The art of ordinal analysis

Michael Rathjen

Abstract. Ordinal analysis of theories is a core area of proof theory whose origins can be traced back to Hilbert's programme – the aim of which was to lay to rest all worries about the foundations of mathematics once and for all by securing mathematics via an absolute proof of consistency. Ordinal-theoretic proof theory came into existence in 1936, springing forth from Gentzen's head in the course of his consistency proof of arithmetic. The central theme of ordinal analysis is the classification of theories by means of transfinite ordinals that measure their 'consistency strength' and 'computational power'. The so-called *proof-theoretic ordinal* of a theory also serves to characterize its provably recursive functions and can yield both conservation and combinatorial independence results.

This paper intends to survey the development of "ordinally informative" proof theory from the work of Gentzen up to more recent advances in determining the proof-theoretic ordinals of strong subsystems of second order arithmetic.

Mathematics Subject Classification (2000). Primary 03F15, 03F05, 03F35; Secondary 03F03, 03-03.

Keywords. Proof theory, ordinal analysis, ordinal representation systems, proof-theoretic strength.

1. Introduction

Ordinal analysis of theories is a core area of proof theory. The origins of proof theory can be traced back to the second problem on Hilbert's famous list of problems (presented at the Second International Congress in Paris on August 8, 1900), which called for a proof of consistency of the arithmetical axioms of the reals. Hilbert's work on axiomatic geometry marked the beginning of his live-long interest in the axiomatic method. For geometry, he solved the problem of consistency by furnishing arithmetical-analytical interpretations of the axioms, thereby reducing the question of consistency to the consistency of the axioms for real numbers. The consistency of the latter system of axioms is therefore the ultimate problem for the foundations of mathematics.

Which axioms for real numbers Hilbert had in mind in his problem was made precise only when he took up logic full scale in the 1920s and proposed a research programme with the aim of providing mathematics with a secure foundation. This was to be accomplished by first formalizing logic and mathematics in their entirety, and then showing that these formalizations are consistent, that is to say free of contra-

Proceedings of the International Congress
of Mathematicians, Madrid, Spain, 2006

dictions. Strong restrictions were placed on the methods to be applied in consistency proofs of axiom systems for mathematics: namely, these methods were to be completely *finitistic* in character. The proposal to obtain finitistic consistency proofs of axiom systems for mathematics came to be called *Hilbert's Programme*.

Hilbert's Programme is a reductive enterprise with the aim of showing that whenever a 'real' proposition can be proved by 'ideal' means, it can also be proved by 'real', finitistic means. However, Hilbert's so-called formalism was not intended to eliminate nonconstructive existence proofs in the practice of mathematics, but to vindicate them.

In the 1920s, Ackermann and von Neumann, in pursuit of Hilbert's Programme, were working on consistency proofs for arithmetical systems. Ackermann's 1924 dissertation gives a consistency proof for a second-order version of primitive recursive arithmetic which explicitly uses a finitistic version of transfinite induction up to the ordinal $\omega^{\omega^{\omega}}$. The employment of transfinite induction on ordinals in consistency proofs came explicitly to the fore in Gentzen's 1936 consistency proof for Peano arithmetic, **PA**. This proof led to the assignment of a *proof-theoretic ordinal* to a theory. This so-called *ordinal analysis* of theories allows one to classify theories by means of transfinite ordinals that measure their 'consistency strength' and 'computational power'.

The subject of this paper is the development of ordinal analysis from the work of Gentzen up to very recent advances in determining the proof-theoretic ordinals of strong subsystems of second order arithmetic.

1.1. Gentzen's result. The most important structure in mathematics is arguably the structure of the natural numbers $\mathfrak{N} = \big(\mathbb{N}; 0^{\mathfrak{N}}, 1^{\mathfrak{N}}, +^{\mathfrak{N}}, \times^{\mathfrak{N}}, E^{\mathfrak{N}}, <^{\mathfrak{N}}\big)$, where $0^{\mathfrak{N}}$ denotes zero, $1^{\mathfrak{N}}$ denotes the number one, $+^{\mathfrak{N}}$, $\times^{\mathfrak{N}}$, $E^{\mathfrak{N}}$ denote the successor, addition, multiplication, and exponentiation function, respectively, and $<^{\mathfrak{N}}$ stands for the less-than relation on the natural numbers. In particular, $E^{\mathfrak{N}}(n, m) = n^m$.

Many of the famous theorems and problems of mathematics such as Fermat's and Goldbach's conjecture, the Twin Prime conjecture, and Riemann's hypothesis can be formalized as sentences of the language of $\mathfrak{N}$ and thus concern questions about the structure $\mathfrak{N}$.

Definition 1.1. A theory designed with the intent of axiomatizing the structure $\mathfrak{N}$ is *Peano arithmetic*, **PA**. The language of **PA** has the predicate symbols $=$, $<$, the function symbols $+$, $\times$, E (for addition, multiplication, exponentiation) and the constant symbols 0 and 1. The *Axioms of* **PA** comprise the usual equations and laws for addition, multiplication, exponentiation, and the less-than relation. In addition, **PA** has the *Induction Scheme*

$$(\mathbf{IND}) \qquad \varphi(0) \wedge \forall x[\varphi(x) \rightarrow \varphi(x+1)] \rightarrow \forall x \varphi(x)$$

for all formulae φ of the language of **PA**.

Gentzen showed that transfinite induction up to the ordinal

$$\varepsilon_0 = \sup\{\omega, \omega^\omega, \omega^{\omega^\omega}, \dots\} = \text{least } \alpha.\ \omega^\alpha = \alpha$$

suffices to prove the consistency of **PA**. To appreciate Gentzen's result it is pivotal to note that he applied transfinite induction up to ε_0 solely to elementary computable predicates and besides that his proof used only finitistically justified means. Hence, a more precise rendering of Gentzen's result is

$$\mathbf{F} + \text{EC-TI}(\varepsilon_0) \vdash \mathsf{Con}(\mathbf{PA}); \tag{1}$$

here **F** signifies a theory that embodies only finitistically acceptable means, EC-TI(ε_0) stands for transfinite induction up to ε_0 for elementary computable predicates, and $\mathsf{Con}(\mathbf{PA})$ expresses the consistency of **PA**. Gentzen also showed that his result was the best possible in that **PA** proves transfinite induction up to α for arithmetic predicates for any $\alpha < \varepsilon_0$. The compelling picture conjured up by the above is that the non-finitist part of **PA** is encapsulated in EC-TI(ε_0) and therefore "measured" by ε_0, thereby tempting one to adopt the following definition of *proof-theoretic ordinal* of a theory T:

$$|T|_{\mathsf{Con}} = \text{least } \alpha.\ \mathbf{F} + \text{EC-TI}(\alpha) \vdash \mathsf{Con}(T). \tag{2}$$

In the above, many notions were left unexplained. We will now consider them one by one. The *elementary computable functions* are exactly the Kalmar *elementary functions,* i.e. the class of functions which contains the successor, projection, zero, addition, multiplication, and modified subtraction functions and is closed under composition and bounded sums and products. A predicate is elementary computable if its characteristic function is elementary computable.

According to an influential analysis of finitism due to W.W. Tait, finistic reasoning coincides with a system known as *primitive recursive arithmetic*. For the purposes of ordinal analysis, however, it suffices to identify **F** with an even more restricted theory known as *Elementary Recursive Arithmetic*, **ERA**. **ERA** is a weak subsystem of **PA** having the same defining axioms for $+$, $\times$, E, $<$ but with induction restricted to elementary computable predicates.

In order to formalize EC-TI(α) in the language of arithmetic we should first discuss ordinals and the representation of particular ordinals α as relations on $\mathbb{N}$.

Definition 1.2. A set A equipped with a total ordering $\prec$ (i.e. $\prec$ is transitive, irreflexive, and $\forall x, y \in A\,[x \prec y \vee x = y \vee y \prec x]$) is a *wellordering* if every non-empty subset X of A contains a $\prec$-least element, i.e. $(\exists u \in X)(\forall y \in X)[u \prec y \vee u = y]$.

An *ordinal* is a transitive set wellordered by the elementhood relation $\in$.

Fact 1.3. Every wellordering $(A, \prec)$ is order isomorphic to an ordinal $(\alpha, \in)$.

Ordinals are traditionally denoted by lower case Greek letters $\alpha, \beta, \gamma, \delta, \dots$ and the relation $\in$ on ordinals is notated simply by $<$. The operations of addition, multiplication, and exponentiation can be defined on all ordinals, however, addition and multiplication are in general not commutative.

We are interested in representing specific ordinals α as relations on $\mathbb{N}$. In essence Cantor [10] defined the first ordinal representation system in 1897. Natural ordinal representation systems are frequently derived from structures of the form

$$\mathfrak{A} = \langle \alpha, f_1, \ldots, f_n, <_\alpha \rangle \tag{3}$$

where α is an ordinal, $<_\alpha$ is the ordering of ordinals restricted to elements of α and the f_i are functions

$$f_i : \underbrace{\alpha \times \cdots \times \alpha}_{k_i \text{ times}} \longrightarrow \alpha$$

for some natural number k_i.

$$\mathbb{A} = \langle A, g_1, \ldots, g_n, \prec \rangle \tag{4}$$

is a *computable* (or *recursive*) *representation* of $\mathfrak{A}$ if the following conditions hold:

1. $A \subseteq \mathbb{N}$ and A is a computable set.
2. $\prec$ is a computable total ordering on A and the functions g_i are computable.
3. $\mathfrak{A} \cong \mathbb{A}$, i.e. the two structures are isomorphic.

Theorem 1.4 (Cantor, 1897). *For every ordinal $\beta > 0$ there exist unique ordinals $\beta_0 \geq \beta_1 \geq \cdots \geq \beta_n$ such that*

$$\beta = \omega^{\beta_0} + \cdots + \omega^{\beta_n}. \tag{5}$$

The representation of β in (5) is called the *Cantor normal form*. We shall write $\beta =_{CNF} \omega^{\beta_1} + \cdots + \omega^{\beta_n}$ to convey that $\beta_0 \geq \beta_1 \geq \cdots \geq \beta_k$.

ε_0 denotes the least ordinal $\alpha > 0$ such that $(\forall \beta < \alpha)\, \omega^\beta < \alpha$. ε_0 can also be described as the least ordinal α such that $\omega^\alpha = \alpha$.

