_{i_1}| \end{aligned}$$

Moreover, $|Z_{i_0} \cup Z_{i_1}| > |Z_{i_1}| \geq \frac{|\alpha_{3i_1+1}|_3 - |\alpha_{3i_1}|_3}{z} > |\alpha_{3i_1-1}|_3 > x > N$. This completes the proof. □

A slight modification of the proof above shows the following.

Theorem 5.2.8 ($I\Sigma_1$)**.** $(\mathrm{KM})^n_{b_{n-2}}$ *implies* $(\mathrm{PH})^n$ *for every* $n \geq 2$.

Corollary 5.2.9. *Let* $n \geq 1$.

(i) $I\Sigma_n \nvdash (\mathrm{KM})^{n+1}_{|\cdot|_{n-2}}$.

(ii) $I\Sigma_n \nvdash (\mathrm{KM})^{n+1}_{b_m}$ *iff* $m \leq n-1$.

Proof. $I\Sigma_n \vdash (\mathrm{KM})^{n+1}_{b_n}$ can be proved similarly as Theorem 4.2.6.(ii). ☐

Remark 5.2.10. Let $n \geq 1$.

(i) At the moment we don't know whether $(\mathrm{KM})^{n+1}_{|\cdot|_{n-1}}$ is $I\Sigma_n$-provable or not. We conjecture that it is unprovable. We believe even that

$$I\Sigma_n \nvdash (\mathrm{KM})^{n+1}_{\sqrt[d]{|\cdot|_{n-1}}}$$

for any $d \geq 1$. As we have seen, this is true for $n = 1$.

(ii) Weiermann has recently shown that in the local level there is a difference between $(\mathrm{PH})^n_f$ and $(\mathrm{KM})^n_f$. Indeed, he could show that

$$I\Sigma_n \vdash (\mathrm{PH})^{n+1}_{|\cdot|_{n-1}}.$$

In the full strength of PA, however, there is no difference between $(\mathrm{PH})_f$ and $(\mathrm{KM})_f$. See the following.

Now we are going to see how fast R_f with respect to the iterated binary functions grows. This will be based upon Weiermann's results with respect to the Paris-Harrington principle. Given $f\colon \mathbb{N} \to \mathbb{N}$ set

$$r^*_f(n, x, c) := \min\{y\colon [x, y] \to^*_f (n+1)^n_c\}.$$

For the rest of this section put

$$f_\alpha := |\cdot|_{H^{-1}_\alpha(\cdot)}.$$

Theorem 5.2.11 (Weiermann [61])**.** *Set* $k^{(s)} := k + 3 + 3^2 + \cdots + 3^s$.

(i) *Let* $n \geq 2$ *and* $k \geq 4$.

$$r^*_{|\cdot|_{n-2}}(n+1, 3_n(n+k+3), k^{(n)}) \geq H_{\omega^k_n}(k-1)$$

(ii) *Let* $n \geq 4$.

$$r^*_{f_{\varepsilon_0}}(n+1, 3_n(2n+3), n^{(n)}) \geq H_{\varepsilon_0}(n-2)$$

Corollary 5.2.12. *Let $n \geq 5$.*

$$r^*_{id}(n, p(n), q(n)) \geq H_{\varepsilon_0}(n-3),$$

where p, $q : \mathbb{N} \to \mathbb{N}$ are some primitive recursive functions.

Theorem 5.2.13. *Let $m \in \mathbb{N}$.*

(i) $(\mathrm{KM})_{|\cdot|_m}$ *is unprovable in* PA.

(ii) $(\mathrm{KM})_{|\cdot|_{H^{-1}_{\varepsilon_0}(\cdot)}}$ *is unprovable in* PA.

Proof. (i) A slight modification of the proof of Lemma 5.2.3 shows

$$R_{|\cdot|_m}(n+m+1, x, k+m+1) \geq R_{id}(n, x, k)$$

if x is sufficiently large. On the other hand, we know by Corollary 5.2.12and the proof of Theorem 5.2.7 that

$$R_{id}(n, p_1(n), q_1(n)) > H_{\varepsilon_0}(n-3),$$

where n is sufficiently large and p_1, q_1 are some primitive recursive functions. So

$$R_{|\cdot|_m}(n+m+1, p_2(n), q_2(n)) > H_{\varepsilon_0}(n-3),$$

where n is sufficiently large and p_2, q_2 are some primitive recursive functions. Hence $R_{|\cdot|_m}$ cannot be provably recursive in PA.

(ii) Using the same notation we have $R_{|\cdot|_m}(n, p_2(n), q_2(n)) > H_{\varepsilon_0}(n-m-4)$ for sufficiently large n. So

$$R_{|\cdot|_n}(2n, p_2(2n), q_2(2n)) > H_{\varepsilon_0}(n-4).$$

We claim for sufficiently large n

$$R_{f_{\varepsilon_0}}(2n, p_2(2n), q_2(2n)) > H_{\varepsilon_0}(n-4).$$

Assume $R_{f_{\varepsilon_0}}(2n, p_2(2n), q_2(2n)) \leq H_{\varepsilon_0}(n-4)$. By definition there is some $|\cdot|_n$-regressive function $G\colon [p_2(2n), H_{\varepsilon_0}(n-4)]^{2n} \to \mathbb{N}$ which has no min-homogeneous set of cardinality $q_2(2n)$. On the other hand, G is $|\cdot|_{H^{-1}_{\varepsilon_0}(\cdot)}$-regressive. In fact, we have for all $i \leq H_{\varepsilon_0}(n-4)$

$$H^{-1}_{\varepsilon_0}(i) \leq n-4,$$

so $|i|_n < |i|_{n-4} \leq |i|_{H^{-1}_{\varepsilon_0}(i)}$. So G has a min-homogeneous set of cardinality $q_2(2n)$. Contradiction! □

A general proof of Theorem 5.2.7

A general proof of the following will be given: Given $n \geq 2$

$$\mathrm{I}\Sigma_1 \vdash (\mathrm{KM})^n_{|\cdot|_{n \dot- 3}} \longrightarrow (\mathrm{PH})^n.$$

Proof. Let $n \geq 2$ be given. Let x, z, N be given and y be such that

$$[x,y] \longrightarrow (m)^n_{|\cdot|_{n\dot-3}-reg},$$

where m comes from

$$m \longrightarrow (\ell + n')^n_4,$$

where

$$n' := \begin{cases} 2n+2 & \text{if } n = 2, \\ 2n+1 & \text{otherwise.} \end{cases}$$

Here $\ell = 3(n!+1)(z+n-2)+1$. Such an m exists by the Finite Ramsey Theorem.

We may assume that x is much larger than $\max\{7, z, N, M(n)\}$, where $M(n)$ is so large that for all $j \geq M(n)$

- $2^{2^{|j|_3}} < j$ and $|j|_3 > 2$ for $n = 2$,
- $f_{n-2}(j) < |j|_{n-3}$, $2^{2^{|j|_n}} < |j|_{n-3}$, and $|j|_n > 2$ for $n \geq 3$,

where

$$f_p(i) := \begin{cases} 2 \cdot \lceil \log(i+1) \rceil + 1 & \text{if } p = 1, \\ 2 \cdot \lceil \log(f_{p-1}(i)+1) \rceil + 1 & \text{if } p \geq 2. \end{cases}$$

These conditions will be used to ensure that the following function g is $|\cdot|_{n\dot-3}$-regressive. (So it depends on n.) Claim

$$[x,y] \longrightarrow^* (N)^n_z\,.$$

For this let $f : [x,y]^n \to z$ and define $g : [x,y]^n \to |y|_{n\dot-3}$ which is a $|\cdot|_{n\dot-3}$-regressive function as follows:

Let $\alpha_1 < \alpha_2 < \cdots < \alpha_n$ be from $[x,y]$. Then we define gradually the following functions $Q^j_{\alpha_n \cdots \alpha_{n-j+1}}$ and $R^j_{\alpha_n \cdots \alpha_{n-j+1}\alpha_1}$ for $j \in \{1, \ldots, n-1\}$.

For a better readability set

$$Q^j_n := Q^j_{\alpha_n \cdots \alpha_{n-j+1}} \quad \text{and} \quad R^j_n := R^j_{\alpha_n \cdots \alpha_{n-j+1}\alpha_1}.$$

Furthermore, we use the abbreviation $\vec{i}_m$ for $i_1, \ldots, i_m$ such that $i_1 < \cdots < i_m$ and $h(\vec{i}_m)$ for $h(i_1), \ldots, h(i_m)$, where $h\colon \mathbb{N} \to \mathbb{N}$.

(1) $j = 1$:

- $Q_n^1(0) := x$, $Q_n^1(1) := x+1$, …, $Q_n^1(n-2) := x+n-2$.

Assume $Q_n^1 \restriction i$ is defined.

- $Q_n^1(i) :=$ the smallest t such that $t > Q_n^1(i-1)$, $t < \alpha_n$, and

$$\forall \vec{i}_{n-1} < i \, [\, f(Q_n^1(\vec{i}_{n-1}), t) = f(Q_n^1(\vec{i}_{n-1}), \alpha_n) \,] \, .$$

- If such a t does not exist, set $Q_n^1 := Q_n^1 \restriction i$.

(2) $1 < j \leq n-1$: Assume $\alpha_i \in \mathrm{dom}(Q_n^{n-i})$ for $i \in \{n-j+1, \ldots, n-1\}$.

- $Q_n^j(0) := x$, …, $Q_n^j(n-j-1) = x+n-j-1$.

Assume $Q_n^j \restriction i$ is defined.

- $Q_n^j(i) :=$ the smallest t such that $t > Q_n^j(i-1)$, $t \in \mathrm{Im}(Q_n^{j-1} \restriction \alpha_{n-j+1})$,

  and for all $i_1 < \cdots < i_{n-j} < i$

$$\begin{aligned} &f(Q_n^j(\vec{i}_{n-j}), t, Q_n^{j-2}(\alpha_{n-j+2}), \ldots, Q_n^1(\alpha_{n-1}), \alpha_n) \\ &\qquad = f(Q_n^j(\vec{i}_{n-j}), Q_n^{j-1}(\alpha_{n-j+1}), \ldots, Q_n^1(\alpha_{n-1}), \alpha_n). \end{aligned}$$

- If such a t does not exist, set $Q_n^j = Q_n^j \restriction i$.

