A New Model Construction by Making a Detour via Intuitionistic Theories I: Operational Set Theory without Choice is Π_1-equivalent to KP[☆]

Kentaro Sato[a,1,*], Rico Zumbrunnen[a,1]

[a] *Institut für Informatik und angewandte Mathematik, Universität Bern, Neubrückstrasse 10, 3012 Bern, Switzerland*

Abstract

We introduce a version of operational set theory, OST^-, without a choice operation, which has a machinery for Δ_0 separation based on truth functions and the separation operator, and a new kind of applicative set theory, so-called weak explicit set theory WEST, based on Gödel operations. We show that both the theories and Kripke-Platek set theory KP with infinity are pairwise Π_1 equivalent. We also show analogous assertions for subtheories with $\in$-induction restricted in various ways and for supertheories extended by powerset, beta, limit and Mahlo operations. Whereas the upper bound is given by a refinement of inductive definition in KP, the lower bound is by a combination, in a specific way, of realisability, (intuitionistic) forcing and negative interpretations. Thus, despite interpretability between classical theories, we make "a detour via intuitionistic theories". The combined interpretation, seen as a model construction in the sense of Visser's miniature model theory, is a new way of construction for classical theories and could be said the third kind of model construction ever used which is non-trivial on the logical connective level, after generic extension à la Cohen and Krivine's classical realisability model.

Keywords: applicative set theory, Feferman's OST, intuitionistic forcing, realisability, inductive definition
2010 MSC: 03F25, 03F50 (Secondary) 03B03, 03B40, 03C65, 03D60, 03E70

1. Introduction

1.1. Operational set theory

Intending to provide a uniform framework for small large cardinal notions used in classical set theory and admissible set theory, Feferman [16] introduced operational set theory OST (see also Feferman [17]). This theory is based on the so-called applicative theory, basically the theory of combinator algebra, shared with explicit mathematics, which Feferman [12, 13, 14] introduced intending to provide a uniform framework for classical mathematics, Russian constructive recursive mathematics and Bishop-style constructive mathematics. For this reason, operational set theory is sometimes considered as a set theoretic analogue of explicit mathematics. Whereas explicit mathematics was originally formulated with intuitionistic logic (naturally for the intention) and has later been investigated in the context of classical logic (for the reason, see Feferman [18, p.3 and p.5]), operational set theory is formulated with classical logic from the beginning. The reason is that Feferman [17] explicitly restricted the intention to classical set theory and admissible set theory "to begin with", although small large cardinal notions used in admissible recursion theory, constructive set theory, constructive type theory, explicit mathematics and recursive ordinal notations are also in his ultimate scope (see [17]), and actually Cantini and Crosilla [9, 10] considered similar theories with intuitionistic logic.

In operational set theory, all the objects are sets and, at the same time, operations, in the sense of the additional binary function symbol $\circ$ called application, which are not total in general. Thus we have two

[☆] The main part of this paper is presented in Zumbrunnen's doctoral dissertation [59], defended on December 2013.
[*] Corresponding author
Email addresses: sato@iam.unibe.ch (Kentaro Sato), rico.zumbrunnen@iam.unibe.ch (Rico Zumbrunnen)
[1] The research is partially supported by Swiss National Science Foundation.

notions of function, set-theoretic functions represented by graphs and operations which form a non-total combinator algebra with $\circ$. Among the features of operational set theory are Δ_0 truth function, uniform Δ_0 separation and uniform operational collection as well as uniform choice operation. By the first, we mean that for any Δ_0 formula $A[x, y]$ of the language of pure set theory (not extended by $\circ$ and others) there is an operation (i.e., a set) c_A such that $(c_A \circ y) \circ x \simeq \mathsf{t}$ if $A[x, y]$ holds and $(c_A \circ y) \circ x \simeq \mathsf{f}$ otherwise. By the second we mean an operation $\mathbb{S}$ such that for any operator f from a set a to the doubleton $\{\mathsf{t}, \mathsf{f}\}$

$$(\mathbb{S} \circ f) \circ a = \{x \in a : f \circ x = \mathsf{t}\}.$$

By the third, we mean an operation $\mathbb{R}$ such that for any operation f totally defined on a set b

$$(\mathbb{R} \circ f) \circ b = \{f \circ x : x \in b\}.$$

And by the last, we mean an operation $\mathbb{C}$ such that if $f \circ x = \mathsf{t}$ for some x then $f \circ (\mathbb{C} \circ f) = \mathsf{t}$.

1.2. Main question

The last feature, the choice operator, is a quite strong form of the axiom of choice, because there is no restriction on the domain, like global choice, and moreover there is no precondition, such as local totality, on the truth function f. Even so, in previous proof theoretic investigations of OST and of the extensions by powerset operation $\mathbb{P}$ and unbounded existential quantifier operation in Jäger [29, 30], he embeds OST and the extensions in Kripke-Platek set theory KP with infinity and corresponding extensions, with the axiom of constructibility $V = L$. Since most of the corresponding extensions can interpret $V = L$ within themselves, it has previously been considered that the strength of choice operator does not matter. However, as shown by Mathias [44], $\mathsf{KP} + (\mathcal{P})$, the extension corresponding to $\mathsf{OST} + (\mathbb{P})$, gets strictly stronger if added to with $V = L$, as opposed to KP or ZF. More recently, Jäger [31] characterises the notion of so-called operational closure, an operational version of admissibility, and the limit axiom based on this notion, with stability and Σ_1 separation in KP, by a proof seemingly based on the unnatural strength of choice operation.

Now it is a natural question if OST^-, defined by dropping the axiom for choice operation from OST, is (proof theoretically) strictly weaker or not. The main purpose of this paper is to give an answer to this. We will see that OST^- has the same proof theoretic strength as OST, and moreover that OST^- is closer to KP than OST, in the sense that OST^- and KP prove the same Π_1 sentences whereas we only know that OST and KP prove the same Δ sentences (or more precisely, the sentences absolute for the constructible hierarchy).

1.3. Lower bound proof

If we are working in OST^-, we cannot deduce the usual Δ_0 collection schema: for a Δ_0 formula $A[x, y]$,

$$(\forall y \in a)\exists x A[x, y] \to \exists b(\forall y \in a)(\exists x \in b)A[x, y].$$

With the choice operator, we can deduce it as follows: taking c_A as mentioned above, from $(\forall y \in a)\exists x A[x, y]$ we can imply $(\forall y \in a)\exists x((c_A \circ y) \circ x = \mathsf{t})$ and $(\forall y \in a)A[\mathbb{C} \circ (c_A \circ y), y]$, whence $(\mathbb{R} \circ (\lambda y.\mathbb{C} \circ (c_A \circ y))) \circ a$ is what we require. This deduction of the usual collection schema from operational collection is the key of the lower bound proof of OST, or more precisely the proof of the interpretability of KP in OST, given both in Feferman [17] and in Jäger [29]. Thus the first sub-question to ask is: can we interpret collection schema in the operational set theory OST^- without choice operator and, if so, how?

In the present paper, we answer the question positively by giving a concrete interpretation of KP in OST^-. First we extract the usual collection schema from operational collection at the cost of classical logic. More precisely, we interpret a version IKP^- of intuitionistic Kripke-Platek set theory in OST^- by realisability interpretation (Section 7). The realisability notion we will use is quite similar to what Feferman [14, IV] introduced in explicit mathematics. Once we can reach at intuitionistic Kripke-Platek set theory, we can follow a general method, due to Avigad [2], to interpret classical theories in constructive ones, with a significant amount of modifications needed for our settings. More concretely, we interpret both a weaker version $\mathsf{IKP}^\sharp$ of intuitionistic Kripke-Platek set theory and so-called Markov's principle in IKP^- by so-called (intuitionistic) forcing interpretation (Section 6), and finally interpret classical but intensional KP^{int} in $\mathsf{IKP}^\sharp$ with Markov's principle by negative interpretation (Section 5). While Avigad combines them with one more interpretation of classical extensional KP in intensional KP^{int}, we do not need this process because we can add the negative interpretation of the axiom of extensionality (N-Ext) to both the intuitionistic theories. Figure 1 shows how we interpret KP in OST^-, where N denotes Gödel-Gentzen's negative interpretation (Definition 35), $\Vdash_{\mathfrak{S}}$ Avigad's forcing interpretation (Definition 56) (which is, strictly speaking, only a local interpretation between described theories, in the sense that $\mathfrak{S}$ depends on proofs to be interpreted; but which can be non-local if we replace theories suitably) and $\mathfrak{r}$ Feferman's realisability interpretation (Definition 77).

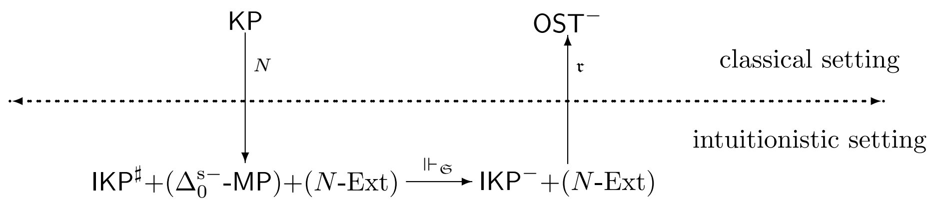

Figure 1: The interpretation process for lower bound

1.4. As a new model construction

Our method for lower bound proof, the combination of three interpretations, is a new kind of model construction, in the following sense. A syntactic interpretation of a theory T_1 in another theory T_2 can be seen as a procedure converting a model of the interpreting theory T_2 into that of the interpreted theory T_1 (or a model of arbitrarily large fragments of T_1, if the interpretation is local) or a way of constructing a model of T_1 (or of arbitrarily large fragments) within in the meta-theory T_2. The meta-theory T_2 could be much weaker, than those employed in the usual model theory (e.g., ZFC). From this viewpoint, theory of interpretation is called *miniature model theory* in Visser [58]. However Visser [58] considers only those interpretations by which the meanings of logical connectives and quantifiers are not changed but only the domain of quantifiers and the meaning of atomic formulae are changed. We call such interpretations *logically trivial*. Among them is the standard interpretation of number theory in set theory, by finite von Neumann ordinals. However, forcing does not satisfy this condition, because for example $\Vdash A \vee B$ is not equivalent to $(\Vdash A) \vee (\Vdash B)$, while we know that forcing is, or at least can be seen as, a model construction. (Actually, as is known and as shown by Sato [47], many applications of forcing method remain possible if we treat the method as a syntactic interpretation; especially we do not need so-called Truth Lemma; notice that a generic extension cannot be a transitive submodel of an original model in any sense.) Thus it is reasonable to include those interpretations which change the meanings of logical connectives and of quantifiers, like forcing interpretation, in the scope, and we call such interpretations or model construction methods *logically non-trivial*. (However, we do not include all the Boolean homomorphisms, as opposed to in Pour-El and Kripke [45]: because they do not have any uniformity on quantifiers, they do not seem to be model constructions.)

It seems that forcing has previously been, practically, the only logically non-trivial way of model construction for theories based on classical logic. Indeed, the inner model method, another major way in set theory, is logically trivial. Even if we extend the scope to the interpretations between intuitionistic theories, there have been only a few logically non-trivial ones, which are, up to the authors' knowledge,

(a) negative interpretations of classical theories in intuitionistic theories, that due to Gödel and Gentzen as well as the variants by Kolmogorov and by Kuroda (for the definitions of the three, see [54, Section 2.3] or [53, Section 2.3]);

(b) realisability interpretation of intuitionistic theories in a broad sense (including *Dialectica interpretation* due to Gödel), which can be seen as a formalisation of Brouwer-Heyting-Kolmogorov interpretation;

(c) less known, intuitionistic forcing interpretation of intuitionistic theories, which can be seen as a straightforward formalisation of Kripke semantics and among which is the forcing that Avigad [2] uses.

Actually, as is known (or as pointed out by Avigad [3, Section 2]), classical forcing (i.e., forcing interpretation for classical theories) can be seen as a combination of (a) and (c). Recently, Krivine [41, 42] has developed a new logically non-trivial model construction for classical theories, and this can be seen as a combination of (a), A-translation due to Friedman [20][2] and (b),[3] where A-translation itself is a logically trivial.[4]

The interpretation we will give is, as explained before, the combination of all of (a), (c) and (b) in this order, and therefore it is different from the previously known two logically non-trivial interpretations for classical theories.[5] Because there seems to be no other practically used logically non-trivial interpretation, our method can be said the third kind of logically non-trivial model construction ever practically used for classical theories. It might be worth mentioning that Beeson [5, 6] and Gordeev [22, 23] use a combination of (b) and (c) in a different way (though in their uses forcing is slightly different from the straightforward formalisation of Kripke semantics of intuitionistic logic), but the final combinations are for intuitionistic theories. Moreover, since cut-elimination method is also combined in Gordeev's results, his final results are reductions rather than interpretations or model construction methods. To repeat, up to the present there seem to be only three kinds of logically non-trivial interpretations or model construction methods for classical theories, the last of which is what we will introduce in the present paper:

(1) forcing interpretation, à la Cohen, a combination of (a) and (c);

(2) Krivine's classical realisability model, a combination of (a) and (b);

(3) our interpretation, a combination of all (a), (c) and (b).

Since both (b) and (c) are quite general methods in the sense that we have variety of interpretations by setting up different "parameters" in the general frameworks of (b) and (c) (e.g., the formulae to be the interpretations of atomic formulae in the both, preorders in (c)), so are the three kinds of interpretations (1)-(3) above for classical theories. As there are many kinds of forcing used in set theory and proof theory, we can define many instances of our method (as well as Krivine's classical realisability model) only by changing the "parameters", and what we will see is only one instance.

1.5. Utility of intuitionistic logic

As discussed above, all the ever used logically non-trivial interpretations between classical theories are combinations of such interpretations between intuitionistic theories. There seems to be no logically non-trivial interpretations or model construction methods ever practically used which is free from intuitionistic logic. This fact might be considered to suggest the special status of intuitionistic logic among many non-classical logics, besides many philosophical and computational motivations for intuitionistic logic.

At least, we could say that our interpretation of the kind (3) is invented by:

- first factoring out (a), (b) and (c) from the previously known two logically non-trivial interpretations (1) and (2) between classical theories, and
- then rearranging the factors (a), (b) and (c) in a different order.

This insight becomes possible only when we extend the scope to intuitionistic theories, since the factors (a), (b) and (c) are for intuitionistic theories by nature. However, the interpretation we finally obtain is between classical theories. Thus we can support the view that intuitionistic logic is a refinement of classical logic, and that this refinement is quite useful even for studies of classical logic and of theories on classical

[2] In some references, the combination of (a) and A-translation in our sense is referred to as A-translation. This combination is basically the negative interpretation with $\perp$ replaced by any fixed formula.

[3] As Krivine pointed out, Krivine's classical realisability generalises classical forcing, because (b) and (c) are almost the same on the negative connectives and because Gödel-Gentzen's negative interpretation converts all connectives into negative ones.

[4] Though the meaning of the logical connective $\perp$ is changed, since $\perp$ is an atomic formula we consider it logically trivial. It is also common, especially in constructive contexts, to consider $\perp$ as the atomic formula $0 = 1$, not as a logical connective.

[5] Actually, our combination generalises Krivine's classical realisability, since A-translation is an instance of (c) intuitionistic forcing with a trivial order. More generally, any logically trivial interpretation is an instance of (c) with a trivial order.

logic. Another application of intuitionistic logic to the study of classical theories, which seems to support this view, is the combination of realisability and the conservation on the geometric fragment, invented by Strahm [51, 52]. Up to now, there seems to be no other non-classical logic proved to be useful in this sense.

From this point of view, we can see our combination of the three interpretations in Figure 1 differently. As will be emphasized in Section 6, Avigad's forcing interpretation $\Vdash_{\mathfrak{S}}$ that we will use becomes trivial in the classical context and hence meaningful only in the intuitionistic (namely, non-classical) context. We could say that the other two interpretations, namely Feferman's realisability interpretation $\mathfrak{r}$ and Gödel-Gentzen's negative interpretation N, are used to make this forcing interpretation meaningful: they switch the classical and intuitionistic situations back and forth. Actually the two interpretations are reversal with each other in a weak sense, namely the combination of them results in the identity or trivial interpretation between classical theories (but the other combination, resulting in an interpretation between intuitionistic ones, is not trivial). In other words, by the pair of Gödel-Gentzen's negative interpretation N and Feferman's realisability interpretation $\mathfrak{r}$, we can "lift" Avigad's forcing interpretation $\Vdash_{\mathfrak{S}}$ into the classical setting, without loosing the meaningfulness.[6] This is the essential of our making "a detour via intuitionistic theories". Here it seems necessary to note that this remark applies only to our particular choices from (a)-(c), namely Gödel-Gentzen's negative interpretation N, Feferman's realisability interpretation $\mathfrak{r}$ and Avigad's forcing interpretation $\Vdash_{\mathfrak{S}}$, but not to general interpretations categorised in (a)-(c).

1.6. Weak explicit set theory

It will turn out that, for our interpretation of KP, not all the machineries featured in OST^- are necessary. Particularly, the truth functions are not essential. The essentials are operational collection and uniform Δ_0 separation. For this reason, we define *weak explicit set theory* WEST directly by uniform Δ_0 separation and by operational collection, and we proceed to the lower bound proof for this theory. Since WEST is a subsystem of OST^-, it is easier to apply, in order to obtain lower bound proofs of other theories.

WEST has, not only such a technical utility, but also some conceptual advantages over the operational set theory OST and its choice-less version OST^-. For example, since it does not feature the truth functions, it is closer to explicit mathematics which does not feature them either. As a consequence, it becomes more natural and easier to define intuitionistic versions of the system, as explicit mathematics. (Notice that the machinery of the truth functions heavily depends on the underlying logics: in OST, the two truth values t and f explicitly occur in the formulation.)

The origin of the name of WEST is the fact that it is closer to explicit mathematics and that it will turn out to be much weaker than the most common system T_0 of explicit mathematics. Actually, as we will see, WEST is Π_1 equivalent to KP, whereas Jäger [24] has shown that T_0 is equiconsistent with $\mathsf{KP} + (\text{Beta})$.

1.7. Extensions to sub- and super-systems

In explicit mathematics, various subtheories and supertheories of the most common system T_0 have been investigated. Among the subtheories are those defined by restricting the induction schema to various classes of formulae and by removing or restricting the fixed-point induction schema associated with the so-called inductive generation operator, for example EETJ_0. Among the supertheories are, on the other hand, those defined by adding a new predicate for universes and some operations which guarantee that the domain of discourse has a large cardinal property in some sense, for example limit axiom and Mahlo axiom (see Jäger, Kahle and Studer [32], Jäger and Studer [37], Jäger and Strahm [35, 36] and Jäger [28]).

It is natural to consider similar subtheories and supertheories of OST, of OST^- and of WEST. Among them are OST^-_0 defined by omitting $\in$-induction schema completely, OST^-_r by restricting $\in$-induction schema to Δ_0 formulae, OST^-_w by restricting $\in$-induction schema to Δ_0 formulae and by adding full induction on natural numbers, the supertheory defined by adding an operation returning an admissible set above, and that by adding so-called Mahlo operation.

Our method for lower bound will turn out to be so versatile that it yields lower bounds of such subtheories and supertheories, as well as the supertheories defined by powerset and beta (Mostowski collapsing)

[6] In the case of Krivine's classical realisability, what is "lifted" by the pair is A-translation, as explained in footnote 5.

operations, of WEST (and hence of OST^-). The proof for these variants, as well as for the original WEST, is quite uniform, and the treatment for additional axioms is completely modular. Therefore we can obtain the lower bound result for the theory defined on the weakest OST_0^- or WEST_0 by adding any combination of the axioms we will consider (Theorem 99 and Table 2).

Such a modularity could be said a benefit from the (local) interpretability result. In general for proof theoretic reduction results (typically based on cut-elimination), on the other hand, such a modularity does not hold, because there might be an unhappy interaction among additional axioms.

1.8. Upper bound proof

Since OST^- is a subsystem of OST, the upper bound proof of OST, previously given, yields also that of OST^-, and hence, together with our lower bound proof, it shows that OST^- is proof theoretically equivalent to Kripke-Platek set theory KP with infinity. Nevertheless, because of the absence of choice operation $\mathbb{C}$, the upper bound proof becomes simpler and, as a result, yields a stronger result. Especially, we do not need the axiom of constructibility $V = L$ to incorporate $\mathbb{C}$ and hence we can interpret OST^- more directly in KP. Actually, the (logically trivial) interpretation of OST, given by Jäger [29], relies on $V = L$, and therefore it is necessary to combine it with an additional (logically trivial) interpretation of $\mathsf{KP} + (V = L)$ in KP, which preserves only Δ sentences. Without choice operation $\mathbb{C}$, on the other hand, we can (logically trivially) interpret OST^- in KP, only by defining the interpretation of the application $\circ$, with Σ inductive definition, and the resulting interpretation preserves all formulae of the language of pure set theory.

Moreover, as we will see, the required inductive definition is available in a subtheory KP_r, which is defined from KP by restricting $\in$-induction schema to Δ_0 formulae (Subsection 9.1). Together with the simplicity from the absence of $V = L$, our upper bound proof also turns out to be so versatile that it yields upper bound of the aforementioned subtheories (but above OST_r^-) and supertheories of OST^- (and therefore of WEST) uniformly. Thus the modularity is available, similarly to the lower bound proof.

Here it is worth mentioning that the Σ inductive definition in KP_r interprets various kinds of inductive definition in corresponding variants of KP uniformly (Remark 103). With some small modifications, it can also interpret variants of the system PA_Ω of numbers and ordinals from Jäger [27] in corresponding variants of KP (Remark 104).

It is also worth mentioning that the composition of the combined interpretation given in the lower bound proof and this inductive model construction yields a model of (variants of) KP within (the same variants of) KP, in the same sense as classical forcing and Krivine's classical realisability yield models of ZF within ZF. Although in the present paper we do not enter into the details of this new model, it is plausible that this would be useful to show the independence or equiconsistency of some mathematical assertions on KP and variants, in the same way as classical forcing and Krivine's classical realisability are.

1.9. Main results

The main results of the present paper follows from the aforementioned lower and upper bound proofs. Namely, both OST^- and WEST are Π_1 equivalent to KP, and this Π_1 equivalence is robust in the sense that it remains true if we restrict $\in$-induction in various ways (but we cannot omit $\in$-induction for Δ_0 formulae) and if we add the aforementioned additional axioms (Corollary 115, and Table 4).

1.10. Results for intuitionistic versions of OST^- *and* WEST

In the lower bound proof, even though we finally obtain an interpretation between classical theories, intuitionistic theories come in the argument. Then it is natural to expect that our method can apply also to intuitionistic versions of OST^- or of WEST, while the proof theoretic investigation of these versions is also well-motivated, from the original intention of operational set theory as explained at the beginning.

However, our method does not work so well for intuitionistic theories as for classical ones. In Appendix, we will consider the reason of this and we will see how restricted the obtainable results for intuitionistic theories are. For this reason, we could claim that our method is mainly for classical theories, despite its intuitionistic nature in the intermediate stages. This seems to support our aforementioned claim: intuitionistic logic is a refinement of classical logic, and this refinement is quite useful even for the studies of classical logic and theories on classical logic.

Part I
Preliminaries

2. Theories OST, OST^- and WEST

In this section, we define the operational set theory OST and variants, from the subsystem with $\in$-induction omitted completely to the supersystems with limit and Mahlo axioms, in a way suitable for our purpose but slightly different from the original formulation in Feferman [16] and from the following formulation in Jäger [29]. The main purpose of the present paper is to investigate the proof theoretic strengths of these systems without choice operation. We also introduce another kind of applicative set theory, called weak explicit set theory, which has a machinery for Δ_0 separation based on Gödel operations, whereas operational set theory has the machinery based on truth functions. We can also consider subsystems and supersystems of weak explicit set theory. These two families are the targets of our proof theoretic investigation in the following sections.

2.1. Language of pure set theory

By $\mathcal{L}_\in$ we denote the language of set theory with the constant ω and by $\mathcal{L}$ we denote a relational extension of it.

Definition 1. The language $\mathcal{L}$ contains only constant ω, no function symbols, variables a, b, c, d, f, g, h, m, n, p, q, u, v, w, x, y, z, ... (possibly with subscripts), the binary relation symbols $=$ and $\in$, countably infinitely many relation symbols R_i of different arities, the connectives $\bot$, $\wedge$, $\vee$ and $\rightarrow$ as well as the quantifiers $\forall$ and $\exists$. We write $\mathcal{L}_\in$ for the *language of set theory*, that is $\mathcal{L}$ but without the relation symbols R_i.

Because we will also work with intuitionistic theories, we consider $\bot$ as a logical symbol, and negation is defined as $\neg A := A \rightarrow \bot$.

The bounded quantifiers $(\forall x \in a)$ and $(\exists x \in a)$ are abbreviations and are defined as usual.

Definition 2. A formula of $\mathcal{L}$ is called Δ_0 *formula* if it does not contain unbounded quantifiers. The classes of Δ, Π, Σ, Π_n and Σ_n formulae (for each natural number n) are defined by prenex forms as usual with respect to this definition of Δ_0 formulae.

Furthermore we also use standard set theoretic notations. For example, the formulae $x = \{x_0, ..., x_n\}$, $x = \langle y, z\rangle$, $x \subseteq y$, $x = \cup y$, $x = y \cup z$ and so on are abbreviations of Δ_0 formulae of $\mathcal{L}$ that are defined as usual (except when otherwise mentioned, i.e., in Section 6). For instance, by $x = \{x_0, ..., x_n\}$ we denote the Δ_0 formula

$$x_0 \in x \wedge ... \wedge x_n \in x \wedge (\forall v \in x)(v = x_0 \vee ... \vee v = x_n).$$

We write $\mathrm{Tran}[a]$ for the Δ_0 formula $(\forall x \in a)(x \subseteq a)$, expressing that a is *transitive*.

For each standard natural number n greater than 0 we write $\mathrm{Tup}_n[a]$ for a Δ_0 formula formalising that a is an ordered n-tuple and $(a)_i = b$, $(a)_i \in b$ ... for Δ_0 formulae formalising that its i-th component is b, is an element of b ... (for the definitions of these formulae see Barwise [4, Chapter I]).

We will write $\mathrm{Rel}[f]$ and $\mathrm{Fun}[f]$ for Δ_0 formulae expressing that f is a set theoretic relation and function, respectively, $\mathrm{Dom}[f, a]$ for the Δ_0 formula expressing that the domain of f is a, $\mathrm{Ran}[f, a]$ for Δ_0 formula expressing that the range of f is a subset of a, and $f'x = y$ for a Δ_0 formula which expresses that the set theoretic function f applied to x is y (for their precise definitions see Barwise [4, Chapter 1]).

2.2. Language of operational set theory

Definition 3. The language $\mathcal{L}^\circ_\in$, is the language $\mathcal{L}_\in$ (which include the constant ω) extended by the constants k, s, t, f, el, non, dis, e, $\mathbb{S}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{K}$, $\mathbb{T}$, $\mathbb{D}$, $\mathbb{U}$, $\mathbb{P}$, $\mathbb{B}$, $\mathbb{A}$, $\mathbb{M}$, $\mathbb{G}_1$, $\mathbb{G}_2$, $\mathbb{G}_3$, $\mathbb{G}_4$, $\mathbb{G}_5$ $\mathbb{G}_=$ and $\mathbb{G}_\in$, the unary relation symbol $\downarrow$ (called the *definedness* predicate) and the binary function symbol $\circ$ called *application*. We write $\mathcal{L}^\circ$ for the analogous extension of $\mathcal{L}$ with ch_R and $\mathbb{G}_R$ for each relation symbol R of $\mathcal{L}$ other than $\in$ and $=$.

We call terms without any occurrence of variables *closed terms*. For terms s, t, t_0, ... , t_n we will often write st for $\circ(s,t)$, and $st_0...t_n$ as well as $s(t_0,...,t_n)$ for $((...((st_0)t_2)...)t_n)$. Furthermore we introduce partial equality $\simeq$ as follows:

$$(s \simeq t) \quad := \quad ((s\downarrow \vee t\downarrow) \to s = t).$$

For an arbitrary $\mathcal{L}^\circ$ formula $A[x]$ and an $\mathcal{L}^\circ$ term t we will write sometimes $t \in \{x : A[x]\}$ for $t\downarrow \wedge A[t]$. In this context we will write $\mathbf{B}$ for $\{x : x = \mathsf{t} \vee x = \mathsf{f}\}$ and $\mathbf{V}$ for $\{x : x\downarrow\}$.

Definition 4. For each standard natural number n and variables a, b, f, x, $x_1, ..., x_{n+1}$ we define

(i) $(f : a \to b) := (\forall x \in a)(fx \in b)$,
(ii) $(f : a^{n+1} \to b) := (\forall x_0, ..., x_n \in a)(f(x_0, ..., x_n) \in b)$.

The variables a and/or b may be replaced by $\mathbf{V}$ and/or $\mathbf{B}$.

We call f an *operation* from a^{n+1} to b if $f : a^{n+1} \to b$ holds. This notion should not be confused with that of set theoretic function: f is called a *set theoretic function* from a^{n+1} to b if $\mathrm{Fun}[f]$, $\mathrm{Dom}[f, a^{n+1}]$ and $\mathrm{Ran}[f, b]$ hold. Note that in our extended language, both notions can be defined.

Definition 5. The Δ_0, Δ, Π, Σ, Π_n and Σ_n formulae of $\mathcal{L}^\circ$ are the Δ_0, Δ, Π, Σ, Π_n and Σ_n formulae of $\mathcal{L}$ but they can contain constants of $\mathcal{L}^\circ$ (but not the function symbol $\circ$ nor the relation symbol $\downarrow$).

2.3. Axioms of operational set theory OST *and* OST$^-$

The logic of OST and OST$^-$ is the *logic of partial terms* due to Beeson [7] but the classical version (which treats $\downarrow$ as a logical symbol) including the common equality axioms for $=$. In this logic each of the formulae $(st)\downarrow$, $(s = t)$ and $(s \in t)$ implies $s\downarrow$ as well as $t\downarrow$.

There are four groups of non-logical axioms of OST and OST$^-$. The axioms of the first three groups are common for both theories.

The so called *applicative axioms* are standard axioms about the combinators k and s:

(I.1) $\mathsf{k}ab = a$,

(I.2) $\mathsf{s}ab\downarrow \wedge \mathsf{s}abc \simeq (ac)(bc)$.

The axioms of the second group, the so called *basic set theoretic axioms*, are standard set theoretic axioms as follows.

(II.1) Axiom of extensionality: $a = b \leftrightarrow (\forall x \in a)(x \in b) \wedge (\forall x \in b)(x \in a)$.

(II.2) Axiom of infinity: $\mathrm{Ind}[\omega] \wedge (\forall x \subseteq \omega)(\mathrm{Ind}[x] \to \omega \subseteq x)$, where $\mathrm{Ind}[x] \equiv \emptyset \in x \wedge (\forall y \in x)(y \cup \{y\} \in x)$.

(II.3) $\in$-induction for arbitrary formulae $A[x]$ of $\mathcal{L}^\circ$: $\forall x((\forall y \in x)A[y] \to A[x]) \to \forall x A[x]$.

The third group of axioms are called *logical operations axioms* or truth function axioms. Axioms in this group describe the truth function of the element relation, the additional relations R_i, the connectives negation and disjunction as well as bounded existential quantification as operations:

(III.1) $\mathsf{t} \neq \mathsf{f}$,

(III.2) $(\mathsf{el} : \mathbf{V}^2 \to \mathbf{B}) \wedge \forall x \forall y(\mathsf{el}(x, y) = \mathsf{t} \leftrightarrow x \in y)$,

(III.3) $(\mathsf{ch}_R : \mathbf{V}^k \to \mathbf{B}) \wedge \forall x_0, ..., x_{k-1}(\mathsf{ch}_R(x_0, ..., x_{k-1}) = \mathsf{t} \leftrightarrow R(x_0, ..., x_{k-1}))$ for each relation symbol R of arity k,

(III.4) $(\mathsf{non} : \mathbf{B} \to \mathbf{B}) \wedge (\forall x \in \mathbf{B})(\mathsf{non}(x) = \mathsf{t} \leftrightarrow x = \mathsf{f})$,

(III.5) $(\mathsf{dis} : \mathbf{B}^2 \to \mathbf{B}) \wedge (\forall x, y \in \mathbf{B})(\mathsf{dis}(x, y) = \mathsf{t} \leftrightarrow (x = \mathsf{t} \vee y = \mathsf{t}))$,

(III.6) $(f : a \to \mathbf{B}) \to (\mathsf{e}(f, a) \in \mathbf{B} \wedge (\mathsf{e}(f, a) = \mathsf{t} \leftrightarrow (\exists x \in a)(fx = \mathsf{t})))$.

The last axioms, the *operational set theoretic axioms*, are given as follows.

(IV.1) Kleene star: $\forall x(\mathbb{K}(x)\downarrow \wedge \mathbb{K}(x) = \{u : (\exists n \in \omega)(\mathrm{Fun}[u] \wedge \mathrm{Dom}[u, n] \wedge \mathrm{Ran}[u, x])\})$.

(IV.2) Transitive closure: $\forall x(\mathbb{T}(x)\downarrow \wedge x \subseteq \mathbb{T}(x) \wedge \mathrm{Tran}[\mathbb{T}(x)] \wedge (\forall y)(x \subseteq y \wedge \mathrm{Tran}[y] \to \mathbb{T}(x) \subseteq y))$.

(IV.3) Doubleton (or unordered pair): $\forall x\forall y(\mathbb{D}(x, y)\downarrow \wedge \mathbb{D}(x, y) = \{x, y\})$.

(IV.4) Union: $\forall x(\mathbb{U}(x)\downarrow \wedge \mathbb{U}(x) = \cup x)$.

(IV.5) Separation for definite operations: $(f : a \to \mathbf{B}) \to (\mathbb{S}(f, a)\downarrow \wedge \forall x(x \in \mathbb{S}(f, a) \leftrightarrow (x \in a \wedge fx = \mathsf{t})))$.

(IV.6) Replacement: $(f : a \to \mathbf{V}) \to (\mathbb{R}(f, a)\downarrow \wedge \forall x(x \in \mathbb{R}(f, a) \leftrightarrow (\exists y \in a)(x = fy)))$.

(IV.7) Choice: $\exists x(fx = \mathsf{t}) \to f(\mathbb{C}f) = \mathsf{t}$.

OST contains all the axioms (IV.1)-(IV.7) of this last group whereas OST^- contains only (IV.1)-(IV.6).

The axioms (IV.1) and (IV.2) are redundant, since there are closed terms, built without $\mathbb{K}$ and $\mathbb{T}$ for forming Kleene star and transitive closure. However, they will be essential in the formulations of subsystems of OST^- in the next subsection.

The system OST was originally formulated in Feferman [16, 17] without the axioms for the operations $\mathbb{K}$, $\mathbb{T}$, $\mathbb{D}$ and $\mathbb{U}$ but with set theoretical axioms which state the existence of unordered pairs and unions. Feferman proved that there are closed terms of his system for forming unordered pairs and unions, respectively (c.f. Feferman [17, Corollary 2]). Therefore, with the aforementioned fact on $\mathbb{K}$ and $\mathbb{T}$, our formulation is equivalent to Feferman's.

The next theorem implies that OST and Kripke-Platek set theory KP with infinity (which will be introduced in the next section) are mutually interpretable and have the same proof theoretic strength. Different proofs of this theorem can be found in Feferman [16, 17] and in Jäger [29].

Theorem 6. The theories OST and KP prove the same absolute sentences of $\mathcal{L}_\in$.

The essential of the lower bound proof is that, by making the use of choice operator $\mathbb{C}$, we can imply the set theoretic collection schema from the operational replacement (see Subsection 1.1). For OST^- without choice, this proof does not work. The main purpose of the present paper is to give a lower bound proof for OST^-, which also gives an alternative lower bound proof of OST immediately.

2.4. Operational set theory with restricted $\in$-induction

Definition 7. The theory OST^-_0 is OST^- but without $\in$-induction and OST^-_ω is OST^-_0 with in addition $\in$-induction restricted to ω:

$$(\forall x \in \omega)((\forall y \in x)A[y] \to A[x]) \to (\forall x \in \omega)A[x].$$

Furthermore, we write OST^-_r for OST^- but $\in$-induction restricted to Δ_0 formulae (Δ_0 $\in$-induction) and OST^-_w for OST^-_r with in addition $\in$-induction restricted to ω. We should be careful about the difference between the subscripts ω and w.

We call the following axiom called *operational $\in$-induction*:

$$(\mathsf{opInd}) \qquad \forall x((\forall y \in x)f(y)\downarrow \to f(x)\downarrow) \to \forall x f(x)\downarrow.$$

Furthermore we write (opInd_ω) for the same axiom but restricted to ω.

Although we remove all the induction axioms completely from OST^-_0, by the axiom of infinity and Δ_0 separation, it proves all the instances of $\in$-induction restricted both to Δ_0 and to ω. For the readers' convenience, Table 1 summarises the meanings of subscripts.

Note that the axioms (IV.1) and (IV.2) are redundant in the formulations of OST^-, OST^-_ω and OST^-_w, because, without $\mathbb{K}$ and $\mathbb{T}$, we can construct closed terms for forming Kleene star and transitive closure. However, they are not redundant in the formulations of OST^-_0 and OST^-_r, since we need some induction (actually operational induction) on ω to prove the totality of such operations.

subscript	0	ω	r	w	(nothing)
$\in$-induction restricted to Δ_0 and to ω	yes	yes	yes	yes	yes
$\in$-induction restricted to ω	no	yes	no	yes	yes
$\in$-induction restricted to Δ_0	no	no	yes	yes	yes
unrestricted $\in$-induction	no	no	no	no	yes

Table 1: The meanings of subscripts

2.5. Weak explicit set theory WEST

Operational set theory has sometimes been considered as a set theoretic counterpart of explicit mathematics (see Feferman [12, 13, 14]). Nevertheless, they differ from each other in many respects as investigated recently by Jäger and Zumbrunnen [39]. One of the biggest differences for our purpose is truth function. Operational set theory has constants and axioms (in the third group) for truth functions, which give the truth values of any given Δ_0 formulae (Lemma 11), and these, together with $\mathbb{S}$, yield uniform Δ_0 separation. Explicit mathematics, on the other hand, has no such machinery but it has name-constructors corresponding to Gödel operations, which give the analogue of Δ_0 separation (see Jäger, Kahle and Studer [32], Jäger and Studer [37], Jäger and Studer [36, 35], and Jäger [28]). Indeed, the treatment by truth functions seems essential for the choice operator, which is also a machinery featured in operational set theory but not in explicit mathematics.

However, we are now dealing with operational set theory without the choice operator. In this context, it might be more natural to design a theory of sets based on Gödel operations like explicit mathematics, rather than on truth functions, and it would definitely be closer to explicit mathematics. That is what we are introducing now and is called *weak explicit set theory*. It would be interesting to extend the ontological comparison in Jäger and Zumbrunnen [39] to weak explicit set theory.

Definition 8. The logic of WEST is the classical logic of partial terms, the same as that of OST and OST^-. The non-logical axioms of WEST are $(I.1)$-$(I.3)$, $(II.1)$-$(II.3)$, $(IV.1)$-$(IV.4)$, $(IV.6)$ and the following.

