THE BULLETIN OF SYMBOLIC LOGIC
Volume 1, Number 4, Dec. 1995

RECENT ADVANCES IN ORDINAL ANALYSIS: Π^1_2 – **CA** AND RELATED SYSTEMS

MICHAEL RATHJEN

§1. Introduction. The purpose of this paper is, in general, to report the state of the art of ordinal analysis and, in particular, the recent success in obtaining an ordinal analysis for the system of Π^1_2-analysis, which is the subsystem of formal second order arithmetic, $\mathbf{Z_2}$, with comprehension confined to Π^1_2-formulae.[1] The same techniques can be used to provide ordinal analyses for theories that are reducible to iterated Π^1_2-comprehension, e.g., Δ^1_3-comprehension. The details will be laid out in [28].

Ordinal-theoretic proof theory came into existence in 1936, springing forth from Gentzen's head in the course of his consistency proof of arithmetic. Gentzen fostered hopes that with sufficiently large constructive ordinals one could establish the consistency of analysis, i.e., $\mathbf{Z_2}$. Considerable progress has been made in proof theory since Gentzen's tragic death on August 4th, 1945, but an ordinal analysis of $\mathbf{Z_2}$ is still something to be sought. However, for reasons that cannot be explained here, Π^1_2-comprehension appears to be the main stumbling block[2] on the road to understanding full comprehension, giving hope for an ordinal analysis of $\mathbf{Z_2}$ in the foreseeable future.

Roughly speaking, *ordinally informative* proof theory attaches ordinals in a recursive representation system to proofs in a given formal system; transformations on proofs to certain canonical forms are then partially mirrored by operations on the associated ordinals. Among other things, ordinal analysis of a formal system serves to characterize its provably recursive ordinals, functions and functionals and can yield both conservation and combinatorial independence results. Since there is no wide familiarity among logicians with ordinally informative proof theory, we begin in §2 with an explanation of its current rationale and goals, which take the place of the original

Received June 1, 1995; revised September 27, 1995.
The author is a Heisenberg Fellow of the German Science Foundation (Deutsche Forschungsgemeinschaft).

[1]I would like to thank Tim Carlson for innumerable and invaluable conversations which have helped me find my way through Π^1_2. Carlson is completing work on an approach to Π^1_2 comprehension which was developed independently of that presented here.

[2]For more background information see [43, p. 259], [11, p. 362], [27, p. 374]

1079-8986/95/0104-0004/$2.80

Hilbert Program, and follow that in §3 with an explanation of its basic technical tools, namely ordinal representation systems and ramified set theory equipped with suitable reflection rules. The connection of the system of Π^1_2-comprehension ($\Pi^1_2 - \mathbf{CA}$ hereafter) with set theory comes through the fact that $\mathbf{KP} + \Sigma_1$-separation is a conservative extension of $\Pi^1_2 - \mathbf{CA} + \mathbf{BI}$, where **BI** is the so-called principle of *Bar Induction*. It is indicated in §4 how this fragment of set theory can be sliced into a hierarchy of reflection principles. Then §5 describes in outline an ordinal representation system which serves for the ordinal analyses of $\mathbf{KP} + \Sigma_1$-separation. Most important in this regard are certain *projection functions* (often called *collapsing functions*). As their existence is rather difficult to prove, in §6 we will indicate a model for them based on very large cardinals, in which they can be construed as inverses of certain partial elementary embeddings.

To set the stage for the following, a very brief history of ordinal-theoretic proof theory since Gentzen reads as follows: In the 1950's proof theory flourished in the hands of Schütte: in [35] he introduced an infinitary system for first order number theory with the so-called ω-rule, which had already been proposed by Hilbert [16]. Ordinals were assigned as lengths to derivations and via cut-elimination he re-obtained Gentzen's ordinal analysis for number theory in a particularly transparent way. Further, Schütte extended his approach to systems of ramified analysis and brought this technique to perfection in his monograph "Beweistheorie" [36]. Independently, in 1964 Feferman [6] and Schütte [37, 38] determined the ordinal bound Γ_0 for theories of autonomous ramified progressions.

A major breakthrough was made by Takeuti in 1967, who for the first time obtained an ordinal analysis of an impredicative theory. In [42] he gave an ordinal analysis of Π^1_1 comprehension, extended in 1973 to Δ^1_2 comprehension in [44] 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 1970's (see [4]). 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 [2]) 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 [5] 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 [25, 24, 26]). With the work of Jäger and Pohlers (see [17, 18, 20]) 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 with Δ^1_2-comprehension plus **BI** [20]. In essence, admissible proof theory is a gathering of cut-elimination techniques for infinitary calculi of ramified set theory with Σ- and/or Π_2-reflection rules[3] that lend itself to ordinal analyses of theories of the form **KP**+ *"there are x many admissibles"* or **KP**+ *"there are many admissibles."* By way of illustration, the subsystem of analysis with Δ^1_2-comprehension and Bar Induction can be couched in such terms, for it is naturally interpretable in the set theory **KPi** := **KP** + $\forall y \exists z(y{\in}z \wedge$ *z is admissible*) (cf. [20]).

After an intermediate step [29], which dealt with a set theory **KPM** which formalizes a recursively Mahlo universe, a major step beyond admissible proof theory was taken in [32]. That paper featured ordinal analyses of extensions of **KP** by Π_n-reflection. A generalization of the methods of [32] underlies the treatment of $\Pi^1_2 - \mathbf{CA}$ sketched below.

§**2. Goals of proof theory.** If proof theory is to be applied to stronger and stronger systems, what are the aims of this research? In the author's view, a fruitful way of looking at what has been accomplished and might be accomplished by ordinal analyses is given by the following list:

(I) Hilbert's Program relativized: reduction (interpretation) of impredicative theories to (in) constructive frameworks
(II) Finding currently untapped resources of set theory: new combinatorial principles
(III) Development of a proof theory of set theory

We shall illustrate the above themes by a few examples.

Ad (I): A modified form of Hilbert's program has, of course, already been pursued by the Hilbert school (Bernays, Gentzen, …). In more elaborate terms, Feferman [9, 10] has put forward a relativized version of Hilbert's program, in which one develops a global picture of *What rests on what?* in mathematics, i.e., what higher-order concepts and principles can be reduced to prima facie more basic terms.

Once one is investigating theories which have at least the strength of $\Delta^1_2 - \mathbf{CA} + \mathbf{BI}$, it is obvious that "securing" ordinary mathematics is not the main concern. Even without knowledge of that program carried out

[3] Recall that the salient feature of admissible sets is that they are models of Δ_0-collection and that Δ_0-collection is equivalent to Σ-reflection on the basis of the other axioms of **KP** (see [1]). Furthermore, admissible sets of the form $\mathbf{L}_\alpha$ also satisfy Π_2-reflection.

under the rubric of "reverse mathematics," it is easily seen that most of ordinary mathematics can be formalized in $\Delta^1_2 - \mathbf{CA} + \mathbf{BI}$ without effort (note that in the context of second order arithmetic **BI** serves the same purpose as $\in$-induction in set theory). So we must be interested in these strong impredicative principles for their own sake.

