Operations, sets and classes

Gerhard Jäger

Institut für Informatik und angewandte Mathematik
Universität Bern, CH-3012 Bern, Switzerland
jaeger@iam.unibe.ch

Abstract

Operational set theory, in the form described below, is an enterprise which consolidates classical set theory with some central concepts of Feferman's *explicit mathematics*. It provides for a careful distinction between operations and set-theoretic functions and as such reconciles set theory with needs arising in constructive environments and even in those enhanced by computer science.

In the following we consider, primarily from a proof-theoretic perspective, the theory OST and some of its most important extensions and determine their consistency strengths by exhibiting equivalent systems in the realm of traditional theories of sets and classes.

Keywords: Operational set theory, explicit mathematics, proof theory, classical and constructive set theories

1 Introduction

Operational set theory, in the form described below, is a comparatively young enterprise which consolidates classical set theory with some central concepts of Feferman's *explicit mathematics*. It provides for a careful distinction between operations and set-theoretic functions and as such reconciles set theory with needs arising in constructive environments and even in those oriented towards computer science.

The general topic of explicit mathematics originated in Feferman's seminal paper [14], where several formal systems, including the famous theory T_0, where introduced. The original aim of explicit mathematics was to provide an appropriate framework for Bishop-style constructive mathematics, and it can be seen as one specific effort in parallel to rather different work by others; see Feferman [16] for a thorough discussion of this aspect.

However, soon it turned out that explicit mathematics also plays an important role in reductive proof theory and as an axiomatic approach to abstract computability. In Buchholz, Feferman, Pohlers and Sieg [11] important subsystems of T_0 are related to subsystems of second order arithmetic; the question concerning the exact proof-theoretic strength of full T_0 is settled by Jäger [26] and Jäger and Pohlers [33].

Feferman [15] lays the foundations for later work about the connections between explicit mathematics and generalized recursion theory and presents some first important results. In Feferman and Jäger [21] and Jäger and Strahm [35] the proof theory of the non-constructive μ-operator and the Suslin operator in an explicit context are studied; Jäger and Strahm [34, 36] deal with various forms of explicit reflections, in particular with Mahloness and analogues of Π_3 reflection.

The proofs of these and many other results about explicit mathematics make heavy use of interesting set-theoretic concepts, at least implicitly. Very often (the intuitive background of) a proof-theoretic argument depends on a subtle interplay between notions in classical set theory and their admissible, constructive or recursive analogues. Operational set theory turns these implicit analogies into explicit generalizations. A central ingredient of this approach is a strict distinction between operations, which may be interpreted as computations or even programs, and functions in the set-theoretic sense, i.e. binary right-unique relations.

Feferman [18] is the starting points of the following considerations and introduces the system OST of operational set theory and a few extensions, motivated by the aim to develop a common language for small large cardinal notions as in classical set theory, admissible set and recursion theory. Feferman [19] presents variants of these systems closer in syntax to original explicit mathematics, and Feferman [20] is a polished up version of parts of [18].

Related work by Cantini and Crosilla [12] is about a constructive set theory with operations COST, which may be considered as a constructive version of OST, and may be regarded as providing a bridge between Aczel's constructive set theory CZF, see Aczel [1, 2, 3], and explicit mathematics. As predecessors of present day operational set theory we may consider Beeson [10], presenting an interesting computation system based on set theory and formulated as a theory of sets and rules, and Feferman [17], where some of the central ideas are outlined.

The present article begins with a rough description of the landscape of set theory and then studies OST and its extensions by operational power set and operational unbounded existential quantification. Afterwards we determine their consistency strengths by exhibiting equivalent systems in the realm of traditional set theory and describe an interesting extension of OST which is conservative over ZFC.

2 Aspects of the set-theoretic landscape

No doubt, Zermelo-Fraenkel set theory with the axiom of choice is the generally accepted framework for wide parts of everyday mathematics. From a foundational perspective, however, the general picture is much richer, and – very roughly – we can distinguish three areas:

I. Classical set theories. Clearly, ZF and ZFC are the two most prominent exponents. But there is also von Neumann-Bernays-Gödel set theory NBG, a theory of sets and classes which yields the same results as ZFC. Morse-Kelley set theory MK is a significant strengthening of NBG permitting highly non-elementary class formation. Further natural strengthenings of NBG are obtained by adding reflection principles such as Π^1_1 reflection or strict Π^1_1 reflection.

On the other hand, theories like ZF, ZFC and their extensions have significant drawbacks from a logical perspective. To mention only a few: (i) They heavily violate the principle of parsimony; most mathematical theorems can be proved from much weaker set existence axioms. (ii) They demand that all mathematical objects (also, for example, computer programs) are realized as sets. (iii) They have only very huge models and, for example, no recursive models. (iv) They do not differentiate between levels of existence and between constructively/recursively and classically valid assertions.

II. Constructive set and type theories. A first possible reaction is to replace classical set theories by their constructive variants. Prominent example of those are:

- Myhill's CST (cf. [42]) and the intuitionistic version IZF of ZF.
- Martin-Löf type theories (cf. [39, 40]).
- Aczel's constructive set theory CZF and its extensions à la Aczel and Rathjen (cf. [1, 2, 3, 4]).
- Proof development systems and proof assistants such as Coq, HOL and Nuprl (cf. [13, 22, 43]).

Most of these systems use intuitionistic logic and have the disjunction and existence property.

III. Admissible set theories. An alternative is to abide with classical logic, but to weaken the set existence axioms dramatically. Most distinctive along these lines is the system $\mathsf{KP}\omega$ of Kripke-Platek set theory with infinity. In calibrating the proof-theoretic strength of subsystems of second order arithmetic and set theory, theories of iterated admissible sets play an important role; for example:

- KPu and KPi for admissible and recursively inaccessible universes (cf. [23, 25, 24, 28]).

- KPm and $\mathsf{KP}\omega + (\Pi_n\text{-}\mathsf{Ref})$ for recursively Mahlo universes and universes satisfying Π_n reflection (cf. [44, 45, 5, 6]).

- admissibility without foundation, i.e. systems like KPi^0, KPm^0 and $\mathsf{KPu}^0 + (\Pi_n\text{-}\mathsf{Ref})$ which are obtained from the above theories by dropping $\in$-induction; they represent the wide spectrum of predicative and metapredicative theories (cf. [27, 28, 31, 34, 36, 49, 50]).

Recent proof-theoretic work of Rathjen (cf.,e.g., [46, 47]) about extensions of Kripke-Platek set theory leading to $\Pi^1_2\text{-}\mathsf{CA}$ may be subsumed under classical and admissible set theories.

It has already been mentioned in the introduction that there exist interesting connections between notions of classical set theory and their recursive or admissible analogues. A typical example is the notion of regular cardinal which collapses to admissible ordinal if we replace arbitrary set-theoretic functions by recursive functions; see Richter and Aczel [48] for more on this from a recursion-theoretic perspective.

Operational set theory is the approach to isolate the basic common principle underlying each of the areas I –III. The central notion is that of operation which may be interpreted classically, constructively or recursively. In this regard it is very much like Feferman's marriage of convenience for explicit mathematics [15].

3 The theory OST and its relatives

The presentation of the theory OST and its extensions follows Jäger [29, 30]; all unexplained notions and further motivation can be found there.

Let $\mathcal{L}_1$ be a typical language of admissible or classical set theory with a symbol for the element relation as its only relation symbol and countably many set variables $a, b, c, f, g, u, v, w, x, y, z, \ldots$ (possibly with subscripts). The formulas of $\mathcal{L}_1$ are defined as usual.

$\mathcal{L}^\circ$, the language of OST and its extensions, augments $\mathcal{L}_1$ by the binary function symbol $\circ$ for partial term application, the unary relation symbol $\downarrow$ (defined) and the following constants: (i) the combinators k and s; (ii) $\top$, $\bot$, $\mathbf{el}$, $\mathbf{non}$, $\mathbf{dis}$, $\mathbf{e}$ and $\mathbf{E}$ for logical operations; (iii) $\mathbb{S}$, $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{P}$ for set-theoretic operations. The meaning of these constants follows from the axioms below.

The *terms* $(r, s, t, r_1, s_1, t_1, \ldots)$ of $\mathcal{L}^\circ$ are inductively generated as follows:

1. The variables and constants of $\mathcal{L}_1$ are terms of $\mathcal{L}^\circ$.

2. If s and t are terms of $\mathcal{L}^\circ$, then so is $\circ(s, t)$.

In the following we often abbreviate $\circ(s, t)$ as $(s \circ t)$, as (st) or – if no confusion arises – simply as st. We also adopt the convention of association to the left so that $s_1 s_2 \ldots s_n$ stands for $(\ldots(s_1 s_2)\ldots s_n)$. In addition, we often write $s(t_1, \ldots, t_n)$ for $st_1 \ldots t_n$ if this seems more intuitive. Moreover, we frequently make use of The vector notation $\vec{s}$ is used as shorthand for a finite string $s_1, \ldots, s_n$ of $\mathcal{L}^\circ$ terms whose length is either not important or evident from the context.

Self-application is possible and meaningful, but it is not necessarily total; there may be terms which do not denote an object. We make use of the definedness predicate $\downarrow$ to single out those which do, and $(t\downarrow)$ is read "t is defined" or "t has a value".

The formulas $(A, B, C, D, A_1, B_1, C_1, D_1, \ldots)$ of $\mathcal{L}^\circ$ are inductively generated as follows:

1. All expressions of the form $(s \in t)$ and $(t\downarrow)$ are formulas of $\mathcal{L}^\circ$; the so-called *atomic* formulas.

2. If A and B are formulas of $\mathcal{L}^\circ$, then so are $\neg A$, $(A \vee B)$ and $(A \wedge B)$.

3. If A is a formula and t a term of $\mathcal{L}^\circ$ which does not contain x, then $(\exists x \in t)A$, $(\forall x \in t)A$, $\exists x A$ and $\forall x A$ are formulas of $\mathcal{L}^\circ$.

We will often omit parentheses and brackets whenever there is no danger of confusion. Since we will be working within classical logic, the remaining logical connectives can be defined as usual; equality of sets is introduced by

$$(s = t) \;:=\; (s\downarrow) \wedge (t\downarrow) \wedge (\forall x \in s)(x \in t) \wedge (\forall x \in t)(x \in s).$$

The free variables of t and A are defined in the conventional way; the closed $\mathcal{L}^\circ$ terms and closed $\mathcal{L}^\circ$ formulas, also called $\mathcal{L}^\circ$ sentences, are those which do not contain free variables.