(3) Let $j \in \{1, \ldots, n-1\}$. Put for $i_1 < \cdots < i_{n-j} \in \alpha_1 \cap \mathrm{dom}(Q_n^j)$

$$R_n^j(i_1, \ldots, i_{n-j}) := f(Q_n^j(\vec{i}_{n-j}), Q_n^{j-1}(\alpha_{n-j+1}), \ldots, Q_n^1(\alpha_{n-1}), \alpha_n).$$

Notice that $Q_n^j \restriction \alpha_1$ can be regained from f, $Q_n^{j-1}(\alpha_{n-j+1})$, …, $Q_n^1(\alpha_{n-1})$, R_n^j, and α_n. Before we define g we need one more assistant function which guarantees some distances between numbers.

Remember the three regressive functions η_j, $j = 1, 2, 3$, from Lemma 5.2.6. Applying Lemma 5.2.2 $n-2$ times define $\bar{\eta}_j : [x,y]^n \to |y|_{n-2}$ such that, if $\bar{H}$ is min-homogeneous for all $\bar{\eta}_j$, then $H := \bar{H} - \{$the last $n-2$ elements of $\bar{H}\}$ is min-homogeneous for all η_j. Let $\vec{\alpha}_n$ denote $\alpha_1, \ldots, \alpha_n \in [x,y]$ such that $\alpha_1 < \cdots < \alpha_n$. We define

$$h(\vec{\alpha}_n) := \begin{cases} 0 & \text{if } \bar{\eta}_j(\vec{\alpha}_n) = 0 \text{ for each } j \in \{1,2,3\}, \\ j & \text{otherwise, where } j \text{ is the least s.t. } \bar{\eta}_j(\vec{\alpha}_n) \neq 0. \end{cases}$$

Then define g on $[x,y]^3$ as follows:

- If $h(\vec{\alpha}_n) = j > 0$, then

$$g(\vec{\alpha}_n) := \bar{\eta}_j(\vec{\alpha}_n).$$

- If $h(\vec{\alpha}_n) = 0$ and unless $x \leq |\alpha_1|_{n+1} < \cdots < |\alpha_n|_{n+1}$, then

$$g(\vec{\alpha}_n) := 0.$$

From now on assume that $h(\vec{\alpha}_n) = 0$ and $x \leq |\alpha_1|_{n+1} < \cdots < |\alpha_n|_{n+1}$. For a better readability we use the following abbreviation: For $\beta_1 < \ldots < \beta_m$ and $j \in \{1, \ldots, m-1\}$ set

$$Q^j_{\beta,m} := Q^j_{|\beta_m|_{n^*} \cdots |\beta_{m-j+1}|_{n^*}}$$

and

$$R^j_{\beta,m} := R^j_{|\beta_m|_{n^*} \cdots |\beta_{m-j+1}|_{n^*} |\beta_1|_{n^*}},$$

where $n^* = 3$ if $n = 2$ and $n^* = n$ otherwise.

- If $n = 2$, then

$$g(\alpha_1, \alpha_2) := \langle R^1_{\alpha,2}, |\alpha_2|_3 \pmod{2z|\alpha_1|_3} \rangle.$$

Let $n > 2$.

- If $|\alpha_j|_n \notin \mathrm{dom}(Q^{n-j}_{\alpha,n})$ and $|\alpha_j|_n \in \mathrm{dom}(Q^{n-j}_{\alpha,n})$ for $j = n-1, \ldots, 2$ and $p = n-1, \ldots, j+1$, then

$$g(\vec{\alpha}_n) := \langle R^1_{\alpha,n}, R^2_{\alpha,n}, \ldots, R^{n-j}_{\alpha,n}, |\alpha_n|_n \pmod{2z|\alpha_1|_n} \rangle.$$

- If $|\alpha_j|_n \in \mathrm{dom}(Q^{n-j}_{\alpha,n})$ for $j = 2, \ldots, n-1$, then

$$g(\vec{\alpha}_n) := \langle R^1_{\alpha,n}, R^2_{\alpha,n}, \ldots, R^{n-1}_{\alpha,n}, |\alpha_n|_n \pmod{2z|\alpha_1|_n} \rangle.$$

Here $\langle \cdot, \cdots, \cdot \rangle$ are suitable coding functions s.t. for all $\alpha_1 < \cdots < \alpha_n$

$$\langle R^1_n, \ldots, R^{n-1}_n, \alpha_n \pmod{2z\alpha_1} \rangle \leq 2^{2^{\alpha_1}} \tag{$*$}$$

if α_1 is large enough. (Here without any iterated binary length function.) This is possible since $\mathrm{dom}(R^j_n) \subseteq \alpha$ and $\mathrm{Im}(R^j_n) \subseteq z$. The finite functions R^j_n are of course to be understood as their codes. (In fact, it is not so important which coding function should be used.) From now on assume that $(*)$ is always satisfied for any $\alpha \geq x$. Then g is $|\cdot|_{n \dot{-} 3}$-regressive since $x > M(n)$.

Let X_0 be min-homogeneous for g and homogeneous for h with $|X_0| \geq \ell + n'$, and set

$$X_1 := X_0 - \{\text{the last n+1 elements of } X_0\},$$

$$X := X_1 - \{\text{the first } n^* \text{ elements of } X_1\},$$

hence $|X| \geq \ell$. And define $Y_1, \ldots, Y_{n_1}$ as follows:

$$\begin{aligned} Y_1 &:= \text{ the set of every } n\text{-th element of } X, \\ Y_2 &:= \text{ the set of every } (n-1)\text{-th element of } Y_1, \\ &\vdots \\ Y_{n-1} &:= \text{ the set of every 2nd element of } Y_{n-2}. \end{aligned}$$

Y_{n-1} is then the set of every $n!$-th element of X, hence

$$|Y_{n-1}| > \ell/(n!+1) > 3z + 3(n-2).$$

Claim 1: $h \restriction [X_1]^n$ is the constant function 0.

Proof of Claim 1: Let $a < b < c < d_1 < \cdots < d_{n-2}$ be the last $n+1$ elements of X_0 and assume $h \restriction [X_1]^n = 1$. Then

$$h \restriction [X_0]^n = 1 \quad \text{and} \quad g \restriction [X_0]^n = \bar{\eta}_1 \restriction [X_0]^n.$$

I.e., X_0 is min-homogeneous for $\bar{\eta}_1$, so $X_0 \setminus \{\vec{d}_{n-2}\}$, $\vec{d}_{n-2} := d_1, \ldots, d_{n-2}$, is min-homogeneous for η_1. By the proof of Lemma 5.2.6 $\eta_1 \restriction [X_0 \setminus \{c, \vec{d}_{n-2}\}]^2 = 0$. Hence $\bar{\eta}_1 \restriction [X_0 \setminus \{c, \vec{d}_{n-2}\}]^n = 0$. Therefore, $\bar{\eta}_1 \restriction [X_0]^n = 0$ contradicting $h \restriction [X_0]^n = 1$. This implies $h \restriction [X_0]^n \neq 1$, so $\bar{\eta}_1 \restriction [X_0]^n = 0$ since it is a constant function. In particular, X_0 is min-homogeneous for $\bar{\eta}_1$, and so $\eta_1 \restriction [X_0 \setminus \{c, \vec{d}_{n-2}\}]^2 = 0$.

Finally, we can iterate the same argument to show $\eta_2 \restriction [X_0 \setminus \{b, c, \vec{d}_{n-2}\}]^2 = 0$ and $\eta_3 \restriction [X_0 \setminus \{a, b, c, \vec{d}_{n-2}\}]^2 = 0$. It follows that $h \restriction [X_1]^n \notin \{1, 2, 3\}$, and so we should have $h \restriction [X_1]^n = 0$. q.e.d.

Then by Lemma 5.2.2 and Lemma 5.2.6 it follows for all $\alpha < \beta \in X_1$ that $2^\alpha < \beta$, hence $|\alpha|_{n^*} < |\beta|_{n^*}$ if $|\alpha|_{n^*} > 2$. And since there are n^* elements from X_1 which are smaller than $\min(X)$ we also have $x < |\alpha|_{n^*}$ for all $\alpha \in X$. Therefore, $g \restriction [X]^n > 0$.

Furthermore, we show the following.

Claim 2: For $\alpha_1 < \alpha_2 < \alpha_3$ from Y_1 we have

$$z|\alpha_1|_{n^*} < |\alpha_3|_{n^*} - |\alpha_2|_{n^*}.$$

Proof of Claim 2: Let $\vec{\tau}_{n-1} := \tau_1, \ldots, \tau_{n-1}$ from X such that $\alpha < \tau_1 < \cdots < \tau_{n-1} < \alpha_2 < \alpha_3$. By min-homogeneity

$$g(\alpha_1, \vec{\tau}_{n-1}) = g(\alpha_1, \vec{\tau}_{n-2}, \alpha_2) = g(\alpha_1, \vec{\tau}_{n-2}, \alpha_3),$$

hence

$$\begin{aligned}|\tau_{n-1}|_{n^*} \ (\mathrm{mod}\ 2z|\alpha_1|_{n^*}) &= |\alpha_2|_{n^*} \ (\mathrm{mod}\ 2z|\alpha_1|_{n^*}) \\ &= |\alpha_3|_{n^*} \ (\mathrm{mod}\ 2z|\alpha_1|_{n^*}).\end{aligned}$$

There are some $k_1, k_2 \in \mathbb{N}$ such that $k_1 < k_2$,

$$|\alpha_2|_{n^*} = |\tau_{n-1}|_{n^*} + k_1 \cdot 2z|\alpha_1|_{n^*},$$

and

$$|\alpha_3|_{n^*} = |\tau_{n-1}|_{n^*} + k_2 \cdot 2z|\alpha_1|_{n^*}.$$

Hence $|\alpha_3|_{n^*} - |\alpha_2|_{n^*} > z|\alpha_1|_{n^*}$. q.e.d.

Let $m \in \{1, \dots, n-1\}$ and $\alpha_1, \dots, \alpha_{m+1}, \beta_1, \dots, \beta_m$ be from Y_m such that $\alpha_1 < \beta_1 < \alpha_2 < \cdots < \beta_m < \alpha_{m+1}$.

Claim 3: For $j = 1, \dots, m$ and $p = 1, \dots, j$ we have

$$|\alpha_j|_{n^*} \in \mathrm{dom}(Q^{m-j+1}_{\alpha,m+1}) \quad \text{and} \quad Q^{m-j+1}_{\alpha,m+1}(|\alpha_p|_{n^*}) < |\beta_p|_{n^*}.$$

Proof of Claim 3: We argue by induction on m.