The reductions we have in mind, underlies a broadened view of "constructivity." There is no once-and-for-all delineated system of constructivism, rather it is an open-ended framework such as Martin-Löf's type theory, where new types along with new forms of inferences might be added as we widen our horizon by reaching higher levels of reflection (a constructivist's large cardinal program). However, space limitations do not permit us to dwell on a constructivist's manifesto of a liberal persuasion.[4]

Constructive theories of functions and sets that relate to Bishop's constructive mathematics as theories like ZFC relate to Cantorian set theory have been proposed by Myhill, Martin–Löf, Feferman and Aczel. Among those are Feferman's constructive theory of operations and classes, $\mathbf{T_0}$ ([7, 8]), and Martin-Löf's intuitionistic type theory of [23] (the latter does not have Russell's notorious *reducibility axiom*). By employing an ordinal analysis for **KPi** it has been shown that **KPi** and consequently $\Delta^1_2 - \mathbf{CA}+\mathbf{BI}$ can be reduced to both these theories.

THEOREM 2.1 (Feferman [7], Jäger [19], Jäger and Pohlers [20]). *$\Delta^1_2 - \mathbf{CA}$ $+\mathbf{BI}$, **KPi**, and $\mathbf{T_0}$ are proof-theoretically equivalent. In particular, these theories prove the same theorems in the negative arithmetic fragment.*

THEOREM 2.2 (Griffor and Rathjen [15]; Setzer [39]). *The soundness of the negative arithmetic fragment of $\Delta^1_2 - \mathbf{CA}+\mathbf{BI}$ and **KPi** is provable in Martin-Löf's 1984 type theory.*

Ad (II): It has been shown that certain large cardinal axioms settle questions about definable sets of reals, in that they imply the axiom of projective determinacy (see [22]). The existence of infinitely many Woodin cardinals implies that every uncountable projective set of reals is Lebesgue measurable and has a perfect subset, and thus possesses the same cardinality as the continuum (see [45]). However, it is as yet to be shown that the strength of large cardinal axioms has interesting repercussions in the context of arithmetic. A hope in connection with ordinal analyses is that they lead to new combinatorial principles which encapsulate considerable proof-theoretic strength. Examples are still scarce. One case where ordinal notations 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 [40]). The gap condition imposed on the embeddings is directly

[4]It is remarkable that Gödel seems to espouse a similar view in an unpublished essay from 1961 (see [14, pp. 382–387]), where he makes interesting remarks on the growth of reflection.

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 [12]).

THEOREM 2.3 (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.*

Ad (III): There is a proof theory for extensions of **KP**. Up till now there does not exist a proof theory of ZFC with its essential use of the power set axiom. As a matter of fact, the flow of ideas has been in the other direction as ordinal-theoretic proof theory has been constructivizing notions of large cardinals. However, if one takes a bold optimistic view that proof theory can be developed up to very large cardinals, then proof theory might shed new light on these lofty regions. For instance, one could hope for new insights as to Kunen's proof of the impossibility of a non-trivial elementary embedding of the universe into itself. All known proofs of $\neg\exists j\ (\ j : V \prec V \ \wedge\ j \neq \mathrm{id}\,)$ have an ad hoc flavor.

§**3. Admissible proof theory.** First order number theory is distinguished by the fact that its realm of discourse is clearly delineated, and moreover, that each element n of its realm possesses a canonical name, the nth numeral $\bar{n}$. By contrast, the universe of a set theory is at most implicitly described by its axioms.[5]

When searching for an analogue of the ω-rule, (for all n: $\vdash\ \phi(\bar{n})\ \Rightarrow \vdash \forall x\phi(x)$), for systems of set theory, one gets confronted with the problem of "naming" sets. However, if a set of ordinals or an ordinal representation system $\mathcal{OR}$ is given at the outset, one can build its associated formal constructible hierarchy, the $RS_{\mathcal{OR}}$-terms, as follows: For each $\alpha \in \mathcal{OR}$, $\mathbb{L}_\alpha$ is an $RS_{\mathcal{OR}}$-term of level α, and, if $s_1, \cdots, s_n$ are $RS_{\mathcal{OR}}$-terms with levels $< \alpha$ and if $F(a, b_1, \cdots, b_n)$ is a formula of the ordinary set-theoretic language, then the formal expression $[x \in \mathbb{L}_\alpha : F(x, s_1, \cdots, s_n)^{\mathbb{L}_\alpha}]$ is an $RS_{\mathcal{OR}}$-term of level α . As for the intended interpretation, $\mathbb{L}_\alpha$ is supposed to "name" $\mathbf{L}_\alpha$ and $[x \in \mathbb{L}_\alpha : F(x, s_1, \cdots, s_n)^{\mathbb{L}_\alpha}]$ to "name" the set $\{x \in \mathbf{L}_\alpha : F(x, a_1, \cdots, a_n)^{\mathbf{L}_\alpha}\}$, where the superscript $\mathbf{L}_\alpha$ makes all unrestricted quantifiers in F become restricted to $\mathbf{L}_\alpha$ and $a_1, \cdots, a_n$ are the sets denoted by $s_1, \cdots, s_n$, respectively. Aside from some notational conventions, we are now prepared to introduce the calculus of ramified set theory, $RS_{\mathcal{OR}}$. The conventions are that for an $RS_{\mathcal{OR}}$-term t, $|\,t\,|$ denotes its level, $t \in \mathcal{T}_\alpha\ :=\ \{t : |\,t\,| < \alpha\}$, and for terms s, t with $|\,s\,| < |\,t\,|$ we set $s \mathring{\in} t\ :=\ B(s)$ if t is $[x \in \mathbb{L}_\beta : B(x)]$ and $s \mathring{\in} t\ :=\ s = s$

[5]This makes ordinal analysis of set theories so intriguing. Ordinals appear both in the guise of ordinal representation systems and as objects of the theory under investigation, and the relation between these appearances is rather subtle.

if t is of the form $\mathbb{L}_\beta$. Observe that $s \dot{\in} t$ and $s \mathring{\in} t$ have the same truth value under the standard interpretation in the constructible hierarchy.

DEFINITION 3.1. The *rules* of $RS_{\mathcal{OR}}$ are:

$$
\begin{array}{ll}
(\wedge) & \dfrac{\Gamma, A \quad \Gamma, A'}{\Gamma, A \wedge A'} \\
(\vee) & \dfrac{\Gamma, A_i}{\Gamma, A_0 \vee A_1} \quad \text{if } i = 0 \text{ or } i = 1 \\
(\forall) & \dfrac{\cdots \Gamma, s \mathring{\in} t \rightarrow F(s) \cdots (s \in \mathcal{T}_{|t|})}{\Gamma, (\forall x \in t) F(x)} \\
(\exists) & \dfrac{\Gamma, s \mathring{\in} t \wedge F(s)}{\Gamma, (\exists x \in t) F(x)} \quad \text{if } s \in \mathcal{T}_{|t|} \\
(\notin) & \dfrac{\cdots \Gamma, s \mathring{\in} t \rightarrow r \neq s \cdots\cdots (s \in \mathcal{T}_{|t|})}{\Gamma, r \notin t} \\
(\in) & \dfrac{\Gamma, s \mathring{\in} t \ \wedge \ r = s}{\Gamma, r \in t} \quad \text{if } s \in \mathcal{T}_{|t|} \\
(\mathrm{Cut}) & \dfrac{\Gamma, A \qquad \Gamma, \neg A}{\Gamma}
\end{array}
$$

The rules of $RS_{\mathcal{OR}}$ are pure infinitary logic and therefore this calculus enjoys cut-elimination. But unlike the case of first order number theory, in general, we won't be able to embed **KP** into $RS_{\mathcal{OR}}$. We will have to add non-logical rules, Σ-reflection rules (cf. [1, I.4.1]), to $RS_{\mathcal{OR}}$ in order to accommodate **KP**. However, the addition of such rules, providing $\mathcal{OR}$ is recursive, will render the infinitary calculus inconsistent, unless we impose uniformity restrictions on infinitary derivations. To explain this kind of uniformity, rather than construing $\mathcal{OR}$ as an initial segment of the ordinals it is important to conceive of $\mathcal{OR}$ as a subset of the ordinals which has gaps.