Ordinals $\beta < \varepsilon_0$ have a Cantor normal form with exponents $\beta_i < \beta$ and these exponents have Cantor normal forms with yet again smaller exponents. As this process must terminate, ordinals $< \varepsilon_0$ can be coded by natural numbers. For instance a coding function

$$\ulcorner\, .\, \urcorner : \varepsilon_0 \longrightarrow \mathbb{N}$$

could be defined as follows:

$$\ulcorner \alpha \urcorner = \begin{cases} 0 & \text{if } \alpha = 0, \\ \langle \ulcorner \alpha_1 \urcorner, \ldots, \ulcorner \alpha_n \urcorner \rangle & \text{if } \alpha =_{CNF} \omega^{\alpha_1} + \cdots + \omega^{\alpha_n} \end{cases}$$

where $\langle k_1, \ldots, k_n \rangle := 2^{k_1+1} \ldots p_n^{k_n+1}$ with p_i being the ith prime number (or any other coding of tuples). Further define:

$$A_0 := \text{range of } \ulcorner . \urcorner, \quad \ulcorner \alpha \urcorner \prec \ulcorner \beta \urcorner \;:\Leftrightarrow\; \alpha < \beta$$

$$\ulcorner \alpha \urcorner \,\hat{+}\, \ulcorner \beta \urcorner := \ulcorner \alpha + \beta \urcorner, \quad \ulcorner \alpha \urcorner \,\hat{\cdot}\, \ulcorner \beta \urcorner := \ulcorner \alpha \cdot \beta \urcorner, \quad \hat{\omega}^{\ulcorner \alpha \urcorner} := \ulcorner \omega^\alpha \urcorner.$$

Then

$$\langle \varepsilon_0, +, \cdot, \delta \mapsto \omega^\delta, < \rangle \cong \langle A_0, \hat{+}, \hat{\cdot}, x \mapsto \hat{\omega}^x, \prec \rangle.$$

$A_0, \hat{+}, \hat{\cdot}, x \mapsto \hat{\omega}^x, \prec$ are computable (recursive), in point of fact, they are all elementary computable.

Finally, we can spell out the scheme EC-TI(ε_0) in the language of **PA**:

$$\forall x \; [\forall y \; (y \prec x \rightarrow P(y)) \rightarrow \; P(x)] \; \rightarrow \; \forall x \, P(x)$$

for all elementary computable predicates P.

1.2. Cut Elimination: Gentzen's Hauptsatz. In the consistency proof, Gentzen used his sequent calculus and employed the technique of *cut elimination*. As this is a tool of utmost importance in proof theory and ordinal analysis, a rough outline of the underlying ideas will be discussed next.

The most common logical calculi are *Hilbert-style* systems. They are specified by delineating a collection of schematic logical axioms and some inference rules. The choice of axioms and rules is more or less arbitrary, only subject to the desire to obtain a *complete* system (in the sense of Gödel's completeness theorem). In model theory it is usually enough to know that there is a complete calculus for first order logic as this already entails the compactness theorem.

There are, however, proof calculi without this arbitrariness of axioms and rules. The *natural deduction calculus* and the *sequent calculus* were both invented by *Gentzen*. Both calculi are pretty illustrations of the symmetries of logic. The sequent calculus since is a central tool in ordinal analysis and allows for generalizations to so-called infinitary logics. Gentzen's main theorem about the sequent calculus is the *Hauptsatz*, i.e. *the cut elimination theorem*.

A *sequent* is an expression $\Gamma \Rightarrow \Delta$ where Γ and Δ are finite sequences of formulae $A_1, \ldots, A_n$ and $B_1, \ldots, B_m$, respectively. We also allow for the possibility that Γ or Δ (or both) are empty. The empty sequence will be denoted by $\emptyset$. $\Gamma \Rightarrow \Delta$ is read, informally, as Γ yields Δ or, rather, the *conjunction* of the A_i yields the *disjunction* of the B_j. In particular, we have:

- If Γ is empty, the sequent asserts the disjunction of the B_j.
- If Δ is empty, it asserts the negation of the conjunction of the A_i.
- if Γ and Δ are both empty, it asserts the *impossible*, i.e. a *contradiction*.

We use upper case Greek letters $\Gamma, \Delta, \Lambda, \Theta, \Xi \ldots$ to range over finite sequences of formulae. $\Gamma \subseteq \Delta$ means that every formula of Γ is also a formula of Δ.

Next we list the axioms and rules of the sequent calculus.

- *Identity Axiom:*

$$A \Rightarrow A$$

  where A is any formula. In point of fact, one could limit this axiom to the case of atomic formulae A.

- *Cut Rule:*

$$\frac{\Gamma \Rightarrow \Delta, A \qquad A, \Lambda \Rightarrow \Theta}{\Gamma, \Lambda \Rightarrow \Delta, \Theta}\ \text{Cut}$$

  The formula A is called the *cut formula* of the inference.

- *Structural Rules:*

$$\frac{\Gamma \Rightarrow \Delta}{\Gamma' \Rightarrow \Delta'} \qquad \text{if } \Gamma \subseteq \Gamma',\ \Delta \subseteq \Delta'.$$

  A special case of the structural rule, known as *contraction*, occurs when the lower sequent has fewer occurrences of a formula than the upper sequent. For instance, $A, \Gamma \Rightarrow \Delta, B$ follows structurally from $A, A, \Gamma \Rightarrow \Delta, B, B$.

- *Rules for Logical Operations:*

Left

$$\frac{\Gamma \Rightarrow \Delta, A}{\neg A, \Gamma \Rightarrow \Delta}$$

$$\frac{\Gamma \Rightarrow \Delta, A \qquad B, \Lambda \Rightarrow \Theta}{A \to B, \Gamma, \Lambda \Rightarrow \Delta, \Theta}$$

$$\frac{A, \Gamma \Rightarrow \Delta}{A \wedge B, \Gamma \Rightarrow \Delta} \qquad \frac{B, \Gamma \Rightarrow \Delta}{A \wedge B, \Gamma \Rightarrow \Delta}$$

$$\frac{A, \Gamma \Rightarrow \Delta \qquad B, \Gamma \Rightarrow \Delta}{A \vee B, \Gamma \Rightarrow \Delta}$$

$$\frac{F(t), \Gamma \Rightarrow \Delta}{\forall x\, F(x), \Gamma \Rightarrow \Delta}\ \forall \text{L}$$

$$\frac{F(a), \Gamma \Rightarrow \Delta}{\exists x\, F(x), \Gamma \Rightarrow \Delta}\ \exists \text{L}$$

Right

$$\frac{B, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, \neg B}$$

$$\frac{A, \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \to B}$$

$$\frac{\Gamma \Rightarrow \Delta, A \qquad \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \wedge B}$$

$$\frac{\Gamma \Rightarrow \Delta, A}{\Gamma \Rightarrow \Delta, A \vee B} \qquad \frac{\Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \vee B}$$

$$\frac{\Gamma \Rightarrow \Delta, F(a)}{\Gamma \Rightarrow \Delta, \forall x\, F(x)}\ \forall \text{R}$$

$$\frac{\Gamma \Rightarrow \Delta, F(t)}{\Gamma \Rightarrow \Delta, \exists x\, F(x)}\ \exists \text{R}$$

In ∀L and ∃R, t is an arbitrary term. The variable a in ∀R and ∃L is an *eigenvariable* of the respective inference, i.e. a is not to occur in the *lower sequent*.

In the rules for logical operations, the formulae highlighted in the premisses are called the *minor formulae* of that inference, while the formula highlighted in the conclusion is the *principal formula* of that inference. The other formulae of an inference are called *side formulae*.

A *proof* (also known as *deduction* or *derivation*) $\mathcal{D}$ is a tree of sequents satisfying the following conditions:

- The topmost sequents of $\mathcal{D}$ are identity axioms.

- Every sequent in $\mathcal{D}$ except the lowest one is an upper sequent of an inference whose lower sequent is also in $\mathcal{D}$.

A sequent $\Gamma \Rightarrow \Delta$ is *deducible* if there is a proof having $\Gamma \Rightarrow \Delta$ as its the bottom sequent.

The Cut rule differs from the other rules in an important respect. With the rules for introduction of a connective on the left or the right, one sees that every formula that occurs above the line occurs below the line either directly, or as a subformula of a formula below the line, and that is also true for the structural rules. (Here $A(t)$ is counted as a subformula, in a slightly extended sense, of both $\exists x A(x)$ and $\forall x A(x)$.) But in the case of the Cut rule, the cut formula A vanishes. Gentzen showed that such "vanishing rules" can be eliminated.

Theorem 1.5 (Gentzen's Hauptsatz). If a sequent $\Gamma \Rightarrow \Delta$ is provable, then it is provable without use of the Cut Rule (called a *cut-free proof*).

The secret to Gentzen's Hauptsatz is the symmetry of left and right rules for the logical connectives. The proof of the cut elimination theorem is rather intricate as the process of removing cuts interferes with the structural rules. The possibility of contraction accounts for the high cost of eliminating cuts. Let $|\mathcal{D}|$ be the *height* of the deduction $\mathcal{D}$. Also, let $\mathrm{rank}(\mathcal{D})$ be *supremum* of the *lengths of cut formulae* occurring in $\mathcal{D}$. Turning $\mathcal{D}$ into a cut-free deduction of the same end sequent results, in the worst case, in a deduction of height $\mathcal{H}(\mathrm{rank}(\mathcal{D}), |\mathcal{D}|)$ where $\mathcal{H}(0, n) = n$ and $\mathcal{H}(k+1, n) = 4^{\mathcal{H}(k,n)}$, yielding hyper-exponential growth.

The *Hauptsatz* has an important corollary which explains its crucial role in obtaining consistency proofs.

Corollary 1.6 (The Subformula Property). *If a sequent* $\Gamma \Rightarrow \Delta$ *is provable, then it has a deduction all of whose formulae are subformulae of the formulae of* Γ *and* Δ.

