THE JOURNAL OF SYMBOLIC LOGIC
Volume 83, Number 3, September 2018

TRUTHS, INDUCTIVE DEFINITIONS, AND KRIPKE-PLATEK SYSTEMS OVER SET THEORY

KENTARO FUJIMOTO

Abstract. In this article we study the systems KF and VF of truth over set theory as well as related systems and compare them with the corresponding systems over arithmetic.

§1. Introduction. How is the content of the term "is true" given to us? Possibly it may be given by an explicit definition in terms of other notions, but Tarski's undefinability theorem imposes a quite stringent restriction on the explicit definability of truth in nonsemantic terms. Some argue that the notion of truth is ultimately to be axiomatically conceived; namely, a certain collection of sentences involving the term "is true", called axioms or meaning postulates of the term, determine its content and use. The present article focuses on such an axiomatic approach toward truth, which takes the term "is true" to be axiomatically understood and studies various axioms for it.

One peculiar feature of the notion of truth is that it can be applied to any sentence about any subject matter but in the same uniform way. While we can talk of truth of two different subjects as separate and independent issues on their own rights, we also regard them as restrictions of a certain general notion of truth to the particular subject matters in question, sharing certain "essential" properties that uniformly permeate through truths of all subject matters.

With this conception of truth, one natural formal setting for theories of truth for a given subject matter is the following. We first pick and fix a formal system of the subject matter, which is called a *base system*. Then we add the axioms of truth on top of the base system. These axioms of truth are given independently of the subject matter and base system; we may need to slightly tweak and adjust their formulation to fit them in the formal structure of the chosen base system, but these axioms should express the same "essential" property of truth from one subject matter to another and from one base system to another. We call the result of this process an *axiomatic system of truth over* the base system (or over the subject matter). One important implication of this view is that the notion of truth is not intrinsically embodied in a chosen subject matter and some general (but informal) conception of truth is somehow taken as given in advance independently of the choice of subject matters. Hence, in principle, we can (and should) consider and investigate axiomatic

Received March 5, 2016.
2010 *Mathematics Subject Classification.* 03B30, 03D70, 03E65, 03E70, 03F35, 03F99.
Key words and phrases. axiomatic theories of truth, second-order set theory, Kripke-Platek set theory, inductive definitions.

0022-4812/18/8303-0002
DOI:10.1017/jsl.2017.77

systems of truth and their axiomatic conceptions of truth over a variety of different subject matters and base systems.

The study of axiomatic systems of truth has so far centered around those over arithmetic, and philosophical debates on the axiomatic approach to truth have been based mostly on the results about those systems over arithmetic. This is not because philosophers are only interested in the truth of arithmetic, but probably because they believe that those systems over arithmetic provide a generic case and most of the fundamental results over arithmetic and philosophical debates based on them can be generalized to other cases. However, in the present article, we will show that there exists some strong disanalogy between axiomatic systems of truth over arithmetic and over set theory, and thereby suggest that axiomatic systems of truth over arithmetic may not be such a generic case.

A distinction is often made between *compositional* and *noncompositional* systems of truth. Halbach explains this distinction as follows:

> I call an axiomatic system [of truth] compositional if, according to its axioms, the semantical status of its expressions (in particular, their truth or falsity) depends only on the semantical status of its constituents. [15, p. 120]

Halbach refers to the Kripke-Feferman system KF as a typical example of a compositional system and to Cantini's VF as an example of a noncompositional system. This distinction is sometimes thought to be fundamental, and Halbach suggests to relate compositionality to predicativity:

> In general, predicativity and compositionality seem closely related. Compositionality is to truth systems what predicativity is to second-order system. [15, p. 120]

In the present article, we will focus on the arch compositional system KF and the arch noncompositional system VF and investigate the relationship between them as well as other relevant systems both over arithmetic and over set theory. Consequently, we will see some strong disanalogy between their behavior and the relationships of them over arithmetic PA and over set theory ZF: in particular, it will be shown that KF and VF are proof-theoretically equivalent over ZF and thus have the same set-theoretic consequences, whereas the former is significantly weaker than the latter over PA.

The structure of the article is as follows. In Sections 2–5, we introduce the main systems we will investigate and show some basic facts about them: namely, the systems KF and VF of truth, the systems $\widehat{\mathsf{ID}}_1$ and ID_1 of fixed-points, and the system SC_1 of stage comparison prewellorderings. We next introduce an intermediate system $\mathsf{KP}\mathbb{V}$, the Kripke-Platek set theory over $\mathbb{V}$, in Section 6, and then give an embedding of $\mathsf{KP}\mathbb{V}$ in SC_1 in Section 7. Finally, by giving an embedding of VF in $\mathsf{KP}\mathbb{V}$, we obtain the equivalence of all those systems in Section 8. After obtaining this main result, we first give two relevant results as an application of our results in Section 9 and then study some variants of those systems in Sections 10–12.

§2. KF **and** VF **over set theory.** Let $\mathcal{L}_\in = \{\in\}$ be the language of first-order set theory with the membership relation $\in$ as its only non-logical symbol. ZF stands for Zermelo-Fraenkel set theory over $\mathcal{L}_\in$. For the sake of systems of truth we also consider an expansion $\mathcal{L}_T$ of $\mathcal{L}_\in$ defined as $\mathcal{L}_T := \mathcal{L}_\in \cup \{T\}$, where T is a unary predicate symbol that is meant to be the truth predicate.

Let $\mathcal{L}$ be either $\mathcal{L}_\in$ or $\mathcal{L}_T$. Within ZF we can formalize the language $\mathcal{L}^\infty$ which consists of $\mathcal{L}$ with constant symbols c_x for each element x of the universe $\mathbb{V}$. This formalization provides us with a coding of the $\mathcal{L}^\infty$-expressions; for an $\mathcal{L}^\infty$-expression e we denote its code by $\ulcorner e \urcorner$; we specially denote the code of the set constant c_x for $x \in \mathbb{V}$ by $\dot{x}$. This formalization also comes with a coding of various syntactic relations and operations on $\mathcal{L}^\infty$. We will use exactly the same notation and definitions for this formalization as in [12]. For instance, we write

$$\mathrm{St}^\infty_\mathcal{L} := \{z \mid z \text{ is a code of an } \mathcal{L}^\infty\text{-sentence}\};$$
$$\mathrm{Fml}^\infty_\mathcal{L} := \{z \mid z \text{ is a code of an } \mathcal{L}^\infty\text{-formula}\};$$

$\underset{\cdot}{\wedge}$ and $\underset{\cdot}{\neg}$ are (class) functions such that, for $\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner \in \mathrm{Fml}^\infty_\mathcal{L}$,

$$\underset{\cdot}{\neg}\ulcorner\varphi\urcorner = \ulcorner\neg\varphi\urcorner \qquad \text{and} \qquad \ulcorner\varphi\urcorner \underset{\cdot}{\wedge} \ulcorner\psi\urcorner = \ulcorner\varphi \wedge \psi\urcorner.$$

We can assume all the syntactic relations and operations that we use are Δ^{ZF}_1. For readability, we write "$\forall\ulcorner\varphi\urcorner$" and "$\exists\ulcorner\varphi\urcorner$" to emphasize that codes of formulae are quantified over and to thereby suppress the syntactical operations; for example, by $(\forall\ulcorner\varphi\urcorner \in \mathrm{St}^\infty_\mathcal{L})(\forall\ulcorner\psi\urcorner \in \mathrm{St}^\infty_\mathcal{L})(T\ulcorner\varphi \wedge \psi\urcorner \leftrightarrow T\ulcorner\neg\neg(\varphi \wedge \psi)\urcorner)$, we mean

$$(\forall x \in \mathrm{St}^\infty_\mathcal{L})(\forall y \in \mathrm{St}^\infty_\mathcal{L})\big(T(x \underset{\cdot}{\wedge} y) \leftrightarrow T(\underset{\cdot}{\neg}(\underset{\cdot}{\neg}(x \underset{\cdot}{\wedge} y)))\big).$$

We will also write $\forall\ulcorner\varphi(v_1 \dots v_k)\urcorner$ to express "for all codes of formulae with at most k variables free"; $\exists\ulcorner\varphi(v_1 \dots v_k)\urcorner$ has the dual meaning for existential quantification. For an $\mathcal{L}^\infty$-formula $\varphi(v)$ with a distinguished free variable v and a set $x \in \mathbb{V}$, we write $\ulcorner\varphi(\dot{x})\urcorner$ for the code of the result of substituting the constant c_x for x for the variable v in φ (i.e., the so-called Feferman's dot convention).

Let $\mathcal{L}$ be any first-order language including $\mathcal{L}_\in$. We will consider the following extensions of the axiom schemata of set theory to $\mathcal{L}$:

$\mathcal{L}$-Ind : $\quad \forall x\big((\forall y \in x)\varphi(y) \to \varphi(x)\big) \to \forall x\varphi(x)$, for each $\varphi \in \mathcal{L}$.

$\mathcal{L}$-Sep : $\quad \forall a\exists b\forall x\big[x \in b \leftrightarrow x \in a \wedge \varphi(x)\big]$, for each $\varphi \in \mathcal{L}$.

$\mathcal{L}$-Repl : $\quad \forall a\big[(\forall x \in a)\exists! y\varphi(x, y) \to \exists b(\forall x \in a)(\exists y \in b)\varphi(x, y)\big]$, for each $\varphi \in \mathcal{L}$.

We can easily show $\mathsf{ZF} + \mathcal{L}\text{-Sep} \vdash \mathcal{L}\text{-Ind}$ for any $\mathcal{L} \supset \mathcal{L}_\in$.

Definition 2.1. The axioms of KF^- comprises those of ZF plus:

K1: $\forall x\forall y[(T\ulcorner\dot{x} = \dot{y}\urcorner \leftrightarrow x = y) \wedge (F\ulcorner\dot{x} = \dot{y}\urcorner \leftrightarrow x \neq y)]$,

K2: $\forall x\forall y[(T\ulcorner\dot{x} \in \dot{y}\urcorner \leftrightarrow x \in y) \wedge (F\ulcorner\dot{x} \in \dot{y}\urcorner \leftrightarrow x \notin y)]$,

K3: $\forall x[(T\ulcorner T\dot{x}\urcorner \leftrightarrow Tx) \wedge (F\ulcorner T\dot{x}\urcorner \leftrightarrow Fx)]$,

K4: $(\forall\ulcorner\sigma\urcorner \in \mathrm{St}^\infty_{\mathcal{L}_T})(T\ulcorner\neg\neg\sigma\urcorner \leftrightarrow T\ulcorner\sigma\urcorner)$,

K5: $(\forall\ulcorner\sigma\urcorner, \ulcorner\tau\urcorner \in \mathrm{St}^\infty_{\mathcal{L}_T})\big[\big(T\ulcorner\sigma \wedge \tau\urcorner \leftrightarrow (T\ulcorner\sigma\urcorner \wedge T\ulcorner\tau\urcorner)\big) \wedge \big(F\ulcorner\sigma \wedge \tau\urcorner \leftrightarrow (F\ulcorner\sigma\urcorner \vee F\ulcorner\tau\urcorner)\big)\big]$,

K6: $(\forall\ulcorner\varphi(v)\urcorner \in \mathrm{Fml}^\infty_{\mathcal{L}_T})\big[\big(T\ulcorner\forall v\varphi(v)\urcorner \leftrightarrow \forall x T\ulcorner\varphi(\dot{x})\urcorner\big) \wedge \big(F\ulcorner\forall v\varphi(v)\urcorner \leftrightarrow \exists x F\ulcorner\varphi(\dot{x})\urcorner\big)\big]$,

where we put $Fx :\Leftrightarrow T\underset{\cdot}{\neg}x$. Then we set $\mathsf{KF} := \mathsf{KF}^- + \mathcal{L}_T\text{-Sep} + \mathcal{L}_T\text{-Repl}$. Some proof-theoretic analyses of KF are already given in [12].

DEFINITION 2.2. The axioms of VF^- comprises those of ZF plus:

V1: $\forall \vec{x}\big(T(\ulcorner \varphi(\dot{\vec{x}})\urcorner) \to \varphi(\vec{x})\big)$, for each $\mathcal{L}_T$-formula $\varphi(\vec{x})$,
V2: $\forall x \forall y\big[\big(T\ulcorner \dot{x} = \dot{y}\urcorner \leftrightarrow x = y\big) \wedge \big(F\ulcorner \dot{x} = \dot{y}\urcorner \leftrightarrow x \neq y\big)\big]$,
V3: $\forall x \forall y\big[\big(T\ulcorner \dot{x} \in \dot{y}\urcorner \leftrightarrow x \in y\big) \wedge \big(F\ulcorner \dot{x} \in \dot{y}\urcorner \leftrightarrow x \notin y\big)\big]$,
V4: $(\forall \ulcorner \varphi(\vec{v})\urcorner \in \mathrm{Fml}^{\infty}_{\mathcal{L}_T})\big(\mathrm{LogAx}_{\mathcal{L}^{\infty}_T}(\ulcorner \varphi(\vec{v})\urcorner) \to T\ulcorner \forall \vec{v}\varphi(\vec{v})\urcorner\big)$,
V5: $(\forall \ulcorner \varphi(v)\urcorner \in \mathrm{Fml}^{\infty}_{\mathcal{L}_T})\big(\forall x T\ulcorner \varphi(\dot{x})\urcorner \to T\ulcorner \forall v\varphi(v)\urcorner\big)$,
V6: $(\forall \ulcorner \sigma\urcorner, \ulcorner \tau\urcorner \in \mathrm{St}^{\infty}_{\mathcal{L}_T})\big(T\ulcorner \sigma \to \tau\urcorner \to (T\ulcorner \sigma\urcorner \to T\ulcorner \tau\urcorner)\big)$,
V7: $(\forall \ulcorner \sigma\urcorner \in \mathrm{St}^{\infty}_{\mathcal{L}_T})(T\ulcorner \sigma\urcorner \to T\ulcorner T\ulcorner \sigma\urcorner\urcorner)$,
V8: $(\forall \ulcorner \sigma\urcorner \in \mathrm{St}^{\infty}_{\mathcal{L}_T})(F\ulcorner T\ulcorner \sigma\urcorner\urcorner \to F\ulcorner \sigma\urcorner)$,
V9: $(\forall \ulcorner \sigma\urcorner \in \mathrm{St}^{\infty}_{\mathcal{L}_T})\ T\ulcorner (T\ulcorner \sigma\urcorner \to \neg T\ulcorner \neg\sigma\urcorner)\urcorner$,

where $\mathrm{LogAx}_{\mathcal{L}^{\infty}_T}(x)$ expresses "x is a logical axiom for $\mathcal{L}^{\infty}_T$"; hence **V3** says "the universal closure of every logical axiom for $\mathcal{L}^{\infty}_T$ is true". Then we set $\mathsf{VF} := \mathsf{VF}^- + \mathcal{L}_T\text{-Sep} + \mathcal{L}_T\text{-Repl}$.

§3. ID_1 and $\widehat{\mathsf{ID}}_1$ over set theory. For a first-order language $\mathcal{L}$, we let $\mathcal{L}^2$ be the second-order language associated with $\mathcal{L}$ with infinitely many unary predicate variables $\mathfrak{X}, \mathfrak{Y}, \mathfrak{Z}, \dots$ but without any new nonlogical symbols added. We call an $\mathcal{L}^2$-formula Φ *elementary* when Φ contains no second-order quantifiers (possibly with second-order free variables); the Π^0_n- and Σ^0_n-formulae are standardly defined. An $\mathcal{L}$-*inductive operator form* is an elementary $\mathcal{L}^2$-formula $\mathcal{A}(x, \mathfrak{X})$ with only one second-order variable $\mathfrak{X}$ and one first-order variable x free in which $\mathfrak{X}$ occurs only positively. We write $\mathfrak{I}(\mathcal{L})$ for the set of $\mathcal{L}$-inductive operator forms.

For an $\mathcal{L}^2$-formula $\mathcal{B}(\mathfrak{X}_1, \dots, \mathfrak{X}_n)$ with designated second-order free variables $\mathfrak{X}_1, \dots, \mathfrak{X}_n$, and for $\mathcal{L}$-formulae $\Psi_1(u_1), \dots, \Psi_n(u_n)$ with designated first-order free variables $u_1, \dots u_n$, an $\mathcal{L}$-formula $\mathcal{B}(\Psi_1(\hat{u}_0), \dots, \Psi_n(\hat{u}_n))$ denotes the result of simultaneously replacing each occurrence of $\mathfrak{X}_i t$ by $\Psi_i(t)$ for each term t $(1 \leq i \leq n)$ with renaming of bound variables in $\mathcal{B}$ and Ψ_i's as necessary to avoid collision; we occasionally suppress '$\hat{u}_i$'s and simply write $\mathcal{B}(\Psi_0, \dots, \Psi_n)$. For an $\mathcal{L}$-formula $\Psi(z)$ and an $\mathcal{L}^2$-formula $\mathcal{C}(x, \mathfrak{X})$ with designated free variables z, and x and $\mathfrak{X}$, respectively, possibly with parameters, we define

$$Clos_{\mathcal{C}}(\Psi(\hat{z})) \ := \ \forall x\big(\mathcal{C}(x, \Psi(\hat{z})) \to \Psi(x)\big).$$

Again we will suppress "$\hat{z}$" when there is no worry of confusion.

A first-order language $\mathcal{L}_{\mathrm{Fix}}$ for systems of inductive definitions is defined as $\mathcal{L}_{\in}$ plus *unary* predicates $J_{\mathcal{A}}$ associated to each $\mathcal{A}(x, \mathfrak{X}) \in \mathfrak{I}(\mathcal{L}_{\in})$. We will occasionally identify a formula $\varphi(x)$, possibly with parameters, and the class $\{x \mid \varphi(x)\}$; e.g., we write $x \in J_{\mathcal{A}}$ for $J_{\mathcal{A}}(x)$ and $J_{\mathcal{A}} \subset \Phi$ for $\forall x(x \in J_{\mathcal{A}} \to \Phi(x))$.

DEFINITION 3.1. The $\mathcal{L}_{\mathrm{Fix}}$-system $\widehat{\mathsf{ID}}_1^-$ comprises ZF plus the following schema:

$$\forall x[\,J_{\mathcal{A}}(x) \leftrightarrow \mathcal{A}(x, J_{\mathcal{A}})\,], \text{ for each } \mathcal{A} \in \mathfrak{I}(\mathcal{L}_{\in}).$$

Then we set $\widehat{\mathsf{ID}}_1 := \widehat{\mathsf{ID}}_1^- + \mathcal{L}_{\mathrm{Fix}}\text{-Sep} + \mathcal{L}_{\mathrm{Fix}}\text{-Repl}$.

The $\mathcal{L}_{\mathrm{Fix}}$-system ID_1^- comprises ZF plus the following schemata:

$Clos_{\mathcal{A}}(J_{\mathcal{A}})$, for each $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_{\in})$.

$Clos_{\mathcal{A}}(\Psi) \to \forall x[J_{\mathcal{A}}(x) \to \Psi(x)]$, for each $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_{\in})$ and $\Psi \in \mathcal{L}_{\mathrm{Fix}}$.

Then we set $\mathsf{ID}_1 := \mathsf{ID}_1^- + \mathcal{L}_{\mathrm{Fix}}\text{-Sep} + \mathcal{L}_{\mathrm{Fix}}\text{-Repl}$.

We can standardly show that $\widehat{\mathsf{ID}}_1^-$ is a sub-theory of ID_1^- (see [4, Lemma 2.1.1]).

Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be first-order languages, and let S and T be systems over $\mathcal{L}_1$ and $\mathcal{L}_2$, respectively. For $\mathcal{L} \subset \mathcal{L}_1 \cap \mathcal{L}_2$, we write $\mathsf{S} \subset_{\mathcal{L}} \mathsf{T}$ when S is conservative over T for $\mathcal{L}$; the relation $\mathsf{S} =_{\mathcal{L}} \mathsf{T}$ means that $\mathsf{S} \subset_{\mathcal{L}} \mathsf{T}$ and $\mathsf{T} \subset_{\mathcal{L}} \mathsf{S}$.

There are a number of similarities and analogies between systems of truth or inductive definitions over set theory ZF and over arithmetic PA, and we will discuss the arithmetical counterparts of VF, ID_1, etc., over PA. Hence, to clearly distinguish them, when mentioning those systems over PA, we will add "$[\![\mathsf{PA}]\!]$" after the names of systems; e.g., $\mathsf{VF}[\![\mathsf{PA}]\!]$, $\mathsf{ID}_1[\![\mathsf{PA}]\!]$, etc.

THEOREM 3.2. 1. $\mathsf{KF} =_{\mathcal{L}_\in} \widehat{\mathsf{ID}}_1$. 2. $\mathsf{ID}_1 \subset_{\mathcal{L}_\in} \mathsf{VF}$.

PROOF. 1. One inclusion $\widehat{\mathsf{ID}}_1 \subset_{\mathcal{L}_\in} \mathsf{KF}$ can be shown in an exactly parallel manner to Cantini's [6] proof of $\widehat{\mathsf{ID}}_1[\![\mathsf{PA}]\!] \subset_{\mathcal{L}_\mathbb{N}} \mathsf{KF}[\![\mathsf{PA}]\!]$ over arithmetic, where $\mathcal{L}_\mathbb{N}$ is the first-order language of arithmetic. The converse can be shown by interpreting the truth predicate T of KF by a fixed-point of an inductive operator form $\mathcal{T}(x, \mathfrak{X}) \in \mathfrak{I}(\mathcal{L}_\in)$ describing the closure condition of the strong Kleene evaluation schema; the proof is exactly parallel to Feferman's [9] proof of $\mathsf{KF}[\![\mathsf{PA}]\!] \subset_{\mathcal{L}_\mathbb{N}} \Sigma^1_1\text{-AC}$. These proofs yield the mutual interpretability of KF^- and $\widehat{\mathsf{ID}}_1^-$ in which the $\mathcal{L}_\in$-part is preserved; hence, we actually have $\mathsf{KF}^- =_{\mathcal{L}_\in} \widehat{\mathsf{ID}}_1^-$.

2. Kahle [18] gives a direct interpretation of $\mathsf{VF}[\![\mathsf{PA}]\!]$ in $\mathsf{ID}_1[\![\mathsf{PA}]\!]$ that preserves the arithmetical part, and this interpretation can be used as it is for our claim: that is, for each $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)$, we can interpret $J_\mathcal{A}(x)$ by an $\mathcal{L}_T$-formula

$$(\forall \ulcorner \varphi(v) \urcorner \in \mathrm{Fml}^\infty_{\mathcal{L}_T})[\, Clos_\mathcal{A}(T \ulcorner \varphi(\hat{v}) \urcorner) \to T \ulcorner \varphi(\dot{x}) \urcorner\,].$$

In fact, this is an interpretation of ID_1^- in VF^- and thus $\mathsf{ID}_1^- \subset_{\mathcal{L}_\in} \mathsf{VF}^-$. ⊣

§4. Systems of stage comparison strict preordering. Let us fix any $\mathcal{L}_\in$-structure $\mathfrak{M} = \langle M, E \rangle$ where E is an interpretation of the symbol $\in$. Each $\mathcal{A}(x, \mathfrak{X}) \in \mathfrak{I}(\mathcal{L}_\in)$ induces a monotone operator $\Phi^\mathfrak{M}_\mathcal{A} \colon \mathcal{P}(M) \to \mathcal{P}(M)$ such that, for $X \subset M$, $\Phi^\mathfrak{M}_\mathcal{A}(X) = \{x \in M \mid \langle M, E, X \rangle \models \mathcal{A}(x, \mathfrak{X})\}$, where $\langle M, E, X \rangle$ is an $(\mathcal{L}_\in \cup \{\mathfrak{X}\})$-structure in which the predicate $\mathfrak{X}$ is interpreted by X. An operator $\Phi \colon \mathcal{P}(M) \to \mathcal{P}(M)$ is called *inductive* on $\mathfrak{M}$, when $\Phi = \Phi^\mathfrak{M}_\mathcal{A}$ for some $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)$.