(1) $m = 1$: Take $\vec{\tau}_{n-1} \in X$ such that $\alpha_1 < \tau_1 < \cdots < \tau_{n-1} < \beta_1$. Then the min-homogeneity implies

$$g(\alpha_1, \vec{\tau}_{n-1}) = g(\alpha_1, \vec{\tau}_{n-2}, \alpha_2).$$

It follows that $R^1_{|\tau_{n-1}|_{n^*}\,|\alpha_1|_{n^*}} = R^1_{|\alpha_2|_{n^*}\,|\alpha_1|_{n^*}}$. That is to say,

$$\mu := |\alpha_1|_{n^*} \cap \mathrm{dom}(Q^1_{\tau,n-1}) = |\alpha_1|_{n^*} \cap \mathrm{dom}(Q^1_{\alpha,2})$$

and

$$Q^1_{\tau,n-1} \restriction \mu = Q^1_{\alpha_2,3} \restriction \mu.$$

Therefore, for all $\vec{i}_{n-1}$ such that $i_1 < \cdots < i_{n-2} < \mu$ we have

$$f(Q^1_{\tau,n-1}(\vec{i}_{n-1}), |\tau_{n-1}|_{n^*}) = f(Q^1_{\tau,n-1}(\vec{i}_{n-1}), |\alpha_2|_{n^*}).$$

If $\mu < |\alpha_1|_{n^*}$ this would mean that $Q^1_{\tau,n-1}(\mu)$ were defined, contradicting the definition of μ. In the same way we can show that $|\alpha_1|_{n^*} \in \mathrm{dom}(Q^1_{\alpha,2})$ and

$$Q^1_{\alpha,2}(|\alpha_1|_{n^*}) \leq |\tau_{n-1}|_{n^*} < |\beta_1|_{n^*}.$$

(2) $m \geq 2$: Take $\vec{\tau}_{n-m}$ from Y_{m-1} such that

$$\alpha_1 < \tau_1 < \cdots < \tau_{n-m} < \beta_1 < \alpha_2 < \beta_2 < \cdots < \alpha_{m+1}.$$

By min-homogeneity

$$g(\alpha_1, \vec{\tau}_{n-m}, \alpha_3, \ldots, \alpha_{m+1}) = g(\alpha_1, \vec{\tau}_{n-m-1}, \alpha_2, \ldots, \alpha_{m+1}).$$

On the other hand, we have by I.H.

$$|\alpha_j|_{n^*} \in \mathrm{dom}(Q^{m-j+1}_{\alpha,m+1})$$

for $j = 2, \ldots, m$ and

$$Q^{m-j+1}_{\alpha,m+1}(|\alpha_p|_{n^*}) < |\beta_p|_{n^*}$$

for $p = 2, \ldots, j$. So $|\tau_{n-m}|_{n^*} \in \mathrm{dom}(Q^{m-1}_{\alpha,m+1})$ and

$$R^m_{\alpha,m+1} = R^m_{|\alpha_{m+1}|_{n^*} \cdots |\alpha_3|_{n^*}\, |\tau_{n-m}|\, |\alpha_1|_{n^*}}.$$

This means that

$$\begin{aligned} \mu &:= |\alpha_1|_{n^*} \cap \mathrm{dom}(Q^m_{\alpha,m+1}) \\ &= |\alpha_1|_{n^*} \cap \mathrm{dom}(Q^m_{|\alpha_{m+1}|_{n^*} \cdots |\alpha_3|_{n^*}\, |\tau_{n-m}|_{n^*}}) \end{aligned}$$

and

$$Q^m_{\alpha,m+1} \restriction \mu = Q^m_{|\alpha_{m+1}|_{n^*} \cdots |\alpha_3|_{n^*}\, |\tau_{n-m}|_{n^*}} \restriction \mu.$$

Then for all $i_1 < \cdots < i_{n-m} < \mu$

$$\begin{aligned} &f(Q^m_{\alpha,m+1}(\vec{i}_{n-m}), Q^{m-1}_{\alpha,m+1}(|\alpha_2|_{n^*}), \ldots, Q^1_{\alpha,m+1}(|\alpha_m|_{n^*}), |\alpha_{m+1}|_{n^*}) \\ &\qquad = f(Q^m_{\alpha,m+1}(\vec{i}_{n-m}), Q^{m-1}_{\alpha,m+1}(|\tau_{n-m}|_{n^*}), \ldots, Q^1_{\alpha,m+1}(|\alpha_m|_{n^*}), |\alpha_{m+1}|_{n^*}). \end{aligned}$$

As in the case (1) we can show that $\mu = |\alpha_1|_{n^*}$ and $|\alpha_1|_{n^*} \in \mathrm{dom}(Q^m_{\alpha,m+1})$. Then $Q^m_{\alpha,m+1}(|\alpha_1|_{n^*}) \le Q^{m-1}_{\alpha,m+1}(|\tau_{n-m}|_{n^*}) < |\beta_1|_{n^*}$ by induction hypothesis. q.e.d.

Let $Y_{n-1} = \{\alpha_i \mid i \le k\}$ be in ascending order and pick sets

$$Z_i \subseteq \mathrm{Im}(Q^{n-1}_{\alpha,k} \restriction [\,|\alpha_{3i}|_{n^*}, |\alpha_{3i+1}|_{n^*})),$$

where $i < \lfloor \frac{k}{3} \rfloor - \lceil \frac{n-2}{3} \rceil$, such that

$$|Z_i| \ge \frac{|\alpha_{3i+1}|_{n^*} - |\alpha_{3i}|_{n^*}}{z}$$

and the function

$$t \mapsto f(t, Q^{n-2}_{\alpha,k}(|\alpha_{k-n+2}|_{n^*}), \ldots, Q^1_{\alpha,k}(|\alpha_{k-1}|_{n^*}), |\alpha_k|_{n^*})$$

is constant on Z_i, with say constant value $c_i < z$. This is possible since $Q^{n-1}_{\alpha,k}$ is strictly increasing. Since $\lfloor \frac{k}{3} \rfloor - \lceil \frac{n-2}{3} \rceil > z$ we have $c_{i_0} = c_{i_1}$for some $i_0 < i_1 < \lfloor \frac{k}{3} \rfloor - \lceil \frac{n-2}{3} \rceil$. We claim that $Z_{i_0} \cup Z_{i_1}$ is a large homogeneous set for f.

- f is constant on $[Z_{i_0} \cup Z_{i_1}]^n$: Let $u_1 < \cdots < u_n$ from $Z_{i_0} \cup Z_{i_1}$. Then

$$\begin{aligned} f(\vec{u}_n) &= f(u_1, \ldots, u_{n-1}, |\alpha_k|_{n^*}) \\ &= f(u_1, \ldots, u_{n-2}, Q^1_{\alpha,k}(|\alpha_{k-1}|_{n^*}), |\alpha_k|_{n^*}) \\ &\vdots \\ &= f(u_1, Q^{n-2}_{\alpha,k}(|\alpha_{k-n+2}|_{n^*}), \ldots, Q^1_{\alpha,k}(|\alpha_{k-1}|_{n^*}), |\alpha_k|_{n^*}), \end{aligned}$$

since $u_i \in \mathrm{Im}(Q^2_{\alpha,k})$ for all $i = 1, \ldots, n$.

On the other hand, we have

$$f(u_1, Q^{n-2}_{\alpha,k}(|\alpha_{k-n+2}|_{n^*}), \ldots, Q^1_{\alpha,k}(|\alpha_{k-1}|_{n^*}), |\alpha_k|_3) \\ = f(u'_1, Q^{n-2}_{\alpha,k}(|\alpha_{k-n+2}|_{n^*}, \ldots, Q^1_{\alpha,k}(|\alpha_{k-1}|_{n^*}), |\alpha_k|_{n^*})$$

for all $u'_1 \in Z_{i_0} \cup Z_{i_1}$. So f is homogeneous on $[Z_{i_0} \cup Z_{i_1}]^n$.

- $Z_{i_0} \cup Z_{i_1}$ is large:

$$\begin{aligned} \min(Z_{i_0} \cup Z_{i_1}) &= \min(Z_{i_0}) \\ &< |\alpha_{3i_0+2}|_{n^*} \\ &< \frac{|\alpha_{3i_1+1}|_{n^*} - |\alpha_{3i_1}|_{n^*}}{z} \\ &\leq |Z_{i_1}| \ < \ |Z_{i_0} \cup Z_{i_1}| \end{aligned}$$

In addition, $|Z_{i_0} \cup Z_{i_1}| > |Z_{i_1}| \geq \frac{|\alpha_{3i_1+1}|_{n^*} - |\alpha_{3i_1}|_{n^*}}{z} > |\alpha_{3i_1-1}|_{n^*} > x > N$. □

Part III

Beyond PA

Chapter 6

Kruskal's theorem and $\vartheta\Omega^\omega$

Simpson [49] presented some Friedman style independence results about of finite rooted trees. They are based in part on the existence of a close relationship between finite trees on the one hand and ordinal notation systems well-known in Gentzen-style proof theory on the other. It is shown that both Kruskal's theorem and its miniaturization are too strong to be provable in ATR_0.

Moreover, Rathjen and Weiermann [43] showed that Kruskal's theorem is much stronger than ATR_0. The proof-theoretic strength of ACA_0 with Kruskal's theorem is *the small Veblen ordinal* $\vartheta\Omega^\omega$, also called *Ackermann's ordinal*, and much bigger than Γ_0, the proof-theoretic strength of ATR_0.

In this chapter we show that Friedman style miniaturization could be an adequate tool for showing phase transitions with respect to provability in theories beyond PA.

6.1 Kruskal's theorem

A *finite rooted tree* is a finite partial ordering $(T, \preceq)$ such that, if T is not empty,

- T has a smallest element called the *root* of T;
- for each $b \in T$ the set $\{a \in T \colon a \preceq b\}$ is totally ordered.

Let $a \wedge b$ denote the infimum of a and b for a, $b \in T$. Given finite rooted trees T_1 and T_2, a *homeomorphic embedding* of T_1 into T_2 is an one-to-one mapping $f \colon T_1 \to T_2$ such that $f(a \wedge b) = f(a) \wedge f(b)$ for all a, $b \in T_1$. We write $T_1 \trianglelefteq T_2$ if there exists a homeomorphic embedding $f \colon T_1 \to T_2$, and T_1 is said to be *homeomorphically embeddable* into T_2.

Theorem 6.1.1 (Kruskal's theorem [35])**.** *For every sequence of finite rooted trees $(T_k)_{k<\omega}$, there are indices $\ell < m$ satisfying $T_\ell \trianglelefteq T_m$.*

Note that Kruskal's theorem is a Π^1_1 sentence. Let $\mathbb{T}$ be the set of all finite rooted trees. Kruskal's theorem says that $\langle \mathbb{T}, \trianglelefteq \rangle$ is a well-partial-ordering.

