Notre Dame Journal of Formal Logic Volume 36, Number 2, Spring 1995

Levels of Truth

ANDREA CANTINI

Abstract This paper is concerned with the interaction between formal semantics and the foundations of mathematics. We introduce a formal theory of truth, TLR, which extends the classical first order theory of pure combinators with a primitive truth predicate and a family of truth approximations, indexed by a directed partial ordering. TLR naturally works as a theory of partial classifications, in which type-free comprehension coexists with functional abstraction. TLR provides an inner model for a well known subsystem ATR₀ of second order arithmetic; indeed, TLR is proof-theoretically equivalent to Predicative Analysis.

1 Introduction It is well known that if we axiomatize the basic closure properties of fixed point models for partial self-referential truth à la Kripke over a given theory (say Peano Arithmetic) we obtain rather extensive systems which are appealing not only for formal semantics but also for the foundations of mathematics (cf. Feferman [14], Reinhardt [27]).

Nevertheless, systems based on self-referential truth, T in short, are far from being satisfactorily closed: they show a limited ability in reflecting negative information and hypothetical reasoning. For instance, the inference from $TA \rightarrow TB$ to $T(A \rightarrow B)$, which corresponds to the usual implication introduction rule, is generally unsound unless we have *the additional information that A is a proposition in the sense of T*, i.e., $TA \lor T \neg A$. (In this introduction we neglect details concerning Gödel numbering, and we simply write TA instead of T[A], where [A] is a suitable encoding of A).

In general, we have no chance to reduce negative external information $\neg TA$ to internal negative information $T\neg A$: *T* is essentially partial. Thus we are naturally confronted with the problem of designing formal frameworks which can reflect, at least to a certain extent, *negative semantic information* and hence can exhibit a higher degree of completeness.

Of course, a number of formal moves are conceivable here. In this paper we choose to develop a formal theory of (abstract) self-referential truth which is supple-

Received June 16, 1994; revised May 9, 1995

mented with *levels of truth*. The basic intuition, which is certainly not new, stems from the observation that, once we have fixed a semantic schema (here Kleene's strong three valued logic), the truth predicate *T* is *parametric*: *T* depends upon a set \mathcal{K}_0 including complete information about given primitive predicates, to be regarded as *the context T is about*. Thus *T* is properly $T(\mathcal{K}_0)$, for some \mathcal{K}_0 . Furthermore, by Tarski's theorem the context cannot include *complete information about T itself*. Hence, if we consider $T(\mathcal{K}_0)$ as completed or fully grasped, we shift from the con*text* \mathcal{K}_0 to a new one \mathcal{K}_1 , which also includes a complete description of $T(\mathcal{K}_0)$ as *primitive*. For instance, if *A* is any sentence such that $A \notin T(\mathcal{K}_0)$ (or $A \in T(\mathcal{K}_0)$), we must have $(\neg T_0 A) \in \mathcal{K}_1$ (or $(T_0 A) \in \mathcal{K}_1$, T_0 being the formal counterpart of $T(\mathcal{K}_0)$). We underline that we must add $(\neg T_0 A)$ to \mathcal{K}_1 and not simply $(\neg A)$: $(\neg A)$ would in general conceal its context dependence, and this might lead to paradoxes.

These considerations are rough, but they naturally suggest that the parametric dependence of truth ought to be made explicit by means of *levels*: the shift from $T(\mathcal{K}_0)$ to $T(\mathcal{K}_1)$ is seen as a step to a higher reflection stage and, formally, from truth of ground level T_0 to truth of higher level T_1 . On the other hand, the step from level 0 to level 1 can actually be understood as a general uniform method for generating new truth predicates from given ones. For the sake of generalization, we simply identify levels with ordinals, and we imagine a language in which, besides T, we also dispose of level dependent truth predicates T_i . Informally, we can sum up the fundamental tenets behind our theory TLR (= truth with levels and reflection) in three points.

- 1. If *i*, *j* are levels and $i \prec j$ (where \prec is the precedence order on levels), T_i is related to T_j in such a way that:
 - (a) whatever is declared true by T_i , is declared true by T_j , i.e., $\forall x(T_i x \rightarrow T_j x)$;
 - (b) T_i is decidable with respect to T_j , i.e., T_jT_iA or $T_j\neg T_iA$ (A arbitrary; we neglect formalization details).
- 2. Each local truth predicate T_i satisfies the closure principles of the general theory of partial truth à la Kripke and Feferman.
- 3. There is also a level-free truth predicate T which is conceived as the "limit" of the local truth predicates; in addition, we still assume that T itself has the self-referential abilities of the T_i s.

Principles (1)–(3) are formally implemented in the theory TLR of Sections 2–3 (actually we consider a more general system without number-theoretic induction up to Section 8). TLR is a first order extension of combinatory logic, expanded with a level-free truth predicate and a binary predicate T_i (truth of level *i*). The level ordering is assumed to be only partial, not well-founded and unbounded; but it also satisfies a nontrivial reflection principle, implying the second part of 3 above.

We underline that TLR has a built-in theory of total untyped operations, which takes care of predicate abstraction and self-referential constructions in a very uniform way. We also postulate an injection of levels into objects in order to codify sentences involving levels. This move puts important constraints for building models of TLR and it also requires non-trivial facts about admissible ordinals (projectibility; see Section 7).

Section 4 surveys elementary closure conditions for truth and truth predicates with levels, whereas Section 5 is concerned with the development in TLR of a theory of partial classifications and type-free abstraction in the sense of Feferman [13]. In particular, we can distinguish *i*-classes, i.e., predicates which are total relative to truth of level *i*, from classes, i.e., total predicates *tout court*; *i*-classes are nicely closed under forms of elementary comprehension and effective disjoint union.

Section 6 investigates the influence of the local structure on the closure properties of the level-free statements. $CL := \{x : x \text{ class}\}$ splits into a directed family $\{CL_i : i \text{ level}\}$, where each $CL_i := \{x : x \text{ class} \text{ of level } i\}$ is itself a class at any higher level j > i. As a consequence, classes are closed under an analog of Weyl's *Iterationsprinzip* (see Weyl [31]), a transfinite recursion principle along *CL*-well-founded linear orderings. We can also recover in the present context a satisfactory notion of *universe* (see Feferman [12], Jäger [21], Martin-Löf [24], Marzetta [25]).

Section 7 describes a model C_i for a strengthening TLR⁺ of TLR, which also contains linearity, well-foundedness of level ordering, and number-theoretic axioms (induction schema included). The model is built up by means of a suitable iterated inductive definition along the first recursively inaccessible ordinal; the step from truth of a given level to truth of higher level essentially corresponds to the hyperjump operation of generalized recursion theory (see Hinman [20]). In the final section 8 we consider the relation with classical subsystems of second order arithmetic: we can produce a model to Friedman's subsystem ATR₀ (Friedman et al. [16]) within TLR (with number-theoretic induction for classes of arbitrary level). Indeed, this interpretation yields a proof-theoretic lower bound on TLR; the lower bound is sharp and TLR is a strong version of Predicative Analysis. In this connection it might be interesting to settle the precise relation between TLR and the theory of iterated admissibility without foundation, which is also known to have the proof-theoretic strength of Predicative Analysis by Jäger [21].

We conclude this section by briefly discussing the relation of the present work to the literature. First of all, the philosophical paper of Burge [4] already contains an interesting approach to semantical paradoxes based on the indexical nature of truth and on the notion of level (see also Parsons [26]; related ideas are independently sketched by Gaifman [17]). In this respect, we might consider TLR as a sort of axiomatization for (a version of) Burge's proposal, in which the problem of extending the construction of truth predicates into the transfinite is explicitly tackled.

The idea of internalizing negative information by means of a reflective process indexed by levels, is already present in earlier work about "type-free logic" (e.g., Schütte [29] and Fitch [15]). In this respect, the paper of Lorenzen and Myhill [23] deserves a special mention (in particular, cf. pp. 47–49), as well as for its applications to foundational issues and to recursion theory.

Subsequently, similar ideas emerge anew in connection with the problem of expanding lambda calculus models with truth by Scott [30]; an earlier version of TLR (outlined in Cantini [6]) was directly inspired by an attempt to investigate Scott's model for a hierarchy of self-referential truth predicates. In Cantini [8] we defined a theory of abstraction based on truth of arbitrary finite levels and a stronger semantical schema.

Recently Aczel, Carlisle, and Mendler [1] introduced a hierarchy of propositions

and truth predicates as a basis for logical theories of constructions to be used in theoretical computer science.

2 *Truth with levels: preliminaries* In this section we describe the formal language of reflective truth with levels, and we summarize the basic facts of combinatory logic needed below; finally we define a suitable Gödel numbering.

The language \mathcal{L}_V includes:

- 1. a denumerable list of individual variables x_1, x_2, x_3, \ldots ;
- 2. a denumerable list of variables for levels i_0, i_1, \ldots (in short L-variables);
- 3. the individual constants *K*, *S* (combinators); the binary function symbol *Ap* (functional application); the unary function symbol *LT* (level injection);
- 4. four binary predicates = (object equality), ≤ (level ordering), =_l (level equality), V (local truth); two unary predicates T (for truth) and N (for natural numbers);
- 5. the logical constants \neg , \land , \forall .
- x, y, u, v, w, z are used as syntactical variables for object variables x_1, x_2, x_3 , etc.

2.1 L(evel)-terms, terms, formulas of \mathcal{L}_V and acceptable formulas

- 1. L-variables are exactly the L-terms (i, j, k metavariables for L-terms);
- 2. the set of \mathcal{L}_V -terms is the least collection which is closed under the following clauses: individual variables and constants are terms; if *j* is an L-term, LT(j) is a term; if *t*, *s* are terms, Ap(t, s) is a term;
- 3. the set of \mathcal{L}_V -formulas is the smallest collection closed under the following clauses: if *j* and *i* are L-terms, $i \leq j$ and $i =_l j$ are atoms (and hence formulas); if *t*, *s* are terms and *i* is an L-term, Nt, t = s, Tt and V(i, t) are atoms (and hence formulas); if *A*, *B* are formulas, $\neg A$, $A \wedge B$ are formulas; if *A* is a formula, *x* an individual variable and *j* is an L-variable, then $\forall xA$ and $\forall jA$ are formulas (where *x*, *j* occur bound).

Atoms of the form t = s and Nt are called e-atoms (e = elementary).

4. The collection Δ^+ of acceptable formulas of \mathcal{L}_V is the smallest collection which includes the atoms Tt, T_it , t = s, Nt and is closed under negation, conjunction and universal quantification on object variables.

The intended meaning of V(i, t) is that "t is true at level i"; we write $T_i t$ for V(i, t); Tt := "t is true"; $i \le j$ ($i =_l j$) := "the level i is less than or equal (equal tout court) to the level j." If i, j are L-terms, i = j is a shortening for $i =_l j$; we also write $i \prec j$ for $(\neg i = j) \land (i \le j)$. The intended meaning of Nt is "t is a natural number"; however, N will not play any active role until Section 8 in the comparison with a subsystem of second order arithmetic.

2.2 The fragments \mathcal{L} and \mathcal{L}_{op} of \mathcal{L}_{V}

- 1. \mathcal{L} is obtained from \mathcal{L}_V by omitting L-variables, $LT, V, \leq =_l$;
- 2. \mathcal{L}_{op} is obtained from \mathcal{L} by omitting the predicate T.

We write (ts) for Ap(t, s) and outer parentheses are omitted, and the missing ones are restored by associating to the left. For clarity, we sometimes use f, g, h for variables occurring in operand position (e.g., fx instead of yx).

The systems we consider in this paper include combinatory logic as a ground theory of (untyped) operations.

2.3 The system TO (= total operations)

TO is the formal theory in the language \mathcal{L}_V which contains:

- 1. classical (two-sorted) predicate logic with identity in the language \mathcal{L}_V ;
- 2. the combinatory axioms,
 - C1 $K \neq S$; C2 $\forall x \forall y \forall z (Kxy = x \land Sxyz = xz(yz))$.

We inductively introduce λ -abstraction according to the standard definitions of combinatory logic, i.e., $\lambda x.x = SKK$; $\lambda x.t = Kt$, provided x is not free in t; $\lambda x.ts = S(\lambda x.t)(\lambda x.s)$. As with the quantifiers, we usually insert a dot between λx and its body, for the sake of readability; occasionally, we use dots as separating symbols in place of parentheses. If E is any expression (term or formula), E[x := t] denotes the result of replacing x with t in E. λ -abstraction satisfies β -conversion and the fixed point theorem provably in **TO**:

Lemma 2.1

- 1. **TO** proves $(\lambda x.t)u = t[x := u]$.
- 2. We can define a closed term Y such that **TO** proves $\forall x(Yx = x(Yx))$.