Let Φ be an inductive operator on $\mathfrak{M}$. By recursion on ordinals α, we define sets $I^{<\alpha}_\Phi, I^\alpha_\Phi \in \mathcal{P}(M)$ as $I^{<\alpha}_\Phi := \bigcup_{\beta<\alpha} I^\beta_\Phi$ and $I^\alpha_\Phi := \Phi(I^{<\alpha}_\Phi)$, respectively. Then there is an ordinal α such that $I^\alpha_\Phi = I^{<\alpha}_\Phi$ and $\Phi(I^\alpha_\Phi) = I^\alpha_\Phi$. We denote the least such α by $||\Phi||$ and simply write I_Φ for $I^{||\Phi||}_\Phi$. For each $x \in I_\Phi$, we set $||x||_\Phi := \min\{\xi \mid x \in I^\xi_\Phi\}$, which induces a strict prewellordering $\prec_\Phi$ on M:

$$x \prec_\Phi y \qquad \Leftrightarrow \qquad \begin{cases} ||x||_\Phi < ||y||_\Phi & \text{if } x, y \in I_\Phi, \\ x \in I_\Phi \wedge y \notin I_\Phi & \text{otherwise.} \end{cases}$$

We call $\prec_\Phi$ the *stage comparison strict prewellordering* of Φ. This is so defined that the field $fd(\prec_\Phi)$ of $\prec_\Phi$ is M and the elements $y \in M \setminus I_\Phi$ are all maximal elements greater than any $x \in I_\Phi$. We have $I_\Phi = \{x \in M \mid \exists y(x \in \Phi(I^{<||y||_\Phi}_\Phi)\}$.

Now the following is easily observed:

$$x \prec_\Phi y \;\Leftrightarrow\; x \in I_\Phi \wedge y \notin \Phi(I_\Phi^{<||x||_\Phi}) \;\Leftrightarrow\; x \in I_\Phi \wedge y \notin \Phi(\{u \mid u \prec_\Phi x\}).$$

We use this equivalence for axiomatizing the stage comparison strict prewellorderings $\prec_\Phi$ of inductive operators (on our intended model $\mathfrak{M} := \langle \mathbb{V}, \in \rangle$).

Let $\mathcal{L}_{\mathrm{SC}}$ be a language defined as $\mathcal{L}_{\mathrm{Fix}}$ plus a further unary predicate $\prec_\mathcal{A}$ associated to each inductive operator form $\mathcal{A}(x, \mathfrak{X}) \in \mathfrak{I}(\mathcal{L}_\in)$, which is meant to express the stage comparison strict prewellordering of $\Phi_\mathcal{A}^\mathfrak{M}$. For readability we will write $x \prec_\mathcal{A} y$ for $\langle x, y\rangle \in \prec_\mathcal{A}$. Given $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)$ we write $\prec_\mathcal{A}\restriction_x$ for the class of $\prec_\mathcal{A}$-predecessors of x, i.e., $\{y \mid y \prec_\mathcal{A} x\}$ (with x as a parameter).

Definition 4.1. The $\mathcal{L}_{\mathrm{SC}}$-system SC_1^- comprises ID_1^- plus: for all $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)$,

(SC0): $\prec_\mathcal{A} \subset$ *Pair*, where *Pair* denotes the class of *ordered* pairs;

(SC1): $\forall x \forall y \big[x \prec_\mathcal{A} y \leftrightarrow \big(x \in J_\mathcal{A} \wedge \neg\mathcal{A}(y, \prec_\mathcal{A}\restriction_x)\big)\big]$;

(SC2): $\forall x \big(\forall y (y \prec_\mathcal{A} x \to \varphi(y)) \to \varphi(x)\big) \to \forall x \varphi(x)$, for all $\varphi(x) \in \mathcal{L}_{\mathrm{SC}}$.

Then we set $\mathsf{SC}_1 := \mathsf{SC}_1^- + \mathcal{L}_{\mathrm{SC}}$-Sep $+ \mathcal{L}_{\mathrm{SC}}$-Repl.

Remark 4.2. SC_1 is equivalent to Sato's [22, p. 106] axiomatization ID_1^+ of stage comparison prewellorderings. The equivalence will be shown in Appendix.

Lemma 4.3. 1. $\mathsf{SC}_1^- \vdash \forall x\, [\, x \in J_\mathcal{A} \leftrightarrow \mathcal{A}(x, \prec_\mathcal{A}\restriction_x)]$, *for each* $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)$.
2. $\mathsf{SC}_1^- \vdash \forall x (x \notin J_\mathcal{A} \leftrightarrow J_\mathcal{A} \subset \prec_\mathcal{A}\restriction_x)$, *for each* $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)$.

Proof. 1. Note that $\prec_\mathcal{A}$ is irreflexive due to (SC2). Hence, if $x \in J_\mathcal{A}$ then $\mathcal{A}(x, \prec_\mathcal{A}\restriction_x)$ by (SC1). Suppose $\mathcal{A}(x, \prec_\mathcal{A}\restriction_x)$. By (SC1) we have $\prec_\mathcal{A}\restriction_z \subset J_\mathcal{A}$ for all z in general. Hence, we obtain $\mathcal{A}(x, J_\mathcal{A})$ by monotonicity and thus $x \in J_\mathcal{A}$.

2. Suppose $x \notin J_\mathcal{A}$. We have $\neg\mathcal{A}(x, J_\mathcal{A})$. Since $\prec_\mathcal{A}\restriction_z \subset J_\mathcal{A}$ for all z, we have $\neg\mathcal{A}(x, \prec_\mathcal{A}\restriction_z)$ for all z by monotonicity and thus $z \prec_\mathcal{A} x$ for all $z \in J_\mathcal{A}$ by (SC1). For the converse, if $x \in J_\mathcal{A}$ then $J_\mathcal{A} \not\subset \prec_\mathcal{A}\restriction_x$ since $x \notin \prec_\mathcal{A}\restriction_x$ by irreflexivity. ⊣

Lemma 4.4. $\mathsf{SC}_1^- \vdash$ *"$\prec_\mathcal{A}$ is transitive", for every* $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)$.

Proof. It suffices to show by $\prec_\mathcal{A}$-induction on x, using (SC2), that

$$\forall y \forall z \big(z \prec_\mathcal{A} y \wedge y \prec_\mathcal{A} x \;\to\; z \prec x\big), \text{ for all } x.$$

Let $z \prec_\mathcal{A} y$ and $y \prec_\mathcal{A} x$. We have $z \in J_\mathcal{A}$ and $\neg\mathcal{A}(x, \prec_\mathcal{A}\restriction_y)$ by (SC1). Since we have $\prec_\mathcal{A}\restriction_z \subset \prec_\mathcal{A}\restriction_y$ by the induction hypothesis, we obtain $\neg\mathcal{A}(x, \prec_\mathcal{A}\restriction_z)$ by monotonicity and thus $z \prec x$ by (SC1). ⊣

Lemma 4.5. $\mathsf{SC}_1^- \vdash \forall x \forall y \big(x \prec_\mathcal{A} y \vee y \prec_\mathcal{A} x \vee \prec_\mathcal{A}\restriction_x = \prec_\mathcal{A}\restriction_y\big)$, *for any* $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)$.

Proof. We can assume $x, y \in J_\mathcal{A}$; otherwise the claim follows from Lemma 4.3.2. We will show by $\prec_\mathcal{A}$-induction on x with side $\prec_\mathcal{A}$-induction on y that

$$x \prec_\mathcal{A} y \vee y \prec_\mathcal{A} x \vee \prec_\mathcal{A}\restriction_x = \prec_\mathcal{A}\restriction_y, \text{ for all } x \in J_\mathcal{A} \text{ and } y \in J_\mathcal{A}.$$

Assume $x \not\prec_\mathcal{A} y$ and $y \not\prec_\mathcal{A} x$. Take any $z \prec_\mathcal{A} y$. By transitivity, $x \prec z$ can't be the case. If $\prec\restriction_x = \prec\restriction_z$ were the case, then we would get $\neg\mathcal{A}(y, \prec\restriction_x)$ and thus $x \prec_\mathcal{A} y$ by (SC1). Hence, we obtain $z \prec_\mathcal{A} x$ by the sub-inductive hypothesis; we have shown $\prec\restriction_y \subset \prec\restriction_x$. The converse is shown parallelly but by using the main-induction hypothesis instead. ⊣

DEFINITION 4.6. We put $\preceq_{\mathcal{A}} := \{\langle x, y\rangle \mid \mathcal{A}(x, \prec_{\mathcal{A}}\restriction_y)\}$ for each $\mathcal{A} \in \Im(\mathcal{L}_\in)$, and write $x \preceq_{\mathcal{A}} y :\Leftrightarrow \langle x, y\rangle \in \preceq_{\mathcal{A}}$. Note that $\prec_{\mathcal{A}}$ occurs in $x \preceq_{\mathcal{A}} y$ only positively.

LEMMA 4.7. $\mathsf{SC}_1^- \vdash \forall x \forall y[x \preceq_{\mathcal{A}} y \leftrightarrow (x \in J_{\mathcal{A}} \wedge y \not\prec_{\mathcal{A}} x)]$; *by Lemma* 4.3.

The next theorem by Sato [22] is of crucial importance for the present article. The proof of the theorem is an ingenious modification of the Stage Comparison Theorem (see [19]) specially for base systems with a certain reflection property.

THEOREM 4.8 (Sato [22]). SC_1 *is a definitional extension of* $\widehat{\mathsf{ID}}_1$.

We close this section with one immediate consequence of Sato's theorem.

Burgess [5] presented an extension of $\mathsf{KF}[\![\mathsf{PA}]\!]$ over arithmetic, which augments $\mathsf{KF}[\![\mathsf{PA}]\!]$ with axioms expressing that T is the least fixed-point of the Kripkean operator with the strong Kleene evaluation schema, namely, the inductive operator form $\mathcal{T}(x, \mathfrak{X}) \in \Im(\mathcal{L}_\in)$ taken in the proof of Theorem 3.2.1. Thereby Burgess's system KF_μ over set theory is defined as KF plus the following schema:

$$Clos_{\mathcal{T}}(\Psi) \to \forall x(Tx \to \Psi(x)), \text{ for each } \Psi \in \mathcal{L}_{\mathrm{Fix}}.$$

Obviously KF_μ is interpretable in ID_1 simply by translating T to $J_{\mathcal{T}}$. Hence, it follows from Sato's Theorem and Theorem 3.2, we have the next theorem.

THEOREM 4.9. KF_μ *and* KF *are proof-theoretically equivalent.*

§5. **Basic facts of inductive classes provable in** SC_1^-**.** We will formalize some basic results of inductive relations (cf. [19]) within SC_1^-.

For a $(k+1)$-tuple $a = \langle a_0, \dots, a_k\rangle$ and $i \leq k$, we denote its $(i+1)$-th component a_i by $(a)_i$. Given a class X and $a \in \mathbb{V}$, we put $X^a = \{x \mid \langle x, a\rangle \in X\}$; note that we do *not* generally have $X^a = \{(z)_0 \mid z \in X\}$ unless $X \subset Pair$. We assume for simplicity that $(a)_i$ is defined for all sets $a \in \mathbb{V}$ and all $i < \omega$.

Until and including Proposition 5.4, we will work within ID_1^-.

DEFINITION 5.1. The following definition is made in ID_1^-. A class X is said to be *inductive*, if there is $\mathcal{A} \in \Im(\mathcal{L}_\in)$ such that $X = J^a_{\mathcal{A}}$ for some $a \in \mathbb{V}$; when this holds we say that X is defined by $\mathcal{A}$ with parameter a. We also say that X is *coinductive* when $-X := \{x \mid x \notin X\}$ is inductive, and that X is *hyperelementary* when X is both inductive and coinductive.

THEOREM 5.2 (Transitivity Theorem). *The following is provable in* ID_1^-. *Let* $\mathcal{A}(x, v_1, \dots, x_l, \mathfrak{X}, \mathfrak{Y}_1, \dots, \mathfrak{Y}_k) \in \mathcal{L}^2_\in$ *be elementary in which only the displayed variables are free and* $\mathfrak{X}, \mathfrak{Y}_1, \dots, \mathfrak{Y}_k$ *occur only positively. For every inductive* $Y_1, \dots, Y_k$ *and parameters* $a_1, \dots, a_l \in \mathbb{V}$, *there is an inductive* X *such that*

$$\forall x\big(\mathcal{A}(x, \vec{a}, X, \vec{Y}) \to x \in X\big). \tag{T1}$$

$$\forall x\big(\mathcal{A}(x, \vec{a}, Z, \vec{Y}) \to x \in Z\big) \to X \subset Z, \textit{ for all classes } Z. \tag{T2}$$

PROOF. Let Y_i be defined by $\mathcal{B}_i$ with b_i $(1 \leq i \leq k)$. Note that $\mathcal{A}(x, \vec{a}, \mathfrak{X}, \vec{Y})$ then contains $\vec{b}$ as parameters besides $\vec{a}$. Put $\mathcal{A}'(x, \mathfrak{X}) \in \Im(\mathcal{L}_\in)$ to be

$$\Big[((x)_1)_0 = 0 \to \mathcal{A}\Big((x)_0, ((x)_1)_1, \dots, ((x)_1)_l, \mathfrak{X}^{(x)_1}, (\mathfrak{X}^1)^{((x)_1)_{l+1}}, \dots, (\mathfrak{X}^k)^{((x)_1)_{l+k}}\Big)\Big]$$

$$\wedge \bigwedge_{1 \leq i \leq k} \Big((x)_1 = i \to \mathcal{B}_i\big((x)_0, \mathfrak{X}^i\big)\Big) \wedge x \in Pair;$$

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

then $J_{\mathcal{A}'} \subset Pair$. We can assume that if $((x)_1)_0 = 0$ then $(x)_1 \neq i$ for all $i > 0$.

We first show $J^i_{\mathcal{A}'} = J_{\mathcal{B}_i}$ for $1 \leq i \leq k$. Let $B_i = (J_{\mathcal{B}_i} \times \{i\}) \cup (\mathbb{V} \times (\mathbb{V} \setminus \{i\}))$. We have $Clos_{\mathcal{A}'}(B_i)$ and thus $J_{\mathcal{A}'} \subset B_i$; hence $J^i_{\mathcal{A}'} \subset B^i_i = J_{\mathcal{B}_i}$. If $\mathcal{B}_i(x, J^i_{\mathcal{A}'})$, then $\mathcal{A}'(\langle x, i\rangle, J_{\mathcal{A}'})$ and thus $\langle x, i\rangle \in J_{\mathcal{A}'}$; hence $Clos_{\mathcal{B}_i}(J^i_{\mathcal{A}'})$ and thus $J_{\mathcal{B}_i} \subset J^i_{\mathcal{A}'}$.

We have shown $Y_i = (J^i_{\mathcal{A}'})^{b_i}$ for $1 \leq i \leq k$. Let $c = \langle 0, a_1, \dots, a_l, b_1, \dots, b_k\rangle$ and $X := J^c_{\mathcal{A}'}$. For (T1), if $\mathcal{A}(x, \vec{a}, X, \vec{Y})$ then $\mathcal{A}'(\langle x, c\rangle, J_{\mathcal{A}'})$ and thus $x \in J^c_{\mathcal{A}'}$. For (T2) suppose $\forall x(\mathcal{A}(x, \vec{a}, Z, \vec{Y}) \to x \in Z)$ for a class Z. Put $Z' := (Z \times \{c\}) \cup \{x \in J_{\mathcal{A}'} \mid (x)_1 \neq c\}$. We have $Clos_{\mathcal{A}'}(Z')$, since $\mathcal{A}'(x, Z')$ and $(x)_1 \neq c$ implies $\mathcal{A}'(x, J_{\mathcal{A}'})$ and thus $x \in Z'$. Hence $J_{\mathcal{A}'} \subset Z'$ and thus $X \subset Z$. ⊣

We say that $\mathcal{C}(x, \mathfrak{X}) \in \mathcal{L}^2_{\mathrm{SC}}$ possibly with parameters is *positive elementary in* classes $Y_1, \dots, Y_k$, when there are some $\vec{a} \in \mathbb{V}$ and $\mathcal{A}(x, \vec{v}, \mathfrak{X}, \mathfrak{Y}_1, \dots, \mathfrak{Y}_k) \in \mathcal{L}^2_{\in}$ with at most the displayed variables free and only with positive occurrences of $\mathfrak{X}, \mathfrak{Y}_1, \dots, \mathfrak{Y}_k$ such that $\mathcal{C}(x, \mathfrak{X}) \leftrightarrow \mathcal{A}(x, \vec{a}, \mathfrak{X}, Y_1, \dots, Y_k)$. Hence, Theorem 5.2 says that every positive elementary $\mathcal{C}(x, \mathfrak{X})$ in inductive classes has an inductive least fixed-point provably in ID^-_1, and we denote it by $J_{\mathcal{C}}$.

We will occasionally treat classes of n-tuples ($n \geq 2$) as if they were n-ary predicates (or relations) and write $P(x_1, \dots, x_n)$, $Q(x_1, \dots, x_m)$, etc.

Corollary 5.3. *In* ID^-_1, *the collection of inductive relations is closed under conjunction, disjunction, existential, and universal quantification.*

A relation R is said to be *elementary on* classes $X_1, \dots, X_k$, if R is constructed from $X_1, \dots, X_k$, $=$, and $\in$, by $\neg$, $\wedge$, $\exists$, and $\forall$. We simply say X is *elementary* if X is elementary on $\mathbb{V}$, which is obviously hyperelementary.

Corollary 5.4. *In* ID^-_1, *if* $X_1, \dots, X_k$ *are hyperelementary and* R *is elementary on* $X_1, \dots, X_k$, *then* R *is hyperelementary.*

From now on, we will work within SC^-_1 in the rest of the present section.

Proposition 5.5. *For* $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_{\in})$ *and* $a \in \mathbb{V}$, *we set* $x \prec_{\mathcal{A},a} y :\Leftrightarrow \langle x, a\rangle \prec_{\mathcal{A}} \langle y, a\rangle$. *This* $\prec_{\mathcal{A},a}$ *strictly prewellorders* $J^a_{\mathcal{A}}$: *i.e., it is irreflexive, transitive, and*

$$(\forall x \in J^a_{\mathcal{A}})\big[\forall y(y \prec_{\mathcal{A},a} x \to y \in Y) \to x \in Y\big] \to J^a_{\mathcal{A}} \subset Y, \textit{ for all classes } Y.$$

We also define $x \preceq_{\mathcal{A},a} y :\Leftrightarrow \langle x, a\rangle \preceq_{\mathcal{A}} \langle y, a\rangle$, *which is transitive and well-founded.*

The way in which $\prec_{\mathcal{A},a}$ prewellorders $X = J^a_{\mathcal{A}}$ depends on the choice of $\mathcal{A}$ and a, but the choice of the pair do not matter for our subsequent argument and so we let $\prec_X$ denote $\prec_{\mathcal{A},a}$ for some fixed $\mathcal{A}$ and a defining X.

Proposition 5.6. *Let* X *be inductive. By Lemmata* 4.3 *and* 4.7, *we have*

1. $x \prec_X y$ *implies* $x \in X$, *and* $x \preceq_X y$ *implies* $x \in X$;
2. $y \notin X$ *implies* $X \subset {\prec_X}\restriction_y$ *and* $X \subset {\preceq_X}\restriction_y$;
3. $x \prec_X y$ *iff* $x \in X \wedge y \not\preceq_X x$, *and* $x \preceq_X y$ *iff* $x \in X \wedge y \not\prec_X x$.

Theorem 5.7 (Stage Comparison Theorem). *The relation* $\prec_{\mathcal{A}}$ *is inductive. Hence, by Corollaries* 5.3 *and* 5.4, $\preceq_{\mathcal{A}}$ *is also inductive.*

Proof. Let $\mathcal{B}(x, \mathfrak{X}) := \mathcal{A}(x, \mathfrak{X}) \vee \mathfrak{X} = \mathbb{V}$, and set $\mathcal{B}'(x, \mathfrak{X}) \in \mathfrak{I}(\mathcal{L}_{\in})$ to be

$$x \in Pair \wedge \neg\mathcal{B}\big((x)_1, \big\{u \mid \neg\mathcal{B}((x)_0, \{v \mid \langle v, u\rangle \in \mathfrak{X}\big)\big\}\big);$$

note that we then have $J_{\mathcal{B}'} \subset Pair$. We will show $\prec_{\mathcal{A}} = J_{\mathcal{B}'}$.

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

For one direction, $J_{\mathcal{B}'} \subset \prec_{\mathcal{A}}$, it suffices to show that $Clos_{\mathcal{B}'}(\prec_{\mathcal{A}})$, which follows from the following equivalences: for every x and y,

$$\begin{aligned}\neg\mathcal{B}\big(y, \{u \mid \neg\mathcal{B}(x, \prec_{\mathcal{A}}\restriction_u)\}\big) &\Leftrightarrow \neg\mathcal{B}\big(y, \{u \mid \neg\mathcal{A}(x, \prec_{\mathcal{A}}\restriction_u)\}\big) \\ &\overset{4.3}{\Leftrightarrow} \big[x \notin J_{\mathcal{A}} \wedge \neg\mathcal{B}(y, \mathbb{V})\big] \vee \big[x \in J_{\mathcal{A}} \wedge \neg\mathcal{B}(y, \prec_{\mathcal{A}}\restriction_x)\big] \\ &\Leftrightarrow x \in J_{\mathcal{A}} \wedge \neg\mathcal{A}(y, \prec_{\mathcal{A}}\restriction_x) \Leftrightarrow x \prec_{\mathcal{A}} y;\end{aligned}$$

the first and third equivalences obtain since $\prec_{\mathcal{A}}\restriction_u \neq \mathbb{V}$ for all u by irreflexivity.