Theorem 6.1.2 (Friedman)**.** Kruskal's theorem *is not provable in* ATR_0*.*

Proof. In ACA_0 the well-foundedness of $\langle \mathbb{T}, \trianglelefteq \rangle$ implies that of Γ_0, see [49]. □

Let $\|T\|$ denote the number of nodes of the finite tree T. Assume further that the set of finite rooted trees is coded primitive recursively into a set of natural numbers as usual. Given $f \colon \mathbb{N} \to \mathbb{N}$ let $\mathrm{SWP}(\mathbb{T}, \trianglelefteq, f)$ be the following Π^0_2 sentence:

> For any k there exists a constant n so large that, for any finite sequence $T_0, \dots, T_n$ of finite rooted trees with $\|T_i\| \leq k + f(i)$ for all $i \leq n$, there are indices $\ell < m \leq n$ satisfying $T_\ell \trianglelefteq T_m$.

Theorem 6.1.3 (Friedman [49], Smith [51])**.** $\mathrm{ATR}_0 \nvdash \mathrm{SWP}(\mathbb{T}, \trianglelefteq, +)$

It is furthermore possible to give a threshold for provability. According to Otter [39] there are two positive real numbers α, β such that

$$t(n) \sim \beta \cdot \alpha^n \cdot n^{-\frac{2}{3}},$$

where $t(n)$ is the cardinality of $\{T \colon \|T\| = n\}$. $\alpha = 2.9557652856\dots$ is called *Otter's tree constant.*

Theorem 6.1.4 (Weiermann [60])**.** *Let* $c = \frac{1}{\log(\alpha)}$ *and* r *be a primitively recursive real number. Set* $f_r(i) := r \cdot |i|$*.*

(i) *If* $r > c$*, then* $\mathrm{SWP}(\mathbb{T}, \trianglelefteq, f_r)$ *is* PA*-unprovable.*

(ii) *If* $r \leq c$*, then* $\mathrm{SWP}(\mathbb{T}, \trianglelefteq, f_r)$ *is* PRA*-provable.*

Using the general approach developed in Simpson [49] and Smith [51], Theorem 1.3.4, and the results from Rathjen and Weiermann [43], especially the proof of Theorem 1.2, it is possible to show that Theorem 6.1.4 can be extended to $\mathrm{ACA}_0 + \Pi^1_2\text{-BI}$. That is, we have phase transitions of provability in $\mathrm{ACA}_0 + \Pi^1_2\text{-BI}$ with respect to Kruskal's theorem:

Theorem 6.1.5. *Let* c*,* $r > c$*,* f_r *be as above.*

$$\mathrm{ACA}_0 + \Pi^1_2\text{-BI} \nvdash \mathrm{SWP}(\mathbb{T}, \trianglelefteq, f_r)$$

Proof. Cf. the proof of Theorem 6.2.8. □

6.2 Ordinal notation systems of $\vartheta\Omega^\omega$

The proof-theoretic ordinal $\vartheta\Omega^\omega$ of $\mathrm{ACA}_0+\Pi^1_2$-BI is given by the study of Rathjen and Weiermann [43]. It is called *the small Veblen ordinal* or *Ackermann's ordinal*. The ordinal notation system for $\vartheta\Omega^\omega$ is based on fixed-point free Veblen functions. We summarize some results from [43].

Let On be the class of ordinals. $AP := \{\alpha \in On\colon \exists\beta \in On(\alpha = \omega^\beta)\}$ is the class of additive principle ordinals, and $E := \{\alpha \in On\colon \alpha = \omega^\alpha\}$ is the class of ε-numbers. Let $\lambda\eta\,.\,\varepsilon_\eta$ enumerate the elements of E. By Cantor's normal form theorem, for every $\alpha \notin E \cup \{0\}$, there are uniquely determined ordinals β and δ such that $\alpha =_{NF} \omega^\beta + \delta$.

Let $\Omega := \aleph_1$. For any $\alpha < \varepsilon_{\Omega_1}$, $E_\Omega(\alpha)$ denotes the ε-numbers below Ω which are needed for the unique representation of α in Cantor normal form:

- $E_\Omega(0) := E_\Omega(\Omega) := \varnothing$;
- $E_\Omega(\alpha) := \{\alpha\}$ if $\alpha \in E \cap \Omega$;
- $E_\Omega(\alpha) := E_\Omega(\beta) \cup E_\Omega(\delta)$ if $\alpha =_{NF} \omega^\beta + \delta$.

Let $\alpha^* := \max(E_\Omega(\alpha) \cup \{0\})$.

Given $\beta < \Omega$ we define $C(\alpha, \beta)$ and $C_n(\alpha, \beta)$ by main recursion $\alpha < \varepsilon_{\Omega+1}$ and subsidiary recursion $n < \omega$ as follows:

- $\{0, \omega\} \cup \beta \subseteq C_n(\alpha, \beta)$;
- if γ, $\delta \in C_n(\alpha, \beta)$ and $\xi =_{NF} \omega^\gamma + \delta$, then $\xi \in C_{n+1}(\alpha, \beta)$;
- if $\delta \in C_n(\alpha, \beta) \cap \alpha$, then $\vartheta\delta \in C_{n+1}(\alpha, \beta)$;
- $C(\alpha, \beta) := \bigcup\{C_n(\alpha, \beta)\colon n < \omega\}$;
- $\vartheta\alpha := \min\{\xi < \Omega\colon C(\alpha, \xi) \cap \Omega \subseteq \xi \text{ and } \alpha \in C(\alpha, \xi)\}$.

Lemma 6.2.1 (Rathjen and Weiermann [43])**.** *Let* $\alpha < \varepsilon_{\Omega+1}$.

(i) $\vartheta\alpha$ *is well-defined and* $\vartheta\alpha \in E$.

(ii) $\alpha \in C(\alpha, \vartheta\alpha)$.

(iii) $\vartheta\alpha \in C(\alpha, \vartheta\alpha) \cap \Omega$ *and* $\vartheta\alpha \notin C(\alpha, \vartheta\alpha)$.

(iv) $\gamma \in C(\alpha, \beta)$ *iff* $\gamma^* \in C(\alpha, \beta)$.

(v) $\alpha^* < \vartheta\alpha$.

(vi) *If* $\vartheta\alpha = \vartheta\beta$, *then* $\alpha = \beta$.

(vii) $\vartheta\alpha = \vartheta\beta$ *iff* $(\alpha < \beta \wedge \alpha^* < \vartheta\beta) \vee (\beta < \alpha \wedge \vartheta\alpha \leq \beta^*)$.

(viii) $\beta < \vartheta\alpha$ *iff* $\omega^\beta < \vartheta\alpha$.

Below we shall frequently draw on the following result.

Lemma 6.2.2 (Rathjen and Weiermann [43]). *Let* $\alpha \in E \cap \vartheta\Omega^\omega$. *Then there exist uniquely defined ordinals* $n \in \omega$ *and* $\alpha_0, \ldots, \alpha_n < \alpha$ *such that*

$$\alpha = \vartheta(\Omega^n \cdot \alpha_n + \cdots + \Omega^0 \cdot \alpha_0),$$

and $\alpha_n \neq 0$ *if* $n \neq 0$.

The question how really strong Kruskal's theorem is answered by:

Theorem 6.2.3 (Rathjen and Weiermann [43]). *In* ACA_0, Kruskal's theorem *and* $\mathrm{WF}(\vartheta\Omega^\omega)$ *are equivalent.*

We are now going to give an ordinal notation system for Ackermann's ordinal $\vartheta\Omega^\omega$ as a set of terms. Let $[\alpha_1, \ldots, \alpha_n]$ denote the multiset of $\alpha_1, \ldots, \alpha_n$.

Assume a constant symbol o and $(j+1)$-ary function symbols f_j are given. The sets S, P, M are defined as follows:

- $o \in S$;
- if $\alpha_0, \ldots, \alpha_j \in S$, then $f_j\alpha_0 \cdots \alpha_j \in P \subseteq S$;
- if $\alpha_0, \ldots, \alpha_{m+1} \in P$, then $[\alpha_0, \ldots, \alpha_{m+1}] \in M \subseteq S$.

Instead of defining an ordering on S directly we give a correspondence σ between S and $\vartheta\Omega^\omega$.

- $\sigma(o) := 0$;
- $\sigma(f_0\alpha) = \vartheta\alpha$;
- assume $j \geq 1$.
  - ➤ if $\sigma(\alpha_0) < \omega$, then
    $\sigma(f_j\alpha_0 \cdots \alpha_j) := \vartheta(\Omega^j \cdot (\sigma(\alpha_0) + 1) + \Omega^{j-1} \cdot \sigma(\alpha_1) + \cdots + \Omega^0 \cdot \sigma(\alpha_j))$;
  - ➤ if $\sigma(\alpha_0) \geq \omega$, then
    $\sigma(f_j\alpha_0 \cdots \alpha_j) := \vartheta(\Omega^j \cdot \sigma(\alpha_0) + \Omega^{j-1} \cdot \sigma(\alpha_1) + \cdots + \Omega^0 \cdot \sigma(\alpha_j))$;
- $\sigma([\alpha_0, \ldots, \alpha_{m+1}]) := \omega^{\sigma(\alpha_0)} \# \cdots \# \omega^{\sigma(\alpha_{m+1})}$.

Here $\#$ denotes the natural sum of ordinals. Then by Lemma 6.2.2 we know that $\sigma\colon S \to \vartheta\Omega^\omega$ is surjective. The intention is that f_j represents a $j+1$-ary fixed-point free version of Veblen function. Defining the binary relation $\prec$ on S canonically, we have just shown following lemma.

Lemma 6.2.4. *$\prec$ is a wpo on S and $o(S) = \vartheta\Omega^\omega$.*

Proof. See also Schmidt [44]. □

The following lemma is easy to show and very useful. We write $f_j\bar{\alpha}$ for $f_j\alpha_0\cdots\alpha_j$. And $\prec_{lex}$ denotes the lexicographic ordering on S^*, the set of all finite sequences of elements from S, which is based on $\prec$.

Lemma 6.2.5.

(i) *$f_{j+1}\bar{0}$ is the first infinite ordinal closed under $+$ and f_k, $k \le j$.*

(ii) *$f_j\bar{\alpha} \prec f_j\bar{\gamma}$ is true if one of the following holds:*

- *$f_j\bar{\alpha} \preceq \gamma_i$ for some $i \le j$;*
- *$(\bar{\alpha}) \prec_{lex} (\bar{\gamma})$ and all $\alpha_i \prec f_j\bar{\gamma}$.*

We define a norm function on S.