Given an $\mathcal{L}^\circ$ formula A and a variable u not occurring in A, we write A^u for the result of replacing each unbounded set quantifier $\exists x(\ldots)$ and $\forall x(\ldots)$ in A by $(\exists x \in u)(\ldots)$ and $(\forall x \in u)(\ldots)$, respectively. Suppose now that $\vec{u} = u_1, \ldots, u_n$ and $\vec{s} = s_1, \ldots, s_n$. Then $A[\vec{s}/\vec{u}]$ is the $\mathcal{L}^\circ$ formula which is obtained from A by simultaneously replacing all free occurrences of the variables $\vec{u}$ by the $\mathcal{L}^\circ$ terms $\vec{s}$; in order to avoid collision of variables, a renaming of bound variables may be necessary. If the $\mathcal{L}^\circ$ formula A is written as $B[\vec{u}]$, then we often simply write $B[\vec{s}]$ instead of $B[\vec{s}/\vec{u}]$. Further variants of this notation will be obvious.

The logic of OST is the classical *logic of partial terms* due to Beeson [8, 9], including the common equality axioms. Partial equality of terms is introduced by

$$(s \simeq t) \;:=\; (s\downarrow \vee\, t\downarrow \;\rightarrow\; s = t)$$

and says that if either s or t denotes anything, then they both denote the same object.

The non-logical axioms of OST comprise axioms about the applicative structure of the universe, some basic set-theoretic properties, the representation of elementary logical connectives as operations and operational set existence axioms. They divide into four groups.

I. Applicative axioms.

(1) $\mathsf{k} \neq \mathsf{s}$,

(2) $\mathsf{k}xy = x$,

(3) $\mathsf{s}xy\downarrow \wedge \mathsf{s}xyz \simeq (xz)(yz)$.

Thus the universe is a partial combinatory algebra. We have λ-abstraction and thus can introduce for each $\mathcal{L}^\circ$ term t a term $(\lambda x.t)$ whose variables are those of t other than x such that

$$(\lambda x.t)\downarrow \ \wedge \ (\lambda x.t)y \simeq t[y/x].$$

As usual we can generalize λ abstraction to several arguments by simply iterating abstraction for one argument. Accordingly, we set for all $\mathcal{L}^\circ$ terms t and all variables $x_1, \ldots, x_n$,

$$(\lambda x_1 \ldots x_n.t) := (\lambda x_1.(\ldots(\lambda x_n.t)\ldots)).$$

Often the term $(\lambda x_1 \ldots x_n.t)$ is also simply written as $\lambda x_1 \ldots x_n.t$. If $\vec{x}$ is the sequence $x_1, \ldots, x_n$, then $\lambda\vec{x}.t$ stands for $\lambda x_1 \ldots x_n.t$ and $t\vec{x}$ for $tx_1 \ldots x_n$.

Furthermore, there exists a closed $\mathcal{L}^\circ$ term fix, a so-called fixed point operator, with

$$\mathsf{fix}(f)\downarrow \ \wedge \ (\mathsf{fix}(f) = g \ \rightarrow \ gx \simeq f(g, x)).$$

II. Basic set-theoretic axioms. They state that: (i) there is the empty set; (ii) there are unordered pairs and unions; (iii) there exists an infinite ordinal; (iv) $\in$-induction is available for arbitrary formulas $A[x]$ of $\mathcal{L}^\circ$,

$$(\mathcal{L}^\circ\text{-}\mathsf{I}_\in) \qquad \forall x((\forall y \in x)A[y] \rightarrow A[x]) \ \rightarrow \ \forall x A[x].$$

To increase readability, we will freely use standard set-theoretic terminology; also, if $A[x]$ is an $\mathcal{L}^\circ$ formula, then $\{x : A[x]\}$ denotes the collection of all sets satisfying A; it may be (extensionally equal to) a set, but this is not necessarily the case. In particular, we set

$$\mathbb{B} \ := \ \{x : x = \top \vee x = \bot\} \quad \text{and} \quad \mathbb{V} \ := \ \{x : x\downarrow\}$$

so that $\mathbb{B}$ stands for the unordered pair consisting of the truth values $\top$ and $\bot$, which is a set by the previous axioms. $\mathbb{V}$ is the collection of all sets but not a set itself. The following shorthand notations, for n an arbitrary natural number,

$$(f : a \rightarrow b) \quad := \quad (\forall x \in a)(fx \in b),$$

$$(f : a^{n+1} \rightarrow b) \quad := \quad (\forall x_1, \ldots, x_{n+1} \in a)(f(x_1, \ldots, x_{n+1}) \in b)$$

express that f, in the operational sense, is a unary and $(n{+}1)$-ary mapping from a to b, respectively. They do not say, however, that f is a unary or $(n{+}1)$-ary function in the set-theoretic sense.

In the previous definition the set variables a and/or b may be replaced by $\mathbb{V}$ and/or $\mathbb{B}$. So, for example, $(f : a \to \mathbb{V})$ means that f is an operation which is total on a, and $(f : \mathbb{V} \to b)$ means that f maps all sets into b. If we have $(f : a \to \mathbb{B})$, we may regard f as a *definite predicate* on a; if we have $(f : \mathbb{V} \to \mathbb{B})$, we call f a *total characteristic operation.*

III. Logical operations axioms.

(L1) $\top \neq \bot$.

(L2) $(\mathbf{el} : \mathbb{V}^2 \to \mathbb{B}) \;\wedge\; \forall x \forall y(\mathbf{el}(x,y) = \top \;\leftrightarrow\; x \in y)$.

(L3) $(\mathbf{non} : \mathbb{B} \to \mathbb{B}) \;\wedge\; (\forall x \in \mathbb{B})(\mathbf{non}(x) = \top \;\leftrightarrow\; x = \bot)$.

(L4) $(\mathbf{dis} : \mathbb{B}^2 \to \mathbb{B}) \;\wedge\; (\forall x, y \in \mathbb{B})(\mathbf{dis}(x,y) = \top \;\leftrightarrow\; (x = \top \,\vee\, y = \top))$.

(L5) $(f : a \to \mathbb{B}) \;\to\; (\mathbf{e}(f,a) \in \mathbb{B} \;\wedge\; (\mathbf{e}(f,a) = \top \;\leftrightarrow\; (\exists x \in a)(fx = \top)))$.

The Δ_0 formulas of $\mathcal{L}^\circ$ are those $\mathcal{L}^\circ$ formulas which do not contain the function symbol $\circ$, the relation symbols $\downarrow$ or unbounded quantifiers. Hence they are the usual Δ_0 formulas of set theory, possibly containing additional constants. The above logical operations make it possible to represent all Δ_0 formulas by constant $\mathcal{L}^\circ$ terms. For a proof of the following see Feferman [18, 20].

Lemma 1 *Let $\vec{u}$ be the sequence of variables $u_1, \ldots, u_n$. For every Δ_0 formula $A[\vec{u}]$ of $\mathcal{L}^\circ$ with at most the variables $\vec{u}$ free, there exists a closed $\mathcal{L}^\circ$ term t_A such that the axioms introduced so far yield*

$$t_A{\downarrow} \;\wedge\; (t_A : \mathbb{V}^n \to \mathbb{B}) \;\wedge\; \forall \vec{x}(A[\vec{x}] \;\leftrightarrow\; t_A(\vec{x}) = \top).$$

IV. Operational set-theoretic axioms.

(S1) Separation for definite operations:

$$(f : a \to \mathbb{B}) \;\to\; (\mathbb{S}(f,a){\downarrow} \;\wedge\; \forall x(x \in \mathbb{S}(f,a) \;\leftrightarrow\; (x \in a \;\wedge\; fx = \top))).$$

(S2) Replacement:

$$(f : a \to \mathbb{V}) \;\to\; (\mathbb{R}(f,a){\downarrow} \;\wedge\; \forall x(x \in \mathbb{R}(f,a) \;\leftrightarrow\; (\exists y \in a)(x = fy))).$$

(S3) Choice:

$$\exists x(fx = \top) \;\to\; (\mathbb{C}f{\downarrow} \;\wedge\; f(\mathbb{C}f) = \top).$$

This finishes the description of the non-logical axioms of OST. A significant strengthening $\mathsf{OST}(\mathbb{P})$ of OST is obtained by adding the operational form of the power set axiom

$$(\mathbb{P}) \qquad (\mathbb{P} : \mathbb{V} \to \mathbb{V}) \;\wedge\; \forall x \forall y (x \in \mathbb{P}y \;\leftrightarrow\; x \subset y).$$

Note that in OST and $\mathsf{OST}(\mathbb{P})$ we cannot treat unbounded existential quantification operationally. For that we use the constant $\mathbf{E}$ and the additional axiom

$$(\mathbf{E}) \qquad (f : \mathbb{V} \to \mathbb{B}) \;\to\; (\mathbf{E}(f) \in \mathbb{B} \wedge (\mathbf{E}(f) = \top \;\leftrightarrow\; \exists x (fx = \top))).$$

In the following we write $\mathsf{OST}(\mathbf{E}, \mathbb{P})$ for $\mathsf{OST}(\mathbb{P}) + (\mathbf{E})$. This is the strongest operational theory we consider in this article.

Call those formulas of $\mathcal{L}^\circ$ which do not contain the function symbol $\circ$ or the relation symbol $\downarrow$ *pure formulas* of $\mathcal{L}^\circ$; they are the same as those of $\mathcal{L}_1$ plus the constants of OST. Then $(\mathbf{E})$ permits the extension of Lemma 1 to pure formulas.

Lemma 2 *Let $\vec{u}$ be the sequence of variables $u_1, \ldots, u_n$. For every pure formula $A[\vec{u}]$ of $\mathcal{L}^\circ$ with at most the variables $\vec{u}$ free, there exists a closed $\mathcal{L}^\circ$ term t_A such that* $\mathsf{OST}(\mathbf{E}, \mathbb{P})$ *proves*

$$t_A\downarrow \;\wedge\; (t_A : \mathbb{V}^n \to \mathbb{B}) \;\wedge\; \forall \vec{x} (A[\vec{x}] \;\leftrightarrow\; t_A(\vec{x}) = \top).$$

Its proof is analogous to the proof of Lemma 1; simply use $(\mathbf{E})$ to deal with unbounded quantifiers.

From Feferman [18] and Jäger [29] we know that, provably in the systems OST, there exist closed $\mathcal{L}^\circ$ terms $\emptyset$ for the empty set, **uopa** for forming unordered pairs, **un** for forming unions, **p** for forming ordered pairs (Kuratowski pairs) and **prod** for forming Cartesian products. In addition, there are closed $\mathcal{L}^\circ$ terms $\mathbf{p}_L$ and $\mathbf{p}_R$ which act as projections with respect to $\mathbf{p}$, i.e.

$$\mathbf{p}_L(\mathbf{p}(a,b)) = a \quad \text{and} \quad \mathbf{p}_R(\mathbf{p}(a,b)) = b.$$

To comply with the set-theoretic conventions, we generally write $\{a, b\}$ instead of $\mathbf{uopa}(a,b)$, $\cup a$ instead of $\mathbf{un}(a)$, $\langle a, b\rangle$ instead of $\mathbf{p}(a,b)$ and $a \times b$ instead of $\mathbf{prod}(a,b)$. Remember that ω is a constant for the first infinite ordinal and belongs to the base language $\mathcal{L}_1$. OST is also fairly strong with respect to definition by cases.