For the converse, $\prec_{\mathcal{A}} \subset J_{\mathcal{B}'}$, it suffices to show $\forall y\big(x \prec_{\mathcal{A}} y \to \langle x, y\rangle \in J_{\mathcal{B}'}\big)$ for all $x \in J_{\mathcal{A}}$ by induction along $\prec_{\mathcal{A}}$. Let $y \succ_{\mathcal{A}} x$. We will show that $\mathcal{B}'(\langle x, y\rangle, J_{\mathcal{B}'})$. Take any $u \not\prec_{\mathcal{A}} x$. By Lemma 4.5 we have $\prec_{\mathcal{A}}\restriction_x \subset \prec_{\mathcal{A}}\restriction_u$. Hence, for all $v \prec_{\mathcal{A}} x$, we have $\langle v, u\rangle \in J_{\mathcal{B}'}$ by IH and thus $\mathcal{A}(x, \{v \mid \langle v, u\rangle \in J_{\mathcal{B}'}\})$ by Lemma 4.3.1 and monotonicity. Since $u \not\prec_{\mathcal{A}} x$ was arbitrary, we obtain

$$\big\{u \mid \neg\mathcal{B}\big(x, \{v \mid \langle v, u\rangle \in J_{\mathcal{B}'}\}\big)\big\} \subset \big\{u \mid \neg\mathcal{A}\big(x, \{v \mid \langle v, u\rangle \in J_{\mathcal{B}'}\}\big)\big\} \subset \prec_{\mathcal{A}}\restriction_x \neq \mathbb{V}.$$

Since $x \prec_{\mathcal{A}} y$ implies $\neg\mathcal{A}(y, \prec_{\mathcal{A}}\restriction_x)$, we obtain the claim by monotonicity. ⊣

THEOREM 5.8 (Hyperelementary selection theorem). *Let $P(x, y)$ be an inductive relation. There are inductive $Q(x, y)$ and coinductive $\check{Q}(x, y)$ such that*

(i) $Q \subset P$; (ii) $\exists y P(x, y) \to \exists y Q(x, y)$; (iii) $\exists y P(x, y) \to \forall y[Q(x, y) \leftrightarrow \check{Q}(x, y)]$.

PROOF. We define

$$Q := \{\langle x, y\rangle \mid \forall z(\langle x, y\rangle \preceq_P \langle x, z\rangle)\} \quad \text{and} \quad \check{Q} := \{\langle x, y\rangle \mid \forall z(\langle x, z\rangle \not\prec_P \langle x, y\rangle)\}.$$

Then (i) and (iii) follow from Proposition 5.6. For (ii), suppose $\exists y P(x, y)$ and put $u \prec_{x,P} v :\Leftrightarrow \langle x, u\rangle \prec_P \langle x, v\rangle$. Then $\prec_{x,P}$ prewellorders $\{w \mid \langle x, w\rangle \in P\} \neq \emptyset$ and we can pick a $\prec_{x,P}$-minimal element y'; hence, we get $\langle x, y'\rangle \in Q$. ⊣

THEOREM 5.9 (Covering theorem). *Let X be inductive but not coinductive, and let Y be coinductive. Let R be a hyperelementary relation such that $dom(R) \supset Y$ and $R[Y] \subset X$, where $R[Y]$ is the image $\{x \mid \exists y[y \in Y \wedge R(y, x)]\}$ of Y by R. Then, $(\exists c \in X)\forall x\big(x \in R[Y] \to x \preceq_X c\big)$.*

PROOF. Otherwise X would become coinductive, since it would hold that

$$c \in X \Leftrightarrow (\exists y \in Y)\exists x\big(R(y, x) \wedge x \not\preceq_X c\big). \qquad ⊣$$

THEOREM 5.10 (Good parametrization theorem for inductive classes). *There exist an inductive class U and elementary function $S : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ such that*

1. *for all inductive classes X, there is some $a \in \mathbb{V}$ such that $U^a = X$, and*
2. *for all inductive classes X and $a \in \mathbb{V}$, if $U^a = X$ then $\forall c(U^{S(a,c)} = X^c)$.*

PROOF. It is known that for each $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)$ there is $\mathcal{B} \in \mathfrak{I}(\mathcal{L}_\in) \cap \Pi^0_2$ such that $J_{\mathcal{A}} = J^p_{\mathcal{B}}$ for some p; see [22, Section 3]. Take a universal Π^0_2-inductive operator form $\mathcal{U}$ such that $\exists q(J^q_{\mathcal{U}} = J_{\mathcal{B}})$ for all $\mathcal{B} \in \mathfrak{I}(\mathcal{L}_\in) \cap \Pi^0_2$. Hence, we have $\exists p \exists q\big((J^q_{\mathcal{U}})^p = J_{\mathcal{A}}\big)$ for all $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)$; we can assume here that $p, q \in \mathbb{N}$ and they can be primitive recursively computed from given $\mathcal{A}$. Then we take

$$U := \big\{\langle x, \langle a, p, q\rangle\rangle \mid \big\langle\big\langle\big\langle\langle x, (a)_0\rangle, (a)_1\big\rangle, p\big\rangle, q\big\rangle \in J_{\mathcal{U}}\big\} \subset Pair.$$

Since the class $\{\langle x, d\rangle \mid x \in X\}$ with any "dummy" index d (e.g., 0) is inductive for all inductive X, we can easily verify the Claim 1. For the Claim 2, let $Q :=$

$\{\langle x, \langle y, z\rangle\rangle \mid \langle\langle x, y\rangle, z\rangle \in U\}$. Q is inductive and thus there are p, q, b such that $Q = ((J^q_{\mathcal{U}})^p)^b$ and fix any such p, q, b. Then,

$$\langle x, c\rangle \in U^a \Leftrightarrow \langle x, \langle c, a\rangle\rangle \in Q \Leftrightarrow \langle x, \langle c, a\rangle\rangle \in ((J^q_{\mathcal{U}})^p)^b \Leftrightarrow \big\langle x, \big\langle\langle\langle c, a\rangle, b\rangle, p, q\big\rangle\big\rangle \in U.$$

Hence we can take $S(a, c) := \big\langle\langle\langle c, a\rangle, b\rangle, p, q\big\rangle$. ⊣

LEMMA 5.11. *The class U taken in Theorem 5.10 is not coinductive.*

PROOF. If U was coinductive, $P := \{x \mid \langle x, x\rangle \notin U\}$ would be inductive and thus there would be some a such that $a \in U^a \Leftrightarrow a \in P \Leftrightarrow a \notin U^a$. ⊣

THEOREM 5.12 (Good parametrization theorem for hyperelementary classes). *There exist inductive classes I and H, and a coinductive class $\check{H}$ such that*

(i) *if $a \in I$ then $H^a = \check{H}^a$ (and thus H^a is hyperelementary for all $a \in I$);*
(ii) *if X is hyperelementary then $X = H^a$ for some $a \in I$;*
(iii) *For any inductive P and coinductive Q, there exists a hyperelementary function $J : \mathbb{V} \to \mathbb{V}$ such that if $P^a = Q^a$ then $J(a) \in I$ and $H^{J(a)} = P^a$.*

PROOF. Let U be the class taken in Theorem 5.10. We take $I, H, \check{H}$ so that

$$\begin{array}{lcl} a \in I & :\Leftrightarrow & (a)_1 \in U \\ \langle x, a\rangle \in H & :\Leftrightarrow & a \in I \wedge \big\langle x, (a)_0\big\rangle \preceq_U (a)_1 \\ \langle x, a\rangle \in \check{H} & :\Leftrightarrow & (a)_1 \not\prec_U \big\langle x, (a)_0\big\rangle. \end{array}$$

The claim (i) is obvious by Proposition 5.6.

For (ii), let X be hyperelementary and $X = U^b$. Then $R := \{\langle x, \langle x, b\rangle\rangle \mid x \in X\}$ is a hyperelementary relation. By Theorem 5.9 and Lemma 5.11 we can pick $c \in U$ with $\forall x(x \in X \to \langle x, b\rangle \preceq_U c)$. Hence we can take $a := \langle b, c\rangle \in I$, since

$$x \in H^a \;\Leftrightarrow\; a \in I \wedge \langle x, b\rangle \preceq_U c \;\Leftrightarrow\; x \in U^b = X.$$

For (iii), let $P = U^b$ be inductive and Q be coinductive. Let us put

$$Z := \big\{\langle x, y\rangle \mid \forall u\big(u \in Q^y \to \langle u, S(b, y)\rangle \preceq_U \langle x, S(x, y)\rangle\big)\big\},$$

which is inductive, and pick c such that $U^c = Z$. Then we define the function J by $J(a) := \big\langle S(b, a), \langle c, S(c, a)\rangle\big\rangle$. Suppose $P^a = Q^a$ $(= U^{S(b,a)})$. We have $c \in U^{S(c,a)} = Z^a$; for, $c \notin U^{S(c,a)}$ and $\langle u, S(b, a)\rangle \not\preceq_U \langle c, S(c, a)\rangle$ implies $\langle u, S(b, a)\rangle \notin U$; hence $J(a) \in I$. Thereby we also get $P^a = \big\{u \mid \langle u, S(b, a)\rangle \preceq_U \langle c, S(c, a)\rangle\big\}$, and thus, for all u,

$$u \in H^{J(a)} \;\Leftrightarrow\; J(a) \in I \wedge \langle u, S(b, a)\rangle \preceq_U \langle c, S(c, a)\rangle \;\Leftrightarrow\; u \in P^a. \qquad ⊣$$

§6. Kripke-Platek set theory over $\mathbb{V}$. We will consider a Kripke-Platek set theory with urelements, where the set-theoretic universe $\mathbb{V}$ (or a fixed model of ZF, more formally) is taken as the domain of urelements, and in which "higher-order" sets are constructed by the KP-axioms and topped up on $\mathbb{V}$, while keeping the distinction of the sets in $\mathbb{V}$ (as "urelements") and the sets added on top of $\mathbb{V}$ (as "sets"); we also assume that the collection of urelements, $\mathbb{V}$, forms a set and we have a constant V for it. In terms of [2, Chapter I.2], our system is KPU^+ with ZF as the theory of urelements augmented with a constant for the set of urelements.

We take the one-sort formulation of KPU^+. Let $\mathcal{L}_{\mathrm{KP}} = \{\in_0, \in_1, \mathcal{U}, \mathsf{V}\}$, where $\mathcal{U}$ is a unary predicate for urelements, $\in_0$ is the membership relation among urelements,

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

$\in_1$ is the membership relation for sets, and V is a constant symbol for the set of urelements. We will write $\mathcal{S}x$ for $\neg\mathcal{U}x$ to express set-hood. As in the previous sections, we will occasionally treat $\mathcal{L}_{\mathrm{KP}}$-formulae as classes, and write $x \in \mathcal{U}$ and $x \in \mathcal{S}$ for example; the use of the symbol "$\in$" here should not be confused with "$\in_0$" or "$\in_1$", which are in the vocabulary of $\mathcal{L}_{\mathrm{KP}}$.

We standardly define the collection of Δ_0-formulae as the smallest collection of $\mathcal{L}_{\mathrm{KP}}$-formulae that contains all atomic $\mathcal{L}_{\mathrm{KP}}$-formulae and is closed under Boolean connectives and bounded quantifiers $(\forall z \in_1 x)$ and $(\exists z \in_1 x)$. The other collections of $\mathcal{L}_{\mathrm{KP}}$-formulae in the Levy hierarchy are standardly defined from Δ_0.

For each $\varphi \in \mathcal{L}_{\in}$, we denote the relativization of φ to $\langle \mathcal{U}, \in_0 \rangle$ by $\varphi^{\mathcal{U}}$ $(\in \mathcal{L}_{\mathrm{KP}})$, where all the quantifiers $\forall x$ and $\exists x$ are restricted to $\mathcal{U}$ and the membership relation $\in$ of $\mathcal{L}_{\in}$ is replaced by $\in_0$; accordingly, $\mathsf{ZF}^{\mathcal{U}}$ means $\{\sigma^{\mathcal{U}} \mid \sigma \in \mathsf{ZF}\}$.

Definition 6.1. The $\mathcal{L}_{\mathrm{KP}}$-system $\mathsf{KP}\mathbb{V}^-$ comprises $\mathsf{ZF}^{\mathcal{U}}$ plus:

(Ext): $(\forall a, b \in \mathcal{S})\big(\forall x(x \in_1 a \leftrightarrow x \in_1 b) \to a = b\big)$,

(Found$_1$): $\forall x\big((\forall y \in_1 x)\varphi(y) \to \varphi(x)\big) \to \forall x \varphi(x)$,

(Pair): $\forall x \forall y (\exists a \in \mathcal{S})\big(x \in_1 a \wedge y \in_1 a\big)$,

(Union): $(\forall a \in \mathcal{S})(\exists b \in \mathcal{S})(\forall x \in_1 a)(\forall y \in_1 x) y \in_1 b$,

(Δ_0-Sep$_1$): $(\forall a \in \mathcal{S})(\exists b \in \mathcal{S})\forall x\big(x \in_1 b \leftrightarrow x \in_1 a \wedge \psi(x)\big)$,

(Δ_0-Coll$_1$): $(\forall a \in \mathcal{S})\big[(\forall x \in_1 a)\exists y \psi(x, y) \to (\exists b \in \mathcal{S})(\forall x \in_1 a)(\exists y \in_1 b)\psi(x, y)\big]$,

(U): $\mathsf{V} \in \mathcal{S} \wedge \forall x \forall y\big((x \in_1 \mathsf{V} \leftrightarrow x \in \mathcal{U}) \wedge (x \in_1 y \to y \in \mathcal{S}) \wedge (x \in_0 y \to x, y \in \mathcal{U})\big)$,

(Eq) : $\forall x(x = x)$ and $\forall x \forall y\big[x = y \to \big(\xi(x) \leftrightarrow \xi(y)\big)\big]$,

where φ is any $\mathcal{L}_{\mathrm{KP}}$-formula, ψ is any Δ_0-formula without b free, and ξ is any atomic $\mathcal{L}_{\mathrm{KP}}$-formula. For each ZF-axiom σ, its relativization $\sigma^{\mathcal{U}}$ is (equivalently) Δ_0 due to the axiom (U).

We also consider the following additional axiom schemata.

$$(\text{Found}_0^+) : \ (\forall x \in \mathcal{U})\big((\forall x \in_0 y)\varphi(y) \to \varphi(x)\big) \to (\forall x \in \mathcal{U})\varphi(x).$$

$$(\text{Sep}_0^+) : \ (\forall a \in \mathcal{U})(\exists b \in \mathcal{U})(\forall z \in \mathcal{U})\big(z \in_0 y \leftrightarrow z \in_0 x \wedge \varphi(z)\big).$$

$$(\text{Repl}_0^+) : \ (\forall a \in \mathcal{U})\big[(\forall x \in_0 a)(\exists! y \in \mathcal{U})\varphi \to (\exists b \in \mathcal{U})(\forall x \in_0 a)(\exists y \in_0 b)\varphi\big],$$

where φ is any $\mathcal{L}_{\mathrm{KP}}$-formula. Then we set $\mathsf{KP}\mathbb{V} := \mathsf{KP}\mathbb{V}^- + (\text{Sep}_0^+) + (\text{Repl}_0^+)$.[1]

We express various sets and classes in $\mathcal{L}_{\in}$, such as $\emptyset$, ω, the class *Tran* of transitive sets, the class *On* of ordinals, etc. Now, $\mathcal{L}_{\mathrm{KP}}$ possesses two different membership relations $\in_0$ and $\in_1$ and bears two different set-theoretic structures $\langle \mathcal{U}, \in_0 \rangle$ and $\langle \mathcal{S}, \mathcal{U}, \in_1 \rangle$ (where $\mathcal{U}$ gives the domain of urelements). Hence, those sets and classes can be expressed in two different ways in terms of $\in_0$ and $\in_1$. We will distinguish them by attaching superscript $\mathcal{U}$ or $\mathcal{S}$; for example, $\emptyset^{\mathcal{U}}$ denotes the empty set in $\langle \mathcal{U}, \in_0 \rangle$ such that $\emptyset^{\mathcal{U}} \in \mathcal{U}$ and $(\forall x \in \mathcal{U}) x \notin_0 \emptyset^{\mathcal{U}}$, and $\emptyset^{\mathcal{S}}$ denotes the empty set in $\langle \mathcal{S}, \mathcal{U}, \in_1 \rangle$ such that $\emptyset^{\mathcal{S}} \in \mathcal{S}$ and $\forall x(x \notin_1 \emptyset^{\mathcal{S}})$; $Tran^{\mathcal{U}}$ denotes the class $\{x \in \mathcal{U} \mid (\forall u \in_0 x)(v \in_0 u) v \in_0 x\}$ of transitive sets in $\langle \mathcal{U}, \in_0 \rangle$, and $Tran^{\mathcal{S}}$ is the class $\{x \in \mathcal{S} \mid (\forall u \in_1 x)(v \in_1 u) v \in_1 x\}$ of transitive sets in $\langle \mathcal{S}, \mathcal{U}, \in_1 \rangle$.

[1] $\mathsf{KP}\mathbb{V}^-$ does not derive the axiom of infinity for sets in $\mathcal{S}$, but $\mathsf{KP}\mathbb{V}$ does due to (Found_0^+).

§7. Reduction of $\mathsf{KP}\mathbb{V}^-$ to SC_1^-. We will give an embedding $*$ of $\mathsf{KP}\mathbb{V}^-$ in SC_1^-. It will be done by formalizing and generalizing the Barwise-Gandy-Moschovakis theorem [3]. We work within SC_1^- throughout the present section.

The interpretations $\mathcal{U}^*$ and $\in_0^*$, of the domain $\mathcal{U}$ of urelements and the membership relation $\in_0$ for urelements in $\mathsf{KP}\mathbb{V}^-$, are given by

$$\mathcal{U}^* := \{\langle a, 0\rangle \mid a \in \mathbb{V}\} \quad \text{and} \quad x \in_0^* y :\Leftrightarrow x \in \mathcal{U}^* \wedge y \in \mathcal{U}^* \wedge (x)_0 \in (y)_0;$$

note that both are elementary. To give the interpretations of $=$ and $\in_1$, we need some preliminary definitions and results that we will explain at length below.

We say a class T is a *tree* when the following holds

$$T \neq \emptyset \wedge T \subset Seq \wedge \forall x \forall y \big[\big(x, y \in Seq \wedge x * y \in T\big) \to x \in T\big],$$

where Seq is the (elementary) class of finite sequences (or tuples), and $x * y$ denotes the concatenation of the two sequences x and y. For $x \in Seq$, we denote its length by $lh(x)$ $(\in \omega)$ and its $(i+1)$-th component $(i < lh(x))$ by $(x)_i$ as in Section 5. We include the empty sequence ϵ in Seq so that ϵ is the unique sequence with length 0, every nonempty $x \in Seq$ is a proper extension of ϵ, and $\epsilon * x = x = x * \epsilon$ for all $x \in Seq$; hence, ϵ is a member of every tree; for technical convenience, we stipulate that $\epsilon \notin \mathcal{U}^*$ and $(u)_{-1} = \epsilon$ for each $u \in Seq$.

For a class Y, we define a strict preordering $\sqsubset_Y$ by

$$x \sqsubset_Y y :\Leftrightarrow x, y \in Y \wedge x, y \in Seq \wedge (\text{"}x \text{ is a proper extension of } y\text{"});$$

note that ϵ is always the maximum element ("root") of $\sqsubset_Y$ if Y is a tree.

For a binary relation R, we let $\mathcal{W}[R](x, \mathfrak{X}, R) \in \mathcal{L}^2_{\mathrm{SC}}$ be $\forall y(yRx \to y \in \mathfrak{X})$. Since $\mathcal{W}[R]$ is positive elementary in $-R$, the inductive class $J_{\mathcal{W}[R]}$ exists for every *coinductive* R by Theorem 5.2 and expresses the *accessible part* of R, which we will denote by $Acc(R)$. For a coinductive tree T, $\sqsubset_T$ is also coinductive and thus its well-foundedness is expressed as $\epsilon \in Acc(\sqsubset_T)$ $(\leftrightarrow \mathbb{V} = Acc(\sqsubset_T))$; when this holds, T is said to be *well-founded*. When T is well-founded, we have

$$(\forall u \in T)\big((\forall v \sqsubset_T u)\, v \in X \to u \in X\big) \to T \subset X, \text{ for all classes } X.$$

Let $\min(\sqsubset_T) := \{u \in T \mid \forall x(u * \langle x\rangle \notin T)\}$ (i.e., the class of "leaves" of T); this class is elementary on T. Then we define two classes both also elementary on T:

$$\mathcal{U}(T) := \{u \in T \mid u \in \min(\sqsubset_T) \wedge (u)_{lh(u)-1} \in \mathcal{U}^*\} \quad \text{and} \quad \mathcal{S}(T) := T \setminus \mathcal{U}(T);$$

note that $(u)_{lh(u)-1}$ is the last component of a sequence $u = \langle u_0, \ldots, u_{lh(u)-1}\rangle$.

For interpreting the domain $\mathcal{S}$ of sets of $\mathsf{KP}\mathbb{V}^-$ in SC_1^-, we make use of the so-called tree representation of well-founded sets: we let each well-founded tree T represent the unique well-founded set b such that $\langle \mathrm{TC}(\{b\}), \in\rangle$ is the Mostowski collapse of $\langle T, \sqsubset_T\rangle$; hence, bisimilar well-founded trees represent the same well-founded set, say, c, and those trees are also bisimlar to the *canonical* tree representation (or, tree picture) of c, defined as $\{\epsilon\} \cup \{\langle c_1, \ldots, c_k\rangle \mid c_k \in \cdots \in c_1 \in c\}$ (see [1] for a detailed exposition). However, since we allow urelements in $\mathsf{KP}\mathbb{V}^-$, the notions of collapse and bisimulation must be so modified as to accommodate urelements; each leaf of a well-founded tree corresponds to an object with no member that is contained in the transitive closure of the set represented by the tree, and we must somehow distinguish the cases where the leaf represents the emptyset and where it represents an urelement, both of which contain no element. For this purpose, we stipulate that, for a leaf u of a tree T, if $u = \langle u_0, \dots, u_k \rangle$ ends with an element of the form $u_k = \langle x, 0 \rangle \in \mathcal{U}^*$ (and thus $u \in \mathcal{U}(T)$), then it represents the urelement $x^{\mathcal{U}} \in \mathcal{U}$ of $\mathsf{KP}\mathbb{V}^-$, and otherwise represents $\emptyset^{\mathcal{S}}$.

We first define an inductive class M so that

$$a \in M \;:\Leftrightarrow\; a \in I \text{ and } H^a \;(= \check{H}^a) \text{ is a well-founded tree,}$$

where I, H, and $\check{H}$ are the inductive and coinductive classes taken in Theorem 5.12. We have to make sure that M can be properly defined. Let us put

$$x \sqsubset y \;:\Leftrightarrow\; x, y \in \mathit{Pair} \wedge (x)_1 = (y)_1 \wedge (x)_0 \sqsubset_{\check{H}^{(x)_1}} (y)_0.$$

We can take $\mathit{Acc}(\sqsubset)$ since $\sqsubset$ is coinductive. Let us write $x \sqsubset_a y$ for $\langle x, a \rangle \sqsubset \langle y, a \rangle$. Then we can show that, for all $a \in I$ and $x, y \in \mathbb{V}$,

$$x \sqsubset_{H^a} y \Leftrightarrow x \sqsubset_a y \qquad \text{and} \qquad \epsilon \in \mathit{Acc}(\sqsubset_a) \Leftrightarrow \langle \epsilon, a \rangle \in \mathit{Acc}(\sqsubset);$$

hence we can take $M = \{a \in I \mid \text{“}H^a \text{ is a tree”} \wedge \langle \epsilon, a \rangle \in \mathit{Acc}(\sqsubset)\}$, and $\sqsubset_a$ prewellorders H^a uniformly for each $a \in M$. The interpretation of the domain $\mathcal{S}$ of sets is thereby given as:

$$\mathcal{S}^* := \{\langle a, 1 \rangle \mid a \in M\};$$

we add the index “1” here, in contrast to “0” added for $\mathcal{U}^*$, to make $\mathcal{U}^*$ and $\mathcal{S}^*$ disjoint. Accordingly, the quantifiers “$\forall v$” and “$\exists v$” of $\mathcal{L}_{\mathrm{KP}}$ are interpreted by $*$ into “$\forall v \in (\mathcal{S}^* \cup \mathcal{U}^*)$” and “$\exists v \in (\mathcal{S}^* \cup \mathcal{U}^*)$”; note that the interpretations $\in_1^*$ and $=^*$ still remain to be defined, and their definitions will be given later.