(V.1) Domain: $\forall x(\mathbb{G}_1(x)\downarrow \wedge \mathbb{G}_1(x) = \{v : \exists w(\langle v, w\rangle \in x\})$.

(V.2) Range: $\forall x(\mathbb{G}_2(x)\downarrow \wedge \mathbb{G}_2(x) = \{w : \exists v(\langle v, w\rangle \in x\})$.

(V.3) Difference: $\forall x, y(\mathbb{G}_3(x, y)\downarrow \wedge \mathbb{G}_3(x, y) = x \setminus y)$.

(V.4) Concatenation: $\forall x, y(\mathbb{G}_4(x, y)\downarrow \wedge \mathbb{G}_4(x, y) = \{\langle u, v, w\rangle : \langle u, v\rangle \in x \wedge w \in y\})$.

(V.5) Permutation: $\forall x, y(\mathbb{G}_5(x, y)\downarrow \wedge \mathbb{G}_5(x, y) = \{\langle u, w, v\rangle : \langle u, v\rangle \in x \wedge w \in y\})$.

(V.6) Diagonalisation: $\forall x, y(\mathbb{G}_=(x, y)\downarrow \wedge \mathbb{G}_=(x, y) = \{\langle v, w\rangle \in x \times y : v = w\})$.

(V.7) Local extension of membership $\forall x, y(\mathbb{G}_\in(x, y)\downarrow \wedge \mathbb{G}_\in(x, y) = \{\langle v, w\rangle \in x \times y : v \in w\})$.

(V.8) $\forall x_0, ..x_{k-1}(\mathbb{G}_R(x_0, ..., x_{k-1})\downarrow \wedge \mathbb{G}_R(x_0, ..., x_{k-1}) = \{\langle v_0, ..., v_{k-1}\rangle \in x_0 \times ... \times x_{k-1} : R(v_0, ..., v_{k-1})\}$ for each relation symbol R of $\mathcal{L}$ of arity k.

The subsystems WEST_0, WEST_ω, WEST_r and WEST_w are defined analogously.

By *applicative set theories* we mean both operational set theory and weak explicit set theory.

2.6. Basic lemmata in applicative set theories

In the following we show that some set theoretic operations, which we will use later, are available in our applicative set theories. In some proofs in Feferman [16, 17] of the analogous lemmata for OST the choice operation $\mathbb{C}$ is used. So let us see what we can do without $\mathbb{C}$.

Since the applicative axioms are available in OST_0^- as in OST and also in WEST_0, we can introduce lambda abstraction and define a closed term **fix** such that the following recursion theorem holds (as for instance in Beeson [7, p. 101 and p. 103]).

Lemma 9 (Recursion theorem)**.** Both OST_0^- and WEST_0 prove for variables x and f:

$$\mathbf{fix}(f)\downarrow \wedge (\mathbf{fix}(f,x) \simeq f(\mathbf{fix}(f,x))).$$

Proposition 10. There are closed $\mathcal{L}_{\in}^{\circ}$ terms $\mathbf{p}$ and $\mathbf{prod}$ such that both OST_0^- and WEST_0 prove

(i) $\forall x, y(\mathbf{p}(x,y)\downarrow \wedge \mathbf{p}(x,y) = \langle x, y\rangle)$,
(ii) $\forall x, y(\mathbf{prod}(x,y)\downarrow \wedge \mathbf{prod}(x,y) = \{\langle v, w\rangle : v \in x \wedge w \in y\})$ (i.e. $\mathbf{prod}(x,y)$ is $x \times y$).

Proof. Let $\mathbf{p} = \lambda x, y.\mathbb{D}(\mathbb{D}(x,x), \mathbb{D}(x,y))$. Then obviously (i) holds.
For the second assertion, let

$$\mathbf{prod} = \lambda x, y.\mathbb{U}(\mathbb{R}(\lambda v.\mathbb{R}(\lambda u.\mathbb{D}(\mathbb{D}(u,u), \mathbb{D}(u,v)), x), y)).$$

Then we can see $\mathbb{R}(\lambda u.\mathbb{D}(\mathbb{D}(u,u), \mathbb{D}(u,v)), x) = \{\langle u, v\rangle : u \in x\}$ and also $\mathbb{R}(\lambda v.\mathbb{R}(\lambda u.\mathbb{D}(\mathbb{D}(u,u), \mathbb{D}(u,v)), x), y) = \{\{\langle u, v\rangle : u \in x\} : v \in y\}$. □

The following lemma is almost trivial and can be proved as in Feferman [16, 17].

Lemma 11. For a Δ_0 formula $A[u_0, ..., u_{n-1}]$ of $\mathcal{L}$, there exists a closed $\mathcal{L}^{\circ}$ term t_A such that OST_0^- proves

$$t_A\downarrow \wedge (t_A : \mathbf{V}^n \to \mathbf{B}) \wedge \forall x_0, ..., x_{n-1}(A[x_0, ..., x_{n-1}] \leftrightarrow t_A(x_0, ..., x_{n-1}) = \mathsf{t}).$$

By the standard argument about Gödel operations (see Barwise [4, 6.1 Lemma]), we have the following lemma. This holds also for OST_0^-, since we can set $s_A = \lambda y, \vec{x}.\mathbb{S}(\lambda v.t_A[v, \vec{x}], y)$.

Lemma 12. For a Δ_0 formula $A[v, u_0, ..., u_{n-1}]$ of $\mathcal{L}$, there exists a closed $\mathcal{L}^{\circ}$ term s_A such that WEST_0 proves

$$s_A\downarrow \wedge (s_A : \mathbf{V}^{n+1} \to \mathbf{V}) \wedge \forall y, x_0, ..., x_{n-1}(s_A(y, x_0, ..., x_{n-1}) = \{v \in y : A[v, x_0, ..., x_{n-1}]\}).$$

Proposition 13. There are closed $\mathcal{L}_{\in}^{\circ}$ terms $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{dom}$, $\mathbf{op}$ and $\mathbf{fun}$ such that WEST_0 prove

(i) $\forall x, y, z(z = \langle x, y\rangle \to (\mathbf{p}_0(z) = x \wedge \mathbf{p}_1(z) = y))$;
(ii) $\mathbf{dom}(f)\downarrow \wedge \mathbf{ran}(f)\downarrow \wedge (\ \mathrm{Rel}[f] \to f \subseteq \mathbf{dom}(f) \times \mathbf{ran}(f)\ )$;
(iii) $\mathbf{op}(f)\downarrow \wedge (\ (\mathrm{Fun}[f] \wedge y \in \mathbf{dom}(f)) \to f'y = \mathbf{op}(f,y)\ )$;
(iv) $(f : a \to \mathbf{V}) \to \mathrm{Fun}[\mathbf{fun}(f,a)] \wedge \mathrm{Dom}[\mathbf{fun}(f,a), a] \wedge (\forall x \in a)(\mathbf{fun}(f,a)'x = fx)$,

The same assertions hold for OST_0.

Proof. Let $A[x,z]$ be the Δ_0 formula $(\exists u \in z)(\exists v \in u)(\langle x, v\rangle = z)$, and s_A the corresponding term due to Lemma 12. Then $s_A(\mathbb{U}(\langle x, y\rangle), \langle x, y\rangle) = \{x\}$ and hence we can take $\mathbf{p}_0 = \lambda z.\mathbb{U}(s_A(\mathbb{U}(z)), z)$. Similarly we can define $\mathbf{p}_1$ from $B[y,z] := (\exists u \in z)(\exists v \in u)(\langle v, y\rangle = z)$.

For $\mathbf{dom}$ and $\mathbf{ran}$, in WEST_0 we can just set $\mathbf{dom} = \mathbb{G}_1$ and $\mathbf{ran} = \mathbb{G}_2$. In OST_0^-, we can use Lemma 11.

For constructing $\mathbf{op}$ without $\mathbb{C}$, let $A[x, y, a, f]$ be the Δ_0 formula $(\exists z \in a)(f'y = z \wedge x \in z)$ and s_A the corresponding term due to Lemma 12. Then WEST_0 proves for any set theoretic function f with y in its domain that $f'y = \{x \in \mathbb{U}(\mathbf{ran}(f)) : A[x, y, \mathbf{ran}(f), f]\} = s_A(\mathbb{U}(\mathbf{ran}(f)), y, \mathbf{ran}(f), f)$. So it proves the stated properties for $\mathbf{op} := \lambda f, y.s_A(\mathbb{U}(\mathbf{ran}(f)), y, \mathbf{ran}(f), f)$.

For the fourth assertion, it is easy to see that $\mathbf{fun} = \lambda f, a.\mathbb{R}(\lambda x.\mathbf{p}(x, fx), a)$ satisfies the stated property. □

Remark 14. In the proof given above, only (iii) requires the axiom of extensionality and the others (i), (ii) and (iv) can be proved in WEST_0 without extensionality, but an additional equation $\mathbb{U}(\{x\}) = x$ is needed for (i). More precisely, for any a such that $(\forall z)(z \in a \leftrightarrow z = x)$, $\mathbb{U}(a) = x$.

Now we are ready to show that WEST_0 can be seen as a subtheory of OST_0^-.

Theorem 15. There exist closed $\mathcal{L}^\circ_\in$ terms $g_1, \ldots, g_5, g_=, g_\in$ and $\mathcal{L}^\circ$-terms g_R for each relational symbol R of $\mathcal{L}$ such that OST_0^- proves the axioms (V.1)-(V.7) with $\mathbb{G}_1, \ldots, \mathbb{G}_5, \mathbb{G}_=, \mathbb{G}_\in$ and $\mathbb{G}_R$ replaced by $g_1, \ldots, g_5, g_=, g_\in$ and g_R respectively.

Proof. We set $g_1 = \mathbf{dom}$ and $g_2 = \mathbf{ran}$, with which we can prove (V.1) and (V.2). We can construct g_3, $g_=$, $g_\in$ and g_R with the operator $\mathbb{S}$ and Lemma 11.

For g_4, first notice that there is a closed term t such that

$$t{\downarrow} \wedge \forall u(t(u){\downarrow} \wedge t(u) = \langle \mathbf{p}_0(\mathbf{p}_0(u)), \mathbf{p}_1(\mathbf{p}_0(u)), \mathbf{p}_1(u)\rangle),$$

and therefore we can set $g_4 = \lambda x, y.\mathbb{R}(t, \mathbf{prod}(x, y))$. We can construct g_5 similarly. □

In what follows, when we are proving lower bound results we work only in WEST_0 and extensions, for the analogous results for OST_0^- immediately follow by this lemma.

2.7. Δ_0 converted into definedness

We introduce the well known λ-terms $\overline{0} := \lambda f, x.x$, $\overline{1} := \lambda f, x.fx$ and iszero $:= \lambda x, y, z.x(\lambda u.z)y$ and we can compute that the next lemma holds.

Lemma 16. The applicative axioms of WEST_0 prove for all y, z that

$$\text{iszero}(\overline{0}, y, z) = y \text{ and } \text{iszero}(\overline{1}, y, z) = z.$$

Lemma 17. For every Δ_0 formula $A[\vec{x}]$ of $\mathcal{L}^\circ$ with at most the variables $\vec{x}$ free, there exists a closed $\mathcal{L}^\circ$ term ite_A such that WEST_0 proves that

$$\text{ite}_A{\downarrow} \;\wedge\; ((A[\vec{x}] \to \text{ite}_A(u, v)(\vec{x}) = u) \;\wedge\; (\neg A[\vec{x}] \to \text{ite}_A(u, v)(\vec{x}) = v)).$$

Proof. Let $A[\vec{x}]$ be a Δ_0 formula of $\mathcal{L}^\circ$ with at most the variables $\vec{x}$ free, $B[y, \vec{x}]$ the Δ_0 formula

$$(y = \overline{0} \wedge A[\vec{x}]) \vee (y = \overline{1} \wedge \neg A[\vec{x}]).$$

and $s_B[y, \vec{x}]$ the term due to Lemma 12. Then it is easy to check that

$$\text{ite}_A := \lambda u, v.\lambda\vec{x}.\text{iszero}(\mathbb{U}(s_B(\mathbb{D}(\overline{0}, \overline{1}), \vec{x})), u, v)$$

has the stated properties. □

Lemma 18. For every Δ_0 formula $A[\vec{x}]$ of $\mathcal{L}^\circ$ with at most the variables $\vec{x}$ free, there exists a closed $\mathcal{L}^\circ$ term iff_A such that WEST_0 proves

$$\text{iff}_A{\downarrow} \;\wedge\; \forall\vec{x}(A[\vec{x}] \leftrightarrow \text{iff}_A(\vec{x}){\downarrow}).$$

Proof. Let $B[x]$ be the formula $x = \overline{0}$, t the term $\lambda yz.\text{ite}_B\overline{1}\,\overline{0}(yz)$ and s the term $\mathbf{fix}(t)$. By Lemma 9 and the previous lemma we have for all z that

$$s(z) \simeq t(s, z) \simeq \begin{cases} \overline{1} & \text{if } s(z) = \overline{0}, \\ \overline{0} & \text{if } s(z){\downarrow} \wedge s(z) \neq \overline{0}. \end{cases}$$

It is easy to see that we have therefore $\neg s(z){\downarrow}$ for every z. Now we define iff_A as the term $\lambda\vec{x}.(\text{ite}_A(\lambda x.\overline{0}, s)(\vec{x}))(\overline{0})$ and we have therefore by the previous lemma $\text{iff}_A{\downarrow}$ as well as

$$\text{iff}_A(\vec{x}) \simeq \begin{cases} (\lambda x.\overline{0})(\overline{0}) = \overline{0} & \text{if } A[\vec{x}], \\ s(\overline{0}) & \text{if } \neg A[\vec{x}], \end{cases}$$

for all $\vec{x}$. We get the stated properties of iff_A by the properties of s. □

As one can see, these proofs heavily depend on classical logic. We do not know if we can obtain the same results in the intuitionistic context.

Lemma 19. Let $\vec{v}$ be a sequence of variables $v_0, ..., v_n$. For every Δ_0 formula $A[x, \vec{v}]$ of $\mathcal{L}^\circ$ with at most the variables $x, \vec{v}$ free and all terms $\vec{t}[x] = t_0[x], ..., t_n[x]$ of $\mathcal{L}^\circ$,

- WEST_0 + (opInd) proves

$$\forall x((\forall y \in x)(\vec{t}[y]\downarrow \wedge A[y, \vec{t}[y]]) \to (\vec{t}[x]\downarrow \wedge A[x, \vec{t}[x]])) \to \forall x(\vec{t}[x]\downarrow \wedge A[x, \vec{t}[x]]),$$

- WEST_0 + (opInd_ω) proves

$$(\forall x \in \omega)((\forall y \in x)(\vec{t}[y]\downarrow \wedge A[y, \vec{t}[y]]) \to (\vec{t}[x]\downarrow \wedge A[x, \vec{t}[x]])) \to (\forall x \in \omega)(\vec{t}[x]\downarrow \wedge A[x, \vec{t}[x]]),$$

where $\vec{t}[x]\downarrow$ abbreviates $t_0[x]\downarrow \wedge ... \wedge t_n[x]\downarrow$.

PROOF. Let s be the term $\lambda x.\mathrm{iff}_A(x, t_0[x], ..., t_n[x])$. Then we have $s\downarrow$ and by the previous lemma and the strictness axioms of the logic of partial terms

$$s(x)\downarrow \;\leftrightarrow\; \bigwedge\nolimits_{i=0}^{n} t_i[x]\downarrow \wedge A[x, \vec{t}[x]]$$

for every x. The stated assertions follow therefore if we apply (opInd) or (opInd_ω) to s. □

Remark 20. Since Lemma 19 depends on Lemma 18, which is based on classical logic, we do not know if this holds intuitionistically. However, we do not need extensionality in the argument, except in Lemma 17 we need it to show $\mathbb{U}(\{x\}) = x$ as in Remark 14.

2.8. Extensions of OST, OST^- *and* WEST

The introduced applicative set theories can be extended by the following two axioms which provide us with operations for creating powersets and collapsing functions of the accessible part, respectively:

($\mathbb{P}$) $(\mathbb{P} : \mathbf{V} \to \mathbf{V}) \wedge \forall y(y \in \mathbb{P}a \leftrightarrow y \subseteq a)$,

($\mathbb{B}$) $(\mathbb{B} : \mathbf{V}^2 \to \mathbf{V}) \wedge \mathrm{Fun}[\mathbb{B}(a, r)] \wedge \mathrm{DwCl}[\mathbf{dom}(\mathbb{B}(a, r)), a, r] \wedge \mathrm{Prog}[\mathbf{dom}(\mathbb{B}(a, r)), a, r]$
$\wedge\, (\forall x \in \mathbf{dom}(\mathbb{B}(a, r)))(\mathbb{B}(a, r)'x = \{\mathbb{B}(a, r)'y : y \in \mathbf{dom}(\mathbb{B}(a, r)) \wedge \langle y, x\rangle \in r\})$),

where $\mathrm{DwCl}[b, a, r]$ is the Δ_0 formula saying that b is downward closed with respect to a and r:

$$(\forall x \in b)(\forall y \in a)(\langle y, x\rangle \in r \to y \in b)$$

and $\mathrm{Prog}[b, a, r]$ is the Δ_0 formula expressing that b is progressive with respect to a and r:

$$b \subseteq a \wedge (\forall x \in a)((\forall y \in a)(\langle y, x\rangle \in r \to y \in b) \to x \in b).$$

In Jäger [29] is proved that the theory of so-called power admissible sets $\mathsf{KP} + (\mathcal{P})$ (for its formulation see the next section) can be embedded into $\mathsf{OST} + (\mathbb{P})$ and that the latter can be embedded into $\mathsf{KP} + (\mathcal{P}) + (V = L)$, where $(V = L)$ denotes the axiom of constructibility. However, this result does not resolve the exact strength of $\mathsf{OST} + (\mathbb{P})$ since $\mathsf{KP} + (\mathcal{P}) + (V = L)$ is strictly stronger than $\mathsf{KP} + (\mathcal{P})$ as shown in Mathias [44, Theorem 6.47].

We can also introduce the set $\mathcal{AD}$ of axioms for the unary relation symbol Ad of $\mathcal{L}$ for expressing that specific sets are admissible. The axioms about the relation Ad are

(AD.i) $\mathsf{Ad}(a) \to \omega \subseteq a \wedge \omega \in a$,

(AD.ii) $\mathsf{Ad}(a) \to (\forall \vec{x} \in a)A^{(a)}[\vec{x}]$,

(AD.iii) $(\mathsf{Ad}(a) \wedge \mathsf{Ad}(b)) \to (a \in b \vee a = b \vee b \in a)$, where $A[\vec{u}]$ is an instance of the axioms pairing, Δ_0 separation, Δ_0 collection (these axioms are introduced in the next section) and $\vec{x}$ denotes in each case its free variables.

Finally, the axioms for the operations $\mathbb{A}$ and $\mathbb{M}$ are given by

($\mathbb{A}$) $x \in \mathbb{A}(x) \land \mathrm{Tran}[\mathbb{A}(x)] \land \mathsf{Ad}(\mathbb{A}(x))$,

($\mathbb{M}$) $(f : \mathbf{V} \to \mathbf{V}) \to (x \in \mathbb{M}(f,x) \land \mathrm{Tran}[\mathbb{M}(f,x)] \land \mathsf{Ad}(\mathbb{M}(f,x)) \land (f : \mathbb{M}(f,x) \to \mathbb{M}(f,x)))$.

Both are natural set theoretic analogues of the axioms for limit and Mahlo operators in explicit mathematics, introduced in Jäger, Kahle and Studer [32], Jäger and Studer [37], Jäger and Strahm [35, 36] and Jäger [28]. Although in the standard formulation of Ad the axiom $\mathsf{Ad}(x) \to \mathrm{Tran}[x]$ is included, in our formulation however $\mathrm{Tran}[x]$ is added to the existence axioms, for some technical convenience.

The former axiom ($\mathbb{A}$) is weaker than the axiom (Inac), which is analysed on the base of operational set theory OST in Jäger and Zumbrunnen [38].

3. Theories KP and KP^{int}

We will use Kripke-Platek set theory with infinity and variants as reference systems, for many of these systems have already been well investigated proof theoretically. In this section, we introduce these systems and their intensional versions, and summarise basic facts on these systems. Both theories are formulated in the language $\mathcal{L}$ and are based on classical first order logic with equality axioms. We will use the same versions of the non-logical axioms, in the later sections, for formulating intuitionistic versions of Kripke-Platek set theory. There it will be important that they have exactly this form.

3.1. Axioms of Kripke-Platek set theory

The theory KP consists of the following non-logical axioms:

(KP.1) $\exists x(a \subseteq x \land \mathrm{Tran}[x])$ (transitive superset).

(KP.2) $\exists x(a \in x \land b \in x)$ (pairing).

(KP.3) $\exists x(\forall y \in a)(\forall z \in y)(z \in x)$ (union).

(KP.4) $\exists x((\forall y \in x)(y \in a \land A[y]) \land (\forall y \in a)(A[y] \to y \in x))$
for all Δ_0 formulae $A[y]$ in which x does not occur (Δ_0 separation).

(KP.5) $(\forall x \in a)\exists y A[x,y] \to \exists z(\forall x \in a)(\exists y \in z)A[x,y]$
for all Δ_0 formulae $A[x,y]$ in which z does not occur (Δ_0 collection).

(KP.6) $\in$-induction for all $\mathcal{L}$ formulae.

(KP.7) $\mathrm{Ind}[\omega] \land (\forall x \subseteq \omega)(\mathrm{Ind}[x] \to \omega \subseteq x)$ (infinity)
where $\mathrm{Ind}[x]$ is the formula $(\exists y \in x)\mathrm{zero}[y] \land (\forall y \in x)(\exists z \in x)\mathrm{succ}[y,z]$ and
where $\mathrm{zero}[y] \equiv (\forall z \in y)\bot$ and $\mathrm{succ}[y,z] \equiv y \in z \land (\forall u \in y)(u \in z) \land (\forall u \in z)(u \in y \lor u = y)$.

(KP.8) $(\exists z)\mathrm{famfun}[\omega, y, z]$ (Kleene star)
where the formula $\mathrm{famfun}[\omega,y,z]$ is the conjunction of $(\forall u \in z)(\exists n \in \omega)(\mathrm{Fun}[u] \land \mathrm{Dom}[u,n] \land \mathrm{Ran}[u,y])$, $(\exists u \in z)\mathrm{zero}[u]$ and $(\forall u \in z)(\forall n \in \omega)(\mathrm{Dom}[u,n] \to (\forall x \in y)(\exists v \in z)(u \subseteq v \land (\exists w \in v)(w = \langle n,x \rangle)))$.

(KP.9) $a = b \leftrightarrow (\forall x \in a)(x \in b) \land (\forall x \in b)(x \in a)$ (extensionality).

The intensional version KP^{int} is the theory KP without the axiom (KP.9).

Remark 21. As in operational set theory, the axiom (KP.1) and (KP.8) are redundant, but they are essential in the formulations of subsystems below.

We do not use the notation "$x = 0$" or "$z = y + 1$" for $\mathrm{zero}[x]$ or $\mathrm{succ}[y,z]$ in the formulation, because we consider also the intensional version KP^{int}. In the absence of extensionality, the emptyset, the successor or Kleene's star are not determined uniquely and therefore we have to be careful by which formula we formalise these notions. "$(\exists w \in v)(w = \langle n,x \rangle)$" means v contains some set corresponding to $\langle n,x \rangle$.

3.2. Kripke-Platek set theory with restricted $\in$-induction

The theories KP_0 and KP_0^{int} are KP and KP^{int}, respectively, but without $\in$-induction. KP_ω is KP_0 with in addition $\in$-induction restricted to ω, KP_r is KP_0 with in addition $\in$-induction restricted to Δ_0 formulae, and KP_w is KP_r with in addition $\in$-induction restricted to ω. KP_ω^{int}, KP_w^{int} and KP_r^{int} are defined analogously.

The rule of assigning subscripts is the same as before (see Table 1).

We write in addition (Σ_1-Ind) for the schema consisting of all instances of $\in$-induction for Σ_1 formulae and (Σ_1-Ind_ω) for (Σ_1-Ind) restricted to ω.

Notice that in these theories, in contrast to similar theories introduced in Jäger [26], we do not have urelements. Instead, we have Kleene star, or the set $x^* = x^{<\omega}$ of all finite sequences of elements from any given set x. In Jäger [26] the letter w (not ω) is used for naming restricted theories in which $\in$-induction for sets (or equivalently for Δ_0 formulae) is available in the whole universe and induction for all formulae is available on natural numbers as urelements. Besides these, our notation concerning subscripts follows Jäger's.

Remark 22. Because of the lack of the structure on urelements assumed in Jäger [26], in order to show that our set theories contain (a part of) arithmetic via the standard interpretation, we have to define the summation and the multiplication on ω as in the usual set theory. This can be done in KP_0 because the existence of $\omega^{<\omega}$ makes the formulae, required in $\in$-induction, be Δ_0. For example, $x + y = z$ can be formalised as the Δ_0 formula

$$(\exists s \in \omega^{<\omega})(\mathrm{Fun}[s] \wedge \mathrm{Dom}[s, y+1] \wedge s'0 = x \wedge (\forall k \in y)(s'(k+1) = (s'k) + 1) \wedge s'y = z).$$

Though our formulation of KP_r contains Kleene star, which is not included in Jäger [25], the same method used there works well for our KP_r and therefore we can reduce KP_r into Peano arithmetic. Conversely, as outlined just above, Peano arithmetic can be interpreted in KP_0, in the standard way. Thus we conclude that both KP_0 and KP_r are conservative over Peano arithmetic.

Remark 23. As is known, KP_0 and (KP.5) imply Σ_1 collection (for the proof, see [4, Chapter I, 4.4 Theorem]). Also, we can define a satisfaction or partial truth predicate for all standard Σ_1 formulae (with respect to a fixed Gödel numbering) by a single Σ_1 formula (e.g., [4, Chapter V, 1.6 Proposition] or by modifying the proof of [50, Definition VII.2.1 and Lemma VII.2.2]; notice that we need only meta-induction, since we concern only formulae of standard length). Thus we have a universal Σ_1 formula $(\exists u)\pi[e, u, x, y, b]$ where π is a Δ_0 formula such that, for any Σ_1 formula $A[x, y, b_1, ..., b_{k-1}]$, $\mathsf{KP}_0 \vdash A[x, y, b_1, ..., b_{k-1}] \leftrightarrow \exists u \pi[\ulcorner A \urcorner, u, x, y, \langle b_0, ..., b_{k-1}\rangle]$. Therefore, the schema (KP.5) can be replaced by a single instance

$$(\forall x \in a)(\exists y)\pi[e, (y)_0, x, (y)_1, b] \rightarrow \exists z(\forall x \in a)(\exists y \in z)\pi[e, (y)_0, x, (y)_1, b].$$

Similarly, (Σ_1-Ind) and (Σ_1-Ind_ω) can be replaced by their single instances. Also, as is well known, (KP.4) can be replaced by finite number of instances (for the proof, see [4, Chapter II, Lemma 6.1]).

Thus the only schema which cannot be replaced by finitely many instances is (KP.6), $\in$-induction.

3.3. Extensions of Kripke-Platek set theory

KP and variants can be extended by the powerset axiom (formulated with a binary relation symbol $\mathcal{P}$ of $\mathcal{L}$) or the well known axiom beta:

($\mathcal{P}$) $\forall x \exists y \mathcal{P}(x, y) \wedge \forall x, y(\mathcal{P}(x, y) \leftrightarrow \forall z(z \in y \leftrightarrow z \subseteq x))$,

(Beta) $\mathrm{WF}[a, r] \rightarrow \exists f(\mathrm{Fun}[f] \wedge \mathrm{Dom}[f, a] \wedge (\forall x \in a)(f'x = \{f'y : y \in a \wedge \langle y, x\rangle \in r\}))$,

where $\mathrm{WF}[a, r]$ is the formula expressing that r is a well founded relation on a:

$$\forall b \subseteq a(\mathrm{Prog}[b, a, r] \rightarrow a \subseteq b).$$

There have been many ways proposed to treat powerset on the base of KP. Mathias [44] compares some of them, which include the adding the predicate $\mathcal{P}$ as above as well as treating $\subseteq$-bounded quantifiers similarly to $\in$-bounded ones.

As for the axiom beta, it is worth mentioning that $\mathsf{KP}_r + \text{(Beta)}$ is the same as Simpson's ATR_0^{set} plus Δ_0 collection (from Simpson [50, Chapter VII]).

We will write $\mathcal{L}_{\mathsf{Ad}}$ for the language $\mathcal{L}$ but restricted to the relation symbols $\in$, $=$ and Ad. We introduce the axiom (Lim), which says that every set is a member of some admissible set, i.e.,

(Lim) $\forall x \exists y (x \in y \wedge \mathrm{Tran}[y] \wedge \mathsf{Ad}(y))$.

The theories KPI_0, KPI_ω, KPI_r, KPI_w and KPI are the theories KP_0, KP_ω, KP_r, KP_w and KP, respectively, with the additional axioms $(\mathcal{AD}) + (\mathsf{Lim})$, where $(\mathcal{AD})$ is from the last section.

We also introduce

(Π_2-Ref) $\forall x \exists y A[x, y, \vec{u}] \to \exists v (\vec{u} \in v \wedge \mathrm{Tran}[v] \wedge \mathsf{Ad}(v) \wedge (\forall x \in v)(\exists y \in v) A[x, y, \vec{u}])$,

for any Δ_0 formula A. The theories KPM_0, KPM_ω, KPM_r, KPM_w and KPM are the theories KP_0, KP_ω, KP_r, KP_w and KP, respectively, with the additional axioms $(\mathcal{AD}) + (\Pi_2\text{-}\mathsf{Ref})$.

Lemma 24. $\mathsf{KP}_r + (\mathcal{AD}) + (\mathsf{Lim})$ proves that for any x there is exactly one y such that

$$x \in y \wedge \mathrm{Tran}[y] \wedge \mathsf{Ad}(y) \wedge (\forall z \in y)(\neg x \in z \vee \neg \mathrm{Tran}[z] \vee \neg \mathsf{Ad}(z)).$$

Proof. For any x, (Lim) gives us a such that $x \in a$, $\mathrm{Tran}[a]$ and $\mathsf{Ad}(a)$. Define

$$b = \{z \in a \cup \{a\} : \neg x \in z \vee \neg \mathrm{Tran}[z] \vee \neg \mathsf{Ad}(z)\}.$$

Suppose for contradiction that there is no y that satisfies the condition. Then we can prove by induction on $z \in a \cup \{a\}$ that $z \in b$, and in particular $a \in b$, a contradiction. Thus there is some y satisfying the condition.

For the uniqueness, assume that both a and b satisfy the condition:

$$\begin{aligned}&(x \in a \wedge \mathrm{Tran}[a] \wedge \mathsf{Ad}(a) \wedge (\forall z \in a)(x \in z \wedge \mathrm{Tran}[z] \to \neg \mathsf{Ad}(z)))\\ \wedge\ &(x \in b \wedge \mathrm{Tran}[b] \wedge \mathsf{Ad}(b) \wedge (\forall z \in b)(x \in z \wedge \mathrm{Tran}[z] \to \neg \mathsf{Ad}(z))).\end{aligned}$$

If $a \neq b$, by $\mathsf{Ad}(a)$ and $\mathsf{Ad}(b)$, (AD.iii) implies $a \in b$ or $b \in a$, a contradiction. □

Lemma 25. For any Π_2 formula $A[x]$, $\mathsf{KP}_r + (\mathcal{AD}) + (\Pi_2\text{-}\mathsf{Ref})$ proves that for any x if $A[x]$ holds then there is exactly one y such that

$$x \in y \wedge \mathrm{Tran}[y] \wedge \mathsf{Ad}(y) \wedge A^{(y)}[x] \wedge (\forall z \in y)(\neg x \in z \vee \neg \mathrm{Tran}[z] \vee \neg \mathsf{Ad}(z) \vee \neg A^{(z)}[x]).$$

3.4. Δ definability of Kleene star and transitive closure

Lemma 26. KP_0 proves that for any b there is exactly one c such that

$$\begin{aligned}&(\forall u \in c)(\exists n \in \omega)(\mathrm{Fun}[u] \wedge \mathrm{Dom}[u, n] \wedge \mathrm{Ran}[u, b])\\ &\qquad\wedge \emptyset \in c \wedge (\forall u \in c)(\forall x \in b)(\forall n \in \omega)(\mathrm{Dom}[u, n] \to u \cup \{\langle n, x\rangle\} \in c).\end{aligned}$$

and this c is specified by

$$c = \{s : (\exists n \in \omega)(\mathrm{Fun}[s] \wedge \mathrm{Dom}[s, n] \wedge \mathrm{Ran}[s, b])\}.$$

Proof. By the axiom (Kleene star), there is c which satisfies the former condition. Since the uniqueness follows from the latter condition, it remains to show the latter.

Since $c \subseteq \{s : (\exists n \in \omega)(\mathrm{Fun}[s] \wedge \mathrm{Dom}[s, n] \wedge \mathrm{Ran}[s, b])\}$ is immediate from the first clause, we prove the converse. Let $\mathrm{Fun}[s] \wedge \mathrm{Dom}[s, m] \wedge \mathrm{Ran}[s, b]$. We prove $(\exists u \in c)(\mathrm{Dom}[u, n] \wedge (\forall k < n)(u'k = s'k))$ by induction on $n \leq m$. For $n = 0$, this is obvious by the second clause $\emptyset \in c$. For the other case, the induction hypothesis yields $u \in c$ with $\mathrm{Dom}[u, n]$ and $(\exists u \in c)(\forall k < n)(u'k = s'k)$ and the last clause yields $u \cup \{\langle n, s'n\rangle\} \in c$. Thus there is $u \in c$ with $\mathrm{Dom}[u, m]$ and $(\forall k < m)(u'k = s'k)$. Then, by (extensionality), $s = u \in c$. □

Lemma 27. KP_0 proves the equivalence between $\forall u(a \subseteq u \wedge \mathrm{Tran}[u] \to x \in u)$ and

$$\exists f(\exists n \in \omega)(\mathrm{Dom}[f] \wedge \mathrm{Dom}[f, n{+}1] \wedge f'0 \in a \wedge (\forall k \in n)(f'(k+1) \in f'k) \wedge x = f'n).$$

Proof. Assume that the latter holds. To show that the former holds, let $a \subseteq u \wedge \mathrm{Tran}[u]$. Then, by induction on k, we have $k \in n{+}1 \to f'k \in u$, and in particular $x \in u$.

Conversely assume $\forall u(a \subseteq u \wedge \mathrm{Tran}[u] \to x \in u)$. By the axiom (transitive superset) we have v such that $a \subseteq v$ and $\mathrm{Tran}[v]$, and by the axiom (Kleene star) we have $v^{<\omega}$. Then we can take

$$u = \{y : (\exists s \in v^{<\omega})(\exists n \in \omega)(\mathrm{Dom}[s, n+1] \wedge s'0 \in a \wedge (\forall k \in n)(s'(k+1) \in s'k) \wedge y = s'n\}.$$

Now $a \subseteq u$ and $\mathrm{Tran}[u]$. Thus $x \in u$ and let s be the witness. This s witnesses the latter. □

Lemma 28. KP_0 proves that for any set b there is exactly one c such that

$$\mathrm{Tran}[c] \wedge b \subseteq c \wedge (\forall x \in c)(\exists f)(\exists n \in \omega)(\mathrm{Fun}[f] \wedge \mathrm{Dom}[f, n{+}1] \wedge (\forall k \in n)(f'(k{+}1) \in f'k) \wedge f'0 \in b \wedge f'n = x)$$

and this c satisfies $c = \{x : \forall u(b \subseteq u \wedge \mathrm{Tran}[u] \to x \in u)\}$.

Proof. By Δ separation, the equivalence in the last lemma shows that $c = \{x : \forall u(b \subseteq u \wedge \mathrm{Tran}[u] \to x \in u)\}$ exists as a set. Again by the same equivalence, c satisfies the condition.

If d satisfies the same condition, namely

$$\mathrm{Tran}[d] \wedge b \subseteq d \wedge (\forall x \in d)(\exists f)(\exists n \in \omega)(\mathrm{Fun}[f] \wedge \mathrm{Dom}[f, n{+}1] \wedge (\forall k \in n)(f'(k{+}1) \in f'k) \wedge f'0 \in b \wedge f'n = x).$$

By the definition of c and the first clause $\mathrm{Tran}[d]$, we have $c \subseteq d$. By the third clause for d and by the definition of c with the previous equivalence, we have $d \subseteq c$. □

3.5. Two kinds of axiom beta

Axiom (Beta) does not have the adequate form that we can prove later its connection to the axiom ($\mathbb{B}$) directly. Therefore we introduce an alternative version of this axiom:

$$(\mathrm{Beta}') \quad (\forall a, r)(\exists f, b)(\mathrm{Fun}[f] \wedge \mathrm{Dom}[f, b] \wedge b \subseteq a \wedge \mathrm{DwCl}[b, a, r] \wedge \mathrm{Prog}[b, a, r] \wedge (\forall v \in b)(f'v = \{f'w : w \in b \wedge \langle w, v\rangle \in r\}))\ ).$$

We prove that this is equivalent to the original formulation (Beta). One direction is easy.

Lemma 29. $\mathsf{KP}_0 + (\mathrm{Beta}')$ proves (Beta).

Proof. Let $\mathrm{WF}[a, r]$. By (Beta′), we have f and b as described above. By $\mathrm{WF}[a, r]$ and $\mathrm{Prog}[b, a, r]$, we can infer $a \subseteq b$ and so $a = b$. Thus f is what is required in (Beta). □

Definition 30. We introduce a Σ formula $\mathrm{WP}[x, a, r]$ as

$$\exists u(\ x \in u \subseteq a \wedge \mathrm{DwCl}[u, a, r] \wedge \exists f(\mathrm{Fun}[f] \wedge \mathrm{Dom}[f, u] \wedge (\forall y \in u)(f'y = \{f'z : z \in u \wedge \langle z, y\rangle \in r\}))\ ).$$

We also introduce the Π formula $\mathrm{WP}'[x, a, r]$ as

$$(\forall v, s)(\forall n \in \omega)(\ \mathrm{Fun}[s] \wedge \mathrm{Dom}[s, n+1] \wedge \mathrm{Ran}[s, a] \wedge s'0 = x \wedge (\forall k \in n)(\langle s'(k+1), s'k\rangle \in r) \to (\mathrm{Prog}[v, a, r] \to s'n \in v)\ ).$$

Both the formulae are intended to express that x is in the accessible part or well founded part of the order structure given by a and r:

Lemma 31.

(i) KP_0 proves that for any $u \subseteq a$, if $\mathrm{DwCl}[u,a,r]$, then (a) $\mathrm{Prog}[v,a,r]$ implies $\mathrm{Prog}[v,u,r]$ and (b) $\mathrm{Prog}[v,u,r]$ implies $\mathrm{Prog}[v \cup (a \setminus u), a, r]$.
(ii) KP_r proves that $\mathrm{Fun}[f] \wedge \mathrm{Dom}[f,b] \wedge (\forall x \in b)(f'x = \{f'y : y \in b \wedge \langle y,x\rangle \in r\})$ implies $\mathrm{WF}[b,r]$.
(iii) KP_r + (Beta) proves that $\mathrm{WP}[x,a,r]$ and $\mathrm{WP}'[x,a,r]$ are equivalent, and that $\{x \in a : \mathrm{WP}[x,a,r]\}$ is a set.
(iv) KP_r + (Beta) proves that $b = \{x \in a : \mathrm{WP}[x,a,r]\}$ implies that $\mathrm{DwCl}[b,a,r] \wedge \mathrm{Prog}[b,a,r]$ and that there is a unique function f with domain b and $(\forall x \in b)(f'x = \{f'y : y \in b \wedge \langle y,x\rangle \in r\})$.