Definition 6.2.6. Inductive definition of $\|\cdot\| \colon S \to \omega$.

- $\|o\| := 0$;
- $\|f_j\alpha_0\cdots\alpha_j\| := 1 + j + \|\alpha_0\| + \cdots + \|\alpha_j\|$;
- $\|[\alpha_0,\ldots,\alpha_{m+1}]\| := \|\alpha_0\| + \cdots + \|\alpha_{m+1}\|$.

Note that $\|\alpha\| > 0$ for any $\alpha \in P$, hence $\|\cdot\|$ is really a norm function. Given $f \colon \mathbb{N} \to \mathbb{N}$ define $\mathrm{SWP}(S, \preceq, f)$ as follows:

> For any k there exists a constant n which is so large that, for any finite sequence $\alpha_0,\ldots,\alpha_n$ from S with $\|\alpha_i\| \le k + f(i)$ for all $i \le n$, there are indices $\ell < m \le n$ satisfying $\alpha_i \preceq \alpha_j$.

Let F_f be the Skolem function of $\mathrm{SWP}(S, \preceq, f)$. By *König's Lemma* F_f is a total function for any f. And by Theorem 1.3.4 we have

Theorem 6.2.7. $\mathrm{ACA}_0 + \Pi^1_2\text{-BI} \nvdash \mathrm{SWP}(S, \preceq, id)$

In particular, F_{id} is not provably total in $\mathrm{ACA}_0 + \Pi^1_2$-BI. We shall even see that there is a phase transition of provability.

Given a real number r let $f_r(i) := r|i|$.

Theorem 6.2.8. *There is a real number r_0 such that the following hold for any primitively recursive real number r.*

(i) $\mathrm{ACA}_0 + \Pi^1_2\text{-BI} \nvdash \mathrm{SWP}(S, \preceq, f_r)$ *if $r > r_0$.*

(ii) $\mathrm{PRA} \vdash \mathrm{SWP}(S, \preceq, f_r)$ *if $r \le r_0$.*

Proof. See Section 6.4. □

6.3 Generating functions

In searching for thresholds for provability via Friedman style miniaturizations, we have seen that norm functions play the crucial role. Here we will demonstrate again the importance of generating function methods to show how one can deal with norm functions.

We start with a classification of certain subsystems of S which build a cumulative hierarchy for S. The following process in dealing with generating functions is based on that of Woods [63].

For any natural number d we define simultaneously S^d, P^d, M^d as follows:

- $o \in S^d$;
- if $j \leq d$ and $\alpha_0, \dots, \alpha_j \in S^d$, then $f_j\alpha_0 \cdots \alpha_j \in P^d \subseteq S^d$;
- if $\alpha_0, \dots, \alpha_{m+1} \in P^d$, then $[\alpha_0, \dots, \alpha_{m+1}] \in M^d \subseteq S^d$.

With S, S^d we associate S_ℓ and S^d_ℓ as follows:

$$S_\ell := \{\alpha \in S \colon \|\alpha\| = \ell\} \quad \text{and} \quad S^d_\ell := \{\alpha \in S^d \colon \|\alpha\| = \ell\}$$

$S_{\leq \ell}$, $S^d_{\leq \ell}$, M_ℓ, P_ℓ, etc. are defined similarly. Let $s_\ell := \overline{\overline{S}}_\ell$, $s^d_\ell := \overline{\overline{S}}{}^d_\ell$ and so on, and $S(z)$, $S^d(z)$, etc. be the corresponding generating functions:

$$S(z) = \sum_{\ell=0}^{\infty} s_\ell \cdot z^\ell, \quad S^d(z) = \sum_{\ell=0}^{\infty} s^d_\ell \cdot z^\ell, \quad \text{etc.}$$

Using the admissible operators from Section 1.3.3, we get the following equations of the generating functions.

$$\begin{aligned} S(z) &= 1 + P(z) + M(z), \\ P(z) &= \sum_{\ell=0}^{\infty} (z \cdot S(z))^{\ell+1} = -1 + \sum_{\ell=0}^{\infty} (z \cdot S(z))^\ell, \\ M(z) &= \mathfrak{M}(P(z)) - (1 + P(z)), \end{aligned} \tag{6.1}$$

where $\mathfrak{M}(f(z)) := \exp(\sum_{\ell=1}^{\infty} f(z^\ell)/\ell)$ is the multiset operator. Furthermore,

$$\begin{aligned} S^d(z) &= 1 + P^d(z) + M^d(z), \\ P^d(z) &= \sum_{\ell=0}^{d} (z \cdot S^d(z))^{\ell+1}, \\ M^d(z) &= \mathfrak{M}(P^d(z)) - (1 + P^d(z)), \end{aligned} \tag{6.2}$$

Then by (6.1) and (6.2)

$$\begin{aligned} S(z) &= 1 + P(z) + M(z) = \mathfrak{M}(P(z)), \\ P(z) &= -1 + \sum_{\ell=0}^{\infty} (z \cdot \mathfrak{M}(P(z)))^\ell, \end{aligned}$$

and

$$\begin{aligned} S^d(z) &= 1 + P^d(z) + M^d(z) = \mathfrak{M}(P^d(z)), \\ P^d(z) &= -1 + \sum_{\ell=0}^{\infty} (z \cdot \mathfrak{M}(P^d(z)))^\ell. \end{aligned}$$

We can use simple bounds on s_n to show that $S(z)$ has a positive radius of convergence (r.o.c.) $\rho < 1$. Note just that the sequence and multiset operator are admissible operators with positive *r.o.c.*s. Note also that S, P, M have the same *r.o.c.* ρ. We shall work with $P(z)$ to get some information about ρ, since it is easier to handle. We won't calculate ρ concretely and just refer to some programs such as Mathematica or Maple.

Since all the coefficients of $P(z)$ are positive, $z = \rho$ is a singularity of $P(z)$ by Pringsheim's lemma, Theorem 1.3.10. Hence for z, $|z| < \rho$, we have

$$P(z) = -1 + \frac{1}{1 - z \cdot \mathcal{M}(P(z))},$$

i.e.

$$\frac{P(z)}{1 + P(z)} = z \cdot \mathcal{M}(P(z)). \tag{6.3}$$

This implies $P(x)$ converges as $x \longrightarrow \rho^-$ for $x \in \mathbb{R}$, hence for all z with $|z| = \rho$ $P(z)$ converges and satisfies (6.3). Let $g(z, w) := (1 + w) \cdot e^w \cdot G(z)$, where

$$G(z) = \exp\left(\sum_{\ell \geq 2} \frac{P(z^\ell)}{\ell} \right).$$

Then we have

$$P(z) = z \cdot g(z, P(z)). \tag{6.4}$$

Since $\rho < 1$ is the *r.o.c.* of $P(z)$, $g(z, w)$ is holomorphic (i.e. analytic in z, w separately and continuous) for $|z| < \rho^{1/2}$. The implicit function theorem says that if $|z_0| \leq \rho$ and $w_0 = P(z_0)$, then unless

$$z_0 \frac{\partial g}{\partial w}(z_0, w_0) = 1, \tag{6.5}$$

there is a neighborhood of z_0 in which the equation $w = z \cdot g(z, w)$ has a unique solution with $w = w_0$ at $z = z_0$, which must be (an analytic continuation of) $w = P(z)$. From (6.4) it follows that

$$\begin{aligned} z \cdot \frac{\partial g}{\partial w} &= z \cdot (e^w \cdot G(z) + (1 + w) \cdot e^w \cdot G(z)) \\ &= z \cdot (2 + w) \cdot e^w \cdot G(z)\,. \end{aligned}$$

By (6.3)

$$\begin{aligned}\rho(e^{P(\rho)} \cdot G(\rho) + (1+P(\rho)) \cdot e^{P(\rho)} \cdot G(\rho)) &= \rho \cdot e^{P(\rho)} \cdot G(\rho) + P(\rho) \\ &= 1\end{aligned} \tag{6.6}$$

and $\rho(2+P(\rho)) \cdot e^{P(\rho)} \cdot G(\rho) = 1$, that is,

$$\rho \cdot e^{P(\rho)} \cdot G(\rho) = \frac{1}{2+P(\rho)}\,. \tag{6.7}$$

By (6.7) and (6.6) $P(\rho)^2 + P(\rho) - 1 = 0$, hence

$$P(\rho) = \frac{-1+\sqrt{5}}{2}\,. \tag{6.8}$$

Notice that (6.8) is true for every z_0, $|z_0| = \rho$, at which $P(z)$ fails to be analytic.

On the other hand, if $|z_0| = \rho$ and $P(z_0) = P(\rho)$, then $|P(\rho)| = P(|z_0|)$. Since, however, all the coefficients p_n, p_{n+1} are positive, it follows that $|p_n + p_{n+1} \cdot z_0| = p_n + p_{n+1} \cdot |z_0|$ which is possible only if $z_0 = |z_0| = \rho$. Therefore, $z = \rho$ is the only singularity on the circle $|z| = \rho$ in the complex plane.