To make the preceding more understandable, we shall expound to some extent on a concrete example, the ordinal analysis of **KPi**. To begin with, we shall explain how a sufficiently strong ordinal representation system comes about. Recall that the least standard model of **KPi** is the structure $\mathbf{L_I}$, where $\mathbf{I}$ denotes the first recursively inaccessible ordinal.

Let $\alpha \mapsto \Omega_\alpha$ be the function that enumerates the admissible ordinals and their limits, and let $\Re$ be the set of admissible ordinals κ so that $\omega < \kappa \leq \mathbf{I}$. Variables κ, π will range over $\Re$. An ordinal representation system for the analysis of **KPi** can be derived from the following sets of ordinals, defined by recursion on α:

$$
(1) \qquad C(\alpha, \beta) \ = \ \left\{ \begin{array}{l} \textit{closure of } \beta \cup \{0, \mathbf{I}\} \textit{ under}: \\ \quad +, \quad (\xi \mapsto \omega^\xi), \\ \quad (\xi \mapsto \Omega_\xi)_{\xi < \mathbf{I}}, \quad (\xi\pi \longmapsto \psi_\pi(\xi))_{\xi < \alpha} \end{array} \right.
$$

$$
(2) \qquad \psi_\pi(\alpha) \ = \ \min\{\rho : \ C(\alpha, \rho) \cap \pi = \rho \ \wedge \ \pi \in C(\alpha, \rho)\}.
$$

It is non-trivial to show that for any $\pi \in \Re$ satisfying $\pi \in C(\alpha, \pi)$, one has $\psi_\pi(\alpha) < \pi$ (see [30, 31]). Note that if $\rho = \psi_\pi(\alpha) < \pi$, then $[\rho, \pi) \cap C(\alpha, \rho) = \emptyset$, thus the order-type of the ordinals below π which belong to the Skolem hull $C(\alpha, \rho)$ is ρ . In more pictorial terms, ρ is the αth *collapse* of π.

Let $\varepsilon_{\mathbf{I}+1}$ be the least ordinal $\alpha > \mathbf{I}$ such that $\omega^\alpha = \alpha$. $C(\varepsilon_{\mathbf{I}+1}, 0)$ gives rise to an ordinal representation system, i.e., there is a primitive recursive ordinal representation system $\langle \mathcal{OR}, \lhd, \hat{\Re}, \hat{\psi}, \dots \rangle$, so that

$$(3) \qquad \langle C(\varepsilon_{\mathbf{I}+1}, 0), <, \Re, \psi, \dots \rangle \;\cong\; \langle \mathcal{OR}, \lhd, \hat{\Re}, \hat{\psi}, \dots \rangle.$$

"..." is supposed to indicate that more structure carries over to the ordinal representation system. The ordinal analysis of **KPi** (cf. [20]) establishes a close connection between **KPi** and $\mathcal{OR}$. Thus, in view of of the isomorphism $\cong$ one is led to ask how **KPi** relates to $C(\varepsilon_{\mathbf{I}+1}, 0)$. We will say that an ordinal $\alpha < \mathbf{I}$ has a *good* Σ_1*-definition in* **KPi** if there is a Σ_1-formula $\phi(u)$ such that $\mathbf{L_I} \models \phi[\alpha]$ and $\mathbf{KPi} \vdash \exists! x \phi(x)$ (see [1, II.5.13]). It turns out that all ordinals $\alpha \in C(\varepsilon_{\mathbf{I}+1}, 0)$ have a canonical good Σ_1-definition in **KPi**, arising from the definition of the Skolem hulls $C(\alpha, \beta)$ (see [30, 31, 34]), but the reverse inclusion doesn't hold. For the time being, the general relation between **KPi** and $C(\varepsilon_{\mathbf{I}+1}, 0)$ remains to be unveiled.

Returning to the main track, to allow for an interpretation of **KPi** in the infinitary calculus, we add the following rules:

$$(\neg\mathrm{Ad}) \qquad \frac{\cdots \Gamma, \mathbb{L}_\pi \neq t \cdots (\pi \in \Re;\ \pi \leq |\,t\,|)}{\Gamma, \neg \mathrm{Ad}(t)}$$

$$(\mathrm{Ad}) \qquad \frac{\Gamma, \mathbb{L}_\pi = t}{\Gamma, Ad(t)} \quad \text{if } \pi \in \Re \text{ and } \pi \leq |\,t\,|$$

$$(\mathrm{Ref}_{\Sigma(\mathbb{L}_\pi)}) \qquad \frac{\Gamma, A^{\mathbb{L}_\pi}}{\Gamma, (\exists z \in \mathbb{L}_\pi) A^z} \quad \text{if } A \in \Sigma(\mathbb{L}_\pi).$$

The rules $(\neg\mathrm{Ad})$ and (Ad) just introduce a new predicate whose meaning is the class $\{\mathbf{L}_\pi : \pi \in \Re\}$. New strength is only gained by the rules $(\mathrm{Ref}_{\Sigma(\mathbb{L}_\pi)})$.

An infinitary derivation $\mathfrak{D}$ is said to be a $C(\alpha, \beta)$–derivation if it belongs to the calculus $RS_{\mathcal{OR}}$, where $\mathcal{OR} = C(\alpha, \beta)$. Variations of the parameters α and β will allow us to convey a notion of uniformity. A $C(\alpha, \beta)$–derivation $\mathfrak{D}$ is called α-β-uniform if for any pair (α', β') such that $C(\alpha, \beta) \subseteq C(\alpha', \beta')$, $\mathfrak{D}$ can be *blown up* (or dilated) to a $C(\alpha', \beta')$–derivation $\mathfrak{D}'$. What is meant by blowing-up is best explained by the reverse process, i.e., *mutilation*. If $\mathfrak{D}$ is obtained from $\mathfrak{D}'$ by deleting in all inferences $(\forall)$, $(\notin)$, and $(\neg\mathrm{Ad})$ of $\mathfrak{D}'$ all those premises (together with the subproofs of which they are conclusions) which involve terms that do not belong to $\mathcal{OR} = C(\alpha, \beta)$, then $\mathfrak{D}$ is said to arise by mutilation from $\mathfrak{D}'$. The finite derivations of **KPi** can be unwound to yield $C(0, 0)$–uniform infinitary derivations. The whole process of (partial) cut-elimination, starting with the embedded **KPi** derivations, can then be carried out with $C(\alpha,0)$–uniform derivations, where α stays in the ordinal representation system $C(\varepsilon_{\mathbf{I}+1},0)$, though, as the cut-elimination process progresses, α propagates to ever bigger ordinals.

Hopefully, the preceding has conveyed some understanding of why the ordinal representation systems should be viewed together with their gaps.