Proof. For (i)(a), let $\mathrm{Prog}[v,a,r]$, $y \in u$ and $(\forall z \in u)(\langle z,y\rangle \in r \to z \in v)$. We have to show $y \in v$. Since $\mathrm{DwCl}[u,a,r]$, this means $(\forall z \in a)(\langle z,y\rangle \in r \to z \in v)$, from which $\mathrm{Prog}[v,a,r]$ implies $y \in v$.

For (i)(b), let $\mathrm{Prog}[v,u,r]$, $y \in a$ and $(\forall z \in a)(\langle z,y\rangle \in r \to z \in v \cup (a \setminus u))$. If $y \in u$, the second assumption is equivalent to $(\forall z \in u)(\langle z,y\rangle \in r \to z \in v)$, from which $\mathrm{Prog}[v,u,r]$ implies $y \in v \subseteq v \cup (a \setminus u)$. Otherwise trivially $y \in a \setminus u \subseteq v \cup (a \setminus u)$.

For (ii) we assume $\mathrm{Fun}[f]$, $\mathrm{Dom}[f,b]$ and

$$(\forall x \in b)(f'x = \{f'y : y \in b \wedge \langle y,x\rangle \in r\}).$$

Furthermore we assume $\mathrm{Prog}[c,b,r]$ for some $c \subseteq b$. It follows for every $x \in b$ that $x \in c$ if $y \in c$ for all $y \in b$ with $f'y \in f'x$. This means, if d is the range of f and a is the set $\{z \in d : (\forall y \in b)(z = f'y \to y \in c)\}$, then

$$(\forall v \in d)((\forall z \in d)(z \in v \to z \in a) \to v \in a).$$

Since d is transitive, $(\forall v \in d)((\forall z \in v)(z \in a) \to v \in a)$ and by $\in$-induction it follows $(\forall v \in d)(v \in a)$, which implies $b \subseteq c$. All in all this implies $\mathrm{WF}[b,r]$.

For (iii), first notice that by (ii) and (Beta) we have

$$\mathrm{WP}[x,a,r] \leftrightarrow \exists u(\, x \in u \subseteq a \wedge \mathrm{DwCl}[u,a,r] \wedge \mathrm{WF}[u,r]).$$

Assume $\mathrm{WP}[x,a,r]$, say $x \in u$, $\mathrm{DwCl}[u,a,r]$ and $\mathrm{WF}[u,r]$. To show $\mathrm{WP}'[x,a,r]$, let s be as in the antecedent and $\mathrm{Prog}[v,a,r]$. By (i)(a), we have $\mathrm{Prog}[v,u,r]$, from which $\mathrm{WF}[u,r]$ implies $u \subseteq v$. Since u is downward closed, $s'n \in v$.

Conversely, assume $\mathrm{WP}'[x,a,r]$. Let u be the downward closure of x, namely

$$u = \{y \in a : (\exists s \in a^{<\omega})(\exists n \in \omega)(\mathrm{Fun}[s] \wedge \mathrm{Dom}[s,n+1] \wedge s'0 = x \wedge (\forall k \in n)(\langle s'(k+1), s'k\rangle \in r \wedge y = s'n)\}.$$

Then obviously $x \in u$ and $\mathrm{DwCl}[u,a,r]$. To show that this u witnesses $\mathrm{WP}[x,a,r]$, it remains to show $\mathrm{WF}[u,r]$. To see this, let $\mathrm{Prog}[v,u,r]$. By (i)(b), $\mathrm{Prog}[v \cup (a \setminus u), a, r]$ and, by $\mathrm{WP}'[x,a,r]$, we have $u \subseteq v \cup (a \setminus u)$, namely $u \subseteq v$.

Thus by Δ separation $\{x \in a : \mathrm{WP}[x,a,r]\}$ is a set.

For (iv), by the equation $b = \{x \in a : \mathrm{WP}'[x,a,r]\}$, we have $\mathrm{DwCl}[b,a,r]$. Now we prove $\mathrm{WF}[b,r]$, from which the axiom (Beta) implies the existence of f. To show this, assume $\mathrm{Prog}[c,b,r]$, which by (i)(b) implies $\mathrm{Prog}[c \cup (a \setminus b), a, r]$. We have to show $b \subseteq c$. Now $\mathrm{WP}'[x,a,r]$ and $\mathrm{Prog}[c \cup (a \setminus b), a, r]$ imply $x \in c \cup (a \setminus b)$. Thus if $x \in b$, we have $x \in c$.

For uniqueness, let $(\forall x \in b)(f'x = \{f'y : y \in b \wedge \langle y,x\rangle \in r\})$ and $(\forall x \in b)(g'x = \{g'y : y \in b \wedge \langle y,x\rangle \in r\})$. Then $v = \{x \in b : f'x = g'x\}$ is progressive with respect to a and r and, by $\mathrm{WF}[b,r]$, we have $b \subset v$.

For the progressiveness of b, we need to show $\mathrm{WP}'[x,a,r]$ by assuming $(\forall y \in a)(\langle y,x\rangle \in r \to \mathrm{WP}'[y,a,r])$. Take s, n as in the antecedent of $\mathrm{WP}'[x,a,r]$ and $\mathrm{Prog}[v,a,r]$. We have to show $s'n \in v$. If $n \geq 1$, then t defined by $t'k = s'(k+1)$ satisfies the antecedent of $\mathrm{WP}'[s'0,a,r]$ and so $s'n = t'(n-1) \in v$. Now, all the cases where $n = 1$ imply $(\forall y \in a)(\langle y,x\rangle \in r \to y \in v)$, from which $\mathrm{Prog}[v,a,r]$ implies $x \in v$, the case of $n = 0$. □

Lemma 32. KP_r proves that (Beta) and (Beta′) are equivalent.

Proof. This follows from Lemma 29 and Lemma 31 (iii) an (iv). □

Part II
Lower Bound: A New Model Construction

4. Intermediate Intuitionistic Theories

Our first goal is to interpret KP and variants in WEST and corresponding variants. We do not do this directly, but by combining three interpretations. The point of our approach is that the three interpretations to be combined concern theories based on intuitionistic logic. In other words, we interpret classical theories in classical theories via two intuitionistic theories (which are not based on classical logic). In this section, we summarise basic facts on intuitionistic logic and introduce these intermediate theories. It must be emphasised that we use these intuitionistic theories for our purpose and we do not claim that these theories are natural from the viewpoints of intuitionism or constructivism.

4.1. Basics of intuitionistic logic

We assume the readers' familiarity with intuitionistic logic whose origin goes back to Brouwer. The readers not familiar with the subject can refer to van Dalen [57, Chapter 5]. As is well known, intuitionistic logic does not have the law of excluded middle $A \vee \neg A$ nor double negation elimination $\neg\neg A \to A$. For the readers' convenience, we list some intuitionistically valid formulae below, whose proofs are straightforward.

Lemma 33. The following are intuitionistically valid for any $\mathcal{L}$ formulae A, A_0, A_1 and A_2.

(1) $(A_0 \to (A_1 \to A_2)) \leftrightarrow ((A_0 \wedge A_1) \to A_2)$,
(2) $(A_0 \to (A_1 \to A_2)) \leftrightarrow (A_1 \to (A_0 \to A_2))$,
(3) $(A_0 \to (A_1 \wedge A_2)) \leftrightarrow ((A_0 \to A_1) \wedge (A_0 \to A_2))$,
(4) $(A_0 \to (A_1 \to A_2)) \leftrightarrow ((A_0 \to A_1) \to (A_0 \to A_2))$,
(5) $(A_0 \to A_1) \to (\neg\neg A_0 \to \neg\neg A_1)$,
(6) $\neg\neg(A_0 \wedge A_1) \;\leftrightarrow\; (\neg\neg A_0 \wedge \neg\neg A_1)$.
(7) $(\neg A_0 \wedge \neg A_1) \leftrightarrow \neg(A_0 \vee A_1)$.
(8) $\neg\neg(A_0 \to A_1) \;\leftrightarrow\; (\neg\neg A_0 \to \neg\neg A_1)$.
(9) $\exists x(A_0 \to A_1[x]) \to (A_0 \to \exists x A_1[x])$ if x is not free in A_0,
(10) $(\exists x \in y)(A_0[y] \to A_1[x,y]) \to (A_0[y] \to (\exists x \in y)A_1[x,y])$ if x is not free in $A_0[y]$,
(11) $\forall x(A_0 \to A_1[x]) \leftrightarrow (A_0 \to \forall x A_1[x])$ if x is not free in A_0,
(12) $(\forall x \in y)(A_0[y] \to A_1[x,y]) \leftrightarrow (A_0[y] \to (\forall x \in y)A_1[x,y])$ if x is not free in $A_0[y]$,
(13) $\exists x A[x] \to \neg\forall x\neg A[x]$,
(14) $(\exists x \in y)A[x,y] \to \neg(\forall x \in y)\neg A[x,y]$,
(15) $\exists y\neg(\forall x \in y)A[x] \to \neg\forall x A[x]$,
(16) $\forall x(A_0[x] \leftrightarrow A_1[x]) \to (\forall x A_0[x] \leftrightarrow \forall x A_1[x])$,
(17) $\exists x(A_0[x,z] \to (\forall y \in z)A_1[x,y,z]) \;\to\; \forall y\exists x(A_0[x,z] \to (y \in z \to A_1[x,y,z]))$ if y is not free in $A_0[x,z]$,
(18) $\forall y\exists x(A_0[x] \to (y \in z \to A_1[y])) \;\to\; (\forall y \in z)\exists x(A_0[x] \to A_1[y])$,
(19) $(\exists x A_0[x] \to A_1) \;\leftrightarrow\; \forall x(A_0[x] \to A_1)$ if x is not free in A_1,

4.2. Negative interpretation

Definition 34 (Negative and strongly negative formulae)**.** We call $\mathcal{L}$ formulae *negative* if they are built up from atomic formulae by means of the connectives $\wedge$ and $\to$ and the quantifier $\forall$. The *strongly negative* formulae are inductively defined as follows:

(i) $\bot$ is a strongly negative formula;
(ii) if A is an atomic formula and B a strongly negative formula, then also $A \to B$ is strongly negative;
(iii) if the formulae A and B are strongly negative, then so are $A \to B$, $A \wedge B$ as well as $\forall x A$.

We assign in the next definition to each formula A a strongly negative formula A^N which is classically equivalent to A. It is well known, that if A is classically valid, then A^N is intuitionistically valid.

Definition 35 (Gödel-Gentzen negative interpretation)**.** The *negative interpretation* A^N of each $\mathcal{L}$ formula A is inductively defined as follows:

(i) If A is an atomic formula, then A^N is the formula $\neg\neg A$.
(ii) If A is a formula $B \wedge C$, then A^N is the formula $B^N \wedge C^N$.
(iii) If A is a formula $B \vee C$, then A^N is the formula $\neg(\neg B^N \wedge \neg C^N)$.
(iv) If A is a formula $B \rightarrow C$, then A^N is the formula $B^N \rightarrow C^N$.
(v) If A is a formula $\exists x B$, then A^N is the formula $\neg \forall x \neg B^N$.
(vi) If A is a formula $\forall x B$, then A^N is the formula $\forall x B^N$.

The next lemma can be proved by induction on the length of B.

Lemma 36. If B is a strongly negative formula, B and B^N are intuitionistically equivalent.

Lemma 37.

(i) The following formulae are intuitionistically valid for an arbitrary $\mathcal{L}$ formula A:
  (a) $(y \in x \rightarrow A^N) \;\leftrightarrow\; (\neg\neg(y \in x) \rightarrow A^N)$,
  (b) $((\forall y \in x)A)^N \;\leftrightarrow\; (\forall y \in x)A^N$,
  (c) $((\exists y \in x)A)^N \;\leftrightarrow\; \neg(\forall y \in x)\neg A^N$.
(ii) If A is a Δ_0 formula of $\mathcal{L}$ then A^N is intuitionistically equivalent to some strongly negative Δ_0 formula.

Proof. Exactly as the proof of Lemma 5.1 in Avigad [2]. □

4.3. The intuitionistic theories $\mathsf{IKP}^\sharp$ *and* IKP^-

We first introduce the theory $\mathsf{IKP}^\sharp$, a weaker version of what Avigad uses in [2, Section 5] under the name $\mathsf{IKP}^{int\sharp}$.

Definition 38. The theory $\mathsf{IKP}^\sharp$ is formulated in the language $\mathcal{L}$. It is based on intuitionistic first-order logic with equality axioms and consists of the following non-logical axioms.

(i) $\exists y((\forall z \in x)\neg\neg(z \in y) \wedge (\forall z \in y)(\forall u \in z)\neg\neg(u \in y))$ (N-transitive superset).
(ii) $\exists x(\neg\neg a \in x \wedge \neg\neg b \in x)$ (N-pairing).
(iii) $\exists x(\forall y \in a)(\forall z \in y)\neg\neg(z \in x)$ (N-union).
(iv) $\exists x((\forall y \in x)(\neg\neg y \in a \wedge A[y]) \wedge (\forall y \in a)(A[y] \rightarrow \neg\neg y \in x))$ for all strongly negative Δ_0 formulae $A[y]$ in which x does not occur ($\Delta_0^{\mathrm{s}-}$ N-separation).
(v) $(\forall x \in a)\exists y A[x,y] \rightarrow \exists z(\forall x \in a)\neg(\forall y \in z)\neg A[x,y]$ for all strongly negative Δ_0 formulae $A[x,y]$ in which z does not occur ($\Delta_0^{\mathrm{s}-}$ collection$^\sharp$).
(vi) $\forall x((\forall y \in x)A[y] \rightarrow A[x]) \rightarrow \forall x A[x]$ for all strongly negative formulae $A[x]$ ($\mathcal{L}^{\mathrm{s}-}$-$\mathsf{Ind}$).
(vii) $\mathrm{Ind}[\omega]^N \wedge (\forall x \subseteq \omega)(\mathrm{Ind}[x] \rightarrow \omega \subseteq x)^N \wedge (\forall y)(\exists z)\mathrm{famfun}[\omega, y, z]^N$ (N-infinity), where $\mathrm{Ind}[x]$ and $\mathrm{famfun}[\omega, y, z]$ are defined as in the last section.

Also here we introduce the theories with weaker induction principles. $\mathsf{IKP}_0^\sharp$ is $\mathsf{IKP}^\sharp$ but without ($\mathcal{L}^{\mathrm{s}-}$-$\mathsf{Ind}$), and $\mathsf{IKP}_\omega^\sharp$ is $\mathsf{IKP}_0^\sharp$ with in addition ($\mathcal{L}^{\mathrm{s}-}$-$\mathsf{Ind}_\omega$), i.e., ($\mathcal{L}^{\mathrm{s}-}$-$\mathsf{Ind}$) restricted to ω.

We also use a slightly stronger intuitionistic theory IKP^-, which will turn out to be proof theoretically equivalent to the classical KP. The difference from $\mathsf{IKP}^\sharp$ is that non-negative formulae are allowed in collection and induction schemata and the axioms of transitive superset, pairing, union and Δ_0^- separation are defined straightforwardly. In Section 6 this theory will play the role of IKP^{int} in Avigad [2].

Definition 39. The system IKP^-, also formulated in the language $\mathcal{L}$, is based on intuitionistic logic with equality axioms and consists of the non-logical axioms of transitive superset, pairing, union, Δ_0 collection and $\in$-induction (all formulated as for KP) as well as N-infinity (as for $\mathsf{IKP}^\sharp$) and Δ_0^- separation:

$$\exists x((\forall y \in x)(y \in a \wedge A[y]) \wedge (\forall y \in a)(A[y] \to y \in x))$$

for all negative Δ_0 formulae $A[y]$ in which x does not occur.

Furthermore IKP_0^- is IKP^- but without $\in$-induction, and IKP_ω^- is IKP_0^- with in addition $\in$-induction restricted to ω.

These theories are in a sense constructive, because they are subsystems of constructive Zermelo-Fraenkel set theory CZF (for the constructive justification of CZF, see Aczel [1]). Also, both the theories are intensional in the sense that they do not include the axiom of extensionality.

4.4. The semi-constructive axiom schema (Δ_0^{s-}-MP)

We will use the weaker theory $\mathsf{IKP}^\sharp$ only with the following axiom (a set theoretic version of Markov's principle) which is outside of CZF.

Definition 40.

($\Delta_0^{\mathrm{s}-}$-MP) $\qquad \neg\forall x A[x] \to \exists y \neg(\forall x \in y)A[x]$, $\qquad$ for all strongly negative Δ_0 formulae $A[x]$ of $\mathcal{L}$.

Roughly speaking, ($\Delta_0^{\mathrm{s}-}$-MP) is a reflection principle for double-negated Σ_1 formulae. Obviously ($\Delta_0^{\mathrm{s}-}$-MP) is classically valid but it is not intuitionistically. ($\Delta_0^{\mathrm{s}-}$-MP) is sometimes called *semi-constructive*, because it is validated from some kinds of constructive viewpoint. Indeed, a similar axiom (under the name Markov's principle) is accepted in so-called *Constructive Recursive Mathematics*, a Russian school of constructivism. Constructive Recursive Mathematics, along with Bishop-style constructive mathematics, is among the starting examples of explicit mathematics (see Feferman [14]).

Another principle called semi-constructive is *double negation shift*

$$\forall x \neg\neg A[x] \to \neg\neg\forall x A[x]$$

where A should be restricted to various classes. However, this principle will not play any essential role in the present paper, as opposed to Markov's principle, while it will be mentioned in AppendixA.4.

4.5. Negatively Σ_1 *and weak* Σ_1

Definition 41 (Negatively Σ_1 and negatively Π_2 formulae)**.** The *(very) negatively* Σ_1 *formulae* of $\mathcal{L}$ are the formulae of the form $\exists x A[x]$ where $A[x]$ is a (strongly) negative Δ_0 formula.

The *(very) negatively* Π_2 *formulae* of $\mathcal{L}$ are the formulae of the form $\forall x A[x]$ where $A[x]$ is a (very) negatively Σ_1 formula.

If a formula A is intuitionistically equivalent to some (very) negatively Σ_1 (or Π_2) formula, we will for a reason of simplification call A itself (very) negatively Σ_1 (or Π_2).

Note that negatively Σ_1 (or Π_2) formulae themselves are not negative.

Definition 42. We write ($\mathsf{n}\Sigma_1$-Ind) for the schema consisting of all instances of $\in$-induction for negatively Σ_1 formulae and ($\mathsf{n}\Sigma_1$-Ind_ω) for ($\mathsf{n}\Sigma_1$-Ind) restricted to ω.

We write ($\mathsf{n}\Pi_2$-Ref) for the schema consisting of

$$A[\vec{u}] \to \exists v(\vec{u} \in v \wedge \mathrm{Tran}[v] \wedge \mathsf{Ad}(v) \wedge A^{(v)}[\vec{u}])$$

for any negatively Π_2 formula A in which v does not occur freely. We also write ($\mathsf{sn}\Pi_2$-Ref) for the analogue for very negatively Π_2 formulae.

Definition 43 (Weak Σ_1 and weak Π_2 formulae)**.** The *weak* Σ_1 *formulae* of $\mathcal{L}$ are the formulae of the form $\exists y \neg (\forall x \in y) A[x]$ where $A[x]$ is a strongly negative Δ_0 formula without any occurrence of the variable y.

The *weak* Π_2 *formulae* of $\mathcal{L}$ are the formulae of the form $\forall x A[x]$ where $A[x]$ is a weak Σ_1 formula.

If a formula A is intuitionistically equivalent to some weak Σ_1 (or Π_2) formula, we will for a reason of simplification call A itself weak Σ_1 (or Π_2).

Definition 44. We write ($\mathsf{w}\Sigma_1$-Ind) for the schema consisting of all instances of $\in$-induction for weak Σ_1 formulae and ($\mathsf{w}\Sigma_1$-Ind_ω) for ($\mathsf{w}\Sigma_1$-Ind) restricted to ω.

We write ($\mathsf{w}\Pi_2$-Ref) for the schema consisting of

$$A[\vec{u}] \to \exists v (\vec{u} \in v \wedge \mathrm{Tran}[v] \wedge \mathsf{Ad}(v) \wedge A^{(v)}[\vec{u}])$$

for any weak Π_2 formula A in which v does not occur freely.

5. Interpreting KP^{int} in $\mathsf{IKP}^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP})$: Negative Interpretation

As announced in Introduction, we will convert the operational replacement into the constructive or intuitionistic collection schema by a realisability interpretation, and therefore interpret IKP^-, which has constructive or intuitionistic collection schema, in our applicative set theories. So one might think that we could finish by combining it with the negative interpretation. However, this does not work directly, because the negative interpretation of the (classical) collection schema is not implied by the constructive collection schema. Actually, the negative interpretation embeds anything into negative fragment, which is preserved under the realisability interpretation that we will give, and therefore the composition of these two interpretations is nothing more than the trivial one. To make use of the intuitionistic machinery, we need some twist before going back to classical context with the negative interpretation.

Avigad [2] solved the problem of interpreting classical collection by intuitionistic collection. His idea is to introduce another intermediate intuitionistic theory with Markov's principle, with which the constructive collection schema implies the negative interpretation of the (classical) collection schema (as we will see in this section), and to interpret this theory in intuitionistic set theory IKP. We will follow his approach in this and the next sections.

In this section, we take the first part of Avigad's approach, which is also the first step towards the combined interpretation of KP in WEST, namely interpret the classical *intensional* set theory KP^{int} and variants in the intuitionistic set theory $\mathsf{IKP}^\sharp$ and corresponding variants, with semi-constructive Markov's principle ($\Delta_0^{\mathrm{s}-}$-MP), by Gödel-Gentzen negative interpretation.

5.1. Interpreting the axioms of KP^{int}

Lemma 45. The negative interpretation of each axiom of KP_0^{int} is provable in $\mathsf{IKP}_0^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP})$. The same assertion holds between KP^{int} and $\mathsf{IKP}^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP})$ and also between KP_ω^{int} and $\mathsf{IKP}_\omega^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP})$.

Proof. Let $\exists x A[x]$ be any of the axioms transitive superset, pairing, union, Δ_0 separation and Kleene star, where $A[x]$ is a Δ_0 formula. Then the corresponding axiom of $\mathsf{IKP}_0^\sharp$ yields $\exists x A^N[x]$ and therefore, Lemma 33 (13), $\neg \forall x \neg A^N[x]$, that is $(\exists x A[x])^N$. We can treat the axiom of infinite more easily.

It remains to show the provability of the negative interpretation of Δ_0 collection. By Lemma 33 (13), it suffices to show

$$(\forall x \in a) \neg \forall y \neg A^N[x, y] \ \to \ \exists w (\forall x \in a) \neg (\forall y \in w) \neg A^N[x, y].$$

Assume the antecedent. By ($\Delta_0^{\mathrm{s}-}$-MP), we have $(\forall x \in a) \exists u \neg (\forall y \in u) \neg A^N[x, y]$ and, by Δ_0 collection$^\sharp$,

$$\exists v (\forall x \in a) \neg (\forall u \in v) \neg\neg (\forall y \in u) \neg A^N[x, y], \tag{1}$$

where v is fresh. Now, letting $w \supseteq \bigcup v$, we have

$$(\forall y \in w) \neg A^N[x, y] \ \to \ (\forall u \in v)(\forall y \in u) \neg A^N[x, y]$$

and, by weakening the conclusion with $\neg\neg$ and taking the contraposition,

$$\neg(\forall u \in v)\neg\neg(\forall y \in u)\neg A^N[x,y] \;\to\; \neg(\forall y \in w)\neg A^N[x,y].$$

Combining this with (1), we have what we have to show. $\square$

Remark 46. Notice that in the proof above the axiom schema (Δ_0^{s-}-MP) is used only for proving the negative interpretation of instances of Δ_0 collection. Therefore, to interpret the particular instance of Δ_0 collection described in Remark 23, we need only one instance of (Δ_0^{s-}-MP) and one of (Δ_0^{s-} collection$^\sharp$). Similarly, we need only finite instances of (Δ_0^{s-} N-separation). The same remark applies to the next lemma as well.

As mentioned in Avigad [2, Section 5], (Δ_0^{s-}-MP) implies that the negative interpretation of each Σ_1 formula is equivalently weak Σ_1. Therefore we get the following lemma.

Lemma 47. (i) $(\Sigma_1\text{-}\mathsf{Ind})^N$ is provable in $\mathsf{IKP}_0^\sharp + (\Delta_0^{s-}\text{-}\mathsf{MP}) + (\mathsf{w}\Sigma_1\text{-}\mathsf{Ind})$.
(ii) $(\Sigma_1\text{-}\mathsf{Ind}_\omega)^N$ is provable in $\mathsf{IKP}_0^\sharp + (\Delta_0^{s-}\text{-}\mathsf{MP}) + (\mathsf{w}\Sigma_1\text{-}\mathsf{Ind}_\omega)$.
(iii) $(\Pi_2\text{-}\mathsf{Ref})^N$ is provable in $\mathsf{IKP}_0^\sharp + (\Delta_0^{s-}\text{-}\mathsf{MP}) + (\mathsf{w}\Pi_2\text{-}\mathsf{Ref})$.

Proof. For the first assertion, by Lemma 37 (i),

$$(((\forall y \in x)A[y] \to A[x]) \to \forall x A[x])^N \leftrightarrow ((\forall y \in x)A^N[y] \to A^N[x]) \to \forall x A^N[x]$$

where if A is Σ_1 then A^N is equivalently weak Σ_1 by (Δ_0^{s-}-MP).

The proof of the second assertion is analogous.

For the third assertion, for any Δ_0 formula $A[x,y,\vec{u}]$, we have to show

$$((\forall x)(\exists y)A[x,y,\vec{u}])^N \to ((\exists v)(\vec{u} \in v \wedge \mathrm{Tran}[v] \wedge \mathsf{Ad}(v) \wedge (\forall x \in v)(\exists y \in v)A[x,y,\vec{u}]))^N.$$

Assume $((\forall x)(\exists y)A[x,y,\vec{u}])^N$, that is $(\forall x)\neg(\forall y)\neg A^N[x,y,\vec{u}]$. By ($\Delta_0^{s-}$-$\mathsf{MP}$), we have

$$(\forall x)(\exists z)\neg(\forall y \in z)\neg A^N[x,y,\vec{u}]$$

and, by ($\mathsf{w}\Pi_2$-Ref), there is v with $\vec{u} \in v$, $\mathrm{Tran}[v]$ and $\mathsf{Ad}(v)$ such that

$$(\forall x \in v)(\exists z \in v)\neg(\forall y \in z)\neg A^N[x,y,\vec{u}].$$

Since $\mathrm{Tran}[v]$, we have $(\forall x \in v)\neg(\forall y \in v)\neg A^N[x,y,\vec{u}]$. $\square$

5.2. Interpretability result

We can define now a class of formulae, for which KP^{int} is conservative over $\mathsf{IKP}^\sharp + (\Delta_0^{s-}\text{-}\mathsf{MP})$.

Definition 48 ($\mathcal{C}_{res}$)**.** The set $\mathcal{C}_{res}$ of $\mathcal{L}$ formulae is inductively defined as follows:

(i) Every weak Σ_1 formula is in $\mathcal{C}_{res}$.
(ii) If A and B are in $\mathcal{C}_{res}$, then also $A \wedge B$ and $A \to B$ are in $\mathcal{C}_{res}$.
(iii) If A is in $\mathcal{C}_{res}$, then also $\forall x A$ is in $\mathcal{C}_{res}$.

In particular A^N is in $\mathcal{C}_{res}$ for an arbitrary $\mathcal{L}$ formula A.

Lemma 49. For each formula A in $\mathcal{C}_{res}$, there is a strongly negative $\mathcal{L}$ formula A' such that KP_0^{int} proves and (Δ_0^{s-}-MP) implies (intuitionistically) that A and A' are equivalent.

Proof. The proof is by induction on the length of A. If A is a very weak Σ_1 formula it is of the form $\exists y \neg(\forall x \in y)C[x]$ where $C[x]$ is a strongly negative Δ_0 formula. Let A' be $\neg\forall x C[x]$ which is strongly negative. Then A is by Lemma 33 (15) equivalent to A' over (Δ_0^{s-}-MP). Furthermore $A \to A'$ is classically valid and KP^{int} proves $A' \to A$ (if there is z for which $C[z]$ does not hold, then KP^{int} proves $\neg(\forall x \in \{z\})C[x]$ for this z). If A is $C \wedge D$, $C \to D$ or $\forall x C$ for C and D in $\mathcal{C}_{res}$, the assertion easily follows from the induction hypothesis. $\square$

Theorem 50. Let $\mathcal{A}$ be a set of $\mathcal{L}$ formulae such that C^N is provable in $\mathsf{IKP}_0^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP}) + \mathcal{A}$ for all $C \in \mathcal{A}$. If an $\mathcal{L}$ formula B is in $\mathcal{C}_{res}$ and provable in $\mathsf{KP}_0^{int} + \mathcal{A}$, then B is also provable in $\mathsf{IKP}_0^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP}) + \mathcal{A}$.

All the assertions, with the subscript 0 replaced by ω or just omitted, also hold.

Furthermore, the analogous assertions hold if (Σ_1-Ind), (Σ_1-Ind_ω) or (Π_2-Ref) are added to $\mathsf{KP}_0^{int} + \mathcal{A}$ and at the same time ($\mathsf{w}\Sigma_1$-Ind), ($\mathsf{w}\Sigma_1$-Ind_ω) or ($\mathsf{w}\Pi_2$-Ref), respectively, are added to $\mathsf{IKP}_0^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP}) + \mathcal{A}$.

The proof of the theorem is based on the fact, that if a formula A is classically valid, then A^N is intuitionistically valid. Therefore it is necessary to have the restriction on the set of axioms $\mathcal{A}$ in the formulation of the theorem. The restriction ensures that every negative interpretation of an axiom of $\mathcal{A}$ is provable in $\mathsf{IKP}^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP}) + \mathcal{A}$. For instance the restriction seems not to be fulfilled if $\mathcal{A}$ is the set of all instances of Δ_0 collection.

Proof of Theorem 50. Assume B is in $\mathcal{C}_{res}$ and provable in $\mathsf{KP}^{int} + \mathcal{A}$. Let B' be as in the previous lemma and let A be the conjunction of all non-logical axioms occurring in some proof in $\mathsf{KP}^{int} + \mathcal{A}$ of B'. Let all the free variables occurring in A be among $\vec{x}$. By the deduction theorem it follows that $(\forall \vec{x} A) \to B'$ is classically valid and therefore $(\forall \vec{x} A^N) \to B'^N$ is intuitionistically valid. By Lemma 45 and the property of $\mathcal{A}$ we know that A^N is provable in $\mathsf{IKP}^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP}) + \mathcal{A}$. The stated assertion follows by modus ponens, Lemma 36 and the properties of B'.

The proof for the analogous assertions is analogous, for some of which Lemma 47 is used. □

6. Interpreting $\mathsf{IKP}^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP})$ in IKP^-: Forcing Interpretation

In this section, we see how to interpret the intermediate theory $\mathsf{IKP}^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP})$ in IKP^-. This is the twist that we need and it is the second step towards the combined interpretation of KP in WEST. The method is the so-called intuitionistic forcing interpretation, which is more or less a formalisation of Kripke semantics in an intuitionistic meta-theory with some modification in the interpretation of $\perp$ (for more details, see Avigad [2, Section 2]). Many instances of forcing commonly used in set theory (see Kunen [43, Chapter VII] or Jech [40, Chapter 3]) and those used in proof theory (see Avigad [3]), that are used to interpret classical theories, can be seen as the combination of forcing method in the present sense and negative interpretation (as explained in Avigad [3, Subsection 2.2]). Unfortunately, in this part, the interpretation depends on proofs to be interpreted and so we have only local interpretability, as Avigad [2]. However this is not so essential because we can replace the intermediate theories so that it becomes non-local (see Remark 74).

We follow Avigad [2, Section 5]. However we need to work in IKP^- while Avigad works in IKP. We have to check carefully whether the argument works in this weaker system in each step, and actually we need to make some changes on defined notions. We already changed the definitions of $\mathsf{IKP}^\sharp$ from Avigad's in Section 4. Since we have separation only for negative Δ_0 formulae we change our definition of the abbreviations concerning $a \cup b$, $\cup a$ and $\{a_0, ..., a_n\}$ for this section as follows:

- $a = \{a_0, ..., a_n\}$ is the formula $a_0 \in a \wedge ... \wedge a_n \in a \wedge (\forall x \in a)\neg(x \neq a_0 \wedge ... \wedge x \neq a_n)$,
- $a = \cup b$ is the formula $(\forall x \in b)(\forall y \in x)(y \in a) \wedge (\forall y \in a)\neg(\forall x \in b)(\neg y \in x)$,
- $a = b \cup c$ is the formula $a = \cup\{b, c\}$.

The axioms pairing, union and Δ_0^- separation assure that IKP^- proves the existence of sets corresponding to $a \cup b$, $\cup a$ and $\{a_0, ..., a_n\}$ due to these abbreviations. We will also use in this section abbreviations of the form $A[\{a_0, ..., a_n\}]$, $A[\cup a]$ and $A[a \cup b]$ in the sense of the redefined abbreviations ($A[\cup a]$, for instance, is an abbreviation for $\forall x(x = \cup a \to A[x])$ which is equivalent to $\exists x(x = \cup a \wedge A[x])$ if x occurs only on the right side of $\in$, as it will be almost always.).

6.1. Definition of the auxiliary relation $\Vdash_{\mathfrak{S}}$

To embed $\mathsf{IKP}^\sharp$ and $(\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP})$ in IKP^- we will use Avigad's forcing method. We will introduce a forcing relation in IKP^- for each proof of some formula in $\mathsf{IKP}^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP})$. For that purpose we need a kind of a truth predicate for a finite set of strongly negative Δ_0 formulae of $\mathcal{L}$.

Definition 51 ($\mathrm{Tr}_{\mathfrak{S}}$)**.** Let $\mathfrak{S}$ be a finite sequence $D_0[z,\vec{y}],...,D_{n-1}[z,\vec{y}]$ of strongly negative Δ_0 formulae with at most the variables z, $\vec{y}=y_0,...,y_m$ free. Then $\mathrm{Tr}_{\mathfrak{S}}[p,u]$ is a strongly negative Δ_0 formula equivalent to

$$\bigwedge_{i=0}^{n-1} \forall \vec{y}(\langle i,y_0,...,y_m\rangle \in p \to (\forall z \in u)D_i[z,\vec{y}]) \ \wedge \ \bigwedge_{i=n}^{2n-1} \forall z \forall \vec{y}(\langle i,z,y_0,...,y_m\rangle \in p \to D_{i-n}[z,\vec{y}]).$$

In the next definition we introduce for each $\mathcal{L}$ formula A the formula $p \Vdash_{\mathfrak{S}} A$. In Avigad's words it means that there is u which is "sufficiently large to witness the fact that A follows from the formulae in p" [2, p. 20], where $\langle i,y_0,...,y_m\rangle \in p$ encodes that the formula $(\forall z \in u)D_i[z,\vec{y}]$ is in p and $\langle i+n,z,y_0,...,y_m\rangle \in p$ encodes that the formula $D_i[z,\vec{y}]$ is in p

Definition 52 ($\Vdash_{\mathfrak{S}}$)**.** Let $\mathfrak{S}$ be a finite sequence $D_0[z,\vec{y}],...,D_{n-1}[z,\vec{y}]$ of strongly negative Δ_0 formulae with at most the variables z, $\vec{y}$ free. For any $\mathcal{L}$ formula A we define

$$p \Vdash_{\mathfrak{S}} A \ := \ \exists u(\mathrm{Tr}_{\mathfrak{S}}[p,u] \to A),$$

where u is a variable different from z, $\vec{y}$ and not occurring in A.

If $B[z,\vec{y}]$ is the $(i+1)$-th formula in the sequence $\mathfrak{S}$ we write, for arbitrary variables $\vec{x}=x_0,...,x_m$ and z,

- $p, \forall z B[z,\vec{x}] \Vdash_{\mathfrak{S}} A$ for $p \cup \{\langle i,x_0,...,x_m\rangle\} \Vdash_{\mathfrak{S}} A$,
- $p, B[z,\vec{x}] \Vdash_{\mathfrak{S}} A$ for $p \cup \{\langle i{+}n,z,x_0,...,x_m\rangle\} \Vdash_{\mathfrak{S}} A$.

Furthermore $p,q \Vdash_{\mathfrak{S}} A$ abbreviates $p \cup q \Vdash_{\mathfrak{S}} A$ and $\Vdash_{\mathfrak{S}} A$ abbreviates $\emptyset \Vdash_{\mathfrak{S}} A$.

Notice that this relation is not interesting in the classical context, for the defining formula $\exists u(\mathrm{Tr}_{\mathfrak{S}}[p,u] \to A)$ is classically equivalent to A if $\exists u \mathrm{Tr}_{\mathfrak{S}}[p,u]$ and to $\top$ otherwise, since u does not occur in A.

Here we restrict D_i's to be strongly negative formulae. Though in Zumbrunnen [59, Definition 2.20] all negative formulae are allowed, the restriction of strong-negativity seems necessary, since otherwise it seems impossible to have the following lemma, which is implicitly used in many places also in Zumbrunnen [59]. The second author admits the error and corrects all the related assertions as in the present paper.

Lemma 53. Let $\mathfrak{S}$ be as in the last definition and B a formula occurring in $\mathfrak{S}$. Then, for any $\mathcal{L}$ formula A, IKP_0^- proves

(i) $p, \forall z B[z,\vec{x}] \Vdash_{\mathfrak{S}} A$ is equivalent to $\exists u(\mathrm{Tr}_{\mathfrak{S}}(p,u) \wedge (\forall z \in u)B[z,\vec{x}] \to A)$,
(ii) $p, B[z,\vec{x}] \Vdash_{\mathfrak{S}} A$ is equivalent to $\exists u(\mathrm{Tr}_{\mathfrak{S}}(p,u) \wedge B[z,\vec{x}] \to A)$.

Proof. Let $\mathfrak{S}$ be the sequence $D_0[z,\vec{y}],...,D_{n-1}[z,\vec{y}]$ and let also $B[z,\vec{y}]$ be $D_k[z,\vec{y}]$ with $k \le n-1$. Let $q = p \cup \{\langle k,x_0,...,x_m\rangle\}$, guaranteed by Δ_0^- separation, namely q satisfies

$$p \subseteq q \wedge \langle k,x_0,...,x_m\rangle \in q \wedge (\forall w \in q)\neg(\neg(w \in p) \wedge \neg(w = \langle k,x_0,...,x_m\rangle)).$$

For the first assertion, we prove the equivalence between $\mathrm{Tr}_{\mathfrak{S}}[q,u]$ and $\mathrm{Tr}_{\mathfrak{S}}(p,u) \wedge (\forall z \in u)B[z,\vec{x}]$.