Theorem 6.3.1. *The generating function $S(z)$ has the positive r.o.c. $\rho < 1$ which is the only singularity on the circle $|z| = \rho$ in the complex plane.*

Proof. It follows directly from (6.1) since the generating function $S(z)$, $P(z)$ and $M(z)$ have the same r.o.c. □

Now we use the Weierstrass' preparation theorem, Theorem 1.3.11, to show that the singularity of $S(z)$ at $z = \rho$ turns out to be a branch point. From (6.1) we know

$$S(z) = \mathcal{M}\left(\sum_{\ell=1}^{\infty}(z \cdot S(z))^\ell\right) = \exp\left(\frac{z \cdot S(z)}{1 - z \cdot S(z)}\right) \cdot H(z),$$

where $H(z) = \exp\left(\sum_{\ell=2}^{\infty} \frac{\left(\sum_{k=1}^{\infty}(z^\ell \cdot S(z^\ell))^k\right)}{\ell}\right)$. Set

$$g(z,w) = \exp\left(\frac{z \cdot w}{1 - z \cdot w}\right) \cdot H(z).$$

This is holomorphic (i.e., analytic in z, w separately and continuous) for $|z| < \rho^{1/2}$, and we have

$$S(z) = g(z, S(z)).$$

Set $F(z,w) = g(z,w) - w$, $z_0 = \rho$, and $w_0 = S(\rho)$. We claim

$$F(z_0,w_0) = 0, \quad F(z_0,w) \not\equiv 0, \quad \frac{\partial F}{\partial w}(z_0,w_0) = 0, \quad \text{and} \quad \frac{\partial^2 F}{\partial w^2}(z_0,w_0) \neq 0.$$

Still to show is $\frac{\partial^2 F}{\partial w^2}(z_0, w_0) \neq 0$. By definition it follows that

$$\begin{aligned} \frac{\partial F}{\partial w}(z,w) &= \frac{z}{(1-z\cdot w)^2}\cdot \exp\left(\frac{z\cdot w}{1-z\cdot w}\right)\cdot H(z) - 1, \qquad (6.9) \\ \frac{\partial^2 F}{\partial w^2}(z,w) &= \frac{z^2}{(1-z\cdot w)^3}\cdot\left(\frac{1}{1-z\cdot w}+2\right)\cdot \exp\left(\frac{z\cdot w}{1-z\cdot w}\right)\cdot H(z) \\ &= \left(\frac{\partial F}{\partial w}(z,w)+1\right)\cdot\frac{z}{1-z\cdot w}\cdot\left(\frac{1}{1-z\cdot w}+2\right). \end{aligned}$$

For $z \neq 0$, $\frac{\partial F}{\partial w}(z,w) = \frac{\partial^2 F}{\partial w^2}(z,w) = 0$ implies $z\cdot w = \frac{3}{2}$. On the other hand,

$$F(z_0, w_0) = \exp\left(\frac{z_0\cdot w_0}{1-z_0\cdot w_0}\right)\cdot H(z_0) - w_0 = 0,$$

so by (6.9)

$$\frac{z_0\cdot w_0}{(1-z_0\cdot w_0)^2} = 1.$$

This means that $\frac{\partial^2 F}{\partial w^2}(z_0, w_0) \neq 0$ if $z_0\cdot w_0 = \frac{3}{2}$.

According to the Weierstrass' preparation theorem, the equation $F(z,w) = 0$ is locally equivalent to the equation

$$A_0(z) + A_1(z)w + w^2 = 0, \qquad (6.10)$$

where $A_0(z)$ and $A_1(z)$ are analytic in some neighborhood of $z_0 = \rho$. Following the arguments in Section 3.12 of [38], we can show that $z_0 = \rho$ is actually a branch point. In fact, in a neighborhood of $z_0 = \rho$, the analytic continuations of $S(z)$ at all points other than $z_0 = \rho$ are given by

$$S(z) = h(\sqrt{\rho - z}) = 1 + h_1\sqrt{\rho - z} + h_2(\rho - z) + h_3(\sqrt{\rho - z})^3 + \cdots, \qquad (6.11)$$

where $h_1 \neq 0$ and

$$h(w) = 1 + h_1 w + h_2 w^2 + h_3 w^3 + \cdots$$

is an analytic function in a neighborhood of $w = 0$. The following lemma asserts that the coefficients s_n of the power series $S(z)$ are asymptotic to those of $h_1\sqrt{\rho - z}$ expanded (by the binomial theorem) about $z = 0$.

Lemma 6.3.2 (Darboux)**.** *Suppose $a(z) = a_0 + a_1 z + a_2 z^2 + \cdots$ has r.o.c. ρ, and has no singularities other than $z = \rho$ on the circle $|z| = \rho$. If in a neighborhood of $z = \rho$*

$$a(z) = h_0 + h_1\sqrt{\rho - z} + h_2(\rho - z) + h_3(\rho - z)^{3/2} + \cdots$$

with $h_1 \neq 0$, where $h(w) = h_0 + h_1 w + h_2 w^2 + \cdots$ is analytic in a neighborhood of $w = 0$, then for each $m \geq 0$,

$$a_\ell = \frac{-h_1}{2\sqrt{\pi\tau}} \frac{\tau^3}{\ell^{3/2}} \left\{ 1 + \frac{c_1}{\ell} + \frac{c_2}{\ell^2} + \cdots + \frac{c_m}{\ell^m} + \mathcal{O}_m\left(\frac{1}{\ell^{m+1}}\right) \right\},$$

where $\tau = \rho^{-1}$, $c_1, c_2, \ldots, c_m$ are constants, and the subscript m indicates that the implied $\mathcal{O}$ constant may depend on m. More generally, if m is the least odd number such that $h_m \neq 0$, but all the other conditions hold, then $a_\ell \sim C \cdot \rho^{-\ell} \cdot \ell^{-(m+2)/2}$ for some constant C.

Proof. See e.g. Wilf [62]. □

Together with this lemma, (6.11) implies that

$$s_\ell \sim C \cdot \rho^{-\ell} \cdot \ell^{-3/2}$$

for some constant $C > 0$. Harary, Robinson, and Schwenk [27] gave an algorithmic way to deal with such arguments above.

Up to now we have only talked about the generating function $S(z)$, i.e., the case which has no restriction on the arity of f_j. We, however, can see that the arguments above can be slightly adapted to the function $S^d(z)$.

Theorem 6.3.3. *Let ρ and ρ_d, $d \geq 1$, be the r.o.c.s of $S(z)$ and $S^d(z)$, resp.*

(i) *The sequence $(\rho_d)_{d\geq 1}$ is decreasing and converges to ρ.*

(ii) *There is a real number $C > 0$ such that*

$$s_\ell \sim C \cdot \rho^{-\ell} \cdot \ell^{-3/2}.$$

(iii) *There are real numbers $C_d > 0$ such that*

$$s_\ell^d \sim C_d \cdot \rho_d^{-\ell} \cdot \ell^{-3/2}.$$

Proof. It remains to show (i). Obviously we have $\rho_d \geq \rho_{d+1} \geq \rho$. Thus $(\rho_d)_{d\geq 1}$ converges, say to $\rho_\infty \geq \rho$. Note that, since $\rho_d < 1$, we have for any z such that $|z| < \rho_d^2$

$$S^d(z) = g_d(z, S^d(z)),$$

where $g_d(z, w) = \exp(zw + z^2w^2 + \cdots + z^{d+1}w^{d+1}) \cdot H_d(z)$ and $H_d(z)$ depends only on z and d.

Put $\alpha_d := S^d(\rho_d)$ and $f(z) := z + 2z^2 \cdot \alpha_d + \cdots + (d+1) \cdot z^{d+1} \cdot \alpha_d^d$. Then

$$\frac{\partial g_d}{\partial w}(\rho_d, \alpha_d) = f(\rho_d) \cdot g_d(\rho_d, \alpha_d) = 1,$$

hence

$$\frac{1}{f(\rho_1)} \leq \alpha_d = S^d(\rho_d) = g_d(\rho_d, \alpha_d) = \frac{1}{f(\rho_d)} \leq \frac{1}{f(\rho_\infty)}.$$

So α_d must be bounded, say by $M > 0$. It also means that

$$\lim_{d\to\infty} S^d(\rho_\infty) \leq M.$$

Now assume $\rho_\infty > \rho$. Then there is an n satisfying

$$\sum_{\ell=0}^{n} s_\ell\, \rho_\infty^\ell > M.$$

This yields a contradiction:

$$M < \sum_{\ell=0}^{n} s_\ell\, \rho_\infty^\ell = \sum_{\ell=0}^{n} s_\ell^n \rho_\infty^\ell \leq \sum_{\ell=0}^{\infty} s_\ell^n \rho_\infty^\ell \leq M$$

So we should have $\rho_\infty = \rho$. □

6.4 Phase transitions in $\mathrm{ACA}_0 + \Pi^1_2$-BI

We prove Theorem 6.2.8 with $r_0 := \frac{1}{\log(\rho^{-1})}$. Given a primitively recursive real number r let $f_r(i) := r|i|$.

Provability

We need the following lemma called Schur's theorem.

Lemma 6.4.1 (Schur)**.** *Let $U(z) = \sum_{\ell=0}^{\infty} u_\ell \cdot z^\ell$ and $V(z) = \sum_{\ell=0}^{\infty} v_\ell \cdot z^\ell$ be two power series such that for some $\tau \geq 0$*

- *$V(z)$ has the r.o.c. τ, and*
- *$U(z)$ has the r.o.c. larger than τ.*

Then

$$\lim_{\ell\to\infty} \frac{[z^\ell](U(z)\cdot V(z))}{v_\ell} = U(\tau).$$

Proof. See e.g. [11]. □

Proof of Theorem 6.2.8: provability

By Cauchy's formula for the product of two power series we have

$$\sum_{\ell=0}^{\infty} s_{\leq \ell} \cdot z^{\ell} = \frac{1}{1-z} \cdot S(z).$$

Employing Lemma 6.4.1, we find a natural number D so large that

$$s_{\leq i} < \frac{1}{1-\eta^{-1}} \cdot \frac{11}{10} \cdot C \cdot \eta^{i} \cdot i^{-3/2}$$

for any $i \geq D$. Let $k > 2$ be given, where $\eta := \rho^{-1}$. Note that $\eta^{r_0} = 2$. Put

$$n := 2^{P^{k+D}},$$

where $P := (n_0+1) \cdot m_0 \cdot \lceil C \cdot 22/10 \rceil$, and n_0, $m_0 \in \mathbb{N}$ satisfying $n_0 \leq \eta < n_0 + 1$ and $r_0 \cdot m_0 > 1$. Assume $\alpha_0 < \cdots < \alpha_n$ is a sequence of elements from S such that

$$\|\alpha_i\| \leq k + r_0 \cdot |i|$$

for any $i \leq n$. Then

$$\|\alpha_i\| \leq k + r_0 \cdot |n| = k + r_0(P^{k+D}+1).$$

Hence

$$\begin{aligned} n &< \frac{1}{1-\eta^{-1}} \cdot \frac{11}{10} \cdot C \cdot \frac{\eta^{k+r_0(P^{k+D}+1)}}{(k+r_0(P^{k+D}+1))^{3/2}} \\ &< \frac{n_0}{n_0-1} \cdot \frac{11}{10} \cdot C \cdot \frac{\eta^{k} \cdot (\eta^{r_0})^{P^{k+D}+1}}{r_0^{3/2} \cdot P^{(k+D)\cdot \frac{3}{2}}} \\ &< \frac{n_0}{n_0-1} \cdot \frac{22}{10} \cdot C \cdot \frac{\eta^{k} \cdot (2)^{P^{k+D}} \cdot m_0^{3/2}}{P^{(k+D)\cdot \frac{3}{2}}} \\ &= \frac{n_0}{n_0-1} \cdot \frac{22}{10} \cdot C \cdot \frac{m_0^{3/2}}{P^{(k+D)\cdot \frac{3}{2}}} \cdot \eta^{k} \cdot 2^{P^{k+D}} \\ &< \frac{n_0}{n_0-1} \cdot \frac{22}{10} \cdot C \cdot \frac{(m_0 \cdot (n_0+1))^{k}}{P^{3(k+D)/2}} \cdot 2^{P^{k+D}} \\ &< 2^{P^{k+D}} = n. \end{aligned}$$

Contradiction! □

Unprovability

This will be done in two steps. Let n_0 is a fixed natural number such that $n_0 > 1 + r_0$.