Of course, the vocabulary we have just used is familiar from Girard's notion of β-proof and the categorical framework of dilators (see [13]). The problem of forging notions of uniformity for infinitary derivations is ubiquitous in this area of proof theory. β-proofs, however, are too weak a tool, if one aims beyond admissible proof theory. The novel feature needed is that as the cut-elimination procedure progresses we not only require the insertion of more ordinals but also of completely new reflection rules. This means that the cut-elimination process amounts to a very complicated transformation process which cannot be subsumed under the heading of β-proofs.

Fortunately, Buchholz has provided us with a very elegant and flexible setting for describing uniformity in infinitary proofs, called *operator controlled derivations* (see [3]).

DEFINITION 3.2. Let $P(On) = \{X : X \text{ is a set of ordinals}\}$. A class function $\mathcal{H} : P(On) \to P(On)$ will be called an *operator* if $\mathcal{H}$ is a *closure operator*, i.e., monotone, inclusive and idempotent, and satisfies the following conditions for all $X \in P(On)$: $0 \in \mathcal{H}(X)$, and, if α has Cantor normal form $\omega^{\alpha_1} + \cdots + \omega^{\alpha_n}$, then $\alpha \in \mathcal{H}(X) \iff \alpha_1, \ldots, \alpha_n \in \mathcal{H}(X)$. The latter ensures that $\mathcal{H}(X)$ will be closed under $+$ and $\sigma \mapsto \omega^\sigma$, and decomposition of its members into additive and multiplicative components. For $Z \in P(On)$, the operator $\mathcal{H}[Z]$ is defined by $\mathcal{H}[Z](X) := \mathcal{H}(Z \cup X)$.

If $\mathfrak{X}$ consists of "syntactic material," i.e., terms, formulae, and possibly elements from $\{0,1\}$, then let $\mathcal{H}[\mathfrak{X}](X) := \mathcal{H}(k(\mathfrak{X}) \cup X)$, where $k(\mathfrak{X})$ is the set of ordinals needed to build this "material." Finally, if s is a term, then define $\mathcal{H}[s]$ by $\mathcal{H}[\{s\}]$.

To facilitate the definition of $\mathcal{H}$–controlled derivations, we assign to each $RS_{\mathcal{OR}}$–formula A, either a (possibly infinite) disjunction $\bigvee(A_\iota)_{\iota\in I}$ or a conjunction $\bigwedge(A_\iota)_{\iota\in I}$ of $RS_{\mathcal{OR}}$–formulae. This assignment will be indicated by $A \cong \bigvee(A_\iota)_{\iota\in I}$ and $A \cong \bigwedge(A_\iota)_{\iota\in I}$, respectively. Define: $r\in t \cong \bigvee(s \mathring{\in} t \wedge r = s)_{s\in\mathcal{T}_{|t|}}$; $\mathrm{Ad}(t) \cong \bigvee(\mathbb{L}_\pi = t)_{\mathbb{L}_\pi\in I}$, where $I := \{\mathbb{L}_\eta : \eta\in\Re; \eta \le |\, t\,|\}$; $(\exists x\in t)F(x) \cong \bigvee(s\mathring{\in} t \wedge F(s))_{s\in\mathcal{T}_{|t|}}$; $A_0 \vee A_1 \cong \bigvee(A_\iota)_{\iota\in\{0,1\}}$; $\neg A \cong \bigwedge(\neg A_\iota)_{\iota\in I}$, if $A \cong \bigvee(A_\iota)_{\iota\in I}$. Using this representation of formulae, we can define the *subformulae* of a formula as follows. When $A \cong \bigwedge(A_\iota)_{\iota\in I}$ or $A \cong \bigvee(A_\iota)_{\iota\in I}$, then B is a subformula of A if $B \equiv A$ or, for some $\iota\in I$, B is a subformula of A_ι.

Since one also wants to keep track of the complexity of cuts appearing in derivations, each formula F gets assigned an ordinal rank $rk(F)$ which is the sup of the level of terms in F plus a finite number.

Using the formula representation, in spite of the many rules of $RS_{\mathcal{OR}}$, the notion of $\mathcal{H}$–controlled derivability can be defined concisely. We shall use $I \restriction \alpha$ to denote the set $\{\iota{\in}I : |\iota| < \alpha\}$.

Definition 3.3. Let $\mathcal{H}$ be an operator and let Γ be a finite set of $RS_{\mathcal{OR}}$–formulae. $\mathcal{H} \,|\!\frac{\alpha}{\rho}\, \Gamma$ is defined by recursion on α. It is always demanded that $\{\alpha\} \cup k(\Gamma) \subseteq \mathcal{H}(\emptyset)$. The inductive clauses are:

$$(\bigvee) \qquad \frac{\mathcal{H} \,|\!\frac{\alpha_0}{\rho}\, \Lambda, A_{\iota_0}}{\mathcal{H} \,|\!\frac{\alpha}{\rho}\, \Lambda, \bigvee(A_\iota)_{\iota\in I}} \qquad \begin{array}{r} \alpha_0 < \alpha \\ \iota_0{\in}I \restriction \alpha \end{array}$$

$$(\bigwedge) \qquad \frac{\mathcal{H}[\iota] \,|\!\frac{\alpha_\iota}{\rho}\, \Lambda, A_\iota \text{ for all } \iota{\in}I}{\mathcal{H} \,|\!\frac{\alpha}{\rho}\, \Lambda, \bigwedge(A_\iota)_{\iota\in I}} \qquad |\iota| \leq \alpha_\iota < \alpha$$

$$(\mathrm{Cut}) \qquad \frac{\mathcal{H} \,|\!\frac{\alpha_0}{\rho}\, \Lambda, B \qquad \mathcal{H} \,|\!\frac{\alpha_0}{\rho}\, \Lambda, \neg B}{\mathcal{H} \,|\!\frac{\alpha}{\rho}\, \Lambda} \qquad \begin{array}{r} \alpha_0 < \alpha \\ rk(B) < \rho \end{array}$$

$$(\mathrm{Ref}_{\Sigma(\mathbb{L}_\pi)}) \qquad \frac{\mathcal{H} \,|\!\frac{\alpha_0}{\rho}\, \Lambda, A^{\mathbb{L}_\pi}}{\mathcal{H} \,|\!\frac{\alpha}{\rho}\, \Lambda, (\exists z{\in}\mathbb{L}_\pi)A^z} \qquad \begin{array}{r} \alpha_0 < \alpha \\ \pi \in \Re \\ A^{\mathbb{L}_\pi}{\in}\Sigma(\mathbb{L}_\pi) \end{array}$$

The specification of the operators needed for an ordinal analysis will of course hinge upon the particular theory and ordinal representation system.

§**4. Beyond admissible proof theory.** The following gives a heuristic plan for approaching the ordinal analysis of a set theory T:

- Study the standard/canonical models of T in set-theoretic terms.
- Carry out an analysis of the impredicative principles of T.
- Determine an ordering of degrees of impredicativity.
- Distill an ordinal representation system from the previous ordering.
- Develop an infinitary proof system, based on a formal constructible hierarchy, in such a way that degrees of impredicativity are mirrored by reflection rules.
- Show (partial) cut-elimination for the proof system.

As for the degrees of impredicativity, all systems from **KP** through **KPi** have degree 1, **KPM** has degree 2, whereas the system of Π_3 reflection in [32] already has transfinitely many degrees of impredicativity.