Lemma 3 *There exist closed $\mathcal{L}^\circ$ terms $\mathsf{d}_=$, $\mathsf{d}_\emptyset$ and $\mathsf{d}_\mathbb{B}$ such that* OST *proves:*

1. $(u = v \;\wedge\; \mathsf{d}_=(a,b,u,v) = a) \;\vee\; (u \neq v \;\wedge\; \mathsf{d}_=(a,b,u,v) = b)$.
2. $(u = \emptyset \;\wedge\; \mathsf{d}_\emptyset(a,b,u) = a) \;\vee\; (u \neq \emptyset \;\wedge\; \mathsf{d}_\emptyset(a,b,u) = b)$.
3. $\mathsf{d}_\mathbb{B}(a,b,\top) = a \;\wedge\; \mathsf{d}_\mathbb{B}(a,b,\bot) = b$.

We end this section with an interesting aspect of our axiom about operational choice: it provides for a form of global choice and as such is of some relevance in connection with von Neumann-Bernays-Gödel set theory NBG to be introduced later.

Theorem 4 (Global choice) *There exists a closed $\mathcal{L}^\circ$ term* choice *such that* OST *proves*

$$(\mathsf{choice} : \mathbb{V} \to \mathbb{V}) \;\wedge\; \forall x(x \neq \emptyset \;\to\; \mathsf{choice}(x) \in x) \;\wedge\; \mathsf{choice}(\emptyset) = \top.$$

The above results about definition by cases and this theorem about global choice are proved in Jäger [30]; see also Feferman [20] where it is pointed out that (AC) is provable in OST, but the proof actually demonstrates our statement of global choice.

Recall the system $\mathsf{KP}\omega$ of Kripke-Platek set theory with infinity. It is formulated in $\mathcal{L}_1$ and based on classical first order predicate calculus with equality. Its non-logical axioms are pair, union, infinity, Δ_0 separation, Δ_0 collection and $\in$-induction for arbitrary formulas of $\mathcal{L}_1$. We write (AC) for the axiom of choice and $(V{=}L)$ for the axiom of constructibility. As well-known from the literature, $\mathsf{KP}\omega$, $\mathsf{KP}\omega + (\mathsf{AC})$ and $\mathsf{KP}\omega + (V{=}L)$ are of the same consistency strength, and $\mathsf{KP}\omega + (V{=}L)$ is conservative over $\mathsf{KP}\omega$ for absolute formulas.

$\mathcal{L}_1(\mathcal{P})$ is first order language obtained from $\mathcal{L}_1$ by adding a new binary relation symbol $\mathcal{P}$. The formulas of $\mathcal{L}_1(\mathcal{P})$ are defined as the formulas of $\mathcal{L}_1$, but with expressions of the form $\mathcal{P}(r,s)$ permitted as additional atomic formulas. The $\Delta_0(\mathcal{P})$ formulas are those formulas of $\mathcal{L}_1(\mathcal{P})$ which do not contain unbounded quantifier; in particular, each $\mathcal{P}(r,s)$ is $\Delta_0(\mathcal{P})$.

The extension $\mathsf{KP}(\mathcal{P})$ of $\mathsf{KP}\omega$ is formulated in $\mathcal{L}_1(\mathcal{P})$ and characterized by: (i) it encompasses pair, union, infinity $\Delta_0(\mathcal{P})$ separation and $\Delta_0(\mathcal{P})$ collection; (ii) $\in$-induction is formulated for arbitrary $\mathcal{L}_1(\mathcal{P})$ formulas; (iii) the new axiom $(\mathcal{P})$ provides the meaning of the relation symbol $\mathcal{P}$,

$$(\mathcal{P}) \qquad \forall x \exists y \mathcal{P}(x,y) \;\wedge\; \forall x \forall y(\mathcal{P}(x,y) \;\leftrightarrow\; \forall z(z \in y \;\leftrightarrow\; z \subset x)).$$

The following theorem, which is proved in Feferman [20] and Jäger [29], is not surprising and establishes a lower bound for the proof-theoretic strength of OST.

Theorem 5 1. *The theory* $\mathsf{KP}\omega + (\mathsf{AC})$ *is contained in* OST.

2. *The theory* $\mathsf{KP}(\mathcal{P}) + (\mathsf{AC})$ *is contained in* $\mathsf{OST}(\mathbb{P})$.

Our next goal is to look for upper proof-theoretic bounds, and the method to find those is to model OST and $\mathsf{OST}(\mathbb{P})$.

4 Modeling OST and $\mathsf{OST}(\mathbb{P})$

There are two principal ways for constructing models of OST: one goes back to Feferman [18, 20] and uses ideas from generalized recursion theory, the other is presented in detail in Jäger [29] and is based on interpreting the application operation of OST via a suitably defined (nonmonotonic) inductive definition. Both will be briefly sketched in the following.

4.1 Feferman's model construction

We follow Feferman [20] and quote from there: The underlying applicative structure of OST is interpreted in the codes for functions that are Σ_1 definable in parameters, obtained by uniformizing the Σ_1 predicates. This proceeds as in Barwise [7], pp. 164- 167, which is applicable since under the assumption $(V{=}L)$, the universe is recursively listed in the sense given there. The treatment in Barwise must be modified slightly to account for parameters; this is done as follows. First one constructs a Σ_1 formula $\psi[w,x,y,z]$ such that for each Σ_1 formula $\theta[x,y,z]$ one can effectively find an $e \in \omega$ with $\theta[x,y,z]$ equivalent to $\psi[e,x,y,z]$. Then one uniformizises ψ with respect to y, i.e. produces a Σ_1 formula $\psi^*[w,x,y,z]$ that satisfies

(1) $\psi^*[w,x,y,z] \to \psi[w,x,y,z]$,

(2) $\exists y \psi^*[w,x,y,z] \to \exists! y \psi[w,x,y,z]$.

Given a set parameter p, one takes $\langle e,p\rangle$ to be the code of the partial function

(3) $\langle e,p\rangle(x) = y \;\leftrightarrow\; \psi^*[e,x,y,p]$.

One can then defined generalized "S-n-m" functions in a straightforward way, and from those give a model of the applicative axioms of OST. The rest of the interpretation proceeds in a straightforward way.

This construction is very elegant and compact. It depends, however, on some features specific of admissible recursion theory, and it seems not so clear whether it can be generalized to a framework for significant strengthenings of OST.

4.2 Inductive model construction

Alternatively, we can provide a direct inductive definition of the application operation. Apart from being more direct, this way of modeling OST has the advantage that it can be directly adapted (see below) to dealing with strong extensions of OST.

We use lower case Greek letters $\alpha, \beta, \gamma, \delta \ldots$ (possibly with subscripts) for ordinals – they are Δ_0 definable in $\mathsf{KP}\omega$ – and write $(\alpha < \beta)$ for $(\alpha \in \beta)$. Furthermore, $(a \in L_\alpha)$ states that the set a is an element of the αth level L_α of the constructible hierarchy, and $(a <_L b)$ means that a is smaller than b

according to the well-ordering $<_L$ on L. It is well-known that the assertions $(a \in L_\alpha)$ and $(a <_L b)$ are Δ over $\mathsf{KP}\omega$; see, e.g., Barwise [7] or Kunen [37].

The following approach is similar to those in Feferman and Jäger [21] and Jäger and Strahm [34] and begins with some notational preparations. For any natural number n greater than 0 we select (i) a Δ_0 formula $\mathit{Tup}_n(a)$ formalizing that a is an ordered n-tuple and (ii) a Δ_0 formula $(a)_n = b$ formalizing that b the projection of a on its nth component so that

$$\mathit{Tup}_n(a) \wedge (a)_1 = b_1 \wedge \ldots \wedge (a)_n = b_n \rightarrow a = \langle b_1, \ldots, b_n \rangle.$$

Then we fix pairwise different sets $\widehat{\mathsf{k}}$, $\widehat{\mathsf{s}}$, $\widehat{\top}$, $\widehat{\bot}$, $\widehat{\mathbf{el}}$, $\widehat{\mathbf{non}}$, $\widehat{\mathbf{dis}}$, $\widehat{\mathbf{e}}$, $\widehat{\mathbf{E}}$, $\widehat{\mathbb{S}}$, $\widehat{\mathbb{R}}$, $\widehat{\mathbb{C}}$ and $\widehat{\mathbb{P}}$ none of which belongs to the collection of ordered pairs and triples; they will later act as the codes of the corresponding constants of $\mathcal{L}^\circ$. We are going to code the $\mathcal{L}^\circ$ terms $\mathsf{k}x$, $\mathsf{s}x$, $\mathsf{s}xy$, ... by the ordered tuples $\langle \widehat{\mathsf{k}}, x \rangle$, $\langle \widehat{\mathsf{s}}, x \rangle$, $\langle \widehat{\mathsf{s}}, x, y \rangle$, ... of the corresponding form. For example, to satisfy $\mathsf{k}xy = x$ we interpret $\mathsf{k}x$ as $\langle \widehat{\mathsf{k}}, x \rangle$, and "$\langle \widehat{\mathsf{k}}, x \rangle$ applied to y" is taken to be x.

Next let R be a fresh 4-place relation symbol and extend $\mathcal{L}_1$ to the language $\mathcal{L}_1(R)$ with expressions $R(\alpha, a, b, c)$ as additional atomic formulas. We also abbreviate

$$R^{<\alpha}(a, b, c) \ := \ (\exists \beta < \alpha) R(\beta, a, b, c).$$

For finding the required interpretation of the application operation of OST within $\mathsf{KP}\omega + (V{=}L)$ we work with a specific $\mathcal{L}_1(R)$ formula, introduced in the following definition. Afterwards, this formula together with Σ recursion will help to provide what we need.