For one direction, assume $\mathrm{Tr}_{\mathfrak{S}}(p,u) \wedge (\forall z \in u)B[z,\vec{x}]$, namely

$$\begin{aligned} &\bigwedge_{i=0}^{n-1} \forall \vec{y}(\langle i,\vec{y}\rangle \in p \to (\forall z \in u)D_i[z,\vec{y}]) \wedge \bigwedge_{i=n}^{2n-1} \forall \vec{y}(\langle i,\vec{y}\rangle \in p \to D_{i-n}[z,\vec{y}]) \\ &\wedge \bigwedge_{i=0}^{n-1} \forall \vec{y}(\langle i,\vec{y}\rangle = \langle k,\vec{x}\rangle \to (\forall z \in u)D_i[z,\vec{y}]) \wedge \bigwedge_{i=n}^{2n-1} \forall \vec{y}(\langle i,\vec{y}\rangle = \langle k,\vec{x}\rangle \to D_{i-n}[z,\vec{y}]). \end{aligned}$$

This implies

$$\begin{aligned} &\bigwedge_{i=0}^{n-1} \forall \vec{y}(\neg(\forall z \in u)D_i[z,\vec{y}] \to \neg\langle i,\vec{y}\rangle \in p) \wedge \bigwedge_{i=n}^{2n-1} \forall \vec{y}(\neg D_{i-n}[z,\vec{y}] \to \neg\langle i,\vec{y}\rangle \in p) \\ &\wedge \bigwedge_{i=0}^{n-1} \forall \vec{y}(\neg(\forall z \in u)D_i[z,\vec{y}] \to \neg\langle i,\vec{y}\rangle = \langle k,\vec{x}\rangle) \wedge \bigwedge_{i=n}^{2n-1} \forall \vec{y}(\neg D_{i-n}[x,\vec{y}] \to \neg\langle i,\vec{y}\rangle = \langle k,\vec{x}\rangle), \end{aligned}$$

which is, Lemma 33 (3), equivalent to

$$\bigwedge_{i=0}^{n-1} \forall \vec{y}(\neg(\forall z \in u) D_i[z,\vec{y}] \to \neg\langle i,\vec{y}\rangle \in p \wedge \neg\langle i,\vec{y}\rangle = \langle k,\vec{x}\rangle)$$
$$\wedge \bigwedge_{i=n}^{2n-1} \forall \vec{y}(\neg D_{i-n}[z,\vec{y}] \to \neg\langle i,\vec{y}\rangle \in p \wedge \neg\langle i,\vec{y}\rangle = \langle k,\vec{x}\rangle).$$

Furthermore, this implies

$$\bigwedge_{i=0}^{n-1} \forall \vec{y}(\neg(\neg\langle i,\vec{y}\rangle \in p \wedge \neg\langle i,\vec{y}\rangle = \langle k,\vec{x}\rangle) \to \neg\neg(\forall z \in u) D_i[z,\vec{y}])$$
$$\wedge \bigwedge_{i=n}^{2n-1} \forall \vec{y}(\neg(\neg\langle i,\vec{y}\rangle \in p \wedge \neg\langle i,\vec{y}\rangle = \langle k,\vec{x}\rangle) \to \neg\neg D_{i-n}[z,\vec{y}]).$$

and finally

$$\bigwedge_{i=0}^{n-1} \forall \vec{y}(\langle i,\vec{y}\rangle \in q \to \neg\neg(\forall z \in u) D_i[z,\vec{y}]) \wedge \bigwedge_{i=n}^{2n-1} \forall \vec{y}(\langle i,\vec{y}\rangle \in q \to \neg\neg D_{i-n}[z,\vec{y}]).$$

Since D_i's are strongly negative, this means $\mathrm{Tr}_{\mathfrak{S}}[q,u]$.

Conversely, $\mathrm{Tr}_{\mathfrak{S}}[q,u]$ means

$$\bigwedge_{i=0}^{n-1} (\forall \vec{y})(\langle i,\vec{y}\rangle \in q \to (\forall z \in u) D_i[z,\vec{y}]) \wedge \bigwedge_{i=n}^{2n-1} (\forall \vec{y})(\langle i,\vec{y}\rangle \in q \to D_{i-n}[z,\vec{y}])$$

which implies

$$\bigwedge_{i=0}^{n-1} (\forall \vec{y})(\langle i,\vec{y}\rangle \in p \to (\forall z \in u) D_i[z,\vec{y}]) \wedge \bigwedge_{i=n}^{2n-1} (\forall \vec{y})(\langle i,\vec{y}\rangle \in p \to D_{i-n}[z,\vec{y}]) \wedge (\forall z \in u) B[z,\vec{x}],$$

namely $\mathrm{Tr}_{\mathfrak{S}}(p,u) \wedge (\forall z \in u) B[z,\vec{x}]$.

Similarly we can prove the second assertion. □

6.2. Forcing-like behaviour of $\Vdash_{\mathfrak{S}}$

Lemma 54. For an arbitrary $\mathcal{L}$ formula A, IKP_0^- proves that $a \subseteq b$ implies:

(i) $\mathrm{Tr}_{\mathfrak{S}}[p,b]$ implies $\mathrm{Tr}_{\mathfrak{S}}[p,a]$ and
(ii) $\mathrm{Tr}_{\mathfrak{S}}[p,a] \to A$ implies $\mathrm{Tr}_{\mathfrak{S}}[p,b] \to A$.

Lemma 55. Let A and B be arbitrary $\mathcal{L}$ formulae, $C[z,\vec{y}]$ a strongly negative Δ_0 formula of $\mathcal{L}$ occurring in $\mathfrak{S}$ and D an arbitrary Δ_0 formula of $\mathcal{L}$. Then IKP_0^- proves the following:

(i) if $p \Vdash_{\mathfrak{S}} A$ and $p \subseteq q$ then $q \Vdash_{\mathfrak{S}} A$;
(ii) $C[z,\vec{x}] \Vdash_{\mathfrak{S}} C[z,\vec{x}]$ for arbitrary variables $\vec{x} = x_0,...,x_m$ and z;
(iii) $p \Vdash_{\mathfrak{S}} (A \wedge B)$ if and only if $p \Vdash_{\mathfrak{S}} A$ and $p \Vdash_{\mathfrak{S}} B$;
(iv) $p \Vdash_{\mathfrak{S}} (C[z,\vec{x}] \to A)$ if and only if $p, C[z,\vec{x}] \Vdash_{\mathfrak{S}} A$ for arbitrary variables $\vec{x} = x_0,...,x_m$ and z;
(v) If $p \Vdash_{\mathfrak{S}} (A \to B)$ and $q \Vdash_{\mathfrak{S}} A$ then $r \Vdash_{\mathfrak{S}} B$ for any r with $r \supseteq p$ and $r \supseteq q$;
(vi) The following are equivalent (if x is a variable not occurring in $\mathrm{Tr}_{\mathfrak{S}}[p,u]$)
   (a) $p \Vdash_{\mathfrak{S}} (\forall x \in z) D$,
   (b) $\forall x(p \Vdash_{\mathfrak{S}} (x \in z \to D))$ and
   (c) $(\forall x \in z)(p \Vdash_{\mathfrak{S}} D)$.

Proof. The proof works exactly as in Avigad [2, Lemma 5.5]. The first assertion follows directly from Definition 52. By Lemma 53 with $p = \emptyset$, $C[z, \vec{x}] \Vdash\!\vdash_{\mathfrak{S}} C[z, \vec{x}]$ is equivalent to $C[z, \vec{x}] \to C[z, \vec{x}]$.

The direction from left to right of the third assertion follows from the fact that the formula $\mathrm{Tr}_{\mathfrak{S}}[p, a] \to (A \wedge B)$ implies $\mathrm{Tr}_{\mathfrak{S}}[p, a] \to A$ and $\mathrm{Tr}_{\mathfrak{S}}[p, a] \to B$. For the converse direction assume $\mathrm{Tr}_{\mathfrak{S}}[p, a] \to A$ and $\mathrm{Tr}_{\mathfrak{S}}[p, b] \to B$. Then it is provable by the previous lemma in IKP_0^- that $\mathrm{Tr}_{\mathfrak{S}}[p, c] \to (A \wedge B)$ if $c \supseteq a \cup b$.

By Lemma 33 (1), IKP_0^- proves that

$$\mathrm{Tr}_{\mathfrak{S}}[p, a] \to (C[z, \vec{x}] \to A) \text{ is equivalent to } (\mathrm{Tr}_{\mathfrak{S}}[p, a] \wedge C[z, \vec{x}]) \to A.$$

By Lemma 53 the latter is equivalent to $\mathrm{Tr}_{\mathfrak{S}}[p \cup \{\langle i{+}n, z, x_0, ..., x_m\rangle\}, a] \to A$ if $C[z, \vec{y}]$ is the $(i{+}1)$-th formula of $\mathfrak{S}$. Therefore the fourth assertion holds.

By the previous lemma and the first assertion, IKP_0^- proves that $\mathrm{Tr}_{\mathfrak{S}}[p, a] \to (A \to B)$ and $\mathrm{Tr}_{\mathfrak{S}}[q, b] \to A$ imply together $\mathrm{Tr}_{\mathfrak{S}}[r, a \cup b] \to B$. Hence the fifth assertion holds.

In the sixth assertion we have by Lemma 33 (17) and (18) that (a) implies intuitionistically (b) and (b) implies intuitionistically (c), respectively. In the following we work informally within IKP_0^- and prove that (c) implies (a). So assume (c), namely

$$(\forall x \in z)\exists u(\mathrm{Tr}_{\mathfrak{S}}[p, u] \to D).$$

By Δ_0 collection it follows that there is w such that

$$(\forall x \in z)(\exists u \in w)(\mathrm{Tr}_{\mathfrak{S}}[p, u] \to D).$$

If $v \supseteq \cup w$, we have that $u \in w$ implies $u \subseteq v$ and hence by the previous lemma

$$(\forall x \in z)(\mathrm{Tr}_{\mathfrak{S}}[p, v] \to D),$$

which is equivalent to $\mathrm{Tr}_{\mathfrak{S}}[p, v] \to (\forall x \in z)D$ by Lemma 33 (2) and (11) since x does not occur in $\mathrm{Tr}_{\mathfrak{S}}[p, v]$. □

6.3. Avigad forcing $\Vdash_{\mathfrak{S}}$

Definition 56 ($\Vdash_{\mathfrak{S}}$)**.** Let $\mathfrak{S}$ be a finite sequence $D_0[z, \vec{y}], ..., D_n[z, \vec{y}]$ of strongly negative Δ_0 formulae with at most the variables z, $\vec{y}$ free. For an arbitrary $\mathcal{L}$ formula A, the $\mathcal{L}$ formula $p \Vdash_{\mathfrak{S}} A$ is defined inductively as follows:

(i) if A is atomic, then $p \Vdash_{\mathfrak{S}} A$ is the formula $p \Vdash\!\vdash_{\mathfrak{S}} \neg\neg A$;
(ii) if A is $B \wedge C$, then $p \Vdash_{\mathfrak{S}} A$ is the formula $(p \Vdash_{\mathfrak{S}} B) \wedge (p \Vdash_{\mathfrak{S}} C)$;
(iii) if A is $B \vee C$, then $p \Vdash_{\mathfrak{S}} A$ is the formula $(p \Vdash_{\mathfrak{S}} B) \vee (p \Vdash_{\mathfrak{S}} C)$;
(iv) if A is $B \to C$, then $p \Vdash_{\mathfrak{S}} A$ is the formula

$$\forall q(p \subseteq q \to ((q \Vdash_{\mathfrak{S}} B) \to (q \Vdash_{\mathfrak{S}} C)));$$

(v) if A is $\forall x B[x]$ then $p \Vdash_{\mathfrak{S}} A$ is the formula $\forall x_0(p \Vdash_{\mathfrak{S}} B[x_0])$ where x_0 is not in $\mathrm{Tr}_{\mathfrak{S}}[p, u]$ nor in $B[x]$;
(vi) if A is $\exists x B[x]$ then $p \Vdash_{\mathfrak{S}} A$ is the formula $\exists x_0(p \Vdash_{\mathfrak{S}} B[x_0])$ where x_0 is not in $\mathrm{Tr}_{\mathfrak{S}}[p, u]$ nor in $B[x]$.

The abbreviations $p, \forall z B[z, \vec{x}] \Vdash_{\mathfrak{S}} A$; $p, B[z, \vec{x}] \Vdash_{\mathfrak{S}} A$ (for B in $\mathfrak{S}$); $p, q \Vdash_{\mathfrak{S}} A$ and $\Vdash_{\mathfrak{S}} A$ are defined in the same way as for $\Vdash\!\vdash_{\mathfrak{S}}$ in Definition 52.

In the case of atomic formulae, there is a change from Avigad's definition, which seems necessary for our weaker interpreting theory IKP^-.

Again, this forcing relation is uninteresting in the classical context, because classically $p \Vdash_{\mathfrak{S}} A$ is equivalent to A if $\exists u \mathrm{Tr}_{\mathfrak{S}}[p, u]$ and to $\top$ otherwise.

The next lemma corresponds to Proposition 2.4 in Avigad [2] and its first assertion is proved as ibidem (it is easy to see that the forcing relation $\Vdash_{\mathfrak{S}}$ is "good" in the sense of Avigad [2, Definition 2.2]). Its second assertion is a direct consequence of the first one.

Lemma 57. For any sequence $\mathfrak{S}$ of strongly negative Δ_0 formulae and $\mathcal{L}$ formulae A, B and $C_0, ..., C_n$,

(i) if A is provable from $\{C_0, ..., C_n\}$ intuitionistically, then $p \Vdash_{\mathfrak{S}} A$ is provable from IKP_0^- and $\{p \Vdash_{\mathfrak{S}} C_0, ..., p \Vdash_{\mathfrak{S}} C_n\}$;

(ii) if B is an intuitionistic consequence of A, then IKP_0^- proves

$$p \Vdash_{\mathfrak{S}} A \quad \text{implies} \quad p \Vdash_{\mathfrak{S}} B.$$

Definition 58 (Prominent formulae)**.** We call an $\mathcal{L}$ formula $A[z, \vec{y}]$ *prominent for* $\mathfrak{S}$ if it is a strongly negative Δ_0 formula such that all subformulae of A, except atomic formulae, and the double negation of all the atomic subformulae (except $\bot$) of A occur in the sequence $\mathfrak{S}$.

In what follows, for simplicity, we always assume that $\neg y_0 \in y_1$ occurs in $\mathfrak{S}$.

6.4. Partial equivalence between auxiliary $\Vvdash_{\mathfrak{S}}$ and forcing $\Vdash_{\mathfrak{S}}$

Lemma 59. Let $A[z, \vec{y}]$ be a prominent formula for $\mathfrak{S}$. Then IKP_0^- proves for arbitrary variables $\vec{x} = x_0, ..., x_m$ and z that

$$p \Vdash_{\mathfrak{S}} A[z, \vec{x}] \text{ is equivalent to } p \Vvdash_{\mathfrak{S}} A[z, \vec{x}].$$

Proof. The proof is similar to the proof of Lemma 5.6 in Avigad [2] by induction on the complexity of A. If $A[z, \vec{x}]$ is atomic, namely $\bot$, the assertion follows by definition and the equivalence between $\bot$ and $\neg\neg\bot$. If $A[z, \vec{x}]$ is a formula of the form $B[z, \vec{x}] \wedge C[z, \vec{x}]$ then the assertion follows by Lemma 55 (iii) and the induction hypothesis.

If $A[z, \vec{x}]$ is of the form $B[z, \vec{x}] \to C[z, \vec{x}]$ where $B[z, \vec{x}]$ is atomic and $C[z, \vec{x}]$ is strongly negative, we know by definition and by the induction hypothesis that $p \Vdash_{\mathfrak{S}} A[z, \vec{x}]$ is equivalent to

$$\forall q(p \subseteq q \to ((q \Vvdash_{\mathfrak{S}} \neg\neg B[z, \vec{x}]) \to (q \Vvdash_{\mathfrak{S}} C[z, \vec{x}]))). \tag{2}$$

On the other hand, by the strong negativeness of C, $\neg\neg B[z, \vec{x}] \to C[z, \vec{x}]$ is intuitionistically equivalent to $B[z, \vec{x}] \to C[z, \vec{x}]$ and hence $p \Vvdash_{\mathfrak{S}} A[z, \vec{x}]$ is equivalent to

$$p \Vvdash_{\mathfrak{S}} (\neg\neg B[z, \vec{x}] \to C[z, \vec{x}]). \tag{3}$$

We assume for one direction that (2) holds. Since $\neg\neg B$ occurs in $\mathfrak{S}$, we have $p, \neg\neg B[z, \vec{x}] \Vvdash_{\mathfrak{S}} \neg\neg B[z, \vec{x}]$ and so, by (2), we have $p, \neg\neg B[z, \vec{x}] \Vvdash_{\mathfrak{S}} C[z, \vec{x}]$. Now, by Lemma 55 (iv), we obtain (3). For the other direction assume (3). By Lemma 55 (v) we have for an arbitrary $q \supseteq p$ with $q \Vvdash_{\mathfrak{S}} \neg\neg B[z, \vec{x}]$ that $q \Vvdash_{\mathfrak{S}} C[z, \vec{x}]$. Thus we get (2) and therefore $p \Vdash_{\mathfrak{S}} A[z, \vec{x}]$.

If $A[z, \vec{x}]$ is of the form $B[z, \vec{x}] \to C[z, \vec{x}]$ where both $B[z, \vec{x}]$ and $C[z, \vec{x}]$ are strongly negative, since we may assume theta B is not atomic (i.e., $\bot$), we know by the induction hypothesis that $p \Vdash_{\mathfrak{S}} A[z, \vec{x}]$ is equivalent to

$$\forall q(p \subseteq q \to ((q \Vvdash_{\mathfrak{S}} B[z, \vec{x}]) \to (q \Vvdash_{\mathfrak{S}} C[z, \vec{x}]))). \tag{4}$$

We assume for one direction that $p \Vdash_{\mathfrak{S}} A[z, \vec{x}]$ holds. We have clearly $p, B[z, \vec{x}] \Vvdash_{\mathfrak{S}} B[z, \vec{x}]$ and therefore it follows by (4) that $p, B[z, \vec{x}] \Vvdash_{\mathfrak{S}} C[z, \vec{x}]$. We get by Lemma 55 (iv)

$$p \Vvdash_{\mathfrak{S}} (B[z, \vec{x}] \to C[z, \vec{x}]).$$

For the other direction assume $p \Vvdash_{\mathfrak{S}} (B[z, \vec{x}] \to C[z, \vec{x}])$. By Lemma 55 (v) we have for an arbitrary $q \supseteq p$ with $q \Vvdash_{\mathfrak{S}} B[z, \vec{x}]$ that $q \Vvdash_{\mathfrak{S}} C[z, \vec{x}]$. Thus we get (4) and therefore $p \Vdash_{\mathfrak{S}} A[z, \vec{x}]$.

If A is of the form $(\forall v \in a)B[v, z, \vec{x}]$, then $p \Vdash_{\mathfrak{S}} A[z, \vec{x}]$ is the formula

$$\forall v_0(p \Vdash_{\mathfrak{S}} (v_0 \in a \to B[v_0, z, \vec{x}]))$$

(where v_0 does not occur in $\mathrm{Tr}_{\mathfrak{S}}[p, u]$) which is by induction hypothesis equivalent to

$$\forall v_0(p \Vvdash_{\mathfrak{S}} (v_0 \in a \to B[v_0, z, \vec{x}])).$$

Since v_0 does not appear in $\mathrm{Tr}_{\mathfrak{S}}[p, u]$, this is by Lemma 55 (vi) equivalent to $p \Vvdash_{\mathfrak{S}} (\forall v_0 \in a)B[v_0, z, \vec{x}]$, that is $p \Vvdash_{\mathfrak{S}} A[z, \vec{x}]$. $\square$

6.5. Bounded quantifiers in forcing

The next lemma corresponds to the Lemmata 5.12 and 5.13 in Avigad [2] and is proved as ibidem.

Lemma 60. Let A be a negative Δ_0 formula, $B[x]$ an arbitrary Δ_0 formula of $\mathcal{L}$ in which y does not occur, $C[x,y]$ an arbitrary $\mathcal{L}$ formula and $D[x,y]$ any strongly negative $\mathcal{L}$ formula. Then IKP_0^- proves:

(i) $A \to \exists x B[x]$ implies $\exists y(A \to (\exists x \in y)B[x])$,
(ii) $p \Vdash_{\mathfrak{S}} (x \in y \to C[x,y])$ implies $x \in y \to (p \Vdash_{\mathfrak{S}} C[x,y])$,
(iii) $x \in y \to (p \Vvdash_{\mathfrak{S}} C[x,y])$ implies $p \Vvdash_{\mathfrak{S}} (x \in y \to C[x,y])$ and
(iv) $x_0 \in x_1 \to (p \Vdash_{\mathfrak{S}} D[x_0,x_1])$ is equivalent to $p \Vdash_{\mathfrak{S}} (x_0 \in x_1 \to D[x_0,x_1])$ for all variables x_0, x_1 not occurring in $\mathrm{Tr}_{\mathfrak{S}}[p,u]$.

Proof. For the first assertion we assume $A \to \exists x B[x]$. By pairing and Δ_0^- separation there is an empty set $\emptyset$ (i.e. $(\forall w \in \emptyset)\bot$) and u which contains $\emptyset$. By Δ_0^- separation there is $\{\emptyset\} = \{w \in u : (\forall v \in w)\bot\}$. In the following we assume that w does not appear in A nor in $B[x]$. By the same axioms there is a set a such that

$$\forall w(w \in a \leftrightarrow (w \in \{\emptyset\} \wedge A)).$$

It follows that $w \in a$ implies A and therefore we have by our assumption $w \in a \to \exists x B[x]$. It follows $(\forall w \in a)\exists x B[x]$ and by Δ_0 collection

$$\exists y(\forall w \in a)(\exists x \in y)B[x].$$

Since A implies $\emptyset \in a$ we can conclude $\exists y(A \to (\exists x \in y)B[x])$.

For the second assertion assume $p \Vdash_{\mathfrak{S}} (x \in y \to C[x,y])$ which is by definition

$$\forall q(p \subseteq q \to ((q \Vdash_{\mathfrak{S}} x \in y) \to (q \Vdash_{\mathfrak{S}} C[x,y]))).$$

Since $x \in y$ implies $\neg\neg x \in y$ and $p \Vdash_{\mathfrak{S}} x \in y$, it follows $x \in y \to (p \Vdash_{\mathfrak{S}} C[x,y])$.

For the third assertion, $x_0 \in x_1 \to (p \Vvdash_{\mathfrak{S}} C[x_0,x_1])$ is by definition $x_0 \in x_1 \to \exists u(\mathrm{Tr}_{\mathfrak{S}}[p,u] \to C[x_0,x_1])$. By the first assertion there is w such that $x_0 \in x_1 \to (\exists u \in w)(\mathrm{Tr}_{\mathfrak{S}}[p,u] \to C[x_0,x_1])$. If $v \supseteq \cup w$ it follows by Lemma 54 $x_0 \in x_1 \to (\mathrm{Tr}_{\mathfrak{S}}[p,v] \to C[x_0,x_1])$ which is equivalent to $\mathrm{Tr}_{\mathfrak{S}}[p,v] \to (x_0 \in x_1 \to C[x_0,x_1])$ by Lemma 33 (2). Hence we have $p \Vvdash_{\mathfrak{S}} (x_0 \in x_1 \to C[x_0,x_1])$.

The direction from right to left of the fourth assertion follows from the second one. The direction from left to right is proved by induction on the complexity of D. When D is atomic, that is $\bot$, we assume $x_0 \in x_1 \to (p \Vdash_{\mathfrak{S}} D[x_0,x_1])$, that is, $x_0 \in x_1 \to (p \Vvdash_{\mathfrak{S}} D[x_0,x_1])$. By the third assertion, we have $p \Vvdash_{\mathfrak{S}} (x_0 \in x_1 \to D[x_0,x_1])$ which is by Lemma 59 equivalent to $p \Vdash_{\mathfrak{S}} (x_0 \in x_1 \to D[x_0,x_1])$, since we always assume $x_0 \in x_1 \to D[x_0,x_1]$, namely $\neg(x_0 \in x_1)$ occurs in $\mathfrak{S}$ (after Definition 58) and hence prominent for $\mathfrak{S}$.

If $D[x_0,x_1]$ is of the form $D_0 \wedge D_1$, then $x_0 \in x_1 \to (p \Vdash_{\mathfrak{S}} D[x_0,x_1])$ is

$$x_0 \in x_1 \to ((p \Vdash_{\mathfrak{S}} D_0) \wedge (p \Vdash_{\mathfrak{S}} D_1))$$

which is by Lemma 33 (3) equivalent to $(x_0 \in x_1 \to (p \Vdash_{\mathfrak{S}} D_0)) \wedge (x_0 \in x_1 \to (p \Vdash_{\mathfrak{S}} D_1))$. This implies by induction hypothesis $(p \Vdash_{\mathfrak{S}} (x_0 \in x_1 \to D_0)) \wedge (p \Vdash_{\mathfrak{S}} (x_0 \in x_1 \to D_1))$. So we have $p \Vdash_{\mathfrak{S}} ((x_0 \in x_1 \to D_0) \wedge (x_0 \in x_1 \to D_1))$ and by the Lemmata 33 (3) and 57, we have $p \Vdash_{\mathfrak{S}} (x_0 \in x_1 \to D[x_0,x_1])$.

If $D[x_0,x_1]$ is of the form $D_0 \to D_1$, then $x_0 \in x_1 \to (p \Vdash_{\mathfrak{S}} D[x_0,x_1])$ is

$$x_0 \in x_1 \to \forall q(p \subseteq q \to ((q \Vdash_{\mathfrak{S}} D_0) \to (q \Vdash_{\mathfrak{S}} D_1)))$$

which is by Lemma 33 (11),(2) and (3), equivalent to

$$\forall q(p \subseteq q \to ((x_0 \in x_1 \to (q \Vdash_{\mathfrak{S}} D_0)) \to (x_0 \in x_1 \to (q \Vdash_{\mathfrak{S}} D_1)))).$$

By the induction hypothesis for D_1 and the second assertion, it follows

$$\forall q(p \subseteq q \to ((q \Vdash_{\mathfrak{S}} (x_0 \in x_1 \to D_0)) \to (q \Vdash_{\mathfrak{S}} (x_0 \in x_1 \to D_1))))$$

and that is $p \Vdash_{\mathfrak{S}} ((x_0 \in x_1 \to D_0)) \to (x_0 \in x_1 \to D_1))$. Hence $p \Vdash_{\mathfrak{S}} x_0 \in x_1 \to D[x_0, x_1]$ by the Lemmata 33 (3) and 57.

If $D[x_0, x_1]$ is of the form $\forall z E[x_0, x_1, z]$, then $x_0 \in x_1 \to (p \Vdash_{\mathfrak{S}} D[x_0, x_1])$ is

$$x_0 \in x_1 \to \forall x_2(p \Vdash_{\mathfrak{S}} E[x_0, x_1, x_2])$$

(where x_2 does not occur in $\mathrm{Tr}_{\mathfrak{S}}[p, u]$) which is equivalent to $\forall x_2(x_0 \in x_1 \to (p \Vdash_{\mathfrak{S}} E[x_0, x_1, x_2]))$ by Lemma 33 (11). By induction hypothesis it follows $\forall x_2(p \Vdash_{\mathfrak{S}} (x_0 \in x_1 \to E[x_0, x_1, x_2]))$. That is $p \Vdash_{\mathfrak{S}} \forall z(x_0 \in x_1 \to E[x_0, x_1, z])$ and by the Lemmata 33 (11) and 57, we have $p \Vdash_{\mathfrak{S}} (x_0 \in x_1 \to D[x_0, x_1])$. □

Note that, in the base case of the proof of the fourth assertion, the strongly negativeness of $x_0 \in x_1 \to D[x_0, x_1]$ is necessary.

Lemma 61. Let A be an arbitrary $\mathcal{L}$ formula, and B any strongly negative $\mathcal{L}$ formula. Furthermore let $\vec{x} = x_0, ..., x_m$ and z be arbitrary variables. Then IKP_0^- proves:

(i) $p \Vdash_{\mathfrak{S}} (\forall x \in y)A[x, y]$ implies $(\forall x_0 \in y)(p \Vdash_{\mathfrak{S}} A[x_0, y])$ where x_0 does not occur in $\mathrm{Tr}_{\mathfrak{S}}[p, u]$ and y is arbitrary,

(ii) $(\forall x_0 \in x_1)(p \Vdash_{\mathfrak{S}} B[x_0, x_1])$ is equivalent to $p \Vdash_{\mathfrak{S}} (\forall x \in x_1)B[x, x_1]$ for all variables x_0, x_1 not occurring in $\mathrm{Tr}_{\mathfrak{S}}[p, u]$.

The proof is as the proofs of the Lemmata 5.7 and 5.13 in Avigad [2].

Proof. For the first assertion assume $p \Vdash_{\mathfrak{S}} (\forall x \in y)A[x, y]$ which is by definition

$$\forall x_0(p \Vdash_{\mathfrak{S}} (x_0 \in y \to A[x_0, y])).$$

By the second assertion of the previous lemma it follows $\forall x_0(x_0 \in y \to p \Vdash_{\mathfrak{S}} A[x_0, y])$ which is $(\forall x_0 \in y)(p \Vdash_{\mathfrak{S}} A[x_0])$.

The second assertion follows by Definition 56 (v) and the fourth of the previous lemma. □

The following corresponds to Lemma 5.7.6 in Avigad [2].

Lemma 62. For a weak Σ_1 formula $A[x, y]$ whose largest Δ_0 subformula is prominent for $\mathfrak{S}$, IKP_0^- proves $p \Vdash_{\mathfrak{S}} (\forall x \in y)A[x, y]$ is equivalent to $(\forall x \in y)(p \Vdash_{\mathfrak{S}} A[x, y])$ if x and y do not occur in $\mathrm{Tr}_{\mathfrak{S}}[p, u]$.

Proof. The direction from left to right is Lemma 61 (i).

For the converse, let $A[x, y]$ be $\exists z \neg(\forall v \in z)B[x, y, v]$ with $B[x, y, v]$ being a strongly negative Δ_0 formula. Suppose $(\forall x \in y)(p \Vdash_{\mathfrak{S}} \exists z \neg(\forall v \in z)B[x, y, v])$, which is, by definition and Lemma 59, equivalent to

$$(\forall x \in y)(\exists z')(p \Vvdash_{\mathfrak{S}} \neg(\forall v \in z')B[x, y, v]).$$

Since $p \Vvdash_{\mathfrak{S}} \neg(\forall v \in z')B[x, y, v]$ is a Σ_1 formula, by Σ_1 collection we have w such that

$$(\forall x \in y)(\exists z' \in w)(p \Vvdash_{\mathfrak{S}} \neg(\forall v \in z')B[x, y, v]).$$

If we set $w' \supseteq \cup w$, we obtain $(\forall x \in y)(p \Vvdash_{\mathfrak{S}} \neg(\forall v \in w')B[x, y, v])$, and $(\forall x \in y)(p \Vdash_{\mathfrak{S}} \neg(\forall v \in w')B[x, y, v])$ again by Lemma 59. Now, by Lemma 61 (ii), we also have

$$p \Vdash_{\mathfrak{S}} (\forall x \in y)\neg(\forall v \in w')B[x, y, v])$$

and so $p \Vdash_{\mathfrak{S}} \exists w'(\forall x \in y)\neg(\forall v \in w')B[x, y, v])$. Thus $p \Vdash_{\mathfrak{S}} (\forall x \in y)(\exists z)\neg(\forall v \in z)B[x, y, v])$ by Lemma 57. □

Lemma 63. For a strongly negative Δ_0 formula $B[x, y]$ which is prominent for $\mathfrak{S}$, $p \Vdash_{\mathfrak{S}} \exists y B[x, y]$ is, equivalently over IKP_0^-, a negatively Σ_1 formula.

Proof. $p \Vdash_{\mathfrak{S}} \exists y B[x, y]$ is, by definition, $\exists y(p \Vdash_{\mathfrak{S}} B[x, y])$, which is, by Lemma 59, equivalent to $\exists y, u(\mathrm{Tr}_{\mathfrak{S}}[p, u] \to B[x, y])$. □

6.6. Forcing the axioms of $\mathsf{IKP}^\sharp$ *and Markov's principle*

Lemma 64. Let $A[z,\vec{y}]$ be a prominent formula for $\mathfrak{S}$. Then IKP_0^- proves:

(i) $\Vdash_{\mathfrak{S}} \exists x A[z,\vec{x}]$ is equivalent to $\exists x A[z,\vec{x}]$,
(ii) $\forall z A[z,\vec{x}] \Vdash_{\mathfrak{S}} \forall z A[z,\vec{x}]$,

Proof. The first assertion follows from the definitions of $\Vdash_{\mathfrak{S}}$ and $|\!\!\Vdash_{\mathfrak{S}}$ and Lemma 59.

For the second assertion notice that $\forall z A[z,\vec{x}] \Vdash_{\mathfrak{S}} \forall z A[z,\vec{x}]$ is $\forall x_0(\forall z A[z,\vec{x}] \Vdash_{\mathfrak{S}} A[x_0,\vec{x}])$ which is by Lemma 59 equivalent to $\forall x_0(\forall z A[z,\vec{x}] \;|\!\!\Vdash_{\mathfrak{S}} A[x_0,\vec{x}])$. This is by Lemma 53 equivalent to $\forall x_0 \exists u((\forall z \in u)A[z,\vec{x}] \to A[x_0,\vec{x}])$ which is provable in IKP_0^- (for each x_0 let $u \supseteq \{x_0\}$). □

The next three lemmata correspond to lemma 5.8 in Avigad [2].

Lemma 65. Let $\exists x A$ be an instance of N-transitive superset, N-pairing, N-union or $\Delta_0^{\mathrm{s}-}$ N-separation. Then IKP_0^- proves $\Vdash_{\mathfrak{S}} \exists x A$ if A is prominent for $\mathfrak{S}$.

Proof. Let $\exists x B$ be the corresponding axiom of IKP_0^-. Then A is the same as B^N and B implies A. Thus $\Vdash_{\mathfrak{S}} \exists x A$ follows by the first assertion of Lemma 64. □

The next lemma states that also instances of Δ_0 collection$^\sharp$ and ($\Delta_0^{\mathrm{s}-}$-MP) are forced if we chose a suitable sequence $\mathfrak{S}$.

Lemma 66 ($\Vdash_{\mathfrak{S}}$ Δ_0 collection$^\sharp$ and $\Vdash_{\mathfrak{S}}$ ($\Delta_0^{\mathrm{s}-}$-MP))**.** If the formulae $(\forall x \in a)\neg(\forall y \in z)\neg A[x,y]$ and $\neg(\forall x \in y)B[x]$ are prominent for $\mathfrak{S}$ (therefore both A and B are strongly negative), then IKP_0^- proves

(i) $\Vdash_{\mathfrak{S}} (\forall x \in a)\exists y A[x,y] \to \exists z(\forall x \in a)\neg(\forall y \in z)\neg A[x,y]$ and
(ii) $\Vdash_{\mathfrak{S}} \neg\forall x B[x] \to \exists y \neg(\forall x \in y)B[x]$.

Proof. For the first assertion assume $p \Vdash_{\mathfrak{S}} (\forall x \in y)\exists z A[x,y]$. Therefore, by Lemma 61 (i), we have $(\forall x' \in y)(p \Vdash_{\mathfrak{S}} \exists z A[x',z])$ and this is $(\forall x' \in y)\exists x_0(p \Vdash_{\mathfrak{S}} A[x',x_0])$ by definition (x' and x_0 do not appear in $\mathrm{Tr}_{\mathfrak{S}}[p,u]$). By Lemma 59 this is equivalent to

$$(\forall x' \in y)\exists x_0(p \;|\!\!\Vdash_{\mathfrak{S}} A[x',x_0])$$

which is by definition $(\forall x' \in y)\exists x_0 \exists u(\mathrm{Tr}_{\mathfrak{S}}[p,u] \to A[x',x_0])$. If we set $w \supseteq \{u,x_0\}$ we get $(\forall x' \in y)\exists w(\exists x_0, u \in w)(\mathrm{Tr}_{\mathfrak{S}}[p,u] \to A[x',x_0])$ and therefore by Δ_0 collection

$$\exists v(\forall x' \in y)(\exists w \in v)(\exists x_0, u \in w)(\mathrm{Tr}_{\mathfrak{S}}[p,u] \to A[x',x_0]).$$

And if we let $v' \supseteq \cup v$ and $v'' \supseteq \cup\cup v$ then $w \subseteq v'$ and $u \subseteq v''$ for every $w \in v$ and $u \in w$. Therefore it follows by Lemma 54

$$\exists v', v''(\forall x' \in y)(\exists x_0 \in v')(\mathrm{Tr}_{\mathfrak{S}}[p,v''] \to A[x',x_0]).$$

By Lemma 33 (9) this implies

$$\exists v', v''(\forall x' \in y)(\mathrm{Tr}_{\mathfrak{S}}[p,v''] \to (\exists x_0 \in v')A[x',x_0])$$

and therefore by the same lemma (13)

$$\exists v', v''(\forall x' \in y)(\mathrm{Tr}_{\mathfrak{S}}[p,v''] \to \neg(\forall x_0 \in v')\neg A[x',x_0]).$$

Again by the same lemma (12), this is equivalent to

$$\exists v', v''(\mathrm{Tr}_{\mathfrak{S}}[p,v''] \to (\forall x' \in y)\neg(\forall x_0 \in v')\neg A[x',x_0]).$$

The latter is by definition and renaming of the bounded variables x' and x_0 the same as

$$\exists v'(p \;|\!\!\Vdash_{\mathfrak{S}} (\forall x \in y)\neg(\forall z \in v')\neg A[x,z])$$

and this is by Lemma 59 and the definition of $\Vdash_{\mathfrak{S}}$ equivalent to

$$p \Vdash_{\mathfrak{S}} \exists w(\forall x \in y)\neg(\forall z \in w)\neg A[x, z].$$

For the second assertion assume $p \Vdash_{\mathfrak{S}} \neg\forall z B[z, \vec{x}]$ which is by definition

$$\forall q(p \subseteq q \to ((q \Vdash_{\mathfrak{S}} \forall z B[z, \vec{x}]) \to (q \Vdash_{\mathfrak{S}} \bot))).$$

By Lemma 64 (ii), we have $p, \forall z B[z, \vec{x}] \Vdash_{\mathfrak{S}} \forall z B[z, \vec{x}]$ and therefore $p, \forall z B[z, \vec{x}] \Vdash_{\mathfrak{S}} \bot$. Since $\bot$ is atomic and since $\bot$ and $\neg\neg\bot$ are equivalent, this is by definition and Lemma 53 equivalent to

$$\exists u((\mathrm{Tr}_{\mathfrak{S}}[p, u] \wedge (\forall z \in u)B[z, \vec{x}]) \to \bot).$$

By Lemma 33 (1) this is equivalent to $\exists u(\mathrm{Tr}_{\mathfrak{S}}[p, u] \to ((\forall z \in u)B[z, \vec{x}] \to \bot))$. It follows (take $x_0 = u$) $\exists x_0 \exists u(\mathrm{Tr}_{\mathfrak{S}}[p, u] \to \neg(\forall z \in x_0)B[z, \vec{x}])$ which is nothing else than

$$\exists x_0(p \Vvdash_{\mathfrak{S}} \neg(\forall z \in x_0)B[z, \vec{x}]).$$

By Lemma 59 and the definition of $\Vdash_{\mathfrak{S}}$, this is equivalent to $p \Vdash_{\mathfrak{S}} \exists y\neg(\forall z \in y)B[z, \vec{x}])$. □

Lemma 67 ($\Vdash_{\mathfrak{S}}$ ($\mathcal{L}^{\mathrm{s}-}$-$\mathsf{Ind}$))**.** Let A be any strongly negative $\mathcal{L}$ formula. Then IKP^- proves

$$\Vdash_{\mathfrak{S}} \forall x((\forall y \in x)A[y] \to A[x]) \to \forall x A[x].$$

Analogous for IKP^-_ω and ($\mathcal{L}^{\mathrm{s}-}$-$\mathsf{Ind}_\omega$).