We also have to modify the notion of the restriction of a tree T to its node u (or “sub-tree of T below u”) so as to accommodate urelements. Preliminarily, for a tree T and $u \in \mathcal{S}(T)$ we put $T_u := \{v \mid u * v \in T\}$, which is also a tree.

Proposition 7.1. *Let T be a coinductive tree and $u \in \mathcal{S}(T)$. If T_u is well-founded then $u \in \mathit{Acc}(\sqsubset_T)$.*

Proof. We can show $(\forall v \in T_u)\big(u * v \in \mathit{Acc}(\sqsubset_T)\big)$ by induction on $\sqsubset_{T_u}$. ⊣

Lemma 7.2. *There exists an elementary function $j : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ such that $(H^a)_u = H^{j(a,u)}$ for all $a \in M$ and $u \in \mathcal{S}(H^a)$.*

Proof. We apply Theorem 5.12 to the following P and Q:

$$P := \{\langle v, \langle a, u \rangle\rangle \mid u * v \in H^a \;\&\; u \in \mathcal{S}(H^a)\} \quad \text{and}$$
$$Q := \{\langle v, \langle a, u \rangle\rangle \mid u * v \in \check{H}^a \;\&\; u \in \mathcal{S}(\check{H}^a)\}.$$

Since we have $P^{\langle a,u \rangle} = Q^{\langle a,u \rangle}$ for all $a \in M$ and $u \in \mathcal{S}(H^a)$, there exists J such that $J(\langle a, u \rangle) \in M$ and $(H^a)_u = P^{\langle a,u \rangle} = H^{J(\langle a,u \rangle)}$ for all $a \in M$ and $u \in \mathcal{S}(H^a)$. So we can take $j(a, u) := J(\langle a, u \rangle)$. ⊣

For each $x = \langle a, 1 \rangle \in \mathcal{S}^*$ and $u \in H^a$, we define $x \downarrow_u \in \mathcal{S}^* \cup \mathcal{U}^*$ so that

$$x \downarrow_u := \begin{cases} (u)_{lh(u)-1} \;\; (\in \mathcal{U}^*) & \text{if } u \in \mathcal{U}(H^a), \\ \langle j(a, u), 1 \rangle \;\; (\in \mathcal{S}^*) & \text{if } u \in \mathcal{S}(H^a). \end{cases}$$

Remark 7.3. Let us give an informal explanation of the definitions given so far. Fix a transitive model $\mathfrak{A} = \langle A, \in \rangle$ of ZF. On the one hand, by treating A as the set

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

of urelements, the universe $\mathbb{V}_A$ of sets *on* A (see [2, p. 42]) and A gives a model of $\mathsf{KP}\mathbb{V}^-$, where we interpret $\mathcal{S}$ and $\mathcal{U}$ by $\mathbb{V}_A$ and A, respectively. On the other hand, we extend $\mathfrak{A}$ to a model of SC_1^- and define $I, H, \check{H}$ $(\subset A)$ on $\mathfrak{A}$ in the standard manner; thereby we define $\mathcal{S}^*$ and $\mathcal{U}^*$ in terms of them. Then, for $x = \langle a, 1\rangle \in \mathcal{S}^*$ and $u \in H^a$, where H^a is a well-founded tree, we define a set $m(H^a, u) \in \mathbb{V}_A \cup A$ by recursion along $\sqsubset_{H^a}$:

$$m(H^a, u) := \begin{cases} p \quad \text{(as an urelement in } A) & \text{if } u = \langle u_0, \dots, \langle p, 0\rangle\rangle \in \mathcal{U}(H^a), \\ \{m(H^a, u * \langle v\rangle) \mid u * \langle v\rangle \in H^a\} & \text{if } u \in \mathcal{S}(H^a). \end{cases}$$

Namely, $\{m(H^a, u) \mid u \in H^a\} = \mathrm{TC}(\{m(H^a, \epsilon)\})$ is the Mostowski collapse (in a modified sense taking urelements into account) of $(H^a, \sqsubset_{H^a})$, and $\mathcal{U}(H^a)$ corresponds to $\mathrm{TC}\big(\{m(H^a, \epsilon)\}\big) \cap A$, i.e., the *support* of $m(H^a, \epsilon)$; see [2, p. 29]. Thereby we let x $(= \langle a, 1\rangle)$ interpret the set $m(H^a, \epsilon) \in \mathbb{V}_A$. Now, the Barwise-Gandy-Moschovakis Theorem [3, 19], generalized to our setting, says that

$$\mathcal{M} := \big\{m(H^a, \epsilon) \mid a \in I \text{ and } H^a \text{ is a well-founded tree}\big\}$$

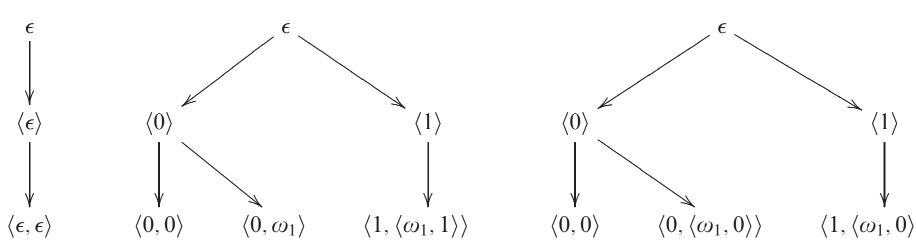
is equal to $\mathbb{HYP}_{\mathfrak{A}}$, i.e., the smallest admissible set above $\mathfrak{A}$, and we will formalize this argument within SC_1^-.

With this interpretation, a *bisimulation* between two trees T and S is defined to be a relation $R \subset T \times S$ satisfying $R(\epsilon, \epsilon)$ and the next four conditions:

(i) if $t\,R\,s$ and $t * \langle u\rangle \in \mathcal{U}(T)$, then $s * \langle u\rangle \in S$ and $R\big(t * \langle u\rangle,\ s * \langle u\rangle\big)$;
(ii) if $t\,R\,s$ and $t * \langle v\rangle \in \mathcal{S}(T)$, then $s * \langle w\rangle \in S$ and $R\big(t*\langle v\rangle,\ s*\langle w\rangle\big)$ for some w;
(iii) if $t\,R\,s$ and $s * \langle u\rangle \in \mathcal{U}(S)$, then $t * \langle u\rangle \in T$ and $R\big(t * \langle u\rangle,\ s * \langle u\rangle\big)$;
(iv) if $t\,R\,s$ and $s * \langle w\rangle \in \mathcal{S}(S)$, then $t * \langle v\rangle \in T$ and $R\big(t*\langle v\rangle,\ s*\langle w\rangle\big)$ for some v;

We say two trees are bisimilar when there is a bisimulation between them. Accordingly, for $x = \langle a, 1\rangle \in \mathcal{S}^*$ and $u \in \mathcal{S}(H^a)$, the well-founded tree $(H^a)_u$ is bisimilar to the canonical tree representation of $m(H^a, u) \in \mathbb{V}_A$, and $x\downarrow_u$ interprets the set $m(H^{j(a,u)}, \epsilon) = m((H^a)_u, \epsilon) = m(H^a, u)$; if $u \in \mathcal{U}(H^a)$, then $x\downarrow_u$ interprets the urelement $p \in A$ such that $(u)_{lh(u)-1} = \langle p, 0\rangle$.

Example 7.4. The tree $\{\epsilon\}$ represents $\emptyset^{\mathcal{S}}$ in the sense that $\{\epsilon\}$ is bisimilar to the canonical tree representation of $\emptyset^{\mathcal{S}}$. The trees $\{\epsilon, \langle\langle 1, 1\rangle\rangle\}$ and $\{\epsilon, \langle\langle 2, 1\rangle\rangle\}$ both represent $\{\emptyset^{\mathcal{S}}\}$, but $\{\epsilon, \langle\langle 1, 0\rangle\rangle\}$ and $\{\epsilon, \langle\langle 2, 0\rangle\rangle\}$ represent $\{1^{\mathcal{U}}\}$ and $\{2^{\mathcal{U}}\}$, respectively. Next, let us call the following trees T_1, T_2, and T_3 from left to right:



T_1 and T_2 are bisimilar and represent the same set $\{\{\emptyset^{\mathcal{S}}\}\}$. We have $\mathcal{U}(T_1) = \mathcal{U}(T_2) = \emptyset$, but $\mathcal{U}(T_3) = \{\langle 0, \langle\omega_1, 0\rangle\rangle,\ \langle 1, \langle\omega_1, 0\rangle\rangle\}$. Hence, whereas T_2 and T_3 have

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

the same shape, they are not bisimilar and represent different sets; T_3 represents $\{\{\emptyset^{\mathcal{S}}, \omega_1^{\mathcal{U}}\}, \{\omega_1^{\mathcal{U}}\}\}$. Let $a_i \in M$ be such that $H^{a_i} = T_i$ and $x_i = \langle a_i, 1\rangle \in \mathcal{S}^*$ for $1 \leq i \leq 3$. Then $x_1 \downarrow_{\langle \epsilon,\epsilon\rangle}$, $x_2 \downarrow_{\langle 0,0\rangle}$, and $x_3 \downarrow_{\langle 0,0\rangle}$ interpret the same set $\emptyset^{\mathcal{S}}$, and $x_3 \downarrow_{\langle 0,\langle \omega_1,0\rangle\rangle}$ and $x_3 \downarrow_{\langle 1,\langle \omega_1,0\rangle\rangle}$ interpret the same urelement $\omega_1^{\mathcal{U}}$.

For defining the interpretations of $=$ and $\in_1$, we first formalize, within SC_1^-, the aforementioned notion of bisimulation of hyperelementary well-founded trees as an inductive relation. Let $\mathcal{B}(\langle a,b,u,v\rangle, \mathfrak{X})$ (with parameters H and $\check{H}$) be

$$\begin{aligned}
&\forall x\big[u * \langle x\rangle \in \check{H}^a \to \exists y\big(v * \langle y\rangle \in H^b \wedge \big\langle a, b,\, u * \langle x\rangle,\, v * \langle y\rangle\big\rangle \in \mathfrak{X}\big)\big]\\
&\wedge \big[u \in \mathcal{U}(\check{H}^a) \to \big(v \in \mathcal{U}(H^b) \wedge (u)_{lh(u)} = (v)_{lh(v)}\big)\big]\\
&\wedge \forall y\big[v * \langle y\rangle \in \check{H}^b \to \exists x\big(u * \langle x\rangle \in H^a \wedge \big\langle a, b,\, u * \langle x\rangle,\, v * \langle y\rangle\big\rangle \in \mathfrak{X}\big)\big]\\
&\wedge \big[v \in \mathcal{U}(\check{H}^b) \to \big(u \in \mathcal{U}(H^a) \wedge (u)_{lh(u)} = (v)_{lh(v)}\big)\big] \;\wedge\; u \in H^a \;\wedge\; v \in H^b;
\end{aligned}$$

in terms of Remark 7.3, the monotone operator on $\mathfrak{A}$ induced by $\mathcal{B}$ inductively lists up the bisimilar pairs $\langle (H^a)_u, (H^b)_v\rangle$ starting from the leaves towards the roots. Since $\mathcal{B}$ is positive elementary in H and $-\check{H}$, the inductive least fixed point $J_{\mathcal{B}}$ of $\mathcal{B}$ exists by Theorem 5.2, and we let $B(a,b,u,v)$ denote $\langle a,b,u,v\rangle \in J_{\mathcal{B}}$; note that $B(a,b,u,v)$ implies $u \in H^a$ and $v \in H^b$.

Lemma 7.5. *Let $a,b,c \in M$ and $u,v,w \in \mathbb{V}$. The following hold*

1. *$B(a,b,u,v)$ iff $B(b,a,v,u)$.*
2. *If $B(a,b,u,v)$ and $B(b,c,v,w)$, then $B(a,c,u,w)$.*
3. *If $H^a = H^b$ and $u \in H^a$, then $B(a,b,u,u)$.*

Each of them is shown by induction along $\sqsubset_a$.

Lemma 7.6. *Let $a \in M$ and $u \in \mathcal{S}(H^a)$. Then $B(a, j(a,u), u * v, v)$ holds for all $v \in (H^a)_u$; hence we have $B(a, j(a,u), u, \epsilon)$ in particular. This claim is shown by induction on v along $\sqsubset_{j(a,u)}$ $(= \sqsubset_{(H^a)_u})$.*

We next define the dual operator $\mathcal{C}(\langle a,b,u,v\rangle, \mathfrak{X})$ of $\mathcal{B}$ by the following:

$$\begin{aligned}
&\exists x\Big[u * \langle x\rangle \in H^a \wedge \forall y\big(v * \langle y\rangle \in \check{H}^b \to \big\langle a, b,\, u * \langle x\rangle,\, v * \langle y\rangle\big\rangle \in \mathfrak{X}\big)\Big]\\
&\vee \Big[u \in \mathcal{U}(H^a) \wedge (v \notin \mathcal{U}(\check{H}^b) \vee (u)_{lh(u)} \neq (v)_{lh(v)})\Big]\\
&\vee \exists y\Big[v * \langle y\rangle \in H^b \wedge \forall x\big(u * \langle x\rangle \in \check{H}^a \to \big\langle a, b,\, u * \langle x\rangle,\, v * \langle y\rangle\big\rangle \in \mathfrak{X}\big)\Big]\\
&\vee \Big[v \in \mathcal{U}(H^b) \wedge (u \notin \mathcal{U}(\check{H}^a) \vee (u)_{lh(u)} \neq (v)_{lh(v)})\Big] \;\vee\; u \notin \check{H}^a \;\vee\; v \notin \check{H}^b;
\end{aligned}$$

the monotone operator induced by $\mathcal{C}$ inductively lists up the nonbisimilar pairs $\langle (H^a)_u, (H^b)_v)\rangle$ starting from the leaves toward the roots. Let C denote the least fixed-point $J_{\mathcal{C}}$ of $\mathcal{C}$; note that $u \notin \check{H}^a$ or $v \notin \check{H}^b$ implies $C(a,b,u,v)$.

Lemma 7.7. *Let $a,b \in M$. Then, for all $u \in H^a$ and $v \in H^b$, it holds*

$$C(a,b,u,v) \;\Leftrightarrow\; \neg B(a,b,u,v).$$

Proof. The claim is shown by induction on u along $\sqsubset_a$. ⊣

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

For interpreting the identity $=$ and the other membership relation $\in_1$ as well as their negations $\neq$ and $\notin_1$, we use the following *inductive* relations.

$$P^+_=(x,y) \;:\Leftrightarrow\; \big[x,y \in \mathcal{U}^* \wedge (x)_0 = (y)_0\big] \vee \big[x,y \in \mathcal{S}^* \wedge B\big((x)_0,(y)_0,\epsilon,\epsilon\big)\big]$$
$$P^-_=(x,y) \;:\Leftrightarrow\; \big[x,y \in \mathcal{U}^* \wedge (x)_0 \neq (y)_0\big] \vee \big[(x)_1 \neq (y)_1\big] \vee \big[x,y \in \mathcal{S}^* \wedge C\big((x)_0,(y)_0,\epsilon,\epsilon\big)\big]$$
$$P^+_{\in_1}(x,y) \;:\Leftrightarrow\; y \in \mathcal{S}^* \wedge \exists z\big(\langle z\rangle \in H^{(y)_0} \wedge P^+_=(x,\, y{\downarrow_{\langle z\rangle}})\big)$$
$$P^-_{\in_1}(x,y) \;:\Leftrightarrow\; y \in \mathcal{U}^* \vee \forall z\big(\langle z\rangle \in \check{H}^{(y)_0} \rightarrow P^-_=(x,\, y{\downarrow_{\langle z\rangle}}\big).$$

Corollary 7.8. *For all $x, y \in \mathcal{U}^* \cup \mathcal{S}^*$, we have*

$$\neg P^+_=(x,y) \leftrightarrow P^-_=(x,y) \qquad \textit{and} \qquad \neg P^+_\in(x,y) \leftrightarrow P^-_\in(x,y).$$

Finally, we define the interpretations of $=$ and $\in_1$ as follows:

$$x =^* y \;:\Leftrightarrow\; P^+_=(x,y) \qquad \text{and} \qquad x \in_1^* y \;:\Leftrightarrow\; P^+_{\in_1}(x,y).$$

In particular, for every $x = \langle a,1\rangle \in \mathcal{S}^*$, $z \in_1^* x$ holds, if and only if either $z = \langle b,1\rangle \in \mathcal{S}^*$ for some b and the tree H^b is bisimilar to some immediate subtree of H^a, or $z = \langle c,0\rangle \in \mathcal{U}^*$ for some c and there is a leaf of H^a immediately below the root ϵ that represents $c^{\mathcal{U}}$ (i.e., $\langle\langle c,0\rangle\rangle \in H^a \cap \min(\sqsubset_a)$); also, $z =^* x$ if and only if $z = \langle d,1\rangle \in \mathcal{S}^*$ for some d such that H^d is bisimilar to H^a.

Lemma 7.9. $\mathsf{SC}_1^- \vdash (\mathrm{Eq})^*$; *by definition and Lemmata* 7.5 *and* 7.6.

Lemma 7.10. *Let $X \subset \mathcal{S}^* \cup \mathcal{U}^*$ be hyperelementary. There exists $y \in \mathcal{S}^*$ such that $(\forall z \in \mathcal{U}^* \cup \mathcal{S}^*)\big[P^+_{\in_1}(z,y) \leftrightarrow \exists x\big(x \in X \wedge P^+_=(z,x)\big)\big]$.*

Proof. Let T be a hyperelementary tree defined by

$$T := \{\epsilon\} \cup \{\langle x\rangle \mid x \in X \cap \mathcal{U}^*\} \cup \{\langle x\rangle * u \mid x \in X \setminus \mathcal{U}^* \wedge u \in H^{(x)_0}\};$$

we have $\{\langle x\rangle \mid x \in X \cap \mathcal{U}^*\} \subset \mathcal{U}(T)$ and $\{\langle x\rangle \mid x \in X \setminus \mathcal{U}^*\} = \{\langle x\rangle \mid x \in X \cap \mathcal{S}^*\} \subset \mathcal{S}(T)$. We first show that T is well-founded. Since $Acc(\sqsubset_T)$ is downward closed, it suffices to show $\langle x\rangle \in Acc(\sqsubset_T)$ for all $x \in X$. This obviously holds for $x \in X \cap \mathcal{U}^*$. Let $x = \langle a,1\rangle \in X \cap \mathcal{S}^*$. We have $\langle x\rangle \in \mathcal{S}(T)$ and $T_{\langle x\rangle} = H^a$, which is well-founded since $\langle a,1\rangle \in \mathcal{S}^*$. Therefore, by Proposition 7.1, the well-foundedness of $T_{\langle x\rangle}$ implies $\langle x\rangle \in Acc(\sqsubset_T)$.

Now, pick $b \in M$ with $H^b = T$. Let $y = \langle b,1\rangle \in \mathcal{S}^*$ and take any $z \in \mathcal{U}^* \cup \mathcal{S}^*$.

Suppose $P^+_{\in_1}(z,y)$. There is some $\langle w\rangle \in H^b$ with $P^+_=(z, y{\downarrow_{\langle w\rangle}})$. If $z \in \mathcal{U}^*$, then $z = y{\downarrow_{\langle w\rangle}}$ and thus $\langle w\rangle \in \mathcal{U}(H^b)$, which entails $z = y{\downarrow_{\langle w\rangle}} = w \in X \cap \mathcal{U}^*$. Assume $z = \langle a,1\rangle \in \mathcal{S}^*$ for some a. Then $y{\downarrow_{\langle w\rangle}} \in \mathcal{S}^*$ and thus $\langle w\rangle \in \mathcal{S}(H^b)$; hence $w \in X \cap \mathcal{S}^*$. Let $w = \langle c,1\rangle$. Since $H^c = T_{\langle w\rangle} = H^{j(b,\langle w\rangle)}$, we have $B(j(b,\langle w\rangle),c,\epsilon,\epsilon)$ by Lemma 7.5.3 and thus $P^+_=(y{\downarrow_{\langle w\rangle}}, w)$; hence $P^+_=(z,w)$.

Let $x \in X$ be such that $P^+_=(z,x)$; then $\langle x\rangle \in H^b$. If $x \in \mathcal{U}^*$ then $y{\downarrow_{\langle x\rangle}} = x$ and thus $P^+_=(z, y{\downarrow_{\langle x\rangle}})$. If $x = \langle c,1\rangle \in \mathcal{S}^*$, then $T_{\langle x\rangle} = H^c$ and thus $B(c, j(b,\langle x\rangle),\epsilon,\epsilon)$ by Lemma 7.5.3. Hence we have $P^+_=(x, y{\downarrow_{\langle x\rangle}})$ and thus $P^+_=(z, y{\downarrow_{\langle x\rangle}})$. ⊣

Lemma 7.11. $\mathsf{SC}_1^- \vdash (\mathrm{Pair})^*$; *apply Lemma* 7.10 *to $X = \{v,w\}$ for $v, w \in \mathcal{S}^* \cup \mathcal{U}^*$.*

Proposition 7.12. *Let $x = \langle a,1\rangle \in \mathcal{S}^*$, $u \in \mathcal{S}(H^a)$, and $v \in (H^a)_u$. Then we have $P^+_=\big((x{\downarrow_u}){\downarrow_v},\, x{\downarrow_{u*v}}\big)$; by definition and Lemma* 7.5.3.

LEMMA 7.13. $\mathsf{SC}_1^- \vdash (\text{Union})^*$.