Definition 6 *We choose $\mathfrak{A}[R, \alpha, a, b, c]$ to be the $\mathcal{L}_1(R)$ formula defined as*

$$\mathfrak{A}[R, \alpha, a, b, c] \quad := \quad c \in L_\alpha \wedge \mathfrak{B}[R, \alpha, a, b, c],$$

where $\mathfrak{B}[R, \alpha, a, b, c]$ is an auxiliary $\mathcal{L}_1(R)$ formula given as the disjunction of the following clauses:

(1) $a = \widehat{\mathsf{k}} \wedge c = \langle \widehat{\mathsf{k}}, b \rangle$,

(2) $\mathit{Tup}_2(a) \wedge (a)_1 = \widehat{\mathsf{k}} \wedge (a)_2 = c$,

(3) $a = \widehat{\mathsf{s}} \wedge c = \langle \widehat{\mathsf{s}}, b \rangle$,

(4) $\mathit{Tup}_2(a) \wedge (a)_1 = \widehat{\mathsf{s}} \wedge c = \langle \widehat{\mathsf{s}}, (a)_2, b \rangle$,

(5) $\mathit{Tup}_3(a) \wedge (a)_1 = \widehat{\mathsf{s}} \wedge$

$$(\exists x, y \in L_\alpha)(R^{<\alpha}((a)_2, b, x) \wedge R^{<\alpha}((a)_3, b, y) \wedge R^{<\alpha}(x, y, c)),$$

(6) $a = \widehat{\mathbf{el}} \wedge c = \langle \widehat{\mathbf{el}}, b \rangle$,

(7) $\mathit{Tup}_2(a) \wedge (a)_1 = \widehat{\mathbf{el}} \wedge (a)_2 \in b \wedge c = \widehat{\top}$,

(8) $\mathit{Tup}_2(a) \wedge (a)_1 = \widehat{\mathbf{el}} \wedge (a)_2 \notin b \wedge c = \widehat{\bot}$,

(9) $a = \widehat{\mathbf{non}} \ \wedge \ b = \widehat{\top} \ \wedge \ c = \widehat{\bot}$,

(10) $a = \widehat{\mathbf{non}} \ \wedge \ b = \widehat{\bot} \ \wedge \ c = \widehat{\top}$,

(11) $a = \widehat{\mathbf{dis}} \ \wedge \ c = \langle \widehat{\mathbf{dis}}, b \rangle$,

(12) $\mathit{Tup}_2(a) \ \wedge \ (a)_1 = \widehat{\mathbf{dis}} \ \wedge \ (a)_2 = \widehat{\top} \ \wedge \ c = \widehat{\top}$,

(13) $\mathit{Tup}_2(a) \ \wedge \ (a)_1 = \widehat{\mathbf{dis}} \ \wedge \ (a)_2 = \widehat{\bot} \ \wedge \ b = \widehat{\top} \ \wedge \ c = \widehat{\top}$,

(14) $\mathit{Tup}_2(a) \ \wedge \ (a)_1 = \widehat{\mathbf{dis}} \ \wedge \ (a)_2 = \widehat{\bot} \ \wedge \ b = \widehat{\bot} \ \wedge \ c = \widehat{\bot}$,

(15) $a = \widehat{\mathbf{e}} \ \wedge \ c = \langle \widehat{\mathbf{e}}, b \rangle$,

(16) $\mathit{Tup}_2(a) \ \wedge \ (a)_1 = \widehat{\mathbf{e}} \ \wedge \ (\exists x \in b) R^{<\alpha}((a)_2, x, \widehat{\top}) \ \wedge \ c = \widehat{\top}$,

(17) $\mathit{Tup}_2(a) \ \wedge \ (a)_1 = \widehat{\mathbf{e}} \ \wedge \ (\forall x \in b) R^{<\alpha}((a)_2, x, \widehat{\bot}) \ \wedge \ c = \widehat{\bot}$,

(18) $a = \widehat{\mathbb{S}} \ \wedge \ c = \langle \widehat{\mathbb{S}}, b \rangle$,

(19) $\mathit{Tup}_2(a) \ \wedge \ (a)_1 = \widehat{\mathbb{S}} \ \wedge \ (\forall x \in b)(R^{<\alpha}((a)_2, x, \widehat{\top}) \vee R^{<\alpha}((a)_2, x, \widehat{\bot})) \ \wedge$
$(\forall x \in c)(x \in b \ \wedge \ R^{<\alpha}((a)_2, x, \widehat{\top})) \ \wedge$
$(\forall x \in b)(R^{<\alpha}((a)_2, x, \widehat{\top}) \ \rightarrow \ x \in c)$,

(20) $a = \widehat{\mathbb{R}} \ \wedge \ c = \langle \widehat{\mathbb{R}}, b \rangle$,

(21) $\mathit{Tup}_2(a) \ \wedge \ (a)_1 = \widehat{\mathbb{R}} \ \wedge \ (\forall x \in b)(\exists y \in c) R^{<\alpha}((a)_2, x, y) \ \wedge$
$(\forall y \in c)(\exists x \in b) R^{<\alpha}((a)_2, x, y)$,

(22) $a = \widehat{\mathbb{C}} \ \wedge \ R^{<\alpha}(b, c, \widehat{\top}) \ \wedge \ (\forall x \in L_\alpha)(x <_L c \ \rightarrow \ \neg R^{<\alpha}(b, x, \widehat{\top})) \ \wedge$
$(\forall \beta < \alpha)(\forall x \in L_\beta) \neg R^{<\beta}(b, x, \widehat{\top})$.

We immediately see that $\mathfrak{A}[R, \alpha, a, b, c]$ is Δ over $\mathsf{KP}\omega$ with respect to the language $\mathcal{L}_1(R)$. It is also easy to verify that $\mathfrak{A}[R, \alpha, a, b, c]$ is deterministic in the following sense: from $\mathfrak{A}[R, \alpha, a, b, c]$ we can conclude that exactly one of the clauses (1)–(22) of the previous definition is satisfied for these α, a, b, c.

For any $\mathcal{L}_1$ formula $B[\alpha, a, b, c]$ with at most the indicated free variables we write $\mathfrak{A}[B, \alpha, a, b, c]$ for the $\mathcal{L}_1$ formula resulting by replacing each occurrence of an atomic formula of the form $R(\alpha, r, s, t)$ in $\mathfrak{A}[R, \alpha, a, b, c]$ by $B[\alpha, r, s, t]$. The following theorem is a special case of "Definition by Σ Recursion" as developed in Barwise [7].

Theorem 7 *There exists a Σ formula $B[\alpha, a, b, c]$ of $\mathcal{L}_1$ with at most α, a, b and c free so that $\mathsf{KP}\omega$ proves*

$$(\Sigma\text{-}\mathsf{Rec}/\mathfrak{A}) \qquad B[\alpha, a, b, c] \ \leftrightarrow \ \mathfrak{A}[B, \alpha, a, b, c].$$

Any such a formula $B[\alpha, a, b, c]$ may be used to describe the αth level of the interpretation of the OST application $(ab \simeq c)$. Accordingly, we proceed as follows.

Definition 8 *Let $B_{\mathfrak{A}}[\alpha, a, b, c]$ be a Σ formula of $\mathcal{L}_1$ associated to the operator form $\mathfrak{A}[R, \alpha, a, b, c]$ according to (Σ-$\mathsf{Rec}/\mathfrak{A}$) of the previous theorem. Then we define*

$$Ap_{\mathfrak{A}}[a, b, c] \quad := \quad \exists \alpha B_{\mathfrak{A}}[\alpha, a, b, c].$$

As can be easily checked, $Ap_{\mathfrak{A}}[a, b, c]$ is functional in its third argument. It is therefore suitable for handling the $\mathcal{L}^\circ$ terms within $\mathsf{KP}\omega + (V{=}L)$. To each term t of $\mathcal{L}^\circ$ we associate a formula $[\![t]\!]_{\mathfrak{A}}(u)$ of $\mathcal{L}_1$ expressing that u is the value of t under the interpretation of the OST-application via the Σ formula $Ap_{\mathfrak{A}}[a, b, c]$.

Definition 9 *For each $\mathcal{L}^\circ$ term t we introduce an $\mathcal{L}_1$ formula $[\![t]\!]_{\mathfrak{A}}(u)$, with u not occurring in t, which is inductively defined as follows:*

1. *If t is a set variable, then $[\![t]\!]_{\mathfrak{A}}(u)$ is the formula $(t = u)$.*
2. *If t is a constant, then $[\![t]\!]_{\mathfrak{A}}(u)$ is the formula $(\widehat{t} = u)$.*
3. *If t is the term (rs), then we set*

$$[\![t]\!]_{\mathfrak{A}}(u) \;:=\; \exists x \exists y([\![r]\!]_{\mathfrak{A}}(x) \,\wedge\, [\![s]\!]_{\mathfrak{A}}(y) \,\wedge\, Ap_{\mathfrak{A}}[x, y, u]).$$

Observe that for every term t of $\mathcal{L}^\circ$ its translation $[\![t]\!]_{\mathfrak{A}}(u)$ is a Σ formula of $\mathcal{L}_1$. By this treatment of the terms of $\mathcal{L}^\circ$, the translation of arbitrary formulas of $\mathcal{L}^\circ$ into formulas of $\mathcal{L}_1$ is predetermined.

Definition 10 *The translation of an $\mathcal{L}^\circ$ formula A into the $\mathcal{L}_1$ formula A^* is inductively defined as follows:*

1. *For the atomic formulas of $\mathcal{L}^\circ$ we stipulate*

$$\begin{aligned} (t\downarrow)^* \quad &:= \quad \exists x [\![t]\!]_{\mathfrak{A}}(x), \\ (s \in t)^* \quad &:= \quad \exists x \exists y([\![s]\!]_{\mathfrak{A}}(x) \,\wedge\, [\![t]\!]_{\mathfrak{A}}(y) \,\wedge\, x \in y). \end{aligned}$$

2. *If A is a formula $\neg B$, then A^* is $\neg B^*$.*
3. *If A is a formula $(B \Diamond C)$ for $\Diamond$ being the binary junctor $\vee$ or $\wedge$, then A^* is $(B^* \Diamond C^*)$.*
4. *If A is a formula $(\exists x \in t)B[x]$, then*

$$A^* \;:=\; \exists y([\![t]\!]_{\mathfrak{A}}(y) \,\wedge\, (\exists x \in y)B^*[x]).$$

5. *If A is a formula $(\forall x \in t)B[x]$, then*

$$A^* := \forall y([\![t]\!]_{\mathfrak{A}}(y) \rightarrow (\forall x \in y)B^*[x]).$$

6. *If A is a formula $QxB[x]$ for a quantifier Q, then A^* is $QxB^*[x]$.*

This translation of $\mathcal{L}^\circ$ formulas leads directly to an interpretation of OST in $\mathsf{KP}\omega + (V{=}L)$. The corresponding interpretation result is proved in Jäger [29].

Theorem 11 *The theory OST is interpretable in $\mathsf{KP}\omega + (V{=}L)$; i.e. for all formulas A of $\mathcal{L}^\circ$ we have*

$$\mathsf{OST} \vdash A \quad \Longrightarrow \quad \mathsf{KP}\omega + (V{=}L) \vdash A^*.$$

As remarked earlier, $\mathsf{KP}\omega + (V{=}L)$ is a conservative extension of $\mathsf{KP}\omega$ for absolute formulas. If we combine this with Theorem 5 and Theorem 11, we obtain the following corollary, which settles the question of the consistency strength of OST. This result was first established in Feferman [18].

Corollary 12 *The theory OST is conservative over $\mathsf{KP}\omega$ for absolute formulas. In particular, OST and $\mathsf{KP}\omega$ are equiconsistent.*

We establish an upper bound for $\mathsf{OST}(\mathbb{P})$ by an easy modification of the argument in the previous section: only extend the disjunction in Definition 6 by a clause taking care of the constant $\mathbb{P}$.

Definition 13 *We choose $\mathfrak{C}[R, \alpha, a, b, c]$ to be the $\Delta(\mathcal{P})$ formula of $\mathcal{L}_1(\mathcal{P}, R)$ defined as*

$$\mathfrak{C}[R, \alpha, a, b, c] \quad := \quad c \in L_\alpha \wedge (\mathfrak{B}[R, \alpha, a, b, c] \vee (a = \widehat{\mathbb{P}} \wedge \mathcal{P}(b, c))),$$

where $\mathfrak{B}[R, \alpha, a, b, c]$ is the formula introduced in Definition 6.