Proof. Assume $p \Vdash_{\mathfrak{S}} \forall x((\forall y \in x)A[y] \to A[x])$, which is by definition

$$\forall x_0 \forall q(p \subseteq q \to ((q \Vdash_{\mathfrak{S}} (\forall y \in x_0)A[y]) \to (q \Vdash_{\mathfrak{S}} A[x_0]))),$$

where x_0 does not occur in $\mathrm{Tr}_{\mathfrak{S}}[p, u]$. This implies

$$\forall x_0((p \Vdash_{\mathfrak{S}} (\forall y \in x_0)A[y]) \to (p \Vdash_{\mathfrak{S}} A[x_0]))$$

which is, by Lemma 61 (ii), equivalent to

$$\forall x_0((\forall x_1 \in x_0)(p \Vdash_{\mathfrak{S}} A[x_1]) \to (p \Vdash_{\mathfrak{S}} A[x_0]))$$

where x_1 does not occur in $\mathrm{Tr}_{\mathfrak{S}}[p, u]$. Applying $\in$-induction this leads us to $\forall x_0(p \Vdash_{\mathfrak{S}} A[x_0])$ which is by definition $p \Vdash_{\mathfrak{S}} \forall x A[x]$. All in all we have proved for an arbitrary p that

$$p \Vdash_{\mathfrak{S}} \forall x((\forall y \in x)A[y] \to A[x]) \text{ implies } p \Vdash_{\mathfrak{S}} \forall x A[x]$$

which is by Definition 56 what we want.

The proof that analogous results hold for IKP^-_ω and ($\mathcal{L}^{\mathrm{s}-}$-$\mathsf{Ind}_\omega$) is analogous, since by Lemma 61 (i) $p \Vdash_{\mathfrak{S}} (\forall x \in \omega)((\forall y \in x)A[y] \to A[x])$ implies $(\forall x \in \omega)(p \Vdash_{\mathfrak{S}} (\forall y \in x)A[y] \to A[x])$ and by Lemma 61 (ii) $(\forall x \in \omega)(p \Vdash_{\mathfrak{S}} A[x])$ implies $p \Vdash_{\mathfrak{S}} (\forall x \in \omega)A[x]$. □

The next lemma is proved very similarly as in the previous one with Lemmata 62 and 63.

Lemma 68 ($\Vdash_{\mathfrak{S}}$ $\mathsf{w}\Sigma_1$-Ind)**.** Let A be any weak Σ_1 formula such that the largest Δ_0 subformula is prominent for $\mathfrak{S}$. Then $\mathsf{IKP}^-_0 + (\mathsf{n}\Sigma_1\text{-}\mathsf{Ind})$ proves

$$\Vdash_{\mathfrak{S}} \forall x((\forall y \in x)A[y] \to A[x]) \to \forall x A[x].$$

The analogous assertion for $\mathsf{IKP}^-_0 + (\mathsf{n}\Sigma_1\text{-}\mathsf{Ind}_\omega)$ and $\Vdash_{\mathfrak{S}}$ ($\mathsf{w}\Sigma_1$-Ind_ω) also holds.

The next lemma corresponds to Theorem 5.15 in Avigad [2]. We refer in its formulation to Lemma 37, which tells us that for every Δ_0 formula A there exists a strongly negative Δ_0 formula intuitionistically equivalent to A^N. We may confuse them in the following discussions.

Lemma 69 ($\Vdash_{\mathfrak{S}}$ N-infinity)**.** Let B and $C[y,z]$ be strongly negative Δ_0 formulae such that

$$B \leftrightarrow \mathrm{Ind}[\omega]^N \wedge (\forall x \subseteq \omega)(\mathrm{Ind}[x] \to \omega \subseteq x)^N \text{ and } C[y,z] \leftrightarrow \mathrm{famfun}[\omega,y,z]^N$$

are intuitionistically valid. If B and $C[y,z]$ are prominent for $\mathfrak{S}$, then IKP_0^- proves $\Vdash_{\mathfrak{S}}$ N-infinity.

Proof. Since N-infinity is available in IKP_0^-, this theory proves B and $(\forall y)(\exists z)C[y,z]$. By Lemma 64 (i) and by the definition of $\Vdash_{\mathfrak{S}}$, this implies both $\Vdash_{\mathfrak{S}} B$ and $\Vdash_{\mathfrak{S}} (\forall y)(\exists z)C[y,z]$ and therefore by definition $\Vdash_{\mathfrak{S}} B \wedge (\forall y)(\exists z)C[y,z]$. □

Lemma 70 ($\Vdash_{\mathfrak{S}}$ $\mathsf{sn}\Pi_2$-Ref)**.** Let $A[x,\vec{u}]$ be any very negatively Σ_1 formula such that the largest Δ_0 subformulae of $A[x,\vec{u}]$ are prominent for $\mathfrak{S}$. Then $\mathsf{IKP}_0^- + (\mathsf{n}\Pi_2\text{-}\mathsf{Ref})$ proves

$$\Vdash_{\mathfrak{S}} (\forall x A[x,\vec{u}]) \to (\exists v)(\vec{u} \in v \wedge \mathrm{Tran}[v] \wedge \mathsf{Ad}(v) \wedge (\forall x \in v)A^{(v)}[x,\vec{u}]).$$

Proof. Let $A[x,\vec{u}]$ be very negatively Σ_1 formula $\exists z B[z,x,\vec{u}]$ where B is a strongly negative Δ_0 formula. Assume $p \Vdash_{\mathfrak{S}} (\forall x \exists z B[z,x,\vec{u}])$, that is by definition $\forall x \exists z (p \Vdash_{\mathfrak{S}} B[z,x,\vec{u}])$. Since Lemma 59 implies $\forall x \exists z (p \Vvdash_{\mathfrak{S}} B[z,x,\vec{u}])$, now $(\mathsf{n}\Pi_2\text{-}\mathsf{Ref})$ provides us some v with $\vec{u} \in v$, $\mathrm{Tran}[v]$ and $\mathsf{Ad}(v)$ such that

$$(\forall x \in v)(\exists z \in v)(p \Vvdash_{\mathfrak{S}} B[z,x,\vec{u}])^{(v)},$$

where $(p \Vvdash_{\mathfrak{S}} B[z,x,\vec{u}])^{(v)}$ implies $p \Vvdash_{\mathfrak{S}} B^{(v)}[z,x,\vec{u}]$ by definition. By Lemma 59 and the definition of $\Vdash_{\mathfrak{S}}$, we have $p \Vdash_{\mathfrak{S}} (\forall x \exists z B[z,x,\vec{u}])^{(v)}$, namely $p \Vdash_{\mathfrak{S}} (\forall x A[x,\vec{u}])^{(v)}$. □

6.7. Interpretability result

Definition 71 ($\mathcal{D}_{res}$)**.** The set $\mathcal{D}_{res}$ of $\mathcal{L}$ formulae is inductively defined as follows:

(i) every strongly negative Δ_0 formula is in $\mathcal{D}_{res}$;
(ii) if A and B are in $\mathcal{D}_{res}$, then also $A \wedge B$ and $A \vee B$ are in $\mathcal{D}_{res}$;
(iii) if A is in $\mathcal{D}_{res}$, then also $\forall x A$ and $\exists x A$ are in $\mathcal{D}_{res}$.

Lemma 72. If A is in $\mathcal{D}_{res}$ and all maximal Δ_0 subformulae of A are prominent for $\mathfrak{S}$, then the theory IKP_0^- proves that $\Vdash_{\mathfrak{S}} A$ is equivalent to A.

Proof. The proof is on the complexity of A. If A is a strongly negative Δ_0 formula, the assertion follows from Lemma 64 (i). If A is more complex, it follows directly from the definition of $\Vdash_{\mathfrak{S}}$ and the induction hypothesis. □

Lemma 73. Let $\mathcal{A} \subseteq \mathcal{D}_{res}$ be a set of $\mathcal{L}$ formulae. If the $\mathcal{L}$ formula A is provable in $\mathsf{IKP}_0^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP}) + \mathcal{A}$, then there exists a finite sequence $\mathfrak{T}$ of strongly negative Δ_0 formulae such that for every finite sequence $\mathfrak{S}$ of strongly negative Δ_0 formulae which contains at least all formulae of $\mathfrak{T}$ the theory $\mathsf{IKP}_0^- + \mathcal{A}$ proves $\Vdash_{\mathfrak{S}}$ A.

The analogous assertions hold if $(\mathsf{w}\Sigma_1\text{-}\mathsf{Ind})$, $(\mathsf{w}\Sigma_1\text{-}\mathsf{Ind}_\omega)$ or $(\mathsf{w}\Pi_2\text{-}\mathsf{Ref})$ is added to $\mathsf{IKP}_0^\sharp$ and at the same time $(\mathsf{n}\Sigma_1\text{-}\mathsf{Ind})$, $(\mathsf{n}\Sigma_1\text{-}\mathsf{Ind}_\omega)$ or $(\mathsf{n}\Pi_2\text{-}\mathsf{Ref})$, respectively, is added to IKP_0^-. Furthermore, all the assertions, with the subscript 0 replaced by ω or just omitted, also hold.

Proof. Assume that A is provable in $\mathsf{IKP}_0^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP}) + \mathcal{A}$ and let $B_0, ..., B_n, A$ be a proof of A in a Hilbert-style system. Let $\mathfrak{T}$ be a finite sequence which contains enough formulae such that we can apply the Lemmata 65-69 to all instances of axioms of $\mathsf{IKP}_0^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP})$ occurring in $B_0, ..., B_n, A$, i.e., $\mathfrak{T}$ contains enough formulae such that IKP^- proves $\Vdash_{\mathfrak{S}} C$ for all axioms C of $\mathsf{IKP}_0^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP})$ occurring in $B_0, ..., B_n, A$. By Lemma 57 we can conclude that $\mathsf{IKP}_0^- + \mathcal{A}$ proves $\Vdash_{\mathfrak{S}} A$. □

Remark 74. Since Lemma 67, the forcibility of $\in$-induction, holds without any condition on $\mathfrak{S}$, if we replace the other axiom schemata by their finite number of instances, $\mathfrak{T}$ in the lemma does not depend on the formula A nor on proofs. By Remark 46, this is enough for the interpretability of variant of KP in those of WEST.

Theorem 75. Let $\mathcal{A} \subseteq \mathcal{D}_{res}$ be a set of $\mathcal{L}$ formulae. If the $\mathcal{L}$ formula B is in $\mathcal{D}_{res}$ and provable in $\mathsf{IKP}_0^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP}) + \mathcal{A}$, then it is also provable in $\mathsf{IKP}_0^- + \mathcal{A}$.

The analogous assertions hold if $(\mathsf{w}\Sigma_1\text{-}\mathsf{Ind})$, $(\mathsf{w}\Sigma_1\text{-}\mathsf{Ind}_\omega)$ or $(\mathsf{w}\Pi_2\text{-}\mathsf{Ref})$ is added to $\mathsf{IKP}_0^\sharp$ and at the same time $(\mathsf{n}\Sigma_1\text{-}\mathsf{Ind})$, $(\mathsf{n}\Sigma_1\text{-}\mathsf{Ind}_\omega)$ or $(\mathsf{n}\Pi_2\text{-}\mathsf{Ref})$, respectively, is added to IKP_0^-. Furthermore, all the assertions, with the subscript 0 replaced by ω or just omitted, also hold.

Proof. Let B be in $\mathcal{D}_{res}$ and provable in $\mathsf{IKP}_0^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP}) + \mathcal{A}$. By Lemma 73 there is a finite sequence $\mathfrak{T}$ of negative Δ_0 formulae such that $\mathsf{IKP}_0^- + \mathcal{A}$ proves $\Vdash_{\mathfrak{S}} B$ for every finite super-sequence $\mathfrak{S}$ of $\mathfrak{T}$. Let $\mathfrak{S}$ be such a sequence which contains, besides the formulae of $\mathfrak{T}$, enough formulae so that all maximal Δ_0 subformulae of B are prominent for $\mathfrak{S}$. We can conclude by Lemma 72 that $\mathsf{IKP}_0^- + \mathcal{A}$ proves B. ☐

Remark 76. Similarly to Remark 74, since π from Remark 23 yields a universal Π_1 formula $\forall u \neg\pi[e, u, x, y, b]$, if B in this theorem is the negative interpretation of a Π_1 formula then $\mathfrak{S} = \mathfrak{T} \cup \{\neg\pi^N\}$ seems enough. However, we cannot prove the equivalence between $B[x, y, \vec{b}]^N$ and $\forall u \neg\pi[\ulcorner B \urcorner, u, x, y, \langle\vec{b}\rangle]^N$ because of the absence of negative interpretation of Δ_0 collection. Even if we postpone obtaining the equivalence until after going through the next interpretation (realisability), it does not seem to work.

7. Interpreting IKP^- in OST^-: Realisability Interpretation

In this section, we take the last step towards the combined interpretation of KP in WEST, namely interpret the intuitionistic set theory IKP^- and variants in the applicative system WEST and corresponding variants. The method is so-called realisability interpretation, which can be seen as a formalisation of Brouwer-Heyting-Kolmogorov interpretation of intuitionistic logic. With this interpretation, we can convert operational replacement, the axiom for replacement operator, into collection schema, at the cost of classical logic.

While there are many variations of realisability, here we will take one of the most naïve one, in which an atomic formula is realised by any object whenever it is true. In this kind of realisability, for negative formulae, being realisable and being true are equivalent (Lemma 80). In this sense, the realisability interpretation does not change the meanings of atomic formulae and of the negative fragment built up by $\bot$, $\wedge$, $\to$ and $\forall$, but it changes those of $\vee$ and $\exists$. Since classical $\vee$ and classical $\exists$ can be defined by negative connectives $\bot$, $\wedge$, $\to$ and $\forall$, we may consider the whole classical logic is formulated in the negative fragment and, from this perspective, our realisability interpretation augments the original language with extra connectives for the constructive disjunction and constructive existential quantifier, if the interpreting system is based on classical logic. In these terms, the operational replacement in OST^- or WEST implies the collection schema formulated with constructive existential quantifier, which is of course different from the collection schema in KP, that is, formulated with classical existential quantifier. We already saw how to recover the original form of collection schema in the last two sections.

7.1. Feferman realisability in applicative systems

Similar as for instance in Feferman [12, 4.3] and [14, IV] we define realisation of formulae of $\mathcal{L}$.

Definition 77 (Realising relation $\mathfrak{r}$)**.** For each $\mathcal{L}$ formula A in which f does not occur, the $\mathcal{L}^\circ$ formula $f \mathrel{\mathfrak{r}} A$, *f realises A* or f is a *realiser* of A, is inductively defined as follows:

(i) if A is atomic, then $f \mathrel{\mathfrak{r}} A$ is the formula $A \wedge f = f$;
(ii) if A is $B \wedge C$, then $f \mathrel{\mathfrak{r}} A$ is the formula $(\mathbf{p}_0(f) \mathrel{\mathfrak{r}} B) \wedge (\mathbf{p}_1(f) \mathrel{\mathfrak{r}} C)$;
(iii) if A is $B \vee C$, then $f \mathrel{\mathfrak{r}} A$ is the formula $(\mathbf{p}_0(f) = \overline{0} \wedge \mathbf{p}_1(f) \mathrel{\mathfrak{r}} B) \vee (\mathbf{p}_0(f) = \overline{1} \wedge \mathbf{p}_1(f) \mathrel{\mathfrak{r}} C))$;
(iv) if A is $B \to C$, then $f \mathrel{\mathfrak{r}} A$ is the formula $\forall g(g \mathrel{\mathfrak{r}} B \to f(g) \mathrel{\mathfrak{r}} C) \wedge f = f$;

(v) if A is $\exists x B[x]$, then $f \mathrel{\mathfrak{r}} A$ is the formula $\mathbf{p}_0(f){\downarrow} \wedge \mathbf{p}_1(f) \mathrel{\mathfrak{r}} B[\mathbf{p}_0(f)]$.
(vi) if A is $\forall x B[x]$, then $f \mathrel{\mathfrak{r}} A$ is the formula $\forall x (f(x) \mathrel{\mathfrak{r}} B[x])$.

We say that the formula $A[\vec{x}]$ with at most the free variables $\vec{x}$ is *realisable in* some theory $\mathcal{T}$, if $\mathcal{T}$ proves the formula $\exists f \forall \vec{x}(f(\vec{x}) \mathrel{\mathfrak{r}} A[\vec{x}])$. We just say that a formula is *realisable* if it is realisable in WEST_0.

The aim of clause "$f = f$" is that, if a term t is substituted, $t{\downarrow}$ follows. We will use the notation $f \mathrel{\mathfrak{r}} A[t]$ also if t is an $\mathcal{L}^\circ$ term but not an $\mathcal{L}$ term. When we do so, we mean by $f \mathrel{\mathfrak{r}} A[t]$ the formula $(f \mathrel{\mathfrak{r}} A[v])[t/v]$.

To show the closure of realisability under the inference rules of intuitionistic logic we need the following lemma which can be proved by a straightforward induction on the formula A.

Lemma 78. If A is an $\mathcal{L}$ formula then the free variables of $f \mathrel{\mathfrak{r}} A$ are the variable f and the free variables of A. Also for any $\mathcal{L}^\circ$ term t, WEST_0 proves that $t \mathrel{\mathfrak{r}} A$ implies $t{\downarrow}$.

We introduce now for each negative $\mathcal{L}$ formula A a term which realises A if A holds.

Definition 79 (Canonical term $r_{A,\vec{x}}$)**.** We assign to each finite sequence $\vec{x} = x_0, ..., x_n$ of variables and each negative $\mathcal{L}$ formula $A[\vec{x}]$, in which at most variables $\vec{x}$ occur freely, an $\mathcal{L}^\circ$ term $r_{A,\vec{x}}$ inductively defined as follows:

(i) if $A[\vec{x}]$ is atomic then $r_{A,\vec{x}}$ is the term $\lambda \vec{x}.\bar{0}$;
(ii) if $A[\vec{x}]$ is the formula $B[\vec{x}] \wedge C[\vec{x}]$ then $r_{A,\vec{x}}$ is the term $\lambda \vec{x}.\mathbf{p}(r_{B,\vec{x}}(\vec{x}), r_{C,\vec{x}}(\vec{x}))$;
(iii) if $A[\vec{x}]$ is the formula $B[\vec{x}] \to C[\vec{x}]$ then $r_{A,\vec{x}}$ is the term $\lambda \vec{x}, g.r_{C,\vec{x}}(\vec{x})$, where g is a variable not occurring in $r_{C,\vec{x}}$;
(iv) if $A[\vec{x}]$ is the formula $\forall y B[x_0, ..., x_i, y, x_{i+1}, ..., x_n]$ then $r_{A,\vec{x}}$ is the term

$$\lambda x_0, ..., x_n, y.r_{B,x_0,...,x_i,y,x_{i+1},...,x_n}(x_0, ..., x_i, y, x_{i+1}, ..., x_n).$$

Lemma 80. For each negative $\mathcal{L}$ formula $A[\vec{x}]$ in which at most the variables $\vec{x}$ occur freely:

(i) $r_{A,\vec{x}}$ is a closed term;
(ii) WEST_0 proves that if there is a realiser of $A[\vec{x}]$, then $r_{A,\vec{x}}(\vec{x})$ is a realiser of $A[\vec{x}]$;
(iii) WEST_0 proves $\forall \vec{x}(A[\vec{x}] \leftrightarrow r_{A,\vec{x}}(\vec{x}) \mathrel{\mathfrak{r}} A[\vec{x}])$.

Proof. The proof is by induction on the length of the formula A. For the case with $\to$ in the third assertion we need the second one. □

The next lemma follows by the second and third assertions of the lemma above.

Lemma 81. For any negative formula A of $\mathcal{L}$, WEST_0 proves that $A \leftrightarrow \exists f(f \mathrel{\mathfrak{r}} A)$.

Lemma 82. All axioms of intuitionistic logic are realisable and realisability (in any theory which contains WEST_0) is closed under its rules of inference.

Proof. We can assume that the equality axioms for $=$ are formulated using only negative formulae. Since these axioms are available in WEST_0, they are therefore realisable by Lemma 81.

Since λ-abstraction is available, the proof of the realisability of all the propositional axioms of a Hilbert-style calculus is straightforward in view of Lemma 16 and Propositions 10 and 13. For instance

$$\lambda f, g, h.\text{iszero}(\mathbf{p}_0(h), f(\mathbf{p}_1(h)), g(\mathbf{p}_1(h))) \text{ realises } (A \to C) \to ((B \to C) \to ((A \vee B) \to C)).$$

Further $\lambda f.f(y)$ realises $\forall x A[x] \to A[y]$ and $\lambda f.\mathbf{p}(y, f)$ realises $A[y] \to \exists x A[x]$ if y is not bounded in A. For proving the closure of realisability under the quantifier rules we assume in both cases

$$\forall x, \vec{y}(f(x, \vec{y}) \mathrel{\mathfrak{r}} (A[x, \vec{y}] \to B[x, \vec{y}])) \tag{5}$$

that is, we have for all $x, \vec{y}$

$$\forall g((g \mathrel{\mathfrak{r}} A[x, \vec{y}]) \to (f(x, \vec{y}, g) \mathrel{\mathfrak{r}} B[x, \vec{y}])). \tag{6}$$

For the ∀-rule we assume in addition that x does not occur freely in $A[x,\vec{y}]$, for which we write $A[\vec{y}]$, and get by Lemma 78

$$\forall g((g \mathrel{\mathfrak{r}} A[\vec{y}]) \to \forall x(f(x,\vec{y},g) \mathrel{\mathfrak{r}} B[x,\vec{y}]))$$

because the rule is also available in the interpreting theory. Therefore

$$\forall \vec{y}(\lambda g, x.f(x,\vec{y},g) \mathrel{\mathfrak{r}} (A[\vec{y}] \to \forall x B[x,\vec{y}])).$$

For the ∃-rule we deduce, by substituting $\mathbf{p}_0(g)$ into x and $\mathbf{p}_1(g)$ into g in (6), for all $\vec{y}$

$$\forall g((\mathbf{p}_1(g) \mathrel{\mathfrak{r}} A[\mathbf{p}_0(g),\vec{y}]) \to (f(\mathbf{p}_0(g),\vec{y},\mathbf{p}_1(g)) \mathrel{\mathfrak{r}} B[\mathbf{p}_0(g),\vec{y}]))$$

and assume in addition that x does not occur freely in $B[x,\vec{y}]$, for which we write $B[\vec{y}]$. Therefore

$$\forall \vec{y}(\lambda g.f(\mathbf{p}_0(g),\vec{y},\mathbf{p}_1(g)) \mathrel{\mathfrak{r}} (\exists x A[x,\vec{y}] \to B[\vec{y}])).$$

That realisability is closed under modus ponens is obvious. □

7.2. Realising the axioms of IKP_0^-

Now we are ready for proving the realisability of all the axioms of IKP^- in WEST_0.

Lemma 83 (transitive superset, pairing, union)**.** The axioms of transitive superset, pairing and union are realisable.

Proof. Let $B[x,a]$ be the negative formula $(a \subseteq x \wedge (\forall y \in x)(\forall z \in y)(z \in x))$. Then we have by Lemma 80 that $r_{B,x,a}(\mathbb{T}(a),a) \mathrel{\mathfrak{r}} B[\mathbb{T}(a),a]$. If we set

$$f = \lambda a.\mathbf{p}(\mathbb{T}(a), r_{B,x,a}(\mathbb{T}(a),a)),$$

we have $\mathbf{p}_1(f(a)) \mathrel{\mathfrak{r}} B[\mathbf{p}_0(f(a)),a]$ for every a. Thus $f(a)$ realises the formula $\exists x(a \subseteq x \wedge \mathrm{Tran}[x])$.

Let $C[x,a,b]$ be the negative formula $(a \in x \wedge b \in x)$. Then we have by Lemma 80 that $r_{C,x,a,b}(\mathbb{D}(a,b),a,b) \mathrel{\mathfrak{r}} C[\mathbb{D}(a,b),a,b])$. Similar as in the last proof, $\lambda ab.\mathbf{p}(\mathbb{D}(a,b), r_{C,x,a,b}(\mathbb{D}(a,b),a,b))$ applied to a,b realises $\exists x C[x,a,b]$.

Let $D[x,a]$ be the negative formula $(\forall y \in a)(\forall z \in y)(z \in x)$. Lemma 80 implies $r_{D,x,a}(\mathbb{U}(a),a) \mathrel{\mathfrak{r}} B[\mathbb{U}(a),a]$. Similar as in the last proofs, $\lambda a.\mathbf{p}(\mathbb{U}(a), r_{D,x,a}(\mathbb{U}(a),a))$ applied to a realises $\exists x D[x,a]$. □

Lemma 84 (Δ_0^- separation)**.** For all negative Δ_0 formulae $A[y,a,\vec{v}]$ of $\mathcal{L}$ in which x does not occur and with at most the variables $y,a,\vec{v}$ free, the following formula is realisable:

$$\exists x((\forall y \in x)(y \in a \wedge A[y,a,\vec{v}]) \wedge (\forall y \in a)(A[y,a,\vec{v}] \to y \in x)).$$

Proof. Let $B[x,a,\vec{v}]$ be the negative formula

$$(\forall y \in x)(y \in a \wedge A[y,a,\vec{v}]) \wedge (\forall y \in a)(A[y,a,\vec{v}] \to y \in x).$$

Let s_A be the term defined in Lemma 12. Then Lemma 80 implies

$$r_{B,x,a,\vec{v}}(s_A(a,a,\vec{v}),a,\vec{v}) \mathrel{\mathfrak{r}} B[s_A(a,a,\vec{v}),a,\vec{v}].$$

As in the last lemma, $\lambda a,\vec{v}.\mathbf{p}(s_A(a,a,\vec{v}), r_{B,x,a,\vec{v}}(s_A(a,a,\vec{v}),a,\vec{v}))$ applied to $a,\vec{v}$ is the realiser we are searching for. □

Here the restriction to negative Δ_0 formulae is essential. If a Δ_0 formula A is not negative, we cannot take the canonical realiser and therefore we cannot provide, for given $y \in \{x \in a : A[x]\}$, the witness of $A[y]$. Tupailo [55, 56] interprets theories with Δ_0 separation (or analogous axiom schema) in applicative systems, by using a different realisability notion. However, for this realisability, a kind of exponentiation is necessary, which is not available in our frameworks WEST or OST^-.

Let us turn to the collection schema. The aim of our use of realisability interpretation is to convert the operational replacement into the constructive collection. Actually, operational replacement yields the realisability of the collection schema for all formulae, not only for Δ_0 formulae.

Lemma 85 (Collection)**.** For all formulae $A[x,y,a,\vec{v}]$ of $\mathcal{L}$ in which z does not occur and with at most the variables $x,y,a,\vec{v}$ free, the following formula is realisable:

$$(\forall x\in a)\exists y A[x,y,a,\vec{v}] \to \exists z(\forall x\in a)(\exists y\in z)A[x,y,a,\vec{v}].$$

PROOF. Assume $f\ \mathfrak{r}\ (\forall x\in a)\exists y A[x,y,a,\vec{v}]$. That is to say

$$\forall x,g(\ (g\ \mathfrak{r}\ (x\in a))\to(\mathbf{p}_1(f(x,g))\ \mathfrak{r}\ A[x,\mathbf{p}_0(f(x,g)),a,\vec{v}])\ ).$$

If t is the term $\lambda x,g.\mathbf{p}(\mathbf{p}_0(f(x,g)),\mathbf{p}(\overline{0},\mathbf{p}_1(f(x,g))))$ and $B[x,y,z,a,\vec{v}]$ is the formula $y\in z\wedge A[x,y,a,\vec{v}]$ we get

$$\forall x,g(x\in a\to(\mathbf{p}_1(t(x,g))\ \mathfrak{r}\ B[x,\mathbf{p}_0(t(x,g)),\mathbb{R}(\lambda x.\mathbf{p}_0(f(x,g)),a),a,\vec{v}])),$$

since $\mathbf{p}_0(f(x,g))\in\mathbb{R}(\lambda x.\mathbf{p}_0(f(x,g)),a)$ if $x\in a$, and $\overline{0}$ realises $y\in z$ if it holds and z does not occur in $A[x,y,a,\vec{v}]$. By substituting $\overline{0}$ to g, we have also

$$\forall x,h(\ (h\ \mathfrak{r}\ (x\in a))\to(\mathbf{p}_1(t(x,\overline{0}))\ \mathfrak{r}\ B[x,\mathbf{p}_0(t(x,\overline{0})),\mathbb{R}(\lambda x.\mathbf{p}_0(f(x,\overline{0})),a),a,\vec{v}])\ ).$$

It follows that

$$\mathbf{p}(\mathbb{R}(\lambda x.\mathbf{p}_0(f(x,\overline{0})),a),\lambda x,h.t(x,\overline{0}))\ \mathfrak{r}\ \exists z(\forall x{\in}a)\exists y B[x,y,z,a,\vec{v}].$$

We can conclude that if $s:=\lambda a,\vec{v},f.\mathbf{p}(\mathbb{R}(\lambda x.\mathbf{p}_0(f(x,\overline{0})),a),\lambda x,h.t(x,\overline{0}))$ we have for all $a,\vec{v}$

$$s(a,\vec{v})\ \mathfrak{r}\ ((\forall x\in a)\exists y A[x,y,a,\vec{v}]\to\exists z(\forall x\in a)(\exists y\in z)A[x,y,a,\vec{v}])$$

since $\exists z(\forall x\in a)\exists y B[x,y,a,\vec{v}]$ is nothing else than $\exists z(\forall x\in a)(\exists y\in z)A[x,y,a,\vec{v}]$. □

Lemma 86 (N-infinity)**.** The following formula is realisable:

$$\mathrm{Ind}[\omega]^N\wedge(\forall x\subseteq\omega)(\mathrm{Ind}[x]\to\omega\subseteq x)^N\wedge(\forall y)(\exists z)\mathrm{famfun}[\omega,y,z]^N.$$

PROOF. Lemma 80 implies the realisability of $\mathrm{Ind}[\omega]^N\wedge(\forall x\subseteq\omega)(\mathrm{Ind}[x]\to\omega\subseteq x)^N$. Let

$$B[y,z]\ \equiv\ \mathrm{famfun}[\omega,y,z]^N.$$

Then Lemma 80 implies $r_{B,y,z}(y,\mathbb{K}(y))\ \mathfrak{r}\ B[y,\mathbb{K}(y)]$. Similar as in the proofs of the Lemmata 83 to 84, $\mathbf{p}(\mathbb{K}(y),r_{B,y,z}(y,\mathbb{K}(y)))$ is a realiser of $\exists z B[y,z]$. □

Lemma 87 (Reflection)**.** For any $\mathcal{L}$ formula $B[x,y,\vec{u}]$ in which v does not occur freely, the formula

$$\forall x\exists y B[x,y,\vec{u}]\to\exists v(\vec{u}\in v\wedge\mathrm{Tran}[v]\wedge\mathsf{Ad}(v)\wedge(\forall x\in v)(\exists y\in v)B[\vec{u}])$$

is realisable in $\mathsf{WEST}_0+(\mathbb{M})$.

PROOF. Assume $f\ \mathfrak{r}\ (\forall x)(\exists y)B[x,y,\vec{u}]$, that is, for any x, $f(x)\!\downarrow$ and $\mathbf{p}_1(f(x))\ \mathfrak{r}\ B[x,\mathbf{p}_0(f(x)),\vec{u}]$. Now $\lambda x.\mathbf{p}_0(f(x)):\mathbf{V}\to\mathbf{V}$, by the axiom $(\mathbb{M})$, we can take $v=\mathbb{M}(\lambda x.\mathbf{p}_0(f(x)),\langle\vec{u}\rangle)$. Then $\vec{u}\in v$, $\mathrm{Tran}[v]$, $\mathsf{Ad}(v)$ and $\lambda x.\mathbf{p}_0(f(x)):v\to v$. For any $x\in v$, $\mathbf{p}_0(f(x))\in v$ and so

$$\mathbf{p}(\overline{0},\mathbf{p}_1(f(x)))\ \mathfrak{r}\ (\mathbf{p}_0(f(x))\in v\wedge B[x,\mathbf{p}_0(f(x)),\vec{u}])$$

which means $\mathbf{p}(\mathbf{p}_0(f(x)),\mathbf{p}(\overline{0},\mathbf{p}_1(f(x))))\ \mathfrak{r}\ (\exists y\in v)B[x,y,\vec{u}]$. Thus

$$g[f]=\lambda x,d.\mathbf{p}(\mathbf{p}_0(f(x)),\mathbf{p}(\overline{0},\mathbf{p}_1(f(x))))$$

realises $(\forall x\in v)(\exists y\in v)B[x,y,\vec{u}]$, and so $\mathbf{p}(r_{D,\vec{u},v}(\vec{u},v),g[f])$ realises $D[\vec{u},v]\wedge(\forall x\in v)(\exists y\in v)B[x,y,\vec{u}]$, where $D[\vec{u},w]$ is the negative formula $\vec{u}\in w\wedge\mathrm{Tran}[w]\wedge\mathsf{Ad}(w)$. Therefore

$$\begin{aligned}&\lambda f.\mathbf{p}(v,\mathbf{p}(r_{D,\vec{u},w}(\vec{u},v),g[f]))\\&\quad=\lambda f.\mathbf{p}(\mathbb{M}(\lambda x.\mathbf{p}_0(f(x)),\langle\vec{u}\rangle),\mathbf{p}(r_{D,\vec{u},w}(\vec{u},\mathbb{M}(\lambda x.\mathbf{p}_0(f(x)),\langle\vec{u}\rangle)),\lambda x,d.\mathbf{p}(\mathbf{p}_0(f(x)),\mathbf{p}(\overline{0},\mathbf{p}_1(f(x))))))\end{aligned}$$

realises $\forall x\exists y B[x,y,\vec{u}]\to(\exists v)(\vec{u}\in v\wedge\mathrm{Tran}[v]\wedge\mathsf{Ad}(v)\wedge(\forall x\in v)(\exists y\in v)B[x,y,\vec{u}])$. □

Note that, if $B[x,y,\vec{u}]$ is a Δ_0 formula, $(\forall x\in v)(\exists y\in v)B[x,y,\vec{u}]$ is equivalent to $(\forall x\exists y B[x,y,\vec{u}])^{(v)}$. Thus, in particular, each instance of $(\mathsf{n}\Pi_2\text{-}\mathsf{Ref})$ is realisable in $\mathsf{WEST}_0+(\mathbb{M})$.

Note also that we do not need $(\mathcal{AD})$ in this result.

Lemma 88 ($\in$-induction)**.** For arbitrary $\mathcal{L}$ formulae $A[x, \vec{v}]$ with at most the variables $x, \vec{v}$ free the formula

$$\forall x((\forall y \in x)A[y, \vec{v}] \to A[x, \vec{v}]) \to \forall x A[x, \vec{v}]$$

is realisable in WEST.

Proof. Let t be the term $\lambda h, g, x.g(x, \lambda y, f.h(g, y))$ and s the term $\mathbf{fix}(t)$. From the recursion theorem it follows $s\downarrow$ and $s(g) \simeq \lambda x.g(x, \lambda y, f.s(g, y))$. Fix a realiser g of the antecedent $\forall x((\forall y \in x)A[y, \vec{v}] \to A[x, \vec{v}])$, i.e.,

$$\forall x(\forall h((h \mathrel{\mathfrak{r}} \forall y(y \in x \to A[y, \vec{v}])) \to (g(x, h) \mathrel{\mathfrak{r}} A[x, \vec{v}]))).$$

In the following we prove by $\in$-induction that $\forall x(s(g, x) \mathrel{\mathfrak{r}} A[x, \vec{v}])$. For that purpose we fix an x and assume $(\forall y \in x)(s(g, y) \mathrel{\mathfrak{r}} A[y, \vec{v}])$ which is equivalent to

$$\forall y(\forall f(y \in x \to ((\lambda f.s(g, y))(f) \mathrel{\mathfrak{r}} A[y, \vec{v}]))).$$

The latter means $\forall y(\lambda f.s(g, y) \mathrel{\mathfrak{r}} (y \in x \to A[y, \vec{v}]))$ and it follows

$$\lambda y, f.s(g, y) \mathrel{\mathfrak{r}} \forall y(y \in x \to A[y, \vec{v}]).$$

By the assumption about g we get therefore $g(x, \lambda y, f.s(g, y)) \mathrel{\mathfrak{r}} A[x, \vec{v}]$, and we can conclude $s(g, x) \mathrel{\mathfrak{r}} A[x, \vec{v}]$ since $s(g, x) \simeq g(x, \lambda y, f.s(g, y))$. We have proved now that $s(g, x) \mathrel{\mathfrak{r}} A[x, \vec{v}]$ follows from $(\forall y \in x)(s(g, y) \mathrel{\mathfrak{r}} A[y, \vec{v}])$, therefore the former holds by $\in$-induction for every x. All in all we get that for all $\vec{v}$ the term $(\lambda \vec{v}.s)(\vec{v})$ is the realiser we are searching for. ☐

Similarly we can prove the following.

Lemma 89 ($\in$-induction$_\omega$)**.** For arbitrary $\mathcal{L}$ formulae $A[x, \vec{v}]$ with at most variables the $x, \vec{v}$ free the formula

$$(\forall x \in \omega)((\forall y \in x)A[y, \vec{v}] \to A[x, \vec{v}]) \to (\forall x \in \omega)A[x, \vec{v}]$$

is realisable in WEST_ω.

Lemma 90 (Negatively Σ_1 induction)**.** For every negative Δ_0 formula $B[x, z, \vec{v}]$ and with at most the variables $x, z, \vec{v}$ free. Then

$$\forall x((\forall y \in x)\exists z B[y, z, \vec{v}] \to \exists z B[x, z, \vec{v}]) \to \forall x \exists z B[x, z, \vec{v}]$$

is realisable in $\mathsf{WEST}_0 + (\mathsf{opInd})$.