PROOF. Take any $x = \langle a, 1\rangle \in \mathcal{S}^*$. Put $X := \{x\downarrow_u \mid u \in H^a \wedge lh(u) = 2\}$. We take $y = \langle b, 1\rangle \in \mathcal{S}^*$ such that $(\forall z \in \mathcal{U}^* \cup \mathcal{S}^*)(P^+_{\in_1}(z, y) \leftrightarrow (\exists x \in X)P^+_{=}(z, x))$ by Lemma 7.10. Take any $v \in \mathcal{S}^*$ and $z \in \mathcal{U}^* \cup \mathcal{S}^*$ such that $P^+_{\in_1}(z, v)$ and $P^+_{\in_1}(v, x)$. We have $P^+_{=}(v, x\downarrow_{\langle w\rangle})$ for some $\langle w\rangle \in \mathcal{S}(H^a)$. Hence, we have $P^+_{\in_1}(z, x\downarrow_{\langle w\rangle})$, and thus there exists $\langle w'\rangle \in H^{(x\downarrow_{\langle w\rangle})_0} = (H^a)_{\langle w\rangle}$ such that $P^+_{=}\big(z, (x\downarrow_{\langle w\rangle})\downarrow_{\langle w'\rangle}\big)$. Then, we obtain $P^+_{=}(z, x\downarrow_{\langle w,w'\rangle})$ by Proposition 7.12. Since $\langle w, w'\rangle \in H^a$, we finally get $x\downarrow_{\langle w,w'\rangle} \in X$ and thus $P^+_{\in_1}(z, y)$. ⊣

LEMMA 7.14. *For each Δ_0-formula $\varphi(\vec{x})$ of $\mathcal{L}_{\mathrm{KP}}$, there are inductive relations $P^+_\varphi(\vec{x})$ and $P^-_\varphi(\vec{x})$ such that, for all $\vec{x} \in \mathcal{U}^* \cup \mathcal{S}^*$,*

$$P^+_\varphi(\vec{x}) \Leftrightarrow \varphi^*(\vec{x}) \qquad \text{and} \qquad P^-_\varphi(\vec{x}) \Leftrightarrow \neg\varphi^*(\vec{x}).$$

PROOF. By induction on φ. For example, if $\varphi = (\forall z \in_1 x)\psi(z, x, \vec{v})$, we take

$$P^+_\varphi(x, \vec{v}) \;:\Leftrightarrow\; x \in \mathcal{U}^* \vee \forall w(\langle w\rangle \in \check{H}^{(x)_0} \to P^+_\psi(x\downarrow_{\langle w\rangle}, x, \vec{v}))$$
$$P^-_\varphi(x, \vec{v}) \;:\Leftrightarrow\; x \notin \mathcal{U}^* \wedge \exists w(\langle w\rangle \in H^{(x)_0} \wedge P^-_\psi(x\downarrow_{\langle w\rangle}, x, \vec{v})). \qquad ⊣$$

LEMMA 7.15. $\mathsf{SC}_1^- \vdash (\Delta_0\text{-Sep}_1)^*$.

PROOF. Let $\varphi(z, x, \vec{v}) \in \Delta_0$. Take $x = \langle a, 1\rangle \in \mathcal{S}^*$ and $\vec{v} \in \mathcal{U}^* \cup \mathcal{S}^*$. Put

$$X := \{x\downarrow_{\langle w\rangle} \mid \langle w\rangle \in H^a \wedge P^+_\varphi(x\downarrow_{\langle w\rangle}, x, \vec{v})\},$$

which is hyperelementary. By Lemma 7.10 there is y such that, for all $z \in \mathcal{U}^* \cup \mathcal{S}^*$,

$$P^+_{\in_1}(z, y) \Leftrightarrow \exists w\big(\langle w\rangle \in H^a \wedge P^+_\varphi(x\downarrow_{\langle w\rangle}, x, \vec{v}) \wedge P^+_{=}(z, x\downarrow_{\langle w\rangle})\big)$$
$$\Leftrightarrow P^+_{\in_1}(z, x) \wedge P^+_\varphi(z, x, \vec{v}). \qquad ⊣$$

LEMMA 7.16. $\mathsf{SC}_1^- \vdash (\Delta_0\text{-Coll}_1)^*$.

PROOF. Let $x = \langle a, 1\rangle \in \mathcal{S}^*$ and suppose $(\forall y \in_1^* x)(\exists z \in \mathcal{S}^* \cup \mathcal{U}^*)\varphi^*(y, z)$ for some $\varphi \in \Delta_0$. By Theorem 5.8, where we take $P(y, z) :\Leftrightarrow z \in \mathcal{S}^* \cup \mathcal{U}^* \wedge P^+_\varphi(y, z)$, there are inductive relation Q and coinductive relation $\check{Q}$ such that

- If $Q(y, z)$ then $z \in \mathcal{S}^* \cup \mathcal{U}^*$ and $P^+_\varphi(y, z)$.
- If $P^+_\varphi(y, z)$ for some $z \in \mathcal{U}^* \cup \mathcal{S}^*$, then $\{z \mid Q(y, z)\} = \{z \mid \check{Q}(y, z)\} \neq \emptyset$.

Let $X = \{z \mid \exists u(\langle u\rangle \in H^a \wedge Q(x\downarrow_{\langle u\rangle}, z)\}$, which is hyperelementary. By Lemma 7.10, we pick $w \in \mathcal{S}^*$ such that $(\forall v \in \mathcal{U}^* \cup \mathcal{S}^*)\big(P^+_{\in_1}(v, w) \leftrightarrow (\exists u \in X)P^+_{=}(v, u)\big)$, and let $w = \langle b, 1\rangle$. Take any y with $P^+_{\in_1}(y, x)$. We have $P^+_{=}(y, x\downarrow_{\langle u\rangle})$ for some $\langle u\rangle \in H^a$ and $\varphi^*(y, z)$ for some $z \in \mathcal{S}^* \cup \mathcal{U}^*$. Hence we get $P^+_\varphi(x\downarrow_{\langle u\rangle}, z)$ and thus $Q(x\downarrow_{\langle u\rangle}, z')$ for some z'. We have $z' \in X$, $z' \in \mathcal{U}^* \cup \mathcal{S}^*$, and $P^+_\varphi(x\downarrow_{\langle u\rangle}, z')$, which finally entails $P^+_{\in_1}(z', w)$ and $\varphi^*(y, z')$. ⊣

LEMMA 7.17. $\mathsf{SC}_1^- \vdash (\text{Found}_1)^*$.

PROOF. Suppose $\forall y(y \in_1^* x \to \varphi^*(y)) \to \varphi^*(x)$ for all $x \in \mathcal{U}^* \cup \mathcal{S}^*$. Take any $z \in \mathcal{U}^* \cup \mathcal{S}^*$. We will show $\varphi^*(z)$. If $z \in \mathcal{U}^*$ we trivially get $\varphi^*(z)$ by the supposition. Let $z = \langle a, 1\rangle \in \mathcal{S}^*$. Since $z\downarrow_\epsilon =^* z$, it suffices to show that

$$\forall v(v \sqsubset_a u \to \varphi^*(z\downarrow_v)) \to \varphi^*(z\downarrow_u), \text{ for all } u \in H^a. \tag{1}$$

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

Let $u \in H^a$ and assume $\forall v(v \sqsubset_a u \to \varphi^*(z \downarrow_v))$. Take any w with $P^+_{\in_1}(w, z \downarrow_u)$. We have $P^+_{=}(w, z \downarrow_v)$ for some $v \sqsubset_a u$ by Proposition 7.12 and thus $\varphi^*(w)$ by the assumption. Since w is arbitrary we obtain $\varphi^*(z \downarrow_u)$ by the supposition. ⊣

LEMMA 7.18. $\mathsf{SC}_1^- \vdash (\mathrm{Ext})^*$.

PROOF. Let $x = \langle a, 1\rangle \in \mathcal{S}^*$ and $y = \langle b, 1\rangle \in \mathcal{S}^*$. When $P^+_{\in_1}(z, x) \leftrightarrow P^+_{\in_1}(z, y)$ for all $z \in \mathcal{U}^* \cup \mathcal{S}^*$, we can show $\mathcal{B}(a, b, \epsilon, \epsilon, B)$ and thus $P^+_{=}(x, y)$. ⊣

Finally, let $T = \{\epsilon\} \cup \{\langle\langle v, 0\rangle\rangle \mid v \in \mathbb{V}\}$. We can give an explicit definition of some object $b \in M$ such that $H^b = T$. We fix one such b and its definition, and put $\mathsf{V}^* := \langle b, 1\rangle$. With this interpretation, we can easily verify $\mathsf{SC}_1^- \vdash (\mathrm{U})^*$.

THEOREM 7.19. *The translation $*$ is an interpretation of* $\mathsf{KP}\mathbb{V}^-$ *in* SC_1^-. *It is also an interpretation of* $\mathsf{KP}\mathbb{V}^- + (\mathrm{Sep}_0^+)$ *in* $\mathsf{SC}_1^- + \mathcal{L}_{\mathrm{SC}}$-Sep, *and* $\mathsf{KP}\mathbb{V}$ *in* SC_1.

PROOF. For the first claim, it remains to be shown that $\mathsf{SC}_1^- \vdash (\mathsf{ZF}^{\mathcal{U}})^*$. In general, for each $\varphi(x_1, \dots, x_k) \in \mathcal{L}_\in$ only with the displayed variables free, we can show the following by induction on φ, which immediately entails the claim:

$$\mathsf{SC}_1^- \vdash \forall x_1 \cdots \forall x_k \left[\varphi(\vec{x}) \leftrightarrow (\varphi^{\mathcal{U}})^*(\langle x_0, 0\rangle, \dots, \langle x_k, 0\rangle) \right]. \tag{2}$$

For the second claim, let $\varphi(z, \vec{v}) \in \mathcal{L}_{\mathrm{KP}}$, $x = \langle a, 0\rangle \in \mathcal{U}^*$, and $\vec{v} \in \mathcal{U}^* \cup \mathcal{S}^*$. In the presence of $\mathcal{L}_{\mathrm{SC}}$-Sep, we can take $b = \{c \in a \mid \varphi^*(\langle c, 0\rangle, \vec{v})\}$. Then we put $y = \langle b, 0\rangle \in \mathcal{U}^*$ and have

$$(\forall z \in \mathcal{U}^*)[z \in_0^* y \leftrightarrow z \in_0 x \wedge \varphi^*(z, \vec{v})].$$

The case for (Repl_0^+) and $\mathcal{L}_{\mathrm{SC}}$-Repl is similarly treated. ⊣

THEOREM 7.20. *For all $\varphi \in \mathcal{L}_\in$, if* $\mathsf{KP}\mathbb{V}^- \vdash \varphi^{\mathcal{U}}$ *then* $\mathsf{SC}_1^- \vdash \varphi$. *The same holds for* $\mathsf{KP}\mathbb{V}^- + (\mathrm{Sep}_0^+)$ *and* $\mathsf{SC}_1^- + \mathcal{L}_{\mathrm{Fix}}$-Sep, *and for* $\mathsf{KP}\mathbb{V}$ *and* SC_1^-. *This is an immediate consequence of the last theorem and* (2).

§8. Reduction of VF **to** KP$\mathbb{V}$**.** Cantini [7] gave an embedding of VF⟦PA⟧ in KPu, which is essentially equal to KPU$^+$ [2] over natural numbers augmented with the arithmetical induction schema extended to the whole language (see [7] or [16] for its definition); note that KPω is a urelement-free formulation of KPu.

Essentially the same embedding works for VF (over ZF) and KP$\mathbb{V}$; in fact, it gives an embedding of $\mathsf{VF}^- + \mathcal{L}_T$-Ind in $\mathsf{KP}\mathbb{V}^- + (\mathrm{Found}_0^+)$. Such an embedding is given by what Cantini calls *provability interpretation*, by which we interpret the truth of an $\mathcal{L}_\in^\infty$-sentence $\sigma \in \mathrm{St}^\infty_{\mathcal{L}_T}$ by the provability of $\sigma^{\mathcal{U}} \in (\mathrm{St}^\infty_{\mathcal{L}_T})^{\mathcal{U}}$ within a certain semiformal infinitary system formalizable within $\mathsf{KP}\mathbb{V}^-$.

In what follows we work within $\mathsf{KP}\mathbb{V}^- + (\mathrm{Found}_0^+)$. As in Section 6, we will occasionally treat $\mathcal{L}_{\mathrm{KP}}$-formulae as classes; e.g., we write $x \in \mathcal{U}$ and $x \in (\mathrm{St}^\infty_{\mathcal{L}_T})^{\mathcal{U}}$. We also assume that formulae of $\mathcal{L}_T$ are all expressed in their negation normal forms; so, they can be seen as constructed from literals by means of $\wedge$, $\vee$, $\forall$, and $\exists$; then, for a formula A in its negation normal form, $\neg A$ is standardly defined.

DEFINITION 8.1. For $\alpha, \rho \in On^{\mathcal{S}}$, and for a finite (in the sense of $\mathcal{U}$) set $\Gamma \subset^{\mathcal{U}} (\mathrm{St}^\infty_{\mathcal{L}_T})^{\mathcal{U}}$, the relation $\mathfrak{S} \vert\!\frac{\alpha}{\rho}\, \Gamma$ holds iff one of the following holds

(a) for some $a, b \in \mathcal{U}$, either $\ulcorner \dot{a} \in \dot{b} \urcorner^{\mathcal{U}} \in_0 \Gamma$ and $a \in_0 b$, or $\ulcorner \dot{a} \notin \dot{b} \urcorner^{\mathcal{U}} \in_0 \Gamma$ and $a \notin_0 b$;

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

(b) for some $a, b \in \mathcal{U}$, either $\ulcorner \dot{a} = \dot{b} \urcorner^{\mathcal{U}} \in_0 \Gamma$ and $a = b$, or $\ulcorner \dot{a} \neq \dot{b} \urcorner^{\mathcal{U}} \in_0 \Gamma$ and $a \neq b$;

(c) for some $a \in \mathcal{U}$, it holds that $\ulcorner T\dot{a} \urcorner^{\mathcal{U}}, \ulcorner \neg T\dot{a} \urcorner^{\mathcal{U}} \in_0 \Gamma$;

(d) $\ulcorner \neg T \ulcorner A \urcorner \urcorner^{\mathcal{U}}, \ulcorner \neg T \ulcorner \neg A \urcorner \urcorner^{\mathcal{U}} \in_0 \Gamma$, for some $\ulcorner A \urcorner^{\mathcal{U}} \in (\mathrm{St}^{\infty}_{\mathcal{L}_T})^{\mathcal{U}}$;

(e) there exist some $\ulcorner A \urcorner^{\mathcal{U}}, \ulcorner B \urcorner^{\mathcal{U}} \in (\mathrm{St}^{\infty}_{\mathcal{L}_T})^{\mathcal{U}}$, and $\alpha_0, \alpha_1 <^{\mathcal{S}} \alpha$ such that $\ulcorner A \wedge B \urcorner^{\mathcal{U}} \in_0 \Gamma$, $\mathfrak{S} \vdash^{\alpha_0}_{\rho} \Gamma, \ulcorner A \urcorner^{\mathcal{U}}$, and $\mathfrak{S} \vdash^{\alpha_1}_{\rho} \Gamma, \ulcorner B \urcorner^{\mathcal{U}}$;

(f) there exist some $\ulcorner A \urcorner^{\mathcal{U}}, \ulcorner B \urcorner^{\mathcal{U}} \in (\mathrm{St}^{\infty}_{\mathcal{L}_T})^{\mathcal{U}}$ and $\alpha' <^{\mathcal{S}} \alpha$ such that $\ulcorner A \vee B \urcorner^{\mathcal{U}} \in_0 \Gamma$, and either $\mathfrak{S} \vdash^{\alpha'}_{\rho} \Gamma, \ulcorner A \urcorner^{\mathcal{U}}$ or $\mathfrak{S} \vdash^{\alpha'}_{\rho} \Gamma, \ulcorner B \urcorner^{\mathcal{U}}$;

(g) there exists some $\ulcorner A(x) \urcorner^{\mathcal{U}} \in (\mathrm{Fml}^{\infty}_{\mathcal{L}_T})^{\mathcal{U}}$ such that $\ulcorner \forall x A(x) \urcorner \in_0 \Gamma$, and for each $a \in \mathcal{U}$ there is $\alpha_a <^{\mathcal{S}} \alpha$ such that $\mathfrak{S} \vdash^{\alpha_a}_{\rho} \Gamma, \ulcorner A(\dot{a}) \urcorner^{\mathcal{U}}$;

(h) there exist some $\ulcorner A(x) \urcorner^{\mathcal{U}} \in (\mathrm{Fml}^{\infty}_{\mathcal{L}_T})^{\mathcal{U}}$ and $\alpha' <^{\mathcal{S}} \alpha$ such that $\ulcorner \exists x A(x) \urcorner^{\mathcal{U}} \in_0 \Gamma$, and $\mathfrak{S} \vdash^{\alpha'}_{\rho} \Gamma, \ulcorner A(\dot{a}) \urcorner^{\mathcal{U}}$ for some $a \in \mathcal{U}$;

(i) there exist some $\ulcorner A \urcorner^{\mathcal{U}} \in (\mathrm{St}^{\infty}_{\mathcal{L}_T})^{\mathcal{U}}$, $\alpha' <^{\mathcal{S}} \alpha$, and $\rho' <^{\mathcal{S}} \rho$ such that $\ulcorner T(\ulcorner A \urcorner) \urcorner^{\mathcal{U}} \in_0 \Gamma$, and $\mathfrak{S} \vdash^{\alpha'}_{\rho'} \ulcorner A \urcorner^{\mathcal{U}}$;

(j) there exist some $\ulcorner A \urcorner^{\mathcal{U}} \in (\mathrm{St}^{\infty}_{\mathcal{L}_T})^{\mathcal{U}}$, $\alpha' <^{\mathcal{S}} \alpha$, and $\rho' <^{\mathcal{S}} \rho$ such that $\ulcorner \neg T(\ulcorner A \urcorner) \urcorner^{\mathcal{U}} \in_0 \Gamma$, and $\mathfrak{S} \vdash^{\alpha'}_{\rho'} \ulcorner \neg A \urcorner^{\mathcal{U}}$;

here, following the convention, we mean $\mathfrak{S} \vdash^{\alpha}_{\rho} \Sigma \cup \{\ulcorner D \urcorner\}$ by "$\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner D \urcorner$".

Due to the axiom (U), each of the ten clauses (a)–(j) above (and the finiteness in $\mathcal{U}$) is Δ_0 with parameters α, ρ, and Γ, and the relation $\mathfrak{S} \vdash^{\alpha'}_{\rho'} \Gamma'$ only occurs positively therein. Hence, the relation $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma$ can be defined as a least fixed-point of a positive Σ-operator and thus a Σ-predicate in $\mathsf{KP}\mathbb{V}^{-}$; see [2, Chapter VI].

Lemma 8.2. *The following basic proof-theoretic properties of the semiformal system $\mathfrak{S}$ are all standardly shown by induction on α (using* (Found$_1$)*).*

- $\mathfrak{S} \vdash^{\alpha}_{\rho} \emptyset^{\mathcal{U}}$ *for no α and ρ (Consistency of $\mathfrak{S}$),*
- *For $\Gamma \subset^{\mathcal{U}} \Delta$, $\alpha \leq^{\mathcal{S}} \beta$, and $\rho \leq^{\mathcal{S}} \tau$, if $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma$ then $\mathfrak{S} \vdash^{\beta}_{\tau} \Delta$ (Structural Lemma),*
- *If $a \notin_0 b$ and $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner \dot{a} \in \dot{b} \urcorner^{\mathcal{U}}$, then $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma$ (Falsity Lemma* 1*),*
- *If $a \in_0 b$ and $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner \dot{a} \notin \dot{b} \urcorner^{\mathcal{U}}$, then $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma$ (Falsity Lemma* 2*),*
- *If $a \neq b$ and $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner \dot{a} = \dot{b} \urcorner^{\mathcal{U}}$, then $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma$ (Falsity Lemma* 3*),*
- *If $a = b$ and $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner \dot{a} \neq \dot{b} \urcorner^{\mathcal{U}}$, then $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma$ (Falsity Lemma* 4*),*
- *If $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner A \wedge B \urcorner^{\mathcal{U}}$ then $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner A \urcorner^{\mathcal{U}}$ and $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner B \urcorner^{\mathcal{U}}$ ($\wedge$-Inversion),*
- *If $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner \forall x A(x) \urcorner^{\mathcal{U}}$ then $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner A(\dot{a}) \urcorner^{\mathcal{U}}$ for all $a \in \mathcal{U}$ ($\forall$-Inversion),*
- *If $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner A \vee B \urcorner^{\mathcal{U}}$ then $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner A \urcorner^{\mathcal{U}}, \ulcorner B \urcorner^{\mathcal{U}}$ ($\vee$-Exportation).*

Lemma 8.3. *If $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner \neg T \ulcorner \neg A \urcorner \urcorner^{\mathcal{U}}, \ulcorner T \ulcorner A \urcorner \urcorner^{\mathcal{U}}$, then $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner \neg T \ulcorner \neg A \urcorner \urcorner^{\mathcal{U}}$.*

Proof. By straightforward induction on α. ⊣

Lemma 8.4 (Cut admissibility). *Suppose $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma, \ulcorner A \urcorner$ and $\mathfrak{S} \vdash^{\beta}_{\rho} \Gamma, \ulcorner \neg A \urcorner$ for some $\ulcorner A \urcorner \in (\mathrm{St}^{\infty}_{\mathcal{L}_T})^{\mathcal{U}}$. Then, we have $\mathfrak{S} \vdash^{\gamma}_{\rho} \Gamma$ for some $\gamma \in On^{\mathcal{S}}$.*

Proof. For each $\ulcorner A \urcorner^{\mathcal{U}} \in (\mathrm{St}^{\infty}_{\mathcal{L}_T})^{\mathcal{U}}$, let $sc^{\mathcal{U}}(\ulcorner A \urcorner^{\mathcal{U}}) \in \omega^{\mathcal{U}}$ be the surface complexity of $\ulcorner A \urcorner$: all atomics are assigned the surface complexity 0 and thus $sc^{\mathcal{U}}(\ulcorner T\dot{a} \urcorner) = 0^{\mathcal{U}}$

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

for all $a \in \mathcal{U}$. Using (Found_1) and (Found_0^+), the claim is shown by quadruple induction on $\rho \in On^{\mathcal{S}}$, $sc(\ulcorner A \urcorner) \in \omega^{\mathcal{U}}$, $\alpha \in On^{\mathcal{S}}$, and $\beta \in On^{\mathcal{S}}$. The details are parallel to the proof of Theorem 4.4 of [7] (or Theorem 62.1 of [8]); we note that Lemma 8.3 is used in the case where the rule (d) is used. ⊣

We will write $\mathfrak{S} \vdash \Gamma$ when $\mathfrak{S} \vdash^{\alpha}_{\rho} \Gamma$ for some α and ρ.

Lemma 8.5 (T-elimination). 1. *If* $\mathfrak{S} \vdash \ulcorner T \ulcorner \varphi \urcorner \urcorner^{\mathcal{U}}$ *then* $\mathfrak{S} \vdash \ulcorner \varphi \urcorner^{\mathcal{U}}$.
2. *If* $\mathfrak{S} \vdash \ulcorner \neg T \ulcorner \varphi \urcorner \urcorner^{\mathcal{U}}$ *then* $\mathfrak{S} \vdash \ulcorner \neg \varphi \urcorner^{\mathcal{U}}$.