In $\mathsf{KP}(\mathcal{P})$ we have $\Sigma(\mathcal{P})$ recursion. Completely in the line of the previous section we apply it now, of course, to the operator form $\mathfrak{C}[R, \alpha, a, b, c]$, yielding the following analogue of Theorem 7.

Theorem 14 *There exists a $\Sigma(\mathcal{P})$ formula $B[\alpha, a, b, c]$ of $\mathcal{L}_1(\mathcal{P}, R)$ with at most α, a, b and c free so that $\mathsf{KP}(\mathcal{P})$ proves*

$$(\Sigma(\mathcal{P})\text{-}\mathsf{Rec}/\mathfrak{C}) \qquad B[\alpha, a, b, c] \leftrightarrow \mathfrak{C}[B, \alpha, a, b, c].$$

Naturally, each $\Sigma(\mathcal{P})$ formula $B[\alpha, a, b, c]$ fulfilling this recursion equation $(\Sigma(\mathcal{P})\text{-}\mathsf{Rec}/\mathfrak{C})$ is now a possible candidate for interpreting the $\mathsf{OST}(\mathbb{P})$ application $(ab \simeq c)$.

Definition 15 *Let $B_{\mathfrak{C}}[\alpha, a, b, c]$ be a $\Sigma(\mathcal{P})$ formula of $\mathcal{L}_1(\mathcal{P})$ associated to the operator form $\mathfrak{C}[R, \alpha, a, b, c]$ according to $(\Sigma(\mathcal{P})\text{-}\mathsf{Rec}/\mathfrak{C})$ of the previous theorem. Then we define*

$$Ap_{\mathfrak{C}}[a, b, c] := \exists \alpha B_{\mathfrak{C}}[\alpha, a, b, c].$$

It only remains to proceed as in the previous section, but with $Ap_{\mathfrak{A}}[a,b,c]$ replaced by $Ap_{\mathfrak{C}}[a,b,c]$. For each $\mathcal{L}^\circ$ term t, an $\mathcal{L}_1(\mathcal{P})$ formula $[\![t]\!]_{\mathfrak{C}}(u)$ is introduced, saying that u is the value of the term t under the interpretation of the $\mathsf{OST}(\mathbb{P})$ application via $Ap_{\mathfrak{C}}[a,b,c]$. Finally, following the pattern of Definition 10 and based on these $[\![t]\!]_{\mathfrak{C}}(u)$, each $\mathcal{L}^\circ$ formula A is canonically translated into a formula $A^\sharp$ of $\mathcal{L}_1(\mathcal{P})$.

The interpretation of $\mathsf{OST}(\mathbb{P})$ in $\mathsf{KP}(\mathcal{P})+(V{=}L)$ is a straightforward extension of Theorem 11. For details see Jäger [29].

Theorem 16 *The theory* $\mathsf{OST}(\mathbb{P})$ *is interpretable in* $\mathsf{KP}(\mathcal{P})+(V{=}L)$*; i.e. for all formulas* A *of* $\mathcal{L}^\circ$ *we have*

$$\mathsf{OST}(\mathbb{P}) \vdash A \quad \Longrightarrow \quad \mathsf{KP}(\mathcal{P}) + (V{=}L) \vdash A^\sharp.$$

Unfortunately, the combination of Theorem 5 and Theorem 16 does not completely settle the question about the consistency strength of $\mathsf{OST}(\mathbb{P})$ yet. So far we have an interesting lower and an interesting upper bound, but it still has to be determined what the relationship between $\mathsf{KP}(\mathcal{P})$ and $\mathsf{KP}(\mathcal{P}) + (V{=}L)$ is.

5 The theories $\mathsf{OST}(\mathbf{E},\mathbb{P})$ and $\mathsf{OST}^r(\mathbf{E},\mathbb{P})$

This section is dedicated to the extension $\mathsf{OST}(\mathbf{E},\mathbb{P})$ of OST – baptized $\mathsf{OST}+(\mathsf{Pow})+(\mathsf{Uni})$ in Feferman [18, 20] – and its subsystem $\mathsf{OST}^r(\mathbf{E},\mathbb{P})$. $\mathsf{OST}(\mathbf{E},\mathbb{P})$ provides for unbounded existential quantification and power set and brings us into the realm of ZFC and beyond.

The subsystem $\mathsf{OST}^r(\mathbf{E},\mathbb{P})$ of $\mathsf{OST}(\mathbf{E},\mathbb{P})$ is designed to be a witness of an operational set theory of the same strength as ZFC, sought in Feferman [18]. It is obtained from $\mathsf{OST}(\mathbf{E},\mathbb{P})$ by simply restricting the schema of $\in$-induction for arbitrary formulas to $\in$-induction for sets.

It should be fairly straightforward to prove in $\mathsf{OST}(\mathbf{E},\mathbb{P})$ that ZFC is consistent, implying that $\mathsf{OST}(\mathbf{E},\mathbb{P})$ is stronger than ZFC; actually this also follows from Theorem 18 below. To characterize it in terms of consistency strength it is natural to turn to theories of sets and classes.

5.1 NBG and a bit more

In surveying von Neumann-Bernays-Gödel set theory NBG and its extension $\mathsf{NBG}_{<E_0}$ we follow their presentation in Jäger [30]. The formalization of NBG there is based on standard literature, for example Levy [38] and Mendelson [41].

NBG is a theory of sets and classes conservative over the system ZFC of Zermelo-Fraenkel set theory with the axiom of choice. NBG is known to be finitely axiomatizable although the version we are going to present below permits axiom schemas and as such is an infinite axiomatization.

$\mathcal{L}_2$, the language of NBG, augments $\mathcal{L}_1$ by a second sort of countably many variables $U, V, W, X, Y, Z, \ldots$ (possibly with subscripts) for classes. The set terms of $\mathcal{L}_2$ are the terms of $\mathcal{L}_1$, as class terms we simply have the class variables.

The *formulas* $(A, B, C, D, A_1, B_1, C_1, D_1, \ldots)$ of $\mathcal{L}_2$ are inductively generated as follows:

1. If s and t are set terms of $\mathcal{L}_2$ and U is a class variable, then all expressions of the form $(s \in t)$ and $(s \in U)$ are (atomic) formulas of $\mathcal{L}_2$.

2. If A and B are formulas of $\mathcal{L}_2$, then so are are $\neg A$, $(A \vee B)$ and $(A \wedge B)$.

3. If A is a formula and t a set term of $\mathcal{L}_2$ which does not contain x, then $(\exists x \in t)A$, $(\forall x \in t)A$, $\exists x A$, $\forall x A$, $\exists X A$ and $\forall X A$ are formulas of $\mathcal{L}_2$.

As before, the remaining logical connectives are introduced as abbreviations, and we will often omit parentheses and brackets whenever there is no danger of confusion. Equalities between sets/sets, sets/classes, classes/sets and classes/classes are not atomic formulas of $\mathcal{L}_2$ but defined as

$$(\mathit{Var}_1 = \mathit{Var}_2) \ := \ \forall x(x \in \mathit{Var}_1 \ \leftrightarrow \ x \in \mathit{Var}_2)$$

where Var_1 and Var_2 denote set or class variables. A formula of $\mathcal{L}_2$ is called *elementary* or a Π^1_0 formula if it does not contain bound class variables; free class variables, however, are permitted. The Σ^1_1 formulas of $\mathcal{L}_2$ are those of the form $\exists X A$ with elementary A.

The logic of NBG is classical two-sorted logic with equality for the first sort. The non-logical axioms of NBG are given in six groups. To increase readability, we freely use standard set-theoretic terminology.

I. Elementary comprehension For any elementary formula $A[u]$ of $\mathcal{L}_2$:

(ECA) $$\exists X \forall y (y \in X \ \leftrightarrow \ A[y]).$$

Hence every elementary NBG formula $A[u]$ defines a class, which is typically written as $\{x : A[x]\}$. It may be (extensionally equal to) a set, but this is not necessarily the case. The intersection of a class with a set, however, is always supposed to produce a set by the following principle of Aussonderung.

II. Aussonderung

(AUS) $$\forall X \forall y \exists z (z = X \cap y).$$

From logical reasons, (ECA) and (AUS) we conclude that there is a unique set which has no members; it is denoted by $\emptyset$.

III. Basic set existence

(Pair) $\forall x \forall y \exists z (z = \{x, y\})$,

(Union) $\forall x \exists y (y = \cup x)$,

(Powerset) $\forall x \exists y \forall z (z \in y \leftrightarrow z \subset x)$,

(Infinity) $\exists x (\emptyset \in x \wedge (\forall y \in x)(y \cup \{y\} \in x))$.

As in $\mathsf{OST}(\mathbf{E}, \mathbb{P})$ we write $\langle a, b \rangle$ for the ordered pair of the sets a and b à la Kuratowski. Class relations are classes which consist of ordered pairs only, and class functions are class relations which assign to every set exactly one set; i.e. for all U we set:

$$
\begin{aligned}
\mathit{Rel}[U] &:= \forall x (x \in U \rightarrow \exists y \exists z (x = \langle y, z \rangle)), \\
\mathit{Fun}[U] &:= \mathit{Rel}[U] \wedge \forall x \exists! y (\langle x, y \rangle \in U).
\end{aligned}
$$

If U is a function we write $U(x)$ for the uniquely determined y associated to x by U. Replacement states that the range of a set under a function is a set.

IV. Replacement

(REP) $\forall X (\mathit{Fun}[X] \rightarrow \forall y \exists z (z = \{X(u) : u \in y\}))$.

Global choice is a very uniform principle of choice which claims the existence of a class function which picks an element of any non-empty set.

V. Global choice

(GC) $\exists X (\mathit{Fun}[X] \wedge \forall y (y \neq \emptyset \rightarrow X(y) \in y))$.

Finally, in NBG it is claimed that the element relation is well-founded with respect to classes.

VI. Class foundation

($\mathsf{C}\text{-}\mathsf{I}_\in$) $\forall X (X \neq \emptyset \rightarrow \exists y (y \in X \wedge (\forall z \in y)(z \notin X)))$.

The axioms (Infinity) and ($\mathsf{C}\text{-}\mathsf{I}_\in$) imply that there exists a least infinite ordinal, which we denote by ω, as usual. The elements of ω are identified with the natural numbers in the sense that $0 := \emptyset$, $1 := \{0\}$, $2 := 1 \cup \{1\}$ and so on.

The axioms in groups I - VI supply one possible axiomatization of NBG. According to a well-known result (cf., e.g., Levy [38]) NBG is a conservative extension of ZFC.

Theorem 17 *A sentence of the language $\mathcal{L}_1$ is provable in* NBG *if and only if it is provable in* ZFC.