Corollary 1.7. *A contradiction, i.e. the empty sequent* $\emptyset \Rightarrow \emptyset$*, is not deducible.*

Proof. According to the Hauptsatz, if the empty sequent were deducible it would have a deduction without cuts. In a cut-free deduction of the empty sequent only empty sequents can occur. But such a deduction does not exist. □

While mathematics is based on logic, it cannot be developed solely on the basis of *pure logic*. What is needed in addition are *axioms* that assert the *existence* of *mathematical objects* and their properties. Logic plus axioms gives rise to (formal) *theories* such as *Peano arithmetic* or the axioms of *Zermelo–Fraenkel set theory*. What happens when we try to apply the procedure of cut elimination to theories? Well, axioms are poisonous to this procedure. It breaks down because the symmetry of the sequent calculus is lost. In general, we cannot remove cuts from deductions in a theory T when the cut formula is an axiom of T. However, sometimes the axioms of a theory are of *bounded syntactic complexity*. Then the procedure applies partially in that one can remove all cuts that exceed the complexity of the axioms of T. This gives rise to *partial cut elimination*. It is a very important tool in proof theory. For example, it works very well if the axioms of a theory can be presented as *atomic intuitionistic sequents* (also called *Horn clauses*), yielding the completeness of *Robinson*s *resolution method*. Partial cut elimination also pays off in the case of *fragments* of **PA** and set theory with *restricted induction schemes*, be it induction on natural numbers or sets. This method can be used to extract bounds from proofs of Π^0_2 statements in such fragments.

Full arithmetic (i.e. **PA**), though, does not even allow for partial cut elimination since the induction axioms have unbounded complexity. However, one can remove the obstacle to cut elimination in a drastic way by going *infinite*. The so-called *ω-rule* consists of the two types of *infinitary inferences*:

$$\frac{\Gamma \Rightarrow \Delta, F(0);\ \Gamma \Rightarrow \Delta, F(1);\ \ldots\ ; \Gamma \Rightarrow \Delta, F(n);\ \ldots}{\Gamma \Rightarrow \Delta, \forall x\, F(x)}\ \omega\mathrm{R}$$

$$\frac{F(0), \Gamma \Rightarrow \Delta;\ F(1), \Gamma \Rightarrow \Delta;\ \ldots\ ; F(n), \Gamma \Rightarrow \Delta;\ \ldots}{\exists x\, F(x), \Gamma \Rightarrow \Delta}\ \omega\mathrm{L}$$

The price to pay will be that deductions become infinite objects, i.e. *infinite well-founded trees*.

The sequent-style version of Peano arithmetic with the *ω-rule* will be termed $\mathbf{PA}_\omega$. $\mathbf{PA}_\omega$ has no use for free variables. Thus free variables are discarded and all *terms* will be closed. All formulae of this system are therefore closed, too. The *numerals* are the terms $\bar{n}$, where $\bar{0} = 0$ and $\overline{n+1} = S\bar{n}$. We shall identify $\bar{n}$ with the natural number n. All terms t of $\mathbf{PA}_\omega$ evaluate to a numeral $\bar{n}$.

$\mathbf{PA}_\omega$ has all the inference rules of the sequent calculus except for $\forall R$ and $\exists L$. In their stead, $\mathbf{PA}_\omega$ has the ωR and ωL inferences. The *Axioms* of $\mathbf{PA}_\omega$ are the following: (i) $\emptyset \Rightarrow A$ if A is a *true* atomic sentence; (ii) $B \Rightarrow \emptyset$ if B is a *false* atomic sentence; (iii) $F(s_1, \ldots, s_n) \Rightarrow F(t_1, \ldots, t_n)$ if $F(s_1, \ldots, s_n)$ is an atomic sentence and s_i and t_i evaluate to the same numeral.

With the aid of the ω-rule, each instance of the induction scheme becomes logically deducible, albeit the price to pay will be that the proof tree becomes infinite. To describe the cost of cut elimination for $\mathbf{PA}_\omega$, we introduce the measures of *height* and *cut rank* of a $\mathbf{PA}_\omega$ deduction $\mathcal{D}$. We will notate this by

$$\mathcal{D} \mathrel{\Big|\!\frac{\alpha}{k}} \Gamma \Rightarrow \Delta\,.$$

The above relation is defined inductively following the buildup of the deduction $\mathcal{D}$. For the *cut rank* we need the definition of the *length*, $|A|$ of a formula: $|A| = 0$ if A is atomic; $|\neg A_0| = |A_0| + 1$; $|A_0 \Box A_1| = \max(|A_0, A_1|) + 1$ where $\Box = \wedge, \vee, \rightarrow$; $|\exists x\, F(x)| = |\forall x\, F(x)| = |F(0)| + 1$.

Now suppose the last inference of $\mathcal{D}$ is of the form

$$\frac{\begin{matrix}\mathcal{D}_0 \\ \Gamma_0 \Rightarrow \Delta_0\end{matrix}\quad \cdots \quad \begin{matrix}\mathcal{D}_n \\ \Gamma_n \Rightarrow \Delta_n\end{matrix}\quad \cdots \quad n < \tau}{\Gamma \Rightarrow \Delta}\ I$$

where $\tau = 1, 2, \omega$ and the $\mathcal{D}_n$ are the immediate subdeductions of $\mathcal{D}$. If

$$\mathcal{D}_n \left|\frac{\alpha_n}{k}\right. \Gamma_n \Rightarrow \Delta_n$$

and $\alpha_n < \alpha$ for all $n < \tau$ then

$$\mathcal{D} \left|\frac{\alpha}{k}\right. \Gamma \Rightarrow \Delta$$

providing that in the case of I being a *cut* with cut formula A we also have $|A| < k$. We will write $\mathbf{PA}_\omega \left|\frac{\alpha}{k}\right. \Gamma \Rightarrow \Delta$ to convey that there exists a $\mathbf{PA}_\omega$-deduction $\mathcal{D} \left|\frac{\alpha}{k}\right. \Gamma \Rightarrow \Delta$. The ordinal analysis of **PA** proceeds by first unfolding any **PA**-deduction into a $\mathbf{PA}_\omega$-deduction:

$$\text{If } \mathbf{PA} \vdash \Gamma \Rightarrow \Delta, \text{ then } \mathbf{PA}_\omega \left|\frac{\omega+m}{k}\right. \Gamma \Rightarrow \Delta \tag{6}$$

for some $m, k < \omega$. The next step is to get rid of the cuts. It turns out that the cost of lowering the cut rank from $k + 1$ to k is an exponential with base ω.

Theorem 1.8 (Cut Elimination for $\mathbf{PA}_\omega$).

$$\textit{If } \mathbf{PA}_\omega \left|\frac{\alpha}{k+1}\right. \Gamma \Rightarrow \Delta\textit{, then } \mathbf{PA}_\omega \left|\frac{\omega^\alpha}{k}\right. \Gamma \Rightarrow \Delta.$$

As a result, if $\mathbf{PA}_\omega \left|\frac{\alpha}{n}\right. \Gamma \Rightarrow \Delta$, we may apply the previous theorem n times to arrive at a cut-free deduction $\mathbf{PA}_\omega \left|\frac{\rho}{0}\right. \Gamma \Rightarrow \Delta$ with $\rho = \omega^{\omega^{\cdot^{\cdot^{\cdot^{\omega^\alpha}}}}}$, where the stack has height n. Combining this with the result from (6), it follows that every sequent $\Gamma \Rightarrow \Delta$ deducible in **PA** has a cut-free deduction in $\mathbf{PA}_\omega$ of length $< \varepsilon_0$. Ruminating on the details of how this result was achieved yields a consistency proof for **PA** from transfinite induction up to ε_0 for elementary decidable predicates on the basis of finitistic reasoning (as described in (1)).

Deductions in $\mathbf{PA}_\omega$ being well-founded infinite trees, they have a natural associated ordinal length, namely: the height of the tree as an ordinal. Thus the passage from finite deductions in **PA** to infinite cut-free deductions in $\mathbf{PA}_\omega$ provides an explanation of how the ordinal ε_0 is connected with **PA**.

Gentzen, however, did not consider infinite proof trees. The infinitary version of **PA** with the ω-rule was introduced by Schütte in [35]. Incidentally, the ω-rule had already been proposed by Hilbert [18]. Gentzen worked with finite deductions in the sequent calculus version of **PA**, devising an ingenious method of assigning ordinals to purported derivations of the empty sequent (inconsistency). It turns out in recent work by Buchholz [9] that in fact there is a much closer intrinsic connection between the way Gentzen assigned ordinals to deductions in **PA** and the way that ordinals are assigned to infinite deductions in $\mathbf{PA}_\omega$.

In the 1950s infinitary proof theory flourished in the hands of Schütte. He extended his approach to **PA** to systems of ramified analysis and brought this technique to perfection in his monograph "Beweistheorie" [36]. The ordinal representation systems necessary for Schütte's work will be reviewed in the next subsection.

1.3. A brief history of ordinal representation systems: 1904–1950. Ordinals assigned as lengths to deductions to keep track of the cost of operations such as cut elimination render ordinal analyses of theories particularly transparent. In the case of **PA**, Gentzen could rely on Cantor's normal form for a supply of ordinal representations. For stronger theories, though, segments larger than ε_0 have to be employed. Ordinal representation systems utilized by proof theorists in the 1960s arose in a purely set-theoretic context. This subsection will present some of the underlying ideas as progress in ordinal-theoretic proof theory also hinges on the development of sufficiently strong and transparent ordinal representation systems.

In 1904, Hardy [17] wanted to "construct" a subset of $\mathbb{R}$ of size $\aleph_1$. His method was to represent countable ordinals via increasing sequence of natural numbers and then to correlate a decimal expansion with each such sequence. Hardy used two processes on sequences: (i) Removing the first element to represent the successor; (ii) Diagonalizing at limits. E.g., if the sequence $1, 2, 3, \ldots$ represents the ordinal 1, then $2, 3, 4, \ldots$ represents the ordinal 2 and $3, 4, 5, \ldots$ represents the ordinal 3 etc., while the 'diagonal' $1, 3, 5, \ldots$ provides a representation of ω. In general, if $\lambda = \lim_{n\in\mathbb{N}} \lambda_n$ is a limit ordinal with $b_{n1}, b_{n2}, b_{n3}, \ldots$ representing $\lambda_n < \lambda$, then $b_{11}, b_{22}, b_{33}, \ldots$ represents λ. This representation, however, depends on the sequence chosen with limit λ. A sequence $(\lambda_n)_{n\in\mathbb{N}}$ with $\lambda_n < \lambda$ and $\lim_{n\in\mathbb{N}} \lambda_n = \lambda$ is called a *fundamental sequence* for λ. Hardy's two operations give explicit representations for all ordinals $< \omega^2$.

Veblen [44] extended the initial segment of the countable for which fundamental sequences can be given effectively. The new tools he devised were the operations of *derivation* and *transfinite iteration* applied to *continuous increasing functions* on ordinals.