In the following, we will gradually slice Π^1_2-comprehension into degrees of reflection. The strength of Π^1_2-comprehension is much greater than that of Δ^1_2-comprehension. In particular, there is no way to describe this comprehension simply in terms of admissibility; on the set–theoretic side, Π^1_2-comprehension corresponds to Σ_1-separation, i.e., the schema of axioms

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

for all Σ_1 formulas ϕ. Assuming *Infinity* to be among the axioms of **KP**, the precise relationship is as follows:

THEOREM 4.1. **KP** $+ \Sigma_1$-*separation and* $(\Pi^1_2 - \mathbf{CA}) + \mathbf{BI}$ *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. They are exactly the nonprojectible ordinals $> \omega$.

DEFINITION 4.2. 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 $\mathbf{L}_\rho \models \phi$ for some set–theoretic sentence ϕ (possibly containing parameters from $\mathbf{L}_\rho$), then there exists a $\gamma < \rho$ such that $\mathbf{L}_\gamma \models \phi$. This is because $\mathbf{L}_\rho \models \phi$ implies $\mathbf{L}_\kappa \models \exists\gamma\phi^{\mathbf{L}_\gamma}$, hence $\mathbf{L}_\rho \models \exists\gamma\phi^{\mathbf{L}_\gamma}$ 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 4.3. 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 4.4. $\alpha > 0$ is said to be Π_n-*reflecting* if $\mathbf{L}_\alpha \models \Pi_n$-reflection. By Π_n-reflection (Σ_n-reflection) we mean the scheme $\phi \to \exists z[Tran(z) \wedge z \neq \emptyset \wedge \phi^z]$, where ϕ is Π_n (Σ_n), and $Tran(z)$ expresses that z is a transitive set.

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

LEMMA 4.5 (cf. [33, 1.18]). $\mathbf{L}_\kappa \prec_1 \mathbf{L}_{\kappa+1}$ *iff* κ *is* Π_n*-reflecting for all n.*

With regard to refining the notion of being δ-stable, however, the constructible hierarchy causes some minor annoyance since we want the Σ_n satisfaction relation on $\mathbf{L}_\beta$ to be uniformly Σ_n-definable in $\mathbf{L}_\beta$ and likewise we want the predicate "$x = \mathbf{L}_\alpha$" to be absolute at all stages. The way out of this dilemma is to use the *Jensen hierarchy* (cf. [21]). Roughly speaking, $\mathbf{J}_\beta$ possesses all the properties of the limit levels of the $\mathbf{L}$-hierarchy. However, in the infinitary proof system we can neglect the difference between the $\mathbf{L}$ and the $\mathbf{J}$ hierarchies.

DEFINITION 4.6. κ is said to be δ-Π_n-reflecting if whenever $\phi(u, \vec{x})$ is a set–theoretic Π_n-formula, which may contain parameters from $\mathbf{J}_\kappa$, and $a_1, \dots, a_r \in \mathbf{J}_{\kappa+\delta}$ and $\mathbf{J}_{\kappa+\delta} \models \phi[\kappa, a_1, \dots, a_n]$, then there exists $\kappa_0, \delta_0 < \kappa$ and $b_1, \dots, b_r \in \mathbf{J}_{\kappa_0+\delta_0}$ such that $\mathbf{J}_{\kappa_0+\delta_0} \models \phi[\kappa_0, a_1, \dots, a_n]$ with the parameters of $\phi(u, \vec{x})$ belonging to $\mathbf{J}_{\kappa_0}$.

Putting the previous definition to work, one gets:

COROLLARY 4.7. *If κ is $\delta + 1$-Π_0-reflecting, then, for all n, κ is δ-Π_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 for $\Pi^1_2 - \mathbf{CA}$, there will be symbols $\mathfrak{E}_n$ and $\mathfrak{E}_\omega$ for ρ_n and ρ, respectively. The associated infinitary proof system $RS_{\mathcal{OR}}$ will have rules

$$(\mathrm{Ref}_{\Pi_k(\mathbb{L}_{\mathfrak{E}_n+\delta})}) \quad \frac{\Gamma, \phi(\mathbf{s})^{\mathbb{L}_{\mathfrak{E}_n+\delta}}}{\Gamma, (\exists z \in \mathbb{L}_{\mathfrak{E}_n})(\exists \mathbf{x} \in \mathbb{L}_{\mathfrak{E}_n})[Tran(z) \wedge \phi(\mathbf{x})^z]}$$

where $\phi(\mathbf{x})$ is Π_n, possibly containing $RS_{\mathcal{OR}}$-terms of levels $< \mathfrak{E}_n$, and $\delta < \mathfrak{E}_\omega$. These rules suffice to bring about the embedding of $\mathbf{KP} + \Sigma_1$-separation in 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 $RS_{\mathcal{OR}}$. Assume now that there is an infinitary proof $\mathfrak{D}$ of an arithmetic statement and that further $\mathfrak{D}$ contains δ-Π_{n+1}-reflection inferences at π. To eliminate cuts in $\mathfrak{D}$ one has to eliminate those reflection inferences, too. Let $\mathfrak{D}_0$ be a subproof of $\mathfrak{D}$ whose conclusion is the premise of a δ-Π_{n+1}-reflection inference at π. Since $\mathfrak{D}_0$ is a uniform proof, we know more than the sheer truth of its conclusion. Hopefully, this uniformity will enable one to "project" the proof tree below π. This, however, is mostly wishful thinking. To aggravate matters, assume further that $\mathfrak{D}_0$ already contains several δ-Π_{n+1}-reflection inferences at π. Well, then it is, in general, impossible to project $\mathfrak{D}_0$ below π, for otherwise we could repeat this procedure to produce an infinite decending chain of ordinals. The remedy is to replace those inferences by a hierarchy of inferences which are between δ-Π_{n+1}-reflection and δ-Π_n-reflection. This leads to the following:

DEFINITION 4.8. Let $A \subseteq \kappa \times \kappa$. κ is said to be δ-Π_n-reflecting in A if in Definition 4.6 the requirement $\langle \kappa_0, \delta_0 \rangle \in A$ is added.

LEMMA 4.9. *Still assuming that π is δ-Π_{n+1}-reflecting, suppose $\vartheta(u)$ is $\Pi_{n+1}(\mathbf{J}_\pi)$ and for every $\alpha < \pi$, $\mathbf{J}_{\pi+\delta} \models \vartheta(\alpha)$. Recursively define*

$$A^\alpha_\vartheta := \{\langle \pi_0, \delta_0 \rangle \in \pi \times \pi : \mathbf{J}_{\pi_0+\delta_0} \models \vartheta(\alpha)\ ; \\ \forall \beta < \alpha\, [\pi_0 \textit{ is } \delta_0\text{-}\Pi_n\textit{-reflecting in } A^\beta_\vartheta \cap \pi_0 \times \pi_0]\}.$$

Then, for all $\alpha < \pi$, π is δ-Π_{n+1}-reflecting in A^α_ϑ.

The proof is by induction on α. As a matter of fact, it will be important to extend the hierarchy beyond π for certain ordinals α. This hierarchy has a counterpart in the ordinal representation system. There the formula $\vartheta(\alpha)$ will refer to the αth Skolem hull of ordinals.

§5. Skolem hulls and projection functions.

It makes no sense to present an ordinal representation system without giving a semantic interpretation. For ordinal representation systems in impredicative proof theory it is essential to understand the collapsing functions which they encapsulate. In this section, the intention is to give a glimpse, in terms of Skolem hulls, of the ordinal representation system needed for $\Pi^1_2 - \mathbf{CA}$.