A first step in extending NBG is to add the principle of $\in$-induction for arbitrary $\mathcal{L}_2$ formulas $A[u]$,

$$(\mathcal{L}_2\text{-}\mathsf{I}_\in) \qquad \forall x((\forall y \in x)A[y] \rightarrow A[x]) \rightarrow \forall x A[x].$$

Besides that, we want to be able to iterate elementary comprehension along all well-orderings which can be constructed from the ordinals and the order type of all ordinals by closing them under addition and ω-exponentiation. For this purpose we introduce a notation system $(E_0, \prec)$ which can be considered as the canonical blowing up of $(\varepsilon_0, <)$ triggered by replacing the natural numbers by the ordinals. In particular:

(i) E_0 is an elementarily definable class, and $\prec$ is an elementarily definable binary class relation on E_0.

(ii) For any ordinal α there exists a code $\overline{\alpha}$ which belongs to E_0.

(iii) E_0 contains an element Ω such that $(\Omega, \prec)$ is an isomorphic copy of the class of all ordinals.

(iv) There are a binary class function $\oplus$ and a unary class function Exp_ω, both elementary, such that E_0 is closed under $\oplus$ and Exp_ω. These two functions are for the addition and ω-exponentiation of elements of E_0 in the expected sense.

In the following we write $(a+b)$ – or often simply $a+b$ – for $(a \oplus b)$ and ω^a for $Exp_\omega(a)$. For all natural numbers n, the ordinal terms Ω_n are inductively defined by

$$\Omega_0 := \Omega + 1 \qquad \text{and} \qquad \Omega_{n+1} := \omega^{\Omega_n}.$$

All additional relevant details concerning $(E_0, \prec)$ are worked out in detail in Jäger and Krähenbühl [32]. Amongst other things it is shown there that, for any standard natural number n, the theory $\mathsf{NBG} + (\mathcal{L}_2\text{-}\mathsf{I}_\in)$ proves transfinite induction along $\prec$ up to any Ω_n.

Let $A[U,V,u,v]$ be an elementary $\mathcal{L}_2$ formula with at most the indicated variables free. Then we write $Hier_A[a,U,V]$ for the elementary $\mathcal{L}_2$ formula

$$(\forall b \prec a)((V)_b = \{x : A[U, \Sigma(V,b), b, x]\})$$

where $\Sigma(V,b)$ stands for the class $\{\langle x,c\rangle \in V : c \prec b\}$ representing the disjoint union of the projections of V up to b. This formula states that V codes the hierarchy generated by iterating comprehension via A with class parameter U along $\prec$ up to a.

$\mathsf{NBG}_{<E_0}$ is defined to be the theory of sets and classes which consists of NBG, full $\in$-induction $(\mathcal{L}_2\text{-}\mathsf{I}_\in)$ plus the additional axioms

$$\forall X \exists Y\, Hier_A[\Omega_n, X, Y]$$

for all standard natural numbers n and all elementary formulas $A[U, V, u, v]$ of $\mathcal{L}_2$ with at most the variables U, V, u, v free. So it permits iteration of elementary class comprehension along each (standard) initial segment of E_0.

In operational set theory we call an operation f a *total characteristic operation* if is totally defined and takes values in $\mathbb{B}$ only;

$$TCO(f) \;:=\; \forall x(fx \in \mathbb{B}).$$

By regarding the class variables of $\mathcal{L}_2$ as variables of $\mathcal{L}^\circ$ ranging over total characteristic operations and by translating an atomic $\mathcal{L}_2$ formula $(a \in U)$ into $(fa = \top)$, where f is the $\mathcal{L}^\circ$ variable associated to the $\mathcal{L}_2$ variable U, every $\mathcal{L}_2$ formula A canonically translates into an $\mathcal{L}^\circ$ formula A°. This translation is such that

$$\begin{aligned} \exists X A[X]^\circ &= \exists x(TCO(x) \land A^\circ[x]), \\ \forall X A[X]^\circ &= \forall x(TCO(x) \to A^\circ[x]), \end{aligned}$$

always modulo a renaming of the variables if necessary. This translation leads to the following interpretation theorem.

Theorem 18 *1. The theory* NBG *is interpretable in* $\mathsf{OST}^r(\mathbf{E}, \mathbb{P})$*.*

2. The theory $\mathsf{NBG}_{<E_0}$ *is interpretable in* $\mathsf{OST}(\mathbf{E}, \mathbb{P})$*.*

The second part of this theorem is proved in Jäger [30]; the first part follows directly from an inspection of this proof. A result equivalent to the first assertion can be found in Jäger [29] where it is shown that ZFC can be embedded into $\mathsf{OST}^r(\mathbf{E}, \mathbb{P})$.

5.2 Inductive extensions of ZF

By the previous theorem we have lower proof theoretic bounds for $\mathsf{OST}^r(\mathbf{E}, \mathbb{P})$ and $\mathsf{OST}(\mathbf{E}, \mathbb{P})$, and what remains is to show that that these bounds are sharp. This will be achieved by utilizing inductive model constructions again. They can be carried out in the auxiliary theories $\mathsf{E}^r(\mathsf{ZFW})$ and $\mathsf{E}^r(\mathsf{ZFW}) + (\mathcal{L}_S\text{-}\mathsf{I}_\in)$, depending on how much induction we have to model, which are reducible to NBG and $\mathsf{NBG}_{<E_0}$, respectively.

When building up the inductive model of $\mathsf{OST}(\mathbf{E}, \mathbb{P})$, we have to handle the choice axiom (S3). For this end it is convenient to have a global well-ordering of the set-theoretic universe at our disposal. Therefore, let $\mathcal{L}_1(\mathcal{W})$ be the extension of $\mathcal{L}_1$ by the fresh binary relation symbol $\mathcal{W}$, and let ZFW be the extension of ZF which comprises all axioms of ZF – formulated, of course, with respect to the new language $\mathcal{L}_1(\mathcal{W})$ – plus the following well-ordering axiom

($\mathcal{W}$) $\quad \forall x \exists! \alpha \mathcal{W}(x, \alpha) \;\land\; \forall x \forall y \forall \alpha(\mathcal{W}(x, \alpha) \land \mathcal{W}(y, \alpha) \;\to\; x = y).$

From axiom $(\mathcal{W})$ the desired well-ordering of the universe of sets is canonically obtained if we set

$$(<_{\mathcal{W}}) \qquad a <_{\mathcal{W}} b \;:=\; \exists\alpha\exists\beta(\mathcal{W}(a,\alpha) \,\land\, \mathcal{W}(b,\beta) \,\land\, \alpha < \beta).$$

Analogously to Section 4.2 we pick an n-ary relation symbol R which does not belong to the language $\mathcal{L}_1(\mathcal{W})$ and write $\mathcal{L}_1(\mathcal{W}, R)$ for the extension of $\mathcal{L}_1(\mathcal{W})$ by R. An $\mathcal{L}_1(\mathcal{W}, R)$ formula which contains at most $a_1, \ldots, a_n$ free is called an n-ary operator form, and we let $\mathfrak{F}[R, a_1, \ldots, a_n]$ range over such forms.

Based on a model $\mathcal{M}$ of ZFW with universe $|\mathcal{M}|$, any n-ary operator form $\mathfrak{F}[R, \vec{a}]$ gives rise to subsets $I^{\zeta}_{\mathfrak{F}}$ of $|\mathcal{M}|^n$ generated inductively for all ordinals ζ (not only those belonging to $|\mathcal{M}|$) by

$$I^{<\zeta}_{\mathfrak{F}} \;:=\; \bigcup_{\eta<\zeta} I^{\eta}_{\mathfrak{F}} \quad \text{and} \quad I^{\zeta}_{\mathfrak{F}} \;:=\; \{\langle\vec{x}\rangle \in |\mathcal{M}|^n : \mathcal{M} \models \mathfrak{F}[I^{<\zeta}_{\mathfrak{F}}, \vec{x}]\}.$$

These sets $I^{\zeta}_{\mathfrak{F}}$ are the *stages* of the inductive definition induced by $\mathfrak{F}[R, \vec{a}]$, relative to $\mathcal{M}$; for many models $\mathcal{M}$, operator forms $\mathfrak{F}[R, \vec{a}]$ and ordinals ζ the $I^{\zeta}_{\mathfrak{F}}$ are not elements of $|\mathcal{M}|$. We now enrich ZFW so that we can speak about such stages.

The theory $\mathsf{E}^r(\mathsf{ZFW})$ is formulated in the language $\mathcal{L}_S$ which extends $\mathcal{L}_1(\mathcal{W})$ by adding a new sort of so called stage variables $\rho, \sigma, \tau, \ldots$ (possibly with subscripts) as well as new binary relation symbols $\prec$ and $\equiv$ for the less and equality relation for stage variables, respectively. Moreover, $\mathcal{L}_S$ includes an $(n+1)$-ary relation symbol $Q_{\mathfrak{F}}$ for each operator form $\mathfrak{F}[R, a_1, \ldots, a_n]$.

The set terms of $\mathcal{L}_S$ are the set terms of $\mathcal{L}_1$, and the atomic formulas of $\mathcal{L}_S$ are the atomic formulas of $\mathcal{L}_1(\mathcal{W})$ plus all expressions $(\sigma \prec \tau)$, $(\sigma \equiv \tau)$ and $Q_{\mathfrak{F}}(\sigma, \vec{s})$ for each n-ary operator form $\mathfrak{F}[R, \vec{a}]$. Usually we write $Q^{\sigma}_{\mathfrak{F}}(\vec{s})$ instead of $Q_{\mathfrak{F}}(\sigma, \vec{s})$.

The formulas $(A, B, C, A_1, B_1, C_1, \ldots)$ of $\mathcal{L}_S$ are generated from these atomic formulas by closure under negation, conjunction and disjunction, bounded and unbounded quantification over sets, bounded stage quantification $(\exists\sigma \prec \tau)$ and $(\forall\sigma \prec \tau)$ as well as unbounded stage quantification $\exists\sigma$ and $\forall\sigma$. The Δ^S_0 formulas are those $\mathcal{L}_S$ formulas that do not contain unbounded stage quantifiers. An $\mathcal{L}_S$ formula A is is called Σ^S if all positive occurrences of unbounded stage quantifiers in A are existential and all negative occurrences of unbounded stage quantifiers in A are universal; it is called Π^{Ω} if all positive occurrences of unbounded stage quantifiers in A are universal and all negative occurrences of unbounded stage quantifiers in A are existential.

Further, we write A^{σ} to denote the $\mathcal{L}_S$ formula which is obtained from A by replacing all unbounded stage quantifiers $Q\tau$ in A by bounded stage quantifiers $(Q\tau \prec \sigma)$. Additional abbreviations are

$$Q^{\prec\sigma}_{\mathfrak{F}}(\vec{s}) \;:=\; (\exists\tau \prec \sigma)Q^{\tau}_{\mathfrak{F}}(\vec{s}) \qquad \text{and} \qquad Q_{\mathfrak{F}}(\vec{s}) \;:=\; \exists\sigma Q^{\sigma}_{\mathfrak{F}}(\vec{s}).$$

Clearly, any formula of $\mathcal{L}_1(\mathcal{W})$ is a (trivial) Δ_0^S formula, and A^σ is Δ_0^S for any $\mathcal{L}_S$ formula A.