Proviso: in (1) u is free for x in t.

A pairing operation with projections can be defined in \mathcal{L}_{op} ; e.g., we can choose:

PAIR :=
$$\lambda xyu.uxy$$
; LEFT := $\lambda x.xK$; RIGHT := $\lambda x.x(KI)$,

where $I := \lambda x.x$.

Henceforth we adopt the familiar notations: $\langle t, s \rangle := \text{PAIR } ts$; $(t)_1 = \text{LEFT } t$ and $(t)_2 := \text{RIGHT } t$. Then by Lemma 2.1 we have the following.

Lemma 2.2 TO $\vdash \forall x_1 \forall x_2((\langle x_1, x_2 \rangle)_i = x_i) \text{ (where } i = 1, 2).$

We can obviously define a coding of *n*-tuples; in particular, we choose $\langle t_1, t_2, t_3 \rangle$: = $\langle t_1, \langle t_2, t_3 \rangle \rangle$. The pairing system based on PAIR is also used to represent standard numerals in **TO**. NUM, the collection of numerals, is the least set *X* of closed \mathcal{L}_{op} terms such that $\overline{0} := I \in X$; if $\overline{n} \in X$, then $\overline{n+1} := (\text{PAIR } \overline{n}K) \in X$. Successor and predecessor on NUM are then defined by the terms $SUC := \lambda x.\text{PAIR } xK$, PRED := LEFT. By Lemma 2.1, we can also find a term *D* representing definition by cases on numbers (for details see Barendregt [2]). In fact, one has, for each *n*, *m*, the following.

Lemma 2.3

$$\mathbf{TO} \vdash \neg (\overline{n+1}) = \overline{0} \land ((\overline{n+1}) = (\overline{m+1}) \to \overline{n} = \overline{m})) \land (\operatorname{PRED}(\overline{n+1}) = \overline{0});$$
$$\mathbf{TO} \vdash D\overline{n} \,\overline{n}xy = x \land D\overline{n} \,\overline{m}xy = y \,(for \, n \neq m).$$

In the following, we ambiguously use n, m, k as symbols ranging both on natural numbers and the corresponding numerals of NUM.

Definition 2.4 (Terms representing acceptable formulas)

1. Terms representing logical operators and predicates of \mathcal{L}_V :

$$\begin{split} \text{ID} &:= \lambda x y. \langle \overline{1}, x, y \rangle; \quad TR := \lambda x. \langle \overline{2}, x \rangle; \quad \text{NAT} := \lambda x. \langle \overline{3}, x \rangle; \\ \text{NEG} &:= \lambda x. \langle \overline{4}, x \rangle; \quad \text{AND} := \lambda x y. \langle \overline{5}, x, y \rangle \rangle; \quad \text{ALL} := \lambda x. \langle \overline{6}, x \rangle; \\ TR_i &:= \lambda x. \langle \overline{7}, LT(i), x \rangle. \end{split}$$

2. We then define the map $A \to [A]$ by induction on the notion of acceptable \mathcal{L}_{V} -formula:

$$[t = s] := (ID t)s; \quad [Ns] := NAT s; [Tt] := TRt;$$

$$[T_it] := TR_it; \quad [\neg A] := NEG[A];$$

$$[A \land B] := AND[A][B]; \quad [\forall xA] := ALL(\lambda x[A]).$$

Observe that [A] has exactly the same free variables as A, and it commutes with substitution ([A][x := t] = [A[x := t]]). Moreover ID, TR, TR_i, NEG, AND, ALL denote distinct objects and enjoy unique readability and independence conditions, at least if the following projectibility axiom **PROJ** is assumed.

Axiom 2.5 $\forall i \forall j (LT(i) = LT(j) \rightarrow i = j).$

Lemma 2.6 (Independence of combinators representing logical constructors)

- 1. If L_1 , $L_2 \in \text{LOG}_1 := \{\text{NAT, NEG, } TR_i, TR, \text{ALL}\}$, then $\mathbf{TO} + \mathbf{PROJ} \vdash L_1 x = L_2 y \rightarrow .L_1 = L_2 \land x = y;$
- 2. *if* $G_1, G_2 \in \text{LOG}_2 := \{\text{ID, AND}\}, then$ **TO** $\vdash G_1 xy = G_2 x' y' \rightarrow .G_2 = G_2 \land x = x' \land y = y';$
- 3. *if* $L_1 \in \text{LOG}_1$, $L_2 \in \text{LOG}_2$, *then* **TO** $\vdash \neg L_1 x = L_2 yz$; *if* L_1 , L_2 *are distinct elements of* $\text{LOG}_1 \cup \text{LOG}_2$, *then* **TO** $\vdash \neg L_1 = L_2$.

Proof: By pairing axioms, β -conversion and the projection axiom **PROJ** in the case where $L_1 = L_2 = TR_i$.

Remark 2.7 The choice of [-] is largely a matter of taste, as soon as the conditions of Lemma 2.6 are met. As a viable alternative, one could assume new constants ID, NAT, *TR*, *TR*, *AND*, NEG, ALL with axioms corresponding to the conditions of Lemma 2.6, or even assume [-] as a primitive term constructor.

2.4 *Conventions* We henceforth adhere to the following conventions:

- 1. *TA* is a shorthand for T[A];
- 2. To enhance readability, we use \neg , \land , \forall , etc., and infix notation instead of the terms NEG, AND, ALL, etc. Thus $t \land s$, $\forall x.t$, $\neg t$, etc. stand for the terms (AND *t*)*s*, ALL($\lambda x.t$), NEG *t* (in the given order).
- 3. We also adopt the obvious shorthands $\neg \neg t := \neg(\neg t)$, and $t \lor s, t \to s$ instead of (respectively) $\neg(\neg t \land \neg s)$, $(\neg t \lor s)$. As to the existential operator, we let

$$\exists (f) := \neg (\forall (\lambda u. (\neg (fu)))) \text{ and } \exists x.t := \exists (\lambda x.t))$$

3 *Truth with levels: axioms* The truth principles are grouped into local truth axioms, level and connection axioms, and a reflection axiom. Number-theoretic axioms will be introduced later in order to investigate the relation with subsystems of analysis.

3.1 Local truth axioms

1. $T_i A \leftrightarrow A$, if A := (x = y), Nx, $(\neg x = y)$, $\neg Nx$; 2. $T_i x \rightarrow T_i T_i x$; $T_i \neg x \rightarrow T_i \neg T_i x$; 3. $T_i \neg \neg x \leftrightarrow T_i x$; 4. $T_i (x \land y) \leftrightarrow T_i x \land T_i y$; $T_i \neg (x \land y) \leftrightarrow T_i \neg x \lor T_i \neg y$; 5. $T_i (\forall f) \leftrightarrow \forall x T_i (fx)$; $T_i \neg (\forall f) \leftrightarrow \exists x T_i \neg (fx)$; 6. $\neg (T_i x \land T_i \neg x)$ (Local consistency).

3.2 *Level axioms* These include standard equality axioms for level equality $(=_l)$ and state that \leq is a directed unbounded partial order.

- 1. $\forall i \forall j \forall k ((i \leq i) \land (i \leq j \land j \leq k \rightarrow i \leq k) \land (i \leq j \land j \leq i \rightarrow i = j)).$
- 2. $\forall i \forall j \exists k (i \prec k \land j \prec k)$.

3.3 *Connection axioms* These are the crucial principles of the theory, relating truth predicates of different level.

- 1. Limit: $Tx \rightarrow \exists i. T_i x; T_i x \rightarrow Tx;$
- 2. Persistence: $i \leq j \wedge T_i x \rightarrow T_j x$;
- 3. Localization: $T_iTx \leftrightarrow T_ix$; $T_i \neg Tx \leftrightarrow T_i \neg x$;
- 4. Potential Completeness: $i \prec j \rightarrow (T_j T_i x \lor T_j \neg T_i x)$;
- 5. Positive Soundness: $T_j T_i x \rightarrow i \leq j \wedge T_i x$;
- 6. Negative Soundness: $T_i \neg T_i x \rightarrow (i = j \land T_i \neg x) \lor (i \prec j \land \neg T_i x)$.

3.4 The Reflection principle (REF)

$$\forall i \forall y \forall z \{\forall x \exists j (x\eta_i y \to x\eta_j z) \to \exists k \forall x \exists j (j \leq k \land (x\eta_i y \to x\eta_j z))\}.$$

Definition 3.1

- TL⁻ is the theory based on two sorted classical predicate logic with equality axioms (for the two sorts of objects and levels), which includes the system TO of Section 2.3, the projectibility axiom PROJ of Axiom 2.5 and the axioms of groups from Sections 3.1–3.3.
- 2. TLR⁻ is TL⁻ plus REF.

N.B. The - sign means that no assumption is made on the predicate N.

A few words of comment. By the principles of group from Section 3.1, truth of any level *i* satisfies an abstract version of the KF-axioms for reflective truth ('KF' = Kripke-Feferman; see [14]). The abstract character of truth predicates is to be found in the fact that *they are not, like the usual formalized truth predicates, attributes of (codes of) sentences*, but, more generally, predicates of objects in a given combinatory

algebra. As to Section 3.3, potential completeness ensures that negative information about any level *i* becomes internal at higher levels, whereas limit and localization axioms imply that *global* truth statements always reduce to *local* truth statements (of sufficiently high level). Finally, by persistence and soundness no information is lost at later levels, and later levels do not conflict with earlier ones, even on negative information. The reflection principle says that there are enough levels for *T*, in order to internalize universal statements about objects, and it exactly implies that *T* itself is a model of KF-axioms (see Proposition 4.3 below and the characterization of the inclusion relation in Lemma 5.5). The level axioms are presented in a general form and do not assume linearity or well-foundedness; however the recursion-theoretic model will interpret \prec as the standard ordering on a suitable segment of countable ordinals.

4 Elementary consequences We consider elementary closure properties of T_i and T.

Definition 4.1

- 1. Let *i* be any L-variable: the *i*-transform of $A \in \mathcal{L}$ is the \mathcal{L}_V -formula A_i , which results from A by substituting each occurrence of T with T_i (e.g., $(\forall xT(ax))_i = \forall xT_i(ax), (TTt)_i = T_iTt)$.
- 2. An \mathcal{L}_V -formula *A* is *T*-positive iff *A* belongs to the least collection which contains expressions of the form t = s, $\neg t = s$, Nt, $\neg Nt$, T_it , $\neg T_it$, Tt and is closed under conjunction, disjunction, and quantifiers (on either sort).
- A is a *k*-formula iff A belongs to the least collection of formulas which is closed under ∧, ¬, universal object quantification, and contains atoms of the form t = s, Nt, T_kt.

Proposition 4.2

- *1*. Global consistency: $\neg(Tx \land T\neg x)$;
- 2. $TL^- \vdash T_i x \leftrightarrow T_i T_i x$ and $T_i \neg x \leftrightarrow T_i \neg T_i x$;
- 3. Δ^+ -soundness: $TL^- \vdash T_i A \rightarrow A(A \in \Delta^+)$;
- 4. $TL^- \vdash T_i A \rightarrow A_i \ (A \in \Delta^+);$
- 5. $TL^- \vdash i \leq j \wedge T_i A \rightarrow T_i A \ (A \in \Delta^+);$
- 6. If A is a k-formula,

$$TL^{-} \vdash k \prec j \rightarrow ((T_{i}A \leftrightarrow A) \land (T_{i}A \lor T_{i}\neg A)),$$

hence:

$$TL^{-} \vdash (TA \leftrightarrow A) \land (TA \lor T \neg A).$$

Proof: (1) If Tx and $T\neg x$ are assumed, then by limit axiom T_ix and $T_k\neg x$, for some i, k; hence there exists by Section 3.2 some $j \succ i, k$, such that by persistence T_jx and $T_j\neg x$, against local consistency.

(2) By local truth axioms 3.1.2, positive and negative soundness.

(3) Induction on *A*. If *A* is an e-atom or has the form T_jt , $\neg T_jt$, we apply Axiom 3.1.1, positive and negative soundness and local consistency 3.1.6. Let $A := \neg Tt$ and assume $T_i \neg Tt$; then $T_i \neg t$ by localization, hence $T \neg t$ by limit and $\neg Tt$ by (1) above. The remaining cases are straightforward by IH and local truth axioms.