Proof. Let $t[\vec{v}]$ be the term

$$\lambda h, g, x.\mathbf{p}(\mathbf{p}_0(g(x, \lambda y, f.h(g, y))), r_{B,x,z,\vec{v}}(x, \mathbf{p}_0(g(x, \lambda y, f.h(g, y))), \vec{v}))$$

and s the term $\mathbf{fix}(t)$. From the recursion theorem it follows $s\downarrow$ and

$$s(g) \simeq \lambda x.\mathbf{p}(\mathbf{p}_0(g(x, \lambda y, f.s(g, y))), r_{B,x,z,\vec{v}}(x, \mathbf{p}_0(g(x, \lambda y, f.s(g, y))), \vec{v})).$$

Now we fix some realiser g of the antecedent $\forall x((\forall y \in x)\exists z B[y, z, \vec{v}] \to \exists z B[x, z, \vec{v}])$. This is by definition

$$\forall x(\forall h((h \mathrel{\mathfrak{r}} \forall y(y \in x \to \exists z B[y, z, \vec{v}])) \to (g(x, h) \mathrel{\mathfrak{r}} \exists z B[x, z, \vec{v}]))).$$

Since $B[x, z, \vec{v}]$ is a negative Δ_0 formula, we know by Lemma 80 and the definition of realiser that

$$(s(g, x) \mathrel{\mathfrak{r}} \exists z B[x, z, \vec{v}]) \leftrightarrow \mathbf{p}_0(g(x, \lambda y, f.s(g, y)))\downarrow \wedge B[x, \mathbf{p}_0(g(x, \lambda y, f.s(g, y))), \vec{v}] \tag{7}$$

for all x. Let us write $C[x,\vec{v}]$ for the formula

$$\mathbf{p}_0(g(x,\lambda y,f.s(g,y)))\downarrow \wedge B[x,\mathbf{p}_0(g(x,\lambda y,f.s(g,y))),\vec{v}]).$$

We prove now $\forall x C[x,\vec{v}]$ by the induction principle introduced in Lemma 19. For that purpose we fix an x and assume $(\forall y\in x)C[y,\vec{v}]$ which is by (7) equivalent to

$$\forall y(y\in x\to (s(g,y)\ \mathfrak{r}\ \exists z B[y,z,\vec{v}])),$$

and this implies

$$\lambda y,f.s(g,y)\ \mathfrak{r}\ \forall y(y\in x\to \exists z B[y,z,\vec{v}])$$

by the definition of $\mathfrak{r}$. By the assumption about g we get therefore $g(x,\lambda y,f.s(g,y))\ \mathfrak{r}\ \exists z B[x,z,\vec{v}]$, and it follows by the definition of $\mathfrak{r}$ and by Lemma 80 that

$$\mathbf{p}_0(g(x,\lambda y,f.s(g,y)))\downarrow \wedge B[x,\mathbf{p}_0(g(x,\lambda y,f.s(g,y))),\vec{v}].$$

So we have $C[x,\vec{v}]$. We have proved now that $C[x,\vec{v}]$ follows from $(\forall y\in x)C[y,\vec{v}]$, therefore the former holds by Lemma 19 for every x. Hence we have $s(g,x)\ \mathfrak{r}\ \exists z B[x,z,\vec{v}]$ for all x by (7). All in all we can conclude that for all $\vec{v}$ the term $(\lambda\vec{v}.s)(\vec{v})$ is the realiser we are searching for. □

7.4. Operational Skolemisation

Definition 91 (Operational Skolemisation)**.** If A is an $\mathcal{L}$ formula of the form

$$\forall x_0\exists y_0...\forall x_n\exists y_n B[x_0,...x_n,y_0,...y_n]$$

where B is a negative formula of $\mathcal{L}$, we write $A^{\mathrm{s}}[f_0,...,f_n]$ for its *operational Skolemisation*, the $\mathcal{L}^\circ$ formula

$$\forall x_0,...,x_n B[x_0,...,x_n,f_0(x_0),f_1(x_0,x_1),...,f_n(x_0,...,x_n)]$$

and $A^{\exists\mathrm{s}}$ for the formula $\exists f_0,...,f_n A^{\mathrm{s}}[f_0,...,f_n]$.

If $\mathcal{A}$ is a set of formulae of the form described above, we write $\mathcal{A}^{\exists\mathrm{s}}$ for the set $\{A^{\exists\mathrm{s}} : A\in\mathcal{A}\}$.

This notion seems to fit quite well with the philosophy of explicit mathematics in a broader sense, presented as the slogan: making everything explicit.

Lemma 92. Let $\vec{x}=x_0,...x_n$, $\vec{y}=y_0,...y_n$ and $\vec{z}=z_0,...,z_m$, and $A[\vec{z}]$ be an $\mathcal{L}$ formula of the form $\forall x_0\exists y_0...\forall x_n\exists y_n B[\vec{x},\vec{y},\vec{z}]$ with at most $\vec{z}$ free where B is a negative formula of $\mathcal{L}$. Then there exists an $\mathcal{L}^\circ$ term t such that WEST_0 proves

$$A^{\mathrm{s}}[f_0,...,f_n,\vec{z}]\to t(f_0,...,f_n)\downarrow \wedge \forall\vec{z}(t(f_0,...,f_n,\vec{z})\ \mathfrak{r}\ A[\vec{z}]).$$

PROOF. Assume $A^{\mathrm{s}}[f_0,...,f_n,\vec{z}]$, i.e.

$$\forall\vec{x}B[\vec{x},f_0(x_0),f_1(x_0,x_1),...,f_n(x_0,...,x_n),\vec{z}].$$

Therefore we have by Lemma 80 that

$$r_{B,\vec{x},\vec{y},\vec{z}}(\vec{x},f_0(x_0),f_1(x_0,x_1),...,f_n(x_0,...,x_n),\vec{z})\ \mathfrak{r}\ B[\vec{x},f_0(x_0),f_1(x_0,x_1),...,f_n(x_0,...,x_n),\vec{z}]$$

for all $\vec{x}$. Now let $s[f_0,...,f_n,\vec{z}]$ be the term

$$\lambda x_0.\mathbf{p}(f_0(x_0),\lambda x_1.\mathbf{p}(f_1(x_0,x_1),\ ...\ \lambda x_n.\mathbf{p}(f_n(x_0,...,x_n),r_{B,\vec{x},\vec{y},\vec{z}}(\vec{x},f_0(x_0),f_1(x_0,x_1),...,f_n(x_0,...,x_n),\vec{z}))...)).$$

Then s realises $\forall x_0\exists y_0...\forall x_n\exists y_n B[\vec{x},\vec{y},\vec{z}]$. Thus $\lambda f_0,...,f_n,\vec{z}.s[f_0,...,f_n,\vec{z}]$ is the term we need. □

7.5. Interpretability result

If $\mathcal{A}$ is a set of formulae of the form $\forall x_0 \exists y_0 ... \forall x_n \exists y_n B$, where B is negative, then Lemma 92 implies that every formula in $\mathcal{A}$ is realisable in $\mathsf{WEST}_0 + \mathcal{A}^{\exists s}$. Therefore the following theorem is a consequence of the Lemmata 81 to 92.

Theorem 93. Let $\mathcal{A}$ be a set of formulae of the form $\forall x_0 \exists y_0 ... \forall x_n \exists y_n B$, where B is a negative $\mathcal{L}$ formula. If A is a negative $\mathcal{L}$ formula provable in $\mathsf{IKP}_0^- + \mathcal{A}$, then A is also provable in $\mathsf{WEST}_0 + \mathcal{A}^{\exists s}$.

All the assertions, with the subscript 0 replaced by ω or just omitted, also hold. Furthermore, the analogous assertions hold if ($\mathsf{n}\Sigma_1$-Ind), ($\mathsf{n}\Sigma_1$-Ind_ω) or ($\mathsf{n}\Pi_2$-Ref) is added to the variants of IKP^- and at the same time (opInd), (opInd_ω) or (Π_2-Ref), respectively, is added to the variants of WEST.

Remark 94. Notice that all the arguments used to prove the Lemmata 81 to 92 are also available in intuitionistic logic, except the proofs of Lemma 86 and of Lemma 90 (since it depends on Lemma 19). Also the axiom of extensionality is not used in the proofs of these lemmata, except lemmata for the existence of the closed terms $\mathbf{p}_0$ and $\mathbf{p}_1$ (i.e., Lemma 13 (i); see Remark 14). Therefore the previous theorem holds also for the weakening of WEST_0 based on intuitionistic logic and without extensionality, if we add some harmless axioms for $\mathbf{p}_0$ and $\mathbf{p}_1$ such that Lemma 13 (i) can be also proved without extensionality (that is $\mathbb{U}(\{x\}) = x$). This observation will be essential in the appendix.

8. Merging the Results

Now we are almost ready to merge the three steps we took in the last three sections, to obtain the lower bound of our applicative set theories.

First, we summarise the local interpretability results we have obtained, in Figure 2, where by "local" we mean that the interpretation depends on proofs to be interpreted (through $\mathfrak{S}$ in $\Vdash_{\mathfrak{S}}$ in the current case). Notice that a weak Σ_1 formula is by definition very negatively Σ_1. As mentioned in Remarks 23, 46 and 74 we can make it non-local interpretability results, by restricting the axiom schemata except $\in$-induction with their finitely many instances in $\mathsf{IKP}_0^\sharp$+(Δ_0^{s-}-MP). However, as remarked in Remark 76, we cannot avoid locality if we require Π_1 conservativity of the interpretability, namely, we have only local Π_1-conservative interpretability. However, this minor difference does not affect out main result, Π_1 equivalences.

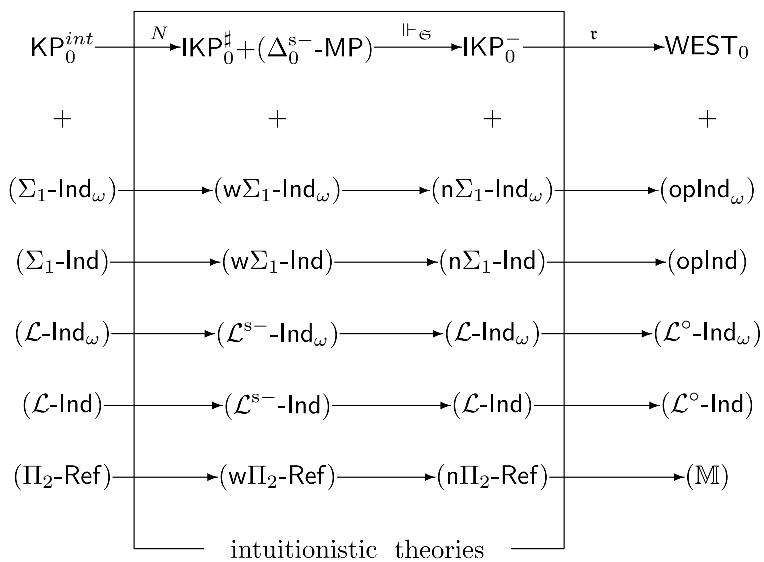

Figure 2: The summary of the last three sections

We introduce the following axiom, a negative interpretation of extensionality, for the next proof:

(N-Ext) $((a = b) \leftrightarrow (\forall x \in a)(x \in b) \wedge (\forall x \in b)(x \in a))^N$.

Theorem 95. Let $\mathcal{A}$ be a set of $\mathcal{L}$ formulae of the form $\forall x_0 \exists y_0 ... \forall x_n \exists y_n \forall x_{n+1} D$, where D is a negative Δ_0 formula. If C is a Π_1 formula of $\mathcal{L}$ provable in $\mathsf{KP}_0 + \mathcal{A}$, then C is provable in $\mathsf{WEST}_0 + \mathcal{A}^{\exists \mathrm{s}}$.

The same assertion, with the subscript 0 replaced by ω or omitted, also hold.

The analogous assertions holds if we add (Σ_1-Ind), (Σ_1-Ind_ω) or (Π_2-Ref) to the variants of KP and, at the same time, add (opInd), (opInd_ω) or ($\mathbb{M}$), respectively, to those of WEST.

Proof. Let C be a Π_1 formula, which is provable in $\mathsf{KP}_0 + \mathcal{A}$. Let $\mathcal{B}$ be the set

$$\{\forall x_0 \exists y_0 ... \forall x_n \exists y_n \forall x_{n+1} D^N : D \text{ is negative } \Delta_0 \text{ and } \forall x_0 \exists y_0 ... \forall x_n \exists y_n \forall x_{n+1} D \in \mathcal{A}\}.$$

By Lemma 33 (13) every formula of $\mathcal{B}$ implies its negative interpretation intuitionistically. Furthermore (N-Ext) is the negative interpretation of axiom (KP.9) of KP_0. Therefore by Lemma 45, $\mathsf{IKP}_0^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP}) + (N\text{-Ext}) + \mathcal{B}$ proves all negative interpretations of axioms of $\mathsf{KP}_0 + \mathcal{A}$. Since C is provable in the latter theory, C^N is provable in $\mathsf{IKP}_0^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP}) + (N\text{-Ext}) + \mathcal{B}$. Because C^N, (N-Ext) and all formulae of $\mathcal{B}$ are in $\mathcal{D}_{res}$, Theorem 75 implies that also $\mathsf{IKP}_0^- + (N\text{-Ext}) + \mathcal{B}$ proves C^N. And since (N-Ext) and $\forall x_{n+1} D^N$ are clearly negative formulae, Theorem 93 implies that C^N is also provable in $\mathsf{WEST}_0 + (N\text{-Ext})^{\exists \mathrm{s}} + \mathcal{B}^{\exists \mathrm{s}}$. Because $(N\text{-Ext})^{\exists \mathrm{s}}$ is identical to (N-Ext) and therefore provable in WEST_0 and because the formulae of $\mathcal{A}^{\exists \mathrm{s}}$ imply the corresponding formulae of $\mathcal{B}^{\exists \mathrm{s}}$ classically, also $\mathsf{WEST}_0 + \mathcal{A}^{\exists \mathrm{s}}$ proves C^N and therefore C. □

In the previous proof it is essential that WEST_0 proves (N-Ext). This proof does therefore not work for an intuitionistic version of WEST. We will consider this point in the appendix.

Theorem 96. The assertion of Theorem 95 holds also if we replace KP_0 by KP_0 + (Beta), KPI_0 or KPM_0 and, at the same time, replace also WEST_0 by $\mathsf{WEST}_0 + (\mathbb{B})$, $\mathsf{WEST}_0 + (\mathcal{AD}) + (\mathbb{A})$ or $\mathsf{WEST}_0 + (\mathcal{AD}) + (\mathbb{M})$ respectively.

All the assertions, with the subscript 0 replaced by ω, r or w, or omitted, also hold. The analogous assertions holds if we add (Σ_1-Ind) or (Σ_1-Ind_ω) to the variants of KP and, at the same time, add (opInd) or (opInd_ω), respectively, to those of WEST.

Proof. It suffices to show that any axiom added to KP_0, other than those mentioned in Theorem 95, is implied in KP_0 from some axiom A such that the corresponding axioms added to WEST_0 implies $A^{\exists \mathrm{s}}$ in WEST_0.

First, consider $\in$-induction for Δ_0 formulae. This is equivalent to the Π_1 formula

$$\forall u, a(x \in u \wedge \mathrm{Tran}[u] \wedge (\forall y \in u)((\forall z \in y)(z \in a) \rightarrow y \in a) \rightarrow x \in a)$$

and so the operational Skolemisation of this is equivalent to itself.

Second, consider (Beta). Since ($\mathbb{B}$) can be seen as the operational Skolemisation of (Beta$'$), Lemma 29 shows that (Beta$'$) is what we need.

The other axiom, namely $(\mathcal{AD}) + (\mathbb{A})$ can be seen as the operational Skolemisation of $(\mathcal{AD}) + (\mathsf{Lim})$. □

We want to apply Theorem 95 to the theories $\mathsf{KP} + (\mathcal{P})$ and $\mathsf{WEST} + (\mathbb{P})$. Doing it directly is not possible, since the theory $\mathsf{WEST} + (\mathbb{P})$ does not contain any axioms about the relation $\mathcal{P}$. Therefore we translate formulae containing $\mathcal{P}$ to formulae of $\mathcal{L}_\in^\circ$. We will write $\mathcal{L}_\mathcal{P}$ for the language $\mathcal{L}$ but restricted to the relation symbols $\in$, $=$ and $\mathcal{P}$.

Definition 97 (Formula $A_\mathbb{P}$)**.** If A is a formula of $\mathcal{L}_\mathcal{P}$, we write $A_\mathbb{P}$ for the $\mathcal{L}_\in^\circ$ formula which we get if we replace any occurrence of $\mathcal{P}(x, y)$ by $\mathbb{P}x = y$.

Lemma 98. If A is an $\mathcal{L}^\circ$ formula provable in $\mathsf{WEST}_0 + (\mathcal{P}) + (\mathbb{P})$ then $A_\mathbb{P}$ is provable in $\mathsf{WEST}_0 + (\mathbb{P})$. The assertion also holds if we delete the subscript 0 or replace it by ω, r or w.

That the previous lemma holds is obvious, since if B is the axiom ($\mathcal{P}$) then $B_\mathbb{P}$ is provable in $\mathsf{WEST}_0 + (\mathbb{P})$.

Theorem 99. For any Π_1 formula A of $\mathcal{L}_\in$, any Π_1 formula B of $\mathcal{L}_\mathcal{P}$ and any Π_1 formula C of $\mathcal{L}_{\mathsf{Ad}}$:

(i) if A is provable in KP_0, it is provable in WEST_0 and so in OST_0^-;
(ii) if A is provable in $\mathsf{KP}_0 + (\text{Beta})$, it is provable in $\mathsf{WEST}_0 + (\mathbb{B})$ and in $\mathsf{OST}_0^- + (\mathbb{B})$;
(iii) if B is provable in $\mathsf{KP}_0 + (\mathcal{P})$, then $B_\mathbb{P}$ is provable in $\mathsf{WEST}_0 + (\mathbb{P})$ and $\mathsf{OST}_0^- + (\mathbb{P})$;
(iv) if C is provable in KPI_0, it is provable in $\mathsf{WEST}_0 + (\mathcal{AD}) + (\mathbb{A})$ and in $\mathsf{OST}_0^- + (\mathcal{AD}) + (\mathbb{A})$;
(v) if C is provable in KPM_0, it is provable in $\mathsf{WEST}_0 + (\mathcal{AD}) + (\mathbb{M})$ and in $\mathsf{OST}_0^- + (\mathcal{AD}) + (\mathbb{M})$.

In these results, we can delete the subscript 0 or replace it by ω, r, w, and also we can add (Σ_1-Ind) or (Σ_1-Ind_ω) to the variants of KP and, at the same time, add (opInd) or (opInd_ω) respectively to those of WEST and of OST^-.

Proof. All four assertions, except the third, are proved in Theorem 96.

For the third assertion let A_0 be the $\mathcal{L}_\mathcal{P}$ formula $\forall x \exists y \mathcal{P}(x, y)$, A_1 the $\mathcal{L}_\mathcal{P}$ formula

$$\forall z(\mathcal{P}(x, y) \to (z \in y \leftrightarrow z \subseteq x))$$

and $\mathcal{A}$ the set $\{A_0, A_1\}$. Then $\mathsf{KP}_0 + \mathcal{A}$ clearly proves that there can be at most one powerset of a given set (by extensionality) and therefore it proves the axiom $(\mathcal{P})$. Furthermore $\mathsf{WEST}_0 + (\mathcal{P}) + (\mathbb{P})$ clearly proves $A_0^{\mathrm{s}}[\mathbb{P}]$ (that is the formula $\forall x \mathcal{P}(x, \mathbb{P}(x))$) and A_1^{s} (it stays the formula A_1). So $\mathsf{WEST}_0 + (\mathcal{P}) + (\mathbb{P})$ contains $\mathsf{WEST}_0 + \mathcal{A}^{\exists\mathrm{s}}$ and proves by Theorem 95 all Π_1 formulae of $\mathcal{L}$ which are provable in $\mathsf{KP}_0 + (\mathcal{P})$. Finally, the assertion follows from the previous lemma. ☐

Remark 100. It is easy to see, that the third line can be generalised in the following way. Let $\mathbb{Q}$ be a new constant symbol which we add to $\mathcal{L}^\circ$, Q a binary relation symbol of $\mathcal{L}$, and $\mathcal{Q}$ some set of Π_1 formulae which imply that for each x there is at most one y such that $Q(x, y)$. Then we can introduce a translation $A_\mathbb{Q}$ of every formula A of $\mathcal{L}_\mathbb{Q}^\circ$ (analogously as we did for $\mathbb{P}$) such that, for any Π_1 formula A of $\mathcal{L}_\mathbb{Q}$,

if $\mathsf{KP} + \mathcal{Q} + (\forall x \exists y Q(x, y))$ proves A then $\mathsf{WEST} + \mathcal{Q} + \forall x Q(x, \mathbb{Q}(x))$ proves $A_\mathbb{Q}$.

This, or more generally the notion of operational Skolemisation, guides us how to define the axiom of applicative set theory, corresponding to a new additional axiom added to (variants of) KP, if the additional axiom is in an appropriate form.

For the readers' convenience, the correspondence between axioms to be added to KP_0 and those to be added to WEST_0 is summarised in Table 2, where the systems on the upper line with any choice of axioms on the upper line are (locally, if Π_1-conservativity is required) interpreted in the corresponding systems on the lower lines with the same choice of corresponding axioms on the lower line.

KP_0	KP_ω	KP_r	KP_w	KP
WEST_0	WEST_ω	WEST_r	WEST_w	WEST
OST_0^-	OST_ω^-	OST_r^-	OST_w^-	OST^-

added by any of

(Σ_1-Ind_ω)	(Σ_1-Ind)	(Beta)	$(\mathcal{P})$	$(\mathcal{AD}) + (\mathsf{Lim})$	$(\mathcal{AD}) + (\Pi_2$-$\mathsf{Ref})$
(opInd_ω)	(opInd)	$(\mathbb{B})$	$(\mathbb{P})$	$(\mathcal{AD}) + (\mathbb{A})$	$(\mathcal{AD}) + (\mathbb{M})$

Table 2: Correspondence between systems and axioms 1

One of the advantages of our result is such modularity of axioms. We have the result for all combinations of all these systems and axioms. This is because we have the (local) interpretability results based on the same (local) interpretation. This kind of modularity is not always available for proof theoretic reducibility results, typically by cut-elimination method, because there might be a problematic interact between additional axioms.

Part III
Upper Bound: Inductive Model

9. Interpreting OST^- in KP: Inductive Definitions

In this section, we will give upper bounds of our applicative set theories. The main task is to define the interpretation of the application $\circ$ by inductive definition. Similar inductive model constructions were presented in Jäger [29] and Jäger and Zumbrunnen [38]. For these constructions, the axiom of constructibility $V = L$ was used. We can use the same method with only a few changes to embed OST^- and WEST (and some extensions) into KP (and corresponding extensions) without the axiom of constructibility, as Zumbrunnen [59]. However, in order to embed some subsystems of OST^- and WEST into the corresponding subsystems of KP, we need to look into more details of inductive definition.

9.1. Σ recursion and Σ inductive definition

If P is an n-ary relation symbol of $\mathcal{L}$, $A[P]$ is a formula of $\mathcal{L}$ and $B[x_0, ..., x_{n-1}]$ is a formula of $\mathcal{L}$ with distinguished free variables $x_0, ..., x_{n-1}$, we write $A[\{(x_0, ..., x_{n-1}) : B[x_0, ..., x_{n-1}]\}]$ for the result of substituting $B[v_0, ..., v_{n-1}]$ for each occurrence of the form $P(v_0, ..., v_{n-1})$ in $A[P]$ and of renaming bound variables if necessary to avoid collision. In Jäger and Zumbrunnen [38] a version of Σ recursion is used. We use here a slightly different version of it. Our formulation follows, in the presence of full $\in$-induction, directly from the *Second Recursion Theorem* in Barwise [4, p.157], as Zumbrunnen [59] used.

Theorem 101 (Σ recursion)**.** Let P be an n-ary relation symbol and $A[x_0, ..., x_{n-1}, P]$ a Σ formula of $\mathcal{L}$ with distinguished free variables $\vec{x} = x_0, ..., x_{n-1}$ in which P occurs only positively. Then there exists a Σ formula $B[\vec{x}]$ of $\mathcal{L}$ such that KP_r proves

$$B[x_0, ..., x_{n-1}] \leftrightarrow A[x_0, ..., x_{n-1}, \{(y_0, ..., y_{n-1}) : B[y_0, ..., y_{n-1}]\}].$$

Although this only asserts that $\{(y_0, ..., y_{n-1}) : B[y_0, ..., y_{n-1}]\}$ is a fixed point of the operator on classes defined by $P \mapsto \{(x_0, ..., x_{n-1}) : A[x_0, ..., x_{n-1}, P]\}$, this theorem is enough to define a *least* fixed point. For by applying the theorem to

$$A'[\alpha, x_0, ..., x_{n-1}, P'] := A[\alpha, x_0, ..., x_{n-1}, \{(y_0, ..., y_{n-1}) : P'[\beta, y_0, ..., y_{n-1}] \wedge \beta \in \alpha\}]$$

with $n{+}1$-ary relation symbol P', we can define a least fixed point from the resulted fixed point B' of A' by

$$\{(x_0, ..., x_{n-1}) : \exists\alpha(\alpha \in \mathrm{Ord} \wedge B'[\alpha, x_0, ..., x_{n-1}])\}.$$

However, the (usual) proof of this theorem is by diagonalisation and we could not have a concrete definition of B or B'. In the context of restricted $\in$-induction, this is unsatisfactory. We need a more detailed definition of required B. The basic idea is to describe it by iterated application of the operator defined by A to the emptyset (and therefore the least fixed point directly, not via a non-least fixed point). More precisely, if Γ denotes the operator on classes $P \mapsto \{(x_0, ..., x_{n-1}) : A[x_0, ..., x_{n-1}, P]\}$ it is known that the least fixed point of Γ can be described by $\bigcup_{\alpha\in\mathrm{Ord}} \Gamma^\alpha(\emptyset)$. However, we cannot formalise this description directly in our context, because there is no way to describe the sequence $(\Gamma^\alpha(\emptyset) : \alpha \in \mathrm{Ord})$ of classes directly. Instead, we can have a "set-size" fragment f of this sequence, in the sense that $f'\alpha \subseteq \Gamma^\alpha(\emptyset)$ for any α in the domain of f. The formula $C[f, \alpha]$ introduced below is intended to mean that f is such a set-size fragment. As we will see, if A is a Σ formula, these fragments are many enough to approximate the sequence $(\Gamma^\alpha(\emptyset) : \alpha \in \mathrm{Ord})$, in the sense that $\Gamma^\alpha(\emptyset)$ is the union of all $f'\alpha$ for such fragments f.

In the following we use Greek letters $\alpha, \beta, \gamma, ...$ to range over ordinal numbers. Since we are working in KP_r, being an ordinal is a Δ_0 predicate.

Theorem 102 (Σ inductive definition)**.** Let P be an n-ary relation symbol and $A[x_0, ..., x_{n-1}, P]$ a Σ formula of $\mathcal{L}$ with distinguished free variables $x_0, ..., x_{n-1}$ in which P occurs only positively. Define two Σ formulae $B[\vec{x}]$ and $C[f, \alpha]$ of $\mathcal{L}$ as follows:

$$C[f,\alpha] := \mathrm{Fun}[f] \wedge \mathrm{Dom}[f, \alpha{+}1] \wedge (\forall \beta \in \alpha{+}1) \left(f'\beta \subseteq \left\{ \langle y_0, ..., y_{n-1} \rangle : A \left[y_0, ..., y_{n-1}, \bigcup_{\gamma \in \beta} f'\gamma \right] \right\} \right),$$
$$B[x_0, ..., x_{n-1}] := (\exists f, \alpha) \left(C[f, \alpha] \wedge \langle x_0, ..., x_{n-1} \rangle \in f'\alpha \right),$$

where $\bigcup_{\gamma \in \beta} f'\gamma$ is substituted into the relation symbol P, as $\{(z_0, ..., z_{n-1}) : (\exists \gamma \in \beta)(\langle z_0, ..., z_{n-1} \rangle \in f'\gamma)\}$.
Then KP_r proves

(i) $(\forall x_0, ..., x_{n-1})(B[x_0, ..., x_{n-1}] \leftrightarrow A\,[x_0, ..., x_{n-1}, \{(y_0, ..., y_{n-1}) : B[y_0, ..., y_{n-1}]\}])$, and
(ii) $(\forall x_0, ..., x_{n-1})(A\,[x_0, ..., x_{n-1}, \{(y_0, ..., y_{n-1}) : D[y_0, ..., y_{n-1}]\}] \to D[x_0, ..., x_{n-1}])$
$\to (\forall x_0, ..., x_{n-1})(B[x_0, ..., x_{n-1}] \to D[x_0, ..., x_{n-1}])$
for any Δ formula D.

Proof. Assume $B[x_0, ..., x_{n-1}]$, say $C[f, \alpha]$ and $\langle x_0, ..., x_{n-1} \rangle \in f'\alpha$. Thus $\langle x_0, ..., x_{n-1} \rangle \in f'\alpha \subseteq \left\{ \langle y_0, ..., y_{n-1} \rangle : A \left[y_0, ..., y_{n-1}, \bigcup_{\gamma \in \alpha} f'\gamma \right] \right\}$, namely

$$A \left[x_0, ..., x_{n-1}, \bigcup_{\gamma \in \alpha} f'\gamma \right]. \tag{8}$$

Obviously $C[f \restriction (\gamma + 1), \gamma]$ for any $\gamma \in \alpha$, and hence $\bigcup_{\gamma \in \alpha} f'\gamma \subseteq \{\langle y_0, ..., y_{n-1} \rangle : B[y_0, ..., y_{n-1}]\}$. Since P occurs only positively in $A[x_0, ..., x_{n-1}, P]$, we can replace $\bigcup_{\gamma \in \alpha} f'\gamma$ by $B[y_0, ..., y_{n-1}]$ in (8), namely

$$A[x_0, ..., x_{n-1}, \{(y_0, ..., y_{n-1}) : B[y_0, ..., y_{n-1}]\}].$$

Conversely, assume $A[x_0, ..., x_{n-1}, \{(y_0, ..., y_{n-1}) : B[y_0, ..., y_{n-1}]\}]$. By Σ *Reflection Principle*, which is provable in KP_r by the same proof given in Barwise [4, p.16], there is a set a such that $\langle x_0, ..., x_{n-1} \rangle \in a$ and $A^{(a)}[x_0, ..., x_{n-1}, \{(y_0, ..., y_{n-1}) : B^{(a)}[y_0, ..., y_{n-1}]\}]$, namely, with the aforementioned terminology,

$$A^{(a)}[x_0, ..., x_{n-1}, \{\langle y_0, ..., y_{n-1} \rangle \in a : B^{(a)}[y_0, ..., y_{n-1}]\}].$$

By Δ_0 separation, we set

$$b = \{\langle y_0, ..., y_{n-1} \rangle \in a : B^{(a)}[y_0, ..., y_{n-1}]\}.$$

Then, for any $\langle y_0, ..., y_{n-1} \rangle \in b$, $B^{(a)}[y_0, ..., y_{n-1}]$ holds and, by upward persistency so does $B[y_0, ..., y_{n-1}]$, which means that $\langle y_0, ..., y_{n-1} \rangle \in f'\beta$ for some f and β with $C[f, \beta]$. That is,

$$(\forall \langle y_0, ..., y_{n-1} \rangle \in b)(\exists f, \beta)(C[f, \beta] \wedge \langle y_0, ..., y_{n-1} \rangle \in f'\beta).$$

Again, by Σ Reflection Principle, there is c such that

$$(\forall \langle y_0, ..., y_{n-1} \rangle \in b)(\exists f, \beta \in c)(C^{(c)}[f, \beta] \wedge \langle y_0, ..., y_{n-1} \rangle \in f'\beta).$$

Define α and g with $\mathrm{Fun}[g]$ and $\mathrm{Dom}[g, \alpha + 2]$ by

$$\alpha = \sup\{\beta + 1 : \beta \in c \cap \mathrm{Ord}\}$$
$$g'\gamma = \begin{cases} \{\langle y_0, ..., y_{n-1} \rangle \in b : (\exists f \in c)(\exists \beta \in \gamma{+}1)(C^{(c)}[f, \beta] \wedge \langle y_0, ..., y_{n-1} \rangle \in f'\beta)\} & \text{if } \gamma \in \alpha \\ b & \text{if } \gamma = \alpha \\ \{\langle x_0, ..., x_{n-1} \rangle\} & \text{if } \gamma = \alpha + 1 \end{cases}.$$

To show $B[x_0, ..., x_{n-1}]$, since $\langle x_0, ..., x_{n-1} \rangle \in g'(\alpha{+}1)$, it remains to show $C[g, \alpha{+}1]$. We shall prove the following for any $\gamma \in \alpha{+}2$:

$$g'\gamma \subseteq \left\{ \langle y_0, ..., y_{n-1} \rangle : A \left[y_0, ..., y_{n-1}, \bigcup_{\delta \in \gamma} g'\delta \right] \right\}.$$

Take $\langle y_0, ..., y_{n-1}\rangle \in g'\gamma$. We shall prove $A\left[y_0, ..., y_{n-1}, \bigcup_{\delta\in\gamma} g'\delta\right]$. If $\gamma \in \alpha$ or $\gamma = \alpha$, there are f in c and β in $\gamma{+}1$ such that $\langle y_0, ..., y_{n-1}\rangle \in f'\gamma$ with $C^{(c)}[f, \beta]$. Note that $f'\delta \subseteq g'\delta$ for any $\delta \in \gamma$. By upward persistency, $C[f, \beta]$, in particular

$$\langle y_0, ..., y_{n-1}\rangle \in f'\gamma \subseteq \left\{\langle z_0, ..., z_{n-1}\rangle : A\left[z_0, ..., z_{n-1}, \bigcup_{\delta\in\gamma} f'\delta\right]\right\},$$

where we can replace $\bigcup_{\delta\in\gamma} f'\delta$ by $\bigcup_{\delta\in\gamma} g'\delta$, resulting in what we need.

Finally, let us consider the case of $\gamma = \alpha{+}1$. By the choice of a, we have $A^{(a)}[x_0, ..., x_{n-1}, g'\alpha]$ and by upward persistency $A[x_0, ..., x_{n-1}, g'\alpha]$, that is $g'(\alpha{+}1) = \{\langle x_0, ..., x_{n-1}\rangle\} \subseteq \{\langle z_0, ..., z_{n-1}\rangle : A[z_0, ..., z_{n-1}, g'\alpha]\}$.

For the second assertion, let

$$(\forall x_0, ..., x_{n-1})(A\left[x_0, ..., x_{n-1}, \{(y_0, ..., y_{n-1}) : D[y_0, ..., y_{n-1}]\}\right] \to D[x_0, ..., x_{n-1}]).$$

Assume further $B[z_0, ..., z_{n-1}]$, say $C[f, \alpha]$ and $\langle z_0, ..., z_{n-1}\rangle \in f'\alpha$. We need to show $D[z_0, ..., z_{n-1}]$. We can prove $f'\beta \subseteq \{\langle y_0, ..., y_{n-1}\rangle : D[y_0, ..., y_{n-1}]\}$ by Δ induction on β:

$$\begin{aligned} f'\beta &\subseteq \left\{\langle x_0, ..., x_{n-1}\rangle : A\left[x_0, ..., x_{n-1}, \bigcup_{\gamma\in\beta} f'\gamma\right]\right\} \\ &\subseteq \{\langle x_0, ..., x_{n-1}\rangle : A\left[x_0, ..., x_{n-1}, \{(y_0, ..., y_{n-1}) : D[y_0, ..., y_{n-1}]\}\right]\} \subseteq \{\langle x_0, ..., x_{n-1}\rangle : D[x_0, ..., x_{n-1}]\}, \end{aligned}$$

where the second inclusion follows from the induction hypothesis. Thus $\langle z_0, ..., z_{n-1}\rangle \in f'\alpha \subseteq \{\langle y_0, ..., y_{n-1}\rangle : D[y_0, ..., y_{n-1}]\}$. □

Note that, in the theorem, B is only Σ and hence it cannot be an instance of D.

9.2. Some remarks on the inductive definitions

Remark 103. Since the Δ-ness of D is needed only when we apply $\in$-induction in the proof of the last theorem, KP proves the assertion (ii) extended to arbitrary formulae, and $\mathsf{KP}_r + (\Sigma_1\text{-}\mathsf{Ind})$ proves that to Σ formulae. Thus, in a completely uniform way, we can *directly* (without cut-elimination method) interpret

- (i) in KP, the first order system ID_1 of non-iterated inductive definition;
- (ii) in KP_w, both the first order system $\widehat{\mathsf{ID}}_1$ of non-iterated *hat* inductive definition or of non-iterated inductive fixed-point from Feferman [15], and the first order system SID_1 of non-iterated *stratified* inductive definition from Ranzi and Strahm [46] and from Jäger and Probst [33];
- (iii) in $\mathsf{KP}_r + (\Sigma_1\text{-}\mathsf{Ind}_\omega)$, the first order system $\mathsf{ID}_1^{\#}$ of non-iterated *sharp* inductive definition from Jäger and Strahm [34] (which, as our argument shows, can be enhanced with the fixed point induction for formulae without the fixed point predicate);
- (iv) in KP_r, the first order system ID_1^0 consisting of fixed point axiom for the fixed point predicate and the induction schema on ω restricted to formulae without the fixed point predicate.

With the known results $|\mathsf{KP}_r| = \varepsilon_0$ and $|\mathsf{KP}_w| = \varphi\varepsilon_0 0$ from Jäger [26], and with $|\mathsf{KP}_r + (\Sigma_1\text{-}\mathsf{Ind}_\omega)| = \varphi\omega 0$ which can be shown similarly, we know the optimality of these interpretations in the sense of proof theoretic strengths.

Remark 104. Despite the proof theoretic optimality of the interpretations in Remark 103, there is actually some redundancy: firstly, the interpreted systems allow only variants of inductive definitions of subsets of ω and hence we have not only a finite approximation of the constructing sequence of a least fixed point but also the sequence itself; and, secondly, since the interpreted systems allow only first order quantifiers, the operator forms A are Δ_0. If we look at this redundancy more closely, we can refine the interpretability results, as follows.

In the proof, we can replace $C[f, \alpha]$ by

$$C'[f, \alpha] := \mathrm{Fun}[f] \wedge \mathrm{Dom}[f, \alpha{+}1] \wedge (\forall \beta \in \alpha{+}1)\left(f'\beta = \left\{\langle y_0, ..., y_{n-1}\rangle \in \omega : A\left[y_0, ..., y_{n-1}, \bigcup_{\gamma\in\beta} f'\gamma\right]\right\}\right),$$

b by $b' = \{\langle y_0, ..., y_{n-1}\rangle \in \omega : B^{(a)}[y_0, ..., y_{n-1}]\}$ and g by h defined in the same way but

$$h'(\alpha+1) = \{\langle z_0, ..., z_{n-1}\rangle \in \omega : A^{(a)}[z_0, ..., z_{n-1}, \{\langle y_0, ..., y_{n-1}\rangle \in \omega : B^{(a)}[y_0, ..., y_{n-1}]\}]\},$$

where $A^{(a)}$ is the same as A, if A is Δ_0. Then since within KP_r there is at most one f with $C'[f,\alpha]$, we can prove the same statement but with every quantifiers restricted to ω. Moreover, the relation

$$(\exists f)(C'[f,\alpha] \wedge \langle y_1, ..., y_{n-1}\rangle \in f'\alpha)$$

saying that the tuple $(y_1, ..., y_{n-1})$ is in the α-th stage, is now Δ formula, for α from $\{\alpha : (\exists f)C'[f,\alpha]\}$.