Let $(\mathfrak{E}_\alpha)_{\alpha<\theta}$ be a chain of Σ_1-elementary ordinals, i.e., $\mathbf{L}_\alpha \prec_1 \mathbf{L}_\beta$ holds for all $\alpha < \beta < \theta$, and let θ be a primitive recursive ordinal. θ will be kept fixed in the sequel.

DEFINITION 5.1. By recursion on α we shall define sets of ordinals $C(\alpha, \gamma)$ (the *α^{th} Skolem Hull generated from γ*), a set $\mathfrak{R}^\alpha$ of *reflection configurations* and a set $\mathfrak{F}^\alpha$ consisting of triples $(\beta, \mathbb{A}, f)$ such that $\beta \leq \alpha$, $\mathbb{A} \in \mathfrak{R}^\alpha$ and f is a partial function; such triples will be called *collapsing functions*.

Let $\mathfrak{R} = \bigcup_\alpha \mathfrak{R}^\alpha$ and $\mathfrak{F} = \bigcup_\alpha \mathfrak{F}^\alpha$. The definition of $\mathfrak{F}$ will actually guarantee that for any α and $\mathbb{A} \in \mathfrak{R}$ there is at most one f satisfying $(\alpha, \mathbb{A}, f) \in \mathfrak{F}$. This justifies the following definition:

$$\Psi^\alpha_{\mathbb{A}} = \begin{cases} f & \text{if } (\alpha, \mathbb{A}, f) \in \mathfrak{F} \text{ for some } f \\ \emptyset & \text{otherwise.} \end{cases}$$

Let's stipulate some conventions for ease of presentation. Instead of $(\alpha, \mathbb{A}, f) \in \mathfrak{F}$ we shall sloppily write $\Psi^\alpha_{\mathbb{A}} \in \mathfrak{F}$. A reflection configuration $\mathbb{A}$ is related to a certain reflection context which determines its *interval*. Formally, $\mathbb{A}$ is a formal expression built up from ordinals, symbols Π_n and parentheses. By $\mathbb{A} \in C(\alpha, \gamma)$ we mean that any ordinal occurring in $\mathbb{A}$ belongs to $C(\alpha, \gamma)$.

Let $\mathfrak{E} = \{\mathfrak{E}_\sigma : \sigma = 0 \text{ or } \sigma = \mathfrak{E}_{\xi+1} \text{ for some } \xi < \theta\}$. $C_n(\alpha, \gamma)$ is defined by recursion on n as follows:

$$
\begin{array}{lrcl}
(4) & C_0(\alpha,\gamma) & = & 1+\gamma \ \cup \ \{\mathfrak{E}_\xi : \xi < \theta\} \\
(5) & C_{n+1}(\alpha,\gamma) & = & C_n(\alpha,\gamma) \ \cup \ \{\beta : \beta =_{NF} \omega^\xi + \eta;\ \xi, \eta \in C_n(\alpha,\gamma)\} \\
 & & & \cup \ \{\Psi^\beta_{\mathbb{B}}(\eta) : \ \mathbb{B}, \beta, \eta \in C_n(\alpha,\gamma); \\
 & & & \Psi^\beta_{\mathbb{B}} \in \mathfrak{F}^{<\alpha};\ \eta \in \mathbf{dom}(\Psi^\beta_{\mathbb{B}})\} \\
(6) & C(\alpha,\gamma) & = & \bigcup_n C_n(\alpha,\gamma).
\end{array}
$$

The burden is now on us to define $\mathfrak{R}^\alpha$ and $\mathfrak{F}^\alpha$. We shall focus on one type of reflection situation. Suppose $\pi \in \mathfrak{E}$ or π is δ-Π_{n+1}-reflecting, but not stronger, and $\pi = \Psi^\beta_{\mathbb{B}}(\kappa)$, $\pi + \delta = \Psi^\beta_{\mathbb{B}}(\kappa + \zeta)$ for some $\mathbb{B}$, where $[\kappa, \kappa + \zeta]$ is the interval of the reflection configuration $\mathbb{B} \in \mathfrak{R}^{<\alpha}$. Then $\mathbb{A} \in \mathfrak{R}^\alpha$, where $\mathbb{A}$ is just a matrix that comprises the information from $\mathbb{B}$ and the data π, δ, Π_{n+1}; its interval is $[\pi, \pi + \delta]$.

The definition of $\Psi^\alpha_{\mathbb{A}}$ is the crucial part. It begins with a search for pairs of ordinals $\langle \pi_0, \delta_0 \rangle$ below π so that

$$C(\alpha, \pi_0) \cap \pi = \pi_0 \tag{7}$$

and π_0 is δ_0-Π_n-reflecting (the latter reflection property is actually way too weak). The meaning of (7) is that there is a gap in $C(\alpha, \pi_0)$ between π_0 and π. The purpose of the function $\Psi^\alpha_{\mathbb{A}}$ is to project the structure of the set $[\pi, \pi + \delta) \cap C(\alpha, \pi)$ below π, where the "structure" comprises all reflection relations belonging to $\mathfrak{R}^{<\alpha}$. Additional requirements are that $\Psi^\alpha_{\mathbb{A}}(\pi) = \pi_0$, $\Psi^\alpha_{\mathbb{A}}(\pi+\delta) = \pi_0 + \delta_0$, and $\Psi^\alpha_{\mathbb{A}}$ be defined everywhere on $C(\alpha, \pi_0) \cap [\pi, \pi+\delta]$. So the domain of $\Psi^\alpha_{\mathbb{A}}$ is contained in $[\pi, \pi + \delta]$, but, in general, it will only be a partial function on this interval. On the other hand, in general, the domain of $\Psi^\alpha_{\mathbb{A}}$ is going to be larger than $C(\alpha, \pi_0) \cap [\pi, \pi + \delta]$.

As already indicated, requiring π just to be δ_0-Π_n-reflecting is not enough. The degree of reflection should be a function of α. The precise amount of reflection involves a hierarchy, redolent of the one in Lemma 4.9.

Necessarily, in the present paper, the description of $\Psi^\alpha_{\mathbb{A}}$ has to fade out at a certain point. But it should be clear by now, that $\Psi^\alpha_{\mathbb{A}}$ will enable us to project certain proofs, which are $C(\alpha, \pi)$-uniform, below π. To conclude, let us exhibit the set of ordinals from which the ordinal representation system is obtained.

Definition 5.2. The set of ordinal representations, $\mathfrak{T}$, is inductively defined as follows:

($\mathfrak{T}$0): $0 \in \mathfrak{T}$ and for any $\eta < \theta$, $\mathfrak{E}_\eta \in \mathfrak{T}$.

($\mathfrak{T}$1): If $\eta =_{NF} \omega^\alpha + \beta$ and $\alpha, \beta \in \mathfrak{T}$, then $\eta \in \mathfrak{T}$.

($\mathfrak{T}2$): If all components of $\mathbb{A}$ are in $\mathfrak{T}$, $\alpha, \rho \in \mathfrak{T}$ and $\rho \in \mathbf{dom}(\Psi^{\alpha}_{\mathbb{A}})$, then $\Psi^{\alpha}_{\mathbb{A}}(\rho) \in \mathfrak{T}$.