(4) Induction on A, using localization if $A = T_i t$.

(5) Apply persistence axiom.

(6) Potential completeness and Δ^+ -soundness yield the first statement, which, in turn, implies the second one by limit, persistence, and unboundedness axioms.

Proposition 4.3

1. TL⁻ proves:

$$TA \leftrightarrow A, \text{ if } A := x = y, Nx, \neg x = y, \neg Nx;$$

$$TTx \leftrightarrow Tx; \quad T\neg Tx \leftrightarrow T\neg x;$$

$$T\neg \neg x \leftrightarrow Tx;$$

$$T(x \wedge y) \leftrightarrow Tx \wedge Ty; \quad T\neg (x \wedge y) \leftrightarrow T\neg x \vee T\neg y;$$

$$T(\forall f) \rightarrow \forall xT(fx); \quad \exists xT\neg (fx) \leftrightarrow T\neg (\forall f);$$

- 2. TLR⁻ $\vdash \forall x \exists i T_i(fx) \rightarrow \exists k \forall x T_k(fx)$ (positive reflection);
- 3. TLR⁻ $\vdash \forall x T(fx) \rightarrow T(\forall f);$
- 4. if A is acceptable and A is T-positive,

$$\mathrm{TLR}^- \vdash A \Leftrightarrow \exists i. T_i A \Leftrightarrow TA \Leftrightarrow \exists iA_i.$$

Proof: (1) By limit, localization, and local truth axioms, together with the directedness of the level ordering.

- (2) Apply reflection with $y := \{u : u = u\}$ and persistence.
- (3) Apply limit, positive reflection, and local truth axiom 3.1.5.

(4) Let us consider the first equivalence. From right to left, it follows from Δ^+ soundness (Proposition 4.2). As to the reverse direction, we argue by induction on *A*. If $A := \neg T_j t$, choose $k \succ j$ by unboundedness of \prec : then $T_k \neg T_j t$ by potential completeness, negative soundness. If $A := \forall x B$, we use IH, positive reflection and the local truth axiom for \forall . The other cases are easy and left to the reader. The second equivalence is just a restatement of the limit axiom. As to the third equivalence, $TA \rightarrow \exists i T_i A \rightarrow \exists i. A_i$ (use Proposition 4.2.4). $A_i \rightarrow TA$ is inductively checked (Proposition 4.3.3 above being used in the case $A := \forall x B$).

By Proposition 4.3, T will satisfy the same basic axioms as the T_i s; there is a "harmony" between global and local structure of truth.

As to the Liar paradox, in the present framework we can distinguish a "local" version ("I am not true at level *i*"), which is formally decidable at any higher level and hence true, from a *T*-undecidable "global" version ("I am not true").

Proposition 4.4

1. Let L(i) be the term such that $TL^- \vdash L(i) = [\neg T_i L(i)]$. Then

$$TL^{-} \vdash i \prec j \rightarrow (T_{j}L(i) \land \neg T_{i}L(i) \land \neg T_{i}(\neg L(i)));$$

$$TL^{-} \vdash \forall i.T(L(i)).$$

2. Let *L* be the term such that $TL^- \vdash L = [\neg TL]$. Then $TL^- \vdash \neg TL \land \neg T \neg L$.

Proof: The existence of L and L(i) is ensured by Lemma 2.1.

(1) $\neg T_iL(i) \land \neg T_i(\neg L(i))$ follows by Axiom 3.1.2, local consistency, and Axiom 3.1.3. On the other hand if $i \prec j$, we have by Proposition 4.2.6 $T_j \neg T_iL(i)$ and hence $T_jL(i)$ by identity logic. The second statement follows by unboundedness, limit and persistence.

(2) Apply Propositions 4.3.1 and 4.2.1.

If we call *i*-proposition any object x such that $T_i x \vee T_i \neg x$, Proposition 4.4 implies the existence of *j*-propositions which are not *i*-propositions, for any $j \succ i$.

5 Truth with levels and abstraction $In TL^-$ and TLR^- we can develop a theory of partial classifications; it is closely related to nonextensional systems based on type-free comprehension.

Definition 5.1

1.

- $F_i t := T_i \neg t; \quad Ft := T \neg t;$ $t\eta_i s := T_i(st) \text{ and } t\bar{\eta}_i s := F_i(st);$ $t\eta s := T(st); \quad t\bar{\eta} s := F(st);$ $Cl_i(t) := \forall x(x\eta_i t \lor x\bar{\eta}_i t) \text{ (} t \text{ is an } i\text{-class}\text{)};$ $Cl(t) := \forall x(x\eta t \lor x\bar{\eta} t) \text{ (} t \text{ is a class}\text{)}.$
- 2. The abstraction operator: if A is acceptable $\{x : A\} := \lambda x.[A]$. 3. $CL := \{x : Cl(x)\}; \quad CL_i := \{x : Cl_i(x)\};$
- 3.

$$R := \{x : \neg x\eta x\}; \quad R(i) := \{x : \neg x\eta_i x\};$$
$$x \equiv_i y := (T_i x \leftrightarrow T_i y) \land (F_i x \leftrightarrow F_i y);$$
$$x \equiv y := (T x \leftrightarrow T y) \land (F x \leftrightarrow F y).$$

4. A formula *B* is *elementary* in the list x_1, \ldots, x_n iff *B* is built up from e-atoms, *T*-atoms of the form $t\eta x_i$ (where $1 \le i \le n$), by means of $\neg, \land, \forall y, (y \notin \{x_1, \ldots, x_n\})$.

Lemma 5.2

- 1. If A is a T-positive *L*-formula, $TL^- \vdash A_k \leftrightarrow T_kA$;
- 2. If A(u, x) is an *L*-formula elementary in x,

$$TL^{-} \vdash Cl_k(x) \rightarrow T_kA(u, x) \lor F_kA(u, x);$$

3. If A(u, x) is an L-formula elementary in x,

$$\mathrm{TL}^- \vdash Cl_k(x) \to A(u, x) \leftrightarrow A_k(u, x) \leftrightarrow T_k A(u, x).$$

Proof: (1–2) Induction on *A*, applying by Proposition 4.2.2 and local truth axioms.

(3) Assume that x is a k-class. The second equivalence is a consequence of (2) and the first equivalence with Proposition 4.2.3. Thus we verify only the first equivalence by induction on A. If A is an atom different from $u\eta x$, the conclusion is trivial. If $A(u, x) := u\eta x$, $u\eta_k x$ implies $u\eta x$ by the limit axiom. In the opposite direction, we get a contradiction from $u\eta x$ and $\neg u\eta_k x$ (apply $Cl_k(x)$, persistence, limit, unboundedness and local consistency). If A(u, x) is a negation, a conjunction or a universal quantification, we simply apply IH.

Suitable forms of untyped comprehension hold provably in TL⁻; it also follows that the notion of *i*-class determines a class at any level j > i, whereas the collection of *i*-classes form a strictly increasing chain. As to the (analogue of the) Russell sentence relativized to level *i*, it becomes true at strictly higher levels.

Proposition 5.3

1. The extended abstraction schema for acceptable formulas: if $A \in \Delta^+$,

$$TL^{-} \vdash \forall u(u\eta\{x : A(x)\} \equiv A[x := u]);$$

2. The local abstraction schema for acceptable formulas: if $A \in \Delta^+$,

 $TL^{-} \vdash \forall i \forall u(u\eta\{x : A\} \equiv_{i} A[x := u]);$

- 3. *if* A(x) *is a j-formula,* $j \prec i \rightarrow \forall u(u\eta_i\{x : A\} \leftrightarrow A[x := u])$; (*u free for x in* A *in* (1)–(3) *above*); *hence:* $TL^- \vdash \forall u(u\eta\{x : A\} \leftrightarrow A[x := u])$;
- 4. TL⁻ $\vdash \forall i.\neg Cl_i(R);$
- 5. $\mathrm{TL}^- \vdash \forall i(i \succ j \rightarrow Cl_i(R(j)) \land \neg Cl_j(R(j)) \land R(j)\eta_i R(j));$
- 6. $\mathrm{TL}^- \vdash i \prec k \rightarrow CL_i \ \eta_k \ CL_k \land CL_i \subset CL_k.$

Proof: (1) By β -conversion and Proposition 4.3.1.

- (2) Immediate by β -conversion and localization axioms.
- (3) By (2) and Proposition 4.2.6.
- (4) By localization and local consistency.

(5) Let $i \succ j$. As to the first conjunct, R(j) is defined by a *j*-formula and hence we apply Proposition 4.2.6 and local abstraction; the second conjunct is Russell's paradox for level *j*; the third conjunct follows from the second one with (3).

(6) Assume $k \succ i$: then $CL_i \subset CL_k$ by persistence and (5). As to $CL_i\eta_k CL_k$, apply local abstraction and Proposition 4.2.6.

Proposition 5.4

1. Closure of CL_k under elementary comprehension:

 $TL^{-} \vdash Cl_{k}(x) \rightarrow (Cl_{k}(\{u : A(u, x)\}) \land \forall v(v\eta\{u : A(u, x)\} \leftrightarrow A(v, x))),$

where A is an \mathcal{L} -formula, elementary in x.

2. Closure of CL_k under join: define $\Sigma(x, f) := \{\langle v, w \rangle : v\eta x \land w\eta(fv)\}$. Then

$$TL^{-} \vdash \forall u(u\eta\Sigma(x, f) \leftrightarrow \exists v \exists w(u = \langle v, w \rangle \land v\eta x \land w\eta(fv)))$$
$$TL^{-} \vdash (Cl_k(x) \land f : x \to CL_k) \to Cl_k(\Sigma(x, f)).$$

(where $f : a \to b := \forall x(x\eta a \to (fx)\eta b)$).

Proof: (1) If x is a k-class, so is $\{u : A(u, x)\}$ by Lemma 5.2.2. The second equivalence is an immediate consequence of Proposition 5.3.1 and Lemma 5.2.3.

(2) Ad (1): apply Proposition 5.3.1 and Proposition 4.3.4, observing that the reflection axiom is not necessary if no universal quantifier is present, and hence we can work in TL⁻. Ad (2): let f be a family of classes indexed by the class

x and $\neg u\eta_k \Sigma(x, f)$. By Proposition 5.3.2 and local truth axioms for T_k , we get $\neg A_k(u, x, f)$, where

$$A(u, x, f) := \exists v \exists w (u = \langle v, w \rangle \land v \eta_k x \land w \eta_k (fv)).$$

If $\neg u = \langle v, w \rangle$ or $\neg v\eta_k x$ is assumed, $T_k \neg A(u, x, f)$ follows by $Cl_k(x)$ and the local truth axioms for T_k ; else we can assume $u = \langle v, w \rangle$ and $v\eta_k x$, which implies $Cl_k(fv)$ by assumption on f, x and hence $w\bar{\eta}_k(fv)$. The axioms for T_k again imply $T_k \neg A(u, x, f)$, whence by abstraction $u\bar{\eta}_k \Sigma(x, f)$.

Elementary comprehension and join ensure that the notion of k-class is nicely closed (e.g., k-classes are closed under boolean operation, generalized products over families of k-classes indexed by a k-class). Elementary comprehension and join were introduced by Feferman [11].

Lemma 5.5

- 1. If A(x) is T-positive and acceptable, $TLR^- \vdash \forall u(u\eta\{x : A\} \leftrightarrow A[x := u]);$
- 2. TLR⁻ \vdash *Cl*(*a*) $\Leftrightarrow \exists iCl_i(a);$
- 3. *let* $a \subseteq b := \forall x(x\eta a \to x\eta b)$; *then* $\text{TLR}^- \vdash a \subseteq b \leftrightarrow \forall i \exists k \forall x(x\eta_i a \to x\eta_k b)$;
- 4. a class of classes is always an i-class, for some level i:

 $\mathrm{TLR}^- \vdash Cl(a) \land a \subseteq CL \to \exists i.a \subseteq CL_i.$

Proof: (1) By Propositions 4.3.4 and 5.3.1.

(2) Apply the limit axiom from right to left. The reverse direction is a consequence of Proposition 4.3.4, as the formula defining Cl is acceptable and T-positive.

(3) \Rightarrow : by limit and reflection; the converse is trivial.