Definition 1.9. Let *ON* be the class of ordinals. A (class) function $f : ON \to ON$ is said to be *increasing* if $\alpha < \beta$ implies $f(\alpha) < f(\beta)$ and *continuous* (in the order topology on *ON*) if

$$f(\lim_{\xi<\lambda} \alpha_\xi) = \lim_{\xi<\lambda} f(\alpha_\xi)$$

holds for every limit ordinal λ and increasing sequence $(\alpha_\xi)_{\xi<\lambda}$. f is called *normal* if it is increasing and continuous.

The function $\beta \mapsto \omega+\beta$ is normal while $\beta \mapsto \beta+\omega$ is not continuous at ω since $\lim_{\xi<\omega}(\xi + \omega) = \omega$ but $(\lim_{\xi<\omega} \xi) + \omega = \omega + \omega$.

Definition 1.10. The *derivative* f' of a function $f : ON \to ON$ is the function which enumerates in increasing order the solutions of the equation $f(\alpha) = \alpha$, also called the *fixed points* of f.

If f is a normal function, $\{\alpha : f(\alpha) = \alpha\}$ is a proper class and f' will be a normal function, too.

Definition 1.11. Now, given a normal function $f : ON \to ON$, define a hierarchy of normal functions as follows:

$$f_0 = f, \quad f_{\alpha+1} = f'_\alpha,$$
$$f_\lambda(\xi) = \xi^{th} \text{ element of } \bigcap_{\alpha<\lambda} (\text{Range of } f_\alpha) \text{ for } \lambda \text{ a limit ordinal.}$$

In this way, from the normal function f we get a two-place function, $\varphi_f(\alpha, \beta) := f_\alpha(\beta)$. Veblen then discusses the hierarchy when $f = \ell$, where $\ell(\alpha) = 1 + \alpha$.

The least ordinal $\gamma > 0$ closed under φ_ℓ, i.e. the least ordinal > 0 satisfying $(\forall \alpha, \beta < \gamma)\ \varphi_\ell(\alpha, \beta) < \gamma$ is the famous ordinal Γ_0 which Feferman [13] and Schütte [37], [38] determined to be the least ordinal 'unreachable' by *predicative means*.

Veblen extended this idea first to arbitrary finite numbers of arguments, but then also to transfinite numbers of arguments, with the proviso that in, for example $\Phi_f(\alpha_0, \alpha_1, \ldots, \alpha_\eta)$, only a finite number of the arguments α_ν may be non-zero. Finally, Veblen singled out the ordinal $E(0)$, where $E(0)$ is the least ordinal $\delta > 0$ which cannot be named in terms of functions $\Phi_\ell(\alpha_0, \alpha_1, \ldots, \alpha_\eta)$ with $\eta < \delta$, and each $\alpha_\gamma < \delta$.

Though the "great Veblen number" (as $E(0)$ is sometimes called) is quite an impressive ordinal it does not furnish an ordinal representation sufficient for the task of analyzing a theory as strong as Π^1_1 comprehension. Of course, it is possible to go beyond $E(0)$ and initiate a new hierarchy based on the function $\xi \mapsto E(\xi)$ or even consider hierarchies utilizing finite type functionals over the ordinals. Still all these further steps amount to rather mundane progress over Veblen's methods. In 1950 Bachmann [3] presented a new kind of operation on ordinals which dwarfs all hierarchies obtained by iterating Veblen's methods. Bachmann builds on Veblen's work but his novel idea was the systematic use of *uncountable ordinals* to keep track of the functions defined by diagonalization. Let Ω be the first uncountable ordinal. Bachmann defines a set of ordinals $\mathfrak{B}$ closed under successor such that with each limit $\lambda \in \mathfrak{B}$ is associated an increasing sequence $\langle \lambda[\xi] : \xi < \tau_\lambda \rangle$ of ordinals $\lambda[\xi] \in \mathfrak{B}$ of length $\tau_\lambda \in \mathfrak{B}$ and $\lim_{\xi<\tau_\lambda} \lambda[\xi] = \lambda$. A hierarchy of functions $(\varphi^{\mathfrak{B}}_\alpha)_{\alpha\in\mathfrak{B}}$ is then obtained as follows:

$$\begin{aligned}
&\varphi^{\mathfrak{B}}_0(\beta) = 1 + \beta, \qquad \varphi^{\mathfrak{B}}_{\alpha+1} = \left(\varphi^{\mathfrak{B}}_\alpha\right)', \\
&\varphi^{\mathfrak{B}}_\lambda \text{ enumerates } \bigcap_{\xi<\tau_\lambda} (\text{Range of } \varphi^{\mathfrak{B}}_{\lambda[\xi]}) \qquad \text{if } \lambda \text{ is a limit with } \tau_\lambda < \Omega, \\
&\varphi^{\mathfrak{B}}_\lambda \text{ enumerates } \{\beta < \Omega : \varphi^{\mathfrak{B}}_{\lambda[\beta]}(0) = \beta\} \qquad \text{if } \lambda \text{ is a limit with } \tau_\lambda = \Omega.
\end{aligned} \tag{7}$$

After the work of Bachmann, the story of ordinal representations becomes very complicated. Significant papers (by Isles, Bridge, Pfeiffer, Schütte, Gerber to mention a few) involve quite horrendous computations to keep track of the fundamental sequences. Also Bachmann's approach was combined with uses of higher type functionals by Aczel and Weyhrauch. Feferman proposed an entirely different method for generating a Bachmann-type hierarchy of normal functions which does not involve fundamental sequences. Buchholz further simplified the systems and proved their recursivity. For details we recommend the preface to [7].

2. Ordinal analyses of systems of second order arithmetic and set theory

Ordinal analysis is concerned with theories serving as frameworks for formalising significant parts of mathematics. It is known that virtually all of ordinary mathematics can be formalized in Zermelo–Fraenkel set theory with the axiom of choice, **ZFC**. Hilbert and Bernays [19] showed that large chunks of mathematics can already be formalized in second order arithmetic. Owing to these observations, proof theory has been focusing on set theories and subsystems of second order arithmetic. Further scrutiny revealed that a small fragment is sufficient. Under the rubric of *Reverse Mathematics* a research programme has been initiated by Harvey Friedman some thirty years ago. The idea is to ask whether, given a theorem, one can prove its equivalence to some axiomatic system, with the aim of determining what proof-theoretical resources are necessary for the theorems of mathematics. More precisely, the objective of reverse mathematics is to investigate the role of set existence axioms in ordinary mathematics. The main question can be stated as follows:

> *Given a specific theorem τ of ordinary mathematics, which set existence axioms are needed in order to prove τ?*

Central to the above is the reference to what is called 'ordinary mathematics'. This concept, of course, doesn't have a precise definition. Roughly speaking, by ordinary mathematics we mean main-stream, non-set-theoretic mathematics, i.e. the core areas of mathematics which make no essential use of the concepts and methods of set theory and do not essentially depend on the theory of uncountable cardinal numbers.

2.1. Subsystems of second order arithmetic. The framework chosen for studying set existence in reverse mathematics, though, is second order arithmetic rather than set theory. Second order arithmetic, $\mathbf{Z_2}$, is a two-sorted formal system with one sort of variables $x, y, z, \ldots$ ranging over natural numbers and the other sort $X, Y, Z, \ldots$ ranging over sets of natural numbers. The language $\mathcal{L}_2$ of second-order arithmetic also contains the symbols of **PA**, and in addition has a binary relation symbol $\in$ for elementhood. Formulae are built from the prime formulae $s = t$, $s < t$, and $s \in X$ (where s, t are numerical terms, i.e. terms of **PA**) by closing off under the connectives $\wedge, \vee, \rightarrow, \neg$, numerical quantifiers $\forall x$, $\exists x$, and set quantifiers $\forall X$, $\exists X$.

The basic arithmetical axioms in all theories of second-order arithmetic are the defining axioms for $0, 1, +, \times, E, <$ (as for **PA**) and the *induction axiom*

$$\forall X(0 \in X \wedge \forall x(x \in X \rightarrow x + 1 \in X) \rightarrow \forall x(x \in X)).$$

We consider the axiom schema of $\mathcal{C}$-*comprehension* for formula classes $\mathcal{C}$ which is given by

$$\mathcal{C}\text{-}\mathbf{CA} \quad \exists X \forall u (u \in X \leftrightarrow F(u))$$

for all formulae $F \in \mathcal{C}$ in which X does not occur. Natural formula classes are the *arithmetical formulae*, consisting of all formulae without second order quantifiers $\forall X$ and $\exists X$, and the Π^1_n-*formulae*, where a Π^1_n-formula is a formula of the form $\forall X_1 \ldots QX_n\, A(X_1, \ldots, X_n)$ with $\forall X_1 \ldots QX_n$ being a string of n alternating set quantifiers, commencing with a universal one, followed by an arithmetical formula $A(X_1, \ldots, X_n)$.

For each axiom scheme **Ax** we denote by $(\mathbf{Ax})_0$ the theory consisting of the basic arithmetical axioms plus the scheme **Ax**. By contrast, $(\mathbf{Ax})$ stands for the theory $(\mathbf{Ax})_0$ augmented by the scheme of induction for all $\mathcal{L}_2$-formulae.

An example for these notations is the theory $(\Pi^1_1\text{-}\mathbf{CA})_0$ which has the comprehension schema for Π^1_1-formulae.