THEOREM 5.3. *Different ordinal representations denote different ordinals.*

THEOREM 5.4. *Each ordinal of $\mathfrak{T}$ can be identified with the unique term denoting it, and thus $\mathfrak{T}$ can be coded as a set of natural numbers.*

Under this coding, $\mathfrak{T}$ is a primitive recursive set and, moreover, the ordering relation inherited from the ordinals is also primitive recursive.

THEOREM 5.5. *Let $\Pi^1_2-\mathbf{CA}_0$ be the system $\Pi^1_2-\mathbf{CA}$ with induction restricted to sets.*

The provably recursive ordinals of $\Pi^1_2-\mathbf{CA}_0$, $\Pi^1_2-\mathbf{CA}+\mathbf{BI}$, and $\Delta^1_3-\mathbf{CA}$ are exactly those representable in $\mathfrak{T}$ when $\theta = \omega$, $\theta = \omega + 1$, $\theta = \varepsilon_0$, respectively.

§6. A model based on indescribable cardinals. In this final section we will indicate a model for the projection functions, employing rather sweeping large cardinal axioms, in that we shall presume the existence of certain strongly indescribable cardinals. 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.

DEFINITION 6.1. Let $V = \bigcup_{\alpha \in On} V_\alpha$ be the cumulative hierarchy of sets. Let $\eta > 0$ and $\mathcal{F}$ be a collection of formulae. A cardinal κ will be called *η-$\mathcal{F}$-shrewd* if for all $P \subseteq V_\kappa$ and every $\mathcal{F}$-formula $\varphi(X, y)$, whenever

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

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

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

κ is said to be *η-$\mathcal{F}$-reducible* if for every $P \subseteq V_{\kappa+\eta}$ there exist $0 < \kappa_0, \eta_0 < \kappa$ and $Q \subseteq V_{\kappa_0+\eta_0}$ such that

$$\langle V_{\kappa_0+\eta_0}; \in; \kappa_0; Q; x\rangle_{x \in V_{\kappa_0}} \equiv_{\mathcal{F}} \langle V_{\kappa+\eta}; \in; \kappa; P; x\rangle_{x \in V_{\kappa_0}}, \tag{8}$$

where $\equiv_{\mathcal{F}}$ means that both structures satisfy the same $\mathcal{F}$ sentences. Observe that (8) implies $V_{\kappa_0} \cap Q = V_{\kappa_0} \cap P$.

κ is *shrewd* (*reducible*) if κ is η-$\mathcal{F}$-shrewd (κ is η-$\mathcal{F}$-reducible) for any $\eta > 0$ with $\mathcal{F}$ being the set-theoretic formulae.

To our knowledge the previous large cardinal notions have not been considered in the set-theoretic literature. With regard to consistency strength, shrewd cardinals are weaker than for instance subtle or ineffable cardinals; in particular is their existence compatible with the constructibility of the universe. The notions of Definition 6.1 are closely related.

LEMMA 6.2. *If κ is ρ-shrewd and $0 < \eta < \rho$, then κ is η-reducible.*

In the situation of (8) there exists a *partial* embedding p from $V_{\kappa_0+\eta_0}$ into $V_{\kappa+\eta}$ satisfying $p \restriction V_{\kappa_0+\eta_0} = \mathrm{id} \restriction V_{\kappa_0+\eta_0}$ and $p(\kappa_0) = \kappa$. Moreover, p can be canonically extended to all elements of $V_{\kappa_0+\eta_0}$ which are $\mathcal{F}$-definable in the structure $\langle V_{\kappa_0+\eta_0}; \in; \kappa_0; Q; x\rangle_{x\in V_{\kappa_0}}$.

We shall use

$$p : \langle V_{\kappa_0+\eta_0}; \in; Q\rangle \xrightarrow[\equiv]{\mathcal{F}} \langle V_{\kappa+\eta}; \in; P\rangle \tag{9}$$

to convey (8) and to indicate that p is the induced map defined on the $\mathcal{F}$-definable elements of $\langle V_{\kappa_0+\eta_0}; \in; \kappa_0; Q; x\rangle_{x\in V_{\kappa_0}}$ and p sends κ_0 to κ.

When rendered as in (9), reducibility bears a resemblance to supercompactness due to the following pullback property of supercompact cardinals:

LEMMA 6.3 (cf. [41, Theorem 5.7]). *κ is supercompact iff for every $\eta > 0$ there exist $\kappa_0, \eta_0 < \kappa$ and an elementary embedding*

$$j : \langle V_{\kappa_0+\eta_0}; \in\rangle \longrightarrow \langle V_{\kappa+\eta}; \in\rangle$$

with critical point κ_0 and $j(\kappa_0) = \kappa$.

The previous cardinal notions can be used to build a notation system by letting the collapsing functions be the inverses of partial elementary embeddings. Instead of being a chain of Σ_1-elementary ordinals, as in Definition 5.1, let $(\mathfrak{E}_\alpha)_{\alpha<\theta}$ be an enumeration of the θ first reducible cardinals. Secondly, replace the reflection notions from Definition 4.6 with the stronger ones from Definition 6.1.

Now, if for instance κ is η-Π^1_n-reducible and $P \subseteq V_{\kappa+\eta}$, then there is "reduction"

$$p : \langle V_{\kappa_0+\eta_0}; \in; Q\rangle \xrightarrow[\equiv]{\Pi^1_n} \langle V_{\kappa+\eta}; \in; P\rangle$$

which insures that the Π^1_n-definable points of the interval $[\kappa, \kappa+\eta]$ in the structure $\langle V_{\kappa+\eta}; \in; P\rangle$ get *collapsed down* to the Π^1_n-definable points of $[\kappa_0, \kappa_0+\eta_0]$ in the structure $\langle V_{\kappa_0+\eta_0}; \in; Q\rangle$ by the partial function

$$f = p^{-1} \restriction [\kappa, \kappa+\eta) \cup \langle \kappa+\eta, \kappa_0+\eta_0\rangle.$$

The following table gives an indication of the analogies between the notions used in §4 and large cardinal notions :

	Reflecting Ordinals	Cardinals
	admissible $> \omega$	regular $> \omega$
	recursively inaccessible	inaccessible
$n \geq 1$	0-Π_{n+2}-reflecting	Π^1_n-indescribable
$\delta, n > 0$	δ-Π_n-reflecting	δ-Π^1_n-reducible

The cardinal analogue that corresponds to the intermediate notions of reflection considered in Definition 4.8 is the following:

Definition 6.4. Suppose $A \subseteq \kappa \times \kappa$. If $\eta > 0$, κ is said to be *η-$\mathcal{F}$-reducible in A* if for every $P \subseteq V_{\kappa+\eta}$ there exist $\langle \kappa_0, \eta_0 \rangle \in A$ with $0 < \eta_0, \kappa_0$ and a $Q \subseteq V_{\kappa_0+\eta_0}$ such that

$$(10) \quad \langle V_{\kappa_0+\eta_0}; \in; \mathfrak{V}; \kappa_0; Q; x\rangle_{x \in V_{\kappa_0}} \equiv_{\mathcal{F}} \langle V_{\kappa+\eta}; \in; \mathfrak{V}; \kappa; P; x\rangle_{x \in V_{\kappa_0}}.$$

REFERENCES

[1] J. Barwise, ***Admissible sets and structures***, Springer-Verlag, Berlin, 1975.

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

[3] ———, *A simplified version of local predicativity*, ***Leeds proof theory 1991***, Cambridge University Press, Cambridge, 1993, pp. 115–147.