(4) Assume that a is a class of classes. Then by (2), a is an *i*-class and

$$a \subseteq CL \implies \forall j \exists k \forall x (x\eta_j a \to x\eta_k CL) \text{ by } (3);$$

$$\Rightarrow \forall j \exists k \forall x (x\eta_j a \to Cl_k(x)), \text{ by localization, local abstraction;}$$

$$\Rightarrow \exists k \forall x (x\eta_i a \to Cl_k(x)) \text{ by logic;}$$

$$\Rightarrow a \subseteq CL_k \text{ for some } k.$$

as $x\eta a \leftrightarrow x\eta_i a$, by assumption on a and global consistency Proposition 4.2.1.

6 Existence of universes and the extended Weyl iteration principle The level structure affects the set of level-free provable statements; the theme is illustrated by two significant statements.

The first principle says that, if we are concerned with logical constructions depending on *classes* as initial data, we can always work within a nicely closed *universe*, which is itself *a class of classes* and to which the initial data belong. We make the idea precise with a lemma and a definition.

Now let $z =_e w := \forall u(u\eta z \leftrightarrow u\eta w)$ (= η -extensional equality). Then we have the following.

Lemma 6.1 There is an *L*-formula Elemclos(y) such that, for every *L*-formula $A(x, u_1, ..., u_n)$ elementary in $u_1, ..., u_n$, TLR⁻ proves

Elemclos(y) $\rightarrow \forall u_1 \dots \forall u_n (u_1 \eta y \land \dots \land u_n \eta y \rightarrow$ $\rightarrow \exists z (z\eta y \land z =_e \{x : A(x, u_1 \dots u_n)\})).$

The lemma is an intensional version of the well known class theorem of Gödel-Bernays set theory; one can show that classes defined by elementary conditions can be generated by a finite number of operations from a finite stock of initial classes. Thus there is a finite number of axioms characterizing the closure of y under elementary comprehension; these axioms are collected into a single formula Elemclos(y), which can be read as "y is elementarily closed." In essence, this result is already in Gordeev [19], pp. 66–67, though in a constructive framework; so we omit the proof (a direct verification relies on the closure conditions of classes, established by Proposition 5.4). Incidentally, we remark that a term witnessing z in (*) can be effectively found from the given formula A.

Definition 6.2

1.
$$y \models J := \forall f \forall c (c\eta y \land f : c \rightarrow y) \rightarrow (\Sigma(c, f)\eta y \land \forall u (u\eta \Sigma(c, f) \leftrightarrow \exists v \exists w (u = \langle v, w \rangle \land v\eta c \land w\eta (fv)))));$$

 $\Sigma(c, f)$ is the term of Proposition 5.4.1 and $y \models J$ states that the join principle of Proposition 5.4.2 holds relativized to *y*.

2. *y* is a universe of classes iff *y* is an elementarily closed class of classes, which is also closed under join; in symbols: Univ(*y*) := $Cl(y) \land y \subseteq CL \land y \models J \land$ Elemclos(*y*), where Elemclos(*y*) is the *L*-formula given by Lemma 6.1.

Theorem 6.3

- *1*. TL⁻ $\vdash \forall k$.Univ(*CL*_k).
- 2. $TLR^- \vdash \forall y(Univ(y) \rightarrow \exists k(y \subseteq CL_k))$.

Proof: (1) That CL_k is closed under join and elementary comprehension already follows from Proposition 5.4; $CL_k \subseteq CL$ and $CL_k\eta CL$ are consequences of the limit axioms of Section 3.3 and Proposition 5.3.6.

(2) immediate by Lemma 5.5.4.

Corollary 6.4 Let $\text{LIM} := \forall x (Cl(x) \rightarrow \exists y (Univ(y) \land x\eta y)); then TLR^- \vdash \text{LIM}.$

Proof: If x is a class, x is already a k-class (Lemma 5.5.2), for some k and CL_k is a universe by the theorem.

Remark 6.5

- 1. LIM is investigated by [12], [21], and [25].
- 2. Each CL_k is closed under the basic type constructors of Martin-Löf's type theory, and Martin-Löf's intuitionistic type theory with arbitrarily many finite universes without *W*-types can be interpreted in the theory TLR.

The second principle we deal with concerns transfinite recursion over well-orderings. It is a priori unclear how to render the notion of well-ordering in the present context: shall we quantify over classes or arbitrary possibly partial predicates? We observe that the two alternatives yield radically different notions and that the sharpest notion is obtained by quantifying over classes (this point can be clarified with the help of the proof-theoretic analysis; details are given in Cantini [9]).

Definition 6.6

- 1. If *w* is used for encoding a binary relation, we use the infix notation $x \prec_w y$ in place of $\langle x, y \rangle \eta w$. Field (\prec_w) stands for the term $\{x : \exists z (x \prec_w z \lor z \prec_w x)\}$ representing the field of \prec_w , whereas the *x*-segment of \prec_w determined by *x* is defined by the term $\prec_w [x := \{u : u \prec_w x\})$.
- 2. $LO(\prec_w)$ states that \prec_w is a linear ordering:

$$LO(\prec_w) := \forall x \forall y \forall z (\neg (x \prec_w x) \land (x \prec_w y \land y \prec_w z \to x \prec_w z) \land Conn(\prec_w)),$$

where $\operatorname{Conn}(\prec_w) := \forall x \forall y (x\eta \operatorname{Field}(\prec_w) \land y\eta \operatorname{Field}(\prec_w) \rightarrow x \prec_w y \lor x = y \lor y \prec_w x)).$

3. Progr $(b, \prec_w) := (\forall x \eta \operatorname{Field}(\prec_w)) \ (\forall y \prec_w x. y \eta b \to x \eta b).$ Progr (\prec_w, b) is to be read "*b* is progressive (relative to \prec_w)." We also define:

$$TI(\prec_w, b) := \operatorname{Progr}(\prec_w, b) \to \operatorname{Field}(\prec_w) \subseteq b.$$

- 4. A linear ordering \prec_w is called a *pseudo-well-ordering* (in symbols $PWO(\prec_w)$, and, for short, \prec_w is a p.w.o.) iff $\forall b(Cl(b) \rightarrow TI(\prec_w, b))$.
- 5. Let A(u, x, y, z) be a formula with the free variables shown;

$$TR(y, A, \prec_w, z) := \forall x \forall u(x\eta \operatorname{Field}(\prec_w) \to (u\eta y(x) \leftrightarrow A(u, x, y[x, z)));$$

here $y(x) := \{v : \langle x, v \rangle \eta y\}$ and $y \lceil x := \{\langle u, v \rangle : u \prec_w x \land v \eta y(u)\}.$

- 6. A formula A of L is *elementary extensional in the list x*₁,..., *x_n* iff A belongs to the least class of formulas inductively generated by means of ∧, ¬, ∀y (*where y* ∉ {*x*₁,..., *x_n*}) from atoms of the form *t* = *s*, *Nt*, *t*η*x_i*, provided *x*₁, ..., *x_n* do not occur in *t*, *s*.
- 7. We recall that $x =_e y := \forall u(u\eta x \leftrightarrow u\eta y)$ (η -extensional equality).

 $TR(y, A, \prec_w, z)$ says that y encodes a sequence of predicates $\{y_x\}$, which begins with an initial class z and is indexed by elements in \prec_w -order; each y_x is recursively computed by application of the functional $a \mapsto \{u : A(u, x, a, z)\}$ to the collection (encoded by) y [x of previously defined predicates.

One may wonder whether there is a class y satisfying $TR(-, A, \prec_w, z)$. If the given z is a class, \prec_w is a p.w.o. and A is elementary extensional in the relevant parameters, the answer is affirmative and makes essential use of the level structure of TL.

Lemma 6.7 If A(u, x, y) is an *L*-formula (see Section 2.2), which is elementary extensional in x, y, then we can prove in pure logic:

$$A(u, x, y) \land x =_e x' \land y =_e y' \to A(u, x', y').$$

Theorem 6.8 (The Weyl Principle WP for p.w.o.s) Let A(u, x, y, z) be an L-formula elementary extensional in y, z with the free variables shown. Then we have, provably in TLR⁻:

1. $Cl(\prec_w) \land PWO(\prec_w) \land Cl(z) \to \exists y(Cl(y) \land TR(y, A, \prec_w, z)).$

2. Uniqueness: if y and y' are two classes satisfying $TR(-, A, \prec_w, z)$, then y and y' are pointwise extensionally equivalent, i.e., TLR^- proves:

$$[Cl(\prec_w) \land PWO(\prec_w) \land Cl(z) \land TR(y, A, \prec_w, z) \land TR(y', A, \prec_w, z) \land \land Cl(y) \land Cl(y')] \rightarrow \forall x(x\eta \operatorname{Field}(\prec_w) \rightarrow (y(x) =_e y'(x))).$$

Proof: (1) Existence: Put $gxzy := \{u : A(u, x, y | x, z)\}$. Then by the fixed point for operations (Lemma 2.1.2) we can find a term $RC[g, \prec_w]$ such that

$$RC[g, \prec_w]zx = gxz\Sigma(\prec_w [x, \lambda u.RC[g, \prec_w]zu).$$

(Σ is the join operation of Proposition 5.4.2).

Also, if z and \prec_w are classes, then z, \prec_w , and Field(\prec_w) are k-classes for some k (by Lemma 5.5.2, directedness and unboundedness of \prec and Proposition 5.3.6). Consider

$$d := \{x : x\eta_k \text{Field}(\prec_w) \land Cl_k(RC[g, \prec_w]zx)\}.$$

If we choose j > k, d is a j-class (its defining condition being a k-formula; see Proposition 4.2.6). Hence d is a class and we can apply induction on \prec_w . Assume $x\eta$ Field(\prec_w) and $\forall y \prec_w x.y\eta d$: then by Proposition 5.3.3, $RC[g, \prec_w]zy$ is a k-class, for each $y \prec_w x$. Hence by closure of CL_k under join (Proposition 5.4.2), the term $t := \Sigma(\prec_w [x, \lambda u. RC[g, \prec_w]zu)$ is a k-class and so is t[x]. Since A(u, x, y, z) is elementary in y and z and CL_k is closed under elementary comprehension, gxzt = $RC[g, \prec_w]zx$ is a k-class, which implies $x\eta_j d$, whence $x\eta d$. Therefore the class d is \prec_w -progressive and we can conclude that $RC[g, \prec_w]zx$ is a k-class for every x in the field of \prec_w , whence, again by join,

 $RC[A, \prec_w, z] := \Sigma(b, \lambda u. RC[g, \prec_w] zu)$ is a k-class (where b is Field(\prec_w)).

If *x* is in the field of \prec_w , we have with Proposition 5.4 and the extensionality property of *A*:

$$u\eta RC[A, \prec_w, z](x) \iff u\eta RC[g, \prec_w]zx$$

$$\Leftrightarrow u\eta gxz(\Sigma(\prec_w [x, \lambda u. RC[g, \prec_w]zu))$$

$$\Leftrightarrow A(u, x, \Sigma(\prec_w [x, \lambda u. RC[g, \prec_w]zu)[x, z))$$

$$\Leftrightarrow A(u, x, RC[A, \prec_w, z][x, z).$$

In the last step, we use the fact that if x is in $b := \text{Field}(\prec_w)$,

$$\langle v, u \rangle \eta(\Sigma(\prec_w [x, \lambda u. RC[g, \prec_w]zu)) [x \leftrightarrow \langle v, u \rangle \eta(\Sigma(b, \lambda u. RC[g, \prec_w]zu)) [x.$$

It follows that $RC[A, \prec_w, z]$ is a class satisfying $TR(-, A, \prec_w, z)$.

(2) The uniqueness (modulo extensional equivalence) follows by applying transfinite induction to $B(x) := x\eta$ Field(\prec_w) $\rightarrow \forall u(u\eta y(x) \leftrightarrow u\eta y'(x))$ (Note that $\{u : B(u)\}$ is a class if y, y' are classes and $x\eta$ Field(\prec_w)).

Remark 6.9

- The existence of ω-sequences of properties, obtained by iterating a given predicative operation (here the map x → {u : A(u, x, y[x, z)}), is stated in [31], "Iterationsprinzip," p. 27. Thus the schema embodied in Theorem 6.8 is an extension of the Iterationsprinzip to p.w.o.s.
- A special form of WP. Assume number-theoretic induction for classes (see also Definition 8.1 below). Observe that Proposition 5.4.2 and Theorem 6.8 are schematic in the choice of the pairing function. In particular, if we interpret (-, -) as a number-theoretic pairing operation (i.e., an injection of N × N into N) and we assume that,
 - (a) the parameters \prec_w and z in the statement of WP are subclasses of \mathbb{N} ;
 - (b) $\{u : A(u, x, b \lceil x, z)\} \subseteq \mathbb{N}$, whenever $x\eta \text{Field}(\prec_w), z, y \lceil x \subseteq \mathbb{N}$ (*A* elementary extensional in *b*, *z*),

then we obtain a subclass *y* of \mathbb{N} such that $TR(y, A, \prec_w, z)$.