Lemma 6.4.2. $\mathrm{ACA}_0 + \Pi^1_2\text{-BI} \nvdash \mathrm{SWP}(S, \preceq, f_{n_0})$.

Proof. Let $\eta_i := \rho_i^{-1}$ and $\eta := \rho^{-1}$. Then $\eta_i \leq \eta_{i+1} \leq \eta$ and $\lim_{i\to\infty} \eta_i$. Since $n_0 > 1+r_0$ there is a primitively recursive real number $r' > r_0$ such that $n_0 > 1+r'$. Then choose d such that $r' > \frac{1}{\log \eta_d}$. By Theorem 6.3.3 there is a natural number E such that for all $i \geq E$

$$s_i^d \geq \frac{9}{10} \cdot C_d \cdot \eta_d^i \cdot i^{-3/2}.$$

Choose also a natural number $D > d+1$ such that for any $i \geq D$

- $\lfloor r'|i| \rfloor \geq E$;
- $\frac{9}{10} \cdot C_d \cdot 2^{\lfloor r'|i| \rfloor \cdot \log(\eta_d)} \cdot (\lfloor r'|i| \rfloor)^{-3/2} \geq 2^{|i|}$.

Let k be given. We may assume w.l.o.g. that

$$k_0 := \lfloor k/2 \rfloor \geq D \text{ and } k_0 + d + D + 4 \leq k.$$

Set

$$M_i := \{\alpha \in S^d \colon \|\alpha\| \leq \lfloor r'|i| \rfloor\}$$

and μ_i be the enumeration function of M_i with respect to a well-ordering on M_i expanding $\prec$. If F_{id} is the Skolem function for $\mathrm{SWP}(S, \preceq, id)$, then by Theorem 6.2.7 F_{id} is not provably recursive in $\mathrm{ACA}_0 + \Pi^1_2$-BI.

Put $n := F_{id}(k_0) - 1$ and $\beta_0, \ldots, \beta_{n-1}$ be a sequence from S such that

$$\sigma(\beta_0) > \cdots > \sigma(\beta_{n-1}) \quad \text{and} \quad \|\beta_i\| \leq k_0 + i$$

for any $i < n$. Then there are no $\ell < m$ such that $\beta_\ell \preceq \beta_m$. Note that all $\beta_i \prec f_{k_0}\bar{0}$ since $\|\beta_0\| \leq k_0$. Define a new sequence as follows.

$$\alpha_i := \begin{cases} f_{k_0+D-i}\bar{0} & \text{if } i \leq D, \\ f_1(f_{d+1}\beta_{|i|}\bar{0})\, \mu_i(2^{|i|} - i) & \text{if } D < i \leq n. \end{cases}$$

Then all α_i, $i \leq n$, are well-defined. Indeed for any $i > D$

$$\bar{\bar{M}}_i \geq s^d_{\lfloor r'|i| \rfloor} \geq \frac{9}{10} \cdot C_d \cdot \eta_d^{\lfloor r'|i| \rfloor} \cdot (\lfloor r'|i| \rfloor)^{-3/2} \geq 2^{|i|}$$

and

$$\begin{aligned}\|\alpha_i\| &\leq \max\{k_0 + D - i + 1, 2 + d + 2 + \|\beta_{|i|}\| + r'|i|\}\\ &\leq \max\{k_0 + D - i + 1, 4 + d + k_0 + (1 + r')|i|\}\\ &< k + n_0|i|.\end{aligned}$$

Using Lemma 6.2.5 we show now $\sigma(\alpha_\ell) > \sigma(\alpha_m)$ for all $\ell < m \leq n$. For simplification we shall make no difference between $\sigma(\alpha)$ and α.

(i) $\ell < m < D$: $\alpha_\ell = f_{k_0+D-\ell}\bar{0} > f_{k_0+D-m}\bar{0} = \alpha_m$.

(ii) $\ell < D \leq m$: $f_{k_0+D-\ell}\bar{0} \geq f_{k_0}\bar{0} > f_{d+1}\beta_{|m|}\bar{0}$.

(iii) $D \leq \ell < m \leq n$:

- $|\ell| < |m|$: Then since $\beta_{|\ell|} > \beta_{|m|}$

  $$f_{d+1}\beta_{|\ell|}\bar{0} > f_{d+1}\beta_{|m|}\bar{0} \quad \text{and} \quad f_{d+1}\beta_{|\ell|}\bar{0} > f_{d+1}\bar{0} > \mu_m(2^{|m|} - m).$$

  Hence $\alpha_\ell > \alpha_m$.

- $|\ell| = |m|$: Then it is obvious since $\beta_{|\ell|} = \beta_{|m|}$ and $\mu_\ell(2^{|\ell|} - \ell) > \mu_m(2^{|m|} - m)$.

The assertion follows now from the fact that $\lambda k\,.\,F_{id}(\lfloor k/2 \rfloor)$ is not provably recursive in PA. $\square$

Proof of Theorem 6.2.8: unprovability

Let $r > r_0$ and F_r be the Skolem function of $\mathrm{SWP}(S, \preceq, f_r)$. And let n_0 and η_i be defined as in Lemma 6.4.2. We choose a rational number r' and a natural number d such that $r > r' > \frac{1}{\log \eta_d}$. By Theorem 6.3.3 there is a natural number E so large that

$$s_i^d \geq \frac{9}{10} \cdot C_d \cdot \eta_d^i \cdot i^{-3/2}$$

for all $i \geq E$. Let $D > d + 1$ be so large that the following inequalities hold for any $i \geq D$:

- $\lfloor r'|i| \rfloor \geq E$;
- $\frac{9}{10} \cdot 2^{\lfloor r'|i| \rfloor \cdot \log(\eta_d)} \cdot C_d \cdot (\lfloor r'|i| \rfloor)^{-3/2} \geq 2^{|i|}$;
- $r'|i| + n_0 \cdot |i|_2 < r|i|$.

Assume that k is given. We may also assume that

$$k_0 := \lfloor k/2 \rfloor \geq D \quad \text{and} \quad k_0 + d + D + 4 \leq k.$$

Let $n := F_{n_0}(k_0) - 1$ and $\beta_0, \dots, \beta_{n-1}$ be a finite sequence from S such that $\sigma(\beta_0) > \cdots > \sigma(\beta_{n-1})$ and all $\|\beta_i\| \leq k_0 + n_0 \cdot |i|$. Set

$$M_i := \{\alpha \in S^d \colon \|\alpha\| \leq \lfloor r'|i| \rfloor\}$$

and μ_i be the enumeration function of M_i with respect to a well-ordering on M_i expanding $\prec$. Define a new sequence of length n by

$$\alpha_i = \begin{cases} f_{k_0+D-i}\bar{0} & \text{if } i \leq D, \\ f_1(f_{d+1}\beta_{|i|}\bar{0})\, \mu_i(2^{|i|} - i) & \text{if } D < i \leq n. \end{cases}$$

Then all α_i are well-defined. Indeed, we have for any $i > D$

$$\bar{\bar{M}}_i \geq S^d_{\lfloor r'|i| \rfloor} \geq \frac{9}{10} \cdot C_d \cdot \eta_d^{\lfloor r'|i| \rfloor} \cdot (\lfloor r'|i| \rfloor)^{-3/2} \geq 2^{|i|}$$

and

$$\begin{aligned} \|\alpha_i\| &\leq \max\{k_0 + D - i + 1, 2 + d + 2 + \|\beta_{|i|}\| + \lfloor r'|i| \rfloor\} \\ &\leq \max\{k_0 + D - i + 1, d + 4 + k_0 + n_0 \cdot |i|_2 + \lfloor r'|i| \rfloor\} \\ &< k + r|i|. \end{aligned}$$

Now assume that $\ell < m \leq n$. Then $\sigma(\alpha_\ell) > \sigma(\alpha_m)$ follows by an argument similar to that of Lemma 6.4.2.

This shows that $F_r(k)$ majorizes $F_{n_0}(\lfloor k/2 \rfloor)$ for large k. Thus F_r is not provably recursive in PA since F_{n_0} eventually dominates every provably recursive function of PA. Thus the claim is proved. □

Remark: Norm sensitivity of phase transitions

Simpson [49] introduced also a certain extension of Kruskal's theorem using finite trees with marks from k. In this case the generating function for the set $\mathcal{T}_k$ of all finite trees does not change so much:

$$\mathcal{T}_k(z) = k \cdot z \cdot \mathcal{M}(\mathcal{T}_k(z))$$

However, if we define a different norm of a tree with marks, then we could get a generating function for $\mathcal{T}_k$ which behaves differently: Let T be a finite tree with marks from k and define

$$\|T\| = \text{the number of nodes} + \text{the total sum of marks in } T.$$

Then $\mathcal{T}_k$ satisfies

$$\mathcal{T}_k(z) = \sum_{\ell=1}^{k} z^\ell \cdot \mathcal{M}(\mathcal{T}_k(z)).$$

This means that we could observe another phase transition and that phase transitions are sensitive to norm functions.

Bibliography

[1] Wilhelm Ackermann. Zur Widerspruchsfreiheit der reinen Zahlentheorie. *Math. Ann.*, 117:162–194, 1940.

[2] Toshiyasu Arai. Variations on a theme by Weiermann. *J. Symbolic Logic*, 63(3):897–925, 1998.

[3] Toshiyasu Arai. Epsilon substitution method for theories of jump hierarchies. *Arch. Math. Logic*, 41(2):123–153, 2002.

[4] Toshiyasu Arai. On the slowly well orderedness of ϵ_0. *MLQ Math. Log. Q.*, 48(1):125–130, 2002.

[5] Paul Bachmann. *Zahlentheorie. II. Teil. Die analytische Zahlentheorie.* Teubner, 1894.

[6] Lev D. Beklemishev. The Worm principle. *Logic Group Preprint Series*, 219, University of Utrecht, 2003. `http://preprints.phil.uu.nl/lgps/`.

[7] Lev D. Beklemishev. Provability algebras and proof-theoretic ordinals. I. *Ann. Pure Appl. Logic*, 128(1-3):103–123, 2004.

[8] Béla Bollobás, Christian Borgs, Jennifer T. Chayes, Jeong Han Kim, and Dabid B. Wilson. The scaling window of the 2-SAT transition. *Random Structures Algorithms*, 18(3):201–256, 2001.