Proof. If $\mathfrak{S} \vdash^{\alpha}_{\rho} \ulcorner T \ulcorner \varphi \urcorner \urcorner^{\mathcal{U}}$, then (i) must be the case for $\Gamma := \{\ulcorner T(\ulcorner \varphi \urcorner) \urcorner^{\mathcal{U}}\}$, and $\mathfrak{S} \vdash^{\alpha'}_{\rho'} \ulcorner \varphi \urcorner^{\mathcal{U}}$ for some $\alpha' <^{\mathcal{S}} \alpha$ and $\rho' <^{\mathcal{S}} \rho$. The Claim 2 is shown similarly. ⊣

Definition 8.6 (Provability interpretation). We define the provability interpretation $A^{\infty} \in \mathcal{L}_{\mathrm{KP}}$ for each $A \in \mathcal{L}_T$ by

$$Tx \mapsto \mathfrak{S} \vdash x, \qquad x \in y \mapsto x \in_0 y, \qquad \text{and} \qquad \forall x \mapsto (\forall x \in \mathcal{U});$$

and all the other vocabulary is unchanged.

Lemma 8.7 (Reflection Lemma). *Let* $A_0(\vec{x}), \dots, A_n(\vec{x}) \in \mathcal{L}_T$ *at most* $\vec{x}$ *free. Then* $\mathsf{KP}\mathbb{V}^- \vdash (\forall \vec{a} \in \mathcal{U})\big(\mathfrak{S} \vdash \{\ulcorner A_0(\vec{a}) \urcorner^{\mathcal{U}}, \dots, \ulcorner A_n(\vec{a}) \urcorner^{\mathcal{U}}\}^{\mathcal{U}} \to \big(A_0^{\infty}(\vec{a}) \vee \cdots \vee A_n^{\infty}(\vec{a})\big)\big)$.

Proof. For each $k \in \mathbb{N}$, we can define within $\mathsf{KP}\mathbb{V}^-$ a partial truth predicate $\mathrm{Tr}_k(x)$ of the $\mathcal{L}_T^{\infty}$-structure $\langle \mathcal{U}, \in_0, T^{\infty}, \{c_u \mid u \in \mathcal{U}\}\rangle$ for all $\ulcorner \psi \urcorner^{\mathcal{U}} \in (\mathrm{St}^{\infty}_{\mathcal{L}_T})^{\mathcal{U}}$ with $sc^{\mathcal{U}}(\ulcorner A \urcorner^{\mathcal{U}}) \leq^{\mathcal{U}} k^{\mathcal{U}}$; c.f., [7, Lemma 5.8.1]. Then we can show

$$(\forall \Gamma \subset (\mathrm{St}^{\infty}_{\mathcal{L}_T})^{\mathcal{U}})\big[\big(\mathfrak{S} \vdash \Gamma \wedge (\forall x \in_0 \Gamma)(sc^{\mathcal{U}}(x) \leq^{\mathcal{U}} k^{\mathcal{U}})\big) \to (\exists x \in_0 \Gamma)\mathrm{Tr}_k(x)\big].$$

by straightforward induction on $\alpha \in On^{\mathcal{S}}$, using ($\text{Found}_1$). ⊣

Lemma 8.8. *If* $\ulcorner \varphi(\vec{v}) \urcorner^{\mathcal{U}} \in \mathrm{LogAx}^{\mathcal{U}}_{\mathcal{L}_T^{\infty}}$, *then* $\mathfrak{S} \vdash \ulcorner \varphi(\vec{a}) \urcorner^{\mathcal{U}}$ *for all* $\vec{a} \in \mathcal{U}$.

Theorem 8.9. *The translation* $A \mapsto A^{\infty}$ *is a relative interpretation.*

Proof. $\mathbf{V1}^{\infty}$ follows from Reflection Lemma. $\mathbf{V2}^{\infty}$ and $\mathbf{V3}^{\infty}$ follow from Falsity Lemma and Consistency of $\mathfrak{S}$. $\mathbf{V4}^{\infty}$ is a consequence of Lemma 8.8. $\mathbf{V5}^{\infty}$ follows from the clause (g) in Definition 8.1, the axiom (U), and Σ-Collection for $\mathcal{S}$ ([2, Chapter I]), which is derivable in $\mathsf{KP}\mathbb{V}^-$. $\mathbf{V6}^{\infty}$ follows by $\vee$-Exportation and Lemma 8.4. $\mathbf{V7}^{\infty}$ is immediate from the clause (i). $\mathbf{V8}^{\infty}$ follows from Lemma 8.5.2. Finally, $\mathbf{V9}^{\infty}$ is immediate from the clause (d). ⊣

Theorem 8.10. *There is an interpretation of* $\mathsf{VF}^- + \mathcal{L}_T$-Ind *in* $\mathsf{SC}_1^- + \mathcal{L}_{\mathrm{SC}}$-Ind *that preserve the* $\mathcal{L}_{\in}$-*part.*[2] *Hence,* $\mathsf{VF}^- + \mathcal{L}_T\text{-Ind} \subset_{\mathcal{L}_{\in}} \mathsf{SC}_1^- + \mathcal{L}_{\mathrm{SC}}\text{-Ind}$.

Proof. By Theorems 7.19 and 8.9, the translation $\varphi \mapsto (\varphi^{\infty})^*$ is an interpretation of $\mathsf{VF}^- + \mathcal{L}_T$-Ind in $\mathsf{SC}_1^- + \mathcal{L}_{\mathrm{SC}}$-Ind, and it maps $\forall x$ to $(\forall x \in \mathcal{U}^*)$, $x \in y$ to $x \in_0^* y$, and Tx to $(\mathfrak{S} \vdash x)^*$. Let $\mathcal{I}$ be a new translation of $\mathcal{L}_T$ to $\mathcal{L}_{\mathrm{SC}}$ that maps Tx to

[2] As a matter of fact, we can embed VF^- in $\mathsf{KP}\mathbb{V}^-$ and thus in SC_1^- by modifying Cantini's embedding of VF^- in $\mathsf{PW}^- + \mathsf{GID}$ in [8]. For this purpose, we need to re-define $(\mathrm{St}^{\infty}_{\mathcal{L}_T})^{\mathcal{U}}$ in terms of inductive definitions within $\mathsf{KP}\mathbb{V}^-$, which makes the use of (Found_0^+) in Theorem 8.4 dispensable; the new definition does not provably equal to the original definition in $\mathsf{KP}\mathbb{V}^-$, though they coincide in $\mathsf{KP}\mathbb{V}^- +$ (Found_0^+). We then introduce an intermediate system that is the same as VF^- except that the quantifiers "$(\forall x \in \mathrm{St}^{\infty}_{\mathcal{L}_T})$" and "$(\forall x \in \mathrm{Fml}^{\infty}_{\mathcal{L}_T})$" in the VF-axioms are replaced by "$\forall x$", which includes VF^-, and embed it in $\mathsf{KP}\mathbb{V}^-$.

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

$(T^\infty)^*(\langle x, 0\rangle)$ and preserves all the rest. In a similar way to (2), we can show that, for all $\varphi(x_1, \ldots, x_k) \in \mathcal{L}_T$ only with the displayed variables free,

$$\mathsf{SC}_1^- \vdash \forall \vec{x}\big[\varphi^{\mathcal{I}}(\vec{x}) \leftrightarrow (\varphi^\infty)^*(\langle x_1, 0\rangle, \ldots, \langle x_k, 0\rangle)\big].$$

Hence, $\mathcal{I}$ is also an interpretation of $\mathsf{VF}^- + \mathcal{L}_T$-Ind in $\mathsf{SC}_1^- + \mathcal{L}_{\mathrm{SC}}$-Ind. ⊣

§**9. Applications.** In the present section, we will present two applications of the results and techniques of the previous sections.

9.1. Answer to an open problem of [12]. It was asked in [12] as an open problem whether Σ^1_1-AC is conservative over KFW (KF with a global well-ordering of sets); we refer the reader to [12] for the definitions of all the systems and axioms of second-order set theory discussed in this subsection. The proof-theoretic equivalence of Σ^1_1-AC⟦PA⟧ and KF⟦PA⟧ over arithmetic is well-known (see [9]), but the known proof of the conservation Σ^1_1-AC⟦PA⟧ $\subset_{\mathcal{L}_\mathbb{N}}$ KF⟦PA⟧ uses a technique that is not yet known to be applicable to those systems over set theory. In this subsection, we will show that the conservation holds also over set theory. Precisely, what we will literally show is that KF is conservative over Π^1_0-Coll; however, Σ^1_1-AC is identical as a theory with Π^1_0-Coll plus a global choice GC ([12, p. 1489]), and the conservation Σ^1_1-AC $\subset_{\mathcal{L}_\in}$ KFW can be shown in an exactly parallel manner, since the addition of a global well-ordering of sets does not affect all the relevant arguments.

Theorem 9.1. Π^1_0-Coll $\subset_{\mathcal{L}_\in}$ KF.

Proof. We make the following definitions in $\mathsf{KP}\mathbb{V}^-$: for $u \in \mathcal{U}$ and $x \in \mathcal{S}$,

$$(x)^u := \{v \in \mathcal{U} \mid \langle v, u\rangle^{\mathcal{U}} \in_1 x\} \quad \text{and} \quad \mathcal{P}^{\mathcal{S}}(x) := \{y \in \mathcal{S} \mid \forall z(z \in_1 y \to z \in_1 x)\}.$$

By interpreting sets and classes of second-order set theory by urelements $u \in \mathcal{U}$ ($= \mathsf{V}$) and sets $x \in \mathcal{P}^{\mathcal{S}}(\mathsf{V})$, respectively, we obviously have a syntactic embedding of $\mathsf{NBG} + \Sigma^1_\infty$-Sep $+ \Sigma^1_\infty$-Repl in $\mathsf{KP}\mathbb{V}$, where Σ^1_∞-Sep and Σ^1_∞-Repl are the separation and replacement schemata extended for all second-order formulae. With this interpretation, each instance of Π^1_0-Coll is translated into

$$(\forall x \in \mathcal{U})(\exists y \in \mathcal{P}^{\mathcal{S}}(\mathsf{V}))\, \varphi^{\mathcal{U}}(x, y) \to (\exists z \in \mathcal{P}^{\mathcal{S}}(\mathsf{V}))(\forall x \in \mathcal{U})(\exists u \in \mathcal{U})\, \varphi^{\mathcal{U}}(x, (z)^u),$$

for some $\varphi \in \mathcal{L}_\in$; note that $\varphi^{\mathcal{U}}$ is Δ_0 for every $\varphi \in \mathcal{L}_\in$. We call this schema ($\Pi^1_0$-$\mathrm{Coll}_{\mathsf{KP}}$). Since we have shown $\mathsf{SC}_1 \subset_{\mathcal{L}_\in}$ KF, it suffices to show $\mathsf{SC}_1 \vdash$ (Π^1_0-$\mathrm{Coll}_{\mathsf{KP}}$)*: the proof is essentially a formalization of Theorem 6D.3 of [19].

Suppose the antecedent of an instance of (Π^1_0-$\mathrm{Coll}_{\mathsf{KP}}$)* holds. We take an inductive relation P so that

$$P(x, y) \;:\Leftrightarrow\; x \in \mathcal{U}^* \to \big((y \in \mathcal{P}^{\mathcal{S}}(\mathsf{V}))^* \wedge (\varphi^{\mathcal{U}})^*(x, y)\big).$$

We have $\forall x \exists y P(x, y)$ by the supposition. It follows by Theorem 5.8 that there is a hyperelementary Q such that $Q \subset P$ and $\forall x \exists y Q(x, y)$. Then we put $B := \{a \mid (\exists x \in \mathcal{U}^*) Q(x, \langle a, 1\rangle)\}$. B is hyperelementary and $\langle a, 1\rangle \in \mathcal{S}^*$ for all $a \in B$. From this B we define another hyperelementary $Z \subset \mathcal{U}^*$ so that

$$Z := \big\{w \mid (\exists v, u \in \mathcal{U}^*)(\exists a \in B)\big((w = \langle v, u\rangle^{\mathcal{U}})^* \wedge u = \langle a, 0\rangle \wedge v \in_1^* \langle a, 1\rangle\big)\big\}.$$

By Lemma 7.10, we pick $z \in \mathcal{S}^*$ such that $(\forall w \in \mathcal{U}^* \cup \mathcal{S}^*)[w \in_1^* z \leftrightarrow w \in Z]$. Now, take any $x \in \mathcal{U}^*$. There exists $y = \langle a, 1\rangle \in (\mathcal{P}^{\mathcal{S}}(\mathsf{V}))^*$ with $Q(x, y)$. We have

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

$a \in B$ and let $u = \langle a, 0\rangle \in \mathcal{U}^*$. Then, for all $v \in \mathcal{U}^*$, $(\langle v, u\rangle^{\mathcal{U}} \in_1 z)^*$ iff $v \in_1^* y$: that is, $((z)^u)^* =^* y$ and thus $(\varphi^{\mathcal{U}})^*(x, ((z)^u)^*)$. ⊣

Remark 9.2. Since $\mathsf{KP}\mathbb{V}$ derives Δ-Separation, Δ^1_1-CA is syntactically embeddable in $\mathsf{KP}\mathbb{V}$. By Theorem 80.2 of [12], we also have Σ^1_1-$\mathsf{Coll} \subset_{\mathcal{L}_\in} \Delta^1_1$-$\mathsf{CA} \subset_{\mathcal{L}_\in} \mathsf{KF}$, which gives an alternative proof of Theorem 9.1.[3]

9.2. Embedding of KPu in $\mathsf{ID}_1[\![\mathsf{PA}]\!]$. It is well-known that KPu (and $\mathsf{KP}\omega$) is proof-theoretically equivalent to $\mathsf{ID}_1[\![\mathsf{PA}]\!]$. The proof-theoretic reduction of $\mathsf{ID}_1[\![\mathsf{PA}]\!]$ to KPu is easily obtained via the standard interpretation (see [21, Chapter 11.5] for example), but the converse reduction was originally obtained by means of ordinal analysis due to Jäger [16]. As far as the author knows, a direct syntactic embedding KPu (or $\mathsf{KP}\omega$) in $\mathsf{ID}_1[\![\mathsf{PA}]\!]$ has not been given in the literature.[4] We will give such an embedding in the present subsection.

It is to be observed that all the proofs in Section 7 can be straightforwardly turned into a proof of embeddability of KPu in $\mathsf{SC}_1[\![\mathsf{PA}]\!]$. Hence, for the purpose of the present section, it suffices to show that $\mathsf{SC}_1[\![\mathsf{PA}]\!]$ is embeddable in $\mathsf{ID}_1[\![\mathsf{PA}]\!]$.

Let $\langle\cdot,\cdot\rangle : \mathbb{N}^2 \to \mathbb{N}$ be a bijective pairing function and $(\cdot)_0$ and $(\cdot)_1$ its associated projections. For a class $X \subset \mathbb{N}$, we write $x <_X y$ for $\langle x, y\rangle \in X$, which is not to be confused with $\prec_X$ (Section 5).

We begin with Sato's lemma in [22]. For each $\mathcal{A}(x, \mathfrak{X}) \in \mathfrak{I}(\mathcal{L}_\mathbb{N})$, we set

$$\mathcal{A}'(x, \mathfrak{X}) \;:=\; \neg\mathcal{A}\big((x)_1, \{u \mid u <_\mathfrak{X} (x)_0\}\big).$$

We call $\mathcal{A}'\big(x, \{z \mid \mathcal{A}'(z, \mathfrak{X})\}\big)$ the *derivative* of $\mathcal{A}$; namely, it is equal to

$$\neg\mathcal{A}\Big((x)_1, \Big\{u \mid \neg\mathcal{A}\big((x)_0, \{v \mid v <_\mathfrak{X} u\}\big)\Big\}\Big).$$

We will abuse the notation and write $\mathcal{A}''$ for the derivative of $\mathcal{A}$.

Lemma 9.3 (Sato). *Let $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\mathbb{N})$. The following is provable in $\mathsf{ID}_1^-[\![\mathsf{PA}]\!]$. Let X be a class with $\forall x[\mathcal{A}''(x, X) \leftrightarrow x \in X]$. Suppose we can take the accessible part of $<_X$, i.e., the least class that is progressive with respect to $<_X$: more precisely, it is the unique (if any) class Y such that $Clos_{\mathcal{W}[<_X]}(Y)$ and $Clos_{\mathcal{W}[<_X]}(Z) \to Y \subset Z$ for all classes Z (see p. 879). Then, there is a class that satisfies the SC_1^--axioms* (SC0)–(SC2) *for $\mathcal{A}$ in place of $\prec_\mathcal{A}$.*[5]

Theorem 9.4. *$\mathsf{SC}_1^-[\![\mathsf{PA}]\!]$ is a definitional extension of $\mathsf{ID}_1^-[\![\mathsf{PA}]\!]$.*

---

[3]The proof of Theorem 80.2 of [12] is flawed, but the statement itself is true and the claim that $\mathsf{PZF}_1 \subset \Delta^1_1$-$\mathsf{CA}$ is also true; for, the class-theoretic counterpart of Σ^1_∞-TI (see [23]) is provable in $\mathsf{NBG} + \Sigma^1_\infty$-Sep $+ \Sigma^1_\infty$-Repl. The proof of Theorem 80.1 of [12] is also flawed, but this statement is an immediate consequence of the main result of [17] and Theorem 18 of [12].

[4]After the submission of the present article, I was informed by Prof. Wolfram Pohlers that the same result was already obtained by Christian Tapp in [24]; but Tapp's thesis is not widely available, and so I keep this result in the present article.

[5]This statement is actually a combination of Theorems 5 and 7 of [22]. Sato gives these theorems for second-order systems of fixed-points, but, as Sato himself notes in [22, Section 8], his proofs can be generalized for first-order cases; there he only considers first-order systems ID_1 ($[\![\mathsf{PA}]\!]$) of fixed-points with the axiom schemata extended to the whole language, but the extension of the schemata is in fact not necessary for the theorems.

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

PROOF. Let $\mathcal{A}(x, \mathfrak{X}) \in \mathfrak{I}(\mathcal{L}_{\mathbb{N}})$. Then we define $\mathcal{B}(x, \mathfrak{X})$ as

$$\mathcal{A}\Big((x)_0, \Big\{u \mid \neg\mathcal{A}\big((x)_1, \{v \mid u \not<_{\mathfrak{X}} v\}\big)\Big\}\Big),$$

where $\mathfrak{X}$ occurs only positively in $\mathcal{B}$ and thus $\mathcal{B} \in \mathfrak{I}(\mathcal{L}_{\mathbb{N}})$. We set $\lhd := \{\langle x, y\rangle \mid \langle y, x\rangle \notin J_{\mathcal{B}}\}$ and we write $x \lhd y$ for $\langle x, y\rangle \in \lhd$; namely, for $x, y \in \mathbb{N}$, we have $x \lhd y \;\Leftrightarrow\; \langle y, x\rangle \notin J_{\mathcal{B}}$. Since $J_{\mathcal{B}}$ is a fixed-point of $\mathcal{B}$, we have

$$\begin{aligned} x \lhd y \;&\Leftrightarrow\; \neg\mathcal{A}\big(y, \{u \mid \neg\mathcal{A}(x, \{v \mid \langle u, v\rangle \notin J_{\mathcal{B}}\})\}\big) \\ &\Leftrightarrow\; \neg\mathcal{A}\big(y, \{u \mid \neg\mathcal{A}(x, \{v \mid v \lhd u\})\}\big) \;\Leftrightarrow\; \mathcal{A}''(\langle x, y\rangle, \lhd). \end{aligned}$$

Namely, $\lhd$ satisfies the first condition of Sato's Lemma. Then, since $\lhd$ is coinductive, we can take its accessible part $Acc(\lhd)$ $(= J_{\mathcal{W}[\lhd]}$, see p. 879) by Theorem 5.2 (modified for arithmetic), and thus the second condition is also satisfied. ⊣

We have a canonical translation of $\mathcal{L}_{\mathbb{N}}$ in the language of KPu in which the translation of each $\varphi \in \mathcal{L}_{\mathbb{N}}$ is of the form of the relativization $\varphi^{\mathcal{U}}$ to the class $\mathcal{U}$ of urelements; note that since KPu has a constant N for the set of urelements, $\varphi^{\mathcal{U}}$ is equivalent to a Δ_0-formulae φ^{N} for all $\varphi \in \mathcal{L}_{\mathbb{N}}$. Now, by Theorem 9.4, we have an embedding of KPu in $\mathsf{ID}_1[\![\mathsf{PA}]\!]$, from which we obtain in a parallel manner to Theorem 7.20 that if $\mathsf{KPu} \vdash \varphi^{\mathcal{U}}$ then $\mathsf{ID}_1[\![\mathsf{PA}]\!] \vdash \varphi$, for all $\varphi \in \mathcal{L}_{\mathbb{N}}$. A parallel argument gives an embedding of $\mathsf{KP}\mathbb{V}^{(-)}$ in $\mathsf{ID}_1^{(-)}$ over set theory too.

$\mathsf{KP}\omega$ is formulated over $\mathcal{L}_{\in}$, and there is a canonical translation φ^{ω} of $\mathcal{L}_{\mathbb{N}}$ in $\mathcal{L}_{\in}$. We can regard $\mathsf{KP}\omega$ as a subsystem of KPu by interpreting $\forall x \mapsto (\forall x \in \mathcal{S})$. Now, KPu proves N and ω are isomorphic (as $\mathcal{L}_{\mathbb{N}}$-structures) and thus $\mathsf{KPu} \vdash \varphi^{\omega} \leftrightarrow \varphi^{\mathsf{N}}$ for all $\varphi \in \mathcal{L}_{\mathbb{N}}$. Hence, we also have an embedding of $\mathsf{KP}\omega$ in $\mathsf{ID}_1[\![\mathsf{PA}]\!]$, which entails that if $\mathsf{KP}\omega \vdash \varphi^{\omega}$ then $\mathsf{ID}_1[\![\mathsf{PA}]\!] \vdash \varphi$, for all $\varphi \in \mathcal{L}_{\mathbb{N}}$.

THEOREM 9.5. KPu *and* $\mathsf{KP}\omega$ *are embeddable in* $\mathsf{ID}_1[\![\mathsf{PA}]\!]$.

§10. On the strength of the replacement axiom. We have seen that the inter-theoretical relation between axiomatic systems of truth changes when we replace the traditional arithmetical base system by a set-theoretic one. It is observed that $\mathcal{L}_{\mathrm{SC}}$-Repl plays a crucial role in the proof of Sato's theorem and thus $\mathcal{L}_T$-Repl is the main cause of this disanalogy. Then, when we drop it, can we still somehow obtain the equivalence of the noncompositional and compositional systems of truth over set theory? The next theorem shows that the answer is affirmative but in a trivial sense.

THEOREM 10.1. 1. $\mathsf{ID}_1^- + \mathcal{L}_{\mathrm{Fix}}$-Sep $\subset_{\mathcal{L}_{\in}}$ ZF. 2. $\mathsf{SC}_1^- + \mathcal{L}_{\mathrm{SC}}$-Sep $\subset_{\mathcal{L}_{\in}}$ ZF. 3. $\mathsf{VF}^- + \mathcal{L}_T$-Sep $\subset_{\mathcal{L}_{\in}}$ ZF. *Hence, in particular,* $\mathsf{VF}^- + \mathcal{L}_T$-Sep $=_{\mathcal{L}_{\in}}$ $\mathsf{KF}^- + \mathcal{L}_T$-Sep.