The theory $\mathsf{E}^r(\mathsf{ZFW})$ is formulated in classical two sorted predicate logic with equality in both sorts; in addition it contains as non-logical axioms all ZFW-axioms of the language $\mathcal{L}_1(\mathcal{W})$, some axioms about stage variables and operator forms, reflection for Σ^S formulas, separation and replacement for Δ_0^S formulas plus induction along $\in$ and along $\prec$ for Δ_0^S formulas.

I. ZFW-axioms. All axioms of the theory ZFW formulated in the language $\mathcal{L}_1(\mathcal{W})$; they do not refer to stage variables or relation symbols associated to operator forms.

II. Linearity axioms. For all stage variables ρ, σ and τ:

$$\sigma \not\prec \sigma \;\wedge\; (\rho \prec \sigma \wedge \sigma \prec \tau \;\rightarrow\; \rho \prec \tau) \;\wedge\; (\sigma \prec \tau \;\vee\; \sigma \equiv \tau \;\vee\; \tau \prec \sigma).$$

III. Operator axioms. For all operator forms $\mathfrak{F}[R, \vec{u}]$ and all set terms $\vec{s}$:

$$Q^\sigma_{\mathfrak{F}}(\vec{s}) \;\leftrightarrow\; \mathfrak{F}[Q^{\prec\sigma}_{\mathfrak{F}}, \vec{s}].$$

IV. Σ^S reflection. For all Σ^S formulas A:

(Σ^S-Ref) $$A \;\rightarrow\; \exists\sigma A^\sigma.$$

V. Δ_0^S Separation. For all Δ_0^S formulas $A[u]$ and all set terms s:

(Δ_0^S-Sep) $$\exists x(x = \{y \in s : A[y]\}).$$

VI. Δ_0^S Replacement. For all Δ_0^S formulas $A[u, v]$ and all set terms s:

(Δ_0^S-Rep) $$(\forall x \in s)\exists! y A[x, y] \;\rightarrow\; \exists z \forall y (y \in z \;\leftrightarrow\; (\exists x \in s) A[x, y]).$$

VII. Δ_0^S induction along $\in$ and $\prec$. For all Δ_0^S formulas $A[u]$:

(Δ_0^S-$\mathsf{I}_\in$) $$\forall x((\forall y \in x)A[y] \rightarrow A[x]) \;\rightarrow\; \forall x A[x],$$

(Δ_0^S-$\mathsf{I}_\prec$) $$\forall \sigma((\forall \tau \prec \sigma)A[\tau] \rightarrow A[\sigma]) \;\rightarrow\; \forall \sigma A[\sigma].$$

The theory $\mathsf{E}^r(\mathsf{ZFW})$ is a restricted system (hence the superscript "r") in the sense that the axioms in groups V, VI and VII are restricted to Δ_0^S formulas. By $\mathsf{E}^r(\mathsf{ZFW}) + (\mathcal{L}_S$-$\mathsf{I}_\in)$ is meant $\mathsf{E}^r(\mathsf{ZFW})$ extended by the schema of $\in$-induction for arbitrary $\mathcal{L}_S$ formulas.

It is important to observe that the stage variables do not belong to the collection of sets; they constitute a different entity which is used to "enumerate" the stages of the inductive definition associated to each operator form. However, in the form of Δ_0^S separation and Δ_0^S replacement they can nevertheless help to constitute new sets in a carefully restricted way.

Following the pattern of Section 4.2 we are now going to introduce a specific inductive definition which will lead to a suitable treatment of the application operation of $\mathsf{OST}^r(\mathbf{E},\mathbb{P})$ and $\mathsf{OST}(\mathbf{E},\mathbb{P})$. The decisive new aspect, see clauses (24) and (25), is the treatment of operational unbounded existential quantification. Also, the well-ordering of the set-theoretic universe generated by the axiom $(V = L)$, as it is used in Definition 6, is replaced by $<_{\mathcal{W}}$.

Definition 19 *The operator form $\mathfrak{F}[R, a, b, c]$ is defined to be the disjunction of the following clauses:*

(1) $a = \widehat{\mathsf{k}} \ \wedge\ c = \langle \widehat{\mathsf{k}}, b \rangle$,

(2) $\mathit{Tup}_2(a) \ \wedge\ (a)_1 = \widehat{\mathsf{k}} \ \wedge\ (a)_2 = c$,

(3) $a = \widehat{\mathsf{s}} \ \wedge\ c = \langle \widehat{\mathsf{s}}, b \rangle$,

(4) $\mathit{Tup}_2(a) \ \wedge\ (a)_1 = \widehat{\mathsf{s}} \ \wedge\ c = \langle \widehat{\mathsf{s}}, (a)_2, b \rangle$,

(5) $\mathit{Tup}_3(a) \ \wedge\ (a)_1 = \widehat{\mathsf{s}} \ \wedge\ \exists x \exists y (R((a)_2, b, x) \ \wedge\ R((a)_3, b, y) \ \wedge\ R(x, y, c))$,

(6) $a = \widehat{\mathbf{el}} \ \wedge\ c = \langle \widehat{\mathbf{el}}, b \rangle$,

(7) $\mathit{Tup}_2(a) \ \wedge\ (a)_1 = \widehat{\mathbf{el}} \ \wedge\ (a)_2 \in b \ \wedge\ c = \widehat{\top}$,

(8) $\mathit{Tup}_2(a) \ \wedge\ (a)_1 = \widehat{\mathbf{el}} \ \wedge\ (a)_2 \notin b \ \wedge\ c = \widehat{\bot}$,

(9) $a = \widehat{\mathbf{non}} \ \wedge\ b = \widehat{\top} \ \wedge\ c = \widehat{\bot}$,

(10) $a = \widehat{\mathbf{non}} \ \wedge\ b = \widehat{\bot} \ \wedge\ c = \widehat{\top}$,

(11) $a = \widehat{\mathbf{dis}} \ \wedge\ c = \langle \widehat{\mathbf{dis}}, b \rangle$,

(12) $\mathit{Tup}_2(a) \ \wedge\ (a)_1 = \widehat{\mathbf{dis}} \ \wedge\ (a)_2 = \widehat{\top} \ \wedge\ c = \widehat{\top}$,

(13) $\mathit{Tup}_2(a) \ \wedge\ (a)_1 = \widehat{\mathbf{dis}} \ \wedge\ (a)_2 = \widehat{\bot} \ \wedge\ b = \widehat{\top} \ \wedge\ c = \widehat{\top}$,

(14) $\mathit{Tup}_2(a) \ \wedge\ (a)_1 = \widehat{\mathbf{dis}} \ \wedge\ (a)_2 = \widehat{\bot} \ \wedge\ b = \widehat{\bot} \ \wedge\ c = \widehat{\bot}$,

(15) $a = \widehat{\mathbf{e}} \ \wedge\ c = \langle \widehat{\mathbf{e}}, b \rangle$,

(16) $\mathit{Tup}_2(a) \ \wedge\ (a)_1 = \widehat{\mathbf{e}} \ \wedge\ (\exists x \in b) R((a)_2, x, \widehat{\top}) \ \wedge\ c = \widehat{\top}$,

(17) $\mathit{Tup}_2(a) \ \wedge\ (a)_1 = \widehat{\mathbf{e}} \ \wedge\ (\forall x \in b) R((a)_2, x, \widehat{\bot}) \ \wedge\ c = \widehat{\bot}$,

(18) $a = \widehat{\mathbb{S}} \ \wedge\ c = \langle \widehat{\mathbb{S}}, b \rangle$,

(19) $Tup_2(a) \;\wedge\; (a)_1 = \widehat{\mathbb{S}} \;\wedge\; (\forall x \in b)(R((a)_2, x, \widehat{\top}) \vee R((a)_2, x, \widehat{\bot})) \;\wedge$

$\forall x(x \in c \leftrightarrow x \in b \wedge R((a)_2, x, \widehat{\top}))$,

(20) $a = \widehat{\mathbb{R}} \;\wedge\; c = \langle \widehat{\mathbb{R}}, b\rangle$,

(21) $Tup_2(a) \;\wedge\; (a)_1 = \widehat{\mathbb{R}} \;\wedge\; (\forall x \in b)(\exists y \in c)R((a)_2, x, y) \;\wedge$

$(\forall y \in c)(\exists x \in b)R((a)_2, x, y)$,

(22) $a = \widehat{\mathbb{C}} \;\wedge\; R(b, c, \widehat{\top}) \;\wedge\; \forall x(x <_{\mathcal{W}} c \rightarrow \neg R(b, x, \widehat{\top})) \;\wedge\; \forall x \neg R(\widehat{\mathbb{C}}, b, x)$,

(23) $a = \widehat{\mathbb{P}} \;\wedge\; \forall x(x \in c \leftrightarrow x \subset b)$,

(24) $a = \widehat{\mathbf{E}} \;\wedge\; \exists x R(b, x, \widehat{\top}) \;\wedge\; c = \widehat{\top}$,

(25) $a = \widehat{\mathbf{E}} \;\wedge\; \forall x R(b, x, \widehat{\bot}) \;\wedge\; c = \widehat{\bot}$.

Clearly, $Q_{\mathfrak{F}}(a, b, c)$ is functional in its third argument. All we have to do now is to follow Section 4.2 again, this time with $Ap_{\mathfrak{A}}[a, b, c]$ replaced by $Q_{\mathfrak{F}}(a, b, c)$. In parallel to Definition 9 an $\mathcal{L}_S$ formula $[\![t]\!]_{\mathfrak{F}}(u)$ is assigned to any $\mathcal{L}^\circ$ term t, saying that u is the value of the term t under the interpretation of the $\mathsf{OST}^r(\mathbf{E}, \mathbb{P})$ and $\mathsf{OST}(\mathbf{E}, \mathbb{P})$ application via $Q_{\mathfrak{F}}$.

Employing these $[\![t]\!]_{\mathfrak{F}}(u)$, each $\mathcal{L}^\circ$ formula A is translated into a formula $A^\diamondsuit$ of $\mathcal{L}_S$ in the obvious way, simply by following Definition 10. Please keep in mind that A and $A^\diamondsuit$ are identical in the case that A is an $\mathcal{L}_1$ formula. With the exception of the treatment of operational $\mathbf{E}$, the following interpretation result is as Theorem 16 and proved in Jäger [30].

Theorem 20 *The theories* $\mathsf{OST}^r(\mathbf{E}, \mathbb{P})$ *and* $\mathsf{OST}(\mathbf{E}, \mathbb{P})$ *are interpretable in* $\mathsf{E}^r(\mathsf{ZFW})$ *and* $\mathsf{E}^r(\mathsf{ZFW}) + (\mathcal{L}_S\text{-}\mathsf{I}_\in)$*, respectively; i.e. for all formulas A of $\mathcal{L}^\circ$ we have:*

1. $\mathsf{OST}^r(\mathbf{E}, \mathbb{P}) \vdash A \quad \Longrightarrow \quad \mathsf{E}^r(\mathsf{ZFW}) \vdash A^\diamondsuit$.
2. $\mathsf{OST}(\mathbf{E}, \mathbb{P}) \vdash A \quad \Longrightarrow \quad \mathsf{E}^r(\mathsf{ZFW}) + (\mathcal{L}_S\text{-}\mathsf{I}_\in) \vdash A^\diamondsuit$.