7 A recursion-theoretic interpretation We produce a model of TLR⁻ plus additional principles on numbers and the level ordering. The construction is carried out within a fragment of (powerless) set theory and it hinges upon admissible set theory and generalized recursion theory. The results we presuppose are covered in Richter and Aczel [28], Barwise [3], and Hinman [20]. However, in order to make the paper reasonably self-contained, we define the basic notions and state the required results.

Step 1: the ground model. We fix a countable model \mathcal{M} of combinatory logic **TO**; \mathcal{M} can be assumed to be arithmetically definable. To be more definite, we identify \mathcal{M} with the closed term model *CTM* of combinatory logic.

Definition 7.1

- Let *CTM* := {*t*: *t* is a closed term in the language *L*_{op}} (thus *LT* does not occur in elements of *CTM*). The *closed term model CTM* is the structure ⟨*CTM*, *, =, *K*, *S*⟩, where *K*, *S* are the basic combinators and
 - *: $CTM^2 \rightarrow CTM$ is the operation of juxtaposition of terms (i.e., t * s = Ap(t, s));
 - $= \subseteq CTM^2$ and t = s holds iff **TO** $\vdash t = s$.
- 2. We also let $N^* := \{t : t \in CTM \text{ and } \mathbf{TO} \vdash t = \overline{n}, \text{ for some } n \in \omega\}.$

That the model is well defined is ensured by the Church-Rosser theorem (see [2]); moreover N^* is isomorphic with ω . Thus we have the following.

Lemma 7.2

- 1. CTM is a nontrivial model of **TO**.
- 2. The expansion $CTM^* := \langle CTM, N^* \rangle$ satisfies the axioms:
- NAT.1 $N\overline{0} \land \forall x(Nx \rightarrow (N(x+1) \land \neg(x+1) = \overline{0} \land \text{PRED}(x+1) = x));$
- NAT.2 $\forall x \forall y \forall u \forall v (Nx \land Ny \land \neg x = y \rightarrow Dxxuv = u \land Dxyuv = v);$
- NIND $A(\overline{0}) \land \forall x(A(x) \to A(x+1)) \to \forall x(Nx \to A(x))$ (A an arbitrary formula).

(For notations, see after Lemma 2.2).

Step 2: set-theoretic preliminaries. The next step is the interpretation of the level axioms: the level ordering is identified with the standard ordering relation on ordinals $< \iota$, the first recursively inaccessible ordinal, and the projection operation *LT* is assigned a suitable projection of ι into ω and hence into *CTM*. We recall the relevant set-theoretic notions.

The pure set-theoretic language \mathcal{L}_s is the standard first-order language with identity and a binary predicate symbol \in for membership. If X is a predicate symbol $\neq \in$, $\mathcal{L}_s(X)$ is $\mathcal{L}_s \cup \{X\}$; $\mathcal{L}_s(X)$ has new atoms of the form Xt; the intended meaning of X is that X is a class (in set-theoretic sense).

 L_{α} is the collection of constructible sets up to the ordinal α , where $L_0 = \emptyset$, $L_{\lambda} = \bigcup \{L_{\alpha} : \alpha < \lambda\}(\lambda \text{ limit})$, and $L_{\alpha+1}$ is the family of subsets of L_{α} first order definable with parameters in the standard set-theoretic language over the structure $\langle L_{\alpha}, \in [L_{\alpha} \rangle$.

 $L := \bigcup \{L_{\alpha} : \alpha \in ON\}$ is the constructible universe. A set-theoretic formula *A* is Σ_1 iff *A* has the form $\exists zB$, for some bounded formula *B*; *B* is bounded if it contains only bounded set quantifiers (i.e., of the form $\forall y \in z, \exists y \in z$). The principle of bounded collection is the schema:

$$\forall x \in u \exists y A(x, y) \rightarrow \exists w \forall x \in u \exists y \in w. A(x, y) \text{ (A bounded)}.$$

When we deal with semantical notions (e.g., definability over *L*), we tacitly assume that the set-theoretic language is expanded with (distinct) constants for (distinct) parameters from a suitably large segment of *L*; but we use the same symbol for the object $a \in L$ and its name. Lower case Greek letters will range over the class *ON* of ordinal numbers.

Definition 7.3

- 1. An ordinal $\alpha > \omega$ is *admissible* iff α is a limit ordinal and L_{α} satisfies the bounded collection schema (equivalently, L_{α} is a model of Kripke-Platek set theory KP plus infinity; cf. [3]).
- 2. An admissible ordinal $\alpha > \omega$ is *recursively inaccessible* iff it is the limit of the admissible ordinals $< \alpha$.
- 3. ι := the smallest recursively inaccessible ordinal.
- 4. If *C* is a class of ordinals and *P* any (set-theoretic) class, an *n*-ary relation *R* is *uniformly* $\Sigma_1(L_\alpha)$ in *P* for $\alpha \in C$, iff there exists a Σ_1 -formula $A(x_1, \ldots, x_n, X)$ of the expanded set-theoretic language $\mathcal{L}_s(X)$, such that, if $\alpha \in C$, then

$$R \cap L_{\alpha} = \{ \langle c_1, \ldots, c_n \rangle : c_1, \ldots, c_n \in L_{\alpha}, \}$$

and

$$\langle L_{\alpha}, P \cap L_{\alpha} \rangle \models A(c_1, \ldots, c_n, X) \},$$

where *X* is interpreted by $P \cap L_{\alpha}$.

If *R* is uniformly $\Sigma_1(L_\alpha)$ in *P* for $\alpha \in C$ together with its complement, we say *R* is uniformly $\Delta_1(L_\alpha)$ in *P* for $\alpha \in C$. An *n*-ary relation $R \subseteq L_\alpha^n$ is $\Sigma_1(L_\alpha)(\Delta_1(L_\alpha))$ iff *R* is uniformly $\Sigma_1(L_\alpha)(\Delta_1(L_\alpha))$ in $P = \emptyset$ for $\alpha \in C$, *C* being $\{\alpha\}$.

- 5. A (possibly partial) function *F* is uniformly $\Sigma_1(L_\alpha)$ in a class *P* for $\alpha \in C$ iff its graph is uniformly $\Sigma_1(L_\alpha)$ in *P* for $\alpha \in C$. Then by Definition 7.3.4, it makes sense to speak of a $\Sigma_1(L_\alpha)$ -, $\Delta_1(L_\alpha)$ -function, etc.
- 6. *LEV*_{*i*} is the structure $\langle \iota, =, \leq \rangle$, where $=, \leq$ are respectively the equality and the less-than-equal relations restricted to ordinals $< \iota$.
- 7. α is *projectible* iff there exists a $\Sigma_1(L_\alpha)$ -injection from α into ω .

Remark 7.4 If a $\Sigma_1(L_\alpha)$ -function $F : L_\alpha \to L_\alpha$ is total, then F is also $\Delta_1(L_\alpha)$. In general, every partial $\Sigma_1(L_\alpha)$ -function $F : C \to L_\alpha$, whose domain $C \subseteq L_\alpha$ is $\Delta_1(L_\alpha)$, can be extended to a total $\Sigma_1(L_\alpha)$ -function. As a consequence, the relation $R_F(a, x) := a \in F(x)$ is $\Delta_1(L_\alpha)$, provided F is $\Sigma_1(L_\alpha)$ and total, or defined on a $\Delta_1(L_\alpha)$ -subset. The same considerations hold for uniform definability.

Now let the level variables range over ordinals below ι , whereas level identity and \leq are realized on ordinal theoretic =, and \leq (in the given order); then we trivially have the following.

Lemma 7.5 *LEV*_{*i*} *is a model of the level axioms of Section 3.2. Indeed, the model satisfies linearity and well-foundedness of* \prec *.*

For convenience, we identify closed terms of \mathcal{L}_{op} with their respective number codes in the arithmetized version of *CTM*, and hence we regard *CTM* as a subset of ω . Since every arithmetically definable subset is definable by a bounded formula on L_{α} , for $\alpha > \omega$, we have the following by inspection of the definition of the term model *CTM*.

Lemma 7.6 The sets CTM, N^* , the application function $*: CTM \times CTM \rightarrow CTM$, and the conversion relation = on CTM are all elements of L_{α} , for every $\alpha > \omega$.

Step 3: satisfying the projectibility axiom. The choice of a denotation for the function symbol LT requires an injection IN of ι into CTM, which is reasonably defined. First, we summarize a few facts, to be applied later.

Lemma 7.7

- 1. The predicate $Ad(\alpha) := ``\alpha is admissible" is uniformly <math>\Delta_1(L_\beta)$ for β limit > ω .
- 2. The operation $\beta \mapsto \beta^+ =$ the least admissible > β is uniformly $\Delta_1(L_{\alpha})$, for α limit of admissibles.
- 3. Let $\tau_0 = \omega$ and $\tau_{\alpha} = \text{least admissible } \gamma > \tau_{\beta}$, for every $\beta < \alpha$, whenever $\alpha > 0$. Then the sequence $\langle \tau_{\alpha} : \alpha < \delta \rangle$ is uniformly $\Delta_1(L_{\tau_{\delta}})$.
- 4. ι is the least α such that $\tau_{\alpha} = \alpha$. In particular the restriction of τ to ι is $\Sigma_1(L_{\iota})$.

Proof: (1) follows by standard techniques of formal set-theoretic semantics and the well-known uniform Δ_1 -definability of the operation $\delta \mapsto L_{\delta}$ (see [3]; Devlin [10]); (2) is immediate by (1), and (3) is a consequence of (1)–(2) and closure of admissible sets under Σ_1 -recursion. (4) is an easy corollary of (3).

Lemma 7.8 (after [28]) There exists a function IN, uniformly $\Sigma_1(L_\beta)$ for β admissible $> \omega$, such that $IN[\tau_\alpha : \tau_\alpha \to \omega$ is total and injective, for every $0 < \alpha \le \iota$ (here $IN[\tau_\alpha \text{ is the restriction of IN to } \tau_\alpha)$.

Proof (sketch): The theorem is implied by the existence of an arithmetical notation system N_{ι} for ι ; this in turn is obtained by iteration of a suitable nonmonotone operator. Let us recall the basic definitions.

1. If $\Gamma : \mathcal{P}(\omega) \to \mathcal{P}(\omega)$ ($\mathcal{P}(\omega) = \text{power set of } \omega$), where Γ is possibly nonmonotone, we recursively define for $\alpha \in ON$:

$$I(\Gamma, \alpha) = \bigcup \{ \Gamma(I(\Gamma, \beta)) : \beta < \alpha \};$$

$$I(\Gamma) = \bigcup \{ I(\Gamma, \alpha) : \alpha \in ON \};$$

if $n \in I(\Gamma), |n| = \text{least } \alpha \text{ such that } n \in I(\Gamma, \alpha + 1).$

Since the sequence $\langle I(\Gamma, \alpha) : \alpha \in ON \rangle$ is nondecreasing with respect to inclusion, it makes sense to define the closure ordinal $|\Gamma|$ of Γ ,

$$|\Gamma| := \text{least } \alpha \text{ such that } I(\Gamma, \alpha) = I(\Gamma, \alpha + 1).$$

 Γ is called arithmetical if there exists an arithmetical formula A(u, X) (i.e., A is built up by means of boolean operations and number quantifiers from atoms of the form t = s and Xt), such that, if $P \subseteq \omega$, then

$$\Gamma(P) := \{m \in \omega : (m, P) \text{ satisfies } A(u, X) \text{ in the standard model of arithmetic} \}.$$

Observe that arithmetical formulas become Δ_1 in L_{α} if $\alpha > \omega$; hence by Δ_1 -recursion we have the following.

- If Γ is arithmetical, the sequence (I(Γ, β) : β < α) is uniformly Σ₁(L_α) for α admissible > ω. In particular, I(Γ, β) ∈ L_α, for each β < α.
 - By corollary 9.4 (i) of [28], we have the following essential result.
- 3. There exists an arithmetical operator Γ whose closure ordinal is ι , i.e., $|\Gamma| = \iota$.