PROOF. The proof is a generalization of Theorem 20 of [12]. Let $\mathcal{L}'_{\in}$ be $\mathcal{L}_{\in} \cup \{c\}$ for a fresh constant symbol c. We define an $\mathcal{L}'_{\in}$-theory T by

$$\begin{aligned} \mathsf{ZF} + \mathcal{L}'_{\in}\text{-Sep} + \mathcal{L}'_{\in}\text{-Repl} + \{\exists\alpha \in On(c = V_\alpha \wedge \text{`}\alpha \text{ is limit'})\} \\ + \big\{(\forall\vec{x} \in c)\big(\varphi^c(\vec{x}) \leftrightarrow \varphi(\vec{x})\big) \mid \varphi(\vec{x}) \in \mathcal{L}_{\in}\big\}, \end{aligned}$$

where $\varphi^c(\vec{x})$ is the relativization of $\varphi(\vec{x})$ to the set c; note that φ here does not contain c. Due to the Reflection Principle, we have $\mathsf{T} \subset_{\mathcal{L}_{\in}}$ ZF. Now we will work within T. For each $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_{\in})$, we can standardly take the least fixed-point $I_{\Phi_{\mathcal{A}}^{\langle c,\in\rangle}}$ of

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

the inductive operator $\Phi_{\mathcal{A}}^{\langle c,\in\rangle} : \mathcal{P}(c) \to \mathcal{P}(c)$. Consider the following translation of $\mathcal{L}_{\mathrm{Fix}}$ to $\mathcal{L}'_{\in}$: $\forall x$ and $\exists x$ are translated to $\forall x \in c$ and $\exists x \in c$; $\in$ is translated to itself; finally $x \in J_{\mathcal{A}}$ is translated to $x \in I_{\Phi_{\mathcal{A}}^{\langle c,\in\rangle}}$. This gives an interpretation of ID_1^- in T. Since $c = V_\alpha$ for some limit $\alpha \in On$, the interpretation of $\mathcal{L}_{\mathrm{Fix}}$-Sep automatically holds. Now, if $\mathsf{ID}_1^- + \mathcal{L}_{\mathrm{Fix}}\text{-Sep} \vdash \sigma$ for $\sigma \in \mathcal{L}_{\in}$, then $\mathsf{T} \vdash \sigma^c$ and thus $\mathsf{T} \vdash \sigma$ due to the reflection axioms postulated for T. The other claims can be proven similarly. $\dashv$

COROLLARY 10.2. $\mathsf{KP}\mathbb{V}^- + (\mathrm{Found}_0^+) + (\mathrm{Sep}_0^+) \subset_{\mathcal{L}_\in} \mathsf{ZF}$.

§11. Schematic reflective closure $\mathsf{VF}^*[\![\mathsf{PA}]\!]$ over arithmetic. Feferman [9] presented the notion of *schematic reflective closure* of schematic systems such as PA and ZF. Feferman's original definition is based on the KF-axioms of truth, but we can generalize this notion with other axiomatizations of truth like VF.

Let $\mathcal{L}_{\mathbb{N}}(P) := \mathcal{L}_{\mathbb{N}} \cup \{P\}$ for a fresh unary predicate symbol P. We define $\mathcal{L}_t(P) := \mathcal{L}_{\mathbb{N}}(P) \cup \{T\}$ as the language of axiomatic systems of truth over arithmetic with a predicate variable P. For a first-order language $\mathcal{L} \supset \mathcal{L}_{\mathbb{N}}$, the $\mathcal{L}$-system $\mathsf{PA}_{\mathcal{L}}$ is the extension of PA with the induction schema extended for $\mathcal{L}$.

DEFINITION 11.1. The $\mathcal{L}_t(P)$-system $\mathsf{VF}(P)[\![\mathsf{PA}]\!]$ is defined as $\mathsf{PA}_{\mathcal{L}_t(P)}$ plus the VF-axioms for $\mathcal{L}_t(P)$, formulated for arithmetic, and the following new axiom:

$$\mathbf{P}: \qquad \forall x(T\ulcorner P\dot{x}\urcorner \leftrightarrow Px).$$

Here we assume P is included in our coding. Another $\mathcal{L}_t(P)$-system $\mathsf{KF}(P)[\![\mathsf{PA}]\!]$ is defined as $\mathsf{PA}_{\mathcal{L}_t(P)}$ plus the KF-axioms for $\mathcal{L}_t(P)$ and $\mathbf{P}$; see [11, Section 3.3].

DEFINITION 11.2 (P-Substitution). Let $\mathcal{L}' \supset \mathcal{L}_{\mathbb{N}}(P)$. A new inference rule, P-substitution (for $\mathcal{L}'$) is defined as

$$\frac{\varphi(P)}{\varphi(\psi(\hat{x}))}\ (P\text{-Subst}_{\mathcal{L}'}),\quad \text{for } \varphi(P) \in \mathcal{L}_{\mathbb{N}}(P) \text{ and } \psi(x) \in \mathcal{L}'.$$

$\mathcal{L}_t(P)$-systems $\mathsf{VF}^*[\![\mathsf{PA}]\!]$ and $\mathsf{KF}^*[\![\mathsf{PA}]\!]$ are defined as $\mathsf{VF}(P)[\![\mathsf{PA}]\!] + (P\text{-Subst}_{\mathcal{L}_t(P)})$ and $\mathsf{KF}(P)[\![\mathsf{PA}]\!] + (P\text{-Subst}_{\mathcal{L}_t(P)})$, respectively. Feferman [9] proved that $\mathsf{KF}^*[\![\mathsf{PA}]\!]$ has the strength of predicative limit and is equivalent to ramified analysis.

We next consider incorporating the P-Substitution rule into first-order systems of inductive definitions. We define $\mathcal{L}_{\mathrm{fix}}(P)$ as $\mathcal{L}_{\mathbb{N}}(P)$ plus unary predicate $J_{\mathcal{A}}$ associated with each inductive operator form $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_{\mathbb{N}}(P))$.

DEFINITION 11.3. The $\mathcal{L}_{\mathrm{fix}}(P)$-system $\mathsf{ID}_1(P)[\![\mathsf{PA}]\!]$ is defined as $\mathsf{PA}_{\mathcal{L}_{\mathrm{fix}}(P)}$ plus

$$Clos_{\mathcal{A}}(J_{\mathcal{A}}), \text{ for each } \mathcal{A} \in \mathfrak{I}(\mathcal{L}_{\mathbb{N}}(P));$$
$$Clos_{\mathcal{A}}(\Psi) \to J_{\mathcal{A}} \subset \Psi, \text{ for each } \mathcal{A} \in \mathfrak{I}(\mathcal{L}_{\mathbb{N}}(P)) \text{ and } \Psi \in \mathcal{L}_{\mathrm{fix}}(P).$$

The $\mathcal{L}_{\mathrm{fix}}(P)$-system $\mathsf{ID}_1^*[\![\mathsf{PA}]\!]$ is defined as $\mathsf{ID}_1(P)[\![\mathsf{PA}]\!] + (P\text{-Subst}_{\mathcal{L}_{\mathrm{fix}}(P)})$.

Also, although we will not study them in the present article, $\mathcal{L}_{\mathrm{fix}}(P)$-systems $\widehat{\mathsf{ID}}_1(P)[\![\mathsf{PA}]\!]$ and $\widehat{\mathsf{ID}}_1^*[\![\mathsf{PA}]\!]$ are defined in an obvious manner, and we can show that $\widehat{\mathsf{ID}}_1(P)[\![\mathsf{PA}]\!] =_{\mathcal{L}_{\mathbb{N}}} \mathsf{KF}(P)[\![\mathsf{PA}]\!]$ and $\widehat{\mathsf{ID}}_1^*[\![\mathsf{PA}]\!] =_{\mathcal{L}_{\mathbb{N}}} \mathsf{KF}^*[\![\mathsf{PA}]\!]$.

The form of P-Substitution rule resembles the Bar Rule and might be seen as a first-order counterpart of the Bar Rule. In fact, the way in which P-Substitution increases the strength of KF or $\widehat{\mathsf{ID}}_1$ up to the predicative limit is pretty much the same as the way in which the Bar Rule increases the strength of second-order systems of arithmetic like Σ^1_1-AC. However, as we will see, it does not add any strength to $\mathsf{VF}[\![\mathsf{PA}]\!]$ and $\mathsf{ID}_1[\![\mathsf{PA}]\!]$.

LEMMA 11.4. $\mathsf{VF}^*[\![\mathsf{PA}]\!] \subset_{\mathcal{L}_{\mathbb{N}}} \mathsf{ID}^*_1[\![\mathsf{PA}]\!]$.

PROOF. We can embed $\mathsf{VF}(P)[\![\mathsf{PA}]\!]$ in $\mathsf{ID}_1(P)[\![\mathsf{PA}]\!]$ in the same manner to Theorem 9.5 (with obvious modifications for arithmetic). This embedding can be extended to an embedding of $\mathsf{VF}^*[\![\mathsf{PA}]\!]$ and $\mathsf{ID}^*_1[\![\mathsf{PA}]\!]$ by a straightforward generalization of Lemma 31 of [11]. We note that this proof actually gives an interpretation of $\mathsf{VF}^*[\![\mathsf{PA}]\!]$ in $\mathsf{ID}^*_1[\![\mathsf{PA}]\!]$ that preserves the $\mathcal{L}_{\mathbb{N}}(P)$-part. ⊣

§12. Analysis of $\mathsf{ID}^*_1[\![\mathsf{PA}]\!]$. We will give ordinal analysis of $\mathsf{ID}^*_1[\![\mathsf{PA}]\!]$, which then gives analysis of $\mathsf{VF}^*[\![\mathsf{PA}]\!]$ via Lemma 11.4. We use the same notation of [13], and the following definitions and results except 12.3–12.6 are all straightforward generalizations of those in [13, Section 6] (or [21, Section 9]) for our current setting.

A general treatment of systems of ν-iterated inductive definitions is aimed for in [13] and thus ζ-ary disjunctions for $\zeta \leq \Omega_\nu$ have to be taken into consideration therein, where Ω_ν is the ν-th uncountable cardinal. However, since we focus on noniterated inductive definitions here, we can restrict all our arguments to $\zeta \leq \Omega_1$ and accordingly simplify some definitions; we will write Ω for Ω_1.

A first-order language $\mathcal{L}^\infty_{\mathrm{fix}}(P)$ is defined as

$$\mathcal{L}_{\mathbb{N}}(P) \cup \{I^{<\xi}_{\mathcal{A}} \mid \xi \leq \Omega \;\&\; \mathcal{A} \in \Im(\mathcal{L}_{\mathbb{N}}(P))\},$$

where $I^{<\xi}_{\mathcal{A}}$ is a unary predicate. As in Section 8, we assume that formulae and sentences are expressed in their negation normal forms in the present section; cf. [13, Section 6.1].

For each $\mathcal{A}(x, \mathfrak{X}) \in \Im(\mathcal{L}(P))$ and $\mathcal{L}_{\mathbb{N}}$-term t, we write $I^{\xi}_{\mathcal{A}}(t)$ for $\mathcal{A}(t, I^{<\xi}_{\mathcal{A}})$; it is to be noted that the $\mathcal{A}$ here and thus $I^{\xi}_{\mathcal{A}}(t)$ may contain P.

We divide the $\mathcal{L}^\infty_{\mathrm{fix}}(P)$-*sentences* into two types, namely, $\bigvee$-type and $\bigwedge$-type, and assign each $\mathcal{L}^\infty_{\mathrm{fix}}(P)$-formula A its characteristic sequence $CS(A) \subset \mathcal{L}^\infty_{\mathrm{fix}}(P)$, rank $rk(A) \in On$, and parameters $\mathsf{par}(A) \subset On$.

A true closed $\mathcal{L}_{\mathbb{N}}$-literal is of $\bigwedge$-type, and a false closed $\mathcal{L}_{\mathbb{N}}$-literal is of $\bigvee$-type. For every closed $\mathcal{L}_{\mathbb{N}}$-term t, both Pt and $\neg Pt$ are neither $\bigwedge$-type nor $\bigvee$-type. For an $\mathcal{L}_{\mathbb{N}}(P)$-literal A, we set $\mathsf{par}(A) = CS(A) = \emptyset$ and $rk(A) = 0$.

For $\mathcal{L}^\infty_{\mathrm{fix}}(P)$-sentences A and B, the sentences $A \wedge B$ and $\forall x A$ are of $\bigwedge$-type, and the sentences $A \vee B$ and $\exists x A$ are of $\bigvee$-type. We define their ranks, parameters and characteristic sequences as follows:

$$\begin{aligned} rk(A \Box B) &= \max\{rk(A), rk(B)\} + 1, & rk(\mathsf{Q}xA) &= rk(A(\underline{0})) + 1, \\ \mathsf{par}(A \Box B) &= \mathsf{par}(A) \cup \mathsf{par}(B), & \mathsf{par}(\mathsf{Q}xA) &= \mathsf{par}(A(\underline{0})), \\ CS(A \Box B) &= \{A, B\}, & CS(\mathsf{Q}xA) &= \{A(\underline{n}) \mid n \in \mathbb{N}\}, \end{aligned}$$

where $\Box \in \{\wedge, \vee\}$, $\mathsf{Q} \in \{\forall, \exists\}$, and $\underline{n}$ is the numeral for $n \in \mathbb{N}$.

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

Let $\mathcal{A}(x, \mathfrak{X}) \in \mathfrak{I}(\mathcal{L}_{\mathbb{N}}(P))$ and $\xi \leq \Omega$. For each closed $\mathcal{L}_{\mathbb{N}}$-term s, $I^{<\xi}_{\mathcal{A}}(s)$ is of $\bigvee$-type and $\neg I^{<\xi}_{\mathcal{A}}(s)$ is of $\bigwedge$-type. Their ranks and parameters are defined by

$$rk\big(I^{<\xi}_{\mathcal{A}}(s)\big) = rk\big(\neg I^{<\xi}_{\mathcal{A}}(s)\big) := \omega \cdot \xi; \quad \mathsf{par}\big(I^{<\xi}_{\mathcal{A}}(s)\big) = \mathsf{par}\big(\neg I^{<\xi}_{\mathcal{A}}(s)\big) := \{\xi\};$$
$$CS\big(I^{<\xi}_{\mathcal{A}}(s)\big) = \{I^{\zeta}_{\mathcal{A}}(s) \mid \zeta < \eta\}; \qquad CS\big(\neg I^{<\xi}_{\mathcal{A}}(s)\big) = \{\neg I^{\zeta}_{\mathcal{A}}(s) \mid \zeta < \eta\}.$$

We define a translation $\Phi^\star$ of $\mathcal{L}_{\rm fix}(P)$-sentences Φ in $\mathcal{L}^{\infty}_{\rm fix}(P)$: for each $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_{\mathbb{N}}(P))$ and closed $\mathcal{L}_{\mathbb{N}}$-term s, we set $J^\star_{\mathcal{A}}(s) := I^{<\Omega}_{\mathcal{A}}(s)$ and $\neg J^\star_{\mathcal{A}}(s) := \neg I^{<\Omega}_{\mathcal{A}}(s)$; all the other atomic $\mathcal{L}_{\rm fix}(P)$-sentences, the boolean connectives, and the quantifiers are preserved. We observe that, for an $\mathcal{L}^{\infty}_{\rm fix}(P)$-sentence A, we have

$$\mathsf{par}(A) := \{\xi \mid I^{<\xi}_{\mathcal{B}} \text{ occurs in } A \text{ for some } \mathcal{B} \in \mathfrak{I}(\mathcal{L}_{\mathbb{N}}(P))\}.$$

We can easily show that $rk\big(I^{\xi}_{\mathcal{A}}(s)\big) < \omega \cdot \xi + \omega$ and $rk\big(J^\star_{\mathcal{A}}(s)\big) = \Omega$ for all $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_{\mathbb{N}}(P))$, and thus $rk(\Phi^\star) < \Omega + \omega$ for every $\mathcal{L}_{\rm fix}(P)$-sentence Φ.

We say that an $\mathcal{L}^{\infty}_{\rm fix}$-*sentence* F is of $\bigvee^{\Omega}$-type, or simply $F \in \bigvee^{\Omega}$, when $\xi < \Omega$ for each occurrence of $\neg I^{<\xi}_{\mathcal{A}}(s)$ in F (but F may contain $I^{<\Omega}_{\mathcal{A}}(s)$).

We define a set $C(\alpha, \beta)$ and the collapsing function $\psi_\Omega(\alpha)$, for $\alpha, \beta \in On$, by simultaneous recursion on α, in exactly the same manner as in [13] and [20].[6]

For an ordinal γ, we define an operator $\mathcal{H}_\gamma \colon \mathcal{P}(On) \to \mathcal{P}(On)$ by

$$\mathcal{H}_\gamma(X) := \bigcap\big\{C(\alpha, \beta) \mid X \subset C(\alpha, \beta) \wedge \gamma < \alpha\big\}.$$

Given $Z \subset On$, we define a new operator $\mathcal{H}_\gamma[Z]$ by putting $\mathcal{H}_\gamma[Z](X) := \mathcal{H}_\gamma(X \cup Z)$. In the following, the letters $\mathcal{H}, \mathcal{H}', \mathcal{H}'', \dots$ will be used as syntactic variables ranging over operators $\mathcal{H}_\gamma[X]$ for some $\gamma \in On$ and $X \in \mathcal{P}(On)$, and the word "operator" will mean such an operator $\mathcal{H}_\gamma[X]$ unless otherwise specified. For $\Delta \subset \mathcal{L}^{\infty}_{\rm fix}(P)$ and $F \in \mathcal{L}^{\infty}_{\rm fix}(P)$, we will write $\mathcal{H}[\Delta]$ for $\mathcal{H}[\bigcup_{A \in \Delta} \mathsf{par}(A)]$ and $\mathcal{H}[F]$ for $\mathcal{H}[\{F\}]$; following the convention, we will also write Δ, F for $\Delta \cup \{F\}$.

Definition 12.1. For an operator $\mathcal{H}$ and a finite set Δ of $\mathcal{L}^{\infty}_{\rm fix}(P)$-*sentences*, the relation $\mathcal{H} \vdash^{\alpha}_{\rho,\tau} \Delta$ holds for $\alpha, \rho \in On$ and $\tau \in \{0, 1\}$, if and only if $\alpha \in \mathcal{H}(\emptyset)$, $\mathsf{par}(\Delta) = \bigcup_{A \in \Delta} \mathsf{par}(A) \subset \mathcal{H}(\emptyset)$, and one of the following holds

(Ax): $Ps, \neg Pt \in \Delta$ for closed $\mathcal{L}_{\mathbb{N}}$-terms s and t with the same value (i.e., $s^{\mathbb{N}} = t^{\mathbb{N}}$);
($\bigwedge$): there are $F \in \Delta \cap \bigwedge$ and $\alpha_G < \alpha$ for each $G \in CS(F)$ such that $\mathcal{H}[G] \vdash^{\alpha_G}_{\rho,\tau} \Delta, G$;
($\bigvee$): there are $F \in \bigvee \cap \Delta$, $G \in CS(F)$, and $\alpha_G < \alpha$ such that $\mathcal{H} \vdash^{\alpha_G}_{\rho,\tau} \Delta, G$;
(cut): there are A with $rk(A) < \rho$ and $\alpha_0 < \alpha$ such that $\mathcal{H} \vdash^{\alpha_0}_{\rho,\tau} \Delta, A$ and $\mathcal{H} \vdash^{\alpha_0}_{\rho,\tau} \Delta, \neg A$;
(cl): $\tau = 1$, and there exist some $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_{\mathbb{N}}(P))$, a closed $\mathcal{L}_{\mathbb{N}}$-term s, and $\alpha_0 < \alpha$ such that $I^{<\Omega}_{\mathcal{A}}(s) \in \Delta$ and $\mathcal{H} \vdash^{\alpha_0}_{\rho,\tau} \Delta, I^{\Omega}_{\mathcal{A}}(s)$.

Note that the new clause (Ax) is added to the semiformal system in [13] to deal with the newly added predicate variable P; also, τ only take either 0 or 1 since we need not consider iterations of inductive definitions.

[6]These are defined in [20] and [13] for the sake of ordinal analyses of impredicative systems up to KPi and Δ^1_2-CA plus bar induction; hence, we here include many redundantly large ordinals for our current purpose, such as $\psi_\Omega(\varepsilon_{I+1})$. We could cut off the redundant ones and simplify the definitions of $C(\alpha, \beta)$ and $\psi_\Omega\alpha$; we could replace I by Ω and drop the closure condition for $\sigma \mapsto \Omega_\sigma$ in the definition of $C(\alpha, \beta)$, and only allow Ω in place of κ in $\psi_\kappa\alpha$. Alternatively, we could define the collapsing function in the manner described in [21, Section 9.4].

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

All the basic proof-theoretic properties are standardly shown in the same manner as in [13, Section 6] (or [20]); in particular,

- Controlled Tautology: $\mathcal{H}[\Delta, F]\vdash^{2\cdot rk(F)}_{0,0}\Delta, \neg F(s), F(t)$, for each $F \in \mathcal{L}^{\infty}_{\text{fix}}(P)$ and closed $\mathcal{L}_{\mathbb{N}}$-terms s and t with $s^{\mathbb{N}} = t^{\mathbb{N}}$.
- Predicative Cut-elimination: If $\Omega \notin [\rho, \rho + \omega^{\beta})$, $\beta \in \mathcal{H}$ and $\mathcal{H}\vdash^{\alpha}_{\rho+\omega^{\beta},\tau}\Delta$, then $\mathcal{H}\vdash^{\varphi_{\beta}\alpha}_{\rho,\tau}\Delta$, where φ here denotes the binary Veblen function.[7]

THEOREM 12.2 (Collapsing Theorem). *Let $X \subset On$. Suppose $\gamma \in \mathcal{H}_{\gamma}[X]$, $\Delta \subset \bigvee^{\Omega}$, and $X \subset C\big(\gamma + 1, \psi_{\Omega}(\gamma + 1)\big)$. Then, we have the following*

$$\text{if } \mathcal{H}_{\gamma}[X]\vdash^{\alpha}_{\Omega+1,\tau} \Delta, \text{ then } \mathcal{H}_{\gamma+3^{\Omega+1+\alpha}}[X]\vdash^{\psi_{\Omega}(\gamma+3^{\Omega+1+\alpha})}_{\psi_{\Omega}(\gamma+3^{\Omega+1+\alpha}),0} \Delta.$$

Let $\Delta(P) = \{A_1(P), \dots, A_n(P)\}$ be a finite set of $\mathcal{L}^{\infty}_{\text{fix}}(P)$-sentences possibly interspersed with P. For $B \in \mathcal{L}^{\infty}_{\text{fix}}(P)$, we denote $\{A_1(B), \dots, A_n(B)\}$ by $\Delta(B)$.

LEMMA 12.3. *Suppose $\mathcal{H}\vdash^{\alpha}_{0,0} \Delta^{\star}(P)$ for a finite set $\Delta(P)$ of $\mathcal{L}_{\mathbb{N}}(P)$-sentences. Then, for any $\mathcal{L}_{\text{fix}}(P)$-formula $\Xi(x)$ with only x free, we have*

$$\mathcal{H}\vdash^{\Omega+\omega+\alpha}_{0,0} \Delta^{\star}\big(\Xi^{\star}(\hat{x})\big).$$

PROOF. We first note that $\mathcal{H}\vdash^{\alpha}_{0,0}\Delta^{\star}(P)$ and $\Delta(P) \subset \mathcal{L}_{\mathbb{N}}(P)$ imply that neither (cut) nor (cl) is used in its derivation and also that no $I^{<\xi}_{\mathcal{A}}$ appears in its derivation for any $\xi \leq \Omega$. The claim is shown by induction on α.