Based on this observation, we can interpret Jäger-style theories of natural numbers and ordinals in pure set theories, by taking $\{\alpha : (\exists f)C'[f,\alpha]\}$ as the domain for the ordinal sort. More precisely, we can interpret

(i) in KP, the first order system PA_Ω of numbers and ordinals with both full induction on numbers and full induction on ordinals, from Jäger [27];
(ii) in KP_w, the first order system PA^w_Ω of numbers and ordinals with full induction on numbers and only Δ^Ω induction on ordinals, also from Jäger [27];
(iii) in $\mathsf{KP}_r + (\Sigma_1\text{-}\mathsf{Ind}_\omega)$, the first order system $\mathsf{PA}^r_\Omega + (\Sigma^\Omega\text{-}\mathsf{I}_\mathbb{N})$ of numbers and ordinals with Σ^Ω induction on numbers and only Δ^Ω induction on ordinals, from Jäger and Strahm [34];
(iv) in KP_r, the first order system PA^r_Ω of numbers and ordinals with both Δ^Ω induction on numbers and only Δ^Ω induction on ordinals, from Jäger [27].

As shown in Jäger and Strahm [34, Theorem 25], it is easy to interpret the variants of systems of inductive definition in the corresponding systems of numbers and ordinals, by making use of so-called Σ^Ω reflection. Thus we have interpretations in Figure 3, which are all optimal in the sense of proof theoretic strengths.

Systems of Inductive Definitions		Systems of Numbers and Ordinals		Set Theory
ID_1	$\longrightarrow$	PA_Ω	$\longrightarrow$	KP
$\widehat{\mathsf{ID}}_1$	$\longrightarrow$	PA^w_Ω	$\longrightarrow$	KP_w
$\mathsf{ID}^\#_1$	$\longrightarrow$	$\mathsf{PA}^r_\Omega + (\Sigma^\Omega\text{-}\mathsf{I}_\mathbb{N})$	$\longrightarrow$	$\mathsf{KP}_w + (\Sigma_1\text{-}\mathsf{Ind}_\omega)$
ID^0_1	$\longrightarrow$	PA^r_Ω	$\longrightarrow$	KP_r

Figure 3: Interpretability for Fixed Point Systems

9.3. Application relation by inductive definition

We fix pairwise different sets $\widehat{\mathsf{k}}$, $\widehat{\mathsf{s}}$, $\widehat{\mathsf{t}}$, $\widehat{\mathsf{f}}$, $\widehat{\mathsf{el}}$, $\widehat{\mathsf{ch}_{\mathsf{Ad}}}$, $\widehat{\mathsf{non}}$, $\widehat{\mathsf{dis}}$, $\widehat{\mathsf{e}}$, $\widehat{\mathbb{S}}$, $\widehat{\mathbb{R}}$, $\widehat{\mathbb{K}}$, $\widehat{\mathbb{T}}$, $\widehat{\mathbb{D}}$, $\widehat{\mathbb{U}}$, $\widehat{\mathbb{P}}$, $\widehat{\mathbb{B}}$, $\widehat{\mathbb{A}}$ and $\widehat{\mathbb{M}}$ which are not ordered pairs nor triples. We are ready to define different versions of a formula $\mathfrak{A}[P,\alpha,a,b,c]$ in a similar way as in Jäger [29] or Jäger and Zumbrunnen [38], where P is a relation symbol of arity 4, which will be used for defining an interpretation of term application. The main difference is, that in our version P occurs only positively and no mention of the constructible hierarchy is made. Both are possible, because we do not have to treat the operation $\mathbb{C}$.

Here, for readability, we consider only $R = \mathsf{Ad}$ is the only relation symbol in $\mathcal{L}$ besides those in $\mathcal{L}_\in$. The following argument easily adapts other relation symbols.

Definition 105. The formula $\mathfrak{A}[a,b,c,P]$ is the disjunction of the formulae (AP.1)-(AP.26), $\mathfrak{A}_\mathbb{B}[a,b,c,P]$ the disjunction of the formulae (AP.1)-(AP.28), $\mathfrak{A}_\mathbb{P}[a,b,c,P]$ the disjunction of the formulae (AP.1)-(AP.26) and (AP.29), $\mathfrak{A}_\mathbb{A}[a,b,c,P]$ the disjunction of the formulae (AP.1)-(AP.26) and (AP.30)-(AP.32) and $\mathfrak{A}_\mathbb{M}[a,b,c,P]$ the disjunction of the formulae (AP.1)-(AP.26) and (AP.30), (AP.31), (AP.33), (AP.34), from Table 3.

(AP.1) $a = \widehat{\mathsf{k}} \wedge c = \langle \widehat{\mathsf{k}}, b \rangle$,
(AP.2) $\mathrm{Tup}_2[a] \wedge (a)_0 = \widehat{\mathsf{k}} \wedge (a)_1 = c$,
(AP.3) $a = \widehat{\mathsf{s}} \wedge c = \langle \widehat{\mathsf{s}}, b \rangle$,
(AP.4) $\mathrm{Tup}_2[a] \wedge (a)_0 = \widehat{\mathsf{s}} \wedge c = \langle \widehat{\mathsf{s}}, (a)_1, b \rangle$,
(AP.5) $\mathrm{Tup}_3[a] \wedge (a)_0 = \widehat{\mathsf{s}} \wedge \ \exists x, y(P((a)_1, b, x) \wedge P((a)_2, b, y) \wedge P(x, y, c))$,
(AP.6) $a = \widehat{\mathsf{el}} \wedge c = \langle \widehat{\mathsf{el}}, b \rangle$,
(AP.7) $\mathrm{Tup}_2[a] \wedge (a)_0 = \widehat{\mathsf{el}} \wedge (a)_1 \in b \wedge c = \widehat{\mathsf{t}}$,
(AP.8) $\mathrm{Tup}_2[a] \wedge (a)_0 = \widehat{\mathsf{el}} \wedge (a)_1 \notin b \wedge c = \widehat{\mathsf{f}}$,
(AP.9) $a = \widehat{\mathsf{non}} \wedge b = \widehat{\mathsf{t}} \wedge c = \widehat{\mathsf{f}}$,
(AP.10) $a = \widehat{\mathsf{non}} \wedge b = \widehat{\mathsf{f}} \wedge c = \widehat{\mathsf{t}}$,
(AP.11) $a = \widehat{\mathsf{dis}} \wedge c = \langle \widehat{\mathsf{dis}}, b \rangle$,
(AP.12) $\mathrm{Tup}_2[a] \wedge (a)_0 = \widehat{\mathsf{dis}} \wedge (a)_1 = \widehat{\mathsf{t}} \wedge c = \widehat{\mathsf{t}}$,
(AP.13) $\mathrm{Tup}_2[a] \wedge (a)_0 = \widehat{\mathsf{dis}} \wedge (a)_1 = \widehat{\mathsf{f}} \wedge b = \widehat{\mathsf{t}} \wedge c = \widehat{\mathsf{t}}$,
(AP.14) $\mathrm{Tup}_2[a] \wedge (a)_0 = \widehat{\mathsf{dis}} \wedge (a)_1 = \widehat{\mathsf{f}} \wedge b = \widehat{\mathsf{f}} \wedge c = \widehat{\mathsf{f}}$,
(AP.15) $a = \widehat{\mathsf{e}} \wedge c = \langle \widehat{\mathsf{e}}, b \rangle$,
(AP.16) $\mathrm{Tup}_2[a] \wedge (a)_0 = \widehat{\mathsf{e}} \wedge (\exists x \in b)(P((a)_1, x, \widehat{\mathsf{t}})) \wedge c = \widehat{\mathsf{t}}$,
(AP.17) $\mathrm{Tup}_2[a] \wedge (a)_0 = \widehat{\mathsf{e}} \wedge (\forall x \in b)(P((a)_1, x, \widehat{\mathsf{f}})) \wedge c = \widehat{\mathsf{f}}$,
(AP.18) $a = \widehat{\mathbb{K}} \wedge (\forall s \in c)(\exists n \in \omega)(\mathrm{Fun}[s] \wedge \mathrm{Dom}[s, n] \wedge \mathrm{Ran}[s, b])$
$\wedge\, \emptyset \in c \wedge (\forall s \in c)(\forall x \in b)(\forall n \in \omega)(\mathrm{Dom}[s, n] \to s \cup \{\langle n, x \rangle\} \in c)$,
(AP.19) $a = \widehat{\mathbb{T}} \wedge \mathrm{Tran}[c] \wedge b \subseteq c \wedge (\forall x \in c)(\exists f)(\exists n \in \omega)(\mathrm{Fun}[f] \wedge \mathrm{Dom}[f, n{+}1]$
$\wedge\, (\forall k \in n)(f'(k{+}1) \in f'k) \wedge f'0 \in b \wedge f'n = x)$,
(AP.20) $a = \widehat{\mathbb{D}} \wedge c = \langle \widehat{\mathbb{D}}, b \rangle$,
(AP.21) $\mathrm{Tup}_2[a] \wedge (a)_0 = \widehat{\mathbb{D}} \wedge c = \{(a)_1, b\}$,
(AP.22) $a = \widehat{\mathbb{U}} \wedge c = \cup b$,
(AP.23) $a = \widehat{\mathbb{S}} \wedge c = \langle \widehat{\mathbb{S}}, b \rangle$,
(AP.24) $\mathrm{Tup}_2[a] \wedge (a)_0 = \widehat{\mathbb{S}} \wedge (\forall x \in c)(x \in b \wedge P((a)_1, x, \widehat{\mathsf{t}})) \wedge (\forall x \in b)(x \notin c \to P((a)_1, x, \widehat{\mathsf{f}}))$,
(AP.25) $a = \widehat{\mathbb{R}} \wedge c = \langle \widehat{\mathbb{R}}, b \rangle$,
(AP.26) $\mathrm{Tup}_2[a] \wedge (a)_0 = \widehat{\mathbb{R}} \wedge (\forall x \in b)(\exists y \in c)(P((a)_1, x, y)) \wedge \ (\forall y \in c)(\exists x \in b)(P((a)_1, x, y))$,
(AP.27) $a = \widehat{\mathbb{B}} \wedge c = \langle \widehat{\mathbb{B}}, b \rangle$,
(AP.28) $\mathrm{Tup}_2[a] \wedge (a)_0 = \widehat{\mathbb{B}} \wedge \exists x(\ x \subseteq (a)_1 \wedge \mathrm{Fun}[c] \wedge \mathrm{Dom}[c, x] \wedge (\forall y \in x)\mathrm{WP}[y, (a)_1, b]$
$\wedge (\forall y \in (a)_1)(\mathrm{WP}'[y, (a)_1, b] \to y \in x) \wedge (\forall y \in x)(c'y = \{c'z : z \in x \wedge \langle z, y \rangle \in b\})\ )$,
(AP.29) $a = \widehat{\mathbb{P}} \wedge \mathcal{P}(b, c)$,
(AP.30) $a = \widehat{\mathsf{ch}_{\mathsf{Ad}}} \wedge \mathsf{Ad}(b) \wedge c = \widehat{\mathsf{t}}$,
(AP.31) $a = \widehat{\mathsf{ch}_{\mathsf{Ad}}} \wedge \neg\mathsf{Ad}(b) \wedge c = \widehat{\mathsf{f}}$,
(AP.32) $a = \widehat{\mathbb{A}} \wedge b \in c \wedge \mathrm{Tran}[c] \wedge \mathsf{Ad}(c) \wedge (\forall x \in c)(\neg b \in x \vee \neg\mathrm{Tran}[x] \vee \neg\mathsf{Ad}(x))$.
(AP.33) $a = \widehat{\mathbb{M}} \wedge c = \langle \widehat{\mathbb{M}}, b \rangle$.
(AP.34) $\mathrm{Tup}_2[a] \wedge (a)_0 = \widehat{\mathbb{M}} \wedge b \in c \wedge \mathrm{Tran}[c] \wedge \mathsf{Ad}(c) \wedge (\forall x \in c)(\exists y \in c)P((a)_1, x, y)$
$\wedge\, (\forall d \in c)(b \in d \wedge \mathrm{Tran}[d] \wedge \mathsf{Ad}(d) \to (\exists x \in d)(\exists y \in c)(P((a)_1, x, y) \wedge \neg y \in d))$.

Table 3: Disjuncts of formula $\mathfrak{A}$

Notice that P occurs only positively in all the clauses and that all of them are Σ formulae. Therefore we can apply Theorem 102: Let $C_{\mathfrak{A}}$ be the formula due to this theorem with respect to $\mathfrak{A}$ and $Ap_{\mathfrak{A}}$ the formula $(\exists f, \alpha)(C_{\mathfrak{A}}[f, \alpha] \wedge \langle a, b, c \rangle \in f'\alpha)$. The formulae $Ap_{\mathfrak{A}_{\mathbb{B}}}$, $Ap_{\mathfrak{A}_{\mathbb{P}}}$ and $Ap_{\mathfrak{A}_{\mathbb{A}}}$ are defined analogously.

Lemma 106. KP_r proves (i) $\exists y Ap_{\mathfrak{A}}[\widehat{\mathbb{K}}, b, y]$ and (ii) $Ap_{\mathfrak{A}}[\widehat{\mathbb{K}}, b, c] \wedge Ap_{\mathfrak{A}_{\mathbb{B}}}[\widehat{\mathbb{K}}, b, d] \to c = d$.

PROOF. By Lemma 26, $Ap_{\mathfrak{A}}[\widehat{\mathbb{K}}, b, c]$ codes nothing else than that $c = \{s : (\exists n \in \omega)(\mathrm{Fun}[s] \wedge \mathrm{Dom}[s, n] \wedge \mathrm{Ran}[s, b])\}$. Therefore both assertions follow from the lemma. □

Lemma 107. KP_r proves (i) $\exists y Ap_{\mathfrak{A}}[\widehat{\mathbb{T}}, b, y]$ and (ii) $Ap_{\mathfrak{A}}[\widehat{\mathbb{T}}, b, c] \wedge Ap_{\mathfrak{A}}[\widehat{\mathbb{T}}, b, d] \to c = d$.

PROOF. Similarly to the last lemma, this follows from Lemma 28. □

Lemma 108. KP_r+(Beta) proves (i) $\exists y Ap_{\mathfrak{A}_{\mathbb{B}}}[\langle \widehat{\mathbb{B}}, a\rangle, b, y]$ and (ii) $Ap_{\mathfrak{A}_{\mathbb{B}}}[\langle \widehat{\mathbb{B}}, a\rangle, b, c] \wedge Ap_{\mathfrak{A}_{\mathbb{B}}}[\langle \widehat{\mathbb{B}}, a\rangle, b, d] \to c = d$.

PROOF. By Lemma 31 (iii), $Ap_{\mathfrak{A}_{\mathbb{B}}}[\langle \widehat{\mathbb{B}}, a\rangle, b, c]$ codes nothing else than that there is an x such that $x = \{y \in a : \mathrm{WP}[y, a, b]\}$ and c is a function with domain x with $(\forall y \in x)(c'y = \{c'z : z \in x \wedge \langle z, y\rangle \in b\})$. Therefore both assertions follow from Lemma 31 (iii) and (iv). □

Lemma 109. $\mathsf{KP}_r + (\mathcal{AD}) + (\mathbb{A})$ proves (i) $\exists y Ap_{\mathfrak{A}_{\mathbb{A}}}[\widehat{\mathbb{A}}, b, y]$ and (ii) $Ap_{\mathfrak{A}_{\mathbb{A}}}[\widehat{\mathbb{A}}, b, c] \wedge Ap_{\mathfrak{A}_{\mathbb{A}}}[\widehat{\mathbb{A}}, b, d] \to c = d$.

PROOF. Similarly to the previous lemmata, this follows from Lemma 24. □

Lemma 110. $\mathsf{KP}_r + (\mathcal{AD}) + (\Pi_2\text{-}\mathsf{Ref})$ proves

(i) $\forall u \exists v Ap_{\mathfrak{A}_{\mathbb{M}}}(f, u, v) \to \exists y Ap_{\mathfrak{A}_{\mathbb{M}}}[\langle \widehat{\mathbb{M}}, f\rangle, x, y]$ and
(ii) $\forall u \exists v Ap_{\mathfrak{A}_{\mathbb{M}}}(f, u, v) \to (Ap_{\mathfrak{A}_{\mathbb{M}}}[\langle \widehat{\mathbb{M}}, f\rangle, b, c] \wedge Ap_{\mathfrak{A}_{\mathbb{M}}}[\langle \widehat{\mathbb{M}}, f\rangle, b, d] \to c = d)$.

PROOF. Assume the antecedent $\forall u \exists v Ap_{\mathfrak{A}_{\mathbb{M}}}(f, u, v)$. By applying Lemma 25 to $A := \forall u \exists v Ap_{\mathfrak{A}_{\mathbb{M}}}(f, u, v)$, we can prove the two succedents. □

The complete description of the fixed point predicate given in Theorem 102 is essential in the following lemma, where the $\in$-induction is restricted.

Lemma 111. The theory KP_r proves

(i) $(C_{\mathfrak{A}}[f, \alpha] \wedge \langle a, b, c\rangle \in f'\alpha) \wedge (C_{\mathfrak{A}}[g, \beta] \wedge \langle a, b, d\rangle \in g'\beta) \to c = d$,
(ii) $Ap_{\mathfrak{A}}[a, b, c] \wedge Ap_{\mathfrak{A}}[a, b, d] \to c = d$.

The analogous assertions for $Ap_{\mathfrak{A}_{\mathbb{B}}}$, $Ap_{\mathfrak{A}_{\mathbb{P}}}$, $Ap_{\mathfrak{A}_{\mathbb{A}}}$ and $Ap_{\mathfrak{A}_{\mathbb{M}}}$ and the respective theories hold too.

PROOF. For the first assertion, let us assume $C_{\mathfrak{A}}[f, \alpha]$ and $C_{\mathfrak{A}}[g, \beta]$. We can prove

$$(\forall x \in f'\gamma)(\forall y \in g'\delta)((x)_0 = (y)_0 \wedge (x)_1 = (y)_1 \to (x)_2 = (y)_2))$$

by double Δ_0 $\in$-induction on γ and δ. The second assertion follows immediately from the first one.

We can prove the analogous assertions in the same way. □

As we can see above, by the description of the least fixed point of the positive Σ operator given in Theorem 102, the *local induction* (i.e., $\in$-induction applied the formula on fixed f) suffices to show the uniqueness of the application. If we have full $\in$-induction, we do not need the description given in Theorem 102 and that given in Theorem 101 is enough as in Zumbrunnen [59, p.88]. In terms of our description, what he did there can be analysed as follows: he applied $\in$-induction to

$$(\forall x, y)((\exists f)(C_{\mathfrak{A}}[f, \alpha] \wedge x \in f'\alpha) \wedge (\exists g)(C_{\mathfrak{A}}[g, \beta] \wedge x \in g'\beta) \to ((x)_0 = (y)_0 \wedge (x)_1 = (y)_1 \to (x)_2 = (y)_2)),$$

and therefore $\in$-induction for Π_1 was needed.

9.4. Definition of interpretation

Using the formula $Ap_{\mathfrak{A}}$ we can define as in Jäger [29] and Jäger and Zumbrunnen [38] for each $\mathcal{L}_{\in}^{\circ}$ term t a Σ formula $[\![t]\!]_{\mathfrak{A}}[u]$ of $\mathcal{L}_{\in}$ expressing that t has the value u via the formula $Ap_{\mathfrak{A}}$ (and analogous for $Ap_{\mathfrak{A}_{\mathbb{B}}}$, $Ap_{\mathfrak{A}_{\mathbb{P}}}$, $Ap_{\mathfrak{A}_{\mathbb{A}}}$ and $Ap_{\mathfrak{A}_{\mathbb{M}}}$).

Definition 112 (Formula $[\![t]\!]_{\mathfrak{A}}$)**.** Let t be an $\mathcal{L}_{\in}^{\circ}$ term with u not occurring in t. We define the Σ formula $[\![t]\!]_{\mathfrak{A}}[u]$ inductively as follows:

(i) if t is a variable or the constant ω, then $[\![t]\!]_{\mathfrak{A}}[u]$ is the formula $(t = u)$;
(ii) if t is another constant, then $[\![t]\!]_{\mathfrak{A}}[u]$ is the formula $(\widehat{t} = u)$;
(iii) if t is the term (rs), then we set

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

The formulae $[\![t]\!]_{\mathfrak{A}_{\mathbb{B}}}$, $[\![t]\!]_{\mathfrak{A}_{\mathbb{P}}}$ (that is an $\mathcal{L}_{\mathcal{P}}$ formula) and $[\![t]\!]_{\mathfrak{A}_{\mathbb{A}}}$ and $[\![t]\!]_{\mathfrak{A}_{\mathbb{M}}}$ (that are $\mathcal{L}_{\mathsf{Ad}}$ formulae) are defined analogously.

Now we can translate $\mathcal{L}_{\in}^{\circ}$ formulae to $\mathcal{L}_{\in}$ formulae.

Definition 113 ($\star$-translations of $\mathcal{L}_{\in}^{\circ}$ formulae)**.** Let A be a formula of $\mathcal{L}_{\in}^{\circ}$. The $\mathcal{L}$ formula $A^{\star}$ is inductively defined as follows:

(i) for the atomic formulae of $\mathcal{L}_{\in}^{\circ}$ we set

$$\begin{aligned} \bot^{\star} &:= \bot, \\ (t{\downarrow})^{\star} &:= \exists x [\![t]\!]_{\mathfrak{A}}[x] \text{ and} \\ (s = t)^{\star} &:= \exists x \exists y([\![s]\!]_{\mathfrak{A}}(x) \wedge [\![t]\!]_{\mathfrak{A}}(y) \wedge x = y), \\ (s \in t)^{\star} &:= \exists x \exists y([\![s]\!]_{\mathfrak{A}}[x] \wedge [\![t]\!]_{\mathfrak{A}}[y] \wedge x \in y); \end{aligned}$$

(ii) if A is the formula $\neg B$, then $A^{\star}$ is $\neg B^{\star}$;
(iii) if A is the formula $(B \wedge C)$, $(B \vee C)$ or $(B \to C)$, then $A^{\star}$ is $(B^{\star} \wedge C^{\star})$, $(B^{\star} \vee C^{\star})$ or $(B^{\star} \to C^{\star})$, respectively;
(iv) if A is the formula $\forall x B[x]$ or $\exists x B[x]$, then $A^{\star}$ is $\forall x B^{\star}[x]$ or $\exists x B^{\star}[x]$, respectively.

The translations $A^{\star\mathbb{B}}$, $A^{\star\mathbb{P}}$ (that is an $\mathcal{L}_{\mathcal{P}}$ formula) and $A^{\star\mathbb{A}}$ and $A^{\star\mathbb{M}}$ (that are $\mathcal{L}_{\mathsf{Ad}}$ formulae) are defined analogously, but we use $[\![t]\!]_{\mathfrak{A}_{\mathbb{B}}}$, $[\![t]\!]_{\mathfrak{A}_{\mathbb{P}}}$, $[\![t]\!]_{\mathfrak{A}_{\mathbb{A}}}$ and $[\![t]\!]_{\mathfrak{A}_{\mathbb{M}}}$, respectively, for the translations of the atomic formulae where the formula $\mathsf{Ad}(t)^{\star\mathbb{A}}$ is defined by $\exists x([\![t]\!]_{\mathfrak{A}_{\mathbb{A}}}[x] \wedge \mathsf{Ad}(x))$

9.5. Interpretability result

Theorem 114. Let A be an arbitrary formula of $\mathcal{L}_{\in}^{\circ}$. Then

(i) if OST_r^- proves A, then KP_r proves $A^{\star}$;
(ii) if $\mathsf{OST}_r^- + (\mathbb{B})$ proves A, then $\mathsf{KP}_r + (\text{Beta})$ proves $A^{\star\mathbb{B}}$;
(iii) if $\mathsf{OST}_r^- + (\mathbb{P})$ proves A, then $\mathsf{KP}_r + (\mathcal{P})$ proves $A^{\star\mathbb{P}}$;
(iv) if $\mathsf{OST}_r^- + \mathcal{AD} + (\mathbb{A})$ proves A, then KPI_r proves $A^{\star\mathbb{A}}$;
(v) if $\mathsf{OST}_r^- + \mathcal{AD} + (\mathbb{M})$ proves A, then KPM_r proves $A^{\star\mathbb{M}}$.

All the assertions, with the subscript r replaced by w or just omitted, also hold.

Furthermore, the analogous assertions hold if (Σ_1-Ind) or (Σ_1-Ind_{ω}) is added to the variants of KP and at the same time (opInd) or (opInd_{ω}), respectively, is added to the variants of OST^-.

PROOF. The assertions follows because the particular $\star$-translated axioms are provable in the respective theories. That the translations of the axioms of the logic of partial terms are provable is straightforward if we can use the previous lemma. That also the $\star$-translated axioms about the operations $\mathbb{K}$, $\mathbb{T}$, $\mathbb{D}$ and $\mathbb{U}$ are provable in KP_r is obvious. We can prove that all five possible $\star$-translations of all the other axioms of OST^-, $\mathsf{OST}^- + (\mathbb{B})$, $\mathsf{OST}^- + (\mathbb{P})$, $\mathsf{OST}^- + (\mathcal{AD}) + (\mathbb{A})$ and $\mathsf{OST}^- + (\mathcal{AD}) + (\mathbb{M})$ are provable in KP, $\mathsf{KP} + (\mathrm{Beta})$, $\mathsf{KP} + (\mathcal{P})$, KPI and KPM, respectively, by using Theorem 102 (i). Particularly, if A is the axiom $(\mathbb{B})$, then $A^{\star\mathbb{B}}$ is provable in $\mathsf{KP} + (\mathrm{Beta})$ by the Lemmata 31 and 108. □

We do not know if OST_0^- or OST_ω^- can be interpreted *directly* in KP_0 or KP_ω respectively.

For the reader's convenience, the correspondence is summarised in Table 4, where the system in the upper lines augmented by any combination of the axioms on the upper line is, by $\star$ and variants, interpreted in the system below on the same column augmented by the axioms below on the same column,

WEST_r	WEST_w	WEST
OST_r^-	OST_w^-	OST^-
KP_r	KP_w	KP

added by any of

(opInd_ω)	(opInd)	$(\mathbb{B})$	$(\mathbb{P})$	$(\mathcal{AD}) + (\mathbb{A})$	$(\mathcal{AD}) + (\mathbb{M})$
$(\Sigma_1\text{-}\mathsf{Ind}_\omega)$	$(\Sigma_1\text{-}\mathsf{Ind})$	(Beta)	$(\mathcal{P})$	$(\mathcal{AD}) + (\mathsf{Lim})$	$(\mathcal{AD}) + (\Pi_2\text{-}\mathsf{Ref})$

Table 4: Correspondence between systems and axioms 2

10. Conclusion

Since the $\star$-translation of an $\mathcal{L}_\in$ formula A is always equivalent to A, we can compound Theorem 114 with Theorem 99 and we get the following.

Corollary 115. The following theories prove in each case the same Π_1 sentences of $\mathcal{L}_\in$:

(i) OST_r^-, WEST_r and KP_r.
(ii) $\mathsf{OST}_r^- + (\mathbb{B})$, $\mathsf{WEST}_r + (\mathbb{B})$ and $\mathsf{KP}_r + (\mathrm{Beta})$.
(iii) $\mathsf{OST}_r^- + (\mathbb{P})$, $\mathsf{WEST}_r + (\mathbb{P})$ and $\mathsf{KP}_r + (\mathcal{P})$.
(iv) $\mathsf{OST}_r^- + (\mathcal{AD}) + (\mathbb{A})$, $\mathsf{WEST}_r + (\mathcal{AD}) + (\mathbb{A})$ and KPI_r.
(v) $\mathsf{OST}_r^- + (\mathcal{AD}) + (\mathbb{M})$, $\mathsf{WEST}_r + (\mathcal{AD}) + (\mathbb{M})$ and KPM_r.

All the assertions, with the subscript r replaced by w or just omitted, also hold.

Furthermore, the analogous assertions hold if $(\Sigma_1\text{-}\mathsf{Ind})$ or $(\Sigma_1\text{-}\mathsf{Ind}_\omega)$ is added to the variants of KP and at the same time (opInd) or (opInd_ω), respectively, is added to the variants of OST^- and of WEST.

It is easy to see that the third line can be generalised as Remark 100: in the same situation as Remark 100, for any Π_1 formula A of $\mathcal{L}_\in$, the following are equivalent:

- $\mathsf{OST}^- + \mathcal{Q} + (\forall x Q(x, \mathbb{Q}(x)))$ proves A;
- $\mathsf{WEST} + \mathcal{Q} + (\forall x Q(x, \mathbb{Q}(x)))$ proves A;
- $\mathsf{KP} + \mathcal{Q} + (\forall x \exists y Q(x, y))$ proves A.

Appendix

AppendixA. Modifications and Difficulties for Intuitionistic Applicative Set Theories

The interpretation of KP in WEST and therefore in OST^-, we gave in Sections 3-5, goes through intuitionistic systems. Furthermore, the original motivation of operational set theory is, at least ultimately, to provide a unified framework both for classical and constructive set theories (especially large cardinal notions) and, therefore, intuitionistic versions are more important for this purpose. Therefore it is natural to wonder if our method can be used for the lower bound proof of intuitionistic versions of operational set theories.

However, we used, in several places, the fact that the final interpreting theory is based on classical logic. Most significantly, in order to interpret KP by negative interpretation, we need (N-Ext), the negative interpretation of extensionality (Theorem 95), which is not usually accepted as a *proper* axiom in the intuitionistic context. Similarly, we use (N-infinity) to interpret the axiom of infinity by negative interpretation. In the previous sections, these intuitionistically unnatural axioms are employed only for technical purposes and, indeed, finally interpreted in the final interpreting theory that is based on classical logic. Moreover, the proof of the general result, Theorem 95, heavily relies on the derivability of $\mathcal{B}^{\exists \mathrm{s}}$ from $\mathcal{A}^{\exists \mathrm{s}}$, which does not generally holds in intuitionistic contexts. Therefore, to have a similar result for intuitionistic versions of applicative set theories, we need a significant amount of modifications, to which this appendix is devoted.

Quite interestingly, we will see how the things were less complicated in the classical case even though we worked through intuitionistic settings. After one sees how restricted the results we can obtain for intuitionistic theories are, he or she might conclude that our approach is only for classical theories although it forces us to deal with intuitionistic theories in the process. This could be said a mystery on our approach.

Whereas some versions of intuitionistic operational set theory were considered in Cantini [8] and Cantini and Crosilla [9, 10, 11], here we concentrate on the intuitionistic versions of weak explicit set theory, in order to avoid a formulation-question.

AppendixA.1. Interpreting KP *in* KP^{int}*: Bisimulation interpretation*

First, we consider how to omit (N-Ext) when we use the negative interpretation. This problem is very historical in the sense that Friedman [19] gave a solution when he interpreted classical Zermelo-Fraenkel set theory in intuitionistic Zermelo-Fraenkel set theory with the negative interpretation. The idea is to combine the negative interpretation with another interpretation, which is of a classical set theory in another classical set theory without extensionality (or an *intensional* classical set theory). We follow this solution, namely, we do not interpret the axiom of extensionality directly by the negative interpretation but combine it with the interpretation of KP in KP^{int}, so-called *bisimulation interpretation* (which was investigated extensively but in slightly different situations by Sato [48, 49] and was used in applicative and intuitionistic settings by Gordeev [21, 22, 23].) However, we have to pay some cost: we need (Σ_1-Ind) on the interpreting side.

The presentation below follows Avigad [2, Section 4]. We will write $x = \{y, z\}$ and $x = \cup y$ also if we work with theories without extensionality. We do this although the x in these abbreviations might not be unique. Namely, this abbreviations mean "x is some set containing exactly y and z" and "x is some set corresponding to the union of y", respectively. Remark immediately before Subsection 6.1 applies to this subsection.

Definition 116. We introduce the following abbreviations for Δ_0 formulae:

$$\begin{aligned} y \sim_a z &:= (\exists x \in a)(x = \{y, z\}), \\ y \in \mathrm{field}(a) &= (\exists x \in a)(\exists z \in x)(x = \{y, z\}), \\ \mathrm{Bis}[a] &:= \forall y, z(y, z \in \mathrm{field}(a) \to (y \sim_a z \leftrightarrow (\forall u \in y)(\exists v \in z)(u \sim_a v) \land (\forall v \in z)(\exists u \in y)(v \sim_a u))). \end{aligned}$$

We also introduce the following abbreviations for Σ_1 formulae:

$$y \sim z := \exists a(\mathrm{Bis}[a] \land y \sim_a z) \qquad\qquad y \in^* x := (\exists z \in x)(y \sim z).$$

The next lemma is proved exactly as Lemma 4.5 in Avigad [2].

Lemma 117. $\mathsf{KP}_0^{int} + (\Sigma_1\text{-}\mathsf{Ind})$ proves that

$$y \sim z \text{ is equivalent to } \forall a(\mathrm{Bis}[a] \wedge y \in \mathrm{field}(a) \wedge z \in \mathrm{field}(a) \to y \sim_a z),$$

and, more general,

$$(\forall y, z \in b)(\ y \sim z \ \leftrightarrow\ \forall a(\mathrm{Bis}[a] \wedge b \subseteq \mathrm{field}(a) \to y \sim_a z)\ ).$$

In the next definition we try to give a sufficient condition on relations R for $\sim$ behaving as an equality relation. The simplest way to do so would be to take the following condition:

$$\bigwedge_{i=0}^{n-1} (x_i \sim x_i') \ \to\ (R(x_0, ..., x_{n-1}) \leftrightarrow R(x_0', ..., x_{n-1}')). \tag{A.1}$$

But if we did so, we would elude the powerset relation $\mathcal{P}$: for $z \neq z'$ but $z \sim z'$, the axiom $(\mathcal{P})$ yields some y that $\mathcal{P}(\{z\}, y)$ and $\neg\mathcal{P}(\{z'\}, y)$ (because $\{z'\}$ is not a subset of $\{z\}$), whereas it proves $\{z\} \sim \{z'\}$ as well as $y \sim y$. In order to designate also the powerset relation, we chose the rather complicated way in the next definition. Notice that in the case $n > 0$ in this definition we demand special properties from the last position of the relation. Just as well we could demand these properties from any other position of the relation (but then we would also have to adjust the case (iii) of Definition 119).

Definition 118 (Bis_R)**.** Let R be an $(n{+}1)$-ary relation symbol of $\mathcal{L}$ and $\vec{x} = x_0, ..., x_{n-1}$; $\vec{x}' = x_0', ..., x_{n-1}'$; y and y' variables. We define a formula Bis_R as follows:

- if $n = 0$ (and therefore R is a unary relation symbol),

$$\mathrm{Bis}_R \ :=\ \forall y, y'(y \sim y' \ \to\ (R(y) \leftrightarrow R(y'))),$$

- or if $n > 0$, Bis_R is the conjunction of
  - (i) $\exists y'(y \sim y' \wedge R(\vec{x}, y')) \ \leftrightarrow \forall y'(R(\vec{x}, y') \to y \sim y')$ as well as
  - (ii) $\bigwedge_{i=0}^{n-1} (x_i \sim x_i') \wedge R(\vec{x}, y) \ \to\ \exists y'(R(\vec{x}', y') \wedge y \sim y')$.

For a set $\mathcal{R}$ of predicate symbols of $\mathcal{L}$, $\mathrm{Bis}_{\mathcal{R}}$ is the set $\{\mathrm{Bis}_R : R \in \mathcal{R}\}$.

An $(n{+}2)$-ary relation R can be seen, in the presence of Bis_R, as the graph of an $(n{+}1)$-ary function up to $\sim$.

Definition 119 (Formula A^*)**.** For any $\mathcal{L}$ formula A we write A^* for the $\mathcal{L}$ formula which we get if we replace in A

- (i) every occurrence of the form $x = y$ by $x \sim y$,
- (ii) every occurrence of the form $x \in y$ by $x \in^* y$ and
- (iii) every occurrence of the form $R(\vec{x}, y)$ by $\exists y'(R(\vec{x}, y') \wedge y \sim y')$ for any relation symbol R of arity two or more other than $=$ and $\in$, where $\vec{x}$ is a string of variables of the correct length.

Furthermore, if $\mathcal{A}$ is a set of $\mathcal{L}$ formulae, we write $\mathcal{A}^*$ for the set $\{A^* : A \in \mathcal{A}\}$.

Notice that $(R(x))^*$ is $R(x)$ for unary relation symbols R. The condition (i) of Definition 118 guarantees that $(R(\vec{x}, y))^*$ is a Δ_1 formula with respect to a theory containing Bis_R. Therefore we can prove the next lemma and theorem as the corresponding assertions are proved in Avigad [2, Section 4].

Lemma 120. Let $\mathcal{R}$ be a set of relational symbols of $\mathcal{R}$. For an $\mathcal{L}_{\mathcal{R}}$ formula A,

- (i) $\mathsf{KP}_0^{int} + (\Sigma_1\text{-}\mathsf{Ind})$ proves $(\forall x \in^* y)A^* \leftrightarrow (\forall x \in y)A^*$ and $(\exists x \in^* y)A^* \leftrightarrow (\exists x \in y)A^*$;

(ii) if A is Δ_0, then A^* is Δ_1 with respect to $\mathsf{KP}_0^{int} + (\Sigma_1\text{-}\mathsf{Ind}) + (\mathrm{Bis}_{\mathcal{R}})$.

Theorem 121. For a set $\mathcal{R}$ of relational symbols of $\mathcal{L}$, a set $\mathcal{A}$ of $\mathcal{L}_{\mathcal{R}}$ formulae and an $\mathcal{L}_{\mathcal{R}}$ formula B,

(i) if B is provable in $\mathsf{KP}_r + (\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{A}$, then B^* is provable in $\mathsf{KP}_r^{int} + (\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{A}^* + (\mathrm{Bis}_{\mathcal{R}})$;
(ii) if B is provable in $\mathsf{KP}_w + (\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{A}$, then B^* is provable in $\mathsf{KP}_w^{int} + (\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{A}^* + (\mathrm{Bis}_{\mathcal{R}})$;
(iii) if B is provable in $\mathsf{KP} + \mathcal{A}$, then B^* is provable in $\mathsf{KP}^{int} + \mathcal{A}^* + (\mathrm{Bis}_{\mathcal{R}})$.

Notice that the $*$-interpretations of the equality axioms are provable with $\mathrm{Bis}_{\mathcal{R}}$ if we work only with relation symbols R with $R \in \mathcal{R}$.

As an example, $\mathsf{KP}_0 + (\mathcal{P})$ proves $\mathrm{Bis}_{\mathcal{P}}$. Unfortunately, however $\mathsf{KP}^{int} + (\mathcal{AD})$ does not prove $\mathrm{Bis}_{\mathsf{Ad}}$. But, as in Remark 23, we could define a truth predicate such that we can express $x \vDash \mathsf{KP}$ (x is a model of KP) in $\mathcal{L}_\in$, and then define a new translation of $\mathcal{L}_{\mathsf{Ad}}$ formulae to $\mathcal{L}_\in$ formulae by replacing $\mathsf{Ad}(x)$ by the formula $(x \vDash \mathsf{KP})^* \wedge \mathrm{Tran}[x]$. Then we can interpret (AD.i) and (AD.ii) in KP_0^{int}. However, we have to give up (AD.iii) the trichotomy for admissible sets. Intuitionistically or constructively, it is well motivated to remove the trichotomy.