Now, if Γ is the operator given by (3), $I(\Gamma, \alpha)$ is a proper subset of $I(\Gamma, \alpha + 1)$ for every $\alpha < \iota$ and the function

 $IN(\alpha) = \text{least } n \in \omega \text{ with } n \in I(\Gamma, \alpha + 1) \text{ and } n \notin I(\Gamma, \alpha)$

is always defined on ι and is trivially injective. By (2) *IN* satisfies the required definability conditions.

Remark 7.9

- 1. In the statement of Lemma 7.8 we can assume that the range of the projection *IN* is *CTM*. Indeed, it is enough to consider the obvious Σ_1 -bijection σ between ω and the set of numerals of *CTM*. Henceforth, we still maintain *IN* as a symbol for the projection of ι into *CTM*.
- Lemma 7.8 still holds if we replace *i* by much larger ordinals, e.g., the first recursively Mahlo ordinal. However, the uniformity of the function *IN* is not shared by all countable admissible projectible ordinals, since there exist non-projectible ordinals below projectible ones (see [3], [20], p. 424).

To sum up, by Lemmas 7.2, 7.5 and 7.8, if we realize the function symbol LT on the map *IN*, we have the following.

Lemma 7.10 The structure $CTM_i = \langle CTM^*, LEV_i, IN \rangle$ is a model of **TO** extended by the level axioms of Section 3.2, the projectibility axiom **PROJ** of Axiom 2.5 and the number-theoretic axioms NAT.1, NAT.2, NIND of Lemma 7.2.

Step 4: satisfying the truth axioms. We expand CTM_{ι} with a family \mathcal{V} of truth predicates indexed by the ordinal ι , such that $\langle CTM_{\iota}, \mathcal{V} \rangle$ is a model of an extension of TLR⁻. The model construction requires only closure of admissible sets under Σ_1 -recursion and Σ_1 -inductive definitions (see [3], p. 26; p. 208) and the fact that ι is an admissible, which is limit of smaller admissibles. We split the construction in a sequence of lemmas.

Henceforth, we assume that the language \mathcal{L}_V of Section 2 is enlarged to a language \mathcal{L}_V^+ with constants for ordinals $< \iota$; lower case Greek letters represent both ordinals $< \iota$ and their names in \mathcal{L}_V^+ . a, b, c are used as metavariables for *arbitrary elements of CTM*, while we keep *i*, *j*, *k* ranging over level variables.

If t is a closed term of \mathcal{L}_{V}^{+} , possibly containing LT and ordinal constants, $CTM_{\iota}(t)$ is the value of t in CTM_{ι} : in other words, $CTM_{\iota}(t)$ is the unique closed term of \mathcal{L} , obtained from t by replacing each subterm of the form $LT(\alpha)$ by $IN(\alpha)$ (which is a term of CTM by Remark 7.9; of course, the first occurrence of α stands for the name of α in the expanded language \mathcal{L}_{V}^{+}).

We lift to the present context the notations and conventions of Section 2; if *a*, *b* are elements of *CTM* and $\alpha < \iota$, then *ab*, $\forall a, \neg a, a \land b, tr(\alpha, a), id(a, b), tr(a), nat(a)$ denote the following elements of *CTM* (in the given order): Ap(a, b), ALL *a*, NEG *a*, AND *ab*, $CTM_{\iota}([T_{\alpha}a]), [a = b], [Ta], [Na].$

Combining Lemma 2.2, Remark 7.4 and Lemma 7.8 yields the following.

Lemma 7.11

- 1. $tr(\alpha, a) = \langle \overline{7}, IN(\alpha), a \rangle$ and the operation $\langle \alpha, a \rangle \mapsto tr(\alpha, a) \in CTM$ is injective (in each coordinate separately);
- 2. *if* $\beta < \tau_{\alpha}$, $c, a \in CTM$, the relation $R(c, \beta, a) := c = tr(\beta, a)$ is uniformly $\Delta_1(L_{\tau_{\alpha}})$ for every $\alpha \leq \iota$.

Hence if $\beta < \tau_{\alpha}$, the function $a \mapsto tr(\beta, a)$ is uniformly $\Delta_1(L_{\tau_{\alpha}})$, for every $\alpha \leq \iota$.

Definition 7.12

1. If $S \subseteq \iota \times CTM$, let, for $\alpha < \iota$:

$$S(\alpha) := \{a : a \in CTM \text{ and } \langle \alpha, a \rangle \in S\}.$$

The structure $\langle CTM_{\iota}, S \rangle$ is the realization of \mathcal{L}_{V^+} , in which T_{α} is interpreted by $S(\alpha)$ ($\alpha < \iota$) and *T* is assigned the set $\cup \{S(\alpha) : \alpha < \iota\}$.

- 2. If $\delta < \iota$, $S \subseteq \delta \times CTM$, $X \subseteq CTM$, $\Gamma(\delta, S, X)$ is the subset of *CTM* such that $a \in \Gamma(\delta, S, X)$ iff for some $b, c \in CTM$, one of the following cases holds:
 - (a) $a = (\neg)tr(\beta, b)$ and $b \in S(\beta)(b \notin S(\beta))$, for some $\beta < \delta$;
 - (b) $a = (\neg)id(b, c)$ and $CTM \models (\neg)b = c$;
 - (c) $a = (\neg)nat(b)$ and $b \in N^*(b \notin N^*; \text{ cf. Definition 7.1});$
 - (d) $a = (\neg)tr(b)$ and $b \in X((\neg b) \in X)$;
 - (e) $a = (\neg)tr(\delta, b)$ and $b \in X((\neg b) \in X)$;

- (f) $a = \neg \neg b$ and $b \in X$;
- (g) $a = (\neg)b \land c$ and $b, c \in X$ (respectively $(\neg b) \in X$ or $(\neg c) \in X$);
- (h) $a = (\neg) \forall b$ and for every $d \in CTM$, $(bd) \in X$ (for some $d \in CTM$, $(\neg bd) \in X$).

Lemma 7.13 Assume $\delta < \iota, S \subseteq \delta \times CTM$.

1. Γ *is monotone in the third variable:*

$$X \subseteq Y \subseteq CTM$$
, then $\Gamma(\delta, S, X) \subseteq \Gamma(\delta, S, Y)$.

2. $\Gamma(\delta, S, X)$ is uniformly $\Delta_1(L_{\alpha})$ in X, S for α admissible with $\omega < \alpha < \iota$ and $\delta < \alpha$. L_{α} is closed under Γ in the following sense: if $\delta < \alpha$, X and S are $\Delta_1(L_{\alpha})$, then $\Gamma(\delta, S, X) \in \mathcal{P}(CTM) \cap L_{\alpha}(\mathcal{P} = the power set operation).$

Proof: (1) Its defining condition positively depends on X.

(2) By inspection of Definition 7.12.2, Lemma 7.11.2, and Remark 7.4, we see that $\Gamma(\delta, S, X)$ is uniformly $\Delta_1(L_{\alpha})$ in X, S and we can apply Δ_1 -separation for L_{α} , since $\Gamma(\delta, S, X) \subseteq CTM \in L_{\alpha}$.

Lemma 7.14 (Inversion) Let $\delta < \iota$, $S \subseteq \delta \times CTM$, $X \subseteq CTM$, $a, b \in CTM$. Then: if A has the form a = b, $\neg a = b$, Na, $\neg Na$, $[A] \in \Gamma(\delta, S, X)$ iff $CTM^* \models A$ (see Lemma 7.2 for CTM^*);

 $tr(\beta, a) \in \Gamma(\delta, S, X) \text{ iff either } \beta = \delta \text{ and } a \in X \text{ or } \beta < \delta \text{ and } a \in S(\beta);$ $(\neg tr(\beta, a)) \in \Gamma(\delta, S, X) \text{ iff either } \beta < \delta \text{ and } a \notin S(\beta) \text{ or } \beta = \delta \text{ and } (\neg a) \in X;$ $(a \land b) \in \Gamma(\delta, S, X) \text{ iff } a \in X \text{ and } b \in X;$ $(\neg(a \land b)) \in \Gamma(\delta, S, X) \text{ iff } (\neg a) \in X \text{ or } (\neg b) \in X;$ $(\forall a) \in \Gamma(\delta, S, X) \text{ iff } (ac) \in X, \text{ for all } c \in CTM;$ $(\neg(\forall a)) \in \Gamma(\delta, S, X) \text{ iff } (\neg(ac)) \in X, \text{ for some } c \in CTM;$ $(tr(a)) \in \Gamma(\delta, S, X) \text{ iff } a \in X;$ $(\neg tr(a)) \in \Gamma(\delta, S, X) \text{ iff } (\neg a) \in X;$ $(\neg \neg a) \in \Gamma(\delta, S, X) \text{ iff } a \in X.$

Proof: From right to left, it holds by definition of Γ . Conversely, we apply the independence Lemma 2.6.

Definition 7.15 Let $\delta < \iota$, $S \subseteq \delta \times CTM$. The β -th iteration $It(\Gamma, \delta, S, \beta)$ of Γ is recursively defined by

 $It(\Gamma, \delta, S, 0) = \emptyset; \text{ for } \lambda \text{ limit, } It(\Gamma, \delta, S, \lambda) = \bigcup \{It(\Gamma, \delta, S, \beta) : \beta < \lambda\};$ $It(\Gamma, \delta, S, \beta + 1) = \Gamma(\delta, S, It(\Gamma, \delta, S, \beta)).$

Clearly $\xi < \zeta$ implies $It(\Gamma, \delta, S, \xi) \subseteq It(\Gamma, \delta, S, \zeta)$ by monotonicity of Γ .

Lemma 7.16 Let $S \subseteq \delta \times CTM$ and $\delta < \iota$.

1. If $a \in CTM$, $\beta < \alpha$, and $\delta < \alpha$, then the relation

$$P(a, \delta, \beta, S) := "a \in It(\Gamma, \delta, S, \beta)"$$

and the function $\beta \mapsto It(\Gamma, \delta, S, \beta)$ are uniformly $\Delta_1(L_\alpha)$ in S for α admissible with $\omega < \alpha < \iota$. Hence if S is $\Delta_1(L_\alpha)$, $It(\Gamma, \delta, S, -) : \alpha \to L_\alpha \cap \mathcal{P}(CTM)$.

- 2. If $\gamma = \alpha^+$, $\delta < \alpha$, α is admissible with $\iota > \alpha > \omega$, S is $\Delta_1(L_\alpha)$ and $I := It(\Gamma, \delta, S, \alpha)$, then
 - (*) $I = \Gamma(\delta, S, I);$
 - (**) $I \in L_{\gamma} \cap \mathcal{P}(CTM)$.

Proof: (1) $It(\Gamma, \delta, S, -)$ is recursively and uniformly defined by means of the operation Γ , which is uniformly $\Delta_1(L_{\alpha})$ in S, and we can apply Lemma 7.13.2 and closure of L_{α} under Δ_1 -recursion.

(2) If L_{α} is admissible, the least fixed point of any given positive $\Sigma_1(L_{\alpha})$ operator is $\Sigma_1(L_{\alpha})$ (this is Gandy's theorem, [3], pp. 208–210). Hence by definition *I* is a $\Delta_1(L_{\gamma})$ -subset of $CTM \in L_{\gamma}$ and (*) is immediate by Lemma 7.13.

Definition 7.17 If $\delta < \iota$, *It* is the functional of Definition 7.15, let

(+)
$$\mathcal{V}(\delta) = It(\Gamma, \delta, \mathcal{V}]\delta, \tau_{\varphi(\delta)}),$$

where $\mathcal{V} \upharpoonright \delta = \{ \langle \beta, a \rangle : \beta < \delta \text{ and } a \in \mathcal{V}(\beta) \}$ and $\varphi(\delta) = \delta$ if δ is a limit; else $\varphi(\delta) = \delta + 1$. \mathcal{V} is well defined on ordinals $< \iota$, by Δ_1 -recursion and Lemmas 7.16 and 7.7.

In the following \mathcal{V} denotes the unique function satisfying (+) above.