[9] George Boolos. *The logic of provability.* Cambridge University Press, 1993.

[10] Wilfried Buchholz. A new system of proof-theoretic ordinal functions. *Ann. Pure Appl. Logic*, 32(3):195–207, 1986.

[11] Stanley N. Burris. *Number theoretic density and logical limit laws*, volume 86 of *Mathematical Surveys and Monographs.* American Mathematical Society, 2001.

[12] Dick H. J. de Jongh and Rohit Parikh. Well-partial orderings and hierarchies. *Nederl. Akad. Wetensch. Proc. Ser. A* **80**=*Indag. Math.*, 39(3):195–207, 1977.

[13] Paul Erdös, András Hajnal, Attila Máté, and Richard Rado. *Combinatorial set theory: partition relations for cardinals*, volume 106 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., 1984.

[14] Paul Erdös and Richard Rado. Combinatorial theorems on classifications of subsets of a given set. *Proc. London Math. Soc. (3)*, 2:417–439, 1952.

[15] Paul Erdös and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.

[16] Matt Fairtlough and Stanley S. Wainer. Hierarchies of provably recursive functions. In *Handbook of proof theory*, volume 137 of *Stud. Logic Found. Math.*, pages 149–207. North-Holland, 1998.

[17] Solomon Feferman. Systems of predicative analysis. *J. Symbolic Logic*, 29:1–30, 1964.

[18] Solomon Feferman. Systems of predicative analysis. II. Representations of ordinals. *J. Symbolic Logic*, 33:193–220, 1968.

[19] Solomon Feferman. Three conceptual problems that bug me. *Lecture text for 7-th Scandinavial Logic Symposium*, 1996. `ftp://math.stanford.edu/pub/papers/feferman/conceptualprobs.ps.gz`.

[20] Ehud Friedgut. Sharp thresholds of graph properties, and the k-sat problem. *J. Amer. Math. Soc.*, 12(4):1017–1054, 1999.

[21] Gerhard Gentzen. Die Widerspruchsfreiheit der reinen Zahlentheorie. *Math. Ann.*, 112:493–565, 1936.

[22] Gerhard Gentzen. *The collected papers of Gerhard Gentzen*. Edited by M. E. Szabo. Studies in Logic and the Foundations of Mathematics. North-Holland, 1969.

[23] Gerhard Gentzen. Der erste Widerspruchsfreiheitsbeweis für die klassische Zahlentheorie. Mit einem Einführung von Paul Bernays. *Arch. Math. Logik Grundlagenforsch.*, 16:97–118, 1974.

[24] J. Willard Gibbs. *Elementary principles in statistical mechanics: developed with especial reference to the rational foundation of thermodynamics*. Dover publications Inc., 1960.

[25] Jean-Yves Girard. *Proof theory and logical complexity*, volume 1 of *Studies in Proof Theory. Monographs*. Bibliopolis, 1987.

[26] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete mathematics*. Addison-Wesley, second edition, 1994.

[27] Frank Harary, Robert W. Robinson, and Allen J. Schwenk. Twenty-step algorithm for determining the asymptotic number of trees of various species. *J. Austral. Math. Soc. Ser. A*, 20(4):483–503, 1975.

[28] Graham Higman. Ordering by divisibility in abstract algebras. *Proc. London Math. Soc. (3)*, 2:326–336, 1952.

[29] Akihiro Kanamori and Kenneth McAloon. On Gödel incompleteness and finite combinatorics. *Ann. Pure Appl. Logic*, 33(1):23–41, 1987.

[30] Laurie Kirby and Jeff Paris. Accessible independence results for Peano arithmetic. *Bull. London Math. Soc.*, 14(4):285–293, 1982.

[31] Menachem Kojman and Saharon Shelah. Regressive Ramsey numbers are Ackermannian. *J. Combin. Theory Ser. A*, 86(1):177–181, 1999.

[32] Dénes König. Über eine Schlußweise aus dem Endlichen ins Unendliche. *Acta Szeged*, 3:121–130, 1927.

[33] Georg Kreisel. On the interpretation of non-finitist proofs. II. Interpretation of number theory. Applications. *J. Symbolic Logic*, 17:43–58, 1952.

[34] Georg Kreisel. Wie die Beweistheorie zu ihren Ordinalzahlen kam und kommt. *Jber. Deutsch. Math.-Verein.*, 78(4):177–223, 1976/77.

[35] Joseph B. Kruskal. Well-quasi-ordering, the Tree Theorem, and Vazsonyi's conjecture. *Trans. Amer. Math. Soc.*, 95:210–225, 1960.

[36] T. D. Lee and C. N. Yang. Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model. *Physical Rev. (2)*, 87:410–419, 1952.

[37] Martin Loebl and Jiří Matoušek. On undecidability of the weakened Kruskal theorem. In *Logic and combinatorics (Arcata, Calif., 1985)*, volume 65 of *Contemp. Math.*, pages 275–280. Amer. Math. Soc., 1987.

[38] Aleksei I. Markushevich. *Theory of functions of a complex variable. Vol. II*. Revised English edition translated and edited by Richard A. Silverman. Prentice-Hall Inc., 1965.

[39] Richard Otter. The number of trees. *Ann. of Math. (2)*, 49:583–599, 1948.

[40] Jeff Paris and Leo Harrington. A mathematical incompleteness in Peano arithmetic. In *J. Barwise, ed., Handbook of Mathematical Logic*, volume 90 of *Studies in Logic and the Foundations of Mathematics*, pages 1133–1142. North-Holland, 1977.

[41] Rózsa Péter. *Recursive functions.* Third revised edition. Translated from the German by István Földes. Academic Press, 1967.

[42] Frank P. Ramsey. On a problem of formal logic. *Proc. London Math. Soc.*, 30(2):264–286, 1929.

[43] Michael Rathjen and Andreas Weiermann. Proof-theoretic investigations on Kruskal's theorem. *Ann. Pure Appl. Logic*, 60(1):49–88, 1993.

[44] Diana Schmidt. *Well-Partial Orderings and Their Maximal order Types.* Habilitationsschrift, Heidelberg, 1979.

[45] Kurt Schütte. Predicative well-orderings. In *Formal Systems and Recursive Functions (Proc. Eighth Logic Colloq., Oxford, 1963)*, pages 280–303. North-Holland, 1965.

[46] Kurt Schütte. *Proof theory.* Springer, 1977.

[47] Kurt Schütte and Stephen G. Simpson. Ein in der reinen Zahlentheorie unbeweisbarer Satz über endliche Folgen von natürlichen Zahlen. *Arch. Math. Logik Grundlag.*, 25(1-2):75–89, 1985.

[48] Robert Sedgewick and Philippe Flajolet. *An introduction to the analysis of algorithms. Foreword by D. E. Knuth.* Addison-Wesley, 1996.

[49] Stephen G. Simpson. Nonprovability of certain combinatorial properties of finite trees. In *Harvey Friedman's research on the foundations of mathematics*, volume 117 of *Stud. Logic Found. Math.*, pages 87–117. North-Holland, 1985.

[50] Stephen G. Simpson. *Subsystems of second order arithmetic.* Perspectives in Mathematical Logic. Springer-Verlag, 1999.

[51] Rick L. Smith. The consistency strengths of some finite forms of the Higman and Kruskal theorems. In *Harvey Friedman's research on the foundations of mathematics*, volume 117 of *Stud. Logic Found. Math.*, pages 119–136. North-Holland, 1985.

[52] Richard P. Stanley. Generating functions. In *Studies in combinatorics*, volume 17 of *MAA Stud. Math.*, pages 100–141. Math. Assoc. America, 1978.

[53] Richard P. Stanley. *Enumerative combinatorics. Vol. 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 1997.

[54] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 1999.

[55] Stan S. Wainer. A classification of the ordinal recursive functions. *Arch. Math. Logik Grundlagenforsch.*, 13:136–153, 1970.

[56] Andreas Weiermann. How to characterize provably total functions by local predicativity. *J. Symbolic Logic*, 61(1):52–69, 1996.

[57] Andreas Weiermann. What makes a (pointwise) subrecursive hierarchy slow growing? In *Sets and proofs (Leeds, 1997)*, volume 258 of *London Math. Soc. Lecture Note Ser.*, pages 403–423. Cambridge Univ. Press, 1999.

[58] Andreas Weiermann. Classifying the phase transition for hydra games and Goodstein sequences. *preprint, to appear*, 2002.

[59] Andreas Weiermann. Anaytical methods in proof theory: Optimal friedman style independence results for PA in terms of binary trees and finite sequences. *prepring, to appear*, 2003.

[60] Andreas Weiermann. An application of graphical enumeration to PA. *J. Symbolic Logic*, 68(1):5–16, 2003.

[61] Andreas Weiermann. A classification of rapidly growing Ramsey functions. *Proc. Amer. Math. Soc.*, 132(2):553–561, 2004.

[62] Herbert S. Wilf. *Generatingfunctionology.* Academic Press Inc., second edition, 1994.

[63] Alan R. Woods. Coloring rules for finite trees, and probabilities of monadic second order sentences. *Random Structures Algorithms*, 10(4):453–485, 1997.

[64] C. N. Yang and T. D. Lee. Statistical theory of equations of state and phase transitions. I. Theory of condensation. *Physical Rev. (2)*, 87:404–409, 1952.

Index

List of symbols

Curriculum Vitae

Surname:	Lee
First name:	Gyesik
Date of birth:	April 15, 1970
Place of birth:	Kyunggi-Do, Korea (ROK)
Citizenship:	Korean
Mother:	Okja Park
Father:	Yeonhun Lee
Family status:	single
Education:	1976–82 Chungbuk elementary school in Pyoungtaek, Korea 1882–85 Chungbuk middle school in Pyoungtaek, Korea 1985–88 Yushin high school in Suwon, Korea 1988–92 Study of Mathematics at Seoul National University, Korea, Grade: Bachelor of science in Mathematics 1992–93 Study of Mathematics at University Dortmund 1993–96 Study of Mathematics and Logic at Westfälische Wilhelms-Universität Münster Grade: Diplom (Master) in Mathematics Since 2000 Study of Ph.D. in Mathematics in Münster at Institut für Mathematische Logik und Grundlagenforschung under supervision of Prof. Dr. A. Weiermann
Military service:	04.1997–05.99 Military service in Korea
Grants:	04.2000-03.2002 Grant of the Konrad-Adenauer-Stiftung Since 04.2002 Grant of the Deutsche Forschungsgemeinschaft