In **PA** one can define an elementary injective pairing function on numbers, e.g $(n, m) := 2^n \times 3^m$. With the help of this function an infinite sequence of sets of natural numbers can be coded as a single set of natural numbers. The n^{th} section of set of natural numbers U is defined by $U_n := \{m : (n, m) \in U\}$. Using this coding, we can formulate the axiom of choice for formulae F in $\mathcal{C}$ by

$$\mathcal{C}\text{-}\mathbf{AC} \quad \forall x \exists Y F(x, Y) \rightarrow \exists Y \forall x F(x, Y_x).$$

For many mathematical theorems τ, there is a weakest natural subsystem $S(\tau)$ of $\mathbf{Z_2}$ such that $S(\tau)$ proves τ. Very often, if a theorem of ordinary mathematics is proved from the weakest possible set existence axioms, the statement of that theorem will turn out to be provably equivalent to those axioms over a still weaker base theory. This theme is referred to as *Reverse Mathematics*. Moreover, it has turned out that $S(\tau)$ often belongs to a small list of specific subsystems of $\mathbf{Z_2}$ dubbed $\mathbf{RCA}_0$, $\mathbf{WKL_0}$, $\mathbf{ACA}_0$, $\mathbf{ATR}_0$ and $(\Pi^1_1\text{-}\mathbf{CA})_0$, respectively. The systems are enumerated in increasing strength. The main set existence axioms of $\mathbf{RCA}_0$, $\mathbf{WKL_0}$, $\mathbf{ACA}_0$, $\mathbf{ATR}_0$, and $(\Pi^1_1\text{-}\mathbf{CA})_0$ are recursive comprehension, weak König's lemma, arithmetical comprehension, arithmetical transfinite recursion, and Π^1_1-comprehension, respectively. For exact definitions of all these systems and their role in reverse mathematics see [40]. The proof-theoretic strength of $\mathbf{RCA}_0$ is weaker than that of **PA** while $\mathbf{ACA}_0$ has the same strength as **PA**. Let $|T| = |T|_{\mathsf{Con}}$. To get a sense of scale, the strengths of the first four theories are best expressed via their proof-theoretic ordinals: $|\mathbf{RCA}_0| = |\mathbf{WKL_0}| = \omega^\omega$, $|\mathbf{ACA}_0| = \varepsilon_0$, $|\mathbf{ATR}_0| = \Gamma_0$. $|(\Pi^1_1\text{-}\mathbf{CA})_0|$, however, eludes expression in the ordinal representations introduced so far. Π^1_1-$\mathbf{CA}$ involves a so-called *impredicative definition*. An impredicative definition of an object refers to a presumed totality of which the object being defined is itself to be a member. For example, to define a set of natural numbers X as $X = \{n \in \mathbb{N} : \forall Y \subseteq \mathbb{N}\; F(n, Y)\}$ is impredicative since it involves the quantified variable 'Y' ranging over arbitrary subsets of the natural numbers $\mathbb{N}$, of which the set X being defined is one member.

Determining whether $\forall Y \subseteq \mathbb{N}\ F(n, Y)$ holds involves an apparent circle since we shall have to know in particular whether $F(n, X)$ holds – but that cannot be settled until X itself is determined. Impredicative set definitions permeate the fabric of Zermelo–Fraenkel set theory in the guise of the separation and replacement axioms as well as the powerset axiom.

A major breakthrough was made by Takeuti in 1967, who for the first time obtained an ordinal analysis of an impredicative theory. In [41] he gave an ordinal analysis of $(\Pi^1_1$-**CA**), extended in 1973 to $(\Pi^1_1$-**AC**) in [43] jointly with Yasugi. For this Takeuti returned to Gentzen's method of assigning ordinals (ordinal diagrams, to be precise) to purported derivations of the empty sequent (inconsistency).

The next wave of results, which concerned theories of iterated inductive definitions, were obtained by Buchholz, Pohlers, and Sieg in the late 1970s (see [7]). Takeuti's methods of reducing derivations of the empty sequent ("the inconsistency") were extremely difficult to follow, and therefore a more perspicuous treatment was to be hoped for. Since the use of the infinitary ω-rule had greatly facilitated the ordinal analysis of number theory, new infinitary rules were sought. In 1977 (see [5]) Buchholz introduced such rules, dubbed Ω-rules to stress the analogy. They led to a proof-theoretic treatment of a wide variety of systems, as exemplified in the monograph [8] by Buchholz and Schütte. Yet simpler infinitary rules were put forward a few years later by Pohlers, leading to the *method of local predicativity*, which proved to be a very versatile tool (see [23]).

2.2. Set theories. With the work of Jäger and Pohlers (see [20], [21]) the forum of ordinal analysis then switched from the realm of second-order arithmetic to set theory, shaping what is now called *admissible proof theory*, after the models of *Kripke–Platek set theory*, **KP**. Their work culminated in the analysis of the system Π^1_1-**AC** plus an induction principle called *Bar Induction* **BI** which is a scheme asserting that transfinite induction along well-founded relations holds for arbitrary formulae (see [21]).

By and large, ordinal analyses for set theories are more uniform and transparent than for subsystems of $\mathbf{Z_2}$. The axiom systems for set theories considered in this paper are formulated in the usual language of set theory (called $\mathcal{L}_\in$ hereafter) containing $\in$ as the only non-logical symbol besides $=$. Formulae are built from prime formulae $a \in b$ and $a = b$ by use of propositional connectives and quantifiers $\forall x, \exists x$. Quantifiers of the forms $\forall x \in a$, $\exists x \in a$ are called *bounded*. *Bounded* or Δ_0-*formulae* are the formulae wherein all quantifiers are bounded; Σ_1-formulae are those of the form $\exists x \varphi(x)$ where $\varphi(a)$ is a Δ_0-formula. For $n > 0$, Π_n-formulae (Σ_n-formulae) are the formulae with a prefix of n alternating unbounded quantifiers starting with a universal (existential) one followed by a Δ_0-formula. The class of Σ-formulae is the smallest class of formulae containing the Δ_0-formulae which is closed under $\wedge$, $\vee$, bounded quantification and unbounded existential quantification.

One of the set theories which is amenable to ordinal analysis is Kripke–Platek set theory, **KP**. Its standard models are called *admissible sets*. One of the reasons that this is an important theory is that a great deal of set theory requires only the axioms of **KP**. An even more important reason is that admissible sets have been a major source of interaction between model theory, recursion theory and set theory (cf. [4]). **KP** arises from **ZF** by completely omitting the power set axiom and restricting separation and collection to bounded formulae. These alterations are suggested by the informal notion of 'predicative'. To be more precise, the axioms of **KP** consist of *Extensionality, Pair, Union, Infinity, Bounded Separation*

$$\exists x\, \forall u\, [u \in x \leftrightarrow (u \in a \;\wedge\; F(u))]$$

for all bounded formulae $F(u)$, *Bounded Collection*

$$\forall x \in a\, \exists y\, G(x, y) \;\rightarrow\; \exists z\, \forall x \in a\, \exists y \in z\, G(x, y)$$

for all bounded formulae $G(x, y)$, and *Set Induction*

$$\forall x\; [(\forall y \in x\; H(y)) \rightarrow H(x)] \;\rightarrow\; \forall x\; H(x)$$

for all formulae $H(x)$.

A transitive set A such that $(A, \in)$ is a model of **KP** is called an *admissible set*. Of particular interest are the models of **KP** formed by segments of Gödel's *constructible hierarchy* **L**. The constructible hierarchy is obtained by iterating the definable powerset operation through the ordinals

$$\begin{aligned} \mathbf{L}_0 &= \emptyset, \\ \mathbf{L}_\lambda &= \bigcup\{\mathbf{L}_\beta : \beta < \lambda\}\ \lambda \text{ limit} \\ \mathbf{L}_{\beta+1} &= \big\{X :\ X \subseteq \mathbf{L}_\beta;\ X \text{ definable over } \langle \mathbf{L}_\beta, \in \rangle\big\}. \end{aligned}$$

So any element of **L** of level α is definable from elements of **L** with levels $< \alpha$ and the parameter $\mathbf{L}_\alpha$. An ordinal α is *admissible* if the structure $(\mathbf{L}_\alpha, \in)$ is a model of **KP**.

Formulae of $\mathcal{L}_2$ can be easily translated into the language of set theory. Some of the subtheories of $\mathbf{Z_2}$ considered above have set-theoretic counterparts, characterized by extensions of **KP**. **KPi** is an extension of **KP** via the axiom

$$(\mathit{Lim}) \quad \forall x \exists y [x \in y \;\wedge\; y \textit{ is an admissible set}].$$

KPl denotes the system **KPi** without Bounded Collection. It turns out that (Π^1_1-**AC**)+ **BI** proves the same $\mathcal{L}_2$-formulae as **KPi**, while (Π^1_1-**CA**) proves the same $\mathcal{L}_2$-formulae as **KPl**.

2.3. Sketches of an ordinal analysis of KP. Serving as a miniature example of an ordinal analysis of an impredicative system, the ordinal analysis of **KP** (see [20], [6]) we will sketched in broad strokes. Bachmann's system can be recast without fundamental sequences as follows: Let Ω be a "big" ordinal, e.g. $\Omega = \aleph_1$. By recursion on α we define sets $C^{\Omega}(\alpha,\beta)$ and the ordinal $\psi_{\Omega}(\alpha)$ as follows:

$$C^{\Omega}(\alpha,\beta)=\begin{cases}\text{closure of } \beta\cup\{0,\Omega\} \text{ under:}\\ +,\ (\xi\mapsto\omega^{\xi})\\ (\xi\longmapsto\psi_{\Omega}(\xi))_{\xi<\alpha}\end{cases} \tag{8}$$

$$\psi_{\Omega}(\alpha)\simeq\min\{\rho<\Omega:\ C^{\Omega}(\alpha,\rho)\cap\Omega=\rho\}. \tag{9}$$

It can be shown that $\psi_{\Omega}(\alpha)$ is always defined and that $\psi_{\Omega}(\alpha)<\Omega$. Moreover, $[\psi_{\Omega}(\alpha),\Omega)\ \cap\ C^{\Omega}(\alpha,\psi_{\Omega}(\alpha))\ =\ \emptyset$; thus the order-type of the ordinals below Ω which belong to the set $C^{\Omega}(\alpha,\psi_{\Omega}(\alpha))$ is $\psi_{\Omega}(\alpha)$. $\psi_{\Omega}(\alpha)$ is also a countable ordinal. In more pictorial terms, $\psi_{\Omega}(\alpha)$ is the α^{th} *collapse* of Ω.

Let $\varepsilon_{\Omega+1}$ be the least ordinal $\alpha>\Omega$ such that $\omega^{\alpha}=\alpha$. The set of ordinals $C^{\Omega}(\varepsilon_{\Omega+1},0)$ gives rise to an elementary computable ordinal representation system. In what follows, $C^{\Omega}(\varepsilon_{\Omega+1},0)$ will be abbreviated to $\mathcal{T}(\Omega)$.