Lemma 7.18

- 1. The relation $R(\delta, a) := a \in \mathcal{V}(\delta)$ is uniformly $\Delta_1(L_{\tau_{\varphi(\delta)+1}})$ for every $\delta < \iota$. Hence $\mathcal{V}(\delta) \in L_{\tau_{\varphi(\delta)+1}}$ and $\mathcal{V} : \iota \to L_{\iota} \cap \mathcal{P}(CTM)$ is $\Delta_1(L_{\iota})$;
- 2. if $\delta < \iota$, $\mathcal{V}(\delta) = \Gamma(\delta, \mathcal{V} \mid \delta, \mathcal{V}(\delta))$;
- 3. if $\delta < \iota$, either $a \notin \mathcal{V}(\delta)$ or $(\neg a) \notin \mathcal{V}(\delta)$, for every $a \in CTM$;

4. for every $\beta < \delta < \iota$, $a \in CTM$, either $tr(\beta, a) \in \mathcal{V}(\delta)$ or $(\neg(tr(\beta, a)) \in \mathcal{V}(\delta);$ 5. if $\beta < \delta < \iota$, $\mathcal{V}(\beta)$ is a proper subset of $\mathcal{V}(\delta)$.

Proof: (1) and (2) follow from Lemmas 7.7 and 7.16 and closure of admissible sets under Δ_1 -recursion.

(3) By main transfinite induction on $\delta < \iota$, and a secondary induction on $\tau_{\varphi(\delta)}$, using (1) and the inversion Lemma 7.14 at the successor step.

(4) Assume $\beta < \delta$ and $a \in \mathcal{V}(\beta)$. It follows that $tr(\beta, a) \in \Gamma(\delta, \mathcal{V} | \delta, \emptyset) \subseteq \Gamma(\delta, \mathcal{V} | \delta, \mathcal{V}(\delta)) \subseteq \mathcal{V}(\delta)$ by definition of Γ , monotonicity and (2)); if $a \notin \mathcal{V}(\beta)$, the argument is similar.

(5) If $\beta < \delta$ and $\mathcal{V}(\beta, \xi) := It(\Gamma, \beta, \mathcal{V} \mid \beta, \xi)$, it is enough to verify by induction on $\xi < \tau_{\varphi(\beta)}$:

$$\mathcal{V}(\beta,\xi) \subseteq \mathcal{V}(\delta) \tag{1}$$

If ξ is 0 or a limit, the proof is trivial. Assume (1) by IH and $a \in \Gamma(\beta, \mathcal{V} | \beta, \mathcal{V}(\beta, \xi))$: we show $a \in \mathcal{V}(\delta)$ as a consequence of the inversion lemma and the property mentioned in Lemma 7.18.2 above. We have to distinguish several cases according to the form of *a*. Let $a = (\neg(tr(\nu, b)))$ for some ν : then by inversion either $\nu < \beta$

and $b \notin \mathcal{V}(v)$ or $v = \beta$ and $(\neg b) \in \mathcal{V}(\beta, \xi)$. In the first case, since $v < \delta$, $a \in \Gamma(\delta, \mathcal{V} | \delta, \emptyset) \subseteq \mathcal{V}(\delta)$ by definition of Γ and Lemma 7.18.2. In the second case, $(\neg b) \in \mathcal{V}(\beta)$ by definition and hence $b \notin \mathcal{V}(\beta)$ by consistency (see Lemma 7.18.3 above). Since $\beta < \delta$, $a \in \Gamma(\delta, \mathcal{V} | \delta, \emptyset) \subseteq \mathcal{V}(\delta)$. Let $a = (b \land c)$; by assumption and inversion $b \in \mathcal{V}(\beta, \xi)$ and $c \in \mathcal{V}(\beta, \xi)$, whence $b, c \in \mathcal{V}(\delta)$ by IH. By definition of Γ , $a \in \Gamma(\delta, \mathcal{V} | \delta, V(\delta))$ and $a \in \mathcal{V}(\delta)$ by Lemma 7.18.2. The extant cases are easily checked as an exercise. As to proper inclusion, consider the term $R(\beta) := \{x : \neg x\eta_{\beta}x\}$ and observe that Lemmas 7.18.2 and 7.18.4 imply $(R(\beta)R(\beta)) \in \mathcal{V}(\beta+1) - \mathcal{V}(\beta)$ (see Proposition 5.3).

Definition 7.19 TLR⁺ is the extension of TLR⁻ that includes the axioms NAT.1–NAT.2, the schema of number-theoretic induction NIND (see Lemma 7.2), plus the schema of transfinite induction on levels TI(lev), where,

$$TI(lev) := \forall i (\forall j \prec i.B(j) \rightarrow B(i)) \rightarrow \forall iB(i)(B \text{ arbitrary}).$$

Theorem 7.20 (Soundness) $C_i = \langle CTM_i, V \rangle \models TLR^+$.

Proof: NIND, NAT.1–NAT.2, the level axioms, TI(lev) and the projectibility axioms hold in the model by Lemma 7.10. The local truth axioms of Section 3.1 and the connection axioms of Section 3.3 are straightforward consequences of the definition of Γ , inversion, Definition 7.12 and the previous lemma. As to the reflection principle, assume,

$$C_{\iota} \models \forall x \exists j (x\eta_{\gamma}a \to x\eta_{j}b) \ (for \ a, b \in CTM \ and \ \gamma < \iota). \tag{2}$$

By $\Delta_1(L_i)$ -definability of $R(\alpha, a) := a \in \mathcal{V}(\alpha)$ (by Lemma 7.18.1), condition 2 is equivalent, by the well-known absoluteness of Δ_1 -conditions, to:

$$L_{\iota} \models (\forall x \in CTM)(\exists \beta) A(x, \gamma, \beta, a, b),$$
(3)

for a suitable Δ_1 -formula A(x, y, z, u, v); hence by Σ_1 -collection (derivable from bounded collection), for some $\xi < \iota$, we have $L_{\iota} \models (\forall x \in CTM)(\exists \beta < \xi)A(x, \gamma, \beta, a, b)$, which yields by equivalence of equations (2) and (3) the required conclusion $C_{\iota} \models \exists k \forall x \exists j (j \leq k \land (x\eta_{\nu}a \rightarrow x\eta_{j}b)).$

Remark 7.21 It is is possible to strengthen the consistency result in two directions. First of all, the arithmetical definability of ground model and the Kleene basis theorem imply the consistency of a reducibility schema for classes:

RPC
$$i \prec k \land Cl_i(x) \land \exists y(Cl(y) \land A(u, x, y)) \rightarrow \exists y(Cl_k(y) \land A(u, x, y))),$$

for every \mathcal{L} -formula A(u, x, y) with the free variables shown, which is elementary extensional in x, y. Hence, at least for elementary predicates, *quantification on arbitrary classes is reducible to quantification on classes of a fixed level.*

TLR + RPC yields an interpretation of the fragment of second order arithmetic based on Π_1^1 -comprehension. According to the second direction, we can consistently assume that *levels are objects* and hence that the projection function *LT* collapses to identity; we can apply reflective truth to expressions containing bounded level quantification. This move implies that classes are closed under Δ_2^1 -comprehension.

8 *Levels of truth and proof theory* We investigate the relation between theories of truth with levels and standard mathematical systems.

Definition 8.1 TLR is obtained from TLR⁺ (see Definition 7.19) by omitting the full transfinite induction schema for levels TI(lev) and replacing the schema of *N*-induction by the *axiom of local N-induction LIND*:

$$Cl_i(x) \wedge Clos_i(x) \rightarrow \forall u(Nu \rightarrow u\eta_i x),$$

where $Clos_i(x) := \bar{0}\eta_i x \wedge \forall v(v\eta_i x \to (v+1)\eta_i x).$

It turns out that, although TLR is based on the logical notions of truth and iteration of the reflection process, TLR is strictly connected with an important subsystem of second order arithmetic Z_2 : by the Weyl iteration principle of Section 6, it is easy to relate TLR with a well known system ATR₀ of Reverse Mathematics ([16]). We define the system ATR₀.

First of all, the language \mathcal{L}_2 of second order arithmetic contains the following elements: a denumerable list of number variables x_1, x_2, x_3, \ldots ; a denumerable list of set variables X_0, X_1, X_2, \ldots ; the individual constant $\overline{0}$; the function symbols ' (successor, 1-ary), + (addition, 2-ary), \cdot (product, 2-ary); the binary predicates < (ordering on ω) and ϵ (membership); classical logical operations (say \neg , \forall , \land); and =.

 \mathcal{L}_2 -terms are inductively generated from number variables and the constant 0 by application of the function symbols ', \cdot , +. Atoms of \mathcal{L}_2 have the form $t = s, t \in X$, t < s, where t, s are terms, X is a set variable. Formulas are inductively generated from atoms by means of negation, conjunction and universal quantification on individual and set variables. A \mathcal{L}_2 -formula A is *arithmetical* if no set variable occurs bound in A.

Definition 8.2 ATR₀ is the theory in the language \mathcal{L}_2 , which contains classical predicate calculus with identity for \mathcal{L}_2 and

1. standard number-theoretic axioms:

$$\begin{aligned} \forall x (\neg x' = 0) \land \forall x \forall y (x' = y' \to x = y); \\ \forall x (\neg x < \bar{0}) \land \forall x \forall y (x < y \leftrightarrow \exists z (z' + x = y)); \\ \forall x (x + \bar{0} = x) \land \forall x (x \cdot \bar{0} = \bar{0}) \land \forall x \forall y (x + y' = (x + y)' \land x \cdot y' = x \cdot y + x); \end{aligned}$$

- 2. the induction axiom Ax-IND: $\overline{0} \in X \land \forall x (x \in X \to x' \in X) \to \forall x (x \in X);$
- 3. arithmetical comprehension schema: $\exists X \forall u(u \in X \leftrightarrow A(u, Y))$, where A(u, Y) is an arbitrary arithmetical \mathcal{L}_2 -formula and X does not occur in A;
- 4. the schema ATR of arithmetical transfinite recursion:

$$\forall X \forall Z \exists Y (WO(<_X) \rightarrow \forall y \forall u(y \in Y_u \leftrightarrow A(y, u, Y [u, Z))),$$

where $WO(<_X)$ is the formula, stating that *X* encodes a linear ordering of ω such that $\forall Y(\forall x(\forall y(y <_X x \rightarrow y \in Y) \rightarrow x \in Y) \rightarrow \forall x(x \in Y))$ {here $y <_X x := \langle y, x \rangle \in X$; $\langle x, y \rangle$ stands for a primitive recursive pairing function}; *A* is an arithmetical formula and Y [u is contextually defined by $\langle v, y \rangle \epsilon Y [u := v <_X u \land \langle v, y \rangle \in Y$, and $y \in Y_u := \langle u, y \rangle \in Y$.

Theorem 8.3 ATR₀ *is interpretable in* TLR.

Proof: We define a translation * of \mathcal{L}_2 into the level free part of the language of TLR. Informally speaking, we simply verify that N plus the subclasses of N is a model of ATR in TLR. Formally, we can choose combinators $\overline{0}$, \cdot , +, ', in order to interpret the basic function symbols of \mathcal{L}_2 (we adopt the same notation). Hence we can inductively assign to each \mathcal{L}_2 -term t a term t* in the language \mathcal{L}_{op} (= the operational fragment of \mathcal{L}_V), with the same free variables. Moreover, if t = s, $t \in X$, t < s are atoms of \mathcal{L}_2 , we put $(t = s)^* := (t^* = s^*)$; $(t \in X)^* := (t^* \eta x)(x \text{ fresh})$; $(t < s)^* := t^* < s^*$ (the second occurrence of < being now a canonically chosen \mathcal{L}_{op} -definition of <). We then extend * to arbitrary formulas of \mathcal{L}_2 by stipulating that the map* commutes with \neg , \wedge and

$$(\forall XA)^* := \forall x (Cl_N(x) \to A^*), (\forall xA)^* := \forall x (Nx \to A^*) = \forall nA^*,$$

where $Cl_N(x) := Cl(x) \land \forall u(u\eta x \to Nx)$. It is clear that * is a well-defined translation of \mathcal{L}_2 into \mathcal{L} . Let A be an \mathcal{L}_2 -formula with free variables in the list $X = X_0, \ldots, X_n, y = y_0, \ldots, y_k$: then we check by induction on the definition of ATR₀-provability:

if
$$ATR_0 \vdash A(y, X)$$
, then $TLR \vdash Ny \land Cl_N(x) \to A^*(y, x)$. (4)

The translation of the number-theoretic axioms is disposed of by means of local class N-induction and the fixed point theorem (Lemma 2.1 takes care of the existence of plus and times, and suitable \mathcal{L}_{op} -instances of N-induction ensure that the definitions are correct). It remains to check that the *-translations of Ax-IND, arithmetical comprehension schema and ATR are provable in TLR. As to the first axiom, we have to verify

$$TLR \vdash Cl(x) \rightarrow (Clos(x) \rightarrow \mathbb{N} \subseteq x).$$

Since x is an *i*-class for some *i* (Lemma 5.5.2), $Clos(x) \leftrightarrow Clos_i(x)$ by Lemma 5.2.3 and hence we can apply local class *N*-induction.