[4] W. Buchholz, S. Feferman, W. Pohlers, and W. Sieg, ***Iterated inductive definitions and subsystems of analysis***, Springer-Verlag, Berlin, 1981.

[5] W. Buchholz and K. Schütte, ***Proof theory of impredicative subsystems of analysis***, Bibliopolis, Naples, 1988.

[6] S. Feferman, *Systems of predicative analysis*, ***Journal of Symbolic Logic***, vol. 29 (1964), pp. 1–30.

[7] ———, *A language and axioms for explicit mathematics*, ***Lecture notes in mathematics 450***, Springer-Verlag, Berlin, 1975, pp. 87–139.

[8] ———, *Constructive theories of functions and classes*, ***Logic colloquium '78***, North-Holland, Amsterdam, 1979, pp. 159–224.

[9] ———, *Proof theory: a personal report*, ***Proof theory*** (G. Takeuti, editor), North-Holland, Amsterdam, second ed., 1987, pp. 445–485.

[10] ———, *Hilbert's program relativized: Proof-theoretical and foundational reductions*, ***Journal of Symbolic Logic***, vol. 53 (1988), pp. 364–384.

[11] ———, *Remarks for "the trends in logic"*, ***Logic colloquium '88***, North-Holland, Amsterdam, 1989, pp. 361–363.

[12] H. Friedman, N. Robertson, and P. Seymour, *The metamathematics of the graph minor theorem*, ***Contemporary Mathematics***, vol. 65 (1987), pp. 229–261.

[13] J.-Y. Girard, *Introduction to Π^1_2–logic*, ***Synthese***, vol. 62 (1985), pp. 191–216.

[14] K. Gödel, *The modern development of the foundations of mathematics in the light of philosophy*, ***K. Gödel, Collected works, vol. III***, Oxford University Press, Oxford, 1995, pp. 374–387.

[15] E. Griffor and M. Rathjen, *The strength of some Martin-Löf type theories*, ***Archive for Mathematical Logic***, vol. 33 (1994), pp. 347–385.

[16] D. Hilbert, *Die Grundlegung der elementaren Zahlentheorie*, ***Mathematische Annalen***, vol. 104 (1931).

[17] G. Jäger, *Beweistheorie von KPN*, ***Archiv für Mathematische Logik und Grundlagenforschung***, vol. 20 (1980), pp. 53–64.

[18] ———, *Zur Beweistheorie der Kripke-Platek Mengenlehre über den natürlichen Zahlen*, ***Archiv für Mathematische Logik und Grundlagenforschung***, vol. 22 (1982), pp. 121–139.

[19] ———, *A well-ordering proof for Feferman's theory T_0*, ***Archiv für Mathematische Logik und Grundlagenforschung***, vol. 23 (1983), pp. 65–77.

[20] G. Jäger and W. Pohlers, ***Eine beweistheoretische Untersuchung von Δ^1_2– CA + BI und verwandter Systeme***, Sitzungsberichte der Bayerischen Akademie der Wissenschaften, Mathematisch–Naturwissenschaftliche Klasse, 1982.

[21] R. B. Jensen, *The fine structure of the constructible hierarchy*, ***Annals of Math. Logic***, vol. 4 (1972), pp. 229–308.

[22] D. Martin and J. Steel, *Projective determinacy*, ***Proceedings of the National Academy of Sciences of the United States of America***, vol. 85, 1988, pp. 6582–6586.

[23] P. Martin-Löf, ***Intuitionistic type theory***, Bibliopolis, Naples, 1984.

[24] W. Pohlers, *Proof-theoretical analysis of ID_ν by the method of local predicativity*, in [4], Springer, Berlin, 1981, pp. 261–357.

[25] ———, *Cut elimination for impredicative infinitary systems, part I: Ordinal analysis of ID_1*, ***Archiv für Mathematische Logik und Grundlagenforschung***, vol. 21 (1981), pp. 69–87.

[26] ———, *Cut elimination for impredicative infinitary systems, part II: Ordinal analysis for iterated inductive definitions*, ***Archiv für Mathematische Logik und Grundlagenforschung***, vol. 22 (1982), pp. 113–129.

[27] ———, *Proof theory and ordinal analysis*, ***Archive for Mathematical Logic***, vol. 30 (1991), pp. 311–376.

[28] M. Rathjen, *An ordinal analysis of Π^1_2 comprehension and related systems*, in preparation.

[29] ———, *Proof-theoretic analysis of KPM*, ***Archive for Mathematical Logic***, vol. 30 (1991), pp. 377–403.

[30] ———, *How to develop proof–theoretic ordinal functions on the basis of admissible sets*, ***Mathematical Quarterly***, vol. 39 (1993), pp. 47–54.

[31] ———, *Collapsing functions based on recursively large ordinals: A well–ordering proof for KPM*, ***Archive for Mathematical Logic***, vol. 33 (1994), pp. 35–55.

[32] ———, *Proof theory of reflection*, ***Annals of Pure and Applied Logic***, vol. 68 (1994), pp. 181–224.

[33] W. Richter and P. Aczel, *Inductive definitions and reflecting properties of admissible ordinals*, ***Generalized recursion theory***, North-Holland, Amsterdam, 1973, pp. 301–381.

[34] A. Schlüter, *Provability in set theories with reflection*, submitted.

[35] K. Schütte, *Beweistheoretische Erfassung der unendlichen Induktion in der Zahlentheorie*, ***Mathematische Annalen***, vol. 122 (1951), pp. 369–389.

[36] ———, ***Beweistheorie***, Springer-Verlag, Berlin, 1960.

[37] ———, *Eine Grenze für die Beweisbarkeit der transfiniten Induktion in der verzweigten Typenlogik*, ***Archiv für Mathematische Logik und Grundlagenforschung***, vol. 67 (1964), pp. 45–60.

[38] ———, *Predicative well-orderings*, ***Formal systems and recursive functions***, North-Holland, 1965, pp. 176–184.

[39] T. Setzer, *Proof theoretical strength of Martin-löf type theory with w-type and one universe*, ***Thesis***, University of Munich, 1993.

[40] S. Simpson, *Nichtbeweisbarkeit von gewissen kombinatorischen Eigenschaften endlicher Bäume*, ***Archiv für Mathematische Logik und Grundlagenforschung***, vol. 25 (1985), pp. 45–65.

[41] R. M. Solovay, W. N. Reinhardt, and A. Kanamori, *Strong axioms of infinity and elementary embeddings*, ***Annals of Mathematical Logic***, vol. 13 (1978), pp. 73–116.

[42] G. Takeuti, *Consistency proofs of subsystems of classical analysis*, ***Annals of Mathematics***, vol. 86 (1967), pp. 299–348.

[43] ———, *Proof theory and set theory*, ***Synthese***, vol. 62 (1985), pp. 255–263.

[44] G. Takeuti and M. Yasugi, *The ordinals of the systems of second order arithmetic with the provably Δ^1_2–comprehension and the Δ^1_2–comprehension axiom respectively*, ***Japanese Journal of Mathematics***, vol. 41 (1973), pp. 1–67.

[45] H. Woodin, *Large cardinal axioms and independence: The continuum problem revisited*, ***Mathematical Intelligencer***, 1994, pp. 31–35.

DEPARTMENT OF MATHEMATICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305, USA

E-mail: rathjen@math.stanford.edu