In the case of **PA** the addition of an infinitary rule restored the possibility of cut elimination. In order to obtain a similar result for set theories like **KP**, one has to work a bit harder. A peculiarity of **PA** is that every object n of the intended model has a canonical name in the language, namely, the n^{th} numeral. It is not clear, though, how to bestow a canonical name to each element of the set-theoretic universe. This is where *Gödel's constructible universe* **L** comes in handy. As **L** is "made" from the ordinals it is pretty obvious how to "name" sets in **L** once one has names for ordinals. These will be taken from $\mathcal{T}(\Omega)$. Henceforth, we shall restrict ourselves to ordinals from $\mathcal{T}(\Omega)$. The set terms and their ordinal levels are defined inductively. First, for each $\alpha\in\mathcal{T}(\Omega)\cap\Omega$, there will be a set term $\mathbb{L}_{\alpha}$. Its ordinal level is declared to be α. If $F(a,\vec{b})$ is a set-theoretic formula (whose free variables are among the indicated) and $\vec{s}\equiv s_1,\ldots,s_n$ are set terms with levels $<\alpha$, then the formal expression $\{x\in\mathbb{L}_{\alpha}:F(x,\vec{s})^{\mathbb{L}_{\alpha}}\}$ is a set term of level α. Here $F(x,\vec{s})^{\mathbb{L}_{\alpha}}$ results from $F(x,\vec{s})$ by restricting all unbounded quantifiers to $\mathbb{L}_{\alpha}$.

The collection of set terms will serve as a formal universe for a theory $\mathbf{KP}_{\infty}$ with infinitary rules. The infinitary rule for the universal quantifier on the right takes the form: From $\Gamma\Rightarrow\Delta,F(t)$ for all RS_{Ω}-terms t conclude $\Gamma\Rightarrow\Delta,\forall x\ F(x)$. There are also rules for bounded universal quantifiers: From $\Gamma\Rightarrow\Delta,F(t)$ for all RS_{Ω}-terms t with levels $<\alpha$ conclude $\Gamma\Rightarrow\Delta,(\forall x\in\mathbb{L}_{\alpha})\ F(x)$. The corresponding rule for introducing a universal quantifier bounded by a term of the form $\{x\in\mathbb{L}_{\alpha}:F(x,\vec{s})^{\mathbb{L}_{\alpha}}\}$ is slightly more complicated. With the help of these infinitary rules it now possible to give logical deductions of all axioms of **KP** with the exception of Bounded Collection. The latter can be deduced from the rule of Σ-Reflection: From $\Gamma\Rightarrow\Delta,C$ conclude $\Gamma\Rightarrow\Delta,\exists z\,C^{z}$ for every Σ-formula C. The class of Σ-formulae is the smallest class of formulae containing the bounded formulae which is closed under $\wedge$, $\vee$, bounded quantification and unbounded existential quantification. C^{z} is obtained from C by replacing all unbounded quantifiers $\exists x$ in C by $\exists x\in z$.

The length and cut ranks of $\mathbf{KP}_\infty$-deductions will be measured by ordinals from $\mathcal{T}(\Omega)$. If

$$\mathbf{KP} \vdash F(u_1, \ldots, u_r)$$

then $\mathbf{KP}_\infty \overset{\Omega\cdot m}{\underset{\Omega+n}{\vdash}} B(s_1, \ldots, s_r)$ holds for some m, n and all set terms $s_1, \ldots, s_r$; m and n depend only on the $\mathbf{KP}$-derivation of $B(\vec{u})$.

The usual cut elimination procedure works unless the cut formulae have been introduced by Σ-reflection rules. The obstacle to pushing cut elimination further is exemplified by the following scenario:

$$\dfrac{\dfrac{\overset{\delta}{\underset{\Omega}{\vdash}} \Gamma \Rightarrow \Delta, C}{\overset{\xi}{\underset{\Omega}{\vdash}} \Gamma \Rightarrow \Delta, \exists z\, C^z}\,(\Sigma\text{-Ref}) \qquad \dfrac{\cdots \overset{\xi_s}{\underset{\Omega}{\vdash}} \Xi, C^s \Rightarrow \Lambda \cdots (|s| < \Omega)}{\overset{\xi}{\underset{\Omega}{\vdash}} \Xi, \exists z\, C^z \Rightarrow \Lambda}\,(\exists L)}{\overset{\alpha}{\underset{\Omega+1}{\vdash}} \Gamma, \Xi \Rightarrow \Delta, \Lambda}\,(\mathrm{Cut})$$

In general, it won't be possible to remove such an instance of the Cut Rule. However, if the complexity of the side formulae is just right, the cut can be removed by a technique called *collapsing of deductions*. This method applies when the formulae in Γ and Ξ are Π-formulae and the formulae in Δ and Λ are Σ-formulae. The class of Π-formulae is the smallest class of formulae containing the bounded formulae which is closed under $\wedge$, $\vee$, bounded quantification and unbounded universal quantification.

For the technique of collapsing one needs the function $\alpha \mapsto \psi_\Omega(\alpha)$ and, moreover, it is necessary to ensure that the infinite deductions are of a very uniform character. The details are rather finicky and took several years to work out. The upshot is that every Σ sentence C deducible in $\mathbf{KP}$ has a cut-free deduction in $\mathbf{KP}_\infty$ of length $\psi_\Omega(\varepsilon_{\Omega+1})$, which entails that $L_{\psi_\Omega(\varepsilon_{\Omega+1})} \models C$. Moreover, the proof-theoretic ordinal of $\mathbf{KP}$ is $\psi_\Omega(\varepsilon_{\Omega+1})$, also known as the *Bachmann–Howard ordinal*.

2.4. Admissible proof theory. **KP** is the weakest in a line of theories that were analyzed by proof theorists of the Munich school in the late 1970s and 1980s. In many respects, **KP** is a very special case. Several fascinating aspects of ordinal analysis do not yet exhibit themselves at the level of **KP**.

Recall that **KPl** is the set-theoretic version of $(\Pi^1_1$-$\mathbf{AC}) + \mathbf{BI}$, while **KPi** is the set-theoretic counterpart to $(\Pi^1_1$-$\mathbf{AC}) + \mathbf{BI}$. The main axiom of **KPl** says that every set is contained in an admissible set (one also says that the admissible sets are cofinal in the universe) without requiring that the universe is also admissible, too. To get a sense of scale for comparing **KP**, **KPl**, and **KPi** it is perhaps best to relate the large cardinal assumptions that give rise to the pertaining ordinal representation systems. In the case of **KPl** the assumptions is that there are infinitely many large ordinals $\Omega_1, \Omega_2, \Omega_3, \ldots$ (where Ω_n can be taken to be $\aleph_n$) each equipped with their own 'collapsing' function $\alpha \mapsto \psi_{\Omega_n}(\alpha)$. The ordinal system sufficient for **KPi** is built using the much bolder assumption that there is an inaccessible cardinal I.

As the above set theories are based on the notion of admissible set it is suitable to call the proof theory concerned with them 'admissible proof theory'. The salient feature of admissible sets is that they are models of Bounded Collection and that that principle is equivalent to Σ Reflection on the basis of the other axioms of **KP** (see [4]). Furthermore, admissible sets of the form $\mathbf{L}_\kappa$ also satisfy Π_2 reflection, i.e., if $\mathbf{L}_\kappa \models \forall x\, \exists y\, C(x, y, \vec{a})$ with $C(x, y)$ bounded and $\vec{a} \in \mathbf{L}_\kappa$, then there exists $\rho < \kappa$ such that $\vec{a} \in \mathbf{L}_\rho$ and $\mathbf{L}_\rho \models \forall x\, \exists y\, C(x, y, \vec{a})$.

In essence, admissible proof theory is a gathering of cut-elimination and collapsing techniques that can handle infinitary calculi of set theory with Σ and/or Π_2 reflection rules, and thus lends itself to ordinal analyses of theories of the form **KP**+ *"there are x many admissibles"* or **KP**+ *"there are many admissibles"*.

A theory on the verge of admissible proof theory is **KPM**, designed to axiomatize essential features of a recursively Mahlo universe of sets. An admissible ordinal κ is said to be recursively Mahlo if it satisfies Π_2-reflection in the above sense but with the extra condition that the reflecting set $\mathbf{L}_\rho$ be admissible as well. The ordinal representation [25] for **KPM** is built on the assumption that there exists a Mahlo cardinal. The novel feature of over previous work is that there are two layers of collapsing functions. The ordinal analysis for **KPM** was carried out in [26]. A different approach to **KPM** using ordinal diagrams is due to Arai [1].

The means of admissible proof theory are too weak to deal with the next level of reflection having three alternations of quantifiers, i.e. Π_3-reflection.

2.5. Rewards of ordinal analysis Results that have been achieved through ordinal analysis mainly fall into four groups: (1) Consistency of subsystems of classical second order arithmetic and set theory relative to constructive theories, (2) reductions of theories formulated as conservation theorems, (3) combinatorial independence results, and (4) classifications of provable functions and ordinals. A detailed account of these results has been given in [31], section 3. An example where ordinal representation systems led to a new combinatorial result was Friedman's extension of Kruskal's Theorem, EKT, which asserts that finite trees are well-quasi-ordered under gap embeddability (see [39]). The gap condition imposed on the embeddings is directly related to an ordinal notation system that was used for the analysis of Π^1_1 comprehension. The principle EKT played a crucial role in the proof of the graph minor theorem of Robertson and Seymour (see [16]).

Theorem 2.1 (Robertson, Seymour). *For any infinite sequence $G_0, G_1, G_2, \ldots$ of finite graphs there exist $i < j$ so that G_i is isomorphic to a minor of G_j.*

3. Beyond admissible proof theory

Gentzen fostered hopes that with sufficiently large constructive ordinals one could establish the consistency of analysis, i.e., $\mathbf{Z_2}$. The purpose of this section is to report on the next major step in analyzing fragments of $\mathbf{Z_2}$. This is obviously the ordinal analysis of the system $(\Pi^1_2\text{-}\mathbf{CA})$.[1] The strength of $(\Pi^1_2\text{-}\mathbf{CA})$ dwarfs that of $(\Pi^1_1\text{-}\mathbf{AC})$. The treatment of Π^1_2 comprehension posed formidable technical challenges (see [30], [32], [33]). Other approaches to ordinal analysis of systems above Π^1_1-**AC** are due to Arai (see [1], [2]) who uses ordinal diagrams and finite deductions, and Carlson [11] who employs patterns of resemblance.