The translation of arithmetical comprehension becomes an instance of elementary comprehension and hence is provable in TLR by Proposition 5.4.1, Lemma 5.5.2, and Lemma 5.2.3.

Note that, if $Cl_N(X) \wedge (WO(<_X))^*$ is assumed, then $<_X$ encodes a subclass of N which is a p.w.o. Hence if z is any subclass of N and A(u, x, Y, Z) is arithmetical, $u\eta \mathbb{N} \wedge A^*(u, x, y, z)$ is elementary extensional in y, z (y, z fresh variables). Now the hypothesis of Theorem 6.8 and Remark 6.9 are trivially met and there exists a subclass of N satisfying the *-translation of the ATR-consequent.

It is well known that ATR_0 has the same proof-theoretic strength as Predicative Analysis (cf. [16]). On the other hand, the lower bound is also an upper bound for TLR. Indeed, we can state a stronger result.

Theorem 8.4 TLR + TI(lev) is proof-theoretically reducible to ATR_0 (i.e., the formal consistency of TLR + TI(lev) is implied by the formal consistency of ATR_0 over, say, Peano arithmetic). The only proof we are aware of is complex and cannot be given here (details are presented in [9], chs. 10–11). However, the gist of the argument can hopefully be gained from an outline of the basic idea.

Step 1: $TLR + TI(lev) \mapsto STLR$. we give a sequent style presentation STLR (= *sequent calculus of truth with levels and reflection*) of a system which contains TLR and *the full transfinite induction schema on levels*.

Step 2: STLR \mapsto STLR^{∞}. STLR is embedded into an infinitary system STLR^{∞} where TI(lev) is dropped in favor of an ω -rule, which forces the level variables to range over *finite standard ordinals*. Since STLR^{∞} contains a reflection principle for levels, STLR^{∞} cannot have ω -standard models; yet, *because of the weak number-theoretic induction*, STLR^{∞} is consistent. STLR^{∞} enjoys a crucial quasi-normalization property: the cut-rule can be restricted to formulas, which contain only unbounded universal or existential level quantifiers.

Step 3: $STLR^{\infty} \mapsto \{STLR_n : n \in \omega\}$. This is the central step of the constructive interpretation. First we define a sequence of finitary approximations $STLR_n$ to $STLR^{\infty}$, in which only bounded level quantifiers are allowed and where we can explicitly refer only to the first n levels. The main fact to establish is that $STLR^{\infty}$ -theorems can be suitably interpreted in the $STLR_n$'s. The result is based on an asymmetric treatment of unbounded quantifiers on levels according to a potentialistic point of view, so that $\forall j$ only refers to arbitrary *finite segments* of the level ordering. Thus the meaning of $\exists j$ depends on the given initial segments, and this dependence is expressed by majorizing functions whose complexity depends upon the *transfinite ordinal height* of the given quasi-normalized $STLR^{\infty}$ -derivations.

Step 4: $STLR_n \mapsto IT_n^{\infty}$. One carries out a complete elimination of bounded level quantification and level structure: each $STLR_n$ -system is embedded in a level free infinitary system IT_n^{∞} , where the number-theoretic induction schema is replaced by an infinitary rule for N.

Step 5: $IT_n^{\infty} \mapsto RS_n$. We design an infinitary ramified system RS_n in which T_n is split into a family $\{T_n^{\alpha} : \alpha < \Gamma_0\}$ of approximations. The T_n^{α} 's are linked together by natural recursive conditions, which can be encoded by symmetric introduction rules with the cut elimination property (see the model construction of Section 7). We embed IT_n^{∞} into RS_n by a modified version of the asymmetric interpretation technique of Step 3 (see Girard [18], Cantini [5]). An analysis of cut free RS_n -derivations readily implies that RS_n -theorems of level < n (i.e., theorems without T_n^{α} -occurrences) are already derivable without T_n -rules and hence in IT_m^{∞} , for some m < n.

By finite iteration of the *T*-elimination procedure, we finally obtain that *T*-free sentences of TLR have *T*-free (infinitary constructively presented) derivations, whose correctness demands only arithmetical principles and suitable instances of the schema $TI(<\Gamma_0)$ of transfinite induction along each $\alpha < \Gamma_0$. Here Γ_0 is the well known ordinal of predicative analysis.

In view of Theorem 8.4, TLR + TI(lev) can be regarded as *a predicatively reducible theory of degrees of predicative evidence*: the higher the level, the lower the predicative evidence; a truth of level j > i is in generally only conditionally predicative relative to the truths of lower level.

A final remark on a possible criticism. It might be objected that we have re-

stored a hierarchy of truth predicates, which is strictly reminiscent of *the Tarskian language/metalanguage hierarchy*, and it seems that we have destroyed the freedom of the original level-free formalism of truth. This is only partly true: indeed the new framework is quite distant from the Tarskian one. In particular by the local truth axioms of Section 3.1 each T_i already encompasses the standard Tarskian predicates, as to closure properties and self-referential ability. Furthermore, the level structure greatly strengthens the deductive force and it can be profitably applied for justifying level-free principles in the context of type-free systems, as it appears from Sections 6–8.

Acknowledgments This research was partially supported by MURST and CNR funds. Part of this paper was presented at the Workshop on Applicative Theories, University of Bern, March 18–19, 1994.

REFERENCES

- Aczel, P., P. Carlisle, and N. Mendler, "Two frameworks of theories and their implementation in Isabelle," pp. 5–39 in *Logical Frameworks*, edited by G. M. Plotkin et al., Cambridge University Press, Cambridge, 1991. Zbl 0756.03013 MR 1139779 1
- [2] Barendregt, H., *The Lambda Calculus: Its Syntax and Semantics*, North-Holland, Amsterdam, 1984. Zbl 0467.03010 MR 83b:03016 2.3, 7
- [3] Barwise, J., Admissible Sets and Structures, Springer-Verlag, Berlin, 1975.
 Zbl 0316.02047 MR 54:12519 7, 1, 7, 2, 7, 7
- [4] Burge, T., "Semantical paradox," *Journal of Philosophy*, vol. 76 (1979), pp. 169-198.
- [5] Cantini, A., "A note on a predicatively reducible theory of elementary iterated induction," *Bollettino dell'Unione Matematica Italiana*, vol. 4-b (1985), pp. 413–430.
 Zbl 0582.03043 MR 87d:03168 8
- [6] Cantini, A., "Su una teoria generale delle proprietà, basata su schemi di comprensione iterati e privi di tipi," pp. 53–57 in *Atti degli Incontri di Logica Matematica*, vol. 3, edited by R. Ferro and A. Zanardo, Cluep, Padova, 1987. Zbl 0628.03017 1
- [7] Cantini, A., "Notes on formal theories of truth," Zeitschrift für Mathematische Logik und Grundlagen der Mathematik, vol. 35 (1989), pp. 97–130. Zbl 0661.03043 MR 90e:03073
- [8] Cantini, A., "A logic of abstraction related to finite constructive number classes," *Archive for Mathematical Logic*, vol. 31 (1991), pp. 69–83. Zbl 0696.03033
 MR 92j:03027 1
- [9] Cantini, A., "Logical frameworks for truth and abstraction," *Studies in Logic*, vol. 135, North-Holland, Amsterdam, 1995. Zbl 0860.03015 MR 97f:03019 6, 8
- [10] Devlin, K., "Constructibility", pp. 453–502 in *Handbook of Mathematical Logic*, edited by J. Barwise, North-Holland, Amsterdam, 1977. 7
- [11] Feferman, S., "Non-extensional type-free theories of partial operations and classifications I," pp. 73–118 in *Proof Theory Symposium*, edited by J. Diller and G. H. Müller, Springer Lecture Notes in Mathematics, vol. 500, 1974. Zbl 0324.02018 MR 53:10581 5

- [12] Feferman, S., "Iterated inductive fixed point theories: application to Hancock's Conjecture," pp. 171–195 in *Patras Logic Symposion*, edited by G. Metakides, North-Holland, Amsterdam, 1982. Zbl 0522.03045 MR 84f:03050 1, 1
- [13] Feferman, S., "Towards useful type-free theories I," *The Journal of Symbolic Logic*, vol. 49 (1984), pp. 75–111. MR 85i:03068 1
- [14] Feferman, S., "Reflecting on incompleteness," *The Journal of Symbolic Logic*, vol. 56 (1991), pp. 1–49. Zbl 0746.03046 MR 93b:03097 1, 3.4
- [15] Fitch, F. B., "Universal metalanguages in philosophy," *Review of Metaphysics*, vol. 17 (1964), pp. 396–402.
- [16] Friedman, H., K. McAloon, and S. G. Simpson, "A finite combinatorial statement which is equivalent to 1-consistency of predicative analysis," pp. 197–230 in *Patras Logic Symposion Logic*, edited by G. Metakides, North-Holland, Amsterdam, 1982. 1, 8, 8
- [17] Gaifman, H., "Paradoxes of infinity and self applications I," *Erkenntnis*, vol. 20 (1983), pp. 131–155.
- [18] Girard, J. Y., "A survey of Π¹₂-logic," pp. 89–107 in *Logic, Methodology and Philoso-phy of Science VI*, edited by L. Cohen, J. Los, H. Pfeiffer, and K. P. Podewski, North-Holland, Amsterdam, 1982. Zbl 0496.03038 MR 84e:03073 8
- [19] Gordeev, L., "Proof theoretic analysis: weak systems of functions and sets," Annals of Pure and Applied Logic, vol. 38 (1989), pp. 1–121.
- [20] Hinman, P., *Recursion-theoretic Hierarchies*, Springer-Verlag, Berlin, 1978. Zbl 0371.02017 MR 82b:03084 1, 7, 2
- [21] Jäger, G., "The strength of admissibility without foundation," *The Journal of Symbolic Logic*, vol. 49 (1984), pp. 867–879. Zbl 0585.03032 MR 86m:03090 1, 1, 1
- [22] Kripke, S., "Outline of a theory of truth," *Journal of Philosophy*, vol. 72 (1975), pp. 690–716. Zbl 0952.03513
- [23] Lorenzen, P., and J. Myhill, "Constructive definitions of certain analytic sets of numbers," *The Journal of Symbolic Logic*, vol. 24 (1959), pp. 37–49. MR 22:14 1
- [24] Martin-Löf, P., Intuitionistic Type Theory, Bibliopolis, Naples, 1984. Zbl 0571.03030 MR 86j:03005 1
- [25] Marzetta, M., "Universes in the theories of types and names," pp. 340–351 in *Computer Science Logic* '92, edited by E. Börger et al., Lecture Notes in Computer Science, vol. 702, Springer-Verlag, Berlin, 1993. Zbl 0794.03083 MR 94i:03114 1, 1
- [26] Parsons, C., "The liar Paradox," *Journal of Philosophical Logic*, vol. 3 (1974), pp. 381–412. Zbl 0296.02001 MR 58:21402 1
- [27] Reinhardt, W., "Some remarks on extending and interpreting theories with a partial predicate for truth," *Journal of Philosophical Logic*, vol. 15 (1986), pp. 219–251.
 Zbl 0629.03002 MR 87i:03007 1
- [28] Richter, W., and P. Aczel, "Inductive definitions and reflecting properties of admissible ordinals," pp. 301–381 in *Generalized Recursion Theory*, edited by J. E. Fenstad and P. G. Hinman, North-Holland, Amsterdam, 1974. Zbl 0318.02042 MR 52:13344 7, 7.8, 2
- [29] Schütte, K., Beweistheorie, Springer-Verlag, Berlin, 1960. Zbl 0102.24704 MR 22:9438

- [30] Scott, D., "Combinators and classes," pp. 1–26 in λ-Calculus and Computer Science, edited by C. Böhm, Springer Lecture Notes in Computer Science, vol. 37, 1975.
 Zbl 0342.02018 MR 58:21489 1
- [31] Weyl, H., *Das Kontinuum*, Von Veit, Leipzig, 1918. Zbl 0005.15403 MR 22:10886 1, 1

Dipartimento di Filosofia Università degli Studi di Firenze via Bolognese 52 I–50139 Firenze Italy email: cantini@mailserver.idg.fi.cnr.it