AppendixA.2. Modification of axioms

In order to make IKP^- be realisable in intuitionistic versions of our applicative set theory, we need at least a closed term s_A satisfying $s_A(x) = \{u \in x : A[u, x]\}$ for each negative Δ_0 formula A. In the context of classical logic, our set of operations is enough. However, in order to have this result in the intuitionistic context, we need more operators, because we have less equivalence between formulae. We can do this by the following procedures.

In OST^-, we can do this by the way employed in Cantini and Crosilla [9, 10], where t and f are identified with $\{\emptyset\}$ and $\emptyset$ respectively and where the subsets of $\{\emptyset\}$ are considered as truth values. In WEST, we can do this by introducing new constants for Gödel operations modified for the intuitionistic context and adding the axioms for such constants. We do not go further into this issue. Instead, we define the intuitionistic version of weak explicit set theory by adding new constants for all negative Δ_0 formulae.

Definition 122. The language of IWEST, is the language $\mathcal{L}_\in^\circ$ extended by the constants $\mathbb{G}_A$ for all negative Δ_0 formulae $A[x, a, \vec{v}]$ of $\mathcal{L}_\in$ with at most $x, a, \vec{v}$ free.

The logic of IWEST is the *intuitionistic* logic of partial terms. The non-logical axioms of IWEST are $(I.1)$-$(I.3)$, $(II.1)$-$(II.3)$, $(IV.1)$-$(IV.4)$, $(IV.6)$ and the following:

$$(\forall a, \vec{v})(\mathbb{G}_A(a, \vec{v})\downarrow \wedge\, \mathbb{G}_A(a, \vec{v}) = \{x \in a : A[x, a, \vec{v}]\}).$$

The theory IWEST^{int} is the theory IWEST without extensionality but with $\mathbb{U}(\{x\}) = x$. IWEST_0, IWEST_ω, IWEST_r, IWEST_w, IWEST_0^{int}, $\mathsf{IWEST}_\omega^{int}$, IWEST_r^{int} and IWEST_w^{int} are defined analogously.

Remark 123. Since we need $(\Sigma_1\text{-}\mathsf{Ind})$ for bisimulation interpretation, our final results will be on the variants of KP with $(\Sigma_1\text{-}\mathsf{Ind})$. For this purpose, we can remove (KP.8) from KP and replace (KP.7) by

$$\mathrm{Ind}[\omega] \wedge (\forall x \in \omega)(\mathrm{zero}[x] \vee (\exists y \in \omega)\mathrm{succ}[y, x]),$$

and change N-infinity in IKP^- accordingly. Actually $(\Sigma_1\text{-}\mathsf{Ind}_\omega)$ is enough to show the equivalence.

Moreover, since we do not have Lemma 19, we can no longer assert that (opInd) corresponds to $(\Sigma_1\text{-}\mathsf{Ind})$. Instead we use the induction schema for the following class of formulae.

Definition 124. An $\mathcal{L}^\circ$ formula A is called a $\Sigma_1(\mathsf{app}^+)$ formula, if it is of the form $\exists x B[x]$ where $B[x]$ has no unbounded quantifiers and all the occurrences of the function symbol $\circ$ and the predicate symbol $\downarrow$ in $B[x]$ are in positive positions. $(\Sigma_1(\mathsf{app}^+)\text{-}\mathsf{Ind})$ consists of all instances of $\in$-induction for $\Sigma_1(\mathsf{app}^+)$ formulae.

An $\mathcal{L}^\circ$ formula is called a $\Pi_2(\mathsf{app}^+)$ formula, if it is of the form $\forall x A[x]$ for some $\Sigma_1(\mathsf{app}^+)$ formula $A[x]$. $(\Pi_2(\mathsf{app}^+)\text{-}\mathsf{Ref})$ is defined analogously.

As we saw in the last subsection, to make $*$-interpretation work, we have to give up trichotomy for admissible sets. We introduce the theories based on non-linear admissibles.

Definition 125. $(\mathcal{AD}')$ consists of (AD.i) and (AD.ii). KPI' and KPM' are the theories $\mathsf{KP} + (\mathcal{AD}')$ augmented by (Lim) and $(\Pi_2\text{-}\mathsf{Ref})$ respectively.

AppendixA.3. Interpretability result

As in the proof of Theorem 93 with Remark 94, we have the following theorem.

Theorem 126. Let $\mathcal{A}$ be a set of $\mathcal{L}$ formulae of the form $\forall x_0 \exists y_0 ... \forall x_n \exists y_n B$, where B is a negative $\mathcal{L}$ formula. If A is a negative $\mathcal{L}$ formula provable in $\mathsf{IKP}_0^- + (\mathsf{n}\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{A}$, then A is also provable in $\mathsf{IWEST}_0^{int} + (\Sigma_1(\mathsf{app}^+)\text{-}\mathsf{Ind}) + \mathcal{A}^{\exists \mathrm{s}}$.

Moreover, the subscript 0 can be replaced by ω or just omitted. The same assertion holds if $(\mathsf{n}\Pi_2\text{-}\mathsf{Ref})$ is added to the variants of IKP^- and at the same time $(\Pi_2(\mathsf{app}^+)\text{-}\mathsf{Ref})$ is added to those of IWEST.

Proof. It is easy to see that the proof of Lemma 90 goes through with $(\Sigma_1(\mathsf{app}^+)\text{-}\mathsf{Ind})$ instead of (opInd). By Remark 94, it suffices to show how the axiom of N-infinity is realised in IWEST_0^{int}. Since Lemma 81 holds in the intuitionistic setting, it remains to show $\mathrm{Ind}[\omega]^N \wedge (\forall x \in \omega)(\mathrm{zero}[x] \vee (\exists y \in \omega)\mathrm{succ}[y, x])^N$ from the axiom of infinity in IWEST_0^{int} by Remark 123.

Now $\mathrm{zero}[y]$, that is $(\forall z \in y)\bot$, is equivalent to its negative interpretation $(\mathrm{zero}[y])^N$, and $\mathrm{succ}[y, z]$ implies $\mathrm{succ}[y, z]^N$, since $\mathrm{succ}[y, z]$ is the conjunction of the formulae $y \in z$ and $(\forall u \in y)(u \in z)$ and $(\forall u \in z)(u \in y \vee u = y)$ where $(u \in y \vee u = y)$ implies $(u \in y \vee u = y)^N$. Thus $\mathrm{Ind}[\omega]$ implies $\mathrm{Ind}[\omega]^N$, and also $(\forall x \in \omega)(\mathrm{zero}[x] \vee (\exists y \in \omega)\mathrm{succ}[y, x])$ implies $(\forall x \in \omega)(\mathrm{zero}[x] \vee (\exists y \in \omega)\mathrm{succ}[y, x])^N$. □

Theorem 127. Let $\mathcal{R}$ be a set of relation symbols of $\mathcal{L}$ and $\mathcal{A}$ a set of $\mathcal{L}_{\mathcal{R}}$ formulae. Define

$$\mathcal{B} = \{\forall x_0 \exists y_0 ... \forall x_n \exists y_n A^N : A \in \Delta_0 \text{ and } \forall x_0 \exists y_0 ... \forall x_n \exists y_n A \text{ is provable in } \mathsf{KP}_0^{int} + (\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{A}\},$$

Assume $\mathcal{A}$ implies $\mathcal{A}^*$ and $\mathrm{Bis}_{\mathcal{R}}$ over $\mathsf{KP}_0^{int} + (\Sigma_1\text{-}\mathsf{Ind})$. Then for a Π_1 formula C of $\mathcal{L}_{\mathcal{R}}$,

(i) if $\mathsf{KP}_r + (\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{A}$ proves C, then $\mathsf{IWEST}_r^{int} + (\Sigma_1(\mathsf{app}^+)\text{-}\mathsf{Ind}) + \mathcal{B}^{\exists \mathrm{s}}$ proves $(C^*)^N$;
(ii) if $\mathsf{KP}_w + (\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{A}$ proves C, then $\mathsf{IWEST}_w^{int} + (\Sigma_1(\mathsf{app}^+)\text{-}\mathsf{Ind}) + \mathcal{B}^{\exists \mathrm{s}}$ proves $(C^*)^N$;
(iii) if $\mathsf{KP} + \mathcal{A}$ proves C, then $\mathsf{IWEST}^{int} + \mathcal{B}^{\exists \mathrm{s}}$ proves $(C^*)^N$.

The same assertions hold, if we add $(\Pi_2\text{-}\mathsf{Ref})$ to the variants of KP and at the same time $(\mathbb{M})$ to those of IWEST^{int}.

Proof. By assumption and by Theorem 121, if C is provable in $\mathsf{KP}_r + (\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{A}$, then C^* is provable in $\mathsf{KP}_r^{int} + (\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{A}^* + \mathrm{Bis}_{\mathcal{R}}$ (equivalently $\mathsf{KP}_0^{int} + (\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{A}^* + \mathrm{Bis}_{\mathcal{R}}$). By the assumption on $\mathcal{A}$, C^* is provable in $\mathsf{KP}_r^{int} + (\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{A}$. Now C^* is provable in $\mathsf{KP}_r^{int} + (\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{B}$, since classically $\mathcal{A}$ is equivalent to $\mathcal{A}^N$ and to $\mathcal{B}$. We may assume that C^* is a Π_1 formula by Lemma 120 (ii). Since $(C^*)^N$ is in $\mathcal{C}_{res}$, and since each element B of $\mathcal{B}$ satisfies $B \to B^N$ trivially, by Theorem 50, $(C^*)^N$ is provable in $\mathsf{IKP}^\sharp + (\Delta_0^{\mathrm{s}-}\text{-}\mathsf{MP}) + (\mathsf{w}\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{B}$. Since all the elements of $\mathcal{B}$ and $(C^*)^N$ are in $\mathcal{D}_{res}$, by Theorem 75, $(C^*)^N$ is provable in $\mathsf{IKP}^- + (\mathsf{n}\Sigma_1\text{-}\mathsf{Ind}) + \mathcal{B}$. By Theorem 126, $(C^*)^N$ is provable in $\mathsf{IWEST}_0^{int} + (\Sigma_1(\mathsf{app}^+)\text{-}\mathsf{Ind}) + \mathcal{B}^{\exists \mathrm{s}}$.

The analogous assertions are proved similarly. □

Note that all the results here are obtained for *intensional* intuitionistic applicative set theories.

AppendixA.4. Difficulties and restricted results for intuitionistic axioms

We have obtained a general interpretability result, Theorem 127, in the last subsection. However, this theorem is not helpful for lower bound proof of intuitionistic applicative set theories, except in the case of $\mathcal{A} = \emptyset$. The reason is the difference between $\mathcal{A}$ and $\mathcal{B}$, in terms of Theorem 127. In order to make the theorem yield the interpretability result for $\mathsf{IWEST} + \mathcal{A}^{\exists \mathrm{s}}$, $\mathcal{A}^{\exists \mathrm{s}}$ must imply $\mathcal{B}^{\exists \mathrm{s}}$ in IWEST, since $\mathcal{B}^{\exists \mathrm{s}}$, which is defined from $\mathcal{A}^{\exists \mathrm{s}}$, is sometimes quite unnatural. For classical theories, this is not a matter because classically $\mathcal{B}$ is equivalent to $\mathcal{A}$. However, this becomes an essential matter in the intuitionistic context. Let us see this more closely in examples.

The axiom (Beta′) requires, for a given binary relational structure, the existence of the accessible part and of Mostowski collapsing function on that part. For $\mathcal{A} = \{(\mathrm{Beta}')\}$,

$$\mathcal{A}^{\mathrm{s}}[f, g] = \{(\forall a, r)(\mathrm{Fun}[f(a, r)] \wedge \mathrm{Dom}[f(a, r), g(a, r)] \wedge ... \wedge \mathrm{Prog}[g(a, r), a, r] \wedge ...)\};$$
$$\mathcal{B}^{\mathrm{s}}[f, g] = \{(\forall a, r)(\mathrm{Fun}^N[f(a, r)] \wedge \mathrm{Dom}^N[f(a, r), g(a, r)] \wedge ... \wedge \mathrm{Prog}^N[g(a, r), a, r] \wedge ...)\}.$$

Here, $\mathrm{Fun}[f(a,r)]$, together with the required property of f, implies $\mathrm{Fun}^N[f(a,r)]$ and $\mathrm{Dom}[f(a,r), g(a,r)]$ implies $\mathrm{Dom}^N[f(a,r), g(a,r)]$, and so on. However, $\mathrm{Prog}[g(a,r), a, r]$ does not imply $\mathrm{Prog}^N[g(a,r), a, r]$.

Similarly, for the axioms ($\mathcal{P}$) and ($\mathcal{AD}$), we need $B \to B^N$ for

- $B := \mathcal{P}(x,y) \leftrightarrow (\forall z)(z \in y \leftrightarrow (\forall u \in z)(u \in x))$ and
- $B := \mathsf{Ad}(x) \to (x \models \mathsf{KP})$ respectively

which do not seem to holds intuitionistically.

All of them seem to require the double negation shift principle, another semi-constructive principle. If we found some trick which can deal with double negation shift, in a similar way as Avigad's forcing deals with Markov's principle, we could solve this problem. However we do not know such a trick.

The source of this kind of problem seems to be the precondition of Theorem 50: $A \to A^N$ is provable in $\mathsf{IKP}_0^\sharp + (\Delta_0^{s-}\text{-}\mathsf{MP})$ for all $A \in \mathcal{A}$.

For this reason, at this point, we do not know if we can obtain the same results for the extensions by ($\mathbb{B}$), by $(\mathcal{AD}') + (\mathbb{A})$ and by $(\mathcal{AD}') + (\mathbb{M})$ in the intuitionistic context. Also, for $*$-interpretation, we cannot omit (Σ_1-Ind) on the KP side. Thus we can have the result only for quite restricted theories.

Let us summarise the restricted results. Since the intuitionistic theory $\mathsf{IWEST}_0^{int} + (\Sigma_1(\mathsf{app}^+)\text{-}\mathsf{Ind})$ is a subsystem of the classical theory $\mathsf{WEST}_0 + (\Sigma_1(\mathsf{app}^+)\text{-}\mathsf{Ind})$, and since $\star$-interpretation (Definition 113) can interpret the latter in $\mathsf{KP}_r + (\Sigma_1\text{-}\mathsf{Ind})$, Theorem 127 gives us the following mutual interpretability results.

Corollary 128. The following theories are in each case mutually interpretable:

(i) $\mathsf{IWEST}_r + (\Sigma_1(\mathsf{app}^+)\text{-}\mathsf{Ind})$, $\mathsf{IWEST}_r^{int} + (\Sigma_1(\mathsf{app}^+)\text{-}\mathsf{Ind})$ and $\mathsf{KP}_r + (\Sigma_1\text{-}\mathsf{Ind})$;

(ii) $\mathsf{IWEST}_w + (\Sigma_1(\mathsf{app}^+)\text{-}\mathsf{Ind})$, $\mathsf{IWEST}_w^{int} + (\Sigma_1(\mathsf{app}^+)\text{-}\mathsf{Ind})$ and $\mathsf{KP}_w + (\Sigma_1\text{-}\mathsf{Ind})$;

(iii) IWEST, IWEST^{int} and KP.

The same assertions hold, if we add (Π_2-Ref) (without ($\mathcal{AD}'$)) to the variants of KP and at the same time ($\mathbb{M}$) to those of IWEST.

Note that since we do not claim any conservation we can have non-local interpretability here.

Although we have not defined the intuitionistic version of OST^-, we would have the same results for the intuitionistic operational set theory as far as it contains IWEST in the similar way as in the classical case.

AppendixA.5. Remark for intuitionistic interpreted theories

One might think that if, giving up the starting interpreted systems to be classical, we can interpret intuitionistic Kripke-Platek set theory in intuitionistic applicative set theories by our method. For this purpose, we do not need to extract the classical collection schema from the constructive collection schema and thus we do not need the two interpretations, negative and forcing interpretations, but only realisability interpretation. However, non-negative Δ_0 separation is not realisable with Feferman realisability notion.

Tupailo [56] interprets constructive Zermelo-Fraenkel set theory CZF, which includes IKP^- with full Δ_0 separation, in explicit mathematics. This is possible because in explicit mathematics we can take "exponential type". If an applicative set theory has exponential in the operational sense $\mathbb{E}(a,b) = \{\mathbf{fun}(f) : (f : a \to b)\}$, then, with a realisation notion similar to that used in Tupailo [56] (and also Tupailo [55]), we can realise full Δ_0 separation (actually all the axioms of CZF) in the applicative set theory.

AppendixB. Interpretability for Classical but Intensional Applicative Set Theories

In the last section, we saw that our main method yields lower bounds for intuitionistic applicative set theories only when $\mathcal{A} = \emptyset$. The method would also be useful if we are interested in the lower bound of classical but intensional applicative set theory, because the trick of (N-Ext), used in Section 8, does not work anymore in the case where the final interpreting system does not have extensionality. Indeed, we are sometimes interested in the lower bound of such theories, since the intensional theory WEST^{int} is easier, than the extensional counterpart WEST, to be embedded into other applicative systems. Here we define WEST^{int} as WEST without extensionality but with the axiom $\mathbb{U}(\{x\}) = x$, because of Remarks 14 and 20. Then $(\Sigma_1(\mathsf{app}^+)\text{-}\mathsf{Ind})$ can be replaced by (opInd) in the proof of Theorem 127.

Theorem 129. Let $\mathcal{A}$ be a set of $\mathcal{L}$ formulae, C a Π_1 formula of $\mathcal{L}_\in$ and D a Π_1 formula of $\mathcal{L}_\mathbb{P}$.

(i) If $\mathsf{KP}_r + (\mathcal{P}) + (\Sigma_1\text{-}\mathsf{Ind})$ proves D, then $\mathsf{WEST}_r^{int} + (\mathbb{P}) + (\mathsf{opInd})$ proves $(D^*)_\mathbb{P}$.
(ii) If $\mathsf{KP}_r + (\text{Beta}) + (\Sigma_1\text{-}\mathsf{Ind})$ proves C, then $\mathsf{WEST}_r^{int} + (\mathbb{B}) + (\mathsf{opInd})$ proves C^*.
(iii) If $\mathsf{KPI}'_r + (\Sigma_1\text{-}\mathsf{Ind})$ proves C, then $\mathsf{WEST}_r^{int} + (\mathcal{AD}') + (\mathbb{A}) + (\mathsf{opInd})$ proves C^*.
(iv) If $\mathsf{KPM}'_r + (\Sigma_1\text{-}\mathsf{Ind})$ proves C, then $\mathsf{WEST}_r^{int} + (\mathcal{AD}') + (\mathbb{M}) + (\mathsf{opInd})$ proves C^*.

The same results holds if we delete the subscript r or replace it by w.

Proof. It suffices to show that, in terms of Theorem 127, $\mathcal{A} = \{A\}$ implies $\mathcal{A}^*$ in KP_0^{int} for $A = (\mathcal{P})$, (Beta), $(\mathcal{AD}')$ and (Lim), since $\mathcal{A}$ and $\mathcal{B}$ are equivalent classically.

Let $A = (\mathcal{P})$. $\forall x \exists y \mathcal{P}(x,y)$ trivially implies $\forall x \exists y \mathcal{P}^*(x,y)$. Let us see $\mathcal{P}^*(x,y) \leftrightarrow (\forall z)(z \in^* y \leftrightarrow z \subseteq^* x)$. If $\mathcal{P}^*(x,y)$, then there is y' such that $\mathcal{P}(x,y')$ and $y \sim y'$ and, by $(\mathcal{P})$, we have $(\forall z)(z \in y' \leftrightarrow z \subseteq x)$ which implies $(\forall z)(z \in^* y \leftrightarrow z \subseteq^* x)$. Conversely, if $(\forall z)(z \in^* y \leftrightarrow z \subseteq^* x)$, then $(\forall z)(z \in^* y \leftrightarrow z \in^* y')$ where y' is such that $\mathcal{P}(x,y')$ as required in $(\mathcal{P})$, and hence $y \sim y'$ with $\mathcal{P}(x,y')$, which means $\mathcal{P}^*(x,y)$.

Let $A = (\text{Beta}')$. For given a, r, by Δ separation, take $r' = \{\langle u,v\rangle \in a \times a : \langle u,v\rangle \in^* r\}$. Then, by (Beta'), we have f and b such that $\text{Fun}[f]$, $\text{Dom}[f,b]$, $b \subseteq a$, $\text{DwCl}[b,a,r']$, $\text{Prog}[b,a,r']$ and, for any $u \in b$,

$$f'u = \{f'v : v \in b \wedge \langle v,u\rangle \in r'\} = \{f'v : v \in b \wedge \langle v,u\rangle \in^* r\}.$$

It is easy to see $\text{Fun}^*[f]$, $\text{Dom}^*[f,b]$, $b \subseteq^* a$ and $(f'u = \{f'v : \langle v,u\rangle \in r\})^*$ for all $u \in b$. Now $\text{DwCl}^*[b,a,r]$ is equivalent to $(\forall u \in b)(\forall v \in a)(\langle v,u\rangle \in^* r \to v \in^* b)$, which is from $(\forall u \in b)(\forall v \in a)(\langle v,u\rangle \in r' \to v \in b)$.

Since $\text{WF}[b,r']$, by induction on $u \in b$ along r', we are proving

$$(\forall u \in b)(\forall u' \in a)(u' \sim u \to u' \in b). \tag{B.1}$$

Let $u' \in a$ and $u' \sim u$. To show $u' \in b$, it suffices to show $(\forall v \in a)(\langle v,u'\rangle \in r' \to v \in b)$, by $\text{Prog}[b,a,r']$. Let $v \in a$ and $\langle v,u'\rangle \in r'$, that is, $\langle v,u'\rangle \in^* r$. Since $u \sim u'$, we have $\langle v,u\rangle \in^* r$, namely $\langle v,u\rangle \in r'$. By $\text{DwCl}[b,a,r']$, we have $v \in b$. This is what we have to show.

To see $\text{Prog}^*[b,a,r]$, let $u \in a$ and $(\forall v \in a)(\langle v,u\rangle \in^* r \to v \in^* b)$. This means $(\forall v \in a)(\langle v,u\rangle \in r' \to (\exists v' \in b) v \sim v')$ and, by (B.1), $(\forall v \in a)(\langle v,u\rangle \in r' \to v \in b)$. Since $\text{Prog}[b,a,r']$, we have $u \in b$.

Let $A = (\mathsf{Lim})$. By (Lim), for any x there is y such that $x \in y$, $\text{Tran}[y]$ and $\mathsf{Ad}(y)$. $\mathsf{Ad}(y)$ implies $y \models \mathsf{KP}^{int}$ and so $y \models (\mathsf{KP})^*$. By the Δ-ness of $\sim$ and $\in^*$ in $\mathsf{KP}_0^{int} + (\Sigma_1\text{-}\mathsf{Ind})$, we also have $(y \models \mathsf{KP})^*$. Since we also have $\text{Tran}^*[y]$, we have $\mathsf{Ad}^*[y]$ in the revised (for Ad) definition of $*$. Thus $(\mathsf{Lim})^*$.

For $A =$(AD.i) and (AD.ii), this is obvious by the revised definition of $*$. □

Now the classical intensional theory $\mathsf{WEST}_0^{int} + (\mathsf{opInd})$ is a subsystem of the classical extensional theory $\mathsf{WEST}_0 + (\mathsf{opInd})$. Moreover, $\star$-interpretation can interpret the latter in $\mathsf{KP}_r + (\Sigma_1\text{-}\mathsf{Ind})$. However, because we lost the trichotomy on admissibles, $\star$-interpretation does not work for $(\mathbb{A})$ or $(\mathbb{M})$. Thus we obtain the similar mutual interpretability results, only for $(\mathbb{P})$ and $(\mathbb{B})$.

Corollary 130. The following theories are in each case mutually interpretable:

(i) $\mathsf{WEST}^{int} + (\mathbb{P})$ and $\mathsf{KP} + (\mathcal{P})$;
(ii) $\mathsf{WEST}^{int} + (\mathbb{B})$ and $\mathsf{KP} + (\text{Beta})$;

The same assertions hold, if we replace WEST^{int} by $\mathsf{WEST}_r^{int} + (\mathsf{opInd})$ or $\mathsf{WEST}_w^{int} + (\mathsf{opInd})$ and at the same time KP by $\mathsf{KP}_r + (\Sigma_1\text{-}\mathsf{Ind})$ or $\mathsf{KP}_w + (\Sigma_1\text{-}\mathsf{Ind})$ respectively.

Actually, in the presence of powerset, we do not need $(\Sigma_1\text{-}\mathsf{Ind})$ any more, because now Δ-ness of $x \sim y$ can be proved as follows. Let c be the transitive closure of $\{x,y\}$ and b the powerset of c. Then, by induction on $u \in c$, we can show $(\exists a \in b)(\text{Bis}[a] \wedge u \in \text{field}(a))$. Particularly, we have $(\forall x,y)(\exists a)(\text{Bis}[a] \wedge x,y \in \text{field}(a))$ which was the only point where we need $(\Sigma_1\text{-}\mathsf{Ind})$ in the last section to show the Δ_1-ness of $\sim$.

Acknowledgments

The authors would like to express their gratitude to the anonymous referee for careful reading. They would also like to thank Solomon Feferman, Gerhard Jäger, Thomas Strahm and Dieter Probst, for their valuable comments on earlier versions of the present paper.

References

[1] Aczel, P., 1978. The type theoretic interpretation of constructive set theory. In: MacIntyre, A., Pacholski, L., Paris, J. B. (Eds.), Logic Colloquium '77, Amsterdam. Vol. 96 of Stud. Logic Foundations Math. North-Holland, Amsterdam, pp. 55–66.

[2] Avigad, J., 2000. Interpreting classical theories in constructive ones. J. Symbolic Logic 65 (4), 1785–1812.
URL http://www.andrew.cmu.edu/user/avigad/Papers/interp.pdf

[3] Avigad, J., 2004. Forcing in proof theory. Bull. Symbolic Logic 10 (3), 305–333.
URL http://www.andrew.cmu.edu/user/avigad/Papers/forcing.pdf

[4] Barwise, J., 1975. Admissible sets and structures: An approach to definability theory. Perspectives in Mathematical Logic. Springer-Verlag, Berlin.
URL http://projecteuclid.org/euclid.pl/1235418470

[5] Beeson, M. J., 1977. Principles of continuous choice and continuity of functions in formal systems for constructive mathematics. Ann. Math. Logic 12 (3), 249–322.
URL http://dx.doi.org/10.1016/S0003-4843(77)80003-X

[6] Beeson, M. J., 1979. Goodman's theorem and beyond. Pacific J. Math. 84 (1), 1–16.
URL http://projecteuclid.org/euclid.pjm/1102784341

[7] Beeson, M. J., 1985. Foundations of constructive mathematics. Vol. 6 of Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]. Springer-Verlag, Berlin.

[8] Cantini, A., 2011. Extending constructive operational set theory by impredicative principles. MLQ Math. Log. Q. 57 (3), 299–322.
URL http://dx.doi.org/10.1002/malq.201010009

[9] Cantini, A., Crosilla, L., 2008. Constructive set theory with operations. In: Logic Colloquium 2004. Vol. 29 of Lect. Notes Log. Assoc. Symbol. Logic, Chicago, IL, pp. 47–83.
URL http://www1.maths.leeds.ac.uk/~pmtmlcr/papers/lc04.pdf

[10] Cantini, A., Crosilla, L., 2010. Elementary constructive operational set theory. In: Schindler, R. (Ed.), Ways of Proof Theory. Ontos Verlag, Heusenstamm, pp. 199–240.
URL http://wwwmath.uni-muenster.de/logik/Personen/rds/pohlers_volume/cantinicrosilla_new.pdf

[11] Cantini, A., Crosilla, L., 2012. Transitive closure is conservative over weak operational set theory. In: Berger, U., Diener, H., Schuster, P., Seisenberger, M. (Eds.), Logic, Construction, Computation. Ontos Verlag, Heusenstamm, pp. 91–122.

[12] Feferman, S., 1975. A language and axioms for explicit mathematics. In: Algebra and logic (Fourteenth Summer Res. Inst., Austral. Math. Soc., Monash Univ., Clayton, 1974). Springer, Berlin, pp. 87–139. Lecture Notes in Math., Vol. 450.
URL http://math.stanford.edu/~feferman/papers/LangAxsExplMaths.pdf

[13] Feferman, S., 1978. Recursion theory and set theory: a marriage of convenience. In: Generalized recursion theory, II (Proc. Second Sympos., Univ. Oslo, Oslo, 1977). Vol. 94 of Stud. Logic Foundations Math. North-Holland, Amsterdam, pp. 55–90.
URL `http://math.stanford.edu/~feferman/papers/RecSetThMarriage.pdf`

[14] Feferman, S., 1979. Constructive theories of functions and classes. In: Logic Colloquium '78 (Mons, 1978). Vol. 97 of Stud. Logic Foundations Math. North-Holland, Amsterdam, pp. 159–224.
URL `http://math.stanford.edu/~feferman/papers/ConstrFns&Classes.pdf`

[15] Feferman, S., 1982. Iterated inductive fixed-point theories: application to Hancock's conjecture. In: Patras Logic Symposion (Patras, 1980). Vol. 109 of Stud. Logic Foundations Math. North-Holland, Amsterdam, pp. 171–196.

[16] Feferman, S., 2001. Notes on operational set theory, I. generalization of "small" large cardinals in classical and admissible set theory, technical notes.
URL `http://math.stanford.edu/~feferman/papers/OperationalST-I.pdf`

[17] Feferman, S., 2009. Operational set theory and small large cardinals. Inform. and Comput. 207 (10), 971–979.
URL `http://dx.doi.org/10.1016/j.ic.2008.04.007`

[18] Feferman, S., viewed in 2014. Introduction [to foundations of explicit mathematics (in progress), by Solomon Feferman, Gerhard Jäger, and Thomas Strahm, with the assistance of Ulrik Buchholtz], draft.
URL `http://www.math.stanford.edu/ feferman/papers/DraftIntroFEM.pdf`

[19] Friedman, H., 1973. The consistency of classical set theory relative to set theory with intuitionitic logic. J. Symbolic Logic 38 (2), 315–319.

[20] Friedman, H., 1978. Classically and intuitionistically provably recursive functions (proc. conf., math. forschungsinst., oberwolfach, 1977). In: Scott, D. S., Muller, G. H. (Eds.), Higher Set Theory. Vol. 699 of Lecture Notes in Mathematics. Springer, Berlin, pp. 21–28.

[21] Gordeev, L., 1982. Constructive models for set theory with extensionality. In: The L.E.J. Brouwer Centenary Symposium (Noordwijkerhout, 1981). Vol. 110 of Stud. Logic Found. Math. North-Holland, Amsterdam, pp. 123–147.

[22] Gordeev, L., 1988. Proof-theoretical analysis: weak systems of functions and classes. Ann. Pure Appl. Logic 38 (1), 1–121.
URL `http://dx.doi.org/10.1016/0168-0072(88)90055-3`

[23] Gordeev, L., 2012. Proof-theoretic conservations of weak weak intuitionistic constructive set theories. Ann. Pure Appl. Logic 164 (12), 1274–1292.
URL `http://dx.doi.org/10.1016/j.apal.2013.06.012`

[24] Jäger, G., 1983. A well-ordering proof for Feferman's theory T_0. Arch. Math. Logik Grundlag. 23 (1-2), 65–77.

[25] Jäger, G., 1984. A version of Kripke-Platek set theory which is conservative over Peano arithmetic. Z. Math. Logik Grundlag. Math. 30 (1), 3–9.

[26] Jäger, G., 1986. Theories for admissible sets: a unifying approach to proof theory. Vol. 2 of Studies in Proof Theory. Lecture Notes. Bibliopolis, Naples.

[27] Jäger, G., 1993. Fixed points in Peano arithmetic with ordinals. Ann. Pure Appl. Logic 60 (2), 119–132.
URL `http://dx.doi.org/10.1016/0168-0072(93)90039-G`

[28] Jäger, G., 2005. Metapredicative and explicit mahlo: a proof-theoretic perspective. In: Logic Colloquium 2000. Logic Colloquium 2000. Assoc. Symbol. Logic, Urbana, IL, pp. 272–293.
URL http://www.iam.unibe.ch/ltgpub/2005/jae05.pdf

[29] Jäger, G., 2007. On Feferman's operational set theory OST. Ann. Pure Appl. Logic 150 (1-3), 19–39.
URL http://www.iam.unibe.ch/ltgpub/2007/jae07.pdf

[30] Jäger, G., 2009. Full operational set theory with unbounded existential quantification and power set. Ann. Pure Appl. Logic 160 (1), 33–52.
URL http://www.iam.unibe.ch/ltgpub/2009/jae09a.pdf

[31] Jäger, G., 2013. Operational closure and stability. Ann. Pure Appl. Logic 164 (7-8), 813 – 821.
URL http://www.iam.unibe.ch/ltgpub/2013/jae13.pdf

[32] Jäger, G., Kahle, R., Studer, T., 2001. Universes in explicit mathematics. Ann. Pure Appl. Logic 109 (3), 141–162.
URL http://www.iam.unibe.ch/ltgpub/2001/jks01.pdf

[33] Jäger, G., Probst, D., 2013. A proof-theoretic analysis of theories for stratified inductive definitions. In: Kahle, R., Rathjen, M. (Eds.), Gentzen's Centenary: The Quest for Consistency. Springer, in press.
URL http://www.iam.unibe.ch/ltgpub/2013/jp13.pdf

[34] Jäger, G., Strahm, T., 1996. Some theories with positive induction of ordinal strength $\varphi\omega 0$. J. Symbolic Logic 61 (3), 818–842.
URL http://www.iam.unibe.ch/ltgpub/1996/js96.ps

[35] Jäger, G., Strahm, T., 2001. Upper bounds for metapredicative Mahlo in explicit methamtics and admissible set theory. J. Symbolic Logic 66 (2), 935–958.
URL http://www.iam.unibe.ch/ltgpub/2001/js01.pdf

[36] Jäger, G., Strahm, T., 2005. Reflections on reflections in explicit mathematics. Ann. Pure Appl. Logic 136 (1-2), 116–133.
URL http://www.iam.unibe.ch/ltgpub/2005/js05.pdf

[37] Jäger, G., Studer, T., 2002. Extending the system T_0 of explicit mathematics: the limit and Mahlo axioms. Ann. Pure Appl. Logic 114 (1–3), 79–101.
URL http://www.iam.unibe.ch/ltgpub/2002/js02b.pdf

[38] Jäger, G., Zumbrunnen, R., 2012. About the strength of operational regularity. In: Berger, U., Diener, H., Schuster, P., Seisenberger, M. (Eds.), Logic, Construction, Computation. Ontos Verlag, pp. 305–324.
URL http://www.iam.unibe.ch/ltgpub/2012/jz12.pdf

[39] Jäger, G., Zumbrunnen, R., 2014. Explicit mathematics and operational set theory: some ontological considerations. Bull. Symbolic Logic, in press.
URL http://www.iam.unibe.ch/ltgpub/2014/jz14.pdf

[40] Jech, T., 1978. Set theory. Academic Press [Harcourt Brace Jovanovich Publishers], New York.

[41] Krivine, J.-L., 2011. Realizability algebras: a program to well order $\mathbb{R}$. Log. Methods Comput. Sci. 7 (2), 3:02, 47 pp.
URL http://arxiv.org/pdf/1005.2395

[42] Krivine, J.-L., 2012. Realizability algebras II: New models of **ZF+DC**. Log. Methods Comput. Sci. 8 (1), 1:10, 28 pp.
URL http://arxiv.org/pdf/1007.0825

[43] Kunen, K., 1980. Set Theory. Vol. 102 of Studies in Logic and the Foundations of Mathematics. North-Holland Publishing Co., Amsterdam, an introduction to independence proofs.

[44] Mathias, A. R. D., 2001. The strength of Mac Lane set theory. Ann. Pure Appl. Logic 110 (1-3), 107–234.
URL http://dx.doi.org/10.1016/S0168-0072(00)00031-2

[45] Pour-El, M. B., Kripke, S., 1967. Deduction-preserving 'recursive isomorphisms' between theories. Fund. Math. 61, 141–163.

[46] Ranzi, F., Strahm, T., 2013. A note on the theory $\mathrm{SID}_{<\omega}$ of stratified induction, submitted.
URL http://www.iam.unibe.ch/ltgpub/2013/rs13.pdf

[47] Sato, K., 2006. Forcing under anti-foundation axiom: an expresion of stalks. MLQ Math. Log. Q. 52 (3), 295–314.
URL http://dx.doi.org/10.1002/malq.200410600

[48] Sato, K., 2009. The strength of extensionality I: weak weak set theory with infinity. Ann. Pure Appl. Logic 157 (2-3), 234–268.
URL http://dx.doi.org/10.1016/j.apal.2008.09.010

[49] Sato, K., 2011. The strength of extensionality II: weak weak set theory without infinity. Ann. Pure Appl. Logic 162 (8), 579–646.
URL http://dx.doi.org/10.1016/j.apal.2010.11.001

[50] Simpson, S. G., 2009. Subsystems of Second Order Arithmetic, 2nd Edition. Perspectives in Logic. Springer, Ithca NY.

[51] Strahm, T., 2003. Theories with self-application and computational complexity. Inform. and Comput. 185 (2), 263–297.
URL http://www.iam.unibe.ch/ltgpub/2003/str03.pdf

[52] Strahm, T., 2004. A proof-theoretic characterization of the basic feasible functionals. Theoret. Comput. Sci. 329 (1-3), 159–176.
URL http://www.iam.unibe.ch/ltgpub/2004/str04.pdf

[53] Troelstra, A. S., Schwichtenberg, H., 2000. Basic Proof Theory, 2nd Edition. Vol. 43 of Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, Cambridge.

[54] Troelstra, A. S., van Dalen, D., 1988. Constructivism in Mathematics. Vol. I. An Introduction. Vol. 121 of Studies in Logic and the Foundations of Mathematics. North-Holland, Amsterdam.

[55] Tupailo, S., 2001. Realization of analysis into explicit mathematics. J. Symbolic Logic 66 (4), 1848–1864.
URL http://www.iam.unibe.ch/ltgpub/2001/tup01.pdf

[56] Tupailo, S., 2003. Realization of constructive set theory into explicit mathematics: a lower bound for impredicative mahlo universe. Ann. Pure Appl. Logic 120 (1-3), 165–196.
URL http://www.iam.unibe.ch/ltgpub/2003/tup03.ps

[57] van Dalen, D., 2004. Logic and Structure, 4th Edition. Universitext. Springer-Verlag, Berlin.
URL www.cin.ufpe.br/~mlogica/livros/Logic%20and%20Structure%20-%20Van%20Dalen.pdf

[58] Visser, A., 2009. The predicative Frege hierarchy. Ann. Pure Appl. Logic 160 (2), 129–153.
URL http://dx.doi.org/10.1016/j.apal.2009.02.001

[59] Zumbrunnen, R., 2013. Contributions to Operational Set Theory. Ph.D. thesis, Universität Bern.
URL http://www.iam.unibe.ch/ltgpub/2013/zum13.pdf