In the following, we will gradually slice Π^1_2 comprehension into degrees of reflection to achieve a sense of scale. There is no way to describe this comprehension simply in terms of admissibility except that on the set-theoretic side, Π^1_2 comprehension corresponds to Σ_1 separation, i.e. the scheme of axioms

$$\exists z(z = \{x \in a : \phi(x)\})$$

for all Σ_1 formulas ϕ. The precise relationship is as follows:

Theorem 3.1. $\mathbf{KP} + \Sigma_1$ *separation and* $(\Pi^1_2\text{-}\mathbf{CA}) + \mathbf{BI}$ *prove the same sentences of second order arithmetic.*

The ordinals κ such that $\mathbf{L}_\kappa \models \mathbf{KP} + \Sigma_1$-*Separation* are familiar from ordinal recursion theory.

Definition 3.2. An admissible ordinal κ is said to be *nonprojectible* if there is no total κ-recursive function mapping κ one-one into some $\beta < \kappa$, where a function $g : \mathbf{L}_\kappa \to \mathbf{L}_\kappa$ is called *κ-recursive* if it is Σ definable in $\mathbf{L}_\kappa$.

The key to the 'largeness' properties of nonprojectible ordinals is that for any nonprojectible ordinal κ, $\mathbf{L}_\kappa$ is a limit of Σ_1-elementary substructures, i.e. for every $\beta < \kappa$ there exists a $\beta < \rho < \kappa$ such that $\mathbf{L}_\rho$ is a Σ_1-elementary substructure of $\mathbf{L}_\kappa$, written $\mathbf{L}_\rho \prec_1 \mathbf{L}_\kappa$.

Such ordinals satisfying $\mathbf{L}_\rho \prec_1 \mathbf{L}_\kappa$ have strong reflecting properties. For instance, if $\mathbf{L}_\rho \models C$ for some set-theoretic sentence C (containing parameters from $\mathbf{L}_\rho$), then there exists a $\gamma < \rho$ such that $\mathbf{L}_\gamma \models C$. This is because $\mathbf{L}_\rho \models C$ implies $\mathbf{L}_\kappa \models \exists\gamma\, C^{\mathbf{L}_\gamma}$, hence $\mathbf{L}_\rho \models \exists\gamma\, C^{\mathbf{L}_\gamma}$ using $\mathbf{L}_\rho \prec_1 \mathbf{L}_\kappa$.

The last result makes it clear that an ordinal analysis of Π^1_2 comprehension would necessarily involve a proof-theoretic treatment of reflections beyond those surfacing in admissible proof theory. The notion of stability will be instrumental.

Definition 3.3. α is *δ-stable* if $\mathbf{L}_\alpha \prec_1 \mathbf{L}_{\alpha+\delta}$.

For our purposes we need refinements of this notion, the simplest being provided by:

Definition 3.4. $\alpha > 0$ *is said to be* Π_n*-reflecting* if $\mathbf{L}_\alpha \models \Pi_n$-reflection. By Π_n-reflection we mean the scheme $C \to \exists z[\mathrm{Tran}(z) \wedge z \neq \emptyset \wedge C^z]$, where C is Π_n, and $\mathrm{Tran}(z)$ expresses that z is a transitive set.

[1] For more background information see [42], p. 259, [15], p. 362, [24], p. 374.

Π_n-reflection for all n suffices to express one step in the $\prec_1$ relation.

Lemma 3.5 (cf. [34], 1.18). $\mathbf{L}_\kappa \prec_1 \mathbf{L}_{\kappa+1}$ iff κ is Π_n-reflecting for all n.

The step of analyzing Kripke–Platek set theory augmented by Π_n-reflection rules was taken in [29]; the ordinal representation system for Π_3-reflection employed a weakly compact cardinal.

A further refinement of the notion of δ-stability will be addressed next.

Definition 3.6. κ is said to be δ-Π_n-reflecting if whenever $C(u,\vec{x})$ is a set-theoretic Π_n formula, $a_1,\dots,a_r \in \mathbf{L}_\kappa$ and $\mathbf{L}_{\kappa+\delta} \models C[\kappa, a_1,\dots,a_n]$, then there exists $\kappa_0, \delta_0 < \kappa$ such that $a_1,\dots,a_r \in \mathbf{L}_{\kappa_0}$ and $\mathbf{L}_{\kappa_0+\delta_0} \models C[\kappa_0, a_1,\dots,a_n]$.

Putting the previous definition to work, one gets:

Corollary 3.7. If κ is $\delta+1$-Σ_1-reflecting, then, for all n, κ is δ-Σ_n-reflecting.

At this point let us return to proof theory to explain the need for even further refinements of the preceding notions. Recall that the first nonprojectible ordinal ρ is a limit of smaller ordinals ρ_n such that $\mathbf{L}_{\rho_n} \prec_1 \mathbf{L}_\rho$. In the ordinal representation system $\mathcal{OR}$ for Π^1_2-**CA**, there will be symbols $\mathfrak{E}_n$ and $\mathfrak{E}_\omega$ for ρ_n and ρ, respectively. The associated infinitary proof system will have rules

$$(\mathrm{Ref}_{\Sigma(\mathbb{L}_{\mathfrak{E}_n+\delta})})\quad \frac{\Gamma \Rightarrow \Delta, C(\vec{s})^{\mathbb{L}_{\mathfrak{E}_n+\delta}}}{\Gamma \Rightarrow \Delta,\ (\exists z \in \mathbb{L}_{\mathfrak{E}_n})(\exists \vec{x} \in \mathbb{L}_{\mathfrak{E}_n})[\mathrm{Tran}(z) \wedge C(\vec{x})^z]},$$

where $C(\vec{x})$ is a Σ formula, $\vec{s}$ are set terms of levels $< \mathfrak{E}_n + \delta$, and $\delta < \mathfrak{E}_\omega$. These rules suffice to bring about the embedding $\mathbf{KP} + \Sigma_1$-Separation into the infinitary proof system, but reflection rules galore will be needed to carry out cut-elimination. For example, there will be "many" ordinals $\pi, \delta \in \mathcal{OR}$ that play the role of δ-Π_{n+1}-reflecting ordinals by virtue of corresponding reflection rules in the infinitary calculus.

4. A large cardinal notion

An important part of ordinal analysis is the development of ordinal representation systems. Extensive ordinal representation systems are difficult to understand from a purely syntactical point of view, often to such an extent that it makes no sense to present an ordinal representation system without giving some kind of semantic interpretation. Large cardinals have been used quite frequently in the definition procedure of strong ordinal representation systems, and large cardinal notions have been an important source of inspiration. In the end, they can be dispensed with, but they add an intriguing twist to the relation between set theory and proof theory. The advantage of working in a strong set-theoretic context is that we can build models without getting buried under complexity considerations.

Such systems are usually generated from collapsing functions. However, from now on we prefer to call them *projection functions* since they will no longer bear any resemblance to Mostowski's collapsing function. In [33], the projection functions needed for the ordinal analysis of Π^1_2 have been construed as inverses to certain partial elementary embeddings. In this final section we shall indicate a model for the projection functions, employing rather sweeping large cardinal axioms, in that we shall presume the existence of certain cardinals, featuring a strong form of indescribability, dubbed *shrewdness*.

To be able to eliminate reflections of the type described in Definition 3.6 requires projection functions which can project intervals $[\kappa, \kappa+\delta]$ of ordinals down below κ.

Definition 4.1. Let $V = \bigcup_{\alpha\in ON} V_\alpha$ be the cumulative hierarchy of sets, i.e.

$$V_0 = \emptyset, \quad V_{\alpha+1} = \{X : X \subseteq V_\alpha\}, \quad V_\lambda = \bigcup_{\xi<\lambda} V_\xi \text{ for limit ordinals } \lambda.$$

Let $\eta > 0$. A cardinal κ is *η-shrewd* if for all $P \subseteq V_\kappa$ and every set-theoretic formula $F(v_0, v_1)$, whenever

$$V_{\kappa+\eta} \models F[P, \kappa],$$

then there exist $0 < \kappa_0, \eta_0 < \kappa$ such that

$$V_{\kappa_0+\eta_0} \models F[P \cap V_{\kappa_0}, \kappa_0].$$

κ is *shrewd* if κ is η-shrewd for every $\eta > 0$.

Let $\mathcal{F}$ be a collection of formulae. A cardinal κ is *η-$\mathcal{F}$-shrewd* if for all $P \subseteq V_\kappa$ and every $\mathcal{F}$-formula $H(v_0, v_1)$, whenever

$$V_{\kappa+\eta} \models H[P, \kappa],$$

then there exist $0 < \kappa_0, \eta_0 < \kappa$ such that

$$V_{\kappa_0+\eta_0} \models H[P \cap V_{\kappa_0}, \kappa_0].$$

We will also consider a notion of shrewdness with regard to a given class.

Let $\mathbf{U}$ be a fresh unary predicate symbol. Given a language $\mathcal{L}$ let $\mathcal{L}(\mathbf{U})$ denote its extension by $\mathbf{U}$. If $\mathcal{A}$ is a class we denote by $\langle V_\alpha; \mathcal{A}\rangle$ the structure $\langle V_\alpha; \in; \mathcal{A} \cap V_\alpha\rangle$.

For an $\mathcal{L}_{\mathrm{set}}(\mathbf{U})$-sentence ϕ, let the meaning of "$\langle V_\alpha; \mathcal{A}\rangle \models \phi$" be determined by interpreting $\mathbf{U}(t)$ as $t \in \mathcal{A} \cap V_\alpha$.

Definition 4.2. Assume that $\mathcal{A}$ is a class. Let $\eta > 0$. A cardinal κ is *$\mathcal{A}$-η-shrewd* if for all $P \subseteq V_\kappa$ and every formula $F(v_0, v_1)$ of $\mathcal{L}_{\mathrm{set}}(\mathbf{U})$, whenever

$$\langle V_{\kappa+\eta}; \mathcal{A}\rangle \models F[P, \kappa],$$

then there exist $0 < \kappa_0, \eta_0 < \kappa$ such that

$$\langle V_{\kappa_0+\eta_0}; \mathcal{A}\rangle \models F[P \cap V_{\kappa_0}, \kappa_0].$$

κ is $\mathcal{A}$-*shrewd* if κ is $\mathcal{A}$-η-shrewd for every $\eta > 0$.

Likewise, for $\mathcal{F}$ a collection of formulae in a language $\mathcal{L}(\mathbf{U})$, we say that a cardinal κ is $\mathcal{A}$-η-$\mathcal{F}$-*shrewd* if for all $P \subseteq V_\kappa$ and every $\mathcal{F}$-formula $H(v_0, v_1)$, whenever

$$\langle V_{\kappa+\eta}; \mathcal{A}\rangle \models H[P, \kappa],$$

then there exist $0 < \kappa_0, \eta_0 < \kappa$ such that

$$\langle V_{\kappa_0+\eta_0}; \mathcal{A}\rangle \models H[P \cap V_{\kappa_0}, \kappa_0].$$

Corollary 4.3. If κ is $\mathcal{A}$-δ-shrewd and $0 < \eta < \delta$, then κ is $\mathcal{A}$-η-shrewd.