If $\Delta^{\star}(P)$ is obtained by (Ax), then $\Delta^{\star}(P)$ contains Ps and $\neg Pt$ for some s and t with $s^{\mathbb{N}} = t^{\mathbb{N}}$. Then $\Delta^{\star}(\Xi^{\star})$ contains $\Xi^{\star}(s)$ and $\neg\Xi^{\star}(t)$, and thus we get $\mathcal{H}\vdash^{\Omega+\omega}_{0,0}\Delta^{\star}(\Xi^{\star})$ by Controlled Tautology, since $\mathsf{par}(\Phi^{\star}) \subset \{\Omega\} \subset \mathcal{H}(\emptyset)$ for every $\Phi \in \mathcal{L}_{\text{fix}}(P)$. If $\Delta^{\star}(P)$ contains a true closed $\mathcal{L}_{\mathbb{N}}$-literal, then so does $\Delta^{\star}(\Xi^{\star})$.

Suppose that the last inference is made by $\bigwedge$-rule and there exists $F(P) \in \Delta^{\star}(P) \cap \bigwedge$ with $CS(F) \neq \emptyset$ such that for all $G \in CS(F)$ there is $\alpha_G < \alpha$ with $\mathcal{H}[G]\vdash^{\alpha_G}_{0,0} \Delta^{\star}(P), G(P)$. Since $\Delta(P) \subset \mathcal{L}_{\mathbb{N}}(P)$, F should be of the form $\forall x\Phi^{\star}(x)$ or $\Phi_0^{\star} \wedge \Phi_1^{\star}$ for some $\mathcal{L}_{\mathbb{N}}(P)$-formulae Φ, Φ_0, and Φ_1. Hence, each $G(P) \in CS(F)$ is equal to $\Psi^{\star}(P)$ for some $\Psi \in \mathcal{L}_{\mathbb{N}}(P)$, $\mathcal{H} = \mathcal{H}[G]$ for all $G(P) \in CS(F)$, and

$$CS\big(F(\Xi^{\star})\big) = \big\{G(\Xi^{\star}) \mid G(P) \in CS(F)\big\}; \tag{3}$$

note that this (3) is not necessarily the case when F is of the form $\neg I^{<\xi}_{\mathcal{A}}(t)$, and so the assumption that $\Delta(P) \subset \mathcal{L}_{\mathbb{N}}(P)$ is crucial here. By the induction hypothesis, for each $G \in CS(F)$, we have $\mathcal{H}\vdash^{\Omega_1+\omega+\alpha_G}_{0,0} \Delta^{\star}(\Xi^{\star}), G(\Xi^{\star})$, and thus we obtain $\mathcal{H}\vdash^{\Omega_1+\omega+\alpha}_{0,0} \Delta^{\star}(\Xi^{\star})$ by $\bigwedge$-rule. The other cases are similarly treated. ⊣

The next is a straightforward generalization of well-known results; see [13, Lemmata 76, 79, and 82] or [21, Section 9].

LEMMA 12.4. *For each axiom Φ of $\mathsf{ID}_1(P)[\![\mathsf{PA}]\!]$, it holds that $\mathcal{H}_0\vdash^{\Omega\cdot 2+\omega}_{\Omega+1,1} \Phi^{\star}$.*

Let $\mathsf{ID}_1^{*}\restriction_n[\![\mathsf{PA}]\!]$ be the system obtained from $\mathsf{ID}_1^{*}[\![\mathsf{PA}]\!]$ by restricting the number of applications of P-$\mathrm{Subst}_{\mathcal{L}_{\text{fix}}(P)}$ to at most n-times.

[7]Due to the new clause (Ax), the proof of Reduction Lemma in [13] (or [20, Lemma 3.4.3.5]) needs slight modification to treat the extra case where the cut-formulae are Ps and $\neg Ps$. Such a modification is well-known and we refer the reader to [21, Section 7.3] and [20, Lemma 2.1.5.7].

THEOREM 12.5. *For each* $n \in \mathbb{N}$, *if* $\mathsf{ID}_1^* \restriction_n [\![\mathsf{PA}]\!] \vdash \Phi(\vec{x})$ *for* $\Phi(\vec{x}) \in \mathcal{L}_{\mathrm{fix}}(P)$, *then there exists some* $\alpha < \psi_\Omega(\varepsilon_{\Omega+1})$ *and* $\gamma < \varepsilon_{\Omega+1}$ *such that* $\mathcal{H}_\gamma \vdash^{\alpha}_{0,0} \Phi^\star(\vec{r})$ *for all closed* $\mathcal{L}_\mathbb{N}$*-terms* $\vec{r}$.

PROOF. The claim is shown by meta-induction on n. Suppose the claim has been shown for m, and let $n = m + 1$. Put T_n to be

$$\mathsf{ID}_1(P)[\![\mathsf{PA}]\!] \cup \{\Theta(\Xi) \mid \mathsf{ID}_1^* \restriction_m [\![\mathsf{PA}]\!] \vdash \Theta(P),\ \Theta(P) \in \mathcal{L}_\mathbb{N}(P), \text{ and } \Xi \in \mathcal{L}_{\mathrm{fix}}(P)\};$$

obviously, $\mathsf{ID}_1^* \restriction_n [\![\mathsf{PA}]\!] \vdash \Phi(\vec{x})$ implies $\mathsf{T}_n \vdash \Phi(\vec{x})$. Let $\mathsf{ID}_1^* \restriction_m [\![\mathsf{PA}]\!] \vdash \Theta(\vec{v}, P)$ and take any $\mathcal{L}_{\mathrm{fix}}(P)$-formula $\Xi(u, \vec{w})$ and closed $\mathcal{L}_\mathbb{N}$-terms $\vec{s}$ and $\vec{t}$. By the induction hypothesis, we have $\mathcal{H}_\gamma \vdash^{\alpha}_{0,0} \Theta^\star(\vec{s}, P)$ for some $\alpha < \psi_\Omega(\varepsilon_{\Omega+1})$ and $\gamma < \varepsilon_{\Omega+1}$. Hence, we get $\mathcal{H}_\gamma \vdash^{\Omega+\omega+\alpha}_{0,0} \Theta^\star\big(\vec{s}, \Xi(\hat{u}, \vec{t})\big)$ by Lemma 12.3. Now assume $\mathsf{T}_n \vdash \Phi(\vec{x})$ and take closed $\mathcal{L}_\mathbb{N}$-terms $\vec{r}$. It follows from the above and Lemma 12.4 that there exist some $n < \omega$ such that $\mathcal{H}_\gamma \vdash^{\Omega\cdot 2+\omega\cdot 2}_{\Omega+1+n,1} \Phi^\star(\vec{r})$. By Predicative Cut-elimination we obtain $\mathcal{H}_\gamma \vdash^{\varphi_0^n(\Omega\cdot 2+\omega\cdot 2)}_{\Omega+1,1} \Phi^\star(\vec{r})$, where $\varphi_0^n(\delta)$ is defined as $\varphi_0^0(\delta) := \varphi_0(\delta+1)$ and $\varphi_0^{k+1} := \varphi_0(\varphi_0^k(\delta))$. Then, by Collapsing Theorem we obtain

$$\mathcal{H}_{\gamma+3^{\Omega+\varphi_0^n(\Omega\cdot 2+\omega\cdot 2)}} \vdash^{\psi_\Omega(\gamma+3^{\Omega+\varphi_0^n(\Omega\cdot 2+\omega\cdot 2)})}_{\psi_\Omega(\gamma+3^{\Omega+\varphi_0^n(\Omega\cdot 2+\omega\cdot 2)}),\,0} \Phi^\star(\vec{r}).$$

We have $\gamma + 3^{\Omega+\varphi_0^n(\Omega\cdot 2+\omega\cdot 2)} < \varepsilon_{\Omega+1}$. By Predicative Cut-elimination we get

$$\mathcal{H}_{\gamma+3^{\Omega+\varphi_0^n(\Omega\cdot 2+\omega\cdot 2)}} \vdash^{\varphi_{\psi_\Omega(\gamma+3^{\Omega+\varphi_0^n(\Omega\cdot 2+\omega\cdot 2)})}\left(\psi_\Omega\left(\gamma+3^{\Omega+\varphi_0^n(\Omega\cdot 2+\omega\cdot 2)}\right)\right)}_{0,0} \Phi^\star(\vec{r}),$$

where $\varphi_{\psi_\Omega(\gamma+3^{\Omega+\varphi_0^n(\Omega\cdot 2+\omega\cdot 2)})}\left(\psi_\Omega\left(\gamma+3^{\Omega+\varphi_0^n(\Omega\cdot 2+\omega\cdot 2)}\right)\right) < \psi_\Omega(\varepsilon_{\Omega+1})$. ⊣

Hence, $\mathsf{ID}_1^*[\![\mathsf{PA}]\!]$ and $\mathsf{ID}_1[\![\mathsf{PA}]\!]$ has the same proof-theoretic ordinal (suitably defined), and the proof gives their proof-theoretic equivalence for $\mathcal{L}_\mathbb{N}(P)$.

THEOREM 12.6. $\mathsf{VF}[\![\mathsf{PA}]\!] =_{\mathcal{L}_\mathbb{N}} \mathsf{VF}^*[\![\mathsf{PA}]\!] =_{\mathcal{L}_\mathbb{N}} \mathsf{ID}_1[\![\mathsf{PA}]\!] =_{\mathcal{L}_\mathbb{N}} \mathsf{ID}_1^*[\![\mathsf{PA}]\!]$.

It is shown in [13] that $(\mathsf{ID}_1^2)_0$ plus the Bar Rule is stronger than $(\mathsf{ID}_1^2)_0$ and its proof-theoretic ordinal is $\psi_\Omega(\varepsilon_{\Omega+\Omega})$. Since $(\mathsf{ID}_1^2)_0$ is the second-order counterpart of ID_1, Theorem 12.6 indicates that P-Substitution does not always behave as an equivalent first-order counterpart of the Bar Rule.

§13. Discussion and conclusion.

The notion of mutual truth-definability between axiomatic systems of truth is introduced in [11] in an attempt to formally capture the "conceptual equivalence" of different axiomatic conceptions of truth, which is a strong equivalence relation of axiomatic systems implying both proof-theoretic equivalence and mutual conservation. The mutual truth-definability of KF and VF over ZF follows from Theorems 3.2, 4.8, and 8.10. In an exactly parallel manner, we can show the mutual truth-definability of $\mathsf{VF}(P)$ and $\mathsf{KF}(P)$ over ZF, and this mutual truth-definability can be extended to that of VF^* and KF^* over ZF by Lemma 31 of [11] (modified for set-theoretic base systems). These make a contrast against the failure of the mutual truth-definability of those systems over arithmetic. Also, although we do not yet know whether KF^* (and VF^*) is stronger than KF (VF resp.) over ZF, Theorem 12.6 gives another disanalogy in either case: if VF^* is stronger than VF, then the schematic reflective closure VF^* adds deductive power over set

https://doi.org/10.1017/jsl.2017.77 Published online by Cambridge University Press

theory while it does *not* over arithmetic; otherwise, the schematic reflective closure KF^* does *not* add deductive power over set theory while it does over arithmetic.

Some results of the present article may also give a new perspective to the so-called conservativeness argument against deflationism about truth. In brief, the argument goes as follows: deflationism about truth requires that the truth predicate and its axioms should not enable any new theorem that is not derivable without them, but adequate axiomatic systems of truth are not conservative over their base systems and thus deflationism is untenable. Traditionally, in the context of the conservativeness argument, only axiomatic systems of truth over arithmetic, such as $\mathsf{KF}[\![\mathsf{PA}]\!]$, are taken into account and referred to as the "evidence" of the claim that adequate axiomatic systems of truth are not conservative over their bases. In reply to this argument, Field [10] points out that the failure of conservativeness is caused by extending the arithmetical induction schema to the truth predicate, and then argues that the extension of the schema is not justifiable solely in virtue of the concept of truth. Theorem 10.1 suggests that different schemata and base systems have different implications for the argument. We leave more philosophical discussions on this issue to [14].

For the future study, we list below two open problems:

1. Do KF^* and VF^* have the same $\mathcal{L}_\in$-theorems as KF and VF?
2. Are KF^- and VF^- mutually truth-definable?

My conjecture is affirmative to the former and negative to the latter.

§**14. Appendix.** In this appendix, we will show that SC_1 is equivalent to Sato's [22, p. 106] original system ID_1^+ of stage comparison prewellorderings.

Definition 14.1. Let $\mathcal{L}'_{\mathrm{SC}}$ be a sublanguage of $\mathcal{L}_{\mathrm{SC}}$ defined by

$$\mathcal{L}'_{\mathrm{SC}} = \mathcal{L}_\in \cup \{R_{\mathcal{A}} \mid \mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)\} = \mathcal{L}_{\mathrm{SC}} \setminus \{J_{\mathcal{A}} \mid \mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)\}.$$

The $\mathcal{L}'_{\mathrm{SC}}$-system $\mathsf{ID}_1^+\!\restriction$ is defined as ZF + (SC0) with (SC2) restricted to $\mathcal{L}'_{\mathrm{SC}}$ plus:
(ID1$^+$)$\exists z\big(\mathcal{A}(x, \prec_{\mathcal{A}}\restriction_z) \wedge \neg\mathcal{A}(y, \prec_{\mathcal{A}}\restriction_z)\big) \to x \prec_{\mathcal{A}} y$, for every $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)$;
(ID2$^+$)$x \prec_{\mathcal{A}} y \leftrightarrow \exists z\big(z \prec_{\mathcal{A}} y \wedge \mathcal{A}(x, \prec_{\mathcal{A}}\restriction_z)\big)$, for every $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)$.
Then we set $\mathsf{ID}_1^+ := \mathsf{ID}_1^+\!\restriction + \mathcal{L}'_{\mathrm{SC}}$-Sep $+ \mathcal{L}'_{\mathrm{SC}}$-Repl.

Sato showed that the transitivity of $\prec_{\mathcal{A}}$ and the converse of (ID1$^+$) are provable in $\mathsf{ID}_1^+\!\restriction$ [22, Lemma 7], and that $\{x \mid \exists y \mathcal{A}(x, \prec_{\mathcal{A}}\restriction_y)\}$ is a least fixed-point of each $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)$ provably in $\mathsf{ID}_1^+\!\restriction$ [22, Lemma 6], which induces an embedding $\flat$ of $\mathcal{L}_{\mathrm{SC}}$ into $\mathcal{L}^-_{\mathrm{SC}}$ in which $J^\flat_{\mathcal{A}}(x) := \exists y \mathcal{A}(x, \prec_{\mathcal{A}}\restriction_y)$.

Lemma 14.2. *Let $\mathcal{A} \in \mathfrak{I}(\mathcal{L}_\in)$. The following are provable in $\mathsf{ID}_1^+\!\restriction$.*

1. *For all x and y, if $y \not\prec_{\mathcal{A}} x$, then $\prec_{\mathcal{A}}\restriction_x \subset \prec_{\mathcal{A}}\restriction_y$.*
2. *For all $x \in J^\flat_{\mathcal{A}}$, it holds that $\mathcal{A}(x, \prec_{\mathcal{A}}\restriction_x)$.*

Proof. 1. Suppose $y \not\prec_{\mathcal{A}} x$. Take any $w \prec_{\mathcal{A}} x$. We have $\mathcal{A}(w, \prec_{\mathcal{A}}\restriction_u)$ for some $u \prec_{\mathcal{A}} x$ by (ID2$^+$). If $w \not\prec_{\mathcal{A}} y$ were the case, we would have $\mathcal{A}(y \prec_{\mathcal{A}}\restriction_u)$ by (ID1$^+$) and thus $y \prec_{\mathcal{A}} x$ by (ID2$^+$).

2. Let $x \in J^\flat_{\mathcal{A}}$ and pick a $\prec_{\mathcal{A}}$-minimal z with $\mathcal{A}(x, \prec_{\mathcal{A}}\restriction_z)$. If $x \prec_{\mathcal{A}} z$, there would be $w \prec_{\mathcal{A}} z$ with $\mathcal{A}(x, \prec_{\mathcal{A}}\restriction_w)$ by (ID2$^+$), which contradicts the minimality of z. Hence, we get $\prec_{\mathcal{A}}\restriction_z \subset \prec_{\mathcal{A}}\restriction_x$ by 1 and thus $\mathcal{A}(x, \prec_{\mathcal{A}}\restriction_x)$ by monotonicity. ⊣

LEMMA 14.3. $\mathsf{ID}_1^+\!\restriction \vdash (\mathrm{SC1})^\flat$. *Hence,* SC_1^- *is a definitional extension of* $\mathsf{ID}_1^+\!\restriction$.

PROOF. If $x \prec_{\mathcal{A}} y$, then $x \in J_{\mathcal{A}}^\flat$ and $\neg\mathcal{A}(y, \prec_{\mathcal{A}}\!\restriction_x)$ by $(\mathrm{ID2}^+)$ and the irreflexivity of $\prec_{\mathcal{A}}$. The converse follows from $(\mathrm{ID1}^+)$ and Lemma 14.2.2. ⊣

LEMMA 14.4. $\mathsf{SC}_1 \vdash (\mathrm{ID1}^+)$ *and* $\mathsf{SC}_1 \vdash (\mathrm{ID2}^+)$.

PROOF. For the first claim, suppose $\mathcal{A}(x, \prec_{\mathcal{A}}\!\restriction_z)$ and $\neg\mathcal{A}(y, \prec_{\mathcal{A}}\!\restriction_z)$ for some z. We have $z \not\prec_{\mathcal{A}} x$ by (SC1) and $x \in J_{\mathcal{A}}$ by $\prec_{\mathcal{A}}\!\restriction_z \subset J_{\mathcal{A}}$. Hence we get $\prec_{\mathcal{A}}\!\restriction_x \subset \prec_{\mathcal{A}}\!\restriction_z$ by Lemma 4.5 and thus $\neg\mathcal{A}(y, \prec_{\mathcal{A}}\!\restriction_x)$ by monotonicity, which implies $x \prec_{\mathcal{A}} y$ by (SC1). For the second claim, let $z \prec_{\mathcal{A}} y$ and $\mathcal{A}(x, \prec_{\mathcal{A}}\!\restriction_z)$. By (SC1) we have $\neg\mathcal{A}(y, \prec_{\mathcal{A}}\!\restriction_z)$, $z \not\prec_{\mathcal{A}} x$, and $x \in J_{\mathcal{A}}$. We get $\prec_{\mathcal{A}}\!\restriction_x \subset \prec_{\mathcal{A}}\!\restriction_z$ by Lemma 4.5 and thus $\neg\mathcal{A}(y, \prec_{\mathcal{A}}\!\restriction_x)$; hence $x \prec_{\mathcal{A}} y$. The converse follows by Lemma 4.3.1. ⊣

§**15. Acknowledgments.** The author is most grateful to Kentaro Sato for valuable and fruitful discussions on his [22] as well as various topics related to the present article. He also thanks Andrea Cantini, Volker Halbach, Leon Horsten, and the anonymous referee for very helpful suggestions and comments.

REFERENCES

[1] P. ACZEL, ***Non-Well-Founded Sets***, CSLI Publication, Stanford, CA, 1988.

[2] J. BARWISE, ***Admissible Sets and Structures***, Springer, Berlin, 1975.

[3] J. BARWISE, R. GANDY, and Y. MOSCHOVAKIS, *The next admissible set*, this JOURNAL, vol. 36 (1971), pp. 108–120.

[4] W. BUCHHOLZ, S. FEFERMAN, W. POHLERS, and W. SIEG, ***Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretic Studies***, Lecture Notes in Mathematics, vol. 897, Springer, Berlin, 1981.

[5] J. BURGESS, *Friedman and the axiomatization of Kripke's theory of truth*, ***Foundational Adventures: Essays in Honor of Harvey M. Friedman*** (N. Tennant, editor), College Publications, London, 2014, pp. 125–148.

[6] A. CANTINI, *Notes on formal theories of truth*. ***Zeitschrift für Mathematische Logik und Grundlagen der Mathematik***, vol. 35 (1989), pp. 97–130.

[7] ———, *A theory of formal truth arithmetically equivalent to ID1*, this JOURNAL, vol. 55 (1990), pp. 244–259.

[8] ———, ***Logical Frameworks for Truth and Abstraction***, Elsevier, Amsterdam, 1996.

[9] S. FEFERMAN, *Reflecting on incompleteness*, this JOURNAL, vol. 56 (1991), pp. 1–49.

[10] H. FIELD, *Deflating the conservativeness argument*. ***The Journal of Philosophy***, vol. 96 (1999), pp. 533–540.

[11] K. FUJIMOTO, *Relative truth definability of axiomatic theories of truth*. ***The Bulletin of Symbolic Logic***, vol. 16 (2010), pp. 305–344.

[12] ———, *Classes and truths in set theory*. ***Annals of Pure and Applied Logic***, vol. 163 (2012), pp. 1484–1523.

[13] ———, *Notes on some second-order systems of iterated inductive definitions and* Π_1^1*-comprehensions and relevant sybsystems of set theory*. ***Annals of Pure and Applied Logic***, vol. 166 (2015), pp. 409–463.

[14] ———, *Deflationism beyond arithmetic*. ***Synthese***, 2017, preprint, doi:10.1007/s11229-017-1495-8.

[15] V. HALBACH, *Truth and reduction*. ***Erkenntnis***, vol. 53 (2000), pp. 97–126.

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

[17] G. JÄGER and J. KRÄHENBÜHL, Σ_1^1 *choice in a theory of sets and classes*, ***Ways of Proof Theory*** (R. Schindler, editor), Ontos Verlag, Frankfurt, 2010, pp. 283–314.

[18] R. KAHLE, *Truth in applicative theories*. ***Studia Logica***, vol. 68 (2001), pp. 103–128.

[19] Y. Moschovakis, ***Elementary Induction on Abstract Structures***, Studies in Logic and the Foundation of Mathematics, no. 77, North Holland, Amsterdam, 1974.

[20] W. Pohlers, *Subsystems of set theory and second order number theory*, ***Handbook of Proof Theory*** (S. Buss, editor), Elsevier, Amsterdam, 1998, pp. 209–336.

[21] ———, ***Proof Theory***, Springer, Berlin, 2009.

[22] K. Sato, *Full and hat inductive definitions are equivalent in NBG*. ***Archive for Mathematical Logic***, vol. 54 (2015), pp. 75–112.

[23] S. G. Simpson, ***Subsystems of Second Order Arithmetic***, Cambridge University Press, Cambridge, 2009.

[24] C. Tapp, ***Eine direkte Einbettung von*** $\mathsf{KP}\omega$ ***in*** ID_1. Diploma thesis, Westfälische Wilhelms-Universität Münster, 1999.

SCHOOL OF MATHEMATICS AND DEPARTMENT OF PHILOSOPHY
UNIVERSITY OF BRISTOL
BRISTOL BS8 1TW, UK
E-mail: kentaro.fujimoto@bristol.ac.uk