If we are able to reduce $\mathsf{E}^r(\mathsf{ZFW})$ and $\mathsf{E}^r(\mathsf{ZFW}) + (\mathcal{L}_S\text{-}\mathsf{I}_\in)$ to NBG and $\mathsf{NBG}_{<E_0}$, respectively, the consistency strength of $\mathsf{OST}^r(\mathbf{E}, \mathbb{P})$ and that of $\mathsf{OST}(\mathbf{E}, \mathbb{P})$ are determined. This task is dealt with in Jäger [29, 30] by means of the following conservation result.

Theorem 21 *Let A be a formula of the language $\mathcal{L}_1$. Then we have:*

1. $\mathsf{E}^r(\mathsf{ZFW}) \vdash A \quad \Longrightarrow \quad \mathsf{NBG} \vdash A$.
2. $\mathsf{E}^r(\mathsf{ZFW}) + (\mathcal{L}_S\text{-}\mathsf{I}_\in) \vdash A \quad \Longrightarrow \quad \mathsf{NBG}_{<E_0} \vdash A$.

All together, Theorem 18, Theorem 20 and Theorem 21 complete the analysis of the theories $\mathsf{OST}^r(\mathbf{E}, \mathbb{P})$ and $\mathsf{OST}(\mathbf{E}, \mathbb{P})$.

Corollary 22 *The theory* $\mathsf{OST}^r(\mathbf{E},\mathbb{P})$ *is equiconsistent with* NBG *and* ZFC*, the theory* $\mathsf{OST}(\mathbf{E},\mathbb{P})$ *equiconsistent with* $\mathsf{NBG}_{<E_0}$*. For all formulas* A *of the language* $\mathcal{L}_1$ *we have:*

1. $\mathsf{OST}^r(\mathbf{E},\mathbb{P}) \vdash A \quad \Longleftrightarrow \quad \mathsf{NBG} \vdash A \quad \Longleftrightarrow \quad \mathsf{ZFC} \vdash A.$
2. $\mathsf{OST}(\mathbf{E},\mathbb{P}) \vdash A \quad \Longleftrightarrow \quad \mathsf{NBG}_{<E_0} \vdash A.$

Actually, Jäger [29] introduces a theory ZFL^r_Ω which is closely related to $\mathsf{E}^r(\mathsf{ZFW})$. It is shown there that $\mathsf{OST}^r(\mathbf{E},\mathbb{P})$ can be embedded into ZFL^r_Ω and that ZFL^r_Ω can be reduced to $\mathsf{ZF} + (V = L)$, thus providing a proof of the first assertion of the previous theorem.

References

[1] P. Aczel, *The type theoretic interpretation of constructive set theory*, Logic Colloquium '77 (A. MacIntyre, L. Pacholski, and J. Paris, eds.), Studies in Logic and the Foundations of Mathematics, North-Holland, 1978, pp. 55–66.

[2] ———, *The type theoretic interpretation of constructive set theory: choice principles*, The L.E.J. Brouwer Centenary Symposium (A. S. Troelstra and D. van Dalen, eds.), Studies in Logic and the Foundations of Mathematics, North-Holland, 1982, pp. 1–40.

[3] ———, *The type theoretic interpretation of constructive set theory: inductive definitions*, Logic, Methodology, and Philosophy of Science VII (R. Barcan Marcus, G. J. W. Dorn, and P. Weingartner, eds.), Studies in Logic and the Foundations of Mathematics, North-Holland, 1986, pp. 17–49.

[4] P. Aczel and M. Rathjen, *Notes on constructive set theory*, Tech. Report 40, Institut Mittag-Leffler, 2001.

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

[6] ———, *Proof theory for theories of ordinals II:* Π_3*-reflection*, Annals of Pure and Applied Logic **129** (2004), 39–92.

[7] K. J. Barwise, *Admissible Sets and Structures*, Perspectives in Mathematical Logic, Springer, 1975.

[8] M. J. Beeson, *Foundations of Constructive Mathematics: Metamathematical Studies*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer, 1985.

[9] ———, *Proving programs and programming proofs*, Logic, Methodology, and Philosophy of Science VII (R. Barcan Marcus, G. J. W. Dorn, and P. Weingartner, eds.), Studies in Logic and the Foundations of Mathematics, North-Holland, 1986, pp. 51–82.

[10] ———, *Towards a computation system based on set theory*, Theoretical Computer Science **60** (1988), 297–340.

[11] W. Buchholz, S. Feferman, W. Pohlers, and W. Sieg, *Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies*, Lecture Notes in Mathematics, vol. 897, Springer, 1981.

[12] A. Cantini and L. Crosilla, *Constructive set theory with operations*, draft, 2005.

[13] COQ, `http://coq.inria.fr/`.

[14] S. Feferman, *A language and axioms for explicit mathematics*, Algebra and Logic (J. N. Crossley, ed.), Lecture Notes in Mathematics, vol. 450, Springer, 1975, pp. 87–139.

[15] ———, *Recursion theory and set theory: a marriage of convenience*, Generalized Recursion Theory II, Oslo 1977 (J. E. Fenstad, R. O. Gandy, and G. E. Sacks, eds.), Studies in Logic and the Foundations of Mathematics, North-Holland, 1978, pp. 55–98.

[16] ———, *Constructive theories of functions and classes*, Logic Colloquium '78 (M. Boffa, D. van Dalen, and K. McAloon, eds.), Studies in Logic and the Foundations of Mathematics, North-Holland, 1979, pp. 159–224.

[17] ———, *Gödel's program for new axioms: Why, where, how and what?*, Gödel '96 (Petr Hájek, ed.), Lecture Notes in Logic, vol. 6, Springer, 1996, pp. 3–22.

[18] ———, *Notes on operational set theory, I. Generalization of "small" large cardinals in classical and admissible set theory*, `http://math.stanford.edu/~feferman/papers/OperationalST-I.pdf`, 2001.

[19] ———, *Operational theories of sets and classes*, 2005, draft.

[20] ———, *Operational set theory and small large cardinals*, `http://math.stanford.edu/~feferman/papers/ostcards.pdf`, 2006.

[21] S. Feferman and G. Jäger, *Systems of explicit mathematics with non-constructive μ-operator. Part I*, Annals of Pure and Applied Logic **65** (1993), 243–263.

[22] HOL, `http://www.cl.cam.ac.uk/research/hvg/HOL/`.

[23] G. Jäger, *Die konstruktible Hierarchie als Hilfsmittel zur beweistheoretischen Untersuchung von Teilsystemen der Mengenlehre und Analysis*, Ph.D. thesis, Mathematisches Institut, Universität München, 1979.

[24] ———, *Iterating admissibility in proof theory*, Logic Colloquium '81. Proceedings of the Herbrand Symposion (J. Stern, ed.), North-Holland, 1982, pp. 137–146.

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

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

[27] ———, *The strength of admissibility without foundation*, Journal of Symbolic Logic **49** (1984), 867–879.

[28] ———, *Theories for Admissible Sets: A Unifying Approach to Proof Theory*, Studies in Proof Theory, vol. 2, Bibliopolis, 1986.

[29] ———, *On Feferman's operational set theory* OST, Annals of Pure and Applied Logic **150** (2007), 19–39.

[30] ———, *Full operational set theory with unbounded existential quantification and power set*, 2007, draft.

[31] G. Jäger, R. Kahle, A. Setzer, and T. Strahm, *The proof-theoretic analysis of transfinitely iterated fixed point theories*, Journal of Symbolic Logic **64** (1999), 53–67.

[32] G. Jäger and J. Krähenbühl, Σ^1_1 *choice in a theory of sets and classes*, 2007, draft.

[33] G. Jäger and W. Pohlers, *Eine beweistheoretische Untersuchung von* $(\Delta^1_2\text{-}\mathsf{CA}) + (\mathsf{BI})$ *und verwandter Systeme*, Sitzungsberichte der Bayerischen Akademie der Wissenschaften, Mathematisch-naturwissenschaftliche Klasse (1982), 1–28.

[34] G. Jäger and T. Strahm, *Upper bounds for metapredicative Mahlo in explicit mathematics and admissible set theory*, Journal of Symbolic Logic **66** (2001), 935–958.

[35] ———, *The proof-theoretic strength of the Suslin operator in applicative theories*, Reflections on the Foundations of Mathematics: Essays in Honor of Solomon Feferman (W. Sieg, R. Sommer, and C. Talcott, eds.), Lecture Notes in Logic, vol. 15, Association for Symbolic Logic, 2002, pp. 270–292.

[36] ———, *Reflections on reflections in explicit mathematics*, Annals of Pure and Applied Logic **136** (2005), 116–133.

[37] K. Kunen, *Set Theory. An Introduction to Independence Proofs*, Studies in Logic and the Foundations of Mathematics, North-Holland, 1980.

[38] A. Levy, *The role of classes in set theory*, Sets and Classes. On the Work by Paul Bernays (G.-H. Müller, ed.), Studies in Logic and the Foundations of Mathematics, North-Holland, 1976, pp. 277–323.

[39] P. Martin-Löf, *An intuitionistic theory of types: predicative part*, Logic Colloquium '73 (H. E. Rose and J. Shepherdson, eds.), Studies in Logic and the Foundations of Mathematics, North-Holland, 1975, pp. 73–118.

[40] P. Martin-Löf, *Intutionistic Type Theory*, Studies in Proof Theory, vol. 1, Bibliopolis, 1984.

[41] E. Mendelson, *Introduction to Mathematical Logic*, Chapmann & Hall, 1997 (fourth edition).

[42] J. Myhill, *Constructive set theory*, Journal of Symbolic Logic **40** (1975), 347–382.

[43] Nuprl, `http://www.cs.cornell.edu/info/projects/nuprl/`.

[44] M. Rathjen, *Proof-theoretic analysis of* KPm, Archive for Mathematical Logic **30** (1991), 377–403.

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

[46] ———, *An ordinal analysis of stability*, Archive for Mathematical Logic **44** (2005), 1–62.

[47] ———, *An ordinal analysis of parameter free* Π^1_2*-comprehension*, Archive for Mathematical Logic **44** (2005), 263–362.

[48] W. Richter and P. Aczel, *Inductive denitions and reflecting properties of admissible ordinals*, Generalized Recursion Theory (J. Fenstad and P.Hinman, eds.), North-Holland, 1974, pp. 301–381.

[49] T. Strahm, *First steps into metapredicativity in explicit mathematics*, Sets and Proofs (S. B. Cooper and J. Truss, eds.), Cambridge University Press, 1999, pp. 383–402.

[50] ———, *Wellordering proofs for metapredicative mahlo*, Journal of Symbolic Logic **67** (2002), 260–278.