There are similarities between the notions of η-shrewdness and η-indescribability (see [12], Ch. 9, §4). However, it should be noted that if κ is η-indescribable and $\rho < \eta$, it does not necessarily follow that κ is also ρ-indescribable (see [12], 9.4.6).

A reason for calling the above cardinals *shrewd* is that if there is a shrewd cardinal κ in the universe, then, loosely speaking, for any notion of large cardinal N which does not make reference to the totality of all ordinals, if there exists an N-cardinal then the least such cardinal is below κ. So for instance, if there are measurable and shrewd cardinals in the universe, then the least measurable is smaller than the least shrewd cardinal.

To situate the notion of shrewdness with regard to consistency strength in the usual hierarchy of large cardinals, we recall the notion of a subtle cardinal.

Definition 4.4. A cardinal κ is said to be *subtle* if for any sequence $\langle S_\alpha : \alpha < \kappa\rangle$ such that $S_\alpha \subseteq \alpha$ and C closed and unbounded in κ, there are $\beta < \delta$ both in C satisfying

$$S_\delta \cap \beta = S_\beta.$$

Since subtle cardinals are not covered in many of the standard texts dealing with large cardinals, we mention the following facts (see [22], §20):

Remark 4.5. Let $\kappa(\omega)$ denote the first ω-Erdös cardinal.

(i) $\{\pi < \kappa(\omega) : \pi \text{ is subtle}\}$ is stationary in $\kappa(\omega)$.

(ii) 'Subtlety' relativises to $\mathbf{L}$, i.e. if π is subtle, then $\mathbf{L} \models$ "π is subtle".

Lemma 4.6. Assume that π is a subtle cardinal and that $\mathcal{A} \subseteq V_\pi$. Then for every $B \subseteq \pi$ closed and unbounded in π there exists $\kappa \in B$ such that

$$\langle V_\pi; \mathcal{A}\rangle \models \text{“}\kappa \text{ is } \mathcal{A}\text{-shrewd''}.$$

Corollary 4.7. Assume that π is a subtle cardinal. Then there exists a cardinal $\kappa < \pi$ such that κ is η-shrewd for all $\eta < \pi$.

References

[1] Arai, T., Proof theory for theories of ordinals I: recursively Mahlo ordinals. *Ann. Pure Appl. Logic* **122** (2003), 1–85.

[2] Arai, T., Proof theory for theories of ordinals II: Π_3-Reflection. *Ann. Pure Appl. Logic* **129** (2004), 39–92.

[3] Bachmann, H., Die Normalfunktionen und das Problem der ausgezeichneten Folgen von Ordinalzahlen. *Vierteljschr. Naturforsch. Ges. Zürich* **95** (1950), 115–147.

[4] Barwise, J., *Admissible Sets and Structures*. Perspectives in Mathematical Logic, Springer-Verlag, Berlin 1975.

[5] Buchholz, W., Eine Erweiterung der Schnitteliminationsmethode. Habilitationsschrift, München 1977.

[6] Buchholz, W., A simplified version of local predicativity. In *Proof theory* (Leeds, 1990), ed. by P. Aczel, H. Simmons, S. S. Wainer, Cambridge University Press, Cambridge 1993, 115–147.

[7] Buchholz, W., Feferman, S., Pohlers, W., Sieg, W., *Iterated inductive definitions and subsystems of analysis*. Lecture Notes in Math. 897, Springer-Verlag, Berlin 1981.

[8] Buchholz, W., and Schütte, K., *Proof theory of impredicative subsystems of analysis.* Stud. Proof Theory Monogr. 2, Bibliopolis, Naples 1988.

[9] Buchholz, W., Explaining Gentzen's consistency proof within infinitary proof theory. In *Computational Logic and Proof Theory, KGC '97* (ed. by G. Gottlob et al.), Lecture Notes in Comput. Sci. 1289, Springer-Verlag, Berlin 1997, 4–17.

[10] Cantor, G., Beiträge zur Begründung der transfiniten Mengenlehre II. *Math. Ann.* **49** (1897), 207–246.

[11] Carlson, T., Elementary patterns of resemblance. *Ann. Pure Appl. Logic* **108** (2001), 19–77.

[12] Drake, F., *Set Theory: An introduction to large cardinals*. Stud. Logic Found. Math. 76, North Holland, Amsterdam 1974.

[13] Feferman, S., Systems of predicative analysis. *J. Symbolic Logic* **29** (1964), 1–30.

[14] Feferman, S., Proof theory: a personal report. In *Proof Theory* (ed. by G. Takeuti), Stud. Logic Found. Math. 81, 2nd edition, North-Holland, Amsterdam 1987, 445–485.

[15] Feferman, S., Remarks for "The Trends in Logic". In *Logic Colloquium '88*, North-Holland, Amsterdam 1989, 361–363.

[16] Friedman, H., Robertson, N., Seymour, P., The metamathematics of the graph minor theorem. *Contemp. Math.* **65** (1987), 229–261.

[17] Hardy, G. H. A theorem concerning the infinite cardinal numbers. *Quart. J. Math.* **35** (1904), 87–94.

[18] Hilbert, D., Die Grundlegung der elementaren Zahlentheorie. *Math. Ann.* **104** (1931), 485–494.

[19] Hilbert, D., and Bernays, P., *Grundlagen der Mathematik II*. Grundlehren Math. Wiss. 50, Springer-Verlag, Berlin 1939.

[20] Jäger, G., Zur Beweistheorie der Kripke–Platek Mengenlehre über den natürlichen Zahlen. *Arch. Math. Logik Grundlag.* **22** (1982), 121–139.

[21] Jäger, G., and Pohlers, W., Eine beweistheoretische Untersuchung von Δ^1_2-**CA** + **BI** und verwandter Systeme. *Bayer. Akad. Wiss. Math.-Natur. Kl. Sitzungsber.* **1982** (1983), 1–28.

[22] Kanamori, A., Magidor, M., The evolution of large cardinal axioms in set theory. In *Higher Set Theory* (ed. by G. H. Müller, D. S. Scott), Lecture Notes in Math. 669, Springer-Verlag, Berlin 1978, 99–275.

[23] Pohlers, W., Cut elimination for impredicative infinitary systems, part II: Ordinal analysis for iterated inductive definitions. *Arch. Math. Logik Grundlag.* **22** (1982), 113–129.

[24] Pohlers, W., Proof theory and ordinal analysis. *Arch. Math. Logic* **30** (1991), 311–376.

[25] Rathjen, M., Ordinal notations based on a weakly Mahlo cardinal. *Arch. Math. Logic* **29** (1990), 249–263.

[26] Rathjen, M., Proof-Theoretic Analysis of KPM. *Arch. Math. Logic* **30** (1991), 377–403.

[27] Rathjen, M., How to develop proof–theoretic ordinal functions on the basis of admissible sets. *Math. Logic Quart.* **39** (1993), 47–54.

[28] Rathjen, M., Collapsing functions based on recursively large ordinals: A well–ordering proof for KPM. *Arch. Math. Logic* **33** (1994), 35–55.

[29] Rathjen, M., Proof theory of reflection. *Ann. Pure Appl. Logic* **68** (1994), 181–224.

[30] Rathjen, M., Recent advances in ordinal analysis: Π^1_2-CA and related systems. *Bull. Symbolic Logic* **1** (1995), 468–485.

[31] Rathjen, M., The realm of ordinal analysis. In *Sets and Proofs* (ed. by S. B. Cooper and J. K. Truss), London Math. Soc. Lecture Note Ser. 258, Cambridge University Press, Cambridge 1999, 219–279.

[32] Rathjen, M., An ordinal analysis of stability. *Arch. Math. Logic* **44** (2005), 1–62.

[33] Rathjen, M., An ordinal analysis of parameter-free Π^1_2 comprehension. *Arch. Math. Logic* **44** (2005), 263–362.

[34] Richter, W. and Aczel, P., Inductive definitions and reflecting properties of admissible ordinals. In *Generalized Recursion Theory* (ed. by J. E. Fenstad and P. G. Hinman), Stud. Logic Found. Math. 79, North Holland, Amsterdam 1973, 301–381.

[35] Schütte, K., Beweistheoretische Erfassung der unendlichen Induktion in der Zahlentheorie. *Math. Ann.* **122** (1951), 369–389.

[36] Schütte, K., *Beweistheorie*. Grundlehren Math. Wiss. 103, Springer-Verlag, Berlin 1960.

[37] Schütte, K., Eine Grenze für die Beweisbarkeit der transfiniten Induktion in der verzweigten Typenlogik. *Arch. Math. Logik Grundlagenforsch.* **67** (1964), 45–60.

[38] Schütte, K., Predicative well-orderings. In *Formal systems and recursive functions* (ed. by J. N. Crossley and M. A. E. Dummett), North-Holland, Amsterdam 1965, 176–184.

[39] Simpson, S., Nichtbeweisbarkeit von gewissen kombinatorischen Eigenschaften endlicher Bäume. *Arch. Math. Logik Grundlag.* **25** (1985), 45–65.

[40] Simpson, S., *Subsystems of second order arithmetic*. Perspect. Math. Logic, Springer-Verlag, Berlin 1999.

[41] Takeuti, G., Consistency proofs of subsystems of classical analysis. *Ann. of Math.* (2) **86** (1967), 299–348.

[42] Takeuti, G., Proof theory and set theory. *Synthese* **62** (1985), 255–263.

[43] Takeuti, G., M. Yasugi, M., The ordinals of the systems of second order arithmetic with the provably Δ^1_2-comprehension and the Δ^1_2-comprehension axiom respectively. *Japan J. Math.* **41** (1973), 1–67.

[44] Veblen, O., Continuous increasing functions of finite and transfinite ordinals. *Trans. Amer. Math. Soc.* **9** (1908), 280–292.

Department of Mathematics, The Ohio State University, Columbus, Ohio 43210, U.S.A.
and
Department of Pure Mathematics, University of Leeds, Leeds LS2 9JT, England
E-mail: rathjen@math.ohio-state.edu